

RELATÓRIO DA ANÁLISE DE DADOS

1. DESCRIÇÃO DO CASO

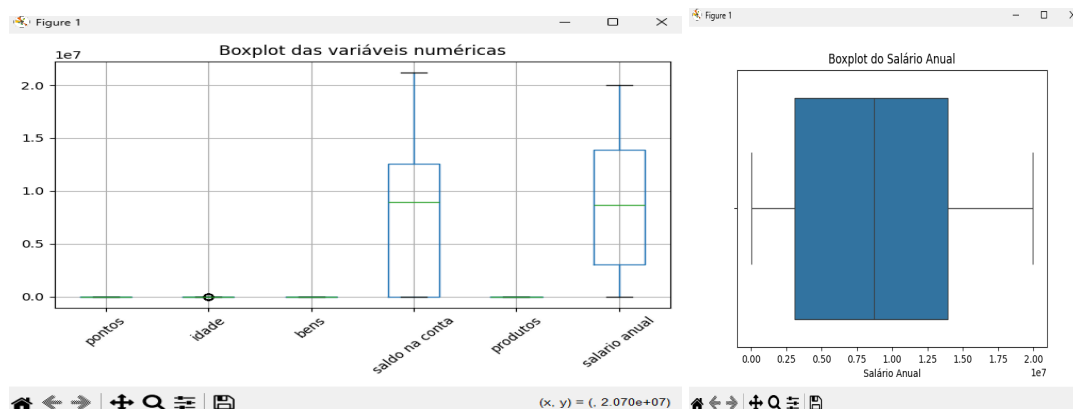
O caso a seguir refere-se a uma análise realizada por Davilson Carvalho, aluno da trilha de Ciência de Dados, que obteve informações sobre os dados por meio de uma tabela no [Google Sheets](#). A presente análise tem como objetivo investigar os principais fatores relacionados à evasão (churn) de clientes em uma base de dados bancária. Os dados utilizados referem-se a uma amostra de clientes contendo informações pessoais, financeiras e comportamentais, incluindo gênero, idade, estado, saldo bancário, bens, produtos e se o cliente saiu ou permaneceu no banco.

2. TRATAMENTO E ANÁLISE

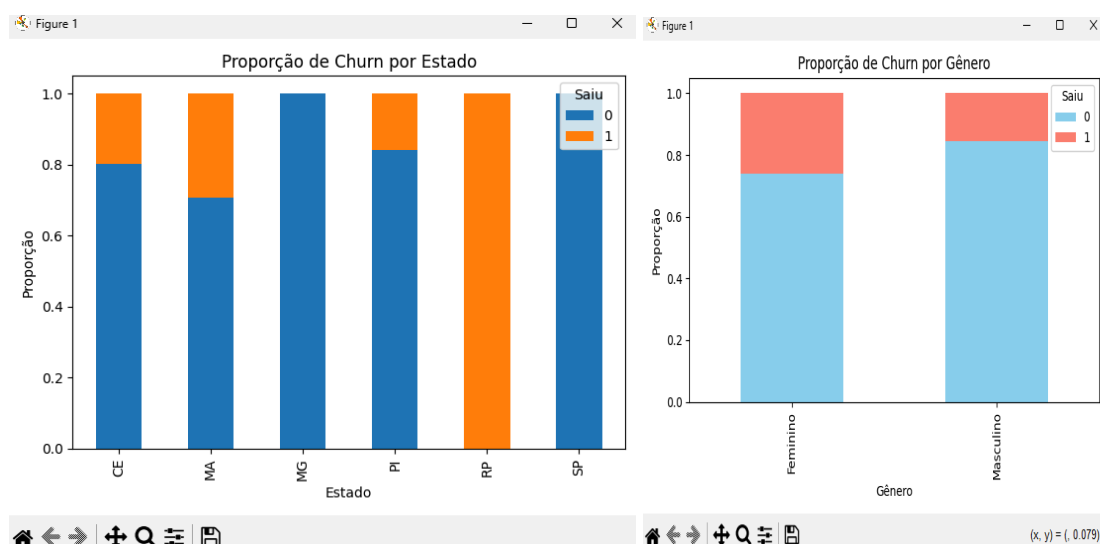
Foi utilizado o Vscod e python para o tratamento e análise dos dados. A primeira etapa consistiu na leitura e tratamento dos dados. As colunas foram padronizadas para letras minúsculas e corrigimos inconsistências nos dados categóricos. Por exemplo, valores como “Mas” e “M” foram unificados como “Masculino”, e “Fem” e “F” como “Feminino”. Também foi tratado valores ausentes nas colunas “gênero” e “salário anual”, substituindo-os por moda e mediana, respectivamente, estratégias eficientes para evitar distorções. Além disso, removi duplicatas

Uma análise descritiva mostrou que a idade média dos clientes é de aproximadamente 43,98 anos. Quando observamos o saldo na conta dos clientes com menos de 40 anos, verificamos uma média de R\$ 6.914.457,68 e uma mediana de R\$ 8.194.148,00, revelando uma distribuição levemente assimétrica. Já entre os clientes com mais de 40 anos, a média do saldo é ainda maior, atingindo R\$ 7.665.165,65, e a mediana chega a R\$ 10.003.347,00. Esses dados reforçam a percepção de que, em geral, clientes mais velhos tendem a acumular mais recursos em suas contas bancárias, o que pode refletir maior estabilidade financeira ou tempo de fidelização ao banco. Essa informação é relevante para estratégias que visem a personalização de produtos financeiros e a fidelização de diferentes perfis etários.

Para visualização, utilizei gráficos boxplots para verificar a distribuição de variáveis como salário, idade e saldo na conta. O boxplot é um gráfico estatístico que exibe a mediana, quartis e possíveis outliers de uma variável, facilitando a identificação de assimetrias e dispersão nos dados.



Também produzi gráficos de barras empilhadas que ilustraram a proporção de churn por gênero e por estado. Estes gráficos deixaram claro que o estado tem influência significativa na evasão, sendo alguns estados mais críticos que outros. O mesmo se aplica ao gênero, embora a diferença não seja tão acentuada.



Com base nesses insights, recomenda-se que a instituição financeira investigue com mais profundidade os motivos pelos quais os clientes do sexo feminino e de determinados estados estão saindo. Estratégias como campanhas personalizadas, maior suporte digital ou reformulação do atendimento presencial podem ser eficazes para reduzir a evasão. Além disso, é importante considerar ações voltadas aos clientes mais jovens, que apresentaram menores saldos em conta, podendo estar mais propensos a migrarem para instituições com melhores vantagens iniciais.