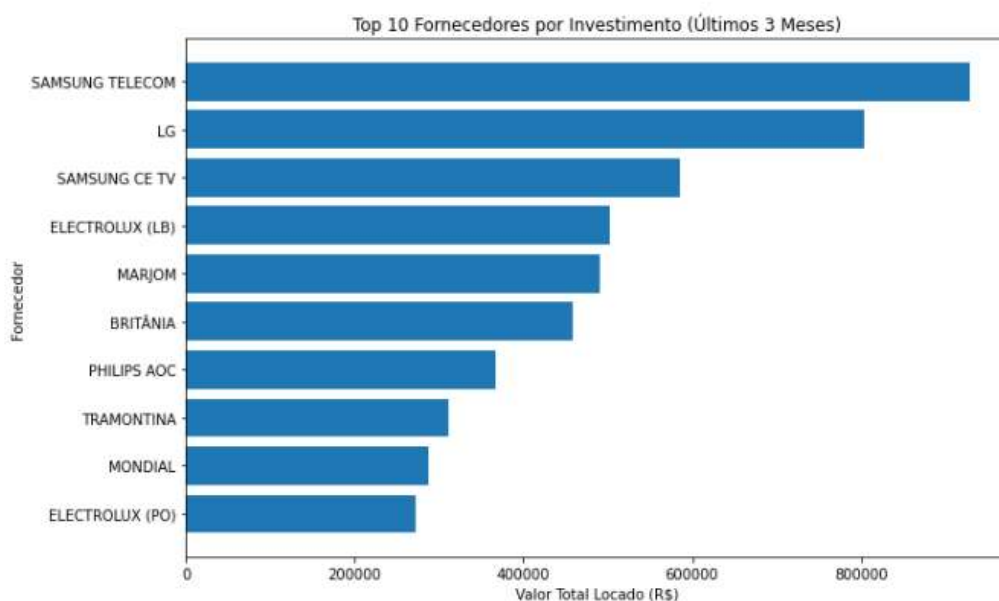




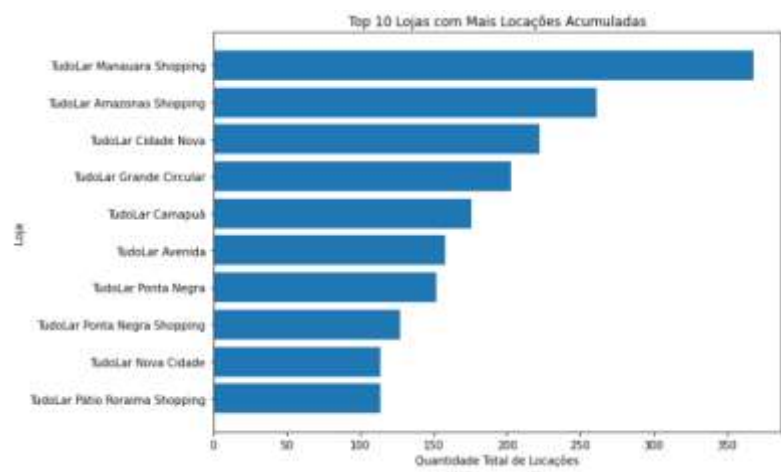
Análises do Problema Proposto e Autoavaliação MVP Engenharia de Dados

Pergunta 1: Quais fornecedores mais investem em espaços nas lojas?



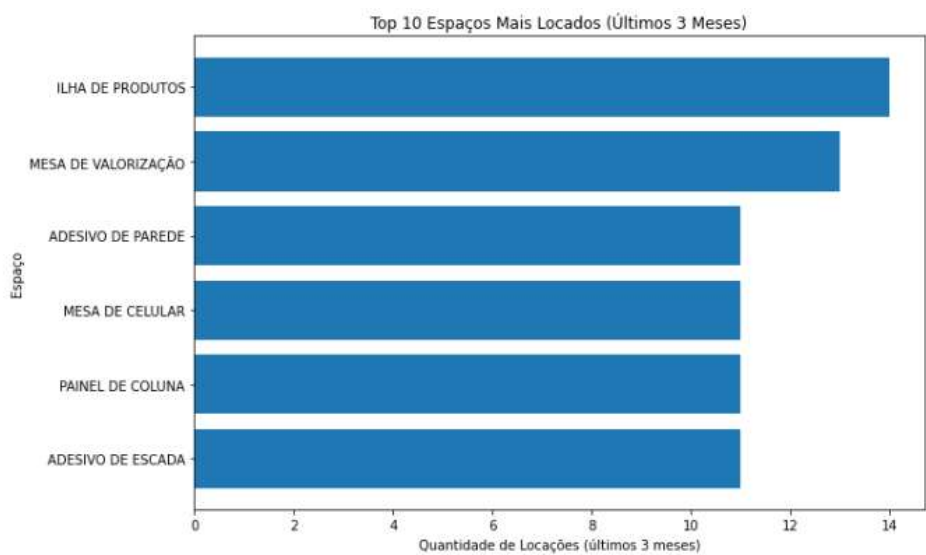
O gráfico mostra que SAMSUNG TELECOM, LG e SAMSUNG CE TV lideram os investimentos em locações, com destaque para o setor de eletroeletrônicos. Isso indica forte presença dessas marcas nas lojas e foco estratégico em visibilidade no ponto de venda.

Pergunta 2: Quais lojas concentram mais locações ao longo do tempo?



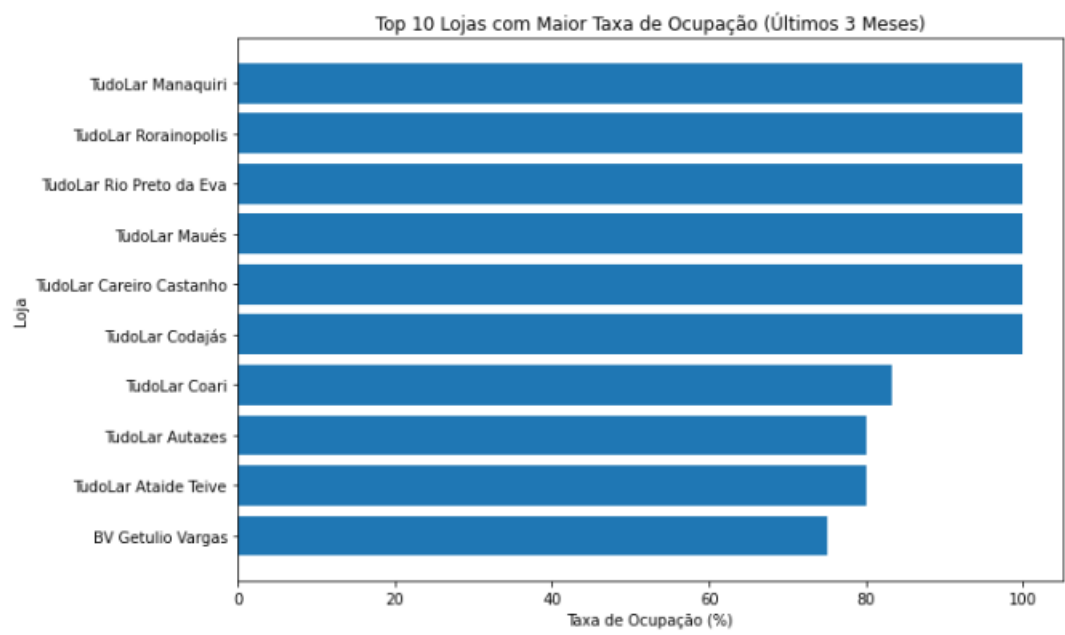
As lojas dos shoppings Manauara e Amazonas lideram em número de locações acumuladas, indicando maior atratividade para os fornecedores. Isso pode estar ligado ao alto fluxo de clientes ou melhor performance nas ativações realizadas.

Pergunta 3: Quais espaços estão sendo mais locados nos últimos 3 meses?



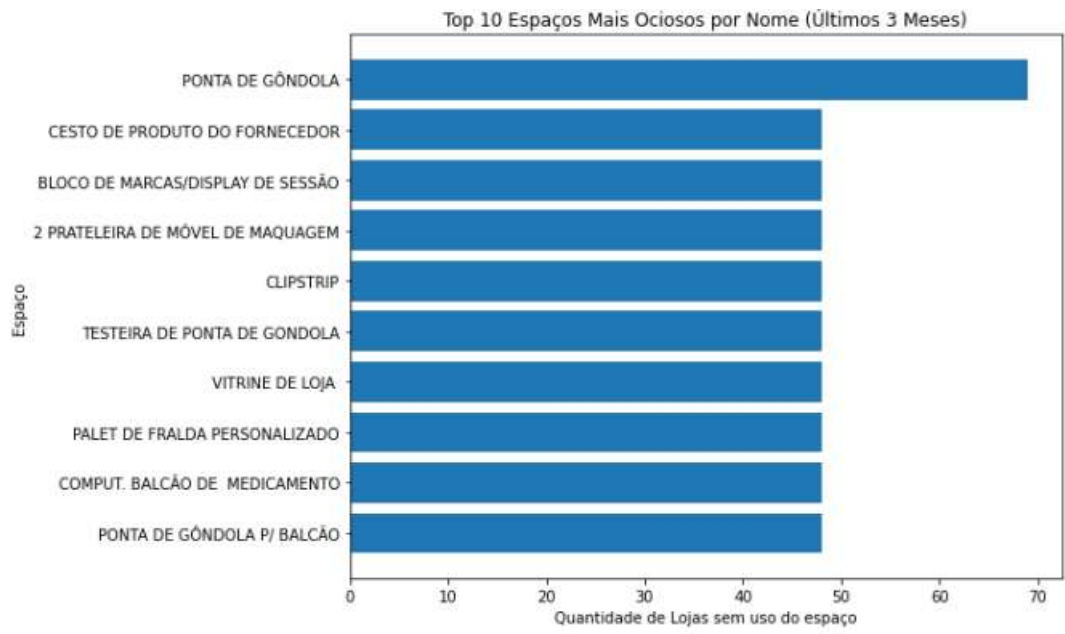
A **Ilha de Produtos** e a **Mesa de Valorização** aparecem como os espaços mais locados recentemente, reforçando sua importância como pontos estratégicos de exposição. Adesivos e mesas temáticas também mostram alta demanda nas ativações.

Pergunta 4: Qual a taxa de ocupação dos espaços por loja?



As lojas do interior, como **Manaquiri** e **Rorainópolis**, apresentam taxa de ocupação próxima de 100%, indicando excelente aproveitamento dos espaços disponíveis. Esses dados destacam oportunidades de expansão de ativações mesmo fora dos grandes centros.

Pergunta 5: Quais espaços estão ociosos nos últimos 3 meses?



A **Ponta de Gôndola** lidera em ociosidade, com ausência de uso em mais de 70 lojas. Outros espaços como **cestos promocionais** e **clipstrips** também aparecem com pouca ocupação, sinalizando a necessidade de revisão na estratégia de exposição ou incentivo para uso desses ativos.

AUTOAVALIAÇÃO

O MVP proposto teve como objetivo desenvolver uma estrutura analítica para gestão e análise de espaços publicitários locados dentro de lojas físicas, com base em dados reais (não sensíveis) do meu ambiente de trabalho.

O objetivo foi plenamente atingido. Consegui estruturar um pipeline completo em camadas (Bronze, Silver e Gold), utilizando o modelo estrela, com tabelas fato e dimensão bem definidas. A partir disso, criei visões analíticas que responderam com clareza às principais perguntas de negócio levantadas no início do projeto.

Entre as principais dificuldades enfrentadas, destaco:

- Problemas de schema nos dados brutos, exigindo atenção com tipos e padronização;
- Conversão de datas em formato numérico (Excel) para `date`, o que exigiu pesquisa e testes;
- Garantir que todas as colunas obrigatórias fossem tratadas corretamente para evitar falhas nas camadas seguintes.

Como oportunidade, esse trabalho me permitiu aprofundar o entendimento de boas práticas em engenharia de dados, especialmente no uso do Databricks com PySpark e Delta Lake, além de aplicar conceitos reais de modelagem dimensional voltada para análise.

O projeto me deixou mais seguro quanto à construção de pipelines robustos e à apresentação clara de resultados de negócio com base em dados confiáveis.