

Técnicas avançadas de planejamento de transportes: estudo de caso da rede de trilhos da Região Metropolitana de Campinas e da ligação Jundiaí – Campinas.

Alexandre Frazão D’Andrea ⁽¹⁾; Geraldo Camargo de Carvalho Junior ⁽²⁾; Luis Fernando Kyono ⁽³⁾,

¹ Sistran Engenharia, Rua Santa Isabel, n.º 160, 3º Andar, CEP:01221-010, São Paulo, SP, (11) 3335-2125, afrazao@sistransp.com.br – Coordenador do Departamento de Modelagem de Transportes

² Sistran Engenharia, Rua Santa Isabel, n.º 160, 3º Andar, CEP:01221-010, São Paulo, SP, (11) 3335-2125, gcarvalho@sistransp.com.br – Gerente da SISTRAN Engenharia

³ Sistran Engenharia, Rua Santa Isabel, n.º 160, 3º Andar, CEP:01221-010, São Paulo, SP, (11) 3335-2125, lkyono@sistransp.com.br – Engenheiro Civil da SISTRAN Engenharia

SÍNTESE

O presente trabalho apresenta novas ferramentas de obtenção e análise de dados utilizadas na elaboração de modelos de previsão de demanda da rede de trilhos da Região Metropolitana de Campinas (RMC) e da ligação Campinas – Jundiaí.

Técnicas multivariadas de análise de clusters foram utilizadas para apoiar o desenvolvimento de modelos de atração de viagens com base domiciliar trabalho considerando a heterogeneidade de atividades econômicas de cada cidade na RMC.

A ferramenta de obtenção de grande quantidade de dados (“*Smart Steps*”) foi utilizada para se obter o universo estatístico utilizado como base para a aplicação dos modelos de escolha modal do novo sistema de transportes entre Jundiaí e Campinas.

Os resultados do trabalho apresentaram correlação significativa com os dados observados, possibilitaram simulações de todas as fases de implantação do sistema sobre trilhos da RMC, além da obtenção das escolhas modais dos passageiros após a implantação da ligação Campinas – Jundiaí.

PALAVRAS-CHAVE

Análise multivariada, “*big-data*”, clusters, previsão de demanda, “*Smart Steps*”,

INTRODUÇÃO

O presente artigo tem o objetivo de apresentar a utilização de técnicas modernas de análise multivariadas de clusters e de “*big data*” utilizadas em previsões de demanda aplicadas ao planejamento de transportes de uma rede de trilhos para a RMC e um serviço da ligação Jundiaí - Campinas.

Análise de clusters e modelos e atração de viagens

Fazer previsões sobre as formas possíveis de deslocamentos de pessoas no futuro é um exercício complexo, ainda mais para a Região Metropolitana de Campinas, onde parte do problema situa-se na forma atual de distribuição espacial das atividades econômicas.

Segundo Ortuzar; Willumsen (2011), para compreender a demanda de transportes, é necessário compreender como as atividades são distribuídas no espaço.

As soluções tradicionais de modelagem de atração de viagens são apoiadas em técnicas de regressões lineares múltiplas que procuram relacionar dados socioeconômicos em cada zona de tráfego, com viagens observadas em pesquisas origem – destino.

No entanto, a influência da distribuição espacial das atividades econômicas na demanda de transportes tem sido um desafio para o planejamento de transportes, principalmente quando

os dados são considerados de forma agregada e sem distinção sobre as características econômicas de cada localidade.

Nesse trabalho são apresentadas técnicas de análise de cluster que possibilitam distinguir grupos de cidades segundo características econômicas e desenvolver modelos específicos de atração de viagens para cada agrupamento.

“Big data” e modelos de divisão modal

Outro desafio relacionado com a previsão de demanda de transporte no contexto dos municípios de Campinas e Jundiaí, foi estimar as decisões modais das pessoas diante da oferta de uma nova tecnologia alternativa aos modos existentes de deslocamento entre as duas localidades.

Essa estimativa foi realizada com o desenvolvimento de modelos de escolha discreta que contemplam as decisões dos usuários frente a alternativas de tempo, custo, acessibilidade e conforto.

A metodologia para o desenvolvimento desses modelos contempla às seguintes etapas:

1. Desenvolvimento de pesquisa de preferência declarada considerando a possibilidade de atração de viagens dos principais modos utilizados entre Jundiaí e Campinas (automóvel, ônibus rodoviário e ônibus fretado).
2. Obtenção de equações de probabilidade segundo modelos de escolha discreta, que possibilitaram compreender a estrutura de decisão dos usuários do novo sistema a partir dos efeitos sobre o comportamento desses usuários perante modificações nesse sistema, tais como tempo de viagem, custo, conforto.
3. Obtenção de curvas futuras de proporção entre modos de transportes, inclusive o modo proposto;
4. Obtenção do universo estatístico por meio de ferramentas computacionais de “big data”;
5. Aplicação das probabilidades no universo estatístico para obtenção dos valores absolutos de demanda e receita.

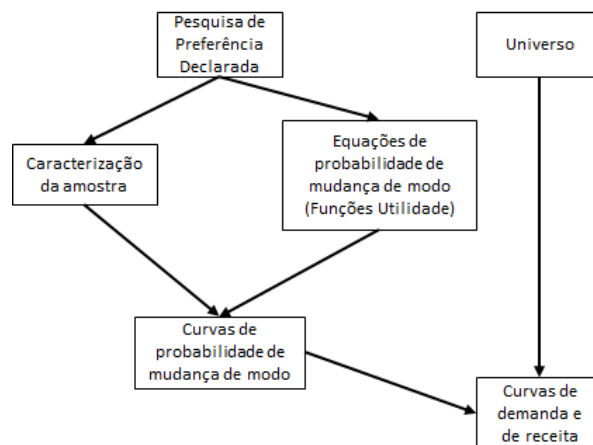


Figura 1 Metodologia de obtenção das curvas de demanda e receita

DIAGNÓSTICO, PROPOSIÇÕES E RESULTADOS

Análise de clusters e modelos e atração de viagens

A RMC é composta por 20 municípios (Americana, Artur Nogueira, Campinas, Cosmópolis, Engenheiro Coelho, Holambra, Hortolândia, Indaiatuba, Itatiba, Jaguariúna, Monte Mor, Morungaba, Nova Odessa, Paulínia, Pedreira, Santa Bárbara d' Oeste, Santo Antônio de Posse, Sumaré, Valinhos, Vinhedo) com um total de aproximadamente 2,8 milhões de habitantes.

Tendo como diferencial com relação as demais regiões metropolitanas brasileiras o fato do processo de industrialização e urbanização da RMC não se restringir ao município-sede, o município de Campinas consolidou no seu entorno importante aglomeração urbana. Além do polo regional, os demais municípios foram também capazes de estabelecer uma base econômica expressiva e dinâmica, tanto industrial, quanto agrícola, configurando uma estrutura diferenciada daquela das demais regiões metropolitanas, onde tradicionalmente é clara a presença de um município forte economicamente, cercado de municípios dormitórios.

A partir dessa consideração foi desenvolvida uma metodologia que procura aprimorar os modelos de atração de viagens com a consideração das características econômicas de cada cidade da RMC.

Foi desenvolvido um método apresentado em D'Andréa; Francoso (2015) composto inicialmente por uma análise de clusters para obter grupos de cidades com características econômicas homogêneas. Posteriormente foram desenvolvidos modelos de atração de viagens para cada agrupamento de cidade e finalmente foram comparados os modelos baseados em cluster com os modelos tradicionais, possibilitando concluir quanto a eficiência do método.

De acordo com Bailey (2014), a técnica de análise de cluster objetiva dividir um conjunto de objetos em um pequeno número de grupos relativamente homogêneos com base em similaridades de um conjunto de variáveis. O algoritmo de agrupamento usa as similaridades ou distâncias entre os objetos ao formar os clusters. Similaridades são um conjunto de regras que servem como critérios para agrupar ou separar itens. Neste caso, as semelhanças têm representado por setores da economia de todas as cidades da RMC. Onde: Setor1: Agricultura, Pecuária, Produção, Florestal, Produção, Pesca, Aquicultura; Setor 2: Indústria; Setor 3: Indústria da Construção; Setor 4: Comércio

Tabela 1 – Quantidade de empregos por setor econômico

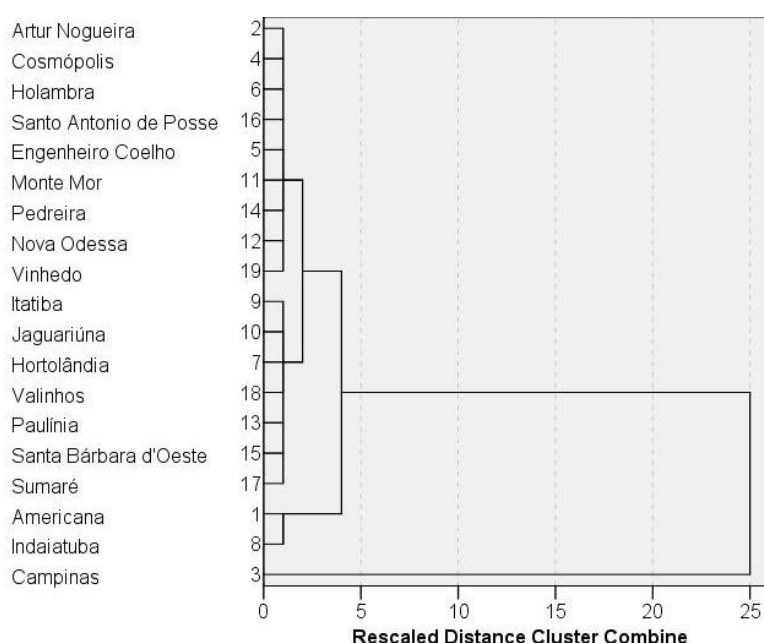
Cidades	Setor 1	Setor 2	Setor 3	Setor 4
Americana	67	28.186	3.376	49.630
Artur Nogueira	815	3.297	193	5.039
Campinas	1.693	61.630	20.758	322.454
Cosmópolis	1.187	1.964	413	6.427
Engenheiro Coelho	496	1.450	31	1.853
Holambra	3.216	1.023	287	2.687
Hortolândia	27	17.318	1.762	25.348
Indaiatuba	613	25.956	3.751	37.061
Itatiba	646	15.051	1.391	19.095
Jaguariúna	503	11.314	570	18.832
Monte Mor	539	5.161	988	4.801
Nova Odessa	130	11.307	869	6.278
Paulínia	226	11.605	5.043	27.558
Pedreira	94	7.080	30	5.924
Santa Bárbara	205	21.389	888	22.870
Santo A.de Posse	1.156	1.903	64	3.904
Sumaré	489	18.645	3.020	29.192
Valinhos	330	14.676	1.536	26.771
Vinhedo	121	6.588	172	13.967

A tabela a seguir apresenta os valores absolutos de empregos por setor de atividade em cada cidade que foram utilizados como base para a análise de cluster.

De acordo com Mooij; Sarstedt (2011), o algoritmo de agrupamento inicia-se com cada objeto representando um conjunto com apenas um elemento. Estes conjuntos são então unidos sequencialmente de acordo com a sua similaridade. Em primeiro lugar, os dois agrupamentos mais semelhantes são unidos para formar um novo agrupamento na parte inferior da hierarquia. No passo seguinte, um outro par de aglomerados é unido e ligado a um nível superior da hierarquia, e assim por diante.

A representação final dos cluster foi representada no dendrograma da Figura 2. Brian (2011) define que o dendrograma, ou diagrama de árvore, é uma representação matemática e pictórica do procedimento de agrupamento completo. Os nós do dendrograma representam clusters, e os comprimentos das hastes (alturas) representam as distâncias em que os clusters são unidos.

Figura 2- Dendrograma



A partir do dendrograma foram elaborados os seguintes agrupamentos:

Cluster 1: Campinas;

Cluster 2: Americana e Indaiatuba;

Cluster 3: Hortolândia, Itatiba, Jaguariúna Paulínia, Santa Barbara, Sumaré, Valinhos;

Cluster 4: Arthur Nogueira, Cosmópolis, Engenheiro Coelho, Holambra, Monte Mor, Nova Odessa, Pedreira, Santo Antônio da Posse e Vinhedo.

Para cada agrupamento foi desenvolvido um modelo de atração de viagens relacionando dados de viagens com dados socioeconômicos obtidos da **Pesquisa Origem-Destino 2011 Região Metropolitana de Campinas - Síntese dos Resultados Pesquisas Domiciliar e Cordon Line** (2012) conforme a seguir:

Tabela 2 – Modelos de atração de viagens baseados em clusters

CLUSTER 1	$V C1_{AM} = -386 + 0,17 \times E_{AM}$	$V C1_B = -475 + 0,16 \times E_{AM}$
CLUSTER 2	$V C2_{AM} = -570 + 0,14 \times E_{AM}$	$V C2_B = -738 + 0,12 \times E_{AM}$
CLUSTER 3	$V C3_{AM} = -276 + 0,11 \times E_{AM}$	$V C3_B = -171 + 0,09 \times E_{AM}$
CLUSTER 4	$V C4_{AM} = -172 + 0,10 \times E_{AM}$	$V C4_B = -203 + 0,09 \times E_{AM}$

Tabela 3 - Modelos de atração de viagens não baseados em clusters

$V_{AM} = -454 + 0,154 \times E_{AM}$	$V_B = -456 + 0,138 \times E_B$
---------------------------------------	---------------------------------

Onde:

V = Viagens atraídas

C1, C2, C3 e C4, representam os clusters 1 a 4;

AM = classe de renda alta e média;

B = classe de renda baixa;

E= empregos formais.

Por fim foram comparados os modelos baseados em cluster com os modelos não baseados em cluster que que contemplam todas as 19 cidades da RMC sem distinção de características econômicas.

Segundo D'Andréa; Francoso (2015), os modelos de atração de viagens estimados segundo agrupamentos de cidades semelhantes produziram menores erros quadrados do que os resultados dos modelos não baseados em clusters.

Observa-se ainda que as comparações gráficas (Figura 3, Figura 4, Figura 5, Figura 6) entre dados modelados e observados apresentam menores dispersões nos modelos baseados em cluster do que nos modelos sem clusters.

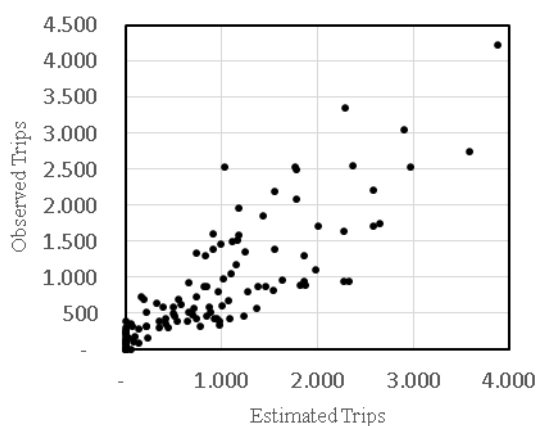


Figura 3 – Modelos sem cluster – Classe média e alta

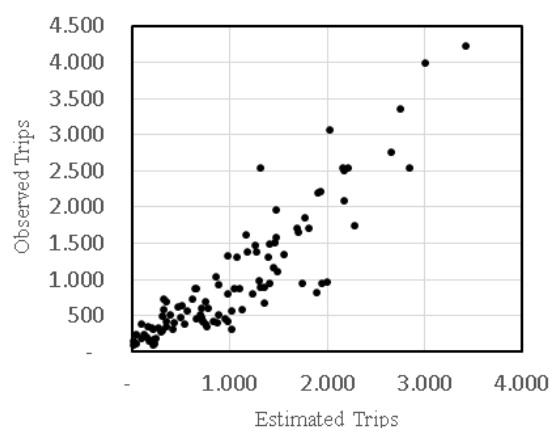


Figura 4 – Modelos com cluster classe média e alta

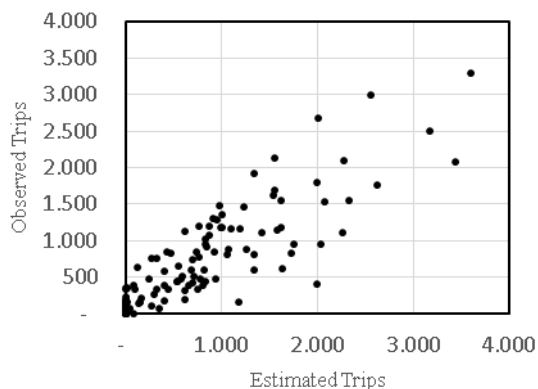


Figura 5 - Modelos sem, cluster – classe baixa

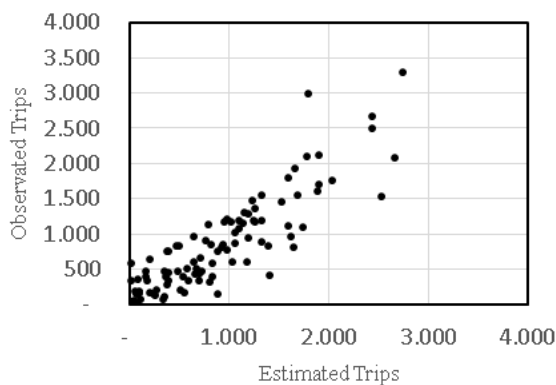


Figura 6 – Modelos com cluster – classe baixa

“Big Data” e modelos de divisão modal

A obtenção tradicional de dados do universo é baseada em pesquisas de campo extensivas, específicas para esse fim (que possuem alcance limitado), e em dados de fontes secundárias – processos estes que em geral elevam consideravelmente o custo e o tempo de elaboração dos projetos.

Na região entre Jundiaí e Campinas esse quadro se acentua por não existirem dados disponíveis atualizados dos deslocamentos de cada modo de transporte entre essas duas cidades e também das cidades intermediárias.

A alternativa encontrada para a obtenção desses dados foi apoiada em uma ferramenta de “big data” denominada “Smart Steps” desenvolvida pela empresa Telefonica.

Essa ferramenta computacional captura dados anônimos e agregados de aparelhos celulares possibilitando estimar as origens e destinos das viagens em um determinado período de tempo. O “Smart Steps” é baseado em obtenção e análise de dados massivos - “big data” e resulta em totais de fluxos de pessoas através de áreas geográficas em qualquer período de tempo e em qualquer nível de escala, tais como, ruas, cidades, regiões metropolitanas, estados e até nacionalmente.

O “Smart Steps” foi utilizado ininterruptamente ao longo de um mês entre Campinas e Jundiaí e resultou em um banco de dados que possibilitou estimar viagens a partir da identificação dos locais de moradia e de trabalho/estudo com base em um histórico das ligações telefônicas realizadas em horários comerciais e não comerciais.

Esta estimativa é feita pela quantidade e pelo local de ligações realizadas por cada aparelho nos horários comerciais e não comerciais. O sistema capta, então, pessoas que fazem viagens com motivo CTE (casa, trabalho, estudo).

Vale notar que o “Smart Steps” quantifica o número de pessoas que realizam viagens com os seguintes motivos de origem e destino:

Tabela 4 – Motivos na origem e motivos no destino

Motivo Origem	Motivo Destino
Casa	Trabalho
Casa	Estudo
Trabalho	Casa
Estudo	Casa

Portanto, o “Smart Steps” fornece respostas às seguintes perguntas:

- Quantos clientes Vivo residem em Campinas?
- Quantos clientes Vivo residem em Campinas e trabalham ou estudam em Jundiaí?
- Quantos clientes Vivo residem em Jundiaí?
- Quantos clientes Vivo residem em Jundiaí e trabalham ou estudam em Campinas?

Conhecendo quantas pessoas residem em uma cidade e quantas delas realizam atividades na outra, foi possível obter a proporção da população que realiza os deslocamentos que o presente estudo deseja compreender.

Uma das hipóteses adotadas é de que o comportamento de viagens entre Jundiaí e Campinas da base de clientes da Vivo é representativo do comportamento de viagens de toda a região pesquisada.

O cálculo do total de pessoas que realizam viagens tipo CTE entre Jundiaí e Campinas foi realizado conforme a Equação 1:

Equação 1 - Total de pessoas que realizam viagens CTE

$$SS = \frac{C_{J \rightarrow C}}{C_J} * P_J + \frac{C_{C \rightarrow J}}{C_C} * P_C$$

Onde:

SS: Total de pessoas que realizam viagens tipo CTE entre Jundiaí e Campinas;

CJ→C: Quantidade de clientes Vivo que moram em Jundiaí e trabalham ou estudam em Campinas;

CJ: Quantidade de clientes Vivo que moram em Jundiaí;

PJ: População de Jundiaí de acordo com o CENSO 2010;

CC→J: Quantidade de clientes Vivo que moram em Campinas e trabalham ou estudam em Jundiaí;

CC: Quantidade de clientes Vivo que moram em Campinas;

PC: População de Campinas de acordo com o CENSO 2010.

Na etapa de determinação do tamanho do universo, estimou-se cerca de 34.096 viagens diárias entre Jundiaí e Campinas e que esta demanda se divide conforme resultados dos modelos de escolha discreta:

- 28.173 viagens diárias por automóvel (83%);
- 985 viagens diárias por ônibus rodoviário (3%);
- 4.937 viagens diárias por ônibus fretado (14%).

Como resultado final do trabalho apresenta-se os valores absolutos de demanda estimada para a alternativa de transportes proposta para os deslocamentos entre Campinas e Jundiaí em função de hipóteses tarifárias variando de R\$10 em R\$10.

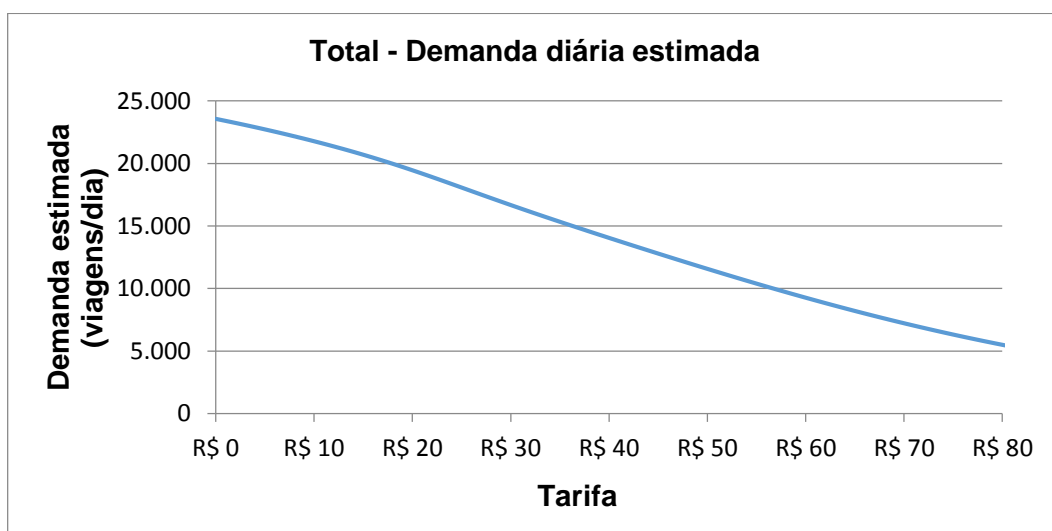


Figura 7 - Demanda diária estimada para o novo serviço de transportes vs tarifa do novo modo – Viagens realizadas por ônibus rodoviário e por automóvel

CONCLUSÕES

Os métodos de obtenção de dados a partir de “*Smart Steps*” assim como os métodos de tratamento de dados de análise multivariada apresentaram fácil aplicação no estudo de caso desse trabalho e resultaram em modelos de previsão de demanda aderentes com os dados observados, além de uma estimativa de universo de demanda entre Campinas e Jundiaí sem a necessidade de se fazer uma pesquisa origem- destino específica.

O objetivo de incorporar elementos relacionados às particularidades econômicas da RMC e distinguir 4 grupos homogêneos de cidades segundo setores econômicos foi atingido com êxito e os resultados satisfatórios desse método de trabalho sugerem a ampliação das possibilidades de aplicação de análises multivariadas em modelos de demanda e constituem um campo aberto de aplicação da matemática moderna na engenharia de transportes.

A utilização dessa alternativa metodológica resultou ainda em economia de tempo durante o desenvolvimento do trabalho, porque os modelos se ajustaram rápida e adequadamente aos dados observados.

A obtenção de dados via “*Smat Steps*” revelou-se um método eficaz na complementação dos dados necessários para aplicar os modelos de escolha discreta. A realização de um mês de coleta de sinais de telefonia celular entre os municípios de Campinas e Jundiaí resultaram em uma quantidade massiva de dados, muitas vezes superior a quantidade de dados obtida por uma pesquisa origem- destino, com a vantagem de apresentar validade estatística na obtenção de estratos da amostra, possibilitando sua utilização de forma adequada como base para aplicação dos modelos de escolha discreta.

REFERÊNCIAS BIBLIOGRÁFICAS

BAILEY, D. Cluster Analysis. , v. 6, n. 1975, p. 59–128, 2014.

BRIAN, S. E. **Cluster Analysis**. 5th ed. London: John Wiley and Sons, Ltd., 2011.

D’ANDRÉA, A. F.; FRANCO SO, M. T. Application of cluster analysis in trip attraction models of Campinas Metropolitan Region. **IACSIT International Journal of Engineering and Technology**, Vol. 7, No. 5, October 2015, p. 5, 2015.

MOOI, E.; SARSTEDT, M. **A Concise Guide to Market Research The Process, Data, and Methods Using IBM SPSS Statistics**. Springer Berlin Heidelberg, 2011.

ORTUZAR, J. DE D.; WILLUMSEN, L. G. **Modelling Transport**. 4th ed. John Wiley and Sons, Ltd., 2011.

Pesquisa Origem-Destino 2011 Região Metropolitana de Campinas - Síntese dos Resultados Pesquisas Domiciliar e Cordon Line. Secretaria dos Transportes Metropolitanos - STM, 2012.