

Mineração de dados no apoio a gestão pública municipal: conhecendo as demandas da cidade de Curitiba pela “Central 156”

RESUMO

Ao analisar a cidade de Curitiba pela “Central 156”, buscou-se identificar padrões na participação da população em reivindicações como: ‘consulta pública’, ‘denúncia’, ‘elogio’, ‘informação’, ‘recadastro’, ‘reclamação’, ‘solicitação’ e ‘sugestão’. Para isso, procedeu-se uma mineração de dados na base da “Central 156”, no ano de 2018, por meio dos aplicativos *Microsoft-Excel* e *Weka*. Primeiramente, coletaram-se dados disponíveis no site da Prefeitura Municipal de Curitiba e uniram-se, mês a mês, em uma única base de dados. A partir da seleção de variáveis pelo método *BestFirst* foram executados algoritmos de associação (*APRIORI*) e classificação (*PART*), bem como análise estatística. As análises revelaram regras como: sempre que o assunto é referente ao transporte público, há a requisição de reclamação, sugerindo um problema neste tipo de serviço; existe uma relação entre o total de demandas por assunto iluminação pública e o total de vítimas por homicídio doloso. Pesquisas futuras podem ainda aumentar o escopo da pesquisa, por meio de mineração de texto no campo de descrição da solicitação.

PALAVRAS-CHAVE: Algoritmos de associação (*APRIORI*) e classificação (*PART*). Central 156. Cidade de Curitiba.

Marcelo Maia
maiamarcelomaia@hotmail.com
Doutorando em Gestão da
Informação - UFPR

Marcos Maia
maia.marcosmaia@gmail.com
Doutorando em Gestão da
Informação – UFPR

Denise Fukumi Tsunoda
dtsunoda@gmail.com
Professora do departamento de
Ciência e Gestão da Informação -
UFPR

INTRODUÇÃO

Dados do Instituto Brasileiro de Geografia e Estatística (IBGE) apontam um crescimento da população residente na cidade de Curitiba de 9,43%, passando de 1.751.907 em 2010 para 1.917.185 em 2018. Este aumento reforça “a crescente complexidade da cidade [...] criando desafios para a gestão urbana e a busca por novas soluções, como o uso de tecnologias que ampliam o entendimento de desafios locais até então invisíveis” (HOJDA, FARINIUK, SIMÃO, 2019, p. 80, tradução nossa). Assim, com o recurso das Tecnologias de Informação (TI), inicia-se a busca por ferramentas de gestão capazes de melhorar e dar maior eficiência administrativa, “fomentado pelo avanço tecnológico e seus impactos sociais, políticos e culturais” (BATISTA *et al*, 2019, p.113)

A “Central 156” de Curitiba, de acordo com o gestor em atuação, na reportagem a Central de Brasileira de Notícias – CBN Curitiba, Oliveira (2019) destaca a capital do Paraná como a pioneira em atendimento centralizado ao cidadão, sendo a primeira capital a ter o serviço de atendimento ao cidadão. As solicitações eram tratadas manualmente, sendo que demorava em torno de três dias para gerar a solicitação. Essas solicitações eram realizadas por telefone, anotadas e enviadas aos órgãos responsáveis através de malotes, para então pudesse se iniciar o processo de realização. Por fim, as respostas eram enviadas por meio de cartas ao cidadão demandante.

Neste contexto, no ano de 2019, a “Central 156” completa 35 anos de atendimento ao Cidadão. Sua “evolução tecnológica seguiu com o lançamento do portal on-line, gravação de chamadas, utilização de telefonia IP, e a transformação da Central (CENTRAL 156, 2019). Desse modo, desde novembro de 2015, a Prefeitura disponibiliza os dados contido nas solicitações (PREFEITURA DE CURITIBA, 2019). Embora a abertura da base tenha ocorrido em 2015, o portal apresenta para *download* só os dados dos últimos 24 meses.

A partir dos dados levantados e apoiados na análise dos mesmos, este estudo tem por objetivo analisar a cidade de Curitiba pela “Central 156”, identificando padrões e associações relacionadas a participação da população no encaminhamento das reivindicações. Estas ações poderão ser utilizadas posteriormente para a tomada de decisão da administração pública.

Este estudo se justifica pelo fato de ser “uma ferramenta de gestão [...] que ajuda a identificar e sanar problemas, verificar tendências e definir rotinas de trabalho” (GONÇALVES, 2019), conforme relato do secretário do governo, sendo “um termômetro do que está acontecendo na cidade” (KOWALSKI, 2019).

Portanto, o presente estudo está dividido em introdução, conceito da gestão pública municipal, histórico da Central 156 na cidade de Curitiba, métodos e materiais a ser utilizada, apresentação dos resultados obtidos, e por fim, discussão e considerações finais.

GESTÃO PÚBLICA

A Constituição da República Federativa do Brasil (1988) dá significativa importância aos municípios na execução e controle de bens e serviços públicos. Conforme determina a Lei de Responsabilidade Fiscal (LRF) que objetiva controlar os gastos dos gestores públicos, tanto os estados quanto os municípios devem

primar pela transparência nos recursos públicos. Assim, Guimarães, Oliveira (2019, p 357) descreve que “a administração pública [...] tratada como gestão pública, compreende todo o universo de instituições, órgãos, serviços e agentes públicos que trabalham buscando suprir as necessidades da população”, sendo uma gestão dos interesses da coletividade para a prestação dos serviços públicos.

Nesse contexto, a participação do cidadão auxilia nos “[...] inúmeros desafios que os gestores públicos enfrentam e se deparam no universo da administração pública para realizar um trabalho correto, eficiente, eficaz e efetivo” (GUIMARÃES, OLIVEIRA, 2019, p 368), sendo a ferramenta de gestão, “a fim de melhorar os serviços prestados, identificar os anseios da população e a necessidades prioritárias dos cidadãos” (OLIVEIRA, OLIVEIRA, 2019, p 291).

Assim a gestão municipal trata “aos aspectos da gestão da estrutura administrativa da prefeitura, por meio de recursos e instrumentos próprios da administração local” (KANUFRE, REZENDE, 2012, p 640), afim de alcançar resultados para atendimento das demandas, criando valor público.

A prefeitura de Curitiba utiliza uma gestão democrática, que estabelece a relação entre a administração pública e a população, constituída de uma base democrática participativa. Essa gestão utiliza diversos serviços de App (Cidades das Ideias), que facilita a vida do cidadão e reúne informações de mais de 600 serviços da prefeitura, notícias e eventos da cidade em um só lugar, podendo agendar serviços da prefeitura pelo celular “Saúde já”, “app 156”, “nota Curitiba”, “fala Curitiba”.

CENTRAL 156

A central de atendimento 156 teve seu início em 08 de fevereiro de 1984, na gestão do prefeito Maurício Fruet. De acordo com o vereador Omar Sabbag Filho, o objetivo era “ampliar a participação da população na definição dos destinos da cidade e no encaminhamento das reivindicações” (CÂMARA MUNICIPAL DE CURITIBA, 19-?). Neste canal direto entre a população e o poder público era possível: solicitar e reclamar de serviços, pedir informações e também algumas certidões negativas. O intuito era “reduzir filas no balcão de atendimento da Prefeitura [...] que funcionava de segunda a sexta-feira, das 7h às 19h, e após esse horário a secretária eletrônica gravava os pedidos” (CENTRAL 156, 2019).

A linha telefônica utilizada pelo serviço recebia pedidos e atendia emergências, sendo os principais atendimentos “o desentupimento de um bueiro ou a poda de uma árvore, a pintura de uma faixa de pedestre ou a interdição do local até que a equipe especializada viesse resolver” (CÂMARA MUNICIPAL DE CURITIBA, 19-?). Tudo isso era resolvido por uma equipe de cinco colaboradores que se dirigiam ao local de atendimento com o veículo Kombi, apelidado de ‘Trovão Azul’. Com o passar dos anos, o serviço ganhou grandiosa importância e se transformou no “principal canal direto de comunicação entre a população e a prefeitura de Curitiba” (CÂMARA MUNICIPAL DE CURITIBA, 19-?).

Até o ano de 2002, a central 156 não realizava atendimento 24 horas por dia, no entanto, o projeto que passava por uma modernização desde 1999 passou a disponibilizar o serviço ininterruptamente a partir de setembro de 2002. Além das mudanças do período de disponibilização do serviço, também ocorreu à integração de todas as secretarias e órgãos municipais. Nesse contexto, há “transmissão em

tempo real de todas as solicitações, informações, elogios, reclamações e sugestões da população diretamente aos responsáveis pelo parecer/ fiscalização/ execução” (CENTRAL 156, 2019).

Desde 2016 é possível realizar o envio de imagens à Central. Em março de 2019 (no aniversário de Curitiba – 326 anos da cidade) foi apresentado o aplicativo ‘Curitiba 156’, serviço disponível para *smartphones* com sistemas Android e IOS. O *app* permite comunicação direta com o órgão da prefeitura, integração ao sistema 156, inclusão de fotos, localização por georreferenciamento, gestão do cartão transporte, posicionamento da rota de ônibus urbano em tempo real e pesquisa de satisfação. O aplicativo inicialmente vai funcionar com apenas 12 dos 270 assuntos disponíveis na Central 156.

Atualmente, a central é gerenciada pelo Instituto das Cidades Inteligentes (ICI) com o objetivo de “viabilizar um sistema de comunicação ágil e eficiente entre o cidadão e a Prefeitura de Curitiba, permitindo o atendimento da demanda de informações e solicitações da população com segurança, confiabilidade e, principalmente qualidade” (CENTRAL 156, 2019).

Assim sendo, a população pode entrar em contato com a “Central 156” pelo telefone, *chat* de atendimento *on line* e aplicativo. O cidadão tem acesso aos serviços: abordagem social de rua, acessibilidade, animais, arruamento, árvore, coleta, dengue, drenagem, edificações, escolas, fiscalização, iluminação pública, informação, IPTU, ITBI, limpeza, lombada física, página da prefeitura, parques, pavimentação, posto de saúde, praças, rua da cidadania, semáforo, sinalização, trânsito, transporte coletivo e outros. Maiores informações destes serviços devem ser obtidas por meio da página <http://www.central156.org.br/servicos>.

O aplicativo dispensa o processo de triagem, encaminhando as demandas diretamente ao órgão responsável para a prestação do serviço. Todavia, quando o atendimento é realizado via *chat* ou telefone, a triagem se faz necessária. De acordo com o gestor da Central 156, a utilização da nova ferramenta “[...] consegue registrar uma ocorrência em 30 ou 40 segundos. Por telefone, esse tempo varia de 1 a 3 minutos dependendo da quantidade de informações que precisam ser coletadas” (RAMOS, 2019).

MÉTODOS E MATERIAIS UTILIZADOS

Nesta seção serão descritos os procedimentos utilizados para obtenção, pré-processamento, mineração, análise e validação do conhecimento na base de dados referente ao ano de 2018 da “Central 156” da prefeitura Municipal de Curitiba, Paraná.

ANÁLISE DOS DADOS

A análise dos dados deste estudo será realizada por mineração e estatística descritiva. Desta forma, optou-se por fazer uma adaptação ao processo de descoberta e conhecimento em bases de dados propostos por Fayyad, Piatetsky-Shapiro e Smyth.

Mineração de dados é parte no processo de descoberta de conhecimento em bases de dados (*Knowledge Discovery in Databases*, ou KDD), conforme Castro e

Ferrari (2016), na qual se utiliza de aplicação de algoritmos específicos para extrair padrões de dados (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996).

Este processo de descoberta e conhecimento em bases de dados envolve algumas etapas (FAYYAD, PIATETSKY-SHAPIRO e SMYTH, 1996), (ADRIAANS, e ZANTINGE, 2011), as quais são sintetizadas por Castro e Ferrari (2016) e apresentadas no quadro 1.

Quadro 1 - Processo de Descoberta de Conhecimento em Bases de Dados

Etapas	Identificação	O que será realizado
1ª Etapa	Base de dados	Conjunto de dados organizados
2ª Etapa	Preparação ou pré-processamento de dados	Etapa destinada à preparação dos dados. Inclui a limpeza (retirada de dados inconsistentes), integração (combinação de dados de várias bases), seleção ou redução (seleção de dados pertinentes a pesquisa) e a transformação dos dados (conversão dos dados de modo que possam ser utilizados na mineração).
3ª Etapa	Mineração de dados	Aplicação de algoritmos para obter-se conhecimento na base de dados. Esses algoritmos podem ser utilizados para classificação, associação e agrupamento.
4ª Etapa	Análise estatística	Estatística descritiva e correlações, como apoio ao resultado obtido junto a 3ª etapa.
5ª Etapa	Avaliação ou validação de conhecimento	Avaliação dos resultados identificando conhecimentos

Fonte: Adaptado de CASTRO; FERRARI (2016).

A mineração de dados, terceira etapa do processo de conhecimento em bases de dados, possui três principais tarefas: classificação, agrupamento e associação. O quadro 2 apresenta estas três tarefas, listando alguns de seus algoritmos.

Quadro 2 – Tarefas de mineração de dados

Tarefa	Tipo de Algoritmo	Algoritmos
Classificação	Bayes	NaiveBayes
		Bayesnet
	Rules	Part
		Decision Tree
		Random Forest
Agrupamentos	Por Densidade	DBSCAN
	Baseado em Protótipo	k-means
		k-medoids
Regras de Associação		Apriori
		FP Growth

Fonte: AMARAL (2016, p. 99).

A quarta etapa, análise estatística, foi introduzida pelos autores, uma vez que não faz parte do processo de descoberta e conhecimento em bases de dados estabelecido por FAYYAD, PIATETSKY-SHAPIRO e SMYTH. No entanto, esta etapa é necessária para se aprimorar o conhecimento da base de dados, compreendendo

o comportamento das variáveis estudadas e auxiliando a validação dos resultados obtidos na mineração.

Utilizou-se nesta pesquisa o aplicativo *Microsoft Excel* na criação, preparação e análise da base de dados, etapas 1, 2 e 4 respectivamente. O aplicativo *Weka* foi utilizado para pré-processamento da base de dados e mineração, etapas 2 e 3.

RESULTADOS E DISCUSSÕES

Nesta seção busca-se demonstrar os principais resultados obtidos por esta pesquisa. Os resultados são apresentados conforme estabelecidos no protocolo de pesquisa descritos no quadro 1.

1ª Etapa: Obtenção da base de dados

Para a coleta de dados, foi utilizado o Portal de Dados Abertos da Prefeitura Municipal de Curitiba – PMC, que [...] disponibiliza através da internet, documentos, informações e dados governamentais de domínio público para a utilização [...] dos dados primários, de forma a ser reutilizado, produzindo informações (CENTRAL 156, 2019). A base de dados referente à “Central 156” apresenta atualização mensal e os dados disponíveis são desde novembro de 2015, no entanto, o estudo compreende apenas o ano de 2018 (janeiro a dezembro), visto que nos anos de 2016 e 2017 há a ausência de alguns meses: janeiro a setembro em 2016 e maio em 2017.

Os dados foram coletados mês a mês (12 arquivos) e reunidos em uma planilha *Microsoft-Excel*. Esta base de dados apresenta as variáveis apresentadas no quadro 3.

Quadro 3 – Variáveis encontradas na base de dados preliminar

Variável	Natureza	O que se pretende medir
Solicitação	Quantitativa discreta. Número formado por 7 algarismos. Exemplo: 7596506	Identificar o número da solicitação efetuada na Central de Atendimento 156.
Tipo	Qualitativa nominal, podendo ser: Consulta Pública, Denúncia, Elogio, Informação, Recadastro, Reclamação, Solicitação e Sugestão	Tipo de demanda solicitada.
Órgão	Qualitativa nominal.	Órgão em que cadastrador está lotado.
Data	Data. Apresentada no formato dia/mês/ano	Data da solicitação.
Horário	Hora. Apresentado no formato hora:minuto:segundo.	Hora em que a demanda foi registrada.
Assunto	Qualitativa nominal. Mais de 200 assuntos preestabelecidos pelo sistema, conforme site: http://www.central156.org.br/servicos	Assuntos disponíveis pelo atendimento.
Subdivisão	Qualitativa nominal. Mais de 800 itens, conforme site da central 156.	Divisão do assunto.

Variável	Natureza	O que se pretende medir
Descrição	Qualitativa nominal. Texto.	Detalhamento da solicitação pelo demandante.
Logradouro_ASS	Qualitativa nominal. Ruas da cidade de Curitiba do assunto da solicitação.	Endereço da demanda.
Bairro_ASS	Qualitativa nominal. Bairros da capital paranaense do assunto da solicitação.	Bairro relacionado a solicitação.
Regional_ASS	Qualitativa nominal. Regionais da cidade de Curitiba do assunto da solicitação: Unidade Regional Bairro Novo, Unidade Regional Boa Vista, Unidade Regional Boqueirão, Unidade Regional Cajuru, Unidade Regional CIC, Unidade Regional Matriz, Unidade Regional Pinheirinho, Unidade Regional Portão, Unidade Regional Santa Felicidade, Unidade Regional Tatuquara.	Regional a qual o bairro está associado.
Meio resposta	Qualitativa nominal. O sistema aceita registros de: e-mail, nenhum, pessoalmente, telefone.	Meio escolhido para resposta.
Observação	Qualitativa nominal. Texto.	Registro falado pelo demandante. Variável não obrigatória.
Sexo	Qualitativa nominal dicotômica, assumindo o M para masculino ou F para feminino.	Gênero do demandante.
Bairro_cidadão	Qualitativa nominal. Apresenta 75 bairros da cidade de Curitiba.	Bairro onde o cidadão reside.
Regional cidadão	Qualitativa nominal. Apresenta 10 regionais da cidade de Curitiba	Regional vinculada ao bairro cidadão.
Data nascimento	Data. Apresentada no formato dia/mês/ano.	Ano em que o demandante nasceu.
Histórico	Qualitativa nominal. Texto.	Último histórico da solicitação (quando possui).
Tipo de cidadão	Qualitativa nominal.	Tipo do cidadão solicitante.
Órgão responsável	Qualitativa nominal. Secretaria Municipal responsável pelo serviço.	Qual órgão é responsável pela resolução da demanda.
Resposta final	Qualitativa nominal. Texto.	Descrição da resposta, quando inserida.

Fonte: Autores.

2ª Etapa: Preparação ou pré-processamento de dados

Com a finalidade de preparar a base de dados para a mineração, excluíram-se primeiramente os dados inseridos na base que não possuíam número de solicitação (31.592 linhas) pois não estavam corretamente preenchidos, apresentando apenas alguns caracteres de texto. Em seguida, foram excluídos os campos: órgão, descrição, observação, tipo cidadão e histórico.

As variáveis: descrição, observação e histórico foram retirados da base de dados, pois são dados textuais que não estão no escopo deste trabalho. O campo

tipo cidadão foi excluído da base visto que ou o campo estava em branco ou apresentava o texto “cidadão”. Por fim, a variável órgão responsável foi retirada, pois além de possuir muitos dados incompletos, apenas servia para relacionar o assunto à secretaria municipal responsável pelo seu atendimento.

Após a exclusão destas variáveis, foi realizada a transformação dos campos: data, horário e data de nascimento. Foram substituídos por: mês, turno e idade. Este procedimento foi necessário em virtude de que os algoritmos *PART* e o *J48*, por exemplo, só são aplicáveis a variáveis nominais.

A data foi alterada pelo mês de solicitação; o horário foi alterado pelos períodos – manhã (06:00:00 às 11:59:59), tarde (12:00:00 às 17:59:59), noite (18:00:00 às 23:59:59) e madrugada (00:00:00 às 05:59:59); e, a idade foi discretizada em: até 18 anos em “criança ou adolescente” pela classificação descrita Lei nº. 8069-90 – Estatuto da Criança e do Adolescente; 60 anos ou mais em “idoso” em decorrência da lei 10741/2003 – Estatuto do Idoso. Para classificar as idades intermediárias, dividiu-se a diferença entre 18 e 60 anos por dois, e classificado como “adulto I” e “adulto II”, conforme classificações descritas pelos autores Meinel (1984), Mosquera (1985) e Kaplan e Sadock (1999).

Por fim, a base de dados passou a apresentar 16 variáveis e um total de 320.057 observações.

Antes de iniciar a etapa de mineração de dados, utilizou-se o algoritmo de Hall (1998) pelo método *BestFirst* para a seleção das variáveis que possuem maior capacidade preditiva individual. Tal análise resultou como principais variáveis: assunto; subdivisão e órgão_resp. Desta forma, estes três atributos foram utilizados na 3ª Etapa, mineração dos dados.

3ª Etapa: Mineração de dados

Nesta etapa, utilizaram-se as tarefas de classificação (algoritmo de *PART*) por meio de validação cruzada de 10 partições e associação (algoritmo *APRIORI*) com confiança de 85% e suporte de 10%.

Com o intuito de identificar prováveis padrões quanto a classificação do tipo de demandas, aplicou-se o algoritmo de *PART* de validação cruzada. Os resultados obtidos estão apresentados nas tabelas 1,2 e no quadro 4. As tabelas 1e 2 apresentam de maneira resumida a eficiência da classificação realizada.

Tabela 1 – Resumo da partição de 10% pelo algoritmo *PART* com a variável solicitação na “Central 156”

Estatísticas	Valor	Porcentagem
Taxa de acerto	299.547	93,59%
Taxa de erro	20.510	6,41%
Estatística Kappa	0,7128	
Erro absoluto médio	0,251	
Erro quadrático médio	0,125	
Erro absoluto relativo	41,58%	
Erro quadrático relativo	64,71%	
Número total de itens	320.057	

Fonte: Autores.

Conforme se pode verificar, a classificação foi assertiva em 93,59% dos casos, apresentando uma coesão (estatística Kappa) de 0,7128, dando legitimidade as interpretações.

Com a matriz de confusão, apresentada na tabela 2, pode-se verificar onde houve mais acerto ou erro na classificação realizada. As variáveis: consulta pública, denúncia e recadastro tiveram erro aproximadamente 100%. Sugestão acertou apenas 9,78%. Solicitação, reclamação, informação e elogio foram as demandas que obtiveram a maior taxa de acerto, as quais foram responsáveis pela alta porcentagem de acerto da classificação.

Tabela 2 – Matriz de confusão pelo algoritmo PART na “Central 156”

a	b	c	d	e	f	g	h	Classificação	Variável
272360	4345	2	11	242	47	5	0	a	solicitação
5369	21239	2	2	2720	4	2	0	b	reclamação
4573	117	4	0	19	0	0	0	c	recadastro
5	6	0	1411	3	0	0	0	d	informação
847	701	0	7	4474	1	0	0	e	elogio
383	204	0	0	16	59	0	0	f	sugestão
730	126	0	0	15	0	0	0	g	denúncia
4	1	0	0	1	0	0	0	h	consulta pública

Fonte: Autores.

As melhores regras encontradas pela classificação foram aquelas em que houve maior porcentagem de acerto, sendo apresentadas no quadro 4. A regra 1 demonstra que das 1271 subdivisão agendamento de consulta básica, 136 referem-se à reclamação.

Quadro 4 – Melhores regras encontradas pelo algoritmo APRIORI na “Central 156”

SUBDIVISÃO	ASSUNTO
AGENDAMENTO DE CONSULTA BÁSICA	RECLAMAÇÃO (1271.0/136.0)
ATRASAR O HORÁRIO DURANTE A OPERAÇÃO	RECLAMAÇÃO (1023.0/133.0)
RECURSOS HUMANOS	QUANTIDADE DE PROFISSIONAIS: RECLAMAÇÃO (878.0/133.0)
DEMORA NO ATENDIMENTO	RECLAMAÇÃO (685.0/39.0)
DEMORA NO ATENDIMENTO PARA CONSULTA MÉDICA/OUTROS PROFISSIONAIS FEAES	RECLAMAÇÃO (588.0/4.0)
DEMORA NO ATENDIMENTO PARA CONSULTA MÉDICA/OUTROS PROFISSIONAIS FEAES/OS	RECLAMAÇÃO (184.0)
CONSULTA BÁSICA	INFORMAÇÃO (22.0)
MEDICAMENTOS	INFORMAÇÃO (11.0)
VACINA	INFORMAÇÃO (10.0)
HORÁRIO DE ÔNIBUS	SOLICITAÇÃO (4.0/1.0)
CADASTRO DE UNIDADE DE SAÚDE	INFORMAÇÃO (4.0)
ASSUNTO = DISQUE SOLIDARIEDADE	DOAÇÕES: SOLICITAÇÃO (6636.0/24.0)

SUBDIVISÃO	ASSUNTO
ASSUNTO = TRÂNSITO	FISCALIZAÇÃO DE ESTACIONAMENTO IRREGULAR: SOLICITAÇÃO (17266.0/61.0)
ASSUNTO = ABORDAGEM SOCIAL DE RUA	ADULTO: SOLICITAÇÃO (17275.0/274.0)
ASSUNTO = ABORDAGEM SOCIAL DE RUA	SOLICITAÇÃO (7229.0/109.0)
ASSUNTO = TRÂNSITO	FISCALIZAÇÃO: SOLICITAÇÃO (5933.0/43.0)
ASSUNTO = ANIMAIS DOMÉSTICOS	SOLICITAÇÃO (5814.0/71.0)
ASSUNTO = FAUNA SINANTRÓPICA	SOLICITAÇÃO (4766.0/68.0)
ASSUNTO = TRÂNSITO	VEÍCULO ESTACIONADO - GUIA REBAIXADA: SOLICITAÇÃO (3684.0/29.0)
ASSUNTO = SEMÁFORO	SOLICITAÇÃO (4866.0/99.0)
ASSUNTO = TRÂNSITO	VEÍCULO ESTACIONADO EM HORÁRIO/LOCAL PROIBIDO: SOLICITAÇÃO (3678.0/21.0)
ASSUNTO = TRÂNSITO	SOLICITA FISCALIZAÇÃO PROGRAMADA: SOLICITAÇÃO (3494.0/70.0)

Fonte: Autores.

Verifica-se que o tipo reclamação está associado basicamente aos serviços de saúde e administrativos fornecidos pela prefeitura. Quando a demanda pelo serviço se refere a informação, constata-se também maiores requerimentos em relação aos serviços de saúde. A maioria das ações de solicitação se refere a questões de trânsito, principalmente no que se refere à fiscalização de irregularidades.

Aplicando o algoritmo de associação APRIORI, chegou-se aos resultados apresentados no quadro 5. A regra 1, representa que: se o assunto for iluminação pública, na subdivisão manutenção de luminárias (com 37.181 itens), as solicitações são encaminhadas ao órgão responsável (Secretaria Municipal de Obras Públicas), que corresponde a 34.979. No entanto, está associado ao tipo de solicitação e o órgão responsável (Secretaria Municipal de obras públicas), e regra com 94% de confiança. Apresentando dependência positiva entre o conseqüente (Secretaria municipal de obras públicas) e o antecedente (Manutenção de luminárias), dados os valores *Lift* maior que 1 (um), o *Levege* maior que 0 (zero) e o valor de convicção 14,84.

Quadro 5 – Melhores regras encontradas pelo algoritmo APRIORI na “Central 156”

ASSUNTO=ILUMINAÇÃO PÚBLICA SUBDIVISAO=MANUTENÇÃO DE LUMINÁRIAS 37181 ==> TIPO=SOLICITAÇÃO ORGAO_RESP=SECRETARIA MUNICIPAL DE OBRAS PÚBLICAS 34979 <conf:(0.94)> lift:(7.78) lev:(0.1) [30485] conv:(14.84)
ASSUNTO=COLETA 74689 ==> TIPO=SOLICITAÇÃO 72201 <conf:(0.97)> lift:(1.12) lev:(0.02) [7557] conv:(4.04)
ASSUNTO=TRÂNSITO 46700 ==> TIPO=SOLICITAÇÃO 46141 <conf:(0.99)> lift:(1.14) lev:(0.02) [5721] conv:(11.22)

Fonte: Autores.

Estas informações corroboram no que se refere às demandas de solicitação, as quais em sua maioria são decorrentes do assunto trânsito. Verificou-se também várias solicitações no que se refere a iluminação pública e coleta de materiais.

Com a finalidade de encontrar outros padrões, excluiu-se o tipo “solicitação”, visto que foi a variável que proporcionou maior taxa de acerto nos algoritmos de classificação e associação. Esta medida almeja encontrar padrões para as demandas com menores frequências.

Isto posto, aplicaram-se novamente os algoritmos de classificação (algoritmo de PART) e associação (algoritmo APRIORI), para os dados sem o tipo ‘solicitação’, obtendo os resultados apresentados nas tabelas 3 e 4.

Tabela 3 – Resumo da partição de 10% pelo algoritmo PART sem a variável solicitação

Estadísticas	Valores	Porcentagem
Taxa de acerto	36.189	84,07%
Taxa de erro	6.855	15,92%
Estatística Kappa	0,6911	
Erro absoluto médio	0,0602	
Erro quadrático médio	0,1762	
Erro absoluto relativo	41,99%	
Erro quadrático relativo	65,79%	
Número total de itens	43.044	

Fonte: Elaborado pelos autores (2019)

De acordo com a tabela 3, a classificação acertou em 84,07% dos casos, apresentando uma coesão (estatística Kappa) de 0,6911, dando legitimidade as interpretações, e um erro médio absoluto de 0,0602.

Tabela 4 - Matriz de confusão pelo algoritmo Part sem a variável solicitação

a	b	c	d	e	f	g	Classificação	Variáveis
0	0	1	0	1	3	0	a	CONSULTA PÚBLICA
0	508	14	0	93	256	0	b	DENÚNCIA
0	7	4522	13	416	1065	7	c	ELOGIO
0	0	2	1414	1	8	0	d	INFORMAÇÃO
0	62	69	0	3874	696	12	e	RECADASTRO
0	115	2700	2	786	25710	25	f	RECLAMAÇÃO
0	5	24	0	59	413	161	g	SUGESTÃO

Fonte: Autores.

A matriz de confusão apresentada na tabela 4, mostra onde houve mais acerto ou erro na classificação realizada. As variáveis: ‘informação’, ‘recadastro’ e ‘reclamação’, foram as que obtiveram maior taxa de acerto. Como as demandas com maiores taxas de acertos foram as mesmas apresentados pelo algoritmo quando as solicitações estavam na base, as melhores regras foram repetidas.

A fim de se verificar se este comportamento se repetia, aplicou-se o algoritmo APRIORI sem a ‘solicitação’. As principais regras encontradas foram similares as

apresentadas no quadro 5. Desta forma, verifica-se que a retirada do item 'solicitação' foi responsável apenas pela diminuição da taxa de acerto dos algoritmos, não apresentando nenhuma regra diferente as já encontradas.

Uma vez realizada a mineração, partiu-se para a análise estatística com o intuito de se realizar análises complementares, bem como corroborar o conhecimento descoberto na etapa de mineração de dados.

4ª Etapa: Análise estatística

Uma vez realizada a mineração de dados, passe-se agora a quarta etapa no processo de descoberta de conhecimento sugerido pelos autores. A análise descritiva dos dados se faz necessária a fim de que se possa identificar o comportamento geral dos dados contidos na base.

Os meses Janeiro e Fevereiro são os que mais apresentam demandas no serviço, correspondendo a 36,04% do total. O período do dia com maior incidência é o período da tarde com 42,62%. Os dez bairros com mais demandas são apresentados na tabela 5:

Tabela 5 – Bairros com maior demanda na “Central 156”

Bairro	%
CENTRO	6,63%
CIDADE INDUSTRIAL	5,17%
CAJURU	4,34%
SITIO CERCADO	4,32%
BOQUEIRAO	3,97%
AGUA VERDE	3,12%
XAXIM	2,86%
UBERABA	2,75%
PORTAO	2,71%
NOVO MUNDO	2,64%

Fonte: Autores.

No centro, os assuntos mais demandados são trânsito, abordagem social de rua – adulto, abordagem social de rua, coleta, motoristas cobradores e porteiros. Na cidade industrial os principais itens são: coleta, iluminação pública, posto de saúde, trânsito, drenagem. No bairro Cajuru, coleta, iluminação pública, posto de saúde, trânsito, unidade de saúde 24h.

Verifica-se que 50,09% dos assuntos com maiores demandas são: coleta, trânsito e iluminação. Os bairros Cajuru, Boqueirão, Bairro Alto, Cidade Industrial, Sítio Cercado e Xaxim são principais responsáveis pela coleta. O assunto trânsito tem a maior incidência nos bairros Centro, Água Verde, Batel, Alto da Glória, Centro Cívico.

Os bairros com maior incidência do assunto Iluminação pública são: Sítio Cercado, Cajuru, Cidade Industrial, Uberaba, Boqueirão. Dada esta informação e considerando que estes bairros são considerados os mais violentos, conforme Secretaria de Segurança Pública do Estado do Paraná, correlacionou o total de demandas por iluminação pública com o total de vítimas homicídio doloso, total

vítimas roubo com resultado morte (latrocínio) e total vítimas lesão corporal com resultado morte. A tabela 6 apresenta a matriz de correlação entre estas variáveis.

Tabela 6 – Matriz de correlação (o total de vítimas homicídio doloso, total vítimas roubo com resultado morte (latrocínio))

	<i>Total Vítimas Homicídio Doloso</i>	<i>Total Vítimas Roubo Com Resultado Morte (Latrocínio)</i>	<i>Total Vítimas Lesão Corporal Com Resultado Morte</i>	<i>Total de demandas por assunto Iluminação Pública</i>
Total Vítimas Homicídio Doloso	1			
Total Vítimas Roubo Com Resultado Morte (Latrocínio)	0,593590215	1		
Total Vítimas Lesão Corporal Com Resultado Morte	0,304702388	0,398862018	1	
Total de demandas por assunto Iluminação Pública	0,702225199	0,596589485	0,365224485	1

Fonte: Autores

Verifica-se, portanto, a presença de associação entre as variáveis total de demandas por assunto iluminação pública com total de vítimas por homicídio doloso.

Outro assunto a se analisar são as demandas quanto aos postos de saúde. Os bairros com maiores elogios, reclamações e solicitações são: Sítio Cercado, Cidade Industrial de Curitiba, Cajuru, Tatuquara e Xaxim. Os elogios correspondem a 20% das reclamações neste assunto.

5ª Etapa: Avaliação ou validação de conhecimento

Os bairros com maior incidência de demandas cujo assunto é iluminação pública: Cidade Industrial de Curitiba, Sítio Cercado e Cajuru, são os bairros com maior número de vítimas ocasionadas por homicídio doloso, conforme pode se verificar no relatório estatístico criminal da Secretaria Estadual de Segurança Pública do Paraná.

Quando o assunto é posto de saúde, os bairros com maiores reclamações, elogios e solicitações são os que possuem um índice de Desenvolvimento Humano Municipal menor em relação aos bairros com menores registros.

A mineração de dados evidenciou que as ações de solicitação têm maior frequência em relação aos assuntos de trânsito. As reclamações e informações se referem e sua maioria nos serviços de saúde. Com estas informações, a prefeitura

poderia focar esforços na melhora da qualidade dos serviços de saúde, o que diminuiria as reclamações. O investimento em educação no trânsito poderia ser uma ótima ferramenta para minimizar as constantes demandas por fiscalização.

CONSIDERAÇÕES FINAIS

A “Central 156” da cidade de Curitiba, pioneira entre as capitais do Brasil e com 35 anos de atendimento ao cidadão, representa o diagnóstico do desempenho da cidade pelo olhar do cidadão, o qual realiza a sua participação no encaminhamento das reivindicações, seja estas: solicitação, informação, elogio, reclamação ou sugestão.

De acordo com o levantamento da história da pesquisa, foi possível verificar a evolução crescente do objeto estudado quando se relaciona a preocupação que a administração pública tem com o habitante da cidade em conjunto com as diversas possibilidades de o mesmo entrar em contato, seja pelo telefone, *chat* de atendimento *on line* e aplicativo.

Foi possível constatar que existe muito a ser estudado sobre o contexto “Central 156”, em especial por ser um assunto que retrata as condições da cidade e que poderá detectar qual o serviço deve haver maior investimentos de recursos pelo gestor, seja este municipal ou estadual, o que permite desenvolver técnica constante de acompanhamento, a fim de melhor entender a cidade no ponto de vista do cidadão.

A partir da configuração metodológica apoiada em um estudo de mineração de dados pela utilização do aplicativo Weka, a pesquisa revelou que mais de 50% dos assuntos com maiores demandas são: coleta, trânsito e iluminação. Os bairros Cajuru, Boqueirão, Bairro Alto, Cidade Industrial, Sítio Cercado e Xaxim são principais responsáveis pela coleta. Sempre que o assunto é referente ao transporte público há a requisição de reclamação, sugerindo um problema neste tipo de serviço. Os bairros com maior incidência de demandas por iluminação pública são: Cidade Industrial de Curitiba, Sítio Cercado e Cajuru.

Quando o assunto é posto de saúde, os bairros com maiores reclamações, elogios e solicitações são Sítio Cercado, Cidade Industrial de Curitiba, Cajuru, Tatuquara e Xaxim. A mineração de dados evidenciou que as ações de solicitação têm maior frequência em relação aos assuntos de trânsito. As reclamações e informações se referem e sua maioria nos serviços de saúde.

Para estudos futuros podem ser desenvolvidos como objetivo de aumentar o escopo da pesquisa, por meio, por exemplo de mineração de texto no campo de descrição de solicitação. Além disso, pode-se realizar comparações da ferramenta com o implementado em outras cidades.

Data mining in support of municipal public management: knowing the demands of the city of Curitiba by "Central 156"

ABSTRACT

Analyzing the city of Curitiba by the "Central 156", we sought to identify patterns in the participation of the population in claims such as: 'public consultation', 'complaint', 'compliment', 'information', 'registration', 'complaint', 'request' and 'suggestion'. To this end, data were mined at the Central 156 base in 2018, using Microsoft-Excel and Weka applications. Firstly, we collected data available on the Curitiba City Hall website and merged, month by month, in a single database. From the selection of variables by the BestFirst method, association (APRIORI) and classification (PART) algorithms were performed, as well as statistical analysis. The analysis revealed rules such as: whenever the subject is related to public transport, there is a complaint request, suggesting a problem in this type of service; There is a relationship between the total demands for public lighting and the total number of victims of intentional homicide. Future searches may further increase the scope of the search by text mining in the request description field.

KEYWORDS: Association (APRIORI) and classification (PART) algorithms. Central 156. Municipality of Curitiba.

REFERÊNCIAS

AMARAL, F. **Introdução a ciência de dados**: mineração de dados e Big Data. Rio de Janeiro: Alta Books, 2016.

ADRIAANS, P.; ZANTINGE, D. **Data mining**. Harlow: Assison-Wesley, 1996; HAN, J.; KAMBER, M.; PEI, J., Data mining: concepts and techniques. 3.ed. São Francisco: Morgan Kaufmann, 2011.

BATISTA, Z. N. et al. Participação pública em processos decisórios sobre políticas de Ciência e Tecnologia: possibilidade de intervenção eficaz contra o risco tecnocientífico. **Revista Tecnologia e Sociedade**, Curitiba, v. 15, n. 36, p. 112-133, abr./jun. 2019. Disponível em: <<https://periodicos.utfpr.edu.br/rts/article/view/8000/6041>>. Acesso em: 17 jun. 2019.

CÂMARA MUNICIPAL DE CURITIBA. Curitiba, 19-?. Disponível em: <<https://www.cmc.pr.gov.br/anexo.php?anexo=17125>>. Acesso em: 15 jun. 2019.

CASTRO, L. N. de.; FERRARI, D. G.. **Introdução à mineração de dados**: conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva, 2016.

CENTRAL 156. Curitiba, 2019. Disponível em: <<http://www.central156.org.br/conteudo/sobre-a-central-156/4>>. Acesso em: 15 jun. 2019.

CENTRAL 156. Curitiba, 2019. Disponível em: <<http://www.central156.org.br/aviso/central-156-completa-35-anos/52>>. Acesso em: 15 jun. 2019.

FAYYAD, U; PIATETSKY-SHAPIO, G; SMYTH, P. The KDD Process for Extracting Useful Knowledge from Volumes of Data. **Communications of the ACM**, 39(11):27–34, November 1996. Disponível em: <https://sceweb.uhcl.edu/boetticher/ML_DataMining/p27-fayyad.pdf>. Acesso em: 23 jun. 2019.

GONÇALVES, A. Serviços públicos na palma da mão: 21 apps para facilitar a vida do curitibano. **Gazeta do Povo**, Curitiba, 09 abr. 2019. Disponível em: <<https://www.gazetadopovo.com.br/curitiba/aplicativos-curitiba-servicos-publicos/>>. Acesso em: 16 jun. 2019.

GUIMARÃES, L.A.; OLIVEIRA, A. F. B. A importância em gestão pública municipal para os gestores públicos no exercício e suas funções. **Id on Line Rev. Mult. Psic.**

v.13, n. 45, 356-372, 2019. Disponível em: <<https://idonline.emnuvens.com.br/id/article/view/1735/0>>. Acesso em: 29 set. 2019.

IBGE, Cidades. Paraná, Curitiba. Disponível em: <<https://cidades.ibge.gov.br/brasil/pr/curitiba/panorama>>. Acesso em: 26 jun. 2019.

HOJDA, A.; FARINIUK, T. M. D.; SIMÃO, M de M. B.. Building a smart city with trust: the case of '156 central' of Curitiba-Brazil. **Economía, Sociedad y Territorio**. vol. XIX, n. 60, mayo-agosto, 79-108, 2019. Disponível: <<https://est.cmq.edu.mx/index.php/est/article/view/1298/1448>>. Acesso em: 23 jun. 2019.

KANUFRE, R. A. M.; REZENDE, D.A. Princípios da gestão orientada para resultados na esfera municipal: o caso da prefeitura de Curitiba. **Revista de Administração**. v. 47, n. 4, outubro/dezembro, 638-652, 2012. Disponível em: <<http://www.scielo.br/pdf/rausp/v47n4/a10v47n4.pdf>>. Acesso em: 29 set. 2019.

KAPLAN, H.I.; SADOCK, B.J. **Tratado de psiquiatria**, v. 1, 6. ed. Porto Alegre: Artmed, 1999.

KOWALSKI, R. L.. No fim de ano, Central 156 de Curitiba tem demanda diferente e bem típica. **Bem Paraná**, Paraná, 19 dez. 2019. Disponível em: <<https://www.gazetadopovo.com.br/curitiba/aplicativos-curitiba-servicos-publicos/>>. Acesso em: 16 jun. 2019.

Hall, M. A.. **Correlation-based Feature Subset Selection for Machine Learning**. 178 f. Tese (Doutorado em Filosofia) – Department of Computer Science, New Zealand, 1998. Disponível em: <https://www.cs.waikato.ac.nz/~mhall/thesis.pdf>>. Acesso em: 23 jun. 2019

MEINEL, K. **Motricidade II**. Rio de Janeiro: Ao Livro Técnico, 1984.

MOSQUERA, J. J. M. Adulto desenvolvimento físico e educação. **Kinesis**, v. 1, n. 2, 1985.

OLIVEIRA, O. 156: A trajetória da principal central de atendimento ao cidadão curitibano. **CBN**. Curitiba, 2019. Disponível em: <<https://cbncuritiba.com/156-a-trajetoria-da-principal-central-de-atendimento-ao-cidadao-curitibano/>>. Acesso em: 23jun. 2019.

OLIVEIRA, D. T de.; OLIVEIRA, A. G. de. As Ouvidorias do SUS: como a Participação Popular se materializa como Ferramenta de Gestão e Política Pública? **R. Technol. Soc.**, Curitiba, v. 15, n. 35, p. 290-310, jan./abr. 2019. Disponível em: <<https://periodicos.utfpr.edu.br/rts/article/view/7269/5913>>. Acesso em: 29 set. 2019.

PREFEITURA MUNICIPAL DE CURITIBA. Curitiba, 2019. Disponível em: <<https://www.curitiba.pr.gov.br/dadosabertos/busca/>>. Acesso em: 23 jun. 2019.

RAMOS, D. Central 156 ganha aplicativo para facilitar solicitações e prefeitura espera explosão de acessos. **Gazeta do Povo**, Curitiba, 29 mar. 2019. Disponível em: <<https://www.gazetadopovo.com.br/curitiba/aplicativo-central-156-curitiba/>>. Acesso em: 16 jun. 2019.

SECRETARIA DA SEGURANÇA PÚBLICA. **Relatório estatístico criminal 4º semestre de 2018**. Curitiba, 2019.

Recebido: 08/07/2019

Aprovado: 17/11/2019

DOI: 10.3895/rts.v16n40.10335

Como citar: MAIA, M.; MAIA, M.; TSUNODA, D. F. Análise de dados no apoio a gestão pública municipal: conhecendo as demandas da cidade de Curitiba pela "Central 156". **R. Technol. Soc.**, Curitiba, v. 16, n. 40, p. 1-18, abr/jun. 2020. Disponível em: <https://periodicos.utfpr.edu.br/rts/article/view/10335>. Acesso em: XXX.

Correspondência:

Direito autoral: Este artigo está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

