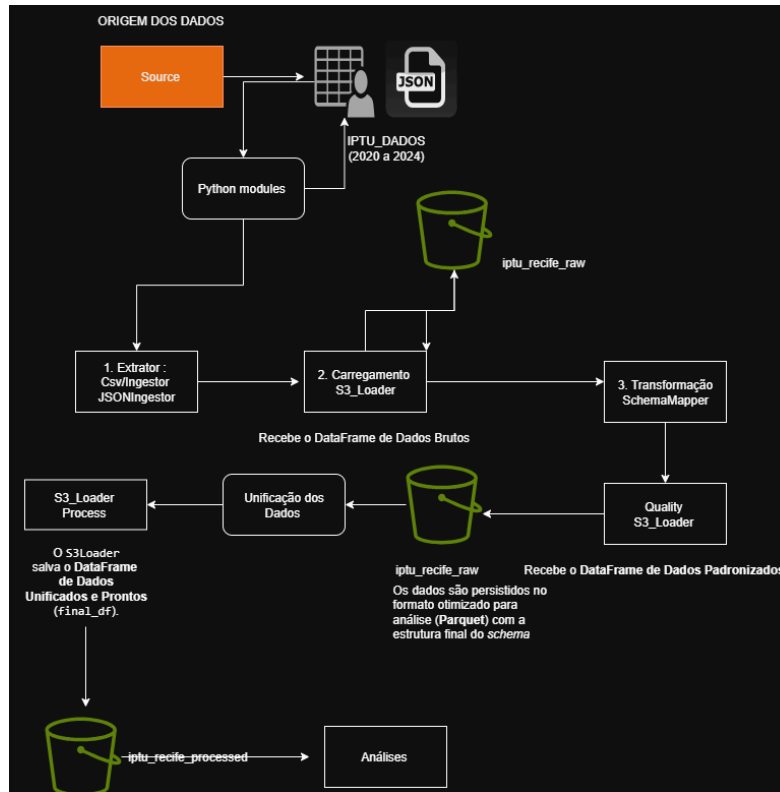


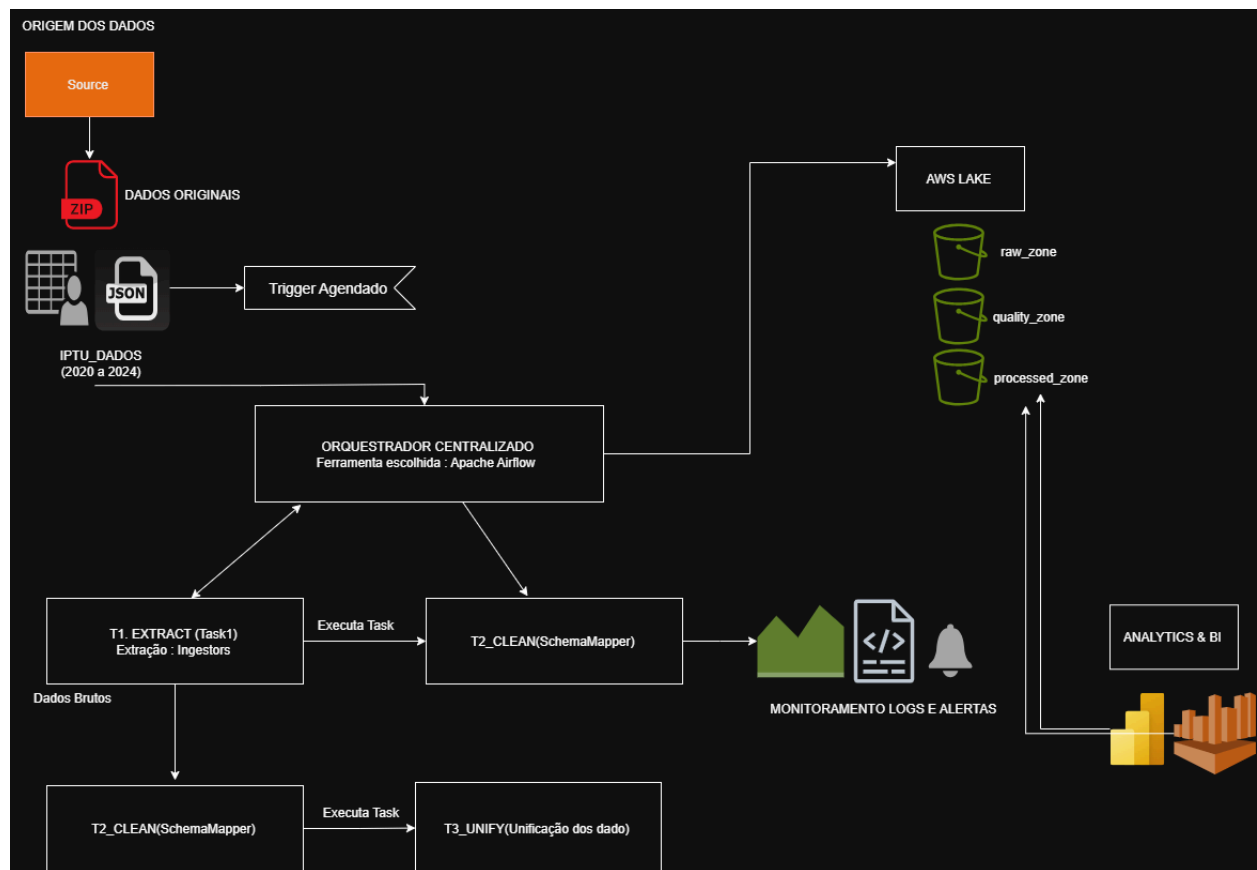
Análises IPTU RECIFE

Arquitetura Atual:



O processo se inicia com a **Extração** dos dados brutos (CSV/JSON) de um arquivo ZIP, utilizando módulos Python (**Ingestors**) que tratam de forma robusta problemas como codificação e leitura em *chunks*. Em seguida, a etapa de **Transformação** é executada pelo **SchemaMapper**, que padroniza nomes de colunas e garante que os dados sigam um *schema* consistente. Finalmente, na etapa de **Carregamento**, os dados limpos são salvos em um **AWS S3 Bucket**, que atua como o seu destino final ou Data Lake, prontos para serem consumidos por ferramentas de BI

Arquitetura Futura / Melhoria com Orquestrador



Seu projeto está estruturado como um pipeline de dados em Python, utilizando pandas e bibliotecas de armazenamento em nuvem (boto3), e o próximo passo lógico para aprimorar sua automação é implementar a Orquestração com Apache Airflow. Para isso, você deve criar um DAG (Directed Acyclic Graph) no Airflow onde cada etapa do seu fluxo (Extração, Limpeza e Unificação) se torna uma Task. A comunicação entre essas Tasks deve ser feita usando o XCom (Cross-Communication) do Airflow para repassar metadados (como o caminho S3 do arquivo recém-salvo) ou, de forma mais eficiente em pipelines de dados, usando os buckets S3 (RAW, QUALITY, PROCESSED) como pontos de entrega: a primeira Task de Extração carrega o arquivo bruto no bucket RAW, e a Task de Limpeza (usando o **SchemaMapper**) é disparada para ler do RAW e salvar no QUALITY, garantindo a gestão de dependências e a resiliência do fluxo.

Sobre as análises de dados dos dados processados :

1. Volume :

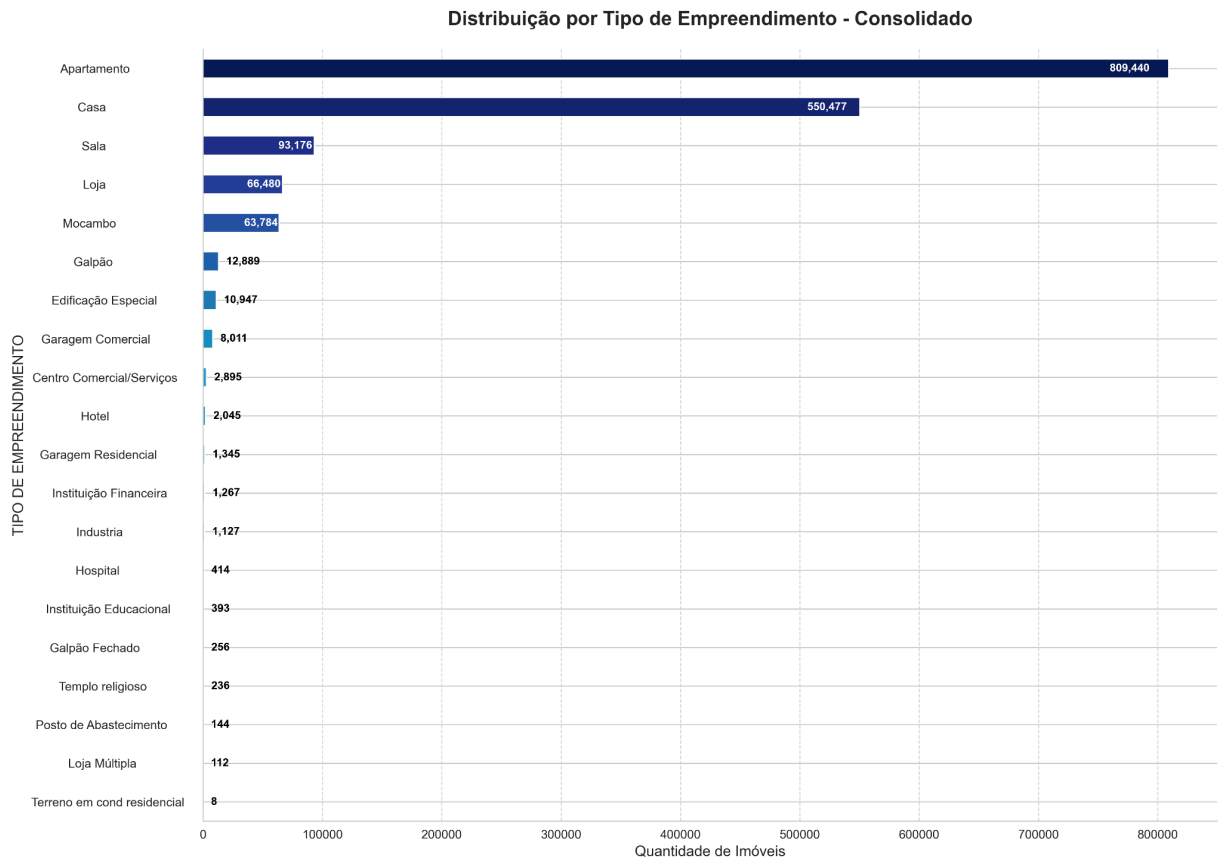
1.1 Qual o total de imóveis e como o inventário está distribuído fisicamente (tipo, bairro etc)?

O volume total de imóveis cadastrados na base de IPTU (consolidando todos os anos disponíveis) é de aproximadamente **1,62 milhão de registros** (1.625.446, conforme o gráfico "Distribuição por Tipo de Uso - Consolidado").



Ao analisar a evolução anual (gráfico "Evolução do Inventário de Imóveis por Ano"), observa-se o seguinte padrão para o volume de cadastros:

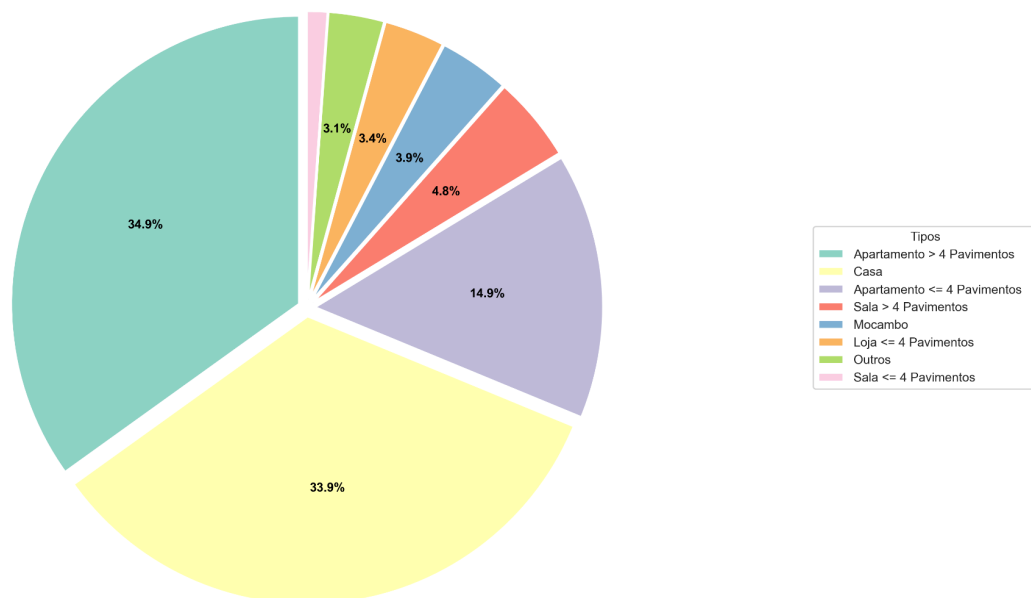
- O inventário se manteve **relativamente estável** entre 2020 e 2023, variando de 401.458 (2020) a 411.152 (2023).
- Há uma **queda abrupta** em 2024, onde o volume cai para apenas 448 imóveis, indicando uma **incompletude acentuada** nos dados de 2024 (provavelmente o ano está em processamento ou apenas dados parciais foram carregados).



Analisando a estrutura do inventário (gráfico "Distribuição por Tipo de Empreendimento - Consolidado"), o domínio é claramente habitacional, dividido principalmente entre:

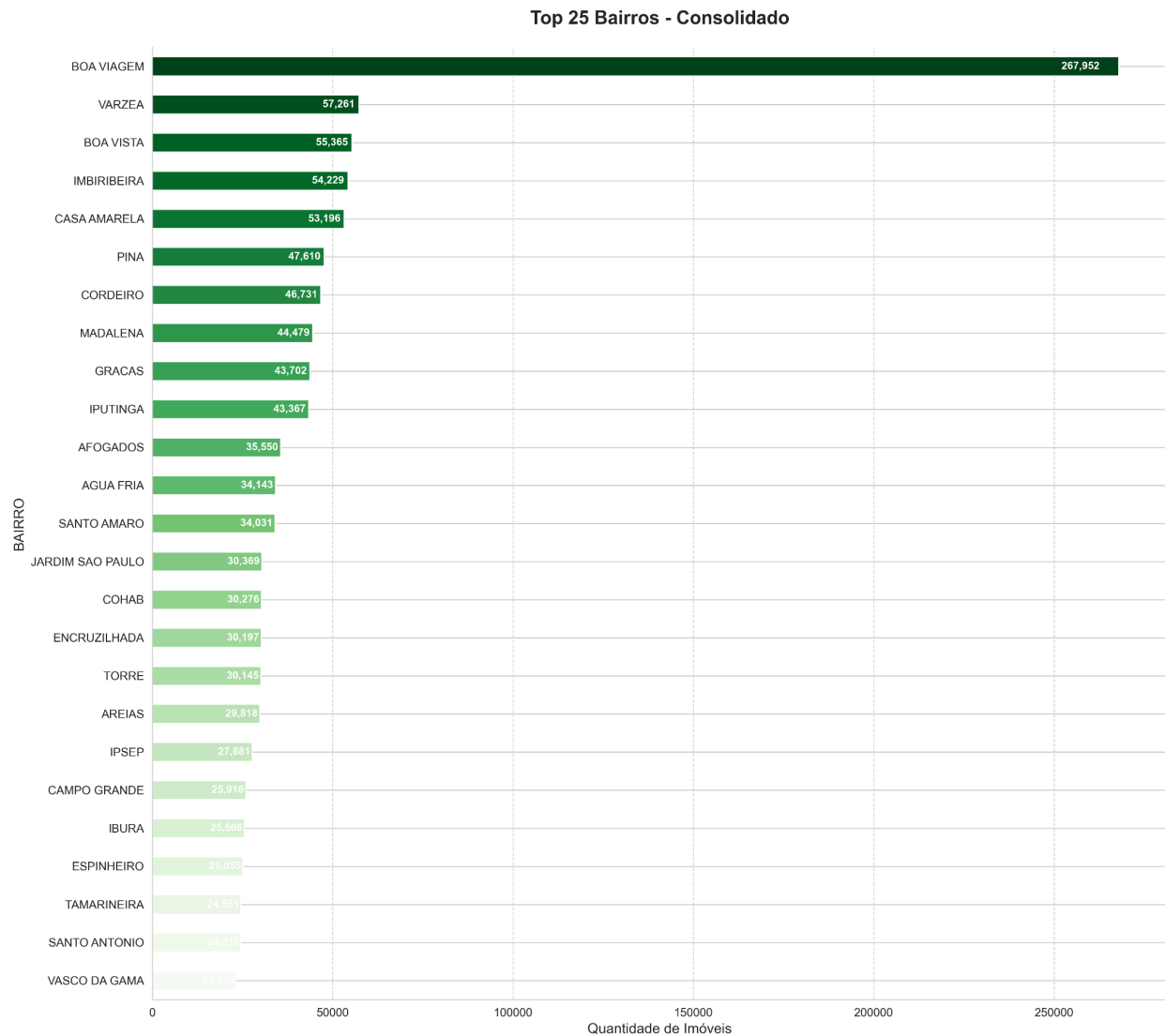
- **Apartamento:** É o tipo de empreendimento mais comum, com **899.440 imóveis**.
- **Casa:** Em segundo lugar, representa **598.477 imóveis**.
- Juntos, esses dois tipos representam a grande maioria do inventário. Outros tipos, como **Sala** (93.176) e **Loja** (66.460), vêm bem abaixo.

Distribuição por Tipo de Construção - Consolidado



A classificação por tipo de construção (gráfico "Distribuição por Tipo de Construção - Consolidado") corrobora a divisão habitacional:

- As duas categorias dominantes são **Apartamento > 4 Pavimentos** (34.9%) e **Casa** (33.9%).
- Em terceiro lugar aparece **Apartamento <= 4 Pavimentos** (14.9%), reforçando que a maior parte do inventário é composta por moradias verticais e horizontais.

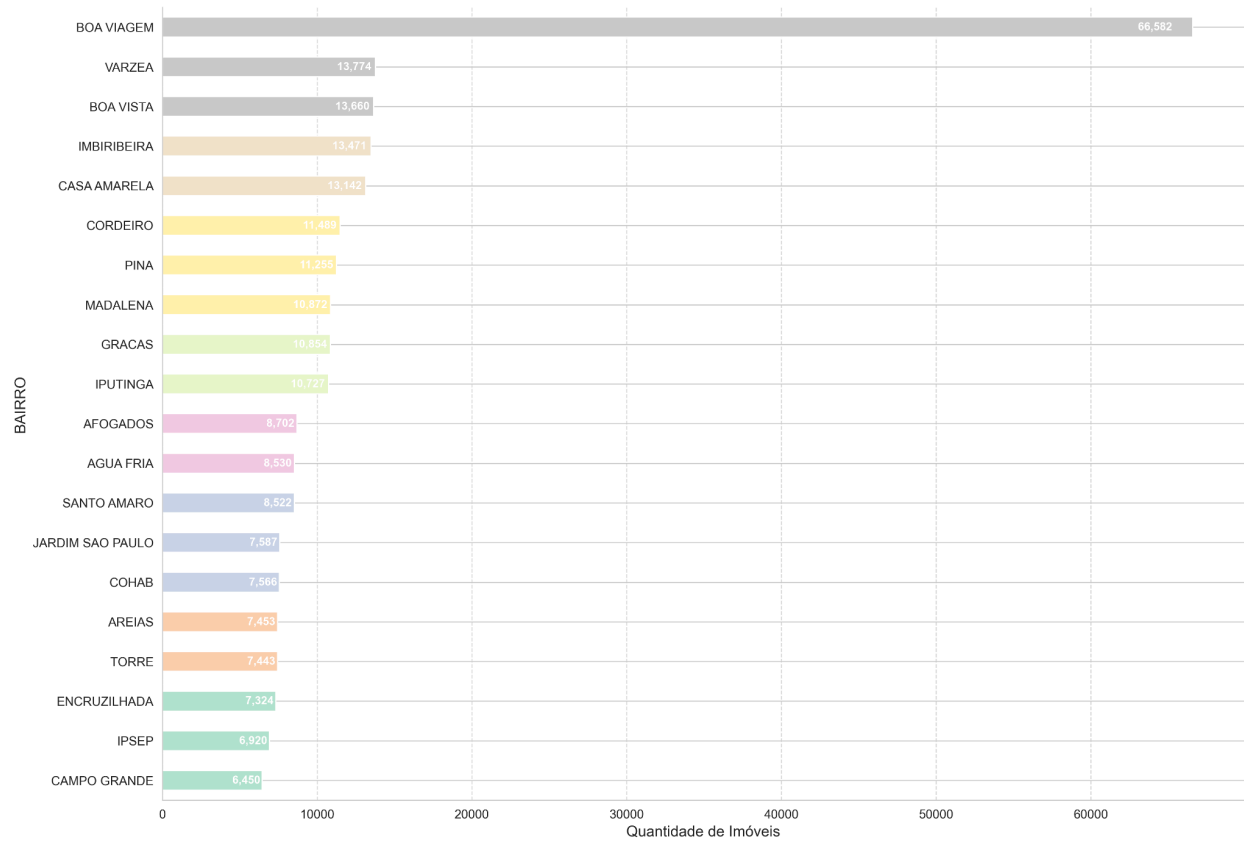


O gráfico "Top 25 Bairros - Consolidado" mostra que o inventário de imóveis está **altamente concentrado** em alguns bairros:

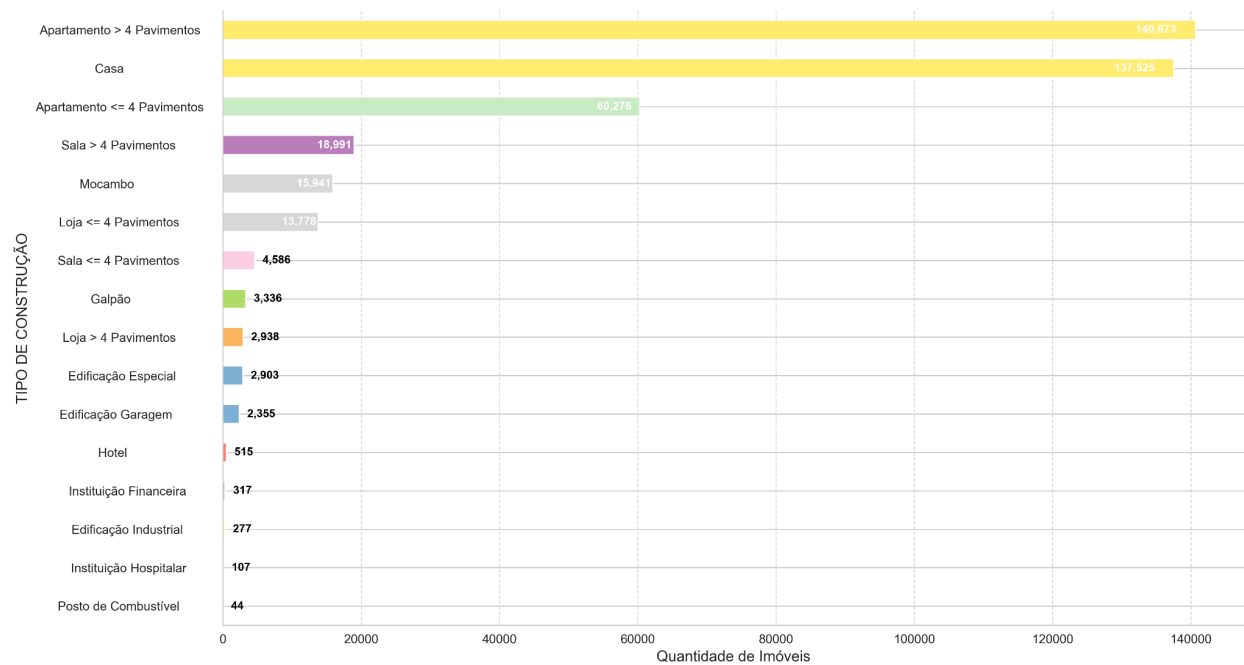
- **Boa Viagem** lidera com uma diferença significativa, totalizando **267.952 imóveis**.
- O segundo e terceiro lugares, **Varzea** (57.281) e **Boa Vista** (55.365), têm um volume quatro a cinco vezes menor que Boa Viagem.
- Essa concentração inicial indica que o maior volume de trabalho de processamento e a maior densidade de valorização estão localizados nos **Top 5 Bairros** (Boa Viagem, Varzea, Boa Vista, Imbiribeira e Casa Amarela), que somam mais de 480 mil imóveis.

Quando a gente vai olhar por ano

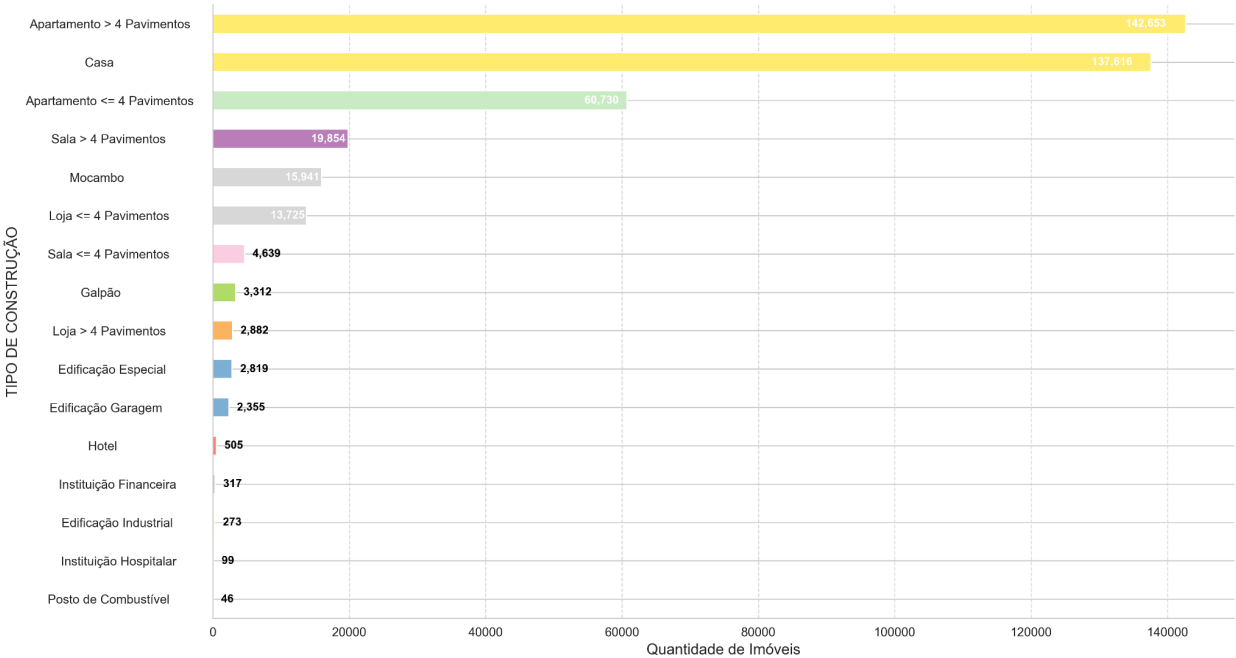
Top 20 Bairros - Ano 2020

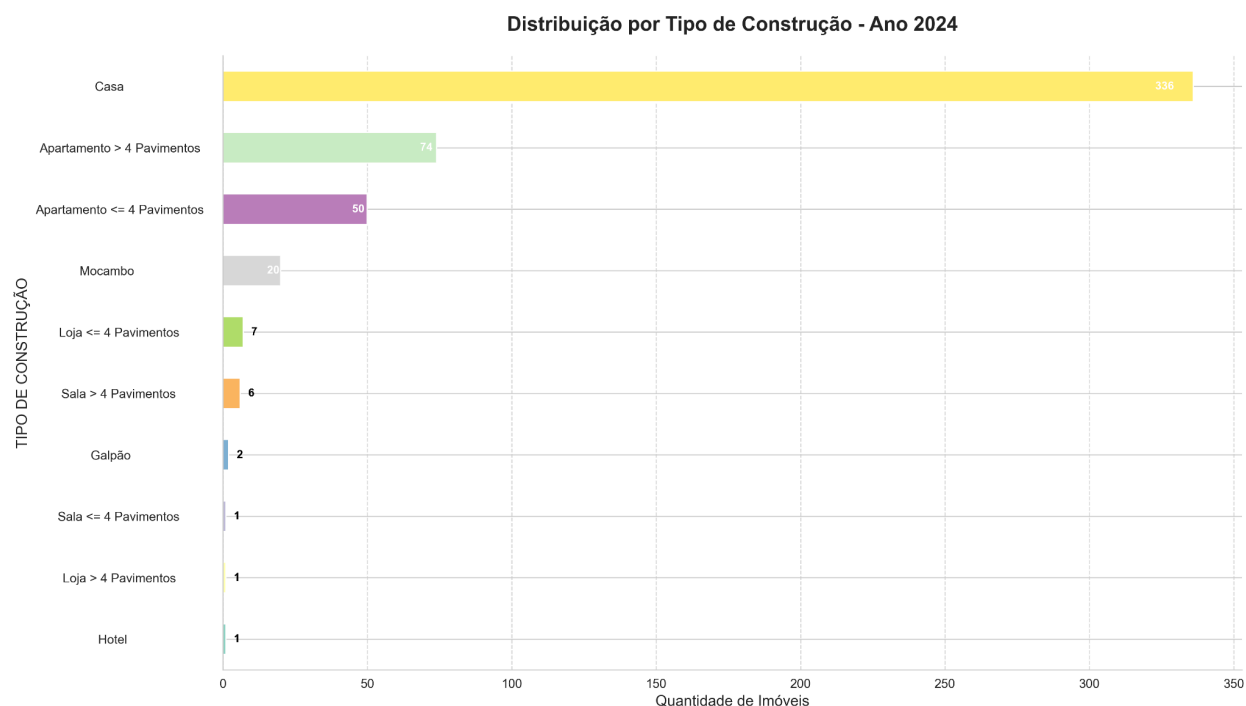
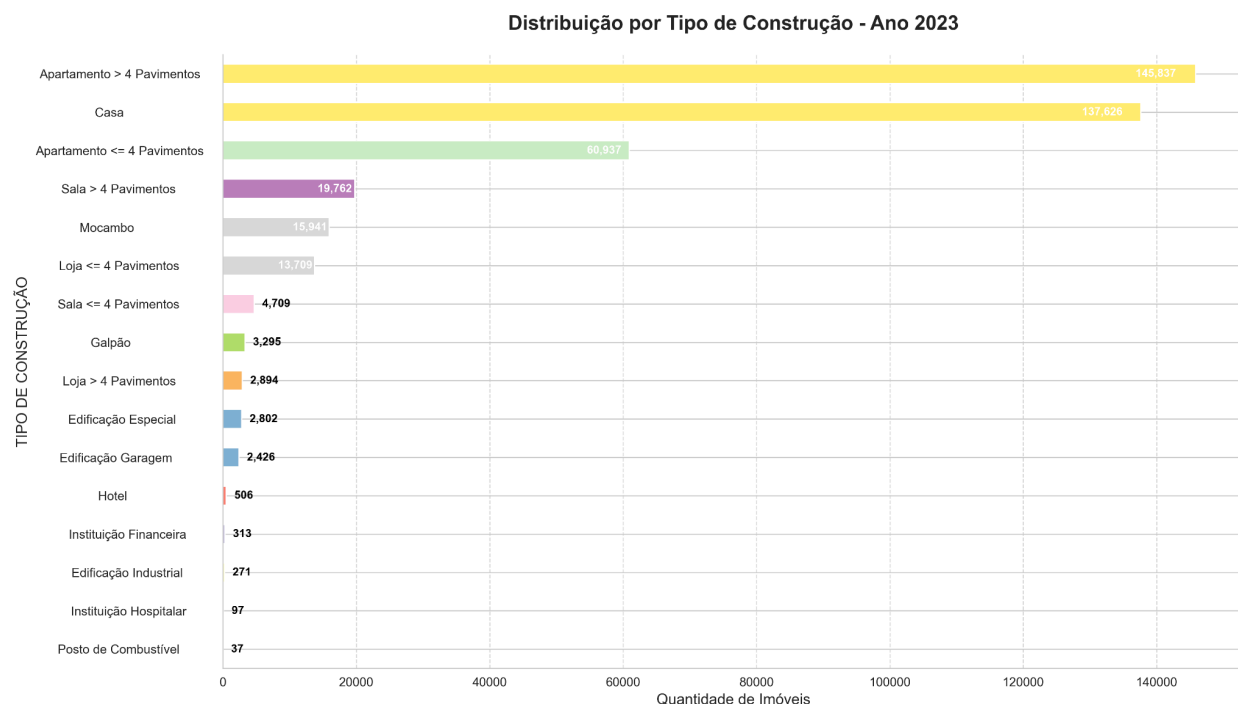


Distribuição por Tipo de Construção - Ano 2021



Distribuição por Tipo de Construção - Ano 2022

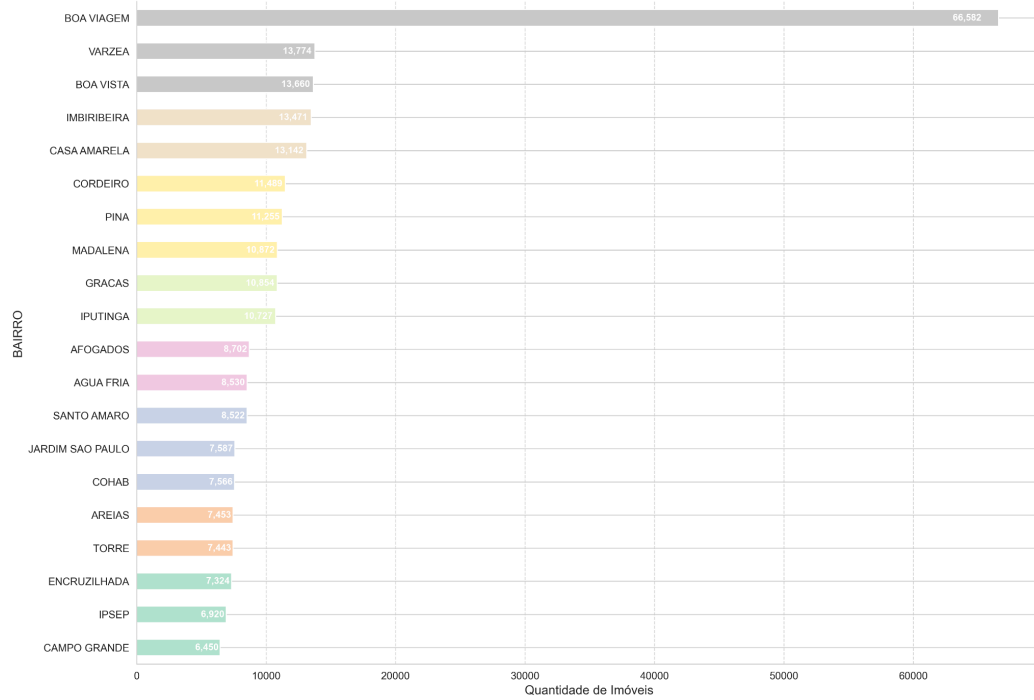




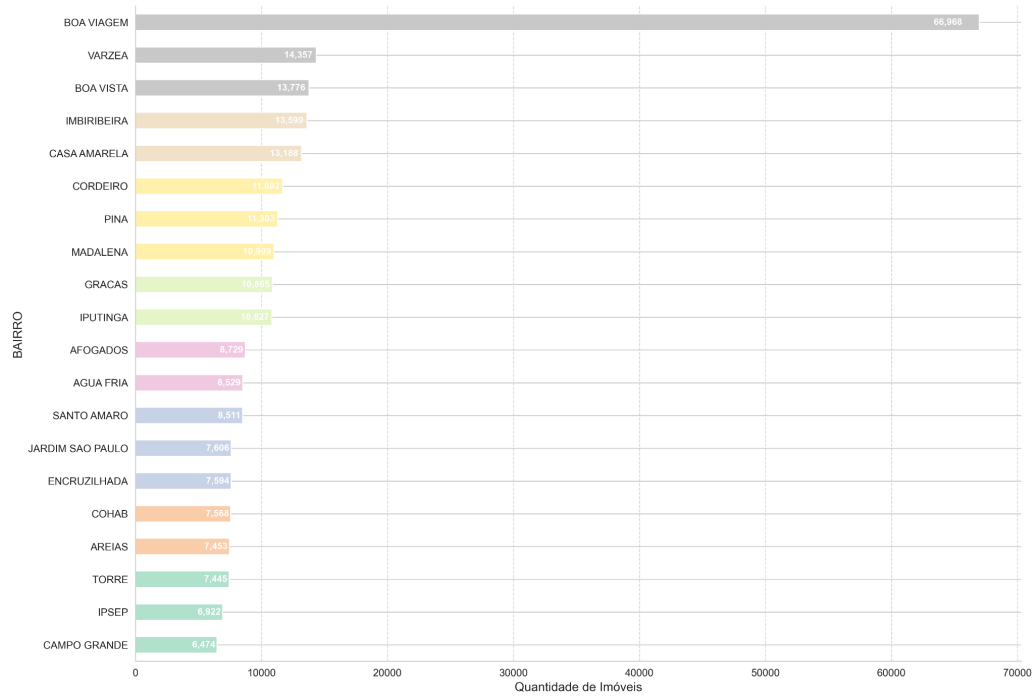
O histórico anual de distribuição por tipo de construção revela uma **estrutura de inventário extremamente estável** entre 2020 e 2023, dominada consistentemente por dois tipos de moradias: **Apartamentos com mais de 4 Pavimentos** e **Casas**. De 2020 a 2023, o volume de Apartamentos acima de quatro pavimentos demonstra um crescimento discreto, passando de **139.873** em 2020 para **143.437** em 2023, confirmando a tendência de **verticalização contínua**. Paralelamente, a categoria **Casa** manteve-se praticamente inalterada nesse mesmo

período, variando de **137.442** em 2020 para **137.526** em 2023, indicando que a base de habitações horizontais não sofreu grandes alterações em volume. Em contraste, o ano de **2024** apresenta uma anomalia grave, com o inventário despencando para apenas 448 imóveis (336 Casas e 74 Apartamentos > 4 Pavimentos), uma redução que sugere fortemente que os dados para este ano estão **incompletos ou parciais**, sendo a análise da distribuição de 2024 estatisticamente irrelevante.

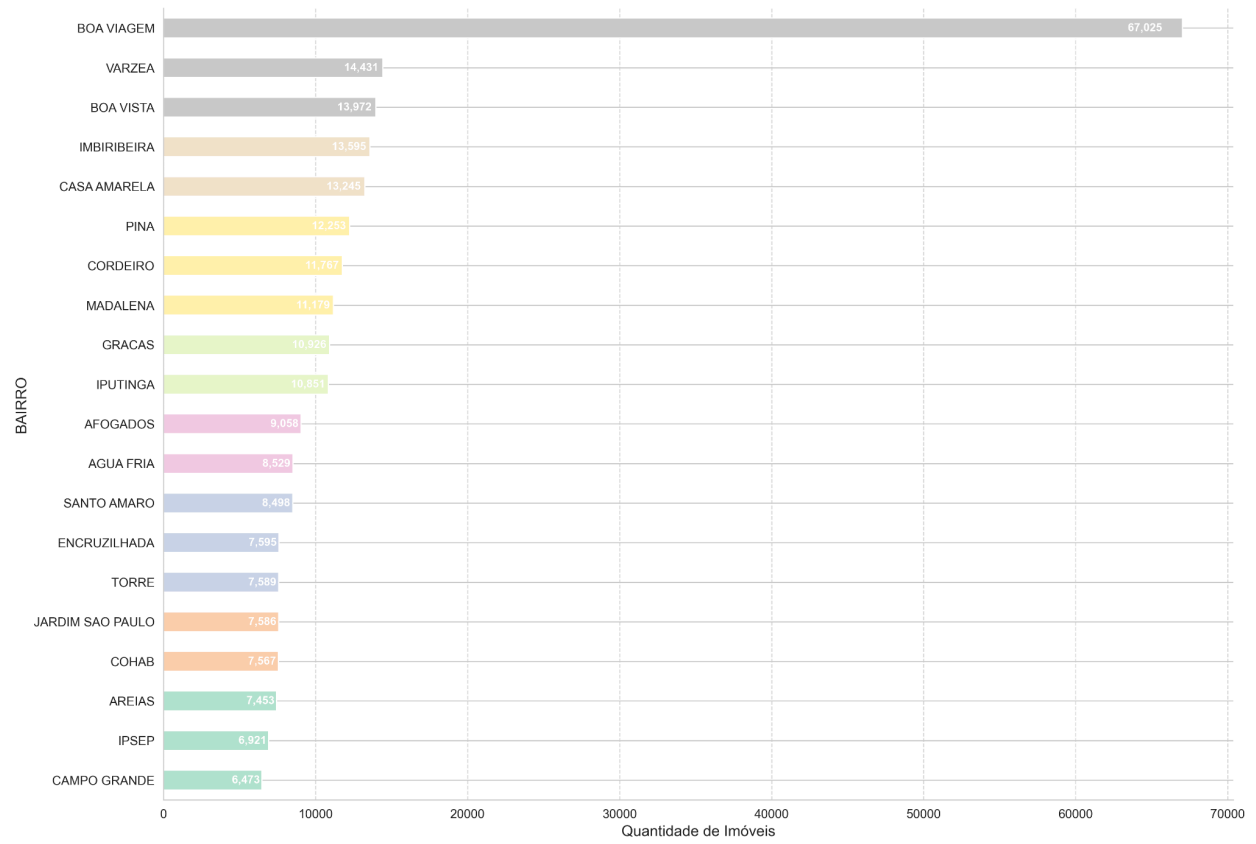
Top 20 Bairros - Ano 2020



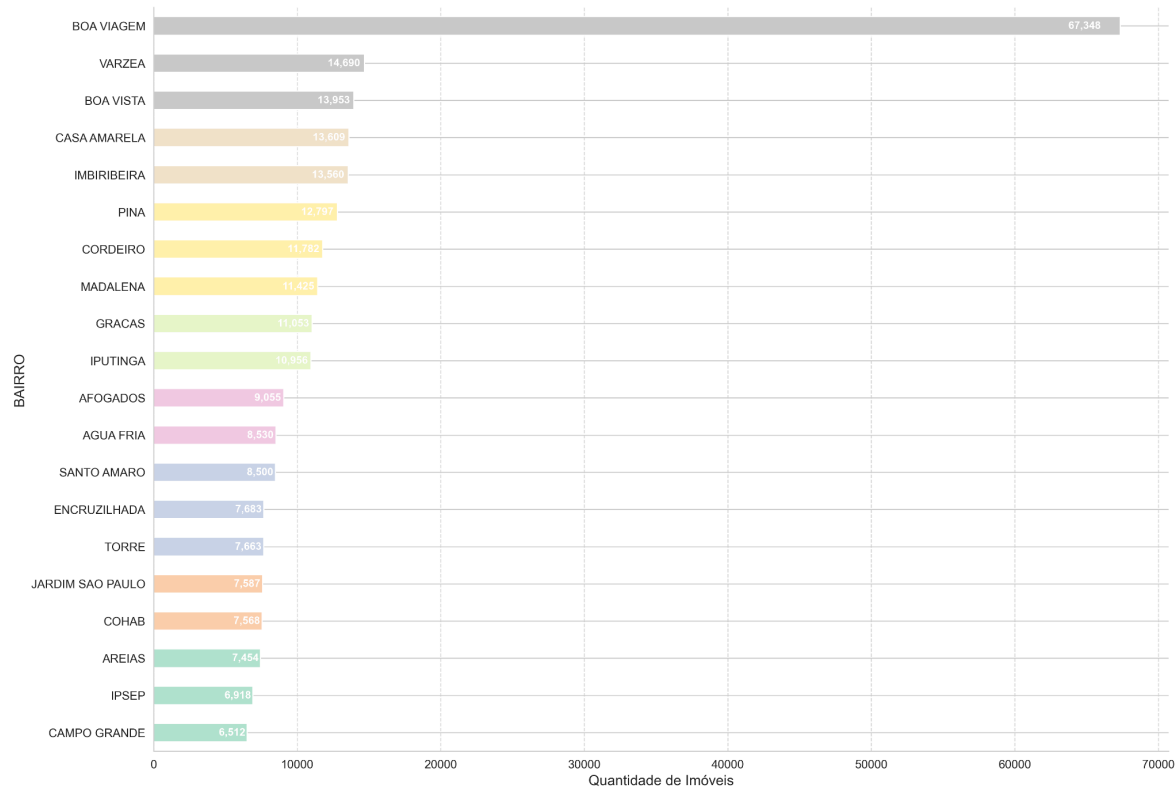
Top 20 Bairros - Ano 2021

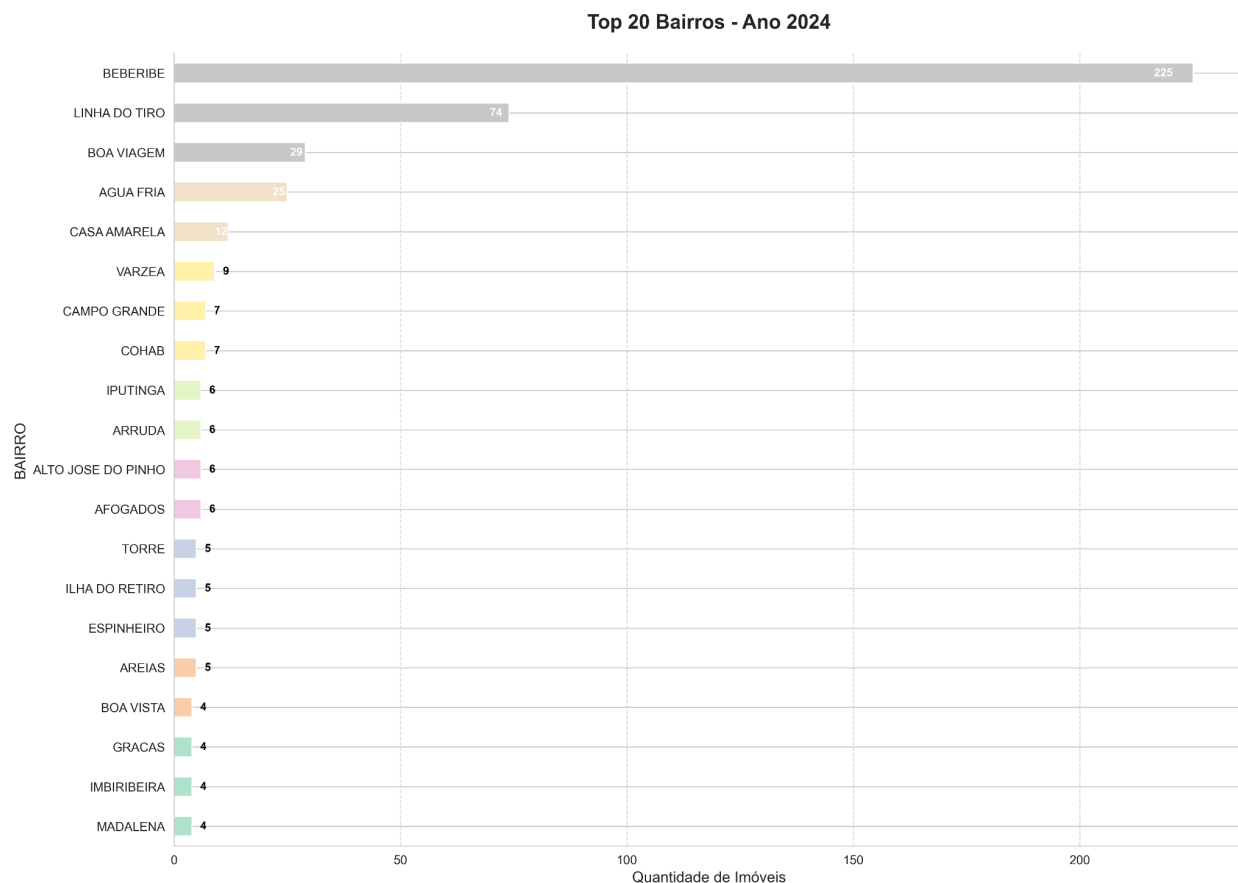


Top 20 Bairros - Ano 2022



Top 20 Bairros - Ano 2023



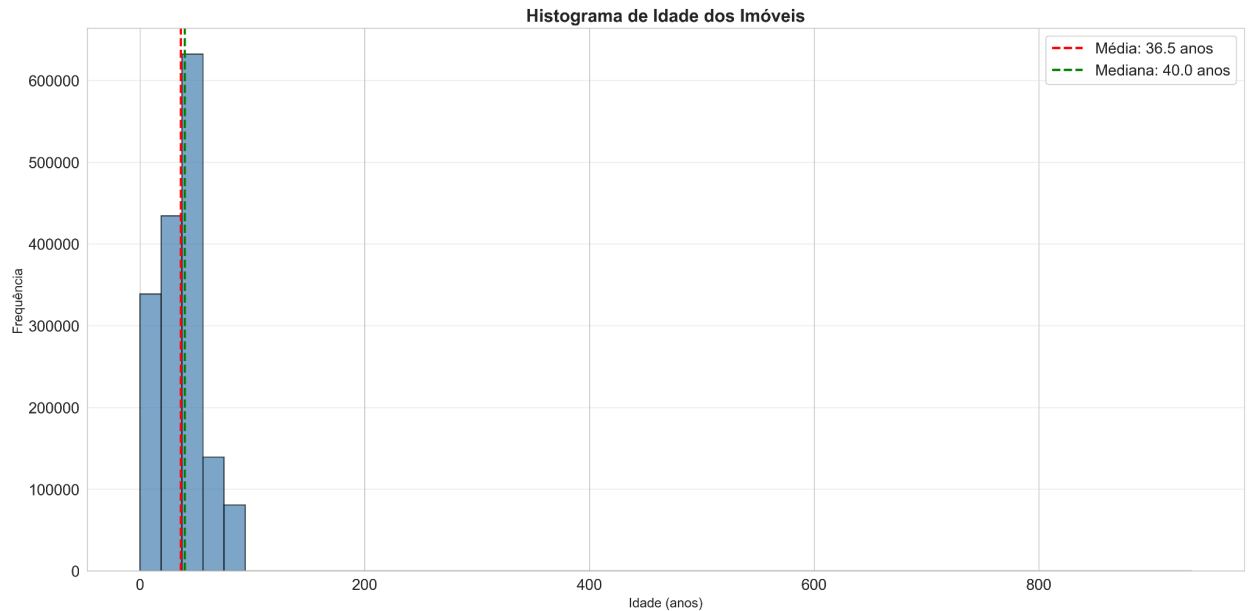


Os gráficos anuais dos **Top 20 Bairros** (2020 a 2023) demonstram uma notável **estabilidade e alta concentração do inventário** em localizações específicas. O bairro **Boa Viagem** manteve-se consistentemente na liderança, com volumes que variaram ligeiramente de **66.582** imóveis em 2020 para **67.348** em 2023. De forma similar, os bairros subsequentes no ranking, como **Várzea**, **Boa Vista**, e **Imbiribeira**, mantiveram suas posições e volumes estáveis ao longo desses anos, confirmando que a distribuição geográfica do inventário de IPTU é **rígida e não sofreu grandes alterações quantitativas** nesse período. Contudo, o gráfico de **2024** apresenta, novamente, uma **discrepância crítica e alarmante**: o volume total de imóveis cai drasticamente, e a lista de bairros é completamente distorcida, com **Beberibe** (225 imóveis) e **Linha do Tiro** (74 imóveis) subitamente no topo, enquanto Boa Viagem aparece com apenas 29 imóveis. Esta anomalia reforça a conclusão anterior de que **os dados de 2024 estão incompletos, corrompidos ou representam apenas uma amostra mínima de teste**, sendo fundamental descartá-los para qualquer análise de evolução ou distribuição histórica.

1.2 Apresente as respostas tanto no histórico consolidado quanto por ano

2. Idade :

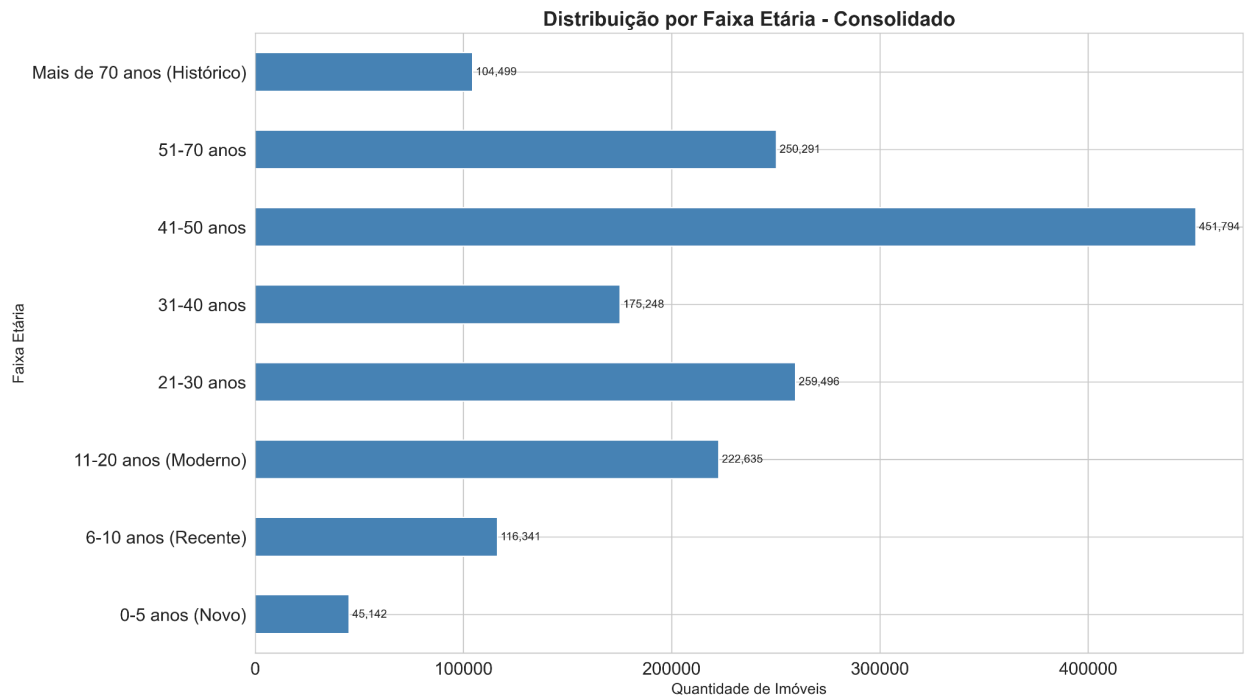
2.1 Como o inventário está distribuído em termos de idade de construção ?



O inventário possui uma idade média e mediana relativamente alta, o que indica uma base patrimonial madura:

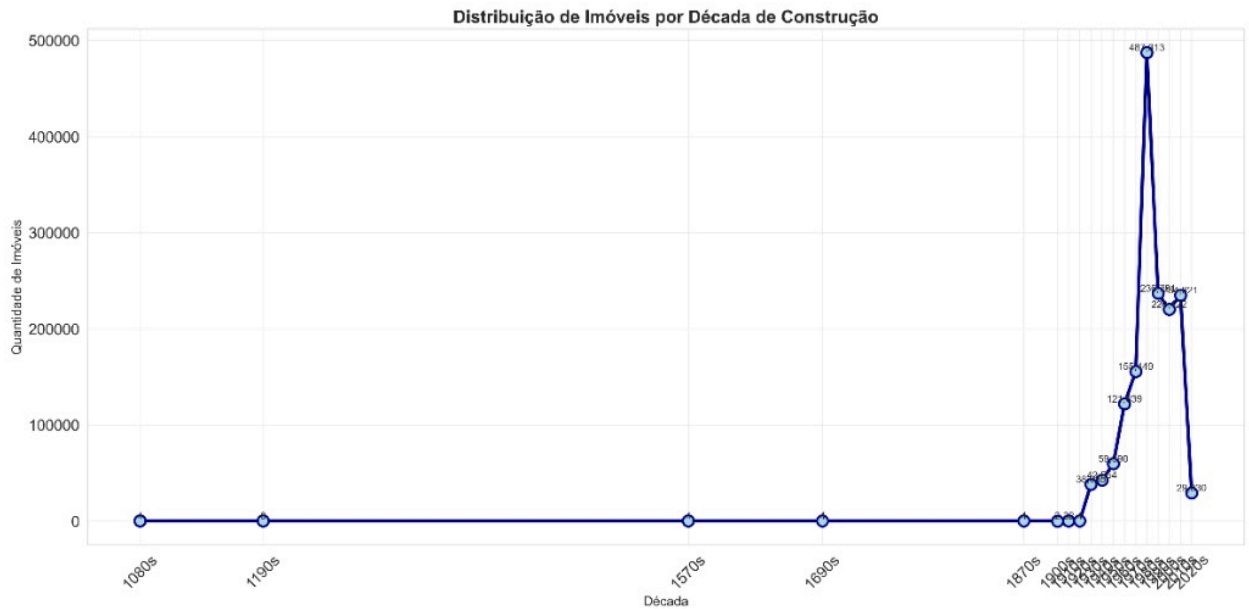
- A **Idade Média** consolidada dos imóveis é de **36.5 anos**.
- A **Idade Mediana** consolidada é ligeiramente superior, em **40.0 anos**.

O histograma (gráfico "Histograma de Idade dos Imóveis") mostra que a distribuição é fortemente **assimétrica positiva (cauda longa)**, com uma grande concentração de imóveis nas primeiras décadas e uma pequena quantidade de imóveis **extremamente antigos** (acima de 200 anos), que elevam levemente a média.



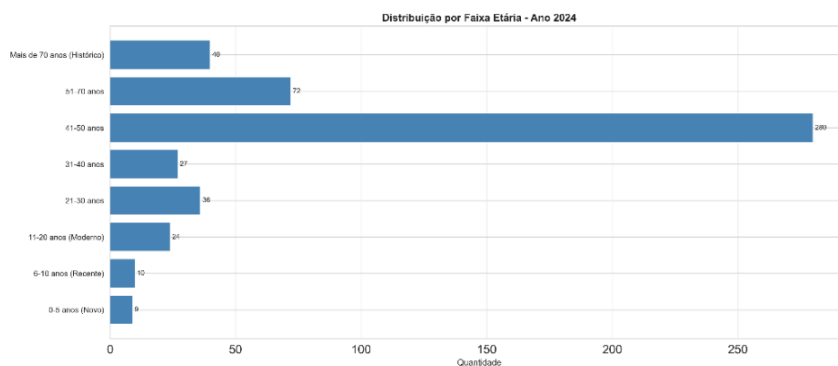
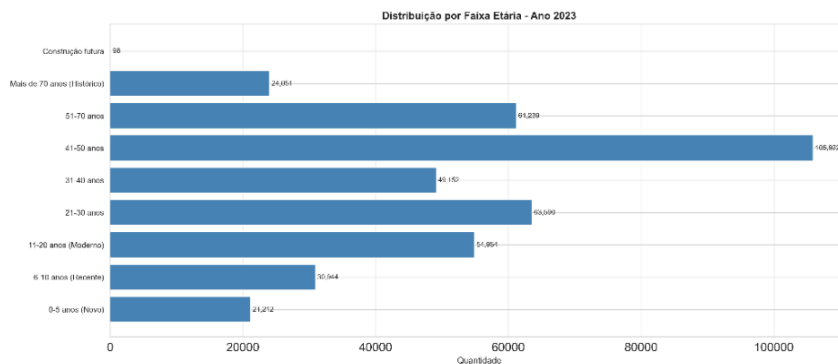
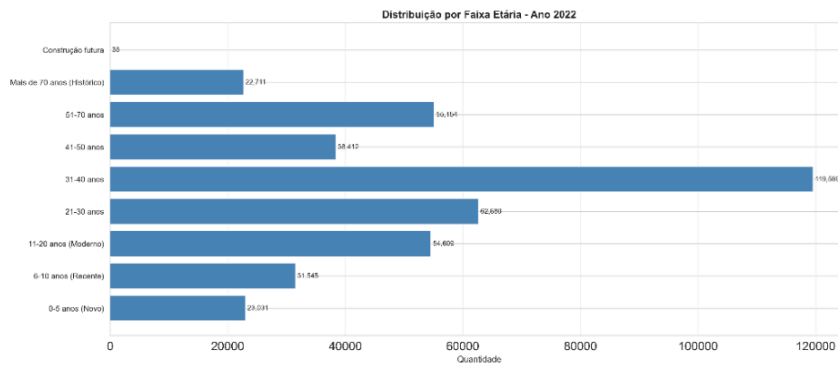
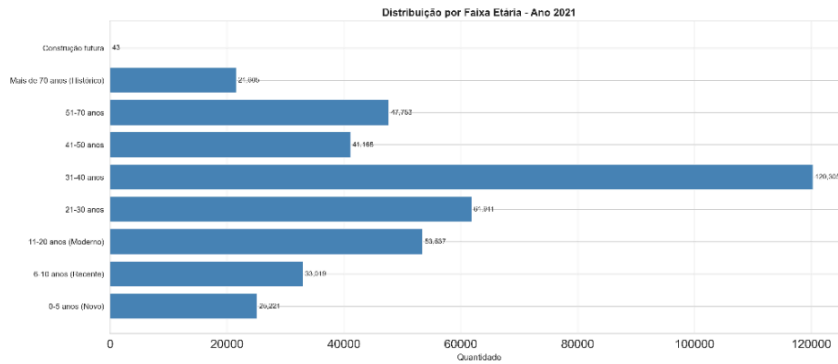
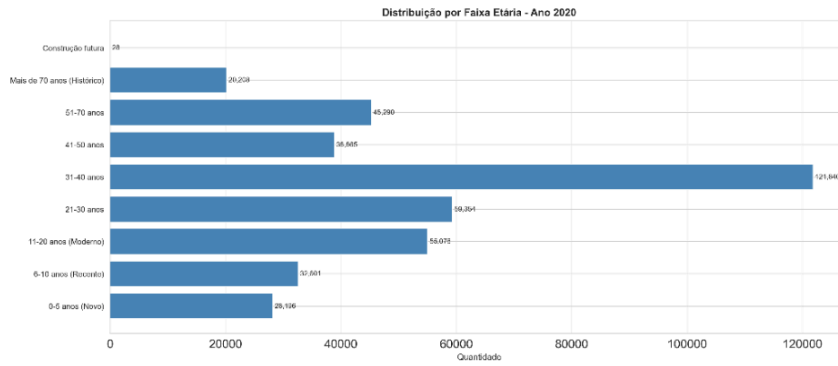
A distribuição detalhada do inventário por faixas etárias (gráfico "Distribuição por Faixa Etária - Consolidado") mostra onde se concentra o maior volume:

- As faixas de **41-50 anos** e **51-70 anos** dominam o inventário. A faixa de **41-50 anos** é a maior, com **27.8%** (451.794 imóveis), seguida pela faixa de **51-70 anos** com **15.4%** (250.291 imóveis).
- Quase metade (aproximadamente **43.2%**) dos imóveis têm **mais de 40 anos**.
- O inventário mais **Novo/Recente** (0-10 anos) representa uma parcela pequena, somando apenas **9.9%** do total (45.142 imóveis de 0-5 anos e 116.341 imóveis de 6-10 anos).



A visão por década (gráfico "Distribuição de Imóveis por Década de Construção") corrobora a maturidade do inventário, mostrando picos de construção recente:

- O maior volume de construção concentrou-se na **década de 2000s**, com **486.130 imóveis**.
- As décadas seguintes (2010s e 2020s) mantiveram um volume elevado, mas decrescente: **228.627 imóveis** na década de 2010s e **70.040 imóveis** na década de 2020s (dados parciais).
- Os picos de construção mostram que a expansão imobiliária mais intensa ocorreu entre os anos 1990 e 2009.

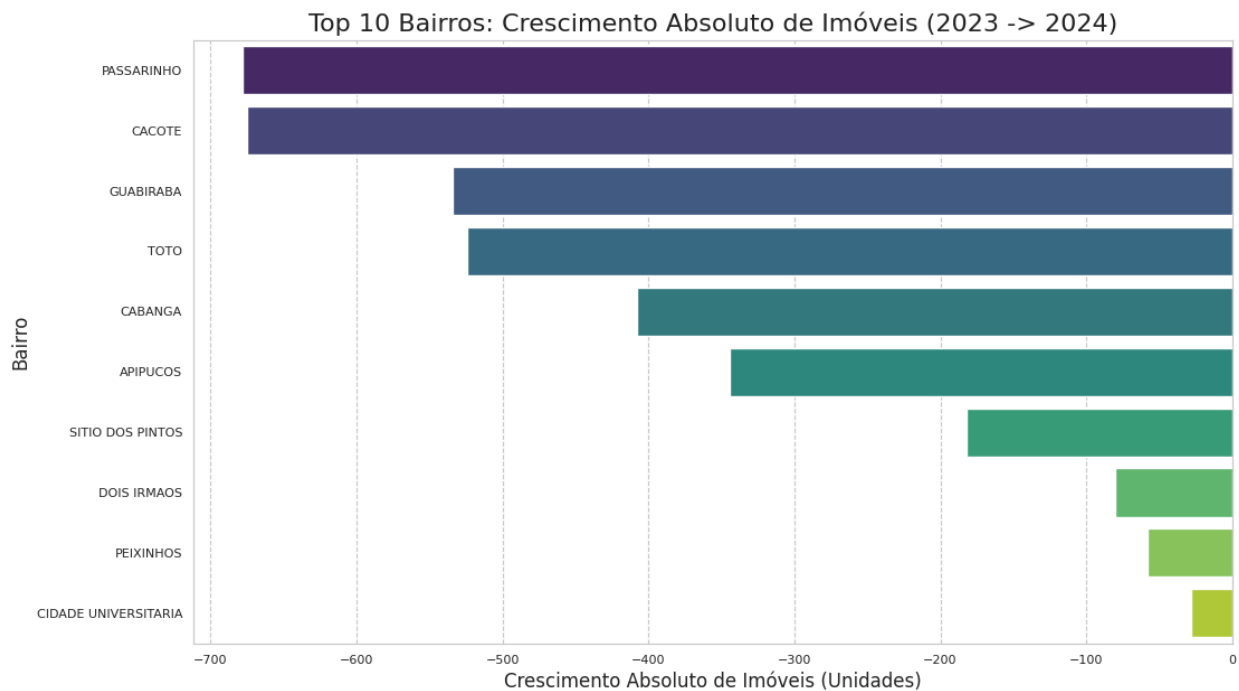


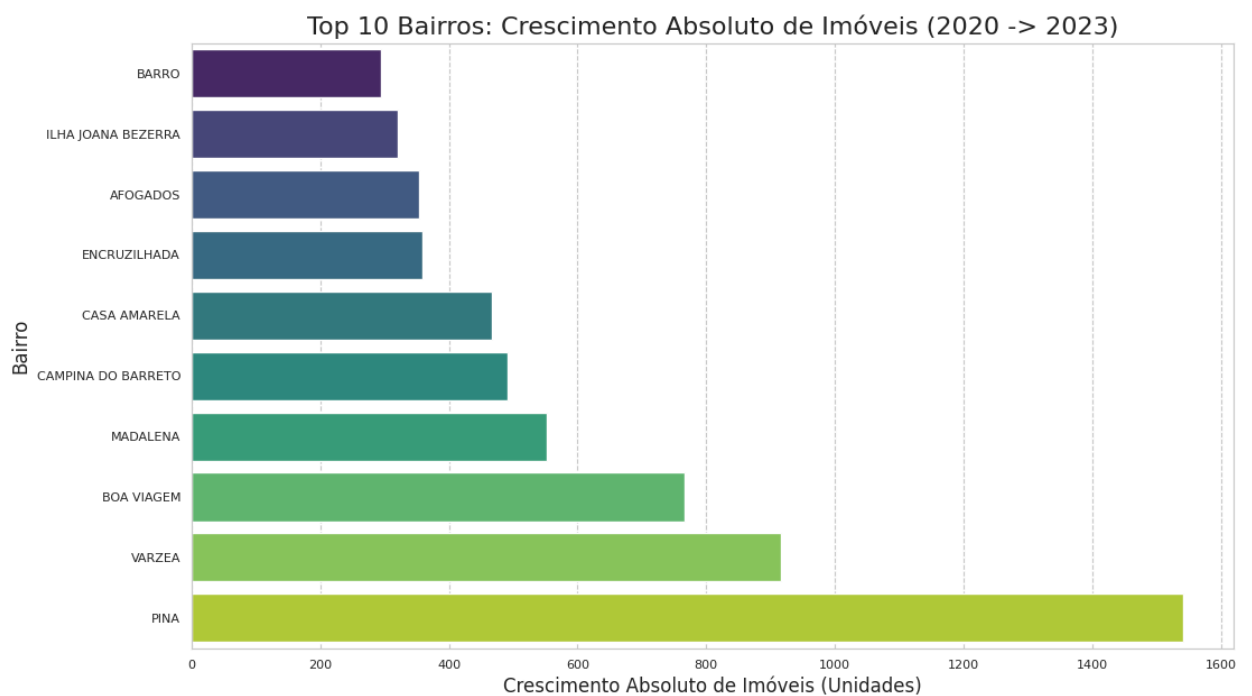
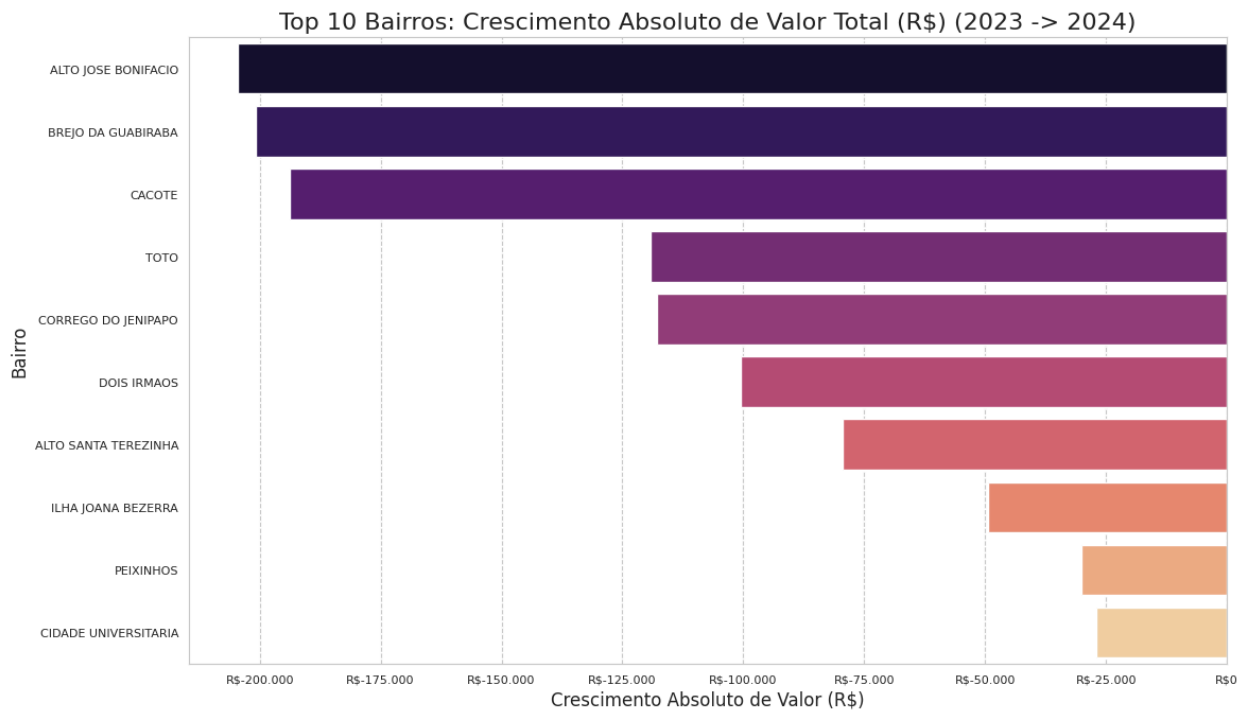
O comparativo anual da idade (gráfico "Comparativo de Idade Média e Mediana por Ano") mostra que o inventário está **envelhecendo progressivamente**:

- A **Idade Média** subiu de **32.8 anos** em 2020 para **35.0 anos** em 2023.
- A **Idade Mediana** também subiu de **37.0 anos** em 2020 para **39.0 anos** em 2023.
- Essa tendência constante de aumento na idade média e mediana, embora sutil, confirma que o ritmo de novas construções (imóveis de 0-5 anos) não é suficiente para rejuvenescer significativamente a base de imóveis mais antigos, levando ao envelhecimento geral do inventário. A grande elevação em 2024 é devida à incompletude dos dados, como observado nas análises anteriores.

2.2 Apresente as respostas tanto no histórico consolidado quanto por ano

3. Valor :
- 3.1 Quais bairros apresentam maior evolução em número de imóveis ? E em relação a valor ?



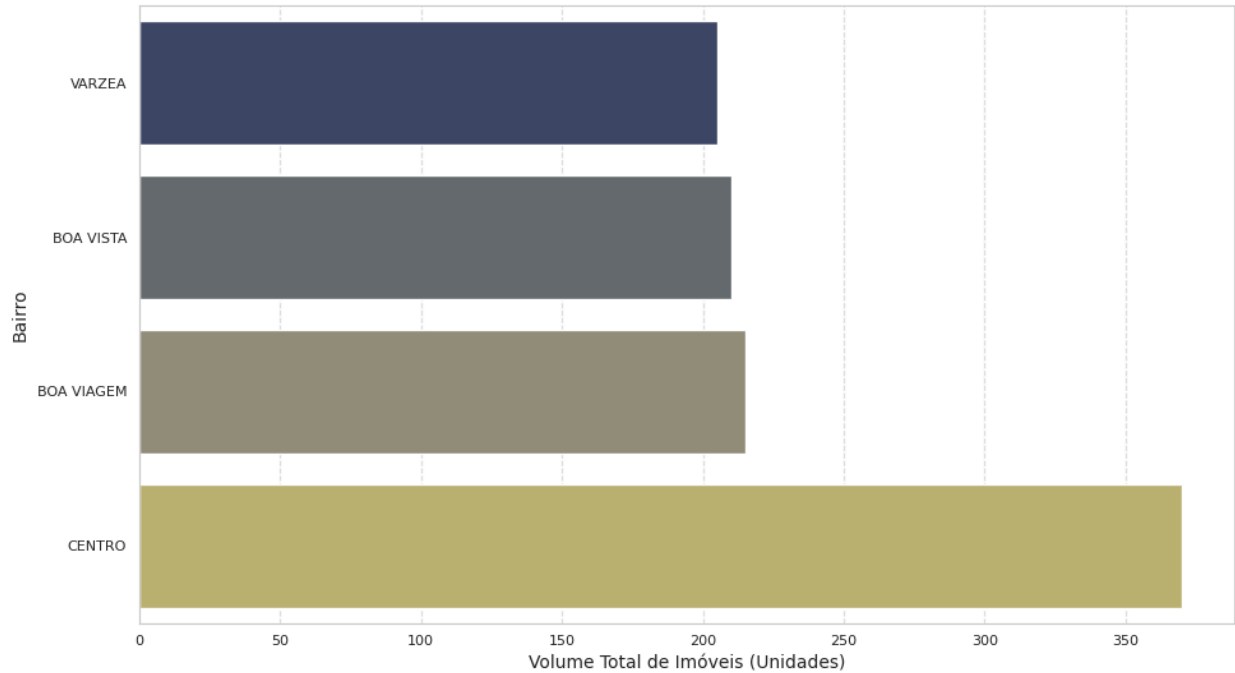


3.2 Apresente as respostas tanto no histórico consolidado quanto por ano

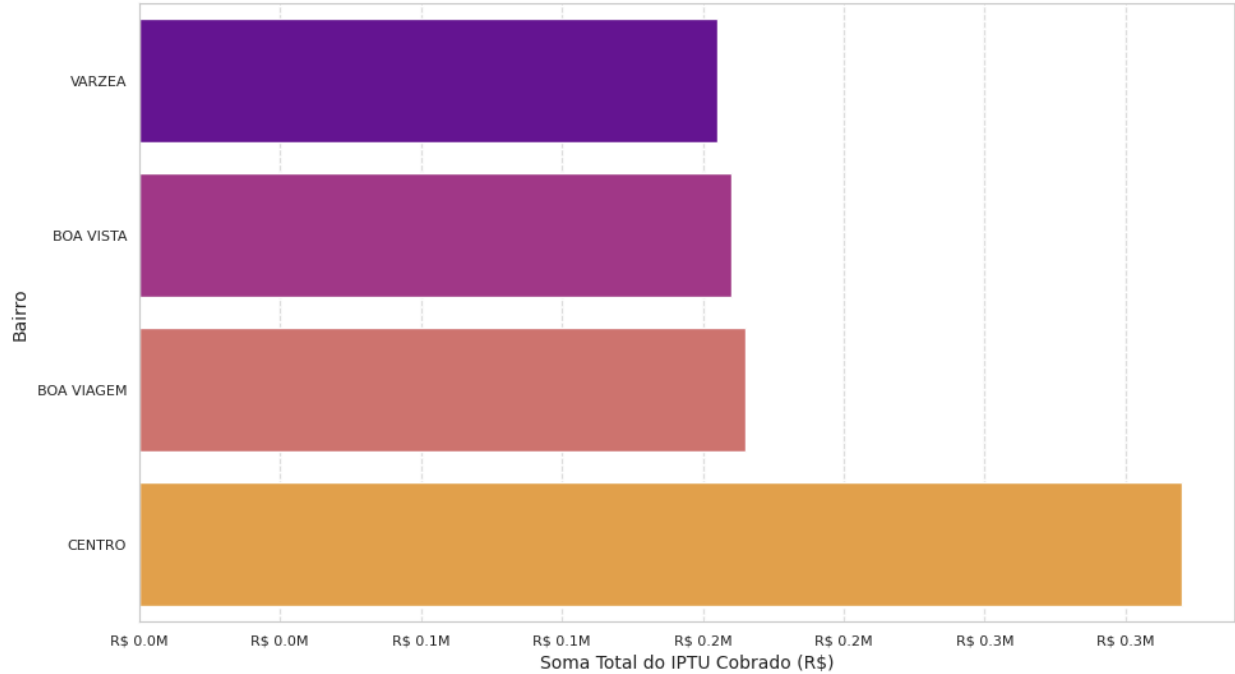
4. Evolução :

4.1 Quais bairros apresentam maior evolução em número de imóveis ? E em relação a valor ?

Top 10 Bairros por Volume Total de Imóveis (Estoque Consolidado 2020 - 2024)



Top 10 Bairros por Soma Total do Valor IPTU Cobrado (Estoque Consolidado 2020 - 2024)



Bairro	Crescimento Absoluto de Imóveis
VARZEA	-4.8
BOA VISTA	-0.5
CENTRO	-0.2
BOA VIAGEM	5.0

A horizontal bar chart titled 'Crescimento Absoluto de Valor (R\$)' on the x-axis. The y-axis is labeled 'Bairro'. The chart displays the absolute value growth for four neighborhoods: VARZEA, BOA VISTA, CENTRO, and BOA VIAGEM. VARZEA has a negative growth of approximately -R\$ 0.01M, represented by an orange bar. BOA VISTA and CENTRO have zero growth, represented by light gray bars. BOA VIAGEM has a positive growth of approximately R\$ 0.01M, represented by a dark brown bar. The x-axis has labels for R\$ -0.0M, R\$ 0.0M, and R\$ 0.0M.

Bairro	Crescimento Absoluto de Valor (R\$)
VARZEA	-R\$ 0.01M
BOA VISTA	R\$ 0.00M
CENTRO	R\$ 0.00M
BOA VIAGEM	R\$ 0.01M

4.2 Apresente as respostas tanto no histórico consolidado quanto por ano

Com base na análise dos gráficos, o inventário de imóveis em Recife demonstra um mercado **maduro e consolidado**, onde **apartamentos em prédios altos (34.9%) e casas (33.9%)** dominam o estoque. A **idade média dos imóveis é de 36.5 anos**, com a maior concentração de construções datando das faixas de **41-50 anos (27.8%) e 21-30 anos (16.8%)**, enquanto imóveis **novos (0-5 anos)** representam uma fatia pequena de apenas **2.8%**. A evolução do inventário de 2020 a 2023 foi **estável, com crescimento marginal** (de 401.453 para 411.162 unidades). Na dinâmica anual mais recente e válida (2022 vs 2023), o crescimento absoluto de estoque e de valor (IPTU) esteve **focado no bairro de Boa Viagem**, que registrou o maior ganho em volume e valor (IPTU), enquanto bairros como Várzea apresentaram leve redução. Em resumo, o mercado é antigo, tem um ritmo lento de renovação e o investimento em novos empreendimentos de alto valor está concentrado em seu principal polo: Boa Viagem.