Objetivos da análise:

Analisar os fatores que influenciam o atraso nas entregas, identificando padrões e variáveis associadas, e construir modelos para prever atrasos, com o intuito de melhorar o desempenho logístico e a satisfação do cliente.

1. **Qual a taxa de atraso nas entregas?**

Quantos pedidos foram entregues fora do prazo? (Comparar dt\_entrega\_pedido com dt\_previsao\_entrega\_cliente e usar tp\_performance\_entrega)

1. **Qual o perfil dos pedidos com atraso?**

Em quais cidades, estados e tipos de praça (cidade\_destinatario, uf, tp\_praca) ocorrem mais atrasos?

Quais unidades de negócio (des\_unidade\_negocio) e centros de distribuição (des\_cd\_origem) têm mais atrasos?

1. **Como o tempo até a entrega (qtd\_dias\_tat) varia entre entregas dentro e fora do prazo?**
2. **Existe padrão nos horários de despacho e entrega nos pedidos atrasados?**
3. Horários de saída do centro de distribuição (hr\_despacho\_pedido) e horário de entrega (hr\_entrega\_pedido) influenciam?
4. **Ocorrências relacionadas a atraso**
5. Existe associação entre flg\_existem\_ocorrencias (ocorrências de entrega) e atrasos?
6. **Evolução temporal**
7. Os atrasos estão aumentando ou diminuindo ao longo do tempo? (Usar dt\_criacao e dt\_entrega\_pedido)
8. **Impacto do pagamento**
9. O tempo entre criação e pagamento (dt\_criacao vs dt\_pagamento\_pedido) influencia no atraso?

**Análise Descritiva**

* 1. Frequências e percentuais de atrasos por categoria (cidade, estado, praça, unidade de negócio).
  2. Médias e medianas de qtd\_dias\_tat por grupo.

1. **Comparações entre grupos**
   1. Testes estatísticos para comparar atrasos entre categorias, como:
      1. Teste Qui-quadrado para associação entre atraso (sim/não) e variáveis categóricas (tp\_praca, uf, flg\_existem\_ocorrencias).
      2. Teste t ou Mann-Whitney para comparar média de dias de entrega (qtd\_dias\_tat) entre grupos.
2. **Análise temporal**
   1. Séries temporais para verificar tendências de atraso ao longo do tempo (dt\_criacao ou dt\_entrega\_pedido).
   2. Gráficos de linha ou barras para visualizar evolução mensal.
3. **Modelagem preditiva**
   1. Regressão logística para prever probabilidade de atraso com base nas variáveis disponíveis.
   2. Regressão linear para modelar os dias até a entrega (qtd\_dias\_tat) como variável contínua.
4. **Análise de correlação**
   1. Entre dias para entrega e tempo entre criação e pagamento.

Teste de Hipótese

Para verificar se diferenças observadas entre grupos (por exemplo, entregas atrasadas vs entregas no prazo) são estatisticamente significativas e não ocorreram por acaso.

1. estar se a proporção de atrasos é diferente entre as regiões (Capital vs Interior vs Região Metropolitana).
2. Verificar se o número médio de dias até a entrega (qtd\_dias\_tat) difere entre unidades de negócio ou centros de distribuição.
3. Testar se a existência de ocorrências (flg\_existem\_ocorrencias) está associada a atrasos.

**Quais testes usar?**

* **Qui-quadrado** para associações entre variáveis categóricas (exemplo: atraso sim/não vs região).
* **Teste t** ou **Mann-Whitney** para comparar médias de variáveis contínuas (exemplo: média de dias para entrega entre atrasados e no prazo).

**Análise de Regressão**

**Por que usar?**  
Para entender e quantificar o impacto de múltiplas variáveis ao mesmo tempo na probabilidade de atraso ou no tempo até a entrega. Além disso, a regressão ajuda a fazer previsões e identificar os principais fatores que influenciam o atraso.

**Tipos de regressão possíveis:**

1. **Regressão logística** (variável dependente binária: atraso sim/não)
   * Avalia como variáveis como cidade, estado, tipo de praça, ocorrências, horários influenciam a chance do pedido atrasar.
   * Exemplo: qual o impacto de ser na Capital vs Interior na probabilidade do atraso?
2. **Regressão linear** (variável dependente contínua: dias até entrega)
   * Estima como os fatores influenciam o tempo em dias para a entrega.
   * Exemplo: como a hora do despacho ou a unidade de negócio afetam o tempo até a entrega?

* Use **teste de hipótese** para investigar diferenças e associações específicas entre grupos.
* Use **regressão** para modelar e entender o efeito conjunto de múltiplas variáveis sobre o atraso.