

RELATÓRIO TÉCNICO – CASE DE VEÍCULOS E CRÉDITO

Analista de Dados: Julianne Lam.

Este desafio teve início na busca por dados públicos. Assim, foram utilizados dados disponibilizados pelo governo Federal, no Portal de Dados Abertos (<https://dados.gov.br/dados/conjuntos-dados>).

Foram priorizados os dados que tinham alguma correlação com veículos e crédito, elucidando alguns pontos sobre crédito, prazo de quitação e inadimplência, além de dados que oferecem um contexto sobre frota (base de 2020).

Os dados baixados eram tabelas que tinham de 2 a 3 colunas. Então, foi feita uma organização: informações segundo o tipo de pessoa (física ou jurídica) foram unificadas conforme o tipo de transação de crédito (para aquisição ou arrendamento de veículo). Ao unir as citadas tabelas, foi acrescentada também coluna de código (para que tivessem chave primária). Além disso, também foi feita uma organização da coluna de data – a partir da separação da coluna de Ano_dia_mês, criando-se coluna só de Anos, depois uma para mês e, por fim, outra com mês e ano.

A base sobre frota não foi categorizada com as tabelas de pessoas porque tratava de temas diferentes. Assim, não passou por união de tabelas.

Os dados foram inicialmente tratados e organizados utilizando-se o Microsoft SQL Server (MSSQLS). Depois, as tabelas selecionadas foram conectadas ao Power BI e foram feitos gráficos. Vale destacar que, em cenários que exigem escalabilidade, é recomendável realizar o ELT na nuvem, permitindo maior flexibilidade. Assim, seria possível extrair todos os dados de uma vez e modelar apenas o que fosse realmente útil para a área de negócio, ganhando agilidade na entrega de informações. No entanto, este desafio foi realizado em um cenário factível à pessoa física com notebook pessoal e, por isso, algumas adaptações foram necessárias.

Algumas tabelas sofreram modificação de nome durante o trabalho no MSSQLS e o código carrega esse legado em alguns trechos. Cogitei a mudança em revisão, mas entendi que poderia afetar a linearidade do entendimento e não seria necessário para o comando (que já tinha sido executado).

Antes de iniciar a transformação dos dados, verifiquei quais eram as colunas e se a importação tinha sido bem sucedida. Em alguns casos foi necessário fazer uma etapa de tratamento antes da importação, evitando problemas com ponto, vírgula e aspas, por exemplo. Foram algumas substituições simples, às vezes feitas no próprio bloco de notas.

Depois da importação, ainda na verificação inicial, observei a quantidade de linhas para entender de que maneira poderia organizar as tabelas unindo-as umas às outras segundo categorias que tinham em comum e que poderiam facilitar para a análise. As colunas intituladas “data” foram renomeadas para não carregar uma palavra reservada (o que poderia gerar vários problemas ou dificuldades evitáveis no código).

Assim, reorganizei e deixei as tabelas prontas para serem usadas no Power BI, na construção do dashboard. Isso incluiu uma verificação de valores nulos – que não foram encontrados. No mais, vale mencionar que os dashboards priorizaram o público que parecia ser mais rentável.

O dashboard traz quantidades de carro por estado e por tipo de veículo em território nacional, além de comparação entre pessoa física e jurídica nas modalidades de aquisição e arrendamento de veículos. Há dados sobre crédito, custo de aquisição e inadimplência.

Foi feita uma comparação entre o crédito disponível para concessão (nos relatórios de instituições financeiras constavam como crédito de modo generalizado, não apenas focado para veículos) e a inadimplência que afetaria o retorno do crédito concedido. Assim, ainda que não haja detalhes sobre os ganhos dessas instituições ou sobre o saldo final, o gráfico serve como orientação geral para instituições que trabalhem ofertando crédito, permitindo uma visualização do que se estima como perda, por exemplo.

Observou-se que as operações de crédito concedidas a arrendamento são significativamente mais baixas que as operações de crédito para aquisição de veículos. Isso parece coerente, pois o leasing em si já é uma alternativa de menor impacto contábil-financeiro e, em tese, não demandaria de empréstimos vultuosos.

Há ainda contagem de veículos por estado e por tipo, para se ter ideia de como estava a preferência e o consumo de veículos no ano estudado e a quantidade por localidade. Os gráficos de frota estão nos primeiros gráficos de barras à esquerda e sua interpretação é bem simples.

Observação: Não foram encontrados dados de 2025 disponíveis em arquivos viáveis, apenas danificados. Assim, entende-se que esses dados de 2020 servem para compreensão de um cenário pré-pandêmico (um marco significativo) e que pode servir de métrica para comparar com o cenário atual. Nesse caso, o processo de análise se daria da seguinte forma: é feita uma comparação entre os dados do mercado (frota nacional) e os dados históricos da empresa no mesmo período (para encontrar padrões, tendências e pontos a serem monitorados). Depois, observa-se como a empresa passou pelo período de maior crise e como se encontra agora. Outras possibilidades de encontrar dados para enriquecer as análises são: contratação de consultoria ou empresa de pesquisa, entrar em contato com os órgãos públicos solicitando dados em arquivos viáveis ou aguardando a publicação de novos dados.

Para separar visualmente os dados de pessoa física e jurídica, foram utilizadas cores diferentes para cada categoria. Azul para pessoa física e roxo para pessoa jurídica. Os gráficos gerais (sobre frota de veículos no país), estão com título em verde, pois não se referem especificamente a um tipo de pessoa.