

.....
.....
.....
.....

Serviço de Pós-Graduação EESC/USP
EXEMPLAR REVISADO
Data de entrada no Serviço: 16.1.04.03
Ass.: *Kawamoto*

**APLICAÇÃO DE MINERADOR DE DADOS NA
OBTENÇÃO DE RELAÇÕES ENTRE PADRÕES
DE ENCADEAMENTO DE VIAGENS
CODIFICADOS E
CARACTERÍSTICAS SÓCIO-ECONÔMICAS**

Sandra Matiko Ichikawa

Dissertação apresentada à Escola de Engenharia
de São Carlos da Universidade de São Paulo,
como parte dos requisitos para obtenção do
título de Mestre em Engenharia Civil com
ênfase em Transportes

Orientador: Prof. Dr. Eiji Kawamoto



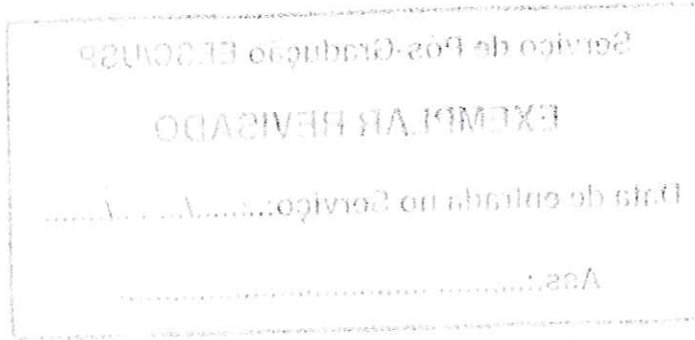
São Carlos
Outubro
2002

DEDALUS - Acervo - EESC



31100053548

Class. TESE-EESC
Cutt. 2755
Tombo T080/03
Sysno 1465371



Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca - EESC/USP

I16a

Ichikawa, Sandra Matiko
Aplicação de minerador de dados na obtenção de
relações entre padrões de encadeamento de viagens
codificados e características sócio-econômicas/ Sandra
Matiko Ichikawa. -- São Carlos, 2002.

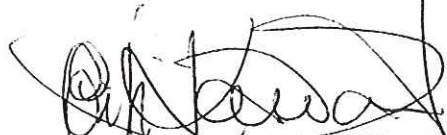
Dissertação (Mestrado) -- Escola de Engenharia de São
Carlos-Universidade de São Paulo, 2002.
Área : Transportes.
Orientador: Prof. Dr. Eiji Kawamoto.

1. Análise de demanda. 2. Teoria de atividades.
3. Encadeamento de viagens. 4. Codificação dos padrões de
viagem. 5. Árvores de decisão. I. Título.

FOLHA DE JULGAMENTO

Candidata: Engenheira SANDRA MATIKO ICHIKAWA

Dissertação defendida e julgada em 29-11-2002 perante a Comissão Julgadora:



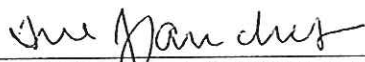
Prof. Dr. **EIJI KAWAMOTO** (Orientador)
(Escola de Engenharia de São Carlos/USP)

APROVADO



Prof. Assoc. **ORLANDO STRAMBI**
(Escola Politécnica/USP)

APROVADO



Profª. Dra. **SUELY DA PENHA SANCHES**
(Universidade Federal de São Carlos/UFSCar)

APROVADA



Prof. Assoc. **JOSÉ REYNALDO ANSELMO SETTI**
Coordenador do Programa de Pós-Graduação em
Engenharia de Transportes



Profª. Assoc. **MARIA DO CARMO CALIJURI**
Presidente da Comissão de Pós-Graduação da EESC

*Há quem diga que todas as noites são de sonhos...
Mas há também quem diga nem todas...
Só as de verão...
Mas no fundo isso não tem muita importância...
O que interessa mesmo não são as noites em si...
São os sonhos...
Sonhos que o homem sonha sempre...
Em todos os lugares, em todas as épocas do ano...
Dormindo ou acordado...*

William Shakespeare

Agradecimentos

Ao Professor Eiji Kawamoto, pela paciência, dedicação e a amizade que marcaram o seu papel de orientador, amigo e principal motivador para meu crescimento profissional. Obrigada pelo tempo que compartilhou comigo.

Ao contribuinte brasileiro e à Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela Bolsa de Estudos concedida.

Aos professores do Departamento de Transportes da EESC/USP, em especial Dr. Edson Martins de Aguiar, Dr. Antônio Nelson Rodrigues da Silva, Dr. João Alexandre Widmer pelo apoio e colaboração.

À Secretaria de Obras, Transportes e Serviços Públicos da Prefeitura Municipal de São Carlos, pelo apoio na conclusão desta pesquisa.

À minha família por tornar possível a realização da minha missão de vida.

Aos amigos que leram os originais e fizeram críticas e comentários que me ajudaram a aprimorar este trabalho. Seria impossível citar todas as pessoas que deram sua parcela de colaboração e que puderam compartilhar com o desenvolvimento do trabalho. A todos, dedico um especial agradecimento.

SUMÁRIO

LISTA DE FIGURAS	i
LISTA DE TABELAS	iii
LISTA DE GRÁFICOS	iv
RESUMO	vi
<i>ABSTRACT</i>	vii
1. INTRODUÇÃO	01
1.1 Objetivos	06
1.2 Motivação	06
1.3 Método.....	07
1.4 Contribuições	09
1.5 Apresentação dos Capítulos	10
2. TEORIA DAS VIAGENS BASEADA NAS ATIVIDADES	12
2.1 Idéias básicas	13
2.2 Estrutura de Decisão de viagens urbanas	14
2.3 Características da demanda de atividades e viagens	17
2.4 Definições.....	20
3. TEORIA DAS VIAGENS BASEADA EM ATIVIDADES E SUAS HIPÓTESES.....	24
4. MINERADOR DE DADOS	34
4.1 Árvores de decisão	37
4.2 Árvores de Classificação e Regressão	39
4.3 Árvores de Classificação e Regressão – S-Plus 2000	40
4.4 Algumas considerações sobre a utilização do modelo de Árvore do S-Plus 2000	42
5. ESTUDO DO CASO DA REGIÃO METROPOLITANA DE SÃO PAULO.....	44
5.1 Pesquisa origem-destino de São Paulo	44

5.2	Tratamento dos Dados	48
5.3	Variáveis extraídas da Pesquisa O/D-87 – RMSP	53
5.4	Codificação dos Dados	54
5.5	Caracterização dos Padrões de Viagens	59
6.	APLICAÇÃO DA ADC E RESULTADOS DO PROCESSAMENTO	63
6.1	Interpretação do Relatório.....	65
6.2	Interpretação do Gráfico	68
7.	ANÁLISE DOS RESULTADOS OBTIDOS	72
8.	CONCLUSÕES E RECOMENDAÇÕES	91
	ANEXOS	100
	REFERÊNCIAS BIBLIOGRÁFICAS	132

LISTA DE FIGURAS

Figura 1.1 -	Evolução das Abordagens de Análise em Atividades	03
Figura 1.2 -	Exemplos de padrões de viagens encontrados para análise de viagens encadeadas	05
Figura 2.1 -	Estrutura de Decisão de Viagens Urbanas	14
Figura 2.2 -	Resposta à política adotada envolvendo o ajuste comportamental motivada pelo desejo de alcançar os objetivos de realizar atividades	16
Figura 2.3 -	Relacionamento entre viagens e atividades por indivíduos de um domicílio	18
Figura 2.4 -	Conceito de "tour" baseado em BOWMAN et al. (1998)	22
Figura 3.1 -	Número de "tours" no padrão diário de viagem	29
Figura 3.2 -	Complexidade de viagens a trabalho	30
Figura 4.1 -	Fases do Minerador de Dados	35
Figura 4.2 -	Esquema de uma Árvore de Decisão	38
Figura 4.3 -	Resultados gráficos da árvore em relação à proporcionalidade	42
Figura 5.1 -	38 municípios da Pesquisa O/D 1987 – RMSP	45
Figura 5.2 -	Área de Abrangência das Pesquisas O/D realizadas em 1967, 1977 e 1987	47
Figura 5.3 -	Distribuição dos Dados da Amostra	52
Figura 5.4 -	Exemplo da forma de codificação dos motivos de viagens por letras	55
Figura 5.5 -	Exemplo da forma de codificação dos motivos de viagens por letras e modos por números	55
Figura 5.6 -	Padrões Utilizados para análise	61

Figura 6.1 -	Árvore de Classificação – Caracterização entre Individuos por padrão de viagem – Desvio mínimo: 0,05	69
Figura 6.2 -	Exemplo das variáveis correspondentes a um grupo.....	70
Figura 6.3 -	Padrões predominantes nos nós da árvore	71
Figura 7.1 -	Grupos formados através da Árvore de Decisão	73
Figura 7.2 -	Primeiros níveis da árvore de decisão - variáveis ESTUDA e TRABALHA	76

LISTA DE TABELAS

Tabela 1.1	- Estimativa do número de alternativas de programação de atividades diárias por indivíduo.....	04
Tabela 2.1	- Atividades e Viagens dos indivíduos do domicílio	18
Tabela 2.2	- Padrões de atividades e de viagens dos indivíduos	19
Tabela 2.3	- Viagens e modos de transporte utilizados pelos indivíduos	19
Tabela 3.1	- Recursos e métodos utilizados nas análises de hipóteses sobre viagens baseadas em atividades	27
Tabela 4.1	- Características gerais do S-Plus 2000	42
Tabela 5.1	- Total de Indivíduos que viajam e não viajam	49
Tabela 5.2	- Total de Indivíduos por Número de Viagens realizadas no dia.....	50
Tabela 5.3	- Viagens Urbanas com origem e destino na Residência	50
Tabela 5.4	- Total de Indivíduos por Número de Viagens no dia	51
Tabela 5.5	- Total de Indivíduos Analisados por Número de Viagens realizadas no dia	53
Tabela 5.6	- Total de Indivíduos Analisados segundo aqueles que viajaram e não viajaram	53
Tabela 5.7	- Ocupação Principal e as Variáveis de Agrupamento "Estuda" e "Trabalha"	57
Tabela 5.8	- Grupos de Indivíduos pelas variáveis "estuda" e "trabalha"	58
Tabela 5.9	- Agrupamento dos Motivos de Viagens	59
Tabela 5.10	- Padrões de Viagens Possíveis	60
Tabela 5.11	- Frequência dos padrões utilizados na análise	62
Tabela 6.1	- Código das variáveis utilizadas e apresentadas na árvore de classificação	63
Tabela 6.2	- Padrões de Viagens em ordem alfabética apresentados no Relatório	67

Tabela 7.1 -	Resumo dos 20 grupos formados	75
Tabela 7.2 -	Probabilidades dos grupos nos primeiros níveis da árvore em relação ao número de viagens	77
Tabela 7.3 -	Variáveis de maior importância para grupos de indivíduos segundo os resultados apresentados pela árvore de decisão	77
Tabela 7.4 -	Distribuição de homens e mulheres nos grupos principais	78
Tabela 7.5 -	Grupos de indivíduos e padrões de viagens – Árvore A-1	79
Tabela 7.6 -	Grupos 1 e Grupos 2 a 6	81
Tabela 7.7 -	Indivíduos que não trabalham e nem estudam, cônjuge ou chefe – padrões de viagens em relação ao número de pessoas na família	82
Tabela 7.8 -	Grupo 3 e Grupos 4 e 5	83
Tabela 7.9 -	Grupos 8, 9 e 10.....	85
Tabela 7.10 -	Grupos 12 e 13	86
Tabela 7.11 -	Grupo 20	89

LISTA DE GRÁFICOS

Gráfico 1 -	Grau de Instrução entre Homens e Mulheres que realizam padrão HWHSH.....	89
-------------	---	----

RESUMO

ICHIKAWA, S. M. (2002). Aplicação de Minerador de Dados na Obtenção de Relações entre Padrões de Encadeamento de Viagens Codificados e Características Sócio-Econômicas. São Carlos, 2002. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo. (137 p.)

O principal objetivo deste trabalho é analisar a aplicabilidade de um minerador de dados para obter relações entre padrões de viagens encadeadas e características sócio-econômicas de viajantes urbanos. Para representar as viagens encadeadas, as viagens correspondentes a cada indivíduo do banco de dados foram codificadas em termos de seqüência de letras que indicam uma ordem cronológica em que atividades são desenvolvidas. O minerador de dados utilizado neste trabalho é Árvore de Decisão e Classificação, uma ferramenta de análise disponível no software S-Plus. A análise foi baseada na pesquisa origem-destino realizada pelo METRÔ-SP na Região Metropolitana de São Paulo, por meio de entrevistas domiciliares, em 1987. Um dos importantes resultados é que indivíduos que têm atributos sócio-econômicos e de viagem similares não se comportam de maneira similar; pelo contrário, eles fazem diferentes padrões de viagens encadeadas, as quais podem ser descritas em termos de probabilidade ou frequência associada a cada padrão. Portanto, o minerador de dados deve possuir a habilidade para representar essa distribuição. A consistência do resultado foi analisada comparando-os com alguns resultados encontrados na literatura referente a análise de viagem baseada em atividades. A principal conclusão é que Árvore de Decisão e Classificação aplicada a dados individuais, contendo encadeamento de viagem codificado e atributos socioeconômicos e de viagem, permite extrair conhecimento e informações ocultas que ajudam a compreender o comportamento de viagem de viajantes urbanos.

Palavras-chave: análise de demanda; teoria de atividades; encadeamento de viagem; codificação dos padrões de viagens; árvores de decisão.

ABSTRACT

ICHIKAWA, S. M. (2002). Aplicação de Minerador de Dados na Obtenção de Relações entre Padrões de Encadeamento de Viagens Codificados e Características Sócio-Econômicas. São Carlos, 2002. 137p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

The main aim of this work is to analyze the applicability of a data miner for obtaining relationships between trip-chaining patterns and urban trip-makers socioeconomic characteristics. In order to represent the trip-chains, trips corresponding to each individual in the data set were coded in terms of letters indicating a chronological order in which activities are performed. Data miner applied in this work is Decision and Classification Tree, an analysis tool available in S-Plus software package. The analysis was based on the origin-destination home-interview survey carried out by METRÔ-SP in São Paulo Metropolitan Area, in 1987. One of the important findings is that individuals having similar socioeconomic and trip attributes do not behave in a similar way; on the contrary, they make different trip-chaining patterns, which may be described in term of probability or frequency associated to each pattern. Therefore, the data miner should have ability to represent that distribution. The consistency of results was analyzed by comparing them with some results found in literature related to activity-based travel analysis. The main conclusion is that Decision and Classification Tree applied to individual data, containing coded trip-chaining and socioeconomic and trip attributes, allows extracting hidden knowledge and information that help to understand the travel behaviour of urban trip-makers.

Keywords: demand analysis; activity based theory; trip chaining; travel pattern coding; decision trees.

CAPÍTULO 1 – INTRODUÇÃO

Os modelos convencionais utilizados em planejamento de transportes, os chamados de modelos de quatro etapas, foram desenvolvidos para avaliar o impacto dos investimentos de capital nos projetos de infra-estrutura durante um período onde o crescimento na oferta de transportes foi direcionado para o crescimento da população e atividade econômica do período pós-guerra. Para as necessidades da época, os modelos de quatro etapas eram suficientes para estimar o desempenho relativo das alternativas de transportes. A partir da década de 70, mudanças nas políticas de energia, de meio ambiente e meio urbano evidenciaram a necessidade de rever as bases de previsão de viagens e foi neste período que as abordagens baseadas em atividades foram estudadas com maior profundidade. As teorias comportamentais, estruturas conceituais, metodologias e estudos empíricos de comportamento de viagens surgiram durante este período. Sendo assim, as abordagens convencionais, no estudo de comportamento de viagens, são recolocadas em estruturas nas quais são analisadas como padrões diários de comportamento, relacionados e derivados de diferentes estilos de vida e participação de atividades entre a população (JONES et. al., 1990). Esta forma de análise tornou-se conhecida como abordagem baseada em atividades. A principal motivação para a utilização de abordagens baseadas em atividades é que as decisões de viagens são baseadas em atividades. Ela identifica a inabilidade dos modelos de quatro etapas quando tenta refletir o comportamento em relação às políticas envolvidas e à expansão da infra-estrutura de transportes e serviços.

O modelo de quatro etapas ainda é o modelo mais aplicado em planejamento de transportes. No entanto, este modelo tem sido bastante criticado devido à imprecisão relativa aos procedimentos de iteração e a descontinuidade entre as etapas (LEAL, 1983). Diversas pesquisas procuraram corrigir alguns pontos fracos, principalmente aqueles relacionados a

sua eficiência computacional, metodologia de previsão de demanda e principalmente o “feedback” que possibilite, dentre outras coisas, a interação entre uso do solo e transportes.

A análise baseada em atividades constitui “mais uma metodologia exploratória..., em que pesquisadores começaram analisando o comportamento do viajante levando-se em conta a sua condição familiar, num contexto mais amplo de padrão de comportamento diário” (JONES et. al., 1983) e esta abordagem “possibilita ao pesquisador fugir dos limites de uma teoria estruturada e de linhas de pensamento, provavelmente erradas.”

Dessa forma, esta análise permite envolver um complexo tratamento das características sócio-econômicas e de viagens e os limites de localização de um indivíduo pelas suas características, pelo seu domicílio e pelo meio ambiente. Enquanto as técnicas de modelagem desagregadas incorporam características específicas dentro de modelos de escolha, por exemplo, a análise baseada em atividades pode associar uma abordagem que avalia as características sócio-econômicas e a combinação entre o papel do indivíduo dentro da família e o estágio de ciclo de vida. Além disso, as características de viagem também podem ser inseridas e avaliadas no padrão total de viagem num dia completo.

Esta teoria fornece, então, uma estrutura coerente para análise do comportamento de viagem e previsão de demanda. A sua idéia parte do novo conhecimento da demanda de viagem pela compreensão de como e porque as atividades são exercidas num espaço de tempo. Outro ponto importante está no reconhecimento de que as viagens não podem ser analisadas independentemente, pois as atividades realizadas sobre um período de tempo são normalmente inter-relacionadas.

Diversos fatores têm tornado os modelos baseados em atividades ferramentas práticas para a previsão de demanda (KITAMURA et. al., 1995). O grande número de resultados de pesquisas baseados em atividades, os avanços nos métodos de pesquisa e métodos estatísticos, além dos avanços na capacidade computacional e programação, todos estes fatores juntos criam um ambiente onde os modelos de comportamento de viagens podem ser desenvolvidos dentro dos princípios dessa abordagem (KITAMURA, 1996).

A Figura 1.1. descreve a variedade das abordagens como um processo crescente com cinco níveis conceituais e de complexidade de dados, começando com as abordagens convencionais - Figura 1.1.(a).

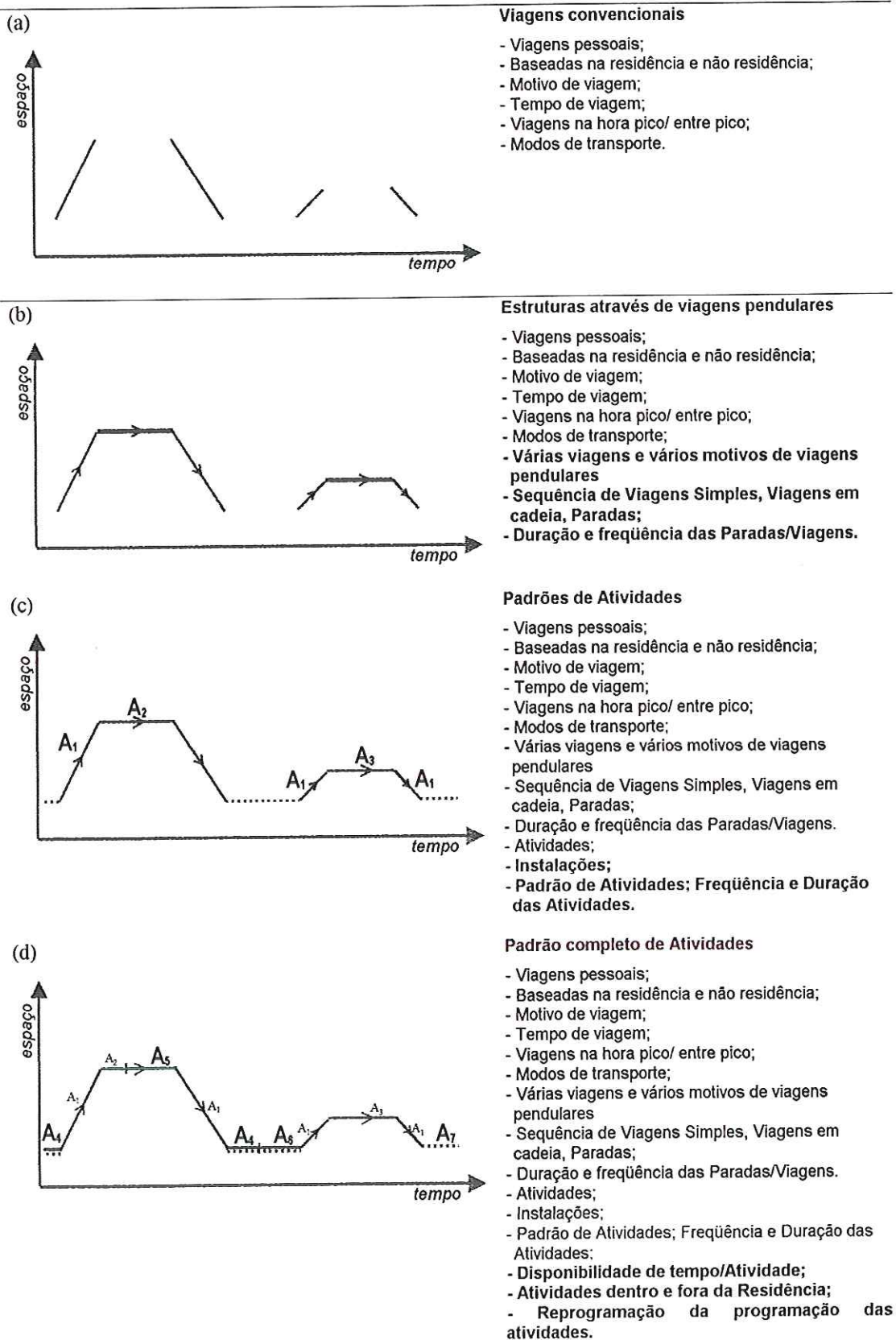


Figura 1.1. Evolução das Abordagens de análise de viagens – Baseado em JONES (1983).

JONES (1983) coloca a representação do comportamento numa estrutura bidimensional de Espaço (distância) e Tempo. As viagens correspondem aos vetores de movimento de um ponto ao outro, usando um dado modo por um propósito de viagem específico. Os períodos de viagem são geralmente analisados nos picos ou nos entre-picos. O primeiro aumento na complexidade é mostrada na Figura 1.1.(b), onde a viagem é analisada dentro de uma estrutura de viagem completa. Viagens casa-trabalho-casa são tratadas como uma entidade, descrita como viagens pendulares ou cadeia de viagem simples. Já nessa análise, a seqüência de viagem torna-se importante, bem como sua composição em relação aos motivos e ao número de viagens. A Figura 1.1.(c) introduz um nível de complexidade para o estudo de comportamento de viagens: a análise de padrões de atividades. A Figura 1.1 (d) representa a análise de um padrão diário completo de atividades e viagens.

Tudo isso implica em uma ampla análise já que o seu objetivo não é limitado pela viagem em si. E, naturalmente, a análise através dos padrões de atividades tende a aumentar os níveis e a complexidade de sua análise. Este é um dos problemas fundamentais na modelagem de viagens baseada em atividades em relação às inúmeras alternativas de programação de atividades. O desafio está na representação adequada do processo de decisão o qual possui muitas possibilidades. Para mostrar a quantidade de combinações possíveis entre as atividades, período do dia, localização, modo e rotas, a Tabela 1.1 apresenta uma estimativa do número de alternativas na realização de atividades por um indivíduo.

Tabela 1.1 - Estimativa do número de alternativas de programação de atividades diárias por indivíduo

Número de Atividades por dia	10	10
Seqüência		10!
Por períodos	10 por atividade	100
Localização ou Destino	1 000 por atividade	10 000
Modo	5 por atividade	50
Rotas	10 por atividade	100
Total		10¹⁷

Fonte: BOWMAN e BEN-AKIVA (1997)

Além disso, o aumento do número de viagens realizadas leva a um acréscimo significativo de combinações possíveis. A Figura 1.2 apresenta exemplos de padrões de viagens, com os algarismos arábicos indicando a seqüência de realização das viagens e as letras

correspondendo a (H) Residência, (W) Trabalho (S) Escola, e (A) outras atividades. Para padrões com cinco viagens, por exemplo, encontram-se 216 combinações.

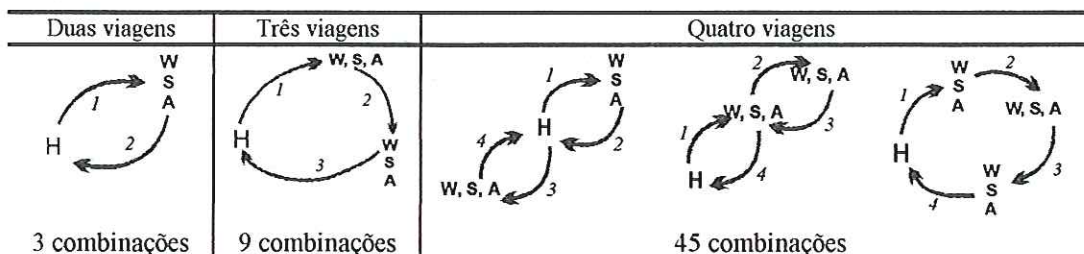


Figura 1.2 – Exemplos de padrões de viagens encontrados para análise de viagens encadeadas

Conseqüentemente, esse aumento de complexidade tem estimulado diversas linhas de pesquisa a aperfeiçoar métodos de avaliação e desenvolver soluções alternativas para estes problemas. Entre as linhas de pesquisa, a modelagem de fenômenos relacionados a transportes foi ocupando espaço juntamente com o desenvolvimento da tecnologia, permitindo o tratamento de problemas mais complexos.

O crescente interesse nos processos da descoberta de conhecimento em bases de dados, ou seja, os mineradores de dados, têm levado ao desenvolvimento de diversas ferramentas para esta análise. Alguns *softwares*, por exemplo, CART (*Classification and Regression Trees*), CHAID (*Chi-square Automatic Interaction Detection*, contido no SPSS), S-PLUS (*Tree-Based Models*) utilizam como forma de análise os modelos de árvore de decisão. Embora estes *softwares* adotem diferentes algoritmos de crescimento de árvore, a idéia básica que fundamenta os algoritmos é semelhante.

O modelo de árvore é um conjunto de regras de classificação que particiona, exaustivamente e sem sobreposição, um conjunto de dados em sub-conjuntos. Considerando um conjunto de dados contendo valores de variável resposta e de variáveis explicativas, estes são analisados para encontrar a melhor regra para classificação ou predição da variável resposta dentro de um conjunto de regras definidas pela divisão hierárquica em sub-conjuntos. Estas regras são definidas de acordo com os valores de um grupo de variáveis explicativas em relação a uma variável de interesse e formar grupos homogêneos (MESA et. al., 2000). Por exemplo, na construção de grupos homogêneos de viajantes de uma população utilizando como ferramenta o modelo de árvore é possível investigar as suas características sócio-econômicas e de viagens, bem como sua estrutura de atributos sócio-econômicos para a realização das viagens.

As vantagens oferecidas pela utilização dos mineradores de dados podem dominar as limitações dos métodos convencionais e permite desenvolver novas opções de políticas que poderão ser importantes no contexto do planejamento em transportes.

Para representar as viagens encadeadas, as viagens correspondentes a cada indivíduo do banco de dados foram codificadas em termos de seqüência de letras que indicam a ordem cronológica em que as atividades são desenvolvidas. Por exemplo, HWH corresponde ao indivíduo que sai de sua residência com destino ao trabalho e volta para casa, ou HSH, aquele indivíduo que sai de sua casa e vai para a escola e retorna para casa. BEN AKIVA e BOWMAN (1996) utilizaram formas semelhantes de codificação, como por exemplo, HWH+, HW+WH. O “+” representa o acréscimo de pelo menos parada adicional para realização de uma atividade. O padrão HW+WH significa duas viagens para o trabalho (início e término da viagem) e a presença de viagens intermediárias baseadas no trabalho. A codificação torna o processamento da árvore mais fácil de ser executado. Além disso, o sequenciamento das atividades realizadas é representado de forma direta.

O padrão encontrado nos nós terminais da árvore corresponde a um conjunto de características sócio-econômicas. Cada nó da árvore apresenta a distribuição de freqüência de ocorrência dos padrões de viagens. Isto significa que indivíduos com atributos sócio-econômicos e de viagens similares não se comportam de maneira similar, pelo contrário, eles fazem diferentes padrões de viagens encadeadas.

1.1 Objetivos

O objetivo fundamental deste trabalho é analisar a aplicabilidade de um minerador de dados para obter relações entre padrões de viagens encadeadas e características sócio-econômicas de viajantes urbanos.

1.2 Motivação

A principal motivação para o estudo da teoria das atividades aplicada na caracterização dos padrões de viagens encadeadas dos indivíduos é que segundo DAMM (1983), as literaturas e evidências empíricas mostram que as decisões de viagem são baseadas em atividades. E, além disso, analisar a possibilidade de relacionar os padrões de viagens às características

sócio-econômicas dos viajantes. A viagem pode ser considerada como um simples mecanismo físico para se atingir um destino com o objetivo de realizar determinado objetivo. A oferta de atividades pode ser caracterizada pela localização no espaço e no tempo e também por atributos inerentes que podem ajudar a limitar sua participação. A demanda por atividades também pode ser expressa dentro do domicílio, não somente pela necessidade de sobrevivência, mas também em função da sociedade (DAMM, 1983).

As políticas voltadas para o controle dos fenômenos como congestionamentos, poluições ou uso do solo, através de programas, tarifas ou regulamentos normalmente não agem pontualmente sobre esses fenômenos, mas afetam diretamente o comportamento dos indivíduos que fazem parte deste contexto. Conseqüentemente, esses indivíduos tentam ajustar seus comportamentos, sempre motivados pelo desejo de alcançar seus objetivos de realizar alguma atividade. Através de um minerador de dados é possível que este comportamento seja observado através dos padrões de viagens encadeadas codificadas.

Embora os modelos de previsão de demanda convencionais encontrados e utilizados hoje em dia possuam boa qualidade e sofisticação, estes modelos ainda deixam a desejar quando se tenta inserir os conceitos da estrutura de decisão e a teoria de viagens baseada em atividades.

Novas metodologias, paradigmas e tecnologias para auxiliar o desenvolvimento do planejamento de transportes são alvos de pesquisa constante por parte dos planejadores e pesquisadores em transportes. Esses novos métodos contribuem na melhoria dos sistemas voltados para análise de previsão de demanda.

1.3 Método

Para atingir os objetivos deste trabalho, os seguintes procedimentos foram adotados:

- i. Revisão bibliográfica relativa à teoria de viagens baseada em atividades e assuntos relacionados ao comportamento dos indivíduos e as características sócio-econômicas;
- ii. Revisão bibliográfica sobre os mineradores de dados, ferramenta utilizada neste estudo;
- iii. Com base no padrão de viagem, avaliar formas de codificação dos dados para análise;

- iv. Verificar a possibilidade de utilizar as árvores de decisão para obter relações entre padrões de viagens e as condições sócio-econômicas;
- v. Analisar as relações resultantes sob o ponto de vista do comportamento dos viajantes;
- vi. Comparar os resultados obtidos com os encontrados na literatura referente à teoria de viagens baseada em atividades.

A pesquisa de literatura sobre o assunto foi baseada em tipos diferentes de procura, como citações, textos e artigos resumidos e completos. Seu foco foi direcionado aos métodos e enfoques observados nas teorias sobre as viagens baseadas em atividade, além da análise do comportamento humano referente à realização de viagens, tipos e padrões de viagens. A tarefa principal na procura e revisão de literatura foi identificar e obter uma seleção manejável de artigos relacionados e úteis que fornecessem uma orientação para a análise e que permitissem comparar os métodos encontrados de acordo com os seus fundamentos.

Uma das dificuldades que surgiu de imediato foi como chegar a tais objetivos: de que forma representar os padrões de viagens e como relacioná-los com as características sócio-econômicas dos indivíduos.

As árvores de decisão foram adotadas para análise em função da sua capacidade de representar a natureza probabilística do objeto analisado, que no caso representa os padrões de viagens. Além disso, através da árvore de decisão é possível analisar um grande conjunto de dados e através das regras hierárquicas e da sua divisão em grupos é possível organizar os dados de maneira compacta e obter uma visão acerca da relação entre padrão de viagem e variáveis sócio-econômicas. Para isso, foi necessário um estudo detalhado da teoria geral dos mineradores de dados, das árvores de decisão e análise do funcionamento do algoritmo da árvore gerada pelo software S-Plus 2000 utilizados neste trabalho. A partir daí, foi possível compreender a forma de seus processamentos, podendo verificar alguns de seus recursos e limitações.

A codificação dos dados foi adotada como forma de viabilizar a representação de padrões de viagens e de relacioná-los a grupos de indivíduos com características semelhantes. A partir dos resultados contendo relações entre indicadores sócio-econômicos e padrões de viagens codificadas, efetuaram-se avaliações sobre os conhecimentos obtidos durante a aplicação do

método de classificação, verificando se alguns conhecimentos disponíveis de antemão tinham sido descobertos pelo minerador.

Os dados utilizados neste trabalho são os da Pesquisa Origem-Destino da Região Metropolitana de São Paulo coletadas pela Companhia do Metropolitano de São Paulo em 1987 e cedidos ao Departamento de Transportes da EESC-USP para o desenvolvimento de pesquisas.

A verificação da consistência dos resultados obtidos através da codificação e do uso do minerador de dados foi realizada com a confrontação de alguns resultados constantes na literatura encontrada referente à teoria de viagens baseada em atividades.

Neste trabalho pretende-se identificar e caracterizar a população da Região Metropolitana de São Paulo a partir desses indicadores sócio-econômicos e relacioná-los a alguns aspectos dos padrões e sequenciamento de viagens. Acredita-se que um diagnóstico analítico e descritivo dos grupos de indivíduos formados e das suas características poderá revelar uma etapa fundamental para aportar conhecimentos que possam ser úteis ao planejamento de transportes.

1.4 Contribuições

A presente pesquisa contribui para a teoria de viagens baseada em atividades com a proposta de uma abordagem mais simplificada de avaliação dos padrões de viagens realizadas pelos indivíduos. A abordagem desenvolvida baseia-se na utilização do modelo de árvore de decisão aplicada às seqüências de viagens individuais codificadas.

Outra contribuição importante deste trabalho é a proposta de uma metodologia para a codificação das viagens necessária para a análise por meio de árvore de decisão. As árvores de decisão permitem extrair conhecimento e informações ocultas que ajudam a compreender o comportamento de viajantes urbanos.

Embora neste trabalho as viagens tenham sido caracterizadas em termos de viagens inter-atividades, outras características, tais como modo, destino e período podem ser acrescentadas.

As abordagens baseadas em atividades tornam-se importantes para o aperfeiçoamento do estado da prática na medida em que mostra que a modelagem baseada em atividades, ainda com aplicações bastante restritas, tem grande potencial no campo da previsão de viagens.

1.5 Apresentação dos Capítulos

Esta dissertação está estruturada em oito capítulos além deste capítulo de introdução.

O capítulo 2 apresenta as idéias básicas da teoria das atividades com a estrutura na qual as decisões de viagem e atividades são realizadas. Em seguida, apresenta as características da demanda de atividades e de viagens. Alguns conceitos importantes encontrados na literatura são apresentados para se familiarizar com a linguagem adotada nas abordagens baseadas em atividades.

O capítulo 3 descreve algumas pesquisas e resumos dos resultados encontrados na literatura.

O capítulo 4 descreve uma introdução aos mineradores de dados, principalmente à teoria das árvores de decisão utilizada nesta análise de dados. As árvores de decisão, ferramenta básica para a análise dos dados deste trabalho, fornecem uma representação simples do conhecimento, extraindo informações desconhecidas a partir de grande base de dados. A ênfase será dada ao modelo de árvores, o qual está disponível no *software* S-plus 2000 desenvolvido pela StatSoft (S-PLUS, 1999 a, b). Segue no capítulo algumas considerações a respeito do modelo de árvore do S-Plus 2000.

O capítulo 5 apresenta um estudo de caso utilizando os dados da Pesquisa Origem-Destino da Região Metropolitana de São Paulo, coletados pela Companhia do Metropolitano de São Paulo em 1987, e cedidos ao Departamento de Transportes da EESC-USP para o desenvolvimento de pesquisas.

O capítulo 6 descreve o tratamento dos dados e a forma de codificação dos mesmos para aplicação e geração da árvore de decisão.

O capítulo 7 apresenta os resultados obtidos através da aplicação da árvore de decisão com a utilização das variáveis sócio-econômicas disponíveis no banco de dados. Este capítulo

mostra também algumas comparações com resultados encontrados na literatura referente à análise de viagem baseada em atividades.

O capítulo 8 apresenta as principais conclusões encontradas, recomendações e sugestões para trabalhos futuros que possam dar continuidade aos desenvolvimentos obtidos nesta dissertação.

Os Anexos e as Referências Bibliográficas citadas nesta pesquisa são apresentados no final deste trabalho.

CAPÍTULO 2 – ABORDAGENS DE VIAGENS BASEADAS EM ATIVIDADES

O ponto de partida da abordagem de viagens baseada em atividades foi o enfoque apresentado por MITCHELL e RAPKIN (1954).

“Parece evidente que alguns indivíduos deslocam-se pela cidade em função de atividades altamente especializadas: médicos atendem os chamados de seus clientes, os vendedores atendem os fregueses, os homens de negócios atendem os banqueiros, as mulheres fazem as compras da casa (...) pode-se imaginar que os papéis conflitantes dos indivíduos podem tornar-se evidentes também nos seus padrões de viagem e podem ter grande influência na sua determinação. Certamente, no projeto de uma pesquisa sobre os motivos dos deslocamentos, a influência do status do indivíduo e dos papéis simples ou múltiplos por ele desempenhados deveriam ser considerados.”

Esta afirmação pode ser considerada como a primeira tentativa de relacionar os estudos de tráfego aos papéis desempenhados pelos indivíduos (VASCONCELLOS, 2000). O estudo dos autores apresentou diversas investigações e uma estrutura compreensiva do comportamento de viagens, mas ainda não estabelecia uma relação direta entre viagens e atividades. Neste período, infelizmente, a visão simplista de uma política que dominava a economia do pós-guerra, levou a formação de modelos de transporte que focavam somente em viagens (quem, o que, onde, e quais motivos das viagens), e a relação entre atividades e viagens ficaram refletidas somente na geração de viagens (MCNALLY, 2000).

A abordagem de atividades começou como uma evolução natural de pesquisa do comportamento humano e, em particular, pesquisas sobre o comportamento de viagens. Críticas dirigidas ao desempenho do modelo de 4 etapas não serviam como um estímulo para

o desenvolvimento das pesquisas sobre viagens baseadas em atividades, até que algumas incoerências desses modelos fossem observadas, e novas políticas surgissem. Particularmente, essas críticas deram lugar a um aprimoramento dos modelos, primeiramente através da introdução de modelos desagregados e modelos de equilíbrio. Nenhuma mudança foi capaz de diminuir o potencial de novas contribuições, já que os modelos desagregados são elementos que compõem as abordagens baseadas em atividades.

O princípio fundamental das abordagens baseadas em atividades é que as decisões de viagens são dirigidas por um conjunto de atividades que formam uma “agenda” de participação nas atividades e que não podem ser analisadas com base nas viagens individuais. Dessa forma, o processo de escolha associado à decisão de viagem pode ser compreendido e modelado somente dentro do contexto de uma programação. O conjunto de atividades e viagens executadas inclui o padrão de atividades do indivíduo, o processo de decisão, as regras de comportamento e o ambiente. Todos estes elementos levam a formação de padrões que caracterizam um comportamento complexo de viagem.

A viagem é um mecanismo de acesso a uma ou várias atividades. Enquanto as abordagens convencionais se satisfazem com os modelos que geram viagens, as abordagens baseadas em atividades focam no que gera a atividade produtora de viagens.

2.1 Idéias Básicas

Os mais importantes elementos da teoria do comportamento do viajante baseada nas atividades podem ser resumidos em quatro idéias básicas.

- a) A demanda por viagem é derivada da demanda por atividades (JONES, 1977).
- b) O comportamento humano está limitado no tempo e no espaço (HAGERSTRAND, 1970). Dentro de limites de tempo e espaço, os homens se movimentam de formas variadas, em diferentes lugares e em diferentes pontos no tempo, sobretudo pela experiência que adquirem sobre o tempo e o custo desse movimento. Considera-se ainda que eles também são limitados pela necessidade de retornar a sua residência para descanso e manutenção pessoais (BOWMAN, 1995).
- c) O domicílio afeta as atividades do indivíduo e a sua decisão de viagem (CHAPIN, 1974; JONES et al, 1983). Normalmente, os indivíduos agem dentro do contexto familiar,

dividindo tarefas e compartilhando recursos com outros membros do domicílio. Muitas das decisões tomadas no domicílio podem ser consideradas como uma unidade. Por exemplo, a participação do filho em um curso de inglês é decidida pelo chefe de família. A decisão de um indivíduo também pode ser influenciada pelos outros membros. A composição do domicílio e a idade dos membros também afetam as decisões individuais e domiciliares (BOWMAN, 1995).

d) A idéia básica é que as decisões acerca das viagens e atividades são dinâmicas (GOODWIN et al., 1990).

2.2. Estrutura de Decisão de Viagens Urbanas

A decisão de viajar está baseada nas atividades em que o indivíduo se dispõe ou necessita realizar.

Na Figura 2.1 observa-se toda a estrutura de decisão para viagens urbanas, com a descrição das importantes categorias de decisão e suas interações (BEN-AKIVA e LERMAN, 1985 e BEN-AKIVA et al., 1995). A estrutura apresentada foca o domicílio e a decisão do indivíduo. Estes elementos levam à demanda por viagem, incluindo a mobilidade, estilo de vida e o horário das atividades e das viagens. Dentro desta estrutura, inclui também o processo de desenvolvimento urbano o qual afeta as decisões dos indivíduos e a interação de todas as decisões com o desempenho do sistema de transporte.



Figura 2.1 – Estrutura de Decisão de Viagens Urbanas – baseado em BOWMAN (1995)

BOWMAN e BEN-AKIVA (1997) descrevem as escolhas do domicílio e do indivíduo numa estrutura de decisão dentro de uma escala de tempo. Eles dividem a estrutura de decisão em três categorias de escolha: (1) mobilidade e estilo de vida, (2) horário das atividades e viagens, e (3) reprogramação dos horários. A decisão sobre a mobilidade e o estilo de vida, tais como a localização da residência e do emprego e a decisão de possuir automóvel ocorrem em intervalos irregulares e não freqüentes, dentro de uma escala de tempo em anos. De acordo com os autores acima, a definição dos horários das atividades e das viagens corresponde a um planejamento na qual as decisões ocorrem com maior freqüência e em intervalos regulares tais como dias ou semanas. O planejamento envolve vários elementos como: um conjunto de atividades; a alocação de atividades de cada membro do domicílio; a seqüência de atividades a ser adotada e a localização das atividades, horários e modos de viagem. A reprogramação ocorre num curto espaço de tempo, no período do dia. Esta reprogramação acontece em função da realização de atividades não programadas, em resposta a eventos não esperados. Isto pode refletir na rota escolhida, velocidade, aceleração, distâncias e estacionamentos ou paradas. A reprogramação pode ser vista como um ajuste de decisão baseada numa revisão de objetivos e limites em virtude de informações recentes.

Segundo BOWMAN e BEN-AKIVA (1997), as decisões de desenvolvimento urbano correspondem à decisão dos governantes, planejadores e empresas. O governo pode fornecer serviços de transporte público, fixar tarifas, o que de certa forma influencia o comportamento dos indivíduos e empresas. Os planejadores fornecem oportunidades de localização para empresas e de localização dos indivíduos. E as empresas determinam os locais de oportunidade de empregos.

Pelas observações de BOWMAN e BEN-AKIVA (1997), o desenvolvimento urbano influencia diretamente as decisões dos indivíduos e domicílios. E ambos, o desenvolvimento urbano e as decisões individuais, afetam o desempenho do sistema de transporte. Isto se manifesta de diversas maneiras, incluindo volumes de tráfego, velocidades, impactos ambientais e congestionamentos. Novamente, estas manifestações afetam simultaneamente o desenvolvimento urbano e as decisões dos indivíduos.

Em resumo, BOWMAN (1998) descreve quatro características importantes da estrutura de decisão. A primeira característica é que a decisão de programar é condicionada por decisões em longo prazo, isto é, pelo estilo de vida escolhido pelo domicílio, como a localização da residência e do emprego, ou pela decisão de comprar automóvel para os deslocamentos. A

segunda característica, em relação à primeira, considera-se que o processo de programar não é seqüencial, mas governado por compromissos e prioridades assumidos, considerando os limites de um dado período de tempo. A terceira característica é que durante o período de programação dentro de um dia deve-se considerar períodos necessários para descanso. E a quarta e última característica corresponde ao processo de programação que interage com o desempenho do sistema de transporte: as escolhas ou decisões dos indivíduos determinam o desempenho do sistema, e essas decisões são então influenciadas pela percepção do desempenho desse sistema.

Os indivíduos ajustam de diversas maneiras seus comportamentos em função do desenvolvimento urbano, sempre motivado pelo desejo de realizar suas atividades. Esta idéia é ilustrada na Figura 2.2.

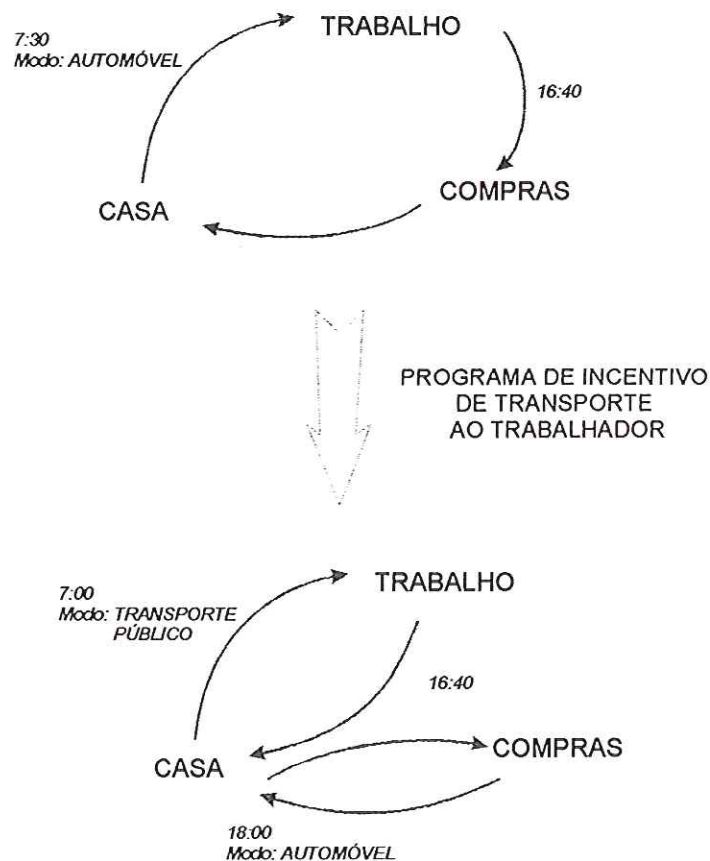


Figura 2.2 – Resposta à política adotada envolvendo o ajuste comportamental motivada pelo desejo de alcançar os objetivos de realizar atividades Baseado em BOWMAN e BEN-AKIVA (1997)

A figura acima representa um padrão de atividade e de viagem diária de um indivíduo que utiliza o seu automóvel para ir ao trabalho. Ele sai da residência às 7:30 da manhã e sai do trabalho às 16:40. Antes de retornar à residência ele pára para fazer compras e em seguida chega à sua residência.

Em resposta ao programa de incentivo para utilização de transporte público para o trabalho, este indivíduo pode fazer a troca por este modo. Isto irá fazer com que ele saia mais cedo de casa (7:00) para chegar ao trabalho. Como, no exemplo apresentado, a rota do transporte público não atende ao local de compras, ele decide retornar à residência depois do trabalho, e somente depois vai às compras.

Esta resposta foi fundamentada na demanda pelas atividades e envolveu um ajuste complexo no padrão diário. Dessa forma, os modelos de previsão serão capazes de obter esse tipo de resposta se esses modelos conseguirem representar a forma como os indivíduos programam suas atividades diárias.

2.3 Características da demanda de atividades e viagens

Uma das características fundamentais da demanda por viagens e um dos princípios mais aceitos é que a demanda por viagem é derivada da demanda por atividades. Este princípio mostra que dentro da estrutura de decisão, a decisão de viajar é um componente da decisão por atividades. CHAPIN (1974) desenvolveu a teoria de que a demanda por atividades é motivada pelo desejo do homem pela sobrevivência ou pela necessidade de encontros sociais ou satisfação pessoal.

Infelizmente, é difícil de modelar fatores baseados na demanda por atividades e pouco progresso tem sido feito para incorporá-los nos modelos de demanda por viagem. Entretanto, uma quantidade significativa de pesquisas está sendo realizada na análise de como os membros do domicílio decidem e programam suas atividades. Observou-se que (1) os domicílios influenciam as decisões sobre atividades, (2) os efeitos diferem por tipo de domicílio, tamanho, relacionamento entre os membros, idade e gênero, e (3) as crianças, em particular, impõem de forma significativa a realização de mais viagens (como por exemplo, escola, médico, atividades de lazer, compras) e ao mesmo tempo limitam os outros membros do domicílio.

As viagens não podem ser vistas como independentes, mas relacionadas às atividades realizadas pelos indivíduos de um domicílio. A figura 2.3 exemplifica o complexo relacionamento entre viagens e atividades (BRÖG e ERL, 1983).

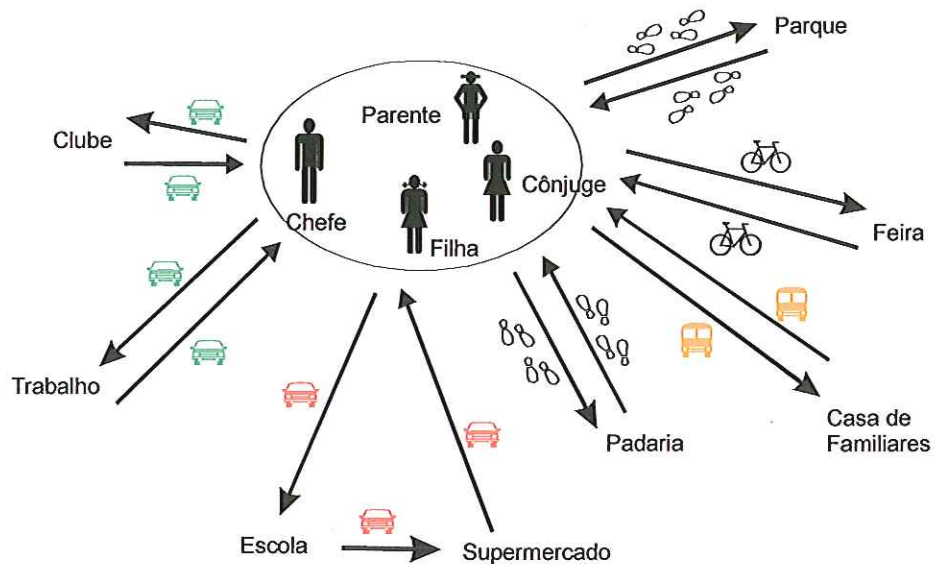


Figura 2.3 – Relacionamento entre viagens e atividades por indivíduos de um domicílio

Esta figura exemplifica os padrões de atividades dos indivíduos de um domicílio. O chefe da família vai para o trabalho de carro. À tarde, ele sai novamente de carro para praticar algum esporte. A esposa realiza três atividades: à padaria (a pé) e à feira (de bicicleta) e visita um parente utilizando o ônibus como meio de transporte. A filha utiliza seu próprio carro para ir à universidade e antes de retornar a casa ela passa no supermercado. O parente faz uma pequena caminhada no parque.

Segundo a figura apresentada por BRÖG e ERL (1983), o domicílio fez no total 8 atividades pelos diferentes membros do domicílio como mostra a Tabela 2.1.

Tabela 2.1. - Atividades e viagens dos indivíduos do domicílio

Membro do Domicílio	Tipo de Atividade			
	Trabalho	Educação	Compras	Lazer/Visitas
Chefe	A	-	-	D ₁
Esposa	-	-	C ₁ , C ₂	D ₂
Filha	-	B	C ₃	-
Avô	-	-	-	D ₃

Baseado em BRÖG e ERL (1983)

Legenda: A – Trabalho; B – Educação; C₁ – Padaria; C₂ – Feira; C₃ – Supermercado; D₁ – Clube; D₂ – Casa de Familiares; D₃ – Parque

A Tabela 2.2 mostra os padrões de atividades dos membros do domicílio. A alteração na programação de atividades pode modificar estes padrões.

Tabela 2.2. – Padrões de atividades e de viagens dos indivíduos

Membro do Domicílio	Tipo de Padrão de Atividade
Chefe	H – A – H – D ₁ – H
Esposa	H – C ₁ – H – C ₂ – H – D ₂ – H
Filha	H – B – C ₃ – H
Avô	H – D ₃ – H

Baseado em BRÖG e ERL (1983)

Legenda: H – Domicílio

Em relação às viagens, a Tabela 2.3 mostra o número de vezes em que o indivíduo saiu da sua residência para alcançar seus objetivos e o número de viagens utilizando os diversos modos de transporte.

Tabela 2.3. – Viagens e modos de transporte utilizados pelos indivíduos

Membro do Domicílio	Número de vezes que sai da residência	Número de viagens			
		Total	Modos individuais	Modos públicos	Modos não motorizados
Chefe	2	4	4	-	-
Esposa	3	6	-	2	4
Filha	1	3	3	-	-
Avô	1	2	-	-	2
Total	7	15	7	2	6

Baseado em BRÖG e ERL (1983)

Segundo BRÖG e ERL (1983), as viagens podem ser alteradas quando políticas de transporte são instituídas. Conseqüentemente, os indivíduos reorganizam as atividades, alteram destinos, ou substituem os membros ou os horários para realização de atividades.

HAGERSTRAND (1970) dirigiu atenção às restrições para realização de atividades. Essas restrições correspondem à dependência, autoridade e capacidade. A restrição de dependência é uma limitação imposta pela natureza ou pela tecnologia. Um exemplo é a necessidade do ser humano em retornar para casa para descanso. As restrições de autoridade são limitações impostas dentro de uma sociedade ou regras adotadas em nível institucional, tais como horários do trabalho ou do comércio, e regulamentos como restrições ao barulho, por exemplo. A restrição de capacidade significa que o indivíduo necessita da presença de outra pessoa ou algum outro recurso para conseguir realizar uma atividade. Como por exemplo, a necessidade do automóvel para realizar uma atividade ou para participar das atividades

comuns do domicílio. Outro exemplo, apresentado por Hagerstrand, é chamado de prisma espaço-tempo, onde as pessoas se deslocam de um ponto a outro baseado na experiência do tempo e no custo que a pessoa percebe desse deslocamento.

Entretanto, nem todas as atividades necessitam de um deslocamento físico. Com o avanço da tecnologia, é possível participar de vários tipos de atividades sem se deslocar fisicamente, além de transferir informações eletronicamente em quantidade e qualidade. Isto leva o indivíduo a realizar atividades como trabalho, compras ou recreação, podendo escolher entre viajar ou não.

Em resumo, as abordagens baseadas em atividades surgiram do desejo dos pesquisadores em modelar o comportamento humano para compreender a natureza da participação em atividades. Isto reforça a idéia de que a viagem é derivada do desejo de participar em atividades dispersas no tempo e no espaço, chamadas de padrões de comportamento diário (HAGERSTRAND, 1970).

2.4 Definições

As atenções dirigidas ao comportamento baseado em atividades ainda são recentes, e como ainda não possui uma teoria universal única adotada por todos os pesquisadores, o assunto está sujeito a uma variedade de abordagens e conceitos desenvolvidos independentemente. Neste contexto de abordagens diferentes, é útil fornecer definições e uma breve explicação dos termos utilizados por diversos pesquisadores.

O conjunto de atividades diárias que as pessoas desejam realizar é referido como *programa de atividades* ou “*activity time budget*”, isto é, orçamento do tempo (tempo de atividade e tempo de viagem). Dado um programa de atividade, para a sua realização dentro de um período de tempo, o indivíduo necessita decidir quando, onde e por quanto tempo as atividades do programa deverão ser desenvolvidas (DAMM, 1983). Estas decisões podem não ser necessariamente idênticas, nem realizadas simultaneamente com a decisão de viagem.

Segundo SCHEUCH (1972), a palavra “*Atividade*” é um tanto difícil de definir considerando que ela pode ser considerada multidimensional, isto é, uma dada atividade pode ser vista como diferentes naturezas. Por exemplo, o termo compras, pode ser considerado como

àquela atividade realizada em função de um trabalho, como também pode ser interpretada como uma atividade de lazer.

Na teoria da atividade humana, a unidade básica de análise é o trabalho do homem (BODKER, 1993). A atividade humana é exercida com o objetivo de alcançar certos propósitos. A atividade é geralmente realizada com a ajuda de um ou mais instrumentos ou ferramentas, e é direcionado para um certo objeto. As ferramentas, bem como as divisões de trabalho, normas e linguagens, podem ser vistas como um *artefato* para a atividade: eles são criados pelos homens e medem as relações entre os seres humanos ou indivíduos e o objeto de sua atividade, o material ou o produto em diferentes estágios. Estes artefatos são introduzidos dentro de uma determinada atividade, como também são produtos de nossas atividades e são constantemente modificadas pela própria atividade (BODKER, 1993).

A *demanda por atividades e viagens* é vista como uma escolha entre todas as possíveis combinações de atividades e viagens dentro do período de um dia. (BOWMAN e BEN-AKIVA, 1997). A *viagem* pode ser considerada um simples mecanismo físico para se atingir um destino com o objetivo de realizar determinado objetivo.

Um “*trip*” corresponde a um segmento de viagem, isto é, a uma viagem com origem em um determinado ponto e destino em um outro ponto.

Um “*Journey*” simples envolve uma viagem individual da residência a um dado destino e o seu retorno à residência. *Jornadas complexas* correspondem a uma seqüência de mais de duas viagens individuais que iniciam e terminam na residência.

Um “*tour*” é definido como uma viagem que tem início em um determinado local para um ou mais lugares de atividades e finalmente o seu retorno ao ponto inicial, podendo ser a residência, o local do trabalho ou o local de estudo, por exemplo.

BOWMAN et al. (1998) definem o “*tour*” como uma seqüência de vários segmentos de viagens (“*trips*”) e o ponto inicial da primeira viagem corresponde ao ponto final da última viagem. Cada “*tour*” pode ter um número de paradas, classificadas segundo três motivos: subsistência (trabalho ou estudo), apoio ou sustento, e lazer. Cada “*tour*” possui um destino primário. O trabalho é considerado o destino primário para os “*tours*” com uma parada no

trabalho. Para os outros “tours”, um conjunto de regras baseado na combinação de motivos e duração das atividades determina o destino primário.

A Figura 2.4, proposta por BOWMAN et al. (1998), define os “tours” e “trips”.

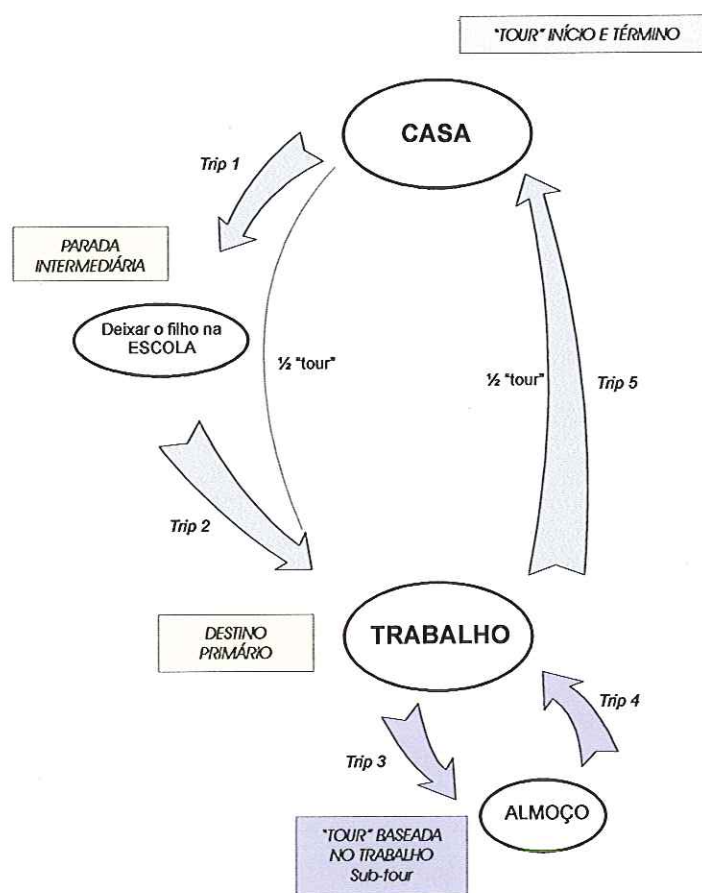


Figura 2.4 – Conceito de “tours” baseado em BOWMAN et. al. (1998)

Segundo a Figura 2.4, a parte do “tour” que corresponde a viagem realizada da residência ao destino primário é considerada $\frac{1}{2}$ “tour”, e a parte do “tour” que corresponde a viagem do destino primário à residência é a outra parte do “tour”. Todas as paradas seguintes diferentes ao destino primário são chamadas de paradas intermediárias. Uma ou mais atividades podem tomar lugar de cada local de parada e sua duração geralmente é sempre menor que a duração no destino primário.

Um “tour” pode ser baseado no trabalho e é definido como uma seqüência de segmentos de viagens (“trips”) que iniciam no trabalho e terminam no trabalho. No exemplo, o “tour” baseado no trabalho é considerado um sub-“tour”. Um exemplo de “tour” baseado no trabalho é o indivíduo saindo do trabalho para almoçar e voltando novamente para o trabalho.

Dentro do amplo conceito de “tour”, alguns autores analisam somente as pessoas com viagens a trabalho, que podem ser consideradas como base para organizar as atividades durante o resto do dia. Se se considerasse o “tour” casa-trabalho-casa como fixo pode-se avaliar em que condições as pessoas adicionam viagens ou atividades nesses “tours”.

Dessa forma, é possível começar a compreender as diferenças entre as pessoas que simplesmente vão ao trabalho e aquelas que, além do trabalho, participam de atividades que poderiam teoricamente ser desempenhadas em outro lugar, em outra hora ou com diferente duração (DAMM, 1983).

O termo *padrão de viagem* é usado para se referir a um conjunto de viagens realizadas pelo indivíduo ou pelo domicílio dentro de um período fixo de tempo (geralmente 24 horas) (ADLER e BEN-AKIVA, 1979). Pode-se considerar também o comprimento da viagem e a freqüência com que são empregados (HANSON, 1979). Alguns pesquisadores definem o *padrão* como o conjunto de coisas que as pessoas fazem no tempo e no espaço (CHAPIN, 1974).

Como as atividades geralmente não são realizadas num mesmo local e nem ao mesmo tempo, a seqüência de deslocamentos no espaço urbano ao longo do dia é necessária. A união de vários segmentos de viagens ou deslocamentos formam as *viagens em cadeia ou viagens encadeadas*. Alguns denominam de cadeia ou seqüências para expressar a mesma idéia. Quando o trecho da viagem é composto de um único destino que não seja a residência a viagem em cadeia é chamada de simples, e quando envolve também vários destinos que não sejam residência, é chamada de complexa.

Além disso, a estrutura de viagens em cadeia permite obter classificações utilizando uma separação espacial do local da residência e do local do trabalho, *como viagens em cadeia simples ao trabalho*.

CAPÍTULO 3 – A TEORIA DAS VIAGENS BASEADAS EM ATIVIDADES E SUAS HIPÓTESES

Este capítulo descreve algumas pesquisas e resumos dos resultados encontrados na literatura.

As idéias básicas da teoria das viagens baseadas em atividades foram tratadas em muitas pesquisas teóricas e descrições empíricas voltadas ao relacionamento entre as atividades humanas e o comportamento de viagem. Muitas hipóteses têm sido testadas e avaliadas, cada qual com sua estrutura e características próprias.

A recente atenção nos estudos de comportamento de viagens baseadas em atividades e a ausência de uma teoria universal permitiram que houvesse o desenvolvimento de várias abordagens e conceitos. São muitos os trabalhos de pesquisa, como foi apresentado por DAMM (1983) através de um resumo de pesquisas empíricas. GOLOB E GOLOB (1983) pesquisaram na literatura 361 trabalhos focando-os através das atividades, atitudes, segmentos sócio-econômicos, experimentos e escolhas. KITAMURA (1988) revisou os trabalhos apresentados por Golob e Golob, categorizando-os em elementos como a participação na atividade, programação, restrições, estrutura de decisão de viagens, estrutura do domicílio, aspectos dinâmicos, políticas, modelos de atividades e desenvolvimentos metodológicos. Trabalhos recentes como de ETTEMA e TIMMERMANS (1997), ARENTZE e TIMMERMANS (2000) e MACNALLY (2000) também mostraram alguns comparativos entre as pesquisas desenvolvidas baseadas nesses conceitos.

DAMM (1983) forneceu uma compilação e revisão de estudos empíricos relacionados às atividades, categorizando grupos de hipóteses que diferem nos escopos, objetivos, métodos de análise, além de listar as variáveis explicativas associadas a cada classe de hipótese.

DAMM (1983) destacou a formação de três grupos distintos de hipóteses que constituem o comportamento dos indivíduos na realização de viagens e atividades: (1) características individuais, (2) características domiciliares e (3) limites ambientais.

Estes três grupos são descritos abaixo através de alguns trabalhos realizados:

a) *Características individuais*: autores como HANSON e HANSON (1980) desenvolveram hipóteses sobre como as variações de comportamento poderiam ser explicadas, questionando se os padrões de atividades entre homens e mulheres eram diferentes. Um exemplo dos resultados encontrados foi que mulheres trabalhadoras fizeram mais paradas por viagens, realizaram distâncias pequenas de viagem, além de muitas paradas para compras. Esta análise foi baseada nas idéias de CHAPIN (1974) que não considerou homens e mulheres isolados do contexto social, mas operando dentro de estruturas impostas por forças externas, como as diferenças culturais, por exemplo. HANSON (1978) postulou que a idade determinava significativamente o padrão de atividade, principalmente quando comparou o comportamento dos mais idosos com os demais grupos etários. Segundo DAMM (1983), em relação às características individuais, o gênero e status de trabalho (isto é, se empregado ou não) foram elementos que também determinavam as variações no comportamento de atividades.

b) *Características domiciliares*: muitas hipóteses foram encontradas por DAMM (1983) no contexto familiar. Muitos pesquisadores empregaram o domicílio como referência principal de análise, embora em alguns casos, o local de trabalho ou da escola, tenham sido considerados. CHAPIN (1974) postulou o “estágio do ciclo de vida” como altamente crucial e influenciou outros pesquisadores que vieram a incluir uma variável similar em seus trabalhos (HAUTZINGER e KESSE, 1977; HOLZAPFEL, 1980).

c) *Limites Ambientais*: Para muitos autores, o comportamento dos indivíduos na realização de viagens seria limitado pelo tipo, quantidade, período, duração e localização das atividades. Como todos estes elementos tendem a agir simultaneamente, a análise se tornaria mais complexa. Fontes ricas de hipóteses sobre estes limites foram levantadas. LENTORP (1976) questionou sobre quais seriam os caminhos que poderiam ser seguidos para realizar um conjunto de atividades quando se encontram restrições temporais e espaciais que limitam um conjunto de alternativas possíveis. HAGERSTRAND (1970) também definiu alguns tipos de limites, conforme apresentados no capítulo anterior. A localização das atividades também foi considerada como uma restrição para completar as necessidades de realização

das atividades. O tipo de área residencial foi uma das restrições colocadas por GODARD (1980) e BONNAFOUS et. al. (1981) como um elemento determinante na variação dos padrões de viagens.

Alguns dos recursos e os métodos empregados por pesquisadores na análise de hipóteses sobre o comportamento de atividades descritas por DAMM (1983) são apresentadas na Tabela 3.1.

Tabela 3.1 – Recursos e métodos utilizados nas análises de hipóteses sobre viagens baseadas em atividades

Pesquisador(es)	Fonte e Tipo de Dados	Ano	Amostra	Análise
Adler	Washington, D.C./U.S.A.	1968	1003 domicílios	logit; utilidade
Ben-Akiva	Home Interview Survey (HIS)			Contagem por padrões
Bentley, Bruce Jones	Watford, England Viagens diárias por semana	1969	1972 indivíduos	Qui-quadrado
Bonnafous et. al.	Lyon, France HIS	1978 1979	283 indivíduos 171 indivíduos	Comparação de tabelas
Chapin	Washington, D.C./U.S.A. Pesquisa de horas de lazer Por observação	1969 1971	424 domicílios 708 domicílios	ANOVA
Damm	Minneapolis-St. Paul. Minesota/U.S.A., HIS.	1970	2345 trabalhadores	Regressão/Duração Probit/Participação Modelos condicionais
Godard	Dijon, França. HIS	1977	1000 indivíduos	Classificação cruzada
Grahan Ogden	Canberra HIS Cobrug HIS Austrália	1975	4% da amostra domiciliar (trabalhadores viajam no pico da manhã)	Método estatístico não linear
Hanson, S.	Uppsala, Sweden 5 semanas diárias com todos os movimentos	1971	300 domicílios	Cluster Análise fatorial
Hautzinger, Kessel	KONTIV/W. Germany	1975	2600 pessoas selecionadas aleatoriamente	MANOVA; análise discrim. generalizada Medida de distâncias
Hemmens	Buffalo, Nova York/ USA HIS	1962	5500 membros de 1600 domicílios	Qui-quadrado; ANOVA; Modelo Markoviano
Holzapfel	KONTIV/W. Berlin	1977	20000 domicílios	Análise de fator e Cluster
Jones, Dix, Clarke, Heggie	HATS Abingdon, Inglaterra	1978 1978 1978	196 domicílios	Tabulação; Cruzamento de Classes; Análise Discriminante e Cluster
Kutter	Braunschweig	1972	-	Análise fatorial
Kutter et. al.	KONTIV/Berlin	1975 1977	9200 indivíduos	Cluster
Lentorp	Stockholm Sweden Pesquisa através do telefone	1967	Viagens realizadas com motivo compras	PESASP (programa de avaliação do conjunto de caminhos alternativos)

Fonte: DAMM (1983)

HAGERSTRAND (1974), LENNTORP (1976), CHAPIN (1974), JONES et. al. (1980), KUTTER (1973), HOLZAPFEL (1980) e HERZ (1980, 1981) definiram “o que as pessoas fazem no tempo e no espaço” como o foco do estudo que permitiu uma diversidade de dimensões para caracterizar o conjunto de atividades (DAMM, 1983). Outros autores destacaram as variáveis para representar a complexidade do comportamento, como ADLER e BEN-AKIVA (1979) que, por exemplo, adotaram um padrão de viagem envolvendo o dia inteiro. A decisão de viagem foi considerada um elemento importante da análise, partindo do pressuposto que qualquer um consegue planejar ou programar simultaneamente suas atividades diárias e que as decisões sobre as viagens são freqüentemente tomadas em conjunto com as decisões sobre viagens prévias e futuras.

BAIN (1976) analisou o tempo total alocado para diversas atividades durante um dia tentando capturar a possibilidade de realização de outras atividades dentro de um padrão de comportamento. Pesquisas sobre o orçamento de tempo mostraram semelhanças nos valores agregados do tempo que grupos de indivíduos com características sociais similares gastam em diferentes tipos de atividades, incluindo viagens (SZALAI, 1972; ZAHAVI, 1976).

As estruturas de tempo e localização geográfica também foram comparadas com o objetivo de mostrar similaridades na maneira como grupos de indivíduos programavam suas atividades ao longo do tempo (JONES et. al., 1983). HANDY (1996) destacou a necessidade de incluir a forma urbana na análise do comportamento das viagens baseadas em atividades.

Os modelos de programação de atividades apareceram nos anos 80. Estes modelos tentaram reproduzir seqüências de atividades entre diferentes grupos de pessoas através da aplicação de modelos estocásticos, com os quais tentavam prever padrões de atividades e de viagens. BURNETT et. al. (1978) utilizaram técnicas das cadeias de Markov para prever quantas são as escolhas e alternativas para a realização ou não de atividades. HANSON (1981) desenvolveu modelos de probabilidades através da observação das seqüências de atividades para modelar padrões de cadeias de viagens associadas com viagens a trabalho. KITAMURA et. al. (1981) verificaram a relação entre cadeias de viagens e número de automóveis e observaram que o número de cadeias de viagens cresce com o número de automóveis disponíveis.

RECKER et. al. (1987) analisaram o padrão de viagem diária em relação ao espaço, tempo e tipo de atividade. Este estudo focou o desenvolvimento de procedimentos para uma análise

da dinâmica relacionada às mudanças nos padrões de atividades e as características do domicílio. Os resultados mostraram que trabalhadores com horários variáveis de trabalho realizaram (1) menos viagens pendulares complexas e (2) mais viagens pendulares por modos a pé, por transporte público e automóvel, em relação aos trabalhadores com horários fixos. As diferenças entre padrões de viagem de trabalhadores com horários fixos e variáveis poderiam ser verificadas através das características sócio-econômicas (renda, gênero, propriedade de automóvel) e através da localização do local do trabalho e do emprego.

BOWMAN e BEN AKIVA (1997) observaram no levantamento dos residentes da área de Boston (1991) a complexidade e a variedade nas programações de atividades e viagens. A análise do primeiro “tour” no padrão de atividade diária mostrou que uma quantidade de pessoas permaneceu na residência o dia inteiro, e que 40% fizeram dois ou mais “tours” durante o dia (Figura 3.1).

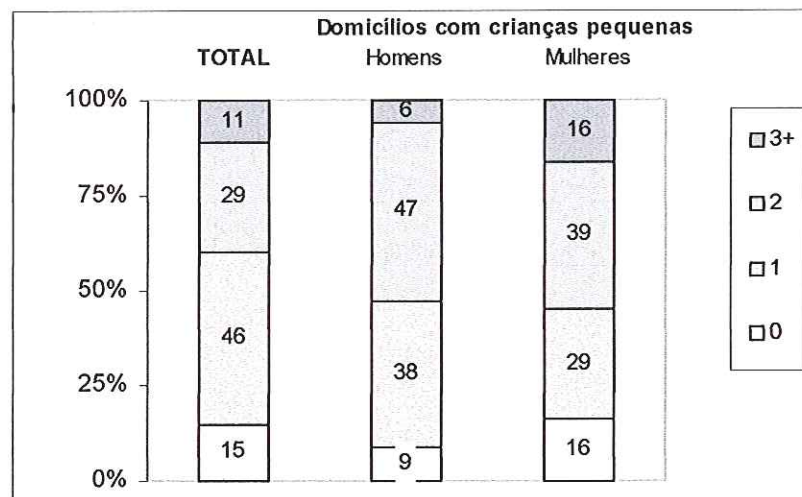


Figura 3.1 – Número de “tours” no padrão diário de viagem – Boston (1991)
Baseado em BOWMAN e BEN-AKIVA (1997)

Além disso, observaram que os padrões de viagens variaram dentro de uma população. Os adultos dos domicílios com crianças pequenas tinham maior tendência em realizar dois ou mais “tours”. Os padrões destes domicílios também apresentaram variações entre homens e mulheres. Os homens apresentaram menor probabilidade de permanecer em casa (9%) e as mulheres maior probabilidade de realizar três ou mais “tours” (16%).

Considerando padrões de viagens a trabalho (Figura 3.2), simples (HWH) e complexas (HWH+ e HW+WH), BOWMAN e BEN-AKIVA (1997) mostraram que 25% dos trabalhadores realizaram atividades fora do local de trabalho em períodos intermediários

(HW+WH), e 36% fizeram viagens ao trabalho (HWH). Nos domicílios com crianças pequenas, aumentam a probabilidade dos homens (40%) em relação às mulheres de realizar viagens simples para o trabalho (23%).

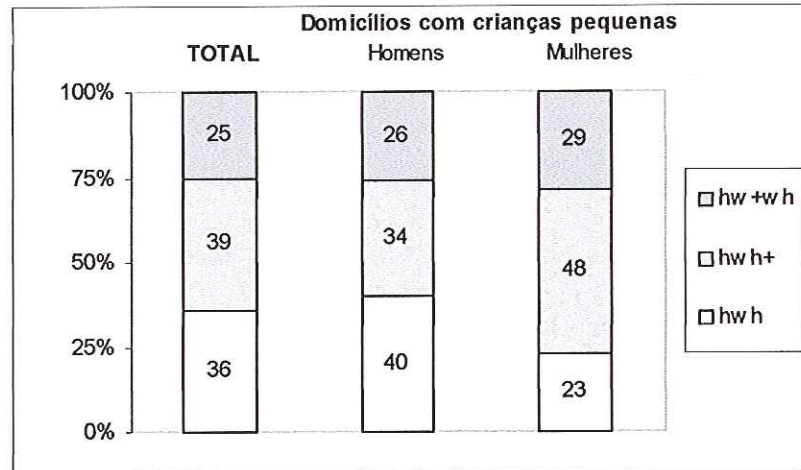


Figura 3.2 – Complexidade de viagens a trabalho
Baseado em BOWMAN e BEN-AKIVA (1997)

Outros estudos indicaram as diferenças de comportamento de viagens pelo gênero, distância de cada viagem realizada, modos, propósito de viagem e distância total de viagem. WACHS (1987) encontrou (1) mulheres que realizaram mais viagens curtas ao trabalho utilizando modos públicos de transporte em relação aos homens, e (2) mulheres que realizaram mais viagens curtas por outros motivos com outros modos de transporte comparadas aos homens. MADDEN (1981) observou que as diferenças nos papéis entre homens e mulheres dentro do domicílio foi o fator que levavam as mulheres a realizar viagens curtas mais próximas das residências. ROSENBLOOM (1987) verificou que mulheres assumiam maiores responsabilidades relacionadas às necessidades de viagens das crianças. BERNARD et. al. (1996) observaram que mulheres trabalhadoras realizaram mais viagens que as não trabalhadoras. GORDON et. al. (1989) verificaram que as viagens não relacionadas ao trabalho são maiores para o grupo de mulheres em relação aos homens. HAMED e MANNERING (1993) encontraram homens com maior probabilidade de ir diretamente para casa depois do trabalho do que as mulheres.

BHAT (1997) observou que as pessoas mais velhas têm maior probabilidade em realizar atividades dentro da residência, isto é, não viajam. ABU-EISHEH e MANNERING (1989) observaram que entre as pessoas que vão para o trabalho de automóvel, os jovens tendem a

correr mais em relação aos mais velhos e estas diferenças alteram a quantidade de tempo gasto no sistema.

MCGUCKIN e MURAKAMI (1999) pesquisaram sobre o comportamento das viagens durante uma semana, entre adultos homens e mulheres, buscando analisar a influência do estágio no ciclo de vida e o sexo no comportamento da viagem diária. Alguns resultados mostraram que, em relação ao homem, mulheres realizam mais viagens ligadas às atividades de manutenção da família, tais como compras. Esses autores observaram também que a participação da mulher na força de trabalho está presente, mas que os padrões de viagens ao trabalho das mulheres são diferentes dos de homens, variando com a situação familiar e a idade dos integrantes do domicílio. Além disso, observaram que a responsabilidade das mulheres dentro do domicílio afeta a escolha do tipo de trabalho e a sua localização.

FOX (1983) destacou que a presença de crianças no domicílio reflete nas atividades das mulheres que trabalham fora de casa, as quais têm padrões distintos de demanda por transporte em relação aos seus companheiros. No entanto, ressalta-se que as diferenças entre chefes e cônjuges se reduzem à medida que seus rendimentos se aproximam (SINGELL e LILLYDAHL, 1986).

Uma das suposições principais da análise baseada em atividades é a de que características particulares dos indivíduos e contextuais estão sistematicamente relacionadas aos seus padrões de viagens (STRATHMAN E DUEKER, 1995). Padrões de viagens representativos estão tipicamente correlacionados com características individuais sócio-econômicas e atributos de viagem.

STRATHMAN e DUEKER (1995), ao verificar o estágio do ciclo de vida, observaram que domicílios com um ou mais adultos correspondem a uma larga porcentagem de cadeia de viagens relacionadas ao trabalho, enquanto que a porcentagem de cadeia de viagens aliadas ao trabalho é menor para os domicílios com um adulto e crianças com idade de 15 anos ou menos. Eles observaram também que, para viagens não relacionadas ao trabalho, a probabilidade relativa de ocorrer viagens complexas é maior para domicílios com um adulto e crianças com idade entre 16 a 21 anos, enquanto a probabilidade de ocorrer viagens complexas é menor para domicílios compostos por dois ou mais adultos com crianças com idade entre 6 a 15 anos e domicílios com dois ou mais adultos. A variável renda foi também analisada por estes autores e verificaram que altas rendas domiciliares levaram a uma maior

tendência de combinação de viagens para o trabalho com outras viagens. No contexto domiciliar, verificaram que domicílios com um indivíduo e domicílios com vários adultos que trabalham têm maior tendência a combinar viagens não-trabalho a viagens a trabalho do que domicílios na qual somente um adulto está empregado. Observaram também que domicílios com um indivíduo têm uma maior tendência de formar padrões de viagens mais complexas.

Em relação ao tipo de ocupação principal, BHAT (2001) comparou resultados entre trabalhadores e não trabalhadores. Um de seus resultados preliminares mostrou que nem sempre os não trabalhadores apresentam irregularidades na realização dos padrões.

CIRILA e AXHAUSEN (2002), através de uma análise empírica, levantaram a idéia de transferibilidade dos modelos de viagens baseados em atividades entre países. Os autores compararam dados de pesquisas americanas, belgas e alemãs. Os aspectos espacial e temporal de mobilidade foram analisados através dos números de “tours” por dia, números de paradas por “tour”, motivos dos “tours” e paradas, a interação entre as paradas e modos de viagem com os diferentes padrões de viagens, a distância entre paradas e a duração das atividades. Segundo a análise dos autores, os resultados apresentaram-se coerentes entre os resultados americanos, belgas e alemães, variando em alguns casos em função de diferenças culturais e reforçando que as políticas adotadas têm papel importante no comportamento dos indivíduos.

A maioria das pesquisas desenvolvidas foi aplicada em dados de levantamento, na sua maioria, de indivíduos de países desenvolvidos. WANG (2001), considerando dados de levantamento de países sub-desenvolvidos, tentou identificar as diferenças nas atividades e comportamento de viagens em diferentes classes sociais e econômicas, analisando os efeitos dos fatores de participação e duração na atividade. Um dos resultados destacou uma relação positiva entre duração e tempo da atividade realizada pelo indivíduo que permitiriam a sua caracterização.

A abordagem baseada em atividades leva a necessidade de uma reestruturação da análise, partindo provavelmente da coleta de dados. Os levantamentos diários de viagem fornecem às pesquisas um conjunto rico de dados, mas existem limitações quando analisados dentro do contexto do encadeamento e de atividades. Alguns autores também levantaram a necessidade de dados mais completos em relação aos dados aplicados aos modelos convencionais. Segundo ALGERS et. al. (2001), para ser capaz de prever o comportamento de viagem seria

necessário incluir como e de que forma as atividades são selecionadas e programadas. A análise incluindo a forma urbana levaria a necessidade de outras variáveis. Além de dados de população e densidade de empregos, a localização das pessoas e atividades, através de sistemas de informação geográfica, por exemplo, seriam necessários (HANDY, 1996). ARENTZE et. al (1997) identificou os dados necessários e a qualidade dos formulários para a coleta dos dados. RICHARDSON et. al. (1995) consideraram essencial determinar o tipo de abordagem adotada, definindo entre viagens realizadas ou atividades que as pessoas fizeram parte. Cada uma requer tipos de questionamentos diferentes e devem ser feitos usando diferentes métodos de levantamento.

A complexidade da análise baseada em atividades também implica em dados com maior gama de informações e em número de pesquisados. Estes dados corresponderiam basicamente às atividades realizadas e suas características (1) aquelas atividades que são realizadas fora do domicílio, (2) localização, (3) a hora de realização e o tipo de atividade, (4) os modos utilizados. Outros elementos como os horários institucionalizados, como horário do comércio, poderiam ser incluídos. No entanto, a coleta poderia ser inviabilizada em função do custo e tempo gasto para o desenvolvimento dos trabalhos.

CAPÍTULO 4 – MINERADOR DE DADOS

Nas últimas duas décadas, com o avanço da tecnologia, houve um crescimento substancial nas formas e na capacidade de armazenamento eletrônico e de processamento de grande quantidade de informações (dados). A conseqüente diminuição no custo de armazenamento desses dados levou esforços e a atenção dirigida na forma de aproveitar melhor estes recursos.

Os dados são muito importantes, pois permitem que os planejadores utilizem esses dados de forma a melhorar a qualidade de suas decisões. A análise dos dados pode fornecer conhecimentos adicionais ao permitir a detecção de tendências e características ocultas, mas implícitas nos dados.

Mineração de dados é uma técnica que busca o conhecimento oculto em grandes bancos de dados. Também é conhecido como KDD, “Knowledge Discovery in DataBases”, ou seja, Descoberta do Conhecimento em Base de Dados (FAYYAD et. al., 1996). A KDD é vista como uma disciplina mais ampla, e o termo em inglês “*Data Mining*” (mineração ou garimpagem de dados) como o componente que trata dos *métodos* de descobrimento do conhecimento (FAYYAD et. al., 1996).

A definição de “data mining” é apresentada abaixo:

Data Mining, DM – ou mineração de dados: tarefa de estabelecer novos padrões de “conhecimento”, geralmente imprevistos, partindo-se de uma massa de dados previamente coletada e preparada para este fim (Tarapanoff, (Org.), 2001).

Basicamente, o minerador de dados está voltado para a análise de dados e o uso de técnicas de “*software*” na procura de padrões em conjuntos de dados. O “*software*” torna-se responsável pela procura de padrões, identificando as regras subjacentes nos dados. Uma das

idéias básicas consiste em encontrar padrões nada óbvios. O processo de análise começa com um conjunto de dados onde é usada metodologia para desenvolver uma representação para a estrutura dos dados, durante o qual conhecimento é adquirido. O conhecimento adquirido pode ser estendido para conjuntos maiores de dados, assumindo que estes têm uma estrutura semelhante aos de dados simples.

A figura seguinte resume algumas das fases em “data mining”.

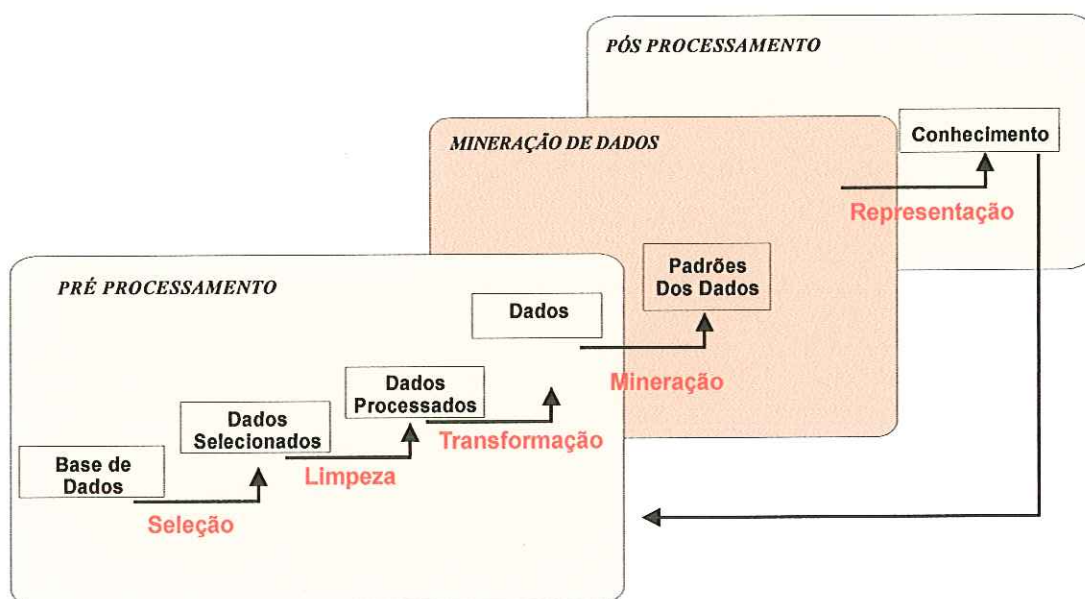


Figura 4.1 – Fases do Minerador de Dados. Fonte: (FAYAAD et. al., 1996)

A seqüência começa com os dados e termina com o conhecimento extraído que foi adquirido em consequência das seguintes fases:

Pré-processamento – fase de preparação e limpeza dos dados, onde é determinada a informação relevante (por exemplo: é desnecessário o campo sexo de um paciente ao estudar gravidez). Nesta fase são executadas as etapas de coleta, armazenagem e “limpeza” dos dados. Para que a fase de pré-processamento seja realizada com sucesso, é necessário que se disponha de conhecimento da base, incluindo o entendimento dos dados, a limpeza e sua preparação para não haver duplicação de conteúdo através de erros de digitação, abreviações diferentes, valores omissos, entre outros.

Nesta fase, os dados são selecionados ou segmentados de acordo com alguns critérios (por exemplo: pessoas que possuem carro). Desta forma, diferentes subconjuntos podem ser

formados. Esta etapa tem a finalidade de identificar todas as fontes internas e externas de informação e selecionar o subconjunto de dados necessários para aplicação de um minerador de dados (QUONIAM et. al, 2001).

Na etapa de processamento, os dados são transformados de forma que possam ser adicionadas *overlays*, isto é, os dados passam a poder ser usados de forma mais coerente e relevante. Os dados são reconfigurados para assegurar um formato consistente, através da eliminação de inconsistências causadas pela diversidade de fontes (por exemplo: sexo pode ser armazenado como “m” e “f” ou “0” e “1”). É a etapa que exige o maior esforço. A etapa é crucial para assegurar a qualidade final dos resultados, por isso as ferramentas utilizadas são tão importantes. Os softwares dedicados a esta etapa devem estar prontos para muitas operações: agregar valor, efetuar conversão, filtrar variáveis, possuir formato de exportação de dados, trabalhar com base de dados relacionais, mapear variáveis de entrada, entre outros (QUONIAM et. al., 2001).

Mineração de Dados – esta fase preocupa-se em extrair padrões dos dados. Um padrão pode ser definido por um conjunto de fatos (dados) F , uma linguagem L , e uma medida de certeza C . Um padrão é uma frase s em L que descreve as relações num subconjunto F_s de F com uma certeza C , de tal forma que s é mais simples em algum sentido que a enumeração de todos os fatos em F_s .

Pós Processamento: corresponde à fase de interpretação, análise e avaliação dos resultados obtidos do processo de mineração de dados. Dois aspectos fundamentais devem ser considerados: informar as novas descobertas e apresentá-las de maneira exploráveis potencialmente (QUONIAM et. al., 2001). Os padrões identificados pelo minerador de dados passam a fazer parte do conhecimento, que podem ser utilizados para o suporte em tomada de decisões, isto é, tarefas de predição e de classificação, resumizando o conteúdo de uma base de dados ou explicando os fenômenos observados.

As técnicas mais utilizadas em “*data mining*” são:

Redes Neurais – Modelos preditivos não lineares que aprendem através de treinamento.

Árvores de Decisão – Estruturas em forma de árvore que representam conjuntos de decisão. Estas decisões geram regras para a classificação de um conjunto de dados.

Alguns métodos de árvores decisão específicos incluem *Classification and Regression Trees* (CART) e *Chi Square Automation Interaction Detection* (CHAID).

Algoritmos Genéticos – Técnicas de otimização que usam processos tais como combinações genéticas, mutações, e seleção natural, baseadas nos conceitos de evolução.

Método Vizinho-mais-próximo – Técnica que classifica cada registro numa base de dados, baseada na combinação das classes dos k registros mais semelhantes a este, num conjunto histórico de dados. Também chamada a técnica dos k vizinhos mais próximos.

Indução de Regras – Extração de regras *if-then* dos dados baseadas em estatísticas significativas. Uma tarefa consiste na descoberta de regras de previsão, do tipo SE...ENTÃO, onde a parte SE (a "condição") da regra especifica alguns valores de atributos previsores e a parte ENTÃO da regra prevê um valor para um determinado atributo cuja previsão é desejada.

Com a aplicação de um minerador de dados pode-se obter vários tipos de descoberta de conhecimento. Dentre eles, descoberta de associações, descoberta de agrupamentos, descoberta de classificações, descoberta de regras de previsão, hierarquias de classificação, descoberta de padrões seqüenciais, descoberta de padrões em séries temporais, categorização e segmentação (ALVARES, 2000).

Este trabalho utilizou como técnica a árvore de decisão disponível no software S-Plus 2000 para obter relações entre as características sócio-econômicas e os padrões de viagens encadeadas. Segue abaixo uma breve descrição da árvore de decisão e sua hierarquia geral, as árvores de classificação e regressão (CART) e as particularidades provenientes do modelo de árvore apresentada no S-Plus 2000.

4.1 Árvores de Decisão

A árvore de decisão é uma forma simples de representação, que classifica exemplos de uma base de dados em um número finito de classes (QUINLAN, 1983).

A árvore de decisão é um diagrama ou gráfico de fluxo que representa um sistema de classificação ou um modelo preditivo. A árvore é estruturada como uma seqüência de questões simples. As respostas para estas questões traçam o caminho da árvore.

O produto final é um conjunto de regras hierárquicas que dividem os dados dentro de grupos, onde uma divisão (classificação ou predição) é feita para cada grupo.

A Figura 4.2 apresenta um esquema de uma árvore de decisão.

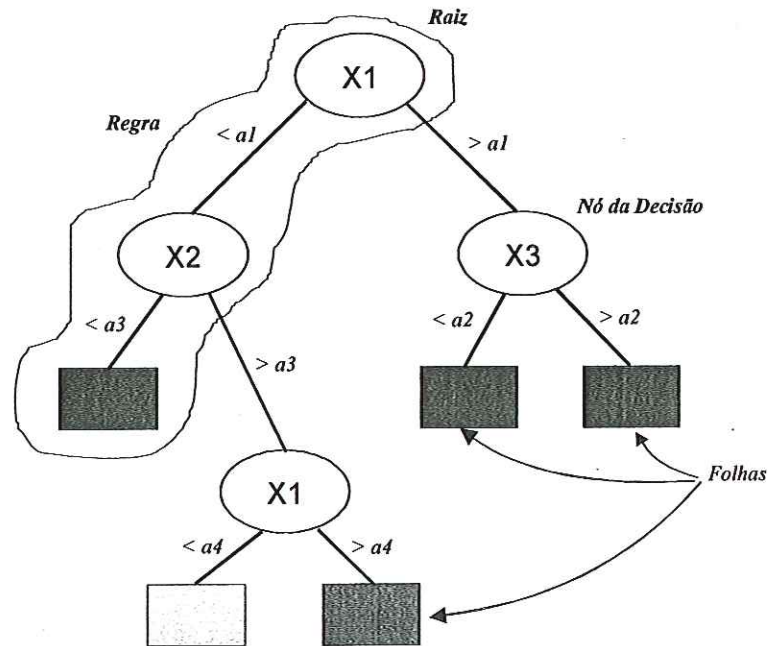


Figura 4.2 – Esquema de uma Árvore de Decisão

A hierarquia é chamada de árvore e cada conjunto de dados é chamado de nó (X1, X2, X3). O conjunto original contém o conjunto completo de dados, referindo-se ao nó raiz da árvore (X1). Este nó contém dados que podem ser subdivididos dentro de outros sub-nós, chamados de nós filhos (X2, X3). Quando os dados do nó não podem ser mais subdivididos em de um outro subconjunto, ele é considerado como um nó final ou terminal (X2, X3). O nó final (nó terminal) é chamado de folha.

Uma árvore de decisão tem a função de particionar recursivamente um conjunto de dados, até que cada subconjunto obtido deste particionamento contenha casos de uma única classe. Para isso, a técnica de árvores de decisão examina e compara a distribuição de classes durante a construção da árvore. Os resultados obtidos, após a construção de uma árvore de decisão, são os dados organizados de maneira compacta.

Muitos são os algoritmos que elaboram árvores de decisão. Não há uma forma de determinar qual é o melhor algoritmo, um pode ter melhor desempenho em determinada situação e outros podem ser mais eficientes em outras situações.

O algoritmo ID3 foi um dos primeiros algoritmos de árvore de decisão, tendo sua elaboração baseada em sistemas de inferência e em conceitos de sistemas de aprendizagem. Logo após foram elaborados diversos algoritmos, sendo os mais conhecidos: C4.5, CART (*Classification and Regression Trees*), CHAID (*Chi Square Automatic Interaction Detection*), entre outros.

Uma das vantagens da árvore de decisão é a sua facilidade de compreensão, podendo produzir diferentes árvores a partir de um conjunto de dados, grande ou pequeno. Normalmente são adequadas para qualquer tipo de dados. A árvore organiza os dados de maneira compacta e fornece uma visão real da natureza do processo. No entanto, quando se utiliza um grande conjunto de dados a análise pode se tornar complexa, dificultando a compreensão da sua representação gráfica.

4.2 Árvores de Classificação e Regressão

As árvores de classificação e regressão - CART (do inglês, *Classification and Regression Tree*) - são baseadas na monografia de Breiman (BREIMAN et. al, 1984) que apresentou um dos mais importantes métodos de árvore de classificação e regressão. O modelo baseado em árvore é um conjunto de regras de classificação que particiona um conjunto de dados em subconjuntos exaustivos e não sobrepostos (MESA et. al., 2000).

As árvores podem ser consideradas um modelo de regressão não-paramétrico e têm por objetivo estabelecer uma relação entre um vetor de variáveis preditoras x e uma única variável resposta y . É ajustado mediante sucessivas divisões binárias no conjunto de dados, de modo a tornar os subconjuntos resultantes cada vez mais homogêneos em relação à variável resposta. Essas divisões são convenientemente representadas por uma estrutura de árvore binária, na qual cada nó corresponde a uma divisão em uma covariável particular (BELL, 1996).

Na CART, as variáveis explicativas e as variáveis resposta podem assumir valores contínuos ou categóricos. Se a variável resposta for contínua, o modelo recebe o nome de **árvore de regressão**; caso contrário, é tratada como uma **árvore de classificação**.

A CART pode ser utilizada como uma alternativa à regressão linear, quando a variável dependente for numérica, e também à regressão logística, quando ela for categórica (FERREIRA et. al., 2001). O resultado apresentado é uma árvore hierárquica de regras de decisão útil para previsão e classificação.

Os componentes básicos de uma CART são os nós e as regras de divisão.

Segundo FERREIRA et. al. (2001), as implementações computacionais consideram que as regras de divisão estejam baseadas em apenas uma das co-variáveis de cada vez, fazendo com que as combinações lineares entre elas não sejam permitidas. Caso essas combinações fossem permitidas, haveria um número grande de possibilidades, tornando o algoritmo lento e, dessa forma, sem utilidade prática.

Apesar dos diversos tipos de programas, modelos e ferramentas disponíveis para análise de dados através dos mineradores de dados, este trabalho utilizou o modelo de árvore apresentando pelo *software* S-Plus 2000 desenvolvido pela *StatSoft*, pois ele está disponível no banco de softwares do Centro de Computação Eletrônica da USP.

4.3 Árvores de Classificação e Regressão – S-Plus 2000

O S-Plus 2000 é um *software* minerador de dados que possui o modelo de árvore de classificação e regressão. A árvore de classificação e regressão do S-Plus 2000 é uma variante do algoritmo do CART, descrita por CLARK E PREGIBON (1993). Ele trata a árvore como modelo de probabilidade e emprega o desvio como o critério de divisão.

O *software* pode ser usado para a construção de árvores proveniente de um conjunto de dados completos (não podendo haver falta de dados em qualquer campo do banco de dados) com no máximo 128 variáveis resposta categórica e 32 níveis de variáveis explicativas categóricas, por que se uma variável explicativa tem k níveis então $2^{(k-1)} - 1$ divisões poderão ser analisadas em função da capacidade do sistema. O *software* automaticamente decide se ajusta a uma árvore de regressão ou classificação de acordo com o tipo de variável resposta. Da mesma forma que o CART, se a variável resposta é categórica, a árvore de classificação é

construída. Se a resposta é contínua, constrói-se a árvore de regressão. A divisão dos dados dentro de cada nó é progressiva tendendo à homogeneidade até que os nós contenham poucas observações (o padrão mínimo corresponde a 5 observações).

O modelo usado pelo S-Plus 2000 para a árvore de classificação assume que a variável resposta segue uma distribuição multinomial. Os pesos não são usados na computação dos desvios da árvore. Para a árvore de regressão assume que a variável numérica de resposta tem uma distribuição normal (Gaussiano). Os pesos são usados se eles estiverem especificados. A cada passo no crescimento da árvore, o particionamento dos dados é feito para minimizar o desvio ou a entropia em todas as divisões permitidas nos nós da árvore (BREIMAN et. al.,1984).

Nas árvores de classificação, a heterogeneidade de cada nó é medida pelo desvio (D) da seguinte forma:

$$D = -2 \sum_t \sum_{j \in t} n(j|t) \log p(j|t) \quad (1)$$

em que: $n(j|t)$ é o número de casos no nó t com a resposta categórica j ;
 $p(j|t)$ denota a proporção dos casos na categoria j em relação a todos os casos pertencentes ao nó t .

Isto significa que na árvore do S-Plus 2000 o desvio foi definido como recíproco da função verossimilhança elevada ao quadrado. Portanto, o desvio significa uma medida da heterogeneidade dos grupos resultantes.

Um dos principais motivos que levaram à escolha da árvore de classificação e regressão é a sua capacidade de representar a natureza probabilística do objeto analisado, que no caso representa os padrões de viagens encadeadas. Ela é um modelo que reconhece que “indivíduos homogêneos” podem tomar diferentes decisões, e associa a probabilidade às diferentes respostas possíveis.

A Tabela 4.1 apresenta um resumo com as principais características do S-Plus 2000.

Tabela 4.1 - Características Gerais do S-Plus 2000

Tipo de variáveis de entrada:	Leitura de variáveis categóricas ou variáveis numéricas
Variáveis explicativas:	Variáveis categóricas ou numéricas
Tipo de Resposta:	Variável Categórica (árvores de classificação) Variável Contínua (árvores de regressão)
Capacidade para controlar grande número de casos:	Sim
Capacidade para controlar falta de dados:	Não
Segmentação:	Binária
Regras de divisão:	Árvores de Classificação: Desvio Árvores de Regressão: Mínimo Quadrado
Regras de poda:	Medida de custo/complexidade
Fornecimento de Árvore ótima:	Não

Fonte: MESA et. al. (2000)

4.4 Algumas considerações sobre a utilização do modelo de árvore do S-Plus 2000

A manipulação dos dados para aplicação da árvore não exige um roteiro específico e as informações relevantes em relação aos resultados são apresentados num único relatório. Em relação ao gráfico básico da árvore, ela apresenta opções de comprimento das ramificações proporcionais à divisão de significância, isto é, para um mesmo conjunto de dados, os resultados podem apresentar formas diferentes. A forma de apresentação vai depender da necessidade das informações.

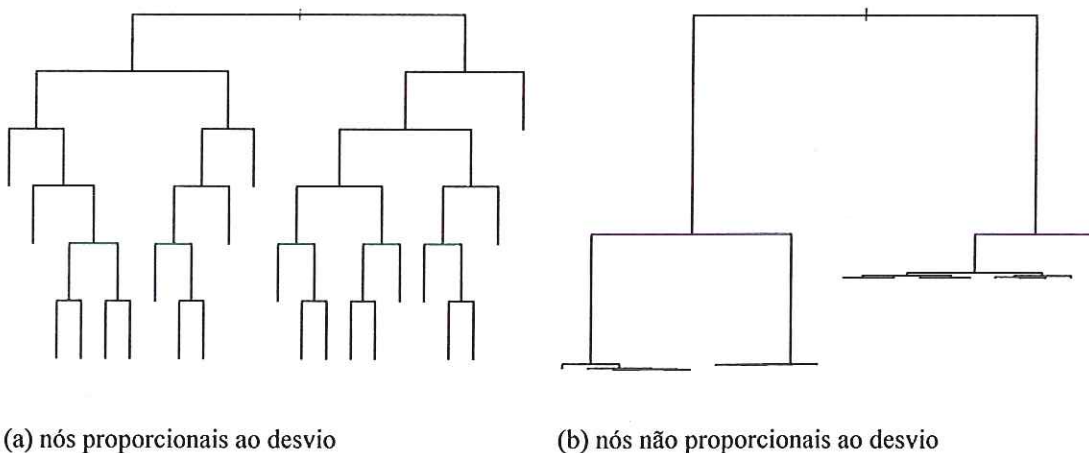


Figura 4.3 – Resultados Gráficos da árvore em relação à proporcionalidade

O modelo básico de avaliação gráfica é baseado nos erros de classificação ou no desvio em relação à classificação verdadeira. Além disso, o programa possui opções para especificar

dados de validação simultâneos aos dados de treinamento. A modelagem não permite dados incompletos (células com valores em branco) para as variáveis explicativas.

Quanto à característica de arquivos de importação, vários tipos de arquivos são reconhecidos e facilmente importados. Entre algumas desvantagens observadas, cita-se a inexistência de uma dinâmica entre o gráfico e o relatório. O relatório é apresentado na forma escrita, sendo necessário a sua formatação em tabela para melhor compreensão dos resultados.

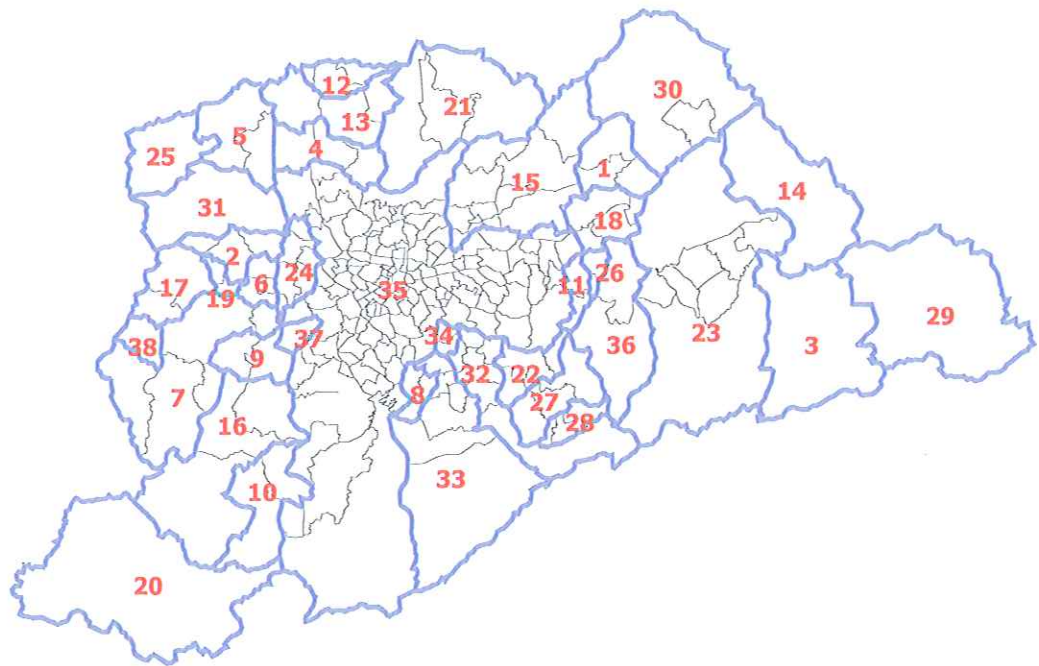
O modelo também não produz uma árvore ótima: o algoritmo constrói uma árvore máxima, isto é, a árvore cresce até formação de todas as classes e o resultado é apresentado de acordo com o desvio e número de observações estabelecido pelo usuário.

Em relação aos desvios adotados, os resultados encontrados com desvios menores aumentam a complexidade dos resultados, podendo apresentar muitos sub-conjuntos. Isto não necessariamente significa resultados melhores. Este trabalho utiliza para análise o desvio de 0,05. O próximo capítulo detalha o desenvolvimento das sub-divisões dos nós em função do desvio mínimo adotado. O ANEXO D apresenta, de forma a ilustrar o número de sub-conjuntos, os resultados gerados a partir de uma árvore com desvio adotado de 0,01.

CAPÍTULO 5 - ESTUDO DO CASO DA REGIÃO METROPOLITANA DE SÃO PAULO

5.1 A Pesquisa Origem e Destino de São Paulo

Este trabalho foi realizado a partir dos dados da Pesquisa Origem Destino da Região Metropolitana de São Paulo de 1987, realizada pela Companhia do Metropolitano de São Paulo (CMSP). A Região Metropolitana de São Paulo (RMSP), segundo o censo do Instituto Brasileiro de Geografia e Estatística (IBGE) de 1987, contava com uma população de 16 milhões de habitantes distribuídas em uma área de aproximadamente 8.050 km², correspondentes aos 38 municípios. São eles: Arujá, Barueri, Biritiba-Mirim, Caieiras, Cajamar, Carapicuíba, Cotia, Diadema, Embu, Embu-Guaçu, Francisco Morato, Ferraz de Vasconcelos, Franco da Rocha, Guararema, Guarulhos, Itapeverica da Serra, Itapevi, Itaquaquecetuba, Jandira, Juquitiba, Mairiporã, Mauá, Mogi das Cruzes, Osasco, Pirapora do Bom Jesus, Poã, Ribeirão Pires, Rio Grande da Serra, Salesópolis, Santa Isabel, Santana de Parnaíba, Santo André, São Bernardo do Campo, São Caetano do Sul, São Paulo, Suzano, Taboão da Serra, Vargem Grande Paulista.



- | | | | |
|--------------------|----------------------------|----------------------------|-----------------------------|
| 1 – ARUJÁ | 11 – FERRAZ DE VASCONCELOS | 21 – MAIRIPORÃ | 31 – SANTANA DO PARNAÍBA |
| 2 – BARUERI | 12 – FRANCISCO MORATO | 22 – MAUÁ | 32 – SANTO ANDRÉ |
| 3 – BIRITIBA MIRIM | 13 – FRANCO DA ROCHA | 23 – MOGI DA S CRUZES | 33 – SÃO BERNARDO DO CAMPO |
| 4 – CAIEIRAS | 14 – GUARACEMA | 24 – OSASCO | 34 – S. CAETANO DO SUL |
| 5 – CAJAMAR | 15 – GUARULHOS | 25 – PIRAPORA DO BOM JESUS | 35 – SÃO PAULO |
| 6 – CARAPICUIBA | 16 – ITAPECERICA DA SERRA | 26 – POÁ | 36 – SUZANO |
| 7 – COTIA | 17 – ITAPEVI | 27 – RIBEIRÃO PIRES | 37 – TABOÃO DA SERRA |
| 8 – DIADEMA | 18 – ITAQUAQUECETUBA | 28 – RIO GRANDE DA SERRA | 38 – VARGEM GRANDE PAULISTA |
| 9 – EMBU | 19 – JANDIRA | 29 – SALESÓPOLIS | |
| 10 – EMBU-GUAÇU | 20 – JUQUITIBA | 30 – SANTA ISABEL | |

Figura 5.1 – 38 municípios da Pesquisa O/D 1987 - RMSP

Os dados considerados na análise referem-se aos dados sócio-econômicos e de viagens, localização do emprego e da escola de cada indivíduo residente entrevistado e dados gerais do domicílio tais como condição de ocupação e número de veículos.

Além disso, os dados correspondem às 254 zonas identificadas pela homogeneidade no uso do solo, obtendo os dados agregados por zona, além dos dados de origem e destino de viagens interzonais. No entanto, neste trabalho foram consideradas apenas as 204 zonas pesquisadas. As demais zonas corresponderam aos grandes equipamentos urbanos sem residentes, algumas com grande poder de atração de viagens e outras com ocupação rarefeita (ANEXO A).

O questionário para a coleta dos dados da pesquisa domiciliar para a RMSP foi dividido em três tipos: o primeiro refere-se a informações relativas ao domicílio (tipo do domicílio, condição de moradia e valor do aluguel), informações relativas à família (tempo de

residência no bairro, tamanho da família, itens de conforto familiar) e informações relativas às características sócio-econômicas de cada pessoa na família (posição na família, idade, sexo, nível de instrução, ocupação profissional e classe de atividade da empresa em que trabalha e a faixa de renda mensal). O segundo tipo corresponde a informações sobre o local de trabalho e/ou estudo, endereço de trabalho e/ou estudo. E o terceiro tipo de questionário corresponde a informações relativas às viagens realizadas no dia anterior à entrevista pelas pessoas do domicílio (origem, destino, hora de início e término da viagem, motivo, modo e duração).

O período da coleta compreendeu os meses de outubro, novembro e os dez primeiros dias de dezembro de 1987, correspondendo a 80% dos questionários. Em março, abril, maio e junho de 1988 foram aplicados os 20% dos questionários restantes.

A pesquisa O/D 1987 coletou basicamente as informações referentes à família (tamanho da família e número de automóveis), e as informações relativas às características sócio-econômicas de cada pessoa na família. As informações relativas às viagens realizadas por cada membro do domicílio no dia anterior à entrevista foram basicamente àquelas relacionadas ao motivo de viagem. Não foram registradas as atividades realizadas dentro do domicílio ou no local de trabalho, por exemplo. Também não foram registradas as viagens realizadas com distâncias menores que 300 metros. Para a definição do modo principal a partir dos dados da pesquisa, utilizou-se como o de maior hierarquia empregado na viagem dada em ordem decrescente a seguir: metrô, trem, tróleibus, ônibus fretado, ônibus diesel, lotação, taxi, dirigindo automóvel, passageiro de automóvel, caminhão, bicicleta-moto, outros (CMSP, 1988). As zonas de origem e de destino, o modo e a duração da viagem não foram considerados na definição dos padrões de viagens por causa da limitação do S-Plus 2000, a qual será comentada adiante. Os dados utilizados na análise foram aqueles desagregados em nível individual.

Na realidade, a pesquisa O/D de 1987 foi a terceira de uma série de quatro pesquisas realizadas na Região Metropolitana de São Paulo até o momento. A primeira pesquisa de 1967 (Figura 5.2), foi realizada em 15 municípios e foi concebida para subsidiar os estudos de planejamento da rede básica de linhas do Metrô. A de 1977 abrangeu 27 municípios da Região Metropolitana, sendo que em alguns municípios a coleta de dados foi feita integralmente e em outros municípios, apenas a área urbanizada. A terceira, de 1987, envolveu todos os municípios da Região Metropolitana, coincidindo os limites da área de

pesquisa com os limites da área de estudo. Na quarta pesquisa, em 1997, foram pesquisados 39 municípios da metrópole (UPP, 2002).

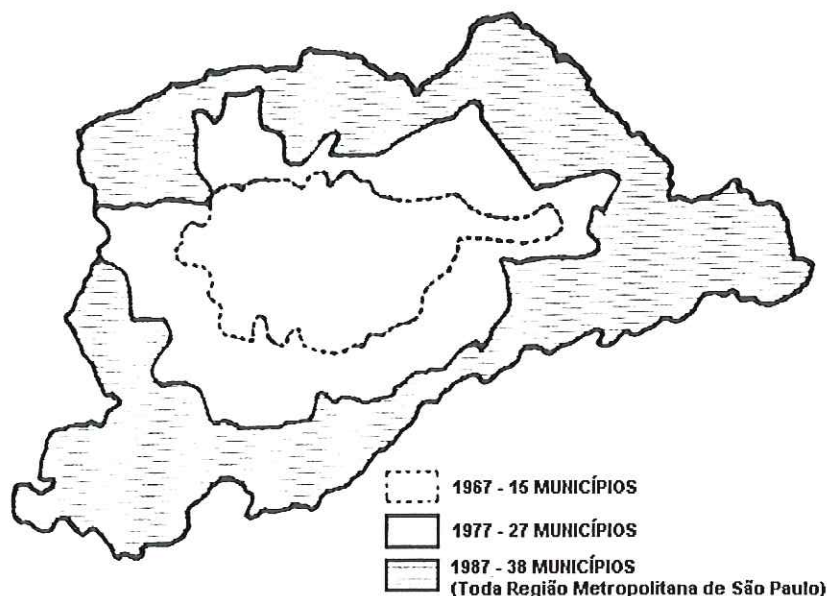


Figura 5.2 – Área de Abrangência das Pesquisas O/D realizadas em 1967, 1977 e 1987

Para se ter uma idéia do ambiente sócio-econômico político em que foi realizada a terceira pesquisa, cabe tecer alguns comentários sobre a realidade da época. Desde o início da década de 80, o país sofria um processo de recessão, o Governo militar se encontrava em crise, e o desemprego aumentava, sobretudo no setor industrial que continuou a ser um dos problemas mais preocupantes do país, atingindo 16,5% em 1997. Com o desemprego formal, cresceu a participação da economia informal, incluindo camelôs, trabalhadores sem carteira assinada e autônomos.

Em 1987, a população da Região Metropolitana de São Paulo era basicamente constituída por jovens com menos de 25 anos (50%) e 15% correspondiam a pessoas com mais de 65 anos. Esta população era constituída por 48% do sexo masculino e 52% do sexo feminino (CMSP, 1988). Em relação a escolaridade, 37% da população era constituída por analfabetos e aqueles com primário incompleto. Desta população, 59% correspondiam a rendas com até 4 salários mínimos e 41% a rendas entre 4 a 8 salários mínimos. Indivíduos que possuíam nível superior de escolaridade significavam somente 7% da população. 31% destes eram encontrados nas famílias com renda acima de 30 salários mínimos e 13% nas faixas de renda entre 15 e 30 salários mínimos. Já nas faixas de renda de 8 a 15 salários mínimos e 15 a 30

salários mínimos predominavam as pessoas com primário completo ou ginásial incompleto (CMSP, 1988).

Pelos dados da Pesquisa OD de São Paulo, de 1987 a 1997, os deslocamentos motorizados individuais subiram de 28% para 32% do total das viagens da RMSP enquanto que o transporte coletivo caiu de 36% para 33%. Considerando o período de 20 anos, essa queda foi muito maior, pois em 1977 o transporte coletivo representava cerca de 44% do total de deslocamentos da região. Outra forma de deslocamento que aumentou significativamente nos últimos 20 anos foi a viagem a pé. Desde 1967 observa-se uma consistente diminuição no ritmo das viagens motorizadas, embora a frota de automóveis particulares tenha crescido consideravelmente, principalmente no período de 1987 a 1997. Atualmente, cerca de um terço das viagens dos paulistanos são deslocamentos a pé. Apesar de ocorrerem em intensidades diferentes, esse fenômeno de aumento do transporte por automóveis e redução dos deslocamentos por transporte coletivo são semelhantes em todas as cidades brasileiras.

5.2 Tratamento dos Dados

Para aplicação do minerador de dados, os esforços foram concentrados na fase de pré-processamento ou de tratamento dos dados, isto é, seleção e limpeza dos dados. Após o tratamento dos dados, foi possível definir a população em estudo e a amostra recolhida.

Esta fase consome mais tempo e esforço. Os dados recebidos pela CMSP (1988) foram extraídos, transformados e limpos. Na extração dos dados estes foram integrados atendo-se a heterogeneidade do banco de dados. A integração foi ser realizada porque normalmente os dados são originados de múltiplas fontes. Isto é, o dado pode estar definido com nomes diferentes em cada banco de dado analisado, ou atributos de mesmos nomes podem ter finalidades diferentes. Isto se faz necessário, já que estes aspectos podem envolver a heterogeneidade em banco de dados (BERSON e SMITH, 1997).

A fase do tratamento dos dados é responsável pela consolidação das informações relevantes para o algoritmo minerador com o objetivo de reduzir a complexidade do problema.

A limpeza dos dados envolve a verificação da consistência das informações, a correção de possíveis erros e o preenchimento ou a eliminação de valores nulos e redundantes. Nessa fase são identificados e removidos os dados duplicados e/ou corrompidos. A execução dessa

fase corrige a base de dados eliminando consultas desnecessárias que seriam executadas pelo algoritmo minerador e que afetariam o seu processamento.

Em relação aos dados dentro das modelagens baseadas em atividades, os dados necessários em geral devem ser mais completos em relação aos dados utilizados nos modelos convencionais. Isto por que, para que os modelos sejam capazes de prever um comportamento de viagem deve-se incluir como e de que forma as atividades são selecionadas e programadas (ALGERS et. al., 2001). Dessa forma, estas abordagens levam a necessidade de uma reestruturação da análise, partindo provavelmente da coleta de dados (ARENTZE et. al., 1997).

Os dados da pesquisa correspondem às informações sobre 110.629 indivíduos entrevistados que residem em 26.056 domicílios.

Estes dados foram divididos primeiramente em dois conjuntos:

- a) Dados completos (Código 6): correspondem aos indivíduos que responderam a todo o questionário da pesquisa;
- b) Dados incompletos (Código 1,2,3,4 e 5): correspondem aos indivíduos que responderam parcialmente ou não responderam o questionário da pesquisa.

Os dados completos (101.411 indivíduos) correspondem a 91,67% de toda a amostra (110.629 indivíduos). A Tabela 5.1 apresenta os indivíduos classificados entre os que realizaram viagens (65,90%) e os que não realizaram viagens (34,10%) no dia anterior à pesquisa. A Tabela 5.2 apresenta o total de indivíduos segundo número de viagens realizadas no dia.

Tabela 5.1 – Indivíduos que viajam e não viajam

Número de Viagens	Total de Indivíduos	%
Não Viajaram	34.579	34,10%
Viajaram	66.832	65,90%
Total	101.411	100,00%

Tabela 5.2 – Total de Indivíduos por Número de Viagens realizadas no dia

Número de Viagens	Total de Indivíduos	% em relação aos dados completos	% em relação à amostra total
0	34.579	34,10%	31,26%
1	677	0,67%	0,61%
2	47.718	47,05%	43,13%
3	4.552	4,49%	4,11%
4	9.522	9,39%	8,61%
5	1.519	1,50%	1,37%
6	1.666	1,64%	1,51%
7	441	0,43%	0,40%
8	385	0,38%	0,35%
9	141	0,14%	0,13%
10	91	0,09%	0,08%
>10	120	0,12%	0,11%
Indivíduos c/ dados completos			
	101.411	100,00%	
Amostra total			91,67%

Indivíduos que haviam realizado duas viagens no dia correspondem a 43,13% de toda a amostra e 71,4% de todos os indivíduos que haviam viajado. Indivíduos que realizaram três viagens corresponderam a 4,11% de toda a amostra e 6,8% dos indivíduos que viajaram. Já os indivíduos que realizaram quatro viagens durante o dia correspondem a 8,61% da amostra e 14,2% dos indivíduos que viajaram.

Neste trabalho, considera-se a residência como a base dos deslocamentos (primeiro e último no dia) realizados pelos indivíduos. Ao observar outros estudos sobre viagens com origem e destino na residência observa-se uma semelhança nos dados (Tabela 5.3.) se comparados com os dados completos da RMSP (91,67%).

Tabela 5.3 – Viagens Urbanas com origem e destino na residência

Local	%
Campina Grande	98 %
Cidades Alemãs	93 %
Lisboa	93 %
Lehndorf (Alemanha)	94 %

Fonte: LEAL (1983)

Para este estudo foram utilizados os dados dos indivíduos que realizaram duas (47,05%), três (4,49%) e quatro viagens (9,39%), além dos indivíduos que não realizaram viagens (34,10%), o que corresponde a 95,03% do total dos indivíduos (Tabela 5.4).

Tabela 5.4 – Total de Indivíduos por Número de Viagens realizadas no dia

Número de Viagens	Total de Indivíduos	Frequência (%)
0	34.579	34.10%
2	47.718	47.05%
3	4.552	4.49%
4	9.522	9.39%
Total	101.411	95.03%

As demais viagens correspondem a 4,97%. Indivíduos que realizaram uma viagem (0,67%) ou mais de quatro viagens (4,30%) no dia correspondem a uma porcentagem relativamente pequena, ou aproximadamente 5.040 pessoas as quais não foram consideradas neste estudo. O aumento do número de viagens leva ao acréscimo no número de combinações possíveis na formação de padrões de viagens e as amostras correspondentes a elas poderiam ter pouca representatividade. Para a análise de viagens baseada na teoria das atividades deve-se ainda classificar esses indivíduos segundo padrões de viagens.

A Figura 5.3 mostra a toda a distribuição e classificação do conjunto de dados que levam ao conjunto de dados de análise.

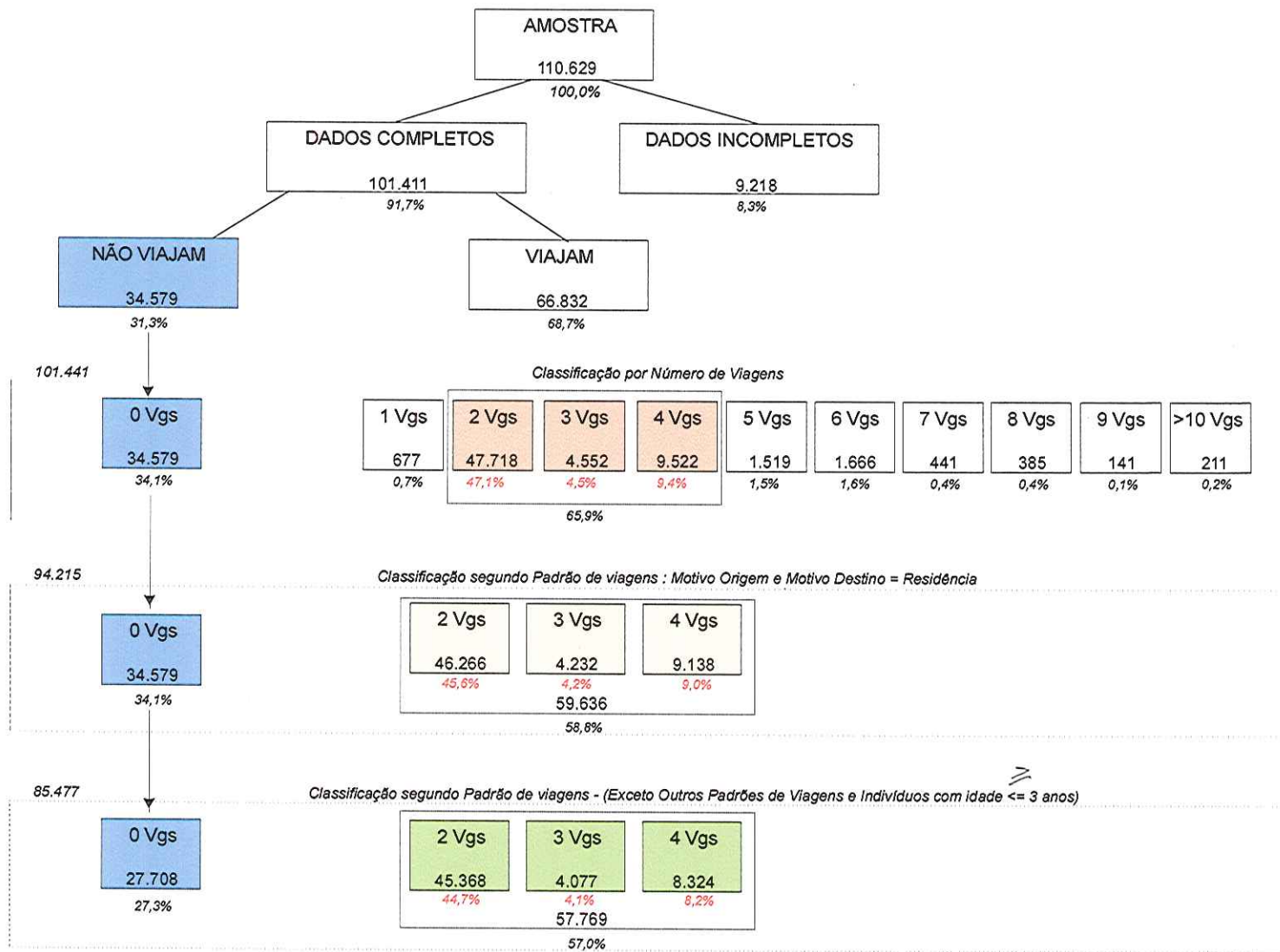


Figura 5.3 – Distribuição dos dados da amostra

Para a classificação, segundo o padrão de viagens, foram considerados os seguintes critérios:

- a) Indivíduos que não realizaram viagens e aqueles que realizaram duas, três e quatro viagens durante o dia. Foram desconsiderados os indivíduos que realizaram uma viagem ou mais de quatro viagens no dia;
- b) somente aqueles indivíduos cuja primeira viagem tem iniciado na residência do entrevistado e a última viagem do dia também terminou na residência; e,
- c) indivíduos com idade acima dos 3 anos.

Em resumo, os dados a serem analisados a partir dessa classificação por padrões de viagens utilizaram dados de indivíduos que realizaram duas (53,08%), três (4,77%) e quatro viagens (9,74%), além dos indivíduos que não realizaram viagens (32,42%), o que corresponde a 84,29% do total dos dados completos (Tabela 5.5 e 5.6).

Tabela 5.5 – Total de indivíduos analisados por número de viagens realizadas no dia

Número de Viagens	Total de Indivíduos	Frequência (%)
0	27.708	32,42%
2	45.368	53,08%
3	4.077	4,77%
4	8.324	9,74%
Total	85.477	100,00%
% em relação ao dados completos		84,29%

Tabela 5.6 – Total de indivíduos analisados segundo aqueles que viajaram e não viajaram

Número de Viagens	Total de Indivíduos	Frequência (%)
Não Viajaram	27.708	32,42%
Viajaram	57.769	67,58%
Total	85.477	100,00%

5.3 Variáveis extraídas da Pesquisa O/D 87 - RMSP

Características como a renda familiar, o número de automóveis na família e o tamanho ou a estrutura das famílias são normalmente utilizadas para descrever as famílias; idade, sexo e ocupação são variáveis típicas para caracterizar os indivíduos (ORTÚZAR e WILLUMSEN, 1994).

As variáveis que podem explicar os padrões de viagens extraídas da Pesquisa O/D 87 são: (a) Características sócio-econômicas (Familiares: número de automóveis por domicílio, número de pessoas na família, renda familiar, e Individuais: idade, sexo, ocupação principal, situação familiar, renda Individual e grau de Instrução); (b) Atividade/ Motivo de viagem; (c) Características de viagens (número de viagens, hora de saída e chegada de cada viagem, zona de origem e destino de cada viagem). De modo a facilitar a análise dos dados e a leitura dos relatórios, foram usadas abreviaturas para designar cada uma das características sócio-econômicas, cujas correspondências podem ser consultadas no ANEXO B.

A partir destas variáveis foram selecionadas aquelas que possuíam atributos referentes às características sócio-econômicas e de atividades/motivo de viagem do conjunto de atributos do banco de dados. O conjunto selecionado constituiu os dados de entrada para o algoritmo de mineração dos dados. Essa seleção teve como objetivo a redução do tempo de processamento do algoritmo minerador, visto que ele trabalhará apenas com dados necessários, diminuindo o seu espaço de busca.

5.4 Codificação dos Dados

Os objetivos básicos da codificação podem ser resumidos em dois: um deles é converter valores quantitativos em valores categóricos e o outro é distribuir os valores contínuos dos atributos (inteiros ou reais) numa lista de intervalos representados por um código, ou seja, cada intervalo pode resultar num valor discreto do atributo.

São dois os objetivos da codificação considerados neste trabalho. O primeiro objetivo é categorizar os atributos, sejam eles individuais, familiares ou de viagens. O segundo refere-se aos padrões de viagens, sintetizando diferentes características de viagens através de um código. As viagens realizadas foram caracterizadas pelo seu motivo de viagem.

A Figura 5.4 mostra a codificação dos valores relativos aos motivos da viagem através de letras e o sequenciamento das atividades ao longo do tempo. As letras indicam as atividades desempenhadas pelos indivíduos no dia e a seqüência das letras significa a ordem cronológica em que as atividades foram desenvolvidas. Para exemplificar a inserção de outras características no padrão de viagem, a Figura 5.5 mostra, além dos motivos de viagem, o modo de viagem adotado para cada viagem. Os números correspondem à codificação dos modos de viagens.

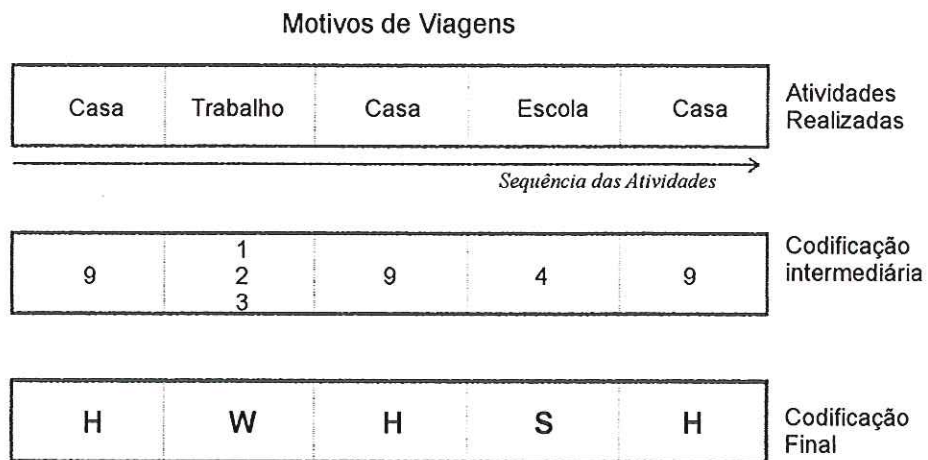


Figura 5.4 – Exemplo da forma de codificação dos motivos de viagens por letras. Os números significam a codificação adotada pela pesquisa O/D 1987.

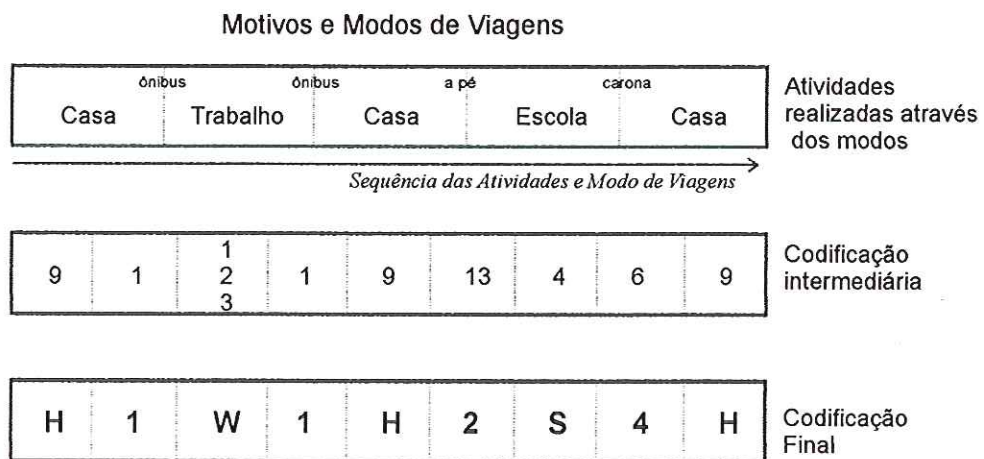


Figura 5.5 – Exemplo da forma de codificação dos motivos de viagens por letras e modos por números. Os números significam a codificação adotada pela pesquisa O/D 1987 e os números finais correspondem a codificação final.

Como vantagens, a codificação de um atributo permite:

- a) melhorar a compreensão do conhecimento descoberto;
- b) reduzir o tempo de processamento para o algoritmo minerador, diminuindo o seu espaço de busca e;
- c) facilitar o algoritmo a tomar decisões globais.

Alguma desvantagem pode ser observada na redução da medida de qualidade de um conhecimento descoberto, perdendo assim detalhes relevantes sobre as informações extraídas.

Em virtude da existência de diversos tipos de codificação (1 a 59) para a ocupação principal declarada pelo entrevistado, as únicas consideradas neste trabalho foram “estuda” e “trabalha” por apresentarem considerável significância na caracterização dos padrões de viagem e pela facilidade na coleta de informações.

A Tabela 5.7 corresponde uma análise detalhada dos quatro grupos de indivíduos formados através da utilização das variáveis “estuda” e “trabalha” e a ocupação principal declarada por estes indivíduos. O objetivo desta análise foi verificar se é possível codificar a ocupação principal declarada em duas variáveis (“estuda” e “trabalha”).

COD	OCUPAÇÃO PRINCIPAL	N TRAB/N EST		N TRAB/ EST		TRAB/ N EST		TRAB/EST		Total	%
		Freq.	%	Freq.	%	Freq.	%	Freq.	%		
1	estudante	52	0,27%	19.196	99,33%	2	0,01%	76	0,39%	19.326	22,61%
2	prendas domesticas	15.451	88,61%	156	0,89%	1.725	9,89%	105	0,60%	17.437	20,40%
3	aposentado	4.932	97,16%	8	0,16%	136	2,68%	-	0,00%	5.076	5,94%
4	sem ocupacao, nunca trabalhou	6.271	83,10%	1.275	16,90%	-	0,00%	-	0,00%	7.546	8,83%
5	desempregado temporariamente	2.590	93,74%	173	6,26%	-	0,00%	-	0,00%	2.763	3,23%
6	em licenca, afastado	209	38,49%	3	0,55%	316	58,20%	15	2,76%	543	0,64%
11	EMPREGADO	30	1,03%	8	0,27%	2.398	82,01%	488	16,69%	2.924	3,42%
12	FIRMA, EMPRESA	38	0,94%	3	0,07%	3.675	90,92%	326	8,07%	4.042	4,73%
13	escriturario	14	0,34%	12	0,30%	2.883	71,04%	1.149	28,31%	4.058	4,75%
14	comercio: vendedor, corretor etc	27	1,38%	6	0,31%	1.666	85,00%	261	13,32%	1.960	2,29%
15	chefe de servicos especializados	10	0,96%	-	0,00%	963	92,42%	69	6,62%	1.042	1,22%
16	nivel supervisor em escritorio	4	0,56%	-	0,00%	593	82,94%	118	16,50%	715	0,84%
17	empregado com nivel de gerencia	4	0,61%	1	0,15%	610	92,85%	42	6,39%	657	0,77%
18	empregado de nivel universitario	14	1,47%	2	0,21%	844	88,66%	92	9,66%	952	1,11%
21	OPERARIO	3	0,27%	1	0,09%	985	87,63%	135	12,01%	1.124	1,31%
22	TRAB. RURAL	15	0,97%	2	0,13%	1.390	89,74%	142	9,17%	1.549	1,81%
23	oper. especializados de oficinas: tecnicos, artesaos	14	1,15%	-	0,00%	1.106	90,73%	99	8,12%	1.219	1,43%
24	oper. especializado (fabrica)	11	0,74%	-	0,00%	1.364	92,04%	107	7,22%	1.482	1,73%
25	chefe de operarios especializados	7	1,63%	-	0,00%	390	90,91%	32	7,46%	429	0,50%
26	trabalhadores rurais	3	2,44%	3	2,44%	107	86,99%	10	8,13%	123	0,14%
27	trabalhador rural nivel capataz	3	18,75%	-	0,00%	13	81,25%	-	0,00%	16	0,02%
31	PROFISSIONAL	50	2,51%	2	0,10%	1.861	93,28%	82	4,11%	1.995	2,33%
32	AUTONOMO	34	1,91%	2	0,11%	1.681	94,23%	67	3,76%	1.784	2,09%
33	nivel de operario especializado	14	1,99%	-	0,00%	664	94,59%	24	3,42%	702	0,82%
34	trabalho intelectual, mental	10	5,75%	2	1,15%	143	82,18%	19	10,92%	174	0,20%
35	tecnico altamente especializado	3	1,94%	-	0,00%	139	89,68%	13	8,39%	155	0,18%
36	nivel universitario	6	1,60%	-	0,00%	358	95,21%	12	3,19%	376	0,44%
41	FUNCIONARIO	3	0,93%	-	0,00%	300	92,59%	21	6,48%	324	0,38%
42	PUBLICO	3	0,51%	-	0,00%	539	90,89%	51	8,60%	593	0,69%
43	escriturario, aux. de escritorio	3	0,49%	2	0,33%	503	82,06%	105	17,13%	613	0,72%
44	nivel universitario, professores	7	0,86%	-	0,00%	723	88,60%	86	10,54%	816	0,95%
45	nivel de diretor	1	0,64%	-	0,00%	141	89,81%	15	9,55%	157	0,18%
46	altos cargos publicos	1	0,52%	-	0,00%	171	88,14%	22	11,34%	194	0,23%
47	militares	5	2,89%	-	0,00%	145	83,82%	23	13,29%	173	0,20%
51	SOCIO OU	5	0,47%	-	0,00%	1.035	97,18%	25	2,35%	1.065	1,25%
52	PROPRIETARIO DE	2	0,49%	-	0,00%	388	94,17%	22	5,34%	412	0,48%
53	micro_empresa industria/comercio	8	1,25%	-	0,00%	614	95,94%	18	2,81%	640	0,75%
54	firma grande/ media de servicos	2	2,11%	-	0,00%	88	92,63%	5	5,26%	95	0,11%
55	empresa media industria/comercio	1	0,63%	-	0,00%	154	96,25%	5	3,13%	160	0,19%
56	grande empresa industria/comercio	-	0,00%	-	0,00%	23	100,00%	-	0,00%	23	0,03%
57	agricultor arrendatario de terra	3	10,00%	-	0,00%	26	86,67%	1	3,33%	30	0,04%
58	agricultor proprietario de terra	5	38,46%	-	0,00%	8	61,54%	-	0,00%	13	0,02%
59	grande fazendeiro	-	-	-	-	-	-	-	-	-	0,00%
		29.868	34,94%	20.857	24,40%	30.870	36,11%	3.882	4,54%	85.477	100,00%

Tabela 5.7 - Ocupação Principal e as Variáveis de Agrupamento "Estuda" e "Trabalha"

A descrição dos quatro grupos da tabela é apresentada na Tabela 5.8.

Tabela 5.8 – Grupos de Indivíduos pelas variáveis “estuda” e “trabalha”

Grupo	Descrição do Grupo
N TRAB/N EST	Não trabalham e não estudam
N TRAB/ EST	Estudam
TRAB/ N EST	Trabalham
TRAB/ EST	Trabalham e estudam

Os resultados mostram que é possível adotar as variáveis “estuda” e “trabalha” como variáveis que possam caracterizar os padrões de viagens.

Observa-se na Tabela 5.7 que 99,33% dos classificados no grupo “N TRAB/EST” são estudantes. Uma pequena porcentagem (0,67%) encontrada nos demais grupos pode corresponder aos erros de digitação ou aos erros durante a entrevista.

Os quatro grupos são compostos basicamente por ocupações principais com características semelhantes. O grupo de não trabalhadores e não estudantes (N TRAB/N EST) são formados por prendas domésticas (2), aposentados (3), sem ocupação (4), desempregados (5), e em licença (6), isto é, ocupações em que não se caracteriza diretamente o estudo ou o trabalho em si. No entanto, no caso das “Prendas Domésticas” observa-se uma pequena porcentagem em “TRAB/ N EST” (9,89%), que pode ser devido às interpretações diferentes, como a “dona de casa” ou a “empregada doméstica”. Também se encontram em “N TRAB/EST”, pessoas que declararam sem ocupação, mas que foram para a escola (16,90%). Isto pode significar que estas pessoas, mesmo sem ocupação, possam estar estudando ou levando os filhos na escola, por exemplo.

A forma de codificação do padrão de viagem refere-se basicamente ao agrupamento dos motivos de viagem em forma de letras (Tabela 5.9) e à apresentação do sequenciamento das mesmas representando o encadeamento da viagem realizada por um determinado indivíduo.

Tabela 5.9 - Agrupamento dos motivos de viagens

Residência (H)	Trabalho (W)	Escola (S)	Outras Atividades (A)
Residência (9)	Indústria (1) Comércio (2) Serviços (3)	Escola ou Educação (4)	Compras (5) Negócios (6) Médico/ Dentista Saúde (7) Recreação ou Visitas (8)

Números de 1 a 9 – Enumeração adotada pelo METRÔ-SP/87

Os motivos de viagens considerados para a caracterização dos padrões foram agrupados em quatro categorias principais (Tabela 6.7): Residência (H), Trabalho (W), Escola (S) e outras atividades (A). O Trabalho (W) corresponde ao trabalho na indústria, comércio ou serviços. A Escola (S) corresponde a viagens realizadas à escola e Outras Atividades (A) corresponde a viagens relacionadas ao lazer, visitas sociais, visitas ao médico ou compras.

Para classificação dos dados, foi destacado o motivo “Escola” (S) como um elemento de análise, na medida em que estas viagens podem ser consideradas tão regulares quanto às viagens ao trabalho. Viagens por motivo compras, visita ao médico, visitas sociais e viagem por lazer, ao contrário do motivo trabalho ou estudo, ocorrem aleatoriamente e apresentam maior dificuldade na sua previsão.

Os padrões de viagens considerados neste trabalho levam em conta apenas o motivo da viagem realizada em cada deslocamento. Outras características não foram consideradas porque o objetivo do trabalho é analisar a viabilidade de aplicar o modelo de árvore às viagens codificadas.

5.5 Caracterização dos Padrões de viagens

O aumento do número de viagens realizadas leva a um acréscimo significativo de combinações possíveis. Em função da limitação do modelo utilizado para geração da árvore de decisão, foram considerados apenas os indivíduos que tivessem realizado zero, dois três ou quatro viagens.

A Tabela 5.10. apresenta o número de combinações de viagens em função dos motivos e do número de viagens realizadas. A terceira e a quarta coluna apresentam os padrões possíveis e aqueles com a origem e destino na residência. Aqueles que realizam duas viagens, com origem e o destino final na residência, pode-se fazê-las de três formas: HWH, HSH, HAH.

Para indivíduos que realizam três viagens, a combinação dos motivos H, W, S e A formam 256 padrões diferentes, dentre os quais somente nove possuem origem e destino final na residência: HWWH, HAWH, HSWH, HWAH, HWSH, HSSH, HSAH, HASH, HAAH. Da mesma forma, para realização de quatro viagens, 45 padrões possuem origem e destino final na residência, no entanto destes somente 13 padrões foram encontrados a partir da amostra. O total de padrões analisados neste trabalho corresponde a 26 padrões, incluindo o padrão H, isto é, indivíduos que não realizam viagens.

Tabela 5.10 - Padrões de viagens possíveis

Número de Viagens	População	Padrões Possíveis	Padrões H_***_H	Padrões Analisados
0		1	1	1
2	45.368	4 ³	3	3
3	4.077	4 ⁴	9	9
4	8.324	4 ⁵	45	13
5	1.390	4 ⁶	216	-
.
.
.
PADRÕES ANALISADOS				26

* O ANEXO C ilustra os padrões H_***_H de duas, três, quatro e cinco viagens

Os padrões a serem analisados foram definidos em função da sua porcentagem em relação à amostra. A Figura 5.6. apresenta os padrões de viagens utilizados neste trabalho, de acordo com as quatro classes de atividades: permanecer na residência (H), viagens ao trabalho (W), escola (S) e outras atividades (A).




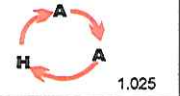
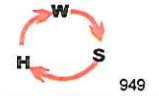
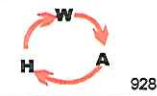

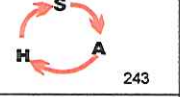
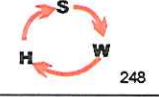
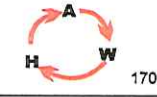
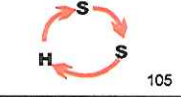

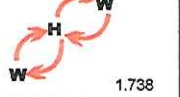

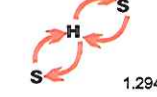
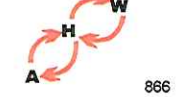
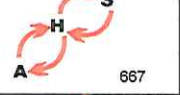
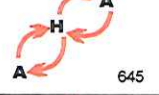
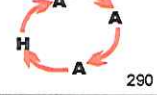

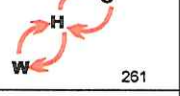
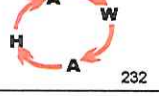
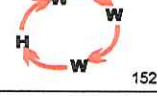
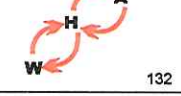
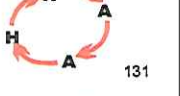
Número de Viagens	PADRÕES DE VIAGENS CONSIDERADOS PARA ANÁLISE				TIPO DE VIAGEM
0 27.708	27.708				SEM VIAGENS
2 45.368	 20.795	 17.262	 7.311		VIAGENS SIMPLS
3 4.077	 1.025	 949	 928	 336	VIAGENS COMPLEXAS
	 243	 248	 170	 105	
	 73				
4 8.324	 1.738	 1.658	 1.294	 866	
	 667	 645	 290	 298	
	 261	 232	 152	 132	
	 131				

Figura 5.6 – Padrões Utilizados na análise

Os padrões utilizados na análise são apresentados também na tabela abaixo (Tabela 5.11). As frequências correspondentes estão em ordem decrescente. Conforme estes resultados, 32,42% dos indivíduos não viajam, o que corresponde ao padrão mais predominante do conjunto de dados.

Tabela 5.11 – Frequência dos padrões utilizados na análise

PADRÃO	FREQUÊNCIA	%
H	27.708	32,42%
HWH	20.795	24,33%
HSH	17.262	20,19%
HAH	7.311	8,55%
HWHWH	1.738	2,03%
HWHS	1.658	1,94%
HSHS	1.254	1,47%
HAAH	1.025	1,20%
HWSH	949	1,11%
HWAH	928	1,09%
HWHAH	866	1,01%
SHAH	667	0,78%
HAHAH	645	0,75%
HWWH	336	0,39%
HAHSH	298	0,35%
HAAAH	290	0,34%
SHHWH	261	0,31%
HSWH	248	0,29%
HSAH	243	0,28%
HWAWH	232	0,27%
HAWH	170	0,20%
HWWWH	152	0,18%
HAHWH	132	0,15%
HWAAH	131	0,15%
HSSH	105	0,12%
HASH	73	0,09%
	85.477	100,00%

6. APLICAÇÃO DA ÁRVORE DE DECISÃO E RESULTADOS DO PROCESSAMENTO

Este capítulo apresenta as variáveis utilizadas para a geração da árvore através do modelo de árvore do S-Plus 2000. Os critérios adotados e a forma de interpretação dos resultados gráficos e do relatório são também apresentados.

Basicamente toda a análise foi realizada com a geração da árvore utilizando as seguintes variáveis sócio-econômicas: renda familiar, renda individual, número de automóveis, situação familiar, total de pessoas na família, grau de instrução, idade do indivíduo, se o indivíduo estuda ou não e se ele trabalha ou não. A representação das variáveis é descrita conforme a Tabela 6.1. e os valores referentes a estas variáveis são apresentadas no ANEXO B.

Tabela 6.1 – Código das variáveis utilizadas e representadas na árvore de classificação

CÓDIGO	DESCRIÇÃO
RF	Renda Familiar
SMED.SM	Salário Médio Individual (sal. Mínimos)
AUT	Número de Automóveis
SF	Situação Familiar
TTPESEF	Total de Pessoas na Família
SEXO	Sexo do Indivíduo
GI	Grau de Instrução
IDADE	Idade
E	Estuda
T	Trabalha

A variável dependente considerada foi o padrão de viagem, conforme codificação apresentada no capítulo anterior.

Em resumo, a árvore A-1 foi gerada a partir dos seguintes dados e critérios:

Características do Banco de Dados:

Nome do Arquivo:	Zdtq.vg.padrao.total.zona.sdd
Número de Entrevistados:	85.447 dados de indivíduos
Tamanho do Arquivo:	20.896.334 bytes

Dados de Entrada:

Nome do Arquivo:	Zdtq.vg.padrao.total.zona.sdd
Método:	Tree Model
Variável Dependente:	PADRAO
Variáveis Independentes:	RF ;SEXO; SF; SMED; IDADE; TTPESF; ESTUDA; TRABALHA; GI; AUTO
Número Mínimo de Observações:	25
Tamanho mínimo do nó:	50
Desvio Mínimo do nó:	0,05
Tipo de Saída Gráfica:	Árvore uniforme

Processamento:

Tempo Calculado de Processamento:	Máximo 1 min
-----------------------------------	--------------

Processador:

Pentium III 600 Mhz – 128 memória RAM e HD 30GB

Apresentação dos Resultados:

Em formato de Árvore:	"Graph." (gráfico)
Em formato de Relatório:	"Report." (relatório)

O relatório do S-Plus 2000 ("Report") apresenta, na forma literal, os detalhes do desenvolvimento da árvore e gráfico ("Graph") através das variáveis em cada nó intermediário. O padrão predominante é apresentado em cada nó terminal.

6.1 Interpretação do Relatório

Parte do relatório da árvore A-1 é apresentado a seguir com a finalidade de descrever alguns detalhes da sua interpretação. O relatório completo é encontrado no ANEXO D.

Para facilitar a compreensão e a explicação dos resultados apresentados, cada linha foi enumerada. As primeiras linhas (1 a 17) correspondem ao cabeçalho do relatório contendo as variáveis adotadas para análise e as variáveis utilizadas na construção da árvore, o número de nós terminais, o desvio residual médio e a taxa de erros de classificação.

```

1      *** Tree Model ***
2
3      Classification tree:
4      tree(formula = PADRAO ~ AUT + RF.SM. + TTPESF + SEXO + SF + IDADE
5      + GI +
6      SMED.SM. + ESTUDA + TRABALHA, data =
7      ZDTQ.VGS.PADRAO.limpo.e.completo,
8      na.action = na.exclude, mincut = 25, minsize = 50, mindev =
9      0.05)
10     Variables actually used in tree construction:
11     [1] "TRABALHA" "ESTUDA" "SF" "TTPESF" "IDADE" "GI"
12     [7] "AUT" "SMED.SM." "SEXO"
13     Number of terminal nodes: 20
14     Residual mean deviance: 2.177 = 186100 / 85460
15     Misclassification error rate: 0.2914 = 24906 / 85477

```

As linhas subsequentes correspondem aos resultados de cada nó da árvore:

```

16     node), split, n, deviance, yval, (yprob)
17     * denotes terminal node
18     1) root 85477 328100 H ( 0.32420 0.0033930 0.011990 0.08553
19     0.0075460 0.0034860 0.00154400 0.0008540 0.0019890 0.00284300 0.201900
20     0.0078030 0.0146700 0.00305300 0.00122800 0.00290100 0.00153300
21     0.0108600 0.00271400 0.2433000 0.0101300 0.01940000 0.0203300
22     0.01110000 0.00393100 0.0017780 )
23     2) TRABALHA<0.5 50725 136200 H ( 0.47340 0.0040410 0.014530
24     0.11180 0.0086550 0.0049480 0.00003943 0.0012420 0.0000000 0.00457400
25     0.333300 0.0126000 0.0227100 0.00027600 0.00195200 0.00019710
26     0.00005914 0.0002563 0.00007886 0.0039630 0.0002366 0.00037460
27     0.0002563 0.00043370 0.00001971 0.0000000 )
28     4) ESTUDA<0.5 29868 56050 H ( 0.73040 0.0061940 0.022570
29     0.17680 0.0138600 0.0012720 0.00000000 0.0006696 0.0000000 0.00100400
30     0.013760 0.0021090 0.0236700 0.00006696 0.00016740 0.00006696
31     0.00010040 0.0004352 0.00006696 0.0056920 0.0004018 0.00010040
32     0.0004352 0.00003348 0.00003348 0.0000000 )
33     8) SF<2.5 18605 41280 H ( 0.65050 0.0075790 0.028490 0.22760

```

```

34 0.0190800 0.0018810 0.00000000 0.0009675 0.0000000 0.00150500 0.018010
35 0.0029020 0.0339700 0.00010750 0.00021500 0.00005375 0.00010750
36 0.0003762 0.00010750 0.0052140 0.0006450 0.00010750 0.0005375
37 0.00005375 0.00005375 0.0000000 )
38      16) TTPESF<2.5 2915 6607 H ( 0.48340 0.0116600 0.052140
39 0.40070 0.0377400 0.0000000 0.00000000 0.0000000 0.0000000 0.00034310
40 0.002058 0.0000000 0.0017150 0.00000000 0.00000000 0.00000000
41 0.00034310 0.0010290 0.00000000 0.0065180 0.0006861 0.00034310
42 0.0013720 0.00000000 0.00000000 0.0000000 ) *
43 ....

```

Através desta parte do relatório da árvore apresentada acima, descreve-se o significado de cada informação apresentada.

O objetivo da enumeração dos nós, no caso no S-Plus 2000, é permitir sua rápida identificação. Os números dos nós apresentados referem-se a uma árvore completa. Isto significa que, para um mesmo conjunto de dados, a geração de árvores com variações no desvio ou no número de observações não modifica a enumeração dos nós referentes à árvore completa, apresentando assim, números de nós correspondentes entre si.

O primeiro número em cada conjunto de linhas (linhas 18 a 22, por exemplo) corresponde ao número do nó.

A linha 23, conforme apresentada na descrição abaixo, corresponde ao nó 2. Ela e os demais conjuntos de linhas podem ser lidos da seguinte forma:

2)	TRABALHA<0.5	50725	136200	H	(0.47340,)	
Número do nó	Variável	Número de Observações	Desvio	Padrão Predominante	Probabilidade Dos Padrões	Tipo De nó

O texto subsequente ao nó 2 (número do nó) corresponde a sua primeira divisão (TRABALHA<0,5). Em seguida, o número 50.725 refere-se ao número de observações. O próximo número, 136.200, corresponde ao desvio, ou seja, à medida de heterogeneidade do nó utilizada no algoritmo da árvore de decisão do S-Plus 2000. O resultado do nó 2 é H, o que significa que o padrão predominante no nó 2 é permanecer em casa. O padrão predominante é apresentado no gráfico da árvore de decisão. Os números subsequentes apresentados "(0.47340 0.0040410 0.014530 0.11180 0.0086550 0.0049480 0.00003943 0.0012420 0.0000000 0.00457400 0.333300 0.0126000 0.0227100 0.00027600 0.00195200 0.00019710 0.00005914 0.0002563 0.00007886 0.0039630 0.0002366 0.00037460 0.0002563 0.00043370 0.00001971 0.0000000)" correspondem à distribuição de probabilidade dos distintos padrões de viagem para os

indivíduos classificados neste nó. A seqüência dos números corresponde à ordem alfabética dos padrões (Tabela 6.2). O tipo de nó corresponde ao nó terminal que é indicado por um asterisco (*).

Tabela 6.2 – Padrões de Viagens em ordem alfabética apresentados no Relatório

ORDEM	PADRÃO	ORDEM	PADRÃO
1	H	14	HSHWH
2	HAAA	15	HSSH
3	HAAH	16	HSWH
4	HAH	17	HWAH
5	HAHAH	18	HWAH
6	HAHSH	19	HWAH
7	HAHWH	20	HWH
8	HASH	21	HWH
9	HAWH	22	HWH
10	HSAH	23	HWH
11	HS	24	HWH
12	HS	25	HWH
13	HS	26	HWH

Basicamente, para o desenvolvimento da árvore de decisão, o algoritmo de partição binária do modelo de árvore do S-Plus 2000 divide recursivamente os dados em cada nó até que ele se torne homogêneo ou que ele contenha poucas observações. Nós com completa homogeneidade apresentam desvio zero. Quanto ao número de observações, adotou-se neste trabalho, o mínimo de 25 para cada nó terminal.

Observa-se que com os critérios adotados, todas as variáveis, com exceção da renda familiar, foram utilizadas para a geração da árvore.

A partir do total de 85.477 observações, as divisões se iniciam considerando a variável de maior importância “TRABALHA”. Dois grupos são subdivididos, de 50.725 indivíduos que não trabalham e de 34.752 indivíduos que trabalham (nó 2 e nó 3). Dos 50.725 indivíduos, a probabilidade de ocorrer o padrão H foi de 0,4734. Para os 34.752 indivíduos a probabilidade de ocorrer HWH foi de 0,5926.

A partir do nó 2 formaram-se dois grupos em função da variável “ESTUDA”: 29.868 e 20.857 indivíduos, entre aqueles que não estudam – $ESTUDA < 0,5$ (nó 4) ou estudam – $ESTUDA > 0,5$ (nó 6), respectivamente.

O grupo observado no nó 4 foi subdividido de acordo com a situação familiar (SF): grupos com cônjuge ou chefe – $SF < 2,5$ (nó 8) e grupos de filhos, parentes ou outros – $SF > 2,5$ (nó 9).

O procedimento de divisões continua até que o número de casos no nó seja menor que o valor mínimo adotado (no caso, 25 observações) ou quando o nó se tornar razoavelmente homogêneo, isto é, o desvio no nó for menor que 0,05.

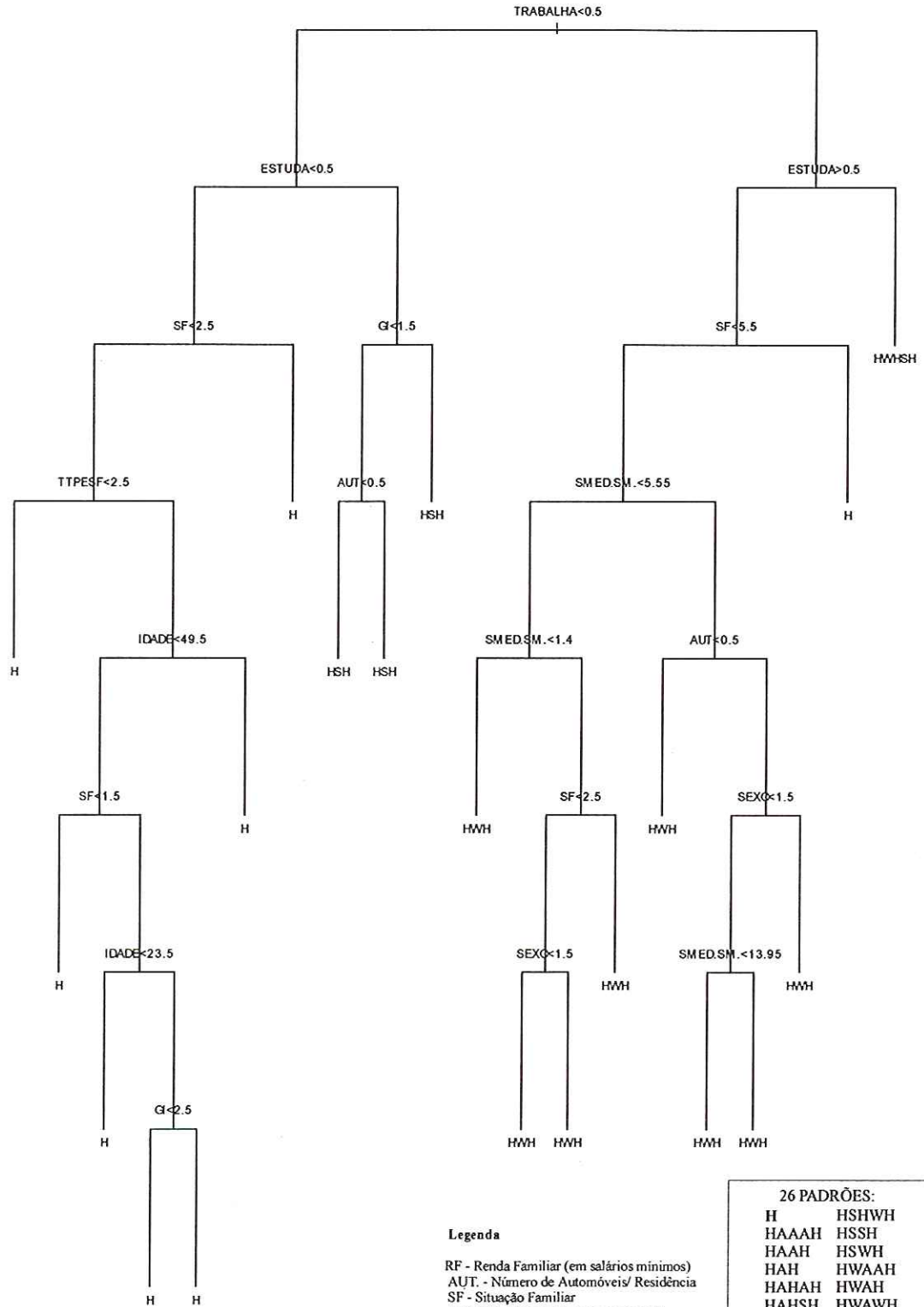
O asterisco significa um nó terminal, isto é, ele não se subdivide. Os nós terminais correspondem aos seguintes nós: 16, 68, 138, 278, 279, 35, 9, 20, 21, 11, 48, 196, 197, 99, 50, 204, 205, 103, 13, 7.

6.2 Interpretação do Gráfico

A Figura 6.1 apresenta o resultado da árvore em forma de gráfico e as ramificações da árvore, com as variáveis utilizadas em cada nó. O nó terminal representa o padrão predominante de cada conjunto de características.

Como definição utilizada neste trabalho, o “grupo” corresponde ao conjunto de valores das variáveis sócio-econômicas que caracterizam um indivíduo ou um conjunto de indivíduos. As variáveis podem estar sub-classificadas de acordo com a sua natureza, “SEXO” (Homem ou Mulher) por exemplo.

Figura 6.1 – A-1 com desvio de 0,05



Tree Model - Classification Tree- S-Plus
 Variables: PADRÃO ~AUT+RF+TTPESF+SEXO+SF+IDADE+GI+SMED+E+T
 Variables used in tree construction: T+E+RF+GI+SMED+SF+SEXO
 Number of terminal nodes: 20
 Mincut=25, minsize=50, mindev=0.05

Legenda

RF - Renda Familiar (em salários mínimos)
 AUT - Número de Automóveis/ Residência
 SF - Situação Familiar
 TTPESF - Total de Pessoas na Família
 GI - Grau de Instrução
 IDADE - Idade
 SMED.SM. - Salário médio Individual (Em salários mínimos)
 ESTUDA - Estuda
 TRABALHA - Trabalha

26 PADRÕES:	
H	HSHWH
HAAA	HSSH
HAAH	HSWH
HAH	HWAAH
HAHAH	HWAH
HAHSH	HWAWH
HAHWH	HWH
HASH	HWHAH
HAWH	HWHSH
HSAH	HWHWH
HSH	HWSH
HSHAH	HWWH
HSHSH	HWWWH

A árvore da Figura 6.1 gerada pelo S-Plus 2000 apresenta em seus nós terminais os padrões de viagem. O modelo ainda permite apresentar, na forma gráfica, o desvio do nó ou o número de observações nos nós terminais.

A árvore apresentou 20 nós terminais a partir do conjunto de dados, isto é, identificou 20 grupos de indivíduos caracterizados segundo os critérios “desvio” e “quantidade mínima no nó” especificados.

A leitura através do gráfico leva a uma certa dificuldade de interpretação, pois nem sempre os valores de cada variável podem representar de forma direta as características do indivíduo. Muitas vezes, as restrições referentes a uma variável são definidas com o uso de vários níveis, como por exemplo: TRABALHA < 0.5; ESTUDA < 0.5; SF < 2.5; TTPESF > 2.5; IDADE < 49,5; SF > 1,5; IDADE < 23,5 (Figura 6.2).

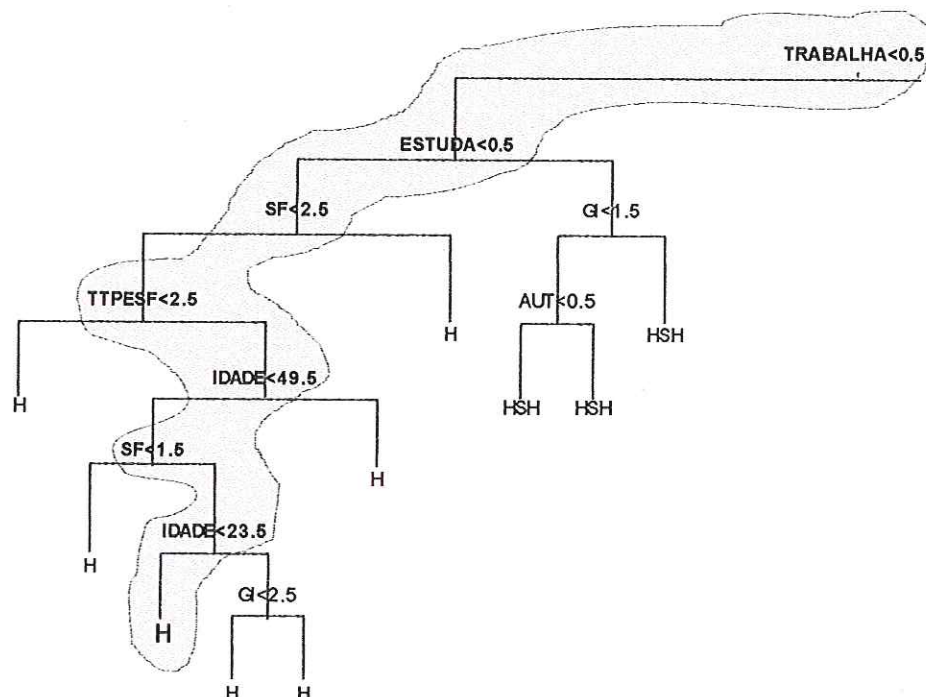


Figura 6.2 – Exemplo das variáveis correspondentes a um grupo

Outra dificuldade é que a árvore apresenta os padrões predominantes, passando para o leitor a idéia de que todos os indivíduos pertencentes a um grupo sócio-econômico realizam mesmo padrão de viagem (Figura 6.3), o que não ocorre. Apesar dos indivíduos

apresentarem mesmas características sócio-econômicas e de viagens, eles podem fazer diferentes padrões de viagens encadeadas.

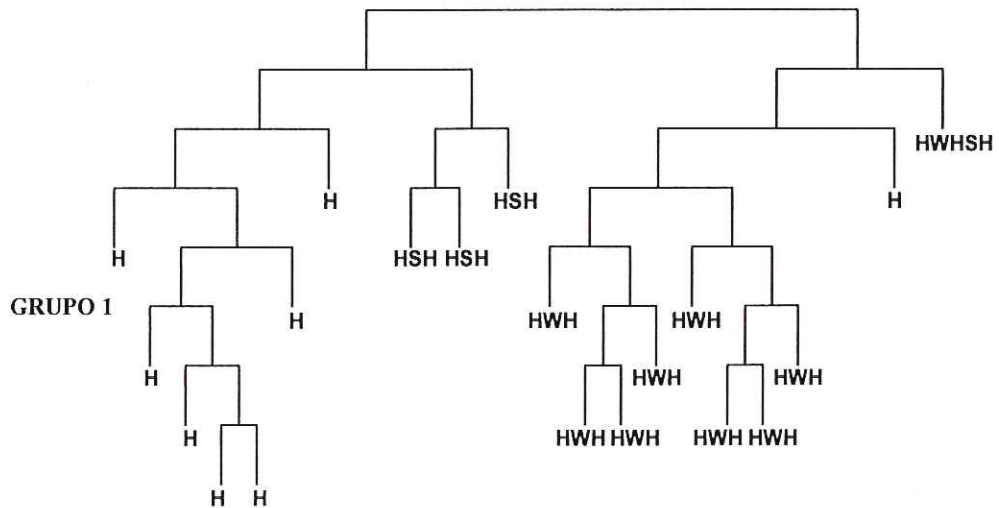


Figura 6.3 – Padrões predominantes nos nós da árvore

A utilização do gráfico, sem a observação dos resultados apresentados no relatório, pode levar a interpretações incompletas. Por exemplo, o Grupo 1 (Figura 6.3) apresentou padrão predominante H, passando a impressão de que todos os indivíduos com as características referentes a este grupo não viajam. No entanto, numa análise a partir do relatório, observa-se que o padrão predominante (H) corresponde a 48,34% do grupo e os demais indivíduos, viajaram realizando outros tipos de padrões de viagem. Isto mostra a importância da análise através dos dados do relatório, pois ele apresenta a probabilidade de ocorrência de cada padrão de viagem para um determinado grupo sócio-econômico em qualquer nível da árvore, conseqüentemente enriquecendo a análise do comportamento do viajante.

Em virtude disso e, principalmente pela riqueza de informações apresentadas no relatório, este trabalho utilizou basicamente o resultado do relatório para análise deste trabalho.

Este trabalho adotou como 0,05 como desvio mínimo. Para se ter uma idéia dos resultados encontrados através da variação do desvio, o ANEXO E apresenta uma árvore com desvio mínimo adotado de 0,01(A-2), e seu respectivo relatório.

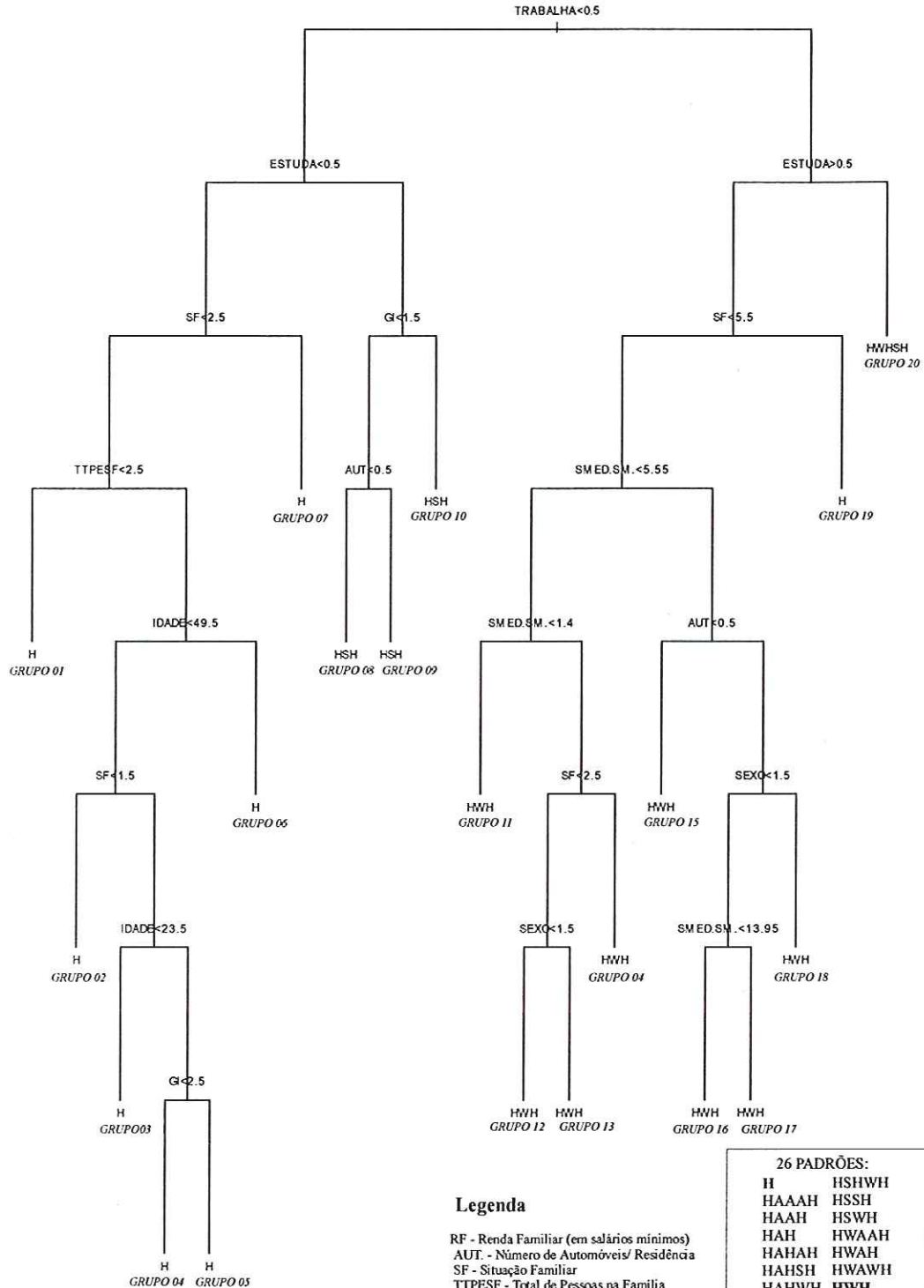
CAPÍTULO 7 - ANÁLISE DOS RESULTADOS OBTIDOS

Os resultados obtidos através da utilização da árvore de decisão são apresentados e confrontados com as pesquisas encontradas na literatura. Estas comparações permitem reforçar a viabilidade de se aplicar o minerador de dados na obtenção das relações entre padrões de viagens e condições sócio-econômicas. Todavia, é preciso ressaltar que o objetivo da comparação não foi o de assegurar que os resultados sejam semelhantes, mesmo porque existem diferenças entre populações, seus hábitos, preferências e restrições que refletem na formação de padrões de viagem. Pretende-se apenas verificar a consistência e a lógica dos resultados.

A árvore de decisão apresentou 20 nós terminais a partir do conjunto de dados, isto é, identificou 20 grupos de indivíduos caracterizados segundo os critérios de “desvio” e “quantidade mínima no nó” adotados. Todas as variáveis foram utilizadas para o processamento do algoritmo. A “Renda Familiar” foi a única variável que não apareceu nos resultados da árvore, o que significa que ela não está entre as variáveis mais relevantes para o caso.

A Figura 7.1 apresenta a árvore de decisão com os padrões predominantes e a identificação dos grupos de indivíduos de acordo com as características socioeconômicas.

Figura 7.1. – Grupos formados através da Arvore de Decisão



Tree Model - Classification Tree- S-Plus
 Variables: PADRÃO ~AUT+RF+TTPESF+SEXO+SF+IDADE+GI+SMED+E+T
 Variables used in tree construction: T+E+RF+GI+SMED+SF+SEXO
 Number of terminal nodes: 20
 Mincut=2.5, minsize=50, mindev=0.05

Legenda

RF - Renda Familiar (em salários mínimos)
 AUT - Número de Automóveis/ Residência
 SF - Situação Familiar
 TTPESF - Total de Pessoas na Família
 GI - Grau de Instrução
 SEXO - Sexo
 IDADE - Idade
 SMED.SM - Salário médio Individual (Em salários mínimos)
 ESTUDA - Estuda
 TRABALHA - Trabalha

26 PADRÕES:	
H	HSHWH
HAAAH	HSSH
HAAH	HSWH
HAH	HWA AH
HAHAH	HWAH
HAHSH	HWAWH
HAHWH	HWH
HASH	HWHAH
HAWH	HWHSH
HSAH	HWHWH
HSH	HWSH
HSHAH	HWWH
HSHSH	HWWWH

Para se ter uma idéia geral dos grupos formados através da árvore de decisão, a Tabela 7.1 resume os grupos encontrados de acordo com as características sócio-econômicas e o padrão predominante de viagem. Os grupos formados, o número de indivíduos em cada nó terminal e o padrão predominante podem ser obtidos através da visualização do gráfico ou através dos resultados do relatório.

Na Tabela 7.1, os valores correspondentes ao padrão predominante e a sua respectiva porcentagem só são obtidos através do relatório, o que mostra a importância da utilização dos dados do relatório para as análises dos resultados.

Tabela 7.1. – Resumo dos 20 grupos formados

Grupo	Características sócio-econômicas	N.º indivíduos classificados no grupo	Padrão Predom.	% da grupo em relação ao total	N.º de Indivíduos no Padrão Predominante	% Indiv.no padrão em relação ao grupo
1	Não trabalha; não estuda; Chefe ou Cônjuge; com 1 ou duas pessoas na família	2.915	H	3,41%	1.409	48,34%
2	Não trabalha; não estuda; Chefe; com 3 pessoas ou mais na família; idade menor que 49 anos	1.123	H	1,31%	664	59,13%
3	Não trabalha; não estuda; Cônjuge; com 3 pessoas ou mais na família; idade menor que 23 anos	901	H	1,05%	689	76,47%
4	Não trabalha; não estuda; Cônjuge; com 3 pessoas ou mais na família; idade entre 23 e 49 anos; Analfabeto até 1.º Grau (4.º série completa)	6.172	H	7,22%	4.272	69,22%
5	Não trabalha; não estuda; Cônjuge; com 3 pessoas ou mais na família; idade entre 23 e 49 anos; Acima do 1.º Grau completo	2.422	H	2,83%	1.460	60,28%
6	Não trabalha; não estuda; Chefe ou Cônjuge; com 3 pessoas ou mais na família; Idade maior que 49 anos	5.072	H	5,93%	3.608	71,14%
7	Não trabalha; não estuda; Filho ou Outros	11.263	H	13,18%	9.715	86,26%
8	Estuda; com 1º Grau (até 4º série); sem automóvel	7.659	HSH	8,96%	6.294	82,18%
9	Estuda; com 1º Grau (até 4º série); com automóvel	5.355	HSH	6,26%	4.256	79,48%
10	Estuda; acima do 1.º Grau completo concluído	7.843	HSH	9,18%	5.947	75,83%
11	Trabalha; Familiares; Salário médio <1.40SM	5.235	HWH	6,12%	3.208	61,28%
12	Trabalha; Chefe; Salário médio maior que 1.40SM e menor que 5;55 SM; Homem	5.742	HWH	6,72%	4.074	70,95%
13	Trabalha; Cônjuge; Salário médio maior que 1.40SM e menor que 5;55 SM; Mulher	3.267	HWH	3,82%	2.061	63,09%
14	Trabalha; Filho, Parente ou Agregado; Salário médio maior que 1.40SM e menor que 5;55 SM	5.381	HWH	6,30%	4.038	75,04%
15	Trabalha; Familiares; Salário médio maior que 5;55SM; sem automóvel na família	3.356	HWH	3,93%	2.364	70,44%
16	Trabalha; Familiares; Salário médio maior que 5;55SM e menor que 13;59SM; com 1 automóvel na família	3.299	HWH	3,86%	2.134	64,69%
17	Trabalha; Familiares; Salário médio maior que 13,95SM; com mais de automóvel na família; Homem	2.167	HWH	2,54%	1.266	58,42%
18	Trabalha; Familiares; Salário médio maior que 5;55SM; com mais de automóvel na família; Mulher	1.645	HWH	1,92%	906	55,08%
19	Trabalha; Empregada/Resid ou visitante não resid.	778	H	0,91%	651	83,68%
20	Trabalha; Estuda	3.882	HWHSH	4,54%	1.555	40,06%
		85.477		100,00%		

Obs.: "Familiares" correspondem a Chefe, Cônjuge, Filho, Parente ou Agregado
 "Outros" correspondem a Parente, Agregado, Empregada ou Visitante

Basicamente, quatro grupos distintos são observados a partir das variáveis ESTUDA e TRABALHA indicando a importância destas variáveis com os padrões de viagens dos indivíduos.

- a) Indivíduos que não trabalham e não estudam;
- b) indivíduos que somente estudam;
- c) indivíduos que somente trabalham; e
- d) indivíduos que estudam e trabalham.

Os padrões predominantes correspondentes são apresentados na Figura 7.2.

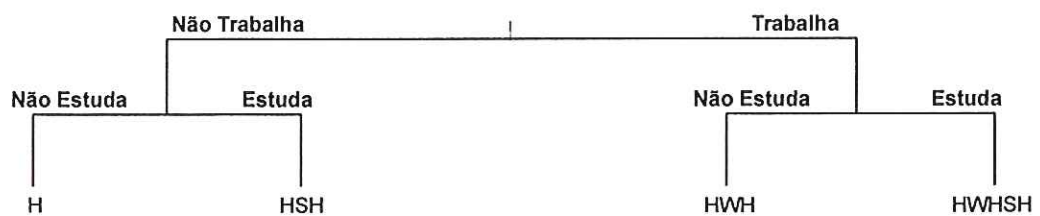


Figura 7.2 – Primeiros níveis da árvore de decisão – variáveis ESTUDA e TRABALHA

Como era de se esperar, a partir desses quatro principais grupos é possível estimar a probabilidade de ocorrência do padrão de viagem a ser realizado. As pessoas que não trabalham nem estudam não viajam na maioria das vezes (H). Indivíduos que estudam apresentam predominantemente o padrão HSH. A maioria dos indivíduos que somente trabalham apresentam o padrão HWH, e aqueles que estudam e trabalham, o padrão HWHS, na maioria das vezes. MITCHELL e RAPKIN (1954) já diziam que o exercício de certas atividades está relacionado à ocupação do indivíduo. De fato, o relatório do S-Plus 2000 mostra que a ocupação dos indivíduos determina, de certa forma, a probabilidade de ocorrência de cada padrão de viagem.

A Tabela 7.2 mostra as probabilidades destes grupos em relação ao número de viagens.

Tabela 7.2. – Probabilidades dos grupos nos primeiros níveis da árvore em relação ao número de viagens

Número de viagens	Grupo de Indivíduos			
	Não Estuda Não trabalha	Estuda	Trabalha	Estuda Trabalha
Total Obs.	29.868	20.857	30.870	3.882
% Total obs.	34,94	24,40	36,12	4,54
0	0,7304	0,1054	0,1162	0,0281
2	0,1963	0,8113	0,7051	0,2107
3	0,0250	0,0210	0,0612	0,2607
4	0,0483	0,0623	0,1175	0,5005

A partir dos quatro principais grupos observa-se nos resultados da árvore a seguinte relação entre as variáveis sócio-econômicas e os grupos de indivíduos (Tabela 7.3).

Tabela 7.3. – Variáveis de maior importância para grupos de indivíduos segundo os resultados apresentados pela árvore de decisão

Variável	Grupo de Indivíduos	Não Estuda Não trabalha	Estuda	Trabalha	Estuda Trabalha
Renda Familiar					
Salário Médio Individual				x	
Número de Automóveis			x	x	
Situação Familiar		x		x	
Total de Pessoas na Família		x			
Sexo do Indivíduo				x	
Grau de Instrução			x		
Idade		x			
Estuda			x		x
Trabalha				x	x

Diferenças significativas dos grupos concentraram-se nas variáveis sócio-econômicas identificadas na análise através da árvore de decisão que poderiam explicar os padrões de viagens. Os indivíduos que não estudam e nem trabalham são classificados basicamente através da “situação familiar”, “número de pessoas na família” e “idade”; estudantes são classificados de acordo com o número de automóveis e grau de instrução e trabalhadores são classificados de acordo com o salário médio individual, número de automóveis, situação familiar e sexo.

Uma análise detalhada dos quatro principais grupos mostra a relação entre os homens e mulheres (Tabela 7.4).

Tabela 7.4. – Distribuição de homens e mulheres nos grupos principais

Sexo	Grupo de Indivíduos			
	Não Estuda Não trabalha	Estuda	Trabalha	Estuda Trabalha
Total Obs.	29.868	20.857	30.870	3.882
% Total obs.	34,94%	24,40%	36,12%	4,54%
Homem	26,65%	48,68%	63,39%	56,25%
Mulher	73,35%	51,32%	36,61%	43,74%
Total	100,00%	100,00%	100,00%	100,00%

Em relação às populações de homens (46,64%) e mulheres (56,36%) da amostra, observa-se que os homens, na sua maioria, trabalham (49,09%), 25,47% somente estudam, 19,97% não trabalham e não estudam e 5,48% estudam e trabalham. Já 48,03% da população das mulheres não estudam e nem trabalham, 24,78% trabalham, 23,47% estudam e somente 3,72% estudam e trabalham. De acordo com o observado por MCGUKIN e MURAKAMI (1999), as mulheres, ainda na sua maioria, realizam mais viagens ligadas às atividades relacionadas à família, tais como compras ou outras atividades como levar os filhos à escola.

Estes resultados mostram que é possível obter relações entre os padrões de viagens realizados pelos indivíduos através de sua ocupação principal, que é coletada durante o recenseamento da população. Logo, parece possível fazer o planejamento do transporte urbano com base em modelos de árvore associados aos dados de recenseamento.

Os 20 grupos de indivíduos da árvore de decisão apresentados na Tabela 7.5 fornecem as probabilidades de ocorrência de todos os padrões de viagem analisados em relação às características socioeconômicas.

O grupo mais freqüentemente observado nesta tabela corresponde aos indivíduos que não trabalham e não estudam, que são filhos ou outros (Grupo 07) com 13,18% de toda a amostra e são indivíduos que predominantemente não realizam viagens (H). O segundo grupo correspondeu aos estudantes com grau de instrução acima do 1.º grau completo (Grupo 10) com padrão HSH (9,18%) e o terceiro grupo, aos estudantes com 1.º grau até a quarta série (Grupo 8) com padrão HSH (8,96%).

Tabela 7.5. – Grupos de indivíduos e padrões de viagens – Árvore A-1

	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5	GRUPO 6	GRUPO 7	GRUPO 8	GRUPO 9	GRUPO 10
1	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha	N Trabalha
2	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda
3	Chefe/Conj	Chefe	Conjuge	Conjuge	Conjuge	Chefe/Conj	Filho/Outro	Estuda	Estuda	Estuda
4	FAM<2	FAM≥3	FAM≥3	FAM≥3	FAM≥3	FAM≥3		GI=1	GI=1	GI≥2
5		Idade≤49	Idade≤23	Idade≥24	Idade≥24	Idade≥24	Idade≥50	AUTO=0	AUTO≥1	
6				Idade≤49	Idade≤49					
7				GI<2	GI≥3					
FREQ	2915	1123	901	6172	2422	5072	11263	7659	5355	7843
DESVIO	6607	2828	1380	13090	6388	9096	12920	10660	9127	15930
%	3,41%	1,31%	1,05%	7,22%	2,83%	5,93%	13,18%	8,96%	6,26%	9,18%
PADRÕES	PROBABILIDADES									
H	0,48340	0,59130	0,76470	0,69220	0,60280	0,71140	0,86260	0,12080	0,10140	0,09282
HAAA	0,01166	0,01603	0,00555	0,00502	0,01156	0,00493	0,00391	0,00065	0,00131	0,00102
HAAH	0,05214	0,04363	0,01665	0,01977	0,02808	0,02445	0,01279	0,00248	0,00205	0,00421
HAH	0,40070	0,25560	0,18310	0,17350	0,18580	0,21550	0,09305	0,01528	0,01942	0,02168
HAHAH	0,03774	0,02137	0,00999	0,01053	0,01569	0,02149	0,00524	0,00065	0,00019	0,00242
HAHSH	-	0,00267	0,00111	0,00275	0,00496	0,00039	0,00027	0,00261	0,00635	0,02027
HAHWH	-	-	-	-	-	-	-	-	-	0,00026
HASH	-	0,00089	-	0,00081	0,00330	0,00079	0,00018	0,00091	0,00205	0,00319
HAWH	-	-	-	-	-	-	-	-	-	-
HSAH	0,00034	0,00089	0,00111	0,00211	0,00330	0,00079	0,00018	0,00496	0,01438	0,01109
HSH	0,00206	0,01069	0,00333	0,02916	0,04088	0,00690	0,00675	0,82180	0,79480	0,75830
HSHAH	-	0,00267	0,00111	0,00454	0,00785	0,00059	0,00080	0,01436	0,03716	0,03404
HSHSH	0,00172	0,02137	0,01110	0,05541	0,09125	0,00592	0,00666	0,00992	0,01046	0,03991
HSHWH	-	-	-	0,00016	-	0,00020	-	0,00052	0,00056	0,00064
HSSH	-	-	-	0,00016	0,00124	-	0,00009	0,00261	0,00579	0,00548
HSWH	-	-	-	-	0,00041	-	0,00009	-	0,00131	0,00013
HWAAH	0,00034	-	-	0,00016	-	-	0,00009	-	-	-
HWAH	0,00103	0,00089	-	0,00016	0,00041	0,00020	0,00053	-	-	-
HWAWH	-	0,00089	0,00111	-	-	-	-	-	0,00019	0,00013
HWH	0,00652	0,02493	0,00111	0,00324	0,00124	0,00513	0,00648	0,00196	0,00037	0,00179
HWHAH	0,00069	0,00445	-	0,00016	-	0,00079	-	-	-	-
HWHSH	0,00034	-	-	-	0,00041	-	0,00009	0,00013	0,00056	0,00153
HWHWH	0,00137	0,00178	-	0,00016	0,00041	0,00039	0,00027	-	-	-
HWSH	-	-	-	-	-	0,00020	-	0,00039	0,00168	0,00115
HWWH	-	-	-	-	0,00041	-	-	-	-	-
HWWWH	-	-	-	-	-	-	-	-	-	-

Tabela 7.5. – Grupos de indivíduos e padrões de viagens – Árvore A-1 – continuação

	GRUPO 11	GRUPO 12	GRUPO 13	GRUPO 14	GRUPO 15	GRUPO 16	GRUPO 17	GRUPO 18	GRUPO 19	GRUPO 20
1	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha	Trabalha
2	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	N Estuda	Trabalha
3	SF≤5	Chefe/Conj.	Chefe/Conj.	Filho/Outro	SF≤5	SF≤5	SF≤5	SF≤5	SF≥6	Estuda
4		Smedio.<5,55	Smedio.<5,55	Smedio.<5,55	Smedio.>5,55	Smedio.>5,55	Smedio.>13,95	Smedio.>5,55		
5	Smedio.<1,4	Smedio.>1,4	Smedio.>1,4	Smedio.>1,4	AUTO=0	Smedio.<13,95	AUTO≥1	AUTO≥1		
6		Homem	Mulher			AUTO≥1	Homem	Mulher		
7						Homem				
FREQ	5235	5742	3267	5381	3356	3299	2167	1645	778	3882
DESVIO	14740	14140	9502	11660	8797	10000	7590	5884	1099	14640
%	6,12%	6,72%	3,82%	6,30%	3,93%	3,86%	2,54%	1,92%	0,91%	4,54%
PADRÕES	PROBABILIDADES									
H	0,16830	0,08063	0,13680	0,07564	0,07151	0,06275	0,05538	0,10330	0,83680	0,02808
HAAA	0,00210	0,00157	0,00276	0,00130	0,00477	0,00485	0,00508	0,00304	-	0,00026
HA	0,00936	0,00749	0,01408	0,00260	0,01043	0,01122	0,01015	0,01398	0,00900	0,00309
HAH	0,06991	0,05904	0,06550	0,02769	0,04559	0,04577	0,03092	0,05471	0,08355	0,01133
HAHA	0,00573	0,00627	0,00765	0,00427	0,00566	0,00849	0,00738	0,01094	0,00514	0,00180
HAHSH	0,00057	0,00035	0,00061	-	-	0,00121	0,00092	0,00182	0,00129	0,00773
HAHWH	0,00439	0,00383	0,00429	0,00149	0,00477	0,00515	0,00785	0,00608	-	0,00077
HA	-	-	0,00092	-	0,00030	-	-	0,00061	-	0,00129
HAWH	0,00287	0,00523	0,00459	0,00372	0,00656	0,00667	0,01384	0,00730	0,00129	0,00077
HSAH	0,00038	-	-	0,00019	-	0,00030	-	0,00061	-	0,00155
HSH	0,00420	0,00244	0,00612	0,00130	0,00149	0,00303	0,00369	0,00547	0,01928	0,06285
HSHA	0,00057	0,00035	0,00122	-	0,00089	-	0,00092	0,00243	0,00386	0,00180
HSHSH	0,00573	0,00105	0,00429	0,00093	0,00119	0,00030	0,00138	0,00243	0,02057	0,00489
HSHWH	0,00191	0,00035	0,00153	0,00074	0,00060	0,00424	0,00462	0,00730	-	0,04843
HSSH	-	-	0,00031	-	-	-	-	-	-	0,00129
HSWH	0,00459	0,00192	0,00459	0,00037	0,00119	0,00940	0,02353	0,00790	0,00129	0,02215
HWA	0,00191	0,00244	0,00429	0,00372	0,00477	0,00364	0,01061	0,00669	-	0,00206
HWAH	0,01662	0,02125	0,03061	0,02918	0,03844	0,03668	0,04569	0,04255	-	0,00773
HWAHSH	0,00554	0,00523	0,00306	0,00520	0,00566	0,01303	0,02215	0,00790	-	0,00206
HWH	0,61280	0,70950	0,63090	0,75040	0,70440	0,64690	0,58420	0,55080	0,01671	0,13650
HWH	0,01471	0,01829	0,02081	0,02472	0,02354	0,04698	0,05261	0,04681	0,00129	0,01159
HWH	0,00115	0,00209	0,00367	0,00260	0,00387	0,00455	0,00231	0,00426	-	0,40060
HWHSH	0,05788	0,05486	0,03887	0,05408	0,03725	0,06244	0,07568	0,07842	-	0,01674
HWSH	0,00191	0,00105	0,00429	0,00149	0,00358	0,00121	0,00415	0,00730	-	0,21950
HWWH	0,00439	0,01045	0,00765	0,00595	0,01698	0,01182	0,02307	0,02188	-	0,00335
HWWW	0,00248	0,00435	0,00061	0,00242	0,00656	0,00940	0,01384	0,00547	-	0,00180

Os grupos de 1 a 7 correspondem a indivíduos, na sua maioria, que não realizam viagens (H) e indivíduos que realizam viagens relacionadas a outras atividades (HAH ou HAAH). Os estudantes correspondem aos Grupos 8, 9 e 10 e os Grupos 11 a 19 correspondem aos trabalhadores. Indivíduos que trabalham e estudam correspondem ao Grupo 20.

Observa-se algumas variações no comportamento dos indivíduos em relação ao tamanho da família (Tabela 7.6). As pessoas que não trabalham nem estudam, chefe ou cônjuge, pertencentes a famílias com uma ou duas pessoas correspondem a 3,41% da amostra (Grupo 1). Já essas mesmas, quando pertencentes a famílias com três ou mais pessoas correspondem a 18,36% da amostra (Grupos 2 a 6).

Tabela 7.6. – Grupo 1 e Grupos 2 a 6 - Não estudam nem trabalham, cônjuge ou chefe
Comparação entre número de pessoas na família

Não estudam Não trabalham Chefe ou Cônjuge	Famílias com 1 ou 2 pessoas (Grupo 1)	Famílias com 3 ou mais pessoas (Grupos de 2 a 6)
Total de Observações	2.915	15.690
% Total obs. amostra	3,41%	18,36%
H	0,4834	0,6815
HAAAH	0,0117	0,0068
HAAH	0,0521	0,0241
HAH	0,4007	0,1954
HAHAH	0,0378	0,0156
HAHSH	-	0,0022
HAHWH	-	-
HASH	-	0,0011
HAWH	-	-
HSAH	0,0003	0,0017
HSH	0,0021	0,0210
HSHAH	-	0,0034
HSHSH	0,0017	0,0399
HSHWH	-	0,0001
HSSH	-	0,0003
HSWH	-	0,0001
HWAAH	0,0003	0,0001
HWAH	0,0010	0,0003
HWAWH	-	0,0001
HWH	0,0065	0,0050
HWHAH	0,0007	0,0006
HWHSH	0,0003	0,0001
HWHWH	0,0014	0,0004
HWSH	-	0,0001
HWWH	-	0,0001
HWWWH	-	-

A Tabela 7.7 coloca os padrões de viagens de forma agrupada em relação às atividades realizadas e as suas respectivas probabilidades. Indivíduos pertencentes a famílias menores (uma ou duas pessoas) viajam mais em relação àqueles pertencentes a famílias maiores (3 ou mais pessoas). Indivíduos de famílias menores ficam menos em casa (0,4834) e realizam mais outros tipos de atividades (0,5023) em relação às famílias maiores (0,2419). A pouca incidência de padrões relacionados à escola pode estar relacionado à não presença de crianças no domicílio, o que mostra que esses indivíduos não possuem restrições para a realização de outras atividades. Padrões de viagens relacionadas à escola nas famílias com três ou mais pessoas (0,0613) reforçam a idéia de que estas famílias provavelmente possuam crianças que são levadas para a escola.

Tabela 7.7. – Indivíduos que não trabalham e nem estudam, cônjuge ou chefe
Padrões de viagem em relação ao número de pessoas na família

Tipo de Padrões de viagem	Famílias ≤ 2	Família ≥ 3
Não viajam (H)	0,4834	0,6815
Somente com realização de Outras Atividades	0,5023	0,2419
Realização de viagens a Escola e de outras Atividades	0,0003	0,0084
Realização de viagens ao Trabalho e outras Atividades	0,0020	0,0011
Realização de viagens somente para a escola	0,0038	0,0613
Realização de viagens para escola e trabalho	0,0003	0,0003
Realização de viagens somente para o trabalho	0,0079	0,0055

O comportamento dos indivíduos pertencentes a famílias com três ou mais pessoas varia em função da faixa etária. A observação entre os Grupos 3, 4 e 5, isto é, indivíduos que não trabalham e não estudam, cônjuge de famílias com 3 ou mais pessoas mostram parte dessas diferenças (Tabela 7.8).

Tabela 7.8. – Grupo 3, Grupos 4 e 5 - Não estudam e não trabalham, cônjuge, famílias com 3 ou mais pessoas - Comparação entre faixas de idade e padrão de viagem

Não estudam/Não trabalham Cônjuge de Famílias com 3 ou mais pessoas	Idade ≤ 23 Grupo 3	24≤Idade≤49 Grupos 4 e 5
% Total obs. Amostra	1,05%	10,05%
H	0,7347	0,6670
HAAA	0,0055	0,0069
HAAH	0,0167	0,0221
HAH	0,1831	0,1770
HAHAH	0,0100	0,0120
HAHSH	0,0011	0,0034
HAHWH	-	-
HASH	-	0,0015
HAWH	-	-
HSAH	0,0011	0,0024
HS	0,0033	0,0325
HSAH	0,0011	0,0055
HS	0,0111	0,0655
HSHWH	-	0,0001
HSSH	-	0,0005
HSWH	-	0,0001
HWAH	-	0,0001
HWAH	-	0,0002
HWAH	0,0011	-
HWH	0,0011	0,0027
HWH	-	0,0001
HWH	-	0,0001
HWH	-	0,0002
HWSH	-	-
HWWH	-	0,0001
HWWH	-	-

Cônjuges com idade menor que 23 anos, que não estudam e nem trabalham, pertencentes a famílias com três ou mais pessoas, correspondem a 1,05% da amostra em relação aquelas com as mesmas características, com idade entre 24 e 49 anos (10,05%). As jovens casadas (idade ≤23 anos) têm maior tendência a permanecerem em casa (0,7347) em relação àquelas mais velhas (0,6670). O que pode indicar que essas jovens ficam mais em casa em relação às mais velhas provavelmente pela presença de crianças pequenas no domicílio. As casadas com idade entre 24 a 49 anos apresentam-se com maior probabilidade de realizar padrão HSH (0,0325) em relação as mais jovens com idade menor que 24 anos (0,0033). Padrões de viagens relacionadas à escola (HSH, HSHSH) encontradas nesses grupos podem corresponder àquelas viagens realizadas com a finalidade de levar ou buscar os filhos ou crianças pequenas para a escola. O mesmo ocorre com o padrão HSHSH (0,0110 e 0,0655 respectivamente).

BHAT (1997) observou que as pessoas mais velhas têm maior probabilidade de realizar atividades dentro da residência. De acordo com BHAT (1997), os resultados também apresentaram maiores probabilidades das mais velhas permanecerem em casa. No entanto, estes resultados mostram que quando comparadas às mais jovens casadas (<23 anos), estas têm maior probabilidade de não viajarem em relação às mais velhas provavelmente em função da presença de filhos pequenos em idade pré-escolar.

A observação dos Grupos 4 e 5 mostra a influência do grau de instrução no comportamento de viagem. Mulheres com idade entre 24 a 49 anos, que não trabalham e nem estudam, cônjuges, pertencentes a famílias com 3 ou mais pessoas (Grupos 4 e 5), correspondem a 10,05% de toda amostra. Parte delas (71,82%), possuem até o primeiro grau até a quarta série concluído (Grupo 4) e as demais (28,18%) acima do primeiro grau completo (Grupo 5).

Cônjuges com grau de instrução baixo têm maior probabilidade de permanecerem em casa (0,6922) em relação àquelas com maior grau de instrução (0,6028). Essas apresentam maior probabilidade de realizar mais viagens e realizar outros padrões de viagens, principalmente HSH e HSHSH. No primeiro caso, provavelmente se trata de mães levando ou indo buscar filho(s) na escola e no segundo caso, levando e indo buscar filho(s) na escola.

A influência do grau de instrução nos padrões também pode ser observada nos estudantes. Os estudantes correspondem a 24,40% da amostra da população, e 62,40% deles possuem grau de instrução até o 1º grau (até a 4ª série) e 37,40% com grau de instrução mais elevado. Os padrões de comportamento dos estudantes variam em função do nível de instrução (Tabela 7.9).

Tabela 7.9. – Grupos 8, 9 e 10 – Estudantes

Comparação entre grupos através do grau de instrução

Estudantes	Analfabeto a 1º Grau até 4º série (Grupos 8 e 9)	Acima do Primeiro Grau completo (Grupo 10)
Total de Observações	13.014	7.843
% Total obs. amostra	13,23%	9,18%
H	0,1128	0,0928
HAAAH	0,0009	0,0010
HAAH	0,0023	0,0042
HAH	0,0170	0,0217
HAHAH	0,0005	0,0024
HAHSH	0,0042	0,0203
HAHWH	-	0,0003
HASH	0,0014	0,0032
HAWH	-	-
HSAH	0,0088	0,0111
HSH	0,8107	0,7583
HSHAH	0,0237	0,0340
HSHSH	0,0102	0,0400
HSHWH	0,0005	0,0006
HSSH	0,0039	0,0055
HSWH	0,0005	0,0001
HWAAH	-	-
HWAH	-	-
HWAWH	0,0001	0,0001
HWH	0,0013	0,0018
HWHAH	-	-
HWHSH	0,0003	0,0015
HWHWH	-	-
HWSH	0,0009	0,0011
HWWH	-	-
HWWWH	-	-

Aparentemente não existem diferenças muito grandes entre estudantes que permanecem em casa em relação ao nível de instrução. Pode-se observar que estudantes acima do primeiro grau completo são um pouco mais assíduos em relação aqueles com grau de instrução baixo. Indivíduos com grau de instrução maior tendem a realizar padrões mais complexos (HSHSH, HSSH), além de fazer mais atividades antes e depois da escola (HASH, HAHSH, HSAH ou HSHAH) em relação aos de grau de instrução baixa.

Variações no comportamento de estudantes em relação à presença de automóvel no domicílio é observada nas análises dos Grupos 8 e 9. Observa-se que o grau de instrução baixo (até a quarta série do primeiro grau) pode indicar que a maioria são crianças. Estudantes que possuem automóvel no domicílio tendem a permanecer menos em casa (0,1014) em relação aos que não possuem automóveis (0,1208). As viagens do tipo HSH,

isto é, viagens diretas à escola, são realizadas mais freqüentemente por estudantes que não possuem automóvel (0,8218) em relação aos que possuem (0,7948).

Quando o padrão de viagem se torna mais complexo, a presença do automóvel pode ser considerada importante, como observado por KITAMURA et. al (1981) que verificaram que o número de viagens encadeadas cresce com o número de automóveis disponíveis na família. De fato, observa-se que a probabilidade de ocorrência do padrão HSHAH aumenta com a presença do automóvel (de 0,01436 para 0,03716), o mesmo ocorre para HSAH (de 0,00496 para 0,01438). Logo, a presença do automóvel no domicílio, provavelmente permite às crianças uma maior facilidade de realizar mais atividades, principalmente depois da escola.

Uma análise mais detalhada mostra que a condição do indivíduo dentro da família indica que indivíduos que não trabalham nem estudam, que são filhos, parentes, agregados ou empregadas residentes, visitantes não residentes normalmente não viajam. Estes apresentam elevada probabilidade de permanecerem em casa (0,8626) e baixa probabilidade na realização de outras atividades - HAH (0,0931), o que reforça de certa forma, com as idéias de MITCHELL e RAPKIN (1954) em relação à ocupação do indivíduo.

Parte das diferenças entre o comportamento de homens e mulheres pode ser observada nos Grupos 12 e 13 (Tabela 7.10). Esses grupos são formados por indivíduos que trabalham, são chefe ou cônjuge e têm salários médios entre 1,40 a 5,55 salários mínimos.

Tabela 7.10. – Grupos 12 e 13 – Trabalhadores, Chefe ou Cônjuge, com salário médio entre 1,40 e 5,55 salários mínimos - Probabilidades de Realização de padrões de viagens entre homens e mulheres

Tipo de Padrões	HOMENS	MULHERES
Sem atividades (H)	0,0806	0,1368
Somente com realização de Outras Atividades	0,0743	0,0899
Realização de viagens a Escola e de outras Atividades	0,0007	0,0028
Realização de viagens ao Trabalho e outras Atividades	0,0563	0,0677
Realização de viagens somente para a escola	0,0035	0,0107
Realização de viagens para escola e trabalho	0,0054	0,0141
Realização de viagens somente para o trabalho	0,7792	0,6780

Observa-se que as mulheres realizam mais atividades em relação aos homens e padrões relacionados ao trabalho são mais freqüentemente realizados pelos homens.

Segundo MACGUCKIN e MURAKAMI (1999), nota-se a participação da mulher na força de trabalho, embora os padrões de viagens das mulheres relacionados ao trabalho variam se comparados aos dos homens.

Os homens têm maior probabilidade de ir diretamente para casa depois do trabalho em relação às mulheres como também observado por HAMED e MANNERING (1993). No entanto, padrões que apresentem viagens relacionadas à escola antes do trabalho são mais frequentes nas mulheres. Os homens com padrão HWHWH também voltam mais para casa (0,05486) em relação às mulheres (0,03887).

As mulheres trabalhadoras têm maior probabilidade de não ir ao trabalho (0,1368) em relação aos homens (0,0806). Os padrões de viagens relacionados a atividades (HAAAH, HAAH, HAH ou HAHAH) também são mais frequentemente realizados pelas mulheres (0,0899) em relação aos homens (0,0743). Essa incidência de padrões de viagens não relacionados ao trabalho nas mulheres pode estar relacionada à sua natureza, pelas suas responsabilidades em relação aos filhos ou à família. Trabalhos temporários também podem indicar a permanência na casa, por exemplo, faxineira, diarista, vendedoras, etc. É possível também que a responsabilidade dos homens na manutenção da família não permitam que eles falem no trabalho.

Os padrões relacionados como levar ou buscar filho(s) na escola são também encontrados em maior proporção nas mulheres (0,0276) em relação aos homens (0,0096). Elas também fazem mais atividades depois do trabalho do tipo HWAH (0,03061) ou HWHAH (0,02081) em relação aos homens (0,02125 e 0,01829, respectivamente). Isto pode indicar a responsabilidade das mulheres dentro do domicílio (ROSENBLOOM, 1987), levando a realizar outros tipos de atividades relacionadas à manutenção da família. Além disso, FOX (1983) mostrou que as diferenças entre os padrões de viagens das mulheres que trabalham quando comparadas aos dos homens podem estar relacionados à presença de crianças no domicílio.

Outra parte das diferenças entre homens e mulheres são analisados através dos Grupos 16, 17 e 18. Esses grupos são formados por indivíduos que trabalham, são membros da família (chefe, cônjuge, filho, parente ou agregado), com salários acima de 5,5 salários mínimos e possuem no mínimo um automóvel na família.

No geral, observa-se que as mulheres faltam mais no trabalho em relação aos homens. Os homens provavelmente realizam outros tipos de atividades depois do trabalho do tipo HWAWH do que as mulheres. A probabilidade do padrão HWAWH também varia em relação aos homens. Homens com rendas entre 5,5 e 13,95 salários mínimos (0,01303) têm maior probabilidade em relação aos de renda maiores de 13,95 salários mínimos (0,02215).

No grupo de trabalhadores, a renda individual representou importância na formação dos grupos, onde observa quatro faixas de salários individuais. O Grupo 11 é basicamente formado por trabalhadores que possuem salário individual baixo (1,40 salários mínimos) e correspondem às pessoas da família (chefe, cônjuge, filho, parente ou agregado). Em linhas gerais, a medida em que os salários vão aumentando, aumenta a complexidade de análise em relação ao número de variáveis que influenciam o comportamento de viagem. Faixas entre 1,40 e 5,55 salários mínimos são classificadas entre a situação familiar e o sexo. Quando os salários aumentam (>5,55 salários mínimos) o número de automóveis e o sexo são variáveis consideradas nas variações de padrão de viagem. Neste caso, a análise abrangeu o conjunto dos Grupos 11 a 18. Em linhas gerais, indivíduos com rendas baixas (<1,40 salários mínimos) não vão trabalhar (0,1683) e apresentam probabilidade do padrão HAH (0,06991) maior em relação aos de rendas mais elevadas. É provável que o aumento do salário permita a participação de mais atividades depois do trabalho, como pode ser observado nos padrões HWAH, HWA AH, HWA AH. Observa-se também que homens que têm salários altos (maior que 13,95 salários mínimos) tendem a realizar padrões do tipo HWAWH (0,02215). De acordo com STRATHMAN e DUEKER (1995), a medida em que os salários aumentam, os padrões de viagens tornam-se mais complexos.

Indivíduos que trabalham e estudam correspondem a 4,54% da população (Grupo 20). Comparando com os grupos que somente trabalham (Grupos de 11 a 19) e grupos que somente estudam (Grupos de 8 a 10), indivíduos que estudam e trabalham têm menor probabilidade de ficar em casa (0,02808). Os quatro padrões mais frequentes encontrados são: HW HSH (0,40060) e HWSH (0,21950), HWH (0,13650) e HSH (0,06285). Uma análise detalhada (Tabela 7.11) mostra as diferenças entre as ocupações de homens e mulheres dos indivíduos com padrão HW HSH (Grupo 20 – Indivíduos que estudam e trabalham).

Tabela 7.11. – Grupo 20 - Principais ocupações entre homens e mulheres que realizam HWHSH

Homens		Mulheres	
	%		%
Escriturário	21,79%	Escriturário	39,63%
Serviços não especializados	18,64%	Serviços semi-especializados	7,89%
Serviços semi-especializados	8,79%	Empregado do comércio	7,07%
Empregado do comércio	6,46%	Prendas domésticas	6,12%
Operário Especializado	4,58%	Serviços não especializados	4,77%
Demais atividades	39,74%	Demais atividades	34,51%

Em relação ao sequenciamento na realização das atividades do Grupo 20, observa-se uma pequena porcentagem de padrões HSHWH (4,83%) e HSWH (2,22%), em relação aos padrões HWHSH (40,06%) e HWSH (21,95%). É provável que o horário de trabalho da maioria dos empregos e serviços ocorram nos períodos diurnos, sendo responsável pela maior incidência de padrões do tipo HWHSH e HWSH. A pequena porcentagem dos padrões em que a atividade escola ocorre antes de ir ao trabalho pode estar relacionado ao tipo de trabalho, como por exemplo, estágios em empresa, que ocorrem em um período do dia.

Em relação ao grau de instrução (Gráfico 7.1) observa-se também neste Grupo que à medida que o grau de instrução aumenta, a realização do padrão HWHSH aumenta.

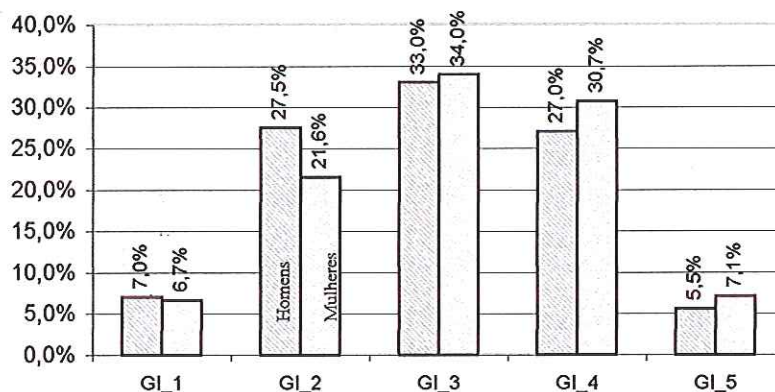


Gráfico 7.1. – Grau de Instrução entre Homens e Mulheres que realizam padrão HWHSH – Grupo 20

Empregadas residentes ou visitantes não residentes que trabalham correspondem a 0,91% da amostra (Grupo 19). 83,68% destes não viajam, 8,20% realizam padrão de viagem do tipo HAH. Apesar desse grupo ser caracterizado por trabalhadores, a maioria não viaja, o que pode estar relacionada ao papel da empregada doméstica, indicando que provavelmente o trabalho seja realizado na maioria das vezes na residência. Os padrões de viagens

relacionados à escola são também encontrados, indicando que as empregadas levam ou buscam as crianças na escola.

Como conclusão geral deste capítulo, as análises a partir das árvores de decisão revelam o comportamento de indivíduos em relação aos padrões de viagens. Os resultados encontrados através da árvore de decisão mostraram as relações entre os padrões de viagens e as condições sócio-econômicas dos indivíduos e conseqüentemente reforçaram os resultados obtidos na literatura. Ainda que algumas variáveis não pudessem ser confrontadas como, por exemplo, o grau de instrução, os resultados apresentaram-se coerentes com os esperados. Além disso, muitas outras variáveis poderiam ser associadas, como características do domicílio, modos de viagem, tempo e distâncias de viagem.

Neste caso, o minerador de dados aplicados aos dados da pesquisa O/D na RMSP, cujas viagens encadeadas foram codificadas, permitiu mostrar as relações entre padrões de viagens e características sócio-econômicas o que permite extrair conhecimento e informações que ajudam a compreender o comportamento dos viajantes.

CAPÍTULO 8 – CONCLUSÕES, RECOMENDAÇÕES E SUGESTÕES PARA TRABALHOS FUTUROS

A principal conclusão foi a constatação de que é possível aplicar um minerador de dados para obter relações entre padrões de viagens encadeadas e características sócio-econômicas de viajantes urbanos. O minerador de dados aplicado a dados individuais, contendo o encadeamento de viagem codificado e atributos sócio-econômicos e de viagem, permite extrair conhecimento e informações desconhecidas que ajudam a compreender o comportamento dos viajantes urbanos.

Os resultados mostraram que indivíduos com atributos sócio-econômicos e de viagens similares não se comportam de maneira semelhante. Ao contrário, muito deles fazem diferentes padrões de viagens, movidos por algumas razões que não estão explícitas nos dados disponíveis. Mas, considerando que esses indivíduos não foram entrevistados no mesmo dia, e que hábito, necessidade, conveniência e preferência variam de pessoa para pessoa a despeito das características sócio-econômicas semelhantes, é de se esperar que certos padrões de encadeamento de viagem estejam mais fortemente relacionados com uns do que com outros. Assim, a rigor, o que difere de um grupo sócio-econômico para outro é a distribuição de probabilidade de ocorrência dos padrões de encadeamento de viagens.

A comparação dos resultados do minerador de dados com os resultados encontrados na literatura reforça a utilidade do minerador de dados como um instrumento de análise do comportamento de viajante urbano. Cabe ressaltar que tal comparação foi feita apenas no sentido lato e não nos detalhes, mesmo porque não se poderia esperar que povos que vivem em diferentes ambientes sócio-econômicos, culturais e institucionais se comportem de forma semelhante. Basicamente, o objeto da comparação foi a relação entre tipos de ocupação dos indivíduos e os padrões predominantes de encadeamento de suas viagens. Nesse sentido, os

resultados mostram que o minerador de dados fornece relações bastante similares às observadas no exterior, ao menos em termos de seqüência de atividades desenvolvidas ao longo de um dia.

A codificação dos padrões de viagens encadeadas constitui uma das principais contribuições deste trabalho. Por meio dela, torna-se possível representar os múltiplos aspectos das viagens encadeadas através de um código único, e de explorar, através de um minerador de dados, as relações entre os padrões de viagens encadeadas e as características sócio-econômicas dos indivíduos e peculiaridades dos locais onde as atividades são desenvolvidas.

Os padrões de viagem considerados neste trabalho são referentes a dois, três e quatro viagens realizadas pelos indivíduos por dia, além do padrão de viagem correspondente a zero viagem. Não foram consideradas para análise os padrões de viagens provenientes de uma viagem ou mais de quatro viagens, uma vez que o número de pessoas que fazem essas viagens é relativamente pequeno (aproximadamente 5,0% do total da amostra) e os padrões derivados de mais de quatro viagens são numerosos. Em função dessas considerações, o número de variáveis resposta aplicados ao minerador correspondeu a um total de 26 padrões de viagens.

Dos 26 padrões de viagens analisados foram encontrados quatro grupos principais de indivíduos caracterizados pelos padrões de viagens: H, HWH, HSH e HWHSH. A primeira variável explicativa considerada na formação desses grupos é aquela relacionada com a ocupação (trabalha ou não trabalha, estuda ou não estuda), o que indica uma importância fundamental desta variável sobre os padrões de viagens dos indivíduos. Dessa forma, estas relações são observadas nos grupos de indivíduos que não estudam e não trabalham que, na maioria das vezes, não viajam (H); trabalhadores apresentam, na sua maioria, padrão HWH; os estudantes padrão HSH; e os que trabalham e estudam apresentam padrão HWHSH.

Estes resultados mostram que é possível obter relações entre os padrões de viagens realizados pelos indivíduos através de sua ocupação principal que é coletada durante o recenseamento da população. Logo, parece possível fazer o planejamento do transporte urbano com base em modelos de árvore associados aos dados de recenseamento.

O reconhecimento de padrões de viagem relacionados às características sócio-econômicas permite desenvolver políticas baseadas nessas características e facilitar o desenvolvimento de planos de ação pontuais na melhoria de qualidade do sistema de transportes.

Apesar de os padrões de viagens considerados neste trabalho levarem em conta apenas o motivo da viagem realizada em cada deslocamento, os padrões de viagem poderiam ser representados também com atributos de viagens, como o período de tempo em que as viagens ou as atividades foram realizadas ou os modos principais utilizados pelo indivíduo durante a viagem ou a localização das atividades. O principal desafio está em considerar todas as dimensões sócio-econômica, demográfica e de viagens relevantes para compreender o comportamento do viajante urbano. No entanto, ressalta-se que nesse trabalho essas características não foram consideradas, já que o objetivo do trabalho é analisar a viabilidade de aplicar o modelo de árvore aos padrões de viagens codificados.

O minerador de dados utilizado nesta pesquisa foi o modelo de árvore de decisão do *software* S-Plus 2000. A escolha entre um ou outro *software* foi feita em função da sua disponibilidade no Centro de Computação Eletrônica da USP e da conveniência. Outros programas (CART, CHAID, etc.) poderiam ter sido utilizados, pois estão disponíveis em algumas unidades da USP, no entanto, isso exigiria deslocamentos a essas unidades.

Os dados da Pesquisa origem-destino (O/D) da Região Metropolitana de São Paulo, realizada em 1987 passaram por um tratamento dos dados que consistiu basicamente na limpeza dos dados e transformação dos dados de viagens nos padrões de viagem codificados. Todo o tratamento de dados teve também como objetivo a redução do tempo de processamento do algoritmo minerador, visto que ele trabalhará apenas com dados necessários, diminuindo o seu espaço de busca. Este tratamento exigiu certo esforço nas correções de possíveis erros de preenchimento ou eliminação de dados inconsistentes e verificação dos formatos dos dados, principalmente nas suas unidades. Além disso, a verificação ou a eliminação de valores nulos foi necessária para a geração da árvore, já que todas as células correspondentes aos valores de atributos (categóricos ou não) não podem apresentar células em branco. A transformação dos dados de viagens em padrões de viagens codificados consolidou as informações do conjunto selecionado, constituindo os dados de entrada para o algoritmo de mineração dos dados.

Através de uma avaliação do modelo de árvore de decisão do S-Plus 2000 foi possível conhecer algumas de suas propriedades. Uma característica importante é que vários tipos de arquivos (xls, dbf, mdb, etc.) inclusive de programas específicos (SPSS, SAS, MATLAB, etc) são reconhecidos e facilmente importados. Outra característica é que aparentemente não há limitação para o número de informações que podem ser inseridos no banco de dados do *software*. Os dados de 85.477 indivíduos puderam ser trabalhados de forma integral, sem a necessidade da sub-divisões em arquivos. E finalmente, os tempos de processamento do S-Plus 2000 na geração da árvore foram pequenos para o conjunto selecionado de dados.

O S-Plus 2000 apresenta capacidade de analisar no máximo 128 variáveis resposta, isto é, 128 padrões de viagens. Se considerar a inserção de outros atributos nas formação dos padrões de viagens, é necessário o cuidado para que o número de padrões de viagens não ultrapasse a capacidade máxima para análise. Uma forma alternativa de reduzir o número de combinações possíveis é o agrupamento codificado de alguns atributos utilizados na formação do padrão, como por exemplo, agrupamento em períodos de viagem os horários da realização das viagens.

O S-Plus 2000 fornece duas formas de apresentação de resultados: a forma gráfica ou através de um relatório. Embora a finalidade principal da forma gráfica seja a de facilitar a interpretação de resultados, esse objetivo é melhor alcançado pelo relatório, pois estes trazem mais informações, tais como a probabilidade ou frequência associada a cada padrão de viagem.

Os resultados deste trabalho corroboram uma das sentenças fundamentais da área de transportes: demanda por viagens é derivada da demanda por atividades. As decisões de viagens são dirigidas por um conjunto de atividades que formam uma “agenda” de participação nas atividades e não podem ser analisadas com base nas viagens individuais. O conjunto de atividades e viagens executadas depende do padrão de atividade do indivíduo, do processo de decisão e das regras de comportamento e do seu ambiente. Todos estes elementos levam a formação de padrões que caracterizam um comportamento complexo de viagem.

Uma característica importante no processo de decisão é que os membros do domicílio influenciam as decisões individuais e as escolhas são limitadas pelo tempo e pelo espaço. Além disso, outros tipos de restrições também afetam essas decisões, como a capacidade do

indivíduo, a dependência entre os membros do domicílio ou restrições impostas por um dos membros do domicílio.

A teoria das viagens baseadas em atividades, os avanços na tecnologia computacional e a necessidade de melhores modelos de previsão de demanda favorecem as análises através das abordagens de viagens baseadas em atividades. Os estudos desenvolvidos avaliam como as pessoas organizam suas atividades no tempo e no espaço. Isto é, tratam das atividades (onde, quando e de que forma elas podem ser realizadas) e as relacionam com as características dos indivíduos e domicílios, as oportunidades, a rede de transportes e as restrições encontradas.

As informações relativas às atividades realizadas pelos indivíduos são elementos importantes para as análises de viagens baseadas em atividades, já que a viagem em si é apenas um mecanismo de acesso para a realização dessas atividades. Dessa forma, considerar essas informações e a reformulação dos modelos de coleta nos levantamentos pode melhorar a representação do comportamento dos indivíduos nos padrões de viagem.

RECOMENDAÇÕES

Recomenda-se a verificação prévia de todos os dados faltantes no conjunto de dados antes da importação do mesmo para o S-Plus 2000. Os dados nulos, isto é, células em branco, não permitem a geração da árvore de decisão pelo *software*. É necessário avaliar a qualidade dos dados e os dados inconsistentes, pois o *software* não possui capacidade de identificá-los, podendo conduzir a erros de análise.

O S-Plus 2000, em seu ambiente, permite o tratamento do banco de dados, podendo inserir ou excluir colunas ou linhas, selecionar dados, realizar cálculos com a utilização de fórmulas, etc., no entanto, esse processamento é lento se comparado com outras planilhas de cálculo. Sendo assim, é interessante que essas tarefas devam ser realizadas em arquivos com maior facilidade de processamento, como o Excel, por exemplo.

Como o número máximo de 128 variáveis resposta considerado no *software* limita a forma de análise através dos padrões de viagens, há a necessidade adotar algumas restrições no caso de incluir outros atributos de viagens, para não ultrapassar a capacidade máxima de análise. Esta medida conseqüentemente leva a redução das combinações possíveis. No caso de inclusão de modos de viagem nos padrões por exemplo, os 13 modos apresentados na Pesquisa O/D 1987 leva a formação de 507 padrões de viagens (duas viagens), o que já de antemão tornaria inviável a utilização do *software*.

Recomenda-se também que os formulários de coleta de dados sejam preenchidos de forma mais precisa, no sentido de permitir identificar algumas características dos indivíduos em relação às viagens realizadas. Um exemplo observado nos resultados e que poderia ser solucionado através da coleta de dados, é que não foi possível diferenciar características de indivíduos que levam os filhos na escola e aqueles que realmente vão à escola.

SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestões para pesquisas futuras, a continuação da pesquisa pode se concentrar em vários pontos:

A realização de análises usando a codificação das viagens encadeadas que considere outros atributos de viagens, tais como modo, distância, períodos e tempo de viagem, etc. O estilo de vida é um fator relacionado aos padrões de viagem. A localização da residência e a distância ao local de trabalho podem influenciar o número de viagens realizadas, e conseqüentemente o padrão de viagem.

Alem das características sócio-econômicas, algumas variáveis como o custo de viagem, tempos de viagem, e dados referentes às políticas existentes, quanto ao meio de transporte, vias de acesso, etc. também são elementos importantes que poderiam incorporar na análise.

A codificação dos modos de viagem incorporados aos padrões de viagens pode levar a resultados que permitem, em primeiro lugar, avaliar a relação entre os padrões de viagens e os modos utilizados dentro de uma população e permitir a comparação com estes resultados. Em segundo lugar, será possível reconhecer grupos distintos de acordo com os modos de transportes e permitir a aplicação de políticas diferenciadas.

A análise no contexto familiar permite verificar diferenças nos padrões de viagens entre os membros de um domicílio. O número de trabalhadores, a posição dentro da família, a idade dos membros do domicílio podem variar a distribuição das tarefas e atividades. Fatores como o número e idade das crianças podem influenciar a forma dos padrões de viagem dos outros membros do domicílio.

A presença de crianças na família tende a mudar o comportamento do domicílio no tempo e no espaço. Para atender as necessidades das crianças pequenas, como ir à escola ou a creche, é necessária a presença de um adulto normalmente membro da família.

A comparação entre padrões de atividades e de viagens de estudantes e trabalhadores, verificando limitações em relação às suas respectivas atividades, já que as viagens com destino ao trabalho são tão freqüentes quanto à escola. Outros tipos de ocupação poderiam ser observadas.

A identificação das diferenças nas atividades e comportamento de viagens em diferentes classes sociais e econômicas, isto é, comparando dados de indivíduos com características diferentes, por exemplo, analisando quais os efeitos resultantes destas comparações. Outra avaliação poderia ser analisada através das mudanças de comportamento da população ao longo do tempo, isto é, a transferibilidade temporal de resultados da análise baseada em atividades. No caso da Região Metropolitana de São Paulo, já foram realizadas a pesquisa O/D 1997.

A aplicação de outros softwares poderá fornecer parâmetros de comparação dos resultados, levantando em conta também diferenças entre o desempenho, implementação e utilização.

Com os avanços nas tecnologias de informação, é possível que alguns indivíduos, para realizar suas atividades, substituam as viagens por outros meios, como, por exemplo, o uso das telecomunicações ou o uso da internet. Esta idéia não é aplicada somente para as viagens a trabalho, mas para todos os tipos de atividades que não necessitam da presença física. Isto leva, de certa forma, uma dificuldade nas modelagens, mas que deve ser levada em consideração nos estudos futuros.

Outros pontos podem ser analisados como a influência do tempo. Os horários das atividades podem determinar a formação de diferentes padrões de viagens.

As variações na duração e o tipo de atividade também podem estar relacionadas com as características individuais e domiciliares. E as características do meio urbano como os horários de trabalho ou os horários de comércio podem afetar o comportamento dos indivíduos no tempo e no espaço.

ANEXOS



ANEXO A – CARACTERÍSTICAS GERAIS DAS ZONAS

Num. Zona	Nome da Zona	Domicílios	Famílias	População	Matriculas	Empregos	Automóvel	Viagens Produzidas
1	Sê	19.186	19.186	50.934	26.572	346.758	5.263	820.769
2	Brás	4.138	4.426	15.848	3.954	24.012	1.505	65.349
3	Tabatinguera	4.589	4.806	14.839	2.456	4.410	518	27.889
4	Glicério	3.740	4.370	15.877	793	7.349	1.936	29.786
5	Liberdade	7.980	9.517	23.523	29.520	22.848	3.254	113.409
6	Bexiga	13.104	13.500	37.368	7.240	40.149	7.141	132.724
7	Consolação	7.396	7.775	13.856	14.920	20.084	2.780	78.344
8	Santa Cecília	6.908	6.908	16.901	4.375	13.334	4.533	50.428
9	Campos Eliseos	9.424	9.305	26.088	8.379	24.247	3.931	77.549
10	Bom Retiro	5.873	6.395	20.975	8.250	38.010	2.204	109.247
11	Ponte Pequena	2.597	2.641	10.011	12.084	24.312	1.007	69.021
12	Pari	6.317	6.576	23.615	11.911	31.661	4.081	100.804
13	Bresser	7.166	8.514	26.943	7.381	61.713	2.846	144.374
14	Moóca	9.868	10.121	37.099	19.133	67.871	7.705	169.269
15	Cambuci	14.826	16.555	49.784	9.321	33.683	11.656	128.864
16	Aclimação	10.538	11.213	40.636	10.630	30.098	8.562	129.867
17	Glória	15.393	15.935	54.456	11.922	22.815	14.226	123.996
18	Paraíso	23.106	23.833	80.841	25.111	68.147	22.113	270.914
19	Paulista	15.711	15.768	44.522	20.146	125.797	12.647	316.705
20	Alamedas	15.028	15.028	51.873	12.517	46.135	20.771	171.993
21	Pacaembu	8.846	8.846	27.878	13.776	32.965	12.844	109.390
22	Veiga Filho	12.531	13.232	37.522	8.056	20.607	10.412	87.220
23	Brigadeiro Galvão	5.880	6.056	14.791	7.901	14.668	3.214	55.745
24	Sumaré	34.594	36.952	122.226	38.202	72.048	35.360	339.190
25	Barra Funda	3.600	3.735	12.400	1.253	40.210	2.315	79.955
26	Casa Verde	10.260	10.921	37.470	14.501	23.185	7.541	113.354
27	Santana	11.574	12.111	42.747	20.040	27.241	8.739	162.024
28	V.Guilherme	20.046	21.013	75.175	19.910	51.698	14.080	252.246
29	Vila Maria	7.870	8.359	27.204	11.906	24.680	4.546	94.985
30	Belenzinho	7.999	8.236	27.163	8.134	47.711	6.178	119.178
31	Quarta Parada	10.769	11.093	41.126	16.178	31.679	8.943	122.543
32	Alto da Moóca	18.487	19.616	70.320	21.489	52.050	21.305	210.935
33	Ipiranga	30.180	30.987	105.931	33.929	71.034	24.348	292.305
34	Saúde	55.195	56.992	192.576	65.393	94.399	48.540	542.567
35	Indianópolis	24.217	24.276	81.624	32.483	74.169	30.727	336.388
36	Itaim I	14.927	15.283	47.136	11.968	47.959	13.689	176.824
37	Jardins I	8.261	8.458	27.874	4.358	25.100	10.074	79.219
38	Clínicas	11.992	12.263	32.799	11.050	38.880	10.517	135.258
39	V.Madalena	13.453	14.320	44.464	8.619	14.147	14.731	98.560
40	Cerro Corá	7.525	7.753	26.254	5.437	8.343	8.963	61.445
41	Pompéia	6.991	7.151	22.320	16.939	40.544	5.255	166.503
42	Água Branca	3.067	3.208	12.935	3.526	32.895	2.991	68.515
43	Freguesia do Ó	15.548	18.597	68.967	19.905	27.795	11.361	157.717
44	Limão	18.644	19.380	70.646	9.497	30.890	16.203	137.758
45	Mandaqui	46.950	51.650	194.335	53.445	40.713	34.235	351.266
46	Água Fria	25.682	27.488	95.150	16.991	26.764	22.967	200.013
47	Tucuruvi - Jaçanã	37.283	39.541	151.712	32.465	32.580	26.777	274.306
48	V.Medeiros	42.972	48.717	182.434	36.751	41.808	24.081	304.220
49	V.Maria Alta	13.583	15.627	61.213	15.415	13.801	12.132	118.329
50	Novo Mundo	7.916	9.149	36.986	4.114	24.048	4.294	76.044

Num. Zona	Nome da Zona	Domicílios	Famílias	População	Matrículas	Empregos	Automóvel	Viagens
								Produzidas
51	Parque São Jorge	11.640	11.946	43.716	21.206	44.048	11.469	165.025
52	Tatuapé	20.679	21.672	77.834	22.604	35.139	15.727	189.059
53	Vila Diva I	10.910	11.251	37.400	10.007	12.473	9.287	82.215
54	Tamanduateí	24.437	26.023	95.376	20.423	34.862	15.537	192.226
55	São João Clímaco	43.440	48.362	197.460	36.646	53.665	28.955	363.183
56	São Judas	22.491	23.420	85.207	24.539	30.972	15.447	198.195
57	Jabaq/Americanópolis	41.238	48.811	197.660	37.662	26.195	13.656	270.847
58	Aeroporto	30.290	33.004	121.541	21.776	55.017	24.880	266.209
59	Campo Belo	27.699	28.204	104.511	23.738	63.701	36.307	284.245
60	Monções	8.946	9.085	33.254	4.219	22.620	8.555	77.642
61	V.Olimpia	5.682	5.682	19.050	3.922	21.722	5.752	62.181
62	Iguatemi	3.396	3.424	12.827	2.277	31.235	5.993	86.565
63	Pinheiros	11.664	11.664	33.368	12.333	52.384	9.012	193.363
64	Estr. da Boiada	6.516	6.579	24.028	5.569	14.631	9.693	68.136
65	Alto da Lapa	5.281	5.544	19.122	7.473	17.854	6.731	65.818
66	V.Anastácio	1.414	1.414	5.086	1.786	14.600	1.008	28.081
67	Ceasa	6.957	7.137	24.227	4.820	39.959	4.484	88.505
68	Vila Jaguaré	27.824	30.505	121.906	26.368	39.433	19.959	223.580
69	Pirituba	17.868	19.396	73.892	19.037	20.181	12.199	134.212
70	Jaraguá	14.788	16.228	72.203	19.674	11.790	5.723	101.147
71	Brasilândia	39.998	43.455	168.255	24.618	18.887	18.564	210.842
72	V.N.Cachoeirinha	14.394	16.786	67.792	16.459	10.982	8.645	106.029
73	Estrada da Parada	9.885	10.668	49.085	11.094	4.868	3.062	52.939
74	Jardim Peri	12.828	14.341	69.111	10.007	8.266	4.025	76.768
75	Barro Branco	17.760	18.347	60.981	17.525	16.949	13.321	130.067
76	Tremembé	17.395	19.683	79.529	12.571	15.934	11.497	119.007
77	Cangaíba	17.770	19.485	76.981	11.927	17.665	15.253	114.625
78	Penha	10.787	12.052	39.951	16.040	21.533	8.934	125.082
79	Vila Esperança	16.537	17.741	63.735	11.063	14.749	13.661	124.251
80	Carrão	19.872	21.241	79.584	22.508	29.661	14.778	173.908
81	Vila Matilde	21.386	23.525	93.236	24.403	20.011	14.188	153.021
82	Vila Formosa	26.847	29.958	112.848	24.044	25.102	17.124	208.762
83	Vila Prudente	53.088	58.053	222.033	45.980	39.612	35.342	380.246
84	E Matarazzo	22.547	25.670	115.462	26.996	20.529	8.230	167.063
85	Vila Ré	42.149	45.826	180.005	33.830	37.425	21.585	283.429
86	Artur Alvim	24.081	25.703	103.168	21.251	16.637	10.561	175.259
87	Líder	11.380	12.659	57.758	11.409	10.935	6.562	79.376
88	Cidade Centenário	15.964	17.970	70.644	20.893	20.029	7.851	134.465
89	Sapopemba	43.054	46.223	199.849	41.006	22.626	14.774	265.861
90	São Miguel	17.344	20.839	87.652	36.092	23.932	5.966	197.061
91	Saudade	25.292	29.589	128.449	26.405	18.975	9.076	167.791
92	Itaquera	18.095	20.052	84.325	30.012	16.593	8.792	149.304
93	Carmo	15.814	16.589	77.261	14.350	14.279	7.874	102.233
94	São Mateus	34.862	39.267	167.072	39.113	27.478	14.747	248.597
95	Parque Paulistano	18.466	21.719	101.957	24.392	16.779	5.668	121.432
96	Vila Curuçá	22.362	24.460	114.613	29.737	13.241	5.723	154.781
97	Itaim Paulista	23.208	27.844	124.136	29.140	20.258	5.246	163.669
98	Parada Quinze	26.266	27.027	123.210	33.525	11.525	6.950	165.302
99	Guaianazes	31.417	36.678	168.883	25.367	16.756	7.247	191.867
100	Passagem Funda	20.250	20.942	89.364	23.863	11.116	6.362	113.096
101	Sete Praias	7.461	8.197	34.042	4.728	3.096	2.045	39.277
102	Pedreira	23.069	23.645	99.914	22.061	15.268	9.417	142.133

Num. Zona	Nome da Zona	Domicílios	Famílias	População	Matrículas	Empregos	Automóvel Produzidas	Viagens
103	Cupecê	14.728	16.031	60.342	15.406	22.513	11.720	115.299
104	Campo Grande	17.704	18.452	77.467	19.073	55.453	16.950	203.492
105	Granja Julieta	16.057	16.643	60.002	23.350	64.858	20.349	236.888
106	Largo Treze	7.098	7.434	24.641	16.527	42.476	7.441	175.624
107	Paraisópolis	1.111	1.203	5.534	2.202	2.305	1.833	14.579
108	Interlagos	64.444	69.935	302.782	82.997	43.792	22.934	450.626
109	Largo do Socorro	5.676	6.063	22.093	6.410	29.072	2.529	74.503
110	Campo Limpo I	31.218	31.821	135.149	32.717	37.832	13.615	224.931
111	Parelheiros	8.512	9.218	38.931	7.501	3.137	2.552	54.099
112	Guarapiranga	45.664	51.855	231.033	48.149	30.684	15.594	310.416
113	Pirajussara	34.853	40.461	176.711	45.095	16.981	9.449	223.180
114	Cidade Jardim	16.756	17.046	61.150	14.792	31.908	22.833	143.371
115	Jaguare	9.633	9.979	41.237	7.102	31.024	8.189	90.254
116	Estrada do Embu	26.790	27.405	94.183	22.196	34.941	21.112	183.083
117	Rio Pequeno	24.082	24.661	96.113	25.803	24.288	19.373	184.536
118	Arpoador	12.965	14.594	68.251	14.387	17.981	4.419	117.070
119	Perus	13.518	17.017	67.664	14.038	7.047	4.337	95.267
120	Cajamar	6.025	6.855	31.875	7.816	8.488	2.096	48.253
121	Caieiras	7.847	8.391	33.893	9.240	8.988	3.814	63.356
122	Franco da Rocha	16.198	17.397	72.389	20.597	14.196	6.399	129.257
123	Francisco Morato	11.494	12.855	57.471	14.217	5.107	2.269	72.572
124	Mariporã	7.566	8.079	33.214	7.299	10.611	5.771	65.399
125	Morros	31.249	34.891	142.493	23.001	17.322	12.616	154.657
126	Timóteo Pentead	38.586	40.456	165.212	58.312	47.038	28.995	317.586
127	Macedo	21.820	24.279	96.379	20.011	20.556	12.920	147.905
128	Guarulhos	4.571	5.072	16.923	12.646	19.765	3.052	87.174
129	Várzea do Palácio	16.896	18.256	67.117	21.884	51.102	12.508	180.008
130	Estrada de Nazaré	15.302	17.233	79.104	18.728	14.322	3.383	107.203
131	Cumbica-Pimentas	24.369	27.787	124.669	29.484	74.442	6.370	248.426
132	Aruja	6.825	7.292	29.356	7.956	7.772	2.618	43.196
133	Sta Isabel	8.099	8.887	34.872	9.334	10.934	3.150	64.058
134	Corredor - Itaquá	2.795	3.188	13.672	2.332	9.279	764	24.748
135	Itaquaquecetuba	20.080	21.579	99.563	24.121	18.076	4.649	155.312
136	Porto de Areia	3.103	3.430	13.670	3.406	6.042	722	30.982
137	Ferraz	17.516	18.753	80.230	19.730	25.183	5.934	120.916
138	Poa	15.183	16.042	67.722	12.410	15.227	4.925	111.493
139	Suzano	29.025	30.917	136.067	32.879	44.077	12.402	180.050
140	Jundiapéba	2.242	2.402	9.991	2.413	2.116	585	13.920
141	Brás Cubas	10.087	10.815	41.520	13.102	10.577	5.314	73.867
142	Mogi das Cruzes	39.890	43.027	170.965	75.220	52.696	25.128	373.281
143	Cocuera	2.092	2.263	9.720	2.334	2.902	634	16.569
144	Cesar de Souza	2.632	2.754	10.799	2.211	9.770	1.239	25.588
145	Sabauna	694	793	3.481	818	755	112	3.943
146	Guararema	3.661	3.896	16.942	3.051	3.649	1.132	19.530
147	Biritiba-Mirim	3.276	3.776	16.247	3.941	4.044	1.218	24.245
148	Salesópolis	2.644	2.821	11.105	3.135	4.250	781	20.774
149	Capuava	36.923	42.907	197.943	45.239	29.520	9.114	257.370
150	Maua	8.210	10.291	37.422	13.357	21.032	6.879	95.407
151	Córrego da Serraria	5.224	6.181	26.256	5.895	7.051	2.224	42.338
152	Ribeirão Pires	15.641	16.866	71.462	18.234	19.377	7.343	123.407
153	Rio Grande da Serra	5.409	6.370	26.314	4.151	2.342	1.147	27.932
154	Parque das Nações	21.556	23.995	95.744	18.896	14.117	12.098	143.629

Num. Zona	Nome da Zona	Domicílios	Famílias	População	Matrículas	Empregos	Automóvel	Viagens Produzidas
155	Utinga	20.183	22.352	86.216	21.450	22.827	15.069	158.222
156	Palmares	18.167	19.428	80.285	18.366	19.843	16.969	150.681
157	Santo André	11.512	11.834	41.792	36.841	65.351	11.380	270.197
158	Vila Pires	33.266	37.137	154.079	37.928	44.504	28.008	312.108
159	Vila Luzita	28.525	32.073	135.496	23.563	13.169	14.652	182.185
160	São Caetano	12.774	13.451	49.870	21.236	62.494	10.412	211.326
161	Vila Gerti	23.956	25.769	103.541	23.187	21.209	21.562	198.370
162	Rudge Ramos	35.726	39.038	148.377	41.928	84.070	29.982	364.580
163	Anchieta	19.486	20.334	80.716	15.933	28.689	18.643	156.571
164	São Bernardo	28.766	31.332	119.152	46.563	60.222	22.797	382.348
165	Demarchi	35.243	37.000	150.605	39.766	70.264	19.702	286.481
166	Riacho Grande	3.248	3.917	17.630	3.660	3.354	1.735	25.794
167	Piraporinha	10.819	13.771	61.224	11.236	36.364	4.858	129.100
168	Diadema	36.112	42.016	178.508	42.298	63.140	13.921	353.234
169	Eldorado	8.352	9.157	37.295	9.651	3.630	1.558	49.904
170	Embu Guaçu	3.320	3.453	16.759	4.678	4.788	951	29.723
171	Cipó	2.827	3.081	13.517	2.797	2.642	739	19.164
172	Taboão da Serra	29.663	33.680	136.606	30.907	35.199	13.369	208.276
173	Embu	27.519	28.211	123.172	19.848	18.011	6.587	144.277
174	Itapecerica	14.623	16.760	67.502	15.853	20.141	5.272	117.698
175	Embu-mirim	1.788	1.858	8.967	1.345	3.001	507	9.438
176	Juquitiba	3.478	4.053	17.081	4.757	4.660	1.437	28.881
177	Granja Viana	6.330	6.641	30.256	3.995	10.171	5.679	45.057
178	Cotia	8.283	8.770	38.646	14.805	19.068	3.680	76.549
179	Capueira	217	227	850	74	47	57	568
180	Caucaia do Alto	1.260	1.392	6.589	1.378	1.930	408	7.430
181	Vargem Grande Paulista	3.519	3.519	13.555	4.140	3.194	1.540	19.737
182	Bussocaba	22.643	25.458	99.167	45.883	39.882	14.022	228.782
183	Osasco	4.463	4.463	16.701	20.814	39.695	3.899	150.188
184	Mutinga	17.429	18.894	76.559	19.146	21.934	9.025	136.786
185	Quitauna	50.292	55.946	245.452	45.497	27.668	22.648	336.506
186	Jardim Helena Maria	21.690	24.180	95.358	16.082	18.133	7.979	135.881
187	Aldeia de Carapicuíba	5.001	5.106	23.618	4.619	4.469	1.449	27.426
188	Vila Dirce	7.920	8.819	39.567	8.734	6.117	1.230	50.853
189	Carapicuíba	39.982	41.458	184.445	39.671	19.802	10.988	229.046
190	Tamboré	5.659	6.104	27.123	5.048	29.775	4.704	74.575
191	Barueri	16.850	19.416	82.291	20.408	27.007	4.985	135.131
192	Jandira	9.892	11.358	52.436	13.157	10.781	3.566	82.244
193	Itapevi	18.291	21.558	86.165	11.633	15.534	3.865	101.969
194	Santana do Parnaíba	5.060	5.479	24.670	6.868	7.573	3.816	41.896
195	Pirapora do Bom Jesus	1.586	1.754	6.768	1.102	2.367	780	10.310
196	Itaim II	3.958	4.006	12.715	3.332	14.593	4.224	48.570
197	Jardins II	328	328	1.546	3.299	2.778	836	10.746
198	Vila Diva II	3.465	3.902	15.601	2.968	3.018	1.411	28.381
199	Vila Diva III	5.272	5.856	20.661	1.382	3.272	4.361	34.840
200	V.N.Cachoeirinha II	3.928	4.002	16.775	1.241	3.243	1.965	23.629
201	Pedra Branca	2.166	2.217	8.388	2.019	1.309	1.216	10.322
202	Guaianazes II	2.250	2.399	10.193	5.364	934	627	17.133
203	Morumbi	8.154	9.221	40.896	8.402	17.613	11.643	89.324
204	Campo Limpo II	19.360	21.092	79.452	13.220	11.594	8.746	111.095
205	C. de Marte - Anhembi				188	2.404		8.237
206	Parque do Ibirapuera				326	4.198		11.922

Num. Zona	Nome da Zona	Domicílios	Famílias	População	Matrículas	Empregos	Viagens	
							Automóvel	Produzidas
207	Parque do Estado				132	2.204		5.209
208	Aerop. de Congonhas					9.495		18.047
209	Horto Florestal				282	306		1.078
210	USP				19.705	19.979		54.634
211	Pico do Jaraguá				119	864		1.860
212	Aeroporto de Cumbica					9.605		18.192
213	Anhanguera				1.186	2.090		6.641
214	Parada Taipas				139	554		1.597
215	Serra da Cantareira					1.577		3.009
216	Eng. Marsilac					185		881
217	Baronesa				1.439	1.248		6.715
218	Jaceguava					-		93
219	Morro do Mateus				258	-		107
220	Monjolinho					795		853
221	Santa Inês					-		389
222	Cristais					562		707
223	Campos de Santo Benedito					-		
224	Sete Voltas					-		
225	Cascatas					-		
226	Represa do Juqueri					128		806
227	Pirucaia					513		896
228	Cabuçu					-		
229	Morro Grande					612		770
230	Fazenda Velha					754		1.287
231	Estrada dos Índios				106	1.783		2.140
232	Reservatório do Jaguari					1.112		2.198
233	Jardim Joseli					-		
234	Guaiú					-		
235	Jardim São José					-		183
236	Palmeiras				2.495	5.220		6.325
237	Itapeti				386	1.305		2.316
238	Taiáçupeba					1.325		1.975
239	V. Aparecida					-		
240	Pilar Velho				2.279	2.502		11.793
241	Jardim Caçula					1.781		1.725
242	Sete Pontes				1.911	191		2.991
243	Parque América					331		208
244	Paranapiacaba				1.096	2.674		6.214
245	Caminhos do Mar				725	1.128		3.921
246	Itararé				90	798		1.098
247	Ressaca				1.137	1.914		6.385
248	São Lourenço				104	323		816
249	Reserva de Caucaia					112		141
250	Jardim das Flores				105	478		831
251	Três Montanhas					3.462		4.620
252	Jardim Belval					208		1.048
253	Ribeirão das Pombas					213		190
254	Quatro Encruzilhadas				63	150		169
Total		3.324.658	3.612.861	14.247.834	3.676.224	5.647.390	2.014.474	29.399.601

ANEXO B – VARIÁVEIS SÓCIO-ECONÔMICAS da Pesquisa O/D 87 - RMSP

Características sócio-econômicas		
Variável	Código *	Descrição
Número de Automóveis	0 de 1 a 9	Sem automóveis (por domicílio) 1 a 9 automóveis (por domicílio)
Número de Pessoas	de 1 a 22	Número de pessoas na família
Renda Familiar	Variável	Número em salários mínimos/1987
Idade	de 4 a 99	Idade em Anos
Sexo	1 2	Homem Mulher
Ocupação Principal		
	01	Estudante
	02	Prendas Domésticas
	03	Aposentado
	04	Sem ocupação, nunca trabalhou
	05	Desempregado temporariamente
	06	Em licença, afastado
	<i>Empregado de firma/ empresa</i>	
	11	Serviços não especializados: office-boy, faxineiro, etc.
	12	Serviços semi-especializados: zelador, telefonista, motorista
	13	Escriturário
	14	Comércio: vendedor, corretor, etc.
	15	Chefe de serviços especializados
	16	Nível superior em escritório
	18	Empregado de nível universitário
	<i>Operário/ trabalhador rural</i>	
	21	Operários sem especialização
	22	Operários semi-especializados
	23	Operários especializados de oficinas: técnicos, artesãos, etc.
	24	Operário especializado (fábrica)
	25	Chefe de operários especializados
	26	Trabalhadores rurais
	27	Trabalhador rural nível capataz
	<i>Profissional autônomo</i>	
	31	Serviços: vendedor, engraxate, etc.
	32	Serviços semi-especializados
	33	Nível de operário especializado
	34	Trabalho intelectual, mental
	35	Técnico altamente especializado
	36	Nível universitário
	<i>Funcionário Público</i>	
	41	Serviço não especializado
	42	Serviço semi-especializado
	43	Escriturário, auxiliar de escritório
	44	Nível universitário, professores
	45	Nível de diretor

Variável	Código *	Descrição
	46	Altos cargos públicos
	47	Militares
	<i>Sócio ou proprietário de</i>	
	51	Pequeno comércio
	52	Firma pequena de serviços
	53	Micro-empresa indústria/ comércio
	54	Firma grande/ média de serviços
	55	Empresa média indústria/ comércio
	56	Grande empresa indústria/ comércio
	57	Agricultor arrendatário de terra
	58	Agricultor proprietário de terra
	59	Grande fazendeiro

Situação Familiar		
	1	Chefe
	2	Cônjuge
	3	Filho
	4	Parente
	5	Agregado
	6	Empregada/Residente
	7	Visitante não Residente

Renda Individual	Variável	Salário médio em salários mínimos (set/1987)
------------------	----------	--

Grau de Instrução		
	1	Analfabeto/ 1.º grau até 4.º série
	2	1.º grau até 4º série concluído
	3	1.º grau concluído
	4	Colegial concluído
	5	Superior concluído

* - Enumeração adotada pelo METRO/87

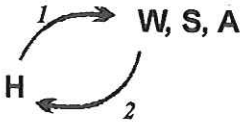
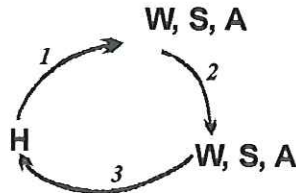
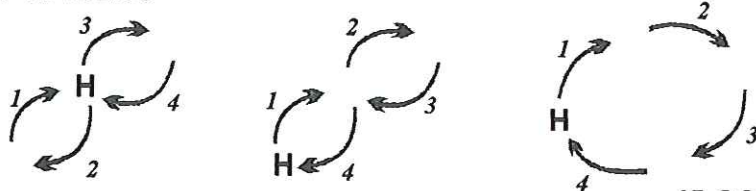
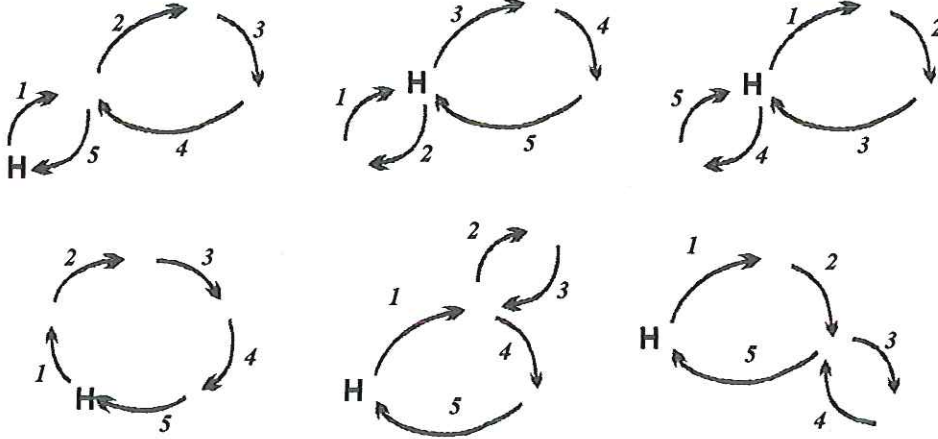
Participação na Atividade / motivo de viagem		
Variável	Código *	Descrição
Motivo origem ou destino	1	Trabalho na indústria
	2	Trabalho no comércio
	3	Trabalho em serviços
	4	Escola ou educação
	5	Compras
	6	Negócios
	7	Médico/dentista/saúde
	8	Recreação ou visitas
	9	Residência

* - Enumeração adotada pelo METRO/87

Características de viagem		
Variável	Código *	Descrição
Número de viagens	0	Número de viagens por dia
Zona de Origem/Destino	1 a 204	Zona de origem/destino
Zona de Trabalho/Escola	1 a 204	Zona do local de Trabalho/Escola
Hora da viagem	Variável	Hora da realização da viagem (minutos)
Modo de Viagem	1	Ônibus diesel
	2	Troleibus
	3	Ônibus Fretado
	4	Escolar
	5	Dirigindo automóvel
	6	Passageiro de Automóvel
	7	Táxi
	8	Lotação
	0	Metrô
	10	Trem
	11	Moto
	12	Bicicleta
	13	A pé
	14	Caminhão
	15	Outros

* - Enumeração adotada pelo METRO/87

ANEXO C – PADRÕES DE VIAGENS E NÚMERO DE COMBINAÇÕES

<p>2 VIAGENS</p>  <p style="text-align: right;">3 COMBINAÇÕES</p>
<p>3 VIAGENS</p>  <p style="text-align: right;">9 COMBINAÇÕES</p>
<p>4 VIAGENS</p>  <p style="text-align: right;">45 COMBINAÇÕES</p>
<p>5 VIAGENS</p>  <p style="text-align: right;">216 COMBINAÇÕES</p>

Observações:

- Padrões de viagens baseadas na Residência;
- Os números correspondem ao sequenciamento da viagem;
- H: Residência; W: Trabalho; S: Escola e A: Atividades.

ANEXO D

RELATÓRIO - S-PLUS ÁRVORE 1 - A-1

Dados de Entrada: 85447 - 005 - 25/50 - TODAS VARIÁVEIS - ESTUDA-TRABALHA

*** Tree Model ***

Classification tree:

tree(formula = PADRAO ~ AUT + RF.SM. + TTPESF + SEXO + SF + IDADE + GI +
SMED.SM. + ESTUDA + TRABALHA, data =
ZDTQ.VGS.PADRAO.limpo.e.completo,
na.action = na.exclude, mincut = 25, minsize = 50, mindev = 0.05)

Variables actually used in tree construction:

[1] "TRABALHA" "ESTUDA" "SF" "TTPESF" "IDADE" "GI"
[7] "AUT" "SMED.SM." "SEXO"

Number of terminal nodes: 20

Residual mean deviance: 2.177 = 186100 / 85460

Misclassification error rate: 0.2914 = 24906 / 85477

node), split, n, deviance, yval, (yprob)

* denotes terminal node

```
1) root 85477 328100 H ( 0.32420 0.0033930 0.011990 0.08553 0.0075460
0.0034860 0.00154400 0.0008540 0.0019890 0.00284300 0.201900 0.0078030
0.0146700 0.00305300 0.00122800 0.00290100 0.00153300 0.0108600 0.00271400
0.2433000 0.0101300 0.01940000 0.0203300 0.01110000 0.00393100 0.0017780 )
2) TRABALHA<0.5 50725 136200 H ( 0.47340 0.0040410 0.014530 0.11180
0.0086550 0.0049480 0.00003943 0.0012420 0.0000000 0.00457400 0.333300
0.0126000 0.0227100 0.00027600 0.00195200 0.00019710 0.00005914 0.0002563
0.00007886 0.0039630 0.0002366 0.00037460 0.0002563 0.00043370 0.00001971
0.0000000 )
4) ESTUDA<0.5 29868 56050 H ( 0.73040 0.0061940 0.022570 0.17680
0.0138600 0.0012720 0.00000000 0.0006696 0.0000000 0.00100400 0.013760
0.0021090 0.0236700 0.00006696 0.00016740 0.00006696 0.00010040 0.0004352
0.00006696 0.0056920 0.0004018 0.00010040 0.0004352 0.00003348 0.00003348
0.0000000 )
8) SF<2.5 18605 41280 H ( 0.65050 0.0075790 0.028490 0.22760
0.0190800 0.0018810 0.00000000 0.0009675 0.0000000 0.00150500 0.018010
0.0029020 0.0339700 0.00010750 0.00021500 0.00005375 0.00010750 0.0003762
0.00010750 0.0052140 0.0006450 0.00010750 0.0005375 0.00005375 0.00005375
0.0000000 )
16) TTPESF<2.5 2915 6607 H ( 0.48340 0.0116600 0.052140 0.40070
0.0377400 0.0000000 0.00000000 0.0000000 0.0000000 0.00034310 0.002058
0.0000000 0.0017150 0.00000000 0.00000000 0.00000000 0.00034310 0.0010290
0.00000000 0.0065180 0.0006861 0.00034310 0.0013720 0.00000000 0.00000000
0.0000000 ) *
17) TTPESF>2.5 15690 33710 H ( 0.68150 0.0068200 0.024090 0.19540
0.0156200 0.0022310 0.00000000 0.0011470 0.0000000 0.00172100 0.020970
0.0034420 0.0399600 0.00012750 0.00025490 0.00006373 0.00006373 0.0002549
0.00012750 0.0049710 0.0006373 0.00006373 0.0003824 0.00006373 0.00006373
0.0000000 )
34) IDADE<49.5 10618 24140 H ( 0.66730 0.0077230 0.023920
0.18580 0.0128100 0.0031080 0.00000000 0.0013190 0.0000000 0.00216600
0.027690 0.0048030 0.0562300 0.00009418 0.00037670 0.00009418 0.00009418
0.0002825 0.00018840 0.0048970 0.0005651 0.00009418 0.0003767 0.00000000
0.00009418 0.0000000 )
68) SF<1.5 1123 2828 H ( 0.59130 0.0160300 0.043630 0.25560
0.0213700 0.0026710 0.00000000 0.0008905 0.0000000 0.00089050 0.010690
0.0026710 0.0213700 0.00000000 0.00000000 0.00000000 0.00000000 0.0008905
```

0.00089050 0.0249300 0.0044520 0.00000000 0.0017810 0.00000000 0.00000000
 0.0000000) *
 69) SF>1.5 9495 21100 H (0.67630 0.0067400 0.021590 0.17760
 0.0118000 0.0031600 0.00000000 0.0013690 0.00000000 0.00231700 0.029700
 0.0050550 0.0603500 0.00010530 0.00042130 0.00010530 0.00010530 0.0002106
 0.00010530 0.0025280 0.0001053 0.00010530 0.0002106 0.00000000 0.00010530
 0.0000000)
 138) IDADE<23.5 901 1380 H (0.76470 0.0055490 0.016650
 0.18310 0.0099890 0.0011100 0.00000000 0.00000000 0.00000000 0.00111000
 0.003330 0.0011100 0.0111000 0.00000000 0.00000000 0.00000000 0.00000000
 0.0000000 0.00111000 0.0011100 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.0000000) *
 139) IDADE>23.5 8594 19590 H (0.66700 0.0068650 0.022110
 0.17700 0.0119900 0.0033740 0.00000000 0.0015130 0.00000000 0.00244400
 0.032460 0.0054690 0.0655100 0.00011640 0.00046540 0.00011640 0.00011640
 0.0002327 0.00000000 0.0026760 0.0001164 0.00011640 0.0002327 0.00000000
 0.00011640 0.0000000)
 278) GI<2.5 6172 13090 H (0.69220 0.0050230 0.019770
 0.17350 0.0105300 0.0027540 0.00000000 0.0008101 0.00000000 0.00210600
 0.029160 0.0045370 0.0554100 0.00016200 0.00016200 0.00000000 0.00016200
 0.0001620 0.00000000 0.0032400 0.0001620 0.00000000 0.0001620 0.00000000
 0.00000000 0.0000000) *
 279) GI>2.5 2422 6388 H (0.60280 0.0115600 0.028080
 0.18580 0.0156900 0.0049550 0.00000000 0.0033030 0.00000000 0.00330300
 0.040880 0.0078450 0.0912500 0.00000000 0.00123900 0.00041290 0.00000000
 0.0004129 0.00000000 0.0012390 0.00000000 0.00041290 0.0004129 0.00000000
 0.00041290 0.0000000) *
 35) IDADE>49.5 5072 9096 H (0.71140 0.0049290 0.024450 0.21550
 0.0214900 0.0003943 0.00000000 0.0007886 0.00000000 0.00078860 0.006901
 0.0005915 0.0059150 0.00019720 0.00000000 0.00000000 0.00000000 0.0001972
 0.00000000 0.0051260 0.0007886 0.00000000 0.0003943 0.00019720 0.00000000
 0.0000000) *
 9) SF>2.5 11263 12920 H (0.86260 0.0039070 0.012790 0.09305
 0.0052380 0.0002664 0.00000000 0.0001776 0.00000000 0.00017760 0.006748
 0.0007991 0.0066590 0.00000000 0.00008879 0.00008879 0.00008879 0.0005327
 0.00000000 0.0064810 0.00000000 0.00008879 0.0002664 0.00000000 0.00000000
 0.0000000) *
 5) ESTUDA>0.5 20857 36310 HSH (0.10530 0.0009589 0.003021 0.01875
 0.0011990 0.0102100 0.00009589 0.0020620 0.00000000 0.00968500 0.791000
 0.0276200 0.0213400 0.00057530 0.00450700 0.00038360 0.00000000 0.00000000
 0.00009589 0.0014860 0.00000000 0.00076710 0.00000000 0.00100700 0.00000000
 0.0000000)
 10) GI<1.5 13014 19960 HSH (0.11280 0.0009221 0.002305 0.01698
 0.0004610 0.0041490 0.00000000 0.0013830 0.00000000 0.00883700 0.810700
 0.0237400 0.0101400 0.00053790 0.00391900 0.00053790 0.00000000 0.00000000
 0.00007684 0.0013060 0.00000000 0.00030740 0.00000000 0.00092210 0.00000000
 0.0000000)
 20) AUT<0.5 7659 10660 HSH (0.12080 0.0006528 0.002481 0.01528
 0.0006528 0.0026110 0.00000000 0.0009140 0.00000000 0.00496100 0.821800
 0.0143600 0.0099230 0.00052230 0.00261100 0.00000000 0.00000000 0.00000000
 0.00000000 0.0019580 0.00000000 0.00013060 0.00000000 0.00039170 0.00000000
 0.0000000) *
 21) AUT>0.5 5355 9127 HSH (0.10140 0.0013070 0.002054 0.01942
 0.0001867 0.0063490 0.00000000 0.0020540 0.00000000 0.01438000 0.794800
 0.0371600 0.0104600 0.00056020 0.00578900 0.00130700 0.00000000 0.00000000
 0.00018670 0.0003735 0.00000000 0.00056020 0.00000000 0.00168100 0.00000000
 0.0000000) *
 11) GI>1.5 7843 15930 HSH (0.09282 0.0010200 0.004208 0.02168
 0.0024230 0.0202700 0.00025500 0.0031880 0.00000000 0.01109000 0.758300
 0.0340400 0.0399100 0.00063750 0.00548300 0.00012750 0.00000000 0.00000000
 0.00012750 0.0017850 0.00000000 0.00153000 0.00000000 0.00114800 0.00000000
 0.0000000) *
 3) TRABALHA>0.5 34752 116600 HWH (0.10630 0.0024460 0.008287 0.04713
 0.0059280 0.0013520 0.00374100 0.0002878 0.0048920 0.00031650 0.010190
 0.0008057 0.0029350 0.00710800 0.00017270 0.00684900 0.00368300 0.0263300

0.00656100 0.5926000 0.0245700 0.04716000 0.0496400 0.02667000 0.00964000
0.0043740)

6) ESTUDA<0.5 30870 88030 HWH (0.11620 0.0027210 0.008941 0.05164
0.0064460 0.0005507 0.00411400 0.0001620 0.0054100 0.00016200 0.003563
0.0006803 0.0026890 0.00191100 0.00003239 0.00492400 0.00388700 0.0286700
0.00712700 0.6500000 0.0262100 0.00272100 0.0537700 0.00243000 0.01043000
0.0046970)

12) SF<5.5 30092 84170 HWH (0.09753 0.0027910 0.008939 0.05081
0.0064800 0.0005317 0.00422000 0.0001662 0.0055160 0.00016620 0.003157
0.0005982 0.0022270 0.00196100 0.00003323 0.00501800 0.00398800 0.0294100
0.00731100 0.6663000 0.0268500 0.00279100 0.0551600 0.00249200 0.01070000
0.0048190)

24) SMED.SM.<5.55 19625 50810 HWH (0.11200 0.0018340 0.007745
0.05442 0.0058090 0.0003567 0.00341400 0.0001529 0.0040760 0.00015290
0.003210 0.0004586 0.0028030 0.00107000 0.00005096 0.00265000 0.00295500
0.0237500 0.00494300 0.6818000 0.0195200 0.00224200 0.0527900 0.00193600
0.00713400 0.0027010)

48) SMED.SM.<1.4 5235 14740 HWH (0.16830 0.0021010 0.009360
0.06991 0.0057310 0.0005731 0.00439400 0.0000000 0.0028650 0.00038200
0.004202 0.0005731 0.0057310 0.00191000 0.00000000 0.00458500 0.00191000
0.0166200 0.00554000 0.6128000 0.0147100 0.00114600 0.0578800 0.00191000
0.00439400 0.0024830) *

49) SMED.SM.>1.4 14390 35710 HWH (0.09152 0.0017370 0.007158
0.04878 0.0058370 0.0002780 0.00305800 0.0002085 0.0045170 0.00006949
0.002849 0.0004170 0.0017370 0.00076440 0.00006949 0.00194600 0.00333600
0.0263400 0.00472600 0.7069000 0.0212600 0.00264100 0.0509400 0.00194600
0.00813100 0.0027800)

98) SF<2.5 9009 23820 HWH (0.10100 0.0019980 0.009879 0.06138
0.0067710 0.0004440 0.00399600 0.0003330 0.0049950 0.00000000 0.003774
0.0006660 0.0022200 0.00077700 0.00011100 0.00288600 0.00310800 0.0246400
0.00444000 0.6810000 0.0192000 0.00266400 0.0490600 0.00222000 0.00943500
0.0029970)

196) SEXO<1.5 5742 14140 HWH (0.08063 0.0015670 0.007489
0.05904 0.0062700 0.0003483 0.00383100 0.0000000 0.0052250 0.00000000
0.002438 0.0003483 0.0010450 0.00034830 0.00000000 0.00191600 0.00243800
0.0212500 0.00522500 0.7095000 0.0182900 0.00209000 0.0548600 0.00104500
0.01045000 0.0043540) *

197) SEXO>1.5 3267 9502 HWH (0.13680 0.0027550 0.014080
0.06550 0.0076520 0.0006122 0.00428500 0.0009183 0.0045910 0.00000000
0.006122 0.0012240 0.0042850 0.00153000 0.00030610 0.00459100 0.00428500
0.0306100 0.00306100 0.6309000 0.0208100 0.00367300 0.0388700 0.00428500
0.00765200 0.0006122) *

99) SF>2.5 5381 11660 HWH (0.07564 0.0013010 0.002602 0.02769
0.0042740 0.0000000 0.00148700 0.0000000 0.0037170 0.00018580 0.001301
0.0000000 0.0009292 0.00074340 0.00000000 0.00037170 0.00371700 0.0291800
0.00520300 0.7504000 0.0247200 0.00260200 0.0540800 0.00148700 0.00594700
0.0024160) *

25) SMED.SM.>5.55 10467 32690 HWH (0.07041 0.0045860 0.011180
0.04404 0.0077390 0.0008598 0.00573200 0.0001911 0.0082160 0.00019110
0.003057 0.0008598 0.0011460 0.00363000 0.00000000 0.00945800 0.00592300
0.0400300 0.01175000 0.6372000 0.0406000 0.00382200 0.0596200 0.00353500
0.01739000 0.0087900)

50) AUT<0.5 3356 8797 HWH (0.07151 0.0047680 0.010430 0.04559
0.0056620 0.0000000 0.00476800 0.0002980 0.0065550 0.00000000 0.001490
0.0008939 0.0011920 0.00059590 0.00000000 0.00119200 0.00476800 0.0384400
0.00566200 0.7044000 0.0235400 0.00387400 0.0372500 0.00357600 0.01698000
0.0065550) *

51) AUT>0.5 7111 23660 HWH (0.06989 0.0045000 0.011530 0.04331
0.0087190 0.0012660 0.00618800 0.0001406 0.0090000 0.00028130 0.003797
0.0008438 0.0011250 0.00506300 0.00000000 0.01336000 0.00646900 0.0407800
0.01463000 0.6055000 0.0486600 0.00379700 0.0701700 0.00351600 0.01758000
0.0098440)

102) SEXO<1.5 5466 17680 HWH (0.05982 0.0049400 0.010790
0.03988 0.0080500 0.0010980 0.00622000 0.0000000 0.0095130 0.00018290
0.003293 0.0003659 0.0007318 0.00439100 0.00000000 0.01500000 0.00640300

0.0402500 0.01665000 0.6220000 0.0492100 0.00365900 0.0676900 0.00237800
 0.01628000 0.0111600)
 204) SMED.SM.<13.95 3299 10000 HWH (0.06275 0.0048500
 0.011220 0.04577 0.0084870 0.0012120 0.00515300 0.0000000 0.0066690
 0.00030310 0.003031 0.0000000 0.0003031 0.00424400 0.00000000 0.00939700
 0.00363700 0.0366800 0.01303000 0.6469000 0.0469800 0.00454700 0.0624400
 0.00121200 0.01182000 0.0093970) *
 205) SMED.SM.>13.95 2167 7590 HWH (0.05538 0.0050760
 0.010150 0.03092 0.0073830 0.0009229 0.00784500 0.0000000 0.0138400
 0.00000000 0.003692 0.0009229 0.0013840 0.00461500 0.00000000 0.02353000
 0.01061000 0.0456900 0.02215000 0.5842000 0.0526100 0.00230700 0.0756800
 0.00415300 0.02307000 0.0138400) *
 103) SEXO>1.5 1645 5884 HWH (0.10330 0.0030400 0.013980
 0.05471 0.0109400 0.0018240 0.00607900 0.0006079 0.0072950 0.00060790
 0.005471 0.0024320 0.0024320 0.00729500 0.00000000 0.00790300 0.006668700
 0.0425500 0.00790300 0.5508000 0.0468100 0.00425500 0.0784200 0.00729500
 0.02188000 0.0054710) *
 13) SF>5.5 778 1099 H (0.83680 0.0000000 0.008997 0.08355
 0.0051410 0.0012850 0.00000000 0.0000000 0.0012850 0.00000000 0.019280
 0.0038560 0.0205700 0.00000000 0.00000000 0.00128500 0.00000000 0.0000000
 0.00000000 0.0167100 0.0012850 0.00000000 0.0000000 0.00000000 0.00000000
 0.0000000) *
 7) ESTUDA>0.5 3882 14640 HWHSH (0.02808 0.0002576 0.003091 0.01133
 0.0018030 0.0077280 0.00077280 0.0012880 0.0007728 0.00154600 0.062850
 0.0018030 0.0048940 0.04843000 0.00128800 0.02215000 0.00206100 0.0077280
 0.00206100 0.1365000 0.0115900 0.40060000 0.0167400 0.21950000 0.00334900
 0.0018030) *

ANEXO E

RELATÓRIO E GRÁFICO – S-PLUS ÁRVORE 2 – A-2

Dados de Entrada: 85447 - 001 - 25/50 - TODAS VARIÁVEIS - ESTUDA-TRABALHA

*** Tree Model ***

Classification tree:

tree(formula = PADRAO ~ AUT + RF.SM. + TPESF + SEXO + SF + IDADE + GI +
SMED.SM. + ESTUDA + TRABALHA, data =

ZDTQ.VGS.PADRAO.limpo.e.completo,

na.action = na.exclude, mincut = 25, minsize = 50, mindev = 0.01)

Number of terminal nodes: 101

Residual mean deviance: 2.107 = 179900 / 85380

Misclassification error rate: 0.2861 = 24455 / 85477

node), split, n, deviance, yval, (yprob)

* denotes terminal node

1) root 85477 328100.00 H (0.32420 0.0033930 0.0119900 0.085530
0.0075460 0.0034860 0.00154400 0.0008540 0.0019890 0.00284300 0.2019000
0.0078030 0.0146700 0.00305300 0.00122800 0.00290100 0.00153300 0.0108600
0.00271400 0.2433000 0.0101300 0.01940000 0.0203300 0.01110000 0.00393100
0.0017780)

2) TRABALHA<0.5 50725 136200.00 H (0.47340 0.0040410 0.0145300
0.111800 0.0086550 0.0049480 0.00003943 0.0012420 0.0000000 0.00457400
0.3333000 0.0126000 0.0227100 0.00027600 0.00195200 0.00019710 0.00005914
0.0002563 0.00007886 0.0039630 0.0002366 0.00037460 0.0002563 0.00043370
0.00001971 0.0000000)

4) ESTUDA<0.5 29868 56050.00 H (0.73040 0.0061940 0.0225700
0.176800 0.0138600 0.0012720 0.00000000 0.0006696 0.0000000 0.00100400
0.0137600 0.0021090 0.0236700 0.00006696 0.00016740 0.00006696 0.00010040
0.0004352 0.00006696 0.0056920 0.0004018 0.00010040 0.0004352 0.00003348
0.00003348 0.0000000)

8) SF<2.5 18605 41280.00 H (0.65050 0.0075790 0.0284900 0.227600
0.0190800 0.0018810 0.00000000 0.0009675 0.0000000 0.00150500 0.0180100
0.0029020 0.0339700 0.00010750 0.00021500 0.00005375 0.00010750 0.0003762
0.00010750 0.0052140 0.0006450 0.00010750 0.0005375 0.00005375 0.00005375
0.0000000)

16) TPESF<2.5 2915 6607.00 H (0.48340 0.0116600 0.0521400
0.400700 0.0377400 0.00000000 0.00000000 0.0000000 0.0000000 0.00034310
0.0020580 0.0000000 0.0017150 0.00000000 0.00000000 0.00000000 0.00034310
0.0010290 0.00000000 0.0065180 0.0006861 0.00034310 0.0013720 0.00000000
0.00000000 0.0000000)

32) SMED.SM.<0.95 1938 3814.00 H (0.58570 0.0061920
0.0387000 0.334900 0.0268300 0.00000000 0.00000000 0.0000000 0.0000000
0.00000000 0.0005160 0.0000000 0.0005160 0.00000000 0.00000000 0.00000000
0.00000000 0.0005160 0.00000000 0.0036120 0.0005160 0.00051600 0.0015480
0.00000000 0.00000000 0.0000000)

64) TPESF<1.5 143 274.80 HAH (0.23080 0.0000000
0.0279700 0.650300 0.0839200 0.00000000 0.00000000 0.0000000 0.0000000
0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.00000000 0.00000000
0.00000000 0.0069930 0.00000000 0.0000000 0.0000000 0.00000000 0.0000000
0.00000000 0.00000000 0.0000000) *

65) TPESF>1.5 1795 3438.00 H (0.61390 0.0066850 0.0395500
0.309700 0.0222800 0.0000000 0.00000000 0.0000000 0.0000000 0.0000000
0.0005571 0.0000000 0.0005571 0.00000000 0.00000000 0.00000000 0.0000000
0.0000000 0.00000000 0.0039000 0.0005571 0.00055710 0.0016710 0.00000000
0.00000000 0.0000000)

130) SEXO<1.5 240 498.10 H (0.48750 0.0125000 0.0583300
0.420800 0.0083330 0.0000000 0.00000000 0.0000000 0.0000000 0.00000000
0.00000000

0.0041670 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0083330 0.0000000
0.0000000 0.0000000) *
131) SEXO>1.5 1555 2904.00 H (0.63340 0.0057880 0.0366600
0.292600 0.0244400 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0006431 0.0000000 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0045020 0.0006431 0.0006431 0.0006431 0.0000000
0.0000000 0.0000000) *
33) SMED.SM.>0.95 977 2521.00 HAH (0.28050 0.0225200
0.0788100 0.531200 0.0593700 0.0000000 0.0000000 0.0000000 0.0000000
0.00102400 0.0051180 0.0000000 0.0040940 0.0000000 0.0000000 0.0000000
0.00102400 0.0020470 0.0000000 0.0122800 0.0010240 0.0000000 0.0010240
0.0000000 0.0000000 0.0000000) *
17) TTPESF>2.5 15690 33710.00 H (0.68150 0.0068200 0.0240900
0.195400 0.0156200 0.0022310 0.0000000 0.0011470 0.0000000 0.00172100
0.0209700 0.0034420 0.0399600 0.00012750 0.00025490 0.00006373 0.00006373
0.0002549 0.00012750 0.0049710 0.0006373 0.00006373 0.0003824 0.00006373
0.00006373 0.0000000) *
34) IDADE<49.5 10618 24140.00 H (0.66730 0.0077230 0.0239200
0.185800 0.0128100 0.0031080 0.0000000 0.0013190 0.0000000 0.00216600
0.0276900 0.0048030 0.0562300 0.00009418 0.00037670 0.00009418 0.00009418
0.0002825 0.00018840 0.0048970 0.0005651 0.00009418 0.0003767 0.0000000
0.00009418 0.0000000) *
68) SF<1.5 1123 2828.00 H (0.59130 0.0160300 0.0436300
0.255600 0.0213700 0.0026710 0.0000000 0.0008905 0.0000000 0.00089050
0.0106900 0.0026710 0.0213700 0.0000000 0.0000000 0.0000000 0.0000000
0.0008905 0.0008905 0.0249300 0.0044520 0.0000000 0.0017810 0.0000000
0.0000000 0.0000000) *
69) SF>1.5 9495 21100.00 H (0.67630 0.0067400 0.0215900
0.177600 0.0118000 0.0031600 0.0000000 0.0013690 0.0000000 0.00231700
0.0297000 0.0050550 0.0603500 0.00010530 0.00042130 0.00010530 0.00010530
0.0002106 0.00010530 0.0025280 0.0001053 0.00010530 0.0002106 0.0000000
0.00010530 0.0000000) *
138) IDADE<23.5 901 1380.00 H (0.76470 0.0055490
0.0166500 0.183100 0.0099890 0.0011100 0.0000000 0.0000000 0.0000000
0.00111000 0.0033300 0.0011100 0.0111000 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0011100 0.0011100 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000) *
139) IDADE>23.5 8594 19590.00 H (0.66700 0.0068650
0.0221100 0.177000 0.0119900 0.0033740 0.0000000 0.0015130 0.0000000
0.00244400 0.0324600 0.0054690 0.0655100 0.00011640 0.00046540 0.00011640
0.00011640 0.0002327 0.0000000 0.0026760 0.0001164 0.00011640 0.0002327
0.0000000 0.00011640 0.0000000) *
278) GI<2.5 6172 13090.00 H (0.69220 0.0050230 0.0197700
0.173500 0.0105300 0.0027540 0.0000000 0.0008101 0.0000000 0.00210600
0.0291600 0.0045370 0.0554100 0.00016200 0.00016200 0.0000000 0.00016200
0.0001620 0.0000000 0.0032400 0.0001620 0.0000000 0.0001620 0.0000000
0.0000000 0.0000000) *
556) IDADE<36.5 3396 7471.00 H (0.67520 0.0047110
0.0215000 0.172600 0.0085390 0.0023560 0.0000000 0.0011780 0.0000000
0.00265000 0.0326900 0.0050060 0.0709700 0.0000000 0.00029450 0.0000000
0.00029450 0.0002945 0.0000000 0.0014720 0.0002945 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000) *
1112) AUT<0.5 2344 4924.00 H (0.69240 0.0029860
0.0256000 0.163000 0.0055460 0.0017060 0.0000000 0.0012800 0.0000000
0.00298600 0.0298600 0.0029860 0.0691100 0.0000000 0.0000000 0.0000000
0.00042660 0.0000000 0.0000000 0.0021330 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000) *
2224) RF.SM.<5.3 1385 2647.00 H (0.71910 0.0028880
0.0209400 0.154500 0.0021660 0.0007220 0.0000000 0.0007220 0.0000000
0.00288800 0.0267100 0.0021660 0.0671500 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
0.0000000 0.0000000 0.0000000) *
2225) RF.SM.>5.3 959 2246.00 H (0.65380 0.0031280
0.0323300 0.175200 0.0104300 0.0031280 0.0000000 0.0020860 0.0000000

0.00312800 0.0344100 0.0041710 0.0719500 0.00000000 0.00000000 0.00000000
 0.00104300 0.00000000 0.00000000 0.0052140 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 1113) AUT>0.5 1052 2501.00 H (0.63690 0.0085550
 0.0123600 0.193900 0.0152100 0.0038020 0.00000000 0.0009506 0.00000000
 0.00190100 0.0389700 0.0095060 0.0751000 0.00000000 0.00095060 0.00000000
 0.00000000 0.0009506 0.00000000 0.00000000 0.0009506 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 557) IDADE>36.5 2776 5557.00 H (0.71290 0.0054030
 0.0176500 0.174700 0.0129700 0.0032420 0.00000000 0.0003602 0.00000000
 0.00144100 0.0248600 0.0039630 0.0363800 0.00036020 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.0054030 0.00000000 0.00000000 0.0003602
 0.00000000 0.00000000 0.00000000) *
 1114) TPESF<6.5 2265 4763.00 H (0.69180 0.0061810
 0.0181000 0.185000 0.0136900 0.0035320 0.00000000 0.0004415 0.00000000
 0.00088300 0.0278100 0.0044150 0.0415000 0.00044150 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.0057400 0.00000000 0.00000000 0.0004415
 0.00000000 0.00000000 0.00000000) *
 2228) TPESF<4.5 1135 2486.00 H (0.67400 0.0079300
 0.0193800 0.197400 0.0141000 0.0052860 0.00000000 0.00000000 0.00000000
 0.00176200 0.0220300 0.0035240 0.0422900 0.00088110 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.0114500 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 2229) TPESF>4.5 1130 2243.00 H (0.70970 0.0044250
 0.0168100 0.172600 0.0132700 0.0017700 0.00000000 0.0008850 0.00000000
 0.00000000 0.0336300 0.0053100 0.0407100 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.0008850
 0.00000000 0.00000000 0.00000000) *
 1115) TPESF>6.5 511 755.60 H (0.80630 0.0019570
 0.0156600 0.129200 0.0097850 0.0019570 0.00000000 0.00000000 0.00000000
 0.00391400 0.0117400 0.0019570 0.0137000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.0039140 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 279) GI>2.5 2422 6388.00 H (0.60280 0.0115600 0.0280800
 0.185800 0.0156900 0.0049550 0.00000000 0.0033030 0.00000000 0.00330300
 0.0408800 0.0078450 0.0912500 0.00000000 0.00123900 0.00041290 0.00000000
 0.0004129 0.00000000 0.0012390 0.00000000 0.00041290 0.0004129 0.00000000
 0.00041290 0.00000000) *
 558) IDADE<42.5 2093 5507.00 H (0.60010 0.0109900
 0.0286700 0.180600 0.0138600 0.0057330 0.00000000 0.0033440 0.00000000
 0.00382200 0.0415700 0.0090780 0.1003000 0.00000000 0.00095560 0.00000000
 0.00000000 0.0004778 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00047780 0.00000000) *
 1116) AUT<1.5 1735 4268.00 H (0.62940 0.0103700
 0.0270900 0.180400 0.0115300 0.0040350 0.00000000 0.0023050 0.00000000
 0.00230500 0.0351600 0.0069160 0.0881800 0.00000000 0.00115300 0.00000000
 0.00000000 0.0005764 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00057640 0.00000000) *
 2232) RF.SM.<5.55 492 982.20 H (0.70120 0.0020330
 0.0162600 0.166700 0.0040650 0.00000000 0.00000000 0.00000000 0.00000000
 0.00609800 0.0365900 0.0040650 0.0630100 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 2233) RF.SM.>5.55 1243 3246.00 H (0.60100
 0.0136800 0.0313800 0.185800 0.0144800 0.0056320 0.00000000 0.0032180
 0.00000000 0.00080450 0.0345900 0.0080450 0.0981500 0.00000000 0.00160900
 0.00000000 0.00000000 0.0008045 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00080450 0.00000000) *
 1117) AUT>1.5 358 1181.00 H (0.45810 0.0139700
 0.0363100 0.181600 0.0251400 0.0139700 0.00000000 0.0083800 0.00000000
 0.01117000 0.0726300 0.0195500 0.1592000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 559) IDADE>42.5 329 820.30 H (0.62010 0.0152000
 0.0243200 0.218800 0.0273600 0.00000000 0.00000000 0.0030400 0.00000000

0.00000000 0.0364700 0.00000000 0.0334300 0.00000000 0.00304000 0.00304000
0.00000000 0.00000000 0.00000000 0.0091190 0.00000000 0.00304000 0.00304000
0.00000000 0.00000000 0.00000000) *
35) IDADE>49.5 5072 9096.00 H (0.71140 0.0049290 0.0244500
0.215500 0.0214900 0.0003943 0.00000000 0.0007886 0.00000000 0.00078860
0.0069010 0.0005915 0.0059150 0.00019720 0.00000000 0.00000000 0.00000000
0.0001972 0.00000000 0.0051260 0.0007886 0.00000000 0.0003943 0.00019720
0.00000000 0.00000000)
70) SMED.SM.<2.55 4122 6652.00 H (0.74920 0.0048520
0.0177100 0.193100 0.0148000 0.0004852 0.00000000 0.0004852 0.00000000
0.00097040 0.0065500 0.0002426 0.0053370 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000 0.0050950 0.0004852 0.00000000 0.0004852
0.00024260 0.00000000 0.00000000)
140) AUT<0.5 2049 2752.00 H (0.78870 0.0024400 0.0102500
0.177200 0.0058570 0.0004880 0.00000000 0.00000000 0.00000000 0.00048800
0.0073210 0.0004880 0.0039040 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.0029280 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000) *
141) AUT>0.5 2073 3827.00 H (0.71010 0.0072360 0.0250800
0.208900 0.0236400 0.0004824 0.00000000 0.0009648 0.00000000 0.00144700
0.0057890 0.00000000 0.0067530 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.0072360 0.0009648 0.00000000 0.0009648 0.00048240
0.00000000 0.00000000)
282) GI<1.5 848 1301.00 H (0.76060 0.0035380 0.0106100
0.189900 0.0129700 0.00000000 0.00000000 0.0023580 0.00000000 0.00117900
0.0023580 0.00000000 0.0094340 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.0058960 0.00000000 0.00000000 0.0011790 0.00000000
0.00000000 0.00000000) *
283) GI>1.5 1225 2482.00 H (0.67510 0.0097960 0.0351000
0.222000 0.0310200 0.0008163 0.00000000 0.00000000 0.00000000 0.00163300
0.0081630 0.00000000 0.0048980 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.0081630 0.0016330 0.00000000 0.0008163 0.00081630
0.00000000 0.00000000) *
71) SMED.SM.>2.55 950 2261.00 H (0.54740 0.0052630
0.0536800 0.312600 0.0505300 0.00000000 0.00000000 0.0021050 0.00000000
0.00000000 0.0084210 0.0021050 0.0084210 0.00105300 0.00000000 0.00000000
0.00000000 0.0010530 0.00000000 0.0052630 0.0021050 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
9) SF>2.5 11263 12920.00 H (0.86260 0.0039070 0.0127900 0.093050
0.0052380 0.0002664 0.00000000 0.0001776 0.00000000 0.00017760 0.0067480
0.0007991 0.0066590 0.00000000 0.00008879 0.00008879 0.00008879 0.0005327
0.00000000 0.0064810 0.00000000 0.00008879 0.0002664 0.00000000 0.00000000
0.00000000)
18) GI<1.5 7795 6405.00 H (0.90920 0.0019240 0.0070560
0.063630 0.0020530 0.0001283 0.00000000 0.0002566 0.00000000 0.00000000
0.0057730 0.0003849 0.0056450 0.00000000 0.00000000 0.00012830 0.00000000
0.0001283 0.00000000 0.0034640 0.00000000 0.00000000 0.0002566 0.00000000
0.00000000 0.00000000)
36) IDADE<12.5 4989 3264.00 H (0.93270 0.0018040 0.0058130
0.046300 0.0014030 0.00000000 0.00000000 0.0004009 0.00000000 0.00000000
0.0054120 0.0004009 0.0040090 0.00000000 0.00000000 0.00020040 0.00000000
0.0002004 0.00000000 0.0014030 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000) *
37) IDADE>12.5 2806 3032.00 H (0.86740 0.0021380 0.0092660
0.094440 0.0032070 0.0003564 0.00000000 0.00000000 0.00000000 0.00000000
0.0064150 0.0003564 0.0085530 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.0071280 0.00000000 0.00000000 0.0007128 0.00000000
0.00000000 0.00000000) *
19) GI>1.5 3468 6043.00 H (0.75780 0.0083620 0.0256600
0.159200 0.0124000 0.0005767 0.00000000 0.00000000 0.00000000 0.00057670
0.0089390 0.0017300 0.0089390 0.00000000 0.00028840 0.00000000 0.00028840
0.0014420 0.00000000 0.0132600 0.00000000 0.00028840 0.0002884 0.00000000
0.00000000 0.00000000)
38) SMED.SM.<1.25 3044 4991.00 H (0.76970 0.0088700
0.0256200 0.157000 0.0115000 0.0006570 0.00000000 0.00000000 0.00000000

0.00000000 0.0098550 0.0019710 0.0095270 0.00000000 0.00032850 0.00000000
 0.00032850 0.0003285 0.00000000 0.0036140 0.00000000 0.00032850 0.0003285
 0.00000000 0.00000000 0.0000000)
 76) GI<2.5 1934 2687.00 H (0.81900 0.0082730 0.0175800
 0.124600 0.0056880 0.0000000 0.00000000 0.0000000 0.0000000 0.0000000
 0.0077560 0.0020680 0.0093070 0.00000000 0.00051710 0.00000000 0.00051710
 0.0005171 0.00000000 0.0031020 0.0000000 0.00051710 0.0005171 0.00000000
 0.00000000 0.0000000) *
 77) GI>2.5 1110 2214.00 H (0.68380 0.0099100 0.0396400
 0.213500 0.0216200 0.0018020 0.00000000 0.0000000 0.0000000 0.0000000
 0.0135100 0.0018020 0.0099100 0.00000000 0.00000000 0.00000000 0.00000000
 0.0000000 0.00000000 0.0045050 0.0000000 0.00000000 0.0000000 0.0000000
 0.00000000 0.0000000) *
 39) SMED.SM.>1.25 424 916.90 H (0.67220 0.0047170 0.0259400
 0.174500 0.0188700 0.0000000 0.00000000 0.0000000 0.0000000 0.00471700
 0.0023580 0.0000000 0.0047170 0.00000000 0.00000000 0.00000000 0.00000000
 0.0094340 0.00000000 0.0825500 0.0000000 0.00000000 0.0000000 0.00000000
 0.00000000 0.0000000) *
 5) ESTUDA>0.5 20857 36310.00 HSH (0.10530 0.0009589 0.0030210
 0.018750 0.0011990 0.0102100 0.00009589 0.0020620 0.0000000 0.00968500
 0.7910000 0.0276200 0.0213400 0.00057530 0.00450700 0.00038360 0.00000000
 0.0000000 0.00009589 0.0014860 0.0000000 0.00076710 0.0000000 0.00100700
 0.00000000 0.0000000)
 10) GI<1.5 13014 19960.00 HSH (0.11280 0.0009221 0.0023050
 0.016980 0.0004610 0.0041490 0.00000000 0.0013830 0.0000000 0.00883700
 0.8107000 0.0237400 0.0101400 0.00053790 0.00391900 0.00053790 0.00000000
 0.0000000 0.00007684 0.0013060 0.0000000 0.00030740 0.0000000 0.00092210
 0.00000000 0.0000000)
 20) AUT<0.5 7659 10660.00 HSH (0.12080 0.0006528 0.0024810
 0.015280 0.0006528 0.0026110 0.00000000 0.0009140 0.0000000 0.00496100
 0.8218000 0.0143600 0.0099230 0.00052230 0.00261100 0.00000000 0.00000000
 0.0000000 0.00000000 0.0019580 0.0000000 0.00013060 0.0000000 0.00039170
 0.00000000 0.0000000)
 40) TTPESF<4.5 2218 3410.00 HSH (0.10500 0.0013530 0.0036070
 0.021640 0.0004509 0.0040580 0.00000000 0.0018030 0.0000000 0.01127000
 0.8142000 0.0216400 0.0076650 0.00135300 0.00270500 0.00000000 0.00000000
 0.0000000 0.00000000 0.0027050 0.0000000 0.00000000 0.0000000 0.00045090
 0.00000000 0.0000000)
 80) IDADE<9.5 1582 2350.00 HSH (0.11250 0.0006321
 0.0037930 0.019600 0.0006321 0.0044250 0.00000000 0.0018960 0.0000000
 0.01391000 0.8154000 0.0164300 0.0044250 0.00000000 0.00379300 0.00000000
 0.00000000 0.0000000 0.00000000 0.0025280 0.0000000 0.00000000 0.0000000
 0.00000000 0.00000000 0.0000000) *
 81) IDADE>9.5 636 1022.00 HSH (0.08648 0.0031450 0.0031450
 0.026730 0.0000000 0.0031450 0.00000000 0.0015720 0.0000000 0.00471700
 0.8113000 0.0345900 0.0157200 0.00471700 0.00000000 0.00000000 0.00000000
 0.0000000 0.00000000 0.0031450 0.0000000 0.00000000 0.0000000 0.00157200
 0.00000000 0.0000000) *
 41) TTPESF>4.5 5441 7186.00 HSH (0.12720 0.0003676 0.0020220
 0.012680 0.0007352 0.0020220 0.00000000 0.0005514 0.0000000 0.00238900
 0.8248000 0.0113900 0.0108400 0.00018380 0.00257300 0.00000000 0.00000000
 0.0000000 0.00000000 0.0016540 0.0000000 0.00018380 0.0000000 0.00036760
 0.00000000 0.0000000)
 82) IDADE<7.5 1546 1957.00 HSH (0.13390 0.0006468
 0.0032340 0.014880 0.0012940 0.0006468 0.00000000 0.0006468 0.0000000
 0.00452800 0.8260000 0.0064680 0.0032340 0.00064680 0.00388100 0.00000000
 0.00000000 0.0000000 0.00000000 0.0000000 0.0000000 0.00000000 0.0000000
 0.00000000 0.00000000 0.0000000) *
 83) IDADE>7.5 3895 5187.00 HSH (0.12450 0.0002567
 0.0015400 0.011810 0.0005135 0.0025670 0.00000000 0.0005135 0.0000000
 0.00154000 0.8244000 0.0133500 0.0138600 0.00000000 0.00205400 0.00000000
 0.00000000 0.0000000 0.00000000 0.0023110 0.0000000 0.00025670 0.0000000
 0.00051350 0.00000000 0.0000000)
 166) RF.SM.<10.35 3287 4388.00 HSH (0.12960 0.0003042
 0.0018250 0.011560 0.0006085 0.0027380 0.00000000 0.0006085 0.0000000

0.00030420 0.8211000 0.0121700 0.0133900 0.00000000 0.00243400 0.00000000
 0.00000000 0.00000000 0.00000000 0.0024340 0.00000000 0.00030420 0.00000000
 0.00060850 0.00000000 0.00000000)
 332) IDADE<13.5 3076 3983.00 HSH (0.12910 0.0003251
 0.0016250 0.011380 0.0006502 0.0019510 0.00000000 0.0006502 0.00000000
 0.00000000 0.8248000 0.0113800 0.0136500 0.00000000 0.00227600 0.00000000
 0.00000000 0.00000000 0.00000000 0.0016250 0.00000000 0.00000000 0.00000000
 0.00065020 0.00000000 0.00000000)
 664) TTPESF<5.5 1074 1421.00 HSH (0.11820 0.0009311
 0.0009311 0.016760 0.0009311 0.0018620 0.00000000 0.00000000 0.00000000
 0.00000000 0.8259000 0.0204800 0.0102400 0.00000000 0.00093110 0.00000000
 0.00000000 0.00000000 0.00000000 0.0018620 0.00000000 0.00000000 0.00000000
 0.00093110 0.00000000 0.00000000) *
 665) TTPESF>5.5 2002 2538.00 HSH (0.13490 0.0000000
 0.0019980 0.008492 0.0004995 0.0019980 0.00000000 0.0009990 0.00000000
 0.00000000 0.8242000 0.0064940 0.0154800 0.00000000 0.00299700 0.00000000
 0.00000000 0.00000000 0.00000000 0.0014990 0.00000000 0.00000000 0.00000000
 0.00049950 0.00000000 0.00000000) *
 333) IDADE>13.5 211 376.20 HSH (0.13740 0.0000000
 0.0047390 0.014220 0.0000000 0.0142200 0.00000000 0.00000000 0.00000000
 0.00473900 0.7678000 0.0237000 0.0094790 0.00000000 0.00473900 0.00000000
 0.00000000 0.00000000 0.00000000 0.0142200 0.00000000 0.00473900 0.00000000
 0.00000000 0.00000000 0.00000000) *
 167) RF.SM.>10.35 608 770.50 HSH (0.09704 0.0000000
 0.00000000 0.013160 0.0000000 0.0016450 0.00000000 0.00000000 0.00000000
 0.00822400 0.8421000 0.0197400 0.0164500 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000 0.0016450 0.00000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.00000000) *
 21) AUT>0.5 5355 9127.00 HSH (0.10140 0.0013070 0.0020540
 0.019420 0.0001867 0.0063490 0.00000000 0.0020540 0.00000000 0.01438000
 0.7948000 0.0371600 0.0104600 0.00056020 0.00578900 0.00130700 0.00000000
 0.00000000 0.00018670 0.0003735 0.00000000 0.00056020 0.00000000 0.00168100
 0.00000000 0.00000000)
 42) RF.SM.<20.1 4064 6416.00 HSH (0.09966 0.0012300
 0.0012300 0.021410 0.0002461 0.0049210 0.00000000 0.0014760 0.00000000
 0.01181000 0.8100000 0.0315000 0.0083660 0.00024610 0.00369100 0.00073820
 0.00000000 0.00000000 0.00024610 0.0002461 0.00000000 0.00073820 0.00000000
 0.00221500 0.00000000 0.00000000)
 84) TTPESF<5.5 2846 4783.00 HSH (0.09663 0.0017570
 0.0017570 0.024240 0.0003514 0.0049190 0.00000000 0.0021080 0.00000000
 0.01335000 0.7980000 0.0383000 0.0098380 0.00035140 0.00527100 0.00070270
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00105400 0.00000000
 0.00140500 0.00000000 0.00000000)
 168) AUT<1.5 2450 4096.00 HSH (0.10120 0.0020410
 0.0020410 0.025710 0.0004082 0.0044900 0.00000000 0.0020410 0.00000000
 0.00898000 0.7967000 0.0379600 0.0093880 0.00040820 0.00530600 0.00081630
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00122400 0.00000000
 0.00122400 0.00000000 0.00000000)
 336) IDADE<14.5 2413 3981.00 HSH (0.10190 0.0020720
 0.0016580 0.024870 0.0004144 0.0037300 0.00000000 0.0020720 0.00000000
 0.00870300 0.7982000 0.0385400 0.0095320 0.00000000 0.00497300 0.00082880
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00124300 0.00000000
 0.00124300 0.00000000 0.00000000)
 672) RF.SM.<3.55 506 673.50 HSH (0.07708 0.0019760
 0.0019760 0.029640 0.0000000 0.0019760 0.00000000 0.00000000 0.00000000
 0.00790500 0.8478000 0.0177900 0.0059290 0.00000000 0.00197600 0.00197600
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00395300 0.00000000
 0.00000000 0.00000000 0.00000000) *
 673) RF.SM.>3.55 1907 3281.00 HSH (0.10850 0.0020980
 0.0015730 0.023600 0.0005244 0.0041950 0.00000000 0.0026220 0.00000000
 0.00891500 0.7850000 0.0440500 0.0104900 0.00000000 0.00576800 0.00052440
 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00052440 0.00000000
 0.00157300 0.00000000 0.00000000) *
 337) IDADE>14.5 37 85.65 HSH (0.05405 0.0000000
 0.0270300 0.081080 0.0000000 0.0540500 0.00000000 0.00000000 0.00000000

0.02703000 0.7027000 0.0000000 0.0000000 0.02703000 0.02703000 0.00000000
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
169) AUT>1.5 396 656.70 HSH (0.06818 0.0000000 0.0000000
0.015150 0.0000000 0.0075760 0.00000000 0.0025250 0.0000000 0.04040000
0.8056000 0.0404000 0.0126300 0.00000000 0.00505100 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00252500
0.00000000 0.00000000) *
85) TTPESF>5.5 1218 1573.00 HSH (0.10670 0.0000000
0.00000000 0.014780 0.00000000 0.0049260 0.00000000 0.00000000 0.00000000
0.00821000 0.8383000 0.0156000 0.0049260 0.00000000 0.00000000 0.00082100
0.00000000 0.00000000 0.00082100 0.0008210 0.00000000 0.00000000 0.00000000
0.00410500 0.00000000 0.00000000) *
43) RF.SM.>20.1 1291 2638.00 HSH (0.10690 0.0015490
0.0046480 0.013170 0.00000000 0.0108400 0.00000000 0.0038730 0.00000000
0.02246000 0.7467000 0.0550000 0.0170400 0.00154900 0.01239000 0.00309800
0.00000000 0.00000000 0.00000000 0.0007746 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
11) GI>1.5 7843 15930.00 HSH (0.09282 0.0010200 0.0042080
0.021680 0.0024230 0.0202700 0.00025500 0.0031880 0.00000000 0.01109000
0.7583000 0.0340400 0.0399100 0.00063750 0.00548300 0.00012750 0.00000000
0.00000000 0.00012750 0.0017850 0.00000000 0.00153000 0.00000000 0.00114800
0.00000000 0.00000000)
22) AUT<1.5 6330 12300.00 HSH (0.09558 0.0009479 0.0047390
0.022430 0.0017380 0.0219600 0.00031600 0.0025280 0.00000000 0.00900500
0.7692000 0.0241700 0.0377600 0.00063190 0.00363300 0.00015800 0.00000000
0.00000000 0.00000000 0.0020540 0.00000000 0.00173800 0.00000000 0.00142200
0.00000000 0.00000000)
44) SMED.SM.<1.25 6258 11900.00 HSH (0.09508 0.0009588
0.0047940 0.022530 0.0017580 0.0220500 0.00031960 0.0025570 0.00000000
0.00878900 0.7728000 0.0242900 0.0375500 0.00047940 0.00335600 0.00015980
0.00000000 0.00000000 0.00000000 0.0012780 0.00000000 0.00031960 0.00000000
0.00095880 0.00000000 0.00000000)
88) IDADE<14.5 3880 6559.00 HSH (0.09253 0.0012890
0.0038660 0.019850 0.0010310 0.0139200 0.00000000 0.0020620 0.00000000
0.00670100 0.7987000 0.0232000 0.0337600 0.00077320 0.00077320 0.00025770
0.00000000 0.00000000 0.00000000 0.0002577 0.00000000 0.00025770 0.00000000
0.00077320 0.00000000 0.00000000)
176) RF.SM.<30.05 3728 6152.00 HSH (0.09281 0.0010730
0.0037550 0.019850 0.0008047 0.0144800 0.00000000 0.0018780 0.00000000
0.00643800 0.8034000 0.0220000 0.0308500 0.00080470 0.00053650 0.00026820
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00026820 0.00000000
0.00080470 0.00000000 0.00000000)
352) TTPESF<5.5 2310 4053.00 HSH (0.08961 0.0017320
0.0038960 0.020780 0.0012990 0.0164500 0.00000000 0.0025970 0.00000000
0.00865800 0.7918000 0.0264100 0.0342000 0.00086580 0.00086580 0.00043290
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
0.00043290 0.00000000 0.00000000)
704) AUT<0.5 1160 2073.00 HSH (0.09828 0.0025860
0.0068970 0.022410 0.0017240 0.0206900 0.00000000 0.0043100 0.00000000
0.00431000 0.7836000 0.0241400 0.0301700 0.00000000 0.00086210 0.00000000
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
705) AUT>0.5 1150 1952.00 HSH (0.08087 0.0008696
0.0008696 0.019130 0.0008696 0.0121700 0.00000000 0.0008696 0.00000000
0.01304000 0.8000000 0.0287000 0.0382600 0.00173900 0.00086960 0.00086960
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00000000
0.00086960 0.00000000 0.00000000) *
353) TTPESF>5.5 1418 2068.00 HSH (0.09803 0.0000000
0.0035260 0.018340 0.00000000 0.0112800 0.00000000 0.0007052 0.00000000
0.00282100 0.8223000 0.0148100 0.0253900 0.00070520 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000 0.00000000 0.00000000 0.00070520 0.00000000
0.00141000 0.00000000 0.00000000) *
177) RF.SM.>30.05 152 363.20 HSH (0.08553 0.0065790
0.0065790 0.019740 0.0065790 0.00000000 0.00000000 0.0065790 0.00000000

0.01316000 0.6842000 0.0526300 0.1053000 0.00000000 0.00657900 0.00000000
0.00000000 0.00000000 0.00000000 0.0065790 0.0000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
89) IDADE>14.5 2378 5246.00 HSH (0.09924 0.0004205
0.0063080 0.026910 0.0029440 0.0353200 0.00084100 0.0033640 0.00000000
0.01220000 0.7304000 0.0260700 0.0437300 0.00000000 0.00756900 0.00000000
0.00000000 0.00000000 0.00000000 0.0029440 0.0000000 0.00042050 0.00000000
0.00126200 0.00000000 0.00000000)
178) IDADE<18.5 1829 3805.00 HSH (0.09677 0.0005467
0.0027340 0.022960 0.0016400 0.0267900 0.00054670 0.0021870 0.00000000
0.01203000 0.7458000 0.0273400 0.0486600 0.00000000 0.00656100 0.00000000
0.00000000 0.00000000 0.00000000 0.0038270 0.0000000 0.00000000 0.00000000
0.00164000 0.00000000 0.00000000)
356) AUT<0.5 945 1834.00 HSH (0.11430 0.0010580
0.0021160 0.024340 0.0010580 0.0370400 0.00000000 0.00000000 0.00000000
0.00634900 0.7534000 0.0222200 0.0285700 0.00000000 0.00423300 0.00000000
0.00000000 0.00000000 0.00000000 0.0052910 0.0000000 0.00000000 0.00000000
0.00000000 0.00000000 0.00000000) *
357) AUT>0.5 884 1916.00 HSH (0.07805 0.0000000
0.0033940 0.021490 0.0022620 0.0158400 0.00113100 0.0045250 0.00000000
0.01810000 0.7376000 0.0328100 0.0701400 0.00000000 0.00905000 0.00000000
0.00000000 0.00000000 0.00000000 0.0022620 0.0000000 0.00000000 0.00000000
0.00339400 0.00000000 0.00000000) *
179) IDADE>18.5 549 1385.00 HSH (0.10750 0.0000000
0.0182100 0.040070 0.0072860 0.0637500 0.00182100 0.0072860 0.00000000
0.01275000 0.6794000 0.0218600 0.0273200 0.00000000 0.01093000 0.00000000
0.00000000 0.00000000 0.00000000 0.0000000 0.0000000 0.00182100 0.00000000
0.00000000 0.00000000 0.00000000) *
45) SMED.SM.>1.25 72 260.10 HSH (0.13890 0.0000000
0.00000000 0.013890 0.00000000 0.0138900 0.00000000 0.00000000 0.00000000
0.02778000 0.4583000 0.0138900 0.0555600 0.01389000 0.02778000 0.00000000
0.00000000 0.00000000 0.00000000 0.0694400 0.0000000 0.12500000 0.00000000
0.04167000 0.00000000 0.00000000) *
23) AUT>1.5 1513 3488.00 HSH (0.08130 0.0013220 0.0019830
0.018510 0.0052880 0.0132200 0.00000000 0.0059480 0.0000000 0.01983000
0.7125000 0.0753500 0.0489100 0.00066090 0.01322000 0.00000000 0.00000000
0.00000000 0.00066090 0.0006609 0.0000000 0.00066090 0.0000000 0.00000000
0.00000000 0.00000000)
46) IDADE<20.5 1328 2888.00 HSH (0.07455 0.0007530 0.0015060
0.015810 0.0045180 0.0113000 0.00000000 0.0037650 0.0000000 0.01581000
0.7259000 0.0805700 0.0512000 0.00000000 0.01355000 0.00000000 0.00000000
0.00000000 0.00000000 0.0007530 0.0000000 0.00000000 0.0000000 0.00000000
0.00000000 0.00000000) *
47) IDADE>20.5 185 550.90 HSH (0.12970 0.0054050 0.0054050
0.037840 0.0108100 0.0270300 0.00000000 0.0216200 0.0000000 0.04865000
0.6162000 0.0378400 0.0324300 0.00540500 0.01081000 0.00000000 0.00000000
0.00000000 0.00540500 0.00000000 0.0000000 0.00540500 0.0000000 0.00000000
0.00000000 0.00000000) *
3) TRABALHA>0.5 34752 116600.00 HWH (0.10630 0.0024460 0.0082870
0.047130 0.0059280 0.0013520 0.00374100 0.0002878 0.0048920 0.00031650
0.0101900 0.0008057 0.0029350 0.00710800 0.00017270 0.00684900 0.00368300
0.0263300 0.00656100 0.5926000 0.0245700 0.04716000 0.0496400 0.02667000
0.00964000 0.0043740)
6) ESTUDA<0.5 30870 88030.00 HWH (0.11620 0.0027210 0.0089410
0.051640 0.0064460 0.0005507 0.00411400 0.0001620 0.0054100 0.00016200
0.0035630 0.0006803 0.0026890 0.00191100 0.00003239 0.00492400 0.00388700
0.0286700 0.00712700 0.6500000 0.0262100 0.00272100 0.0537700 0.00243000
0.01043000 0.0046970)
12) SF<5.5 30092 84170.00 HWH (0.09753 0.0027910 0.0089390
0.050810 0.0064800 0.0005317 0.00422000 0.0001662 0.0055160 0.00016620
0.0031570 0.0005982 0.0022270 0.00196100 0.00003323 0.00501800 0.00398800
0.0294100 0.00731100 0.6663000 0.0268500 0.00279100 0.0551600 0.00249200
0.01070000 0.0048190)
24) SMED.SM.<5.55 19625 50810.00 HWH (0.11200 0.0018340
0.0077450 0.054420 0.0058090 0.0003567 0.00341400 0.0001529 0.0040760

0.00015290 0.0032100 0.0004586 0.0028030 0.00107000 0.00005096 0.00265000
 0.00295500 0.0237500 0.00494300 0.6818000 0.0195200 0.00224200 0.0527900
 0.00193600 0.00713400 0.0027010)
 48) SMED.SM.<1.4 5235 14740.00 HWH (0.16830 0.0021010
 0.0093600 0.069910 0.0057310 0.0005731 0.00439400 0.0000000 0.0028650
 0.00038200 0.0042020 0.0005731 0.0057310 0.00191000 0.00000000 0.00458500
 0.00191000 0.0166200 0.00554000 0.6128000 0.0147100 0.00114600 0.0578800
 0.00191000 0.00439400 0.0024830)
 96) SF<2.5 3157 9727.00 HWH (0.18340 0.0028510 0.0101400
 0.089960 0.0079190 0.0009503 0.00570200 0.0000000 0.0025340 0.00063350
 0.0053850 0.0009503 0.0088690 0.00285100 0.00000000 0.00696900 0.00221700
 0.0177400 0.00633500 0.5654000 0.0126700 0.00158400 0.0525800 0.00316800
 0.00570200 0.0034840)
 192) SEXO<1.5 1705 4930.00 HWH (0.14130 0.0046920
 0.0052790 0.075660 0.0082110 0.0011730 0.00469200 0.0000000 0.0023460
 0.00000000 0.0023460 0.0000000 0.0029330 0.00000000 0.00000000 0.00879800
 0.00293300 0.0181800 0.00879800 0.6141000 0.0140800 0.00117300 0.0680400
 0.00293300 0.00703800 0.0052790)
 384) GI<2.5 1098 3164.00 HWH (0.17400 0.0045540
 0.0054640 0.096540 0.0091070 0.0018210 0.00546400 0.0000000 0.0009107
 0.00000000 0.0036430 0.0000000 0.0018210 0.00000000 0.00000000 0.00273200
 0.00273200 0.0127500 0.00637500 0.5801000 0.0145700 0.00182100 0.0664800
 0.00000000 0.00728600 0.0018210) *
 385) GI>2.5 607 1668.00 HWH (0.08237 0.0049420
 0.0049420 0.037890 0.0065900 0.0000000 0.00329500 0.0000000 0.0049420
 0.00000000 0.0000000 0.0000000 0.0049420 0.00000000 0.00000000 0.01977000
 0.00329500 0.0280100 0.01318000 0.6755000 0.0131800 0.00000000 0.0708400
 0.00823700 0.00659000 0.0115300) *
 193) SEXO>1.5 1452 4648.00 HWH (0.23280 0.0006887
 0.0158400 0.106700 0.0075760 0.0006887 0.00688700 0.0000000 0.0027550
 0.00137700 0.0089530 0.0020660 0.0158400 0.00619800 0.00000000 0.00482100
 0.00137700 0.0172200 0.00344400 0.5083000 0.0110200 0.00206600 0.0344400
 0.00344400 0.00413200 0.0013770)
 386) IDADE<47.5 1080 3415.00 HWH (0.21850 0.0009259
 0.0157400 0.087040 0.0083330 0.0009259 0.00740700 0.0000000 0.0037040
 0.00092590 0.0101900 0.0027780 0.0213000 0.00740700 0.00000000 0.00648100
 0.00000000 0.0157400 0.00463000 0.5361000 0.0120400 0.00185200 0.0305600
 0.00277800 0.00277800 0.0018520)
 772) RF.SM.<3.05 622 1904.00 HWH (0.18650 0.0016080
 0.0160800 0.075560 0.0048230 0.0016080 0.00482300 0.0000000 0.0048230
 0.00000000 0.0128600 0.0048230 0.0176800 0.00803900 0.00000000 0.00643100
 0.00000000 0.0241200 0.00643100 0.5804000 0.0096460 0.00321500 0.0225100
 0.00321500 0.00160800 0.0032150) *
 773) RF.SM.>3.05 458 1465.00 HWH (0.26200 0.0000000
 0.0152800 0.102600 0.0131000 0.0000000 0.01092000 0.0000000 0.0021830
 0.00218300 0.0065500 0.0000000 0.0262000 0.00655000 0.00000000 0.00655000
 0.00000000 0.0043670 0.00218300 0.4760000 0.0152800 0.00000000 0.0414800
 0.00218300 0.00436700 0.0000000) *
 387) IDADE>47.5 372 1169.00 HWH (0.27420 0.0000000
 0.0161300 0.164000 0.0053760 0.0000000 0.00537600 0.0000000 0.0000000
 0.00268800 0.0053760 0.0000000 0.0000000 0.00268800 0.00000000 0.00000000
 0.00537600 0.0215100 0.00000000 0.4274000 0.0080650 0.00268800 0.0457000
 0.00537600 0.00806500 0.0000000) *
 97) SF>2.5 2078 4840.00 HWH (0.14530 0.0009625 0.0081810
 0.039460 0.0024060 0.0000000 0.00240600 0.0000000 0.0033690 0.00000000
 0.0024060 0.0000000 0.0009625 0.00048120 0.00000000 0.00096250 0.00144400
 0.0149200 0.00433100 0.6848000 0.0178100 0.00048120 0.0659300 0.00000000
 0.00240600 0.0009625)
 194) GI<1.5 531 1099.00 HWH (0.19960 0.0000000 0.0000000
 0.048960 0.0018830 0.0000000 0.00000000 0.0000000 0.0000000 0.00000000
 0.0037660 0.0000000 0.0000000 0.00188300 0.00000000 0.00000000 0.00000000
 0.0075330 0.00000000 0.6742000 0.0113000 0.00000000 0.0470800 0.00000000
 0.00188300 0.0018830) *
 195) GI>1.5 1547 3684.00 HWH (0.12670 0.0012930 0.0109900
 0.036200 0.0025860 0.0000000 0.00323200 0.0000000 0.0045250 0.00000000

0.0019390 0.0000000 0.0012930 0.0000000 0.0000000 0.0012930 0.0019390
0.0174500 0.00581800 0.6884000 0.0200400 0.00064640 0.0724000 0.0000000
0.00258600 0.0006464)
390) TTPESF<9.5 1467 3463.00 HWH (0.11320 0.0013630
0.0115900 0.036810 0.0020450 0.0000000 0.00340800 0.0000000 0.0047720
0.0000000 0.0013630 0.0000000 0.0013630 0.0000000 0.0000000 0.0013630
0.00204500 0.0184000 0.00613500 0.6980000 0.0211300 0.00068170 0.0729400
0.00000000 0.00272700 0.0006817)
780) IDADE<26.5 1022 2310.00 HWH (0.10470 0.0019570
0.0117400 0.032290 0.0029350 0.0000000 0.00293500 0.0000000 0.0009785
0.00000000 0.0019570 0.0000000 0.0009785 0.00000000 0.00000000 0.00000000
0.00293500 0.0195700 0.00391400 0.7094000 0.0264200 0.00000000 0.0763200
0.00000000 0.00097850 0.0000000) *
781) IDADE>26.5 445 1110.00 HWH (0.13260 0.0000000
0.0112400 0.047190 0.0000000 0.0000000 0.00449400 0.0000000 0.0134800
0.00000000 0.0000000 0.0000000 0.0022470 0.00000000 0.00000000 0.00449400
0.00000000 0.0157300 0.01124000 0.6719000 0.0089890 0.00224700 0.0651700
0.00000000 0.00674200 0.0022470) *
391) TTPESF>9.5 80 173.70 HWH (0.37500 0.0000000
0.00000000 0.025000 0.0125000 0.0000000 0.00000000 0.0000000 0.0000000
0.00000000 0.0125000 0.0000000 0.0000000 0.00000000 0.00000000 0.00000000
0.00000000 0.0000000 0.00000000 0.5125000 0.0000000 0.00000000 0.0625000
0.00000000 0.00000000 0.0000000) *
49) SMED.SM.>1.4 14390 35710.00 HWH (0.09152 0.0017370
0.0071580 0.048780 0.0058370 0.0002780 0.00305800 0.0002085 0.0045170
0.00006949 0.0028490 0.0004170 0.0017370 0.00076440 0.00006949 0.00194600
0.00333600 0.0263400 0.00472600 0.7069000 0.0212600 0.00264100 0.0509400
0.00194600 0.00813100 0.0027800)
98) SF<2.5 9009 23820.00 HWH (0.10100 0.0019980 0.0098790
0.061380 0.0067710 0.0004440 0.00399600 0.0003330 0.0049950 0.00000000
0.0037740 0.0006660 0.0022200 0.00077700 0.00011100 0.00288600 0.00310800
0.0246400 0.00444000 0.6810000 0.0192000 0.00266400 0.0490600 0.00222000
0.00943500 0.0029970)
196) SEXO<1.5 5742 14140.00 HWH (0.08063 0.0015670
0.0074890 0.059040 0.0062700 0.0003483 0.00383100 0.0000000 0.0052250
0.00000000 0.0024380 0.0003483 0.0010450 0.00034830 0.00000000 0.00191600
0.00243800 0.0212500 0.00522500 0.7095000 0.0182900 0.00209000 0.0548600
0.00104500 0.01045000 0.0043540)
392) IDADE<50.5 4669 11040.00 HWH (0.06854 0.0017130
0.0070680 0.054620 0.0059970 0.0004284 0.00342700 0.0000000 0.0053540
0.00000000 0.0023560 0.0002142 0.0012850 0.00042840 0.00000000 0.00235600
0.00235600 0.0231300 0.00535400 0.7301000 0.0188500 0.00149900 0.0490500
0.00107100 0.01049000 0.0042840)
784) AUT<0.5 3698 8255.00 HWH (0.06787 0.0016220
0.0070310 0.050570 0.0045970 0.0002704 0.00324500 0.0000000 0.0048670
0.00000000 0.0027040 0.0002704 0.0002704 0.00054080 0.00000000 0.00189300
0.00216300 0.0224400 0.00324500 0.7461000 0.0165000 0.00108200 0.0486700
0.00081120 0.00865300 0.0045970)
1568) GI<2.5 3163 6956.00 HWH (0.07240 0.0015810
0.0060070 0.053750 0.0047420 0.0000000 0.00284500 0.0000000 0.0041100
0.00000000 0.0025290 0.0003162 0.0003162 0.00031620 0.00000000 0.00158100
0.00158100 0.0196000 0.00347800 0.7458000 0.0154900 0.00126500 0.0490000
0.00031620 0.00822000 0.0047420)
3136) IDADE<28.5 762 1649.00 HWH (0.05643
0.00000000 0.0052490 0.060370 0.0026250 0.0000000 0.00262500 0.0000000
0.0065620 0.00000000 0.0000000 0.0000000 0.0000000 0.00131200 0.00000000
0.00262500 0.00131200 0.0223100 0.00918600 0.7467000 0.0157500 0.00131200
0.0564300 0.00000000 0.00918600 0.0000000) *
3137) IDADE>28.5 2401 5270.00 HWH (0.07747
0.0020820 0.0062470 0.051650 0.0054140 0.0000000 0.00291500 0.0000000
0.0033320 0.00000000 0.0033320 0.0004165 0.0004165 0.00000000 0.00000000
0.00124900 0.00166600 0.0187400 0.00166600 0.7455000 0.0154100 0.00124900
0.0466500 0.00041650 0.00791300 0.0062470)
6274) IDADE<35.5 920 1953.00 HWH (0.05217
0.0032610 0.0097830 0.041300 0.0054350 0.0000000 0.00434800 0.0000000

0.0032610 0.0000000 0.0043480 0.0010870 0.0010870 0.0000000 0.0000000
 0.0010870 0.0021740 0.0163000 0.0010870 0.7717000 0.0163000 0.0021740
 0.0434800 0.0000000 0.0108700 0.0086960) *
 6275) IDADE>35.5 1481 3285.00 HWH (0.09318
 0.0013500 0.0040510 0.058070 0.0054020 0.0000000 0.0020260 0.0000000
 0.0033760 0.0000000 0.0027010 0.0000000 0.0000000 0.0000000 0.0000000
 0.0013500 0.0013500 0.0202600 0.0020260 0.7292000 0.0148500 0.00067520
 0.0486200 0.00067520 0.00607700 0.0047270)
 12550) IDADE<38.5 380 837.80 HWH (0.10530
 0.0026320 0.0000000 0.063160 0.0078950 0.0000000 0.0000000 0.0000000
 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
 0.0052630 0.0000000 0.0157900 0.0026320 0.7237000 0.0210500 0.0026320
 0.0315800 0.0026320 0.0131600 0.0026320) *
 12551) IDADE>38.5 1101 2413.00 HWH (0.08901
 0.0009083 0.0054500 0.056310 0.0045410 0.0000000 0.0027250 0.0000000
 0.0045410 0.0000000 0.0036330 0.0000000 0.0000000 0.0000000 0.0000000
 0.0000000 0.00181700 0.0218000 0.00181700 0.7312000 0.0127200 0.0000000
 0.0545000 0.0000000 0.0036330 0.0054500) *
 1569) GI>2.5 535 1256.00 HWH (0.04112 0.0018690
 0.0130800 0.031780 0.0037380 0.0018690 0.0056070 0.0000000 0.0093460
 0.0000000 0.0037380 0.0000000 0.0000000 0.0018690 0.0000000 0.0037380
 0.0056070 0.0392500 0.0018690 0.7477000 0.0224300 0.0000000 0.0467300
 0.0037380 0.0112100 0.0037380) *
 785) AUT>0.5 971 2725.00 HWH (0.07106 0.0020600
 0.0072090 0.070030 0.0113300 0.0010300 0.0041190 0.0000000 0.0072090
 0.0000000 0.0010300 0.0000000 0.0051490 0.0000000 0.0000000 0.0041190
 0.0030900 0.0257500 0.0133900 0.6694000 0.0278100 0.0030900 0.0504600
 0.0020600 0.0175100 0.0030900) *
 393) IDADE>50.5 1073 2999.00 HWH (0.13330 0.0009320
 0.0093200 0.078290 0.0074560 0.0000000 0.0055920 0.0000000 0.0046600
 0.0000000 0.0027960 0.0009320 0.0000000 0.0000000 0.0000000 0.0000000
 0.0027960 0.0130500 0.0046600 0.6198000 0.0158400 0.0046600 0.0801500
 0.0009320 0.0102500 0.0046600) *
 197) SEXO>1.5 3267 9502.00 HWH (0.13680 0.0027550
 0.0140800 0.065500 0.0076520 0.0006122 0.0042850 0.0009183 0.0045910
 0.0000000 0.0061220 0.0012240 0.0042850 0.0015300 0.00030610 0.0045910
 0.0042850 0.0306100 0.0030610 0.6309000 0.0208100 0.00367300 0.0388700
 0.0042850 0.0076520 0.0006122)
 394) GI<3.5 2713 7387.00 HWH (0.14340 0.0022120
 0.0121600 0.064140 0.0073720 0.0007372 0.0025800 0.0007372 0.0036860
 0.0000000 0.0051600 0.0007372 0.0044230 0.0007372 0.00036860 0.00405500
 0.00405500 0.0283800 0.0022120 0.6480000 0.0180600 0.0022120 0.0361200
 0.0025800 0.0058980 0.0000000)
 788) TPESF<2.5 566 1458.00 HWH (0.09364 0.0017670
 0.0159000 0.084810 0.0070670 0.0000000 0.0000000 0.0000000 0.0035340
 0.0000000 0.0017670 0.0000000 0.0017670 0.0000000 0.0000000 0.0000000
 0.0053000 0.0424000 0.0000000 0.6625000 0.0371000 0.0000000 0.0388700
 0.0000000 0.0035340 0.0000000) *
 789) TPESF>2.5 2147 5867.00 HWH (0.15650 0.0023290
 0.0111800 0.058690 0.0074520 0.0009315 0.0032600 0.0009315 0.0037260
 0.0000000 0.0060550 0.0009315 0.0051230 0.0009315 0.00046580 0.0051230
 0.0037260 0.0246900 0.0027950 0.6442000 0.0130400 0.0027950 0.0354000
 0.0032600 0.0065210 0.0000000)
 1578) RF.SM.<7.55 1061 2655.00 HWH (0.14040
 0.0018850 0.0056550 0.054670 0.0009425 0.0000000 0.0009425 0.0018850
 0.0037700 0.0000000 0.0065980 0.0009425 0.0065980 0.0009425 0.0009425
 0.00471300 0.0037700 0.0226200 0.0018850 0.6833000 0.0132000 0.00282800
 0.0292200 0.0065980 0.0056550 0.0000000) *
 1579) RF.SM.>7.55 1086 3158.00 HWH (0.17220
 0.0027620 0.0165700 0.062620 0.0138100 0.0018420 0.0055250 0.0000000
 0.0036830 0.0000000 0.0055250 0.0009208 0.0036830 0.0009208 0.0000000
 0.0055250 0.0036830 0.0267000 0.0036830 0.6059000 0.0128900 0.0027620
 0.0414400 0.0000000 0.0073660 0.0000000) *
 395) GI>3.5 554 2033.00 HWH (0.10470 0.0054150
 0.0234700 0.072200 0.0090250 0.0000000 0.0126400 0.0018050 0.0090250

0.00000000 0.0108300 0.0036100 0.0036100 0.00541500 0.00000000 0.00722000
 0.00541500 0.0415200 0.00722000 0.5469000 0.0343000 0.01083000 0.0523500
 0.01264000 0.01625000 0.0036100) *
 99) SF>2.5 5381 11660.00 HWH (0.07564 0.0013010 0.0026020
 0.027690 0.0042740 0.0000000 0.00148700 0.0000000 0.0037170 0.00018580
 0.0013010 0.0000000 0.0009292 0.00074340 0.00000000 0.00037170 0.00371700
 0.0291800 0.00520300 0.7504000 0.0247200 0.00260200 0.0540800 0.00148700
 0.00594700 0.0024160)
 198) IDADE<35.5 4836 10360.00 HWH (0.06886 0.0012410
 0.0020680 0.024810 0.0047560 0.0000000 0.00144700 0.0000000 0.0037220
 0.00000000 0.0012410 0.0000000 0.0008271 0.00082710 0.00000000 0.00041360
 0.00413600 0.0308100 0.00579000 0.7564000 0.0254300 0.00268800 0.0543800
 0.00144700 0.00599700 0.0026880)
 396) AUT<0.5 3225 6261.00 HWH (0.06884 0.0003101
 0.0009302 0.021710 0.0031010 0.0000000 0.00093020 0.0000000 0.0027910
 0.00000000 0.0006202 0.0000000 0.0006202 0.00124000 0.00000000 0.00031010
 0.00403100 0.0300800 0.00527100 0.7802000 0.0179800 0.00186000 0.0505400
 0.00124000 0.00434100 0.0031010)
 792) SEXO<1.5 1803 3318.00 HWH (0.06822 0.0000000
 0.0011090 0.021080 0.0049920 0.0000000 0.00000000 0.0000000 0.0027730
 0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.00000000 0.00000000
 0.00332800 0.0249600 0.00554600 0.7909000 0.0160800 0.00166400 0.0488100
 0.00000000 0.00499200 0.0055460)
 1584) SMED.SM.<2.35 636 1053.00 HWH (0.05818
 0.00000000 0.0015720 0.014150 0.0078620 0.0000000 0.00000000 0.0000000
 0.0015720 0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.00000000
 0.00000000 0.00000000 0.0204400 0.00628900 0.8129000 0.0078620 0.00471700
 0.0566000 0.00000000 0.00000000 0.0078620) *
 1585) SMED.SM.>2.35 1167 2231.00 HWH (0.07369
 0.00000000 0.0008569 0.024850 0.0034280 0.0000000 0.00000000 0.0000000
 0.0034280 0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.00000000
 0.00000000 0.00514100 0.0274200 0.00514100 0.7789000 0.0205700 0.00000000
 0.0445600 0.00000000 0.00771200 0.0042840) *
 793) SEXO>1.5 1422 2892.00 HWH (0.06962 0.0007032
 0.0007032 0.022500 0.0007032 0.0000000 0.00211000 0.0000000 0.0028130
 0.00000000 0.0014060 0.0000000 0.0014060 0.00281300 0.00000000 0.00070320
 0.00492300 0.0365700 0.00492300 0.7665000 0.0203900 0.00211000 0.0527400
 0.00281300 0.00351600 0.0000000) *
 397) AUT>0.5 1611 4025.00 HWH (0.06890 0.0031040
 0.0043450 0.031040 0.0080700 0.0000000 0.00248300 0.0000000 0.0055870
 0.00000000 0.0024830 0.0000000 0.0012410 0.00000000 0.00000000 0.00062070
 0.00434500 0.0322800 0.00682800 0.7089000 0.0403500 0.00434500 0.0620700
 0.00186200 0.00931100 0.0018620)
 794) SEXO<1.5 832 2199.00 HWH (0.06130 0.0036060
 0.0060100 0.042070 0.0096150 0.0000000 0.00120200 0.0000000 0.0072120
 0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.00000000 0.00000000
 0.00240400 0.0336500 0.00961500 0.6815000 0.0528800 0.00360600 0.0685100
 0.00240400 0.01082000 0.0036060) *
 795) SEXO>1.5 779 1785.00 HWH (0.07702 0.0025670
 0.0025670 0.019260 0.0064180 0.0000000 0.00385100 0.0000000 0.0038510
 0.00000000 0.0051350 0.0000000 0.0025670 0.00000000 0.00000000 0.00128400
 0.00641800 0.0308100 0.00385100 0.7381000 0.0269600 0.00513500 0.0552000
 0.00128400 0.00770200 0.0000000) *
 199) IDADE>35.5 545 1235.00 HWH (0.13580 0.0018350
 0.0073390 0.053210 0.0000000 0.0000000 0.00183500 0.0000000 0.0036700
 0.00183500 0.0018350 0.0000000 0.0018350 0.00000000 0.00000000 0.00000000
 0.00000000 0.0146800 0.00000000 0.6972000 0.0183500 0.00183500 0.0513800
 0.00183500 0.00550500 0.0000000) *
 25) SMED.SM.>5.55 10467 32690.00 HWH (0.07041 0.0045860
 0.0111800 0.044040 0.0077390 0.0008598 0.00573200 0.0001911 0.0082160
 0.00019110 0.0030570 0.0008598 0.0011460 0.00363000 0.00000000 0.00945800
 0.00592300 0.0400300 0.01175000 0.6372000 0.0406000 0.00382200 0.0596200
 0.00353500 0.01739000 0.0087900)
 50) AUT<0.5 3356 8797.00 HWH (0.07151 0.0047680 0.0104300
 0.045590 0.0056620 0.0000000 0.00476800 0.0002980 0.0065550 0.00000000

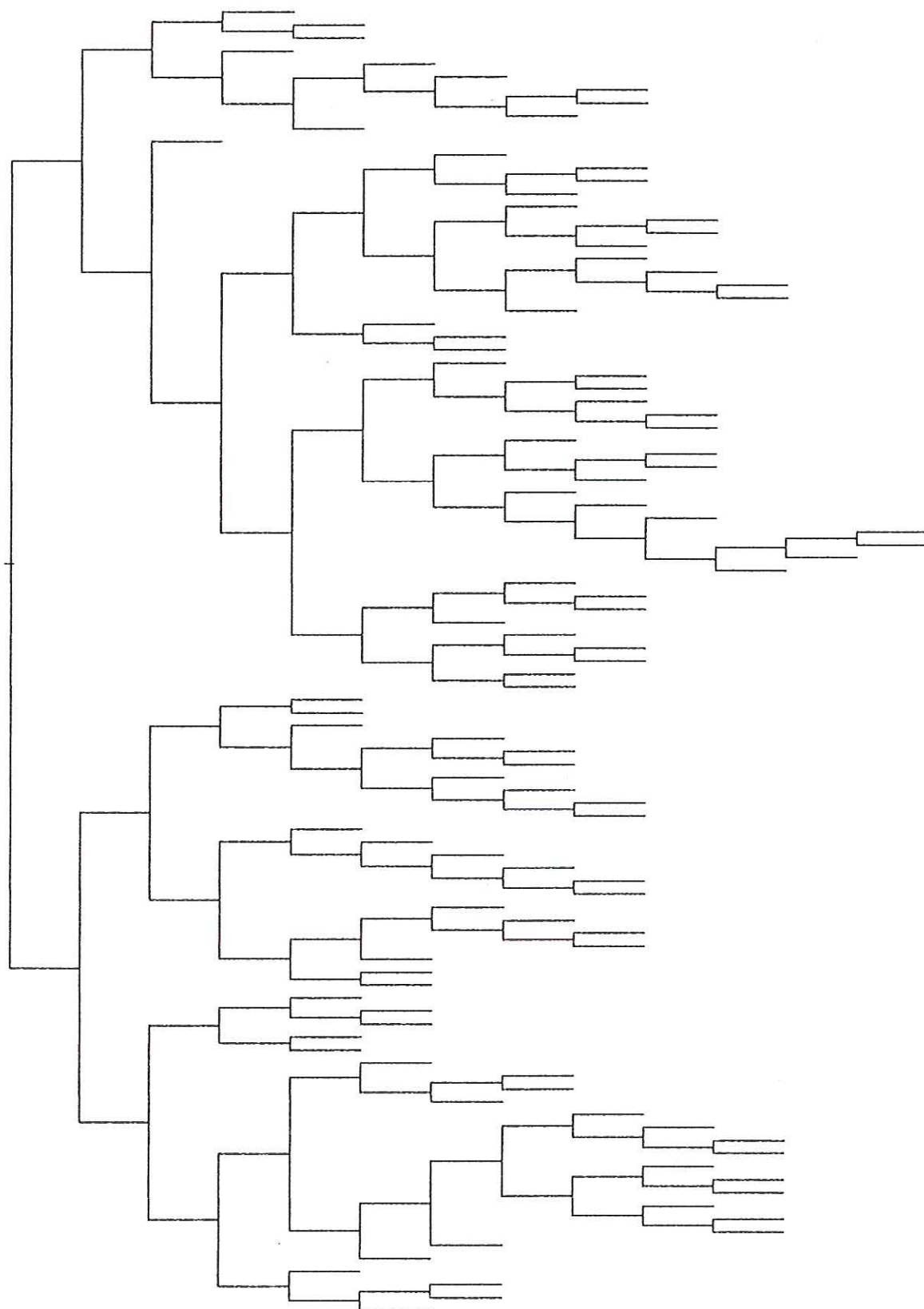
0.0014900 0.0008939 0.0011920 0.00059590 0.00000000 0.00119200 0.00476800
 0.0384400 0.00566200 0.7044000 0.0235400 0.00387400 0.0372500 0.00357600
 0.01698000 0.0065550)
 100) GI<3.5 2377 5827.00 HWH (0.07993 0.0058900 0.0096760
 0.048800 0.0037860 0.0000000 0.00336600 0.0004207 0.0046280 0.00000000
 0.0012620 0.0012620 0.0008414 0.00042070 0.00000000 0.00042070 0.00336600
 0.0286100 0.00546900 0.7215000 0.0201900 0.00378600 0.0353400 0.00210300
 0.01178000 0.0071520)
 200) IDADE<35.5 1207 2932.00 HWH (0.06628 0.0099420
 0.0099420 0.042250 0.0016570 0.00000000 0.00580000 0.00000000 0.0008285
 0.00000000 0.00000000 0.0016570 0.0008285 0.00000000 0.00000000 0.00082850
 0.00331400 0.0372800 0.00828500 0.7274000 0.0240300 0.00580000 0.0356300
 0.00331400 0.01160000 0.0033140) *
 201) IDADE>35.5 1170 2832.00 HWH (0.09402 0.0017090
 0.0094020 0.055560 0.0059830 0.00000000 0.00085470 0.0008547 0.0085470
 0.00000000 0.0025640 0.0008547 0.0008547 0.00085470 0.00000000 0.00000000
 0.00341900 0.0196600 0.00256400 0.7154000 0.0162400 0.00170900 0.0350400
 0.00085470 0.01197000 0.0111100) *
 101) GI>3.5 979 2890.00 HWH (0.05107 0.0020430 0.0122600
 0.037790 0.0102100 0.00000000 0.00817200 0.00000000 0.0112400 0.00000000
 0.0020430 0.00000000 0.0020430 0.00102100 0.00000000 0.00306400 0.00817200
 0.0623100 0.00612900 0.6629000 0.0316600 0.00408600 0.0418800 0.00715000
 0.02962000 0.0051070) *
 51) AUT>0.5 7111 23660.00 HWH (0.06989 0.0045000 0.0115300
 0.043310 0.0087190 0.0012660 0.00618800 0.0001406 0.0090000 0.00028130
 0.0037970 0.0008438 0.0011250 0.00506300 0.00000000 0.01336000 0.00646900
 0.0407800 0.01463000 0.6055000 0.0486600 0.00379700 0.0701700 0.00351600
 0.01758000 0.0098440)
 102) SEXO<1.5 5466 17680.00 HWH (0.05982 0.0049400 0.0107900
 0.039880 0.0080500 0.0010980 0.00622000 0.00000000 0.0095130 0.00018290
 0.0032930 0.0003659 0.0007318 0.00439100 0.00000000 0.01500000 0.00640300
 0.0402500 0.01665000 0.6220000 0.0492100 0.00365900 0.0676900 0.00237800
 0.01628000 0.0111600)
 204) SMED.SM.<13.95 3299 10000.00 HWH (0.06275 0.0048500
 0.0112200 0.045770 0.0084870 0.0012120 0.00515300 0.00000000 0.0066690
 0.00030310 0.0030310 0.00000000 0.0003031 0.00424400 0.00000000 0.00939700
 0.00363700 0.0366800 0.01303000 0.6469000 0.0469800 0.00454700 0.0624400
 0.00121200 0.01182000 0.0093970)
 408) IDADE<28.5 658 1856.00 HWH (0.04407 0.0030400
 0.0091190 0.053190 0.0030400 0.00000000 0.00607900 0.00000000 0.0060790
 0.00000000 0.0030400 0.00000000 0.00000000 0.00152000 0.00000000 0.00000000
 0.00455900 0.0531900 0.01520000 0.6641000 0.0638300 0.00607900 0.0486300
 0.00304000 0.00607900 0.0060790) *
 409) IDADE>28.5 2641 8095.00 HWH (0.06740 0.0053010
 0.0117400 0.043920 0.0098450 0.0015150 0.00492200 0.00000000 0.0068160
 0.00037860 0.0030290 0.00000000 0.0003786 0.00492200 0.00000000 0.01174000
 0.00340800 0.0325600 0.01250000 0.6426000 0.0427900 0.00416500 0.0658800
 0.00075730 0.01325000 0.0102200)
 818) RF.SM.<19.25 2198 6596.00 HWH (0.05960 0.0054600
 0.0104600 0.041860 0.0086440 0.0018200 0.00500500 0.00000000 0.0068240
 0.00045500 0.0022750 0.00000000 0.0004550 0.00409500 0.00000000 0.01365000
 0.00364000 0.0318500 0.01137000 0.6556000 0.0450400 0.00500500 0.0623300
 0.00090990 0.01274000 0.0109200)
 1636) SMED.SM.<8.85 1167 3428.00 HWH (0.06769
 0.0068550 0.0085690 0.043700 0.0102800 0.0008569 0.00771200 0.0000000
 0.0068550 0.00085690 0.0017140 0.00000000 0.0008569 0.00685500 0.00000000
 0.00856900 0.00171400 0.0282800 0.01457000 0.6650000 0.0377000 0.00171400
 0.0548400 0.00171400 0.01371000 0.0094260)
 3272) TTPESF<3.5 357 1012.00 HWH (0.05042
 0.0028010 0.0112000 0.064430 0.0084030 0.00000000 0.00840300 0.0000000
 0.0056020 0.00280100 0.00000000 0.00000000 0.0028010 0.00000000 0.00000000
 0.00000000 0.00560200 0.0420200 0.02241000 0.6667000 0.0308100 0.00000000
 0.0532200 0.00000000 0.01401000 0.0084030) *
 3273) TTPESF>3.5 810 2373.00 HWH (0.07531
 0.0086420 0.0074070 0.034570 0.0111100 0.0012350 0.00740700 0.00000000

0.0074070 0.0000000 0.0024690 0.0000000 0.0000000 0.00987700 0.00000000
0.01235000 0.00000000 0.0222200 0.01111000 0.6642000 0.0407400 0.00246900
0.0555600 0.00246900 0.01358000 0.0098770) *
1637) SMED.SM.>8.85 1031 3123.00 HWH (0.05044
0.0038800 0.0126100 0.039770 0.0067900 0.0029100 0.00194000 0.0000000
0.0067900 0.00000000 0.0029100 0.0000000 0.0000000 0.00096990 0.00000000
0.01940000 0.00582000 0.0358900 0.00775900 0.6450000 0.0533500 0.00872900
0.0708100 0.00000000 0.01164000 0.0126100) *
819) RF.SM.>19.25 443 1458.00 HWH (0.10610 0.0045150
0.0180600 0.054180 0.0158000 0.0000000 0.00451500 0.0000000 0.0067720
0.00000000 0.0067720 0.0000000 0.0000000 0.00902900 0.00000000 0.00225700
0.00225700 0.0361200 0.01806000 0.5779000 0.0316000 0.00000000 0.0835200
0.00000000 0.01580000 0.0067720) *
205) SMED.SM.>13.95 2167 7590.00 HWH (0.05538 0.0050760
0.0101500 0.030920 0.0073830 0.0009229 0.00784500 0.0000000 0.0138400
0.00000000 0.0036920 0.0009229 0.0013840 0.00461500 0.00000000 0.02353000
0.01061000 0.0456900 0.02215000 0.5842000 0.0526100 0.00230700 0.0756800
0.00415300 0.02307000 0.0138400) *
410) IDADE<48.5 1655 5789.00 HWH (0.04955 0.0060420
0.0090630 0.030210 0.0048340 0.0012080 0.00725100 0.0000000 0.0139000
0.00000000 0.0042300 0.0012080 0.0012080 0.00483400 0.00000000 0.02900000
0.01148000 0.0519600 0.02356000 0.5909000 0.0561900 0.00302100 0.0568000
0.00543800 0.02417000 0.0139000) *
820) RF.SM.<22.75 597 1933.00 HWH (0.03350 0.0050250
0.0117300 0.036850 0.0067000 0.0033500 0.00670000 0.0000000 0.0050250
0.00000000 0.0033500 0.0033500 0.0000000 0.00335000 0.00000000 0.02178000
0.01508000 0.0435500 0.02010000 0.6332000 0.0536000 0.00670000 0.0586300
0.00502500 0.01843000 0.0050250) *
821) RF.SM.>22.75 1058 3813.00 HWH (0.05860 0.0066160
0.0075610 0.026470 0.0037810 0.0000000 0.00756100 0.0000000 0.0189000
0.00000000 0.0047260 0.0000000 0.0018900 0.00567100 0.00000000 0.03308000
0.00945200 0.0567100 0.02552000 0.5671000 0.0576600 0.00094520 0.0557700
0.00567100 0.02741000 0.0189000) *
1642) TTPESF<3.5 293 943.30 HWH (0.03072 0.0102400
0.0136500 0.030720 0.0000000 0.0000000 0.00682600 0.0000000 0.0102400
0.00000000 0.0034130 0.0000000 0.0000000 0.00000000 0.00000000 0.00341300
0.01024000 0.0750900 0.01365000 0.6143000 0.0648500 0.00341300 0.0546100
0.01024000 0.02048000 0.0238900) *
1643) TTPESF>3.5 765 2823.00 HWH (0.06928 0.0052290
0.0052290 0.024840 0.0052290 0.0000000 0.00784300 0.0000000 0.0222200
0.00000000 0.0052290 0.0000000 0.0026140 0.00784300 0.00000000 0.04444000
0.00915000 0.0496700 0.03007000 0.5490000 0.0549000 0.00000000 0.0562100
0.00392200 0.03007000 0.0169900) *
411) IDADE>48.5 512 1727.00 HWH (0.07422 0.0019530
0.0136700 0.033200 0.0156200 0.0000000 0.00976600 0.0000000 0.0136700
0.00000000 0.0019530 0.0000000 0.0019530 0.00390600 0.00000000 0.00585900
0.00781200 0.0253900 0.01758000 0.5625000 0.0410200 0.00000000 0.1367000
0.00000000 0.01953000 0.0136700) *
103) SEXO>1.5 1645 5884.00 HWH (0.10330 0.0030400 0.0139800
0.054710 0.0109400 0.0018240 0.00607900 0.0006079 0.0072950 0.00060790
0.0054710 0.0024320 0.0024320 0.00729500 0.00000000 0.00790300 0.00668700
0.0425500 0.00790300 0.5508000 0.0468100 0.00425500 0.0784200 0.00729500
0.02188000 0.0054710) *
206) IDADE<39.5 1107 3834.00 HWH (0.07769 0.0018070
0.0144500 0.042460 0.0099370 0.0018070 0.00632300 0.0000000 0.0063230
0.00090330 0.0054200 0.0036130 0.0027100 0.00542000 0.00000000 0.00722700
0.00722700 0.0496800 0.00903300 0.5827000 0.0533000 0.00632300 0.0677500
0.00903300 0.02439000 0.0045170) *
412) RF.SM.<9.45 111 343.70 HWH (0.03604 0.0000000
0.0270300 0.063060 0.0000000 0.0000000 0.00900900 0.0000000 0.0000000
0.00000000 0.0090090 0.0000000 0.0270300 0.00000000 0.00000000 0.01802000
0.00900900 0.0270300 0.00000000 0.6306000 0.0180200 0.00900900 0.0450500
0.00900900 0.06306000 0.0000000) *
413) RF.SM.>9.45 996 3446.00 HWH (0.08233 0.0020080
0.0130500 0.040160 0.0110400 0.0020080 0.00602400 0.0000000 0.0070280

0.00100400 0.0050200 0.0040160 0.0000000 0.00602400 0.0000000 0.00602400
 0.00702800 0.0522100 0.01004000 0.5773000 0.0572300 0.00602400 0.0702800
 0.00903600 0.02008000 0.0050200)
 826) GI<4.5 449 1430.00 HWH (0.08018 0.0022270
 0.0178200 0.046770 0.0089090 0.0022270 0.00668200 0.0000000 0.0044540
 0.00000000 0.0044540 0.0089090 0.0000000 0.00668200 0.00000000 0.00000000
 0.00222700 0.0490000 0.01114000 0.6147000 0.0668200 0.00445400 0.0423200
 0.00668200 0.00890900 0.0044540) *
 827) GI>4.5 547 1975.00 HWH (0.08410 0.0018280
 0.0091410 0.034730 0.0128000 0.0018280 0.00548400 0.0000000 0.0091410
 0.00182800 0.0054840 0.0000000 0.0000000 0.00548400 0.00000000 0.01097000
 0.01097000 0.0548400 0.00914100 0.5466000 0.0493600 0.00731300 0.0932400
 0.01097000 0.02925000 0.0054840) *
 207) IDADE>39.5 538 1983.00 HWH (0.15610 0.0055760
 0.0130100 0.079930 0.0130100 0.0018590 0.00557600 0.0018590 0.0092940
 0.00000000 0.0055760 0.0000000 0.0018590 0.01115000 0.00000000 0.00929400
 0.00557600 0.0278800 0.00557600 0.4851000 0.0334600 0.00000000 0.1004000
 0.00371700 0.01673000 0.0074350) *
 13) SF>5.5 778 1099.00 H (0.83680 0.0000000 0.0089970 0.083550
 0.0051410 0.0012850 0.00000000 0.0000000 0.0012850 0.00000000 0.0192800
 0.0038560 0.0205700 0.00000000 0.00000000 0.00128500 0.00000000 0.0000000
 0.00000000 0.0167100 0.0012850 0.00000000 0.0000000 0.00000000 0.00000000
 0.00000000) *
 7) ESTUDA>0.5 3882 14640.00 HWHSH (0.02808 0.0002576 0.0030910
 0.011330 0.0018030 0.0077280 0.00077280 0.0012880 0.0007728 0.00154600
 0.0628500 0.0018030 0.0048940 0.04843000 0.00128800 0.02215000 0.00206100
 0.0077280 0.00206100 0.1365000 0.0115900 0.40060000 0.0167400 0.21950000
 0.00334900 0.0018030)
 14) GI<3.5 2529 8997.00 HWHSH (0.02768 0.0003954 0.0035590
 0.009095 0.0011860 0.0071170 0.00039540 0.0011860 0.0011860 0.00118600
 0.0735500 0.0007908 0.0047450 0.05496000 0.00158200 0.01542000 0.00079080
 0.0047450 0.00276800 0.1356000 0.0090950 0.47010000 0.0170000 0.15380000
 0.00039540 0.0015820)
 28) SF<5.5 2473 8631.00 HWHSH (0.02467 0.0004044 0.0036390
 0.008087 0.0012130 0.0068740 0.00040440 0.0012130 0.0012130 0.00040440
 0.0594400 0.0008087 0.0040440 0.05621000 0.00161700 0.01577000 0.00080870
 0.0048520 0.00283100 0.1387000 0.0093000 0.48080000 0.0173900 0.15730000
 0.00040440 0.0016170)
 56) SMED.SM.<0.95 538 2048.00 HWHSH (0.03532 0.0000000
 0.0018590 0.018590 0.0037170 0.0111500 0.00185900 0.0000000 0.0000000
 0.00000000 0.1431000 0.0000000 0.0167300 0.11710000 0.00743500 0.02602000
 0.00000000 0.0000000 0.00000000 0.0873600 0.0037170 0.41820000 0.0130100
 0.09294000 0.00000000 0.0018590) *
 57) SMED.SM.>0.95 1935 6363.00 HWHSH (0.02171 0.0005168
 0.0041340 0.005168 0.0005168 0.0056850 0.00000000 0.0015500 0.0015500
 0.00051680 0.0361800 0.0010340 0.0005168 0.0005168 0.03928000 0.00000000 0.01292000
 0.00103400 0.0062020 0.00361800 0.1530000 0.0108500 0.49820000 0.0186000
 0.17520000 0.00051680 0.0015500)
 114) IDADE<21.5 1521 4891.00 HWHSH (0.02367 0.0000000
 0.0019720 0.005260 0.0006575 0.0052600 0.00000000 0.0013150 0.0013150
 0.00065750 0.0348500 0.0000000 0.0006575 0.04339000 0.00000000 0.01249000
 0.00065750 0.0065750 0.00460200 0.1499000 0.0092040 0.52330000 0.0223500
 0.14990000 0.00065750 0.0013150)
 228) AUT<2.5 1490 4730.00 HWHSH (0.02416 0.0000000
 0.0020130 0.005369 0.0006711 0.0053690 0.00000000 0.0013420 0.0013420
 0.00000000 0.0355700 0.0000000 0.0000000 0.03893000 0.00000000 0.01074000
 0.00067110 0.0067110 0.00402700 0.1490000 0.0093960 0.52750000 0.0228200
 0.15230000 0.00067110 0.0013420)
 456) IDADE<15.5 289 936.80 HWHSH (0.04498 0.0000000
 0.0034600 0.010380 0.0000000 0.0034600 0.00000000 0.0034600 0.0000000
 0.00000000 0.0311400 0.0000000 0.0000000 0.08304000 0.00000000 0.00346000
 0.00000000 0.0034600 0.00346000 0.1557000 0.0138400 0.52600000 0.0311400
 0.08304000 0.00000000 0.0000000) *
 457) IDADE>15.5 1201 3749.00 HWHSH (0.01915 0.0000000
 0.0016650 0.004163 0.0008326 0.0058280 0.00000000 0.0008326 0.0016650

0.00000000 0.0366400 0.0000000 0.0000000 0.02831000 0.00000000 0.01249000
0.00083260 0.0074940 0.00416300 0.1474000 0.0083260 0.52790000 0.0208200
0.16900000 0.00083260 0.0016650)
914) RF.SM.<11.65 649 1831.00 HWHSH (0.01541
0.00000000 0.00000000 0.001541 0.0015410 0.0030820 0.00000000 0.0015410
0.00000000 0.00000000 0.0169500 0.0000000 0.0000000 0.02311000 0.00000000
0.01233000 0.00154100 0.0061630 0.00308200 0.1464000 0.0046220 0.55930000
0.0200300 0.18180000 0.00000000 0.0015410) *
915) RF.SM.>11.65 552 1877.00 HWHSH (0.02355
0.00000000 0.0036230 0.007246 0.0000000 0.0090580 0.00000000 0.0000000
0.0036230 0.00000000 0.0597800 0.0000000 0.0000000 0.03442000 0.00000000
0.01268000 0.00000000 0.0090580 0.00543500 0.1486000 0.0126800 0.49090000
0.0217400 0.15400000 0.00181200 0.0018120) *
229) AUT>2.5 31 105.50 HWHSH (0.00000 0.0000000
0.00000000 0.0000000 0.0000000 0.0000000 0.00000000 0.0000000 0.0000000
0.03226000 0.0000000 0.0000000 0.0322600 0.25810000 0.00000000 0.09677000
0.00000000 0.0000000 0.03226000 0.1935000 0.0000000 0.32260000 0.0000000
0.03226000 0.00000000 0.0000000) *
115) IDADE>21.5 414 1402.00 HWHSH (0.01449 0.0024150
0.0120800 0.004831 0.0000000 0.0072460 0.00000000 0.0024150 0.0024150
0.00000000 0.0410600 0.0048310 0.0000000 0.02415000 0.00000000 0.01449000
0.00241500 0.0048310 0.00000000 0.1643000 0.0169100 0.40580000 0.0048310
0.26810000 0.00000000 0.0024150) *
29) SF>5.5 56 113.40 HSH (0.16070 0.0000000 0.0000000
0.053570 0.0000000 0.0178600 0.00000000 0.0000000 0.0000000 0.03571000
0.6964000 0.0000000 0.0357100 0.00000000 0.0000000 0.0000000 0.00000000
0.00000000 0.00000000 0.0000000 0.0000000 0.00000000 0.0000000 0.00000000
0.00000000 0.0000000) *
15) GI>3.5 1353 5325.00 HWSH (0.02882 0.0000000 0.0022170
0.015520 0.0029560 0.0088690 0.00147800 0.0014780 0.00000000 0.00221700
0.0428700 0.0036950 0.0051740 0.03622000 0.00073910 0.03474000 0.00443500
0.0133000 0.00073910 0.1382000 0.0162600 0.27050000 0.0162600 0.34220000
0.00886900 0.0022170)
30) GI<4.5 1111 4167.00 HWSH (0.02700 0.0000000 0.0027000
0.013500 0.0027000 0.0090010 0.00000000 0.0018000 0.0000000 0.00000000
0.0414000 0.0045000 0.0045000 0.03960000 0.00090010 0.03780000 0.00540100
0.0117000 0.00000000 0.1188000 0.0144000 0.28440000 0.0108000 0.36450000
0.00360000 0.0009001)
60) SMED.SM.<3.95 459 1820.00 HWHSH (0.04575 0.0000000
0.00000000 0.017430 0.0043570 0.0130700 0.00000000 0.0000000 0.0000000
0.00000000 0.0522900 0.0087150 0.0108900 0.05229000 0.00217900 0.04575000
0.00000000 0.0021790 0.00000000 0.1046000 0.0217900 0.31150000 0.0108900
0.28980000 0.00653600 0.0000000) *
61) SMED.SM.>3.95 652 2272.00 HWSH (0.01380 0.0000000
0.0046010 0.010740 0.0015340 0.0061350 0.00000000 0.0030670 0.0000000
0.00000000 0.0337400 0.0015340 0.0000000 0.03067000 0.00000000 0.03221000
0.00920200 0.0184000 0.00000000 0.1288000 0.0092020 0.26530000 0.0107400
0.41720000 0.00153400 0.0015340) *
31) GI>4.5 242 1066.00 HWSH (0.03719 0.0000000 0.0000000
0.024790 0.0041320 0.0082640 0.00826400 0.0000000 0.0000000 0.01240000
0.0495900 0.0000000 0.0082640 0.02066000 0.00000000 0.02066000 0.00000000
0.0206600 0.00413200 0.2273000 0.0247900 0.20660000 0.0413200 0.23970000
0.03306000 0.0082640) *

Figura – Árvore A-2 com desvio de 0,01



* Os dados e os critérios para a geração da árvore foram os mesmos adotados em A-1, com exceção ao desvio.
** Os valores das variáveis e nós finais não estão apresentados na figura em função do seu tamanho.

REFERÊNCIAS BIBLIOGRÁFICAS

REFERÊNCIAS BIBLIOGRÁFICAS

ADLER T.; BEN-AKIVA M. E. (1979). A theoretical and Empirical Model of Trip Chaining Behavior. *Transportation Research B*. 13B. p.243-257.

ALGERS, S.; ELIASSON J.; MATTSSON L. (2001). Activity-based Model Development to Support Transport Planning in the Stockholm Region. Preliminar Version. Division of Transport and Location Analysis. Department of Infrastructure and Planning. Royal Institute of Technology. Stockholm. Sweden. 35p. (to be presented at the *5th Workshop of the TLE Network*. Nynäshamn).

ALVARES, L. (2000) *Aplicação de Data Mining em Bases de Dados Especializadas em Ciência da Informação para Obtenção de Informações sobre a Estrutura de Pesquisa e Desenvolvimento em Ciência da Informação no Brasil*. Brasília. Monografia (Especialização) UFRJ/ECO, MCT/INT/IBICT.

ARENTZE, T., HOFMAN, F., KALFS, N., and TIMMERMANS, H. (1997). Data needs, data collection and data quality requirements of activity-based transport models. *Proceedings, International Conference on Transport Survey Quality and Innovation (Transport Surveys: Raising the Standard)*, 24-30 May 1997, Grainau, Germany.

ARENTZE, T.; TIMMERMANS, H. (2000). *ALBATROSS: A Learning Based Transportation Oriented Simulation System*. European Institute of Retailing and Service Studies.

BAIN, J. H (1976). *Activity Choice Analysis, Time Allocation and Disaggregate Travel Demand Modelling*. Unpublished Masters Thesis. Department of Civil Engineering. MIT. Citado por DAMM (1983)

BELL, J. B. (1996). Application of Classification Tree to the Habitat Preference of Upland Birds. *Journal of Applied Statistics*. 23(2-3): 349-359.

BEN-AKIVA, M.; BOWMAN J. L. (1996). *Activity-based Disaggregate Travel Demand Model System with Activity Schedules*. Massachusetts Institute of Technology. Cambridge.

BEN-AKIVA, M.; BOWMAN J. L. ; GOPINATH D. (1995). Travel Demand Model System for the Information Era. publication pending in *Transportation*.

- BEN-AKIVA, M.; LERMAN, S. (1985). *Discrete choice analysis: Theory and Application to Travel Demand*. Cambridge, Massachusetts. MIT Press.
- BERNARD, A, et al. (1996). Household Structure and Mobility Patterns of Women in O-D Surveys: Methods and Results Based on the Case Studies of Montreal and Paris. A paper presented at the *Second National Conference on Women's Travel Issues*, in Baltimore, Maryland.
- BERSON, A.; SMITH, S. J. (1997). *Data Warehousing, Data Mining & OLAP*. McGraw-Hill. Estados Unidos.
- BHAT, C. R. (1997). Work travel mode choice and number of nowork commute stops. *Transportation Research 31B*. 41-54.
- BHAT, C. R.; GUO, J. Y. (2001). Representation and Analysis Plan and Data Needs Analysis for the Activity-Travel System. Research Project. 0-8040. *Activity based Travel Demand Modeling For Metropolitan Area in Texas*. Center of Transportation Research. Texas.
- BODKER, S. *Historical Analysis and Conflicting Perspectives – contextualizing HCI* (1993). East-west HCI Conference.
- BONNAFOUS, A.; PATIER-MARQUE, D. e PLASSARD, F. (1981). *Mobilité et Vie Quotidienne* (Suivi experimental du metro de Lyon). Presses universitaires de Lyon. France. CITADO POR DAMM (1983).
- BOWMAN, J. L. (1995). *Activity Based Travel Demand Model System with Daily Activity Schedules*. 92p. Dissertação (Mestrado) – Massachusetts Institute of Technology.
- BOWMAN, J. L. (1998). *The Day Activity Schedule Approach to Travel Demand Analysis*. 185 p. Tese (Doutorado) – Massachusetts Institute of Technology.
- BOWMAN, J. L. ; BEN-AKIVA M. (1997). Activity based travel forecasting. *Activity-based Travel Forecasting Conference*. June 2-5. 1995: Summary, Recommendations and Compendium of Papers. New Orleans. Louisiana. USDOT report # DOT-t-97-17. 32 pages.
- BOWMAN, J. L.; BRADLEY M.; SHIFTAN, Y; LAWTON, T.K.; BEN-AKIVA, M. (1998) Demonstration of an Activity-Based Model System for Portland. *8th World Conference on Transport Research*, July 12-17, 1998, Antwerp, Belgium.
- BREIMAN, L; FRIEDMAN, J. OLSHEN, R. e STONE, C. (1984). *Classification and Regression Trees*, Chapman and Hall. New York. NY.
- BROG, W.; ERL E. (1983). Application of a Model of Individual Behaviour (situational approach) to Explain Household Activity Patterns in a Urban Area and to Forecast Behavioural Changes. *Recent advances in travel demand analysis*. CARPENTER S. and JONES P. ed. Aldershot. England. Gower.
- BURNETT, R. P., HANHAM R. Q. and COOK. A. (1978) Choice and Constraints Oriented Modelling, Alternative Approaches to Travel Behaviour. In: D.E. Skinner (ed.). *Directions to Improve Urban Travel Demand Forecasting*, Conference Summary and White Papers. 343-411. U.S. Department of Transportation. Federal Highways Administration. Washington D.C.

- CHAPIN, F.S. (1974) *Human Activity Patterns in the City: Things People Do in Time and Space*, New York: J. Wiley and Sons.
- CIRILA, C.; AXHAUSEN, K. W. (2002). Comparing Urban Activity Travel Behaviour. *Transportation Research Board. 81 th Annual Meeting*. January 13-17, 2002. Washington, D.C.
- CLARK, L. A.; PREGIBON, D. (1992). *Tree-based Models*, In *Statistical Models in S* (eds. J. M. Chambers and T. J. Hastie), AT&T Bell Laboratories and Wadsworth & Books/Cole.
- CMSP - Companhia do Metropolitano de São Paulo (1988). Pesquisa Origem Destino São Paulo 1987, São Paulo.
- DAMM, D. (1983). Theory and Empirical Results: A Comparison of Recent Activity-based Research. *Recent advances in travel demand analysis*. CARPENTER S. and JONES P. ed. Aldershot. England. Gower.
- ETTEMA, D.F. AND TIMMERMANS, H.J.P. (1997) (eds) *Activity-Based Approaches to Travel Analysis*, Elsevier Science Ltd.
- FAYYAD, U. M., PIATETSKY-SHAPIRO, G., SMYTH, P., UTHURUSAMY, R. (1996). *Advances in Knowledge Discovery And Data Mining*. AAAI Press,
- FERREIRA, C. A. ; SOARES J .F.; CRUZ F. R. B. (2001). Reconhecimento de Padrões em Estatística: Uma Abordagem Comparativa, *Proceedings of the V Brazilian Conference on Neural Networks – V Congresso Brasileiro de Redes Neurais*. Pp. 409-414. RJ. Brazil.
- FOX, M. B (1983) Working Women and Travel: The Access of Women to Work and Community Facilities. *Journal of The American Association of Planners*, 49 (2): 156-170.
- GODARD, X. (1980). Analyse de la Mobilité comme Traduction des Pratiques d'activités. Vol. 4 on the report on the project *Recherche sur la mobilité des personnes en zone urbaine*. Arcueil. France: Institut de Recherche des Transports. CITADO POR DAMM (1983).
- GOLOB, J. M.; GOLOB T. F. (1983) Classification of Approaches to Travel-Behavior Analysis. TRB Special Report 201. *Transportation Research Board*.
- GORDON, P; Kumar A.; Richardson. H. W. (1989). Gender Differences in Metropolitan Travel Behavior. *Regional Studies*. 23(6):499-510.
- GOODWIN, P.; KITAMURA R.; MEURS H. (1990). Some Principles of Dynamic Analysis of Travel Behavior, in *Development in Dynamic and Activity-based Approaches to Travel Analysis*, Aldershot, England, Gower. Citado por BOWMAN (1995).
- HAGERSTRAND, T. (1970). What About People in Regional Science? *Paper of the Regional Science Association*, Vol.24.
- HAGERSTRAND, T. (1974). The Impact of Transport on the Quality of Life. Lund University. Department of Geography. Rapport och Notiser. N.º10. Lund.
- HAMED, M. M.; MANNERING F. L. (1993). Modeling Travelers' Postwork Activity Involvement: Toward A New Methodology. *Transportation Science*. 27(4): 381-394.

HANDY, S. (1996). Methodologies for Exploring the Link Between Urban Form and Travel Behavior. *Transportation Research D*. Vol. 1. N.º 2. Pp. 151-165.

HANSON, S. (1978). The Activity Patterns of Elderly Households. *Geographiska Annaler*, Series B. 109-124. Citado por DAMM (1983).

HANSON, S. (1979). Urban Travel Linkages: A Review, in D. Hensher and P. Stopher (eds.), *Behavioral Travel Modeling*. London. Croom-Helm. 81-100. Citado por CARPENTER E JONES (1983).

HANSON, S. (1981). The Importance of Multi-purpose Journey to Work in Urban Travel Behaviour. *Transportation*. Vol. 9. (3). 229-248.

HANSON, S. e HANSON, P. (1980). The Impact of Social Roles on Urban Activity Patterns: the Case of Working Men and Women. *Geographical Review*, 70.3, 291-299.

HAUTZINGER, H e KESSEL, P. (1977) Mobilität im Personenverkehr. *Strassenbau und Strassenverkehrstechnik*. Vol. 231. Bonn. W. Germany. Citado por DAMM (1983).

HERZ, R. (1980). Empirical Findings Regarding Rhythms in the Use of Time. Report for the German Research Society's Project on *Time budget Allocation and Use of Infrastructure*. Karlsruhe. F.R. Germany.

HERZ, R. (1981). Transferability of Daily Behaviors over Space. Report for the German Research Society's Project on *Regional Differences in Households' activity patterns*. Karlsruhe. F.R. Germany.

HOLZAPFEL, H. (1980). Time, Space and People's Behavior Over Distance. Paper presented at the *PTRC's Annual Summer Meeting*. Warwick, England.

JONES, P. M. (1977). "New Approaches to Understanding Travel Behaviour: The Human Activity Approach". *Working Paper 28. Transport Studies Unit*. University of Oxford. Oxford. England

JONES, P. M. (1983). The Practical Application of Activity-based Approaches in Transport Planning: An Assessment. *Recent Advanced Travel Demand Analysis*. Aldershot. England. Gower.

JONES, P. M.; DIX M.C.; CLARKE M. I. HEGGIE I. G. (1980) *Understanding Travel Behavior*, Aldershot, England, Gower. Citado por BOWMAN (1995) e DAMM (1983).

JONES, P.; KOPPELMAN F.; ORFEUIL J.P. (1990) "Activity analysis: State-of-the-art and future directions." in P. Jones (ed.) *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*. Aldershot, U.K.: Gower.

KITAMURA, R. (1988). An Evaluation of Activity Based Travel Analysis. *Transportation*, 15.

KITAMURA, R. (1996) *Applications of Models of Activity Behavior for Activity Based Demand Forecasting*. Kyoto University. Japan.

- KITAMURA, R.; KOSTYNIUK, L.P.; UYENO, M.J. (1981) Basic Properties of Urban Time-space Paths: empirical tests. *Transportation Research Record*. 794. 1-7.
- KITAMURA, R.; PENDYALA, R. M.; PAS E.I.; REDDY P. (1995) Application of AMOS, an Activity-Based TCM Evaluation Tool, to the Washington, D.C., Metropolitan Area. 23rd *European Transport Forum: Proceedings of Seminar E Transportation Planning Methods*. PTRC Education and Research Services, Ltd., London. Pp. 177-190.
- KUTTER, E. (1973) A Model for Individual Travel Behavior. *Urban Studies* 10. 235-258.
- LEAL, J. E. (1983). Bases para um modelo desagregado de geração repartição modal integrado. *Modelos Urbanos, Regionais e de Transportes*. Ed. IPT.
- LENTORP, B. (1976). *Paths in Space-Time Environments*. Lund studies in Geography. Series B. N.º 44. Lund: CWK Gleerup. Citado por DAMM (1983)
- MADDEN, J. F. (1981). Why Women Work Closer to Home. *Urban Studies*. 8:181-194.
- MCGUCKIN N.; MURAKAMI E. (1999) Examining Trip-Chaining Behavior - A Comparison of Travel by Men and Women. *Transportation Research Record* 1683, Transportation Research Board, Washington, D.C.
- MCNALLY M. G. (2000). The Activity-Based Approach. Paper of Institute of Transportation Studies and Department of Civil & Environmental Engineering. UCI-ITS-AS-WP-00-4. California. USA.
- MESA, D.M.; TSAI, P.; CHAMBERS R.L. (2000). *Using Tree-Based Models for Missing Data Imputation: An Evaluation Using UK Census Data*. Forthcoming in Research Papers of Statistics Netherlands. Currently available in the Autlmp website.
- MITCHELL, R.; RAPKIN, C. (1954). *Urban Traffic – a Function of Land Use*. Columbia University Press.
- ORTÚZAR. J. D.; WILLUMSEN L.G. (1994) *Modelling Transport* (2.Ed.) Wiley. Chichester.
- QUINLAN, I. R. (1983) Learning efficient classification procedures and their application to chess end-games. *Machine Learning: An Artificial Intelligence Approach*, Michalski R.S., Carbonell J.G. and Mitchell T.M. (Eds), Morgan Kaufmann, pp. 463-482.
- QUONIAN L., TARAPANOFF K., ARAÚJO JÚNIOR R.H., ALVARES, L. (2001) Inteligência Obtida pela Aplicação de “Data Mining” em Base de Teses Francesas sobre o Brasil. *Ci. Inf.*, Brasília, v. 30, n. 2, p. 20-28.
- RECKER, W.; GOLOB, T. F.; MACNALLY M. G.; LEONARD J. D. (1987) Dynamic Tests of a time-space model of complex travel behaviour. *Travel Behaviour Research. Fifth International Conference on Travel Behaviour*. Avebury. Aldershot. Gower.
- RICHARDSON A. J.; AMPT E. S.; MEYBURG A. H. (1995). *Survey Methods for Transport Planning*. Eucalyptus Press. Melbourne.
- ROSENBLOOM, (1987). The impact of growing children on their parent’s behavior: a comparative analysis. *Transportation Research Board*. 1135: 17-25.

SCHEUCH, E. K. (1972) *The Use of Time*. The Hague: Mouton. The time-budget interview, in Alexander Szalai (ed.).

SINGELL, L. D.; LILLYDAHL J.H. (1986). An Empirical Analysis of the Commute to Work Patterns of Males and Females in Two-Earner Households. *Urban Studies*, 2: 119-129.

STRATHMAN J. G.; DUEKER K. J. (1995) *Understanding Trip Chaining*. 1990 NPTS Special Reports on Trip and Vehicle Attributes. Report FHWA-PL-95-033. FHWA, U.S. Department of Transportation.

S-PLUS 2000 (1999a). User's Guide. Data Analysis Products Division. *MathSoft*. Seattle. WA.

S-PLUS 2000 (1999b) . Programmer's Guide. Data Analysis Products Division. *MathSoft*. Seattle. WA.

SZALAI, A. ed. (1972). *The Use of Time*. Mouton Press. The Hague. Citado por DAMM (1983).

TARAPANOFF, K. (Org.). (2001) *Inteligência Organizacional e Competitiva*. Brasília. Editora Universidade de Brasília.

UPP (2002). UPP 254 Piloto. *Instituto Lidas*. <http://www.lidas.org.br/upp/oquee.html> (18/07/2002).

VASCONCELLOS, E. (2000). *Transporte Urbano nos países em Desenvolvimento – reflexões e propostas*. 3.º Ed. São Paulo. AnnaBlume. 284p.

WACHS, M. (1987). Men, Women, and Wheels: The Historical Basis of Sex Differences in Travel Patterns. In *Transportation Research Record 1135*, TRB, National Research Council, Washington, D. C., pp. 10-16.

WANG, D. (2001). *An Exploratory Analysis and Travel Patterns in Hong Kong*. Department of Geography. Hong Kong Baptist University.

ZAHAVI, Y. (1974). Travel Time Budgets and Mobility in Urban Areas. Report in FWHA PL8183. U.S. Department of Transportation. Washington. D.C. Citado por DAMM (1983).