



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CAMPUS RECIFE
CENTRO DE INFORMÁTICA
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

RAPHAEL FARIAS DE FIGUEIREDO

**ANÁLISE E VISUALIZAÇÃO DE DADOS OPERACIONAIS EM
AMBIENTES DE MOBILIDADE URBANA:** Um Estudo de Caso em uma
Empresa Brasileira de Transporte Privado

Recife
2025

RAPHAEL FARIAS DE FIGUEIREDO

**ANÁLISE E VISUALIZAÇÃO DE DADOS OPERACIONAIS EM
AMBIENTES DE MOBILIDADE URBANA:** Um Estudo de Caso em uma
Empresa Brasileira de Transporte Privado

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Sistemas de Informação do Campus Recife da Universidade Federal de Pernambuco – UFPE, na modalidade de monografia, como requisito parcial para a obtenção do grau de bacharel em Sistemas de Informação.

Área de concentração: Análise de Dados.

Orientador (a): Prof. Dr. Adiel Teixeira de Almeida Filho

Recife

2025

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Figueiredo, Raphael Farias de .

Análise e visualização de dados operacionais em ambientes de mobilidade urbana : Um estudo de caso em uma empresa brasileira de transporte privado / Raphael Farias de Figueiredo. - Recife, 2025.

56 p : il.

Orientador(a): Adiel Teixeira de Almeida Filho

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado, 2025.

Inclui referências.

1. Análise de Dados. 2. Visualização de Dados. 3. Mobilidade Urbana. 4. Engenharia de Software. 5. Usabilidade. I. Almeida Filho, Adiel Teixeira de. (Orientação). II. Título.

000 CDD (22.ed.)

RAPHAEL FARIAS DE FIGUEIREDO

**ANÁLISE E VISUALIZAÇÃO DE DADOS OPERACIONAIS EM
AMBIENTES DE MOBILIDADE URBANA: Um Estudo de Caso em uma
Empresa Brasileira de Transporte Privado**

Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso de
Sistemas de Informação do Campus
Recife da Universidade Federal de
Pernambuco – UFPE, na modalidade de
monografia, como requisito parcial para a
obtenção do grau de bacharel em
Sistemas de Informação.

Aprovado em: 18/12/2025

BANCA EXAMINADORA

Prof. Dr. Adiel Teixeira de Almeida Filho (Orientador)
Universidade Federal de Pernambuco

Prof. Dr. Victor Wanderley Costa de Medeiros
Universidade Federal de Pernambuco

RESUMO

Este trabalho investiga, por meio de um estudo de caso em uma empresa brasileira de mobilidade urbana que atua em modelos híbridos B2B e B2C, como técnicas de análise e visualização de dados podem ser aplicadas para melhorar a eficiência e a gestão operacional. O problema abordado consiste na dificuldade de consolidar e interpretar informações provenientes de múltiplas fontes, como registros de corridas, cancelamentos, tempos de espera, canais de solicitação e meios de pagamento, o que limita a capacidade dos gestores de identificar padrões, gargalos e oportunidades de melhoria em tempo quase real. O estudo foi conduzido em três etapas: (i) identificação e definição dos principais indicadores de desempenho (KPIs) relevantes para o contexto da mobilidade urbana; (ii) construção de um pipeline de análise utilizando a biblioteca Pandas (Python) para extração, transformação e integração dos dados; e (iii) desenvolvimento de dashboards interativos no Looker Studio, voltados à visualização intuitiva dos resultados e ao acompanhamento contínuo das métricas operacionais. Os resultados evidenciam que a combinação entre um pipeline analítico estruturado em Pandas e o uso de uma ferramenta gratuita, de boa usabilidade e low code como o Looker Studio tornou viável, para uma empresa de pequeno porte, a implementação de um ambiente de business intelligence alinhado às necessidades reais dos stakeholders, com ganhos concretos em capacidade analítica e suporte à decisão. Assim, além de oferecer uma solução prática para a organização analisada, o trabalho demonstra que BI é acessível também a empresas de menor porte e propõe um modelo metodológico replicável em outros contextos organizacionais que lidam com dados operacionais e demandam suporte analítico para decisões estratégicas.

Palavras-chave: Análise de Dados. Visualização de Dados. Mobilidade Urbana. Engenharia de Software. Usabilidade.

ABSTRACT

This study investigates, through a case study in a Brazilian urban mobility company operating in hybrid B2B and B2C models, how data analysis and visualization techniques can be applied to improve operational efficiency and management. The core problem lies in the difficulty of consolidating and interpreting information from multiple sources, such as trip records, cancellations, waiting times, request channels, and payment methods, which limits managers' ability to identify patterns, bottlenecks, and improvement opportunities in near real time. The study was conducted in three stages: (i) identification and definition of key performance indicators (KPIs) relevant to the urban mobility context; (ii) construction of an analytical pipeline using the Pandas (Python) library for data extraction, transformation, and integration; and (iii) development of interactive dashboards in Looker Studio, aimed at intuitive visualization of results and continuous monitoring of operational metrics. The results show that the combination of a structured analytical pipeline in Pandas with the use of a free, user-friendly, low-code tool such as Looker Studio made it feasible, for a small-sized company, to implement a business intelligence environment aligned with the real needs of stakeholders, yielding concrete gains in analytical capability and decision support. Thus, in addition to providing a practical solution for the organization under study, this work demonstrates that BI is also accessible to smaller companies and proposes a methodological model that can be replicated in other organizational contexts that deal with operational data and require analytical support for strategic decision-making.

Keywords: Data Analysis. Data Visualization. Urban Mobility. Software Engineering. Usability.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Fluxo do Processo e Ferramentas	32
Figura 2 –	Seleção de Variáveis	37
Figura 3 –	Conversão dos tipos de dados	37
Figura 4 –	Processo de autenticação para upload dos dados	38
Figura 5 –	Índice dos dados das corridas no Looker Studio	39
Figura 6 –	Criação de filtros no Looker Studio	40
Figura 7 –	Configuração dos gráficos no LookerStudio	41
Figura 8 –	Visão Geral do Dashboard Mobile	42
Figura 9 –	Gráfico em barra das corridas realizadas mensalmente	42
Figura 10 –	Gráfico em linha das corridas realizadas mensalmente	43
Figura 11 –	Gráfico em barra das corridas canceladas mensalmente	43
Figura 12 –	Gráfico em linha das corridas canceladas mensalmente	43
Figura 13 –	Gráfico em barra das corridas realizadas mensalmente para PJ	44
Figura 14 –	Gráfico em linha das corridas realizadas mensalmente para PJ	44
Figura 15 –	Gráfico em barra das corridas canceladas mensalmente para PJ	44
Figura 16 –	Gráfico em linha das corridas canceladas mensalmente para PJ	44
Figura 17 –	Relatório Geral e Origem dos Chamados	45
Figura 18 –	Representação Gráfica do Operacional	46
Figura 19 –	Representação Gráfica do Financeiro	47
Figura 20 –	Representação Gráfica do Comercial	47
Figura 21 –	Representação Gráfica do Gerente Geral	48
Figura 22 –	Representação Gráfica do Diretor Presidente	48

SUMÁRIO

1	INTRODUÇÃO	10
2	CONTEXTO	15
2.1	FUNDAMENTOS DE DADOS, INFORMAÇÃO E CONHECIMENTO	15
2.2	ANÁLISE DE DADOS: PROCESSOS, KPI's, TÉCNICAS E FERRAMENTAS	18
2.3	VISUALIZAÇÃO DE DADOS E DESIGN DE DASHBOARDS	20
2.4	MOBILIDADE URBANA E INDICADORES OPERACIONAIS	
3	METODOLOGIA	27
3.1	ESTRATÉGIA DO ESTUDO DE CASO	28
3.2	CONTEXTO ORGANIZACIONAL E UNIDADE DE ANÁLISE	28
3.3	PERGUNTAS DE PESQUISA	29
3.4	PROCEDIMENTOS DE COLETA E PREPARAÇÃO DOS DADOS	30
3.5	PROCEDIMENTOS DE ANÁLISE DOS DADOS	30
3.5.1	Análise dos dados operacionais (PP1)	31
3.5.2	Análise do pipeline e das visualizações (PP2)	31
3.6	AMEAÇAS À VALIDADE	31
4	RESULTADOS	35
4.1	CARACTERIZAÇÃO DA EMPRESA X	35
4.2	DESENHO DA SOLUÇÃO	35
4.3	DEFINIÇÃO DE INDICADORES E REQUISITOS	36
4.4	PIPELINE ANALÍTICO: EXTRAÇÃO, TRANSFORMAÇÃO E CARGA DOS DADOS	37
4.5	CONSTRUÇÃO DOS DASHBOARDS	39
4.6	RESULTADOS VISUAIS	41
5	DISCUSSÃO	49
5.1	INDICADORES OPERACIONAIS E CONVERGÊNCIA COM A LITERATURA DE MOBILIDADE	50
5.2	DO GRÁFICO AO SISTEMA: IMPLICAÇÕES DO PIPELINE EM PANDAS PARA CONFIABILIDADE E GOVERNANÇA	50

5.3	DASHBOARDS COMO MEDIAÇÃO COGNITIVA: ESCOLHAS DE VISUALIZAÇÃO E ESTRUTURA ANALÍTICA	51
5.4	DIVERGÊNCIAS RELEVANTES E OPORTUNIDADES DE FORTALECIMENTO ACADÊMICO	53
6	CONCLUSÃO	55
	REFERÊNCIAS	57

1 INTRODUÇÃO

A transformação digital e o avanço das tecnologias de informação têm redefinido a forma como organizações e cidades gerenciam seus dados e tomam decisões. No cenário contemporâneo, a geração de dados em larga escala (provenientes de sensores, sistemas de informação, dispositivos móveis e interações digitais) tem se consolidado como um dos fenômenos centrais da sociedade informacional.

Nesse contexto, a análise e a visualização de dados emergem como práticas essenciais para converter volumes massivos de registros em conhecimento acionável, promovendo inteligência organizacional e eficiência operacional. A análise de dados envolve o processamento sistemático de informações para identificar padrões, relações e tendências, transformando dados brutos em insights significativos que subsidiam decisões estratégicas.

Complementarmente, a visualização de dados constitui um meio de comunicação entre a complexidade analítica e a compreensão humana, permitindo que fenômenos multidimensionais sejam percebidos de maneira intuitiva por meio de gráficos, painéis e representações visuais interativas (MUNZNER, 2014; NONATO; AUPETIT, 2018). Assim, enquanto a análise estrutura o raciocínio inferencial sobre os dados, a visualização dá forma à narrativa que os torna inteligíveis.

Na literatura contemporânea, essa relação entre análise e visualização é compreendida como parte de um ciclo cognitivo de descoberta, em que o ser humano e o sistema computacional interagem continuamente para interpretar e validar hipóteses. Autores como Börner, Bueckle e Ginda (2019) argumentam que o domínio da alfabetização em visualização de dados (Data Visualization Literacy) é um requisito fundamental para que indivíduos e organizações possam extrair valor da informação.

Esse domínio envolve competências técnicas, perceptuais e cognitivas para interpretar corretamente representações visuais e compreender as estruturas estatísticas subjacentes. Além disso, a literatura recente evidencia que a visualização eficaz não depende apenas do volume ou da precisão dos dados, mas também de critérios de clareza, legibilidade e contexto (HAROZ; COSGROVE; WHITAKER, 2015).

A eficiência perceptual e a escolha apropriada de técnicas visuais são determinantes para garantir que os usuários compreendam as relações apresentadas e tomem decisões embasadas. Nesse sentido, a integração entre métodos estatísticos robustos, como os discutidos por Bickel e Li (2016), e princípios de design de visualização torna-se essencial para a construção de sistemas analíticos confiáveis.

Essa intersecção entre análise, estatística e design visual adquire especial relevância em domínios de alta complexidade, como o da mobilidade urbana. Nas últimas décadas, a urbanização acelerada e o crescimento dos serviços de transporte sob demanda têm intensificado a produção de dados relacionados a deslocamentos, horários de pico, solicitações de corridas, cancelamentos e modos de pagamento.

Esses dados, quando devidamente analisados, permitem compreender dinâmicas urbanas, otimizar recursos e aprimorar a qualidade dos serviços oferecidos. A aplicação de métodos de análise de dados em contextos de mobilidade tem se mostrado uma estratégia eficaz para apoiar a gestão e o planejamento de sistemas de transporte.

Estudos internacionais demonstram que o uso de dados provenientes de GPS, aplicativos e bilhetagem eletrônica pode revelar padrões de deslocamento e auxiliar na tomada de decisões sobre oferta, demanda e sustentabilidade (LENG et al., 2016; ANDRIYANOV, 2020).

Em paralelo, pesquisas recentes têm evidenciado que a visualização de dados aplicada à mobilidade urbana possibilita monitorar, em tempo real, o desempenho operacional, facilitando a identificação de gargalos e o redesenho de estratégias de atendimento (VILA et al., 2016; MASSOBRIO; NESMACHNOW, 2020).

É nesse contexto que se insere o presente trabalho, cujo ponto de partida é um problema concreto vivenciado pela Empresa X, como volume moderado de dados operacionais (por se tratar de uma empresa de pequeno porte), a organização enfrenta dificuldades em consolidá-los, analisá-los e visualizá-los de forma integrada e compreensível para seus diferentes setores.

Os relatórios de desempenho são produzidos de maneira manual e fragmentada, o que impede uma visão sistêmica do comportamento dos serviços e limita a capacidade de acompanhar, em tempo quase real, indicadores como número

de corridas atendidas, cancelamentos, tempo médio de espera, origem dos chamados e métodos de pagamento. Essa dissociação entre o volume de dados disponíveis e a capacidade analítica instalada configura uma lacuna crítica para a gestão, pois compromete a eficiência do processo decisório e a agilidade de resposta da empresa.

Essa lacuna evidencia a necessidade de transformar dados operacionais em conhecimento acionável. A mera existência de registros sobre corridas, cancelamentos ou meios de pagamento não garante, por si só, a geração de insights relevantes para o negócio. É indispensável que esses dados sejam tratados, organizados e sintetizados em informações que possam ser interpretadas pelos gestores e incorporadas aos processos decisórios.

Nesse sentido, o desafio não se restringe ao armazenamento de dados, mas envolve a construção de um fluxo analítico que os converta em evidências capazes de sustentar decisões mais rápidas, fundamentadas e alinhadas às dinâmicas de demanda e desempenho da operação.

Diante desse cenário, coloca-se o seguinte problema de pesquisa: como técnicas de análise e visualização de dados podem ser aplicadas aos dados operacionais de uma empresa de mobilidade urbana B2B para apoiar a tomada de decisão e melhorar a eficiência operacional?

Por isso, a adoção de dashboards, indicadores-chave de desempenho (KPIs) e visualizações interativas assume papel central. A proposta do estudo é implementar um modelo de análise e visualização de dados operacionais baseado em ferramentas abertas e acessíveis, utilizando Pandas (Python) para tratamento e integração das diversas fontes de dados e Looker Studio para a construção de painéis visuais.

Espera-se que, por meio de dashboards interativos, os gestores possam identificar tendências, gargalos e oportunidades de melhoria com maior rapidez e precisão, superando a limitação dos relatórios manuais e fragmentados. Os KPIs passam, assim, a desempenhar a função de síntese estruturada do desempenho da operação, permitindo monitorar, de forma contínua, dimensões críticas da qualidade do serviço.

Neste trabalho, tais KPIs são definidos priorizando indicadores diretamente associados à qualidade e ao desempenho do serviço já registrados nos sistemas operacionais da Empresa X, em especial número de corridas atendidas,

cancelamentos, tempo médio de espera, origem dos chamados e métodos de pagamento.

A motivação para adotar técnicas de análise de dados e visual analytics decorre, portanto, da necessidade de reposicionar os dados operacionais como ativo estratégico. Ao estruturar um modelo que integra tratamento automático de dados, geração de indicadores e visualizações interativas, o trabalho busca promover uma cultura organizacional orientada por dados (data-driven), em que decisões sejam crescentemente ancoradas em evidências e não apenas em percepções subjetivas ou em relatórios pontuais.

A integração entre análise, visualização e avaliação de usabilidade é vista como eixo de inovação, na medida em que possibilita não apenas produzir informações, mas garantir que elas sejam efetivamente compreendidas e utilizadas pelos diferentes stakeholders da empresa.

Por fim, a Empresa X configura-se como um caso particularmente relevante no contexto da mobilidade urbana, por reunir características típicas de organizações do setor que operam com dados operacionais valiosos para análise e múltiplos canais de atendimento.

As dificuldades em consolidar e explorar analiticamente seus dados, a dependência de relatórios manuais e a necessidade de responder a variações de demanda e desempenho são desafios que extrapolam a realidade específica da organização e dialogam com problemas recorrentes em empresas de transporte urbano.

Ao propor e avaliar um modelo de análise e visualização de dados aplicado a esse contexto, o estudo pretende não apenas oferecer uma solução prática para a Empresa X, mas também contribuir para o avanço das práticas de engenharia e visualização de dados na mobilidade urbana, evidenciando o potencial da integração entre análise, visualização e avaliação de usabilidade como vetor de eficiência em serviços de transporte inteligentes.

Os objetivos deste trabalho estruturam o encadeamento lógico das atividades metodológicas e definem o escopo da investigação empírica conduzida na Empresa X. O objetivo geral consiste em investigar, por meio de um estudo de caso, como técnicas de análise e visualização de dados podem ser aplicadas aos dados operacionais de uma empresa de mobilidade urbana B2B para apoiar a tomada de decisão e melhorar a eficiência operacional.

A partir desse propósito central, derivam-se quatro objetivos específicos que orientam cada etapa do método. O primeiro consiste em identificar e estruturar as métricas e indicadores operacionais mais relevantes para a avaliação de desempenho na mobilidade corporativa, incluindo tempo médio de espera, cancelamentos, padrões de origem dos chamados e meios de pagamento, com base em sua relevância para a qualidade do serviço e em sua presença nos registros operacionais já coletados pela Empresa X.

O segundo objetivo envolve aplicar técnicas de extração, transformação, integração e limpeza de dados, utilizando scripts em Python (Pandas), assegurando padronização e consistência das bases operacionais. O terceiro contempla o desenvolvimento de dashboards interativos no Looker Studio, capazes de representar visualmente os KPIs identificados.

Finalmente, o quarto objetivo consiste em analisar os impactos das visualizações no processo decisório, observando de que modo os painéis interativos contribuem para a maturidade analítica da organização. Esses objetivos são interdependentes e formam um ciclo metodológico coerente, no qual os resultados de cada etapa subsidiam a próxima, garantindo alinhamento entre a estratégia de investigação e os fundamentos teóricos do estudo de caso adotado neste trabalho.

A estrutura deste trabalho está organizada da seguinte forma. A seção seguinte, Contexto e Trabalhos Relacionados, apresenta o referencial teórico construído durante a fase exploratória, abrangendo conceitos, métodos e estudos prévios sobre análise de dados, visualização e mobilidade urbana. Em seguida, a seção Metodologia detalha o delineamento do estudo de caso, descrevendo as etapas de coleta, tratamento e visualização dos dados.

A seção Resultados apresenta a aplicação prática do modelo proposto, incluindo os dashboards desenvolvidos e a análise dos indicadores obtidos. Posteriormente, a seção Discussão confronta os resultados empíricos com o referencial teórico, destacando contribuições e limitações do estudo. Por fim, a Conclusão sintetiza os principais achados e aponta caminhos para trabalhos futuros e aplicações no campo da análise e visualização de dados em mobilidade urbana.

2 CONTEXTO E TRABALHOS RELACIONADOS

Esta subseção estabelece a base conceitual do TCC ao explicitar distinções entre dado, informação e conhecimento, bem como os modelos que descrevem sua transformação e o papel da visualização nesse processo. Com base no referencial teórico que reunimos, apresentamos os conceitos que articulam fundamentos em visualização de dados, fundamentos estatísticos e princípios de projeto em visualização, sustentando a passagem do dado bruto ao conhecimento acionável que orienta a tomada de decisão nas organizações.

2.1 FUNDAMENTOS DE DADOS, INFORMAÇÃO E CONHECIMENTO

A tríade dado–informação–conhecimento é apresentada como eixo central da ciência de dados contemporânea, na qual o “dado” é compreendido como um registro observacional bruto, produto de processos específicos de medição, instrumentação e contexto; a “informação” corresponde ao conjunto de relações inferidas e validadas sobre esses registros; e o “conhecimento” emerge do uso pragmático dessas relações para resolver problemas reais e orientar decisões (BICKEL; LI, 2016).

Nessa perspectiva, os dados não constituem verdades autônomas, mas representações permeadas por incerteza e variabilidade, que apenas se tornam informativas quando submetidas a processos de filtragem, modelagem e inferência estatística rigorosa. A passagem de dado a informação exige, portanto, estrutura inferencial; a passagem de informação a conhecimento exige a aplicação dessas inferências em contextos decisórios concretos (BICKEL; LI, 2016).

Ao discutir alfabetização em visualização de dados (Data Visualization Literacy – DVL-FW), Börner et al. (2019) reforçam essa mesma cadeia conceitual, mostrando que “dados” só se convertem em “informação” e, posteriormente, em “conhecimento acionável” quando o sujeito é capaz de ler e construir visualizações de modo competente.

A visualização é concebida como elemento mediador entre o dado bruto e a decisão, articulando um caminho que parte de necessidades de insight, passa pela escolha adequada de escalas de mensuração e codificações visuais, e culmina na interpretação de evidências em narrativas decisórias. Assim, a distinção entre dado,

informação e conhecimento não é meramente terminológica: depende de decisões metodológicas explícitas sobre como os dados são estruturados, analisados e representados visualmente (BÖRNER et al., 2019).

A visualização é definida como um processo de transformar dados em representações visuais interativas que ampliam a cognição humana, tornando padrões, exceções e relações mais facilmente perceptíveis do que por inspeção tabular ou textual. Nessa formulação, os dados são registros simbólicos do mundo, enquanto a informação corresponde à interpretação situada desses registros em função de metas e tarefas analíticas específicas (MUNZNER, 2014).

O conhecimento, por sua vez, emerge quando as leituras visuais, apoiadas em abstrações corretas de dados (tipos de atributos, estruturas como tabelas, redes, campos e conjuntos) e em escolhas adequadas de marcas e canais visuais, sustentam decisões e modificam fluxos e resultados organizacionais. A autora sintetiza esse ciclo em um fluxo de trabalho iterativo — coletar, processar, analisar, visualizar, interagir, interpretar, decidir — no qual a passagem dado → informação → conhecimento é continuamente refinada (MUNZNER, 2014).

Haroz et al., (2015) demonstram que a ciência da comunicação visual de dados explicita que a transformação de dados em conhecimento depende de mecanismos atencionais, perceptivos e de memória. A atenção visual seleciona, em meio ao ruído gráfico, quais elementos serão efetivamente percebidos; a percepção organiza esses elementos em padrões com base em princípios como proximidade, similaridade e continuidade; e a memória de trabalho limita quantas comparações e inferências podem ser realizadas simultaneamente.

Dessa forma, a informação não está “contida” nos dados, mas é construída cognitivamente a partir da forma como são codificados e arranjados. Visualizações que respeitam esses limites cognitivos — por exemplo, por meio de eixos alinhados, escalas compartilhadas e redução de ruído estético — maximizam a clareza informacional e favorecem a formação de conhecimento confiável para apoiar decisões (HAROZ et al., 2015).

A tomada de decisão mediada por visualizações é descrita por meio de um modelo cognitivo dual, em que um sistema rápido, intuitivo e automático (*Type 1*) convive com um sistema lento, deliberativo e analítico (*Type 2*). Nesse enquadramento, o dado é a representação externa (gráficos, mapas, dashboards); a informação surge do processamento perceptivo que organiza elementos visuais em

padrões significativos; e o conhecimento se consolida quando o sistema deliberativo interpreta esses padrões à luz de esquemas mentais, experiências prévias e objetivos decisórios (PADILLA et al., 2018).

Em ambientes organizacionais intensivos em dados, esse encadeamento revela a informação como ativo estratégico: visualizações bem projetadas reduzem a carga cognitiva, alinham forma gráfica e tarefa (cognitive fit) e, assim, aumentam a qualidade e a velocidade das decisões baseadas em dados (PADILLA et al., 2018).

A análise de dados de alta dimensionalidade, característica da era digital e dos cenários usualmente associados a big data, demanda técnicas de projeção que traduzam estruturas matemáticas complexas em representações bidimensionais cognitivamente acessíveis. Nonato e Aupetit (2018) apresentam projeções multidimensionais como instrumentos que convertem volumes consideráveis de variáveis e observações em configurações espaciais nas quais relações de similaridade, agrupamento e distância se tornam visíveis.

Essa compressão controlada transforma excesso de dados em informação relevante e, quando combinada a enriquecimento de layout (metadados, rótulos, uso criterioso de cor e hierarquias visuais), sustenta processos analíticos típicos de ambientes de business intelligence e analytics, nos quais se busca separar sinal de ruído para apoiar decisões em contextos complexos (NONATO; AUPETIT, 2018).

No contexto de arquiteturas de business intelligence e analytics, a visualização é concebida como componente central de sistemas de apoio à decisão, em que dados operacionais são coletados, preparados e agregados em métricas e indicadores de desempenho. A escolha de estruturas de dados (tabelas, campos, redes), de transformações (filtragem, agregação, derivação de atributos) e de estratégias de facetamento, redução e foco+contexto determina quais aspectos do fenômeno organizacional se tornam informacionais e, portanto, passíveis de serem convertidos em conhecimento (MUNZNER, 2014).

Em ambientes de mobilidade urbana, em que operações intensivas em dados geram registros de corridas, cancelamentos, tempos de espera, origens de chamada e formas de pagamento, esse enquadramento teórico permite compreender dashboards e painéis analíticos como materializações da cadeia dado → informação → conhecimento, na qual grandes massas de dados brutos são transformadas em KPIs visualmente organizados para apoiar o planejamento e a gestão (MUNZNER, 2014).

A noção de alfabetização em visualização de dados complementa essa visão ao enfatizar que, sem competências para ler e construir visualizações alinhadas às escalas de mensuração e às tarefas analíticas, organizações intensivas em dados permanecem incapazes de explorar plenamente o potencial estratégico de seus ativos informacionais (BÖRNER et al., 2019).

Em empresas de mobilidade urbana que dependem de decisões rápidas sobre alocação de recursos, monitoramento de desempenho e resposta a flutuações de demanda, a capacidade de transformar logs operacionais em visualizações perceptualmente corretas e interpretativamente válidas torna-se condição para uma cultura data-driven. A informação, nesse cenário, deixa de ser apenas subproduto do registro de eventos e passa a configurar um recurso organizacional crítico, estruturado e ativado por meio de práticas de visual analytics (BÖRNER et al., 2019).

2.2 ANÁLISE DE DADOS: PROCESSOS, KPI's, TÉCNICAS E FERRAMENTAS

Börner et al., (2019) estabelecem que o processo de análise de dados no contexto organizacional deve ser compreendido como um ciclo cognitivo estruturado em múltiplas etapas: aquisição, preparação, transformação, visualização e interpretação. Os autores mostram que a análise é o núcleo do processo analítico, pois é nesse estágio que operações como limpeza, transformação, agregação e modelagem são aplicadas para converter dados brutos em informações consistentes.

Eles argumentam que cada tipo de dado (nominal, ordinal, intervalar ou racional) demanda técnicas analíticas distintas, articulando uma correspondência entre o tipo de variável, o propósito da análise e o insight esperado. Assim, a análise de dados se torna um processo iterativo, no qual o produto de uma etapa retroalimenta as seguintes, estruturalmente alinhado à preparação dos dados operacionais realizada neste TCC (BÖRNER et al., 2019).

Bickel e Li (2016) complementam essa visão ao enfatizar que a análise de dados no ambiente corporativo depende do rigor estatístico e da estruturação formal de hipóteses. Para os autores, a análise só adquire validade quando há controle das incertezas e aplicação correta de modelos estatísticos. A limpeza e normalização dos dados (tarefas centrais do pipeline operacional) representam o elo crítico entre o

dado bruto e a informação estatisticamente confiável. Dessa perspectiva, a análise descritiva, diagnóstica e preditiva é vista como um ciclo contínuo de modelagem, inferência e validação, garantindo que os KPIs corporativos reflitam fenômenos reais e sejam interpretáveis pelos gestores.

Munzner (2014) reforça que a análise organizacional depende de múltiplos níveis de abstração, nos quais o dado passa por transformações progressivas até atingir uma representação visual compreensível. A autora propõe um modelo hierárquico no qual a análise envolve decompor, derivar, filtrar, agregar e estruturar dados para alinhá-los às tarefas cognitivas dos usuários.

Essa abordagem fundamenta o uso de KPIs como representações visuais de alto nível, criadas a partir de operações analíticas que reduzem complexidade e evidenciam padrões. No contexto organizacional, isso significa que métricas como tempo médio de atendimento, taxa de cancelamento e distribuição de chamadas derivam de um pipeline de abstração analítica semelhante ao defendido por Munzner.

Haroz et al. (2015) mostram que o raciocínio analítico depende das capacidades perceptuais e cognitivas dos usuários, o que torna a apresentação dos KPIs parte integrante do próprio processo de análise. As limitações da atenção visual, da percepção de padrões e da memória de trabalho demonstram que relatórios e dashboards devem ser construídos para minimizar a carga cognitiva.

No âmbito organizacional, isso justifica a adoção de boas práticas de análise descritiva (voltada a identificar tendências e magnitude), diagnóstica (orientada à explicação de causas) e preditiva (dedicada à estimativa de comportamentos futuros). Essas práticas, quando sustentadas por visualizações adequadas, ampliam a capacidade dos gestores de interpretar dados e reagir às variações operacionais.

Padilla et al. (2018) destacam que a tomada de decisão baseada em dados ocorre em três níveis: perceptual, cognitivo e decisório. No nível perceptual, dashboards e relatórios facilitam o reconhecimento visual de padrões; no nível cognitivo, os KPIs são integrados ao conhecimento prévio dos stakeholders; e no nível decisório, as inferências orientam escolhas estratégicas.

Desse modo, a análise de dados no contexto empresarial não é apenas técnica, mas um processo cognitivo distribuído que transforma métricas operacionais em conhecimento acionável, justificando a centralidade dos KPIs na gestão moderna. Nonato e Aupetit (2018), por sua vez, contribuem ao explicar que análises

mais complexas podem exigir técnicas de redução e projeção multidimensional para revelar estruturas ocultas entre variáveis.

Embora esse TCC não aplique diretamente projeções de alta dimensionalidade, o princípio é essencial: operações de agregação, filtragem e transformação (equivalentes, em menor escala, às técnicas de redução) são fundamentais para permitir que gestores interpretem dados corporativos. A ideia de “preservação de estrutura informacional” proposta pelos autores orienta a escolha criteriosa dos KPIs, garantindo que resumam o fenômeno sem distorções.

sistematiza a função do Pandas no pipeline analítico, destacando que a biblioteca é responsável pelas etapas de extração, limpeza, transformação e integração dos dados. A leitura de arquivos, a criação de DataFrames, o tratamento de valores ausentes (dropna, fillna), a padronização de tipos (astype), e as operações de junção e agregação (merge, groupby, agg) compõem o núcleo operacional da análise organizacional. (CHEN, 2023)

No contexto do presente estudo, essas operações permitiram harmonizar dados provenientes de várias fontes, remover inconsistências, consolidar métricas e derivar KPIs acionáveis. A natureza declarativa do Pandas possibilita transformar perguntas analíticas em comandos replicáveis, garantindo reprodutibilidade e rastreabilidade, bem como princípios fundamentais para qualquer processo de análise corporativa. (CHEN, 2023)

McKinney (2011), criador do Pandas, reforça que a biblioteca foi concebida para substituir pipelines analíticos fragmentados por um ecossistema unificado, capaz de manipular dados estruturados de maneira expressiva e eficiente. Os mecanismos de alinhamento automático, tratamento de valores faltantes e suporte nativo a operações tabulares complexas tornam o Pandas adequado para pipelines corporativos em que dados heterogêneos precisam ser consolidados rapidamente. Assim, a biblioteca materializa, em nível computacional, os princípios de análise estruturada e redução informacional discutidos por autores anteriores.

2.3 VISUALIZAÇÃO DE DADOS E DESIGN DE DASHBOARDS

A visualização de dados é apresentada como um processo cognitivo estruturado, no qual o framework DVL-FW articula sete dimensões fundamentais (necessidades de insight, escalas de dados, tipos de análise, tipos de visualização,

símbolos gráficos, variáveis gráficas e interações) que conectam teoria, prática e avaliação de competência visual (BÖRNER et al., 2019).

Ao enfatizar padrões de eficácia perceptual (por exemplo, a maior precisão da posição e do comprimento em relação à área ou à cor), o modelo demonstra que uma visualização eficaz depende da adequação entre a natureza dos dados, a tarefa cognitiva e a forma gráfica adotada. Nesse sentido, o design de visualizações para contextos corporativos e de mobilidade urbana não deve ser guiado por critérios estéticos, mas por evidências empíricas sobre percepção visual e literacia em visualização (BÖRNER et al., 2019).

Munzner (2014) afirma que a visualização é concebida como resultado de uma cadeia de decisões que conecta dados, tarefas e percepção humana, organizada metodologicamente pelo modelo *Why–What–How* e pelo pipeline de visualização. O “por quê” corresponde às perguntas analíticas dos stakeholders, o “o quê” à estrutura dos dados (tabelares, temporais, geográficos, relacionais) e o “como” às escolhas de codificação visual (posição, cor, forma, tamanho, interação).

Ao descrever a passagem de dados brutos para abstrações de dados, abstrações visuais e codificação, a autora fundamenta as boas práticas de design de dashboards: consistência gráfica, redução de elementos redundantes, preservação da estrutura dos dados e uso de interações (filtros, *drill-down*, *overview + detail*, *focus + context*) como parte do raciocínio analítico, não como mera camada de interface (MUNZNER, 2014).

Franconeri et al. (2021). A ciência da comunicação visual de dados reforça que a eficácia de um dashboard depende da sua capacidade de transmitir mensagens com clareza, precisão e economia cognitiva. A partir de princípios da psicologia da Gestalt (proximidade, similaridade, continuidade, fechamento) e de estudos sobre atenção visual e recursos pré-atentivos (cor, tamanho, orientação, movimento), os autores mostram que layouts bem estruturados, alinhamento espacial e consistência cromática reduzem o esforço mental necessário para detectar padrões.

Em dashboards corporativos, isso se traduz em organizar métricas relacionadas em blocos próximos, usar cores consistentes para categorias equivalentes e destacar exceções ou alertas com contrastes visuais controlados, criando painéis resilientes a interpretações equivocadas mesmo entre usuários com baixa literacia visual (FRANCONERI et al., 2021).

Fan et al., 2020. Ao tratar a visualização como etapa exploratória e comunicativa do ciclo estatístico, a obra enfatiza que diferentes representações visuais traduzem propriedades específicas dos dados: séries temporais revelam tendências e sazonalidades; gráficos de dispersão evidenciam correlações; gráficos de densidade e histogramas descrevem distribuições; e boxplots destacam variabilidade e outliers.

A escolha entre esses tipos de gráficos é, portanto, um ato metodológico. Para séries temporais, recomenda-se o uso de gráficos de linha, que maximizam a percepção de continuidade; para categorias, gráficos de barras favorecem comparações discretas; para proporções, gráficos de setores podem ser utilizados com poucas categorias, preferencialmente complementados por rótulos e valores absolutos; e, para padrões espaciais, mapas e heatmaps são indicados para representar distribuições geográficas ou intensidades regionais.

Essa visão converge com os princípios de codificação visual de Munzner (2014) e com a ênfase de Haroz, Kosara e Franconeri em alinhar tipo de dado, canal visual e tarefa cognitiva. Padilla et al. (2018), por outro lado, diz que a visualização é definida como ferramenta cognitiva para tomada de decisão, e não apenas como produto final da análise.

O framework proposto mostra que diferentes estruturas visuais favorecem raciocínios distintos (gráficos de linha para tendências, barras para comparação, mapas e dispersões para correlações espaciais) e introduz o conceito de feedback cognitivo, no qual a interação com visualizações interativas (filtros, drill-down, reconfiguração de eixos) permite que o usuário teste hipóteses, refine interpretações e aprenda com o próprio painel.

Além disso, os autores destacam a importância de representar explicitamente a incerteza (faixas de variação, intervalos de confiança, bandas de erro) e de organizar visualizações em narrativas analíticas coerentes, que conduzam o usuário de visões agregadas para detalhes contextuais, alinhando storytelling com suporte à decisão em ambientes data-driven.

Os resultados da análise e projeções multidimensionais de Nonato e Aupetit (2018), acrescenta uma camada importante à discussão dos tipos de visualização e de suas limitações perceptuais. Ao classificar distorções globais, locais e

semânticas, os autores mostram que toda visualização implica compromissos entre fidelidade e compreensibilidade, especialmente em cenários com muitas variáveis.

Técnicas de layout enrichment (como uso de cor para categorias, tamanho para magnitude e rótulos contextuais) demonstram como enriquecer visualizações sem sacrificá-las à sobrecarga cognitiva. Em dashboards para mobilidade urbana, esse princípio sustenta, por exemplo, o uso de mapas com camadas adicionais de informação (empresa, horário de pico, tipo de corrida), mantendo ao mesmo tempo a legibilidade das relações principais.

A integração entre análise com Pandas, visualização com Matplotlib e Seaborn e exportação de dados em formatos interoperáveis é apresentada como um pipeline prático para prototipagem visual e posterior construção de dashboards em plataformas como o Looker Studio (CHEN, 2017).

A visualização é tratada como componente contínuo do ciclo analítico: gráficos de linha, barras, dispersão e densidade são utilizados tanto para inspeção exploratória quanto para comunicação de resultados, articulando limpeza, transformação (pivot, melt, groupby, resample) e padronização de tipos (datas, categorias, valores ausentes) como etapas prévias a representações cognitivamente fiéis (CHEN, 2023).

Essa abordagem garante que os dados que alimentam os painéis mantenham coerência semântica, reduzindo o risco de distorções gráficas decorrentes de problemas de curadoria. McKinney et al., (2011) descrevem o Pandas como uma biblioteca fundacional que integra processamento analítico e visualização, ao oferecer estruturas de dados com índices e metadados que se propagam diretamente para gráficos.

Métodos como plot() e pivot_table() permitem gerar visualizações tabulares e gráficas reproduzíveis, que podem ser conectadas a sistemas de dashboard. O uso de hierarchical indexing habilita representações visuais que respeitam hierarquias (por exemplo, empresa–cidade–período), convergindo com os princípios de enriquecimento de layout e coordenação de múltiplas visões discutidos em visual analytics.

Essa camada de preparação e organização dos dados é fundamental para que o Looker Studio receba fontes consistentes, facilitando a criação de painéis estáveis e semanticamente claros.

Além disso, a aplicação do Looker Studio integrada ao BigQuery na análise do volume de resíduos por distrito ilustra, de forma concreta, como uma ferramenta de visualização pode operacionalizar princípios metodológicos e cognitivos de visualização em um contexto aplicado (YANTO et al., 2023).

Gráficos de barras, linhas, setores e mapas são combinados para permitir comparações por região, análise de tendências temporais e leitura espacial de padrões, sempre orientados por tarefas analíticas específicas. O estudo evidencia que a integração entre backend analítico e frontend visual reduz a complexidade perceptual ao agregar e filtrar dados em níveis adequados, reforçando a ideia de dashboards como instrumentos de redução perceptual e de mediação cognitiva entre especialistas e gestores (YANTO et al., 2023).

Dessa forma a construção de dashboards de vendas em tempo real com o Looker Studio, organizada em múltiplas páginas interativas, demonstra o potencial da ferramenta como plataforma de business intelligence em nuvem. A combinação de gráficos de barras, linhas e pizza, estruturados em hierarquias top-down (de indicadores agregados a detalhes por produto, cliente ou cidade), mostra como princípios de clareza informacional, consistência perceptual e storytelling visual podem ser incorporados ao design de painéis (SANDHIYASA et al., 2025).

A alta taxa de aprovação em testes de aceitação de usuários (UAT) ressalta que a eficácia da visualização depende não apenas da correção técnica, mas da usabilidade e da compreensão do público-alvo. Esses resultados sustentam a escolha do Looker Studio neste trabalho como ferramenta central para construção de dashboards de mobilidade urbana, nos quais interatividade, atualização contínua e coerência visual são requisitos para apoiar decisões baseadas em dados (SANDHIYASA et al., 2025).

2.4 MOBILIDADE URBANA E INDICADORES OPERACIONAIS

Hoque et al. (2012) mostram que os principais indicadores operacionais da mobilidade urbana derivam da análise integrada de padrões espaço-temporais e do comportamento dos veículos em circulação. Entre os indicadores fundamentais, destacam-se o tempo médio das viagens, o tempo de cruzeiro vazio, a densidade dos hotspots de embarque e desembarque, a evolução temporal da velocidade média e a conectividade entre veículos.

Esses elementos estruturam a compreensão da eficiência da frota e revelam que métricas como tempo médio de espera, taxas de cancelamento e padrões de origem das corridas são essenciais para avaliar a responsividade e o aproveitamento operacional dos sistemas de transporte.

A identificação de regiões com maior demanda, assim como a observação da frequência com que novos passageiros são encontrados, oferece um paralelo direto com métricas adotadas em plataformas corporativas, nas quais a origem dos chamados e o tempo de resposta são indicadores equivalentes da eficiência informacional (HOQUE et al., 2012).

Zhou et al. (2016) reforçam a importância estratégica da análise desses indicadores ao demonstrar que o monitoramento contínuo de trajetórias permite detectar desvios operacionais, estimar tempos médios de rota e identificar anomalias de comportamento em tempo real. Métricas como *trip time*, *route deviation ratio* e *detection latency* revelam gargalos e variações inesperadas que comprometem a qualidade do serviço.

No âmbito organizacional, a mesma lógica se aplica ao monitoramento de fluxos de dados e chamadas corporativas: compreender a origem dos chamados, a frequência dos cancelamentos e os padrões de processamento é essencial para garantir fluidez, previsibilidade e confiabilidade no desempenho do sistema (ZHOU et al., 2016).

Leng et al. (2016) evidenciam que indicadores como número médio de corridas, tempo ocioso, distância média percorrida e distribuição espacial de embarques permitem detectar distorções e identificar sazonalidades de comportamento no sistema urbano. No contexto de mobilidade por aplicativos, esses indicadores revelam padrões de deslocamento influenciados por incentivos, horários de pico e concentração espacial.

Assim como nas plataformas digitais de transporte, onde os métodos de pagamento revelam preferências dos usuários e afetam o tempo de atendimento, métricas operacionais em sistemas corporativos também dependem da análise detalhada de frequência, sazonalidade e distribuição das requisições, refletindo necessidades de ajuste dinâmico da operação.

Andriyanov (2020) acrescenta uma dimensão preditiva ao discutir indicadores derivados de séries temporais, como número de pedidos por unidade de tempo, tempo de atendimento e sazonalidade da demanda. Esses indicadores antecipam

sobrecargas, revelam padrões cíclicos de comportamento e subsidiam decisões sobre redistribuição de recursos.

A lógica de previsão aplicada aos pedidos de táxi encontra um paralelo direto na análise dos fluxos corporativos da Empresa X, em que padrões de chamada, latência operacional e recorrência de eventos formam a base para prever gargalos e adaptar estratégias de resposta (ANDRIYANOV, 2020).

Liu e Dijk (2022) enfatizam que indicadores como tempo de espera, padrões origem-destino, fluxo de congestionamento e níveis de acessibilidade constituem dimensões essenciais para medir sustentabilidade, eficiência e equidade dos sistemas urbanos. A análise conjunta de dados estatísticos, GIS e big data melhora a capacidade de compreender padrões urbanos e de apoiar decisões operacionais.

Na Empresa X, os mesmos princípios orientam a seleção de KPIs que representam o fluxo contínuo de dados: origem das operações, padrões de processamento, métodos de interação e recorrência de falhas constituem uma tradução organizacional dos indicadores urbanos, reforçando a importância da análise multiescalar.

Senaratne et al. (2018) demonstram que a análise de grandes volumes de dados digitais (como conexões a antenas de telecomunicação) permite identificar hotspots urbanos, padrões de deslocamento, áreas de residência e trabalho, e variações temporais de uso. A integração entre análise automática e visualização analítica possibilita detectar mudanças abruptas, ciclos de atividade e padrões anômalos.

Esses resultados fundamentam a compreensão de que indicadores operacionais não se limitam a medir desempenho, mas operam como sensores de mudanças estruturais. No contexto da Empresa X, a análise de origem dos chamados, horários de maior volume e padrões de resposta desempenha papel equivalente, permitindo identificar hotspots informacionais e sazonalidades que afetam sua operação (SENARATNE et al., 2018).

Vila et al. (2016) reforçam que a integração entre dados espaciais, operacionais e transacionais é fundamental para compreender gargalos, sobreposições de rotas e picos de demanda. Métricas como densidade de rotas, frequência de utilização, centralidade e inconsistências de registros evidenciam a complexidade do sistema urbano e a necessidade de mecanismos de qualidade e validação de dados.

A Empresa X enfrenta desafios semelhantes ao integrar múltiplas fontes de dados corporativos, demandando rastreabilidade, consistência e interoperabilidade para que indicadores como tempo médio de espera e cancelamentos produzam diagnósticos confiáveis.

Sobral et al. (2019) destacam que indicadores operacionais como velocidade média, densidade de tráfego, confiabilidade do serviço e tempo de espera tornam-se analiticamente úteis quando integrados a técnicas visuais que permitem explorar padrões, anomalias e tendências. A literatura demonstra que a visualização é parte indissociável da análise de mobilidade, pois traduz dados multivariados em estruturas cognitivamente interpretáveis.

Como no caso dos sistemas ITS, a Empresa X depende da capacidade de representar visualmente indicadores como origem dos chamados, tempos médios de atendimento e padrões de cancelamento, transformando métricas operacionais em conhecimento acionável para decisões estratégicas.

3 METODOLOGIA

O presente estudo adota uma abordagem metodológica de natureza aplicada, com caráter exploratório e método misto (qualitativo e quantitativo), alinhada às diretrizes contemporâneas para estudos empíricos no contexto da Engenharia de Software. Conforme discutido por Wohlin (2021), o estudo de caso caracteriza-se por investigar um fenômeno contemporâneo inserido em seu contexto real, utilizando múltiplas fontes de evidência e reconhecendo que a fronteira entre o fenômeno e o contexto é, frequentemente, difusa.

Essa definição fundamenta a escolha metodológica adotada, uma vez que a análise e visualização de dados operacionais de mobilidade urbana exigem, necessariamente, a compreensão das práticas, processos, fluxos informacionais e percepções dos stakeholders envolvidos no ambiente real da Empresa X.

A combinação de técnicas qualitativas e quantitativas permite capturar tanto evidências numéricas (referentes ao tratamento de dados, cálculo de métricas e uso de dashboards) quanto evidências interpretativas, relacionadas às percepções de usabilidade e utilidade das visualizações analíticas. Dessa forma, a abordagem

mista amplia a robustez da investigação e possibilita triangulação metodológica, recomendada em estudos de caso rigorosos (WOHLIN, 2021).

3.1 ESTRATÉGIA DO ESTUDO DE CASO

A estratégia de estudo de caso foi selecionada por atender aos critérios essenciais propostos por Wohlin (2021):

1. investigar um fenômeno contemporâneo,
2. imerso em um contexto real organizacional,
3. com coleta de dados a partir de múltiplas fontes e
4. sem intervenção ativa do pesquisador nos processos internos da empresa.

O estudo de caso é apropriado para compreender fenômenos complexos, como a transformação de dados operacionais dispersos em indicadores gerenciais visualmente compreensíveis, especialmente quando a fronteira entre o sistema tecnológico, os artefatos de dados e o processo decisório não é claramente delimitada.

Assim, a estratégia metodológica possibilita analisar o comportamento dos dados no ambiente vivo da organização, respeitando a relação entre fenômeno e contexto defendida pela literatura de estudos de caso.

3.2 CONTEXTO ORGANIZACIONAL E UNIDADE DE ANÁLISE

A unidade de análise selecionada é uma empresa brasileira de mobilidade urbana, atuante desde a década de 1980, prestando serviços de transporte corporativo e individual no modelo B2B (em sua maioria) e B2C.

A organização opera com centenas de motoristas, sistemas de rastreamento por GPS, centrais de atendimento e múltiplos canais de solicitação, o que gera um volume expressivo de dados operacionais heterogêneos; incluindo registros de corridas, cancelamentos, tempos de espera, meios de pagamento e padrões de demanda diária.

Esse contexto corporativo configura o estudo de caso como um ambiente adequado para examinar o fluxo completo dos dados: coleta, tratamento, curadoria, integração, cálculo de indicadores e visualização por gestores.

O fenômeno investigado está diretamente relacionado ao funcionamento contemporâneo da empresa, atendendo ao critério de “fenômeno atual em contexto real” estabelecido por Wohlin (2021).

3.3 PERGUNTAS DE PESQUISA

As perguntas de pesquisa estruturam o percurso analítico da investigação e se articulam diretamente com os objetivos anteriormente apresentados. A formulação dessas questões segue a orientação de que, em estudos empíricos, o processo investigativo deve partir de um problema claramente definido.

Desdobrando-se em perguntas que permitam compreender o fenômeno a partir de múltiplas perspectivas, conforme argumenta Wohlin (2021) ao tratar da importância do encadeamento lógico entre problema, questões e método .

PP1 — Quais indicadores e métricas operacionais são mais relevantes para avaliar o desempenho e a qualidade dos serviços em empresas de mobilidade urbana B2B?

A primeira pergunta de pesquisa (PP1) visa identificar os indicadores operacionais mais relevantes para avaliar o desempenho da Empresa X. Essa etapa envolve análise exploratória dos registros operacionais, identificação de variáveis críticas, padronização semântica dos campos e estruturação inicial dos KPIs.

PP2 — Como os dados operacionais podem ser transformados e visualizados para gerar insights estratégicos e apoiar a tomada de decisão em tempo real?

A segunda pergunta de pesquisa (PP2) examina como esses dados podem ser transformados e visualizados para gerar insights estratégicos, compondo a parte central do estudo ao integrar processamento técnico e design analítico. Aqui, constrói-se o pipeline de dados: extração, padronização, limpeza, enriquecimento, cálculo das métricas e elaboração dos dashboards.

As perguntas formam um processo sequencial coerente: identificar → transformar e visualizar → avaliar. Essa lógica progressiva caracteriza o estudo como um ciclo empírico iterativo, alinhado às orientações de pesquisas de caso descritas por Wohlin (2021), onde a compreensão do fenômeno é construída progressivamente por meio de coleta, análise e interpretação sistemática de múltiplas fontes de evidência.

3.4 PROCEDIMENTOS DE COLETA E PREPARAÇÃO DOS DADOS

A coleta de dados seguiu uma estrutura metodológica em três etapas, concebida para responder de forma direta às perguntas de pesquisa. As atividades foram planejadas de modo a permitir a triangulação entre evidências quantitativas, evidências qualitativas e evidências contextuais, como recomendado por Wohlin (2021) para fortalecer a validade do estudo de caso, seguindo as etapas abaixo:

- 1. Etapa 1 — Coleta e extração de dados operacionais:** A primeira etapa consistiu na extração de registros diretamente do sistema corporativo responsável por armazenar dados de mobilidade, incluindo tempos de espera, cancelamentos, origens dos chamados e tipos de pagamento. Esses dados foram exportados em formato bruto e preparados para análise por meio de scripts em Pandas. Seguindo a lógica metodológica de um estudo de caso, esta etapa fornece a principal fonte de evidência quantitativa.
- 2. Etapa 2 — Transformação, integração e construção de dashboards:** Após a coleta inicial, os dados foram submetidos a um ciclo de tratamento que incluiu limpeza, correção de inconsistências, padronização semântica de categorias, integração entre diferentes tabelas e derivação de novos indicadores, como médias, proporções e séries temporais. Esse processo materializa o elo entre análise técnica e significado organizacional, típico de estudos aplicados.

As visualizações foram então construídas no Looker Studio, permitindo representar graficamente padrões e anomalias de forma dinâmica. O uso de dashboards atende ao requisito de investigação contextual real, pois insere os resultados diretamente no ambiente de tomada de decisão.

3.5 PROCEDIMENTOS DE ANÁLISE DOS DADOS

A análise foi conduzida em múltiplas camadas, cada uma diretamente vinculada a uma das perguntas de pesquisa, o que caracteriza um desenho metodológico integrado e coerente. A imagem do fluxo do processo e ferramentas ilustra como ocorre a coleta, extração, transformação e visualização dos dados; utilizando as ferramentas mencionadas anteriormente.

Figura 1 - Fluxo do Processo e Ferramentas



Fonte: O autor.

Complementarmente, a seleção dos KPIs obedeceu a critérios explícitos que articularam literatura e necessidades organizacionais, de modo a garantir que cada indicador incluído nos dashboards tivesse função clara no suporte às decisões.

Em primeiro lugar, foram priorizadas métricas diretamente relacionadas às perguntas de pesquisa e aos objetivos do estudo, assegurando alinhamento entre o que é monitorado e o que se pretende compreender sobre o desempenho da Empresa X.

Em segundo lugar, considerou-se a disponibilidade, qualidade e consistência histórica dos dados nos sistemas operacionais, de forma que apenas variáveis passíveis de extração recorrente, com baixo nível de incompletude e sem ambiguidade semântica, fossem promovidas a indicadores.

Em terceiro lugar, adotou-se como critério a relevância percebida pelos diferentes perfis gerenciais mapeados na etapa de levantamento de requisitos: para que um KPI fosse incorporado ao modelo, precisava responder a uma pergunta concreta de negócio (por exemplo, “estamos atendendo dentro de um tempo de

espera aceitável?” ou “quais clientes apresentaram queda de demanda no período?”) e ser compreensível para os stakeholders que o utilizariam.

Em quarto lugar, avaliou-se a possibilidade de cálculo automatizado e atualização periódica no pipeline construído em Pandas e Looker Studio, evitando indicadores cuja manutenção dependesse de manipulações manuais ou processos ad hoc.

Por fim, buscou-se garantir capacidade de síntese e comunicação visual, selecionando KPIs que, em conjunto, oferecessem uma visão equilibrada de volume (número de corridas e chamados), qualidade do serviço (tempo de espera, cancelamentos), perfil de demanda (origem das solicitações, tipo de cliente) e receita (valores e meios de pagamento), sem redundâncias que sobrecarregasse a interface analítica.

3.5.1 Análise dos dados operacionais (PP1)

Os dados coletados foram submetidos a estatísticas descritivas, análises de distribuição, cálculos de proporções, identificação de padrões temporais e verificação de consistência. A estruturação dos KPIs emergiu desse processo, fornecendo uma base sólida para etapas posteriores do estudo.

3.5.2 Análise do pipeline e das visualizações (PP2)

A segunda camada analítica avaliou o fluxo de dados desde a extração até a visualização final. Nessa fase, o foco recaiu sobre a capacidade dos dashboards de revelar tendências, padrões de sazonalidade, variações operacionais e eventos críticos, assegurando que as visualizações fossem aderentes aos objetivos gerenciais da Empresa X.

3.6 AMEAÇAS À VALIDADE

A condução de um estudo de caso exige tratamento sistemático da validade e da confiabilidade, dimensões centrais discutidas por Wohlin (2021) para garantir rigor científico em pesquisas empíricas.

A **validade de construção** foi assegurada pela definição de KPIs fundamentados na literatura e nas necessidades da organização.

A **validade interna** foi tratada por meio da triangulação entre dados operacionais, percepções de usuários e evidências observadas durante a implementação dos dashboards.

A **validade externa** foi considerada pela descrição detalhada do contexto e das condições organizacionais, permitindo avaliar a transferibilidade dos resultados para empresas com características semelhantes.

A confiabilidade do estudo foi fortalecida pela documentação sistemática de todos os scripts, operações de limpeza e métodos de visualização, garantindo que o processo possa ser replicado por outros pesquisadores ou pela própria empresa.

Do ponto de vista ético, todos os dados analisados foram previamente anonimizados, e nenhuma informação sensível ou capaz de identificar indivíduos foi utilizada.

4 RESULTADOS

4.1 CARACTERIZAÇÃO DA EMPRESA X

Os resultados deste estudo de caso têm como base o estudo de caso da Empresa X, uma organização brasileira de mobilidade urbana com mais de quatro décadas de atuação contínua nos mercados B2B e B2C. A empresa opera de forma ininterrupta, 24 horas por dia, intermediando corridas entre passageiros e motoristas parceiros por meio de múltiplos canais digitais e tradicionais, como aplicativo móvel, call center, website e WhatsApp.

Atualmente, a Empresa X atende mais de 300 empresas clientes e mantém uma estrutura organizacional enxuta, composta por aproximadamente 40 colaboradores distribuídos entre os núcleos de Operações, Tecnologia da Informação, Comercial, Financeiro, Call Center e Administrativo.

Antes da implementação do modelo de análise e visualização de dados proposto neste trabalho, a empresa enfrentava limitações significativas em sua capacidade analítica. Os dados operacionais, embora abundantes, eram explorados de forma manual, fragmentada e predominantemente reativa, por meio de planilhas eletrônicas offline.

Essa abordagem resultava em análises mensais demoradas, dificultando a identificação tempestiva de tendências, gargalos operacionais e variações de desempenho, o que comprometia a agilidade do processo decisório gerencial.

A adoção de dashboards interativos e pipelines automatizados de dados passou a representar uma mudança estrutural na forma como a Empresa X observa e interpreta sua operação, ampliando substancialmente sua capacidade de monitoramento contínuo e tomada de decisão baseada em evidências.

4.2 DESENHO DA SOLUÇÃO

A idealização das visualizações partiu do princípio de que os gráficos deveriam permitir leitura rápida, associação semântica imediata e interpretação intuitiva, considerando o perfil heterogêneo dos stakeholders, que incluem desde gerentes operacionais até o diretor-presidente. Nesse sentido, optou-se

prioritariamente por gráficos de barras, linhas e pizza, cada um empregado de acordo com o tipo de métrica e objetivo analítico.

Os gráficos de linha foram utilizados para evidenciar tendências e padrões temporais, especialmente em séries históricas de corridas realizadas e canceladas. Os gráficos de barras foram empregados para comparações quantitativas entre períodos, categorias ou segmentos de clientes, enquanto os gráficos de pizza foram reservados para representar composições percentuais, como a proporção entre corridas realizadas e canceladas em um determinado intervalo temporal.

A escolha das cores seguiu uma lógica semântica consistente com o contexto operacional da empresa. A cor verde foi associada a valores financeiros, em função de sua relação cultural com receita e resultado econômico. A cor vermelha foi utilizada para corridas canceladas, simbolizando situações de alerta ou potencial problema operacional. Já a cor azul foi adotada para corridas realizadas, estabelecendo um contraponto visual claro e evitando sobreposição semântica com o verde já utilizado.

4.3 DEFINIÇÃO DE INDICADORES E REQUISITOS

O processo de construção dos dashboards iniciou-se com um levantamento sistemático de requisitos junto aos principais stakeholders da organização. Esse processo incluiu conversas estruturadas, observação do fluxo de trabalho e análise das necessidades específicas de cada perfil gerencial, conforme documentado no relatório de resultados.

Para o gerente operacional, destacaram-se indicadores como número de corridas realizadas, corridas canceladas, total de chamados, origem das solicitações e formas de pagamento. Para o gerente comercial, o foco recaiu sobre o acompanhamento gráfico do desempenho das empresas clientes, com ênfase na identificação de tendências de queda ou crescimento.

O gerente financeiro demandou métricas relacionadas a valores absolutos, ticket médio e distinção entre pessoas físicas e jurídicas. Já o gerente geral e o diretor-presidente priorizaram uma visão macro, com possibilidade de aprofundamento pontual e acesso facilitado via dispositivos móveis.

Essa etapa foi fundamental para garantir que os KPIs definidos servissem efetivamente de suporte às tarefas decisórias desempenhadas no cotidiano organizacional.

4.4 PIPELINE ANALÍTICO: EXTRAÇÃO, TRANSFORMAÇÃO E CARGA DOS DADOS

A etapa técnica do processo foi estruturada segundo um fluxo ETL (Extract, Transform, Load), implementado integralmente em Python. A extração dos dados operacionais foi automatizada por meio da biblioteca *mechanize*, responsável por simular o login no sistema da empresa e realizar o download periódico dos relatórios em formato *.xlsx*. O módulo **schedule** foi utilizado para agendar a execução diária do processo, enquanto o **logging** garantiu rastreabilidade e auditoria de todas as execuções.

Após a extração, os dados foram carregados no Python utilizando a biblioteca *Pandas*, sendo organizados em um *DataFrame*. A partir desse ponto, realizou-se um processo rigoroso de limpeza e seleção de variáveis (*feature selection*), reduzindo o conjunto original de 97 colunas para 28 colunas consideradas essenciais para as análises e visualizações.

Figura 2 - Seleção de Variáveis

```
#...(Após o processo de abrir a planilha no dataframe)

# Lista das colunas que serão removidas:
colunas_para_remover = [ 'Desconto %', 'Valor Desconto R$', 'eVoucher', 'Motorista',
    'Usuário Cad.', 'Transbordo', 'No.Matricula', 'No.CC', 'No.CPF',
    'Info Extra 1', 'Info Extra 2', 'Info Extra 3', 'Solicitante',
    'Passageiro', 'No.Telefone', 'Tel.Operadora', 'No.Retorno',
    'Ret.Operadora', 'No.Registro', 'Prefixo Motorista', 'Carro Moto',
    'Prev.Atend', 'Logradouro Origem', 'No.Origem', 'Compl.Origem',
    'Cep Origem', 'Bairro Origem', 'Cidade Origem', 'Ref.Origem',
    'Logradouro Destino', 'No.Destino', 'Compl.Destino', 'Cep Destino',
    'Bairro Destino', 'Cidade Destino', 'Ref.Destino', 'Pagto Extra',
    'Obs.Interna', 'Obs.Moto', 'Filtros', 'Cod Externo 1', 'Cod Externo 2',
    'Antecedencia', 'Preco Fechado', 'Cartao Titular', 'Cartao Bandeira',
    'Cartao No.', 'Cartao Validade', 'Transbordo Codinome', 'Transbordo Nome',
    'Transbordo Cliente', 'No.Prioridade', 'Bina Tronco', 'No.Boleto',
    'Hora Parada R$', 'Pedágio R$', 'Estacionamento R$', 'Algum Comentário',
    'Deduções Compart. R$', 'Deduções Compart. KM', 'Rodoanel', 'Tempo HP',
    'Distância HP', 'Taxa %', 'Valor Taxa R$', 'Desc. Tempo Parado',
    'Recebido por', 'Taxa Fixa', 'Valor Repasse Motorista',
    'Tempo HP(embarque atrasado)']

# Remove as colunas, ignorando se alguma não existir
df_final = df_final.drop(columns=[col for col in colunas_para_remover if col in df_final.columns])
```

Fonte: O Autor

Além da eliminação de colunas irrelevantes, foi necessário converter tipos de dados originalmente representados como texto para formatos numéricos e temporais

adequados. Valores monetários e quilometragem foram convertidos para float, enquanto campos de data foram transformados para o padrão aceito pelo Looker Studio (YYYY/MM/DD), garantindo compatibilidade com filtros temporais e agregações analíticas.

Figura 3 - Conversão dos tipos de dados

```
tabela = pd.read_excel(novo_diretorio_arquivo_apos_cortar_colunas)#

# Converter as colunas para valores numéricos
tabela['Prefixo Veiculo'] = pd.to_numeric(tabela['Prefixo Veiculo'], errors='coerce')
tabela['Valor'] = pd.to_numeric(tabela['Valor'], errors='coerce')
tabela['Valor Preco Fechado'] = pd.to_numeric(tabela['Valor Preco Fechado'], errors='coerce')
tabela['KMl'] = pd.to_numeric(tabela['KMl'], errors='coerce')

# Converter a coluna 'Dt.Agenda' para datetime informando o formato original
tabela['Dt.Agenda'] = pd.to_datetime(tabela['Dt.Agenda'], format='%d/%m/%Y', errors='coerce')

# Formatar a coluna 'Dt.Agenda' para o formato 'yyyy/mm/dd' como string
tabela['Dt.Agenda'] = tabela['Dt.Agenda'].dt.strftime('%Y/%m/%d')
```

Fonte: O autor.

Concluída a transformação, os dados foram carregados automaticamente em uma planilha central no Google Sheets, utilizando a API do Google por meio da biblioteca **gsread** e autenticação via **google.oauth2.service_account**. Esse procedimento assegurou atualização contínua da base histórica e integração direta com a camada de visualização.

Figura 4 - Processo de autenticação para upload dos dados

```
# Carrega a planilha
df = pd.read_excel(novo_diretorio_arquivo_apos_cortar_colunas_e_converter_dados)

# Preenche valores NaN com string vazia
df = df.fillna('')

# Configuração de autenticação
SCOPES = ['https://www.googleapis.com/auth/spreadsheets']
CREDENTIALS_FILE = '/home/admroot/2025/Google Sheets/lexical-raceway-458410-f9-4d147aa0bc8e.json'
SPREADSHEET_ID = '1k5gu51668fwgfsfs9onkbd6ab9vbwcsyntocy'
WORKSHEET_NAME = 'Sheet1'

# Autenticação
creds = Credentials.from_service_account_file(CREDENTIALS_FILE, scopes=SCOPES)
client = gsread.authorize(creds)
planilha = client.open_by_key(SPREADSHEET_ID)
worksheet = planilha.worksheet(WORKSHEET_NAME)

# Converte o DataFrame em uma lista de listas
linhas = df.values.tolist()

# Adiciona todas as linhas de uma só vez
worksheet.append_rows(linhas)
```

Fonte: O autor.

4.5 CONSTRUÇÃO DOS DASHBOARDS

Com os dados consolidados no Google Sheets, procedeu-se à criação dos dashboards no Looker Studio. O processo iniciou-se com a criação de um relatório em branco, seguida pela conexão da fonte de dados por meio da opção de planilhas Google, conforme apresentado em Conexão da Base de Dados no Looker Studio.

Figura 5 - Índice dos dados das corridas no Looker Studio

Campo ↑		Tipo
Agendamento	:	123 Número
Avaliacao Obs	:	ABC Texto
Avaliacao Pergunta	:	ABC Texto
Avaliacao Resposta	:	ABC Texto
Cadastro	:	Data e hora
Cliente	:	ABC Texto
Dt.Agenda	:	Data
Hr_Fim	:	ABC Texto
Hr_Inicio	:	ABC Texto
Hr.Agenda	:	ABC Texto
KM	:	123 Número
KM1	:	123 Número
No.Boleto	:	123 Número
No.Avaliacao	:	123 Número
No.Chamado	:	123 Número
No.Contrato	:	ABC Texto
Origem	:	ABC Texto
PA Destino	:	ABC Texto
PA Origem	:	ABC Texto
Pagto.	:	ABC Texto
Prefixo Veiculo	:	123 Número

Qtd.Chamado Anteriores ...	:	123 Número
Status	:	ABC Texto
Tipo	:	ABC Texto
Tranbordo	:	ABC Texto
Valor	:	123 Número
Valor Corrida R\$	:	123 Número
Valor Preco Fechado	:	123 Número

Fonte: O autor.

Na sequência, foram definidos filtros globais e específicos, permitindo segmentar os dados por status da corrida, tipo de cliente, forma de pagamento, tipo de veículo e período temporal. O processo de criação e gerenciamento desses filtros é fundamental para viabilizar análises personalizadas e comparativas.

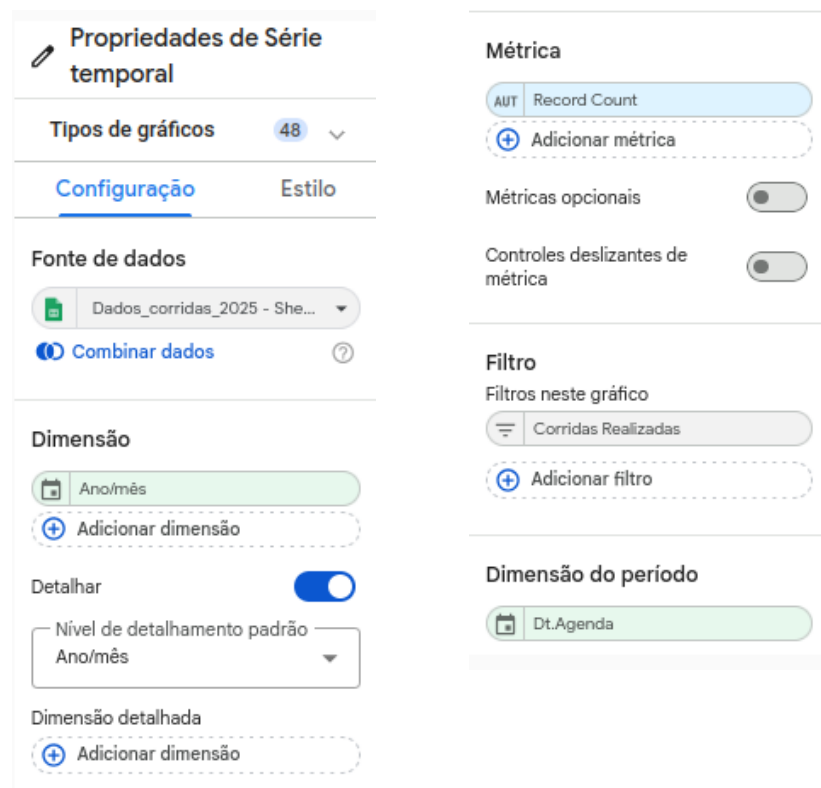
Figura 6 - Criação de filtros no Looker Studio

Filtros		
Nome	Usados no relatório	Descrição
Corridas Realizadas	25 gráficos	Incluir Status Igual a (=) Final
Corridas Canceladas	9 gráficos	Incluir Status Igual a (=) Cancelado
Pessoa Jurídica	12 gráficos	Incluir Pagto. Igual a (=) WPET
Pessoa Física	2 gráficos	Excluir Pagto. Igual a (=) WPET
Táxi	4 gráficos	Incluir Prefixo Veiculo Menor ou igual a (<=) 500
Veículo Particular	4 gráficos	Incluir Prefixo Veiculo Maior que (>) 500

Fonte: O autor.

A etapa final consistiu na criação dos gráficos propriamente ditos, utilizando os recursos nativos do Looker Studio. A seleção do tipo de gráfico foi realizada de acordo com a natureza da métrica analisada, conforme demonstrado nas imagens abaixo.

Figura 7 - Configuração dos gráficos no LookerStudio



Fonte: O autor.

4.6 RESULTADOS VISUAIS

Os dashboards resultantes possibilitaram a visualização integrada de indicadores operacionais críticos. A versão mobile destinada ao diretor-presidente, ilustrada na imagem abaixo, apresenta de forma resumida o total de chamados, a proporção entre corridas realizadas e canceladas e a distribuição por tipo de cliente e categoria de motorista.

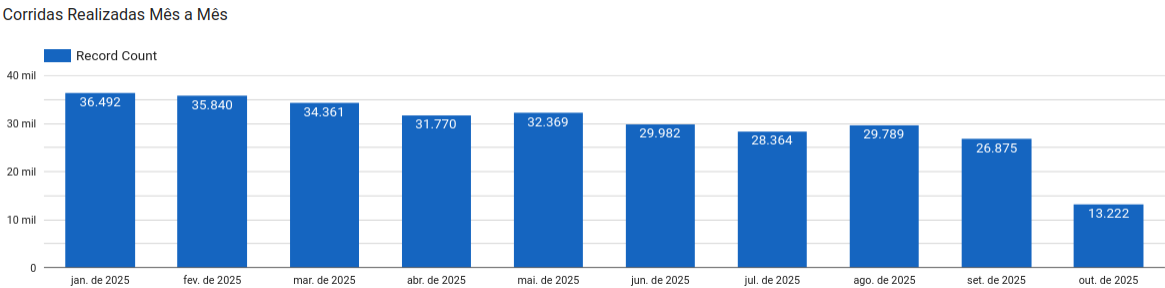
Figura 8 - Visão Geral do Dashboard Mobile



Fonte: O autor.

No nível operacional, destacam-se os gráficos das figuras abaixo, que permitem identificar tendências temporais, sazonalidades e variações diárias no volume de serviços.

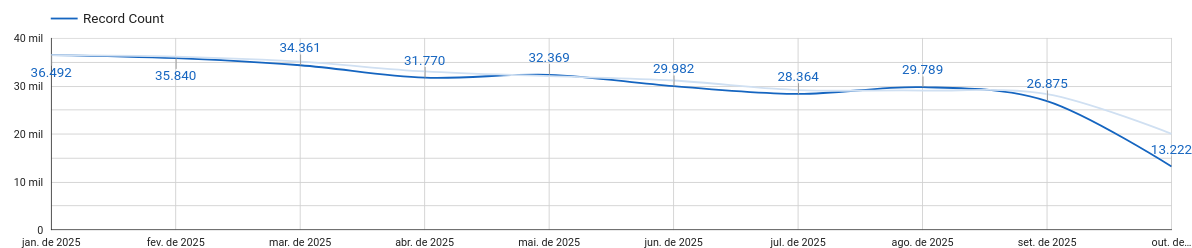
Figura 9 - Gráfico em barra das corridas realizadas mensalmente



Fonte: O autor.

Figura 10 - Gráfico em linha das corridas realizadas mensalmente

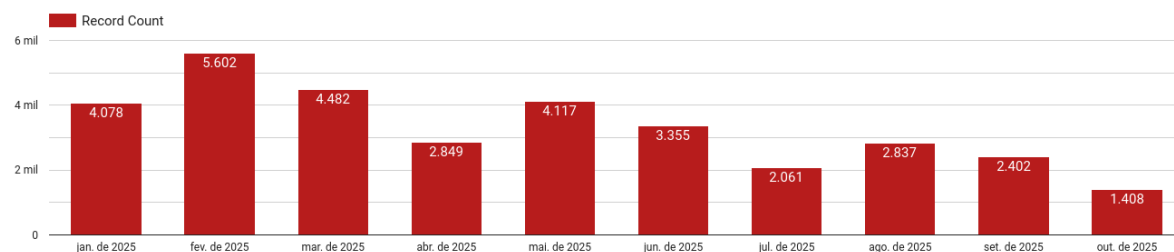
Corridas Realizadas Mês a Mês



Fonte: O autor.

Figura 11 - Gráfico em barra das corridas canceladas mensalmente

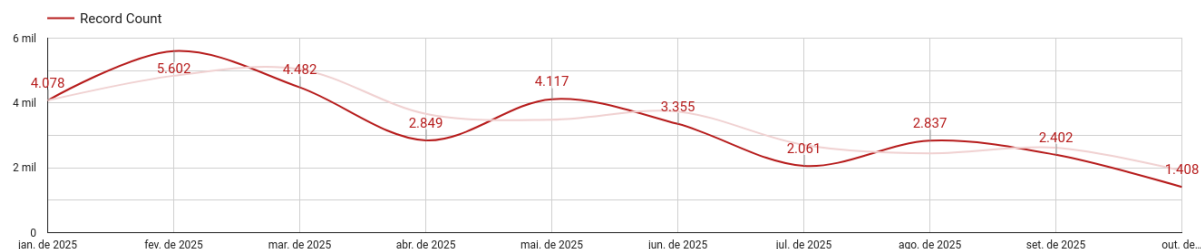
Corridas Canceladas Mês a Mês



Fonte: O autor.

Figura 12 - Gráfico em linha das corridas canceladas mensalmente

Corridas Canceladas Mês a Mês

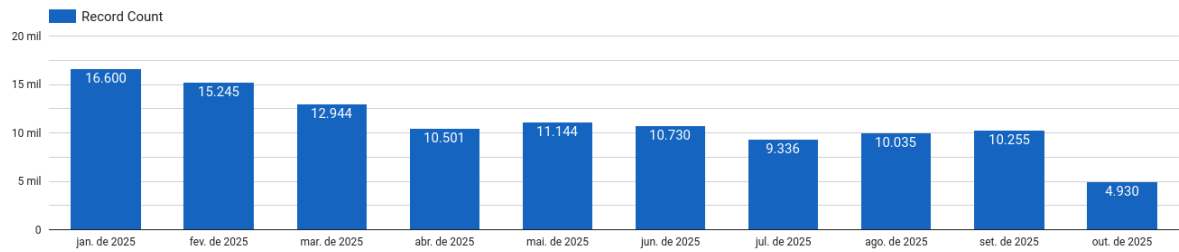


Fonte: O autor.

As análises segmentadas por pessoa jurídica são apresentadas nas figuras abaixo, possibilitando acompanhamento específico do desempenho do mercado corporativo.

Figura 13 - Gráfico em barra das corridas realizadas mensalmente para PJ

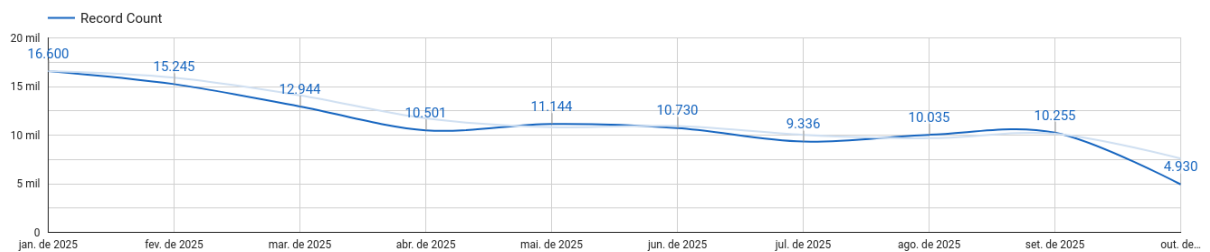
PJ - Corridas Realizadas Mês a Mês



Fonte: O autor.

Figura 14 - Gráfico em linha das corridas realizadas mensalmente para PJ

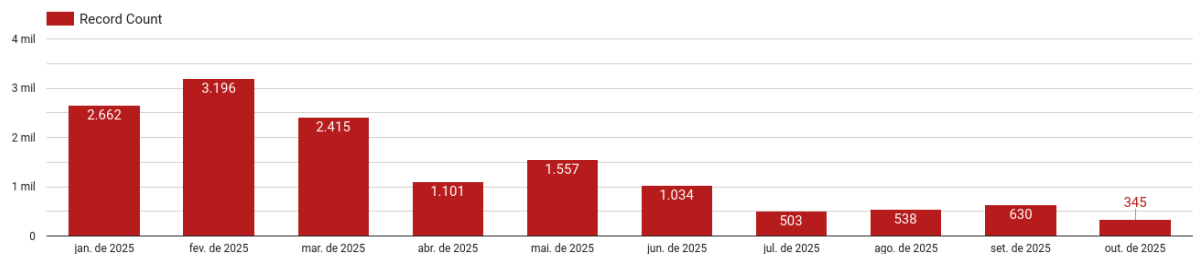
PJ - Corridas Realizadas Mês a Mês



Fonte: O autor.

Figura 15 - Gráfico em barra das corridas canceladas mensalmente para PJ

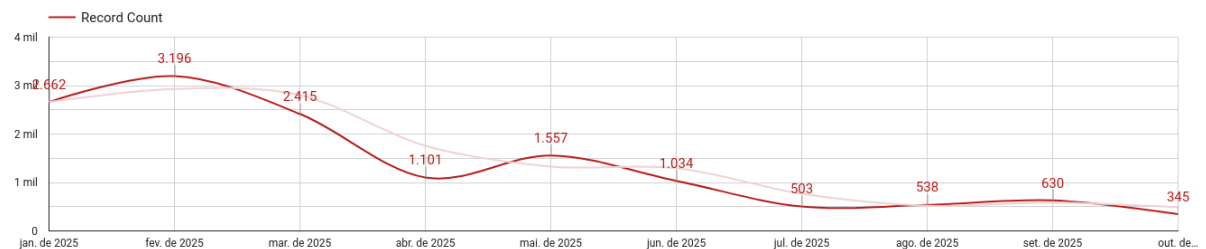
PJ - Corridas Canceladas Mês a Mês



Fonte: O autor.

Figura 16 - Gráfico em linha das corridas canceladas mensalmente para PJ

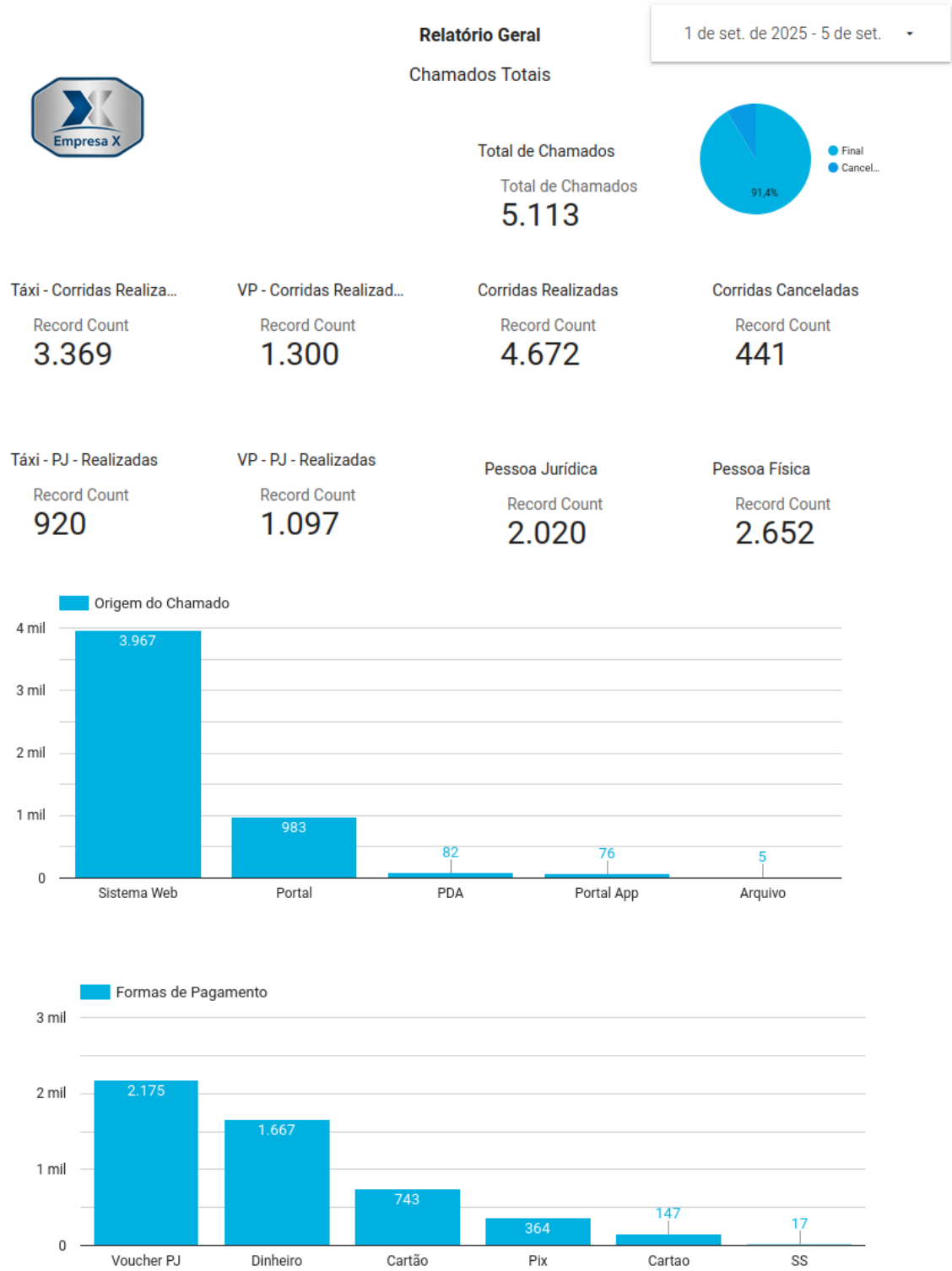
PJ - Corridas Canceladas Mês a Mês



Fonte: O autor.

Adicionalmente, os dashboards incluem visualizações de origem dos chamados e formas de pagamento, ampliando a compreensão sobre o comportamento dos usuários e a eficiência dos canais de atendimento.

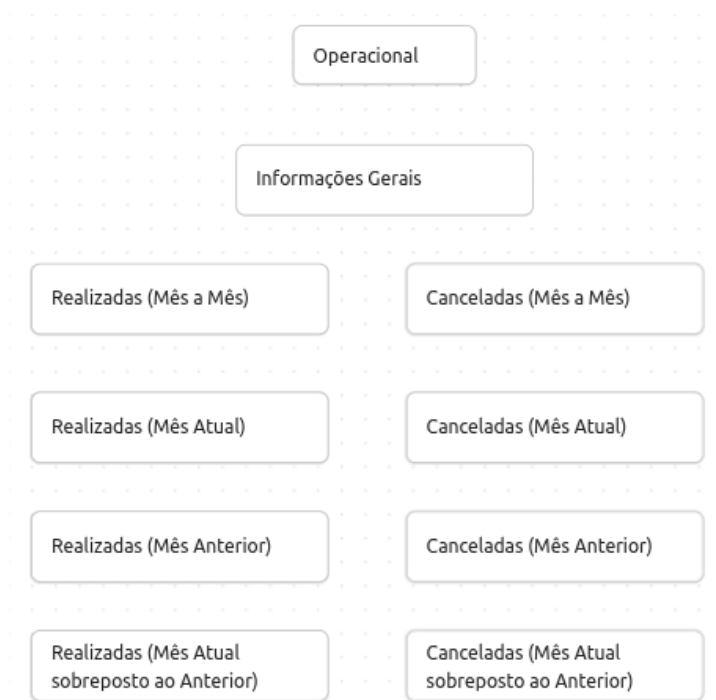
Figura 17 - Relatório Geral e Origem dos Chamados



Fonte: O autor.

Por fim, a organização dos dashboards (e relatórios) por perfil gerencial (operacional, financeiro, comercial, gerente geral e diretor-presidente) é sintetizada nas imagens de representação gráfica dos dashboards por perfil, evidenciando a adequação da solução às diferentes necessidades decisórias da empresa.

Figura 18 - Representação Gráfica dos Relatórios do Gerente Operacional



Fonte: O autor.

Figura 19 - Representação Gráfica dos Relatórios do Gerente Financeiro



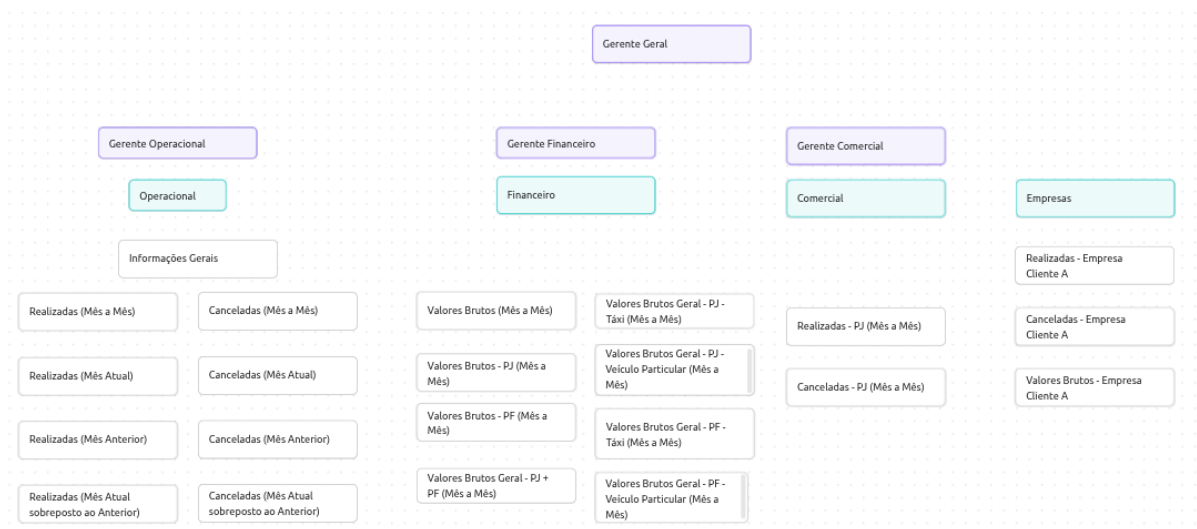
Fonte: O autor.

Figura 20 - Representação Gráfica dos Relatórios do Gerente Comercial



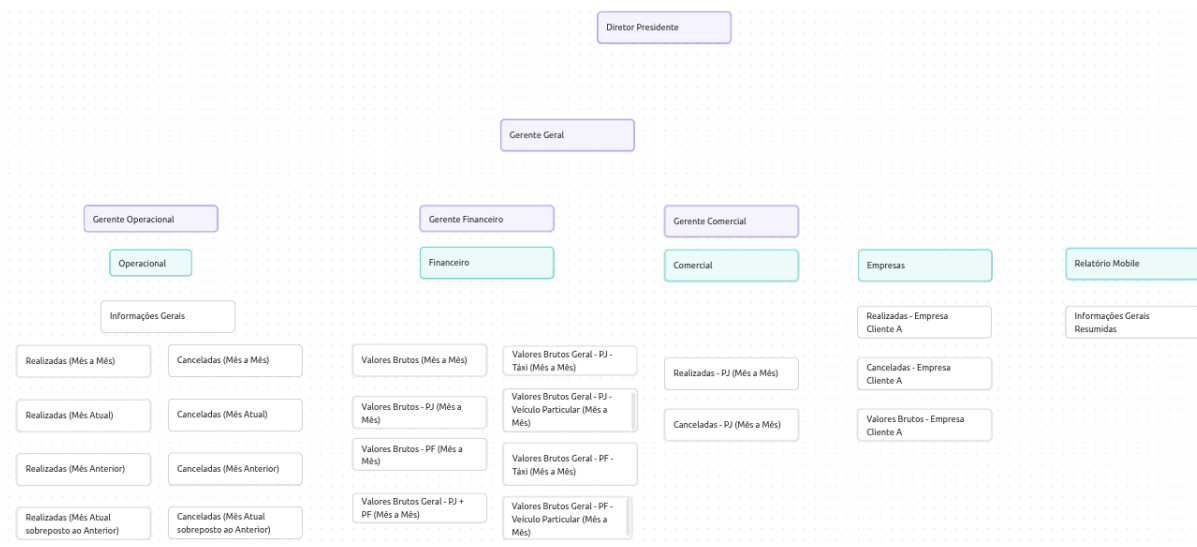
Fonte: O autor.

Figura 21 - Representação Gráfica dos Relatórios do Gerente Geral



Fonte: O autor.

Figura 22 - Representação Gráfica dos Relatórios do Diretor Presidente



Fonte: O autor.

5 DISCUSSÃO

Wohlin (2021) define o estudo de caso como uma estratégia adequada quando o fenômeno investigado é contemporâneo, ocorre em contexto real e demanda múltiplas fontes de evidência, com fronteira difusa entre tecnologia, processos e decisões.

Essa caracterização é consistente com o problema enfrentado pela Empresa X, a dificuldade de consolidar e interpretar dados operacionais heterogêneos (corridas, cancelamentos, tempos de espera, canais de solicitação e meios de pagamento) para identificar padrões e gargalos em tempo hábil para gestão.

Börner et al. (2019) sustentam que dados só se convertem em conhecimento acionável quando há competências para ler e construir visualizações de forma competente, o que posiciona dashboards não como “saída estética”, mas como infraestrutura cognitiva para a organização.

Na prática do caso, os resultados indicam uma tentativa explícita de materializar essa cadeia dado → informação → conhecimento por meio de um modelo em três etapas: (i) definição de KPIs; (ii) pipeline analítico com Pandas; e (iii) dashboards interativos no Looker Studio.

Munzner (2014) propõe que o design de visualizações deve ser orientado pelo encadeamento Why–What–How: tarefas (por quê), estrutura dos dados (o quê) e escolhas de codificação/interação (como).

Ao confrontar esse referencial com a seção de resultados, observa-se alinhamento metodológico na medida em que a definição de indicadores (o “what”) é tratada como etapa antecedente e condicionante das decisões de dashboard, evitando que o painel seja um artefato “supply-driven” (guiado pela disponibilidade dos campos) e aproximando-o de uma lógica “task-driven” (guiada pelas perguntas dos stakeholders).

Esse acoplamento entre tarefa gerencial e estrutura de dados aparece quando o texto explicita perfis distintos de uso (ex.: visão macro com aprofundamento pontual e acesso móvel) como condicionantes de quais métricas foram priorizadas e de como seriam exploradas por filtros e recortes.

5.1 INDICADORES OPERACIONAIS E CONVERGÊNCIA COM A LITERATURA DE MOBILIDADE

Sobral et al. (2019) argumentam que indicadores operacionais (como tempo de espera, confiabilidade e padrões de demanda) tornam-se úteis quando integrados a técnicas visuais que possibilitam explorar padrões, anomalias e tendências. Em consonância, a proposta do caso adota precisamente a premissa de que os KPIs precisam ser “operacionalizados” como elementos visualmente exploráveis, e não apenas listados em relatórios.

Isso aparece no próprio racional do estudo, ao enfatizar a necessidade de identificar padrões, gargalos e oportunidades “em tempo real” (ou em acompanhamento contínuo), o que desloca a análise de um uso retrospectivo para um uso de suporte à decisão.

Wohlin (2021) também reforça que a robustez analítica, em estudos de caso, decorre de encadear coleta/análise às perguntas de pesquisa e de triangular evidências. A estrutura RQ1–RQ2–RQ3 (identificar → transformar/visualizar → avaliar) reforça essa progressão e ajuda a explicar por que, no caso, os “resultados” não são apenas gráficos isolados, mas um sistema analítico que tenta sustentar continuidade e reuso dos indicadores.

5.2 DO GRÁFICO AO SISTEMA: IMPLICAÇÕES DO PIPELINE EM PANDAS PARA CONFIABILIDADE E GOVERNANÇA

Wohlin (2021) ressalta que, em estudo de caso, a credibilidade do achado depende de rastreabilidade e tratamento explícito de validade e confiabilidade. No caso, o pipeline ETL automatizado em Python, com extração periódica, agendamento e registro de execuções, funciona como mecanismo de rastreabilidade operacional do dado que alimenta o dashboard, isto é, não se trata apenas de “produzir um gráfico”, mas de criar condições para repetição e auditoria do processo.

A seção de resultados descreve uma etapa crítica que dialoga diretamente com “qualidade do dado” e “coerência semântica” discutidas no referencial: a redução de dimensionalidade/seleção de variáveis, que diminui o conjunto de colunas de 97 para 28 variáveis essenciais.

Do ponto de vista interpretativo, essa decisão é coerente com a literatura de visual analytics e com o princípio de reduzir ruído para preservar sinal: quanto mais atributos redundantes/irrelevantes são carregados ao ambiente de visualização, maior a chance de inconsistências, filtros incoerentes e interpretações equivocadas. Assim, a **Figura 2 - Seleção de Variáveis** é um resultado metodológico em si: ela evidencia a etapa de curadoria como pré-condição para confiabilidade visual.

O mesmo raciocínio se aplica à normalização de tipos e formatos de data (compatíveis com agregações e filtros temporais), bem como à conversão de valores numéricos (moeda e quilometragem) para tipos adequados. Essas transformações não são apenas detalhes de implementação: elas sustentam a validade de construção ao reduzir distorções decorrentes de parsing incorreto, agrupamentos inconsistentes e séries temporais inválidas. Nesse sentido, a **Figura 3 - Conversão dos tipos de dados** atua como evidência de que a preparação do dado foi tratada como parte do “método”, e não como etapa implícita.

Finalmente, o carregamento automatizado para uma base central (Google Sheets) e a integração direta com o Looker Studio reforçam a noção de “linha de produção” do indicador: o KPI deixa de ser um cálculo pontual e passa a existir como artefato continuamente atualizável e consumível pela camada gerencial.

Isso contribui para o argumento de ganhos operacionais esperados (redução de retrabalho e tempo de consolidação manual), ainda que, na discussão, esses ganhos devam ser tratados como efeitos plausíveis do desenho implementado; não como resultados mensurados, caso a seção de resultados não tenha apresentado métricas de before/after.

5.3 DASHBOARDS COMO MEDIAÇÃO COGNITIVA: ESCOLHAS DE VISUALIZAÇÃO E ESTRUTURA ANALÍTICA

Munzner (2014) defende que dashboards eficazes preservam consistência, reduzem redundâncias e usam interação (filtros, drill-down, overview+detail) como parte do raciocínio analítico. Ao comparar isso com os resultados, observa-se que a construção do relatório no Looker Studio começa pela conexão estruturada da base e pela definição de filtros globais e específicos para segmentação por status de corrida, tipo de cliente, forma de pagamento, tipo de veículo e recorte temporal.

Essa decisão é teoricamente consistente com a ideia de “feedback cognitivo” e exploração iterativa: o usuário não recebe um painel estático, mas um ambiente para formular e testar hipóteses operacionais por recortes sucessivos. **A Figura 5 - Índice dos dados das corridas no Looker Studio e a referência à Conexão da Base de Dados no Looker Studio** sustentam, no texto, a evidência de implementação dessa camada de interação.

Franconeri et al. (2021) enfatizam que organização espacial (Gestalt) e consistência cromática reduzem carga cognitiva e aumentam robustez interpretativa, especialmente entre usuários com baixa literacia visual.

Embora a seção de resultados precise explicitar com maior precisão quais foram as convenções de cor adotadas (por exemplo, se cores representam categorias fixas como “status”, “tipo de cliente” e “meio de pagamento”, ou se foram usadas para alertas e exceções), o próprio referencial estabelece critérios objetivos para justificar tais escolhas: cores devem ser consistentes por categoria e contrastes devem ser controlados para destacar exceções sem gerar ruído.

Assim, quando os resultados descrevem o painel como instrumento de identificação de tendências, sazonalidade e eventos críticos, o fundamento teórico pertinente é que a codificação visual deve ser deliberada e orientada a tarefas (comparar, identificar picos, localizar anomalias), e não meramente decorativa.

Padilla et al. (2018) reforçam a noção de “cognitive fit”: forma gráfica e tarefa precisam estar alinhadas para acelerar decisões e reduzir interpretações incorretas. A discussão, portanto, deve interpretar os resultados considerando se os tipos de gráficos escolhidos efetivamente correspondem às tarefas gerenciais predominantes: séries temporais para tendências e sazonalidades, barras para comparações discretas, e gráficos de proporção apenas quando há poucas categorias e com rótulos/valores que reduzam ambiguidade.

Quando a seção de resultados apresenta dashboards que permitem comparar períodos, segmentar perfis e identificar anomalias, ela se alinha com esse princípio de ajuste cognitivo, sobretudo ao estruturar o painel de modo que o usuário transite entre visão agregada e detalhe filtrado.

5.4 DIVERGÊNCIAS RELEVANTES E OPORTUNIDADES DE FORTALECIMENTO ACADÊMICO

Haroze, Kosara e Franconeri (e textos correlatos do referencial) destacam a importância de representar explicitamente incerteza (bandas, faixas, intervalos) e de reconhecer limites perceptuais e distorções inerentes à visualizações.

Ao confrontar esse ponto com o que foi descrito nos resultados (pipeline, filtros, dashboards), emerge uma lacuna potencial: o estudo parece focar fortemente na “construção e disponibilização” dos indicadores, mas não evidencia, pelo menos nos trechos citados, mecanismos explícitos de visualização/gestão de incerteza (por exemplo, dados faltantes, discrepâncias de registros, janelas de atualização, atrasos de extração, ou confiabilidade por fonte).

Isso não invalida os resultados; ao contrário, cria uma oportunidade de discussão acadêmica: em serviços de mobilidade urbana, onde cancelamentos, tempos e canais podem variar por qualidade de registro e latência de atualização, explicitar incerteza pode elevar a maturidade analítica do painel e reduzir decisões baseadas em falso senso de precisão.

Wohlin (2021) também recomenda discutir ameaças à validade de forma transparente. O trabalho já encaminha essa discussão ao explicitar validade de construção (KPIs fundamentados em literatura e necessidades), validade interna (triangulação) e tratamento sistemático de confiabilidade.

Nesta seção, esse material pode ser articulado com os resultados técnicos para sustentar que a rastreabilidade (logging), automação e curadoria (seleção de variáveis e tipagem) são mecanismos concretos de mitigação de ameaças, em vez de apenas declarações metodológicas.

Wohlin (2021) e Munzner (2014) sustentam que, em estudos aplicados, o valor científico emerge quando o artefato produzido (pipeline + dashboard) é interpretado como resposta sistemática a um problema real e como modelo potencialmente replicável.

Nesse sentido, os resultados do caso da Empresa X devem ser valorizados como uma demonstração de ponta a ponta: desde a extração automatizada, passando por curadoria e padronização, até a operacionalização no Looker Studio com filtros e visão gerencial.

Ao integrar essa cadeia, o estudo reforça o argumento de que a eficiência operacional e a capacidade decisória tendem a se beneficiar não apenas de “mais dados”, mas de dados mais tratáveis, rastreáveis e visualmente exploráveis, com redução de dependência de consolidações manuais e percepção subjetiva.

Börner et al. (2019) e Padilla et al. (2018) por sua vez, permitem concluir no plano teórico, que a principal contribuição do caso não está somente na criação de gráficos, mas na criação de condições para uma cultura data-driven: acesso recorrente, compreensão facilitada e suporte à formulação de hipóteses por interação.

A discussão, portanto, pode enquadrar os resultados como evidência de amadurecimento da “mobilidade informacional” da empresa, isto é, da capacidade de fazer os dados “circularem” com integridade e significado até o ponto de decisão; ainda que a mensuração final dessa adoção cultural demande evidências complementares.

6 CONCLUSÃO

Os resultados obtidos neste estudo também evidenciam um impacto concreto na forma como a Empresa X acessa e utiliza inteligência de negócios. Antes da construção do pipeline analítico com Python (Pandas) e dos dashboards no Looker Studio, a realidade da organização era marcada pela dependência de relatórios manuais e pela ausência de uma camada de BI sustentável, em parte porque soluções corporativas robustas, como SAP, Power BI ou Tableau, não se mostravam financeiramente viáveis nem proporcionalmente justificáveis para uma empresa de pequeno porte.

A solução desenvolvida demonstra que é possível estruturar um ambiente analítico consistente sem recorrer a plataformas de alto custo, desde que exista um fluxo bem definido de tratamento de dados e um desenho visual coerente com as necessidades de decisão.

Do ponto de vista prático, a adoção de uma ferramenta gratuita, de boa usabilidade e com características low code, como o Looker Studio, configurou-se como alternativa tecnicamente adequada e organizacionalmente realista. Os dashboards construídos alinham-se diretamente às demandas dos diferentes stakeholders da Empresa X, oferecendo visualizações simples, intuitivas e alinhadas ao vocabulário de negócio dos gestores operacional, comercial, financeiro e estratégico.

Em vez de uma solução genérica e superdimensionada, o estudo entregou um conjunto de painéis focalizados, capaz de responder a perguntas concretas — como comportamento de clientes ao longo do tempo, variação de demanda, concentração de cancelamentos e padrões de pagamento — com o mínimo de fricção no uso cotidiano.

Nesse sentido, o caso da Empresa X reforça a ideia de que “business intelligence é para todos”: ao combinar ferramentas acessíveis com um processo cuidadoso de seleção de KPIs e desenho de visualizações, este trabalho mostra que pequenas empresas também podem se beneficiar de práticas de análise de dados avançadas, tradicionalmente associadas a grandes corporações.

A experiência relatada indica que a barreira principal não é apenas tecnológica ou financeira, mas metodológica: quando os dados são estruturados, os indicadores são bem definidos e as visualizações são pensadas a partir do usuário

final, a inteligência de negócios deixa de ser um privilégio de organizações com grande capacidade de investimento.

Por fim, mesmo não sendo uma empresa de grande porte, a Empresa X passa a operar com um nível superior de maturidade analítica, o que potencializa vantagem competitiva em seu mercado. A possibilidade de acompanhar indicadores críticos em tempo quase real, identificar tendências com antecedência e ajustar operações com base em evidências coloca a organização em posição diferenciada frente a concorrentes que ainda dependem exclusivamente de relatórios manuais ou percepções intuitivas.

Assim, os resultados deste estudo sugerem que a combinação entre análise de dados, visual analytics e ferramentas acessíveis não apenas resolve um problema local da Empresa X, mas também aponta um caminho viável para que outras empresas de mobilidade urbana de pequeno e médio porte incorporem práticas de BI de forma gradual, sustentável e alinhada à sua realidade operacional.

REFERÊNCIAS

- BÖRNER, Katy; BUECKLE, Andreas; GINDA, Michael. **Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments**. Proceedings of the National Academy of Sciences, v. 116, n. 6, p. 1857-1864, 2019.
- FAN, Jianqing et al. **Statistical foundations of data science**. Chapman and Hall/CRC, 2020.
- MUNZNER, Tamara. **Visualization analysis and design**. In: Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Courses. 2025. p. 1-2.
- FRANCONERI, Steven L. et al. **The science of visual data communication: What works**. Psychological Science in the public interest, v. 22, n. 3, p. 110-161, 2021.
- NONATO, Luis Gustavo; AUPETIT, Michael. **Multidimensional projection for visual analytics: Linking techniques with distortions, tasks, and layout enrichment**. IEEE Transactions on Visualization and Computer Graphics, v. 25, n. 8, p. 2650-2673, 2018.
- PADILLA, Lace M. et al. **Decision making with visualizations: a cognitive framework across disciplines**. Cognitive research: principles and implications, v. 3, n. 1, p. 29, 2018.
- CHEN, Daniel Y. **Pandas for everyone: Python data analysis**. Addison-Wesley Professional, 2017.
- MCKINNEY, Wes et al. **pandas: a foundational Python library for data analysis and statistics**. Python for high performance and scientific computing, v. 14, n. 9, p. 1-9, 2011.
- LI, Shaomeng et al. **Data reduction techniques for simulation, visualization and data analysis**. In: Computer graphics forum. 2018. p. 422-447.
- YANTO, Budi; SUDARYANTO, Aris; PRATIWI, Hasri Ainun. **Data Visualization Analysis of Waste Production Volume in Every District of Tangerang Regency in 2021 Using Looker Studio and Big Query Platforms**. 2023.
- SANDHIYASA, I. Made Subrata et al. **Real Time Sales Data Visualization at Agung Wastra Using Looker Studio**. TECHNOVATE: Journal of Information Technology and Strategic Innovation Management, v. 2, n. 2, p. 96-103, 2025.
- HOQUE, Mohammad Asadul; HONG, Xiaoyan; DIXON, Brandon. **Analysis of mobility patterns for urban taxi cabs**. In: 2012 international conference on computing, networking and communications (ICNC). IEEE, 2012. p. 756-760.
- ZHOU, Zuojian et al. **A method for real-time trajectory monitoring to improve taxi service using GPS big data**. Information & Management, v. 53, n. 8, p. 964-977, 2016.

LENG, Biao et al. **Analysis of taxi drivers' behaviors within a battle between two taxi apps**. IEEE Transactions on Intelligent Transportation Systems, v. 17, n. 1, p. 296-300, 2015.

ANDRIYANOV, Nikita A. **Development of prediction methods for taxi order service on the basis of intellectual data analysis**. In: Science and Information Conference. Cham: Springer International Publishing, 2020. p. 652-664.

LIU, Xu; DIJK, Marc. **The role of data in sustainability assessment of urban mobility policies**. Data & Policy, v. 4, p. e2, 2022.

SENARATNE, Hansi et al. **Urban mobility analysis with mobile network data: A visual analytics approach**. IEEE Transactions on Intelligent Transportation Systems, v. 19, n. 5, p. 1537-1546, 2017.

MASSOBRIO, Renzo; NESMACHNOW, Sergio. **Urban mobility data analysis for public transportation systems: a case study in Montevideo, Uruguay**. Applied Sciences, v. 10, n. 16, p. 5400, 2020.

SOBRAL, Thiago; GALVÃO, Teresa; BORGES, José. **Visualization of urban mobility data from intelligent transportation systems**. Sensors, v. 19, n. 2, p. 332, 2019.

VILA, Juan Jose Rodriguez et al. **Urban mobility challenges—an exploratory analysis of public transportation data in Curitiba**. Revista de Informática Aplicada, v. 12, n. 1, 2016.

WOHLIN, Claes. **Case Study Research in Software Engineering—It is a Case, and it is a Study, but is it a Case Study?**. Information and Software Technology, v. 133, p. 106514, 2021.