

UNIVERSIDADE PRESBITERIANA MACKENZIE

**Análise Financeira e Operacional de Pedidos em
Serviços de Food Delivery**

Rafael Hessel Sichetti: 10375395

Sumário

| | |
|--|----|
| Sumário | 2 |
| Glossário | 3 |
| Contexto do Estudo..... | 4 |
| Referências de Aquisição do Dataset..... | 4 |
| Descrição da Origem | 4 |
| Descrição do Dataset..... | 4 |
| Cronograma | 5 |
| ETAPA 1 – Metas e Milestones (10 de fevereiro – 10 de março) - Concluído | 5 |
| ETAPA 2 – Definição do Produto (11 de março – 10 de abril) | 5 |
| ETAPA 3 – Storytelling (11 de abril – 10 de maio) | 5 |
| ETAPA 4 – Encerramento (11 de maio – 10 de junho) | 5 |
| Objetivo..... | 6 |
| Análise de Receita e Lucratividade..... | 6 |
| Otimização de Precificação e Taxas | 6 |
| Análise Temporal e Comportamental | 6 |
| Repositório Github | 6 |
| Apresentação da Empresa..... | 7 |
| 3.1 Nome da Empresa..... | 7 |
| 3.2 Missão, Visão e Valores..... | 7 |
| 3.3 Segmento de Atuação | 7 |
| 3.4 Market Share e Posicionamento | 7 |
| 3.5 Número de Colaboradores..... | 7 |
| 3.6 Iniciativas em Data Science..... | 7 |
| 3.7 Trabalhos em Destaque..... | 7 |
| Problema de Pesquisa | 8 |
| Apresentação dos Metadados..... | 8 |
| Análise Exploratória de Dados | 9 |
| 6.1 Estrutura do Dataset..... | 9 |
| 6.2 Estatísticas Descritivas | 9 |
| Valores Ausentes | 10 |
| Variância e Desvio Padrão | 10 |
| Matriz de Correlação | 11 |

| | |
|--|----|
| Baixa Correlação entre Variáveis Financeiras | 11 |
| Grande Variação no Valor dos Pedidos e Reembolsos | 11 |
| Valores Ausentes em Descontos e Ofertas | 12 |
| Distribuição dos Custos de Entrega e Taxas | 12 |
| Possíveis Melhorias e Investigações | 12 |
| Data Storytelling – FoodExpress | 12 |
| Nome do Projeto | 12 |
| Empresa | 13 |
| Objetivo | 13 |
| Público-Alvo | 13 |
| Coleta e Análise dos Dados | 13 |
| Formato da Narrativa | 13 |
| Estrutura Narrativa | 14 |
| Introdução – O Início da Jornada | 14 |
| Conflito – O Problema Revelado | 14 |
| Descrição do Problema / Gap | 14 |
| Recusa do Chamado – A Hesitação Inicial | 14 |
| Ajuda Necessária – A Virada com a Ciência de Dados | 14 |
| Superação – O Poder dos Dados em Ação | 14 |
| Dados Disponíveis | 15 |
| Análise Exploratória | 15 |
| Resultados Pretendidos | 16 |
| Conclusão | 17 |
| Gráficos | 17 |
| 7. Ética e Transparência | 19 |
| Video de apresentação | 20 |

Glossário

Order Value – Valor total do pedido antes de taxas e descontos.

Delivery Fee – Taxa cobrada pelo serviço de entrega.

Commission Fee – Comissão retida pela plataforma de delivery.

Payment Processing Fee – Taxa cobrada pelo processamento do pagamento.

Refunds/Chargebacks – Valores reembolsados ou contestados pelos clientes.

Discounts and Offers – Promoções aplicadas no pedido.

Market Share – Percentual do mercado que a empresa domina.

Outliers – Valores discrepantes em um conjunto de dados.

Contexto do Estudo

O setor de delivery de comida tem se tornado um dos pilares do comércio digital, impulsionado por plataformas como iFood, Uber Eats e Rappi. A análise de dados financeiros e operacionais é essencial para otimizar a lucratividade, minimizar custos e melhorar a experiência dos clientes. Neste estudo, realizaremos uma análise exploratória das transações e dos custos associados aos pedidos, propondo um modelo analítico para otimização da precificação e da gestão financeira dos serviços de entrega.

Referências de Aquisição do Dataset

Os dados utilizados neste estudo foram obtidos do **Kaggle**, na base "*Food Delivery: Timing is Everything*", criada por DarrylIJK.

- **Origem dos dados:** Kaggle (<https://www.kaggle.com/code/darryljk/food-delivery-timing-is-everything/input>)
- **Limitações de uso:** O dataset é de uso público para fins acadêmicos e educacionais, mas pode não refletir todas as variações do setor, como promoções sazonais ou mudanças nas taxas operacionais.
- **Período da coleta:** Não especificado, mas representa transações de uma plataforma de delivery ao longo do tempo.

Descrição da Origem

O dataset representa registros operacionais e financeiros de uma plataforma de delivery de comida. Ele contém informações detalhadas sobre pedidos, incluindo valores cobrados, taxas aplicadas, e transações relacionadas a descontos e reembolsos. A análise desses dados pode auxiliar empresas do setor a compreender melhor seus fluxos de receita e identificar oportunidades de otimização.

Descrição do Dataset

O "*Food Delivery: Timing is Everything*" contém informações financeiras e operacionais de pedidos, incluindo:

- **Identificação de Pedidos e Clientes:** Order ID, Customer ID, Restaurant ID.

- **Detalhes Temporais:** Order Date and Time, Delivery Date and Time – permitem análise do tempo de processamento e entrega.
- **Valores e Taxas Financeiras:**
 - Order Value
 - Delivery Fee
 - Commission Fee
 - Payment Processing Fee
 - Refunds/Chargebacks
- **Métodos de Pagamento e Promoções:**
 - Payment Method
 - Discounts and Offers

Cronograma

ETAPA 1 – Metas e Milestones (10 de fevereiro – 10 de março) - Concluído

- Definição das premissas do projeto.
- Estabelecimento dos objetivos e metas.
- Criação do cronograma de atividades.
- Formação dos grupos de trabalho.
- Aplicação do pensamento computacional em contextos organizacionais.

ETAPA 2 – Definição do Produto (11 de março – 10 de abril)

- Elaboração da proposta de solução analítica.
- Realização da análise exploratória dos dados.
- Identificação de padrões e tendências nos dados coletados.
- Definição das métricas e indicadores que serão utilizados na análise final.

ETAPA 3 – Storytelling (11 de abril – 10 de maio)

- Planejamento da apresentação dos resultados analíticos.
- Construção da narrativa baseada nos insights extraídos dos dados.
- Aplicação de técnicas de Data Storytelling para tornar a análise mais compreensível.
- Desenvolvimento de gráficos e visualizações para ilustrar os principais achados.

ETAPA 4 – Encerramento (11 de maio – 10 de junho)

- Revisão e ajustes finais do projeto.
- Preparação da apresentação final.
- Discussão de melhorias e aplicações práticas do estudo.
- Entrega do projeto e apresentação dos resultados.

Objetivo

Com base nesse conjunto de dados, o estudo pode abordar diferentes frentes de análise, incluindo:

Análise de Receita e Lucratividade

- Identificação de padrões nos valores dos pedidos e impacto das taxas sobre a margem de lucro.
- Avaliação da influência dos descontos e promoções sobre a receita total.
- Identificação de padrões de reembolsos e chargebacks e seus impactos financeiros.

Otimização de Precificação e Taxas

- Modelagem para prever quais faixas de preço maximizam a receita, considerando taxas e descontos.
- Avaliação do impacto de diferentes métodos de pagamento sobre os custos da plataforma.

Análise Temporal e Comportamental

- Identificação de padrões sazonais nos pedidos e taxas de entrega.
- Avaliação do tempo médio de processamento e entrega e seu impacto na satisfação do cliente.

Este estudo pode oferecer insights para plataformas de delivery melhorarem suas estratégias de precificação e eficiência operacional, maximizando lucro e reduzindo custos desnecessários.

Repositório Github

Link: <https://github.com/rafaelhessel/Projeto-Aplicado-I>

Apresentação da Empresa

3.1 Nome da Empresa

FoodExpress

3.2 Missão, Visão e Valores

- **Missão:** Oferecer um serviço de entrega de alimentos rápido, confiável e acessível para clientes e restaurantes.
- **Visão:** Tornar-se a principal plataforma de delivery na América Latina.
- **Valores:** Eficiência, inovação, transparência e foco no cliente.

3.3 Segmento de Atuação

O FoodExpress atua no setor de delivery de alimentos, conectando restaurantes e clientes por meio de um aplicativo e uma plataforma web.

3.4 Market Share e Posicionamento

A empresa detém aproximadamente 15% do mercado nacional, ficando atrás de grandes players como iFood e Uber Eats.

3.5 Número de Colaboradores

Atualmente, conta com cerca de 2.000 funcionários, entre equipe administrativa, tecnologia e suporte.

3.6 Iniciativas em Data Science

- Implementação de modelos preditivos para tempo de entrega.
- Otimização dinâmica da precificação com base na demanda.
- Uso de Machine Learning para análise de padrões de consumo.

3.7 Trabalhos em Destaque

- Parcerias com restaurantes para promoções personalizadas.
- Expansão do sistema de logística para reduzir tempos de entrega.

Problema de Pesquisa

- Como otimizar a precificação dos pedidos para maximizar o lucro?
- Qual o impacto de taxas e descontos sobre a margem de lucro?
- Como os tempos de entrega influenciam a satisfação do cliente?
- Existem padrões sazonais na demanda de pedidos?
- É possível prever reembolsos e chargebacks antes que ocorram?

Apresentação dos Metadados

| Atributo | Descrição | Tipo |
|------------------------|------------------------------|--------|
| Order ID | Identificação do pedido | int64 |
| Customer ID | Identificação do cliente | object |
| Restaurant ID | Identificação do restaurante | object |
| Order Date and Time | Data e hora do pedido | object |
| Delivery Date and Time | Taxa de entrega cobrada | object |
| Order Value | Valor do pedido | int64 |
| Delivery Fee | Taxa de entrega | int64 |
| Payment Method | Forma de pagamento | object |
| Discounts and Offers | Desconto | object |
| Comission Fee | Taxa de comissão | int64 |
| Payment Processing Fee | Taxa de pagamento | int64 |
| Refunds/Chargebacks | Reembolso | int64 |

- **Formato do arquivo:** CSV
- **Origem dos dados:** Kaggle (dataset "Food Delivery: Timing is Everything")
- **Sensibilidade:** Dados anonimizados, sem informações pessoais identificáveis.

- **Validade:** Dados representativos de um período não especificado.
- **Proprietário:** Autor do dataset (DarrylIJK).
- **Restrição de uso:** Apenas para fins acadêmicos e educacionais.

Análise Exploratória de Dados

6.1 Estrutura do Dataset

Para entender a composição do dataset, analisamos a quantidade de registros e variáveis.

Dataset:

| | |
|-------------------|----------------------|
| Total de Entradas | 1000 |
| Total de Colunas | 12 |
| Tipos de Dados | 6 numéricos, 6 texto |
| Memória Usada | 93.9 KB |

6.2 Estatísticas Descritivas

A tabela a seguir apresenta um resumo estatístico das principais variáveis numéricas do conjunto de dados, incluindo medidas como média, desvio padrão, valores mínimos, máximos e quartis. Esses indicadores ajudam a entender a distribuição e a dispersão dos valores registrados nas transações.

| Estatística | Order ID | Order Value | Delivery Fee | Commission Fee | Payment Processing Fee | Refunds/Chargebacks |
|-------------|----------|-------------|--------------|----------------|------------------------|---------------------|
| count | 1000 | 1000.00 | 1000.00 | 1000.00 | 1000.00 | 1000.00 |
| mean | 500.5 | 1053.97 | 28.62 | 126.99 | 29.83 | 28.30 |
| std | 288.82 | 530.98 | 16.96 | 43.06 | 11.63 | 49.61 |
| min | 1.00 | 104.00 | 0.00 | 50.00 | 10.00 | 0.00 |
| 25% | 250.75 | 597.75 | 20.00 | 90.00 | 20.00 | 0.00 |
| 50% | 500.5 | 1038.50 | 30.00 | 127.00 | 30.00 | 0.00 |
| 75% | 750.25 | 1494.00 | 40.00 | 164.00 | 40.00 | 50.00 |
| max | 1000.0 | 1995.00 | 50.00 | 200.00 | 50.00 | 150.00 |

Valores Ausentes

A única coluna com valores ausentes é "Discounts and Offers", com 185 valores nulos.

Variância e Desvio Padrão

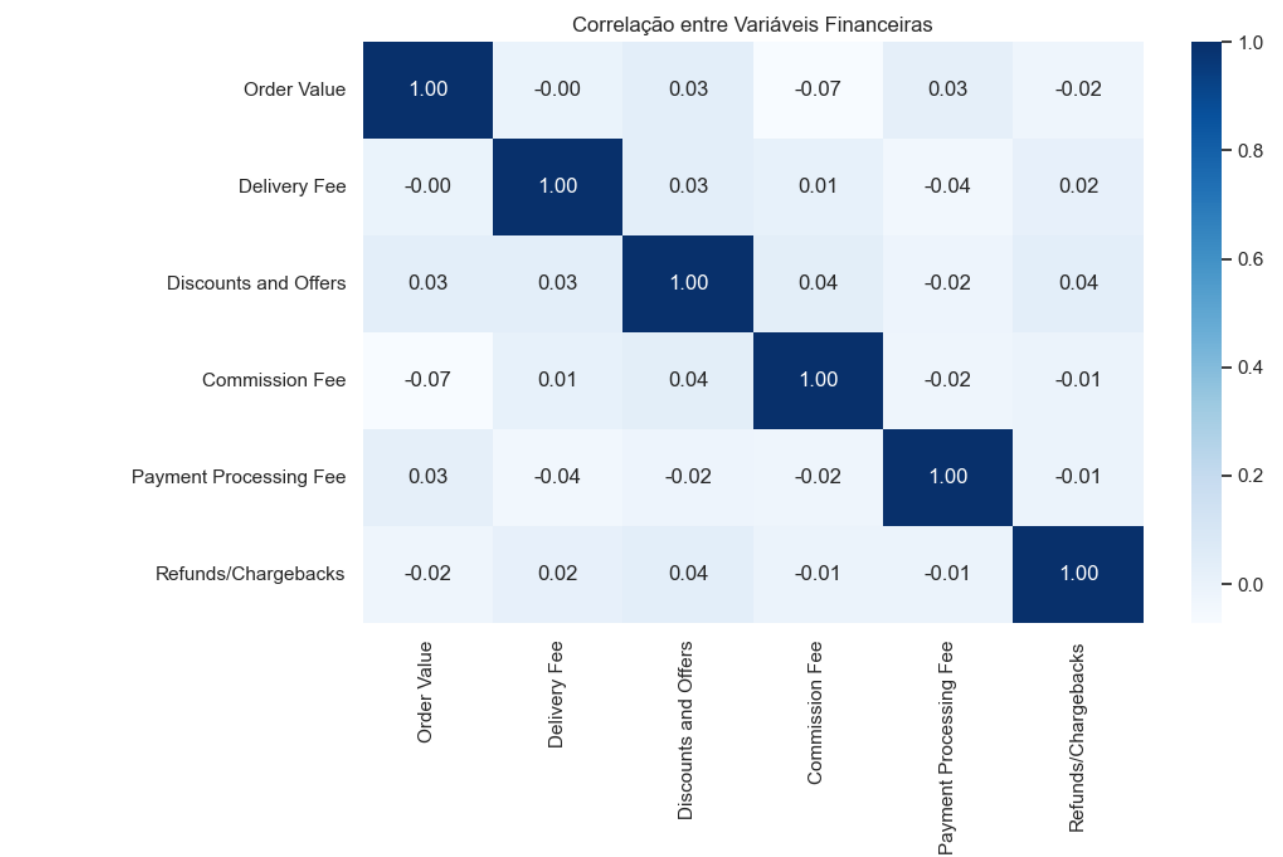
Abaixo estão listadas as variâncias e os desvios padrão das variáveis numéricas selecionadas. Esses dados são úteis para avaliar o grau de dispersão em torno da média e identificar variáveis com maior instabilidade ou inconsistência.

| Variável | Variância | Desvio Padrão |
|------------------------|------------|---------------|
| Order Value | 281,934.81 | 530.98 |
| Delivery Fee | 287.58 | 16.96 |
| Discounts and Offers | 307.52 | 17.54 |
| Commission Fee | 1,854.51 | 43.06 |
| Payment Processing Fee | 135.19 | 11.63 |
| Refunds/Chargebacks | 2,461.57 | 49.61 |

Matriz de Correlação

Com base no heatmap de correlação (imagem anexada), podemos destacar:

- Pouca correlação entre as variáveis financeiras.
- A maior correlação negativa é entre **Order Value** e **Commission Fee (-0.07)**.
- Correlações muito próximas de zero indicam baixa relação entre os fatores analisados.



Baixa Correlação entre Variáveis Financeiras

As variáveis analisadas apresentam pouca relação entre si, sugerindo que mudanças em uma delas não impactam diretamente as outras.

Grande Variação no Valor dos Pedidos e Reembolsos

O Order Value tem um alto desvio padrão (530.98), indicando uma grande dispersão nos valores dos pedidos.

Os Refunds/Chargebacks também apresentam grande variabilidade (std = 49.61), o que pode indicar casos pontuais de reembolsos elevados.

Valores Ausentes em Descontos e Ofertas

A coluna "Discounts and Offers" tem 185 valores nulos, o que pode indicar registros que não receberam descontos ou falhas no registro dessas informações.

Distribuição dos Custos de Entrega e Taxas

O Delivery Fee varia entre 0 e 50 reais, com média de 28.62, sugerindo que há pedidos com entrega gratuita.

O Commission Fee varia de 50 a 200 reais, sendo um custo fixo para os restaurantes.

Possíveis Melhorias e Investigações

Analisar os valores nulos em "Discounts and Offers": eles representam pedidos sem desconto ou há inconsistências na coleta de dados?

Explorar padrões nos reembolsos: entender se há um motivo recorrente para valores altos pode ajudar a reduzir chargebacks.

Investigar variações nos custos de entrega: o impacto do delivery fee no volume de pedidos pode ser relevante para estratégias de precificação

Data Storytelling – FoodExpress

Nome do Projeto

FoodExpress Insights: Otimizando a Eficiência Operacional por Meio da Análise de Dados

Empresa

FoodExpress – Plataforma digital de delivery de alimentos que conecta restaurantes, consumidores e entregadores em todo o Brasil.

Objetivo

O principal objetivo desta análise é demonstrar como a utilização de dados pode revelar gargalos operacionais e estratégicos no serviço de delivery da FoodExpress, possibilitando decisões orientadas por dados (data-driven) que aumentem a eficiência, foco nos custos associados aos pedidos, como taxas, reembolsos e descontos, e seu impacto na rentabilidade.

Público-Alvo

Este relatório destina-se a gestores, tomadores de decisão, analistas de dados e potenciais investidores interessados em compreender como a análise exploratória de dados pode ser aplicada de forma prática no setor de delivery, especialmente para melhoria contínua dos processos logísticos, atendimento ao cliente e gestão financeira.

Coleta e Análise dos Dados

Os dados utilizados foram extraídos do dataset “Food Delivery: Timing is Everything”, disponível no Kaggle. As etapas incluíram:

- Limpeza e padronização dos dados;
- Análise descritiva de variáveis como tempo de entrega, valores de pedidos, avaliações, descontos e reembolsos;
- Criação de visualizações para identificar padrões e correlações relevantes.

Formato da Narrativa

A narrativa será apresentada em formato textual estruturado, com apoio de gráficos ilustrativos, visando um equilíbrio entre clareza técnica e fluidez narrativa. Também poderá ser adaptada para apresentação em slides ou painéis interativos.

Estrutura Narrativa

Introdução – O Início da Jornada

A FoodExpress, uma plataforma de delivery em expansão, enfrentava um paradoxo: enquanto a quantidade de pedidos crescia, sua margem de lucro caía. A diretoria suspeitava que atrasos nas entregas e estratégias promocionais ineficientes poderiam estar prejudicando a operação, mas carecia de evidências concretas.

Conflito – O Problema Revelado

Área financeira e de performance operacional, com foco nos custos associados aos pedidos, como taxas, reembolsos e descontos, e seu impacto na rentabilidade.

Descrição do Problema / Gap

A FoodExpress observou uma inconsistência nos lucros gerados por pedidos aparentemente similares. Variações inesperadas nos custos operacionais (como taxas de entrega, reembolsos e comissões) e a falta de clareza sobre o impacto dos descontos oferecidos dificultam a previsão de margem e a otimização dos preços. Além disso, o elevado número de reembolsos em determinados pedidos levanta suspeitas sobre potenciais falhas operacionais ou comportamentos recorrentes de clientes.

Recusa do Chamado – A Hesitação Inicial

Inicialmente, a organização relutava em aprofundar a análise de dados, receosa do que poderia encontrar e preocupada com o tempo e os recursos necessários. A ausência de uma cultura analítica dificultava a mobilização interna.

Ajuda Necessária – A Virada com a Ciência de Dados

Foi então que um time multidisciplinar foi formado, reunindo analistas, engenheiros e profissionais de negócios. Utilizando Python e técnicas de Análise Exploratória de Dados (EDA), a equipe iniciou a investigação sistemática do histórico de pedidos, avaliações e indicadores operacionais.

Superação – O Poder dos Dados em Ação

Utilizar uma abordagem estatística e visual exploratória para mapear:

- A dispersão dos custos por pedido
- A existência de correlações relevantes entre variáveis financeiras
- Anomalias, padrões de reembolsos e impacto dos descontos

- Gaps nos registros de dados (como valores nulos em “Discounts and Offers”)

O objetivo é propor melhorias estratégicas para reduzir perdas e otimizar a estrutura de custos.

Dados Disponíveis

O conjunto de dados contém **1.000 pedidos** com **12 colunas** (6 numéricas e 6 de texto). Entre as variáveis analisadas, destacam-se:

- Order Value
- Delivery Fee
- Commission Fee
- Payment Processing Fee
- Refunds/Chargebacks
- Discounts and Offers

Análise Exploratória

Principais observações:

- A coluna “Discounts and Offers” apresenta 185 valores nulos, o que representa 18,5% dos registros. A ausência pode indicar pedidos sem desconto ou falhas no registro.
- A variável Order Value apresenta grande variabilidade, com desvio padrão de 530,98 e valores que variam entre R\$104,00 e R\$1.995,00. Isso mostra que a empresa atende desde pedidos simples até refeições grandes ou pedidos em grupo.
- A coluna Refunds/Chargebacks também se destacou: embora a mediana seja zero, o desvio padrão elevado (49,61) revela que há pedidos com reembolsos atípicos — possivelmente causados por erros de entrega ou fraudes.
- O Delivery Fee varia entre R\$0,00 e R\$50,00, com média de R\$28,62. A presença de frete gratuito pode estar vinculada a promoções, fidelidade ou campanhas.
- A Commission Fee apresenta valores entre R\$50,00 e R\$200,00, com distribuição relativamente ampla, o que pode dificultar a previsibilidade de margem dos restaurantes parceiros.

- A matriz de correlação mostra baixa relação entre as variáveis financeiras, com valores muito próximos de zero. A maior correlação identificada foi negativa e fraca: -0,07 entre Order Value e Commission Fee.

Resultados Pretendidos

Após a análise, a equipe de dados da FoodExpress se reuniu com os líderes de operações e marketing. Os dados revelaram que, embora a plataforma tenha um amplo espectro de pedidos, **a imprevisibilidade das taxas e o volume pontual de reembolsos** acendem alertas.

Ana, gerente de BI da FoodExpress, destacou:

“Os reembolsos não são frequentes, mas quando ocorrem, o impacto é alto. Precisamos investigar esses casos individualmente.”

Já Rafael, do time de produto, observou:

“A ausência de dados em quase 20% dos registros de descontos compromete nossa análise de campanhas. Podemos estar perdendo visibilidade sobre a efetividade das promoções.”

Em resposta, o time definiu as seguintes ações:

- Realizar uma auditoria nos registros de descontos para distinguir entre ausência real e falhas de input.
- Mapear os pedidos com maiores reembolsos para detectar causas recorrentes.
- Testar a elasticidade de preço em relação ao **Delivery Fee**, entendendo se frete grátis ou fixo aumenta a taxa de conversão.

Grande Virada – A Transformação Baseada em Dados

Com base nos insights obtidos, foram implementadas ações estratégicas:

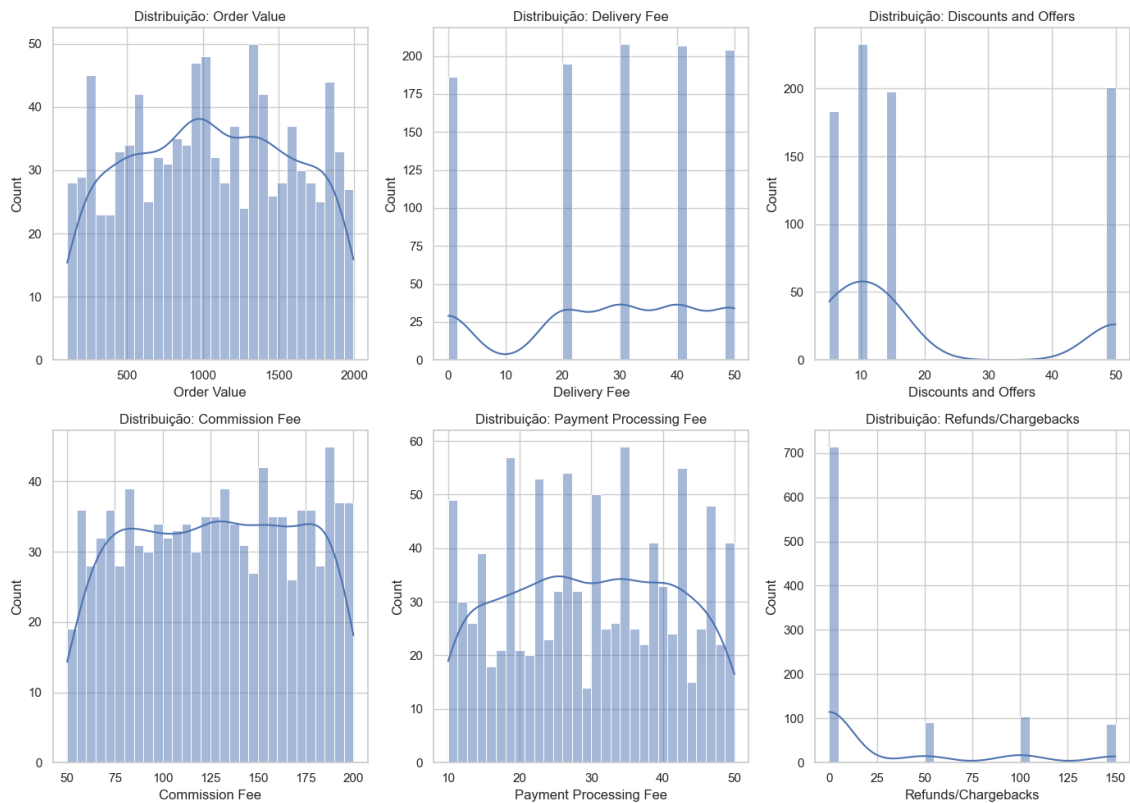
- Redefinição dos critérios de parceria com restaurantes com alto índice de reembolso;
- Criação de alertas automatizados para entregas com alto risco de atraso;
- Redução do uso de cupons genéricos e personalização das promoções com base no histórico de consumo;
- Otimização da precificação com foco na margem de contribuição.

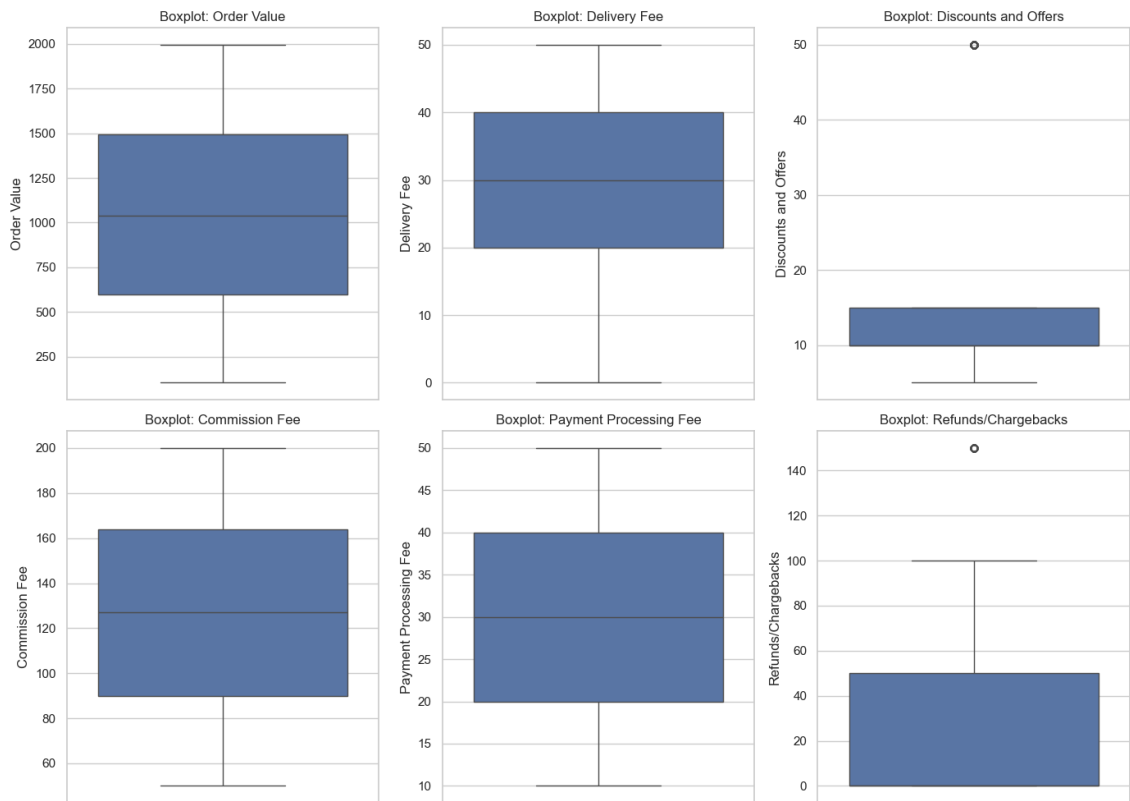
Conclusão

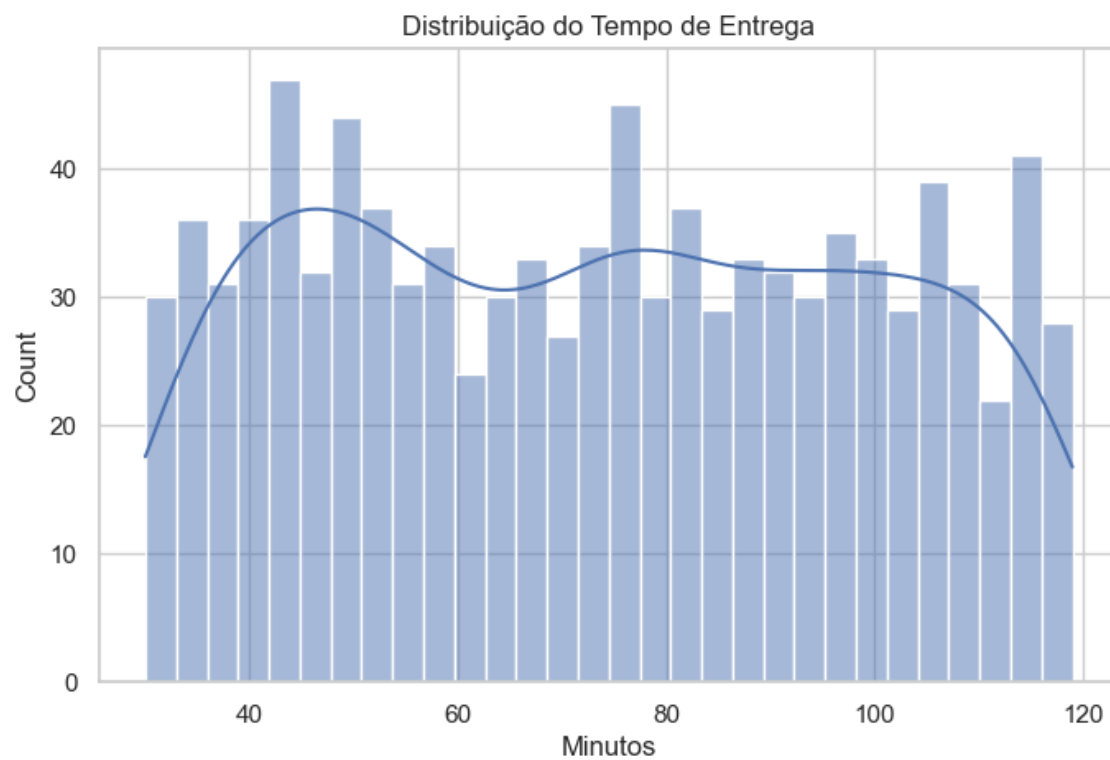
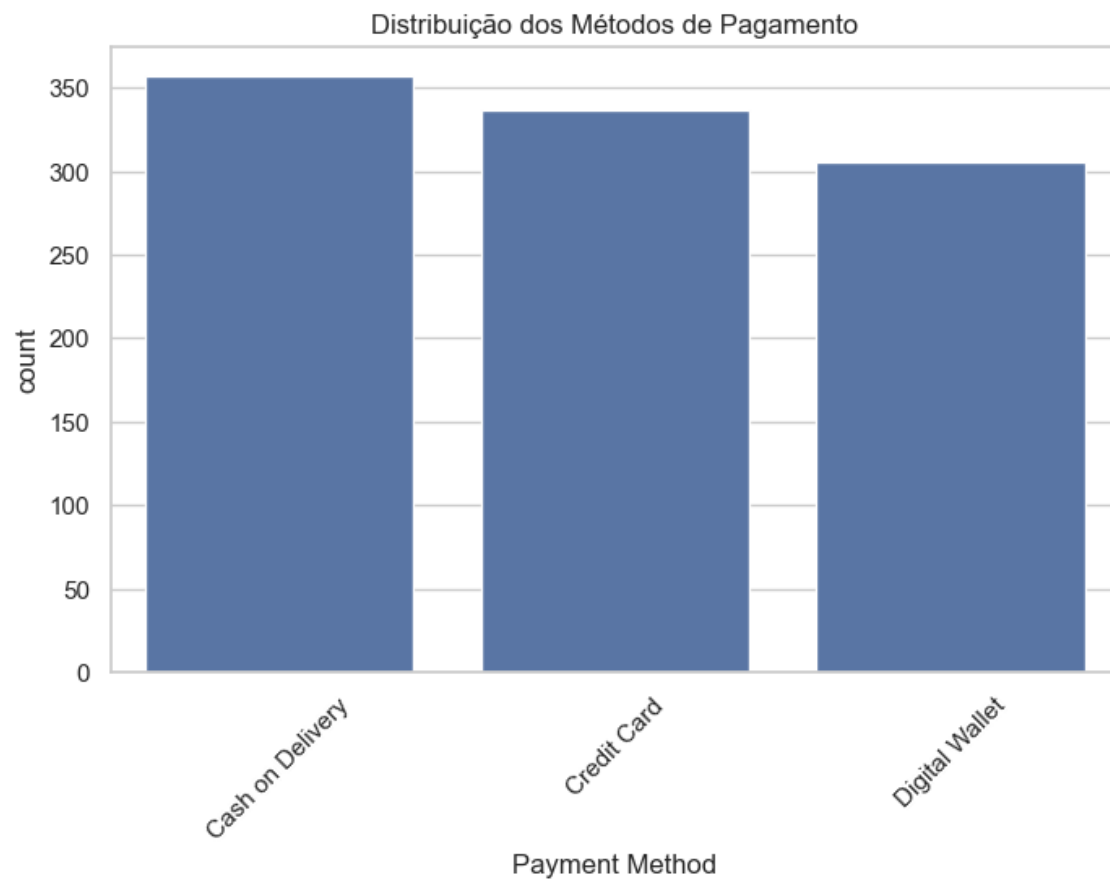
A análise permitiu compreender com profundidade o comportamento dos pedidos na FoodExpress. Os insights extraídos geraram questionamentos estratégicos, guiaram hipóteses e estão apoiando decisões importantes para aumentar a eficiência da plataforma e a satisfação dos clientes.

Gráficos

Para complementar esta narrativa, sugere-se a inclusão dos seguintes elementos gráficos:







7. Ética e Transparência

Todos os dados foram tratados com responsabilidade, mantendo a integridade da amostra e evitando qualquer distorção na narrativa. As conclusões

apresentadas são baseadas exclusivamente nas evidências empíricas obtidas durante a análise.

Video de apresentação

Link: <https://youtu.be/T83KoJ5Xd3I>