

---

## **ANÁLISE DOS BOLETINS DE OCORRÊNCIAS: um estudo sobre a Criminalidade na Baixada Santista à luz da Ciência de Dados**

*ANALYSIS OF INCIDENT REPORTS: a study on Crime in Baixada Santista in the light of Data Science*

Carlos Eduardo França Amador  
(Fatec Baixada Santista - Rubens Lara)  
carlos.amador@fatec.sp.gov.br

Jefferson Leonardo dos Santos Sepulveda  
(Fatec Baixada Santista - Rubens Lara)  
jefferson.sepulveda@fatec.sp.gov.br

Wesley Vinícius Rodrigues da Silva  
(Fatec Baixada Santista - Rubens Lara)  
wesley.silva21@fatec.sp.gov.br

### **RESUMO**

O artigo propõe uma análise exploratória da situação da criminalidade na Baixada Santista, considerando estratégias de combate implementadas e suas implicações sociais e econômicas. Utilizando dados da Secretaria de Segurança Pública de São Paulo (SSP-SP), busca-se compreender a evolução do cenário de segurança na região de 2022 a 2024. O estudo visa em proporcionar recursos informativos para subsidiar políticas públicas e auxiliar na tomada de decisão a fim de orientar ações mais assertivas visando promover a segurança e o bem-estar dos residentes e visitantes na Região da Baixada Santista.

**Palavras-chave:** Ciência de dados, Análise de Dados; Baixada Santista; SSP-SP; criminalidade.

### **ABSTRACT**

*The article proposes an exploratory analysis of the crime situation in Baixada Santista, considering innovative combat strategies and their social and economic implications. Using data from the São Paulo Public Security Secretariat (SSP-SP), it seeks to understand the evolution of the security scenario in the region from 2022 to 2024. The study aims to provide informative resources for public policies and assertive actions to promote security and safety. well-being of residents and visitors.*

**KEYWORDS:** Crime; Data Science; Data Analytics; Baixada Santista; SSP-SP.

---

## 1 INTRODUÇÃO

Os Objetivos de Desenvolvimento Sustentável (ODS) são um plano global em ação adotada pela Organização das Nações Unidas (ONU) para acabar com a pobreza, proteger o meio ambiente e o clima e garantir que as pessoas, em todos os lugares, possam desfrutar de paz e de prosperidade. Estes são os objetivos para os quais a ONU faz estudos e publicações dessas causas e que através de dados adquiridos orientam as ações governamentais a fim de que atinjam resultados positivos até 2030.

Tratando-se do Brasil, um dos grandes desafios para que a maioria das cidades alcancem o Objetivo de Desenvolvimento Sustentável (ODS) de número 16, que preconiza “promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos e construir instituições eficazes, responsáveis e inclusivas em todos os níveis”, alinhando-se com o item 16.1 da ODS, que visa “a redução significativa de todas as formas de violência e taxas de mortalidade relacionadas em todos os lugares”, tem como objetivo, reduzir o índice de criminalidade na Região. (A ESTRATÉGIA ODS, 2024)

A Baixada Santista, região Litorânea do Estado de São Paulo, é composta por nove municípios como Cubatão, Guarujá, Santos, Bertioga, São Vicente, Praia Grande, Mongaguá, Itanhaém e Peruíbe. As características econômicas de cada cidade são bem variadas. Essas regionalidades incluem desde o Polo industrial agregando setores petroquímicos, siderúrgicos e químicos, o Complexo Portuário, considerado o maior Porto da América Latina, demandando o armazenamento e transporte de cargas de âmbito rodoviário, ferroviário e aquaviário, até a forte área do turismo, presente em todas as cidades do Litoral Paulista que atraem uma quantidade expressiva de turistas aumentando o fluxo de pessoas na região. (ESTADO DE SÃO PAULO, 2020)

Com toda essa amplitude de atividades na Região, é relevante buscar entender, com uma visão abrangente, a dinâmica de segurança pública, pois o índice de criminalidade na área tem despertado atenção tanto da comunidade acadêmica, quanto das autoridades governamentais, motivando estudos e análises sobre suas causas e consequências.

De acordo com a extração e a manipulação da base de dados de boletins de ocorrências da Baixada Santista, em 2022 foram registradas 54.523 ocorrências, em 2023 foram 55.021, e nos três primeiros meses de 2024 foram registradas 13.994 ocorrências. Estes números destacam a urgência de medidas eficazes para lidar com a questão da segurança pública na região." (ESTADO DE SÃO PAULO, 2024)

Sendo assim, este artigo propõe um estudo exploratório sobre a situação da criminalidade na Baixada Santista, além de detalhar as implicações sociais e econômicas resultantes dessas ações.

Compreender esses aspectos é essencial para subsidiar políticas públicas e ações assertivas, que promovam a segurança e o bem-estar dos residentes e dos visitantes que frequentam a região, utilizando dados obtidos do Portal de Transparência da Secretaria de Segurança Pública do Estado de São Paulo, pretende-se analisar a evolução do cenário de segurança na Baixada Santista no período de 2022 a março de 2024.

Como objetivos específicos, inicialmente serão apresentadas algumas características da região da Baixada Santista e descrevendo peculiaridades de cada cidade quanto à criminalidade. A seguir, a partir da base de dados Portal da Transparência do Estado de São Paulo (SSP-SP), os dados serão devidamente tratados e inseridos na ferramenta Power BI para a análise; por fim, haverá a elaboração de *dashboards* interativos para auxiliarem na apresentação das informações relevantes apuradas após análise dos resultados deste trabalho.

## 2 REFERENCIAL TEÓRICO

Neste tópico, são apresentados os principais conceitos da área de Segurança Pública relativos à criminalidade. Estudar-se-ão características da segurança pública, violência, classes de crimes, objetivando identificar aspectos relativos à região da Baixada Santista para embasar a análise exploratória dos dados de cada cidade. Do mesmo modo, foi explorado o campo da Ciência de Dados, ferramentas e plataformas disponíveis e a suas aplicações dentro do contexto do tema da pesquisa.

### 2.1 CONCEITOS DE VIOLÊNCIA

A expressão violência pode ser compreendida e abordada sob várias perspectivas no âmbito da filosofia e das ciências sociais e não há dúvida de que, no Brasil contemporâneo, ela é portadora de um núcleo de sentido claramente determinado como: a violação da lei penal (o cometimento de condutas definidas legalmente como crimes), que atravessa sistematicamente o cotidiano da maioria das nossas cidades, trazendo implicações negativas (lesões ou ameaças de lesão) para bens como a vida, a integridade física e o patrimônio de indivíduos (vítimas) (SILVA, 2018).

O aumento de episódios de violência e seu desdobramento para áreas antes razoavelmente protegidas, que vêm se acumulando nas últimas décadas, transformou-o em um dos principais problemas da agenda pública, feito a partir da identificação de ameaças à integridade pessoal e patrimonial passou a dominar amplas camadas da população.

Além de aumentarem em número, os crimes agora contemplariam modalidades e meios cada vez mais violentos como roubos, sequestros e homicídios, em geral acompanhados pelo uso de armas de fogo de grosso calibre, bem como maior grau de complexidade, atrelando-se, por exemplo, a quadrilhas, organizações ou facções criminosas. Como consequência, emerge entre as famílias do país uma sensação de cada vez mais inseguras. (CALDEIRA, 2000).

### 2.2 CONCEITO DE CRIME E SUA CLASSIFICAÇÃO

O crime trata-se de uma conduta humana que infringe as normas estabelecidas pelo ordenamento jurídico penal. No ramo penal, o crime é estabelecido como uma ação ou omissão que contraria uma norma jurídica penal e é punível com sanção criminal. Em resumo, trata-se de um ato proibido por lei que pode resultar em punição, como uma prisão, multa ou ambas, alternativa ou cumulativa. (BRASIL, 1941)

A definição de crime varia de acordo com os códigos penais de cada país, mas em geral, um crime é uma atitude que infringe a lei de uma maneira que prejudica ou ameaça os interesses da sociedade.

Embora no Brasil há uma vasta quantidade de crimes estipulados por lei, neste artigo abordar-se-ão somente alguns deles relacionados às explorações feitas através de dados extraídos pelo portal de transparência “Números sem Mistérios”, portal esse da Secretaria de Segurança Pública de São Paulo (SSP-SP).

No presente trabalho, serão abordados dois tipos específicos: crimes contra a pessoa e crimes contra o patrimônio.

- a) **CRIME CONTRA A PESSOA:** Arruda (2023) afirma que "Os crimes contra a pessoa são infrações penais que atentam contra a integridade física, a vida e a liberdade individual". Sendo eles Homicídio, Lesão corporal e Estupro.
- b) **CRIMES CONTRA O PATRIMÔNIO:** O Código Penal, em sua Parte Especial, Título II, aborda os crimes contra o patrimônio. O principal bem jurídico protegido é o patrimônio; no entanto, em alguns crimes, como o roubo, também se tutelam a integridade física, mental e a vida da vítima, sendo eles Furto, Roubo, Extorsão, Latrocínio. (BRASIL, 1940)

## 2.3 PERCEPÇÃO DE SEGURANÇA NAS CIDADES

A percepção de segurança refere-se à sensação que as pessoas têm em relação ao ambiente em que vivem, não necessariamente relacionada à presença real de perigos, mas sim à sua segurança pessoal e ao local onde residem. Essa percepção pode ser medida não apenas por dados objetivos, mas também pela confiança e sensação subjetiva que as pessoas têm em relação à segurança em sua comunidade.

A percepção de segurança tem sido diretamente influenciada pela qualidade dos serviços públicos disponíveis, como iluminação, pavimentação, áreas de lazer e, notavelmente, o policiamento. Um estudo conduzido em 2021 pelo IBGE destacou que a avaliação positiva destes serviços está vinculada a um aumento significativo na sensação de segurança entre os moradores. Por outro lado, crimes como extorsão, tráfico de armas, roubos e tiroteios reduzem drasticamente a sensação de segurança, enquanto o consumo e venda de drogas ilegais, embora com um impacto menor, também afetam negativamente a percepção de segurança (IBGE, 2021).

O nascimento das grandes concentrações urbanas trouxe uma mudança radical do nível de insegurança. Se a cidade murada medieval era um perigo fora dos muros, nas cidades modernas o perigo reside na própria cidade. Nesse sentido, o surgimento das grandes cidades instalou o medo dentro da cidade. A multidão passou a ser vista como uma ameaça em potencial. Assim, paulatinamente o controle do crime deixa de ser um controle social informal, feito pela própria sociedade, para se tornar um controle formal, levado a efeito pela polícia, tribunais e prisões (MOLERO, 2002).

Com essa ideia de mudança de hábitos cotidianos e exacerbação de conflitos sociais em mente disseminou-se o “medo” na sociedade. Entende-se por medo como um estado de retraimento, individual ou coletivo, diante de situações que representem lesões ou perigo de lesões a bens jurídicos importantes, tais como vida, integridade

física e patrimônio. Esse medo se apresenta como algo capaz de influenciar e modificar o comportamento humano tanto individualmente quanto em sociedade.

Entretanto, o ato de sentir medo também acaba despertando uma maior sensação de cuidado, assim aumentando a sua segurança e proteção, porém vivendo em sociedade quando o medo se relaciona, principalmente à violência, acaba gerando desconforto social, levando o sentimento que deveria causar sua prevenção a provocar um desconforto na vida das pessoas.

## 2.4 ASPECTOS DA SEGURANÇA PÚBLICA DE RESPONSABILIDADE DO ESTADO

A segurança pública, dever do Estado, direito e responsabilidade de todos, é exercida para a preservação da ordem pública e da incolumidade das pessoas e do patrimônio, sob a égide dos valores da cidadania e dos direitos humanos, através dos órgãos instituídos pela União e pelos Estados. previsto no art. 144 da CF. (BRASIL, 1988).

De acordo com COSTA (2014), atualmente, uma das maiores preocupações tanto populacional quanto governamental é a segurança pública, tanto pelo lado do aumento de crimes, como crimes contra a vida e ao patrimônio, tráfico de entorpecentes, corrupção e crime organizado. A segurança pública enfrenta desafios que vão além da criminalidade, como a desigualdade social, o desemprego e a precariedade dos serviços públicos, que contribuem para a complexidade do problema. Isso demonstra que o problema da segurança pública não está apenas centralizado à criminalidade e sim a outros fatores: políticos, sociais, econômicos, éticos, culturais, administrativos e históricos.

Ainda segundo o autor, o organograma do aparelho estatal encontra-se estruturado de forma a distribuir responsabilidades no exercício de suas atribuições: os vinculados à União, crimes federais, nas contravenções e ilícitos penais no âmbito dos Estados, polícias estaduais, bem como a complementação do poder de polícia dos Estados através da segurança privada em particular, quanto aos crimes contra o patrimônio e as guardas municipais no âmbito dos Municípios, particularmente quanto aos atos antissociais.

Exercendo em uma democracia íntegra, a missão da Segurança Pública é garantir a proteção dos direitos individuais e assegurar o pleno exercício da população trazendo qualidade de vida ao seu cotidiano. Sendo assim passou a ser o principal desafio ao de estado de direito no Brasil, ganhando maior visibilidade pública e está presente nos debates tanto de especialistas, como do público em geral. (VEDOVA, 2018).

## 2.5 INDICADORES DE SEGURANÇA PÚBLICA

De acordo com Costa (2011), a segurança pública pode ser avaliada por meio de uma variedade de indicadores e métodos. Entre eles aqui estão alguns dos aspectos considerados pelos Órgãos Responsáveis para aferir a segurança pública:

- a) TAXAS DE CRIMINALIDADE: Análise metódica de dados estatísticos relacionados aos delitos perpetrados em uma localidade específica,

abrangendo categorias como homicídios, roubos, furtos, entre outros. A avaliação dessas taxas é crucial para aferir a frequência e gravidade dos crimes, bem como para avaliar a efetividade das intervenções e políticas de segurança pública implementadas.

- b) **EFICIÊNCIA OPERACIONAL:** A avaliação da eficiência das operações policiais compreende diversos aspectos, tais como o tempo de resposta a ocorrências, medido tanto pela instituição policial quanto pela percepção da população, a resolução de casos, e o uso adequado de recursos, incluindo viaturas, câmeras e outros equipamentos auxiliares.

Essa vertente dá mensuração de segurança pública não será objeto desta pesquisa, portanto não será detalhado neste artigo.

- c) **AVALIAÇÃO DE PROGRAMAS E POLÍTICAS:** É crucial realizar uma análise dos resultados de programas e políticas públicas implementadas na área de segurança, a fim de verificar se alcançaram os objetivos propostos e se tiveram impacto positivo na redução da criminalidade. Um exemplo emblemático é a 'Operação Verão' na Baixada Santista, que é executada durante períodos de aumento exponencial no fluxo de pessoas na região.

Essa vertente da mensuração de segurança pública não será objeto desta pesquisa, portanto não será detalhado neste artigo.

## **2.6 - EXPLORANDO A CIÊNCIA DE DADOS: UMA ABORDAGEM PRÁTICA COM FERRAMENTAS ESSENCIAIS**

AWARI (2023) relata que a Ciência de Dados é um campo multidisciplinar que utiliza matemática, estatística, programação especializada, e conhecimento em áreas específicas para extrair insights valiosos dos dados. Este campo é essencial para a tomada de decisões e planejamento estratégico, especialmente com o aumento exponencial do volume de dados. O processo de Estudo Técnico Preliminar (ETP) é crucial na ciência de dados, pois auxilia a identificar e analisar as necessidades antes de iniciar projetos, garantindo que as soluções propostas sejam viáveis e assertivas.

O processo de ETL (*Extract, Transform, Load*) neste artigo científico focado em segurança pública e percepção de segurança é crucial. Conforme descrito em um material fornecido pela plataforma *Databricks* (2024), o processo de ETL é caracterizado pela extração, transformação, carga de dados e desempenha um papel fundamental na preparação de conjuntos de dados confiáveis e utilizáveis para análises. Essa etapa crítica permite que os dados sejam posteriormente explorados em aplicações *downstream*, proporcionando insights valiosos para a resolução de problemas.



- a) **EXCEL:** O Microsoft Excel é uma ferramenta de planilha eletrônica lançada pela primeira vez em novembro de 1985 para computadores *Macintosh*, e sua primeira versão para sistemas Windows, a 2.0, foi lançada em 1987.

Desde então, o Microsoft Excel tornou-se uma das ferramentas mais populares para análise e manipulação de dados em formato de planilha. Sua ampla gama de recursos e sua interface familiar fizeram dele uma escolha preferida para uma variedade de usuários, desde estudantes até profissionais de negócios e cientistas de dados. A contínua evolução do Excel ao longo dos anos tem garantido sua relevância e utilidade em diversas áreas.

- b) **C:** A linguagem de programação C, desenvolvida por Dennis Ritchie no Bell Labs em 1972, para o uso no Sistema Operacional Unix, é uma linguagem de propósito geral destacada por sua eficiência e controle de baixo nível sobre o hardware. Utilizada amplamente em sistemas operacionais, software de sistemas e aplicações de alto desempenho, o C oferece características como manipulação direta de memória, portabilidade, sintaxe simples, uma robusta biblioteca padrão. Estas características oferecem o máximo em controle e eficiência, facilitando o aprendizado para programadores iniciantes.

Desde sua criação, o C tornou-se uma linguagem popular tanto entre profissionais quanto iniciantes, sendo amplamente utilizada no meio acadêmico para desenvolvimento de pesquisas científicas e como instrumento de aprendizado para o desenvolvimento de algoritmos.

Para exemplificar o quanto a linguagem C é difundida, podemos citar alguns softwares escritos em C: *Matlab*, *Windows*, *Kernel* do *Linux*, *Nedit* (editor de texto), *Octave* (similar do *Matlab* em código aberto), o servidor de vídeo *RI*, *Skype* (escrito em C/C++) e *Network Simulator* (NS2). (de Sá, 2015)

Além disso, muitas das linguagens de programação líderes atuais, como C#, *Java*, *JavaScript*, *PHP* e *Python*, derivam do C, destacando sua influência duradoura na computação. (ATRA, 2024)

- c) **POWER BI:** O *Power BI* é uma ferramenta poderosa de análise de dados desenvolvida pela Microsoft. Ele oferece uma variedade de serviços de software, aplicativos e conectores que trabalham juntos para transformar dados em visualizações claras e fáceis de entender. Algumas das fontes de dados suportadas pelo *Power BI* incluem o Excel e uma variedade de data “warehouses” híbridos locais ou baseados na nuvem.

Com o *Power BI*, os usuários podem criar *dashboards* interativos que tem como definição painéis de informações, indicadores, métricas importantes para o funcionamento de uma atividade, relatórios dinâmicos e análises avançadas de forma visual em relação a gráficos e tabelas.

Uma das principais vantagens do *Power BI* é sua capacidade de fornecer insights em tempo real, permitindo que os usuários visualizem e analisem

dados conforme são atualizados, o que facilita a compreensão dos envolvidos e a tomada de decisões ágeis e informadas.

Como os arquivos tratados possuem um volume muito alto, o Excel será utilizado para a delimitação e redução do arquivo conforme as métricas propostas anteriormente.

### 3 PROCEDIMENTOS METODOLÓGICOS

Por meio dos processos de ETL, os dados brutos foram extraídos da base de dados da Secretaria de Segurança Pública do Estado de São Paulo, incluindo os registros de crimes, cobrindo um período de três anos. Esses dados foram submetidos ao processo de transformação, que envolveu a padronização dos formatos, a limpeza de dados ausentes e a agregação de informações relevantes. Assim, os dados foram preparados para análise estatística e interpretação.

A aplicação dos processos facilitou a integração e análise desses dados heterogêneos, fornecendo uma base sólida para a formulação de políticas públicas mais eficazes em segurança. Essa abordagem metodológica robusta, combinada com técnicas de análise de dados, contribui para uma compreensão mais completa das complexidades envolvidas na segurança pública e na percepção do público.

#### 3.1 EXTRAÇÃO DOS DADOS

A coleta dos dados foi feita através da plataforma "Números sem Mistérios" da SSP-SP, base que disponibiliza dados criminais referentes ao estado de São Paulo. (fonte, ano). Para fins do estudo, foram selecionados os dados criminais dos anos 2022, 2023 e 2024 (jan-mar), no formato XLSX.

#### 3.2 TRANSFORMAÇÃO DE DADOS

Usando o MS Excel, os *datasets* passaram por um processo de filtragem com a premissa de reduzir o arquivo para conter apenas dados necessários conforme o escopo da pesquisa. Para tanto, foram filtradas as cidades da Baixada Santista para afunilar o escopo. Com base na coluna NATUREZA\_APURADA, foram excluídas as linhas com conteúdo diferente de Santos, São Vicente, Guarujá, Praia Grande, Bertioga, Peruíbe, Mongaguá, Itanhaém e Cubatão limitando-se a região da Baixada Santista.

Além das linhas excluídas na etapa anterior, algumas colunas foram excluídas por não ser necessárias conforme o teor da pesquisa ou por ser uma coluna de dados redundantes. As colunas excluídas são demonstradas na Figura 1:

Figura 1: Lista de colunas excluídas do *dataset*



Colunas Excluídas
NUM_BO
ANO_BO
LOGRADOURO
NUMERO_LOGRADOURO
RUBRICA
DESCR_CONDUTA
MÊS_ESTADISTICA,ANO_ESTADISTICA
NOME_DELEGACIA_CIRCUNSCRIÇÃO
NOME_DEPARTAMENTO_CIRCUNSCRIÇÃO
NOME_SECCIONAL_CIRCUNSCRIÇÃO
NOME_MUNICIPIO_CIRCUNSCRIÇÃO

Fonte: Elaborado pelos autores

Utilizando a coluna “NATUREZA\_APURADA” foram selecionados os crimes citados no referencial teórico deste trabalho, sendo eles organizados conforme apresentado na Figura 2. Nesta etapa, foi utilizada a ferramenta de filtro do MS Excel.

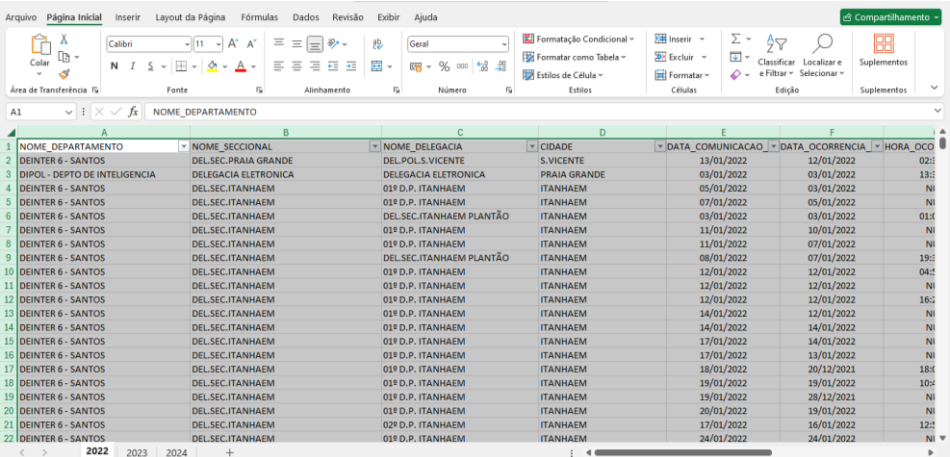
Figura 2: Organização da classificação por tipificação da ocorrência

Crime contra a Pessoa	Crime contra o Patrimônio
Homicídio	Furto
Lesão Corporal	Extorsão
Estupro	Latrocínio
	Roubo

Fonte: Elaborado pelos autores

Estes procedimentos de seleção de linhas foram realizados nos *datasets* dos 3 anos selecionados para a pesquisa. Após, os *datasets* foram unificados, para que assim fosse possível o tratamento de todos os dados em um só arquivo, resultando em 3 planilhas. Este resultado pode ser observado na Figura 3.

Figura 3: Unificando os 3 anos de dados em único arquivo



The screenshot shows the Microsoft Excel interface with a data table. The table has columns labeled A through F. Column A contains a list of crime records, including 'NOME\_DEPARTAMENTO', 'NOME\_SECCIONAL', 'NOME\_DELEGACIA', 'CIDADE', 'DATA\_COMUNICACAO', 'DATA\_OCORRENCIA', and 'HORA\_OCO'. The data rows show various crime types and locations, such as 'DEINTER 6 - SANTOS', 'DEINTER 6 - SANTOS', 'DEINTER 6 - SANTOS', etc., with corresponding dates and times.

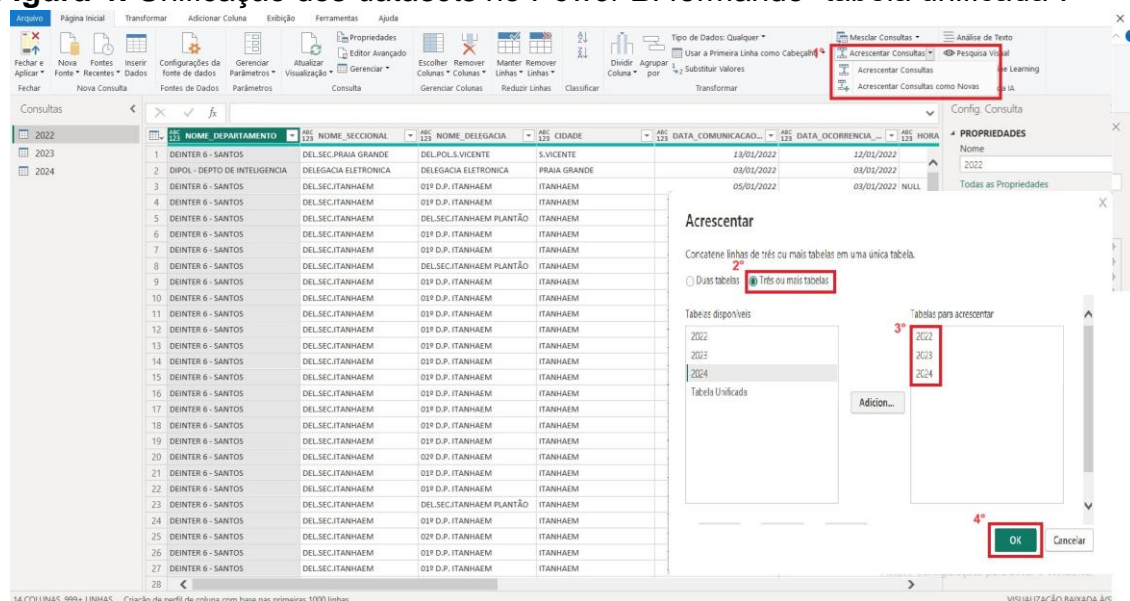
Fonte: Elaborado pelos autores

Sequencialmente foi criado um arquivo no *Power BI*, inserindo a planilha “Dados Baixada” com as informações referente a região da Baixada Santista visando a utilização de ferramentas exclusivas do programa.

Na seção de transformação de dados no *Power BI* foram realizados ajustes, inicialmente sendo a exclusão de células vazias que podem interferir na fidedignidade dos processos de análise.

Para não haver divergências quando for feito a mesclagem de tabelas, todas as colunas foram padronizadas com os mesmos títulos. Visando melhor manipulação nas análises e *dashboards*, foram unificadas as 3 tabelas usando a função “acrescentar consultas como nova”. Na Figura 4, são destacadas as ferramentas utilizadas nestas etapas, sequencialmente numeradas.

**Figura 4:** Unificação dos *datasets* no *Power BI* formando “tabela unificada”.



Fonte: Elaborado pelos autores

Devido à ausência de dados específicos sobre “hora” ou “período” na maioria dos registros, foi desenvolvido um programa em linguagem C, utilizando a ferramenta *Power Query*, com o objetivo de converter as colunas que continham dados em horas para dados em períodos, uniformizando o *dataset*. Nessa conversão, foi obtida uma coluna completa que agrupa todas as ocorrências em períodos definidos: madrugada (00:00 à 05:59), manhã (06:00 à 11:59), tarde (12:00 à 17:59), noite (18:00 à 23:59), e 'hora incerta' para os dados que não possuem horários específicos. Na Figura 5, é possível observar a situação encontrada no *dataset* original e a situação após o tratamento realizado

**Figura 5 e Figura 6:** Unificação de períodos das ocorrências

ABC 123	HORA_OCORRENCIA_...	ABC 123	DESCR_PERIODO	ABC C	Personalizar.1
	31/12/1899 02:30:00				Madrugada
	31/12/1899 13:30:00				Tarde
NULL			DE MADRUGADA		Madrugada
NULL			PELA MANHÃ		Manhã
	31/12/1899 01:00:00				Madrugada
NULL			A NOITE		Noite
NULL			EM HORA INCERTA		Hora Incerta
	31/12/1899 19:30:00				Noite
	31/12/1899 04:50:00				Madrugada
NULL			PELA MANHÃ		Manhã
	31/12/1899 16:21:00				Tarde
NULL			EM HORA INCERTA		Hora Incerta
NULL			PELA MANHÃ		Manhã
NULL			EM HORA INCERTA		Hora Incerta
NULL			EM HORA INCERTA		Hora Incerta

Fonte: Elaborado pelos autores

Após ser feito todo o processo para obtenção do exato do dia foram excluídas as seguintes colunas por trazerem redundância: “HORA\_OCORRENCIA\_BO”, “DESCR\_PERIODO”, “Personalizar”, “Mesclado”.

Com o intuito de corrigir possíveis erros e inconsistências na nomenclatura dos crimes, e para generalizar os crimes conforme apresentado na Figura 2, por meio do uso novamente da função de Substituir Valores, foram realizados os agrupamentos e substituições. Na Figura 7 é possível verificar a situação original e a final do *dataset* neste ponto de tratamento.

Figura 7: Agrupamento dos crimes

Pré	Pós
<div><div>Selecionar Tabo</div><div><div><input checked="" type="checkbox"/> ESTUPRO</div><div><input checked="" type="checkbox"/> ESTUPRO DE VULNERÁVEL</div><div><input checked="" type="checkbox"/> FURTO - OUTROS</div><div><input checked="" type="checkbox"/> FURTO DE CARGA</div><div><input checked="" type="checkbox"/> FURTO DE VEÍCULO</div><div><input checked="" type="checkbox"/> HOMICÍDIO CULPOSO OUTROS</div><div><input checked="" type="checkbox"/> HOMICÍDIO CULPOSO POR ACIDENTE DE TRAN</div><div><input checked="" type="checkbox"/> HOMICÍDIO DOLOSO</div><div><input checked="" type="checkbox"/> HOMICÍDIO DOLOSO POR ACIDENTE DE TRAN</div><div><input checked="" type="checkbox"/> LATROCÍNIO</div><div><input checked="" type="checkbox"/> LESÃO CORPORAL CULPOSA POR ACIDENTE D</div><div><input checked="" type="checkbox"/> LESÃO CORPORAL CULPOSA – OUTRAS</div><div><input checked="" type="checkbox"/> LESÃO CORPORAL DOLOSA</div><div><input checked="" type="checkbox"/> LESÃO CORPORAL SEGUIDA DE MORTE</div><div><input checked="" type="checkbox"/> ROUBO - OUTROS</div><div><input checked="" type="checkbox"/> ROUBO A RANCO</div></div></div>	<div><div><input checked="" type="checkbox"/> ESTUPRO</div><div><input checked="" type="checkbox"/> ESTUPRO DE VULNERÁVEL</div><div><input checked="" type="checkbox"/> EXTORSÃO</div><div><input checked="" type="checkbox"/> FURTO</div><div><input checked="" type="checkbox"/> HOMICÍDIO</div><div><input checked="" type="checkbox"/> LATROCÍNIO</div><div><input checked="" type="checkbox"/> LESÃO CORPORAL</div><div><input checked="" type="checkbox"/> ROUBO</div></div>

Fonte: Elaborado pelos autores

### 3.3 FERRAMENTA DE VISUALIZAÇÃO

Ao selecionar o *Power BI Desktop* para as visualizações, capitaliza-se sua amplitude de recursos. São adotadas estratégias como criação de listas de valores possíveis, expressões condicionais e consultas auxiliares. Essas abordagens garantem análises precisas e dinâmicas, possibilitando insights impactantes para o estudo.

## 4 ANÁLISE E RESULTADOS

### ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

As informações a seguir apresentam uma análise utilizando a Análise de Componentes Principais (PCA) para detalhar as características das ocorrências de polícia na Baixada Santista. O PCA simplifica e interpreta eficientemente dados multidimensionais, extraindo componentes principais que capturam a variação nas medidas geográficas. Isso proporciona uma visão clara das características distintivas entre os locais das ocorrências, identificando padrões e variações.

#### Dimensão dos Dados

Número de instâncias: Número total de ocorrências após remoção de dados ausentes. Número de atributos: 2 atributos numéricos preditivos: latitude e longitude.

#### Matriz de Covariância

A matriz de covariância é uma métrica estatística que caracteriza a interdependência entre as variáveis de um conjunto de dados, revelando as correlações entre suas características. A seguir está a Matriz Quadrada Reduzida da Matriz de Covariância:

$$\begin{bmatrix} 1.00008163 & 1.00007313 \\ 1.00007313 & 1.00008163 \end{bmatrix}$$

#### Autovalores e Autovetores

Os autovalores e autovetores, derivados da matriz de covariância, oferecem insights sobre as direções primárias dos dados e suas relevâncias. Os autovetores

delineiam as direções principais, ou componentes principais, enquanto os autovalores indicam a variância explicada por cada componente.

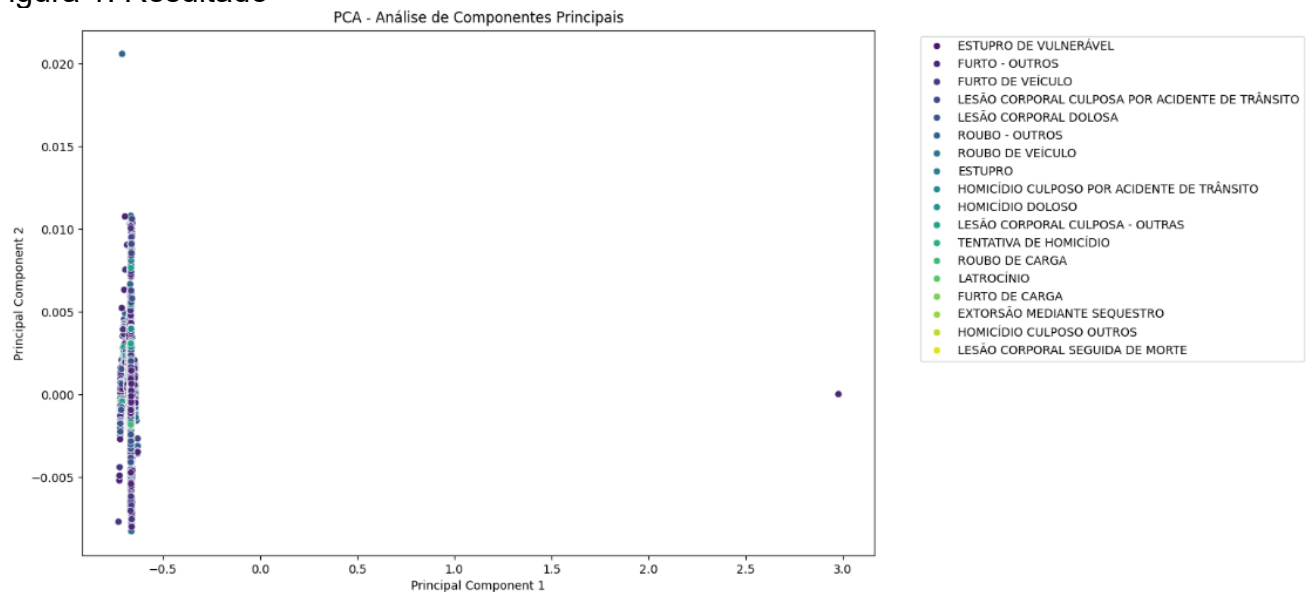
#### Autovalores:

$$[2.00015475e+00 \ 8.49825139e-06]$$

#### Autovetores:

$$\begin{bmatrix} 0.70710678 & 0.70710678 \\ 0.70710678 & -0.70710678 \end{bmatrix}$$

Figura 1: Resultado



Fonte: Elaborado pelos autores

A plotagem do PCA das ocorrências de polícia na Baixada Santista revelou que a maioria dos incidentes tem características geográficas semelhantes, concentradas próximas à origem do gráfico. No entanto, alguns incidentes, como um ponto de "Furto de Veículo", estão dispersos, indicando ocorrências geograficamente distintas. Essa visualização facilita a identificação de padrões e áreas específicas de interesse para a segurança pública, ajudando no planejamento estratégico.

REFERÊNCIAS

AMAZON, A. **O que é Python**. Disponível em: <https://aws.amazon.com/pt/what-is/python/#:~:text=O%20Python%20%C3%A9%20uma%20linguagem,executada%20em%20muitas%20plataformas%20diferentes>. Acesso em: 10 abr. 2024.

ARRUDA, A. **Crimes contra a pessoa**. Disponível em: <https://www.jusbrasil.com.br/artigos/crimes-contr-a-pessoa/2106936388>. Acesso em: 09 abr. 2024.

ATRA. (2024). **História das Linguagens de Programação**. ATRA INFORMÁTICA. Disponível em: <https://www.atrainformatica.com.br/2023/04/05/historia-das-linguagens-de-programacao/>. Acesso em 14 jun. 2024

AWARI. **Como Se Tornar um Analista de Dados: Guia Completo de Formação**. Disponível em: <https://awari.com.br/como-se-tornar-um-analista-de-dados-guia-completo-de-formacao/>. Acesso em: 9 maio. 2024.

A ESTRATÉGIA ODS In: **ODS 16**. Disponível em: <https://www.estrategiaods.org.br/os-ods/ods16/>. Acesso em: 08 mar. 2024.

A TRIBUNA. **Entendendo a Violência na Região**. Disponível em: <https://www.atribuna.com.br/opiniao/tribuna-livre/entendendo-a-violencia-na-regiao#:~:text=O%20IECV%20Vida%20da%20Baixada,05%2C%20entre%202021%20e%202022>. Acesso em: 08 mar. 2024.

BRASIL. BRASIL. (org.). **DECRETO-LEI No 2.848, DE 7 DE DEZEMBRO DE 1940: código penal. Código Penal**. 1940. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/decreto-lei/del2848compilado.htm](https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm). Acesso em: 14 abr. 2024.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Promulgada em 5 de outubro de 1988. Diário Oficial da União, Brasília, DF, 5 de outubro de 1988.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/projetos/PEC/msg975-970902.htm](https://www.planalto.gov.br/ccivil_03/projetos/PEC/msg975-970902.htm). Acesso em: 10 abr. 2024.

BRASIL. **Crime**. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/decreto-lei/del3914.htm](https://www.planalto.gov.br/ccivil_03/decreto-lei/del3914.htm). Acesso em: 11 abr. 2024.

CALDEIRA, T. P. **Cidade de muros: crime, segregação e cidadania em São Paulo**. São Paulo: 34, 2000. Disponível em: <https://csociais.wordpress.com/wp-content/uploads/2014/05/caldeira-teresa-a-cidade-de-muros-completo.pdf>

SIPS 2011. - **Sistema de Indicadores de Percepção Social – Segurança Pública**. Gov.br. Disponível em:



[https://portalantigo.ipea.gov.br/agencia/index.php?option=com\\_content&view=article&id=24416](https://portalantigo.ipea.gov.br/agencia/index.php?option=com_content&view=article&id=24416) Acesso em: 10 maio. 2024.

COSTA, Arthur Trindade; LIMA, Renato Sérgio. Segurança pública. **Crime, polícia e justiça no Brasil. São Paulo: Contexto**, 2014. Acesso em: 05 maio. 2024.

DANTAS, George Felipe de Lima; SILVA JÚNIOR, Álvaro Pereira da; PERSIJN, Annik. **O Medo do Crime**. 2006. Disponível em: <https://ibsp.org.br/wp-content/uploads/2018/06/O-Medo-do-Crime.pdf>. Acesso em: 03 abr. 2024.

DATABRICKS. **O que é ETL**. Disponível em: <https://www.databricks.com/br/glossary/extract-transform-load>. Acesso em: 11 maio. 2024.

DA SILVA, Luiz Antônio Machado. "Violência urbana", **Segurança Pública e favelas - o caso do Rio de Janeiro atual. Caderno Crh**, v. 23, n. 59, p. 283-300, 2010. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 07 abr. 2024.

DE SÁ, Marcela Rocha Tortureli. (2015). **Introdução à Linguagem C**. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/4211311/mod\\_resource/content/1/Apostila-de-Introdu%C3%A7%C3%A3o-%C3%A0-Linguagem-C.pdf](https://edisciplinas.usp.br/pluginfile.php/4211311/mod_resource/content/1/Apostila-de-Introdu%C3%A7%C3%A3o-%C3%A0-Linguagem-C.pdf). Acesso em: 14 jun. 2024.

EQUIPE, L. F. G. **Conheça melhor os crimes mais cometidos no Brasil**. Disponível em: <https://blog.lfg.com.br/estudos/crimes-mais-cometidos-no-brasil/>. Acesso em: 12 abr. 2024.

ESTADO DE SÃO PAULO. Planejamento SP. **Região Metropolitana da Baixada Santista RMBS**. Disponível em: [http://planejamento.sp.gov.br/static/arquivos/audiencias/caracterizacao2020/BAIXADA\\_SANTISTA\\_Caracterizacao.pdf](http://planejamento.sp.gov.br/static/arquivos/audiencias/caracterizacao2020/BAIXADA_SANTISTA_Caracterizacao.pdf). Acesso em: 08 mar. 2024.

FERRARO, Kenneth F. **Fear of Crime: interpreting victimization risk**. State University of New York Press.: 1995. <https://www.jusbrasil.com.br/artigos/o-que-e-seguranca-publica/586735267>. Acesso em: 09 abr. 2024.

IBGE. **40% da população acredita ter chance média ou alta de ser roubada na rua**. Agência IBGE Notícias, 2021. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/35760-40-da-populacao-acredita-ter-chance-media-ou-alta-de-ser-roubada-na-rua>. Acesso em: 07 abr. 2024.

MICROSOFT. **Power BI**. Disponível em: <https://learn.microsoft.com/pt-br/power-bi/fundamentals/power-bi-overview>. Acesso em: 11 abr. 2024.

MOLERO, María Naredo. **Seguridad urbana y miedo al crimen**. Boletín CF+S 22. Madrid: Instituto Juan de Herrera, 2002. Disponível em: <https://polired.upm.es/index.php/boletincfs/article/viewFile/2400/2481>

SILVA, Fábio de Sá. **Violência e Segurança Pública**. 2018. Disponível em: <https://fpabramo.org.br/publicacoes/wp-content/uploads/sites/5/2017/05/colecaoquesaber-03-com-capa.pdf>

VEDOVA, Daiane. **O que é Segurança Pública**. Disponível em: <https://www.jusbrasil.com.br/artigos/o-que-e-seguranca-publica/586735267>. Acesso em: 09 abr. 2024

WIKIPEDIA. **Crime**. Disponível em: <https://pt.wikipedia.org/w/index.php?title=Crime&oldid=67654077>. Acesso em: 10 abr. 2024.

WIKIPEDIA. **Microsoft Excel**. Disponível em: [https://pt.wikipedia.org/w/index.php?title=Microsoft\\_Excel&oldid=67674744](https://pt.wikipedia.org/w/index.php?title=Microsoft_Excel&oldid=67674744). Acesso em: 10 abr. 2024.