# **Desafio**

Neste desafio você vai gerar um relatório com indicadores de performance para responder às demandas do BanVic. Para isso você deve criar indicadores de negócio e sugerir como a empresa pode agir em cima deles para melhorar seus resultados. Você pode usar as ferramentas que tiver maior facilidade ou achar mais interessante tais como Excel, Google Sheets, Python, R, dbt (data build tool) etc.  Lembre que este relatório será apresentado para a CEO, então você deve formatá-lo de forma profissional e objetiva.

# **Entregas**

**1. Nem todos os envolvidos estão convictos com a ideia de uma frente de BIs no BanVic. Como podemos convencer Camila Diniz, diretora comercial,  que as nossas soluções serão úteis para a empresa? Apresente sua análise exploratória realizada a partir dos dados apresentados e cite quais perguntas de negócio podemos responder a fim de mostrar o valor que dados possuem para a empresa, levando em consideração o desafio apresentado. Pergunta de nível teórico. O seu projeto final não necessariamente precisa trazer os mesmos ou todas as análises/perguntas de negócio citados nessa entrega.**

Minha análise exploratória revelou as seguintes características das tabelas:

1. Agências (agencias.csv)

Dimensão: 10 linhas × 7 colunas.

Colunas: Código da agência, nome, endereço, cidade, estado, data de abertura e tipo de agência.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

2. Clientes (clientes.csv)

Dimensão: 998 linhas × 10 colunas.

Colunas: código do cliente, primeiro nome, último nome, e-mail, tipo cliente, data inclusão, cpfcnpj, data nascimento, endereço, cep.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

3. Relacionamento entre colaboradores e agências (colaborador\_agencia.csv)

Dimensão: 100 linhas × 2 colunas.

Colunas: código do colaborador e código da agência.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

4. Colaboradores (colaboradores.csv)

Dimensão: 100 linhas × 8 colunas.

Colunas: código do colaborador, primeiro nome, ultimo nome, email, cpf, data de nascimento, endereço, cep.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

5. Contas (contas.csv)

Dimensão: 999 linhas × 9 colunas.

Colunas: número da conta, código do cliente, código da agência, código do colaborador, tipo da conta, data da abertura, saldo total, saldo disponível, data do último lançamento.

Valores ausentes: Nenhum.

Formato dos dados Todos estão formatados como texto

6. Propostas de Crédito (propostas\_credito.csv)

Dimensão: 2.000 linhas × 12 colunas.

Colunas: código da proposta, código do cliente, código do colaborador, data da entrada da proposta, taxa de juros mensal, valor da proposta, valor do financiamento, valor da entrada, valor da prestação, quantidade de parcelas, carência e status da proposta.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

7. Transações (transações.csv)

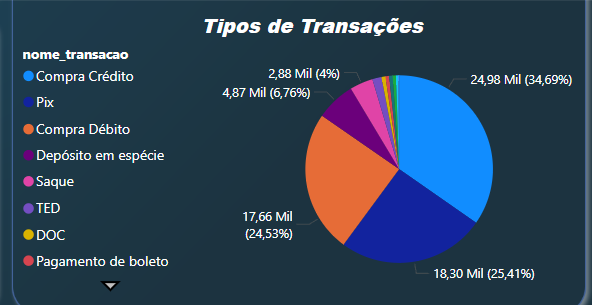
Dimensão: 71.999 linhas × 5 colunas.

Colunas: código da transação, número da conta, data da transação, nome da transação, valor da transação.

Valores ausentes: Nenhum.

Formato dos dados: Todos estão formatados como texto

**Propostas de perguntas de negócio**  
  
**Quais são os principais tipos de transações realizadas?**



De acordo com o gráfico de pizza os principais tipos de transações são:

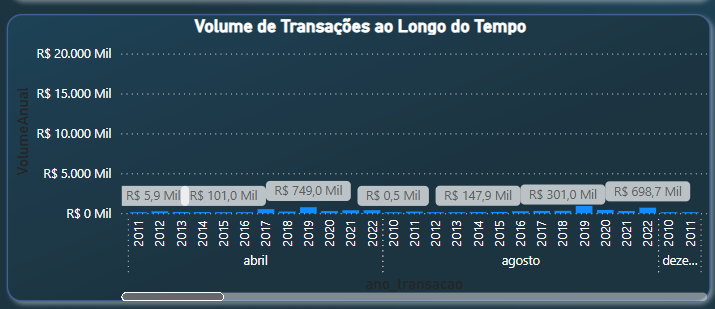
Compra Crédito: 24977 (34,69%).

Pix: 18298 (25,41%).

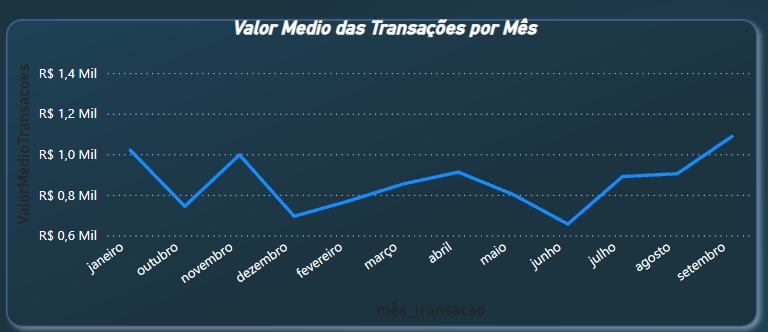
Compra Débito: 17662 (24,53%).

Saque: 4886 mil (6,76%).

**Existe sazonalidade no volume de transações ao longo do tempo (mensal)?**



De acordo com o gráfico de colunas agrupadas e linha os meses de Novembro e Dezembro tendem a movimentar um volume maior de transações.

**Quais são os meses que concentram os maiores valores médios de transações ?**

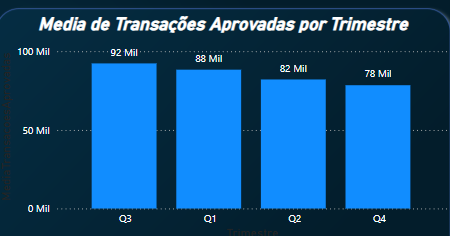
De acordo com o gráfico de linhas os meses são: Setembro, Janeiro e Novembro.

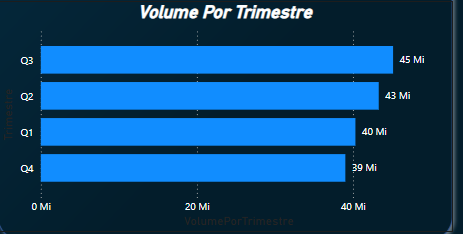
**2. Quais serão os indicadores ou análises utilizadas para responder às perguntas de negócio citadas anteriormente? Pergunta de nível teórico. O seu projeto final não necessariamente precisa trazer os mesmos ou todos os indicadores/análises citados nessa entrega.**

| **Pergunta** | **Indicador/Análise** |
| --- | --- |
| Quais são os principais tipos de transações realizadas? | Gráfico de pizza com a distribuição e participação de cada tipo de transação. |
| Existe sazonalidade no volume de transações ao longo do tempo (mensal)? | Gráfico de colunas agrupadas e linha com o volume total de transações mensalmente. |
| Quais são os meses que concentram os maiores valores médios de transações ? | Gráfico de linhas com o valor médio das transações por mês. |

**3. Uma dimensão de datas bem construída viabiliza a elaboração de análises mais robustas com relação ao tempo. O BanVic quer começar seu Data Warehouse (DW) com uma dim\_dates que atenda seus requisitos de análise. Utilizando uma dimensão de datas, responda às duas perguntas a seguir. Por fim, proponha outras duas análises utilizando a dim\_dates que possam ter valor de negócio para o BanVic. Pergunta prática e teórica.**

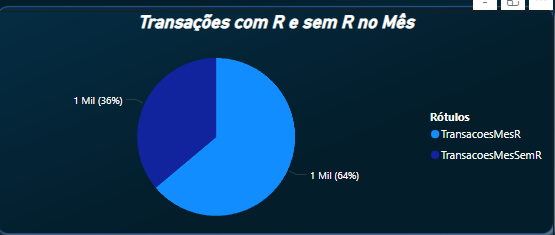
**Qual trimestre tem, em média, mais transações aprovadas e qual tem, também em média,  maior volume movimentado?**





De acordo com os gráficos apresentados, o trimestre que possui em média mais transações aprovadas é o Q3(Julho, Agosto, Setembro) e o trimestre que tem em média maior volume movimentado é também o Q3(Julho, Agosto, Setembro).

**Um analista sugeriu que, meses que contém R no seu nome (Ex: janeiRo), tem um padrão de quantidade de transações diferentes dos meses que não apresentam R no nome (Ex: junho). Apresente suas análises validando ou não essa afirmação.**



De acordo com as informações fornecidas pelo gráfico de pizza e pelos cartões há um indicativo de que meses com "R" apresentam um padrão maior de transações em relação a meses sem "R".

**4. André Tech, o especialista em tecnologia, solicitou que você buscasse fontes de dados públicas que poderiam ser utilizadas para enriquecer a base de dados do BanVic e ampliar as possibilidades de análise, considerando principalmente a necessidade atual apresentada. Pergunta prática e teórica.**

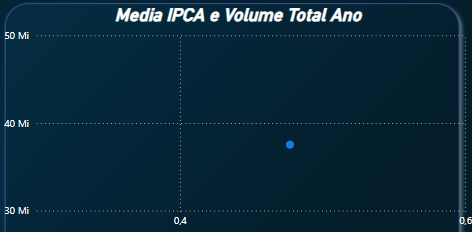
* **Traga uma proposta de possíveis dados públicos a serem utilizados no Data Warehouse apresentando o valor que eles podem trazer para o desafio apresentado.**

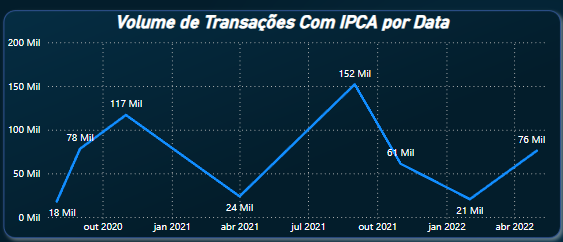
IPCA (Índice de Preços ao Consumidor Amplo)

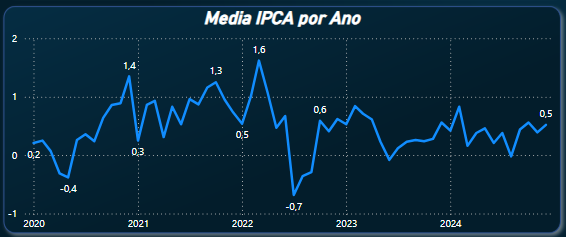
Fonte: SGS - Sistema Gerenciador de Séries Temporais

Valor para o negócio: O IPCA mede a inflação e pode ajudar a entender se a variação de preços impacta o número de transações e o volume movimentado.

Possível análise: Correlação entre variação do IPCA e volume de transações. Se a inflação sobe, o número de transações cai? Ou o valor médio das transações aumenta?

* **Um exemplo de dado externo é o Índice Nacional de Preços ao Consumidor Amplo. Considerando o IPCA no Brasil para o período apresentado na base de dados da BanVic, avalie se existe alguma relação do índice de preços com o volume de transações e/ou com o valor transacionado no período. Apresente suas análises e justifique sua opinião.**





A análise sugere que, quando a inflação (IPCA) aumenta, o volume de transações tende a cair, indicando uma relação inversa entre os dois.

* **Quais decisões poderiam ser facilitadas com o uso desses novos dados? Pergunta de nível teórico. O seu projeto final não necessariamente precisa trazer esses dados externos.**

A inclusão de novos dados, pode enriquecer as análises do BanVic e facilitar decisões estratégicas em várias áreas de negócio. Como a precificação, tendências, avaliação de riscos, previsões de lucro e receita, etc.

**5. Processos de Transformação e Tratamento de Dados**

Importação e Integração: Os dados foram coletados de múltiplas fontes, planilhas internas (transações, clientes, agências, propostas de crédito, colaboradores, colaborador\_agencia, contas) e fontes públicas (IPCA).

Limpeza de Dados: Foi feita a remoção de valores nulos, duplicados, e padronização de formatos (datas, números, textos).

Transformações: Criação de colunas calculadas para facilitar análises (média e quantidade de transações aprovadas, transações com e sem R no mês, volume por trimestre, volume total anual, etc ).

Conversão de valores para formatos apropriados (ex: transformação de números em texto para decimal e arredondamento para duas casas decimais).

Criação de uma dimensão de datas (dim\_dates) para análises temporais detalhadas.

Integração com Dados Públicos: Foi adicionada a tabela IPCA, associando-a ao volume de transações para análises.

Relacionamentos: Criação de relacionamentos entre tabelas de propostas\_credito e dimensões, entre outras.

**Conclusão Geral:**

O BanVic apresenta um desempenho sólido, com sazonalidade previsível em transações. Entretanto, os dados mostram vulnerabilidade em períodos de alta inflação, exigindo ações estratégicas para mitigar esses impactos.

**Recomendações:**

Intensificar campanhas promocionais no início do ano para equilibrar sazonalidades, criar incentivos para aumentar o volume de transações em períodos de baixa movimentação, implementar programas de fidelidade ou benefícios exclusivos para clientes de alta relevância, propor novos produtos ou serviços para clientes que realizam transferências de alto valor e ajustar taxas e tarifas com base na inflação, garantindo competitividade.

**6. Justificativa das Ferramentas Adotadas**

Power BI: Escolhido pela capacidade de criar dashboards interativos e visualizações dinâmicas além de Permitir fácil conexão entre tabelas e implementação de DAX para cálculos avançados.

Fontes de Dados Públicas: Adição do IPCA trouxe mais contexto às análises, enriquecendo os insights.

Modelo Dimensional: Estruturado com tabelas de fatos e dimensões (dim\_dates) para análises mais eficientes e detalhadas.