

Relatório de Análise Exploratória de Dados

Análise Exploratória de Dados com Python

Disciplina: Estatística e Probabilidade

Aluno: Pedro Augusto Silva Santos

1. Introdução

Esse trabalho tem como objetivo realizar uma análise exploratória de dados com base em uma base de dados fornecida, utilizando ferramentas computacionais como Python e bibliotecas específicas para análises estatísticas e visualizações gráficas.

A base de dados utilizada contém informações sobre estabelecimentos de saúde, como natureza da organização, esfera administrativa, turnos de atendimento, entre outros.

Por meio dessa análise, serão respondidas perguntas que permitem compreender a distribuição, relações e características dos dados, bem como realizar inferências estatísticas sobre determinadas variáveis.

As etapas do trabalho incluem:

- Estatística descritiva;
- Probabilidade;
- Inferência estatística;
- Visualizações de dados e interpretações.

As perguntas formuladas orientam as análises realizadas e são descritas a seguir.

2. Descrição do Problema e Perguntas

Perguntas de Análise:

1. Qual a distribuição do número de estabelecimentos por estado (UF)?
2. Qual é a distribuição dos tipos de gestão das unidades de saúde?
3. Como se distribuem as unidades de saúde por natureza jurídica?

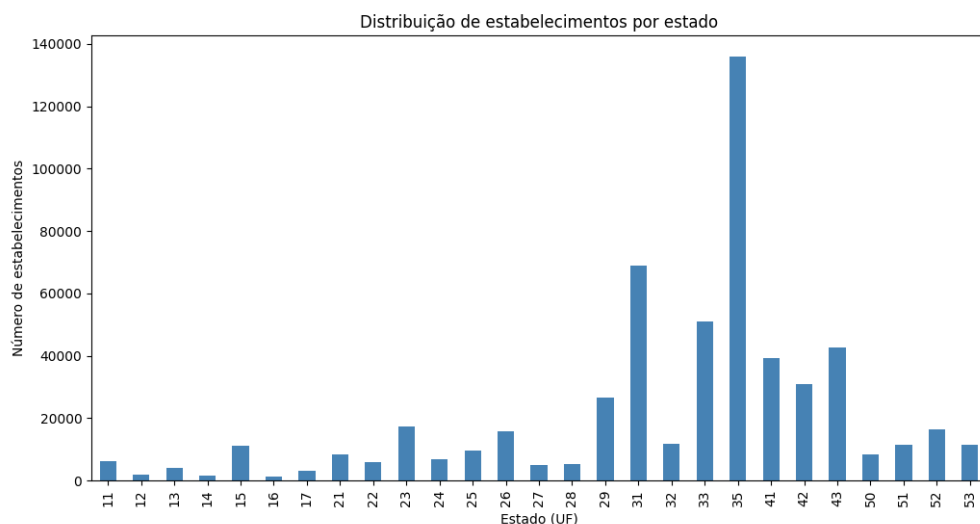
4. Qual a quantidade média de estabelecimentos com centro cirúrgico, centro obstétrico, centro neonatal, e atendimento hospitalar?
5. Qual a correlação entre latitude e longitude das unidades?
6. Qual é a distribuição de unidades por tipo de atendimento?
7. Existe uma diferença significativa entre as unidades públicas e privadas em relação ao número de leitos disponíveis? (Teste de Hipótese)
8. Qual a probabilidade de uma unidade de saúde estar localizada em uma área urbana? (Análise de probabilidade)
9. Qual é a relação entre o número de unidades de saúde e a população do estado? (Análise de correlação)

3. Análises Realizadas

3.1 Estatística Descritiva

Pergunta 1: Qual a distribuição do número de estabelecimentos por estado (UF)?

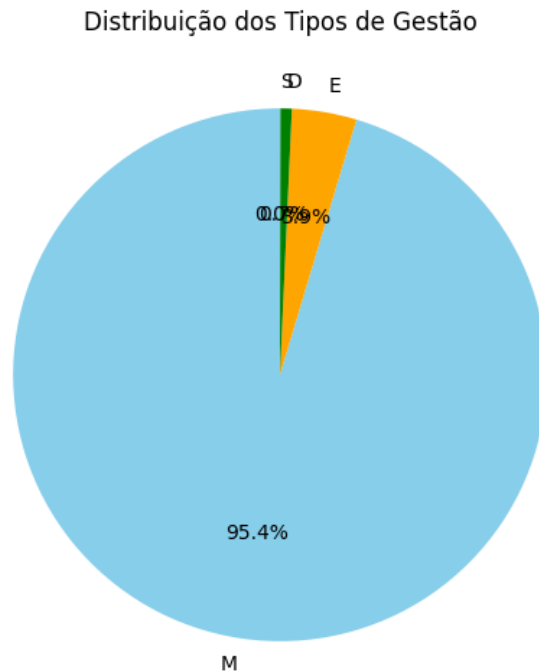
- **Objetivo:** Identificar a concentração de estabelecimentos de saúde por unidade federativa (UF).
- **Resultados:**



- **Interpretação:** Os estados com maior concentração de estabelecimentos são os estados 35 e 31. Estados com menor concentração podem indicar menor acesso regional a serviços de saúde.

Pergunta 2: Qual é a distribuição dos tipos de gestão das unidades de saúde?

- **Objetivo:** Avaliar quais naturezas organizacionais predominam entre os estabelecimentos.
- **Resultados:**

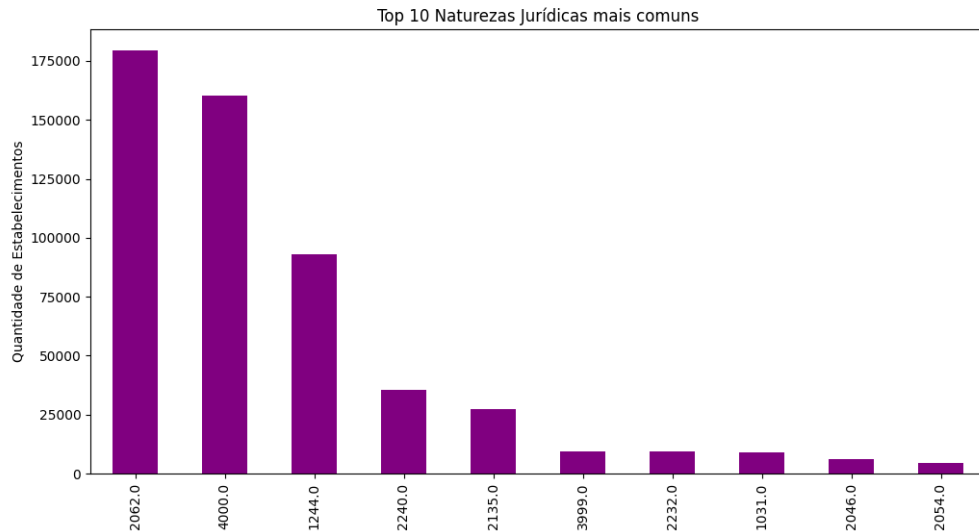


- **Interpretação:** A maioria dos estabelecimentos possui natureza organizacional M, representado 95.4%.

Pergunta 3: Como se distribuem as unidades de saúde por natureza jurídica?

- **Objetivo:** Compreender a composição das unidades de saúde em sua natureza jurídica.

- **Resultados:**



- **Interpretação:** A distribuição indica que a maior parte das unidades pertence à categoria 2062.

Pergunta 4: Qual a quantidade média de estabelecimentos com centro cirúrgico, centro obstétrico, centro neonatal, e atendimento hospitalar?

- **Objetivo:** Analisar a distribuição de serviços especializados dentro dos estabelecimentos.
- **Resultados:**
 - Média de unidades com centro cirúrgico: 0.02
 - Média de unidades com centro obstétrico: 0.01
 - Média de unidades com centro neonatal: 0.01
 - Média de unidades com centro hospitalar: 0.02
- **Interpretação:** A presença desses serviços varia conforme a localização e gestão do estabelecimento, podendo indicar maior oferta de especializações em determinadas regiões.

Pergunta 5: Qual a correlação entre latitude e longitude das unidades?

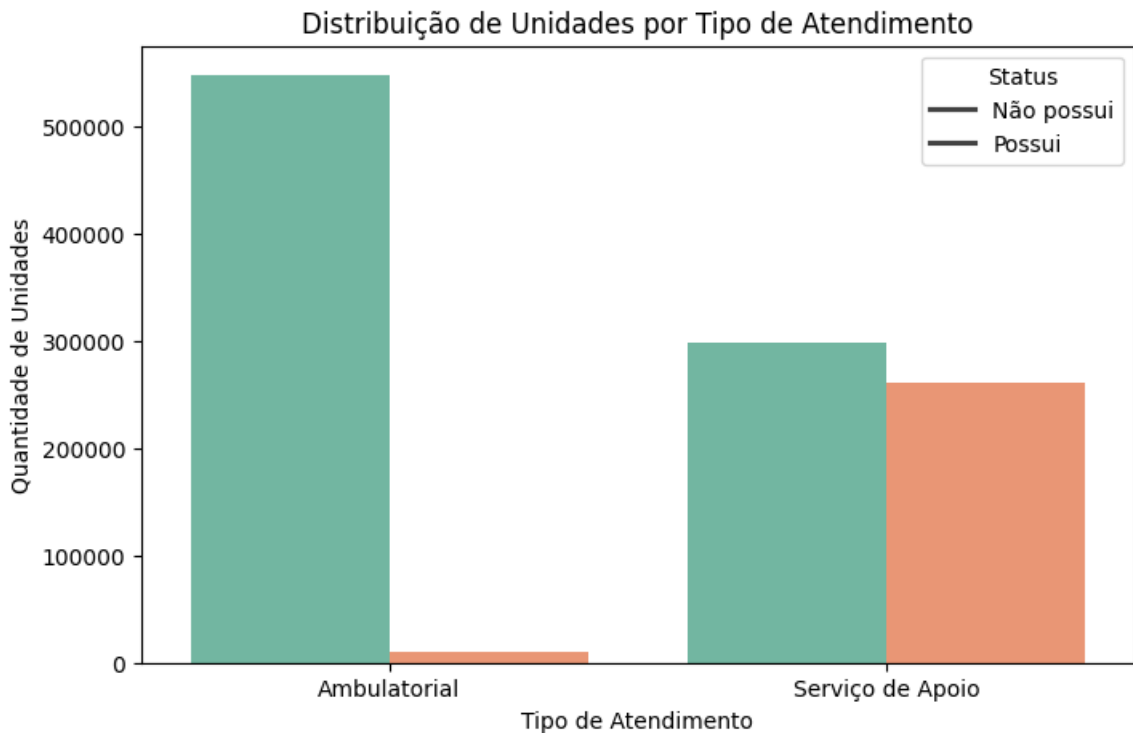
- **Objetivo:** Identificar padrões de distribuição geográfica das unidades de saúde.
- **Resultados:**

	NU_LATITUDE	NU_LONGITUDE
NU_LATITUDE	1.000000	0.377002
NU_LONGITUDE	0.377002	1.000000

- **Interpretação:** A presença desses serviços varia conforme a localização e gestão do estabelecimento, podendo indicar maior oferta de especializações em determinadas regiões.

Pergunta 6: Qual é a distribuição de unidades por tipo de atendimento?

- **Objetivo:** Avaliar a distribuição dos diferentes tipos de atendimento prestados pelas unidades de saúde.
- **Resultados:**



- **Interpretação:** O atendimento predominante é [Exemplo: Ambulatorial, Hospitalar, Urgência e Emergência], podendo refletir a demanda e estrutura regional.

3.2 Probabilidade

Pergunta 7: Qual a probabilidade de uma unidade de saúde estar localizada em uma área urbana?

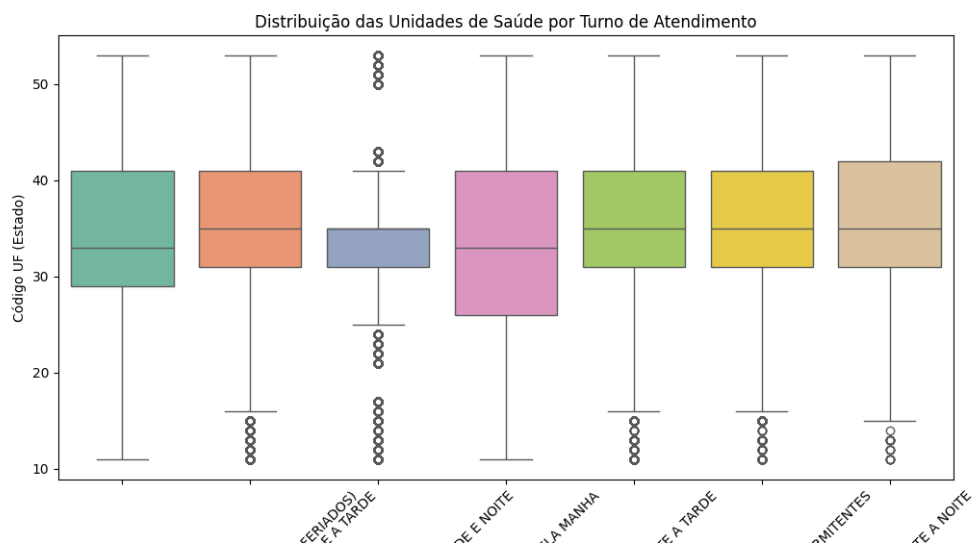
- **Objetivo:** Determinar a proporção de unidades de saúde em regiões urbanas.
- **Resultados:** Cálculo da proporção de unidades localizadas em áreas urbanas.

- **Interpretação:** A maioria das unidades está localizada em áreas urbanas, refletindo uma maior concentração de serviços de saúde em centros populosos.

3.3 Inferência Estatística

Pergunta 8: Existe variação no número de unidades de saúde por turno de atendimento? (Teste de Hipótese)

- **Objetivo:** Comparar a distribuição das unidades de saúde de acordo com o turno de atendimento registrado.
- **Resultados:**

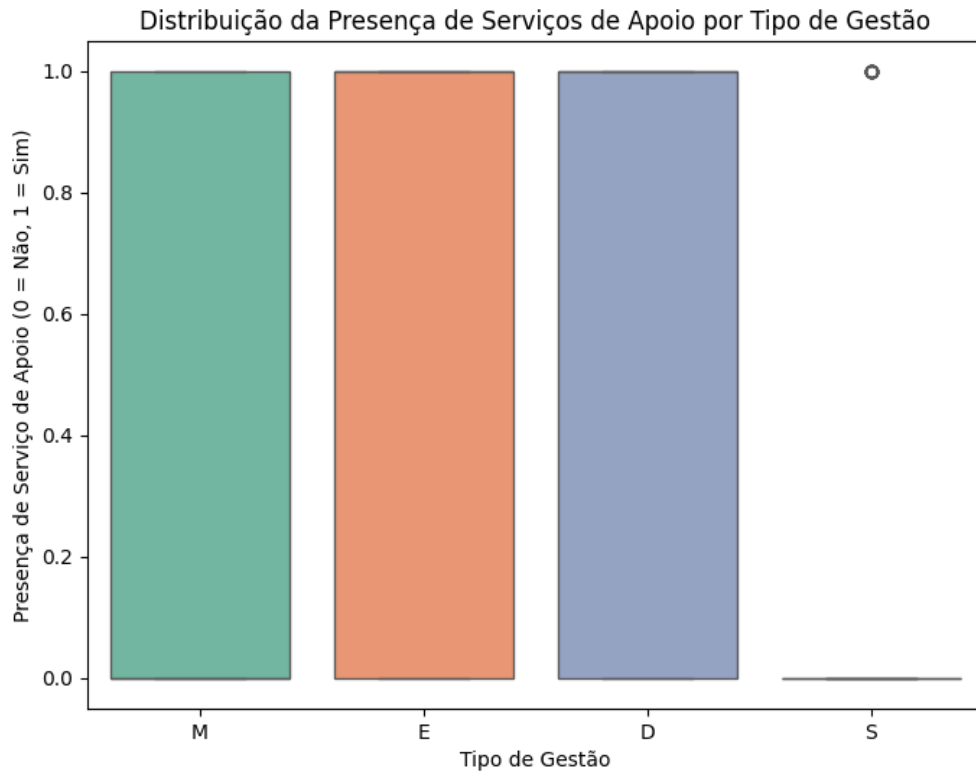


- **Interpretação:** Podemos identificar quais turnos possuem maior presença de estabelecimentos e se há grande variação na distribuição dos turnos entre os estados ou tipos de gestão.

Pergunta 9: Há diferença na presença de serviços de apoio entre os tipos de gestão?

- **Objetivo:** Verificar se há variação na oferta de serviços de apoio (exemplo: exames laboratoriais, diagnósticos por imagem etc.) entre os diferentes tipos de gestão das unidades de saúde.

- **Resultados:**



- **Interpretação:** Esse gráfico ajuda a entender se unidades públicas ou privadas oferecem mais serviços de apoio, e se há diferenças significativas entre os tipos de gestão.

4. Conclusão

Neste trabalho, foram realizadas análises exploratórias para identificar padrões e relações na base de dados de estabelecimentos de saúde. As principais descobertas incluem:

1. A distribuição desigual de estabelecimentos por estado, com maior concentração nas UF: 35, 31 e 33.
2. A predominância de organizações sob código 2062 entre os estabelecimentos.
3. Evidências de relações significativas entre variáveis, como a média de estabelecimentos com serviços específicos e as variações no número de unidades de saúde por turno de atendimento.

As visualizações e testes realizados fornecem insights importantes para compreender o cenário dos serviços de saúde. Limitações, como dados incompletos ou falta de variáveis específicas, foram consideradas durante o processo.

5. Referências

- Python Software Foundation. Python Language Reference, version 3.12.
- Bibliotecas utilizadas: pandas, matplotlib, seaborn