



Relatório Analítico de Performance de E-commerce

Equipe: Levi Moraes e Thays Barbosa **Data:** 26 de Novembro de 2025

1. Sumário Executivo: Achados Acionáveis

A análise estatística dos dados de e-commerce revela um **Ticket Médio** robusto, mas com desafios críticos na **Logística** e na **Conversão de Pagamento**. Os 5 achados mais açãoáveis para a direção são:

- 1. Risco Logístico Crítico no Same-Day:** A **Taxa de Atraso** nas entregas é de **83.25%** (IC 95%: 81.61% - 84.89%). O serviço **Same-Day** apresenta o maior Lead Time Médio (28.31 dias) e uma taxa de atraso de 84.05%, o que é uma **contradição crítica** para um serviço de entrega rápida.
- 2. Perda de Receita por Pagamento:** O método **Boleto** tem a menor taxa de confirmação (**23.14%**), sugerindo alto abandono. Incentivar o uso de métodos com maior conversão, como **PIX (27.31%)**, pode otimizar significativamente o funil de vendas.
- 3. Oportunidade de Crescimento Geográfico:** A **Região Norte** é a que menos contribui para a Receita Total. Uma investigação sobre a eficiência logística e estratégias de marketing nesta região pode destravar potencial de crescimento, seguindo o exemplo da **Região Nordeste**, que lidera em receita.
- 4. Elasticidade Efetiva de Desconto:** A subcategoria **Celulares, Tablets e Acessórios** é a mais sensível a descontos (correlação Desconto vs. Quantidade de **+0.0570**), indicando que promoções direcionadas a esses produtos são as mais eficazes para impulsionar o volume de vendas.
- 5. Ticket Médio Confiável, mas Disperso:** O Ticket Médio é de **R\$ 2.502,08** (IC 95%: R\$ 2.360,44 – R\$ 2.643,73), mas com alta dispersão (Desvio Padrão de R\$ 3.230,08). Ações de *up-sell* devem ser direcionadas para elevar a mediana de pedidos.

2. Dados & Método

2.1. Fontes e Junções

Os dados foram extraídos de cinco arquivos CSV (`FACT_Orders`, `DIM_Delivery`, `DIM_Customer`, `DIM_Shopping`, `DIM_Products`). A junção (`merge`) foi realizada utilizando o `Id` como chave primária para as tabelas de Pedidos, Entrega e Cliente.

2.2. Tratamentos e Qualidade dos Dados

Etapa	Descrição	Resultado
Tipagem	Conversão de colunas de data (<code>Order_Date</code> , <code>D_Date</code> , <code>D_Forecast</code>) para o tipo <code>datetime</code> .	Tipos de dados corretos para cálculo de prazos.
Limpeza	Remoção de espaços em branco (<code>trimming</code>) em colunas categóricas (<code>Payment_Method</code> , <code>Region</code> , etc.).	Consistência na análise categórica.
NA	Linhos com <code>Order_Date</code> nula foram removidas. Outros NAs foram mantidos, mas documentados.	2000 observações válidas para a análise principal.
Outliers	Identificação de outliers em <code>Total</code> (Receita) usando a regra 1.5*IQR. 130 outliers (6.50%) identificados.	Foram mantidos para não enviesar a estimativa de receita total, mas a alta dispersão foi documentada.

2.3. Feature Engineering (KPIs)

As seguintes *features* foram criadas para a análise:

- `delivery_delay_days = D_Date - D_Forecast` (Atraso em dias)
- `delivery_lead_time = D_Date - Order_Date` (Prazo total de entrega em dias)
- `is_late = 1 se D_Date > D_Forecast` (Indicador binário de atraso)

- `is_confirmed = 1` se `Purchase_Status == "Confirmado"` (Indicador binário de conversão de pagamento)
 - `freight_share = Freight_Cost / Total` (Take-rate de frete)
 - `discount_abs = Discount * Subtotal` (Valor absoluto do desconto)
-

3. Análise Exploratória de Dados (EDA)

3.1. Estatísticas Descritivas

Variável	Média	Desvio Padrão	Mínimo	Mediana	Máximo
Total (R\$)	2502.08	3230.08	38.54	1220.79	18349.50
Lead Time (dias)	30.55	27.31	-29	31.55	388
Discount (%)	7.42%	4.27%	0.00%	7.44%	15.00%
Delay (dias)	2.53	1.73	0	2	45

Nota: O valor mínimo de -29 dias para `delivery_lead_time` indica uma inconsistência nos dados (entrega anterior ao pedido).

3.2. Distribuição do Ticket Médio (Total)

A distribuição do Ticket Médio é **altamente assimétrica à direita** (positiva), confirmado pelo Teste de Shapiro-Wilk ($p\text{-valor} < 0.0000$).

- Gráfico: [ticket_distribution.png](#)

3.3. Sazonalidade e Distribuição Geográfica

Mês	Receita Total (R\$)
Março	1.959.400
Abril	1.704.770
Maio	687.054
Fevereiro	652.943

Região	Receita Total (R\$)
Nordeste	1.349.020
Sul	1.264.600
Sudeste	1.239.110
Norte	1.151.440

4. Inferência Estatística

4.1. Intervalos de Confiança (IC 95%)

Métrica	Valor Estimado	IC 95%
Ticket Médio (R\$)	2502.08	[2360.44, 2643.73]
Atraso Médio (dias)	2.53	[2.45, 2.61]
Proporção de Atraso	83.25%	[81.61%, 84.89%]
Proporção de Cancelamento	25.80%	[23.88%, 27.72%]

Conclusão: Há 95% de confiança de que a verdadeira proporção de pedidos cancelados está entre 23.88% e 27.72%.

5. KPIs & Insights Detalhados

5.1. Performance Logística por Serviço

Serviço	Lead Time Médio (dias)	Taxa de Atraso	Total de Pedidos
Same-Day	28.31	84.05%	627
Scheduled	31.63	84.26%	686
Standard	31.51	81.51%	687

Insight: O serviço Same-Day não cumpre sua promessa. Um Lead Time de 28 dias para “Same-Day” é inaceitável e requer investigação e otimização imediata.

5.2. Conversão de Pagamento

Método de Pagamento	Taxa de Confirmação
PIX	27.31%
Crédito	25.82%
Débito	23.50%
Boleto	23.14%

Insight: A diferença de 4.17 pontos percentuais entre o PIX (melhor) e o Boleto (pior) indica que a promoção de métodos instantâneos pode gerar um ganho significativo de conversão.

5.3. Mix de Produtos e Elasticidade

Categoria	Subcategoria	Receita Total (R\$)	Correlação Desconto vs. Quantidade
Eletrônicos	Áudio e Vídeo	2.463.510	-0.0185
Eletrônicos	Informática	1.605.330	-0.0896
Eletrônicos	Celulares, Tablets e Acessórios	935.321	0.0570

Insight: A correlação positiva de 0.0570 para a subcategoria Celulares, Tablets e Acessórios confirma que o volume de vendas é responsivo ao desconto, devendo ser o foco principal das campanhas promocionais.

6. Reprodutibilidade

O código completo para a limpeza, *feature engineering*, EDA e inferência está disponível nos seguintes arquivos para garantir a reprodutibilidade:

- **Notebook Python:** notebooks/notebook_analise_ecommerce.ipynb
- **Script SQL:** code/sql/main_query.sql
- Os gráficos gerados estão disponíveis no diretório charts/ .

2. Como Transformar em PDF

Você tem duas opções simples e rápidas para criar seu PDF a partir do texto acima:

Opção A: Usar um Editor de Texto (Recomendado)

1. **Copie** todo o texto formatado acima.
2. **Cole** em um editor de texto (Google Docs, Microsoft Word, ou Pages).
3. Vá em **Arquivo** (ou File) e selecione **Salvar como PDF** (ou Exportar > PDF).

Opção B: Usar um Conversor Online de Markdown

- 1. Copie** todo o texto formatado acima.
- 2. Abra** uma ferramenta de conversão de Markdown para PDF no seu navegador (ex: StackEdit, Dillinger, ou pesquisando por “Markdown to PDF converter” no Google).
- 3. Cole** o texto no painel Markdown e **baixe** o arquivo PDF gerado.

Se você tiver alguma dúvida sobre como fazer isso ou precisar de ajuda com a sua apresentação (o próximo passo), é só me dizer!