



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE ENGENHARIAS E TECNOLOGIA
CURSO DE ENGENHARIA DE SOFTWARE

ERIKY ABREU VELOSO

ANÁLISE DO MERCADO DE LOCAÇÃO DE VEÍCULOS NO BRASIL:
UMA ABORDAGEM BASEADA EM DADOS

PAU DOS FERROS

2026

ERIKY ABREU VELOSO

**ANÁLISE DO MERCADO DE LOCAÇÃO DE VEÍCULOS NO BRASIL:
UMA ABORDAGEM BASEADA EM DADOS**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Engenharia de Software.

Orientador: Reudismam Rolim de Sousa Prof. Dr.

Co-orientador: João Batista de Souza Neto Prof. Dr.

PAU DOS FERROS

2026

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

V432a Veloso, Eriky Abreu.
Análise do mercado de locação de veículos no
Brasil: uma abordagem baseada em dados / Eriky
Abreu Veloso. - 2026.
68 f. : il.

Orientador: Reudismam Rolim de Sousa.
Coorientador: João Batista de Souza Neto.
Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Engenharia de
Software, 2026.

1. Locação de veículos. 2. Ciência de dados. 3.
Análise exploratória. 4. Dashboard. I. Sousa,
Reudismam Rolim de, orient. II. Souza Neto, João
Batista de, co-orient. III. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo(a) autor(a).

Biblioteca Campus Pau dos Ferros
Bibliotecária: Erlanda Maria Lopes da Silva
CRB: 15/966

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

ERIKY ABREU VELOSO

**ANÁLISE DO MERCADO DE LOCAÇÃO DE VEÍCULOS NO BRASIL:
UMA ABORDAGEM BASEADA EM DADOS**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Engenharia de Software.

Defendida em: 06 / 07 / 2026

BANCA EXAMINADORA

Reudismam Rolim de Sousa, Prof. Dr. (UFERSA)
Presidente

João Batista de Souza Neto Prof. Dr., Prof. Dr. (IFRN)
Co-orientador

George Felipe Fernandes Vieira, Bel. (UFERSA)
Membro Examinador

Letícia Maria Bandeira de Lucena, Bel.a. (Examinadora Externa)
Membro Examinador

DEDICATÓRIA

Dedico este trabalho à memória de minha querida avó, Isabel Alexandre da Rocha. Guardarei para sempre a lembrança de sua alegria contagiante, de seu amor pela dança e de todos os momentos especiais que tivemos juntos (*In Memoriam*).

AGRADECIMENTOS

Gostaria de iniciar agradecendo aos três pilares da minha vida. Sem eles, nenhuma de minhas conquistas seria possível. À Edna da Rocha Abreu, uma mulher guerreira e exemplo de mãe, que sempre se dedicou ao máximo para realizar os sonhos de seus filhos, muitas vezes colocando-os acima dos seus próprios. À Érica Abreu Veloso, minha irmã mais velha, que sempre esteve ao meu lado, guardando meus segredos, compartilhando minhas alegrias e me apoiando em todas as minhas decisões. Ao Nivaldo da Costa Veloso, um pai que nunca mediu esforços para proporcionar a melhor educação possível aos seus filhos. A esses três, deixo minha eterna gratidão.

Posteriormente, gostaria de agradecer às minhas outras mães, Edneide Abreu e Neide Abreu. Sou profundamente grato por cada conselho, por todo o auxílio ao longo da minha trajetória e pela confiança que sempre depositaram em mim. Agradeço também aos meus primos, Leonardo Abreu e Ana Clara, que sempre estiveram presentes tanto nos momentos de descontração em João Pessoa quanto nas inúmeras reflexões sobre os desafios da vida adulta.

Também gostaria de expressar minha gratidão a Rhemo Guedes, por todos os momentos de alegria, raiva e angústia que vivenciamos juntos. Ao longo dessa caminhada, você se mostrou um verdadeiro ponto de apoio, um conselheiro e um incrível parceiro. Agradeço também a Jussier David que, apesar da distância, continua ocupando um lugar muito especial em minha vida. Mesmo longe, seu amor de irmão sempre esteve presente, e sou profundamente grato por isso.

Expresso meu carinho e minha gratidão a todos os meus amigos de Areia Branca, que ao longo dos anos sempre me proporcionaram momentos de alegria e companheirismo. Em especial, agradeço a Amélia Sena, minha maior confidente, alguém em quem deposito minha confiança e meu amor; a Antônio Neto, sempre responsável pelas melhores conversas e pelos melhores *memes*, um amigo com quem posso contar em qualquer momento; a Camila Kelly, a personificação da bondade e da alegria, um exemplo de ser humano que inspira todos ao seu redor; a Jennifer Vieira, dona das maiores viagens e teorias, parceira das conversas mais improváveis, por quem meu amor só aumentam a cada dia e, por fim, mas não menos importante, a Vinícius Nascimento, companheiro de incontáveis aventuras e aquele amigo que consigo compreender apenas com um olhar.

Não poderia deixar de agradecer por todos os laços construídos em Pau dos Ferros. Sou eternamente grato às amigadas que tornaram minha trajetória nesta cidade mais leve e repleta de boas memórias. Em especial, agradeço a Diego Pinheiro, um amigo que

compartilhou comigo momentos de extrema felicidade e também de profunda tristeza. Responsável por me apresentar ao universo de One Piece e por pilotar uma moto pela cidade para caçar Pokémon, transformando momentos simples em memória. Geovan Gomes, que me acompanha desde o início da graduação, proporcionando momentos de alegria, algumas doses de raiva e inúmeras experiências que jamais serão esquecidas.

Também ressalto Andrey Sabino, sempre trazendo as histórias mais loucas e engraçadas, além de ser uma fonte constante de energia, apoio e boas risadas. Pedro Guilherme, uma das pessoas mais humildes e engraçadas que tive o prazer de conhecer, cujo bom humor tornaram muitos dias melhores. Edglê Santana, sinônimo de tranquilidade, sempre disposto a encarar qualquer tipo de “rolê”. Por fim, agradeço a Stheffany Câmara, uma patricinha simpática, obrigado por todo o apoio, pela amizade e pela disposição em ajudar nos mais diversos momentos dessa trajetória.

Desejo testemunhar a minha gratidão ao meu orientador, Professor Reudismam Rolim de Sousa, e ao meu coorientador, Professor João Batista de Souza Neto, por toda a orientação, dedicação e apoio ao longo do desenvolvimento deste trabalho. Agradeço pela confiança depositada em mim, pelos ensinamentos compartilhados, pela disponibilidade em esclarecer dúvidas e pelas valiosas contribuições que foram fundamentais para a realização desta pesquisa.

Também gostaria de agradecer à banca examinadora, pela disponibilidade, atenção e comprometimento em avaliar este trabalho, contribuindo com observações e sugestões valiosas para seu aperfeiçoamento. Por fim, deixo registrada minha gratidão a todos os familiares, professores, amigos e colegas que, de alguma forma, fizeram parte da minha trajetória e contribuíram para minha formação pessoal, acadêmica e profissional.

EPÍGRAFE

“Eu acredito que não existam coincidências no mundo. Tudo acaba acontecendo por algum motivo. O nosso destino vai sendo moldado aos poucos.”

(Silers Rayleigh)

RESUMO

O mercado de locação de veículos possui grande relevância econômica no cenário nacional, sendo um segmento que apresenta uma produção contínua de dados. No entanto, a forma como essas informações são disponibilizadas nos relatórios oficiais do setor, frequentemente de maneira fragmentada e sem recursos visuais auxiliares, eleva a complexidade da leitura analítica e dificulta a identificação de padrões e tendências de mercado. Diante desse cenário, o presente estudo teve como objetivo desenvolver uma base de dados unificada e um *dashboard* interativo voltados à análise do mercado de locação de veículos no Brasil, com um enfoque maior para a região Nordeste, compreendendo o período de 2004 a 2025. A metodologia adotada consistiu na extração automatizada dos dados dos anuários oficiais com a biblioteca *pdfplumber*, seguida pelas etapas de limpeza, tratamento e estruturação das tabelas nas camadas da Arquitetura Medallion, utilizando como ferramenta para esses processos a biblioteca Pandas. Como forma de promover a reprodutibilidade, os conjuntos de dados foram disponibilizados publicamente na plataforma Kaggle, servindo como um *dataset* para pesquisas na área de locação veicular. Em relação aos resultados gerados pelas análises, percebe-se uma grande expansão e resiliência financeira do setor. No que diz respeito à distribuição de mercado, os dados revelam uma acentuada polarização geográfica no estado de São Paulo, que concentra a maior parcela de locadoras do país. Além disso, a análise de frotas identificou uma clara preferência de compra voltada aos subsegmentos de *Hatches* Pequenos e SUVs, alinhando-se diretamente às demandas de consumo da população. Conclui-se que a centralização dos dados e a estruturação do *dashboard* converteram informações dispersas em um produto real, auxiliando na otimização do processo de tomada de decisão estratégica e possibilitando novas investigações no setor.

Palavras-chave: locação de veículos; ciência de dados; análise exploratória; *dashboard*.

ABSTRACT

The vehicle rental market holds major economic relevance on the national stage, representing a segment that features a continuous production of data. However, the way this information is provided in official industry reports, often in a fragmented manner and lacking auxiliary visual resources, increases the complexity of analytical reading and hinders the identification of market patterns and trends. Given this scenario, this study aimed to develop a unified database and an interactive dashboard for analyzing the vehicle rental market in Brazil, focusing more deeply on the Northeast region and covering the period from 2004 to 2025. The methodology consisted of automated data extraction from official yearbooks using the pdfplumber library, followed by the stages of cleaning, processing, and structuring tables within the Medallion Architecture layers, employing the Pandas library as the tool for these processes. To promote reproducibility, the datasets were made publicly available on the Kaggle platform, serving as a dataset for research in the vehicular rental field. Regarding the results generated by the analyses, a major expansion and financial resilience of the sector can be observed. With respect to market distribution, the data reveal a sharp geographical polarization in the state of São Paulo, which concentrates the largest share of rental companies in the country. In addition, fleet analysis identified a clear purchase preference toward the Small Hatchback and SUV subsegments, directly aligning with population consumption demands. It is concluded that data centralization and dashboard structuring converted scattered information into a tangible product, aiding in the optimization of the strategic decision-making process and enabling further investigations in the sector.

Keywords: vehicle rental; data science; exploratory data analysis; dashboard.

LISTA DE FIGURAS

Figura 1	– Interface inicial do DMLV	43
Figura 2	– Dashboard com as informações nacionais (Parte 1)	44
Figura 3	– Dashboard com as informações nacionais (Parte 2)	46
Figura 4	– Dashboard com as informações estaduais (Parte 1)	48
Figura 5	– Dashboard com as informações estaduais (Parte 2)	49

LISTA DE GRÁFICOS

Gráfico 1	– Oscilações do faturamento líquido ao longo dos anos	51
Gráfico 2	– Oscilações do faturamento bruto ao longo dos anos	51
Gráfico 3	– Evolução da quantidade de locadoras ao longo dos anos	52
Gráfico 4	– Distribuição geográfica das locadoras no Brasil	53
Gráfico 5	– Evolução da quantidade de usuários ao longo dos anos	57
Gráfico 6	– Distribuição dos empregados por gênero ao longo dos anos	58
Gráfico 7	– Número de emplacamentos ao longo dos anos	59
Gráfico 8	– Número de emplacamentos por fabricante ao longo dos anos	60
Gráfico 9	– Número de emplacamentos por subsegmento ao longo dos anos	61

LISTA DE QUADROS

Quadro 1	–	Comparação entre os trabalhos relacionados e o presente trabalho	35
----------	---	--	----

LISTA DE TABELAS

Tabela 1	– Distribuição geográfica das locadoras - Sudeste (2025)	54
Tabela 2	– Distribuição geográfica das locadoras - Nordeste (2025)	55
Tabela 3	– Distribuição geográfica das locadoras - Sul (2025)	55
Tabela 4	– Distribuição geográfica das locadoras - Norte (2025)	56
Tabela 5	– Distribuição geográfica das locadoras - Centro-Oeste (2025)	56

LISTA DE ABREVIATURAS E SIGLAS

ABLA	Associação Brasileira das Locadoras de Automóveis
ABSLV	Anuário Brasileiro do Setor de Locação de Veículos
API	Application Programming Interface
BI	Business Intelligence
CD	Ciência de Dados
covid-19	Corona Virus Disease 2019
CSS	Cascading Style Sheets
DMLV	Dashboard do Mercado de Locação de Veículos
GTF	Gestão e Terceirização de Frotas
KPI	Key Performance Indicator
OCR	Optical Character Recognition
PDF	Portable Document Format
RaC	Rent a Car
SUV	Sport Utility Vehicle
UI	Interfaces de usuário
UX	Experiências de usuário

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Justificativa e problemática	15
1.2	Objetivos e contribuições	16
1.3	Organização do trabalho	17
2	REFERENCIAL TEÓRICO	18
2.1	Locação de veículos	18
2.1.1	Locação de veículos no brasil	18
2.1.2	Associação Brasileira das Locadoras de Automóveis	19
2.2	Ciência de Dados	20
2.2.1	Aquisição dos dados	21
2.2.2	Limpeza dos dados	22
2.2.3	Uso dos dados	23
2.2.4	Tipos de gráficos	23
2.3	Ferramentas e tecnologias	24
3	TRABALHOS RELACIONADOS	28
3.1	Ciência de Dados aplicada à tomada de decisão	28
3.2	Desafios técnicos da extração de dados	31
3.3	<i>Interfaces</i> de suporte à decisão baseada em dados	32
3.4	Mercado de locação de veículos	32
3.5	Síntese analítica dos trabalhos relacionados	33
4	METODOLOGIA	37
4.1	Aquisição dos dados	37
4.2	Limpeza dos dados	38
4.3	Uso dos dados	39
4.4	Publicação dos dados	40
5	RESULTADOS	41
5.1	Estruturação e disponibilização dos dados coletados	41
5.2	<i>Dashboard</i> para análise do mercado de locação de veículos	42
5.3	Insights e conclusões das análises gráficas	50
6	CONCLUSÕES E TRABALHOS FUTUROS	63
	REFERÊNCIAS	66

1 INTRODUÇÃO

O mercado de locação de veículos no Brasil iniciou sua trajetória na década de 1950, impulsionado pela expansão automotiva do Plano de Metas, e profissionalizou-se em 1956 com a fundação da primeira empresa exclusiva do setor (Luiz, 2018). O segmento evoluiu de um serviço de nicho para uma indústria madura, dividida entre o atendimento ao varejo e a terceirização de frotas corporativas (Dolle *et al.*, 2025). Esse crescimento consolidou o setor como um pilar estratégico para a economia do país, gerando forte arrecadação tributária e centenas de milhares de empregos (Luiz, 2018).

Atualmente, a relevância do setor reflete uma transformação na mobilidade urbana, posicionando as locadoras como as maiores compradoras das montadoras de veículos instaladas no país (Esteve *et al.*, 2020). Diante da magnitude desse mercado, torna-se indispensável analisar seus indicadores para auxiliar o planejamento estratégico nesse segmento.

Para isso, a fonte de dados de maior relevância é o Anuário Brasileiro do Setor de Locação de Veículos (ABSLV), principal publicação estatística do mercado de aluguel de carros no Brasil, elaborado e divulgado anualmente pela Associação Brasileira das Locadoras de Automóveis (ABLA) (Pucci, 2026). A ABLA é a entidade de classe oficial que representa o setor de locação de veículos no cenário nacional. Atuando na defesa dos interesses do segmento perante órgãos governamentais e a sociedade, a associação desempenha um papel estratégico ao promover o desenvolvimento do mercado de locação (Evelyn, 2026).

1.1 Justificativa e problemática

O crescimento exponencial na geração de dados têm impulsionado a necessidade de métodos eficientes para coleta, organização e análise de informações, especialmente em contextos que demandam suporte à tomada de decisão (Morettin; Singer, 2020), como o mercado de locação de veículos. Nesse cenário, a Ciência de Dados (CD) surge como uma área fundamental, integrando técnicas estatísticas, computacionais e analíticas com o objetivo de transformar dados brutos em conhecimento útil (Morettin; Singer, 2020).

Nesse contexto, percebe-se que a organização e a estruturação apropriada dos dados são essenciais para interpretar as informações corretamente. Informações não estruturadas ou distribuídas, por não aderirem a um formato padrão, trazem uma complexidade maior para análise e entendimento, demandando um esforço adicional para a extração de sentido (Santos,

2024). Na CD, a etapa de exibição possui extrema importância. A falta de técnicas apropriadas de visualização pode afetar diretamente a interpretação dos resultados, já que representações visuais bem construídas ajudam os usuários a entender e atribuir significado aos dados (Cairo, 2019).

Posto isto, a maneira como os dados do Anuário Brasileiro do Setor de Locação de Veículos são apresentados, frequentemente de forma fragmentada ou sem o auxílio de recursos visuais apropriados, pode aumentar a complexidade de uma leitura analítica. Isso não só compromete a compreensão das informações, mas também dificulta a identificação de padrões, tendências e relações importantes para o setor.

Considerando o crescimento e uma significativa importância econômica do mercado de locação de veículos, e compreendendo a importância das técnicas aplicadas ao tratamento e exposição das informações, torna-se evidente a importância de estabelecer uma base de dados unificada que possa integrar as informações, que hoje estão dispersas nos ABSLVs.

1.2 Objetivos e contribuições

O presente trabalho tem como objetivo geral desenvolver uma base de dados unificada e um *dashboard* interativo para analisar o mercado de locação de veículos. Com isso, o estudo oferece uma contribuição direta ao centralizar e estruturar esses registros históricos em um repositório unificado, reduzindo o esforço manual de busca em múltiplos documentos isolados. Entretanto, o maior impacto da pesquisa consiste na criação do *dashboard* interativo, que atua como uma ferramenta de inteligência de dados capaz de converter tabelas estáticas em visualizações dinâmicas e de fácil interpretação. Esse painel otimiza o tempo de análise de tendências temporais e regionais, servindo como um suporte confiável para tomadores de decisão, pesquisadores e investidores que buscam compreender as oscilações e o comportamento do mercado de locação nacional, compreendendo o recorte temporal de 2004 a 2025.

Para alcançar a solução proposta, a pesquisa adota uma abordagem estruturada em etapas. Inicialmente, os dados brutos extraídos dos anuários passam por um *pipeline* de engenharia de dados utilizando a linguagem *Python*, em que são organizados seguindo os preceitos da arquitetura *Medallion*; nessa estrutura, as informações brutas são armazenadas na camada *Bronze*, refinadas e padronizadas na camada *Silver*, e consolidadas em uma base de dados unificada na camada *Gold* (Strengtholt, 2025).

Por fim, para atingir os objetivos anteriormente descritos, a base de dados e o protótipo desenvolvido, alimentaram a construção do *dashboard* interativo via *framework Streamlit*, que recebeu estilizações personalizadas em CSS, obtendo como produto uma ferramenta para a geração de *insights* estratégicos sobre o setor de locação de carros.

1.3 Organização do trabalho

Além desta Introdução, este trabalho está estruturado em outros cinco capítulos. O Capítulo 2 traz o Referencial Teórico com os conceitos essenciais que fundamentam a pesquisa, seguido pelo Capítulo 3, que aborda os Trabalhos Relacionados ao apresentar estudos da área e suas respectivas comparações com este projeto. A Metodologia é detalhada no Capítulo 4, expondo todo o processo de engenharia de dados e o desenvolvimento do painel visual. Na sequência, o Capítulo 5 demonstra e discute os Resultados obtidos por meio dos produtos gerados e das análises realizadas. Por fim, o Capítulo 6 encerra o trabalho com as Conclusões e Trabalhos Futuros, sintetizando as contribuições do estudo e propondo novos caminhos de investigação e desenvolvimento.

2 REFERENCIAL TEÓRICO

Neste capítulo, serão abordados os principais conceitos e fundamentos teóricos necessários para a compreensão e contextualização do presente estudo. Com esse intuito, a seção 2.1 contextualiza o panorama da locação de veículos, detalhando suas modalidades de atuação e sua evolução histórica no cenário brasileiro. Posteriormente, na seção 2.2, discute-se o campo da CD, explorando seu ciclo de vida por meio das etapas sequenciais de aquisição, limpeza e uso dos dados, além de detalhar os tipos de representações gráficas mais adequadas para a análise de informações. Por fim, a seção 2.3 apresenta as principais ferramentas e tecnologias computacionais selecionadas para a execução prática deste trabalho, destacando os *softwares*, bibliotecas e linguagens que auxiliaram o *pipeline* de extração, processamento e visualização dos dados.

2.1 Locação de veículos

A atividade de locação veicular consiste no modelo de negócio em que a geração de receita ocorre por meio de contratos de aluguel da frota, estruturando-se comercialmente em dois eixos principais: o *Rent a Car* (RaC) e a Gestão e Terceirização de Frotas (GTF). O segmento de RaC é caracterizado pelo aluguel de curto prazo, geralmente diário ou semanal, voltado tanto para pessoas físicas quanto jurídicas (Dolle *et al.*, 2025).

Compensatoriamente, a Gestão e Terceirização de Frotas concentra-se em contratos de longo prazo, com média de 24 a 36 meses, atendendo a empresas privadas e órgãos públicos por meio da locação e administração completa dos veículos, incluindo serviços de manutenção, seguros e documentação, além de abranger os modelos de carro por assinatura (Dolle *et al.*, 2025).

2.1.1 Locação de veículos no Brasil

O setor de locação de veículos no Brasil começou a se desenvolver na década de 1950, impulsionado pelo crescimento da indústria automobilística do país durante o Plano de Metas de Juscelino Kubitschek (Luiz, 2018). O movimento teve início de maneira espontânea no centro de São Paulo, no qual comerciantes de automóveis usados começaram a disponibilizar o aluguel como um serviço adicional. Contudo, foi em 1956 que o setor se profissionalizou de fato, com a criação da Auto Drive S.A. por Adalberto Camargo e outros sócios (Luiz, 2018).

Em sua monografia, Luiz (2018), afirma que o setor de aluguel de carros estabeleceu-se como um pilar estratégico da economia brasileira, tendo um impacto significativo tanto no progresso financeiro quanto no bem-estar social do país. Ao longo dos últimos anos, o setor alcançou um nível elevado de maturidade, evidenciando um crescimento constante ao decorrer do tempo. Além disso, sua importância é evidenciada pela significativa contribuição fiscal e pelo papel essencial na criação de oportunidades de trabalho, mantendo uma ampla rede de empregos diretos e indiretos.

Na pesquisa realizada por Esteve *et al.* (2020), o mesmo alega que o crescimento rápido do setor de locação tem atraído o interesse direto de montadoras e investidores que veem no setor um futuro promissor. Além de um crescimento no mercado, nota-se uma transformação nos padrões de consumo: a procura tradicional pelo aluguel voltado ao turismo ou à gestão de frotas está sendo substituída por uma nova perspectiva sobre a mobilidade urbana. Essa mudança redefine a relação dos brasileiros com a posse de veículos, priorizando o acesso ao serviço em vez da propriedade privada e consolidando uma verdadeira revolução no comportamento do consumidor.

2.1.2 Associação Brasileira das Locadoras de Automóveis

A Associação Brasileira das Locadoras de Automóveis é a entidade de classe oficial que representa o setor de locação de veículos no cenário nacional. Fundada com o propósito de profissionalizar e fortalecer o segmento, a associação atua como um órgão representativo central na defesa dos interesses das empresas locadoras perante instâncias governamentais, entidades reguladoras e a sociedade civil. A entidade desempenha um papel estratégico ao promover o desenvolvimento do mercado e ao consolidar o aluguel de automóveis como um pilar indispensável para a mobilidade urbana e a infraestrutura econômica do Brasil (Evelyn, 2026).

Além de sua função institucional, a associação atua como importante difusora de informações sobre o mercado, coletando, consolidando e divulgando indicadores estatísticos relacionados à frota, faturamento e empregabilidade do setor por meio de seus relatórios e anuários.

2.2 Ciência de Dados

Singer (2020) define Ciência de Dados como a progressão prática da Estatística, fortalecida pela combinação de modelos matemáticos tradicionais com a atual e poderosa capacidade de processamento computacional. Conjuntamente, os autores alegam que apesar da CD ser comumente descrita como uma nova área de estudo, ela se baseia em conceitos e algoritmos, como árvores de decisão e análise exploratória de dados, que foram desenvolvidos há várias décadas, mas que, na época, enfrentavam restrições de processamento.

Desse modo, a CD é caracterizada por um fluxo de trabalho que prioriza a preparação, descrição e visualização dos dados antes da modelagem confirmatória. Isso é feito por meio do uso da computação atual para aplicar técnicas estatísticas complexas em larga escala, conferindo nova visibilidade e aplicabilidade ao ofício estatístico tradicional (Morettin; Singer, 2020).

Segundo Carvalho, Menezes e Bonidia (2024), a CD caracteriza-se como uma área do conhecimento voltada ao estudo e à aplicação de princípios e técnicas, operacionalizados por meio de algoritmos, com o objetivo de extrair conhecimentos novos, relevantes e úteis a partir de um conjunto de dados. Os autores apontam que o foco central dessa disciplina reside na busca por respostas eficientes sobre como extrair conhecimento de grandes volumes ou fluxos de dados de forma a apoiar o processo de tomada de decisão.

Embora frequentemente confundido com o termo *Big Data*, Carvalho, Menezes e Bonidia (2024) esclarecem que os conceitos possuem propósitos diferentes, enquanto o *Big Data* diz respeito ao estudo e desenvolvimento de tecnologias para a infraestrutura de dados, a CD dedica-se à formulação de soluções computacionais voltadas para a transformação, o pré-processamento e a modelagem desses dados.

Desta forma, a CD pode ser compreendida tanto como um campo de estudo cujo objetivo é extrair ao máximo todos os *insights* provenientes dos dados independentemente de sua fonte, quanto como uma ferramenta única que integra conceitos complementares da Computação, da Estatística e da Matemática (Carvalho; Menezes; Bonidia, 2024). Essa natureza abrangente e dinâmica permite que a área incorpore múltiplas etapas sequenciais em problemas teóricos e aplicados, às quais se iniciam no entendimento do problema de negócio, passam pelas transformações, tratamentos e modelagens necessárias, e culminam na etapa de validação dos resultados obtidos.

Para que essa extração de conhecimento ocorra de maneira efetiva, Carvalho, Menezes e Bonidia (2024) destacam a relevância de compreender o ciclo de vida dos dados, o qual é

composto por cinco etapas principais: a aquisição, a limpeza, o uso e reúso, a publicação e a preservação ou destruição dos dados. Cada uma dessas fases engloba atividades específicas que operam sob a influência de um ambiente que envolve aspectos éticos, políticos, gerenciais, plataformas tecnológicas e o domínio do problema.

Os autores ainda ressaltam que o desenvolvimento de ciclos de vida eficientes demanda a união de esforços de diferentes profissionais, integrando a visão dos geradores de dados com o foco dos profissionais de Tecnologia da Informação e Comunicação (TIC), responsáveis pelas plataformas, técnicas de processamento, modelagem e visualização, além da contribuição de engenheiros, estatísticos e cientistas da informação.

2.2.1 Aquisição dos dados

No âmbito da CD, a raspagem de dados é definida como um conjunto de processos automatizados destinados à coleta sistemática de informações acessíveis na internet. No fluxo de um projeto de CD, essa técnica é essencial nas fases de coleta e organização, possibilitando que os pesquisadores acessem bases que não proporcionam transparência ativa (Trecenti, 2023).

Pela perspectiva operacional, a raspagem opera de forma semelhante a uma consulta em bancos de dados remotos ou ao uso de uma *Application Programming Interface* (API), atuando como uma ponte tecnológica fundamental para converter dados públicos dispersos em conjuntos estruturados e prontos para análise (Trecenti, 2023).

Neste contexto, um dos desafios recorrentes é a extração de informações contidas em documentos no formato *Portable Document Format* (PDF). Segundo Yang, Cao e Zhao (2024), embora o PDF seja o padrão global para armazenamento e transmissão de documentos eletrônicos devido à sua fidelidade visual, sua estrutura não nativa para edição dificulta a reutilização de dados críticos, especialmente aqueles organizados em tabelas.

Para mitigar esse problema, a literatura recente aponta o uso de bibliotecas de *parsing* baseadas em *Python*. Diferente de métodos tradicionais que tratam o PDF apenas como um fluxo de texto contínuo, essa abordagem converte o documento em objetos *Python* analisáveis, permitindo que o algoritmo identifique a geometria da página, como linhas verticais, horizontais e retângulos, para reconstruir a estrutura lógica das células de uma tabela (Yang; Cao; Zhao, 2024).

2.2.2 Limpeza dos dados

O processo de extração de dados de PDFs para fins de análise de dados exige um *pipeline* estruturado. Yang, Cao e Zhao (2024) detalham que, após a localização da tabela nas páginas do documento, os dados brutos extraídos frequentemente contêm ruídos, como tabulação, quebras de linha indevidas e espaços em branco. Portanto, uma etapa fundamental de pré-processamento é a aplicação de técnicas de limpeza, em que caracteres inválidos são removidos ou substituídos programaticamente.

Neste contexto, uma abordagem para guiar esse fluxo de tratamento e organização é a arquitetura *Medallion*. De acordo com Strenghtolt (2025), esse padrão de *design* é utilizado para organizar logicamente as informações. O modelo estrutura a plataforma de dados em três camadas sequenciais, *Bronze*, *Silver* e *Gold*, com o objetivo de melhorar progressivamente a estrutura e a qualidade dos dados à medida que eles fluem pela arquitetura.

A primeira etapa desse fluxo ocorre na camada *Bronze*, que funciona como o armazenamento inicial, sendo responsável por receber e guardar os dados brutos vindos do processo de extração em sua estrutura nativa, ainda contendo os ruídos originais (Strenghtolt, 2025). Na sequência, os dados avançam para a camada *Silver*, que atua como o estágio de transição em que as correções e padronizações necessárias são executadas. Nessa fase, aplicam-se checagens de qualidade, de duplicações e as técnicas de limpeza programática, gerando um conjunto de dados granular, consistente e livre de caracteres inválidos (Strenghtolt, 2025).

Por fim, a camada *Gold* destina-se à otimização desses dados já refinados para atender a necessidades de negócios e tomadas de decisões específicas. Nela, as informações são agregadas, sumarizadas e enriquecidas para relatórios e análises de alto nível, priorizando o desempenho e a escalabilidade para fornecer acesso rápido a métricas e *insights* centrais (Strenghtolt, 2025).

Apesar de fornecer rótulos intuitivos e uma divisão clara para o negócio, Strenghtolt (2025) pondera que a implementação eficaz desse *design* exige maturidade técnica e compreensão da evolução das arquiteturas de dados, visto que a escolha de uma plataforma tecnológica não resolve por si só os desafios de integração e automação.

2.2.3 Uso dos dados

A fase de visualização é crucial para a análise exploratória, pois possibilita que os pesquisadores utilizem o raciocínio analítico para reconhecer padrões, fazer comparações e entender causalidades de maneira intuitiva. A visualização de dados desempenha um papel fundamental ao usar gráficos estatísticos e tabelas bem organizadas para tornar o conhecimento técnico mais acessível e utilizável, facilitando a comunicação precisa dos testes e descobertas da CD de modo a torná-la mais compreensível para os gestores (Silva, 2019).

Diante dessa necessidade de traduzir dados complexos em insumos acionáveis, Carvalho e Melo (2018) apontam que as organizações têm adotado cada vez mais uma abordagem de tomada de decisão baseada em dados para aprimorar seu desempenho e obter vantagens competitivas no mercado.

Nesse contexto, o *dashboard* consolida-se como uma ferramenta essencial de auxílio aos gestores para viabilizar e otimizar a tomada de decisões. Conceitualmente, trata-se de uma tela visual única que reúne e representa graficamente um conjunto de dados e informações essenciais, tanto presentes quanto históricas, orientadas para o alcance de objetivos específicos. A finalidade central dessa *interface* é disponibilizar ao tomador de decisão, de maneira clara e sintética, todos os indicadores necessários para as resoluções chave de sua respectiva área de atuação (Carvalho; Melo, 2018).

2.2.4 Tipos de gráficos

Para que a interface seja realmente eficaz, Carvalho e Melo (2018) ressaltam que o *dashboard* precisa de uma visualização de informação de qualidade, estruturada de forma fácil de ler e interpretar, permitindo que o usuário identifique o panorama exposto em poucos segundos.

Nesse contexto, Camm, Cochran e Fry (2026) realizam um estudo detalhado sobre os tipos de gráficos e suas melhores aplicações no contexto analítico. Segundo os autores, o gráfico de linha, por exemplo, liga com pontos um par de valores de variáveis quantitativas (um valor ao longo do eixo horizontal e outro ao longo do vertical), sendo apropriado para a representação de dados de séries temporais, isto é, informações que são coletadas de forma consecutiva ao longo do tempo, em intervalos que podem ser de minutos, horas, dias, anos, entre outros.

Para a análise comparativa entre categorias, o gráfico de colunas exibe uma variável quantitativa por meio de barras verticais que demonstram a magnitude do dado para cada grupo ou período de tempo específico. Em cenários que exigem um nível maior de detalhamento, pode-se aplicar o gráfico de colunas empilhadas, cuja estrutura utiliza cores distintas dentro de uma mesma barra para indicar visualmente a contribuição de cada subcategoria em relação ao valor total (Camm; Cochran; Fry, 2026).

Quando a análise demanda uma contextualização espacial, emprega-se o mapa coroplético. Essa solução consiste em um mapa geográfico que utiliza tons de uma mesma cor, cores únicas ou símbolos para indicar a distribuição de variáveis quantitativas ou categóricas por região ou área geográfica delimitada (Camm; Cochran; Fry, 2026).

Entretanto, caso o objetivo seja evidenciar o efeito cumulativo de alterações positivas e negativas ocorridas em uma variável específica, o gráfico de cascata é a representação visual ideal. Nesse modelo, as variações são apresentadas ao longo de uma série de categorias sucessivas, como intervalos temporais, sendo a magnitude de cada alteração expressa por meio de uma coluna cuja base se ancora na altura cumulativa resultante das categorias precedentes (Camm; Cochran; Fry, 2026).

2.3 Ferramentas e tecnologias

Para a viabilização e execução deste trabalho, fez-se necessário o uso de ferramentas e tecnologias que auxiliassem nas etapas de prototipação, coleta, armazenamento, processamento, análise e visualização dos dados. Nos próximos parágrafos, apresentam-se a descrição dos principais instrumentos utilizados no desenvolvimento da pesquisa.

O *Figma*¹ é uma plataforma de *design* colaborativo baseada em nuvem, amplamente utilizada para a criação de *interfaces* de usuário (UI), experiências de usuário (UX), protótipos interativos e sistemas de design. A ferramenta permite que equipes de *designers*, desenvolvedores e *stakeholders* trabalhem simultaneamente no mesmo arquivo em tempo real, facilitando o processo de *brainstorm*, prototipagem e desenvolvimento de produtos digitais.

O *Visual Studio Code*² é um editor de código-fonte gratuito, multiplataforma e desenvolvido pela Microsoft. A ferramenta destaca-se por otimizar o fluxo de trabalho de desenvolvimento de software ao integrar recursos nativos como controle de versão e suporte a

¹ Disponível em: <https://www.figma.com/pt-br/>

² Disponível em: <https://code.visualstudio.com/>

compilação e depuração diretas na interface de usuário. Adicionalmente, sua arquitetura leve é potencializada por um amplo ecossistema de extensões de terceiros, o que confere versatilidade ao editor e viabiliza seu uso em múltiplas vertentes computacionais, incluindo o desenvolvimento de aplicações *web*, automação por *scripts* e engenharia de dados.

O *Git*³ é um sistema de controle de versão distribuído, gratuito e de código aberto, projetado para gerenciar projetos de desenvolvimento de *software* de qualquer escala com alto desempenho e eficiência. A ferramenta permite o rastreamento sistemático de alterações no código-fonte ao longo do tempo, viabilizando o trabalho colaborativo e a coordenação síncrona entre equipes de desenvolvimento.

O *Google Sheets*⁴ é uma aplicação de planilhas eletrônicas baseada em computação em nuvem, desenvolvida pela empresa *Google* como parte integrante do pacote de ferramentas corporativas *Google Workspace*. A plataforma permite a criação, edição e manipulação de dados diretamente em navegadores *web*, viabilizando o desenvolvimento de rotinas de trabalho colaborativas e simultâneas entre múltiplos usuários sem a necessidade de instalação de *softwares* locais.

O *Python*⁵ é uma linguagem de programação de alto nível, amplamente reconhecida por sua legibilidade, eficiência e capacidade de resolução de problemas complexos de forma ágil. A plataforma destaca-se por apresentar uma sintaxe limpa e um ecossistema robusto, composto por um amplo conjunto de bibliotecas e *frameworks* especializados. Em virtude dessa versatilidade, a linguagem é empregada em múltiplos domínios computacionais, abrangendo desde o desenvolvimento *web* e aplicações corporativas até a computação científica.

O *pdfplumber*⁶ é uma biblioteca da linguagem *Python* voltada para a extração e análise detalhada de dados de arquivos PDF. A ferramenta permite que os desenvolvedores obtenham informações sobre cada elemento gráfico e textual presente no documento, como caracteres, retângulos, linhas e curvas. Entre suas principais funcionalidades, destacam-se a extração de texto com preservação do *layout* original, a identificação e conversão de tabelas complexas em dados estruturados, além de recursos de depuração visual.

O *Pandas*⁷ é uma biblioteca de código aberto desenvolvida originalmente em 2008 para a linguagem *Python*, projetada com o escopo de fornecer ferramentas robustas e flexíveis

³ Disponível em: <https://git-scm.com/>

⁴ Disponível em: <https://workspace.google.com/intl/pt-BR/products/sheets/>

⁵ Disponível em: <https://www.python.org/>

⁶ Disponível em: <https://pypi.org/project/pdfplumber/>

⁷ Disponível em: <http://pandas.pydata.org/about/index.html>

para a análise e manipulação de estruturas de dados. O objetivo central dessa tecnologia é atuar como um componente fundamental para o processamento de dados do mundo real, oferecendo funcionalidades avançadas que otimizam as etapas de tratamento, limpeza e transformação de grandes volumes de dados.

A *Plotly*⁸ é uma biblioteca de código aberto para a linguagem *Python*, amplamente utilizada para a criação de visualizações de dados interativos e de alta qualidade. A ferramenta oferece uma vasta gama de tipos de gráficos, incluindo dispersão, linhas, barras, mapas, gráficos 3D, animações e representações financeiras. Entre suas principais características, destaca-se a interatividade nativa dos gráficos gerados, que permitem zoom, seleção de dados e exibição de detalhes ao passar o mouse.

O *Streamlit*⁹ é um *framework* de código aberto desenvolvido para a linguagem *Python*, projetado com o escopo de permitir a construção simplificada de aplicações *web* interativas voltadas para a visualização e a análise de dados. A ferramenta viabiliza que cientistas de dados transformem *scripts* de dados em painéis dinâmicos e compartilháveis de forma otimizada, sem a necessidade de conhecimentos aprofundados ou experiência prévia em desenvolvimento de interfaces de usuário.

O *Cascading Style Sheets*¹⁰ (CSS) é uma linguagem de estilo utilizada para descrever a apresentação e a formatação visual de documentos estruturados em linguagens de marcação, como o HTML e o XML. Essa tecnologia atua na definição de como os elementos estruturais de uma página *web* devem ser exibidos em diferentes telas e mídias, permitindo a separação entre o conteúdo documental e o seu *design* visual.

O *Gemini*¹¹ é um modelo de linguagem de grande escala (*Large Language Model* - LLM) desenvolvido e mantido pela *Google*. Baseado na arquitetura de redes neurais *Transformer*, o modelo foi projetado de forma nativamente multimodal, o que o capacita a processar, compreender e operar de maneira integrada diferentes modalidades de informação, incluindo texto, código, imagens e áudio. Válido ressaltar que, para o presente estudo, o *Gemini* foi utilizado exclusivamente como ferramenta de suporte na correção gramatical e aprimoramento da fluidez da redação técnica.

O *LanguageTool*¹² é uma ferramenta multilíngue de verificação ortográfica, gramatical e de estilo, projetada para auxiliar usuários na correção e melhoria da qualidade de seus

⁸ Disponível em: <https://plotly.com/python/>

⁹ Disponível em: <https://streamlit.io/>

¹⁰ Disponível em: <https://developer.mozilla.org/pt-BR/docs/Web/CSS>

¹¹ Disponível em: <https://gemini.google/about/>

¹² Disponível em: <https://languagetool.org/pt-BR>

textos. A plataforma atua na identificação e sugestão de correções para uma ampla gama de erros, incluindo falhas de ortografia, problemas de concordância verbal e nominal, entre outros.

3 TRABALHOS RELACIONADOS

Neste capítulo, serão apresentados os trabalhos relacionados, destacando seus aspectos essenciais e sua relevância para o presente estudo. Para a seleção, utilizou-se o método de Análise de Conteúdo proposto por Bardin em 1977. Essa metodologia é composta por um conjunto de técnicas sistemáticas e objetivas para descrever o conteúdo das mensagens e das comunicações, possibilitando a geração de indicadores que sustentam a compreensão e a interpretação de conhecimentos (Bardin, 1977 *apud* Quadros; Silveira, 2025). Todo o processo de mapeamento e classificação dos artigos seguiu estritamente os três pontos de análise que o método propõe: a pré-análise, a exploração do material e, por último, o tratamento dos resultados através da análise e interpretação.

Na fase de pré-análise, procedeu-se à seleção do material a ser analisado, foram definidos os objetivos da revisão da literatura e foi elaborada uma leitura dinâmica dos documentos. Para a constituição do acervo bibliográfico, foi necessário observar as quatro regras básicas de Bardin: a homogeneidade, a representatividade, a exaustividade, e a pertinência (Quadros; Silveira, 1977). A princípio, foi realizada uma busca detalhada nas bases de dados acadêmicas *Google Acadêmico*, *IEEE Xplore* e *SciELO*, considerando palavras-chave estritamente alinhadas ao objeto de estudo, tais como "Ciência de Dados", "Business Intelligence", "Extração de dados", "Dashboards" e "Locação de veículos".

Posteriormente, na fase de exploração do material, os documentos foram lidos detalhadamente para identificar suas ideias centrais. Esses pontos importantes foram agrupados e organizados nas quatro categorias que dividem este capítulo: Ciência de Dados aplicada à tomada de decisão, Desafios técnicos da extração de dados, Interfaces de suporte à decisão baseada em dados e o Mercado de locação de veículos.

Por fim, foi realizado o tratamento e a interpretação dos resultados obtidos, estabelecendo relações entre os trabalhos selecionados e os objetivos do presente estudo. Dessa forma, buscou-se identificar aproximações, diferenças e contribuições relevantes para a presente pesquisa, permitindo compreender como os estudos analisados se relacionam com a proposta do trabalho, conforme exposto na Seção 3.5.

3.1 Ciência de Dados aplicada à tomada de decisão

O estudo de Wagner, Brandt e Neumann (2015) evidencia que a CD e o *Big Data Analytics* deixaram de ser apenas ferramentas de suporte técnico para se tornarem pilares

essenciais na formulação de estratégias de mercado competitivas, especialmente no setor de mobilidade e transporte.

Os autores alegam que em mercados caracterizados por rápidas transformações e crescimento dinâmico, os métodos tradicionais ou decisões baseadas em abordagens empíricas e generalistas mostram-se insuficientes para responder às variações de demanda e para orientar planos de expansão de longo prazo. Nesse cenário, a análise detalhada de dados surge como um diferencial estratégico indispensável para transformar registros brutos em conhecimento útil.

O referido estudo propõe uma abordagem analítica voltada para o suporte à tomada de decisão no mercado de compartilhamento de veículos (*carsharing*) na cidade de Berlim. Os autores demonstram que, ao processar um amplo conjunto de dados reais de locação, é possível derivar indicadores precisos sobre a atratividade de determinadas regiões urbanas. Essa modelagem orientada a dados permite que gestores identifiquem com maior exatidão o comportamento dos usuários e otimizem decisões operacionais críticas, como o balanceamento da capacidade da frota e a delimitação geográfica ideal das áreas de atuação das empresas.

Além disso, o trabalho de Wagner, Brandt e Neumann (2015) ressalta o papel fundamental da visualização de dados para o gerenciamento executivo. Diante de fluxos informacionais que atendem aos critérios de volume e variedade do *Big Data*, a representação gráfica e cartográfica da demanda diária torna-se um fator-chave de sucesso, pois reduz a complexidade da leitura analítica e permite que os tomadores de decisão reajam de maneira ágil às oscilações do mercado. Dessa forma, a transição de um modelo de gestão puramente intuitivo para uma cultura orientada por dados consolida-se como o caminho para a sustentabilidade econômica e eficiência operacional de serviços modernos de transporte.

Assim como observado no estudo supracitado, a aplicação de técnicas de CD tem se consolidado como um importante suporte para a tomada de decisões estratégicas no setor de mobilidade e locação de veículos. Complementando essa perspectiva, Jeon *et al.* (2015) destacam que o avanço da internet e a disponibilidade de informações em mercados abertos possibilitam que empresas obtenham grandes volumes de dados de referência, especialmente relacionados à precificação de ativos.

Segundo os autores, essas informações podem ser utilizadas para estimar de maneira mais precisa os preços futuros de veículos usados e os valores de locação, contribuindo diretamente para decisões operacionais e financeiras no setor.

No referido estudo, os autores propõem uma estrutura analítica capaz de coletar dados públicos de plataformas *online* de venda de automóveis, utilizando-os como referência para estimar valores de mercado. A pesquisa demonstra que fatores temporais, como o tempo de uso do veículo, exercem influência significativa sobre a depreciação dos ativos e, consequentemente, sobre os preços de revenda e locação.

Dessa forma, a utilização de modelos analíticos e regressões estatísticas permite que empresas do setor obtenham previsões mais precisas em comparação a métodos baseados exclusivamente em transações históricas internas.

Além disso, o trabalho de Jeon *et al.* (2015), evidencia como a coleta e análise de dados provenientes de fontes abertas podem ser utilizadas para apoiar decisões estratégicas relacionadas à gestão de ativos e definição de preços. Os autores ressaltam que a utilização de sistemas analíticos permite transformar dados disponíveis na internet em informações relevantes para o negócio, aumentando a capacidade das empresas de responder às mudanças do mercado de maneira mais eficiente.

Conforme enfatizado no estudo realizado por Leal (2025), a necessidade de transição para uma cultura orientada por dados é ainda mais crítica quando analisada sob a ótica do mercado automotivo brasileiro de veículos pesados e operações de repasse. A ausência de processos estruturados e de uma análise sistemática de informações nesse segmento tende a ampliar ineficiências operacionais, comprometer margens financeiras e elevar os riscos associados à imobilização de ativos.

O autor argumenta que negligenciar o ciclo analítico e basear a tomada de decisão predominantemente em experiências empíricas, limita drasticamente a agilidade das organizações frente às rápidas oscilações da demanda nacional. Para mitigar esses riscos, Leal (2026) alega que a avaliação automotiva deve transcender a mera precificação pontual e ser estabelecida como um mecanismo contínuo de inteligência de mercado.

Sob essa perspectiva, decisões de repasse e o *timing* comercial de desmobilização de frotas, quando fundamentados em históricos de liquidez e dados estruturados, reduzem drasticamente as perdas financeiras. A verdadeira transformação analítica ocorre quando os modelos estatísticos deixam de ser recursos eventuais e passam a guiar as decisões operacionais rotineiras, integrando-se organicamente às estruturas e à cultura decisória da empresa.

Em seu trabalho Leal (2025) comprova que a estruturação de *pipelines* de dados e processos orientados por análise de dados resulta em impactos práticos expressivos, como a

redução do tempo médio de comercialização dos veículos e o aumento da previsibilidade comercial. Esse alinhamento preciso entre precificação, canais de escoamento e demanda real não apenas otimiza a eficiência do fluxo de caixa, mas também garante maior estabilidade financeira para as operações do setor.

Portanto, a literatura valida que a consolidação, organização e visualização de bases de dados descentralizadas constituem o alicerce fundamental para a sobrevivência e expansão de empresas inseridas no dinâmico ecossistema de mobilidade e locação de veículos.

3.2 Desafios técnicos da extração de dados

Segundo Pappula e Rusum (2023), a extração de dados estruturados a partir de documentos não estruturados, como arquivos PDF, imagens digitalizadas ou relatórios visuais representa um dos principais desafios técnicos em processos de análise de dados. Embora ferramentas tradicionais de reconhecimento óptico de caracteres (*Optical Character Recognition* - OCR) sejam eficientes na digitalização de textos, elas apresentam limitações na interpretação de estruturas complexas, tabelas e elementos semânticos presentes nos documentos. Essa problemática torna-se ainda mais evidente em cenários nos quais os arquivos possuem diferentes formatos, baixa padronização visual e grande diversidade estrutural.

Pappula e Rusum (2023), ainda ressaltam que elementos como a mudança de *layout*, tabelas complexas, cabeçalhos, notas de rodapé e gráficos dificultam a localização e a organização automática dos dados. Além disso, as variações na resolução, no ruído e na estrutura visual dos documentos tornam a extração e o pré-processamento dos dados muito mais complexos. Nesse contexto, a simples conversão textual torna-se insuficiente, sendo necessária a utilização de abordagens capazes de compreender também a organização visual e o contexto semântico das informações extraídas.

Buscando solucionar essas limitações, o estudo propõe a implementação de estratégias multimodais, combinando informações textuais e visuais para melhorar a precisão na interpretação dos documentos. A pesquisa demonstra que a combinação entre modelos de linguagem, análise visual e mecanismos de interpretação semântica possibilita uma extração de dados mais robusta e contextualizada, principalmente em documentos complexos e heterogêneos.

3.3 Interfaces de suporte à decisão baseada em dados

No âmbito do desenvolvimento de sistemas voltados para a CD, o fechamento do ciclo analítico não se limita à extração e ao tratamento da informação, mas consolida-se na forma como esse conhecimento é disponibilizado para os tomadores de decisão. Conforme elucida Jain (2021), as *interfaces* de *Business Intelligence* (BI), especificamente os *dashboards* interativos, surgiram como suportes para potencializar decisões estratégicas a partir de *insights* visuais.

O autor ainda alega que ao agregar dados dispersos de múltiplas fontes em uma interface unificada, elimina-se a necessidade de um processo historicamente lento e propenso a falhas que são os resumos manuais em formato de relatórios, mitigando erros operacionais e elevando a precisão na identificação de tendências e riscos de mercado. Logo, o valor dessas plataformas reside na capacidade intrínseca da visualização de dados de decodificar e simplificar conjuntos de dados massivos e altamente complexos.

De acordo com Jain (2021), a transformação de registros brutos em elementos gráficos digeríveis, como gráficos interativos, mapas de calor e Indicadores-Chave de Desempenho (*Key Performance Indicator* - KPIs), permite que partes interessadas em qualquer nível organizacional percebam discrepâncias e padrões ocultos que passariam despercebidos em análises tabulares tradicionais. Outro ponto destacado pelo autor refere-se à evolução dos *dashboards* de BI a partir da integração com tecnologias como Inteligência Artificial (IA) e análise preditiva, permitindo que as organizações não apenas monitorem informações históricas, mas também antecipem tendências de mercado e possíveis desafios operacionais.

3.4 Mercado de locação de veículos

Conforme documenta Luiz (2018), a compreensão do mercado de locação de veículos no Brasil exige o resgate de sua trajetória histórica e de sua consolidação como um dos pilares de maior importância no cenário econômico e social do país.

Surgindo de forma suplementar na região central de São Paulo, por meio de revendedores de carros usados, profissionalizou-se em 1956 com a fundação da primeira empresa voltada exclusivamente para essa finalidade. Desde então, o setor evoluiu de um serviço de nicho para uma indústria robusta que ultrapassou a marca de dezenas de milhões de usuários e consolidou-se como um vetor de alta relevância socioeconômica, tanto pelo

faturamento e arrecadação tributária quanto pela geração de centenas de milhares de empregos diretos e indiretos.

Luiz (2018) também alega que além do impacto social, a relevância científica de se investigar o setor reside na sua magnitude operacional e na sua íntima ligação com a cadeia produtiva automotiva global. Indicadores cruciais como o tamanho da frota total circulante e a taxa de aquisição de novos veículos demonstram o papel central das locadoras como o principal cliente das montadoras instaladas no país, absorvendo parcelas expressivas da produção nacional. O uso desses indicadores também permite identificar tendências, avaliar a evolução organizacional e apoiar projeções futuras, tornando-se uma ferramenta essencial para a tomada de decisão estratégica.

Corroborando com o trabalho de Luiz (2018), no qual destaca o mercado de locação veicular como um dos pilares da economia brasileira, Milhome (2022) destaca a elevada importância que o segmento de locação de automóveis possui, tanto pelo faturamento quanto pela geração de empregos e arrecadação tributária. Entretanto, o autor enfatiza que o setor apresenta características típicas de um oligopólio, no qual poucas empresas concentram grande parte da participação de mercado e possuem capacidade de exercer influência significativa sobre preços e estratégias competitivas.

Para a realização do estudo, Milhome (2022) utilizou dados secundários relacionados à quantidade de diárias vendidas pelas empresas do setor para avaliar o nível de concentração do mercado entre 2014 e 2021. Os resultados demonstram que o segmento apresenta forte concentração em torno de grandes empresas, com destaque para a liderança da Localiza, seguida por outras organizações de menor participação relativa.

Além disso, a pesquisa evidencia que o setor passou por um processo de intensificação da concentração ao longo do período analisado, reforçando barreiras de entrada relacionadas à necessidade de elevado capital e vantagens competitivas associadas à escala operacional.

3.5 Síntese analítica dos trabalhos relacionados

Diante do panorama teórico apresentado anteriormente, esta seção dedica-se a evidenciar as convergências e as divergências que interligam os trabalhos supracitados ao presente estudo. Para tanto, a discussão é estruturada por meio de uma análise comparativa, cujos pontos centrais encontram-se resumidos e organizados no Quadro 1.

Os trabalhos de Wagner, Brandt e Neumann (2015), Jeon *et al.* (2015) e Leal (2025), expostos na seção 3.1, compartilham a premissa fundamental de que a transição para uma

cultura orientada por dados é o principal fator de eficiência e vantagem competitiva no setor automotivo. Os três estudos repudiam abordagens gerenciais puramente intuitivas ou empíricas, provando que a consolidação de bases de dados em modelos analíticos rotineiros mitiga riscos financeiros e otimiza processos operacionais críticos. Essa visão integrada valida diretamente o propósito do presente trabalho, que busca utilizar a CD para estruturar informações descentralizadas e transformá-las em um ativo estratégico de tomada de decisão.

Apresentado na seção 3.2, o estudo de Pappula e Rusum (2023) aborda os obstáculos que surgem no processo de extração de dados estruturados em documentos não estruturados, como PDFs e arquivos com distintos *layouts*. Os autores ressaltam que os métodos de OCR convencionais falham ao lidar com tabelas, elementos gráficos e contextos semânticos, principalmente quando aplicados a documentos variados e não padronizados.

Nesse sentido, o trabalho de Pappula e Rusum (2023) aproxima-se da presente pesquisa ao evidenciar os desafios envolvidos na extração e estruturação de dados provenientes de documentos PDF. De maneira semelhante, o presente estudo também enfrenta limitações relacionadas à heterogeneidade estrutural dos ABSLVs, tornando necessária a realização de processos de seleção, tratamento e padronização das informações antes de sua análise.

Entretanto, enquanto os autores concentram-se na aplicação de IA multimodal para extração automatizada de dados complexos, a presente pesquisa utiliza ferramentas de raspagem, tratamento e organização de dados voltadas à construção de uma base histórica estruturada para análise do mercado de locação de veículos no Brasil.

Na seção 3.3, em que aborda a temática de visualização dos dados, o presente trabalho menciona o autor Jain (2021), no qual destaca o uso de *dashboards* como ferramenta fundamental no suporte à tomada de decisão estratégica, ao transformar grandes volumes de dados em informações visuais e interativas de fácil interpretação. Dessa forma, o trabalho de Jain (2021) aproxima-se da presente pesquisa ao demonstrar a importância da visualização de dados e dos *dashboards* interativos na transformação de dados em informações estratégicas.

Entretanto, enquanto o estudo concentra-se em *dashboards* de BI voltados ao *marketing* estratégico, a presente pesquisa busca aplicar técnicas de visualização e análise exploratória aos dados históricos do setor brasileiro de locação de veículos, utilizando *dashboards* interativos para facilitar a interpretação de tendências, padrões e anomalias.

Quadro 1 - Comparação entre os trabalhos relacionados e o presente trabalho

Trabalho	Principais contribuições	Diferencial do presente trabalho
Wagner <i>et al.</i> (2015)	Aplicação de <i>Big Data Analytics</i> e visualização de dados para apoiar decisões estratégicas em serviços de carsharing, especialmente na definição de áreas operacionais e expansão de mercado	Estruturação e análise histórica dos dados do setor brasileiro de locação de veículos, com foco em séries temporais, <i>dashboards</i> interativos e geração de <i>insights</i> de mercado
Jeon <i>et al.</i> (2015)	Desenvolvimento de modelos analíticos para precificação de veículos e serviços de locação a partir da coleta e análise de dados públicos de mercado	Utilização de dados consolidados do ABSLVs para análises exploratórias e identificação de padrões relacionados à evolução do mercado
Leal (2025)	Demonstração do impacto da análise orientada por dados na eficiência operacional, previsibilidade comercial e gestão de ativos no setor automotivo	Construção de uma base de dados unificada e aplicação de técnicas de visualização e análise temporal para auxiliar na compreensão estratégica do setor de locação de veículos no Brasil
Pappula e Rusum (2023)	Aplicação de IA multimodal para extração estruturada de dados em documentos PDF complexos.	Construção e padronização de uma base histórica do setor de locação de veículos a partir de dados extraídos dos ABSLVs.
Jain (2021)	Demonstração do uso de <i>dashboards</i> de BI e visualização de dados no apoio à tomada de decisão estratégica.	Aplicação de <i>dashboards</i> interativos para análise histórica e identificação de tendências no setor brasileiro de locação de veículos.
Luiz (2018)	Análise da relevância econômica e do desempenho financeiro do setor de locação de veículos no Brasil.	Consolidação e análise histórica dos dados do setor com técnicas de CD e <i>dashboards</i> interativos.
Milhorne (2022)	Investigação da concentração de mercado e da estrutura oligopolista do setor de locação de veículos.	Identificação de tendências e padrões do mercado por meio de análise temporal e regionais.

Fonte: Autoria Própria (2026).

Os estudos de Luiz (2018) e Milhorne (2022), apresentados na seção 3.4 evidenciam a relevância econômica e estratégica do setor de locação de veículos no Brasil, destacando o seu crescimento e sua influência sobre diferentes segmentos da economia. Válido destacar que ambos os trabalhos reforçam a importância da análise sistemática de dados e indicadores para compreensão do comportamento do mercado e suporte à tomada de decisão. Nesse contexto, as pesquisas aproximam-se do presente trabalho ao demonstrarem o potencial analítico dos dados do setor de locação de veículos. Entretanto, diferentemente dos estudos relacionados, a

presente pesquisa propõe a consolidação, tratamento e análise histórica dos dados do ABSLVs por meio de técnicas de CD, buscando proporcionar a identificação de tendências e *insights* estratégicos sobre a evolução do mercado brasileiro de locação de veículos.

4 METODOLOGIA

A presente pesquisa caracteriza-se como um estudo de natureza aplicada com abordagem quantitativa, baseada na utilização de métodos para analisar os dados dos ABSLVs. Em relação aos seus objetivos, o estudo tem natureza descritiva, pois procura reconhecer padrões, tendências e relações essenciais que determinam o comportamento desse mercado.

O presente estudo foi conduzido com base em uma metodologia estruturada em quatro etapas principais, sendo elas: aquisição dos dados (seção 4.1) descreve o processo inicial de obtenção e centralização das informações extraídas dos ABSLVs. Posteriormente, a limpeza dos dados (seção 4.2) aborda os procedimentos de tratamento, padronização e estruturação das variáveis através das camadas da arquitetura *Medallion*. Em seguida, o uso dos dados (seção 4.3) detalha as técnicas de análise aplicadas para a identificação de padrões, tendências e geração de *insights* sobre o mercado de locação. Por fim, a publicação dos dados (seção 4.4), documentando o desenvolvimento do *dashboard* interativo utilizando o *framework Streamlit*.

4.1 Aquisição dos dados

A etapa de aquisição de dados constitui o ponto de partida desta metodologia, sendo responsável por fornecer os insumos necessários para alimentar todo o processo de análise. Para compreender o mercado de locação de veículos, fez-se necessária a seleção de uma base de dados que apresentasse abrangência nacional e legitimidade da informação.

Sob esse critério, a base de dados utilizada para a execução deste estudo foi extraída dos ABSLVs, um documento oficial de periodicidade anual institucionalizado pela ABLA. Os anuários são considerados os principais levantamentos estatísticos e analíticos sobre o mercado de aluguel de veículos no Brasil. O documento mapeia o desempenho, o tamanho da frota, o faturamento, os investimentos e o impacto do setor na economia, demonstrando como a locação se tornou protagonista na indústria automotiva, no turismo e na mobilidade urbana (ABLA, 2026).

Para a coleta, foram selecionadas as edições dos anuários da ABLA publicadas entre os anos de 2012 e 2026, a partir desse conjunto de publicações, realizou-se à seleção prévia das páginas que continham as informações pertinentes ao escopo do estudo. Essa delimitação manual funcionou como uma etapa de filtragem qualitativa, assegurando que o processo de

extração automatizada fosse direcionado exclusivamente às seções de interesse, minimizando a captura de ruídos informacionais e otimizando o processamento dos dados.

Concluída a seleção das seções de relevância para o estudo, realizou-se a extração automatizada dos dados contidos nos arquivos em formato PDF utilizando a biblioteca *pdfplumber*, da linguagem *Python*, gerando-se arquivos correspondentes em formato XLSX. Após a extração e a conversão inicial das tabelas, os dados passaram por um processo de validação manual com o auxílio do *Google Sheets*. Nessa etapa, as informações textuais que estavam embutidas exclusivamente como imagens nos relatórios originais foram coletadas e inseridas sistematicamente nas tabelas, buscando garantir a integridade da base de dados.

4.2 Limpeza dos dados

A etapa de extração resultou em dois conjuntos distintos de dados. No âmbito nacional (Brasil), coletou-se um panorama amplo que inclui o faturamento líquido e bruto do setor, o número de usuários por segmento (RaC ou GTF) e a quantidade de locadoras que operam com ou sem motorista. Adicionalmente, foram obtidos dados sobre o número de funcionários diretos por gênero e nível de escolaridade, além do volume de emplacamentos por subsegmento e fabricante.

Por outro lado, a disponibilidade de dados para as unidades federativas mostrou-se mais restrita, para os estados, as variáveis mapeadas limitaram-se à divisão de locadoras, com ou sem motorista, ao número total de empregados diretos e ao volume de emplacamentos por subsegmento e montadora. Além disso, com o intuito de delimitar o escopo do estudo, a extração dos dados estaduais restringiu-se à Região Nordeste, essa escolha justifica-se pela relevância de contextualizar a realidade socioeconômica regional, dado que a instituição de ensino no qual a presente pesquisa foi desenvolvida localiza-se no estado do Rio Grande do Norte.

Após a aquisição, os dados obtidos apresentavam ruídos, registros vazios e inconsistências de formatação que inviabilizavam uma análise direta, conforme o descrito pela camada *Bronze* da arquitetura *Medallion*. Para o tratamento dessas informações, utilizou-se a biblioteca *Pandas*, da linguagem *Python*, que permitiu realizar a limpeza, a correção dos formatos e a reorganização estrutural das tabelas, elevando os dados para a camada *Silver*.

Na etapa final, correspondente à camada *Gold*, realizou-se a agregação e o enriquecimento das variáveis, gerando a base de dados consolidada para a construção dos

gráficos. Cabe destacar que os *scripts* computacionais de processamento foram desenvolvidos de forma separada, com um bloco dedicado aos indicadores de nível nacional e outro para as unidades federativas, visto que as informações obtidas apresentavam estruturas e níveis de detalhamento distintos.

4.3 Uso dos dados

Dando início à fase de uso e análise dos dados, os *scripts* computacionais voltados para a visualização gráfica foram estruturados de forma independente para os indicadores nacionais e estaduais, mantendo a lógica de segmentação anterior. Utilizando a biblioteca *Plotly* da linguagem *Python*, foram criados múltiplos modelos de gráficos adaptados às particularidades de cada conjunto de dados. O objetivo central dessa etapa foi buscar garantir que cada variável fosse representada pelo formato visual que melhor comunicasse seu comportamento e suas tendências.

O gráfico de cascata foi o modelo escolhido para ilustrar o comportamento do faturamento líquido e bruto, destacando visualmente os fatores de crescimento ou redução desse indicador ao decorrer dos anos. Em paralelo, utilizou-se o gráfico de colunas empilhadas para mapear variáveis que exigiam a análise de subcategorias em relação ao todo. Essa abordagem visual foi empregada no detalhamento do volume de usuários por segmento (RaC ou GTF), na quantidade de locadoras operantes com ou sem motorista, e no perfil dos colaboradores diretos por gênero e escolaridade em âmbito nacional.

Devido à ausência de dados detalhados por gênero e escolaridade no âmbito das unidades federativas, a análise temporal do número de empregados diretos por estado apresentou-se mais simples, justificando a escolha do gráfico de colunas para sua representação. Já para ilustrar a distribuição geográfica do setor, utilizou-se o mapa coroplético, cuja finalidade é evidenciar, por meio de diferentes tonalidades de cores, as regiões que concentram a maior quantidade de locadoras.

Por fim, o gráfico de linha foi o modelo selecionado para acompanhar o comportamento histórico do volume de emplacamentos, bem como o seu desdobramento por subsegmento e fabricante ao decorrer dos anos. A escolha desse e dos demais formatos de gráficos seguiu as recomendações analíticas propostas na obra denominada: “Visualização de Dados: Explorando e Explicando com Dados”, de Camm, Cochran e Fry (2026).

Para a definição das cores, buscando obter um contraste adequado para a visualização das informações, utilizou-se uma paleta disponibilizada pelo *Chart Color Generator*¹³. Essa ferramenta *web* gratuita, é voltada especificamente para a geração de esquemas cromáticos apropriados para gráficos, documentos e apresentações.

4.4 Publicação dos dados

No âmbito do uso e da publicação dos dados, projetou-se a construção de um *dashboard* interativo com o propósito de disponibilizar aos tomadores de decisão, de maneira clara e sintética, os indicadores essenciais do setor (Carvalho; Melo, 2018). Em um primeiro momento, a *interface* gráfica foi prototipada com o auxílio da ferramenta *Figma*, com o intuito de estruturar um *dashboard* que prezasse pela clareza visual e exibisse estritamente as informações necessárias para o segmento de locação.

Para o desenvolvimento computacional da plataforma, utilizou-se o *framework* de código aberto *Streamlit*, baseado na linguagem *Python*, cuja escolha justificou-se pela capacidade de converter *scripts* de dados em painéis dinâmicos de forma ágil e integrada. Contudo, devido a limitações estéticas identificadas nos padrões nativos da ferramenta, fez-se necessária a incorporação de CSS, atuando na definição e personalização dos elementos estruturais para garantir a máxima fidelidade ao protótipo construído.

A construção do *dashboard* em questão foi projetada para responder a uma série de questionamentos sobre o panorama atual e histórico do setor de locação. Por meio das visualizações propostas, busca-se compreender: em que ritmo se deu o crescimento econômico do mercado sob a ótica dos faturamentos históricos, como se comportou a expansão dos polos de atuação e quais regiões concentram a maior força do setor, qual o real impacto da atividade na geração de empregos diretos ao longo dos anos. Além disso, sob uma perspectiva mais técnica, identificar quais fabricantes e subsegmentos são efetivamente mais demandados pelas locadoras. São essas lacunas que as análises subsequentes buscam preencher.

¹³ Disponível em: <https://colors.design4u.jp/>

5 RESULTADOS

No presente capítulo, apresentam-se os resultados alcançados neste estudo, os quais serão analisados e discutidos nos tópicos subsequentes para detalhar cada fase do projeto. Na seção 5.1, apresenta-se a estruturação e a disponibilização dos dados coletados sobre o setor de locação. Na seção 5.2, realiza-se uma descrição detalhada do *dashboard* desenvolvido para a análise do mercado de veículos, evidenciando suas principais funcionalidades. Por fim, na seção 5.3, detalham-se os *insights* obtidos das análises gráficas, englobando a evolução histórica de usuários, os impactos no faturamento, a geração de empregos com recorte de gênero e o volume de emplacamentos do setor.

5.1 Estruturação e disponibilização dos dados coletados

Como resultado do processo de extração e centralização dos dados, foi possível consolidar um conjunto de sete tabelas no âmbito nacional, as quais cobrem diferentes dimensões do mercado automobilístico e socioeconômico de locação. Esse conjunto de tabelas compreende indicadores de desempenho do setor, organizados nos seguintes arquivos estruturados: “*faturamento_brasil*”, que mapeia o crescimento econômico bruto e líquido do setor; “*br_automoveis*”, contendo a série histórica de emplacamentos anuais; e “*br_quant_usuarios*”, detalhando o volume de usuários e a divisão percentual entre os eixos *Rent a Car* e Gestão e Terceirização de Frotas.

Complementarmente, a base de dados abrange as dinâmicas operacionais e mercadológicas por meio das tabelas “*br_emplacamento_por_fabricante*” e “*br_emplacamento_por_subsegmento*”, voltadas à análise de marcas e categorias de veículos, além de registrar a estrutura do mercado de trabalho e o panorama empresarial do segmento através dos arquivos “*br_empregos_diretos*” e “*br_numero_de_locadoras*”.

Em contrapartida, as informações disponibilizadas nos anuários da ABLA em escopo regional apresentam-se de forma mais restrita, resultando na obtenção de apenas 4 tabelas por estado após a etapa de extração. Buscando garantir a organização e a rastreabilidade desse volume de dados, adotou-se um padrão de nomenclatura estruturado pelo prefixo *{uf}_nome_do_arquivo*, no qual a abreviação inicial corresponde à sigla da Unidade Federativa pesquisada.

Esse conjunto de dados regionalizados subdivide-se nos arquivos “*{uf}_carros_por_fabricante*” e “*{uf}_carros_por_subsegmento*”, desenvolvidos para identificar as marcas predominantes e o perfil de consumo da frota local, e nas tabelas

“*{uf}_empregos_diretos*” e “*{uf}_numero_locadoras*”, que refletem, respectivamente, a evolução da força de trabalho e o histórico da quantidade de locadoras na região.

Importante salientar que o presente estudo realizou a extração de dados de todas as unidades federativas da região Nordeste do Brasil: Alagoas (AL), Bahia (BA), Ceará (CE), Maranhão (MA), Paraíba (PB), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN) e Sergipe (SE), totalizando uma base de dados regional de 36 tabelas. Todo esse *dataset*, tanto em âmbito estadual quanto nacional, foi disponibilizado publicamente na plataforma *Kaggle*, almejando garantir a transparência e a reprodutibilidade dos dados obtidos nesta pesquisa.

A escolha do *Kaggle*¹⁴ como repositório para o *dataset* desenvolvido justifica-se por sua relevância como uma plataforma global voltada à comunidade de CD e IA, atuando como um ecossistema colaborativo que abriga milhões de desenvolvedores e pesquisadores. A plataforma destaca-se pelo acesso a um vasto catálogo de conjuntos de dados públicos e uma biblioteca com dezenas de milhares de modelos pré-treinados prontos para uso.

Por fim, é válido ressaltar que os conjuntos de dados denominados como "Locação de Veículos no Brasil"¹⁵, âmbito nacional, e "Mercado de Locação de Veículos por Estado"¹⁶, âmbito estadual, disponibilizados no *Kaggle* apresentam-se na camada *Bronze* da Arquitetura *Medallion*, possuindo como único tratamento prévio o agrupamento das informações brutas em suas respectivas tabelas. Os processos subsequentes de limpeza e padronização para a estruturação da camada *Silver*, bem como o enriquecimento das variáveis que compõem a camada *Gold*, constituem etapas de responsabilidade exclusiva do pesquisador.

5.2 *Dashboard* para análise do mercado de locação de veículos

Araújo (2019) enfatiza que o uso de *dashboards* e elementos gráficos vai além do simples apelo visual, tornando-se um componente fundamental para a tomada de decisões estratégicas e para a compreensão aprofundada de dados complexos. Essas ferramentas transformam registros brutos em representações visuais intuitivas, diminuindo significativamente o esforço mental necessário para interpretar métricas. Isso permite que os analistas detectem tendências e anomalias de forma mais rápida e precisa.

Reconhecendo as vantagens da análise visual por meio de painéis dinâmicos, o presente estudo propôs o desenvolvimento do *Dashboard* do Mercado de Locação de Veículos

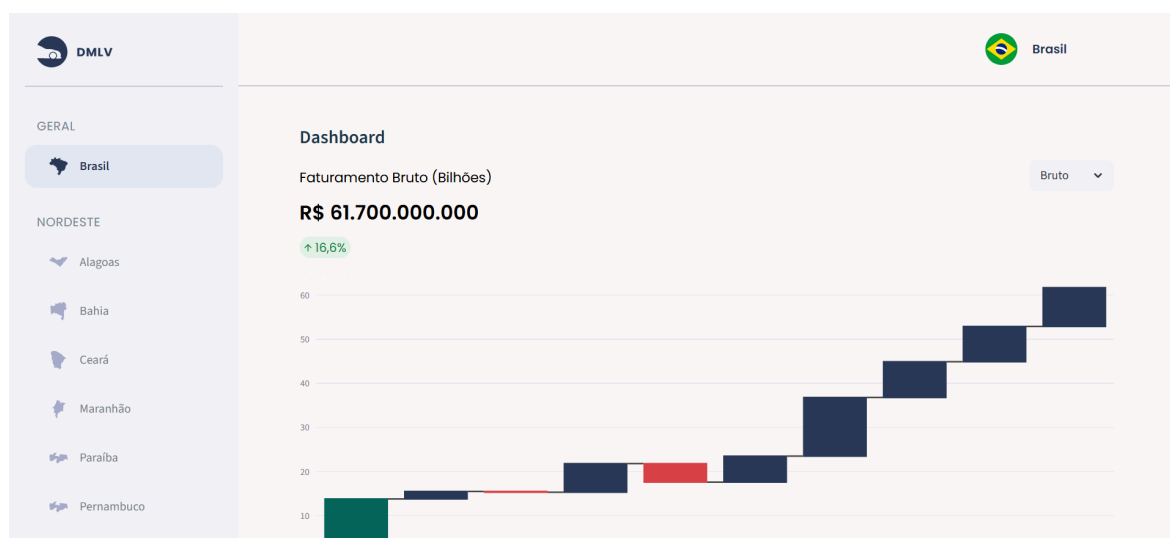
¹⁴ Disponível em: <https://www.kaggle.com/>

¹⁵ Disponível em: <https://www.kaggle.com/datasets/erikyveloso/locao-de-veculos-no-brasil>

¹⁶ Disponível em: <https://www.kaggle.com/datasets/erikyveloso/mercado-de-locao-de-veculos-por-estado>

(DMLV). Conforme ilustrado na Figura 1, a *interface* é estruturada a partir de um menu lateral que permite ao usuário alternar entre a exibição de dados de âmbito nacional ou de estados específicos da região Nordeste. Além de centralizar os filtros de navegação, este menu cumpre a função de indicar de forma clara a localização referente às variáveis em exibição no momento.

Figura 1 - Interface inicial do DMLV

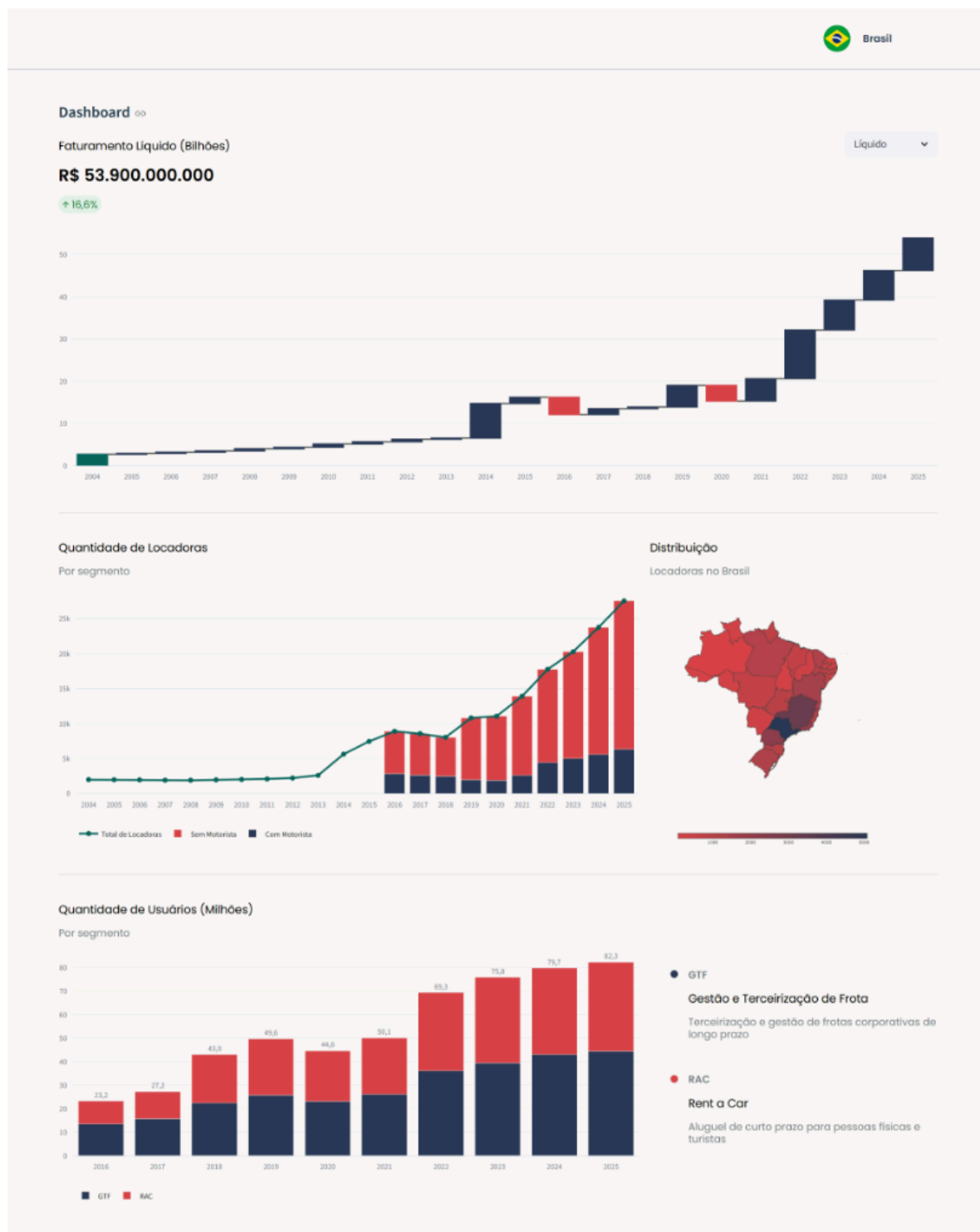


Fonte: Autoria Própria (2026).

Visando otimizar a visualização dos gráficos, o usuário tem a opção de ocultar o menu lateral, ampliando o espaço dedicado ao conteúdo principal. Para assegurar que essa ação não prejudique a informação da região de origem dos dados, a parte superior da página conta com um cabeçalho fixo (Figura 1). Esse elemento exibe continuamente o nome e a bandeira da localidade selecionada, permitindo que o usuário identifique rapidamente a origem das informações em análise, mesmo quando o menu de navegação estiver recolhido.

Em relação às informações em âmbito nacional, o *dashboard* inicia-se com um gráfico em cascata (Figura 2), projetado para evidenciar o efeito cumulativo das variações anuais do setor. Essa visualização permite analisar tanto o faturamento líquido, no período de 2004 a 2025, quanto o faturamento bruto, em um recorte temporal mais recente, de 2016 a 2025, cabendo ao usuário selecionar a métrica desejada. Complementando a análise gráfica, a interface exibe previamente o valor aproximado do faturamento no último ano mapeado, acompanhado do respectivo percentual de crescimento em comparação ao período anterior.

Figura 2 - Dashboard com as informações nacionais (Parte 1)



Fonte: Autoria Própria (2026).

Na sequência, a Figura 2 apresenta a quantidade de locadoras por segmento em um gráfico de colunas empilhadas, no qual a tonalidade vermelha identifica as empresas com

motorista e a azul, aquelas sem motorista. Sobreposto a essa visualização, inseriu-se um gráfico de linhas que exibe o total geral de locadoras ativas. Essa decisão de *design* foi tomada para maximizar a exposição dos dados históricos, visto que, no período de 2004 a 2015, a fonte de dados original não realizava a segregação entre as categorias de atuação do setor.

Ao lado do gráfico da quantidade de locadoras, foi *plotado* um mapa coroplético com o objetivo de evidenciar a distribuição geográfica e a concentração dessas empresas pelo território nacional. A visualização utiliza uma escala cromática na qual a intensidade das cores indica a densidade do mercado: estados com tonalidades mais próximas ao azul apresentam o maior número de locadoras ativas, enquanto as regiões que se aproximam do vermelho indicam uma menor quantidade de estabelecimentos.

Posteriormente, conforme ilustrado na seção inferior da Figura 2, apresenta-se o volume de usuários das locadoras por meio de um gráfico de colunas empilhadas, segregado entre as modalidades RaC e GTF. Considerando que se trata de siglas estritamente específicas do setor, foi inserida uma breve seção explicativa adjacente ao gráfico, fornecendo as definições conceituais de cada segmento com o intuito de facilitar a compreensão e a acessibilidade da leitura para o público geral.

Ainda em relação ao panorama nacional, apresenta-se uma tabela que detalha o quantitativo de trabalhadores do setor segregados por nível de instrução (Figura 3). Essa estrutura reflete o cenário com base nos dados consolidados de 2025 (último período disponível), distribuídos entre as seguintes categorias de escolaridade: fundamental incompleto e completo, médio incompleto e completo, superior incompleto e completo.

Complementarmente, adjacente à tabela de escolaridade, apresenta-se um gráfico de colunas empilhadas que mapeia a evolução do número de trabalhadores no setor, segregado por gênero. A visualização utiliza a tonalidade azul representando o agrupamento masculino e a vermelha, o feminino, permitindo que o usuário analise o crescimento do volume total de colaboradores ao longo dos anos, bem como as variações na representatividade e na igualdade de gênero dentro do mercado de locação.

Posteriormente, para evidenciar a participação anual do setor no total de emplacamentos do país, foi construído um gráfico de área empilhada. Esse gráfico compara o número de emplacamentos das locadoras com o total acumulado do Brasil ao longo dos anos, permitindo enxergar visualmente o tamanho do setor dentro do mercado automobilístico. Logo acima, o painel destaca o número exato de emplacamentos realizados pelas locadoras no

Por fim, com o objetivo de aprofundar a análise e gerar novos *insights*, o *dashboard* apresenta dois gráficos de linhas com o volume de emplacamentos anuais distribuídos por fabricante e por subsegmento, como ilustrado ao final da Figura 3. Cabe destacar que, para mitigar a poluição visual e preservar a clareza do *layout*, esses gráficos foram configurados com filtros ativos por padrão. Dessa forma, a *interface* exibe inicialmente apenas as três fabricantes e os três subsegmentos com maior volume de emplacamentos no último ano. Contudo, o gráfico mantém sua propriedade interativa, permitindo que o usuário habilite as demais marcas e subcategorias diretamente na legenda, de acordo com o seu interesse de análise.

No que se refere ao painel de informações estaduais, todas as unidades federativas compartilham a mesma estrutura visual e configurações do exemplo demonstrado nas Figuras 4 e 5, que utilizam os dados do Rio Grande do Norte como modelo.

Conforme ilustrado na Figura 4, a seção estadual inicia-se com um gráfico de colunas empilhadas que apresenta a série histórica do quantitativo de locadoras ativas no estado entre os anos de 2017 e 2025. Essa visualização adota o mesmo padrão de segmentação do panorama nacional, diferenciando as empresas que oferecem serviços com motorista, na tonalidade azul, daquelas sem motorista, na tonalidade vermelha. Auxiliando à exibição gráfica, a *interface* destaca o número de locadoras ativas no estado no último ano analisado. Esse indicador é acompanhado pelo percentual de variação em relação ao período anterior, permitindo identificar de forma direta o ritmo de crescimento ou retração do mercado local.

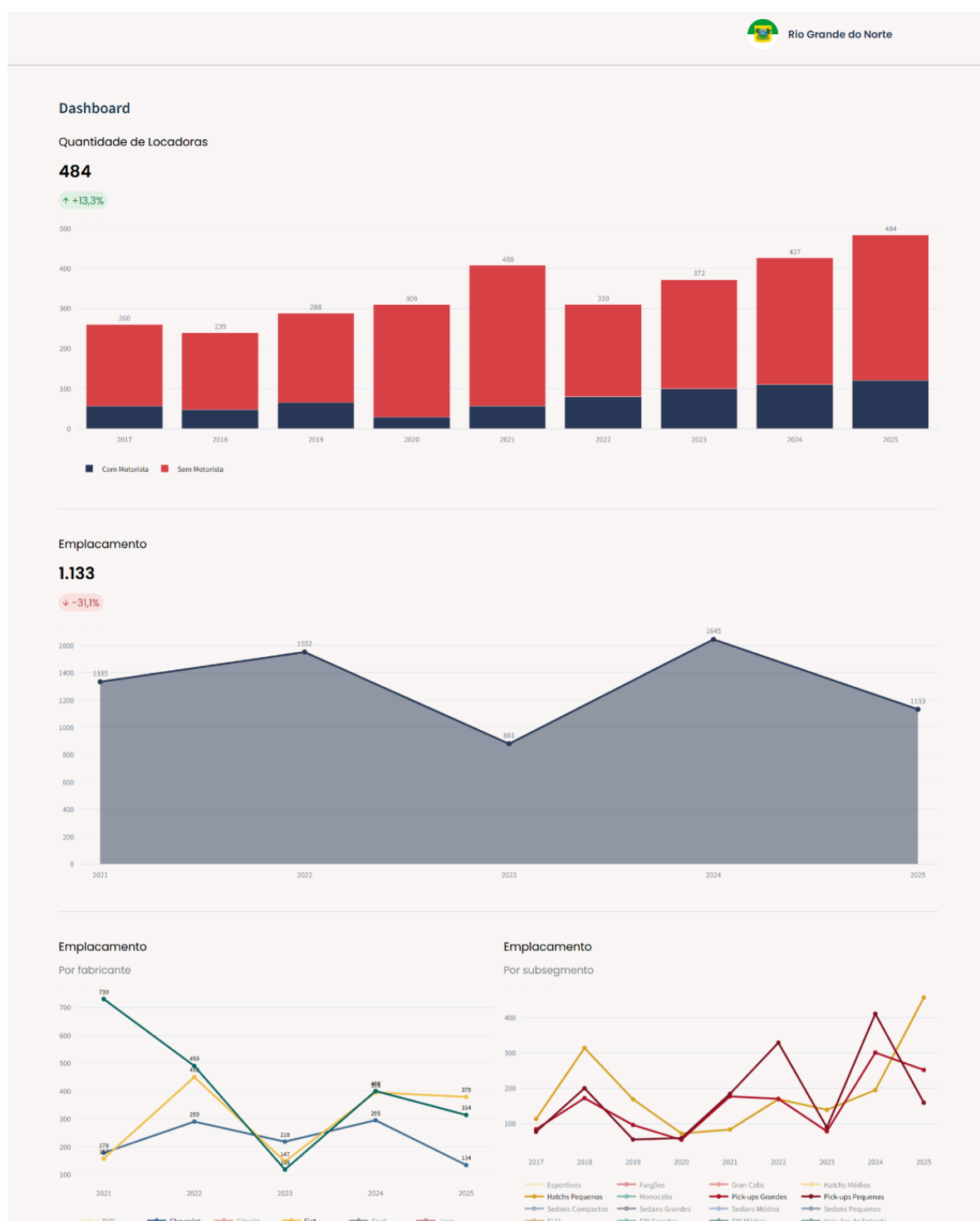
Posteriormente, mantendo o padrão visual adotado no âmbito nacional, os dados sobre o volume de emplacamentos anuais no estado são apresentados por meio de um gráfico de área. Para facilitar a leitura rápida das informações, o painel destaca o número exato de emplacamentos do último ano no canto superior esquerdo da visualização, acompanhado pelo percentual de variação em relação ao período anterior, como evidenciado na Figura 4.

Além disso, com o objetivo de possibilitar que os analistas identifiquem a preferência de mercado das locadoras estaduais em relação às montadoras e às modalidades de atuação, a *interface* disponibiliza dois gráficos de linhas posicionados logo abaixo do gráfico de área supracitado. Essas visualizações exibem o histórico de emplacamentos anuais, segmentadas individualmente por fabricante e por subsegmento.

Para otimizar a experiência do usuário e evitar a sobreposição de linhas na tela, os painéis contam com um filtro aplicado por padrão que exibem inicialmente apenas as três marcas e os três subsegmentos que registraram o maior volume de emplacamentos no último

ano da base de dados, como exposto na Figura 4. Essa abordagem busca preservar a clareza visual do *layout*, sem limitar o poder de exploração do analista, que pode reativar as demais categorias a qualquer momento simplesmente clicando sobre os elementos da legenda.

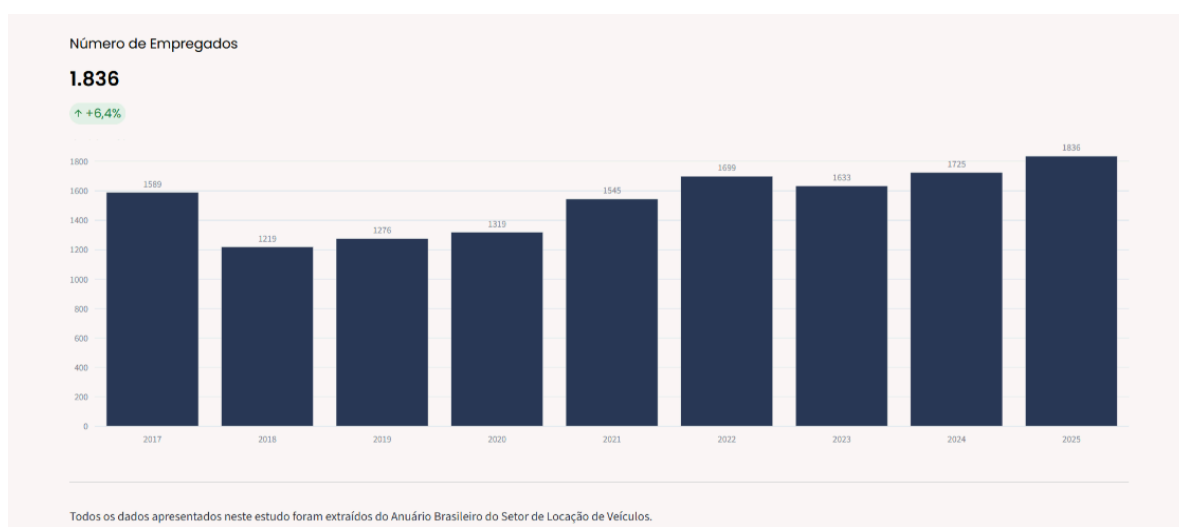
Figura 4 - Dashboard com as informações estaduais (Parte 1)



Fonte: Autoria Própria (2026).

Para finalizar a seção de dados estaduais, a *interface* apresenta um gráfico de colunas que ilustra a série histórica do número de empregos diretos gerados pelo setor na região, evidenciado na Figura 5. Cabe ressaltar que, diferentemente do panorama nacional, não foi possível realizar a segmentação desses dados por gênero ou por nível de instrução, devido a uma limitação na base de dados original, que não disponibiliza essas variáveis de forma detalhada em âmbito estadual.

Figura 5 - Dashboard com as informações estaduais (Parte 2)



Fonte: Autoria Própria (2026).

Seguindo o padrão de usabilidade adotado nos componentes anteriores, o volume total de empregos referente ao último ano ganha destaque na área superior do gráfico, sendo acompanhado pelo percentual de variação em relação ao período anterior para indicar a tendência de crescimento ou retração na geração de vagas (Figura 5).

Conforme observado nas Figuras de 1 a 5, o DMLV utiliza como identidade visual variações de branco e de azul, evitando o uso do branco puro com o objetivo de reduzir a fadiga e o cansaço visual do usuário durante a navegação. Além disso, cabe destacar que a maioria dos gráficos gerados pela plataforma possui recursos nativos de interatividade.

A *interface* permite que o usuário realize o *download* das visualizações em formato *.png*, selecione áreas específicas da tela para ampliação e recorte temporal personalizado, e movimente o eixo dos gráficos para navegar entre diferentes períodos históricos. O painel também disponibiliza comandos para ajuste automático de *zoom*, redução de escala e retorno

ao enquadramento inicial, além da opção de expandir o gráfico para o modo de tela cheia, facilitando uma análise mais minuciosa por parte do analista.

Além de todas as funcionalidades mencionadas, a plataforma conta com recursos de suporte visual para buscar solucionar eventuais dúvidas sobre os dados apresentados. Ao posicionar o cursor do mouse sobre qualquer elemento dos gráficos, seja um ponto, barra ou linha, o sistema exibe uma caixa de texto explicativa contendo o detalhamento completo e os valores exatos daquela informação específica.

Contudo, por questões de otimização e desempenho do sistema, o mapa coroplético da distribuição de locadoras por estados não conta com essas ferramentas de interação. Essa decisão de desenvolvimento foi necessária para assegurar uma plataforma mais leve e fluida, visto que o processamento dessas renderizações interativas no mapa tornava a navegação lenta, ocasionada pela necessidade de leitura do arquivo GeoJSON, comprometendo a usabilidade do *dashboard*.

5.3 Insights e conclusões das análises gráficas

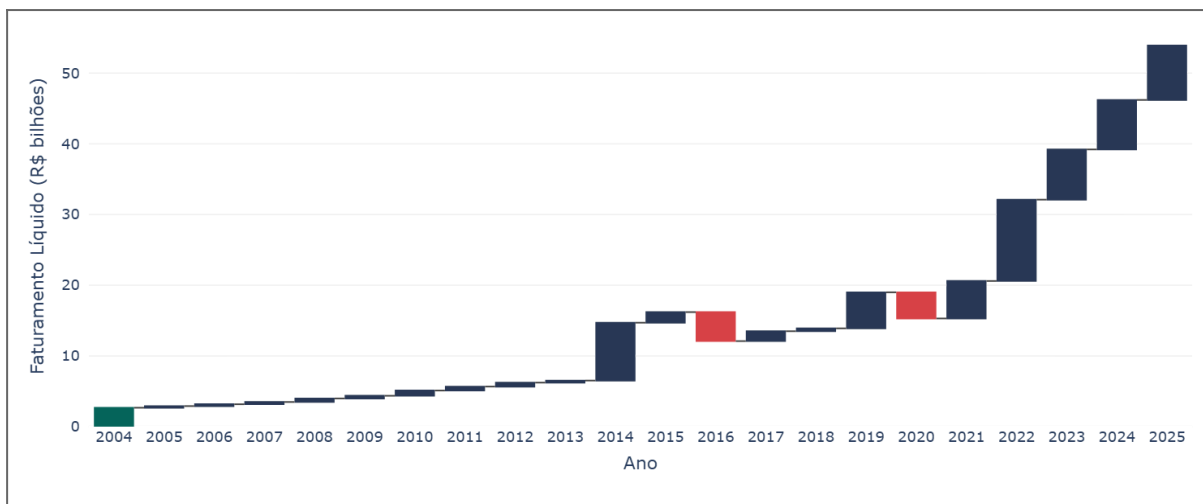
Para o presente trabalho, é fundamental compreender a relevância econômica e a grande capacidade de geração de receita que o setor de locação de veículos possui no mercado nacional. Segundo a ABLA, o segmento opera com um faturamento expressivo e mantém uma trajetória de crescimento constante, mostrando-se um negócio sólido e lucrativo mesmo diante de cenários econômicos instáveis. Esse forte desempenho financeiro se reflete no montante investido anualmente na ampliação das frotas e no volume de faturamento bruto gerado pelas empresas, o que impulsiona diretamente o comércio (ABLA, 2026).

Corroborando a afirmação exposta no anuário da ABLA de 2026, os Gráficos 1 e 2, que apresentam o histórico de faturamento líquido e bruto, demonstram a evolução financeira do setor. O gráfico em modelo cascata evidencia um crescimento financeiro constante na maioria dos períodos analisados, representado pelas barras na cor azul. As únicas retrações no faturamento líquido ocorreram no ano de 2016, em relação ao ano anterior, e em 2020, quando comparado ao resultado obtido em 2019.

Uma forte hipótese consiste em afirmar que a principal motivação para a queda em 2020 foi a pandemia da covid-19, declarada pela Organização Mundial da Saúde em março do ano em análise. A crise sanitária impôs um choque severo à economia brasileira, interrompendo as projeções de crescimento e afetando severamente diversos mercados. Especialmente, as atividades de turismo e viagens corporativas, que foram os segmentos mais

afetados e os últimos a retomarem a normalidade (Gullo, 2020), impactando diretamente o mercado de locação, visto a existência de um forte vínculo entre os setores.

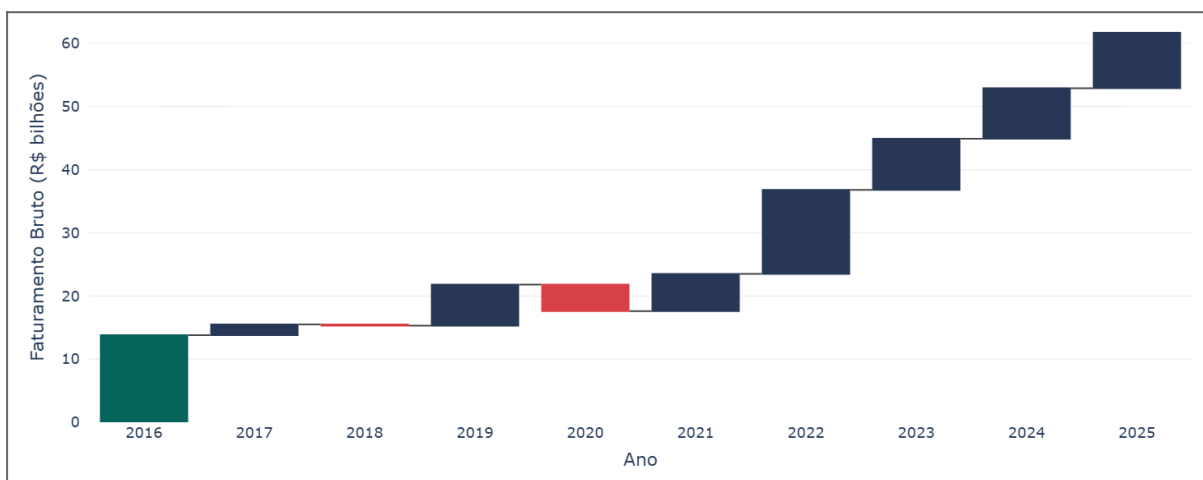
Gráfico 1 - Oscilações do faturamento líquido ao longo dos anos



Fonte: Autoria Própria (2026).

Entretanto, um ponto digno de destaque é a rápida recuperação tanto do faturamento líquido quanto do bruto logo no ano seguinte, em 2021. Mesmo com o Brasil ainda vivenciando o cenário pandêmico e suas restrições, os Gráficos 1 e 2 demonstram uma retomada expressiva nos ganhos. Esse comportamento evidencia a solidez que o mercado de locação de veículos alcançou no país, reforçando a capacidade de adaptação e a relevância do setor para a economia nacional, conforme apontado pelo anuário da ABLA.

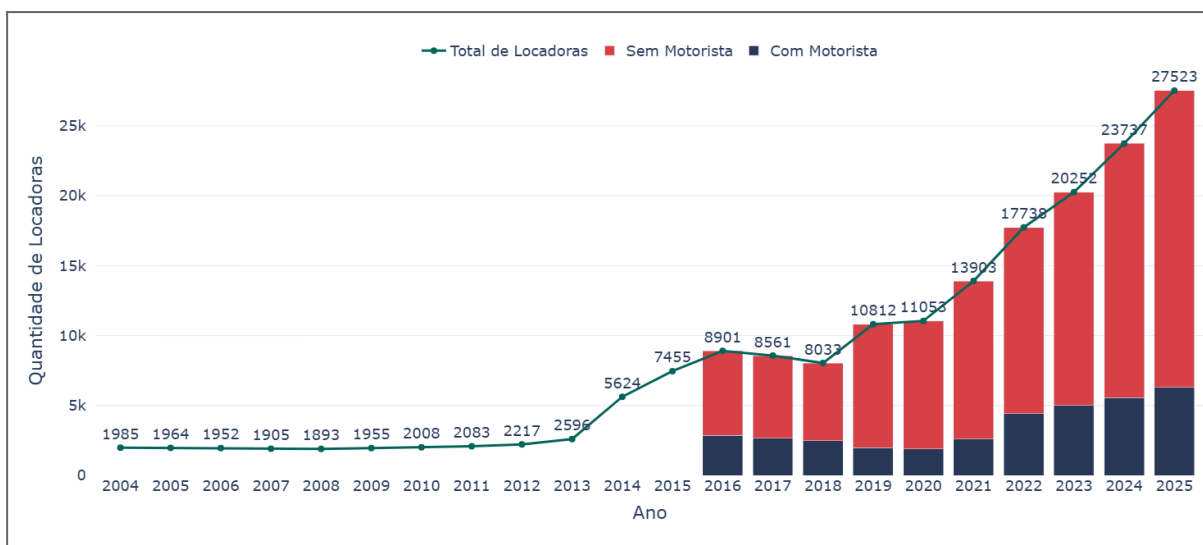
Gráfico 2 - Oscilações do faturamento bruto ao longo dos anos



Fonte: Autoria Própria (2026).

Outra forma de comprovar a solidez do mercado destacada pela ABLA é a evolução na quantidade de empresas ativas no país. Apresentado no Gráfico 3, esse indicador aponta para um crescimento contínuo desde 2018, culminando no total de 27.523 locadoras de veículos em território brasileiro ao final de 2025.

Gráfico 3 - Evolução da quantidade de locadoras ao longo dos anos



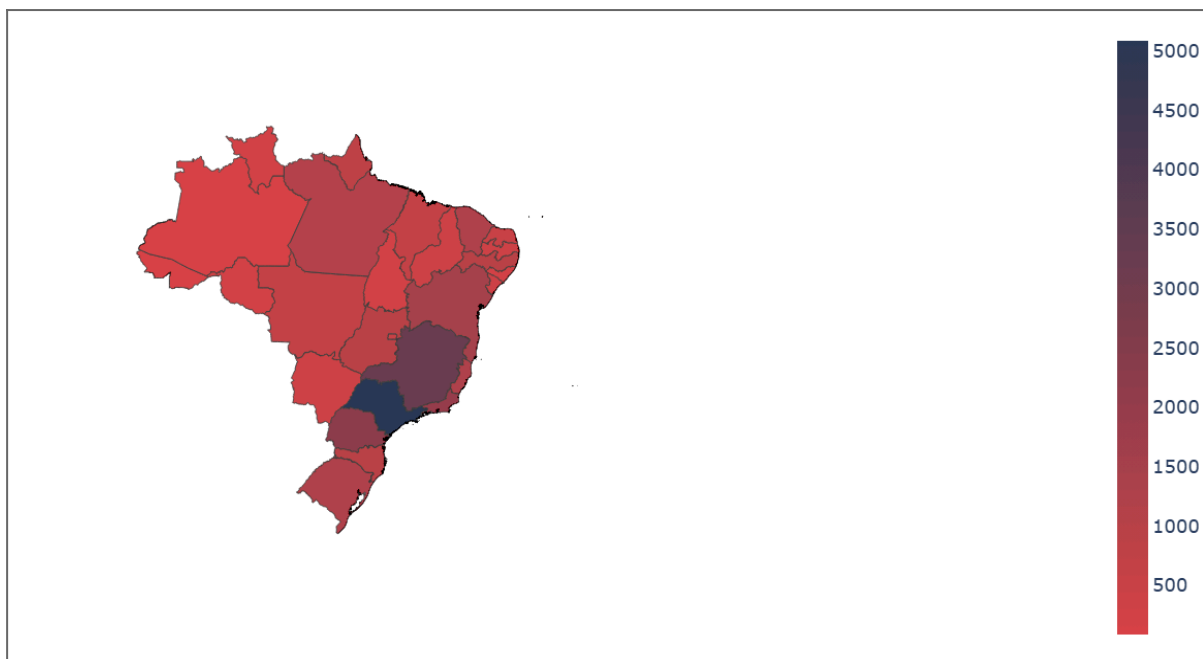
Fonte: Autoria Própria (2026).

O Gráfico 3 também revela a ampla predominância, em todos os anos da série histórica, das locadoras que realizam aluguel de veículos sem motorista (representadas pela cor vermelha) em comparação com as que oferecem o serviço com motorista (representadas pela cor azul). No ano de 2025, por exemplo, foram identificadas 21.217 empresas atuando sem motorista contra apenas 6.306 com motorista, o que significa que o segmento sem motorista concentrou aproximadamente 77,09% do mercado total de empresas do setor.

Entretanto, apesar do crescimento geral no número de locadoras no país demonstrado no Gráfico 3, o Gráfico 4 evidencia uma forte concentração dessas empresas na Região Sudeste, com especial relevância para o estado de São Paulo. No mapa coroplético, o território paulista se diferencia visualmente de todas as demais unidades federativas por ser o único a atingir a tonalidade azul, que representa valores iguais ou superiores a 5.000 locadoras. Historicamente, o estado possui um vínculo com o setor, visto que o mercado de locação nacional teve início justamente no centro da cidade de São Paulo. A região também

foi o berço da profissionalização da atividade no país, iniciada em 1956 com a fundação da Auto Drive S.A. (Luiz, 2018).

Gráfico 4 - Distribuição geográfica das locadoras no Brasil



Fonte: Autoria Própria (2026).

Complementando os dados apresentados no mapa coroplético, a Tabela 1 evidencia numericamente a liderança da Região Sudeste no mercado nacional de locação, em comparação com as demais regiões do país. O Sudeste concentra 2.417 estabelecimentos de locação com motorista, o que representa cerca de 38,33% do total nacional deste segmento. No mercado de locadoras sem motorista, a participação regional é ainda mais expressiva, somando 9.053 empresas ativas, o que equivale a aproximadamente 42,67% do total do país.

Analisando especificamente o estado de São Paulo, que já havia se destacado no mapa coroplético, nota-se que o território paulista possui 3.980 estabelecimentos no segmento de locação sem motorista. Este valor representa cerca de 18,76% do total nacional da categoria, superando, individualmente, o total de quase todas as outras regiões brasileiras. Conforme observado na Tabela 3, a Região Sul conta com 3.838 empresas, aproximadamente 18,09% do segmento, enquanto a Tabela 4 aponta que a Região Norte possui 1.966 locadoras, cerca de 9,27%. Por fim, a menor participação no mercado de locação sem motorista pertence à Região Centro-Oeste, detalhada na Tabela 5, que concentra 1.933 empresas ativas, representando aproximadamente 9,11% do segmento.

Semelhantemente, no segmento de locação com motorista, o estado de São Paulo também ocupa uma parcela significativa do mercado. Com 1.104 estabelecimentos, o território paulista representa aproximadamente 17,51% do total nacional da categoria, superando toda a Região Norte (Tabela 4), que concentra 880 unidades e responde por cerca de 13,95% do segmento. Na sequência, aparece a Região Sul (Tabela 3), com 593 locadoras com motorista, o que representa aproximadamente 9,40% do setor. Por fim, a Região Centro-Oeste apresenta a menor participação, com apenas 576 unidades registradas, equivalente a 9,13% do segmento.

Tabela 1 - Distribuição geográfica das locadoras - Sudeste (2025)

Regiões	Estados	Locadoras de Carro	
		Com Motorista	Sem Motorista
Sudeste	São Paulo	1.104	3.980
	Minas Gerais	556	2.646
	Rio de Janeiro	462	1.458
	Espírito Santo	295	969
	Total	2.417	9.053

Fonte: Adaptado de ABLA (2026).

A Região Nordeste, escolhida para ser detalhada no DMLV, é a única que, em sua totalidade, consegue superar o estado de São Paulo em quantidade de locadoras ativas em ambos os segmentos. Conforme apresentado na Tabela 2, a região possui 4.427 locadoras ativas sem motorista, o que representa aproximadamente 20,86% do total nacional deste segmento. Essa parcela de participação se torna ainda maior quando analisado o mercado com motorista, no qual o Nordeste concentra 1.840 estabelecimentos, equivalente a cerca de 29,18% das unidades do país. Os estados que se destacam na Região Nordeste são a Bahia, o Ceará e Pernambuco, com 1.529, 1.253 e 988 empresas de locação de carros, respectivamente.

Tabela 2 - Distribuição geográfica das locadoras - Nordeste (2025)

Regiões	Estados	Locadoras de Carro	
		Com Motorista	Sem Motorista
Nordeste	Bahia	509	1.020
	Ceará	348	905
	Pernambuco	243	755
	Maranhão	203	433
	Rio Grande do Norte	121	363
	Piauí	138	276
	Sergipe	150	243
	Paraíba	75	262
	Alagoas	53	170
	Total	1.840	4.427

Fonte: Adaptado de ABLA (2026).

No que diz respeito à Região Sul, a Tabela 3 aponta o estado do Paraná como o principal destaque, contabilizando um total de 2.244 locadoras de veículos. Da mesma forma, na Região Norte, conforme exposto na Tabela 4, os estados que assumem maior protagonismo são o Pará e o Amapá, registrando 1.114 e 769 estabelecimentos, respectivamente.

Tabela 3 - Distribuição geográfica das locadoras - Sul (2025)

Regiões	Estados	Locadoras de Carro	
		Com Motorista	Sem Motorista
Sul	Paraná	216	2.028
	Rio Grande do Sul	250	966
	Santa Catarina	127	844
	Total	593	3.838

Fonte: Adaptado de ABLA (2026).

Tabela 4 - Distribuição geográfica das locadoras - Norte (2025)

Regiões	Estados	Locadoras de Carro	
		Com Motorista	Sem Motorista
Norte	Pará	338	776
	Amapá	226	543
	Roraima	113	156
	Rondônia	59	169
	Tocantins	56	170
	Amazonas	52	101
	Acre	36	51
	Total	880	1.966

Fonte: Adaptado de ABLA (2026).

Por fim, a Região Centro-Oeste, exposta na Tabela 5, que apresenta a menor participação nacional em quantidade de locadoras de veículos, tem como principal destaque o estado de Goiás, com 927 estabelecimentos ativos. Desse total, 705 atuam no segmento sem motorista, o que representa aproximadamente 36,47% do mercado regional nessa categoria. Já as outras 222 empresas operam na modalidade com motorista, sendo responsáveis por cerca de 38,54% do total da região para esse segmento.

Tabela 5 - Distribuição geográfica das locadoras - Centro-Oeste (2025)

Regiões	Estados	Locadoras de Carro	
		Com Motorista	Sem Motorista
Centro-Oeste	Goiás	222	705
	Mato Grosso	156	572
	Distrito Federal	121	350
	Mato Grosso do Sul	77	306
	Total	576	1.933

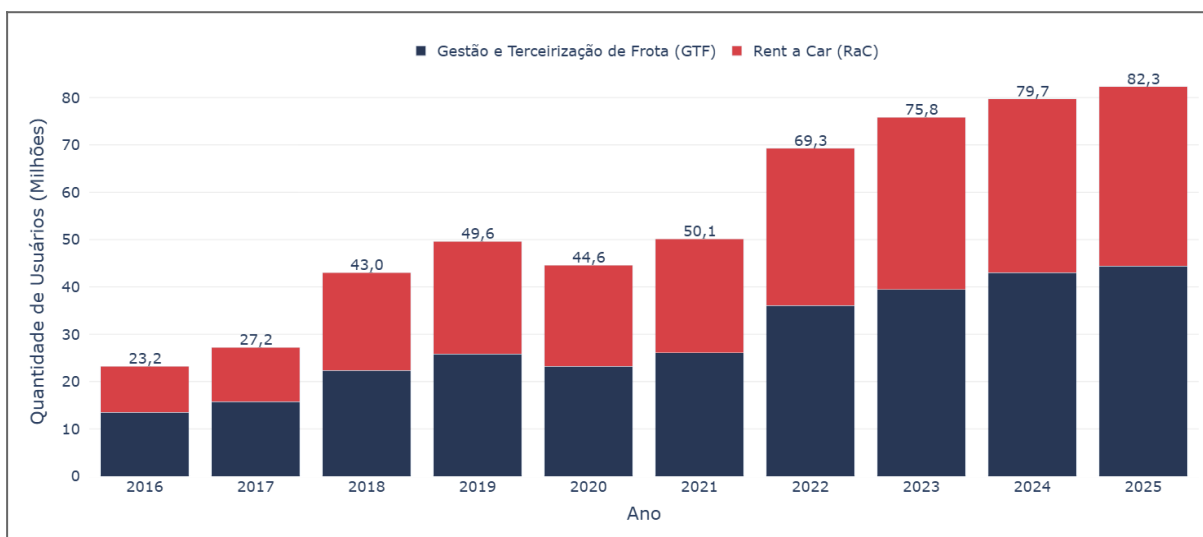
Fonte: Adaptado de ABLA (2026).

O crescimento na quantidade de locadoras de veículos, analisado anteriormente, mostra-se alinhado à demanda do mercado, visto que a busca por esse serviço apresenta uma tendência de alta. Conforme demonstrado no Gráfico 5, que considera o recorte histórico de 2016 a 2025, o número de usuários expandiu-se anualmente. A única exceção a essa trajetória ocorreu no ano de 2020, tendo como possível fator causador a necessidade de isolamento social imposta pela pandemia da covid-19.

Ao final de 2025, o setor contabilizou mais de 82 milhões de clientes. Esse volume expressivo, somado ao crescimento observado ao longo da série histórica, corrobora com a veracidade da afirmativa de Luiz (2018), ao enfatizar que o mercado de locação evoluiu de um serviço de nicho para uma indústria robusta, consolidando-se na marca de dezenas de milhões de usuários.

Outra análise histórica possibilitada por meio do Gráfico 5 refere-se à distribuição dos tipos de clientes do setor. A partir dos dados, nota-se o equilíbrio entre os segmentos, sem a predominância absoluta de uma modalidade específica. No último ano analisado (2025), o segmento de Gestão e Terceirização de Frotas, focado no atendimento corporativo, finalizou o período com 44.442 usuários, o que representa 54,00% do total do mercado. Complementarmente, o modelo *Rent a Car*, voltado a alugueis de curto prazo para pessoas físicas e turistas, encerrou o ano com 37.858 clientes, respondendo por 46,00% da base total. Essa pequena margem de diferença evidencia a relevância e a consolidação de público para ambas as modalidades de negócio.

Gráfico 5 - Evolução da quantidade de usuários ao longo dos anos

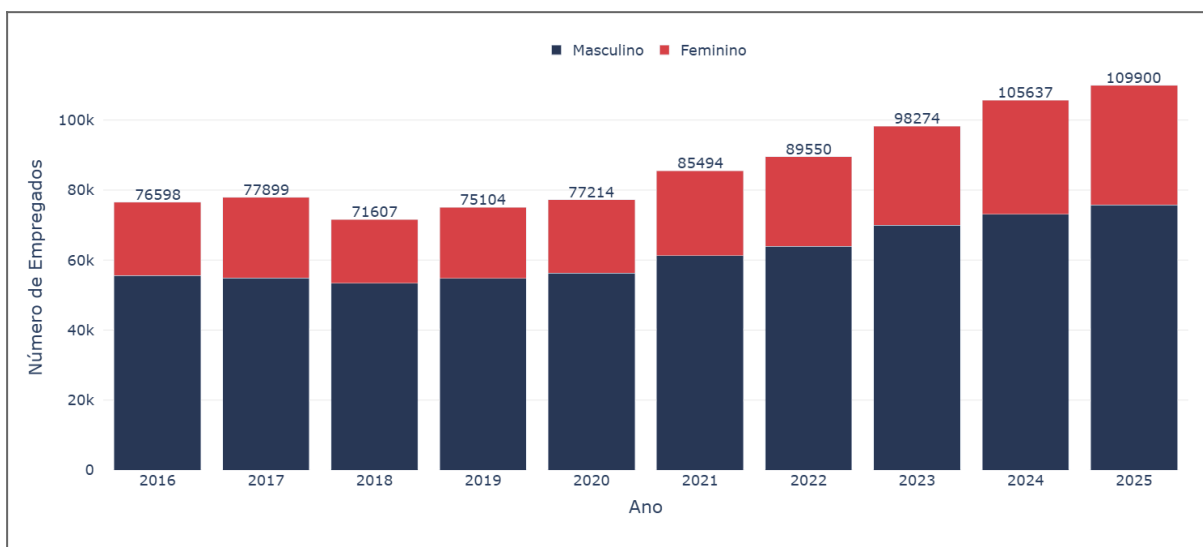


Fonte: Autoria Própria (2026).

Para suportar a expansão na quantidade de empresas de locação no Brasil, tornou-se necessária a ampliação no número de colaboradores diretos. Esse movimento pode ser observado por meio do Gráfico 6, que aponta uma trajetória de crescimento na maior parte do recorte histórico, registrando apenas uma leve retração no quadro de funcionários no ano de 2018 em relação a 2017. Essa expansão constante corrobora a ideia de Milhome (2022), que destaca a relevância socioeconômica do segmento de locação de automóveis, evidenciada tanto pelo seu expressivo faturamento quanto pelo seu papel na geração de postos de trabalho.

Ao final do recorte temporal analisado, os dados revelam que o setor atingiu a expressiva marca de 109.900 trabalhadores formais, esse número representa aproximadamente 0,28% do total de pessoas empregadas no setor privado com carteira assinada no país, estimado em 38.941.000 trabalhadores no mesmo período (IBGE, 2025). Embora percentualmente discreto no cenário macroeconômico, esse indicador torna-se significativo quando analisado para um único segmento econômico específico, reforçando o impacto direto e a relevância desta atividade na economia nacional.

Gráfico 6 - Distribuição dos empregados por gênero ao longo dos anos



Fonte: Autoria Própria (2026).

Entretanto, o Gráfico 6 também revela uma expressiva desigualdade de gênero, evidenciada pela predominância masculina no quadro de funcionários ao longo de todo o período analisado. Essa disparidade manifesta-se de forma ainda mais acentuada em momentos de retração do mercado, como observado na transição de 2017 para 2018, nesse intervalo, enquanto o segmento masculino sofreu uma contração de apenas 2,87%, passando

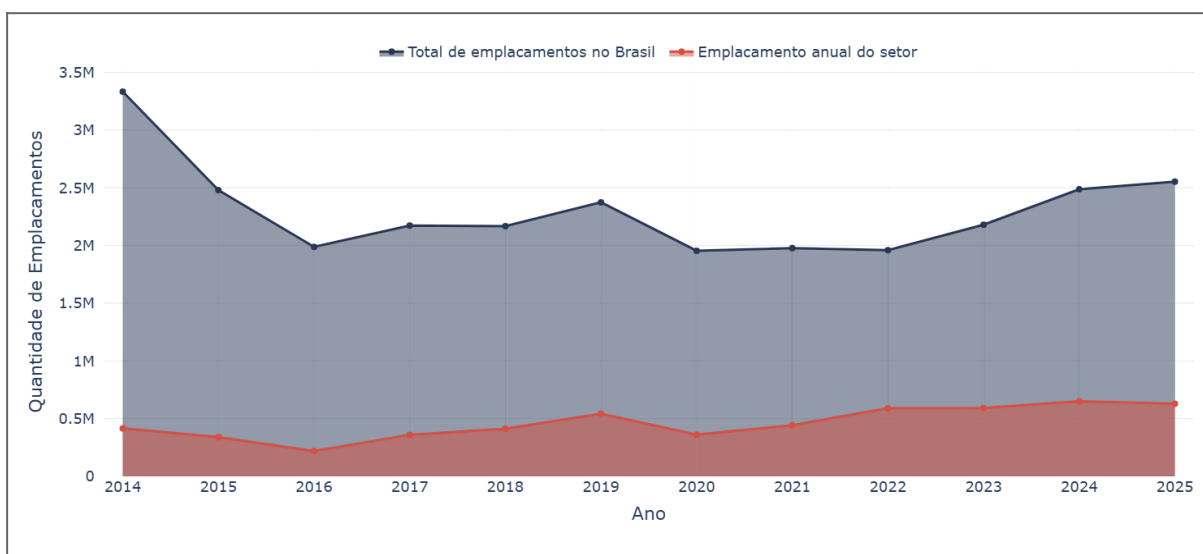
de 54.909 para 53.331 profissionais, a força de trabalho feminina foi desproporcionalmente afetada, registrando uma redução severa de 20,50%, com a perda de 4.714 postos de trabalho, reduzindo de 22.990 para 18.276.

No ano de 2025, o grupo de homens trabalhando no setor somou 75.726 profissionais, o que representa aproximadamente 68,90% do total de colaboradores. Além disso, ao comparar o aumento de contratações de 2024 para 2025, observa-se que foram admitidos 2.613 profissionais do gênero masculino contra 1.650 do gênero feminino, um cenário que reforça a disparidade de gênero no mercado de locação.

Ao analisar o período crítico da pandemia da covid-19, entre 2020 e 2023, o Gráfico 6 torna evidente a maturidade alcançada pelo mercado de locação. Mesmo enfrentando uma retração no faturamento líquido logo no início da crise sanitária em 2020, conforme apresentado no Gráfico 1, as empresas do setor demonstraram resiliência ao preservar a quantidade de funcionários e, gradativamente, expandir o número de postos de trabalho diretos nos anos seguintes.

Diante da expansão no número de locadoras no Brasil e da constante necessidade de atualização de suas frotas, o volume anual de emplacamentos realizados pelo setor alcançou números expressivos.

Gráfico 7 - Número de emplacamentos ao longo dos anos



Fonte: Autoria Própria (2026).

Em seu trabalho, Luiz (2018) enfatiza que as locadoras são as principais clientes das montadoras instaladas no país, absorvendo parcelas significativas da produção nacional, esse

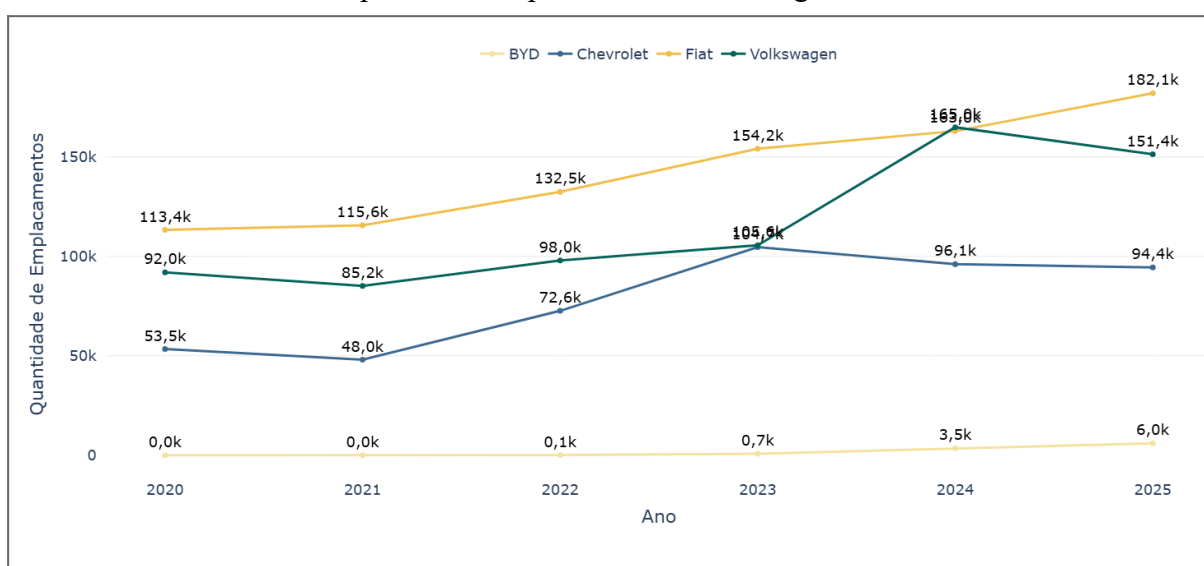
argumento é embasado no tamanho da frota total circulante e a taxa de aquisição de novos automóveis pelo setor.

Corroborando essa perspectiva, o Gráfico 7 evidencia a participação anual do setor no total de emplacamentos de veículos no Brasil. No ano de 2025, o mercado nacional registrou o licenciamento de 2.552.201 unidades, sendo o segmento de locação responsável por cerca de 24,64% desse montante, o que representou um total de 628.970 veículos emplacados.

Buscando obter informações mais detalhadas sobre os emplacamentos anuais realizados pelo setor, desenvolveu-se a análise apresentada nos Gráficos 8 e 9. Essas representações evidenciam tanto as fabricantes mais demandadas pelas locadoras quanto os subsegmentos de veículos que possuem maior aceitação entre os usuários do serviço.

A partir dos dados apresentados no Gráfico 8, observa-se que as montadoras com maior participação nos emplacamentos do setor no ano de 2025 foram a Fiat, a Volkswagen e a Chevrolet, respectivamente. Juntas, essas empresas concentraram aproximadamente 68,03% dos emplacamentos realizados pelo mercado de locação. A Fiat liderou o período com 28,95% de participação, totalizando 182.086 veículos emplacados, a Volkswagen ocupou a segunda posição com cerca de 24,07% (151.370 unidades), e a Chevrolet garantiu o terceiro lugar com uma parcela de 15,01% (94.410 automóveis).

Gráfico 8 - Número de emplacamentos por fabricante ao longo dos anos



Fonte: Autoria Própria (2026).

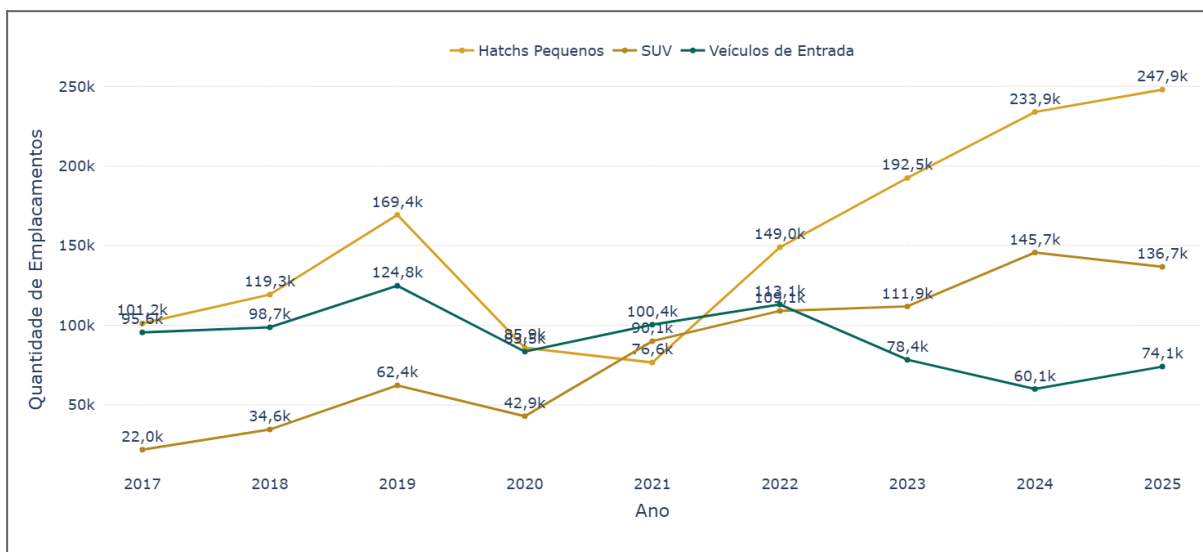
Além disso, sob uma perspectiva histórica, nota-se por meio do Gráfico 8 um crescimento consistente no interesse das locadoras pelos modelos da Volkswagen. A

montadora, que iniciou a série histórica na segunda colocação, alcançou o topo do ranking como a fabricante com o maior número de carros emplacados pelo setor no ano de 2024.

Outra possível análise a ser realizada através do Gráfico 8, é o crescente interesse das locadoras pelos carros elétricos, tendência evidenciada pelo aumento do número de veículos emplacados vinculados à montadora BYD¹⁷. A fabricante, detentora da maior fábrica voltada à transportes elétricos na América Latina, realizou o lançamento de seus primeiros modelos no país em 2021. A análise dos dados aponta que, no ano seguinte à sua estreia, o volume de emplacamentos da marca pelas locadoras foi de aproximadamente 100 unidades. Esse número saltou significativamente ao longo da série histórica, atingindo o total de 5.982 veículos emplacados pelas locadoras em todo o território brasileiro em 2025, o último período analisado.

Em relação aos subsegmentos, o Gráfico 9 evidencia que a maior participação nos emplacamentos do setor no ano de 2025 pertenceu às categorias de *Hatchs* Pequenos, SUVs e Veículos de Entrada, respectivamente. Juntos, esses três modelos concentraram aproximadamente 72,94% dos licenciamentos realizados pelo mercado de locação.

Gráfico 9 - Número de emplacamentos por subsegmento ao longo dos anos



Fonte: Autoria Própria (2026).

Observando o Gráfico 9, percebe-se que os *Hatchs* Pequenos lideraram o ranking em 2025 com 39,42% de participação, totalizando 247.939 veículos emplacados. Os SUVs (*Sport Utility Vehicles*) ocuparam a segunda posição, respondendo por cerca de 21,74% do segmento

¹⁷ Disponível em: <https://www.byd.com/br/byddobrasil>

(136.726 unidades), enquanto os veículos de entrada garantiram o terceiro lugar, com uma parcela de 11,78% (74.081 automóveis).

Vale salientar que, de maneira semelhante ao panorama observado no mercado de locação, os *hatchs* e os SUVs também exercem papéis de destaque no total de veículos no Brasil. Os autores Oliveira, Loureiro e Pessoa (2022) enfatizam que, no ano de 2022, os SUVs já constituíam a maioria no mercado nacional, correspondendo a 47,50% dos licenciamentos. Na sequência, aparecem os sedãs, com 25,80%, e os *hatchs*, na terceira posição, com 18,20%. O estudo citado corrobora a premissa de preferência da população brasileira por essas categorias de veículos, justificando o comportamento de compra das locadoras para atender a essa demanda.

6 CONCLUSÕES E TRABALHOS FUTUROS

O mercado de locação de veículos no Brasil consolida-se como um segmento de magnitude impressionante e de extrema importância para a economia nacional, tendo o segmento de carros como seu principal protagonista (Pucci, 2026). A estabilidade e a proporcionalidade dessa operação manifestam-se de forma nítida em seus indicadores macroeconômicos recentes: com uma frota que atingiu o patamar recorde superior a 1,7 milhão de veículos e um faturamento que ultrapassou a marca de R\$ 61 bilhões em 2025.

Diante da relevância e do crescimento contínuo do setor de locação de veículos, este estudo propôs a consolidação dos dados disponibilizados pela Associação Brasileira das Locadoras de Automóveis em uma base estruturada e unificada, visando ampliar o acesso às informações e apoiar o desenvolvimento de futuras pesquisas na área. Complementarmente, foi desenvolvido um *dashboard* interativo para facilitar a análise dos indicadores do setor, permitindo a identificação de tendências, padrões e variações ao longo do tempo.

O trabalho baseou o seu desenvolvimento em um processo estruturado de análise de dados fundamentado na Arquitetura *Medallion*. A aplicação dessa abordagem permitiu compreender o comportamento do mercado de locação de veículos no Brasil, possibilitando um estudo sobre o histórico do setor, além do fornecimento de *insights* para a geração futura de locatários.

Como primeira etapa do processo, a extração dos dados dos ABSLVs foi executada de forma bem-sucedida, utilizando a biblioteca *pdfplumber* para automatizar a leitura dos relatórios e a extração das tabelas. Da mesma forma, a fase subsequente de limpeza e tratamento das inconsistências na camada *Bronze*, realizada por meio da biblioteca Pandas, mostrou-se fundamental para eliminar os ruídos existentes. Ambos os processos, possibilitaram a criação de uma base de dados unificada e estruturada.

Suprindo um dos objetivos do presente trabalho, a base de dados foi disponibilizada publicamente por meio da plataforma Kaggle. Esse compartilhamento é fundamental, pois transforma os dados extraídos em um recurso acessível, preenchendo uma lacuna de dados centralizados sobre o setor. Ao disponibilizar os dois conjuntos de dados, compostos por 7 tabelas nacionais e 36 tabelas estaduais focadas na região Nordeste, o trabalho possibilita que outros pesquisadores e analistas realizem seus próprios estudos, promovendo a transparência e a reprodutibilidade na comunidade. No momento da escrita do presente trabalho, as métricas

da própria plataforma já registraram visualizações e *downloads* dos dados por outros usuários, o que reforça o valor prático e a importância dessa iniciativa para a comunidade de CD.

Outro marco importante para o atual trabalho foi a criação do *dashboard*, por transformar arquivos brutos em um produto visual e interativo. Esse painel facilitou a análise do mercado de locação, permitindo identificar tendências de forma mais ágil. Para a pesquisa, o desenvolvimento da ferramenta foi indispensável pois demonstrou, na prática, o potencial de construção e exploração que a base de dados organizada oferece, transformando o processo iniciado na extração em uma entrega real e palpável.

Considerando as análises realizadas através dos gráficos e tabelas construídos, observou-se uma trajetória de forte expansão financeira no setor, com um crescimento constante tanto no faturamento líquido quanto no bruto. Esse desempenho corrobora as projeções da ABLA (2026) de que o mercado de locação se consolidou como um negócio sólido e lucrativo. A solidez do setor também foi confirmada pelo aumento no número de empresas ativas, com uma ampla predominância do segmento de locação sem motorista.

No entanto, as análises evidenciaram uma forte concentração geográfica na Região Sudeste, impulsionada pelo estado de São Paulo. O território paulista polariza o setor ao concentrar isoladamente cerca de 18,76% das locadoras sem motorista e 17,51% das empresas com motoristas de todo o país, superando o total acumulado de quase todas as outras regiões brasileiras.

No aspecto social, as análises confirmaram a força do setor na geração de empregos diretos. Contudo, os dados acenderam um alerta sobre a desigualdade de gênero, visto que os homens representam aproximadamente 68,90% dos profissionais em 2025. No mercado automotivo, o peso das locadoras ficou evidente, sendo elas responsáveis por quase um quarto de todos os emplacamentos do país. Entre as marcas, Fiat, Volkswagen e Chevrolet continuam sendo as favoritas do setor, que também começa a abrir espaço para os modelos elétricos da BYD. Por fim, a preferência das empresas por modelos *Hatch* e SUV confirmou que as locadoras alinham suas frotas diretamente com o gosto e a demanda de consumo da população brasileira.

Como sugestão para trabalhos futuros, recomenda-se a expansão do escopo geográfico da pesquisa, realizando a extração e o tratamento dos dados das demais regiões do país, visto que este estudo delimitou-se aos panoramas nacional e da região Nordeste. Além disso, no âmbito do desenvolvimento técnico, aponta-se como oportunidade a busca por alternativas de otimização para a renderização do mapa coroplético do *dashboard*. Atualmente configurado

como uma imagem estática devido ao peso do processamento de arquivos GeoJSON, encontrar caminhos para tornar essa visualização interativa, sem comprometer a fluidez e a usabilidade da plataforma, enriquecerá significativamente a experiência de uso.

Adicionalmente, ressalta-se a importância de manter a base de dados atualizada de forma contínua através do acompanhamento dos novos anuários publicados anualmente pela ABLA. Para além dessa atualização periódica, futuras pesquisas podem enriquecer o *dataset* ao buscar e integrar outras fontes de dados de confiança sobre o setor de locação, tanto em âmbito nacional quanto internacional. Essa expansão de fontes permitirá traçar comparativos globais mais robustos, consolidando análises ainda mais profundas sobre o comportamento e a evolução do mercado.

REFERÊNCIAS

- ASSOCIAÇÃO BRASILEIRA DAS LOCADORAS DE AUTOMÓVEIS (ABLA). **Anuário Brasileiro do Setor de Locação de Veículos 2026**. São Paulo: ABLA, 2026. Disponível em: <<https://www.virapagina.com.br/abla2026/>>. Acesso em 23 abr. 2026.
- ARAÚJO, Gabriel Tonini de. **Elaboração de dashboards para análises de Big Data como vantagem competitiva para o planejamento estratégico em uma organização**. Trabalho de Conclusão de Curso. Universidade Federal de Ouro Preto - UFOP, Ouro Preto – MG, 2019.
- BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 1977. 229 p.
- CAIRO, Alberto. Por que a visualização de dados falha se não pensamos sobre seu propósito. **LUMINA**, Miami, v. 13, n. 2, p. 117-125, 2019. DOI: doi.org/10.34019/1981-4070.2019.v13.27762
- CAMM, Jeffrey D.; COCHRAN, James J.; FRY, Michael J.; et al. **Visualização de Dados: Explorando e Explicando com Dados**. Porto Alegre: Bookman, 2026. 382 p.
- CARVALHO, André C. P. L. F de; MENEZES, Angelo G.; BONIDIA, Robson P. **Ciência de Dados - Fundamentos e Aplicações**. Rio de Janeiro: LTC, 2024. 350 p.
- CARVALHO, Rafael de C.; MELO, Claudia de O.. Tomada de decisão baseada em dados: avaliando a visualização de informação em dashboards. In: WORKSHOP DE INICIAÇÃO CIENTÍFICA EM SISTEMAS DE INFORMAÇÃO - SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO (SBSI), 14. , 2018, Caxias do Sul. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2018. p. 24-27. ISSN 3086-545X.
- DOLLE, Camilla; *et al.* Setorial Locação de Veículos. **XP RESEARCH**, 2025. Disponível: <<https://conteudos.xpi.com.br/wp-content/uploads/2025/05/Setorial-Locacao-de-Veiculos-XP-Research-Renda-Fixa-2.pdf>>. Acesso em 04 abr. 2026.
- ESTEVE, Alain Bruhn; *et al.* **Tendências de consumo e mobilidade: um estudo para o desenvolvimento de um novo negócio em locação de veículos**. Trabalho de Conclusão de Curso. Fundação Dom Cabral - FDC, Recife - PE, 2020.
- EVELYN, Francine. O papel da ABLA no fortalecimento das locadoras de veículos. **ABLA**, São Paulo, 13 de mar. de 2026. Disponível: <<https://www.abla.com.br/artigo/o-papel-da-abla-no-fortalecimento-das-locadoras-de-veiculos>>. Acesso em 14 abr. 2026.
- GULLO, Maria Carolina R. A economia na pandemia covid-19: algumas considerações. **Rosa dos Ventos –Turismo e Hospitalidade**, v. 12, n. 3, p. 1-8, 2020. DOI: doi.org/10.18226/21789061.v12i3a05.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Indicadores Econômicos do Brasil 2025**. Rio de Janeiro: IBGE, 2025. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv102279_informativo.pdf>. Acesso em 07 jun. 2026.

JAIN, Arpit. A Decision Dashboards for Strategic Marketing. **International Journal of Research Publications in Engineering, Technology and Management - IJRPETM**, v. 4, n. 6, p. 5888-5896, 2021. DOI: doi.org/10.15662/IJRPETM.2021.0406004.

JEON, Seongmin; *et al.* Business Analytics in Action: The Case of Asset Pricing for a Rent-a-Car Company. **International Conference on Electronic Commerce**, New York, NY, USA, n. 9, p. 1-6, 2015. DOI: doi.org/10.1145/2781562.2781572.

LEAL, Guilherme Ehros. Transformação operacional no setor automotivo: metodologias de gestão de repasses e caminhões baseadas em dados e eficiência comercial. **International Integralize Scientific**, v. 5, n. 54, 2025. DOI: doi.org/10.63391/zma99761.

LUIZ, Simone Macedo Paulista. **Análise da performance empresarial das empresas do setor de locação de veículos listadas na BM&FBOVESPA**. Trabalho de Conclusão de Curso. Universidade Federal de Minas Gerais - UFMG, Belo Horizonte - MG, 2018.

MILHOME, Davi Martins. **Análise da concentração do mercado de locação de automóveis no período entre 2014 e 2021**. Trabalho de Conclusão de Curso. Universidade Federal do Ceará - UFC, Fortaleza - CE, 2022.

MORETTIN, Pedro A.; SINGER, Julio M. **Introdução à Ciência de Dados: Fundamentos e Aplicações**. São Paulo: Departamento de Estatística, 2020.

OLIVEIRA, Sérgio Levi Souza de; LOUREIRO, João Carlos de Freitas; PESSÔA, Luiggi Cavalcanti. A atuação do veículo híbrido no mercado brasileiro: análise das oportunidades de inserção no mercado automotivo. **Revista Brasileira de Gestão Ambiental e Sustentabilidade**, v. 9, n. 21, p. 317-328, 2022. DOI: doi.org/10.21438/rbgas(2022)092120.

PAPPULA, Kiran Kumar; RUSUM, Guru Pramod. Multi-Modal AI for Structured Data Extraction from Documents. **International Journal of Emerging Research in Engineering and Technology**, USA, v. 4, n. 3, p. 75-86, 2023. DOI: doi.org/10.63282/3050-922X.IJERET-V4I3P109.

PUCCI, Olivo. Anuário ABLA 2026: frota do setor de locação ultrapassa 1,7 milhão de automóveis. 2026. **ABLA**, São Paulo, 13 de mar. de 2026. Disponível em <<https://www.abla.com.br/noticia/anuario-abla-2026-frota-do-setor-de-locacao-ultrapassa-17-milhao-de-automoveis->>. Acesso em: 21 de mar. de 2026.

QUADROS, Janeslei Pereira Vaz de; SILVEIRA, Rosemari Monteiro Castilho Fogiatto. Método de análise de conteúdo de bardin na pesquisa educacional em ciência, tecnologia e sociedade. **ARACÊ**, [S. l.], v. 7, n. 3, p. 13849-13861, 2025. DOI: doi.org/10.56238/arev7n3-215.

SANTOS, Claudinéia Correia dos. Tomada de decisão orientada por dados estruturados. **Revista Tópicos**, v. 2, n. 3, 2024. DOI: doi.org/10.5281/zenodo.10720288.

SILVA, Fabiano Couto Corrêa da. Visualização de dados: passado, presente e futuro. **LIINC**, Rio de Janeiro, v.15, n.2, p. 205-223, 2019.

STRENGHOLT, Piethein. **Building Medallion Architectures**: Designing with Delta Lake and Spark. Santa Rosa, CA: O'Reilly Media, 2025. 396 p.

TRECENTI, Julio Adolfo Zucon. **Resolvendo Captchas**: usando raspagem de dados e aprendizado fracamente supervisionado. Tese de Doutorado. Universidade de São Paulo - USP, São Paulo - SP, 2023.

WAGNER, Sebastian; BRANDT, Tobias; NEUMANN, Dirk. Data Analytics in Free-Floating Carsharing: Evidence from the City of Berlin. **Hawaii International Conference on System Sciences**, Kauai, HI, EUA, p. 897-907, 2015. DOI: 10.1109/HICSS.2015.112.

YANG, Wen; CAO, Feifei; ZHAO, Xueli. Extraction of PDF Table Data Based on the Pdfplumber Method. In: Proceedings of the 2024 4th International Joint Conference on Robotics and Artificial Intelligence (JCRAI '24). **Association for Computing Machinery**, New York, NY, USA, p. 103–106, 2025. DOI: doi.org/10.1145/3696474.3696731