

## **Relatório do processo de AED e Limpeza - Paraciclos**

A base escolhida foi de paraciclos na cidade de niterói, extraída desse link: [Paraciclos](#)

### **Visão geral:**

Inicialmente, os dados passaram por um processo de descrição para entender qual a tipologia de cada variável e, a partir disso, tomar decisões de como trabalharia com cada uma.

A seguir, identifiquei valores nulos e ausentes e, a partir disso, deu-se início ao processo de limpeza.

### **Limpeza:**

As colunas foram renomeadas com nomes mais intuitivos a fim de facilitar o entendimento.

Padronizei as variáveis categóricas. Primeiro observando como os dados apareciam e, a partir disso, defini como cada uma ficaria nas colunas material, tipo e órgão responsável.

NA:

Os NA's das variáveis numéricas foram substituídas pela mediana (robusta contra outliers);

Os NA's das variáveis categóricas foram substituídas pela moda (retorna o valor mais frequente, ideal para categóricas).

Feito isso, visualizamos os dados novamente para garantir que tudo foi ajustado até aqui.

O tratamento de outliers das variáveis numéricas não foi necessário. Essa decisão veio após o processo de observação da disposição dos dados ao longo do trabalho e, além disso, o que tive de outcome como classificado como outliers na verdade não fazia sentido para o contexto.

Com os dados categóricos e numéricos tratados, chegamos à visualização.

Com as estatísticas descritivas das variáveis numéricas, foi interessante observar que:

- Localização geográfica (longitude\_utm, latitude\_utm) Os paraciclos estão concentrados entre 692.966 e 695.475 UTM de longitude, e entre 7.464.173 e 7.467.020 UTM de latitude.

Isso mostra uma área geográfica relativamente compacta, como esperado para um município.

- Ano de instalação (ano\_instalacao) Mediana = 2018, moda = 2017 → indica um pico de instalações entre 2017–2018.

Desvio padrão = 2,73 → variação pequena, ou seja, os paraciclos foram instalados em um período relativamente concentrado.

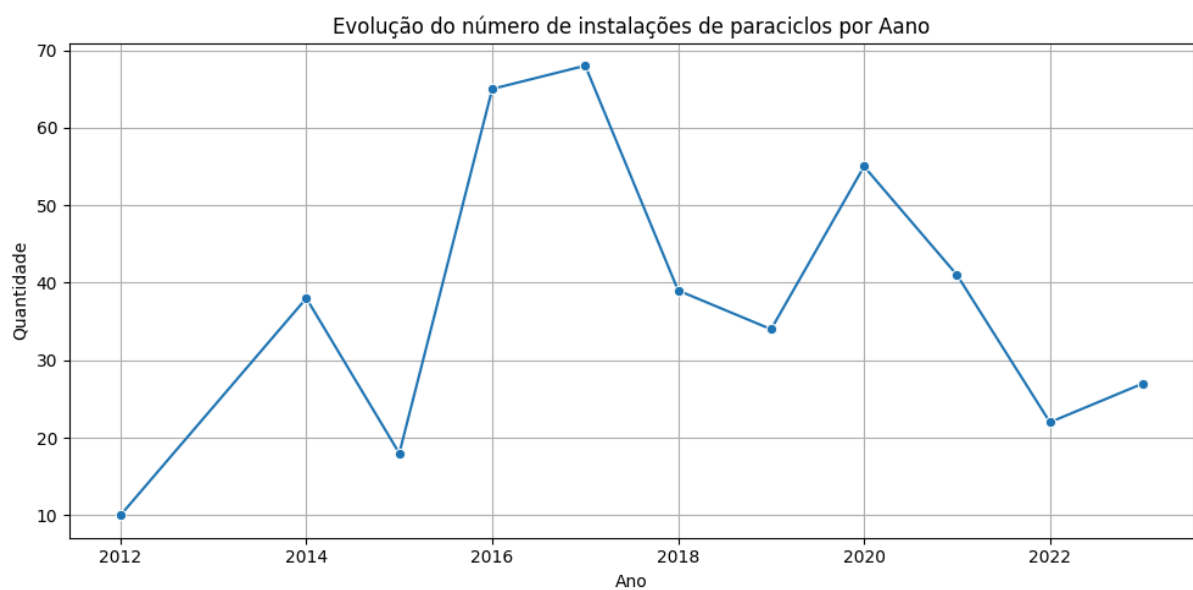
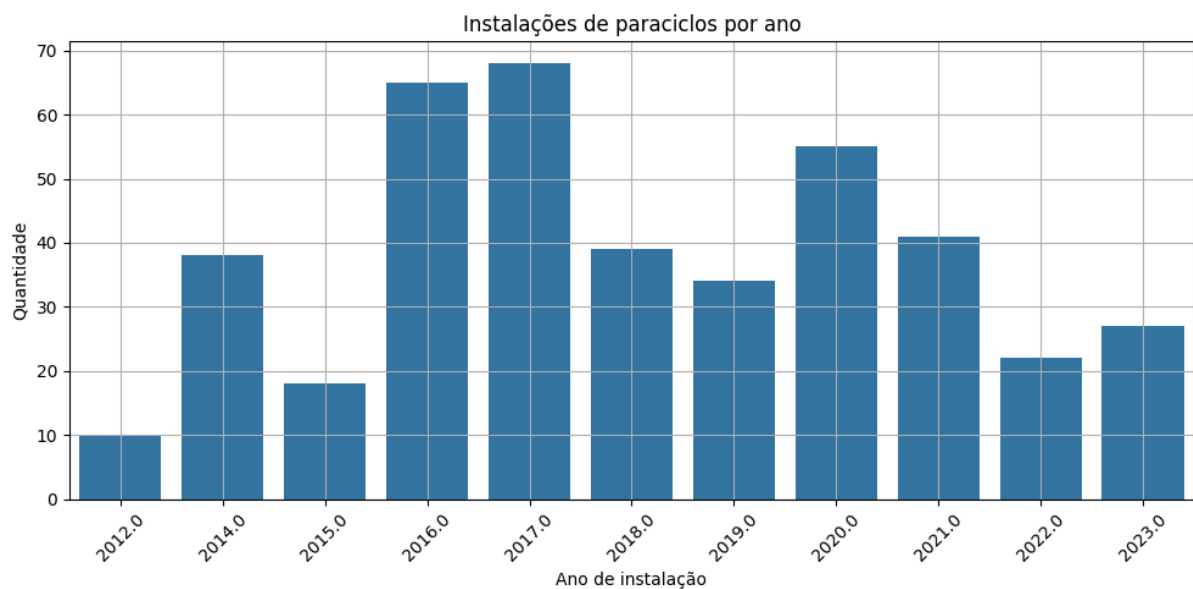
- Número de vagas (num\_vagas) Média = 8,5, mediana = 6, moda = 4, Q3 = 10

Distribuição assimétrica à direita: poucos pontos com alto número de vagas puxam a média para cima.

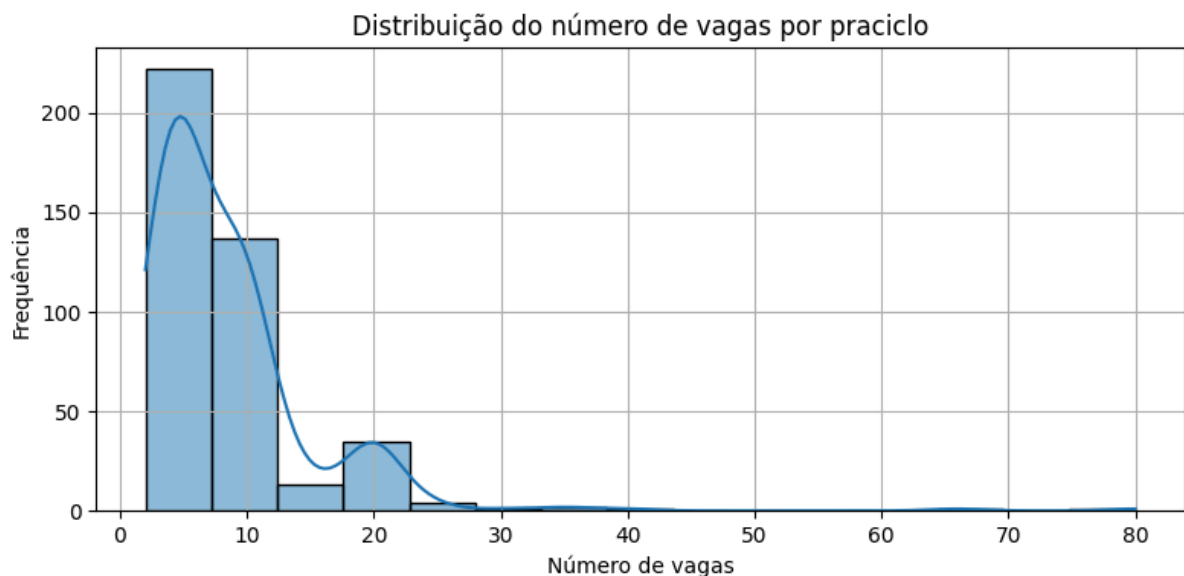
Desvio padrão = 7,27 → variação alta entre o número de vagas dos paraciclos.

A partir desses dados, a decisão foi de fazer:

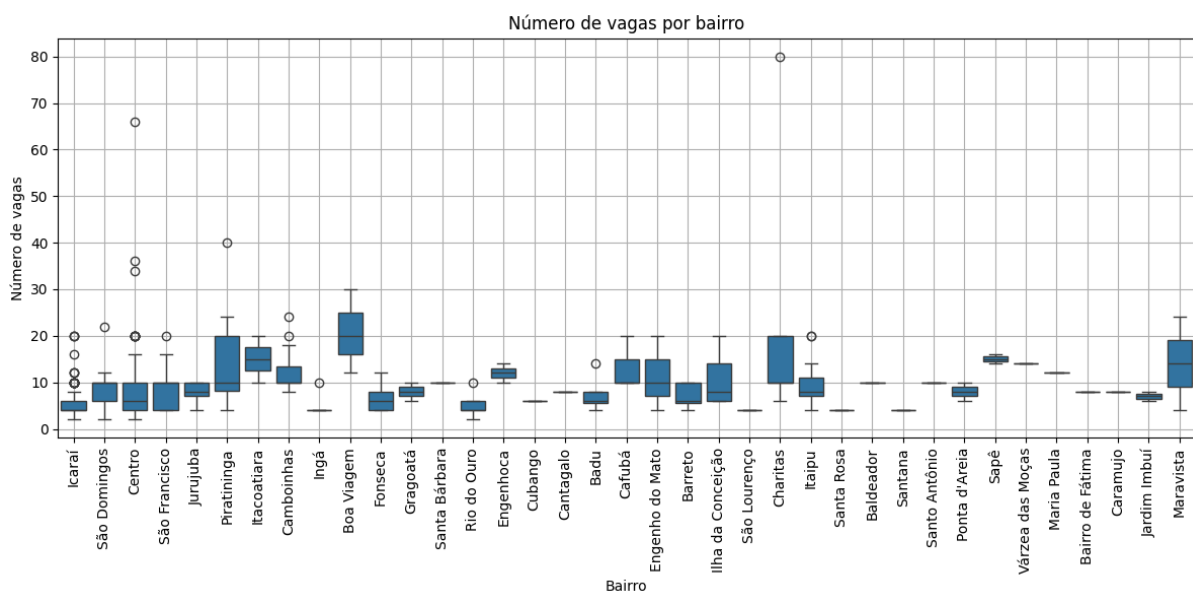
### Histogramas, boxplots e mapa



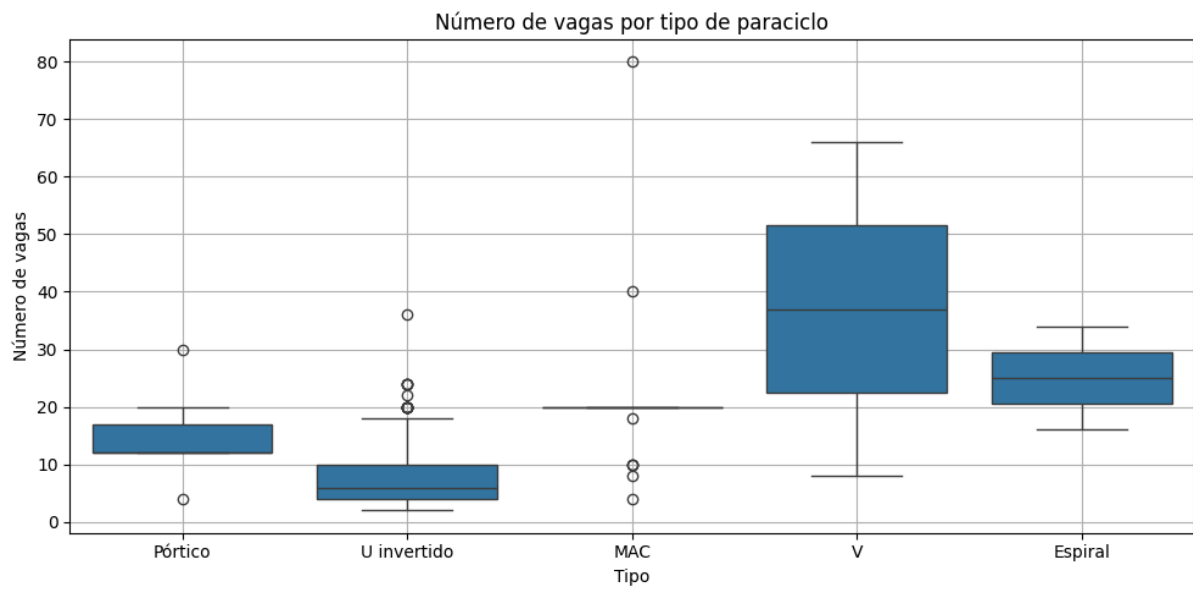
Com esse histograma e gráfico de linhas, fica evidente um pico de instalação de paraciclos em 2016 e 2017. Entender o motivo disso pode ser uma boa estratégia para se repetir ou aumentar significativamente o processo novamente caso faça sentido, ou aplicar esse esforço em políticas assessórias, mas que estejam no tema de ciclovias, mobilidade ativa.



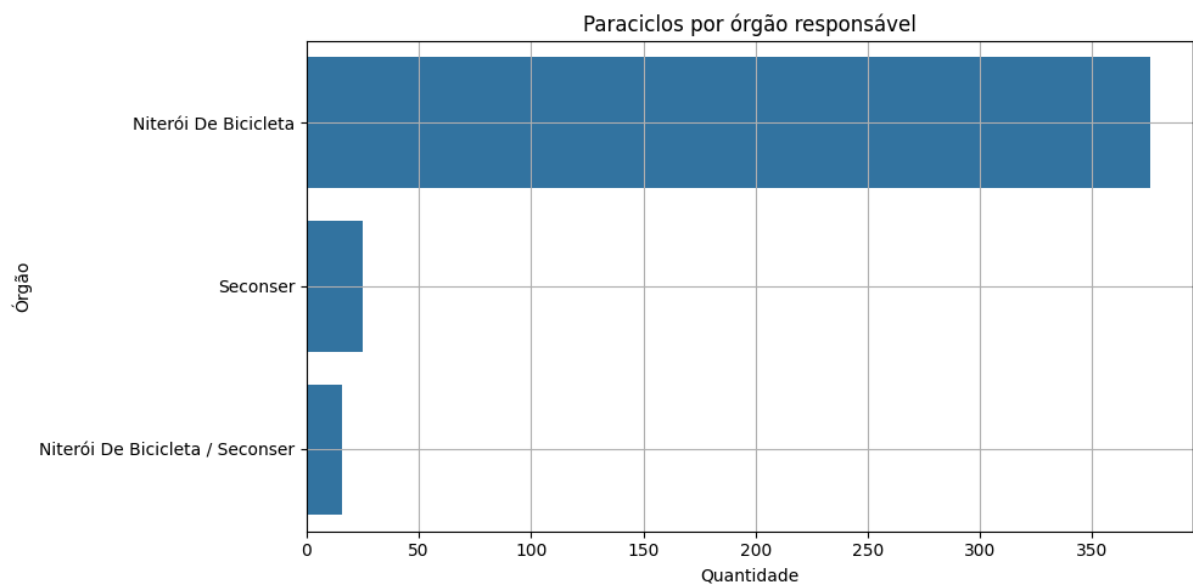
Esse histograma aparenta ter um teor meramente descritivo. Sem maiores análises.



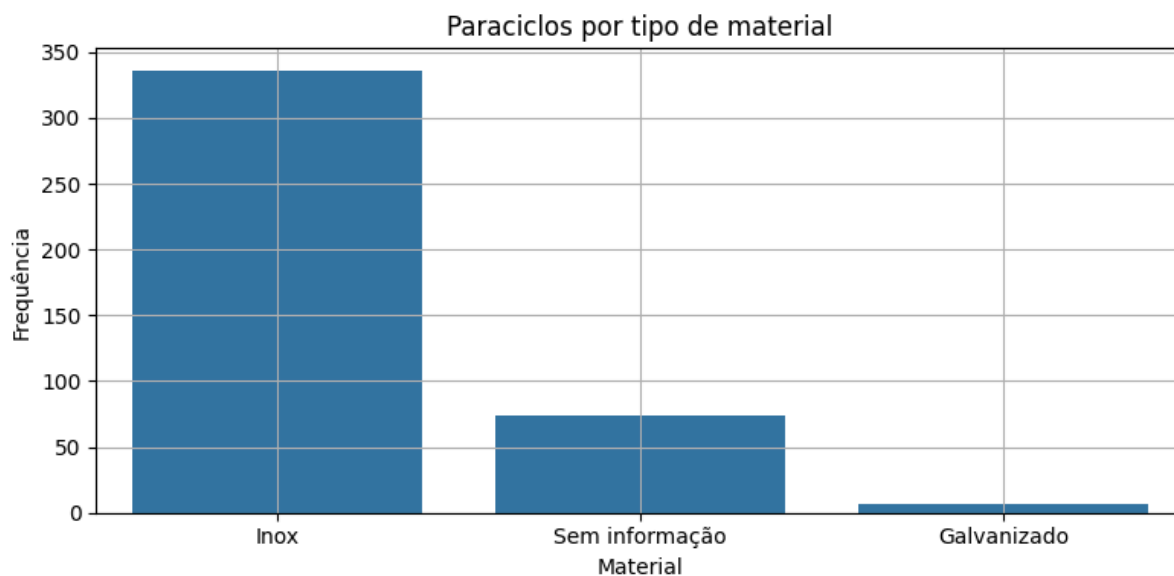
Esse gráfico é interessante para pensarmos em análises do tipo: a disposição dos paraciclos é feita de que forma? Volume populacional na área? Volume de ciclorotas? Com esse gráfico e o mapa que veremos mais abaixo fica evidente a disposição “à beira da praia” e sim, o centro de Niterói (e aqui pensamos em centro comercial, educacional) passa por essa geografia, mas ela é suficiente e exaustiva para justificar essa escolha?



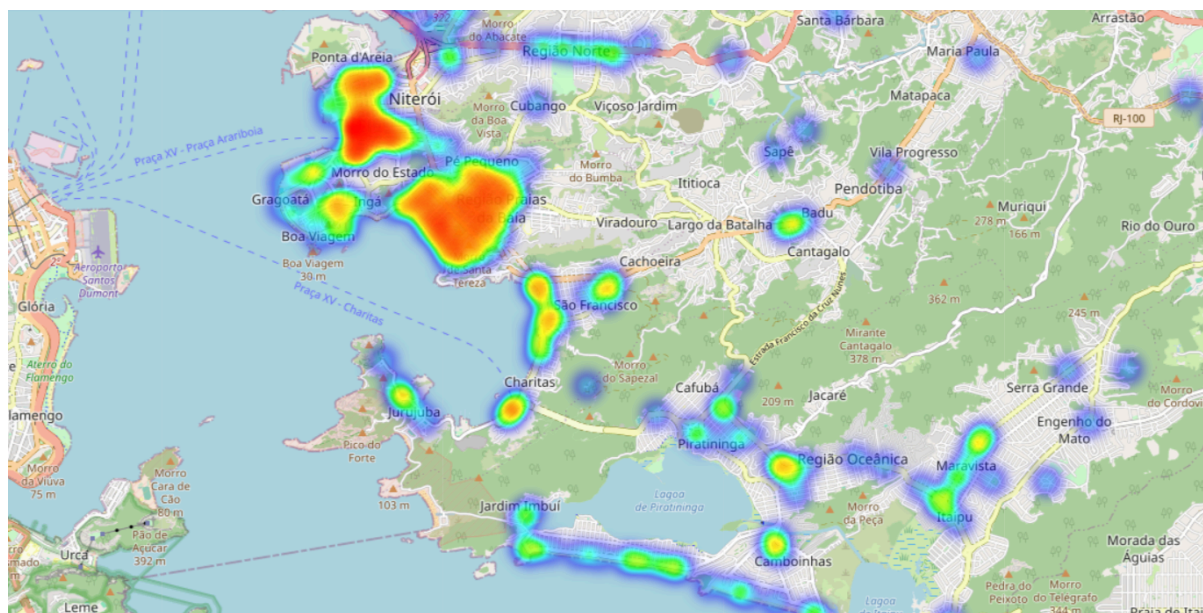
Motivação apenas descritiva.



Motivação apenas descritiva, mas também pode ser interessante a nível investigativo da forma que esses órgãos se relacionam e se engajam.



Motivação apenas descritiva.



Por fim, o mapa de calor que reforça o gráfico de número de vagas por bairro.