



VAREJO: MODELO DE ANÁLISE E PREDIÇÃO

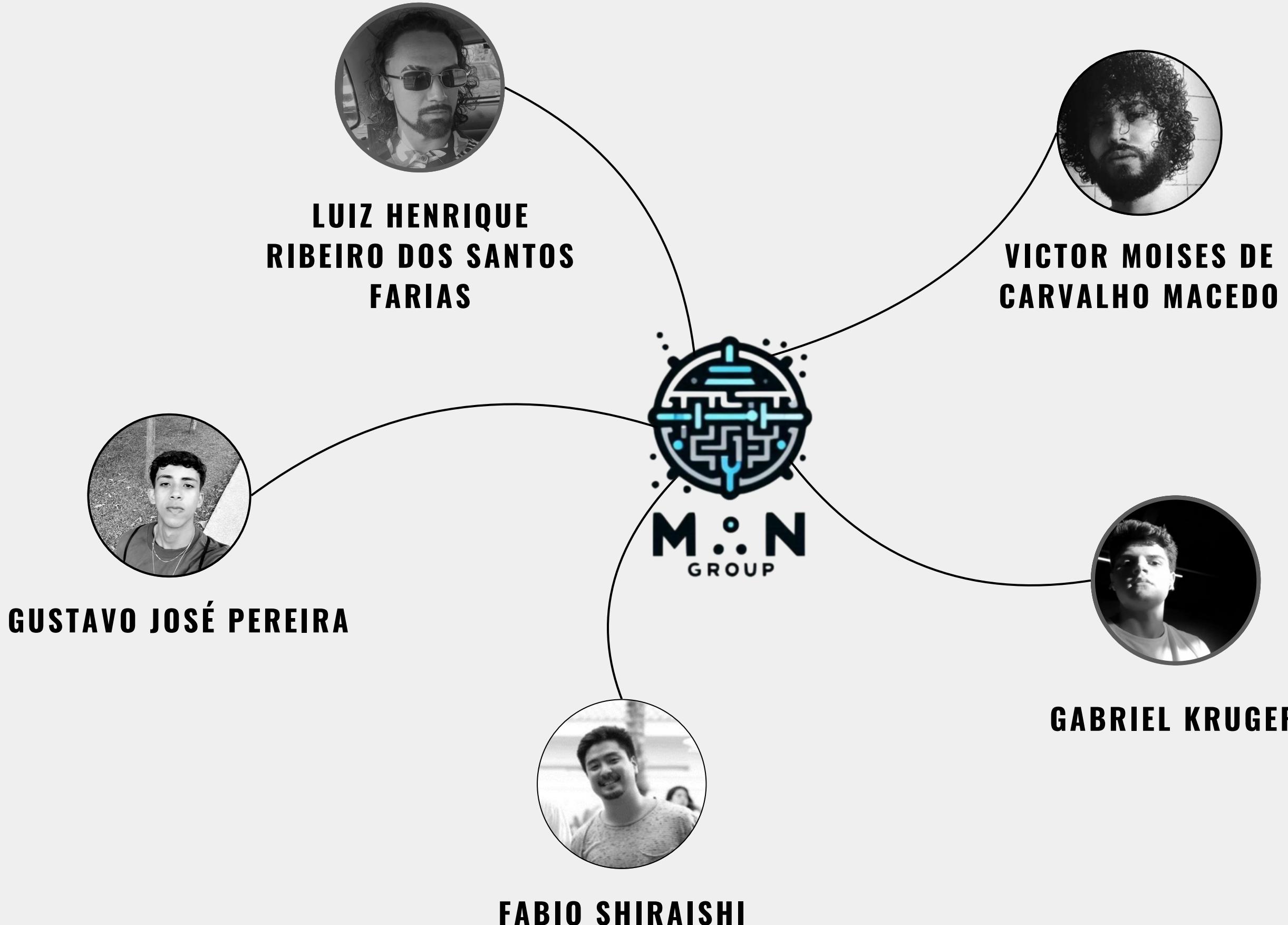
COMPONENTES



TURMA : 1TSCOR

LUIZ HENRIQUE RIBEIRO DOS SANTOS FARIAS	RM559962
GUSTAVO JOSÉ PEREIRA	RM561067
FABIO SHIRASHI	RM559725
GABRIEL KRUGER	RM559493
VICTOR MOISES DE CARVALHO MACEDO	RM559904

APRESENTAÇÃO



CONTEÚDO

- 01** PROBLEMA A SER RESOLVIDO
- 02** PÚBLICO ALVO
- 03** PROPOSTA DE SOLUÇÃO, ESG E LGPD
- 04** DIAGRAMA DA ARQUITETURA DA SOLUÇÃO
- 05** DESCRIÇÃO DA ARQUITETURA
- 06** TECNOLOGIAS USADAS
- 07** PROTÓTIPOS DE SOLUÇÃO
- 08** ANÁLISE EXPLORATÓRIA
- 09** DASHBOARD DA ANÁLISE EXPLORATÓRIA
- 10** FINALIZAÇÃO E AGRADECIMENTOS

PROBLEMA A SER RESOLVIDO - MUDANÇA DE HÁBITO DO CONSUMIDOR

Oliveira (2023) aponta que, recentemente, o mercado de varejo mudou rapidamente, impulsionado pela transformação digital, a pandemia e a busca por sustentabilidade e personalização.

91%



dos consumidores têm maior probabilidade de comprar de marcas que se lembram de suas preferências e oferecem recomendações relevantes.
(MARINHO, 2024)

PROBLEMA A SER RESOLVIDO - MUDANÇA DE HÁBITO DO CONSUMIDOR

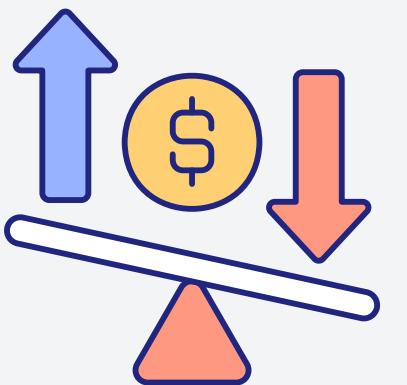
Como a mudança de hábito é constante, as empresas precisam se atualizar sempre.



Melhor do que acompanhar o hábito de consumo, é prever ele.

40%

das empresas que não inovam ficam estagnadas ou perdem terreno em relação ao mercado (ROESLER, 2023)

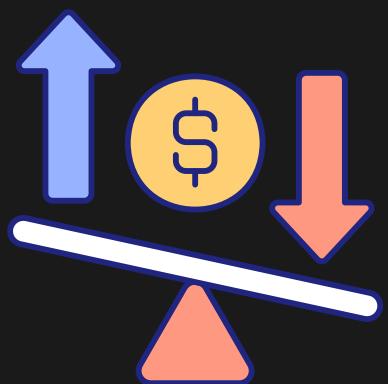


PROBLEMA A SER RESOLVIDO - MUDANÇA DE HÁBITO DO CONSUMIDOR

Segundo uma pesquisa feita pelo IBM Institute for Business Value, os usuários buscam mais por produtos sustentáveis e estão dispostos a pagar mais

62%

das pessoas estão dispostas a apagar a mais por produtos sustentáveis. (G1)



PÚBLICO ALVO



Busca por produtos mais personalizados, bem avaliados, com qualidade e com base em compras anteriores.

CONSUMIDOR EXIGENTE



Cientes em busca de uma entrega confiável e eficiente.
Busca por produtos com prazo de entrega rápido e barato.

CONSUMIDOR ÁGIL



Busca por produtos sustentáveis, conhece o mercado e faz comparativos entre marcas

ECOMMERCE



Buscamos entregar dados qualificados para que a empresa tome decisões estratégicas e assertivas.

ECOMMERCE

PROPOSTA DE SOLUÇÃO

Com base no contexto, estamos propondo uma solução de B.I., onde será disponibilizado um dashboard amigável e interativo, no qual os funcionários terão uma base para tomadas de decisão estratégicas.

Orientado ao dashboard podemos responder perguntas sobre como otimizar processos de entrega, promover os melhores produtos e vendedores com base em indicadores, e também, entender o comportamento do consumidor em relação a quando compra, o que compra, e quais são suas opiniões sobre os produtos ofertados.

A partir da necessidade iremos utilizar processamento de linguagem natural (NPL) para analisar a opinião do usuário em relação aos produtos.

Após coletar, tratar, analisar e processar os dados utilizaremos técnicas de Machine Learning para prever o comportamento de compra e assim ter uma melhor gestão de estoque.



ESG E LGPD

Ao realizar análises preditivas, as empresas, de acordo com o dashboard, podem evitar estoque excessivo. Sendo assim, pensando na **sustentabilidade**, diminui-se os descartes de produtos vencidos ou não vendidos, além de diminuir, dependendo da situação, o espaço físico que esses produtos ocupariam.

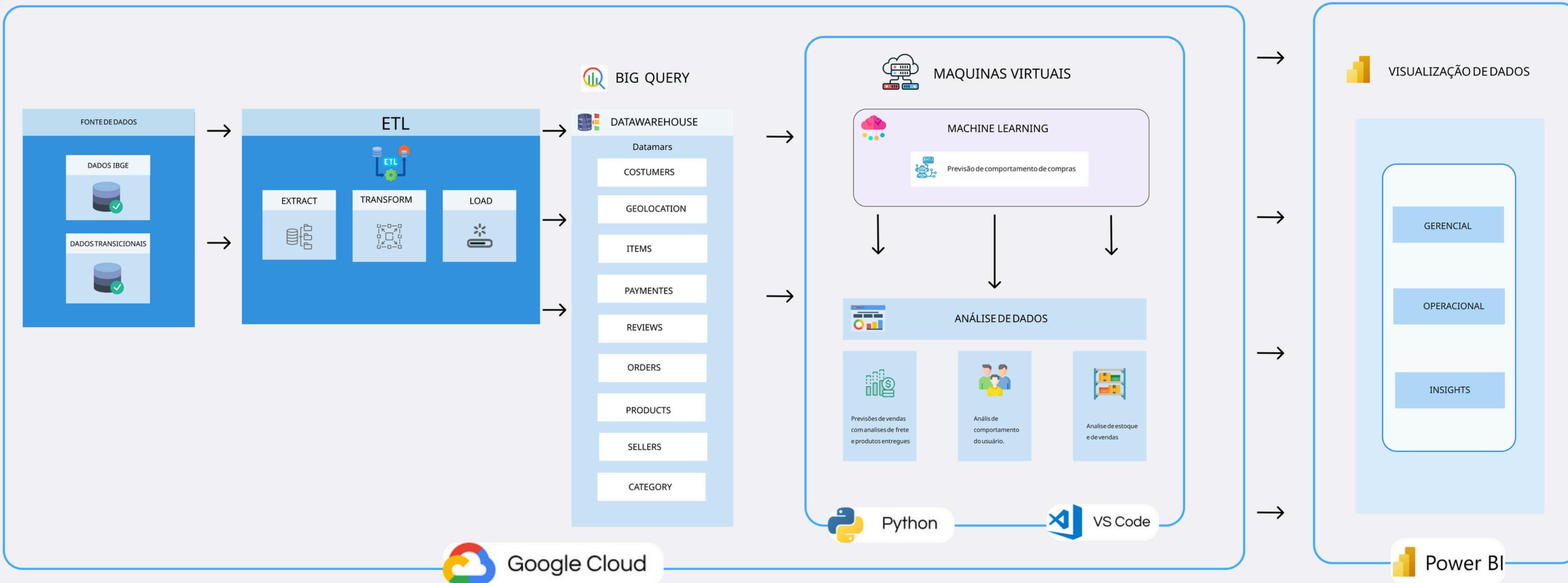
Ao integrar dados sobre preferências dos consumidores em uma determinada região, o mercado do varejo tem a possibilidade de desenvolver produtos e serviços alinhados às necessidades daquela comunidade, algo indispensável para a **sociedade**.

A partir do momento em que as empresas tomam melhores decisões baseado nos dados disponíveis, as corporações tendem a diminuir os riscos e reduzir os processos administrativos. Portanto, visando o longo prazo, o projeto, com todos os recursos que possui, melhora a **governança** das companhias.

Além disso, estamos preocupados com o bem-estar de nossos clientes e a qualidade dos dados que oferecemos. Por isso, trabalhamos com dados anonimizados, como rege a **Lei Geral de Proteção de Dados**



DIAGRAMA DA ARQUITETURA DA SOLUÇÃO



DESCRIÇÃO DA ARQUITETURA

Etapa 1: Extração de Dados

Os dados são extraídos de fontes diversas como IBGE e dados transacionais, convertidos para formato CSV para facilitar o processamento.

Etapa 2: Processo de ETL

Os arquivos CSV são importados para a Nuvem da Google Cloud, onde ocorre o processo de ETL: extração, limpeza, transformação e carga dos dados para o Data Warehouse.

Etapa 3: Data Warehouse e Data Marts

Os dados transformados são armazenados no Data Warehouse (BigQuery). São criados Data Marts específicos para diferentes áreas como vendas, marketing e finanças.

Etapa 4: Machine Learning

Máquinas virtuais e técnicas de Machine Learning são usadas para prever o comportamento dos usuários, analisando tendências e realizando processamento de linguagem natural (NLP) para análise de sentimento.

Etapa 5: Análise e Visualização de Dados

Os dados processados são visualizados e analisados no Power BI, permitindo a criação de dashboards interativos e suporte à tomada de decisões.

Resumo da Arquitetura

1. Extração de Dados: Coleta de dados (IBGE, transacionais), convertidos para CSV.
2. ETL: Importação, limpeza, transformação e carga na Google Cloud.
3. Data Warehouse e Data Marts: Armazenamento no BigQuery, criação de Data Marts.
4. Machine Learning: Previsão de comportamentos e análise preditiva.
5. Visualização: Análise e visualização no Power BI.

Essa arquitetura assegura a gestão eficiente dos dados, desde a extração até a visualização, facilitando a análise e a tomada de decisões baseadas em dados precisos e bem estruturados.



TECNOLOGIAS USADAS

TRELLO

Gerenciamento de projetos e implantação.



TRELLO

GOOGLE CLOUD PLATFORM

BigQuery
Armazenamento de dados transicionais e dados do IBGE e o processo de ETL



GOOGLE CLOUD

MISCRSOFT VISUAL STUDIO CODE

Utilizaremos o **VS Code** para criação dos scripts e modelos de dados que serão aplicados na solução.



Visual Studio Code

POWER BI

Criação de dashboards para a visualização dos indicadores.



POWER BI

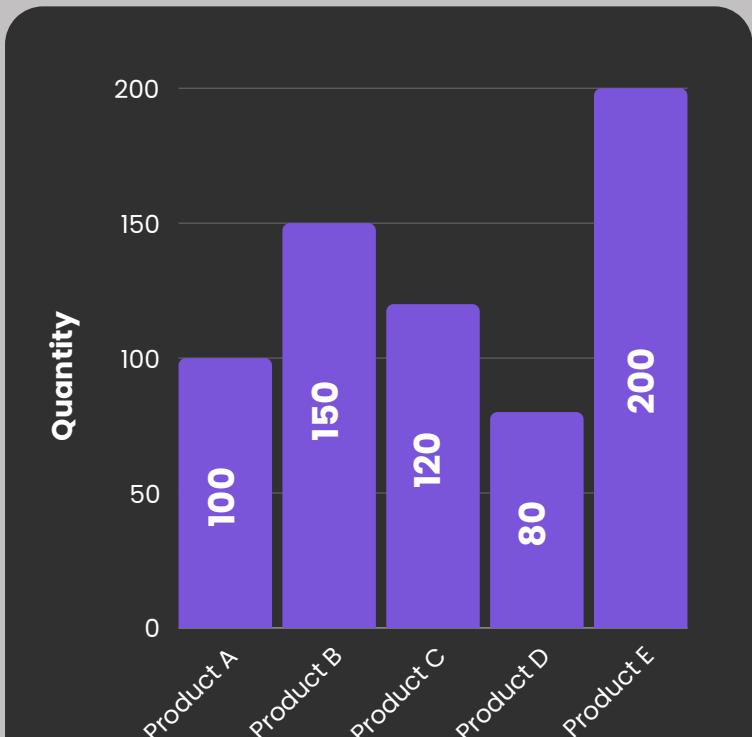
PROTÓTIPO DA SOLUÇÃO I

Planejamento de Estoque x Vendas

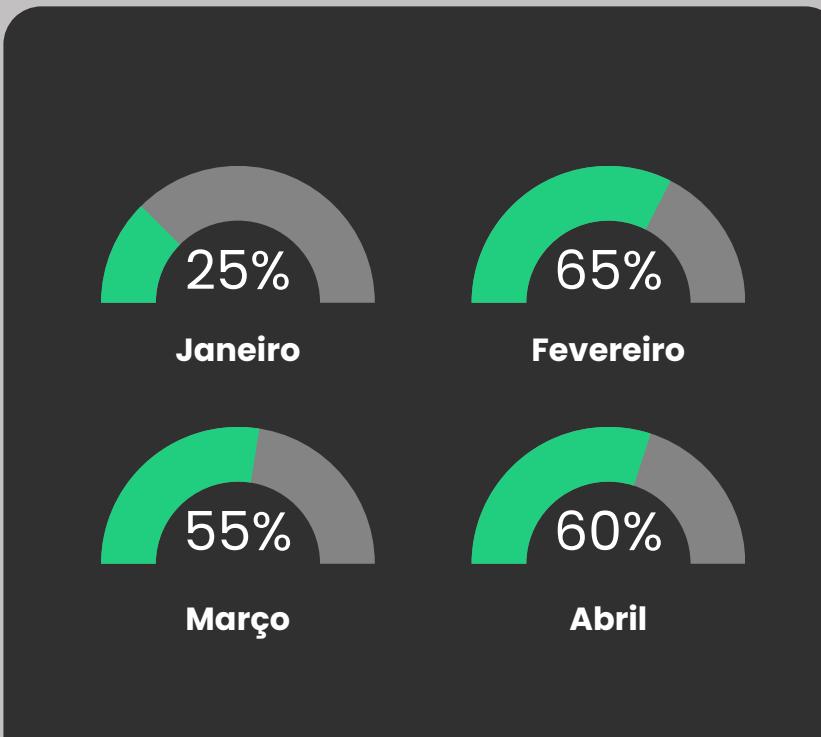
8 de dez à 14 de dez de 2024



Nível de estoque por produto



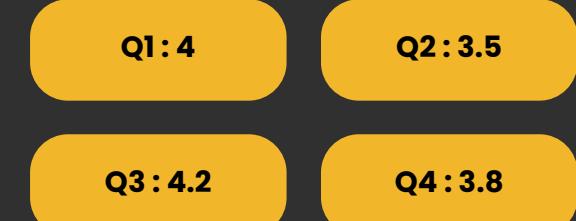
Índice de Ruptura (Últimos 4 meses)



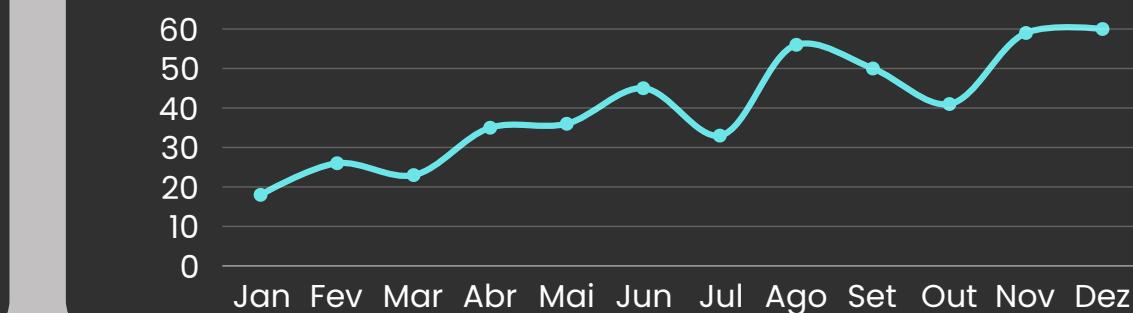
Custos de Estoque



Taxa de Rotatividade de Estoque



Análise preditiva de vendas



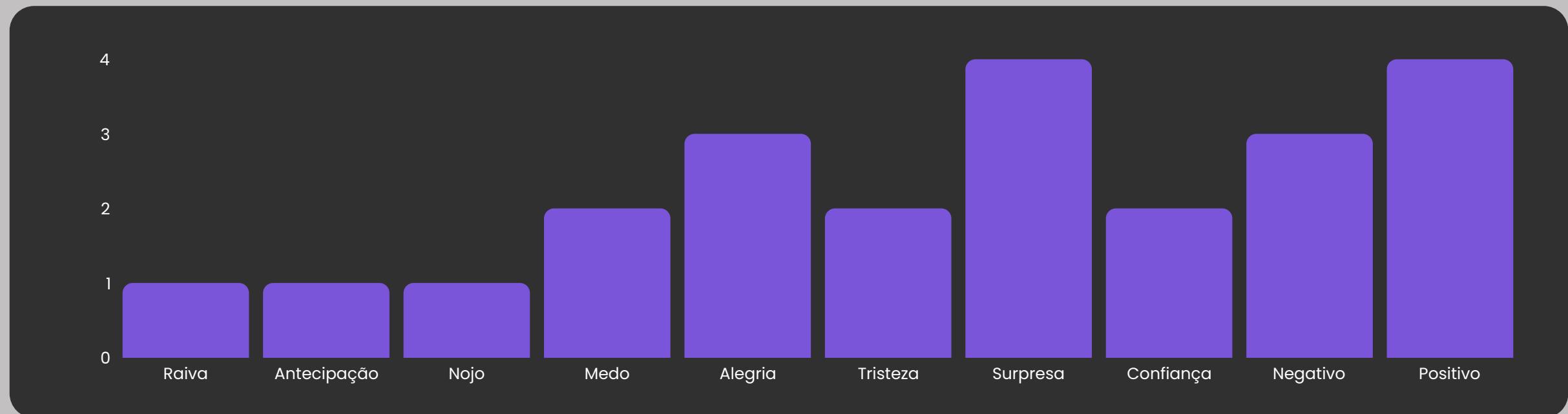
PROTÓTIPO DA SOLUÇÃO II

Análise de Sentimento e opinião dos usuários

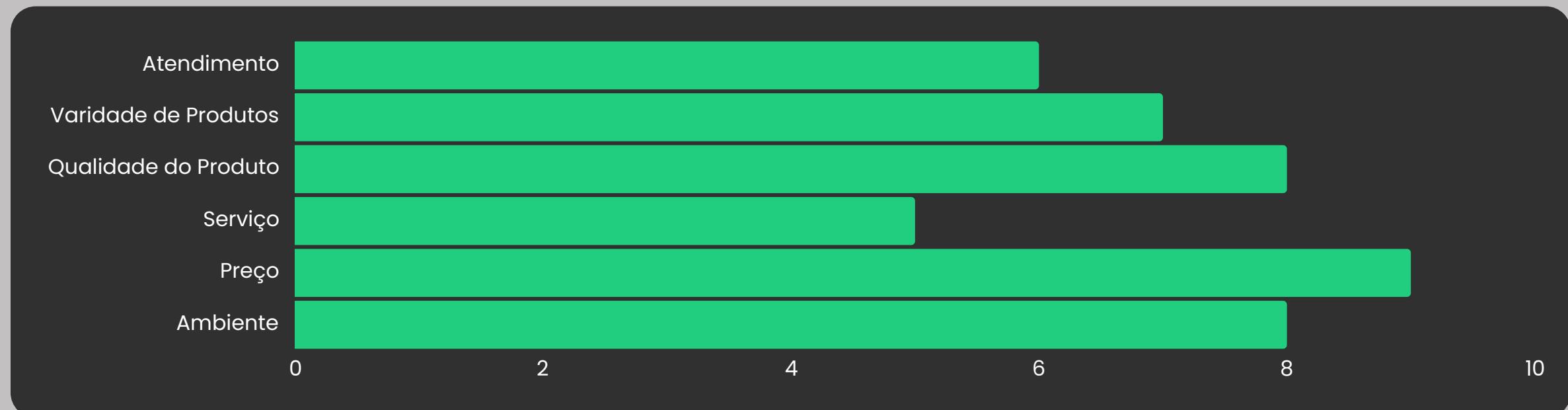
8 de dez à 14 de dez de 2024



Análise de Sentimento



Oponião dos uuários (0 a 10)



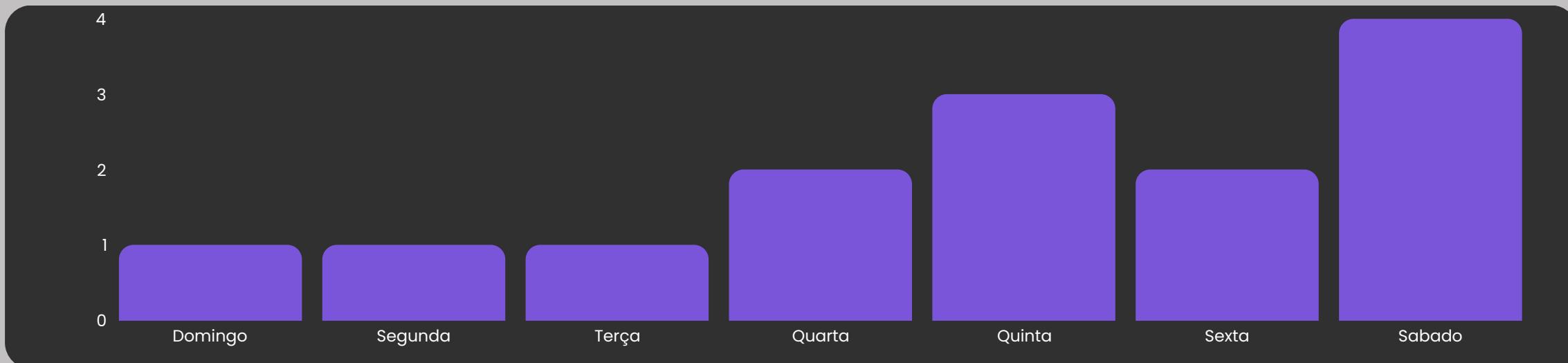
PROTÓTIPO DA SOLUÇÃO II

Análise de frete e produtos entregues

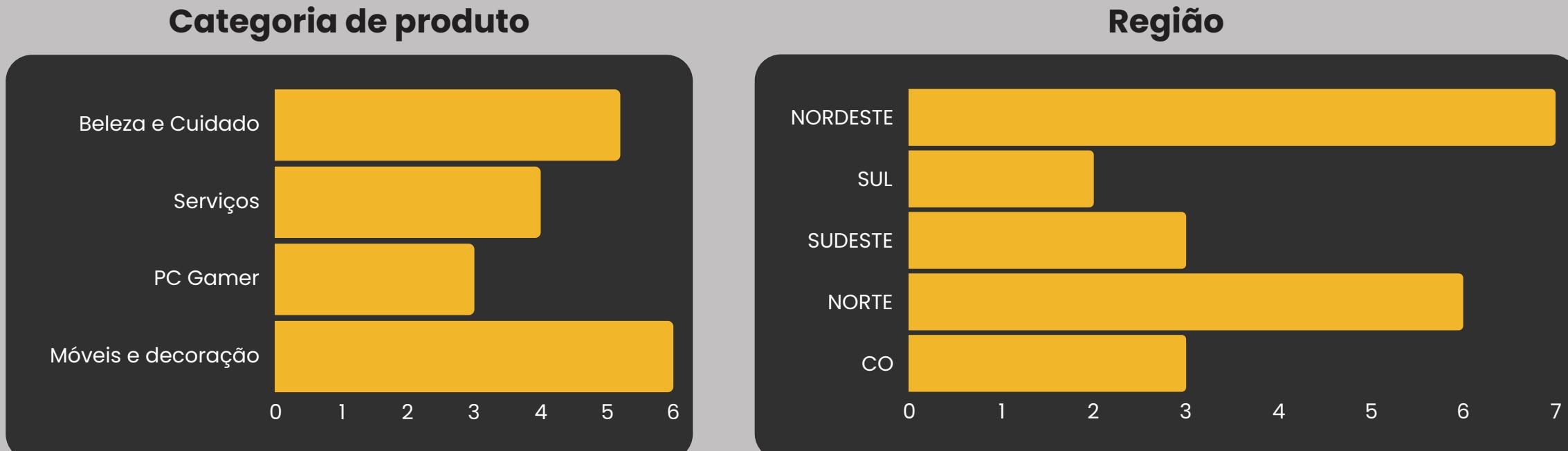
8 de dez à 14 de dez de 2024



Análise de frete



Análise de produtos entregues (média de dias para entrega)



PLANEJAMENTO E GESTÃO DO PROJETO



LINK: FRAMEWORK ÁGIL

Trello Áreas de trabalho Recente Marcado como favorito Templates Criar Pesquisar Atlassian uses cookies to improve your browsing experience, perform analytics and research, and conduct advertising. Accept all cookies to indicate that you agree to our use of cookies on your device. [Atlassian cookies and tracking notice](#) Preferences Only necessary ✓ Accept all

Framework Ágil Visível à Área de trabalho Quadro Power-up do Calendário Power-Ups Automação Filtros Compartilhar ...

Legenda de Sprints

- Sprint 1 Sprint 1 (Ideação)
- Sprint 2 Sprint 2 (Arquitetura, Desenho inicial, e Análise Exploratória de Dados)
- Sprint 3 Sprint 3 - Solução para o cliente
- Sprint 4 Sprint 4 - Estratégias e Insights com cliente

Backlog

- Sprint 4 Sprint 4 - Estratégias e Insights com cliente
- Sprint 4 Definição de agenda com cliente
- Sprint 4 Análise Avançada e identificação de insights
- Sprint 4 Formulação de estratégias e recomendações
- Sprint 4 Apresentação Final

Em Progresso

+ Adicionar um cartão

Next-up

- Sprint 3 Sprint 3 - Solução para o cliente
- Sprint 3 Coleta de Dados e Mapeamento de dados do varejo
- Sprint 3 ETL - Tratamento dos dados
- Sprint 3 Análise de dados - Agrupamento de dados e indicadores
- Sprint 3 Análise Preditiva de Demandas e Vendas
- Sprint 3 Análise de Sentimento - Processamento de Linguagem Natural
- Sprint 3 Dashboard com análises

Sprint Pronta

- Sprint 2 Sprint 2 (Arquitetura, Desenho inicial, e Análise Exploratória de Dados)
- Sprint 2 Protótipos de Solução
- Sprint 2 Análise Exploratória
- Sprint 2 Gerenciamento do Projeto (framework ágil)
- Sprint 2 Aquisição e preparação de dados
- Sprint 2 Arquitetura e desenho inicial da solução

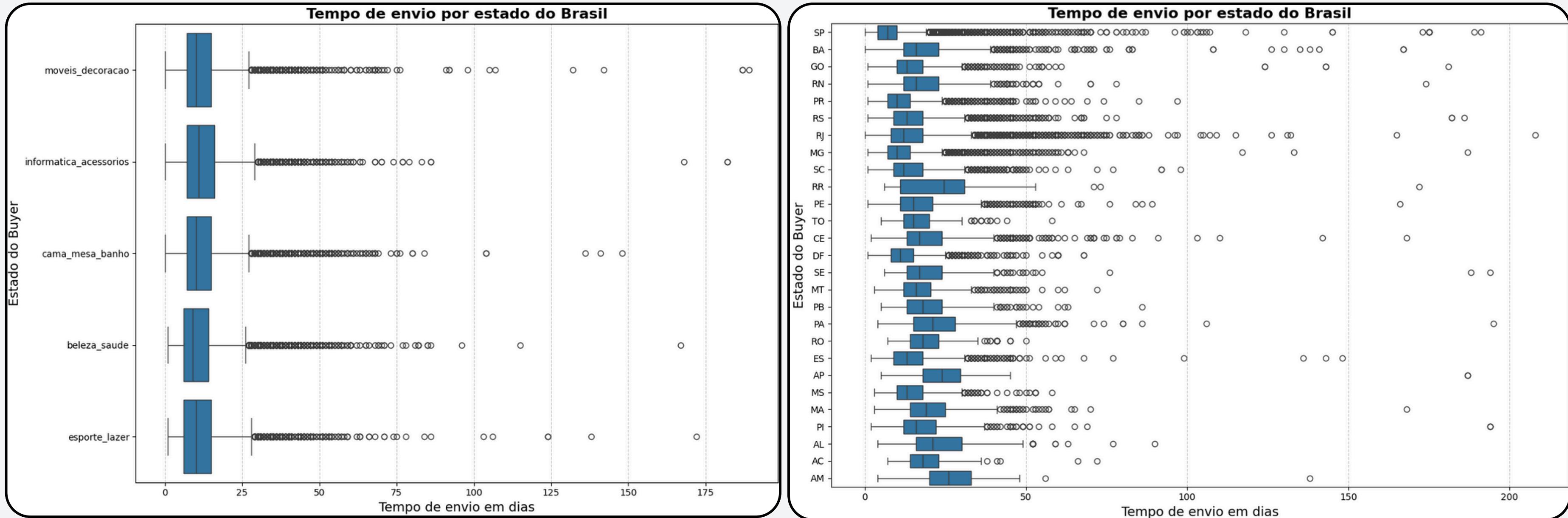
Entregue

- Sprint 1 Sprint 1 - Ideação
- Sprint 1 Nome do Projeto
- Sprint 1 Contextualização do Problema
- Sprint 1 Problema a ser Resolvido
- Sprint 1 Público Alvo
- Sprint 1 Proposta de Solução
- Sprint 1 Impacto da Solução

Agendas

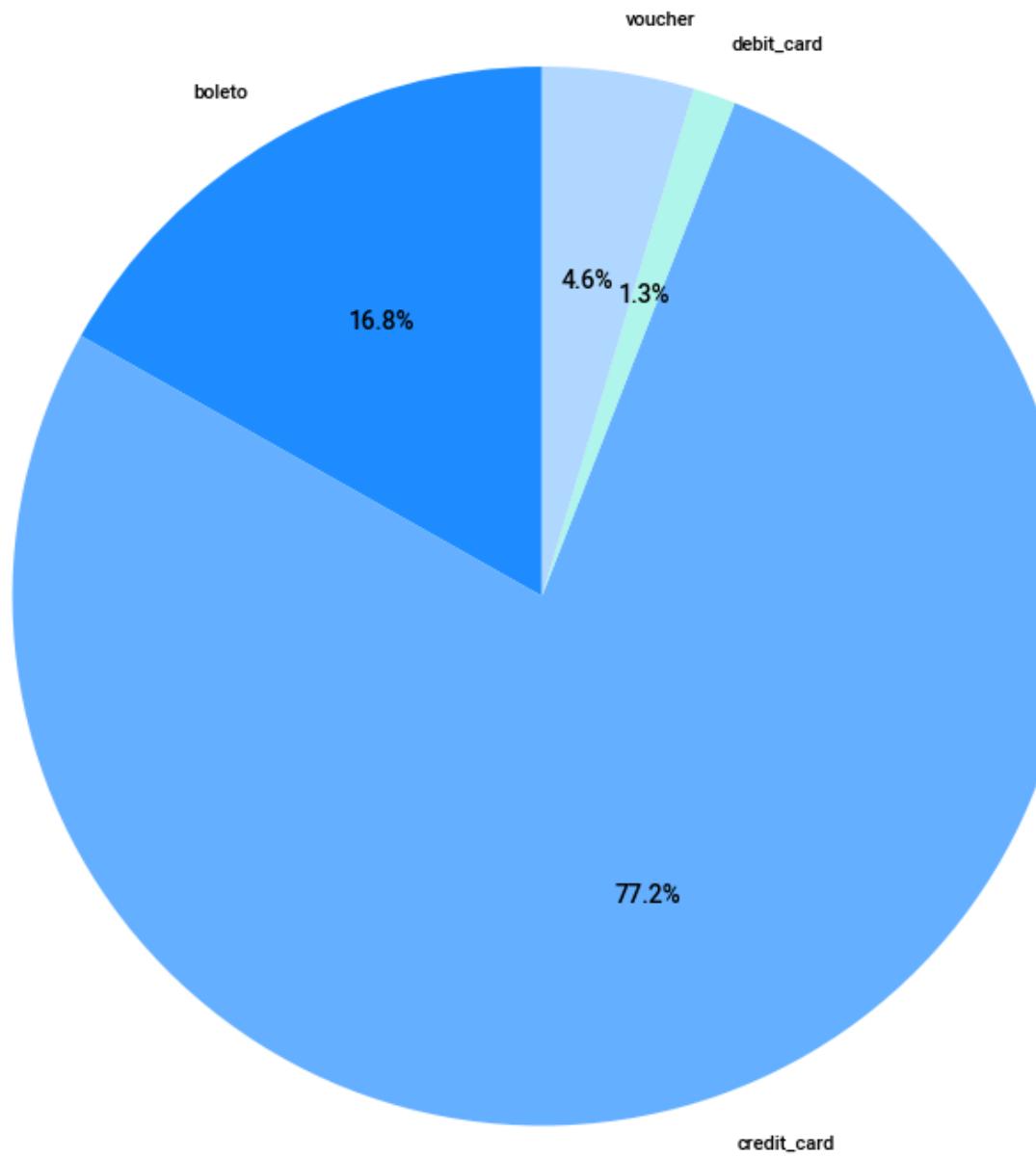
- Alinhamento Semanal - 4a
- + Adicionar um cartão

ANÁLISE EXPLORATÓRIA

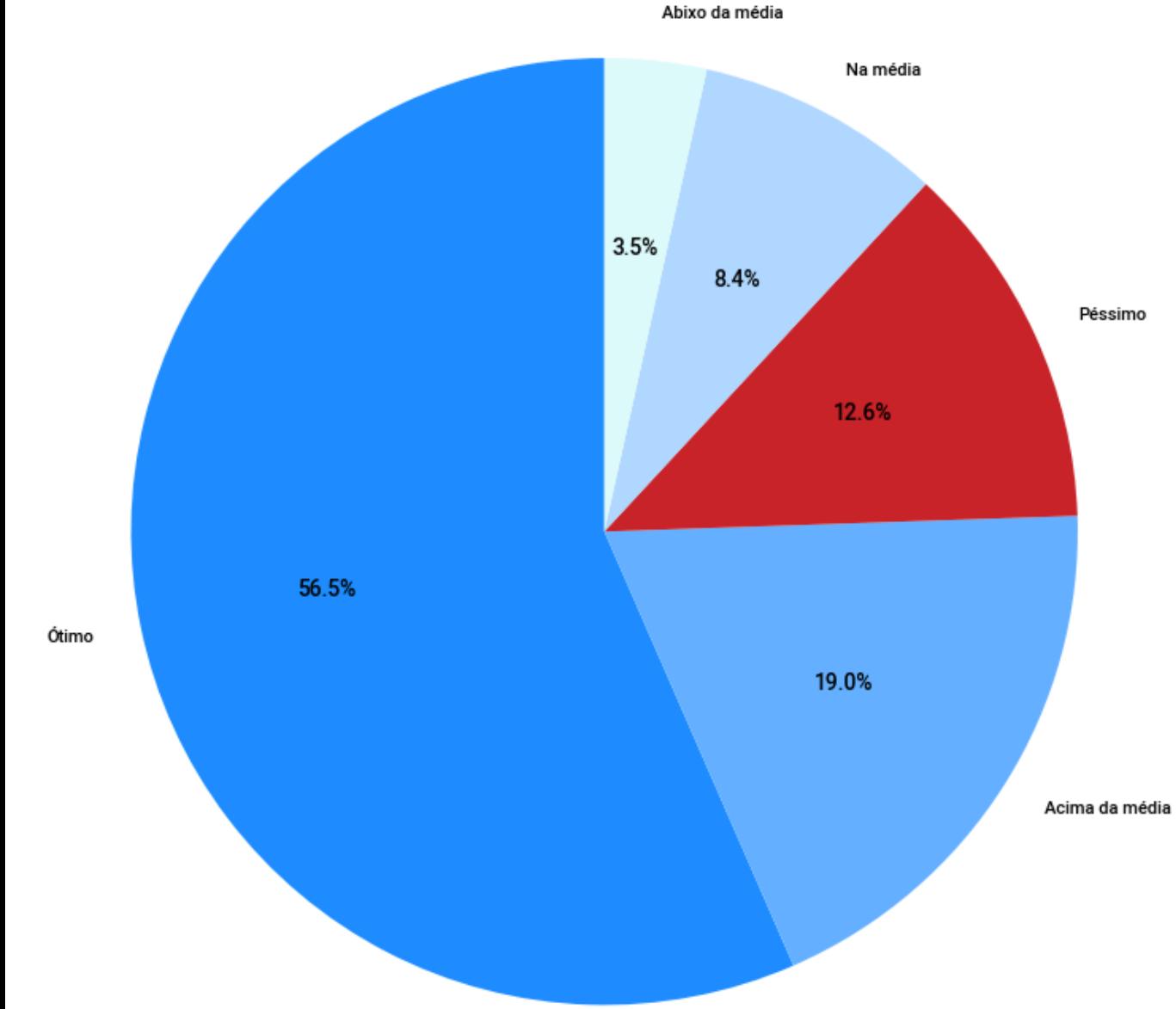


ANÁLISE EXPLORATÓRIA

Distribuição de vendas por tipo de pagamento



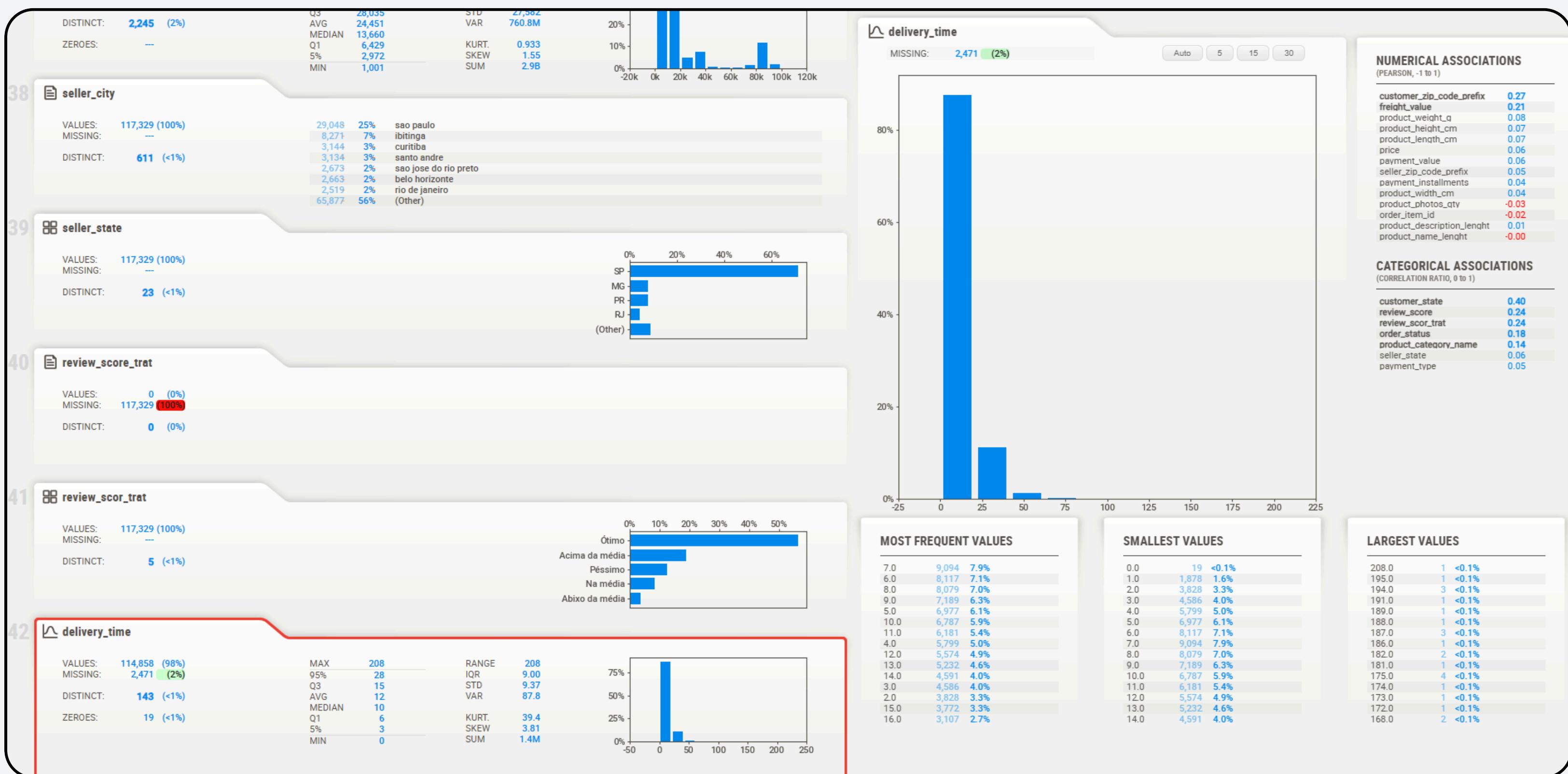
Distribuição de avaliação dos produtos





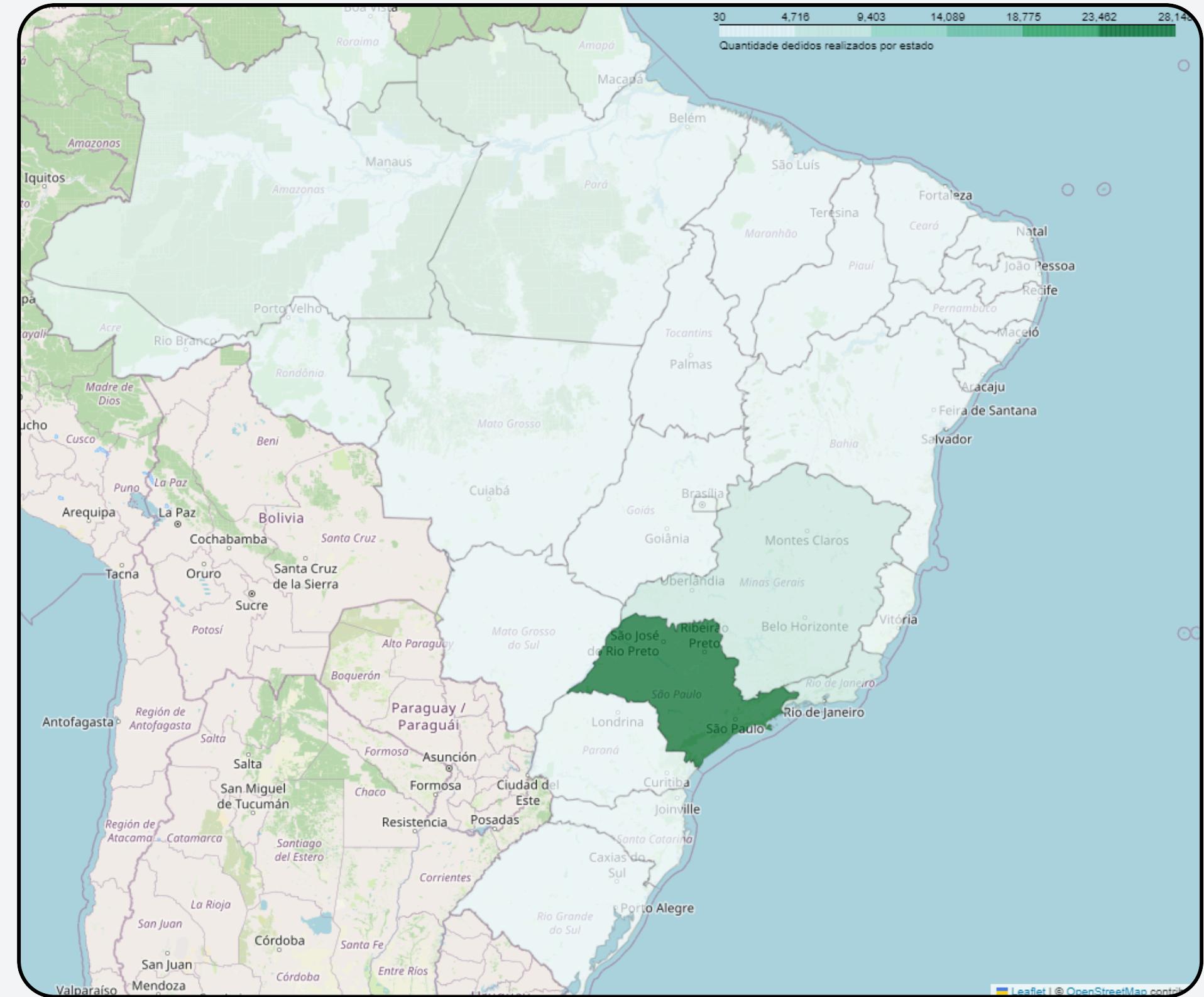
M•N
GROUP

DASHBOARD DA ANÁLISE EXPLORATÓRIA



ANÁLISE EXPLORATÓRIA

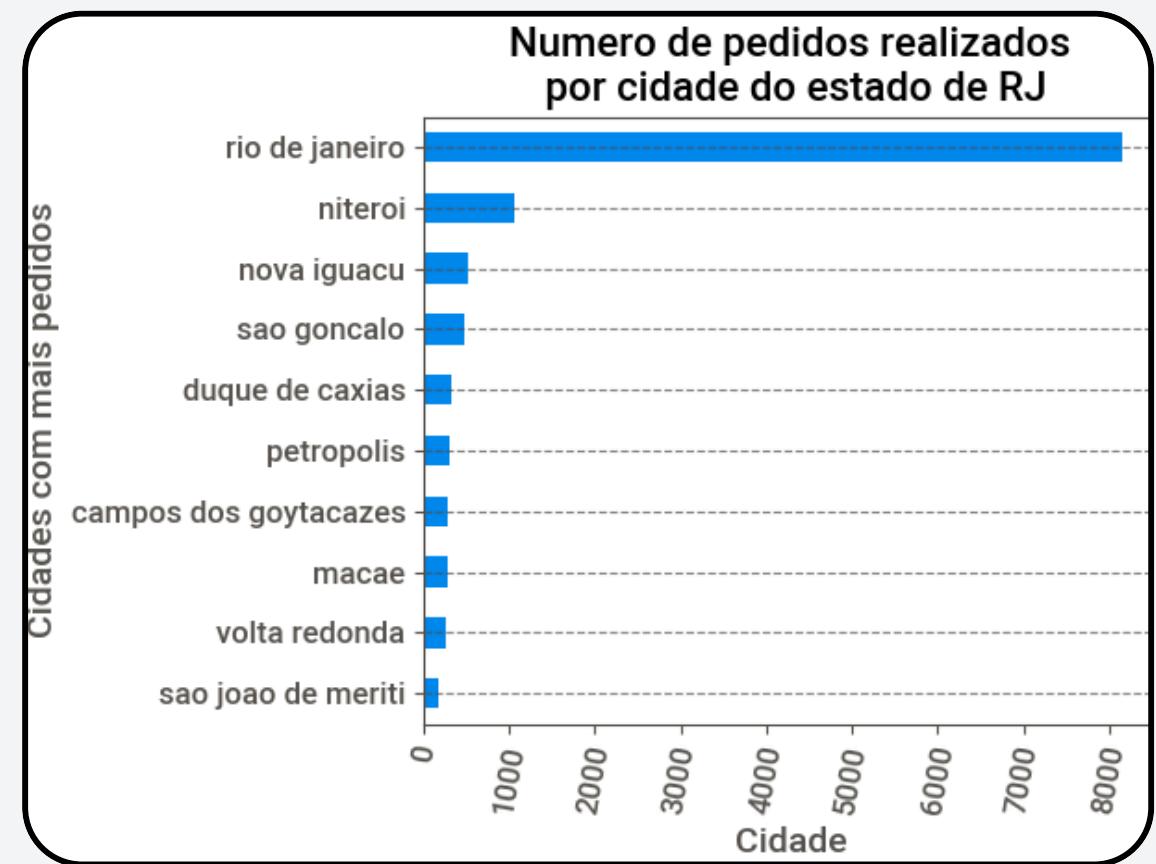
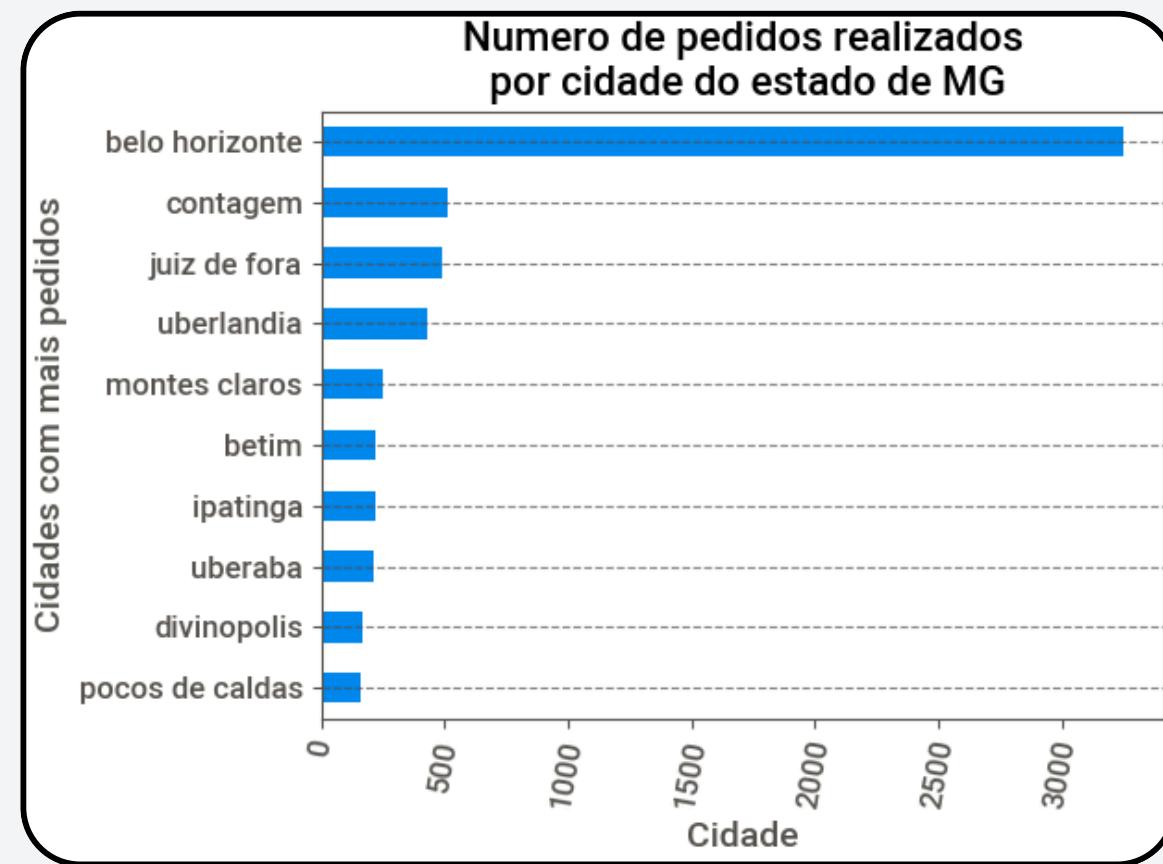
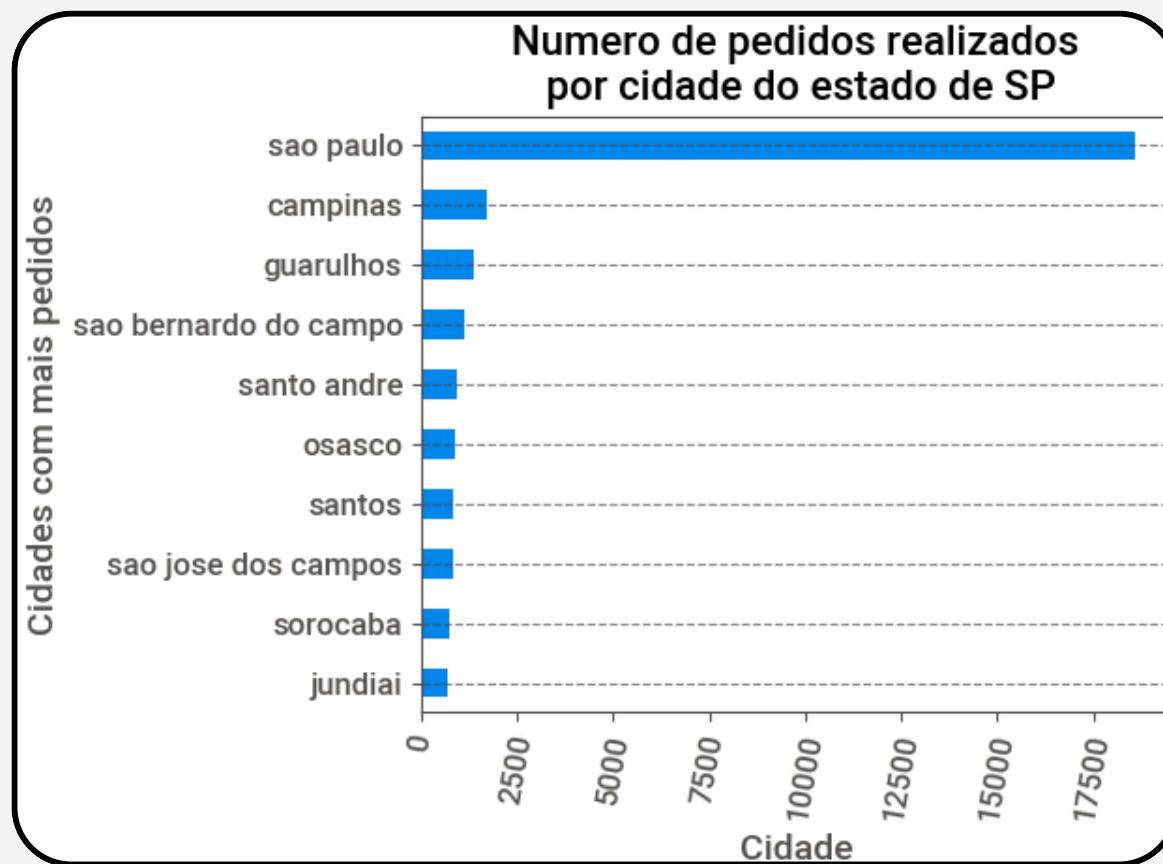
DISTRIBUIÇÃO DE PEDIDOS POR ESTADO



ANÁLISE EXPLORATÓRIA

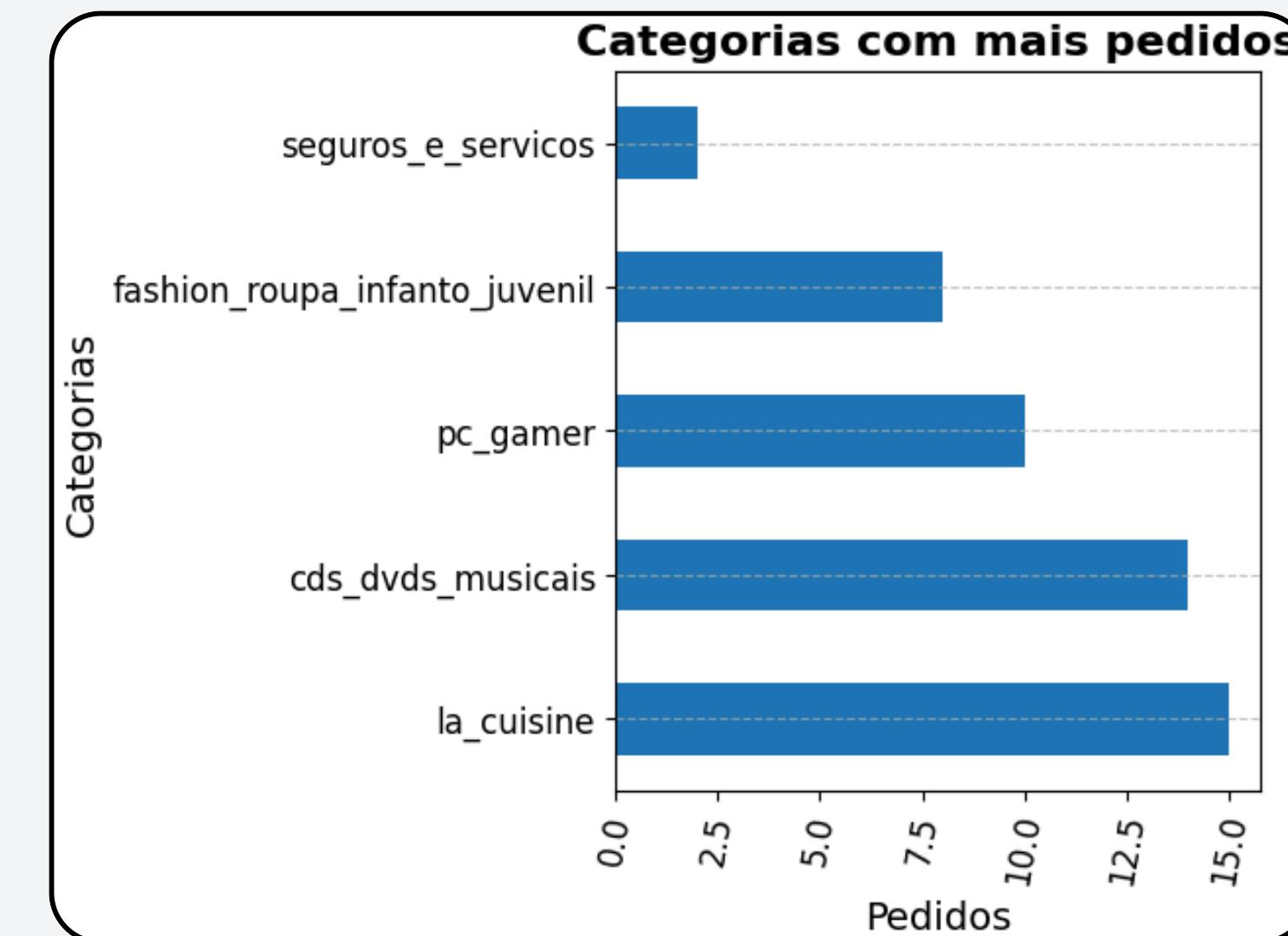
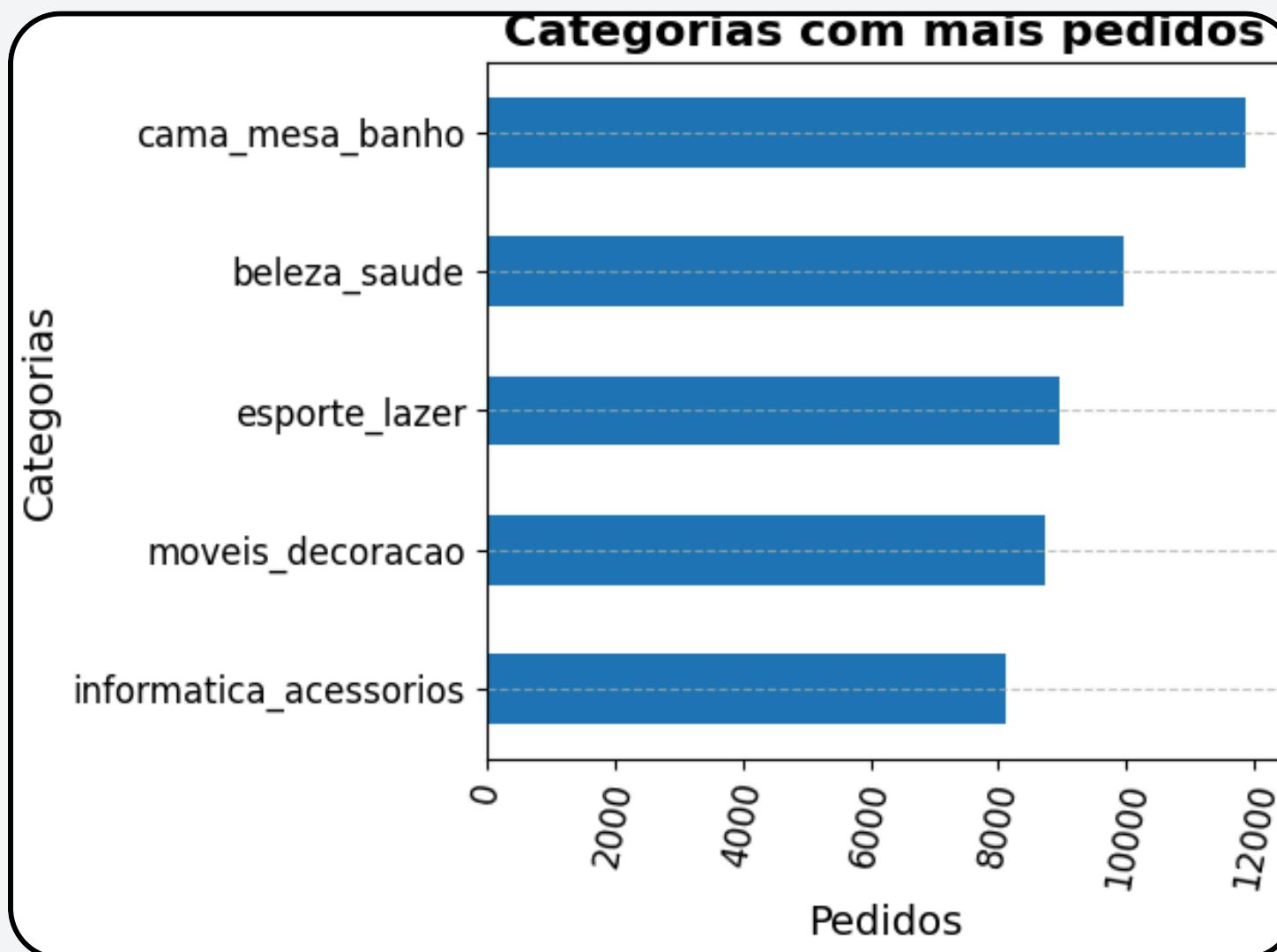
DISTRIBUIÇÃO DE PEDIDOS POR CIDADE

(ESTADOS COM MAIS PEDIDOS)



ANÁLISE EXPLORATÓRIA

DISTRIBUIÇÃO DE PEDIDOS POR CATEGORIA



DASHBOARD DA ANÁLISE EXPLORATÓRIA

38 seller_city

VALUES: 117,329 (100%)
MISSING: —
DISTINCT: 611 (<1%)

	VALUES	PERCENTAGE	CITY
29,048	25%	sao paulo	
8,271	7%	ibitinga	
3,144	3%	curitiba	
3,134	3%	santo andre	
2,673	2%	sao jose do rio preto	
2,663	2%	belo horizonte	
2,519	2%	rio de janeiro	
65,877	56%	(Other)	

39 seller_state

VALUES: 117,329 (100%)
MISSING: —
DISTINCT: 23 (<1%)

40 review_score_trat

VALUES: 0 (0%)
MISSING: 117,329 (100%)
DISTINCT: 0 (0%)

41 review_score_trat

VALUES: 117,329 (100%)
MISSING: —
DISTINCT: 5 (<1%)

42 delivery_time

VALUES: 114,858 (98%)
MISSING: 2,471 (2%)
DISTINCT: 143 (<1%)
ZEROES: 19 (<1%)

	STATISTIC	VALUE	
MAX	208		
95%	28	RANGE	208
Q3	15	IQR	9.00
AVG	12	STD	9.37
MEDIAN	10	VAR	87.8
Q1	6	KURT.	39.4
5%	3	SKEW	3.81
MIN	0	SUM	1.4M

review_scor_trat

MISSING: —

TOP CATEGORIES

CATEGORY	COUNT	PERCENTAGE
Ótimo	66,264	56%
Acima da média	22,286	19%
Péssimo	14,854	13%
Na média	9,840	8%
Abixo da média	4,085	3%

CATEGORICAL ASSOCIATIONS (UNCERTAINTY COEFFICIENT, 0 to 1)

review_scor_trat PROVIDES INFORMATION ON...

FEATURE	ASSOCIATION
review_score	1.00
order_status	0.18
customer_state	0.00
product_category_name	0.00
seller_state	0.00
payment_type	0.00

THESE FEATURES GIVE INFORMATION ON review_scor_trat:

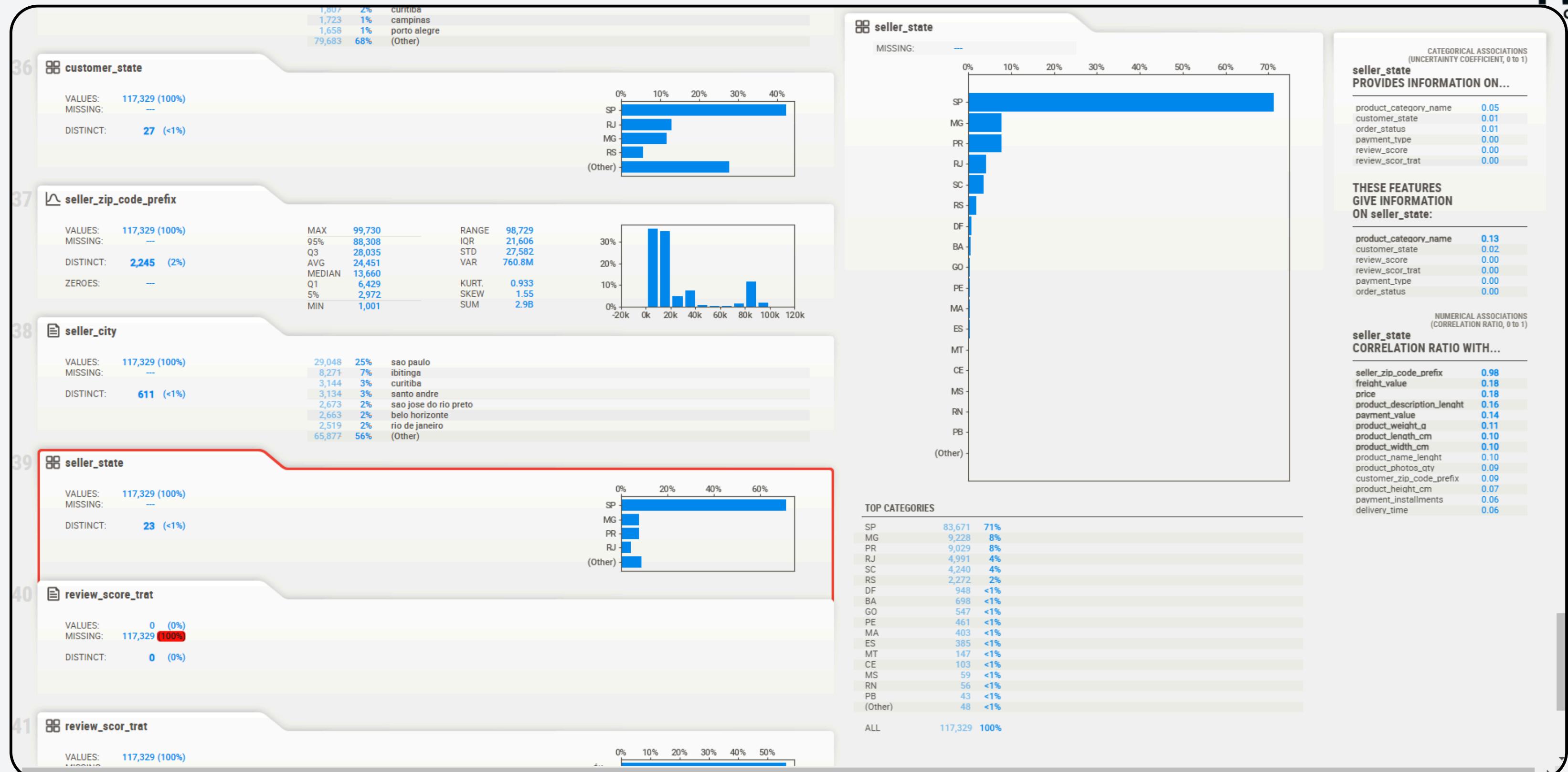
FEATURE	ASSOCIATION
review_score	1.00
order_status	0.02
product_category_name	0.01
customer_state	0.00
seller_state	0.00
payment_type	0.00

NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1)

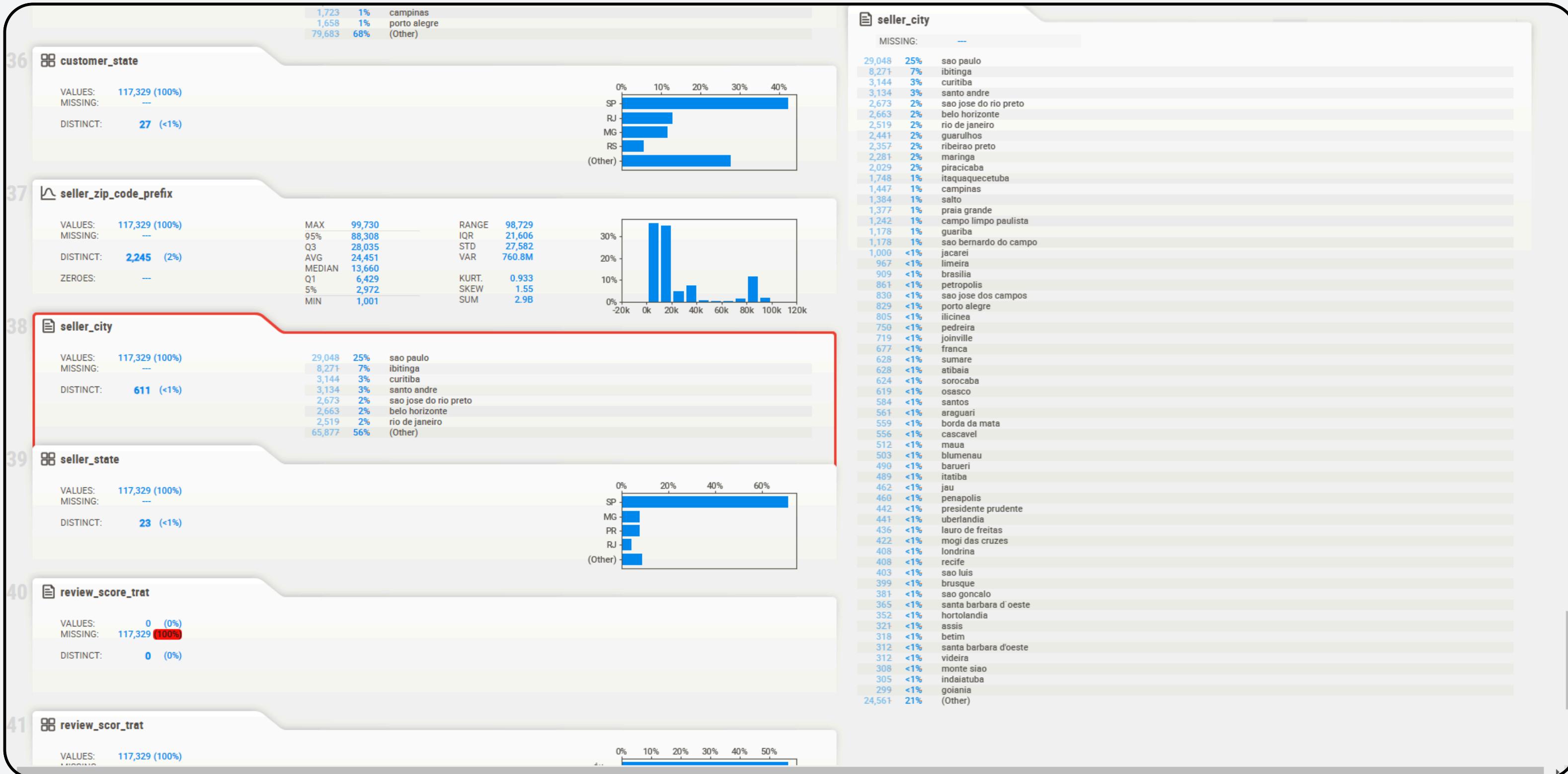
review_scor_trat CORRELATION RATIO WITH...

FEATURE	ASSOCIATION
delivery_time	0.24
order_item_id	0.13
payment_value	0.09
payment_installments	0.05
freight_value	0.04
product_photos_qty	0.03
product_weight_g	0.03
seller_zip_code_prefix	0.03
product_height_cm	0.03
customer_zip_code_prefix	0.03
product_length_cm	0.02
price	0.02
product_name_length	0.02
product_description_length	0.02

DASHBOARD DA ANÁLISE EXPLORATÓRIA



DASHBOARD DA ANÁLISE EXPLORATÓRIA





MUITO OBRIGADO :)

REFERÊNCIAS

MARINHO, Silva. **Atendimento Comum x Personalizado: O Que Seus Clientes Preferem?** Inhouse, 2024. Disponível em: <https://www.inhouse.com.br/atendimento-comum-x-personalizado-o-que-seus-clientes-preferem/> Acesso em: 11 nov. 2024.

OLIVEIRA, Josele Delazeri. **Como o varejo mudou para sempre.** E-Commerce Brasil, 2023. Disponível em: <https://www.ecommercebrasil.com.br/artigos/como-o-varejo-mudou-para-sempre> Acesso em: 11 nov. 2024.

ROESLER, Leonardo. **40% das empresas que não inovam estagnam ou perdem mercado.** Investing, 2023. Disponível em: <https://br.investing.com/analysis/40-das-empresas-que-nao-inovam-estagnam-ou-perdem-mercado-200460059> Acesso em: 11 nov. 2024.

62% das pessoas afirmam pagar mais caro por produtos sustentáveis. G1, 2024. Disponível em: <https://g1.globo.com/sc/santa-catarina/especial-publicitario/engie-brasil/noticia/2024/08/27/62percent-das-pessoas-affirmam-pagar-mais-caro-por-produtos-sustentaveis.ghml> Acesso em: 06 dez. 2024.

