

ANÁLISE DE CIDADES PARA ABERTURA DE CLÍNICAS DE FISIOTERAPIA NO BRASIL

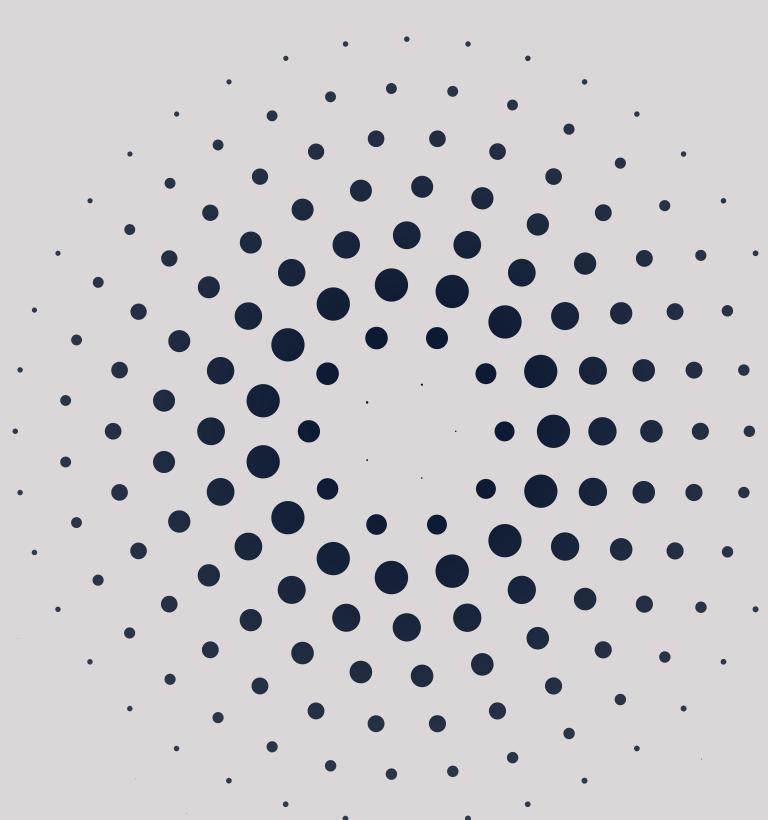
PROJETO EM GRUPO - PARTE 1

25/06/23

**Brenda Farias
Fabricio Leal**

**ADM01007 - Introdução à Data
Science**

**UNIVERSIDADE FEDERAL DO RIO
GRANDE DO SUL**



Etapas

- | | |
|----|--------------------------------------|
| 01 | Seleção da área e objetivo |
| 02 | Coleta de dados |
| 03 | Pré-processamento e limpeza de dados |
| 04 | Análise exploratória |

Objetivos

- Analisar cidades no Brasil e identificar aquelas que indicam melhores oportunidades para a abertura de clínicas de fisioterapia.
- Por meio dessa análise, busca-se fornecer informações para orientar a abertura ou expansão de clínicas de fisioterapia em regiões que apresentem maior necessidade ou oportunidade desses serviços.

Contexto

- O contexto da análise engloba todos os municípios do Brasil, embora dados sobre todos os municípios não foram encontrados para todas as variáveis.
- A área escolha é saúde.

Relevância

1

Impacto na saúde pública: A fisioterapia desempenha um papel crucial na reabilitação e tratamento de lesões e condições físicas.



2

Oportunidades de negócios: Ao determinar as cidades com demanda insuficiente ou subatendida de clínicas de fisioterapia, empresários e investidores podem identificar oportunidades de negócios.



Relevância

3

Planejamento de recursos de saúde: Ao analisar os dados e identificar as cidades que precisam de mais clínicas de fisioterapia, espera-se obter informações valiosas para o planejamento estratégico de recursos de saúde.



4

Melhoria na qualidade de vida: A disponibilidade de serviços de fisioterapia adequados em uma comunidade pode contribuir para melhorar a qualidade de vida dos seus moradores.



Dados

COLETA

Os seguintes locais foram utilizadas para a coleta de dados:

Ministério da Saúde (TABNET):

Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>

Ministério da Fazenda:

Disponível em: <https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>

Agência Nacional de Saúde Suplementar:

Disponível em: <https://www.ans.gov.br/perfil-do-setor/dados-gerais>

Dados

COLETA

Os dados foram importados para o Github e carregados no Google Colab.

Os seguintes dataframes foram carregados:

- Acidentes_de_transito.csv
- Estabelecimentos.csv
- Municipios.csv
- Populacao_residente.csv
- Procedimentos_hospitalares_ARTICULACOES.csv
- Procedimentos_hospitalares_AVC.csv
- Procedimentos_hospitalares_CONDICOESNAUROLOGICAS.csv
- Procedimentos_hospitalares_DOENCASCRONICAS.csv
- Procedimentos_hospitalares_ORTOPEDICAS.csv
- QtdBeneficiarios_planosdesaude.csv

Dados

INFORMAÇÕES

| Dataframe | Tamanho | Linhas |
|--|---------|--------|
| • Acidentes_de_transito.csv | 32,9KB | 1.478 |
| • Estabelecimentos.csv | 11,1MB | 58.840 |
| • Municipios.csv | 100KB | 5.572 |
| • Populacao_residente.csv | 430KB | 5.567 |
| • Procedimentos_hospitalares_ARTICULACOES.csv | 5KB | 200 |
| • Procedimentos_hospitalares_AVC.csv | 64KB | 2.888 |
| • Procedimentos_hospitalares_CONDICOESNAUROLOGICAS.csv | 7KB | 334 |
| • Procedimentos_hospitalares_DOENCASCRONICAS.csv | 11KB | 510 |
| • Procedimentos_hospitalares_ORTOPEDICAS.csv | 4KB | 177 |
| • QtdBeneficiarios_planosdesaude.csv | 429KB | 5.582 |

Dados

INFORMAÇÕES

Os dados coletados refere-se ao ano de 2011, esse filtro foi realizado devido aos dados de acidentes de trânsito mais atuais serem desse ano especificamente.

Todos os dados estão carregados no Github com seus nomes originais, as informações são buscadas para o Google Colab por meio de integração.

Pré-processamento e limpeza de dados

Foram realizados o pré=processamento e limpeza em blocos, conforme cada um dos dataframes armazenados.

Após cada processamento um df novo foi criado para comportar os ajustes realizados, bem como salvo na pasta criada para tal.

Devido a baixa capacidade do Github não foi possivel incluir a base de CNPJs completa, tendo sido incluido somente a base com um pré-ajuste apresentando somente os dados com o CNAE 8650003.

Pré-processamento e limpeza de dados

União dos Dataframes de Procedimentos Hospitalares

As seguintes etapas foram realizadas

```
# Lista dos dataframes a serem unidos  
# Inicializa o dataframe final com o primeiro dataframe da lista  
# Renomeia a coluna 'AIH_aprovadas' para o nome do dataframe e realiza a união  
# Remove a string 'Procedimentos_hospitalares_' dos nomes das outras colunas  
# Split the 'Municipio' column into 'Codigo' and 'Nome'  
# Remove the original 'Municipio' column  
# Reorder the columns to put 'Codigo' and 'Nome' first  
# Salva o dataframe ajustado em um arquivo CSV no diretório especificado
```

Pré-processamento e limpeza de dados

Ajuste no df de população residente

As seguintes etapas foram realizadas

```
# Acessar o dataframe 'Populacao_residente'  
# Printar as primeiras dez linhas para entender a estrutura  
# Retirar a coluna 'total'  
# Separar a coluna 'Municipio' em 'Codigo' e 'Nome'  
# Excluir a coluna 'Municipio'  
# Reordenar as colunas 'Codigo' e 'Nome' para o inicio  
# Cria as novas colunas somando as colunas correspondente  
# Salva o dataframe ajustado em um arquivo CSV no diretório especificado
```

Pré-processamento e limpeza de dados

Ajuste do df de acidentes

As seguintes etapas foram realizadas

```
# Acessar o dataframe 'Populacao_residente'  
# Separar a coluna 'Municipio' em 'Codigo' e 'Nome'  
# Excluir a coluna 'Municipio'  
# Renomear a coluna 'Total' para 'Acidentes de transito'  
# Reordenar as colunas 'Codigo' e 'Nome' para o inicio  
# Cria as novas colunas somando as colunas correspondente  
# Salva o dataframe ajustado em um arquivo CSV no diretório especificado
```

Pré-processamento e limpeza de dados

Ajuste do df de quantidade de beneficiários

As seguintes etapas foram realizadas

```
# Acessar o dataframe 'Qtd_beneficiarios_planosdesaude'  
# Separar a coluna 'Municipio' em 'Codigo' e 'Nome'  
# Excluir a coluna 'Municipio'  
# Criar novas colunas somando as colunas designadas  
# Renomear a coluna '80 anos ou mais'  
# Excluir as colunas original  
# Salva o dataframe ajustado em um arquivo CSV no diretório especificado
```

Pré-processamento e limpeza de dados

Ajuste do df de estabelecimentos "Parte 1"

As seguintes etapas foram realizadas

```
# Acessar o dataframe 'Qtd_beneficiarios_planosdesaude'  
# Exluir as colunas designadas  
# Criar novas colunas de 'MUNICIPIO' com as informações de nomes de municípios  
'NOME MUNICIPIO' em 'Municipios.csv'  
# Fazer o replace dos valores na coluna 'MUNICIPIO' utilizando o dataframe  
designado  
# Fazer o replace da coluna 'SITUACAO CADASTRAL'
```

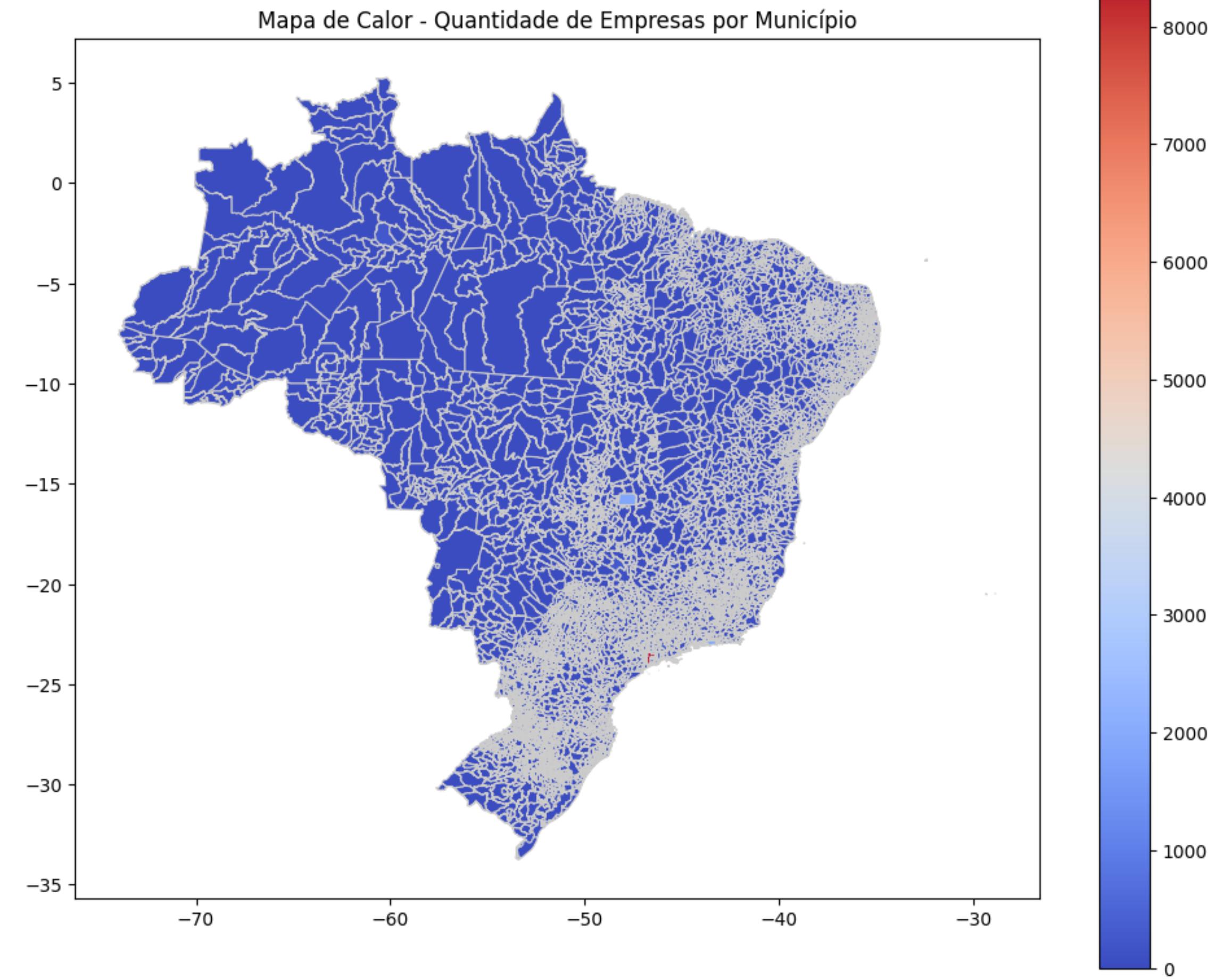
Pré-processamento e limpeza de dados

Ajuste do df de estabelecimentos "Parte 2"

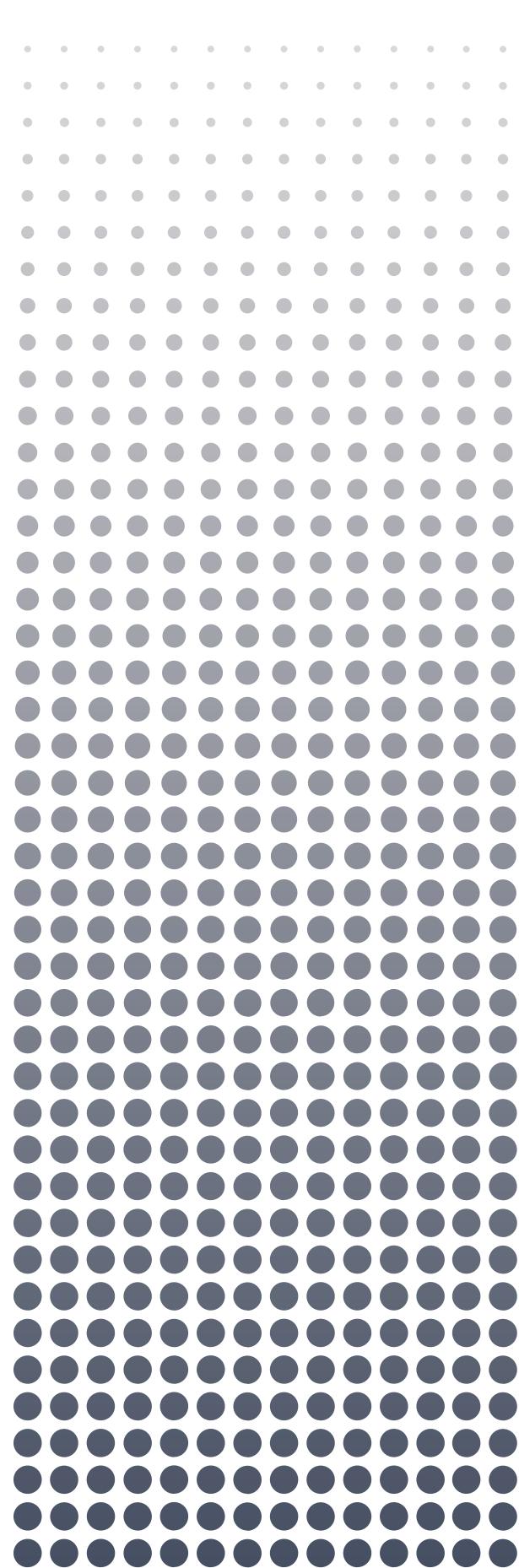
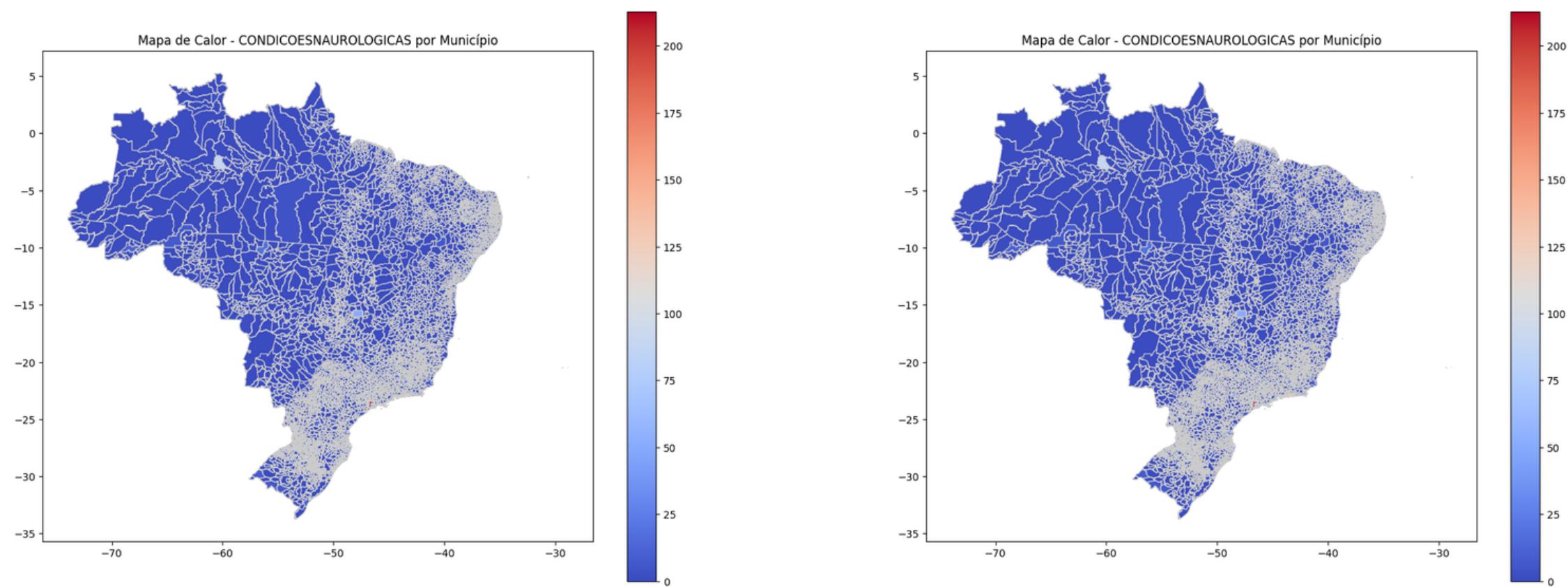
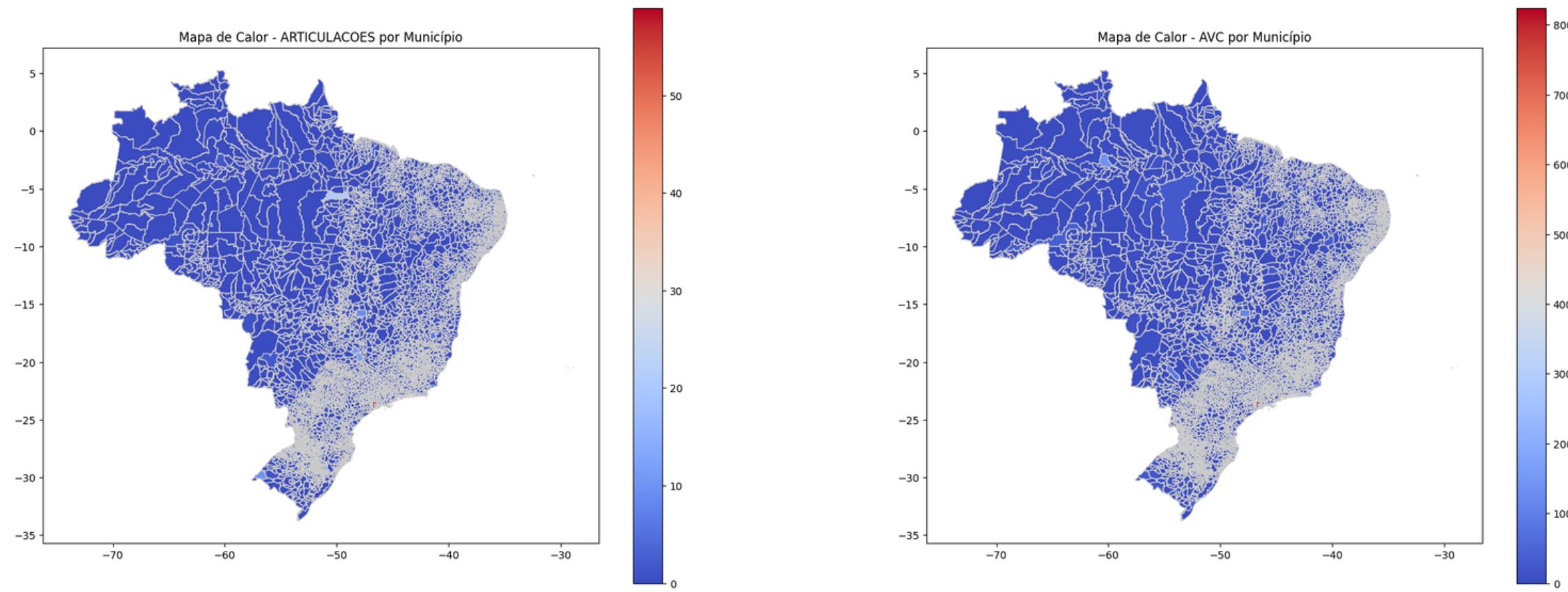
As seguintes etapas foram realizadas

- # Excluir linhas com valores 'NULA', 'SUSPENSA', 'INAPTA', e 'BAIXADA' na coluna 'SITUACAO CADASTRAL'
- # Excluir a coluna 'DATA SITUACAO'
- # Excluir a coluna 'SITUACAO CADASTRAL'
- # Excluir a coluna 'CNAE PRINCIPAL'
- # Converter a coluna 'DATA INICIO ATIVIDADE'
- # Calcular anos de atividades em anos e sem decimais
- # Printar as primeiras dez linhas
- # Salva o dataframe ajustado em um arquivo CSV no diretório especificado

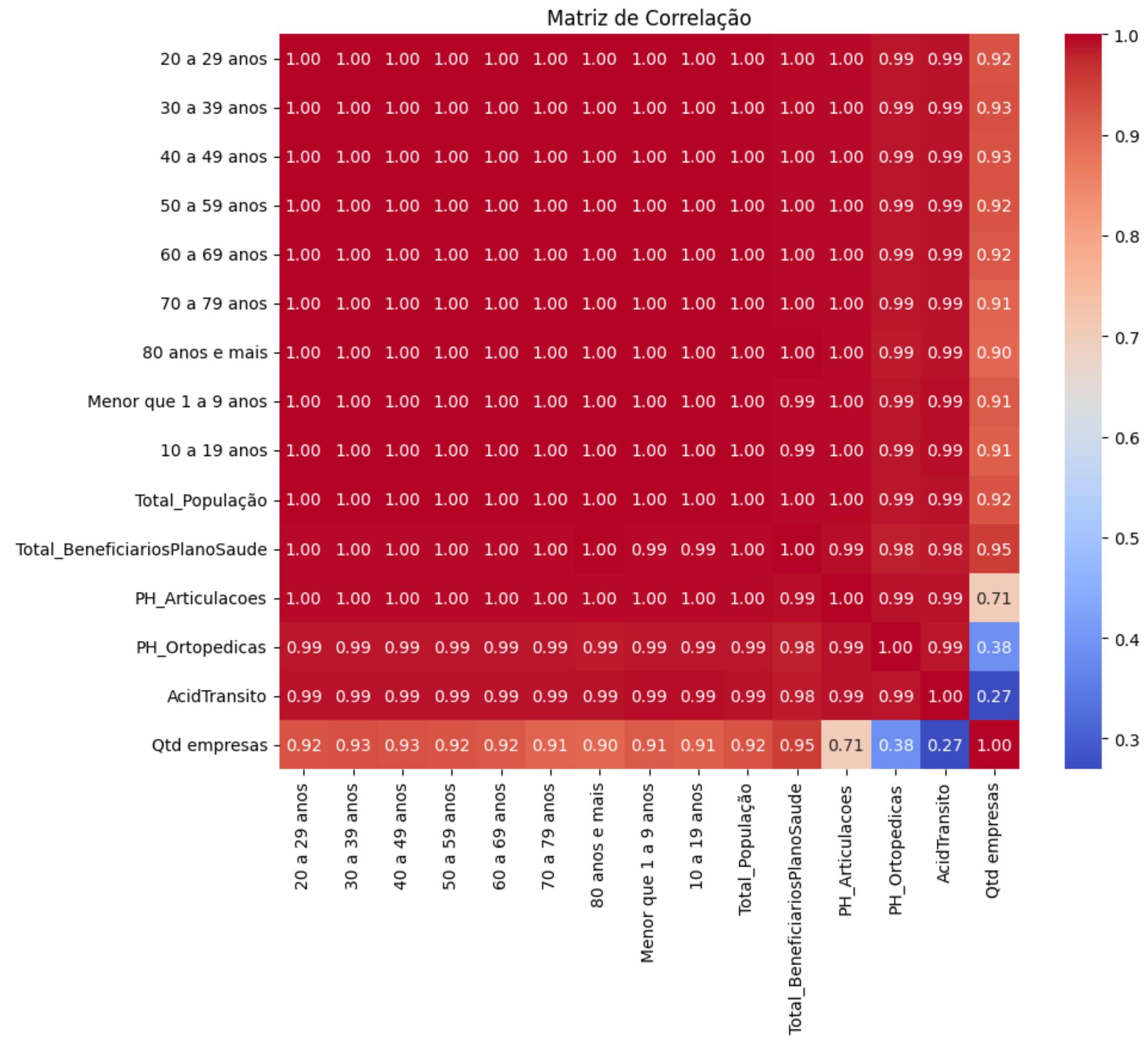
Análise Exploratória



Análise Exploratória

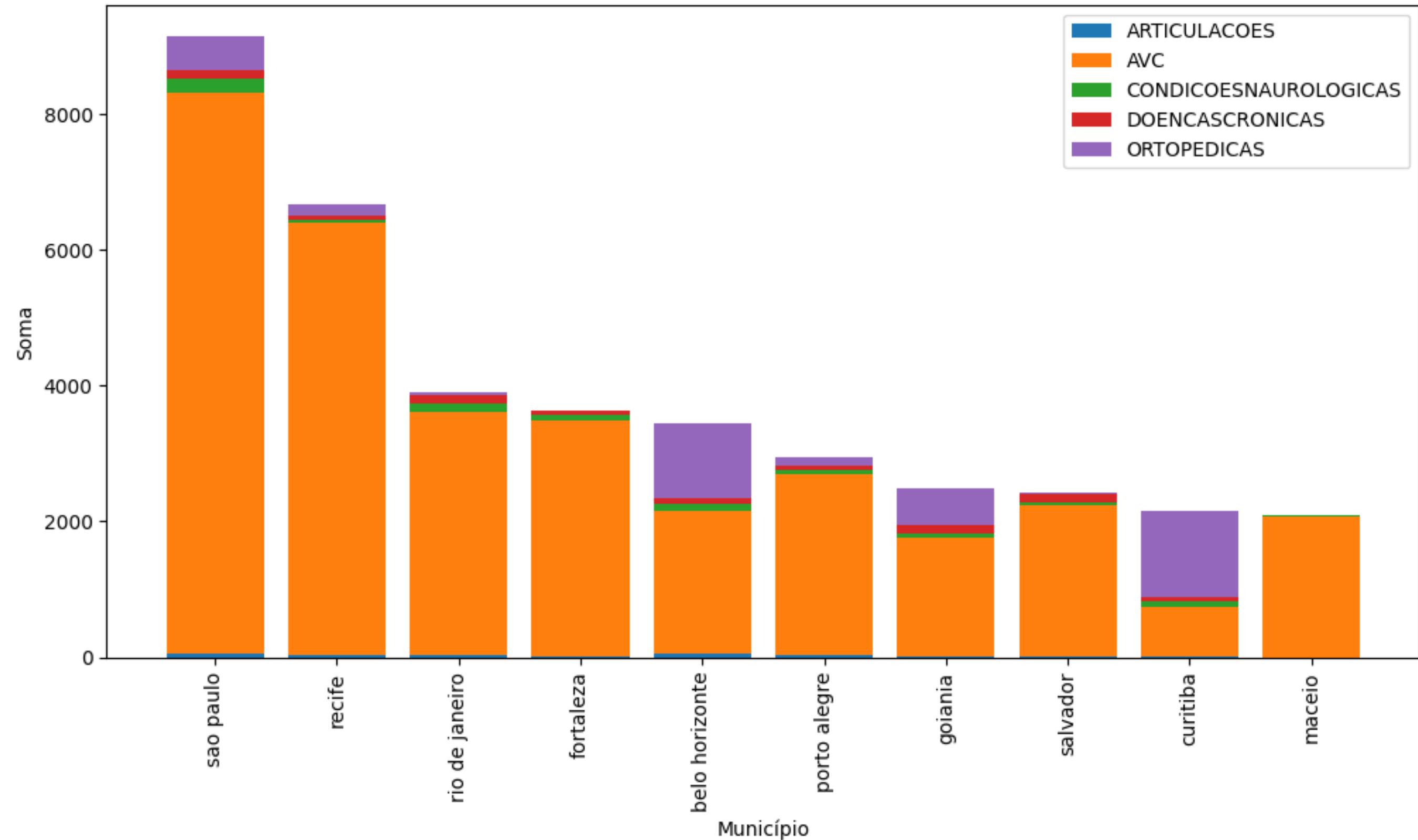


Análise Exploratória

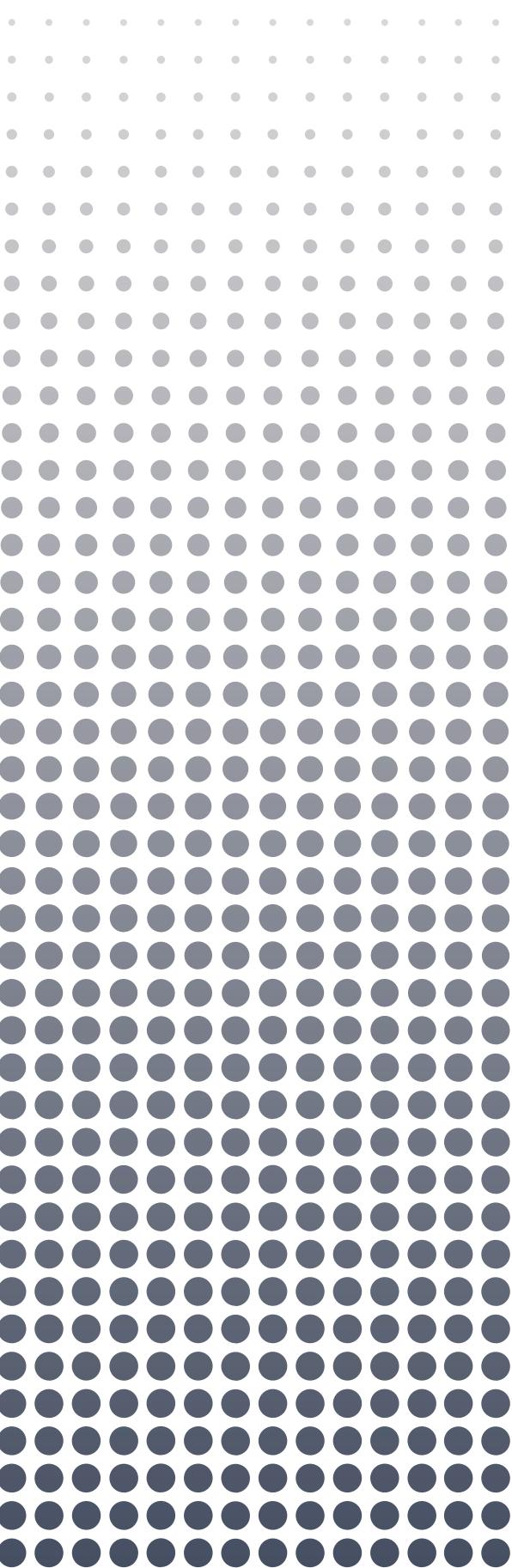
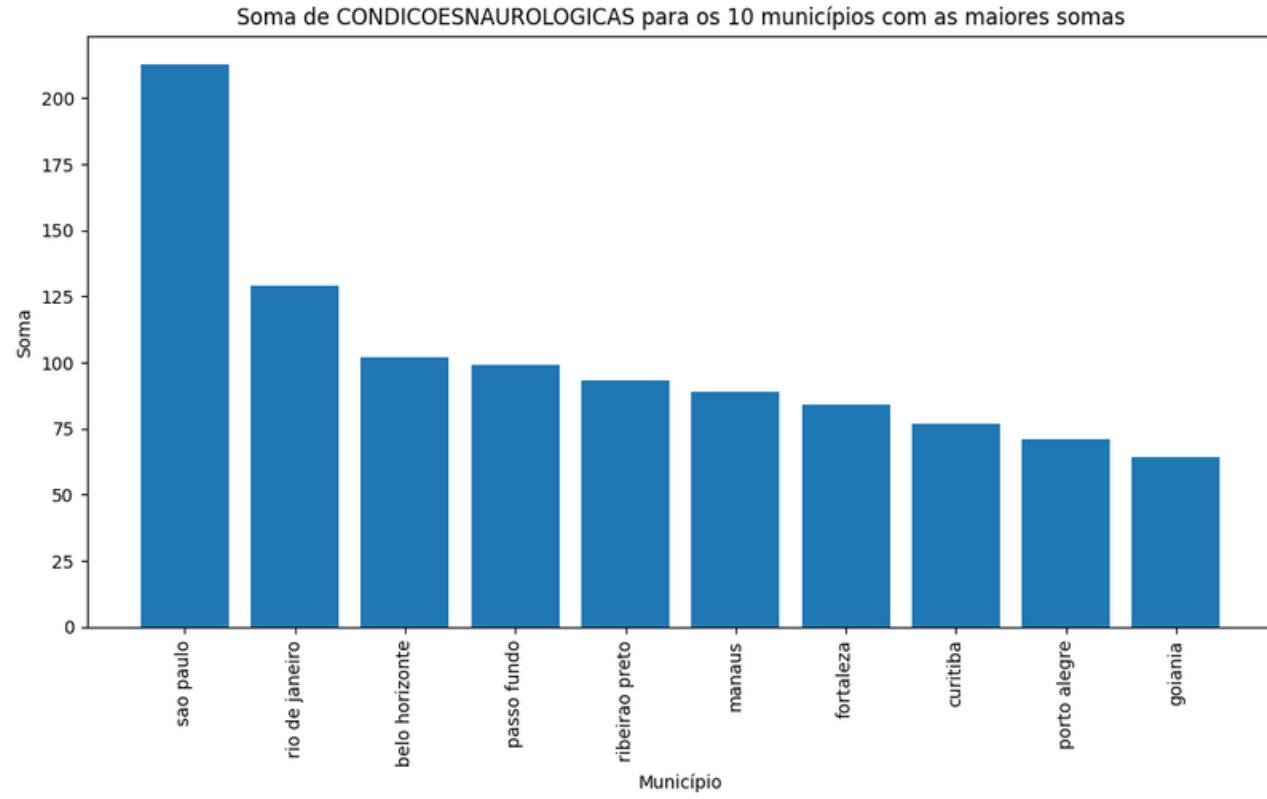
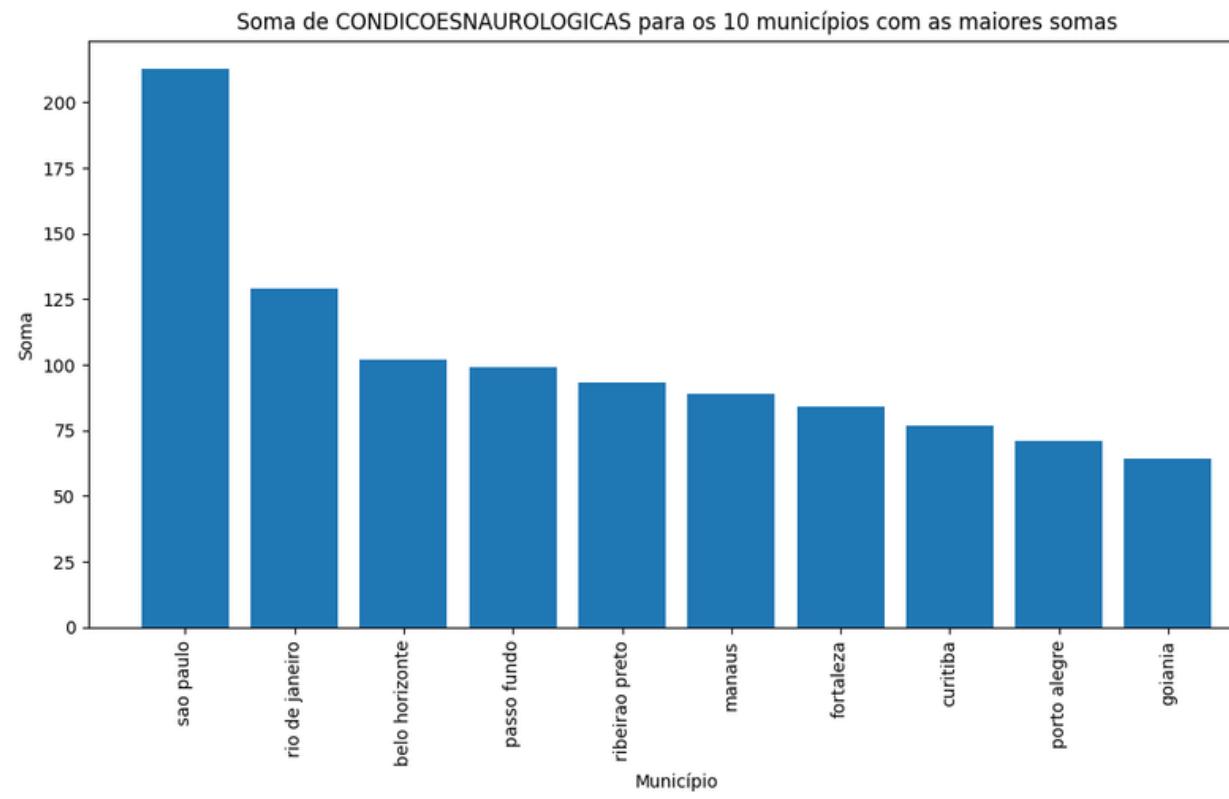
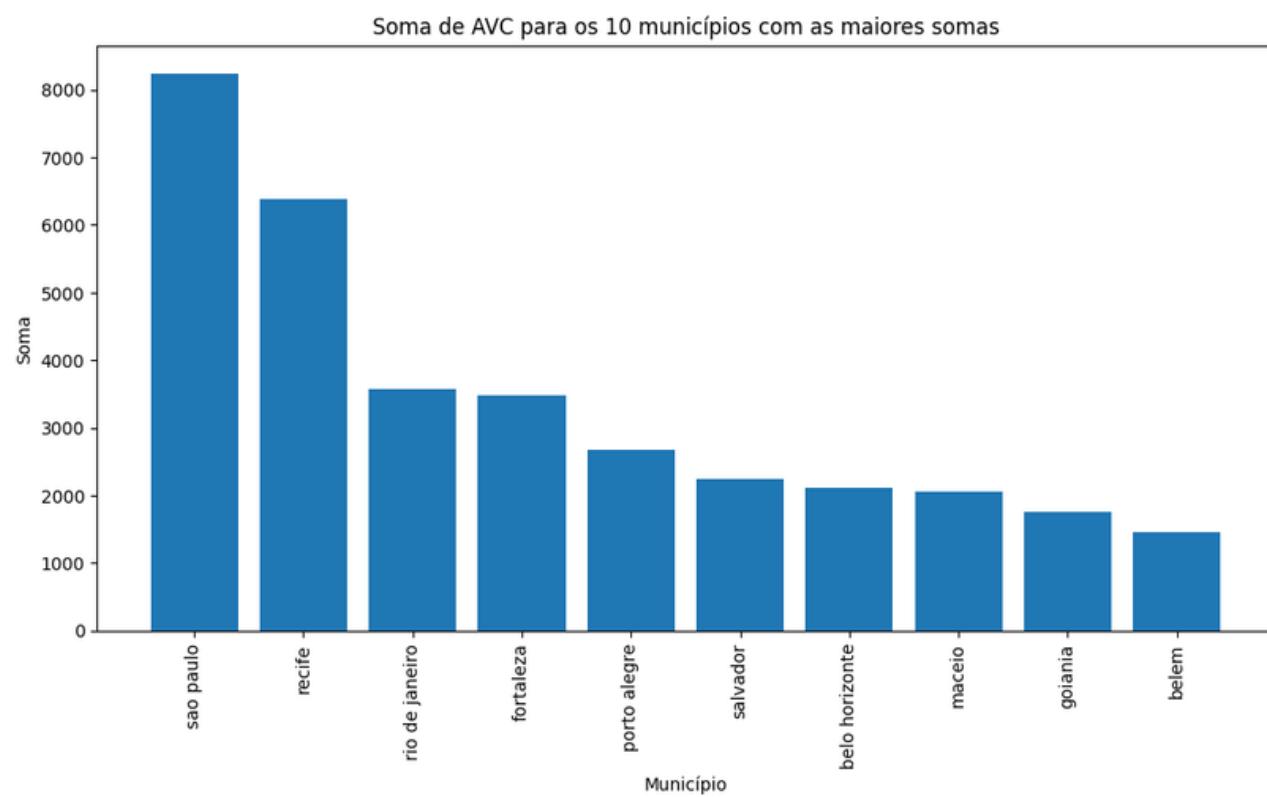
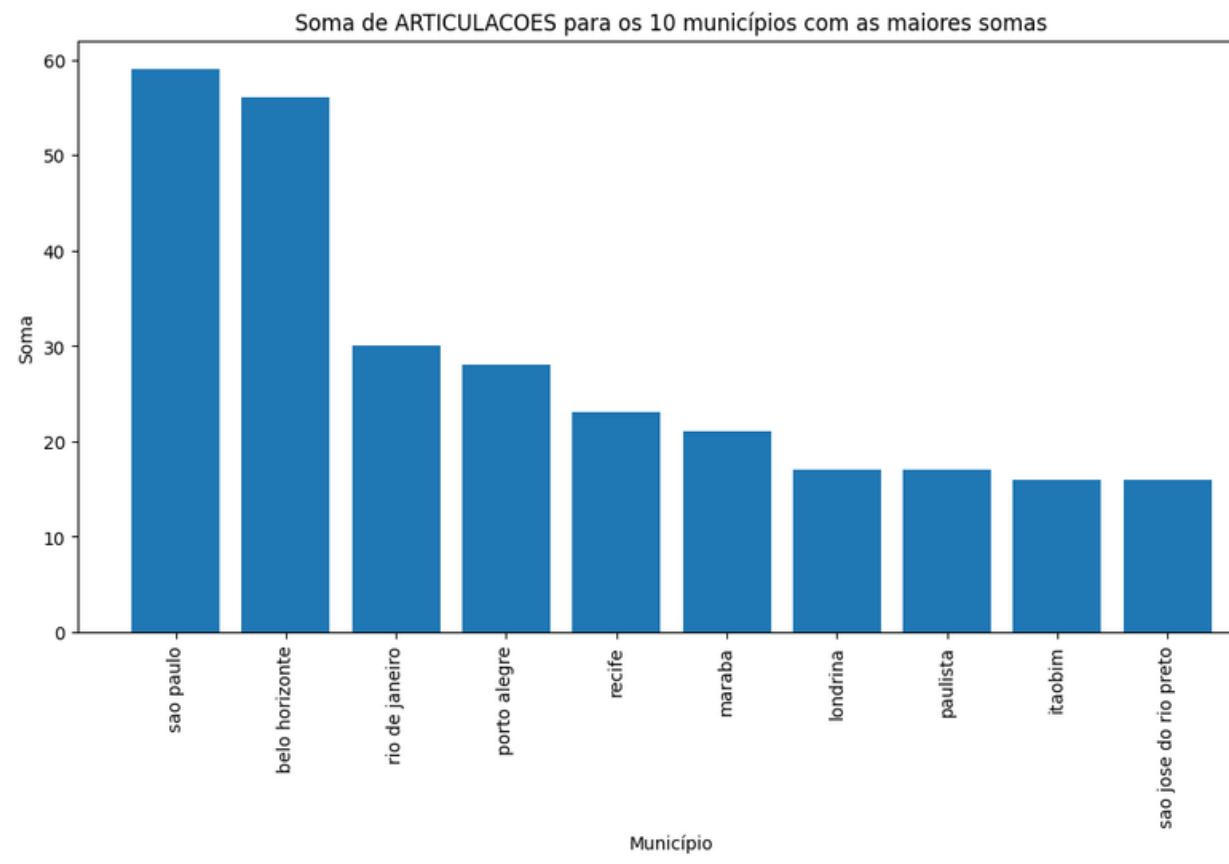


Análise Exploratória

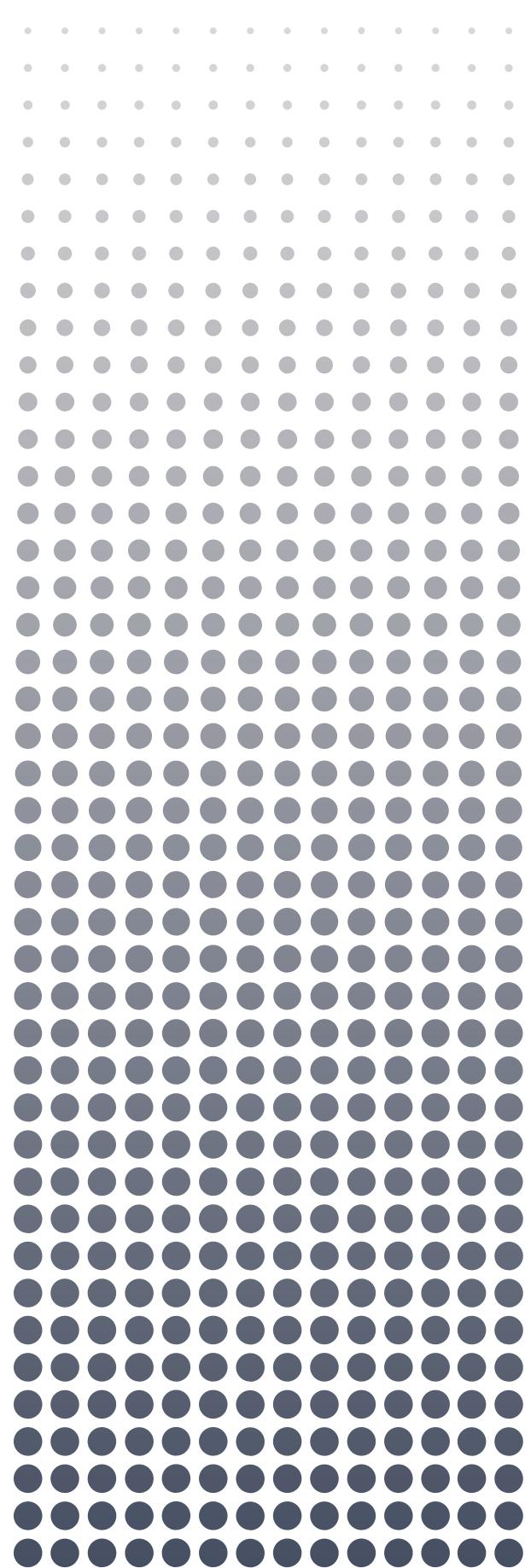
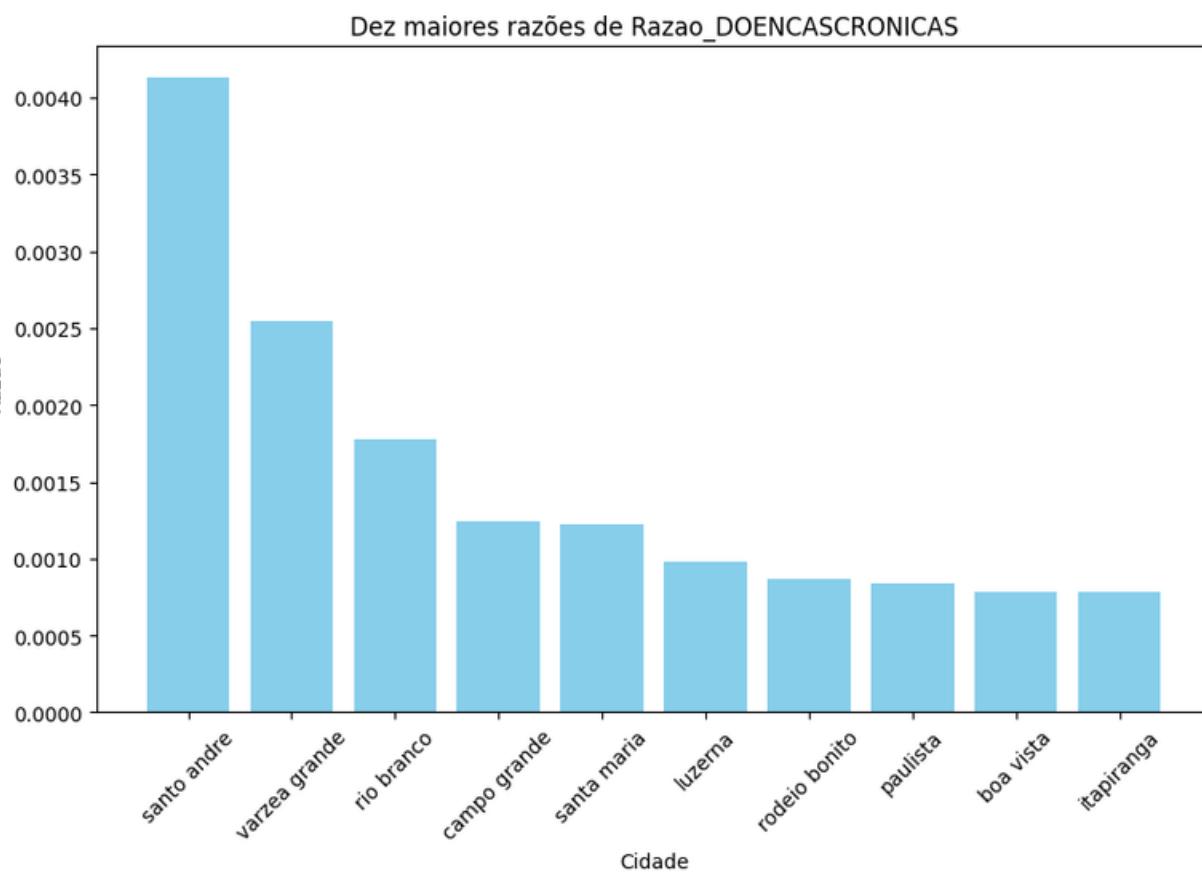
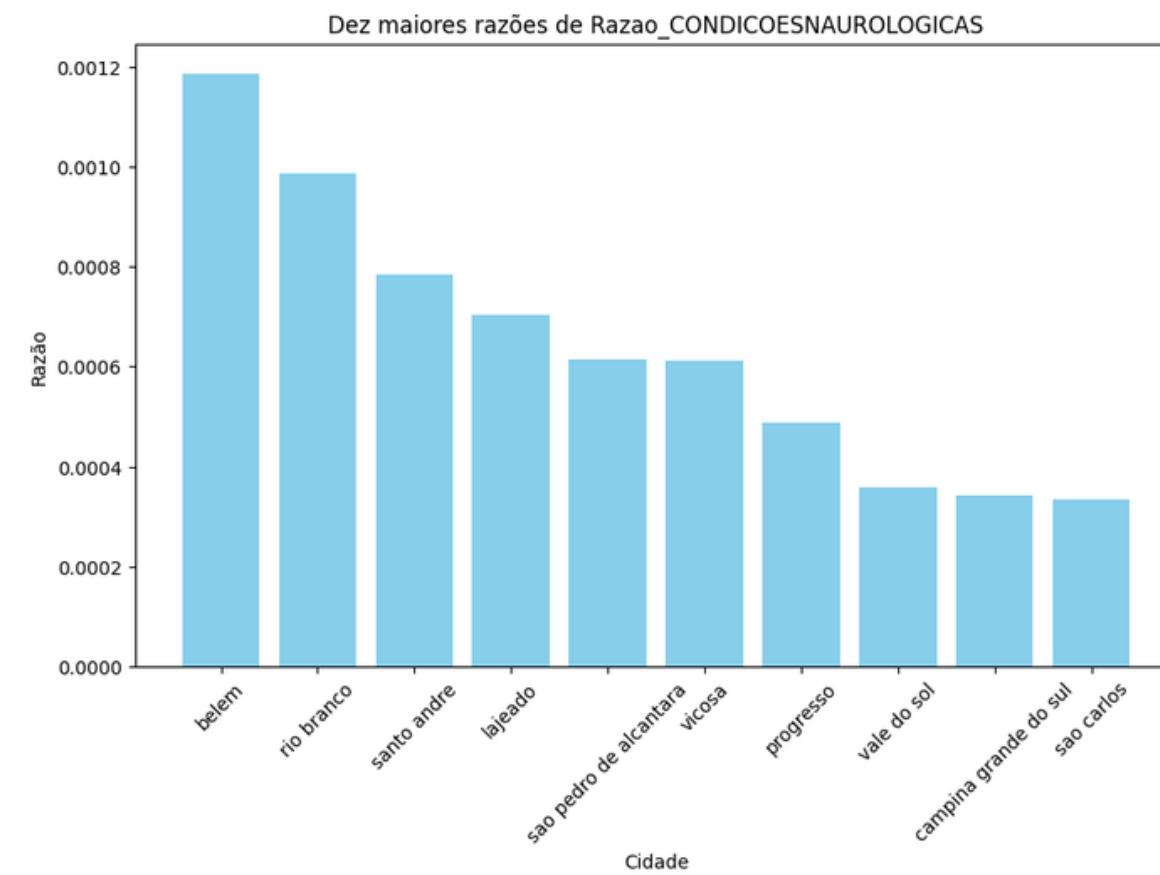
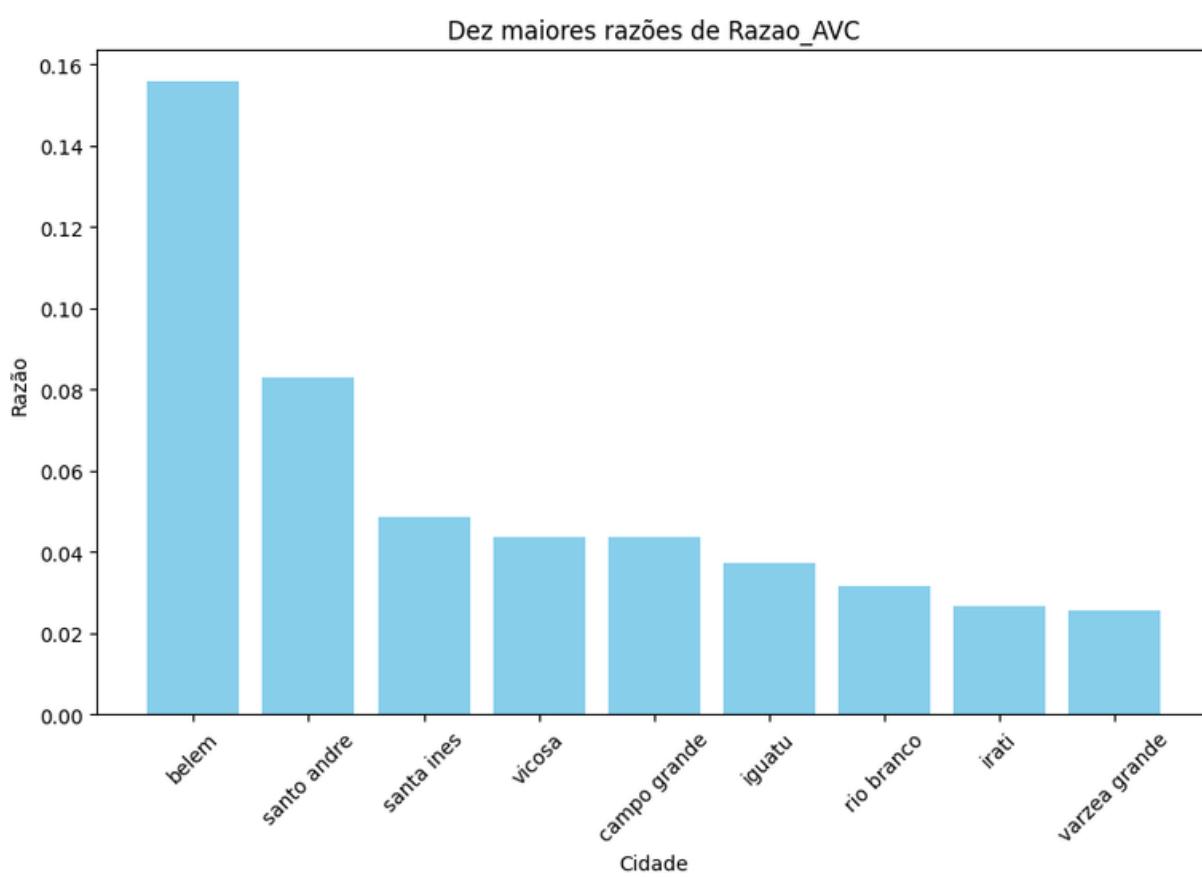
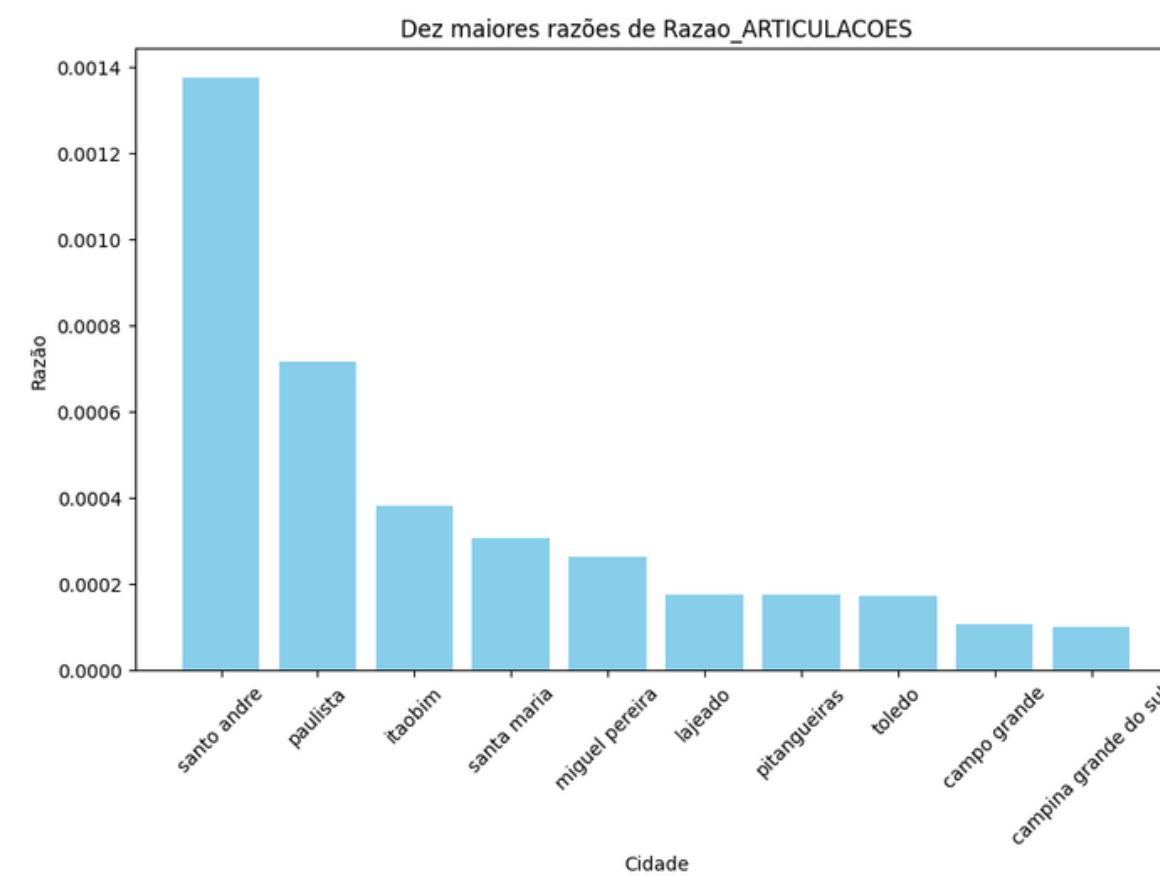
Soma de todos os procedimentos hospitalares para os 10 municípios com as maiores somas



Análise Exploratória



Análise Exploratória



Links úteis



Colab

<https://colab.research.google.com/drive/1Pfkxd eyRwC3OKCHQRGuqK8v DaDTr3zWI?usp=sharing>



Github

[https://githubuserconte nt.com/bsf94/trabgrupo
DS/main](https://githubuserconte nt.com/bsf94/trabgrupo DS/main)