

# Causas de desaparecimento no estado de São Paulo entre 2013 e 2014: uma análise automatizada de boletins de ocorrência

*DISAPPEARANCE CAUSES IN THE STATE OF SÃO PAULO BETWEEN 2013 AND 2014: AN AUTOMATED ANALYSIS OF POLICE REPORTS*

**Mateus Borges\***

**Rafael Izbicki\*\***

**Rafael B. Stern\*\*\***

**Julio A. Z. Trecenti\*\*\*\***

**RESUMO** Este artigo estuda as causas de desaparecimento no estado de São Paulo por idade e sexo dos desaparecidos. Para tanto, usou o PLID, um banco de dados com os boletins de ocorrência dos desaparecimentos de 2013 e 2014. A metodologia consiste no uso de aprendizado de máquina para rotular automaticamente a causa de desaparecimento nos boletins, assim como métodos de quantificação para mensurar as proporções relativas de causa de desaparecimento estratificadas por idade e sexo. Os resultados indicam que grande parte dos boletins de ocorrência não possuem informações suficientes para inferir a causa do desaparecimento. Dentre os boletins com causa de desaparecimento clara, foi possível obter algumas classes de motivos mais frequentes para desaparecimento. A causa de desaparecimento mais frequente é a voluntária, sendo mais comum entre mulheres e homens menores de 18 anos. Também, cerca de 20% dos desaparecimentos ocorrem por usuários de drogas ou álcool, sendo mais comum entre menores de 50 anos. A partir de 50 anos, essa categoria torna-se menos frequente, sendo substituída pelo desaparecimento não intencional. Também se observou que menos de 5% dos indivíduos foram vítimas de crimes. Essas conclusões podem auxiliar na elaboração de políticas públicas mais assertivas.

\* Graduando em Estatística na UFSCAR - Universidade Federal de São Carlos. - mateus\_borges@hotmail.com

\*\* PhD em Estatística, UFSCAR - Universidade Federal de São Carlos. - rafaelizbicki@gmail.com

\*\*\* PhD em Estatística, UFSCAR - Universidade Federal de São Carlos. - rbstern@gmail.com

\*\*\*\* Associação Brasileira de Jurimetria - jtrecenti@abj.org.br

Agradecemos Eliana Vendramini pelo fornecimento dos dados à Associação Brasileira de Jurimetria e Márcia Silva, que pela análise inicial dos dados. Esse projeto foi parcialmente financiado pelo projeto FAPESP 2019/11321-9 e CNPq 306943/2017-4.

**PALAVRAS-CHAVE:** Pessoas desaparecidas, Aprendizado de máquina, São Paulo.

**ABSTRACT** This article studies the causes of disappearance in the state of São Paulo by age and sex of the disappeared. For this purpose, it used the PLID, a database with the disappearance reports of 2013 and 2014. The methodology consists in applying machine learning to automatically classify the cause of disappearance in the reports, as well as quantification methods to measure the relative proportions causes of disappearance stratified by age and sex. The results indicate that many of the reports do not have enough information to infer the cause of disappearance. Among the reports with clear cause of disappearance, it was possible to obtain some classes of more frequent reasons for disappearance. The most frequent cause is voluntary, being more common among women and men under 18 years. Also, about 20% of disappearances occur by drug or alcohol users, being more common among individuals under 50 years of age. From age 50 and up, this category becomes less frequent, being replaced by unintentional disappearance. It was also observed that less than 5% of individuals were victims of crimes. These conclusions can help with the development of more assertive public policies.

**KEYWORDS:** Missing people, Machine Learning, São Paulo.

---

## INTRODUÇÃO

De acordo com o Fórum Brasileiro de Segurança Pública (2017, 2018, 2019), entre 2017 e 2019 foram registrados 869.182 casos de desaparecimento no Brasil, com o estado de São Paulo representando 33,6% dos casos, sendo aquele com maior número de desaparecidos. Tal informação considera como desaparecidos todos os indivíduos com paradeiro desconhecido pelas suas famílias ou que foram dados como desaparecidos por fontes fidedignas. A fim de criar políticas públicas que diminuam o número de casos de desaparecimento, é importante entender as causas que levam indivíduos a desaparecerem.

No Brasil há poucos trabalhos científicos que abordam esse tema (OECHSLER, 2018). Dentre esses, poucos analisam a causa dos desaparecimentos de forma quantitativa. Uma das exceções a isso é o estudo “O desaparecimento nas burocracias do Estado”, realizado pelo Ministério Público do estado do Rio de Janeiro em 2018, que estudou as causas de 4861 boletins de ocorrência relativos a desaparecimentos que foram solucionados. Como esse estudo se restringe a casos elucidados, esse estudo pode apresentar viés de seleção, visto que algumas causas podem ser mais predominantes em pessoas que ainda não foram encontradas. O presente trabalho identificou os principais motivos que levam ao desaparecimento de indivíduos no estado de São

Paulo e como esses motivos variam de acordo com a idade e o sexo do indivíduo desaparecido, e realizou a comparação das causas com o estado do Rio de Janeiro.

Para tanto, utilizou-se os boletins de ocorrência registrados no estado de São Paulo nos anos de 2013 e 2014. Este banco de dados foi criado devido ao Programa de Localização e Identificação de Desaparecidos (PLID), ao qual o Estado de São Paulo aderiu em novembro de 2013. Trata-se de uma rede articulada que tem como objetivo prevenir, apurar e localizar desaparecimentos (POLIANO et al., 2016). O banco de dados é confidencial (visto que possui dados sensíveis dos desaparecidos, como o nome, a idade, o sexo, o endereço, etc.) e não podem ser disponibilizados. Evidencia-se que a causa do desaparecimento não está presente no banco.

Como há um grande volume de relatos (24.401 boletins de ocorrência), é inviável ler todos os textos para classificar o motivo do desaparecimento. Assim, o primeiro autor deste artigo classificou manualmente 1000 boletins. Os restantes 23.401 foram classificados automaticamente por meio de métodos de aprendizado de máquina. Utilizamos também métodos de quantificação (FORMAN, 2006; VAZ ET AL, 2019) para evitar que erros no classificador automático gerassem vieses na estimação das proporções de interesse. Por meio dos resultados obtidos, este artigo pode auxiliar na elaboração de políticas públicas assertivas a fim de combater o número de desaparecidos no estado de São Paulo.

O restante do trabalho é dividido da seguinte forma: na Seção 2, descrevemos a metodologia utilizada, a Seção 3 apresenta resultados e discussões e a Seção 4 conclui o artigo.

## **1. METODOLOGIA**

### **1.1 Classificação de textos**

Foram adotados dois sistemas de categorização das causas de desaparecimento. Essas categorizações foram definidas de modo que cada boletim pertença apenas a uma única categoria em cada sistema. Evidencia-se que os desaparecimentos devido à prisão ou detenção não noticiadas não estão contemplados em nenhum sistema, visto que o conhecimento do evento não resultaria na criação do boletim de ocorrência. Assim sendo, esse motivo está contemplado nos casos em que o relato não possui informações suficientes para inferir a causa.

Abaixo estão definidas as categorias dos sistemas adotados. O primeiro sistema foi retirado de BIEHAL, MITCHELL e WADE (2003) e possui cinco categorias (veja a Tabela 1 para alguns exemplos de boletins de ocorrência com cada uma das causas):

- **Determinado:** Casos em que o desaparecido fez a escolha ativa de desaparecer. Estes são casos nos quais o indivíduo desapareceu do domicílio levando roupas extras ou avisou sobre a decisão ou escaparam de instituições, como crianças institucionalizadas em espera de adoção ou adolescentes em cumprimento de medidas socioeducativas, tratamento psicológicos ou que sumiram de hospitais onde estavam internados;
- **Desviado:** Casos em que o indivíduo desapareceu devido ao uso de drogas ilícitas ou álcool, ou possui histórico de uso;
- **Sem intenção:** Casos em que o indivíduo não teve a intenção de desaparecer, mas sumiu devido a um problema de doença mental, como Alzheimer ou depressão, ou por ter sofrido algum acidente, como desastres naturais;
- **Vítima de crime:** Refere-se a casos em que o indivíduo foi alvo de um crime que o levou a desaparecer. São em geral casos de crimes, como sequestros, ou em casos em que o pai ou mãe sem guarda levou sua criança sem comunicar ao cônjuge;
- **Motivo desconhecido:** Essa categoria não é uma causa de desaparecimento, casos em que não há informações suficientes para inferir a causa no boletim de ocorrência.

O sistema 1 possui categorias que podem ser divididas em subcategorias, uma vez que a origem do problema em cada uma delas pode ser bastante distinta. Por isso, foi criado o sistema 2 pelos autores, com nove categorias que representam uma refinação do sistema 1:

- **Vítima de crime:** Definição idêntica à categoria de mesmo nome do sistema 2.
- **Desastre natural:** Casos em que o indivíduo foi vítima de desastres naturais. É uma subcategoria de Sem intenção, do sistema 1.
- **Problemas de saúde mental:** Casos referentes a pessoas que residem em domicílio e não tiveram a intenção de desaparecer, mas sumiram devido a um problema de doença mental, como Alzheimer ou depressão. É uma subcategoria de Sem intenção, do sistema 1.
- **Drogas ou álcool:** Casos em que o indivíduo reside em domicílio e desapareceu devido ao uso de drogas ilícitas ou álcool, ou possui histórico de uso. É uma subcategoria de Desviado, do sistema 1.
- **Saída voluntária do domicílio:** Casos em que o desaparecido fez a escolha ativa de desaparecer do domicílio em que reside. Estes são casos nos quais o indivíduo desapareceu levando roupas extras ou avisou sobre a decisão. É uma subcategoria de Determinado, do sistema 1.
- **Evadiu de instituição:** Casos referentes a pessoas que escaparam de instituições de reabilitação. Por exemplo, essa categoria engloba casos em que o indivíduo

saiu sem permissão do hospital em que estava internado, assim como de clínicas de reabilitação por uso de drogas ou tratamento psicológico. Também inclui casos de crianças institucionalizadas em espera de adoção ou adolescentes em cumprimento de medidas socioeducativas. Portanto, essa causa está relacionada com a categoria Determinado, Desviado e Sem intenção, do sistema 1.

- **Erro no registro:** Boletim de ocorrência que foram registrados com erro e não são sobre desaparecimento. Por exemplo, boletins de ocorrência sobre furtos se enquadram nesta categoria. Essa categoria não é uma causa de desaparecimento e é uma subcategoria de Motivo Desconhecido, do sistema 1.
- **Sem relato:** Boletins de ocorrência sem texto, possuindo apenas “SEM BO CORRESPONDENTE” no relato. Essa categoria não é uma causa de desaparecimento e é uma subcategoria do Motivo Desconhecido, do sistema 1
- **Não informativo:** São boletins de ocorrência que possuem um relato no registro, mas não se encaixam em nenhuma outra categoria, logo são casos em que não há informações suficientes para inferir a causa. Essa categoria é uma subcategoria do Motivo Desconhecido do sistema 1.

É importante destacar que, em ambas as classificações, os indivíduos com até dois anos de idade não desaparecem voluntariamente, por drogas ou álcool, por problemas de saúde mental ou evasão de instituição. Porém, esses indivíduos podem ser classificados nos motivos citados porque o responsável, que pode se enquadrar nos motivos, desapareceu levando a criança. Por exemplo, se um usuário de drogas desaparece levando seu filho, ambos se encaixam na categoria Desviado (sistema 1) ou Drogas ou álcool (sistema 2), embora a criança não seja usuária. Esses casos não são problemas de classificação, visto que o motivo do desaparecimento do responsável causou o desaparecimento da criança.

**Tabela 1:** Exemplos de classificação de relatos dos boletins de ocorrência pelos sistemas 1 e sistema 2. Os erros de português presentes nos relatos originais foram mantidos.

RELATO	SISTEMA 1	SISTEMA 2
“Presente nesta unidade policial o declarante informando que sua esposa aqui como desaparecida saiu da casa de sua mãe, onde sabe-se que teve uma discussão, levando consigo seus filhos Victor Miguel (3 anos) e Vinicius Gabriel (10 meses), informa o...”	Determinado	Saída voluntária do domicílio

RELATO	SISTEMA 1	SISTEMA 2
“.. a mesma saiu de casa para vir a esta cidade estudar na escola...e não mais retornou ... e que deixou uma carta dizendo que iria morar em Ribeirão Preto com uma amiga e tia desta...”	Determinado	Saída voluntária do domicílio
“comparece nesta especializada a declarante juliana informando ser irmã de Margarete aparecida rocha a qual saiu de casa na manhã do dia 28 /11/2012 e até hoje não retornou. informa ainda que a irmã é usuária de crack, e passou a noite toda consumindo a droga e que já desapareceu outras vezes...”	Desviado	Drogas ou álcool
“..a data de 02/02/2013, foi informada pelos amigos do seu filho que vários dias não sabem do paradeiro de Bruno. Esclarece que Bruno é usuário de entorpecentes e que não é primeira vez que desaparece. Deixa claro também que Bruno tem por costume passar dias fora de casa e que comparece nesta...”	Desviado	Drogas ou álcool
“... Ocorre que, logo no início do voo, após o primeiro contato de seu pai com a torre de comunicação de Jacarepaguá, cerca de 5 a 7 minutos depois, não houve mais qualquer contato do mesmo, havendo a perda de sinal do avião, sendo que desde então seu paradeiro permanece incerto...”	Sem intenção	Desastre natural

RELATO	SISTEMA 1	SISTEMA 2
<p>“...nesta madrugada, quando levantou-se, percebeu que o mesmo não estava mais no quarto. imediatamente saiu a procura de seu padrasto pela casa, uma vez que o mesmo caminha com dificuldades e sofre de Alzheimer, mas não o encontrou. declara que, não sabendo como, mas que possivelmente tenha sido Edvaldo, quem furou o pneu da motocicleta da declarante, causando assim um atraso para iniciar a procura de Edvaldo. que seu padrasto abriu o portão e saiu para a rua... que Edvaldo tem problemas de memória, possivelmente saiu para a rua e não consegue lembrar-se de como voltar...”</p>	Sem intenção	Problemas de saúde mental
<p>“... informando que sua filha é especial e ontem brincava na frente de sua casa, vindo a desaparecer. ... A declarante passou então a mostrar a foto de sua filha às pessoas moradoras no caminho que dá acesso à sua casa, tendo recebido a informação de um vizinho de nome AFONSO e, ele falou que vira a vítima conversando com um desconhecido na viela e, que o tal desconhecido abraçou Ketlyn e subiram tomando sentido a uma favela existente no local...”</p>	Vítima de crime	Vítima de crime
<p>“... combinaram que a criança fosse entregue no domingo à tarde, contudo, o autor não devolveu a criança. Esclarece que desde aquela oportunidade não tem notícias de seus filhos e também não conseguiu contato com nenhum familiar do autor...”</p>	Vítima de crime	Vítima de crime
<p>“SEM BO CORRESPONDENTE”</p>	Motivo desconhecido	Sem relato
<p>“presente nesta delegacia de polícia, a declarante informa que seu irmão Marcelo da silva reside no município de Ibiraci, sendo que na data de 22/02/2013 veio ao centro da cidade de Franca e depois simplesmente desapareceu, não dando notícias até a presente data.”</p>	Motivo desconhecido	Não informativo

RELATO	SISTEMA 1	SISTEMA 2
“por volta de 10:45hrs, a paciente estava no pátio com outras cliente, evadiu-se pulando a cerca de arame, saiu correndo até a portaria, fomos comunicados pelo porteiro, feito busca sem sucesso...”	Desviado	Evadiu de instituição
“Comparece a vítima neste DP e informa que deixou seu veículo estacionado aberto por aproximadamente por 20 minutos e quando retornou percebeu que haviam furtado sua carteira que continha...”	Motivo desconhecido	Erro no registro

## 1.2 Classificação automática via aprendizado de máquina

Mil boletins de ocorrência escolhidos aleatoriamente foram classificados manualmente segundo os dois sistemas descritos na seção anterior. Esses boletins foram então usados para treinar modelos de aprendizado de máquina a fim de classificar automaticamente os outros 23.401 boletins. Os modelos de aprendizado de máquina supervisionado treinados foram Random forest, Penalized Logistic Regression, Support Vector Classifier, k-Nearest Neighbors, XGBoost e Naive Bayes (HASTIE, TIBSHIRANI, FRIEDMAN, 2008; IZBICKI E SANTOS, 2020). Esses métodos foram escolhidos por conseguirem modelar bem diversas relações entre a variável resposta (o motivo do desaparecimento) e as covariáveis extraídas do texto. Por exemplo, random forests, penalized logistic regression e XGBoost fornecem bom desempenho preditivo em cenários em que há muitas covariáveis irrelevantes (isto é, não associadas ao desfecho), uma vez que eles fazem implicitamente seleção de variáveis. Já Support Vector Classifier e k-Nearest Neighbor possuem bom desempenho quando há redundância entre as covariáveis (IZBICKI et al., 2017). Além disso, cada um dos métodos funciona de maneira bastante diferente dos demais, de modo que com esse leque de escolhas é possível modelar bem diversos problemas preditivos. Esses modelos foram ajustados utilizando o módulo scikit-learn da linguagem Python (PEDREGOSA et al., 2011), e os valores de seus tuning parameters foram definidos através da validação cruzada (IZBICKI E SANTOS, 2020).

Escolheu-se então o modelo com maior valor de estatística F1 (a média harmônica entre precisão e recall) para fazer as classificações dos boletins, uma vez que essa estatística fornece um balanço entre erros tipo 1 e 2 (IZBICKI E SANTOS, 2020). No problema de classificação de boletins de ocorrência, precisão é a proporção de



classificações corretas dentre todas as classificações de uma determinada classe estimada pelo modelo e recall é a proporção de classificações corretas dentre todas as classificações rotuladas de uma determinada classe. A estimativa desta estatística foi feita por meio da técnica de validação cruzada a fim de evitar o super-ajuste (JAMES et al., 2013). Após treinar os modelos, a proporção de cada um dos motivos de desaparecimento no estado de São Paulo foi estimada utilizando o método desenvolvido por Vaz et al. (2019) (por meio do estimador combinado). Este método corrige a proporção estimada dada pela contagem de quantos relatos foram classificados em uma determinada categoria de acordo com as taxas de erro do classificador automático. Assim, ele elimina vieses que ocorrem devido ao fato de que classificadores automáticos podem cometer erros. Para mais detalhes, veja o Apêndice A.1. Além de estimar as proporções de cada uma das causas, também mostramos intervalos com 95% de confiança para essas quantidades.

## 2. RESULTADOS E DISCUSSÃO

As Seções 2.1 e 2.2 mostram os resultados para os sistemas de classificação 1 e 2, respectivamente.

### 2.1 Sistema de classificação 1

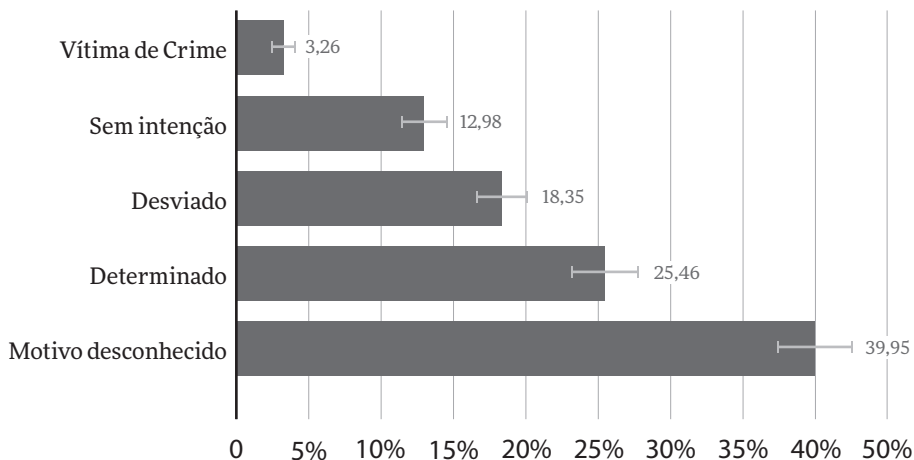
A Tabela 2 mostra a acurácia e estatística F1 de cada modelo de aprendizado de máquina. A Penalized Logistic Regression com penalização l2 foi o melhor modelo segundo ambas as métricas. Assim, esse foi o modelo escolhido para o restante das análises. Mais detalhes sobre o desempenho do modelo podem ser encontrados no Apêndice A.2, que também inclui os dados brutos utilizados para fazer os gráficos apresentados aqui.

**Tabela 2:** Acurácia e valor da estatística F1 dos modelos de aprendizado de máquina treinados para classificar boletins de ocorrência segundo o sistema 1.

MODELO	ACURÁCIA (%)	ESTATÍSTICA F1
XGBClassifier	61,9 ± 1,5	0,642
Naive Bayes	58,7 ± 1,6	0,581
Extra Tree Classifier	65,3 ± 1,5	0,608
<b>Penalized Logistic Regression</b>	<b>69,1 ± 1,5</b>	<b>0,681</b>
Support Vector Classifier	52,3 ± 1,6	0,453

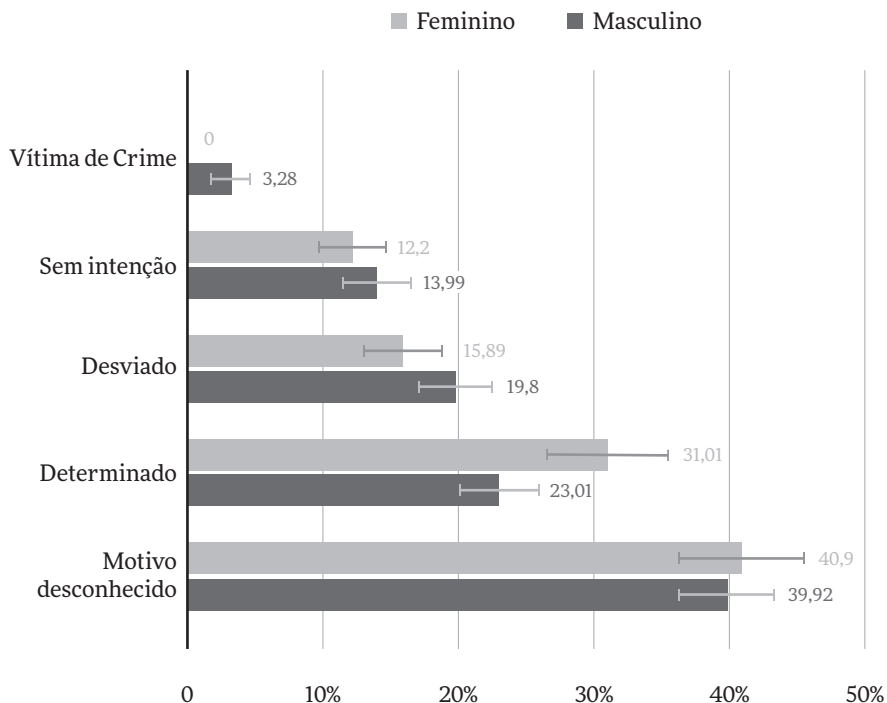
Entre as categorias dos boletins de ocorrência (B.O.), o Motivo desconhecido tem a maior prevalência, como mostrado na Figura 1, indicando que há muitos B.O. sem informações relevantes para inferir a causa. A causa Determinado apresenta a maior prevalência, seguido de Desviado e Sem intenção, enquanto Vítima de crime apresenta pouca prevalência.

**Figura 1:** Porcentagem de cada causa de desaparecimento usando o sistema 1.



A seguir, analisamos a prevalência de cada uma das causas separando a análise por diferentes fatores. A Figura 2 mostra que, entre pessoas do sexo feminino há uma chance maior de o desaparecimento ser da causa Determinado, do que entre homens. Assim, mulheres têm maior probabilidade de ter desaparecido por decisão própria que homens. Não foram encontradas diferenças significativas entre os gêneros com relação às outras causas.

**Figura 2:** Porcentagem das categorias do B.O pelo sexo do indivíduo, usando o sistema 1.

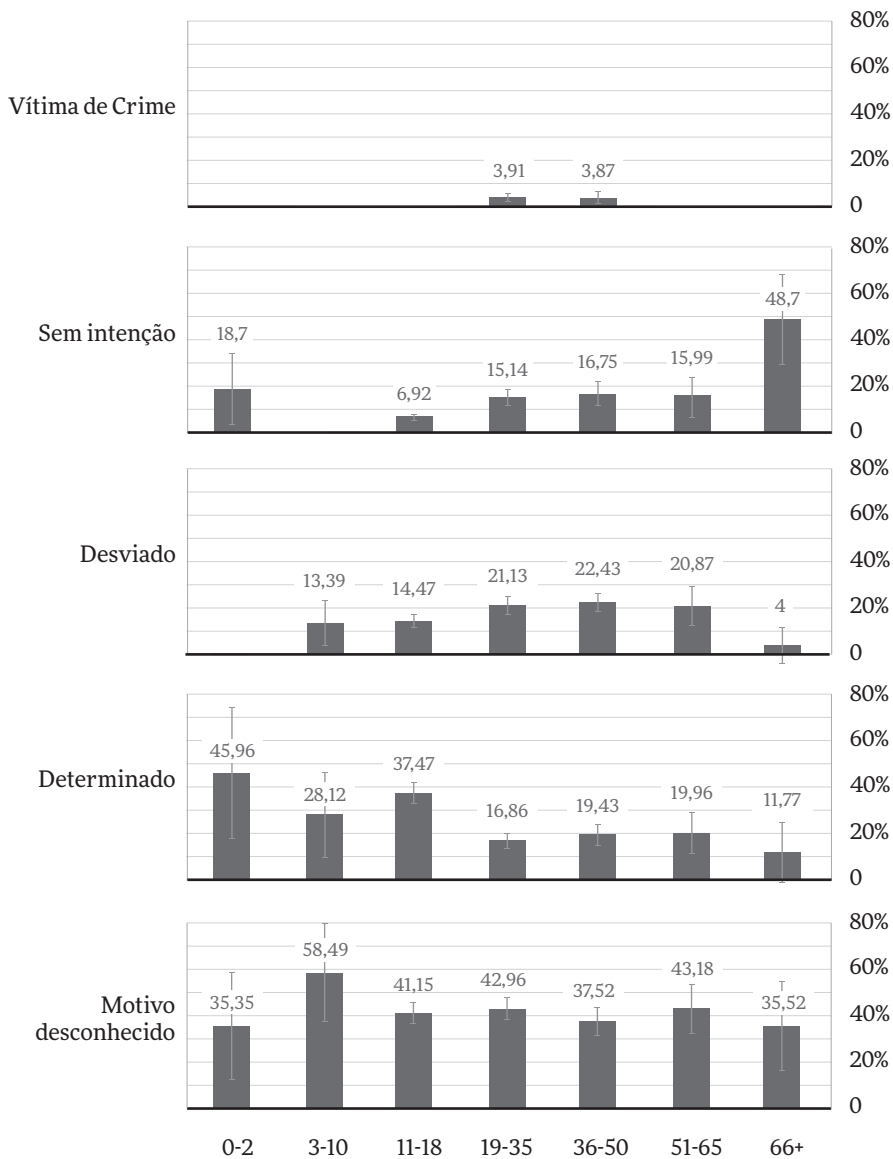


A Figura 3 mostra como as diferentes causas de desaparecimento variam com a idade. Entre os indivíduos com mais de 66 anos, a causa mais frequente é a sem intenção. Além disso, em geral indivíduos entre 3 e 10 anos de idade não desaparecem sem intenção.

A prevalência do desaparecimento do tipo Desviado varia entre as faixas de idade: a porcentagem aumenta até a faixa entre 36 e 50 anos e volta a diminuir logo em seguida, sendo as menores porcentagem para indivíduos menores que 2 anos e maiores que 65 anos, respectivamente.

Indivíduos até 2 anos de idade são os mais afetados pelo desaparecimento do tipo Determinado. As faixas de idade entre 3 a 10 anos e entre 11 a 18 anos possuem as maiores porcentagens respectivamente para essa categoria. Para indivíduos maiores de 18 anos, essa porcentagem decaí.

**Figura 3:** Porcentagem das categorias dos B.O. pelo faixa de idade do indivíduo, usando o sistema 1.



## 2.2 Sistema de classificação 2.

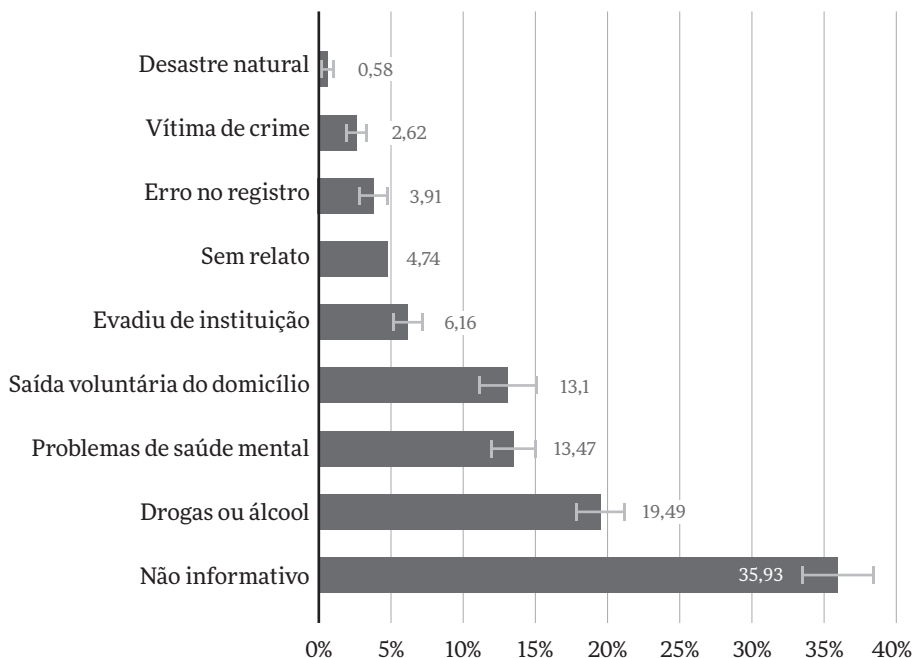
A acurácia e estatística F1 de cada modelo de aprendizado de máquina são apresentadas na Tabela 3. Utilizamos o modelo de Penalized Logistic Regression com penalização l2, por este ter apresentado melhor desempenho.

**Tabela 3:** Acurácia e valor da estatística F1 de modelos de aprendizado de máquina treinados para classificar boletins de ocorrência segundo o sistema 2.

MODELO	ACURÁCIA (%)	ESTATÍSTICA F1
XGBClassifier	68,6 ± 1,5	0,664
KNN	26,6 ± 1,4	0,231
Extra Tree Classifier	57,4 ± 1,6	0,572
<b>Penalized Logistic Regression</b>	<b>68,3 ± 1,5</b>	<b>0,668</b>
Support Vector Classifier	67,7 ± 1,5	0,661
Naive Bayes	62,2 ± 1,5	0,604

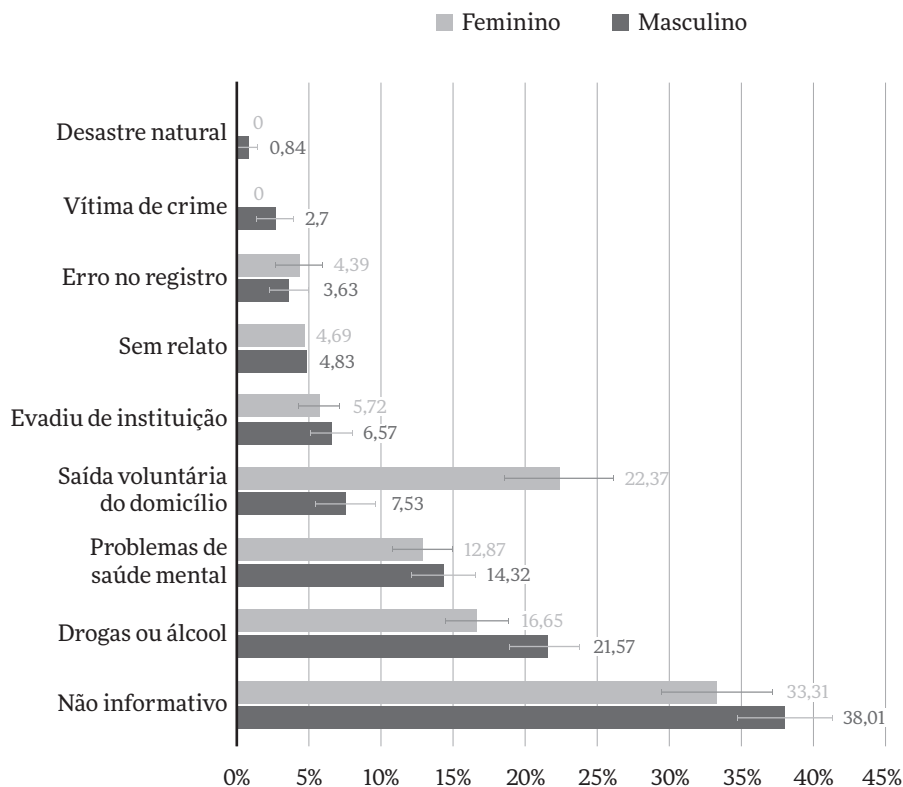
A Figura 4 mostra que Não informativo foi a categoria com maior prevalência, o que indica que há muitos relatos sem informações suficientes para inferir a causa. Entre as causas propriamente ditas, Drogas ou álcool apresentou maior prevalência, seguida de Problemas de saúde mental e Saída voluntária do domicílio, enquanto que Desastre natural apresenta prevalência quase nula.

**Figura 4:** Porcentagem de cada causa usando o sistema 2.



A Figura 5 indica que mulheres têm mais probabilidade de desaparecer do domicílio por causa voluntária que homens. Além disso, indivíduos do sexo masculino apresentam maior prevalência do tipo Drogas ou álcool que mulheres. Não foram encontradas diferenças significativas entre os gêneros com relação às outras causas.

**Figura 5:** Porcentagem das categorias do B.O. pelo sexo do indivíduo, usando o sistema 2.



A Figura 6 mostra como as diferentes causas de desaparecimento variam com a idade. O desaparecimento do tipo Problemas de saúde mental tem um pico para indivíduos maiores de 66 anos e representam, para os boletins de ocorrência lidos, problemas de esquecimento por causa da idade avançada. Em geral, indivíduos entre 3 e 10 anos de idade não desaparecem por problemas de saúde mental.

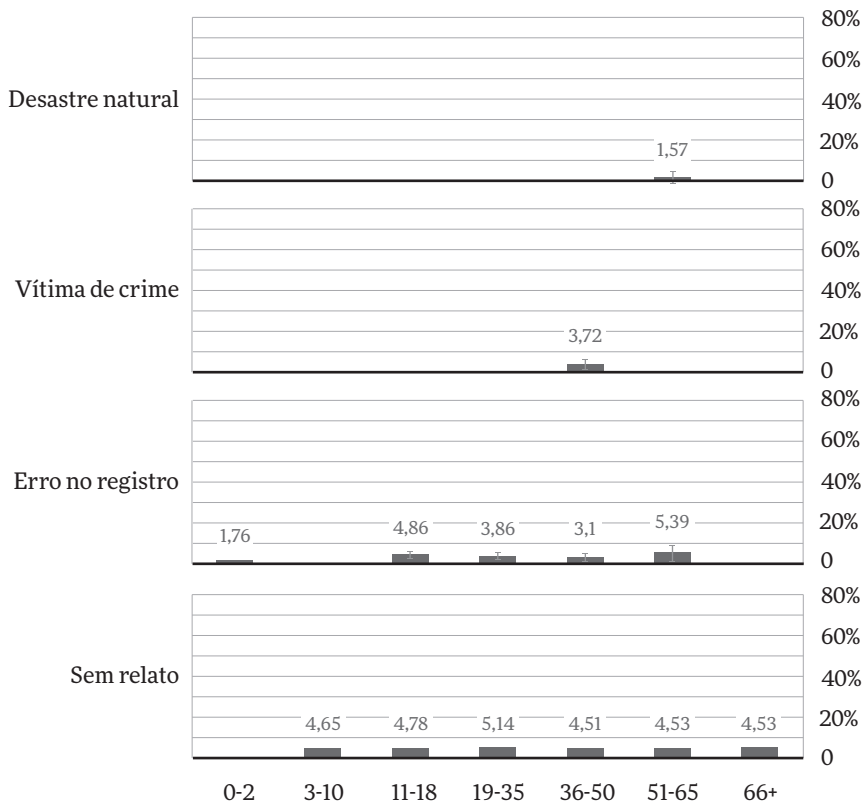
O desaparecimento do tipo Drogas ou álcool varia entre as faixas de idade; a porcentagem aumenta até a faixa entre 36 e 50 anos e volta a diminuir logo após, sendo as menores porcentagem para indivíduos de 65 anos.

A porcentagem da causa de Saída voluntária do domicílio decai após os 18 anos de idade, e possui maior prevalência entre os 11 e 18 anos. Durante a leitura dos relatos, constatamos que indivíduos menores de 2 anos eram levados pelos responsáveis que decidiram desaparecer. Para indivíduos entre 3 e 18 anos, tratava-se de casos em que o

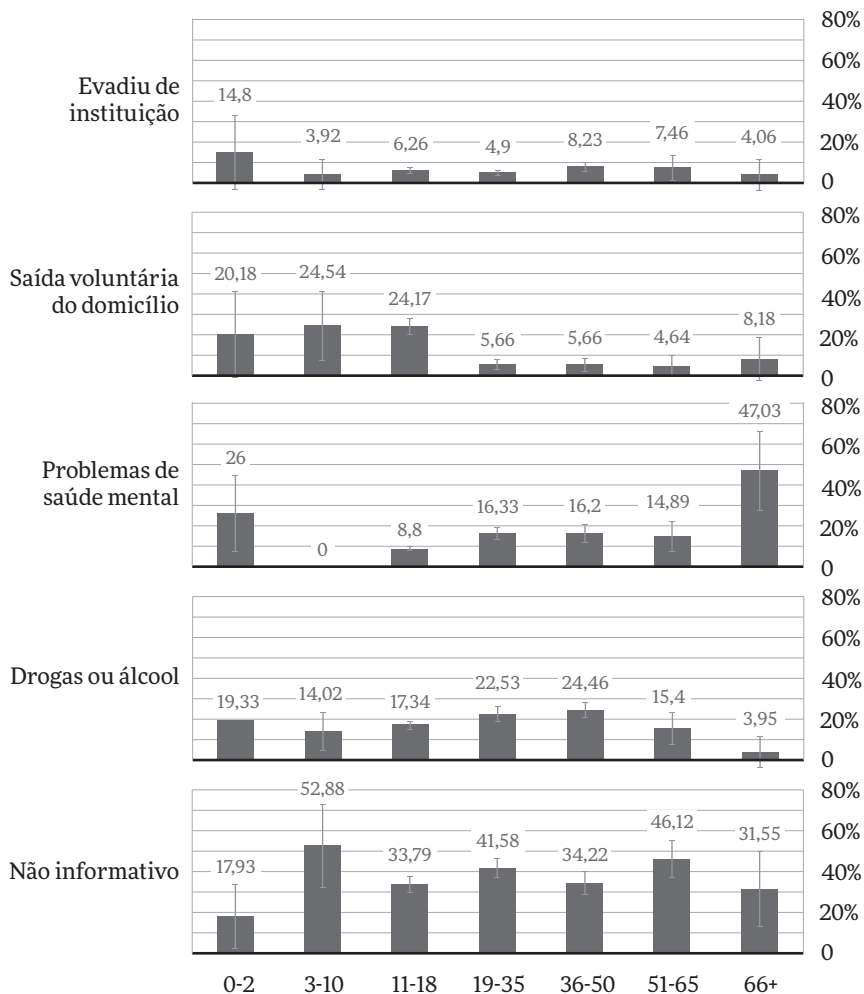
menor de idade tinha problemas familiares e decidia sair do domicílio de seu responsável. Para indivíduos maiores de idade, os relatos representam casos de indivíduos que decidiram desaparecer por brigas com o companheiro e/ou abandonando o lugar onde moravam.

A menor porcentagem da causa do tipo Não informativo é para indivíduos menores de 2 anos de idade. Para indivíduos entre 3 e 10 anos, não é possível inferir o motivo na maioria dos boletins.

**Figura 6:** Porcentagem das categorias dos B.O. pela faixa de idade do indivíduo, usando o sistema 2.







### 2.3 Comparação com Rio de Janeiro

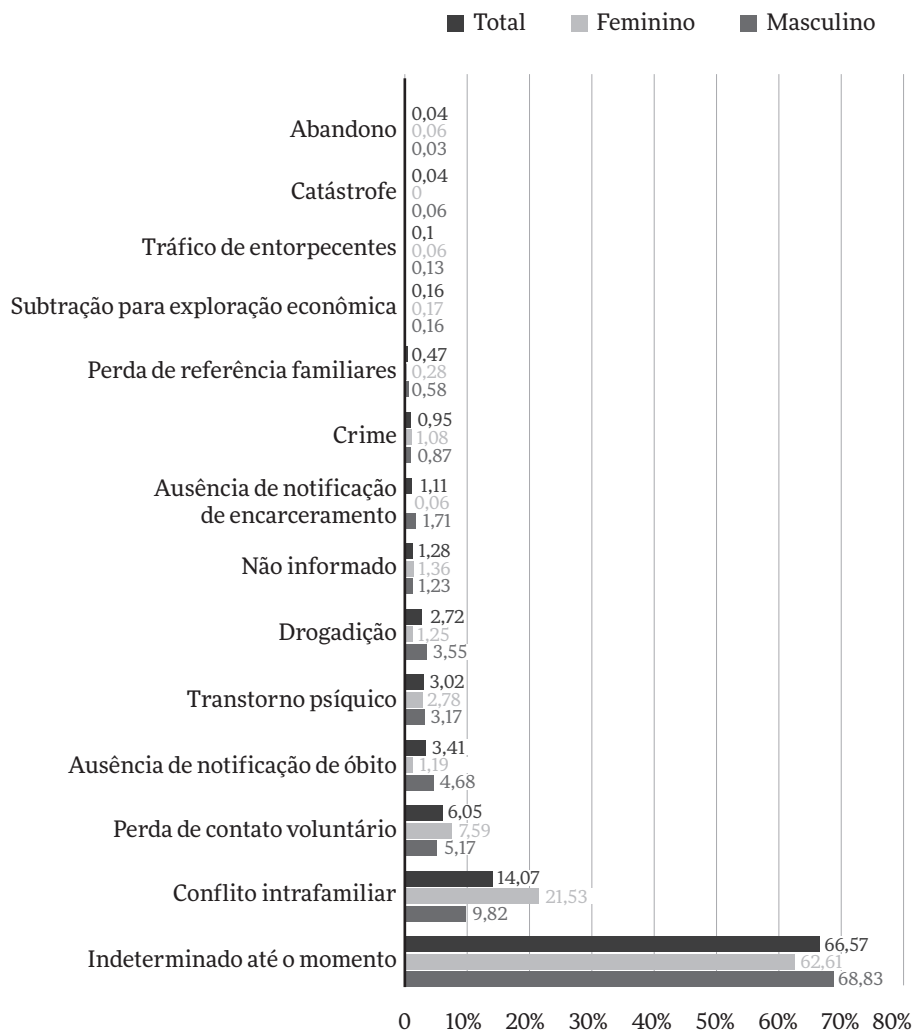
Nesta seção, comparamos os resultados aqui encontrados com os obtidos no estudo “O desaparecimento nas burocracias do Estado”, realizado pelo Ministério Público do estado do Rio de Janeiro em 2018. A categorização foi realizada pelos profissionais do PLID a partir do relato de familiares, para casos positivados. A definição das

categorias para identificar o motivo do desaparecimento, retirada daquele estudo, são as seguintes:

- Conflito intrafamiliar: casos em que o desaparecimento é motivado por problemas na dinâmica familiar.
- Perda de contato voluntário: casos em que a pessoa decide voluntariamente se abster do contato com a família e/ou com outras pessoas de seu convívio.
- Ausência de notificação de óbito.
- Transtorno psíquico: casos em que o desaparecimento é motivado por sofrimento psíquico.
- Drogadição: casos em que o desaparecimento ocorreu em função do uso de drogas.
- Ausência de notificação de encarceramento: casos em que uma pessoa é presa e sua família não é notificada.
- Crime: casos em que a razão do desaparecimento foi um crime contra a pessoa desaparecida.
- Perda de referências familiares: casos em que a pessoa, ainda que sem intenção, perde o contato com os familiares.
- Subtração para exploração econômica.
- Tráfico de entorpecentes: casos em que o desaparecimento ocorreu em função do envolvimento da pessoa com o tráfico de drogas ilícitas.
- Abandono.
- Catástrofe.
- Não informado.
- Indeterminado até o momento

O estudo carece de explicação de algumas categorias e dá a entender que possui sobreposição entre elas. Por exemplo, o desaparecimento por subtração para exploração econômica também são crimes, perda de contato voluntário pode ocorrer por conflito intrafamiliar, o desaparecimento por drogadição possui casos de tráfico de entorpecentes e não há diferença clara entre não informado e indeterminado até o momento. Todavia, como os casos de sobreposição serão agrupados, isso não afeta a análise feita neste artigo. Os resultados gerais e pelo sexo do desaparecido são apresentados na Figura 7.

**Figura 7:** Prevalência do motivo de desaparecimento geral e pelo sexo do desaparecido no estado do Rio de Janeiro. Fonte: Ministério Público do estado do Rio de Janeiro, 2018.



As semelhanças entre os sistemas de classificação são apresentadas na Tabela 4. Há categorias que não possuem similaridade entre os sistemas. Por exemplo, as categorias perda de referência familiares e abandono não são diretamente comparáveis com nenhuma categoria utilizada para São Paulo. Similarmente, a categoria evadido de

instituição, do sistema 2, pode representar casos da categoria desviado, sem intenção ou determinado (no sistema 1) e transtorno psíquico, drogadição ou perda de contato voluntário (no sistema do Rio de Janeiro). Além disso, o estudo do Rio de Janeiro contempla casos de desaparecimento apenas. Assim, não há nenhum caso de erro durante o registro. Desta forma, as categorias perda de referência familiar e abandono (Rio de Janeiro), erro no registro e Evadiu de instituição (São Paulo - sistema 2) não são comparáveis entre esses estudos.

O mesmo ocorre nas categorias “ausência de notificação de óbito” e “ausência de notificação de encarceramento”, uma vez que essas são informações coletadas após a escrita do boletim de ocorrência.

**Tabela 4:** Semelhança entre as classificações do sistema 1 e 2 com as utilizadas no Rio de Janeiro.

SÃO PAULO - SISTEMA 1	SÃO PAULO - SISTEMA 2	RIO DE JANEIRO	COMPARÁVEL ENTRE OS ESTADOS?
Determinado	Saída voluntária do domicílio	Conflito intrafamiliar	SIM
		Perda de contato voluntário	SIM
Desviado	Drogas ou álcool	Drogadição	SIM
		Tráfico de entorpecentes	SIM
Sem intenção	Problemas de saúde mental	Transtorno psíquico	SIM
	Desastre natural	Catástrofe	SIM
Vítima de crime	Vítima de crime	Crime	SIM
		Subtração para exploração econômica	SIM
Motivo desconhecido	Erro no registro	Não há categoria semelhante.	NÃO
	Sem relato e Não informativo	Não informado, Indeterminado até o momento, Ausência de notificação de óbito e Ausência de notificação de encarceramento.	SIM

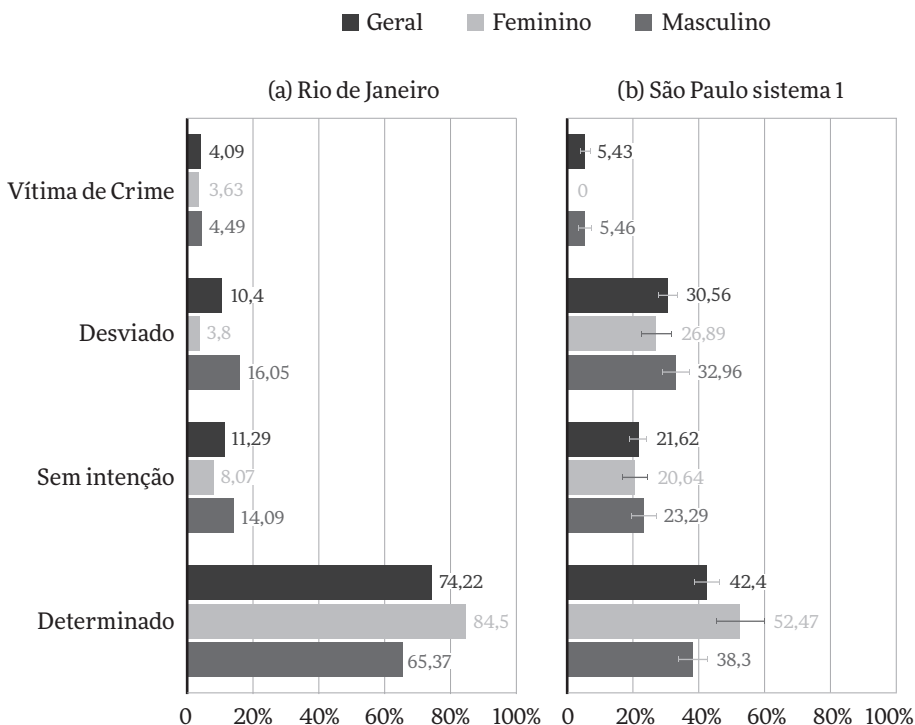
SÃO PAULO - SISTEMA 1	SÃO PAULO - SISTEMA 2	RIO DE JANEIRO	COMPARÁVEL ENTRE OS ESTADOS?
Não há categoria semelhante.	Não há categoria semelhante.	Perda de referência familiar	NÃO
Não há categoria semelhante.	Não há categoria semelhante.	Abandono	NÃO
Casos de Determinado, Desviado e Sem intenção.	Evadiu de instituição	Casos de Drogadição, Perda de contato voluntário e Transtorno psíquico.	NÃO

As porcentagens de casos com motivo desconhecido é 67,85% no Rio de Janeiro (somando os motivos indeterminados até o momento e não informado), muito maior que os valores encontrados para São Paulo: 39,92% e 40,67%, segundo os sistemas 1 e 2, respectivamente. Este fato é surpreendente, uma vez que o estudo do Rio de Janeiro se restringiu a casos já elucidados.

Com a finalidade de tornar a comparação das causas de desaparecimento entre os estados mais direta, a Figura 8 apresenta a prevalência dos motivos de desaparecimento dentre os casos cuja causa é conhecida e comparável entre os estados. Para isso, optou-se por não utilizar o sistema 2 por se tratar de categorias mais gerais e utilizou-se da Tabela 4 para converter as categorias do Rio de Janeiro para o sistema 1 de São Paulo.

Entre os motivos conhecidos, a prevalência do desaparecimento determinado é maior no Rio de Janeiro e, em ambos os estados, mais comuns em mulheres. Por outro lado, em São Paulo é maior a porcentagem de indivíduos que desaparecem por motivos relacionados a drogas ou álcool (desviado), quando comparado ao Rio de Janeiro. Em ambos os estados homens desaparecem mais do que mulheres devido a esse motivo. As causas de desaparecimento sem intenção são maiores em São Paulo e, apenas no Rio de Janeiro, são mais comuns entre os homens. A ocorrência do desaparecimento por crime é semelhante em ambos os estados e não possui diferença significativa entre os sexos.

**Figura 8:** Prevalência dos motivos de desaparecimento geral e pelo sexo do desaparecido (utilizando sistema 1) dentre os casos cuja causa é conhecida e comparável entre os estados para o Rio de Janeiro (a) e São Paulo (b).



## CONCLUSÕES

Em todos os bancos de dados analisados observou-se que, em uma parcela substancial dos boletins de ocorrência, não havia informação suficiente para inferir a causa do desaparecimento. Dentre estes casos, o boletim de ocorrência muitas vezes era composto por apenas um parágrafo. Como não é possível inferir o motivo de desaparecimento em uma parcela substancial de casos tanto em São Paulo quanto no Rio de Janeiro, as conclusões deste trabalho devem ser tomadas com cautela. Por exemplo, talvez o fato de não ser possível identificar o motivo de um desaparecimento com base

em um relato esteja associado ao motivo em si, de modo que as prevalências estimadas podem ser viesadas.

Esta constatação indica a importância de que sejam adotados critérios para a uniformização de boletins de ocorrência. Por exemplo, as tipologias adotadas para classificar desaparecimentos poderiam ser adotadas e utilizadas explicitamente. Também poderiam ser criadas diretrizes sobre o conteúdo mínimo que deve existir em boletins de ocorrência. Tais medidas permitiriam uma melhor compreensão do fenômeno do desaparecimento, bem como o delineamento de estratégias mais eficazes para lidar com estes casos.

Apesar da existência de muitos boletins de ocorrência com poucas informações, os métodos utilizados permitiram obter um panorama geral das causas de desaparecimento em São Paulo. Uma vez determinadas as principais causas de desaparecimento, é possível delinear diretrizes para as estratégias utilizadas pelos investigadores em cada um dos casos.

A causa mais frequente é aquela em que o indivíduo desaparece intencionalmente. Esta decisão é tomada mais por mulheres do que por homens e é especialmente comum entre indivíduos menores de 18 anos. Tal composição é similar, por exemplo, àquela encontrada por JAMES, ANDERSON e PUTT (2008) para desaparecidos na Austrália. Dentro da categoria de menores desaparecidos intencionalmente, estudos como FINKELHOR, SEDLAK e HOTALING (1990) e POSNER (1992) indicam que é comum que a criança ou adolescente desaparecido tenha um núcleo familiar disfuncional, com taxas mais elevadas de depressão, abuso sexual, e problemas escolares. Assim, para efetivamente lidar com o caso, frequentemente não é suficiente encontrar o desaparecido e retorná-lo à sua família.

Uma outra categoria frequente é a de desaparecidos por uso de drogas ou álcool, ultrapassando 20% dos casos analisados. Tal porcentagem é substancialmente maior do que os 10% encontrados por HENDERSON e HENDERSON (1998) na Austrália. Esta discrepância pode ser explicada, por exemplo, pelas diferenças sociais entre o Brasil e a Austrália ou pela dificuldade de separar a drogadição de doenças relacionadas à saúde mental. A causa de desaparecimento por uso de drogas ou álcool revelou-se mais comum em São Paulo entre menores de 50 anos.

Para indivíduos maiores de 66 anos, a causa mais comum para o desaparecimento é a não-intencional. A leitura dos boletins de ocorrência revelou que estes casos se compõem principalmente de situações em que o desaparecido sofria de problemas de esquecimento por causa da idade avançada. Tal caracterização é similar àquela encontrada em outros lugares do mundo (HENDERSON e HENDERSON, 1998; BIEHAL, MITCHELL e WADE, 2003).

Finalmente, encontramos que cerca de 5% dos indivíduos são vítimas de crimes, sendo está a causa com a menor das prevalências em São Paulo.

A caracterização das causas de desaparecimento em São Paulo encontrada ressalta que este é um fenômeno social complexo, sendo frequentemente insuficiente apenas encontrar o desaparecido. O desaparecimento é comumente consequência de outros males que recaem sobre o desaparecido em sua vida cotidiana. Assim, a não ser que o desaparecido receba assistência em relação a estas causas, é comum que o fenômeno volte a ocorrer (STEVENSON et. al, 2013). O entendimento das principais causas de desaparecimento pode ajudar a estabelecer estratégias eficazes. Além disso, as categorias encontradas permitem ao investigador utilizar dados demográficos do desaparecido para levantar hipóteses razoáveis sobre a causa do desaparecimento quando está ainda é incerta.

*Recebido: 23 de outubro de 2020.*

*Aprovado: 09 de dezembro de 2020.*



## REFERÊNCIA BIBLIOGRÁFICAS

- BELLA, A. et al. Quantification via probability estimators. 2010 IEEE International Conference on Data Mining. p. 737-742, 2010.
- BIEHAL, N.; et al. Lost from view: Missing persons in the UK. Policy Press, 2003.
- CHOLLET, F. *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. MITP-Verlags GmbH & Co. KG, 2018.
- FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA- FBSP. *Anuário Brasileiro de Segurança Pública 2019*. São Paulo, 2019.
- FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA- FBSP. *Anuário Brasileiro de Segurança Pública 2018*. São Paulo, 2018.
- FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA- FBSP. *Anuário Brasileiro de Segurança Pública 2017*. São Paulo, 2017.
- COMITÊ INTERNACIONAL DA CRUZ VERMELHA. *Relatório da Avaliação das Necessidades dos Familiares de Pessoas Desaparecidas durante o Regime Militar no Brasil: "Caso Vala de Perus"*, 2017.
- FINKELHOR D.; SEDLAK A.; HOTALING G.T. *Missing, abducted, runaway, and throwaway children in America: First report, numbers and characteristics national incidence studies: Executive summary*. DIANE Publishing; 1990.
- FORMAN, G. et al. Quantifying counts and costs via classification. *Data Mining and Knowledge Discovery*, v. 17, n. 2, p. 164-206, 2008.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. *The elements of statistical learning*. New York: Springer series in statistics, 2001.
- HENDERSON, M.; HENDERSON, P. *Missing people: Issues for the Australian community*. Canberra: Australian Bureau of Criminal Intelligence, 1998.
- IZBICKI, R.; SANTOS, T. M. *Aprendizado de máquina: uma abordagem estatística*. 2020.
- IZBICKI, Rafael et al. Converting high-dimensional regression to high-dimensional conditional density estimation. *Electronic Journal of Statistics*, v. 11, n. 2, p. 2800-2831, 2017.
- JAMES, G. et al. *An introduction to statistical learning*. New York: springer, 2013.
- JAMES, M.P.; ANDERSON J.; PUTT J.; Australian Institute of Criminology. *Missing persons in Australia*. Canberra: Australian Institute of Criminology, 2008.
- MINISTÉRIO PÚBLICO DO ESTADO DO RIO DE JANEIRO. *O desaparecimento nas burocracias do estado*, 2018.
- OECHSLER, D. M. *Crianças desaparecidas: análise dos registros de ocorrências no Estado de Santa Catarina*. 2018. Dissertação (Mestrado em Criminologia) - Faculdade de Ciências Sociais e Humanas, Universidade Fernando Pessoa. Porto, 2018.
- PEDREGOSA, Fabian et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, v. 12, p. 2825-2830, 2011.
- POLIANO, F. et al. *Perfil de pessoas desaparecidas no Estado de São Paulo*, 2016.

- POSNER M. *The runaway risk reduction project assessment report*. Newton, MA: Education Development Center; 1992.
- RIBEIRO, M. et al. *Model-agnostic interpretability of machine learning*. arXiv preprint arXiv:1606.05386, 2016.
- SAERENS, M. et al. Adjusting the outputs of a classifier to new a priori probabilities: a simple procedure. *Neural computation*, v. 14, n. 1, p. 21-41, 2002.
- SALTON, G. et al. A vector space model for automatic indexing. *Communications of the ACM*, v. 18, n. 11, p. 613-620, 1975.
- SRIVASTAVA, A. N.; SAHAMI, M. *Text mining: Classification, clustering, and applications*. CRC Press, 2009.
- STEVENSON, O., et. al. *Geographies of Missing People: Processes, Experiences, Responses, Online*, 2013.
- VAZ, A. et al. Quantification Under Prior Probability Shift: the Ratio Estimator and its Extensions. *J. Mach. Learn. Res.*, v. 20, p. 79:1-79:33, 2019.

## APÊNDICE

### A.1

Antes de usar métodos de aprendizado de máquina, é necessário pré-processar os boletins de ocorrência, que são dados não estruturados. Para isso, utilizamos a representação Modelo Espaço Vetorial (Vector Space Model - VSM) para documentos, introduzida por Salton et al (1975). O objetivo do método é converter o texto não estruturado em recursos estruturados que podemos usar em tarefas de aprendizado de máquina.

A representação do documento como um vetor foi feita usando bag-of-words (SRI-VASTAVA & SAHAMI, 2009). Os textos são transformados em uma matriz documento-termo, onde cada coluna é um termo presente no documento e cada linha é um documento (no caso, um boletim de ocorrência). A matriz é preenchida pela frequência de cada termo em cada boletim de ocorrência ou um valor binário (com 1 indicando que o termo ocorreu no documento e 0 indicando que não ocorreu).

A matriz documento-termo dos dados classificados manualmente foi então usada para treinar modelos de aprendizado de máquina. De modo geral, o objetivo de um modelo de classificação é determinar uma relação entre uma variável aleatória qualitativa, como as causas de desaparecimento do estado de São Paulo, e covariáveis, como a idade, o sexo, o relato, entre outras informações presentes nos boletins de ocorrência.

O principal objetivo deste trabalho é estimar a proporção de cada um dos motivos de desaparecimento no estado de São Paulo. Assim, os classificadores são utilizados apenas como uma forma de inferir os rótulos dos dados não classificados manualmente. Estes rótulos são então usados para estimar as proporções de interesse. Formalmente, o objetivo principal deste artigo é resolver um problema de quantificação (FORMAN, 2008, TASCHE, 2017).

Utilizamos aqui o método desenvolvido por Vaz et al. (2019) para executar essa correção. Este método combina o estimador rotulado (proporções de cada causa nos dados classificados manualmente) com estimador proporção (proporções de cada causa nos dados classificados via aprendizado de máquina, após fazer as devidas correções).

A fim de estudar como a causa varia por idade e sexo, foi utilizado o método de regressão, introduzido por Vaz et al. (2019).

## A.2

As tabelas abaixo mostram os percentuais de incerteza atribuídos a cada categoria ou combinações de categoria das variáveis analisadas.

**Tabela 5:** Porcentagem e Incerteza das causas do desaparecimento usando o sistema 2

CAUSA	PORCENTAGEM (%)
Desastre natural	0,58 ± 0,47
Vítima de crime	2,62 ± 0,99
Problemas de saúde mental	13,47 ± 1,60
Drogas ou álcool	19,49 ± 1,73
Erro no registro	3,91 ± 1,09
Evadiu de instituição	6,16 ± 1,08
Sem relato	4,74 ± 0,00
Não informativo	35,93 ± 2,56
Saída voluntária do domicílio	13,1 ± 2,06

**Tabela 6:** Porcentagem e Incerteza (%) das causas do desaparecimento pelo sexo dos indivíduos usando o sistema 2.

CAUSA	FEMININO	MASCULINO
Desastre natural	0,00 ± 0,00	0,84 ± 0,75
Vítima de crime	0,00 ± 0,00	2,70 ± 1,33
Problemas de saúde mental	12,87 ± 2,17	14,32 ± 2,28
Drogas ou álcool	16,65 ± 2,28	21,57 ± 2,51
Erro no registro	4,39 ± 1,70	3,63 ± 1,42
Evadiu de instituição	5,72 ± 1,50	6,57 ± 1,52
Sem relato	4,69 ± 0,00	4,83 ± 0,00
Não informativo	33,31 ± 3,94	38,01 ± 3,38
Saída voluntária do domicílio	22,37 ± 3,85	7,53 ± 2,15

**Tabela 7:** Porcentagem e Incerteza (%) das causas do desaparecimento pelo faixa de idade dos indivíduos usando o sistema 2.

CAUSA	0-2	3-10	11-18	19-35	36-50	51-65	66+
Desastre natural	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	1,57 ± 3,07	0,00 ± 0,00
Vítima de crime	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	3,72 ± 2,76	0,00 ± 0,00	0,00 ± 0,00
Problemas de saúde mental	26 ± 18,76	0,00 ± 0,00	8,8 ± 1,28	16,33 ± 3,27	16,2 ± 4,75	14,89 ± 7,43	47,03 ± 19,51
Drogas ou álcool	19,33 ± 0,00	14,02 ± 9,42	17,34 ± 2,38	22,53 ± 3,76	24,46 ± 4,06	15,4 ± 7,86	3,95 ± 7,75
Erro no registro	1,76 ± 0,00	0,00 ± 0,00	4,86 ± 1,89	3,86 ± 2,09	3,1 ± 2,40	5,39 ± 4,31	0,00 ± 0,00
Evadiu de instituição	14,8 ± 18,19	3,92 ± 7,70	6,26 ± 1,65	4,9 ± 1,50	8,23 ± 2,34	7,46 ± 6,22	4,06 ± 7,75
Sem relato	0,00 ± 0,00	4,65 ± 0,00	4,78 ± 0,00	5,14 ± 0,00	4,51 ± 0,00	4,53 ± 0,00	5,23 ± 0,00
Não informativo	17,93 ± 15,92	52,88 ± 20,43	33,79 ± 4,11	41,58 ± 4,81	34,22 ± 5,94	46,12 ± 9,23	31,55 ± 18,60
Saída voluntária do domicílio	20,18 ± 21,24	24,54 ± 17,09	24,17 ± 4,10	5,66 ± 2,80	5,56 ± 3,32	4,64 ± 5,24	8,18 ± 10,73

**Tabela 8:** Porcentagem e Incerteza das causas do desaparecimento usando o sistema 1.

CAUSA	PORCENTAGEM (%)
Determinado	25,46 ± 2,41
Desviado	18,35 ± 1,88
Vítima de crime	3,26 ± 1,01
Motivo desconhecido	39,95 ± 2,70
Sem intenção	12,98 ± 1,69

**Tabela 9:** Porcentagem e Incerteza das causas do desaparecimento pelo sexo dos indivíduos usando o sistema 1.

CAUSA	FEMININO	MASCULINO
Desviado	0±	0,8±
Determinado	31,01 ± 4,37	23,01 ± 2,82
Desviado	15,89 ± 2,79	19,80 ± 2,59
Vítima de crime	0,00 ± 0,00	3,28 ± 1,33
Motivo desconhecido	40,90 ± 4,54	39,92 ± 3,41
Sem intenção	12,20 ± 2,41	13,99 ± 2,38

**Tabela 10:** Porcentagem e Incerteza das causas do desaparecimento pela faixa de idade dos indivíduos usando o sistema 1.

CAUSA	0-2	3-10	11-18	19-35	36-50	51-65	66+
Desviado	0,0	0,0	0,0	0,0	0,0	1,6	0,0
Determinado	45,96 ± 27,59	28,12 ± 17,15	37,47 ± 4,60	16,86 ± 3,37	19,43 ± 4,54	19,96 ± 8,44	11,77 ± 12,65
Desviado	0,00 ± 0,00	13,39 ± 9,71	14,47 ± 2,81	21,13 ± 3,92	22,43 ± 4,10	20,87 ± 8,75	4,00 ± 7,87
Vítima de crime	0,00 ± 0,00	0,00 ± 0,00	0,00 ± 0,00	3,91 ± 1,78	3,87 ± 2,87	0,00 ± 0,00	0,00 ± 0,00
Motivo desconhecido	35,35 ± 23,39	58,49 ± 21,30	41,15 ± 4,57	42,96 ± 4,90	37,52 ± 6,06	43,18 ± 10,78	35,52 ± 19,32
Sem intenção	18,70 ± 15,55	0,00 ± 0,00	6,92 ± 1,55	15,14 ± 3,24	16,75 ± 5,07	15,99 ± 8,66	48,70 ± 19,59

**Tabela 11:** Prevalência do motivo de desaparecimento geral e pelo sexo do desaparecido no estado do Rio de Janeiro. Fonte: Ministério Público do estado do Rio de Janeiro, 2018.

MOTIVO DE DESAPARECIMENTO	TOTAL (%)	FEMININO (%)	MASCULINO (%)
Indeterminado até o momento	66,57	62,61	68,83
Conflito intrafamiliar	14,07	21,53	9,82
Perda de contato voluntário	6,05	7,59	5,17
Ausência de notificação de óbito	3,41	1,19	4,68
Transtorno psíquico	3,02	2,78	3,17
Drogadição	2,72	1,25	3,55
Não informado	1,28	1,36	1,23
Ausência de notificação de encarceramento	1,11	0,06	1,71
Crime	0,95	1,08	0,87
Perda de referência familiares	0,47	0,28	0,58
Subtração para exploração econômica	0,16	0,17	0,16
Tráfico de entorpecentes	0,10	0,06	0,13
Abandono	0,04	0,06	0,03
Catástrofe	0,04	0,00	0,06
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>

**Tabela 12:** Prevalência do motivo de desaparecimento geral e pelo sexo do desaparecido no estado do Rio de Janeiro, após transformar para o sistema 1.

MOTIVO DE DESAPARECIMENTO	TOTAL (%)	FEMININO (%)	MASCULINO (%)
Vítima de crime	4,09	3,63	4,49
Desviado	10,40	3,80	16,05
Sem intenção	11,29	8,07	14,09
Determinado	74,22	84,50	65,37
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>

**Tabela 13:** Prevalência do motivo de desaparecimento geral e pelo sexo do desaparecido no estado de São Paulo segundo o sistema 1, considerando apenas as categorias que são de fato causas de desaparecimento (isto é, removendo a classe Sem informação).

<b>MOTIVO DE DESAPARECIMENTO</b>	<b>TOTAL (%)</b>	<b>FEMININO (%)</b>	<b>MASCULINO (%)</b>
Determinado	42,4 ± 4,01	52,47 ± 7,39	38,3 ± 4,69
Desviado	30,56 ± 3,13	26,89 ± 4,72	32,96 ± 4,31
Vítima de crime	5,43 ± 1,68	0,00 ± 0,00	5,46 ± 2,21
Sem intenção	21,62 ± 2,81	20,64 ± 4,08	23,29 ± 3,96
<b>Total</b>	<b>100</b>	<b>100</b>	<b>100</b>