

APLICAÇÃO CONJUNTA DE TÉCNICAS DE ANÁLISE MULTIVARIADA DE DADOS E ANÁLISE ESPACIAL EXPLORATÓRIA PARA AVALIAÇÃO DO DESEMPENHO DE TRANSPORTE PÚBLICO POR ÔNIBUS

Cira Souza Pitombo

Vivian de Oliveira Fernandes

Mauro José Alixandrini Junior

Departamento de Transportes/ Escola Politécnica/ Universidade Federal da Bahia

Denise Vaz de Carvalho Santos

Mestrado em Engenharia Ambiental Urbana/ Universidade Federal da Bahia

Tiago Caetano da Silva Teixeira

Tecnologia em Transporte Terrestre – Escola Politécnica/ UFBA

RESUMO

O principal objetivo do presente trabalho foi fazer uma análise do desempenho do sistema de transporte público por ônibus da cidade de Salvador sob a ótica do usuário. Foram utilizadas ferramentas espaciais e não espaciais para apoio à análise. As técnicas não espaciais utilizadas foram (1) Análise de *Cluster*, a qual identificou grupos de usuários completamente insatisfeitos e que consideram o serviço regular; (2) Árvore de decisão, que distinguiu as principais variáveis independentes que influenciaram a satisfação do usuário; e (3) Regressão Logística, com a finalidade de corroborar resultados obtidos pelas técnicas exploratórias. A análise espacial permitiu a visualização dos bairros críticos por região principal. Resultados apontaram como variáveis críticas a confiabilidade do sistema, a segurança, tarifa e locais de parada. A análise por região identificou a população residente nos locais mais pobres (A- Subúrbio e B – Miolo) como aquela mais descontente.

ABSTRACT

The main aim of the present work was analyzing the performance of public transportation system for bus of the Salvador city concerning the user view. Spatial and not spatial support tools had been used. The not spatial techniques used were: (1) Cluster Analysis, which identified groups of completely unsatisfied users and regular satisfaction regarding the service; (2) Decision Tree, that distinguished the main independent variables that influenced the satisfaction of the user; and (3) Logistic Regression, with the purpose to corroborate gotten results from the exploratory techniques. The spatial analysis allowed the visualization of the critical quarters for main regions. Results had pointed as critical variables the constancy of the system, the security, tariff and stop places. The analysis for region identified that the residents in poorer places (A -*Subúrbio* and B - *Miolo*) are the more unsatisfied.

1. INTRODUÇÃO

Atualmente, grande parte das cidades do mundo promove diferentes políticas urbanas de gestão e planejamento da mobilidade, priorizando o gerenciamento da demanda das viagens, em detrimento da criação de novas infra-estruturas para o transporte. A nova tendência da mobilidade urbana vai de encontro a políticas utilizadas no passado que favoreciam a construção ou ampliação permanente da infra-estrutura viária, incentivando o uso do automóvel, aumentando, conseqüentemente os congestionamentos. Desta forma, a atual dificuldade de circulação prejudicou inclusive o transporte público que é realizado junto com o trânsito em geral. A nova abordagem mundial está voltada ao incentivo ao uso de modos de transporte não motorizados e público.

Vale ressaltar que nos países desenvolvidos a maioria da população tem recursos financeiros para efetuar viagens por automóvel, como também dispõe de sistemas de transportes coletivos de boa qualidade, ao se contemplar atributos tais como regularidade, segurança e confiabilidade, por exemplo. Assim, por em prática políticas de incentivo ao uso de transporte público e não motorizado, aliadas a estratégias de restrições ao uso do automóvel torna-se mais viável. Em países em desenvolvimento, no entanto, especificamente no Brasil, deve-se atentar para fatores como falta de segurança nos transportes coletivos, bem como má qualidade no sistema de transporte público como o todo. Aliado a este fato, o recente

incentivo e facilidade de aquisição de automóveis e motocicletas em detrimento de melhorias na qualidade do transporte público fez com que o país caminhasse na “contra-mão” atual do mundo, aumentando a frota de automóveis e agravando os congestionamentos.

Para reverter esta situação é necessário que haja o incentivo ao uso do transporte público, não somente do usuário cativo (em geral população de baixa renda), como especialmente também do atual usuário de transporte individual motorizado. Uma das principais dificuldades então seria tornar o transporte público atraente, apesar da sua falta de flexibilidade, do tempo necessário de espera devido a baixas frequências ou atrasos causados por congestionamentos, necessidade de transbordos, conduções lotadas, etc. Levando-se em conta tais considerações, torna-se necessário avaliar a qualidade do transporte público através de indicadores como tempo de viagem, frequência, conectividade, acessibilidade, lotação e etc., para que propostas de melhorias possam ser seguidas e políticas implementadas.

O objetivo do presente trabalho é avaliar o desempenho do transporte público por ônibus na cidade de Salvador através de ferramentas como técnicas de Análise Multivariada de Dados (AM) e Análise Espacial Exploratória (AEE). A análise não espacial utilizou três técnicas de forma conjunta: (1) Análise de *Cluster* – a fim de discriminar grupos de indivíduos totalmente insatisfeitos daqueles que avaliam o atual sistema como regular; (2) Árvore de Decisão – com objetivo de identificar regiões da cidade que apresentam piores desempenhos por indicador considerado bem como as variáveis mais importantes para o desempenho global do transporte público; e (3) Regressão Logística – com a finalidade de corroborar os resultados obtidos com a técnica de Árvore de Decisão. Para a representação dos dados espaciais foi utilizada a técnica de mapeamento coroplético a fim visualizar espacialmente o desempenho do sistema.

Trata-se de uma avaliação preliminar da qualidade de transporte público por ônibus sob a ótica dos usuários. O trabalho foi inicialmente desenvolvido em caráter didático pela primeira turma do curso superior de Tecnologia em Transporte Terrestre – Escola Politécnica da UFBA. As principais etapas do trabalho, sumariadas nas próximas seções do presente texto encontram-se ilustradas no esquema representado na Figura 1.

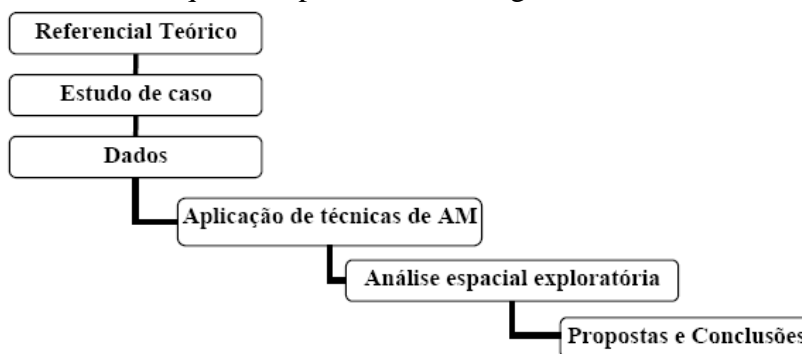


Figura 1 Etapas do trabalho

2. QUALIDADE DO TRANSPORTE

A qualidade do transporte público por ônibus deve ser avaliada sob diversos aspectos. A satisfação diante da qualidade do serviço deve contemplar usuários, comunidade, governo, trabalhadores do setor e empresários do ramo. Ferraz e Torres (2004) propõem como método de avaliação do transporte público, doze fatores principais para caracterização de qualidade para o usuário, tais como: acessibilidade, frequência de atendimento, tempo de viagem, lotação, confiabilidade, segurança, etc. Rodrigues (2006) realizou avaliação do transporte público urbano por ônibus na cidade de São Carlos, sob a ótica da qualidade dos serviços. O autor utilizou três métodos de avaliação: (1) aplicação de pesquisa aos usuários do sistema;

(2) metodologia utilizada pela Prefeitura Municipal de São Carlos no acompanhamento do serviço prestado pela empresa operadora do transporte; e (3) entrevistas a três especialistas que procederam à avaliação dos fatores considerando os padrões estabelecidos por Ferraz e Torres (2004). Através dos resultados, foram identificados os principais aspectos positivos e negativos do transporte coletivo.

Budiono (2009) faz um estudo sobre a satisfação do usuário quanto à qualidade do serviço de transporte público em duas cidades da Indonésia. A investigação se baseia na aplicação de questionários que visam investigar o grau de satisfação com o serviço prestado e os fatores que mais influenciam a satisfação dos usuários de transporte público. Os atributos para avaliação foram frequência, tempo de viagem, pontualidade, preço, sistema de informações, limpeza, comportamento dos operadores, conforto, disponibilidade de assentos, condições dos pontos de parada e segurança. O estudo concluiu que os usuários de transporte público na Indonésia não estão satisfeitos com o serviço. A segurança durante a viagem – quanto a atividades criminosas dentro do ônibus - foi uma das questões críticas apontadas pelos usuários. Outra pesquisa, realizada na Georgia, concluiu que problemas com o horário - como confiabilidade e frequência - somados à falta de conforto e segurança no transporte por ônibus estão levando a população de Tbilisi a adotar de forma crescente a cultura do automóvel (Grdzlishivili, 2010). Na tendência oposta está a cidade de Kaohsiung, em Taiwan, onde as taxas de uso de automóveis têm sido reduzidas consideravelmente. A estratégia mais importante foi implementar sistemas de transporte de massa rápidos. A pesquisa de Wei e Kao (2010) se dedica a avaliar o grau de envolvimento dos usuários com o sistema de transportes, indicando estratégias de marketing para atingir diferentes grupos da sociedade.

3. ESTUDO DE CASO

O município de Salvador, dividido em 22 subdistritos, possui atualmente uma população total estimada em 2.998.056 habitantes (IBGE 2007). Nos últimos anos, Salvador se deparou com um processo de transformação nos planos físico, econômico e social. Tais mudanças acarretaram significativas variações no padrão de mobilidade da cidade. Observou-se o aparecimento de múltiplos centros, maior diversificação das atividades, aumento na economia informal, assim como uma falta de controle do uso do solo. Além disso, verificou-se o aumento de 94% da frota de automóveis em dez anos. Conseqüentemente, houve redução das viagens pendulares e aumento das viagens encadeadas, maior dispersão espacial e temporal dos fluxos de viagens e diluição dos horários de pico.

Segundo relatório anual da TRANSALVADOR (2009), Salvador possui uma frota de 2445 ônibus, com idade média de 4,97 anos, dezoito empresas de ônibus, 445 linhas e uma rede viária de 1223 km. Em termos de transporte público, o Sistema é operado basicamente por ônibus (467,9 milhões de passageiros por ano), micro-ônibus e uma linha de trem. Ainda possui quatro ascensores com 11,6 milhões de passageiros por ano.

Em relação ao transporte público por ônibus, atualmente a cidade apresenta problemas como sobreposição dos atendimentos, ou seja, excesso de linhas concorrentes nos principais corredores, baixa frequência de partidas, falta de regularidade entre as partidas (confiabilidade inexistente), atendimento temporal deficiente, extensões e tempos de viagens críticos, saturação das estações e infra-estrutura deficiente. Outro problema grave, fora do âmbito de Planejamento de Transportes, mas estritamente ligado à segurança pública, é a atual falta de segurança nos coletivos. Segundo Grupo Especial de Repressão a Roubos em Coletivos (GERRC), dois ônibus são assaltados por dia em Salvador e Região Metropolitana.

Para efeito de organização do sistema integrado de linhas, a cidade foi dividida em quatro grandes áreas pelo Sistema de Transporte Coletivo por Ônibus (SCTO). A cidade foi separada em quatro regiões (Subúrbio, Miolo, Orla e Centro), as quais serão utilizadas no presente trabalho. A Figura 2 ilustra as citadas regiões assim como a Tabela 1 apresenta as principais características de cada uma delas.

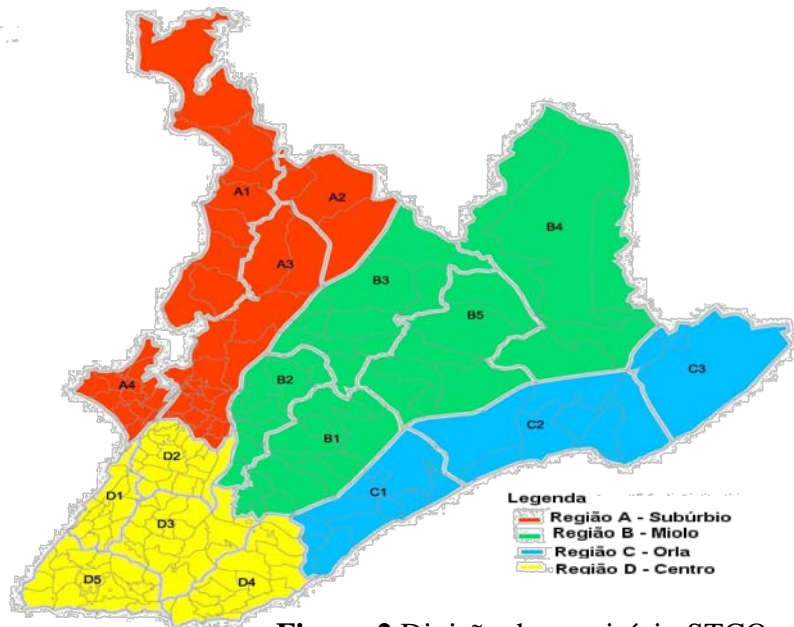


Figura 2 Divisão do município STCO

Tabela 1 Descrição das regiões

A	Subúrbio	A área correspondente ao Subúrbio Ferroviário compõe um dos maiores territórios de pobreza de Salvador. Teve sua ocupação iniciada pela construção da linha férrea. O subúrbio tem aproximadamente 500 mil habitantes, em sua maioria pobres e com baixa escolaridade. Há predominância de habitações precárias e deficientes. A região A é localizada na periferia da cidade, formada por trinta bairros. A população encontra-se segregada espacialmente, geralmente com longos percursos de viagem, com necessidade de transbordos. É atendida por 108 linhas e onze empresas.
B	Miolo	A região situa-se, em termos geográficos, na parte central do município de Salvador, possuindo cerca de 11.500 ha. O miolo urbano de Salvador começou a ser ocupado pela implantação de conjuntos residenciais para a “classe média baixa”, tendo a sua expansão continuada por loteamentos populares e sucessivas invasões coletivas. A área do miolo é formada por cerca de 41 bairros que ocupam aproximadamente 36% da superfície da cidade, em uma densidade demográfica no intervalo de 15.000 – 25.000 hab/km ² . A área considerada do miolo urbano teve uma ocupação mista, com predominância de áreas residenciais, apresentando importantes pólos de atração de viagens. A região é atendida por 182 linhas e dezessete empresas.
C	Orla	A orla, especialmente a parte norte (orla marítima norte) constitui a “área nobre” da cidade, local de moradia, serviços e lazer, onde se concentram a riqueza, os investimentos públicos, os equipamentos urbanos e os interesses de produção imobiliária. A região também apresenta bairros populares com alta densidade demográfica como Boca do Rio. A região é composta de treze bairros, sendo operada por 8 empresas e 30 linhas regulares.
D	Centro	A área do centro, composta em geral por bairros bem atendidos em serviços e equipamentos urbanos, difere das áreas do subúrbio ferroviário e do miolo urbano que podem ser considerados como territórios de pobreza e violência. Nesta região também são encontrados vários pólos geradores, como: centro da cidade, <i>shoppings</i> , comércio, estações, centro histórico, hospitais, faculdades, escolas, praias e vários pontos turísticos e repartições públicas. Tendo a maior densidade demográfica e de atividades econômicas, é atendida por 18 empresas de ônibus, 124 linhas e contém 4 terminais de ônibus.

4. DADOS

Para obtenção da amostra final, foram seguidas algumas etapas: (1) elaboração dos questionários; (2) entrevistas em campo; (3) tabulação de dados e; (4) tratamento da amostra.

A elaboração do questionário foi feita com auxílio dos alunos do curso superior de Tecnologia em Transporte Terrestre, utilizando alguns parâmetros propostos por Rodrigues (2006). Baseando-se ainda na metodologia do mesmo autor, o usuário deveria responder o seu grau de satisfação para cada parâmetro/indicador, bem como a importância de cada parâmetro para a qualidade do transporte público por ônibus na cidade de Salvador. O questionário é composto por 25 questões relativas aos seguintes parâmetros: Acessibilidade (2 questões); Segurança (3 questões); Frequência (1 questão); Tempo de viagem (3 questões); Lotação (1 questão); Confiabilidade (1 questão); Características dos veículos (3 questões); Locais de parada (3 questões); Sistema de Informação (1 questão); Conectividade (3 questões); Operadores (1 questão); Vias (1 questão); Tarifa (2 questões).

As etapas de entrevistas e tabulação dos dados foram realizadas pelos alunos da disciplina. O banco de dados final foi composto por 430 usuários de transporte público por ônibus residentes nas quatro regiões (Miolo, Subúrbio, Centro e Orla), 25 variáveis numéricas discretas (escala *Likert* com cinco posições) e uma variável contínua que representava o desempenho global do transporte público, variando de 1-25. A Tabela 2 traz um quadro com as médias das notas para cada parâmetro por região. Pode-se verificar nesta análise preliminar da amostra que as regiões Orla e Centro possuem usuários um pouco menos insatisfeitos quando comparados aos usuários residentes nas regiões do Miolo e Subúrbio.

Tabela 2 Dados regiões

Fatores de qualidade	Regiões			
	A	B	C	D
Acessibilidade	2,6	2,9	3,2	2,9
Segurança	2,4	2,2	2,4	2,8
Frequência	2,5	2,2	2,8	2,3
Tempo de viagem	2,2	2,1	3,8	2,6
Lotação	2,0	1,8	2,2	2,3
Confiabilidade	2,6	2,4	2,9	2,7
Veículos	2,7	2,4	2,7	2,6
Locais de parada	2,4	2,0	2,3	2,3
Sistema de Informação	2,2	2,0	2,1	2,2
Conectividade	2,6	2,6	2,5	2,5
Operadores	2,8	2,6	2,9	2,9
Vias	2,8	2,6	3,0	2,9
Tarifas	2,5	2,7	2,8	2,8
GLOBAL	9,9	9,6	10,5	10,0

5. APLICAÇÕES DAS TÉCNICAS DE AM

Hair *et al.* (1998) conceituam AM como um conjunto de técnicas estatísticas utilizadas com o objetivo de explicar e prever o grau de relações entre diversas variáveis independentes (inclusive entre si) e a variável dependente. Neste trabalho, foram utilizadas três técnicas de AM em conjunto, descritas nas próximas subseções.

5.1 Análise de *Cluster* (AC)

A Análise de *Cluster* (AC) é um conjunto de técnicas estatísticas cujo objetivo é agrupar objetos segundo suas características, formando grupos homogêneos. Os objetos em cada grupo tendem a ser semelhantes entre si, porém diferentes dos demais objetos dos outros conglomerados. Os conglomerados obtidos devem apresentar tanto uma homogeneidade interna (dentro de cada conglomerado), como uma grande heterogeneidade externa (entre conglomerados) (Hair *et al.*, 1998).

Com a finalidade de agrupar usuários segundo seu grau de satisfação global do serviço de transporte público por ônibus na cidade de Salvador foi feita a aplicação da AC. Assim, aplicou-se a AC (*software* SPSS 15.0), considerando a variável contínua “Satisfação global”, o método de aglomeração *TwoStep Cluster*, a medida de distância *Log-likelihood*, sem

determinar número fixo de *clusters*. A variável “Satisfação global” é uma média ponderada de todos os indicadores de qualidade. O seu cálculo consiste simplesmente na média dos produtos entre a satisfação dos indicadores (parâmetros) e a importância atribuída pelo usuário. O valor da variável contínua varia entre 1 e 25.

Foram obtidos dois grupos de usuários, homogêneos segundo a variável “Satisfação global” indicados na Tabela 3. A mesma tabela mostra as médias não somente da variável “Satisfação global”, como também dos demais indicadores considerados. As médias do grau de importância atribuído pelos usuários por indicador também estão representadas. Além disso, verifica-se a distribuição das regiões (A, B, C e D) em cada grupo bem como uma nomenclatura adotada pelos autores para caracterizar os dois grupos.

Tabela 3: Características gerais dos dois aglomerados obtidos

Grupos	SATISFAÇÃO GLOBAL				REGIÃO				Acessibilidade		Segurança	
	Média	Desvpad	Mínimo	Máximo	A (%)	B (%)	C (%)	D (%)	Satisf	Import	Satisf	Import
0	8,26	1,34	4,20	10,20	27,35	31,84	20,82	20,00	2,60	3,90	2,20	4,30
1	12,32	1,73	10,60	20,00	23,91	19,57	31,52	25,00	3,20	4,20	2,90	4,70
Grupos	Frequencia		Tempo de viagem		Lotação		Confiabilidade		Veículos		Locais de parada	
	Satisf	Import	Satisf	Import	Satisf	Import	Satisf	Import	Satisf	Import	Satisf	Import
0	2,10	3,90	2,20	3,90	1,80	3,90	2,30	4,10	2,30	3,90	2,00	3,90
1	2,80	4,50	2,60	4,30	2,50	4,30	3,40	4,50	2,90	4,40	2,50	4,40
Grupos	SI		Conectividade		Operador		Vias		Tarifa		Nomenclatura	
	Satisf	Import	Satisf	Import	Satisf	Import	Satisf	Import	Satisf	Import		
0	1,90	4,00	2,40	3,90	2,60	4,10	2,60	4,20	2,40	4,10	Completamente insatisfeitos	
1	2,30	4,40	2,80	4,30	3,00	4,50	3,10	4,60	2,90	4,40	Satisfação regular	

A nomenclatura foi adotada considerando a média, desvio padrão, valor mínimo e máximo da “Satisfação global” por grupo. Assim, o primeiro grupo (grupo 0) caracteriza os usuários “**Completamente insatisfeitos**”, com valores de “satisfação global” muito baixos. Este grupo é composto predominantemente por usuários residentes nas regiões A (subúrbio) e B (Miolo). Tais regiões são as mais pobres e carentes da cidade. Usuários residentes no subúrbio, por exemplo, enfrentam altos tempos de viagem, já que os mesmos encontram-se espacialmente segregados e precisam percorrer altas distâncias, além da necessidade de transbordo. Considerando valores médios de satisfação e importância por indicador, o grupo dos usuários “**Completamente insatisfeitos**” apresenta valores baixos (abaixo ou igual a 2,50) de desempenho e altos de importância para a maioria dos indicadores, exceto “Acessibilidade” e “Comportamento dos operadores”.

O grupo 1, dos usuários que julgam o serviço de transporte público por ônibus de Salvador “regular”, denominados “**Usuários com satisfação regular**”, é composto por usuários que na sua maioria, possuem valores médios de “Satisfação global”. Mais frequentemente são residentes das regiões C (Orla) e D (Centro). Conforme descrição sucinta da Tabela 1, ambas as regiões são mais desenvolvidas e bem atendidas em relação a serviços e equipamentos urbanos. No entanto, vale ressaltar que, de uma maneira geral, o usuário está descontente com o serviço de ônibus, independente da região que reside. Fatores como “Lotação”, “Locais de parada” e “Sistemas de Informação” merecem uma atenção especial para o grupo 1.

5.2 Árvore de Decisão

Outra técnica utilizada neste trabalho, em conjunto com a AC, é Árvore de Decisão (AD), considerada uma forma simples de representação de relações existentes em um conjunto de dados. Os dados são divididos em subgrupos, com base nos valores das variáveis. O resultado é uma hierarquia de declarações do tipo “Se ... então ...” que são utilizadas, principalmente, para classificar dados. Para processamento da árvore foi utilizado o *Software* SPSS 15.0.

Utilizou-se o algoritmo do CART, que 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 dependente. Essas divisões são representadas por estrutura de árvore binária, sendo que cada nó corresponde a uma divisão.

A árvore foi gerada a partir da amostra final com 430 usuários, variável dependente (*clusters* – variável com duas categorias – grupos descritos na sub-seção anterior), adotando-se o mínimo de 25 observações por nós filhos. As variáveis independentes foram aquelas que representavam a satisfação dos indicadores considerados no trabalho, bem como a região que reside (variável com quatro categorias: A, B, C e D).

A variável de maior importância (que melhor explica a variabilidade dos dados) é “Confiabilidade”. A partir da raiz, a árvore se ramifica em dois grupos principais: (1) “Confiabilidade” menor ou igual a 2.5 (nó 1 – 77.9% completamente insatisfeitos) e (2) “Confiabilidade” superior a 2.5 (nó 2 – 59.2% satisfação regular). Tal variável relaciona-se ao grau de certeza dos usuários de que o veículo de transporte público vai passar na origem e chegar ao destino no horário previsto (Ferraz e Torres, 2004). Diante deste conceito, pode-se afirmar que a “Confiabilidade” engloba critérios como “Tempo de viagem”, “Frequência” e “Conectividade”. Este fato pode explicar a sua importância quanto à variabilidade em relação à variável dependente (indivíduos completamente insatisfeitos e indivíduos que acham o transporte público regular). Posteriormente ocorrem novas segmentações do conjunto de dados considerando variáveis como “Segurança”, “Tarifa”, “Locais de parada” e “Região”. Ao final da segregação dos dados foi encontrado um total de 6 folhas. A Figura 3 representa a árvore gerada. Nas folhas encontram-se ilustrados as categorias da variável dependente (Grupo 0 e Grupo 1) e a frequência de cada grupo em cada nó.

Observando-se os nós terminais, pode-se afirmar que “Tarifa” pode interferir na satisfação do transporte público por ônibus. As folhas 3 e 4 apresentam grau de satisfação quanto à confiabilidade baixa (menor ou igual a 2.5) e são diferenciadas pelo grau de satisfação da tarifa (nó 3 – “Tarifa” ≤ 3.2 ; nó 4 – “Tarifa” > 3.2). O grupo de indivíduos que compõem o nó 3 predominantemente é formado por usuários completamente insatisfeitos em relação ao serviço de transporte público por ônibus (Grupo 0 – 85.1%) enquanto que a maioria dos indivíduos classificados no nó 4 julga o serviço regular (Grupo 1 – 62.1%).

“Segurança” foi uma variável que se mostrou de extrema importância para discriminar os grupos de indivíduos quanto à sua satisfação global. A segmentação dos dados a partir do nó 2 mostra que quando o valor de satisfação atribuído para o parâmetro “Segurança” é inferior ou igual a 2.8, a maioria dos indivíduos permanece extremamente insatisfeita, mesmo que o grau de “Confiabilidade” seja relativamente alto (nó 5 – Grupo 0 – 55.4%). No entanto, aliando-se alto grau de satisfação relativo à “Confiabilidade” e alto grau de satisfação relacionado à “Segurança”, verifica-se que a maior parte dos usuários considera o transporte regular (nó 6 – Grupo 1 – 80.2%). Finalmente, observa-se a influência da variável “Locais de parada” (nó 7 e 8) na satisfação do usuário, bem como da variável “Região” (nó 9 e 10). Verifica-se novamente que indivíduos residentes nas regiões A e B (Subúrbio e Miolo, respectivamente) estão menos satisfeitos do que aqueles residentes nas regiões melhor atendidas C e D (Orla e Centro, respectivamente). Vale ressaltar ainda que, apesar do alto grau de satisfação atribuído a “Confiabilidade” (superior a 2.5) e “Segurança” (superior a 2.8), a região que usuário reside, e conseqüentemente, inicia e finaliza pelo menos a primeira e última viagem do dia pode influenciar seu grau de satisfação.

5.3 Regressão Logística

A regressão logística é uma forma especializada de regressão que é formulada para prever e explicar uma variável categórica binária, e não uma medida dependente numérica. Como na regressão linear, os coeficientes indicam o impacto relativo de cada variável preditora. O modelo de regressão não linear logístico é utilizado para prever diretamente a probabilidade de um evento ocorrer. Assim, a interpretação dos seus coeficientes está relacionada ao aumento ou diminuição da probabilidade de um evento ocorrer. A curva logística é em forma de S e os coeficientes são expressos em logaritmos. (Hair et al, 1998). Seu modelo calcula a probabilidade de ocorrência do evento pela seguinte fórmula:

$$P(X) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} \quad (1), \text{ Em que, } P(x) = \text{Probabilidade de ocorrência; } \alpha, \beta_i \text{ parâmetros.}$$

No presente trabalho, o uso de regressão logística deu-se a fim de corroborar resultados inicialmente obtidos com a Árvore de Decisão. Assim, através dos coeficientes das variáveis, bem como parâmetros que mensurem o ajuste do modelo geral (-2LL, R^2 de *Cox e Snell*, R^2 de *Nagelkerke*), é possível avaliar a qualidade de ajuste do modelo, verificar a importância das variáveis independentes no cálculo da probabilidade dos usuários estarem no grupo dos usuários que consideram o transporte público regular (Grupo 0 (completamente insatisfeitos) = 0; Grupo 1 (satisfação regular) = 1), bem como identificar se o modelo é bem explicado pelas variáveis independentes. Assim, para estimação do modelo logístico foram consideradas as variáveis descritas na Tabela 4.

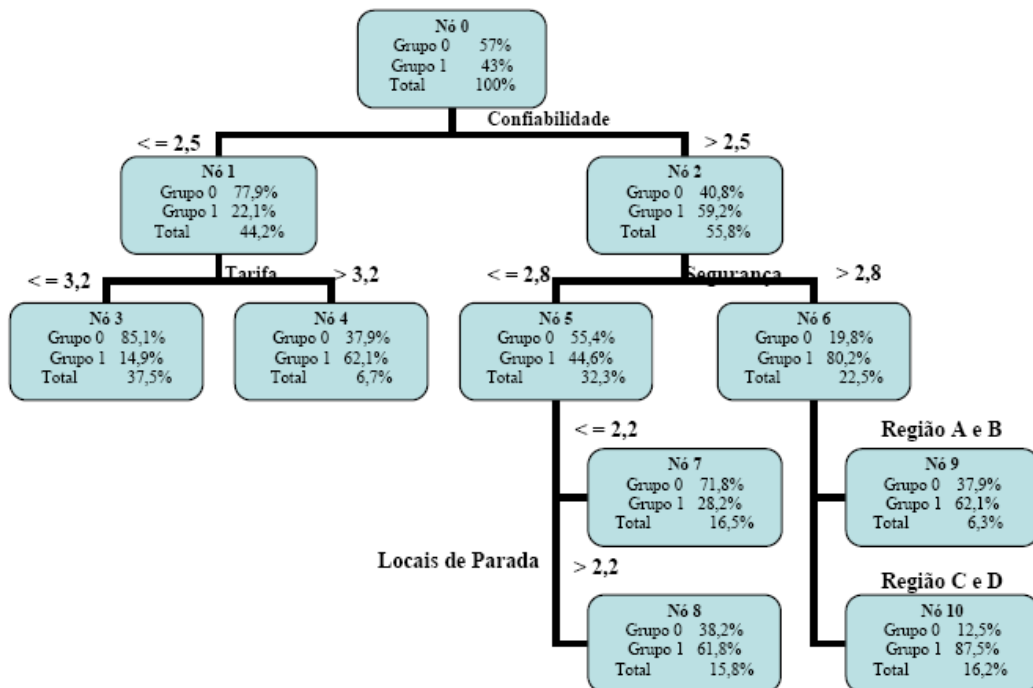


Figura 3 Árvore de Decisão – Satisfação do Transporte Público

O modelo de regressão logístico foi estimado utilizando o método *Forward Stepwise: Wald*. Neste método cada variável é incluída separadamente em cada passo do modelo e analisado seu desempenho em estimar a “satisfação” do usuário. O processo termina até as variáveis que melhor explicam conjuntamente a variável binária “satisfação” forem selecionadas, chegando ao modelo final. O processo foi formado por 7 passos. O sétimo modelo apresentou valores menores de -2LL, bem como valores maiores de R^2 de *Cox e Snell*, R^2 de *Nagelkerke*. Além disso, o modelo apresentou uma frequência de acerto de 81,6% para o grupo dos

“completamente insatisfeitos” (Grupo 0) e 73,2% para o grupo daqueles que consideram o transporte regular (“satisfação regular – Grupo 1).

Tabela 4 Variáveis consideradas no modelo logístico

Dependente	Descrição	Independentes		Descrição
Satisfação	0 - Completamente insatisfeitos 1 - Satisfação regular	Acessibilidade - Satisfação	Segurança - Satisfação	1 - péssimo; 2 - ruim;
		Frequência - Satisfação	Tempo de viagem - Satisfação	
Independentes	Descrição	Lotação - Satisfação	Confiabilidade - Satisfação	3 - regular; 4 - bom; 5 - ótimo
Região A	0 - não reside no subúrbio 1 - reside no subúrbio	Veículo - Satisfação	Locais de parada - Satisfação	
Região B	0 - não reside no miolo 1 - reside no miolo	Sistema de informação - Satisfação	Conectividade - Satisfação	
Região C	0 - não reside na orla 1 - reside na orla	Operadores - Satisfação	Vias - Satisfação	
Região D	0 - não reside no centro 1 - reside no centro	Tarifa - Satisfação		

A Tabela 5 mostra as variáveis que entraram no modelo (modelo 7), gerada pela ferramenta de regressão logística bem como os seus coeficientes e os resultados do teste de significância das variáveis independentes. Na coluna “B (coeficientes)”, são apresentados, os coeficientes do modelo para cada variável explicativa. Já na coluna Erro Padrão, é apresentada uma medida de variabilidade destes coeficientes. A estatística *Wald* é utilizada para testar a hipótese de cada coeficiente do modelo de regressão. Tal teste irá identificar o quanto à variável explicativa participa individualmente da explicação da variável dependente (“satisfação”). A estatística de *Wald*, testa para cada coeficiente das variáveis do modelo a seguinte hipótese de que o coeficiente da variável explicativa é igual a zero contra uma hipótese alternativa de que não são iguais a zero. A coluna “Exp(B)” é o exponencial dos coeficientes estimados de cada categoria variável do modelo.

Tabela 5 Variáveis explicativas do modelo

Variáveis do modelo		B (coef)	Erro padr	Wald	G.	Sig.	Exp(B)
Modelo 7	AcessSatisf	0,3968	0,1585	6,2716	1	0,01	1,49
	SegurSatisf	0,7438	0,1971	14,2481	1	0,00	2,10
	FreqSatisf	0,3639	0,1366	7,0963	1	0,01	1,44
	ConfSatisf	0,6513	0,1513	18,5373	1	0,00	1,92
	Locais de paradaSatisf	0,6291	0,1865	11,3780	1	0,00	1,88
	OpeSatisf	0,2996	0,1409	4,5250	1	0,03	1,35
	TarifaSatisf	0,6588	0,1744	14,2712	1	0,00	1,93
	Constant	-10,0734	1,0007	101,3258	1	0,00	0,00

Considerando os parâmetros apresentados na tabela, verifica-se que as variáveis com maiores valores de coeficientes são, em ordem decrescente “Segurança”, “Tarifa”, “Confiabilidade” e “Locais de Parada”. Todas estas variáveis também foram selecionadas pela Árvore de Decisão e têm um poder discriminatório de grupos de satisfação de usuários de transporte público. Valores altos de grau de satisfação especialmente destas quatro variáveis aumentam a probabilidade de determinado usuário considerar o sistema de transporte público regular. A variável “Região”, por sua vez, não entrou no modelo logístico. Pode-se afirmar que possivelmente “satisfação” não está apenas relacionada a quesitos estritamente relacionados ao transporte público como frequência, tempo de viagem, etc, mas, sobretudo a problemas de segurança pública (como segurança, por exemplo) e sociais (valor da tarifa). A próxima seção traz uma análise espacial exploratória das principais variáveis selecionadas tanto pela Árvore de Decisão como pelo modelo logístico. É possível visualizar, dentro de cada uma das quatro regiões, padrões diferentes relacionados ao grau de satisfação.

6. ANÁLISE ESPACIAL EXPLORATÓRIA

Uma série de técnicas estatísticas vem sendo definidas e adaptadas aos dados espaciais. Estas técnicas, combinadas com funções de visualização, formam, um conjunto de ferramentas que também são conhecidas como análise exploratória de dados espaciais (Anselin e Bao, 1997). No presente trabalho, com os dados tabulares foram gerados vários mapas de diferentes temas os quais foram analisados neste trabalho. Foi utilizada uma base de dados vetoriais com os limites de bairros de Salvador. Verifica-se que o levantamento de dados não foi realizado em todos os bairros da cidade, uma vez que o município foi dividido em regiões, porém a análise espacial foi realizada levando em consideração os bairros onde moravam os entrevistados, ocasionando algumas regiões em branco. Recomenda-se que em futuros estudos, uma amostra maior de dados seja levantada para maior efetividade da representação.

Para a representação dos dados espacialmente foi utilizada a técnica de mapeamento coroplético. Os mapas coropléticos são elaborados com dados quantitativos e apresentam sua legenda ordenada em classes conforme as regras próprias de utilização da variável visual valor por meio de tonalidades de cores. Levando-se em consideração que a variável investigada estima os valores de satisfação podem ser positivos, com conceito bom e excelente e negativo, com conceito ruim e péssimo. Para uma analogia ao semáforo, utilizaram-se as cores vermelho para avaliações negativas, verde para as positivas e amarela para as avaliações intermediárias. Os mapas no modo de implantação zonal são os mais adequados para representar distribuições espaciais de dados que se refiram as áreas, que no caso deste trabalho a análise foi realizada levando-se em consideração os bairros onde os usuários entrevistados residiam.

Foram confeccionados mapas temáticos para todos os parâmetros avaliados na pesquisa. Para exemplificação desta análise são apresentados nas Figuras 4 e 5 os mapas referentes aos critérios: “Segurança” e “Satisfação global”. Na Figura 4, é apresentado o mapa de satisfação do usuário quanto à segurança no transporte público no município de Salvador/BA. A legenda mostra a correspondência entre as notas atribuídas na segurança do transporte público e os conceitos. Aos bairros onde houve residentes entrevistados são atribuídas as tonalidades graduadas em vermelho, amarelo e verde, uma analogia ao semáforo e os bairros em branco, são aqueles em que não houve entrevistados com residência nos mesmos. Ao analisar as tendências apresentada no mapa da Figura 4, verifica-se que a região do miolo apresenta alternância entre avaliações positivas e negativas, sem uma tendência homogênea, revelando a necessidade de uma investigação mais aprofundada e uma amostra mais representativa de dados para esta região. Ao contrário da região da orla que apresenta uma tendência predominantemente negativa (vermelhos e amarelos). No subúrbio verifica-se uma avaliação regular da segurança no transporte público e no centro predominantemente positiva, com exceção do bairro Rio Vermelho.

Através da análise da Figura 5, o mapa apresenta o comportamento das avaliações por bairro ponderadas pela importância dada pelo usuário a cada um dos critérios. Neste mapa a região da orla apresenta avaliações regulares e positivas, isto reflete que para os critérios de maior importância dos usuários naquela região, a avaliação é predominantemente positiva. Ao contrário do que se percebe nas regiões do miolo e do subúrbio e onde a avaliação é mais negativa, porque nos critérios mais importantes aos usuários, o transporte obteve uma avaliação desfavorável. Percebe-se que a avaliação dos usuários, levando-se em consideração da importância dada aos critérios, reflete um resultado mais negativo para estas duas regiões, do que no mapa apresentado na Figuras 4. A região do centro, no entanto, apesar de apresentar uma avaliação global regular na análise não espacial, apresenta um “grupo de bairros litorâneos” com avaliações negativas.

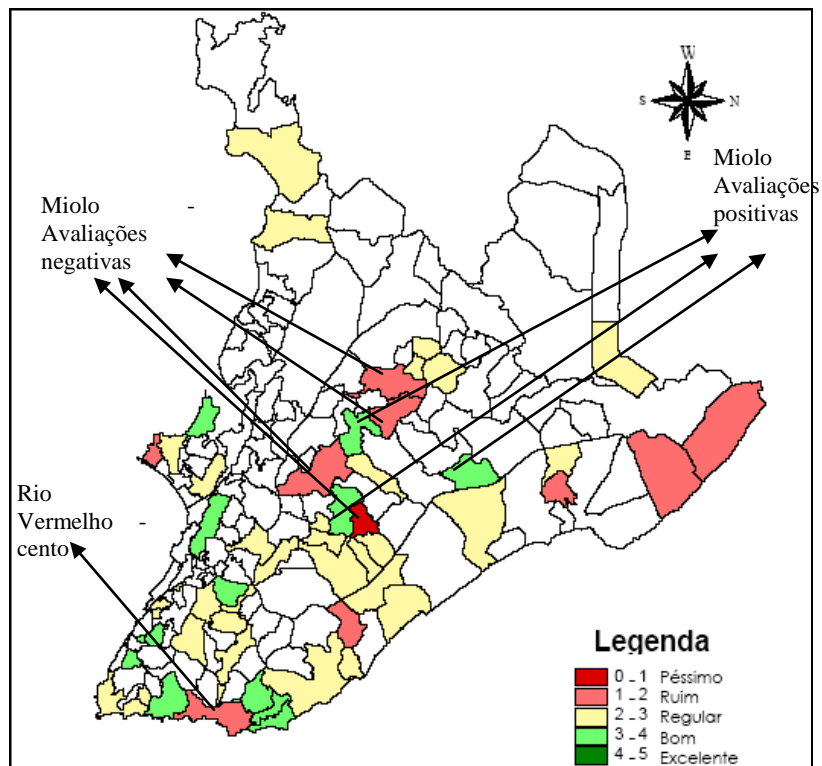


Figura 4 - Mapa de satisfação do usuário quanto à segurança –Salvador/BA

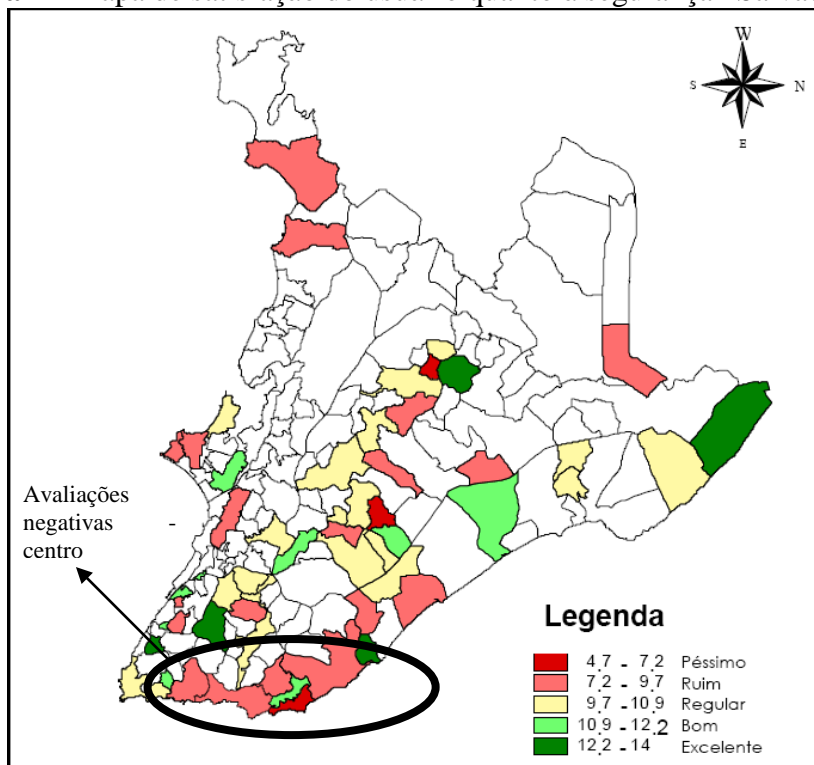


Figura 5 – Satisfação global do usuário no Município de Salvador/BA.

7. CONCLUSÕES

O presente trabalho se propôs a fazer uma análise de desempenho do Sistema de Transporte Público por ônibus na cidade de Salvador, sob a ótica do usuário. Através do uso de Técnicas

de Análise Multivariada de dados foi possível detectar grupos de usuários homogêneos em relação à satisfação global do serviço de transporte público (Análise de *Cluster*), bem como identificar variáveis importantes (dentre as 25 originais) que influenciam a variabilidade dos dados em relação à variável resposta, ou seja grupos de indivíduos completamente insatisfeitos ou que consideram o serviço regular (Árvore de Decisão). A técnica comprobatória utilizada foi a Regressão Logística. Tanto o modelo de árvore quanto o logístico identificaram os parâmetros “Confiabilidade”, “Segurança”, “Tarifa” e “Locais de parada” como os mais importantes para discriminar usuários pertencentes ao grupo de indivíduos completamente insatisfeitos daqueles que consideram o transporte regular.

Dos quatro parâmetros apontados pelas técnicas, dois deles não estão necessariamente ligados ao sistema de transportes. Pode-se dizer que “Segurança” e “Tarifa” são indicadores que caracterizam também problemas sociais e de segurança pública das principais capitais brasileiras. O indicador “Região”, identificado pela Árvore de Decisão como uma variável importante na segregação dos dados apontou um maior descontentamento do usuário nas regiões mais carentes da cidade (A-Subúrbio; B – Miolo).

A Análise espacial exploratória identificou os graus de satisfação dos parâmetros considerados. No presente trabalho foram apresentados mapas relacionados à “Segurança” e “Satisfação Global”. Tal análise desagregada permitiu que fossem identificados bairros com avaliações positivas ou negativas de todas as regiões.

Vale ressaltar que se trata de um trabalho preliminar, com finalidade inicialmente didática. A coleta de dados foi realizada pela turma do curso de Tecnologia em Transporte Terrestre (UFBA), justificando o tamanho da amostra. Para uma análise espacial mais rica é necessário um banco de dados mais diversificado, contemplando residentes de boa parte dos bairros da cidade. A presente coleta da região do subúrbio, por exemplo, foi restrita a apenas cinco bairros. Entretanto, através do trabalho, pode-se apontar, por exemplo, que políticas de transportes poderiam contemplar critérios como Confiabilidade, por exemplo, através de uma melhor operacionalização, frequência de atendimento condizendo com as demandas atuais, dimensionamento dos itinerários de forma a evitar transbordos, etc.

Agradecimentos À primeira turma do curso superior de Tecnologia em Transporte Terrestre – Escola Politécnica da Universidade Federal da Bahia

REFERÊNCIAS BIBLIOGRÁFICAS

- Anselin, L. & Bao, S. Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In: Fischer, M. M. & Getis, A. *Recent developments in spatial analysis*. New York: Springer, 1997, p. 35-59.
- Budiono, O. A. (2009) *Customer Satisfaction in Public Bus Transport - A study of travelers' perception in Indonesia*. Dissertação de Mestrado, Karlstad University.
- Ferraz, A. C. P.; Torres, I. G. E. (2004). Transporte Público Urbano. 2a. ed. São Carlos, Rima.
- Grdzlishvili, I. e Sathre, R. (2010) Understanding the urban travel attitudes and behavior of Tbilisi residents. *Transport Policy*, disponível online desde 07/06/2010 em www.sciencedirect.com.
- Hair, J.F.; R.E Anderson.; R.L Tatham.; W.C. Black (1998). *Multivariate Data Analysis*. 5ª ed. Prentice-Hall. Upper Saddle River, New Jersey, 730p.
- IBGE (2007) Instituto Brasileiro de Geografia e Estatística. Banco de dados. <http://www.ibge.gov.br/home/>
- Rodrigues, M. O. (2006). Avaliação da qualidade do transporte coletivo da cidade de São Carlos. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.
- TRANSALVADOR (2009) Superintendência de Trânsito e Transportes do Salvador. Relatório anual. <http://www2.set.salvador.ba.gov.br/>
- Wei, C.H. e Kao, C.Y. (2010) Measuring traveler involvement in urban public transport services: The case of Kaohsiung. *Transport Policy*, disponível online desde 18/05/2010 em www.sciencedirect.com.

Cira Souza Pitombo cira.pitombo@ufba.br;
Vivian de Oliveira Fernandes vivian.fernandes@ufba.br;
Mauro José Alixandrini Junior mauro.alixandrini@ufba.br;
Denise Vaz de Carvalho Santos denisevaz@gmail.com;
Tiago Caetano da Silva Teixeira caetanooo@gmail.com

Escola Politécnica da Universidade Federal da Bahia – Departamento de Transportes
Rua Aristides Novis, 02, Federação, CEP 40210-630, Salvador, Bahia