

DASH LOGISTICS

TRABALHANDO COM ANÁLISE DE DADOS PARA SETOR DE
LOGÍSTICA

Propósito: Desenvolver um dashboard com foco em análise e visualização de dados, definindo uma estrutura-base (indicadores, métricas e layout) e refazendo essa mesma solução em diferentes plataformas (Power BI, Excel, Python/Plotly, Tableau, Google Data Studio)





O que é Logística?

Descrição

A logística é o conjunto de atividades e processos envolvidos no planejamento, implementação e controle eficiente do fluxo de produtos, serviços e informações, desde o ponto de origem até o ponto de consumo. Seu objetivo é garantir que os recursos certos (produtos, materiais, informações) estejam no local certo, no momento certo, nas quantidades certas e ao menor custo possível.

Áreas

- Transporte: O movimento físico de mercadorias de um lugar para outro, por meio de diferentes modais, como rodoviário, ferroviário, aéreo e marítimo.
- Armazenagem: O processo de armazenar produtos em depósitos ou armazéns, de forma eficiente, para garantir que eles estejam disponíveis quando necessário.
- Gestão de Inventário: A administração do estoque, assegurando que haja a quantidade adequada de produtos disponíveis para atender à demanda, sem excessos ou faltas.
- Distribuição: A entrega dos produtos aos consumidores finais, que pode envolver centros de distribuição, centros de atendimento ao cliente, pontos de venda e até mesmo entregas diretas.
- Processamento de Pedidos: A administração dos pedidos recebidos, desde a sua recepção até a sua execução, garantindo que o cliente receba o produto conforme solicitado.
- Embalagem e Manuseio: O processo de proteger os produtos para transporte, armazenagem e manuseio de forma eficiente e segura.
- Gestão de Fluxo de Informações: A logística também lida com a troca de informações entre os diferentes elos da cadeia de suprimentos, para garantir que todas as partes envolvidas tenham acesso às informações necessárias em tempo real.



Six Steps of DATA ANALYTICS



ASK

Antes de coletar ou analisar dados, é essencial entender o que você quer saber e quais são os problemas ou oportunidades que precisam ser abordados. Definir os objetivos de análise claramente é fundamental.



PROCESS

Nesta fase, os dados são transformados e preparados para análise. Isso pode incluir a normalização, a conversão de formatos e a extração de informações relevantes.



SHARE

Após a análise, os resultados precisam ser compartilhados com as partes interessadas de forma clara e acessível. Isso pode ser feito por meio de relatórios, dashboards, visualizações ou apresentações, dependendo do público-alvo.



PREPARE

Após definir as perguntas, o próximo passo é preparar os dados. Isso inclui a coleta, limpeza e organização dos dados necessários para análise. A preparação também pode envolver a integração de diferentes fontes de dados e a transformação para formatos apropriados.



ANALYZE

Após a análise de dados é onde as perguntas feitas no início são respondidas. Nessa fase, técnicas estatísticas, algoritmos de aprendizado de máquina ou análise exploratória são usados para extrair padrões, tendências ou relações nos dados.



ACT

A fase final envolve a aplicação dos insights para tomar decisões ou agir. Isso pode resultar em mudanças nos processos de negócios, estratégias ou implementação de melhorias com base nas conclusões obtidas durante a análise.





Perguntando

Entendendo 'O que' e 'Porque'

Dentro desta base de dados obtida pela **Kaggle** em formato CVS ou xlsx, vamos levantar algumas perguntas relevantes que nos proporcionarão informações importantes. Neste momento, já temos a base de dados e vamos trabalhar com ela, mas é crucial que começemos a refletir sobre o que realmente precisamos. Caso não tenhamos as informações necessárias, é importante buscá-las, seja por meio de coleta de dados ou por outras fontes.

Perguntas

O principal objetivo desta análise será a melhoria dos prazos de entrega e devoluções. Para alcançar esse objetivo, formularemos perguntas que ajudem a entender os principais fatores que afetam esses prazos. Algumas perguntas importantes podem incluir:

- Quantos pedidos foram feitos?
- Quantos Itens foram pedidos?
- Quanto itens foram Devolvidos?
- Qual o valor total?
- Qual % de Entregas Atrasadas?
- Qual a % de Itens Devolvidos?
- Quantos Pedidos Estão Atrasados?
- Quem são nossos clientes e como estamos de status com eles?
- Porque nossos pedidos são devolvidos?
- Qual a nosso LeadTime Real e Prevista por Região?
- Como nosso Motoristas estão com os números de Pedidos?

Fonte de dados em CVS

kaggle

Preparando

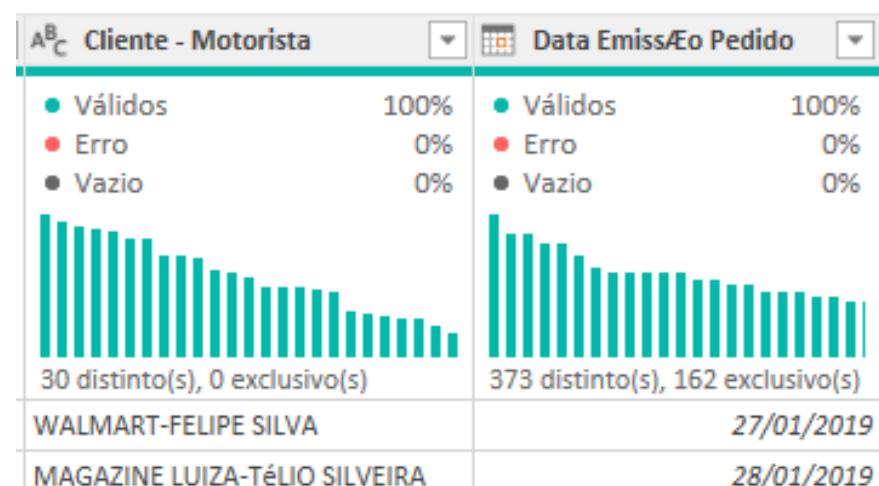
Após definirmos as perguntas corretas, temos um objetivo claro, e assim podemos iniciar o trabalho com os dados. Na segunda etapa — Preparação — o foco será deixar os dados limpos e organizados. Esta fase também inclui a coleta de dados adicionais, caso tenha sido identificado na etapa anterior que faltam informações relevantes.

É importante destacar que, durante o processo de limpeza dos dados, devemos estar atentos à qualidade dos mesmos. A qualidade dos dados refere-se à confiabilidade e precisão das informações, e para garantir isso, precisamos verificar questões como:

- **Presença de erros nos dados** (por exemplo, inconsistências ou falhas nos registros).
- **Células vazias ou dados ausentes** que podem afetar a análise.
- **Valores extremos ou absurdos**, que podem indicar erros de digitação ou de coleta de dados.
- **Formatação inconsistente**, que pode dificultar a análise (por exemplo, diferentes formas de registrar datas ou categorias).

Power Query

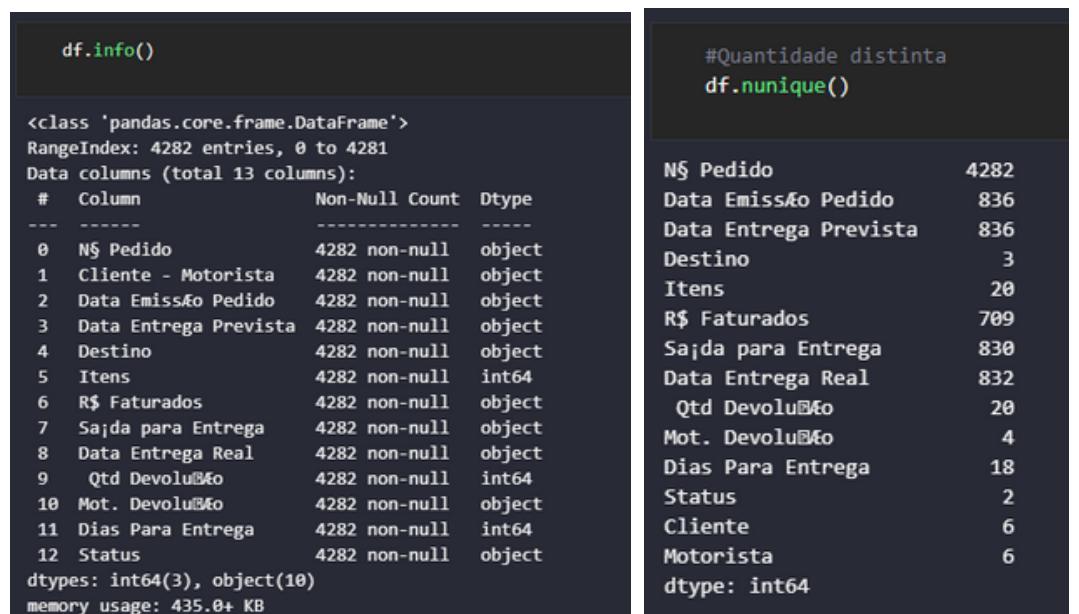
Quando trabalhamos com **Excel** ou **Power Bi** temos acesso a uma ferramenta muito poderosa o **Power Query** ela faz a Transformação e a preparação de dados. Com ela podemos Verificar Rapidamente a qualidade dos dados



Fazendo a verificação Coluna a Coluna, percebemos que todos os dados estão não apresentam erros ou células ausentes, porém a coluna "Cliente - Motorista" precisa ser separada.

Python

Em **Python** podemos fazer toda a verificação com a biblioteca **pandas**, com comandos como **info()** e **nunique()** e fazer a separação das colunas com **split**. Além da Biblioteca em questão foi também utilizado a biblioteca **Polars**, que obtém um ganho de performance quando comparada a Pandas.



The figure shows two code snippets in a Jupyter Notebook. The first snippet, `df.info()`, displays the DataFrame structure with 13 columns, all non-null objects. The second snippet, `#Quantidade distinta df.nunique()`, shows the count of unique values for each column, ranging from 2 to 4282. The output for `df.nunique()` is as follows:

Column	Count
Nº Pedido	4282
Data Emissão Pedido	836
Data Entrega Prevista	836
Destino	3
Itens	20
R\$ Faturados	709
Saída para Entrega	830
Data Entrega Real	832
Qtd Devolução	20
Mot. Devolução	4
Dias Para Entrega	18
Status	2
Cliente	6
Motorista	6





Analizando

Com os **dados devidamente limpos, organizados e processados**, passamos à fase de análise, que tem como objetivo explorar os dados para responder às perguntas definidas na primeira etapa. Nesta fase, transformamos dados brutos em informações úteis, padrões relevantes e insights estratégicos.

A análise pode assumir diferentes abordagens, dependendo do objetivo do projeto e da natureza dos dados:

- **Análise descritiva:** resume as principais características dos dados, como médias, totais, proporções, frequências e variações. Essa abordagem ajuda a entender o comportamento geral dos dados.
- **Análise exploratória:** utiliza gráficos, visualizações interativas e técnicas estatísticas para identificar padrões, outliers, correlações e tendências que possam não ser visíveis à primeira vista.
- **Modelagem preditiva ou inferencial:** quando o projeto exige previsões ou estimativas, podemos aplicar modelos de regressão, classificadores, séries temporais ou algoritmos de machine learning, como árvores de decisão ou redes neurais.
- **Análise comparativa:** avalia o desempenho entre grupos ou períodos diferentes para identificar melhorias ou quedas de desempenho, causas prováveis e possíveis hipóteses.
- É fundamental que, durante a análise, o foco continue sendo as perguntas de negócio ou hipóteses levantadas inicialmente. A análise precisa ser objetiva, clara e gerar informações que possam orientar decisões práticas ou revelar oportunidades escondidas nos dados.

Power Bi, Excel e Locker

A análise será conduzida por meio de um conjunto de visualizações integradas em um **dashboard**, projetado com base em métricas e indicadores definidos anteriormente. A estrutura e o layout desse dashboard estão detalhados no slide seguinte.

Python

Em Python, o objetivo da análise permanece o mesmo: representar visualmente os dados por meio de gráficos e indicadores relevantes. No entanto, diferentemente de ferramentas como Power BI ou Excel, onde a construção visual é feita de forma mais interativa, em Python todo o processo de criação dos gráficos é realizado por meio de linhas de código.

Para isso, utilizamos bibliotecas plotly, que oferecem recursos avançados para customização e automação das visualizações. Essa abordagem exige um conhecimento mais técnico, mas oferece maior flexibilidade, controle sobre os elementos gráficos.



Dashboard Layout

PEDIDOS
0.000

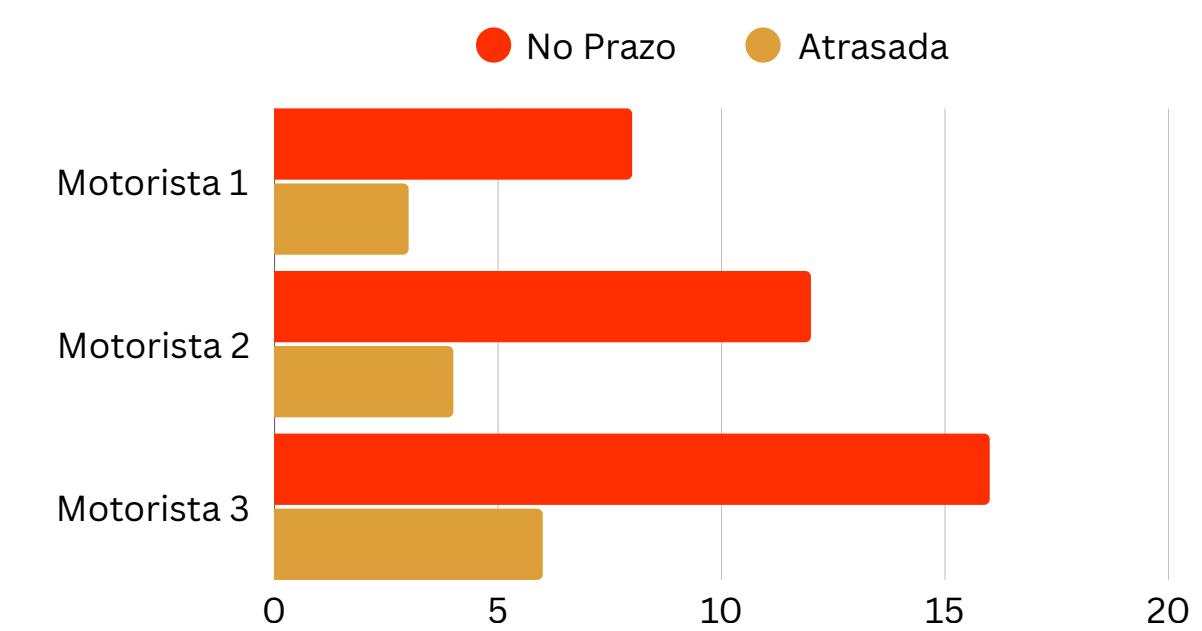
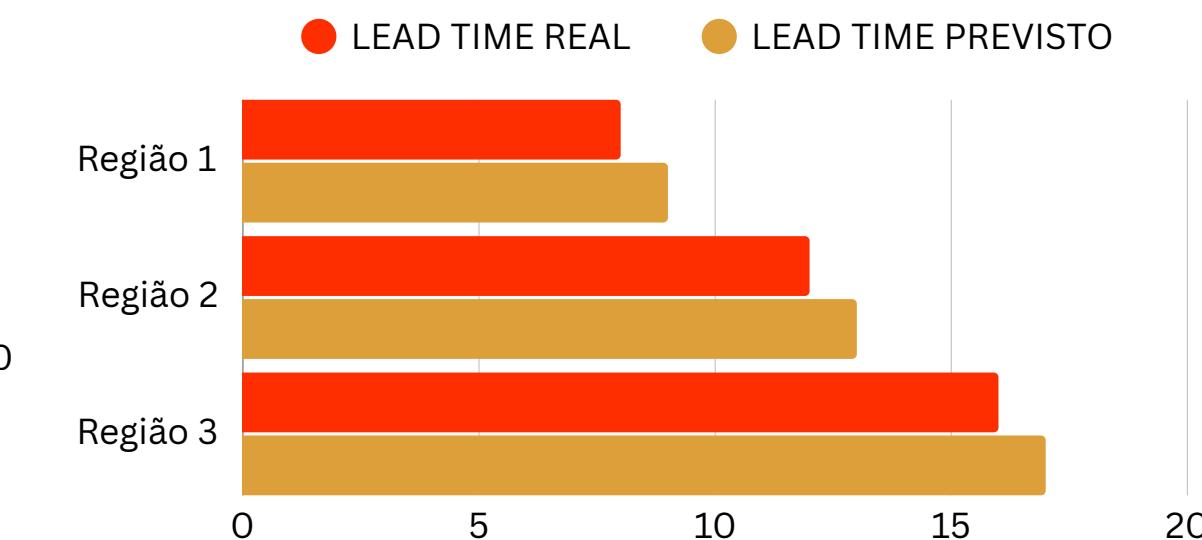
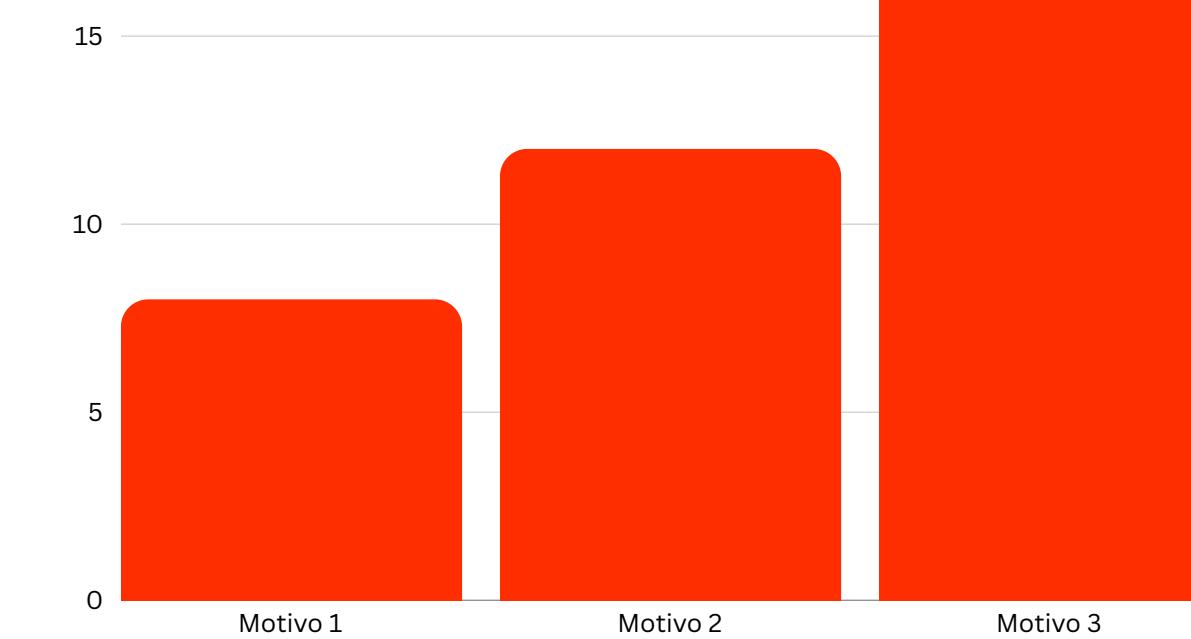
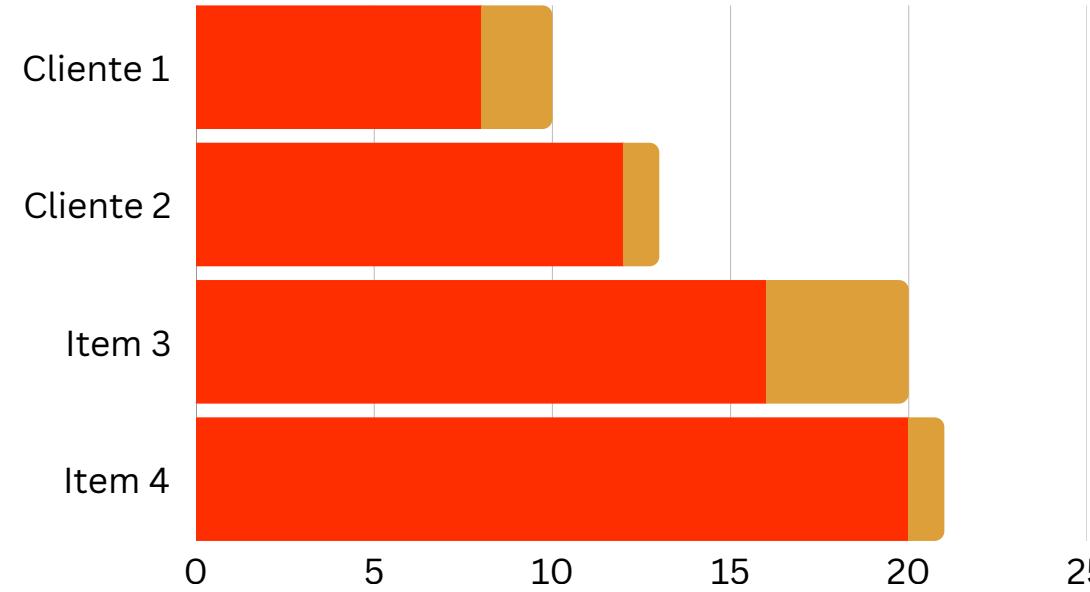
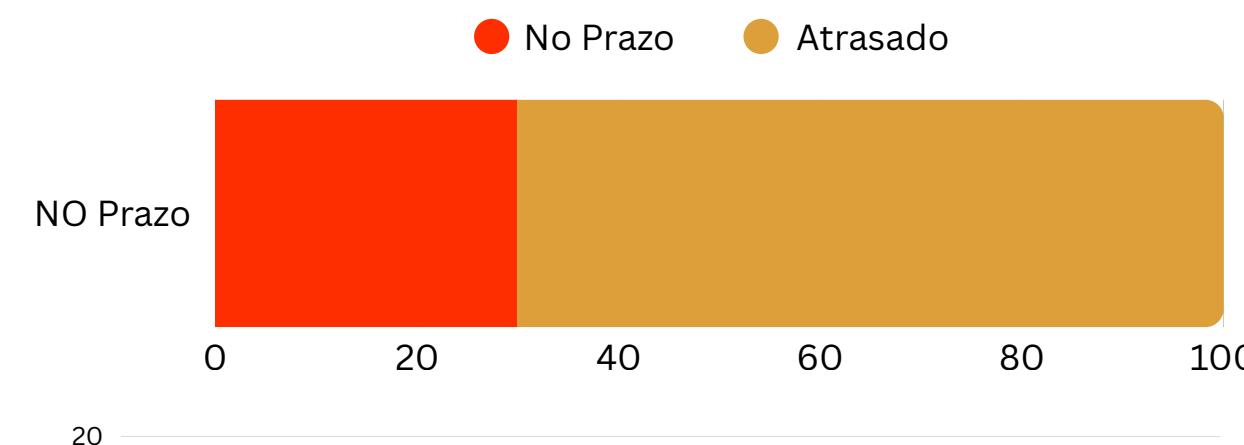
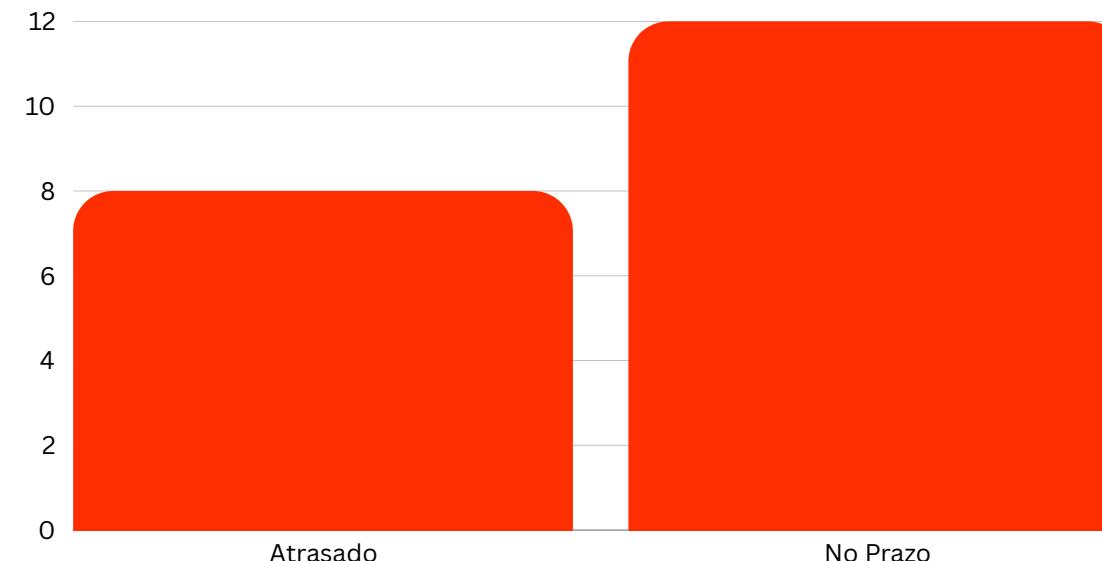
ITENS
0.000

ITENS
DEVOLUÇÃO
0.000

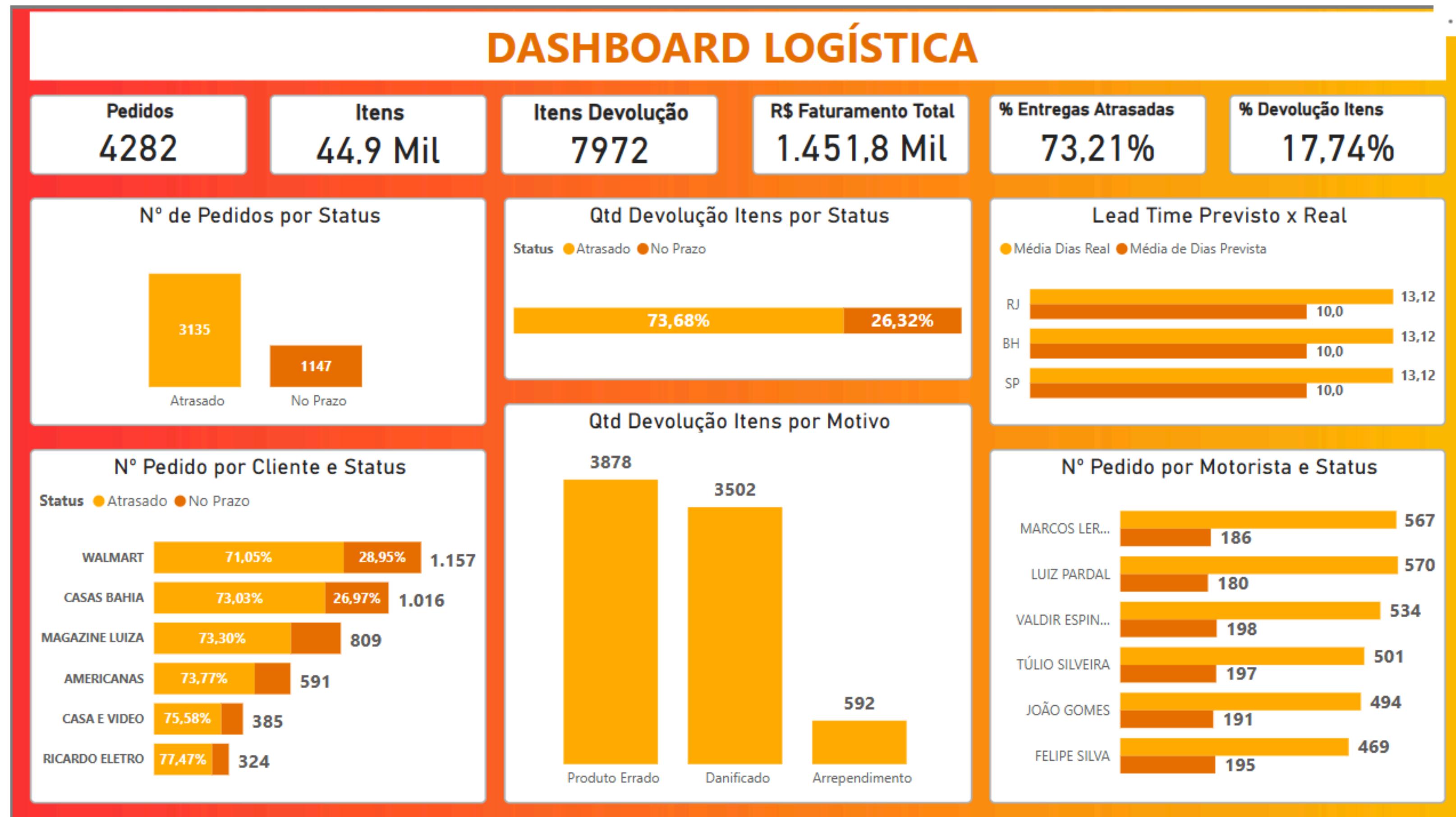
R\$ PEDIDOS
R\$ 0.000,00

ENTREGAS
ATRASADAS
0,00%

ITENS
DEVOLVIDOS
0,00%



POWER BI



DASHBOARD LOGISTICA

PEDIDOS
4.282

ITENS
44.931

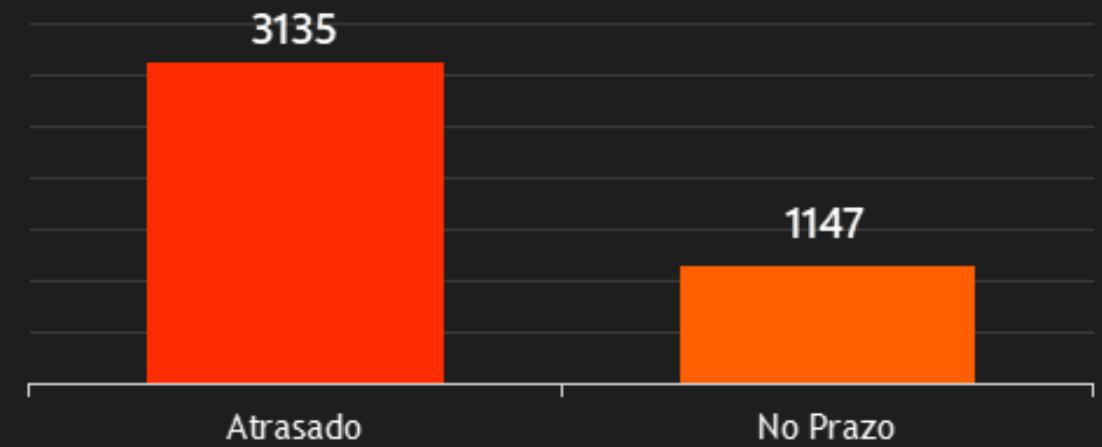
ITENS DEVOLUÇÃO
7.972

PEDIDOS
R\$ 1.451.762,

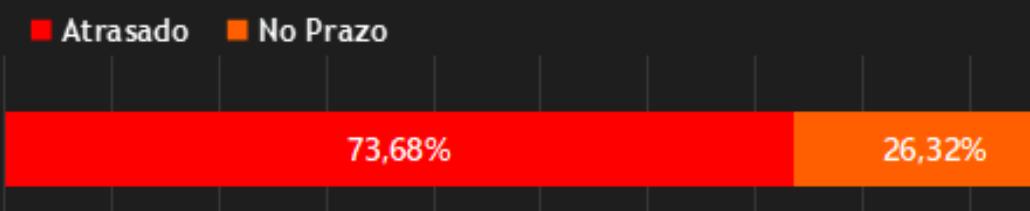
ENTREGAS ATRASADAS
73,21%

ITENS DEVOLVIDOS
17,74%

Nº de Pedido po Status



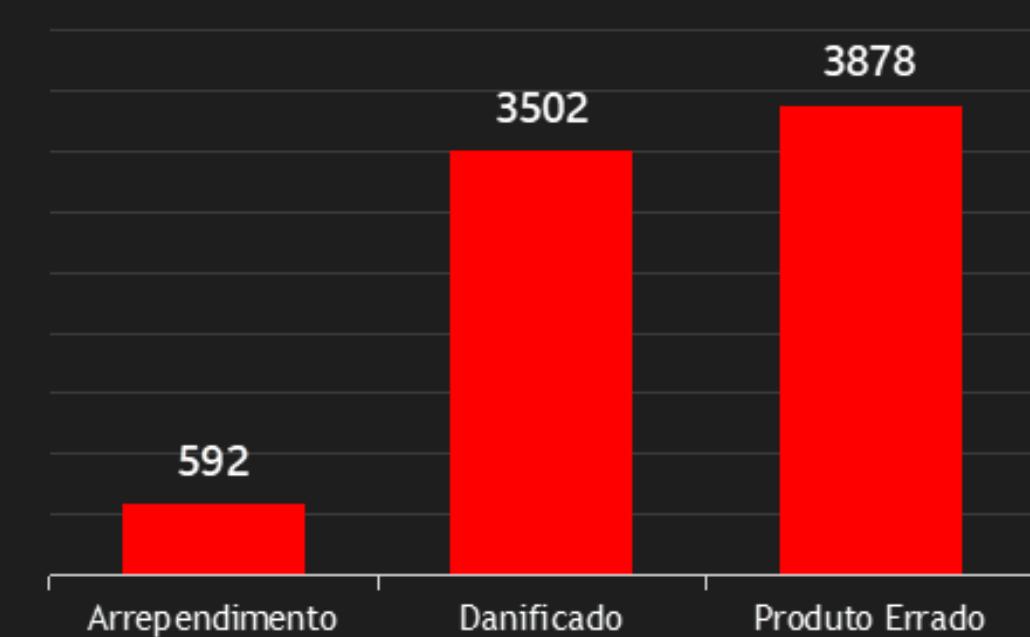
Devolução de Itens por Status



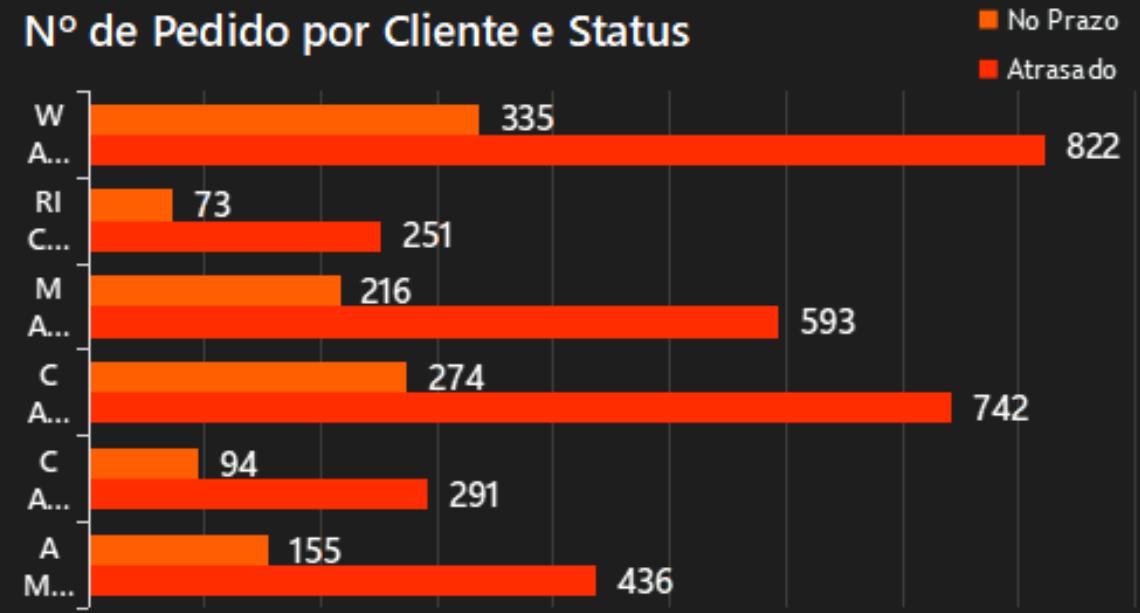
Lead Time Previsto x Real



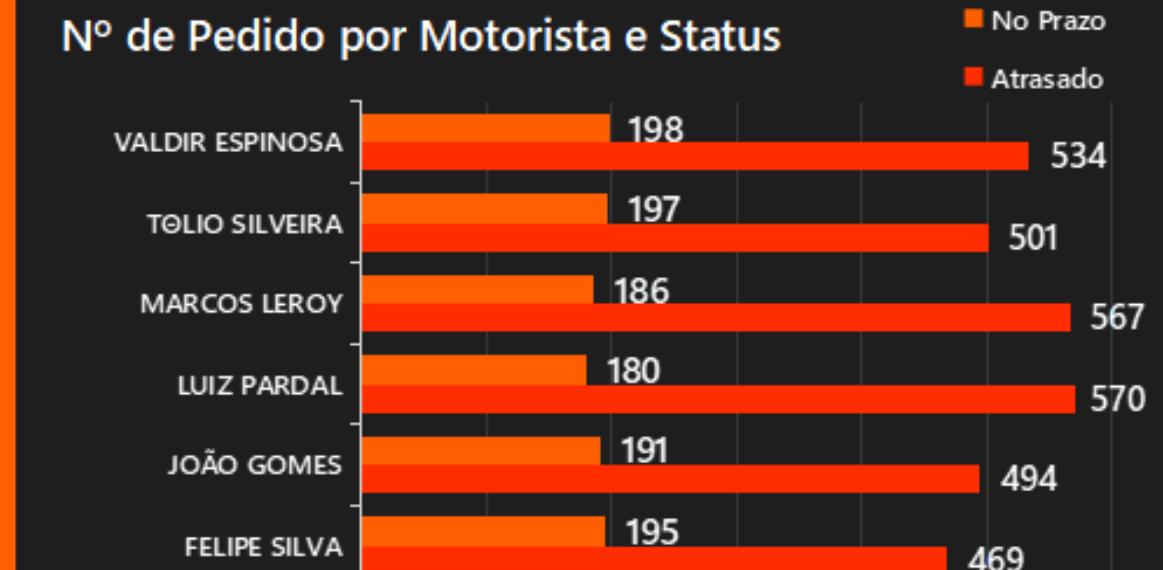
Quantidade de Devolução de Itens por Motivo



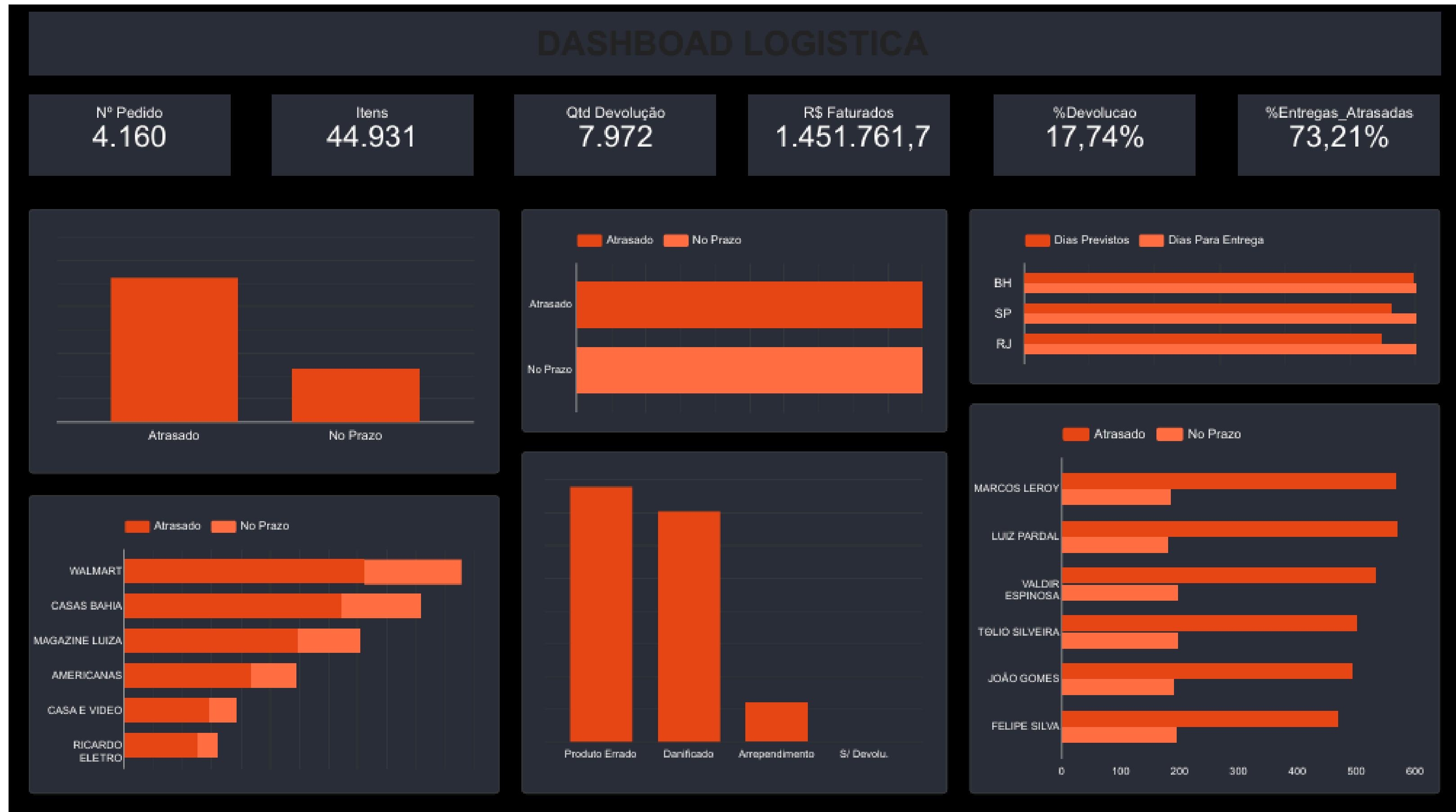
Nº de Pedido por Cliente e Status



Nº de Pedido por Motorista e Status



POWER BI



PYTHON (PLOTLY+STREAMLIT)





RESULTADOS & AGRADECIMENTOS



No geral, todas as plataformas testadas — **Python, Power BI, Looker Studio e Excel** — entregaram **bons resultados** na construção do dashboard proposto. Algumas se mostraram mais rápidas e intuitivas de utilizar do que outras, mas, com a concepção do layout e dos indicadores já definida previamente, as dificuldades encontradas estiveram mais relacionadas ao **nível de familiaridade** com cada ferramenta do que a limitações técnicas propriamente ditas. Embora o Python ofereça um nível de customização muito superior em comparação com as demais plataformas, essa vantagem não teve um impacto significativo neste projeto específico. **Ao final, o fator determinante para a fluidez da execução foi a experiência do usuário com cada plataforma, e não necessariamente suas capacidades individuais.**