

# Relatório de Análise de Indicadores de Performance – BanVic

Autor: Odon Henrique Pacheco Rocha Cedrim

## 1. Introdução

Este relatório tem como objetivo demonstrar como a implementação de uma solução de Business Intelligence (BI) pode agregar valor ao BanVic. Utilizando os dados fornecidos (agencias.csv, clientes.csv, colaborador\_agencia.csv, colaboradores.csv, contas.csv, propostas\_credito.csv e transacoes.csv) e dados externos (IPCA, SELIC, ICC), foram desenvolvidos indicadores que suportam decisões estratégicas e operacionais. O trabalho foi desenvolvido em duas instancias, a primeira utilizando Python e bibliotecas como pandas, matplotlib, seaborn e scikit-learn, e a segunda criando o dashboard interativo no Power BI. Sendo assim, é possível dividir o projeto em quatro etapas principais:

**Tratamento e transformação dos dados (tratamento\_dados.py);**

**Identificação e remoção de inconsistências (inconsistencias.py);**

**Análise exploratória e geração de indicadores (analise\_dados.py);**

**Dashboard interativo (LH\_EA\_AD\_ODONHENRIQUE.pbit).**

## 2. Metodologia

### 2.1. Ferramentas Utilizadas

**Python:** Linguagem escolhida para realizar todo o processo de ETL, sendo possível desenvolver scripts robustos de manipulação, transformação e integração de dados.

**Pandas:** Manipulação e transformação dos datasets.

**Seaborn e Matplotlib:** Criação de gráficos e visualizações.

**APIs para dados externos:** Responsáveis pela obtenção dos dados externos (IPCA, SELIC, ICC).

**Scikit-learn:** Normalização dos dados para facilitar a comparação entre séries temporais.

**Power BI:** Ferramenta escolhida para a criação de dashboards interativos e visualizações dinâmicas.

## 2.2. Organização dos Scripts

### **tratamento\_dados.py:**

- Leitura dos arquivos originais.
- Aplicação de funções para padronizar CEP, calcular idade, extrair ano/mês das datas e categorizar transações.
- Salvamento dos dados processados para serem utilizados na próxima etapa.

### **inconsistencias.py:**

- Validação dos dados a partir de colunas em comum entre diferentes arquivos.
- Separação dos registros consistentes e dos que apresentam inconsistências, garantindo uma base de dados confiável para as análises.

### **analise\_dados.py:**

- Conversão e padronização dos formatos de data.
- Análises exploratórias com gráficos que demonstram a evolução de transações, propostas de crédito, contas, clientes e colaboradores.
- Criação de uma dimensão de datas (dim\_dates) para análises temporais detalhadas.
- Integração com dados externos (IPCA, SELIC, ICC) para avaliar a influência de indicadores macroeconômicos no volume e na quantidade de transações.

## 3. Processamento e Tratamento dos Dados

### 3.1. Limpeza e Transformação (tratamento\_dados.py)

#### **Padronização de Campos:**

- CEP: O CEP foi formatado para o padrão “XXXXX-XXX” através da função **format\_cep** e extraído de endereços completos através da função **extrair\_cep\_endereco**.
- Idade: A idade dos clientes e colaboradores foi calculada de forma precisa usando a data de nascimento através da função **calcular\_idade**.

#### **Extração e Transformação de Datas:**

- Extração do ano e mês de datas importantes através da função **extrair\_ano\_mes**, para facilitar análises temporais.

#### **Checação de valores nulos:**

- Exibe o número de valores nulos em cada coluna de um DataFrame através da função **checar\_valores\_nulos**.

### Categorização das Transações:

- As transações foram classificadas em “Entrada”, “Saída” ou “Outro” com base em um dicionário de mapeamento.
- A simplificação dos nomes das transações permitiu uma análise mais clara dos tipos de operações realizadas.

## 3.2. Remoção de Inconsistências (inconsistencias.py)

Foram aplicadas regras de validação para:

- Garantir que contas possuam referências válidas a clientes, agências e colaboradores.
- Validar que propostas de crédito estejam associadas a clientes e colaboradores existentes.
- As transações foram verificadas para assegurar que os números de conta sejam consistentes com a versão limpa da tabela de contas.

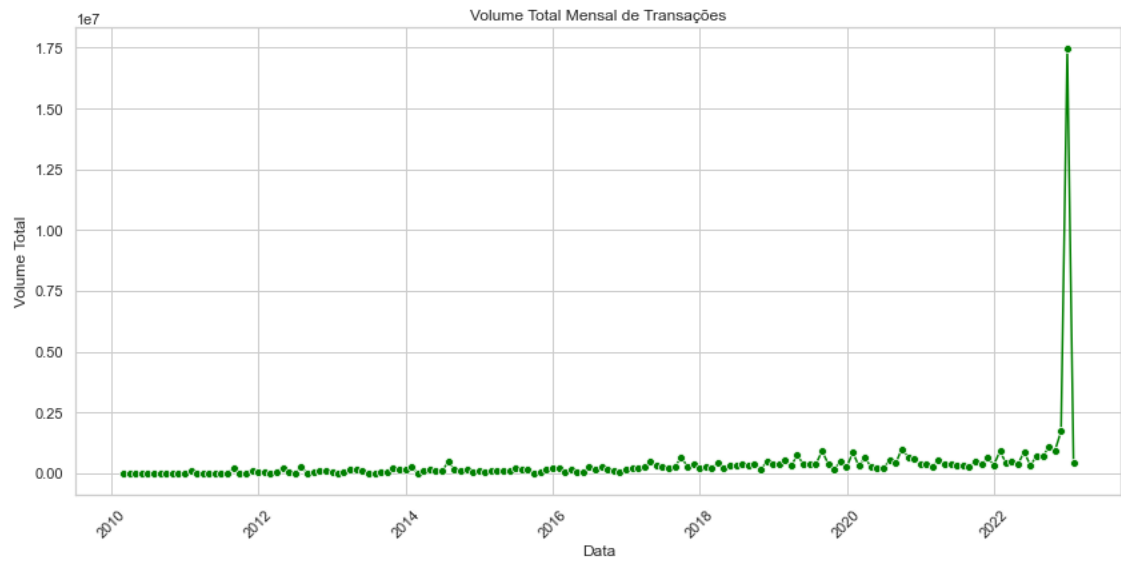
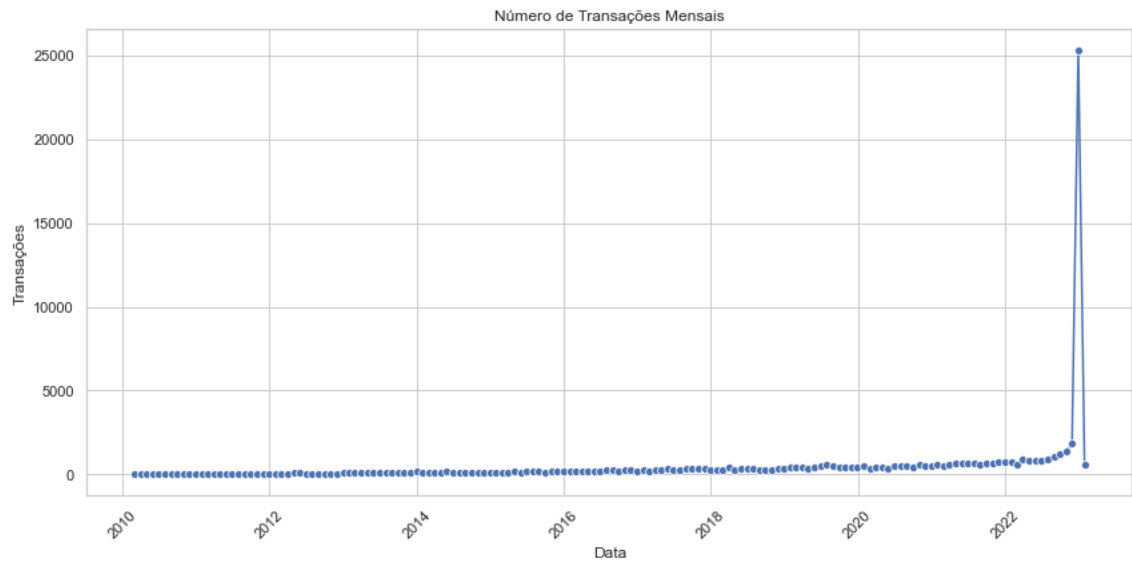
Os registros foram particionados em “sem inconsistências” (utilizados nas análises) e “inconsistentes” (para análise futura ou correção). Essa etapa identificou que a conta de número 528 não consta na base de dados dos clientes, então ela foi removida da base de dados das contas e suas transações removidas da base de dados das transações. Essas observações inconsistentes foram salvas em **contas\_inconsistentes.csv** e **transecoes\_inconsistentes.csv** respectivamente.

## 4. Análise Exploratória e Geração de Indicadores (analise\_dados.py)

### 4.1. Análise de Transações

As transações foram agrupadas por mês, e foram gerados gráficos de linha mostrando a evolução do número de transações e do volume total, a partir desse gráfico é possível obter insights como:

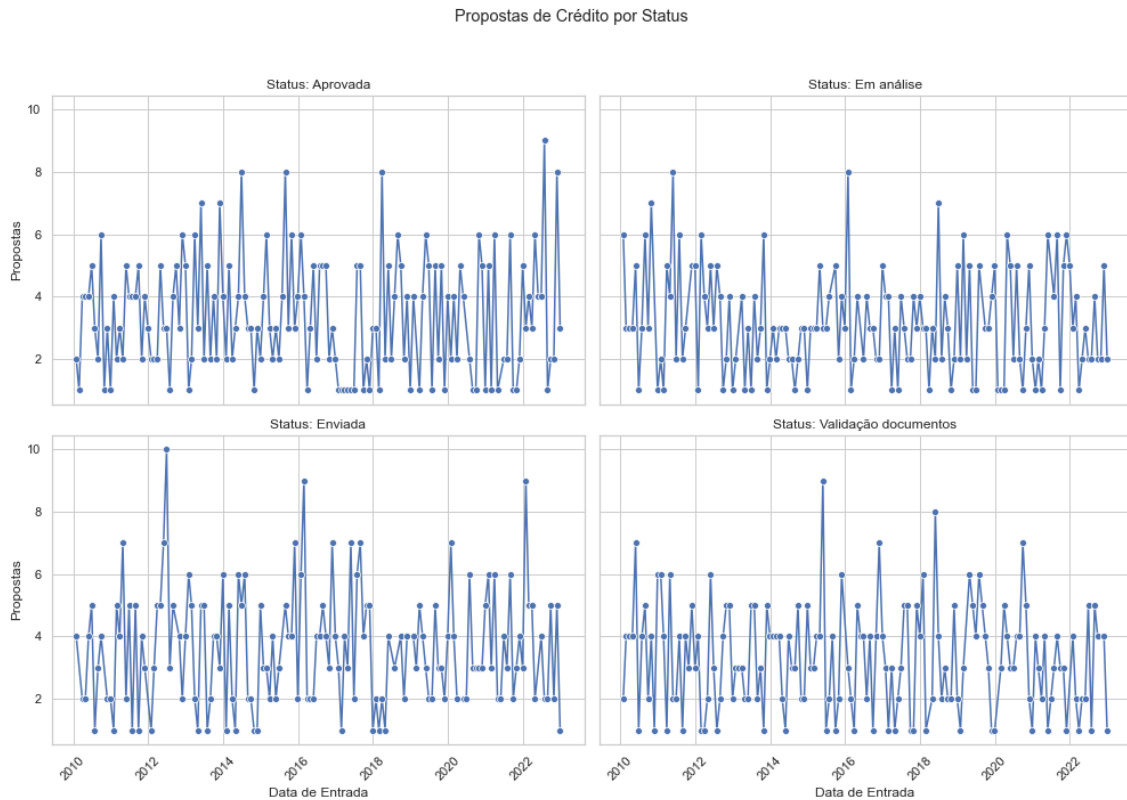
- Sazonalidade e Tendências: Ao agrupar as transações por mês, é possível identificar picos e quedas na atividade. Períodos com alta movimentação podem indicar sazonalidade, eventos promocionais ou ações de marketing bem-sucedidas.
- Planejamento Operacional: A identificação de meses de baixa atividade pode orientar campanhas para estimular a movimentação e melhorar a eficiência operacional.
- Dezembro de 2022 foi o mês que teve tanto o maior volume de movimentações, quanto a maior quantidade de transações



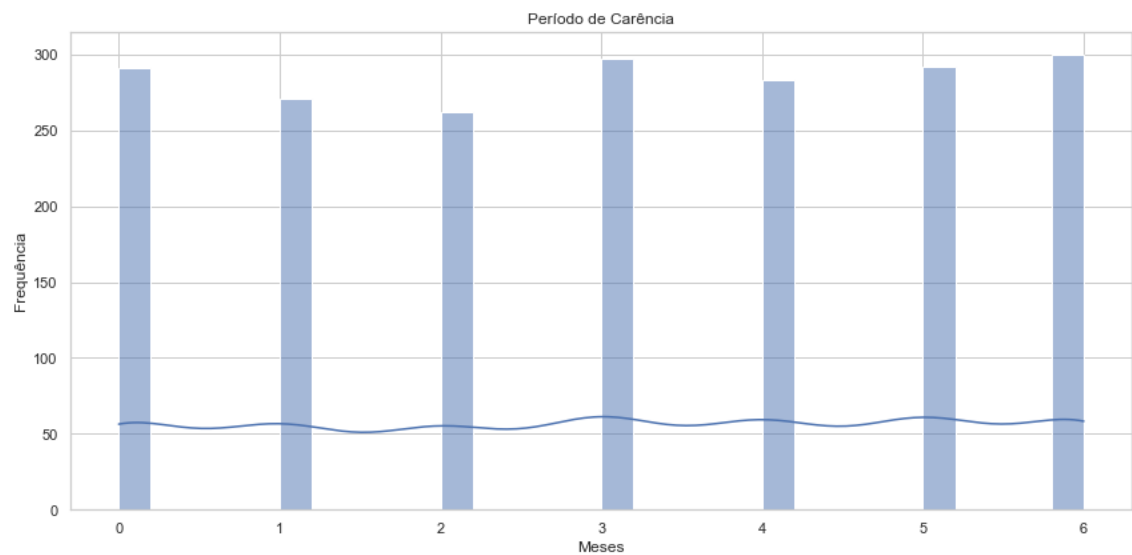
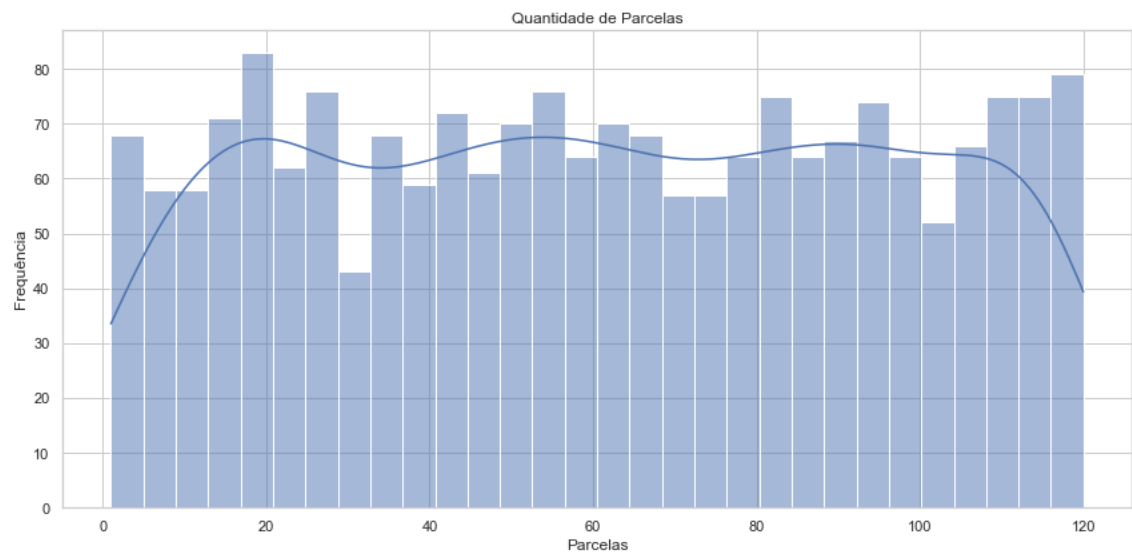
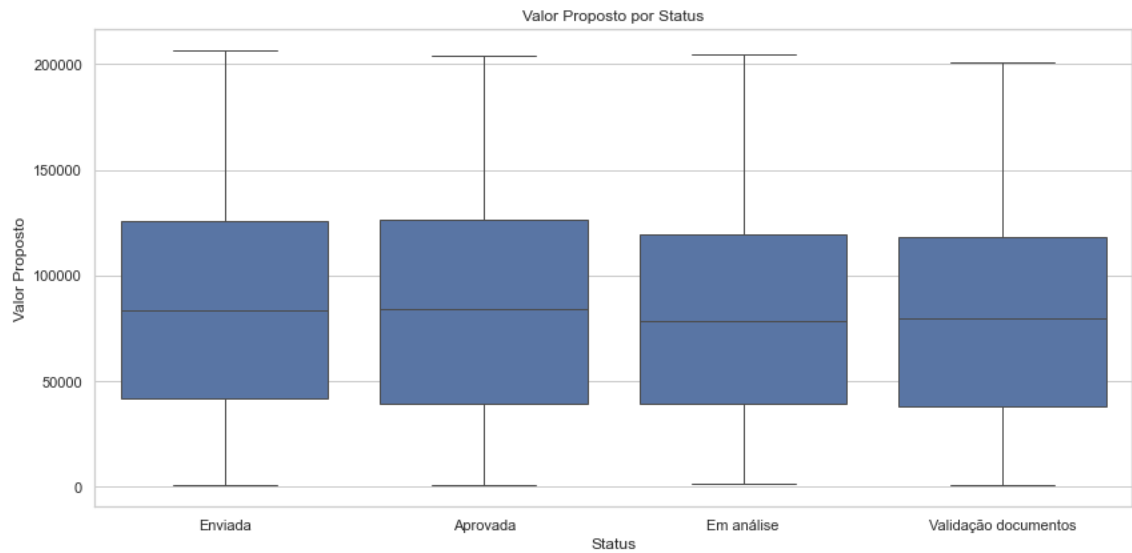
## 4.2. Análise de Propostas de Crédito

Propostas foram agrupadas por data de entrada e status, com visualizações que incluem lineplots, boxplots e histogramas para analisar valores propostos, número de parcelas e período de carência:

- **Propostas de crédito por status:** Permite visualizar a evolução do número de propostas conforme seu status, possibilitando identificar se há algum gargalo ou padrão que precise ser ajustado.



- **Valor proposto por status:** O boxplot do valor proposto, juntamente com histogramas da quantidade de parcelas e período de carência, fornece insights sobre a variabilidade dos pedidos, auxiliando na definição de políticas de crédito e na avaliação do risco.

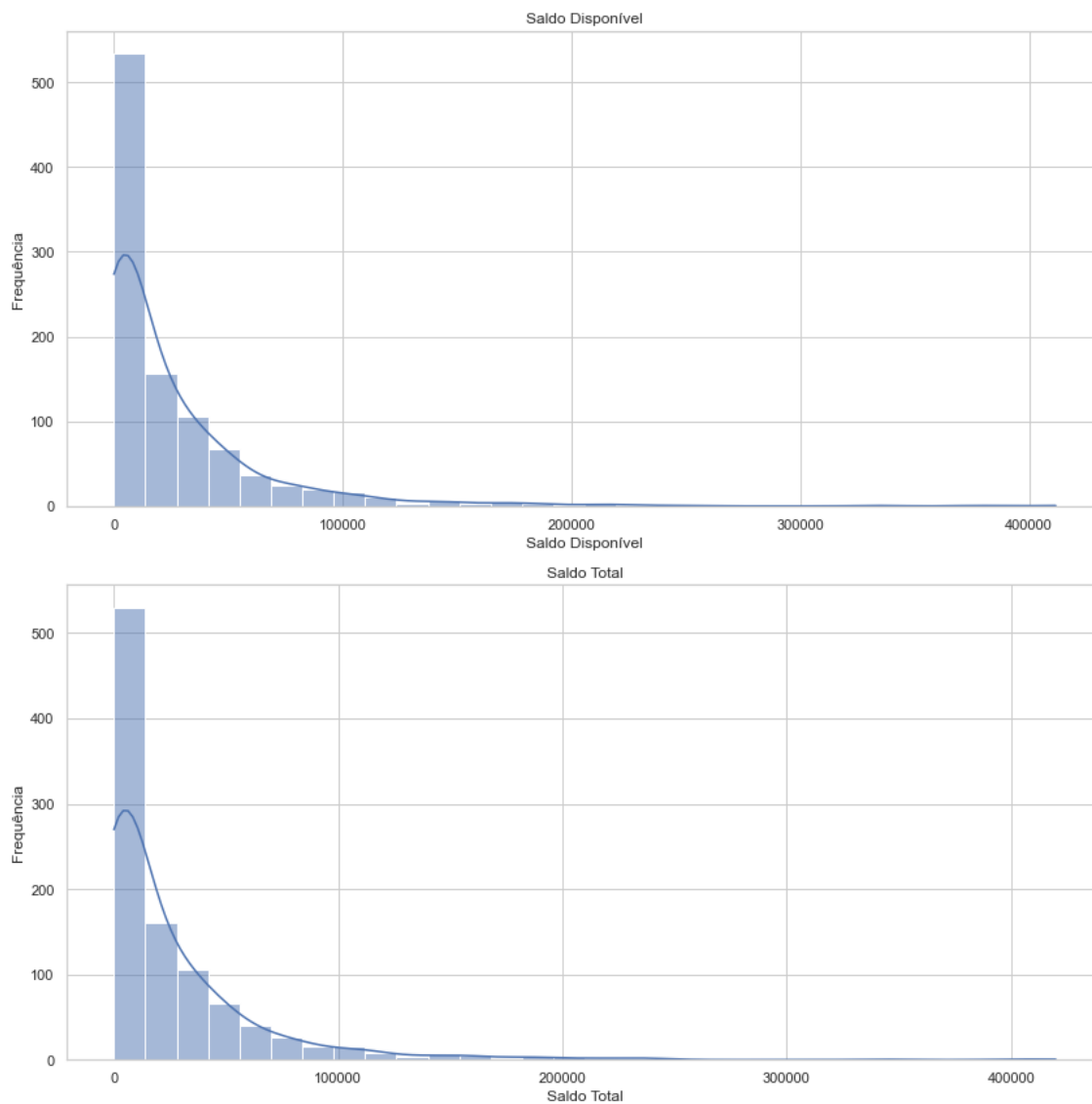


A taxa de aprovação de propostas de crédito calculada foi de 25.70%

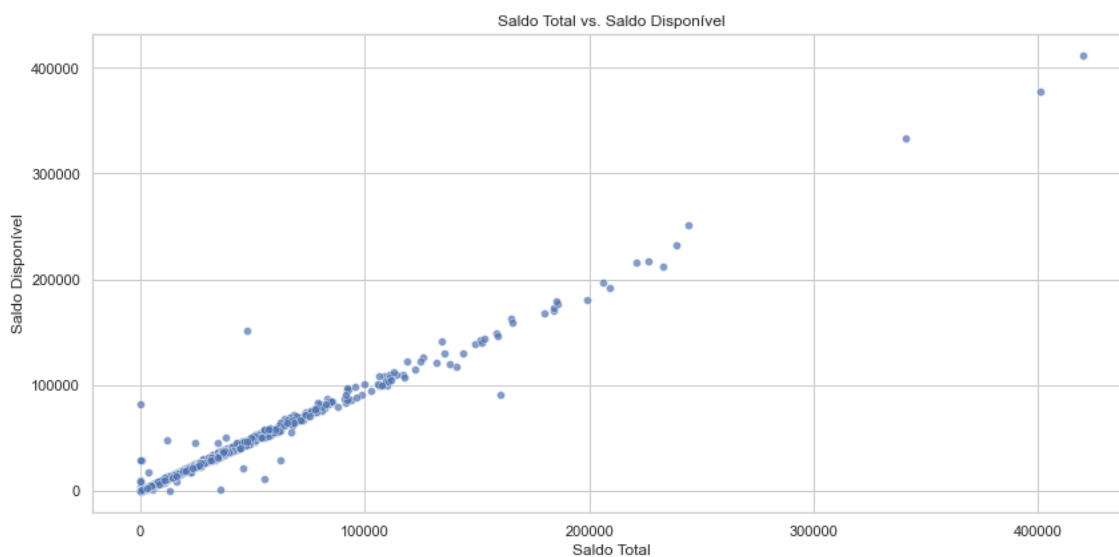
### 4.3. Análise de Contas

Histogramas e scatterplots demonstram a distribuição dos saldos disponíveis e totais:

- A distribuição dos saldos disponíveis e totais permite identificar a saúde financeira dos clientes, apontando, por exemplo, se a maioria possui saldos baixos (indicando possível oportunidade de cross-selling) ou altos (potencial para produtos de investimento).

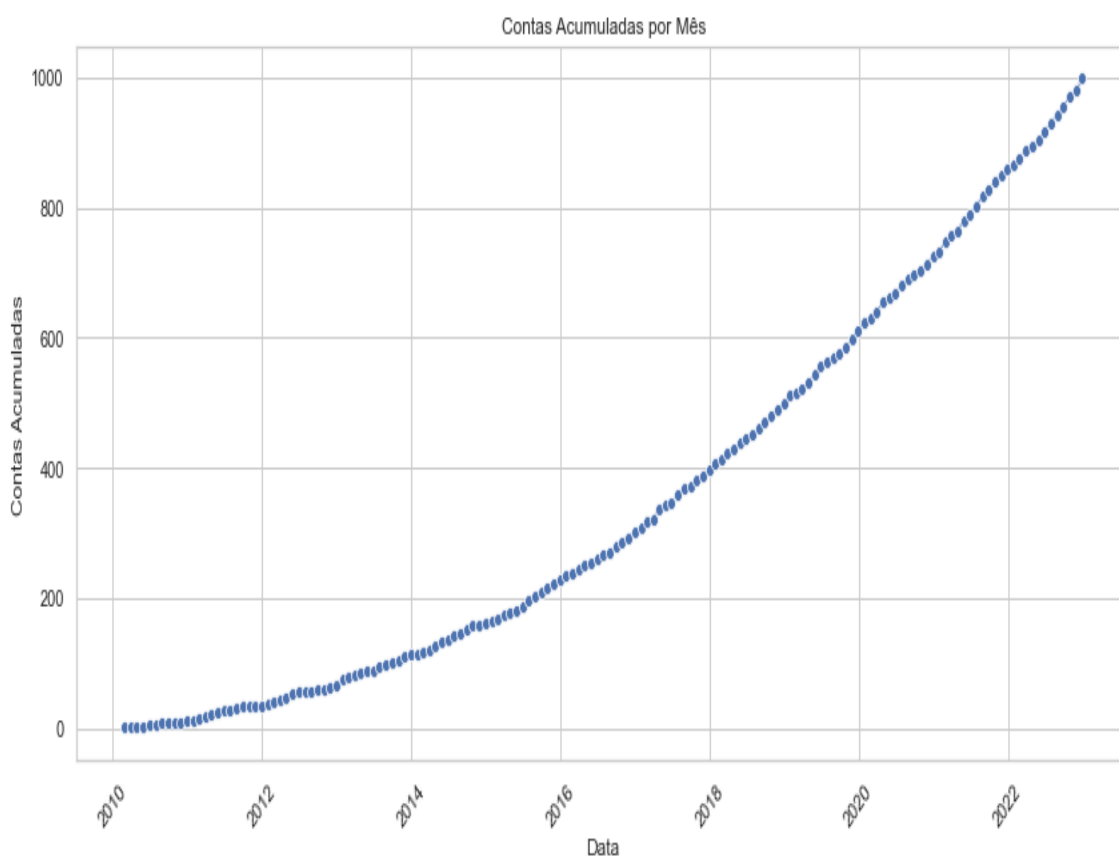


- A análise da relação entre saldo total e saldo disponível pode ajudar a entender o comportamento financeiro dos clientes e identificar inconsistências ou oportunidades de ajustes em produtos oferecidos.



A correlação entre saldo total e saldo disponível é de 0.9893

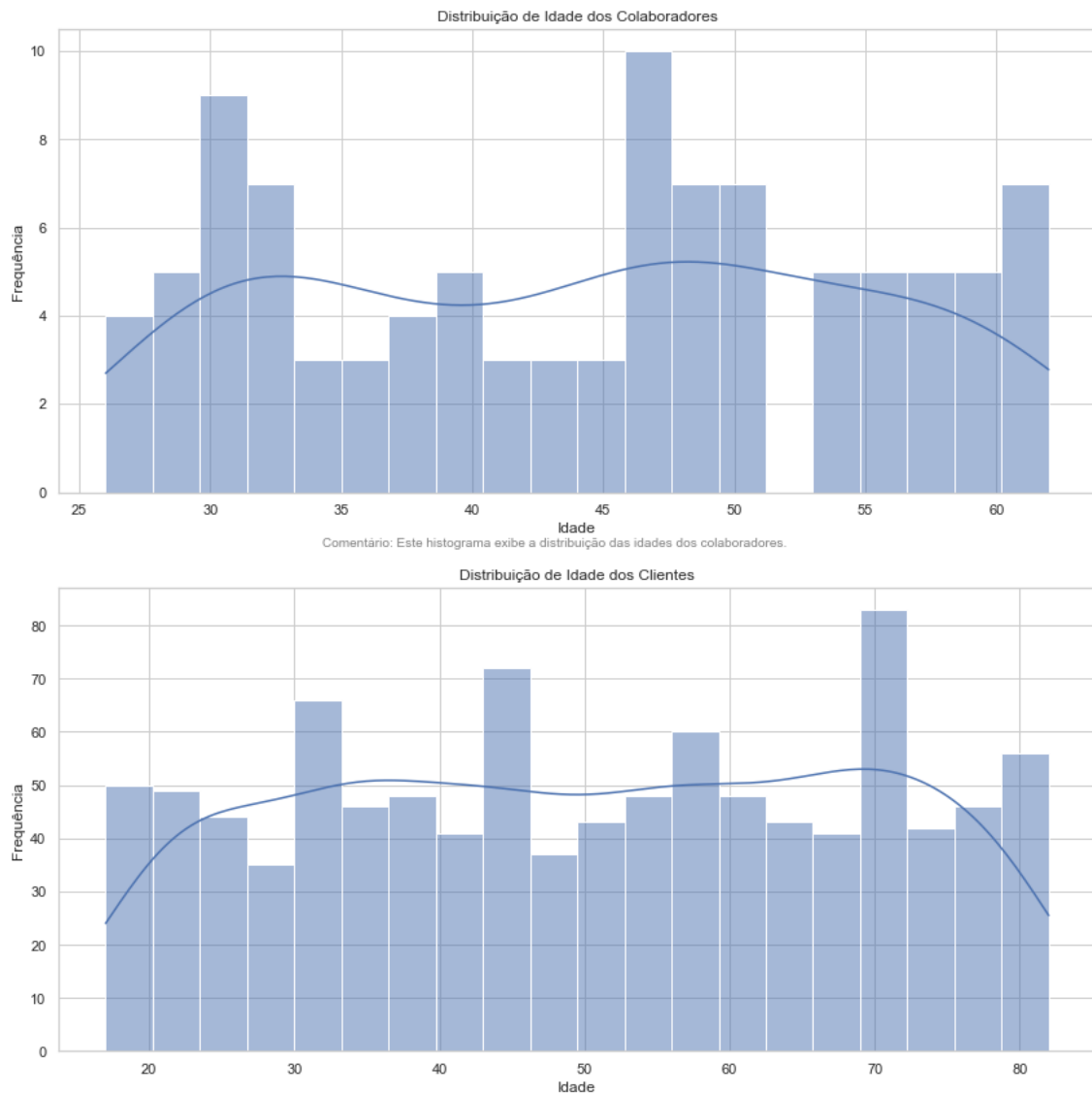
O crescimento do número de contas ao longo do tempo foi avaliado e visualizado por meio de gráficos de linha:





## 4.4. Análise de Colaboradores e Clientes

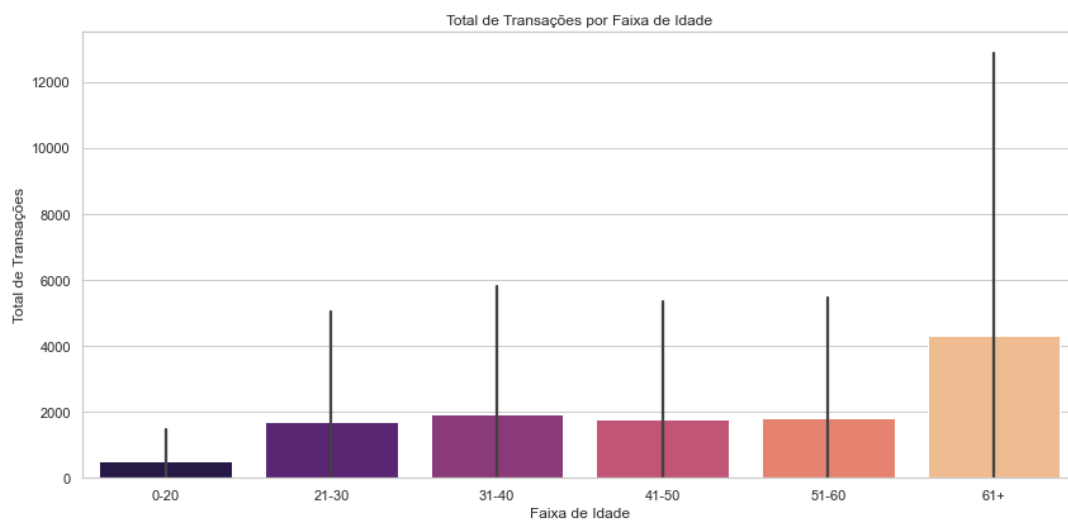
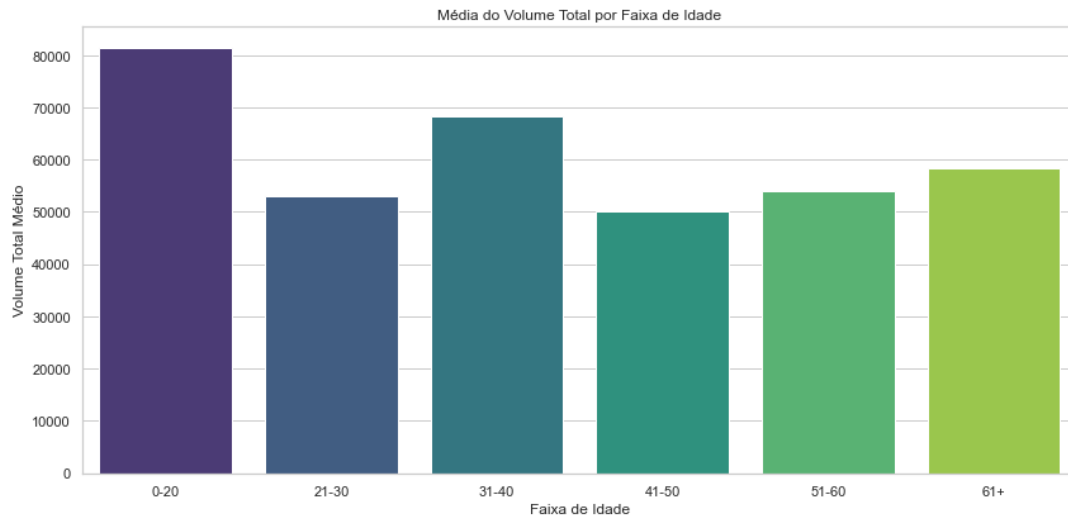
Foram gerados histogramas para visualizar a distribuição etária de colaboradores e clientes.



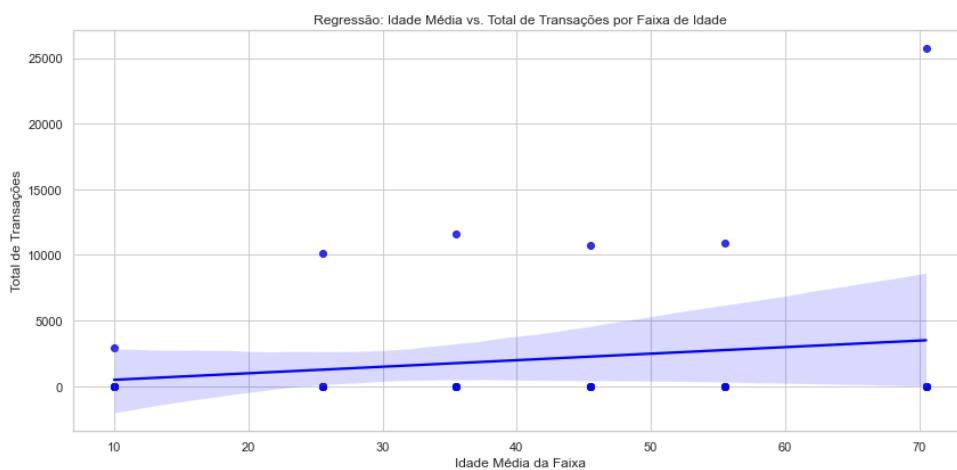
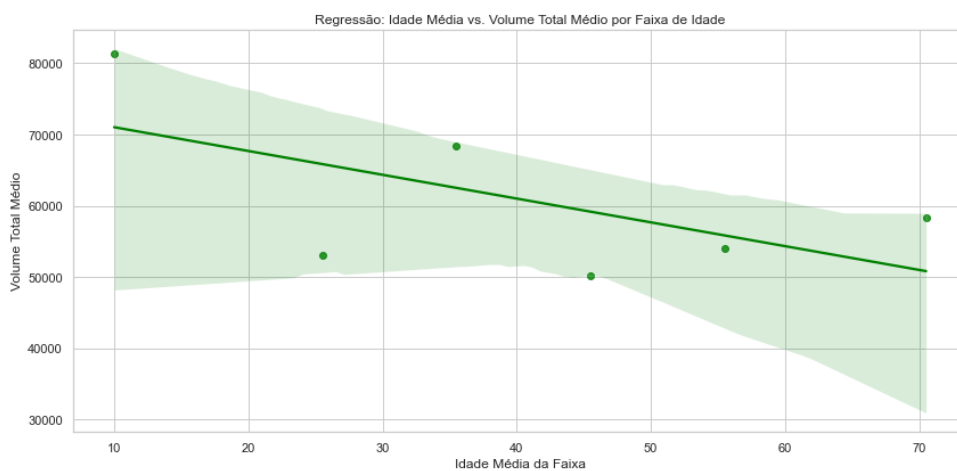
Ao analisar a distribuição das idades dos colaboradores e dos clientes, é possível segmentar os grupos por faixa etária. Isso pode embasar estratégias de comunicação e campanhas direcionadas, além de indicar necessidades de capacitação ou reestruturação.

A relação entre idade e comportamento transacional foi explorada, segmentando os clientes em faixas etárias e aplicando análises de regressão para identificar tendências. Ao cruzar os dados de clientes com os das transações, é possível identificar quais faixas etárias têm maior volume de transações ou maior frequência, permitindo uma segmentação mais precisa e estratégias de marketing personalizadas.

A análise de regressão que relaciona a idade (valor médio da faixa) com o volume total e o número de transações pode revelar tendências de comportamento que influenciam a definição de metas e a oferta de novos produtos.



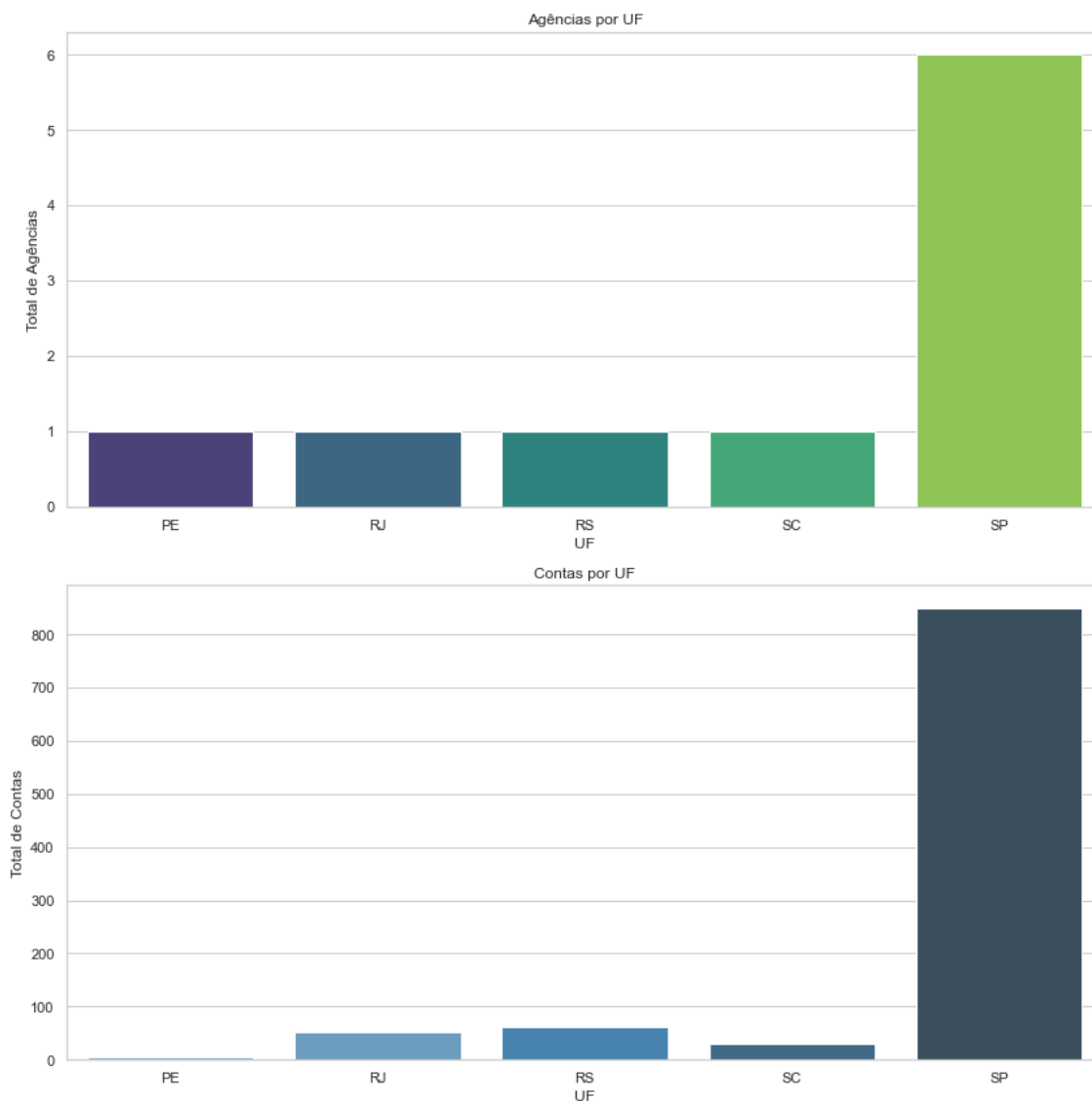
Comentário: O gráfico superior mostra a média do volume total por faixa etária, enquanto o gráfico inferior exibe o total de transações por faixa etária.



Comentário: O gráfico superior de regressão mostra a relação entre a idade média e o volume total médio, enquanto o gráfico inferior exibe a relação entre a idade média e o total de transações por faixa etária.

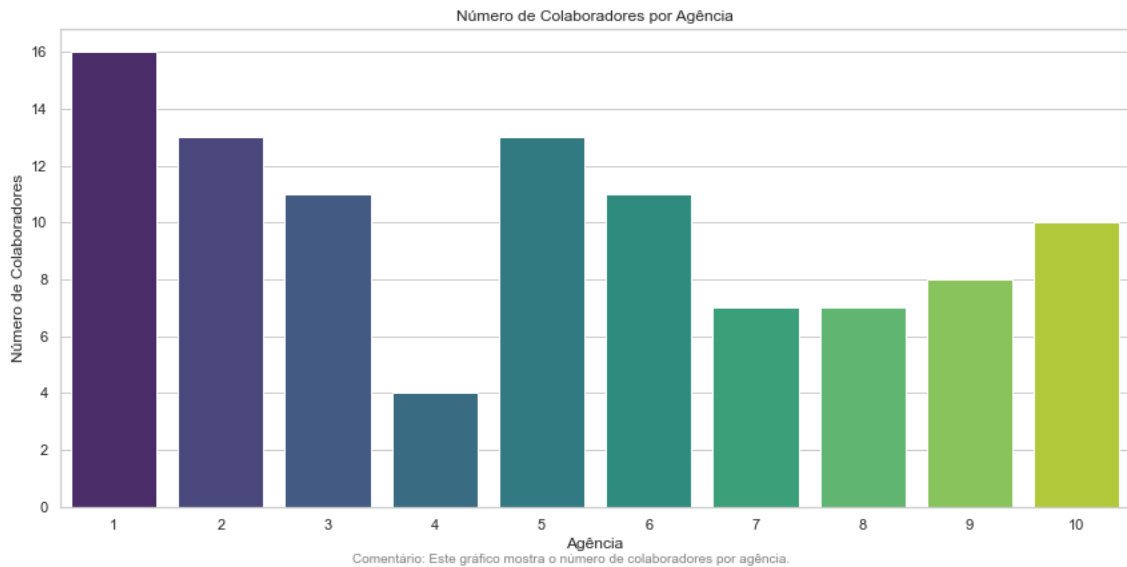
## 4.5. Análise das Agências

Os gráficos demonstram a contagem de agências e o número de contas por UF ajudam a identificar regiões com maior potencial de crescimento e aquelas que podem precisar de uma reavaliação estratégica. Comparar a performance entre agências permite otimizar a alocação de recursos e identificar melhores práticas que possam ser replicadas em outras unidades.



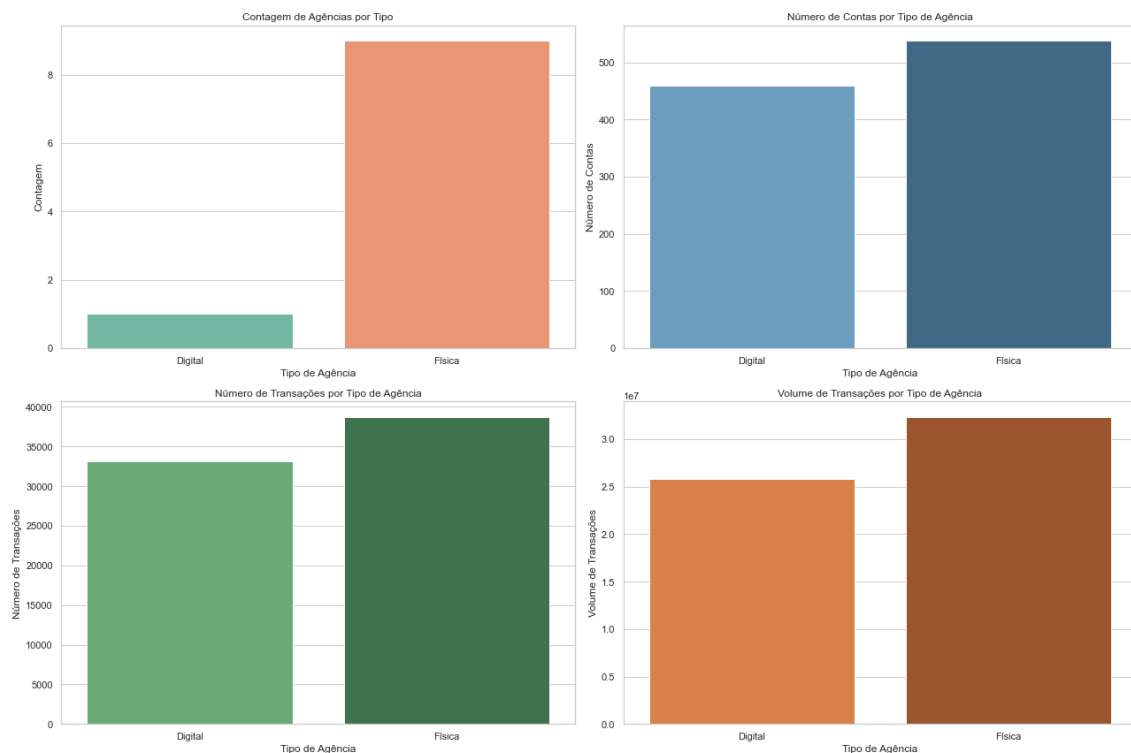
É possível concluir que SP possui o maior número de agências e de contas.

Os gráficos de barras que mostram o número de colaboradores por agência auxiliam na avaliação da estrutura operacional e na identificação de agências que podem necessitar de ajustes ou investimentos.



É possível concluir que a agência 1 (Agência Matriz) possui o maior número de colaboradores.

Esses quatro gráficos em conjunto tem o objetivo de comparar detalhadamente as diferenças entre a agência digital e as físicas. Essa comparação é fundamental para que o BanVic alinhe suas estratégias de negócio, otimize a operação de seus canais de atendimento e ofereça soluções cada vez mais adequadas às expectativas dos clientes.



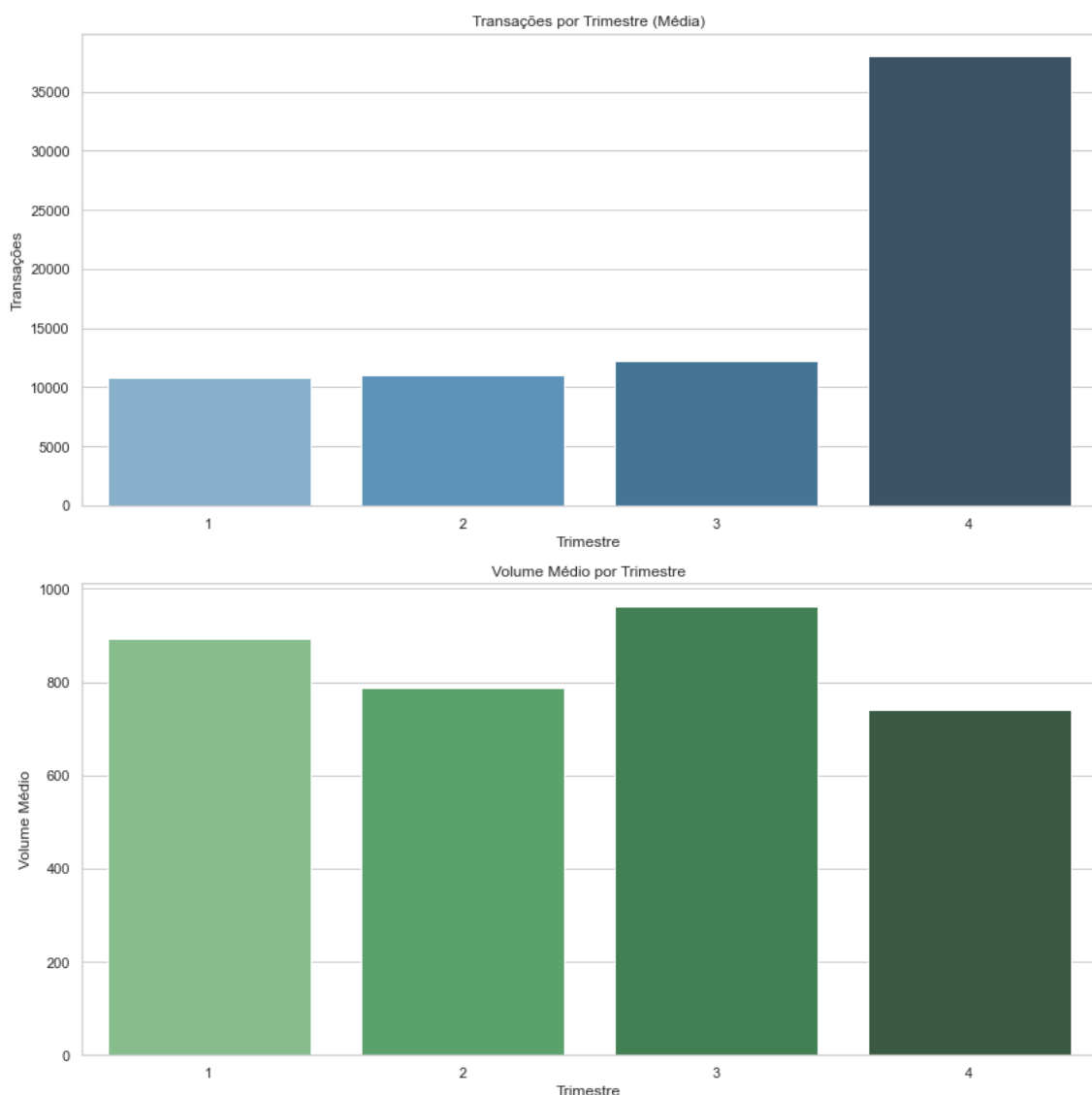
É possível concluir que as agencias físicas possuem o maior número de agencias, de contas, volume e transações, porém está claro que a agência digital merece a devida atenção.

## 4.6. Dimensão de Datas e Análises Temporais Avançadas

Uma tabela com datas diárias foi criada, extraindo atributos como ano, mês, nome do mês e trimestre. Isso permitiu que as análises trimestral e de meses com “r” possam ser feitas de forma eficiente.

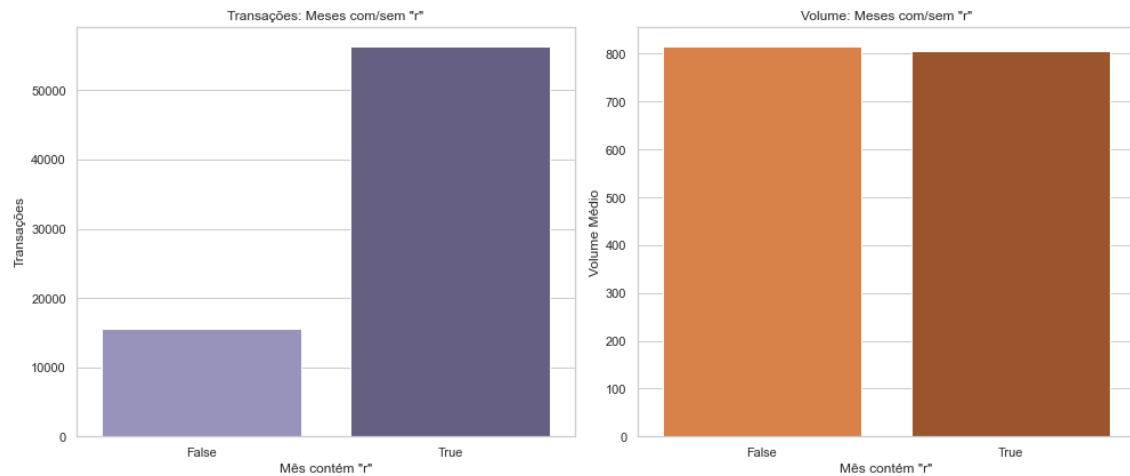
O agrupamento dos dados por trimestre e a comparação entre meses que contêm “r” e os que não contêm podem revelar padrões sazonais ou comportamentais, oferecendo subsídios para ações de marketing e ajustes operacionais específicos para determinados períodos do ano. Esses insights permitem alinhar campanhas promocionais ou ajustar estratégias de atendimento de acordo com as variações sazonais identificadas.

Agrupamento dos dados por trimestre permitiu identificar qual período apresenta, em média, mais transações e maior volume movimentado.



É possível concluir que, em média, o quarto trimestre possui o maior número de transações e o terceiro o maior volume de transações.

Foi realizada uma análise comparativa entre meses que contêm a letra “r” e aqueles que não contêm, verificando diferenças na quantidade e no volume de transações.

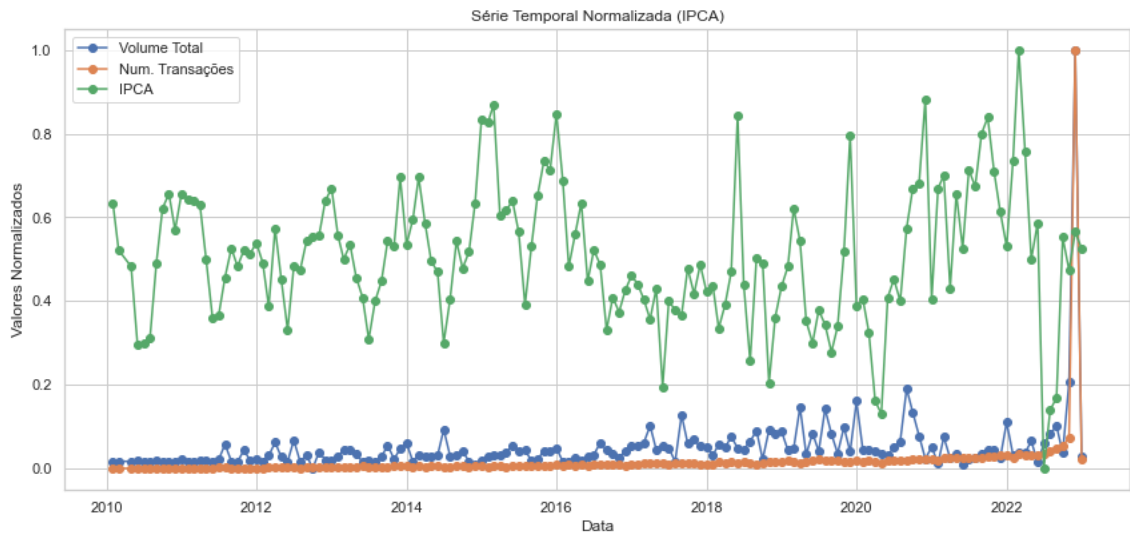


É possível concluir que os meses com “r” tem o maior número de transações e o volume de transações um pouco menor que os meses sem “r”.

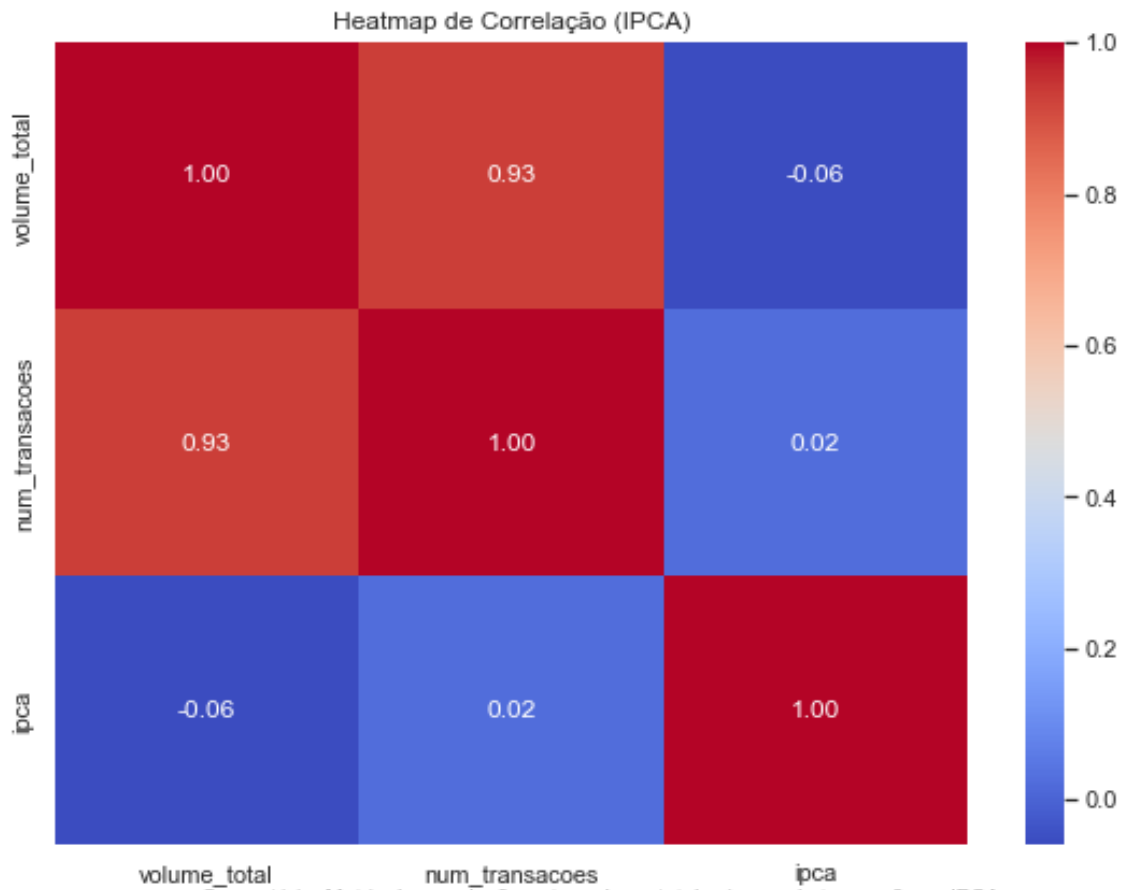
## 4.7. Integração de Dados Externos

Ao comparar os dados internos com indicadores externos, os gráficos normalizados permitem visualizar se há correlação entre a economia (medida pelo IPCA, SELIC e ICC) e o volume ou número de transações. Esses insights podem fundamentar decisões estratégicas, como a revisão de taxas de juros e a definição de ofertas de crédito, considerando o ambiente econômico.

Os dados externos foram obtidos a partir de APIs do Governo, permitindo correlacionar IPCA, SELIC e ICC com o volume total das transações. Foram construídos gráficos de séries temporais normalizadas (através do MinMaxScaler) para comparar o volume total, número de transações e os indicadores macroeconômicos.

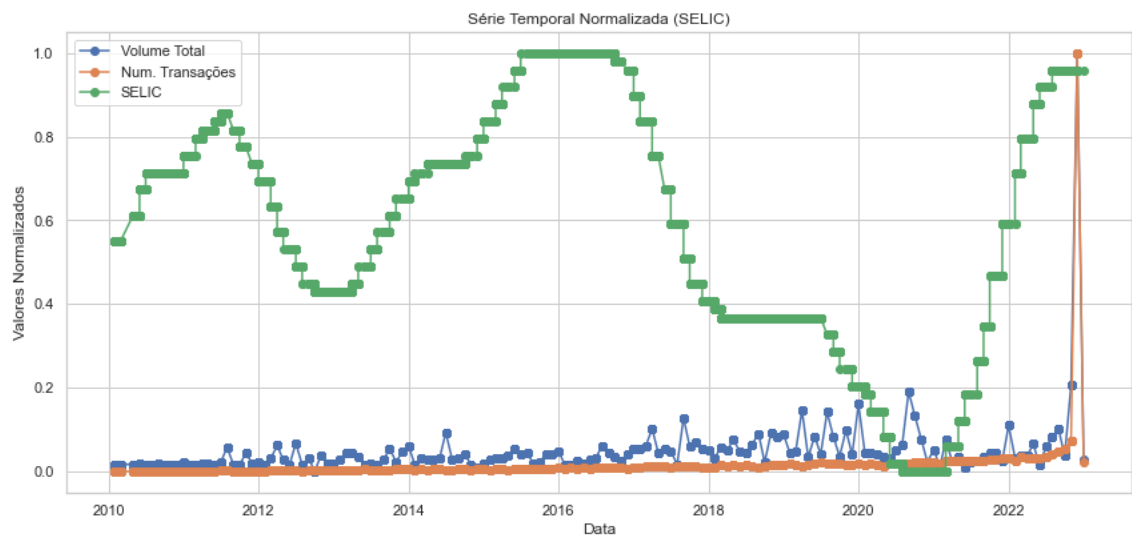


Comentário: Série temporal normalizada comparando volume total, número de transações e IPCA.

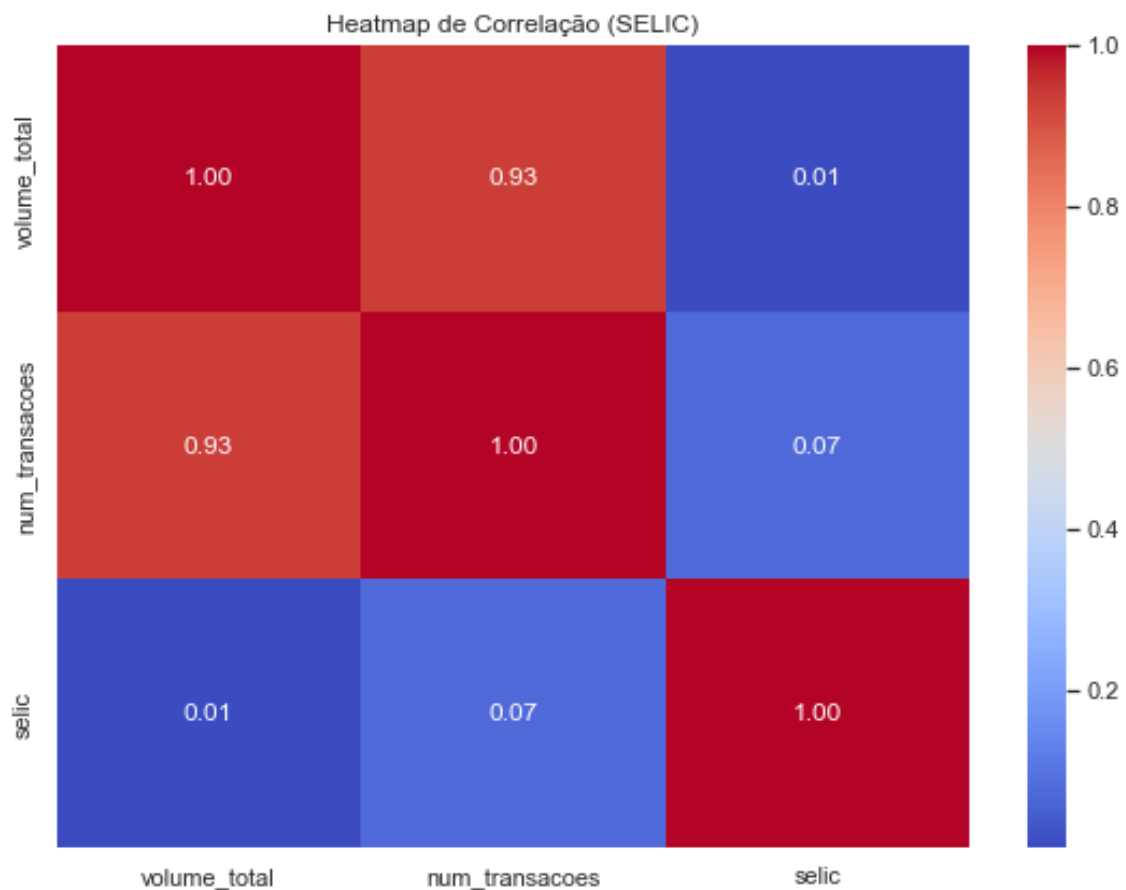


Comentário: Matriz de correlação entre volume total, número de transações e IPCA.

A correlação entre IPCA e volume total é -0.0576



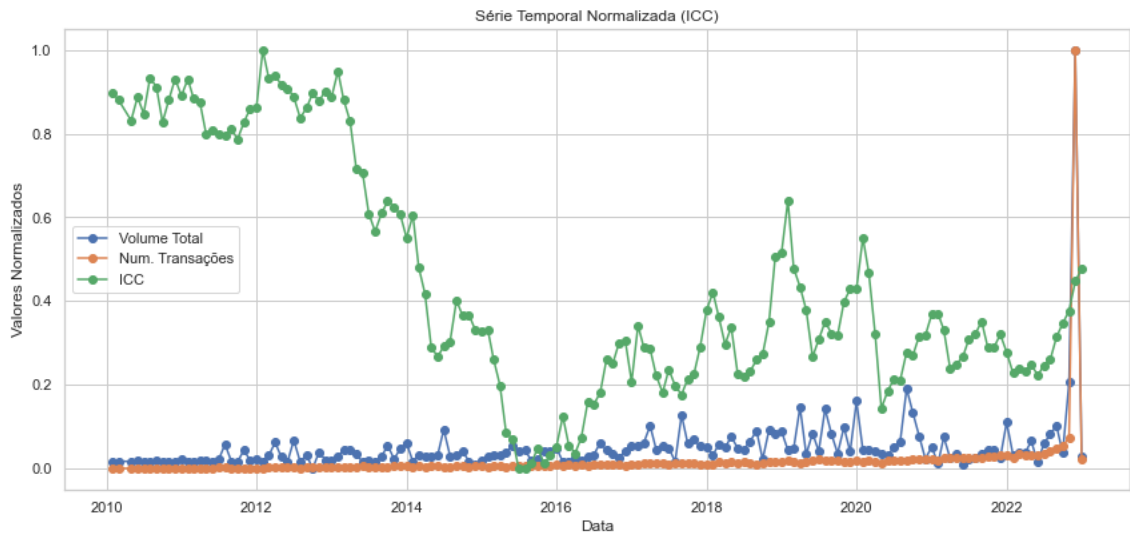
Comentário: Série temporal normalizada comparando volume total, número de transações e SELIC.



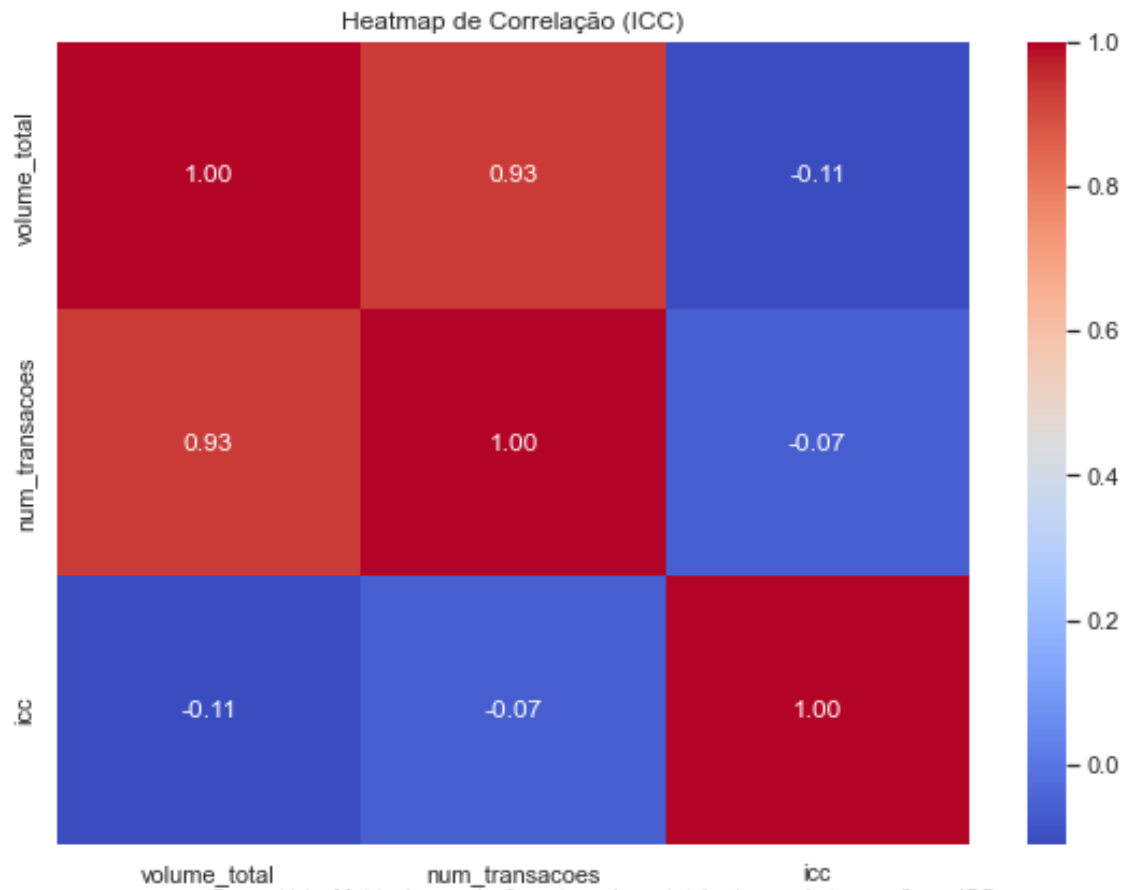
Comentário: Matriz de correlação entre volume total, número de transações e SELIC.

A correlação entre SELIC e volume total é 0.0064





Comentário: Série temporal normalizada comparando volume total, número de transações e ICC.



Comentário: Matriz de correlação entre volume total, número de transações e ICC.

A correlação entre ICC e volume total é -0.1101

## 5. Dashboard interativo

Além dos scripts de ETL, tratamento, e análise exploratória, foi desenvolvido um dashboard interativo que consolida e apresenta os principais indicadores e insights do BanVic de forma dinâmica e visual. Este dashboard foi estruturado para oferecer uma visão ampla e detalhada do desempenho operacional e financeiro, além de integrar dados internos com indicadores macroeconômicos externos.

### 5.1 Filtros

O dashboard dispõe de diversos filtros que permitem ao usuário personalizar a visualização dos dados e realizar análises dinâmicas, tais como:

- Ano e Trimestre: Permite selecionar períodos específicos para análise temporal, facilitando a identificação de tendências e sazonalidades.
- Agência: Filtra os dados por agência, possibilitando comparações detalhadas entre diferentes unidades.
- UF (Unidade Federativa): Oferece a segmentação regional dos dados, auxiliando na avaliação de desempenho por região.
- Idade: Permite segmentar clientes e colaboradores em diferentes faixas etárias, possibilitando análises demográficas.
- Proposta: Filtra os dados relacionados às propostas de crédito, permitindo avaliar a performance dos pedidos.
- Categoria da Transação: Facilita a seleção e comparação entre os diferentes tipos de transações (entrada, saída etc.).
- Outliers: Opção para exibir ou excluir dados considerados outliers, garantindo que análises não sejam distorcidas por valores extremos.

### 5.2 Cards de Indicadores

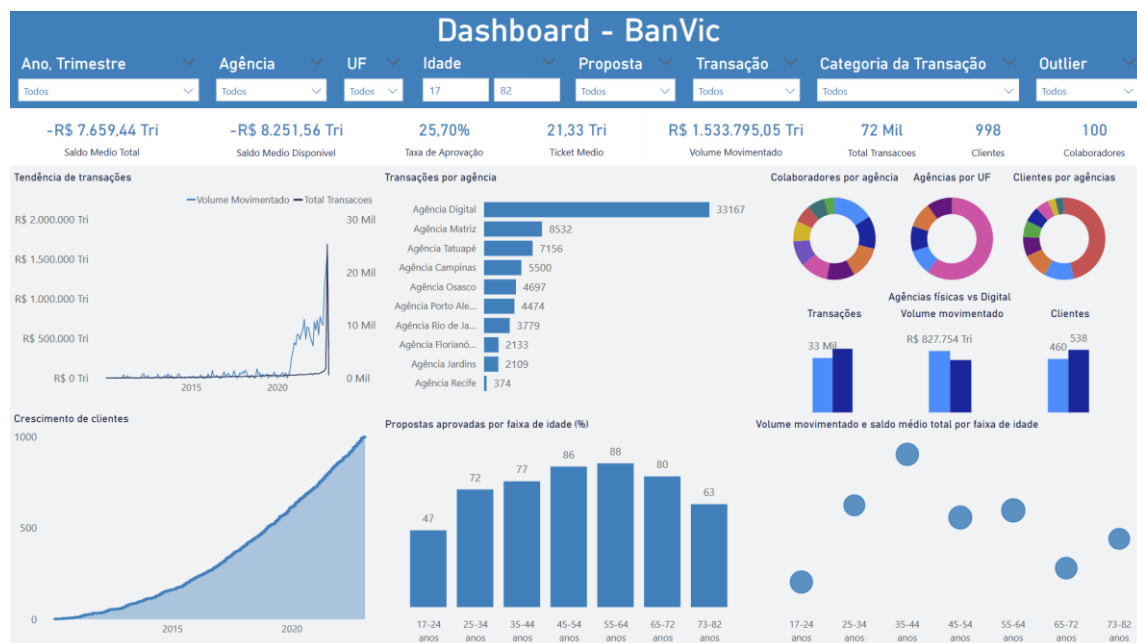
Na parte superior do dashboard, são apresentados diversos cards que destacam os principais indicadores de performance:

- Saldo Médio Total e Disponível: Mostra a média dos saldos das contas, permitindo uma rápida avaliação da saúde financeira dos clientes.
- Taxa de Aprovação: Indica o percentual de propostas de crédito aprovadas, fornecendo insights sobre a eficácia do processo de crédito.
- Ticket Médio: Exibe o valor médio das transações realizadas, contribuindo para a análise do comportamento de consumo dos clientes.
- Volume Movimentado: Demonstra o total de recursos movimentados, refletindo a atividade financeira da instituição.
- Total de Transações: Apresenta a quantidade total de transações, permitindo acompanhar o nível de atividade operacional.
- Clientes e Colaboradores: Indicadores que mostram o total de clientes atendidos e colaboradores envolvidos, oferecendo uma visão do alcance e da estrutura operacional do BanVic.

## 5.3 Gráficos

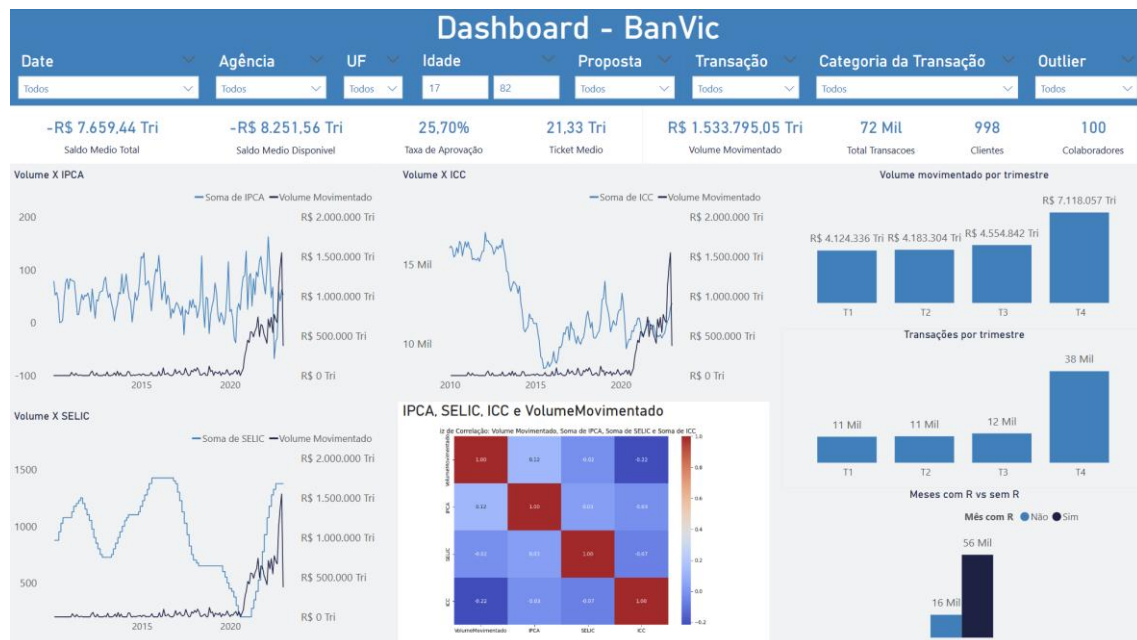
A primeira página do dashboard é dedicada à análise operacional, consolidando gráficos que permitem uma compreensão detalhada das atividades internas:

- **Gráfico de Linhas Duplas:** Vai exibir, ao longo dos anos, o volume movimentado e o total de transações, permitindo a identificação de tendências e variações sazonais.
- **Gráfico de Barras:** Vai comparar o número de transações por agência, evidenciando o desempenho de cada unidade
- **Gráfico de Área:** Vai demonstrar o crescimento do número de clientes ao longo do tempo, destacando a evolução da base de clientes.
- **Gráfico de Colunas (Taxa de Aprovação por Faixa de Idade):** Visualizar a taxa de aprovação das propostas de crédito segmentada por faixa etária, permitindo identificar quais grupos possuem maior ou menor aprovação.
- **Gráfico de Dispersão:** Vai analisar a relação entre o volume movimentado e o saldo médio por faixa etária, oferecendo insights sobre o comportamento financeiro dos clientes.
- **Gráficos de Colunas Comparativos:** Vai comparar o desempenho entre agências físicas e digitais, destacando as diferenças de comportamento e eficiência entre os canais.
- **Gráficos de Rosca:** Vai apresentar a distribuição dos colaboradores por agência, o número de agências por UF e a distribuição de clientes por agência, facilitando a visualização comparativa dos principais agrupamentos.



A segunda página foca na integração dos dados internos com indicadores macroeconômicos e na análise temporal aprofundada:

- Três Gráficos de Linhas Duplas (Volume Movimentado vs. Indicadores Econômicos): Cada gráfico apresenta o volume movimentado ao longo dos anos correlacionado com um indicador macroeconômico específico – um para o IPCA, um para a SELIC, e um para o ICC. Essa comparação permite avaliar a influência do ambiente econômico sobre as operações do BanVic.
- Mapa de Calor: Gerado com Python, este mapa de calor correlaciona o volume movimentado com os indicadores IPCA, SELIC e ICC, facilitando a identificação de relações e padrões entre essas variáveis.
- Gráficos de Colunas (Volume e Transações por Trimestre): Vão exibir o volume movimentado e o número de transações segmentados por trimestre, permitindo uma análise temporal mais detalhada e a identificação de sazonalidades.
- Gráfico de Colunas Comparativo (Transações entre Meses com e sem “r”): Vai comparar a quantidade de transações entre meses que contêm a letra “r” e os que não contêm, oferecendo insights sobre possíveis variações comportamentais e sazonais.



## 6. Indicadores-Chave e Insights para o Negócio

- **Transações e Volume:** O monitoramento mensal e trimestral das transações permite identificar períodos de pico e sazonalidade, auxiliando no planejamento operacional e em campanhas promocionais.
- **Taxa de Aprovação de Crédito:** A análise das propostas de crédito possibilita entender a eficiência do processo e identificar oportunidades para ajustes nas políticas de crédito.
- **Crescimento de Contas e Distribuição Geográfica:** O acompanhamento do crescimento das contas e a análise por região (UF) permitem otimizar a estratégia de expansão e realocar recursos conforme a demanda regional.
- **Perfil dos Clientes:** A segmentação por faixas etárias revela comportamentos diferentes quanto ao volume e à frequência de transações, permitindo campanhas direcionadas e ofertas personalizadas.
- **Integração com Indicadores Externos:** As correlações entre o volume de transações e os indicadores macroeconômicos (IPCA, SELIC, ICC) fornecem subsídios para decisões estratégicas, como ajustes em taxas de juros e ofertas de crédito, considerando o cenário econômico atual.

## 6. Recomendações

Para que o BanVic possa aproveitar ao máximo os insights gerados a partir dos dados, é fundamental investir em soluções de Business Intelligence (BI) e dashboards interativos. Ao desenvolver painéis dinâmicos que consolidem indicadores-chave em tempo real, os gestores terão a possibilidade de monitorar o desempenho da instituição com agilidade e precisão. Esse investimento viabiliza a integração dos dados internos e externos, permitindo análises dinâmicas e a personalização das visualizações conforme as necessidades do negócio. A utilização de ferramentas como Power BI, Tableau ou frameworks em Python, como Dash e Streamlit, contribui para a criação de um ambiente visualmente intuitivo e informativo, onde filtros customizáveis e cards de indicadores facilitam a identificação de tendências e a tomada de decisões estratégicas.

Simultaneamente, o aprimoramento do processo de crédito se revela uma medida estratégica crucial. A análise detalhada da taxa de aprovação e do perfil dos clientes permite identificar segmentos com maior volume de operações e menor risco, direcionando campanhas específicas e personalizadas para esses grupos. Revisar e ajustar os critérios de avaliação das propostas de crédito, com base em modelos preditivos e análises estatísticas, pode reduzir a inadimplência e aumentar a eficiência na concessão de crédito. Essa abordagem torna o processo mais resiliente e alinhado às necessidades do mercado, contribuindo para uma política de crédito mais robusta e eficaz.

Além disso, a integração sistemática de dados externos, como os indicadores macroeconômicos (IPCA, SELIC e ICC), aos processos de análise interna é essencial para antecipar mudanças no comportamento dos clientes e ajustar a estratégia comercial e de risco. Ao incorporar esses dados, o BanVic passa a ter uma visão mais ampla do cenário econômico, possibilitando identificar correlações que podem influenciar o volume de transações, a taxa de aprovação e outros indicadores-chave. Esse monitoramento contínuo do ambiente macroeconômico permite que a instituição responda de forma proativa às variações do mercado, otimizando suas operações e decisões estratégicas.

Por fim, a capacitação da equipe é um pilar indispensável para transformar dados brutos em insights estratégicos e promover uma cultura orientada por dados. Investir em treinamentos e programas de capacitação para as áreas de tecnologia e de negócio assegura que todos os colaboradores estejam aptos a utilizar plenamente as ferramentas de BI disponíveis. Essa formação contínua, aliada a um ambiente de suporte técnico e a iniciativas que incentivem o pensamento data-driven, potencializa a capacidade analítica da equipe e fortalece a tomada de decisões baseada em evidências. Dessa forma, o BanVic estará mais preparado para inovar, responder rapidamente às mudanças do mercado e sustentar um crescimento consistente e sustentável.

## 7. Resultados e conclusão

A análise dos dados revelou que, ao agrupar as transações por trimestre, é possível identificar diferenças marcantes tanto no número de transações aprovadas quanto no volume movimentado. Especificamente, os resultados indicam que o quarto trimestre apresenta, em média, um maior número de transações. Esse comportamento pode estar relacionado a fatores sazonais, como campanhas de final de ano e a intensificação das atividades comerciais nesse período. Em contrapartida, verificamos que o terceiro trimestre demonstrou, em média, o maior volume movimentado, sugerindo que, embora a quantidade de transações seja menor, o valor das operações é significativamente mais elevado. Essa discrepância entre o número de transações e o volume movimentado é fundamental para orientar estratégias diferenciadas tanto na análise de risco quanto na definição de metas comerciais.

Além disso, a hipótese levantada por um dos analistas sobre a influência da presença da letra “r” no nome dos meses foi cuidadosamente avaliada. A análise comparativa dos dados revelou que os meses que contêm a letra “r” tendem a apresentar sim um padrão distinto de quantidade de transações em comparação aos meses que não possuem esse caractere. Essa diferença pode refletir variações comportamentais ou sazonais específicas, embora seja recomendada uma investigação estatística mais aprofundada para confirmar a robustez dessa observação. De qualquer forma, o padrão identificado sugere que esses fatores podem ser levados em consideração para ajustar estratégias de marketing e operações, permitindo uma melhor adaptação às variações sazonais.

A integração dos novos dados externos com os indicadores internos também facilita uma série de decisões estratégicas. Com a inclusão de indicadores macroeconômicos, como o IPCA, SELIC e ICC, o BanVic pode antecipar mudanças no ambiente econômico e ajustar suas estratégias comerciais e de risco de forma proativa. Esse conhecimento permite, por exemplo, que a instituição otimize suas campanhas de marketing, aloque recursos operacionais de maneira mais eficiente e revise suas políticas de crédito para refletir tanto as condições do mercado quanto o comportamento histórico dos clientes. Dessa forma, a análise integrada dos dados não só embasa a tomada de decisão com informações atualizadas, mas também possibilita uma abordagem mais dinâmica e orientada por dados.

Através deste trabalho, foi demonstrado que é possível integrar dados internos e externos para gerar indicadores robustos que suportam decisões estratégicas no BanVic. As análises exploratórias e a criação de uma dimensão de datas possibilitam uma visão detalhada do comportamento transacional ao longo do tempo, enquanto a incorporação de dados macroeconômicos oferece uma perspectiva abrangente do ambiente econômico. Com a implementação das recomendações apresentadas e o investimento em ferramentas de BI, o BanVic estará mais bem preparado para identificar oportunidades, mitigar riscos e aprimorar seus resultados operacionais e financeiros. Esses insights, que incluem a identificação dos trimestres com maior desempenho em transações e volume, bem como a validação de padrões sazonais associados aos meses com e sem “r”, reforçam a importância de uma análise contínua e integrada, capaz de sustentar decisões estratégicas e operacionais com base em dados precisos e atualizados.