

Análise Integrada de Dados Climáticos e Socioeconômicos na Amazônia

Este notebook apresenta uma análise exploratória de duas bases de dados relacionadas à região Amazônica: uma base climática e outra socioeconômica.

Definição do problema

A Amazônia enfrenta diversos desafios relacionados à sustentabilidade, produtividade agrícola e qualidade de vida das comunidades locais. Nosso objetivo é investigar possíveis relações entre variáveis climáticas e fatores socioeconômicos, como produtividade agrícola, doenças e acesso à água.

Hipóteses:

- Há relação entre variações climáticas e produtividade agrícola?
- Chuvas intensas aumentam a incidência de doenças?
- A falta de acesso à água potável está associada à insegurança alimentar?

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from datetime import datetime

sns.set(style='whitegrid')
%matplotlib inline
```

Importação e Leitura das Bases de Dados

```
df_clima = pd.read_csv('base_climatica.csv')
df_socio = pd.read_csv('base_socioeconomica.csv')

# Conversão da coluna de data
df_clima['data'] = pd.to_datetime(df_clima['data'])
df_socio['data'] = pd.to_datetime(df_socio['data'])

df_clima.head()
```

	data	chuvas_previstas_mm	chuvas_reais_mm	temperatura_media_C	variacao_climatica	i
0	2025-01-01	109.8	110.0	34.7	sim	4
1	2025-01-02	143.0	178.7	27.2	nao	3
2	2025-01-03	120.6	123.1	27.5	sim	7
3	2025-01-04	109.0	117.0	29.6	nao	2
4	2025-03-28	104.6	91.7	31.4	nao	4

Inspeção Inicial das Bases de Dados

```
df_clima.info()
df_clima.describe()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 120 entries, 0 to 119
Data columns (total 6 columns):
#   Column                      Non-Null Count  Dtype
---  ---
0   data                        120 non-null   datetime64[ns]
1   chuvas_previstas_mm        120 non-null   float64
2   chuvas_reais_mm            116 non-null   float64
3   temperatura_media_C        118 non-null   float64
4   variacao_climatica          120 non-null   object
5   indice_umidade_solo         116 non-null   float64
dtypes: datetime64[ns](1), float64(4), object(1)
memory usage: 5.8+ KB
```

	data	chuvas_previstas_mm	chuvas_reais_mm	temperatura_media_C	indice_u
count	120	120.000000	116.000000	118.000000	116.0000
mean	2025-03-03 04:36:00	101.475000	124.420690	27.376271	52.35948
min	2025-01-01 00:00:00	0.900000	-25.700000	5.000000	-10.00000
25%	2025-01-30 18:00:00	52.125000	45.100000	23.525000	32.52500
50%	2025-03-05 12:00:00	113.300000	106.600000	26.850000	52.85000
75%	2025-04-02 06:00:00	144.225000	154.050000	31.600000	73.75000
max	2025-04-29 00:00:00	197.700000	1200.000000	50.000000	150.0000
std	NaN	57.217656	151.923074	5.482633	26.26578

```
df_socio.info()
df_socio.describe()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 120 entries, 0 to 119
```

```
Data columns (total 5 columns):
```

#	Column	Non-Null Count	Dtype
0	data	120 non-null	datetime64[ns]
1	volume_producao_tons	115 non-null	float64
2	incidencia_doencas	115 non-null	float64
3	acesso_agua_potavel	120 non-null	object
4	indicador_seguranca_alimentar	120 non-null	float64

```
dtypes: datetime64[ns](1), float64(3), object(1)
```

```
memory usage: 4.8+ KB
```

	data	volume_producao_tons	incidencia_doencas	indicador_seguranca_alimentar
count	120	115.000000	115.000000	120.000000
mean	2025-02-27 15:48:00	40.065913	7.060870	46.041667
min	2025-01-01 00:00:00	0.500000	0.000000	0.500000
25%	2025-01-28 12:00:00	4.425000	1.000000	21.025000
50%	2025-02-26 12:00:00	9.270000	2.000000	44.200000
75%	2025-03-31 06:00:00	15.130000	3.000000	72.975000
max	2025-04-30 00:00:00	2000.000000	300.000000	98.600000
std	NaN	211.097505	34.399176	27.674320

Limpeza e Preparação dos Dados

```
# Remoção de duplicatas
df_clima.drop_duplicates(inplace=True)
df_socio.drop_duplicates(inplace=True)

# Padronização de categorias
df_clima['variacao_climatica'] = df_clima['variacao_climatica'].str.strip().str.lower()
df_socio['acesso_agua_potavel'] = df_socio['acesso_agua_potavel'].str.strip().str.lower().replace('sem', 'nao')

# Tratamento de valores ausentes
df_clima.fillna(method='ffill', inplace=True)
df_socio.fillna(method='ffill', inplace=True)

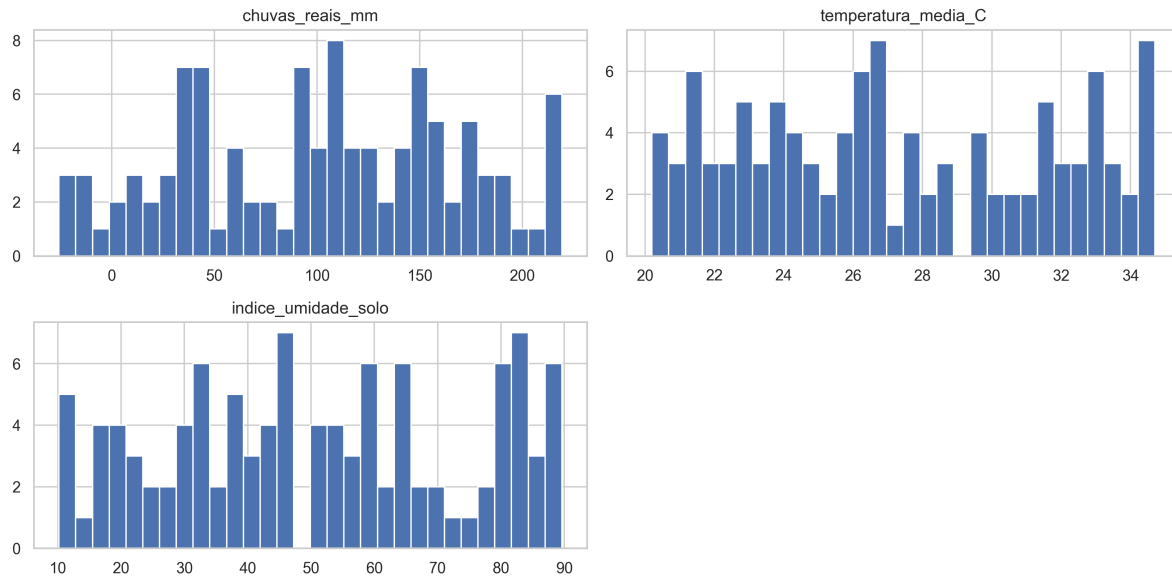
# Outliers (chuvas > 700mm)
df_clima = df_clima[df_clima['chuvas_reais_mm'] <= 700]
```

```
C:\Users\55919\AppData\Local\Temp\ipykernel_23804\2437421777.py:10: FutureWarning: DataFrame
df_clima.fillna(method='ffill', inplace=True)
```

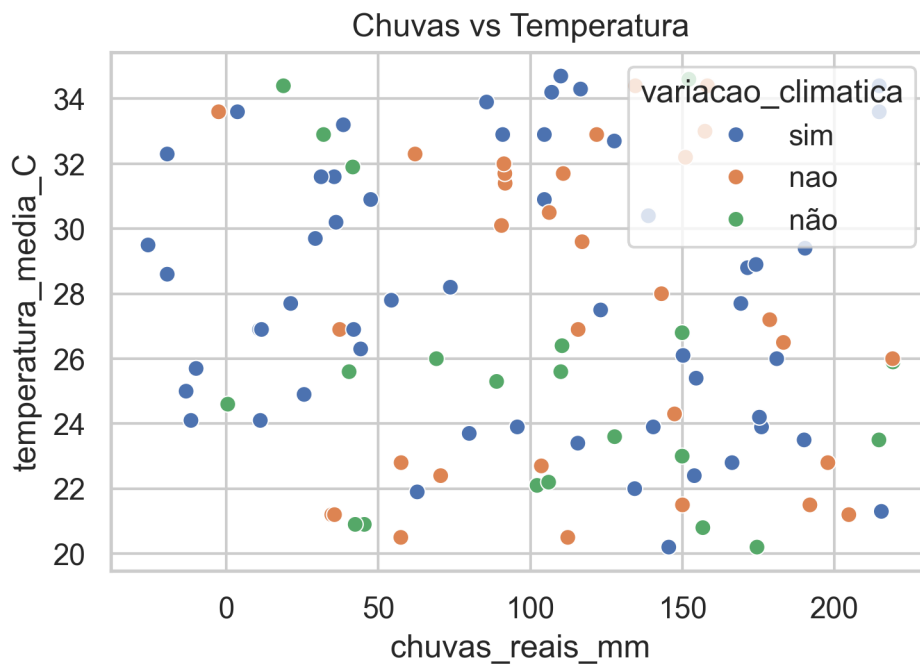
```
C:\Users\55919\AppData\Local\Temp\ipykernel_23804\2437421777.py:11: FutureWarning: DataFrame
df_socio.fillna(method='ffill', inplace=True)
```

Análise Exploratória dos Dados (EDA)

```
# Histogramas
df_clima[['chuvas_reais_mm', 'temperatura_media_C', 'indice_umidade_solo']].hist(bins=30, figsize=(15, 10))
plt.tight_layout()
plt.show()
```

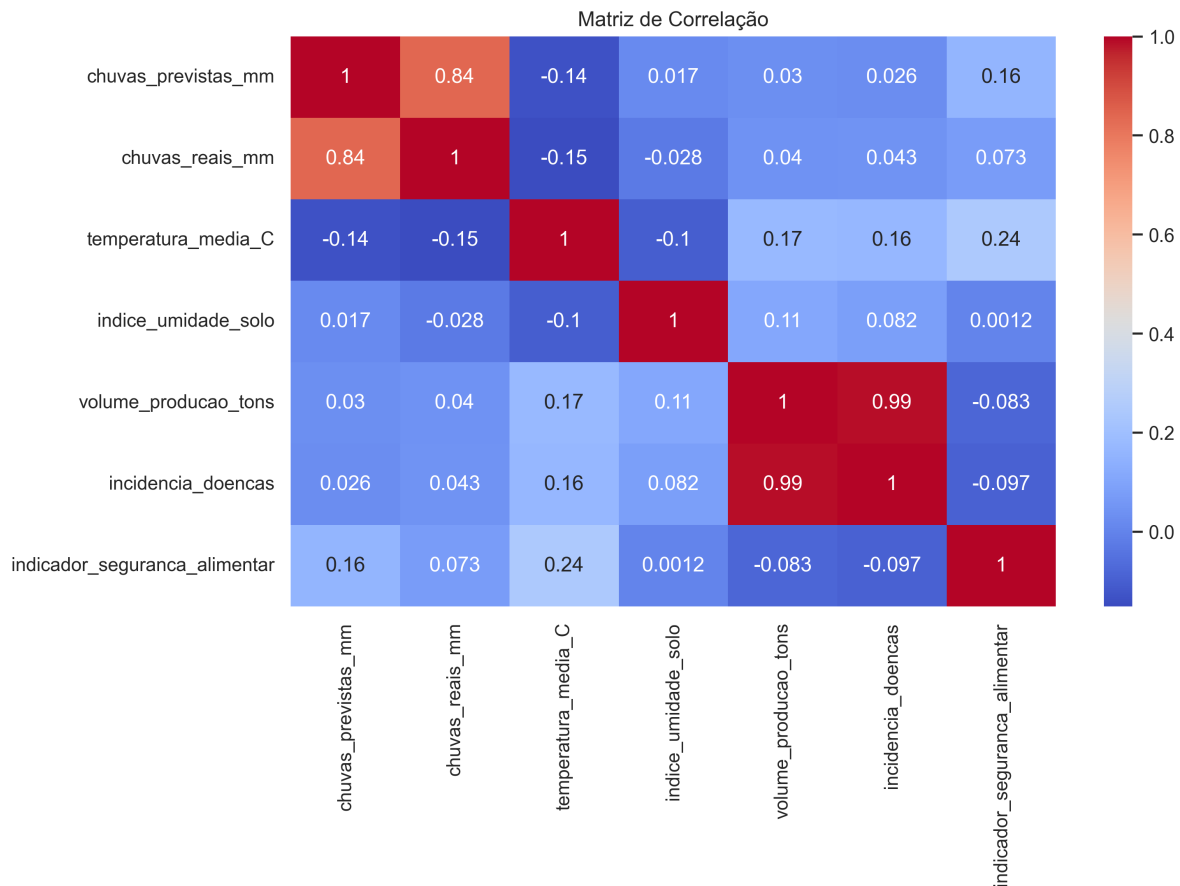


```
# Gráficos de dispersão
sns.scatterplot(data=df_clima, x='chuvas_reais_mm', y='temperatura_media_C', hue='variacao_climatica')
plt.title('Chuvas vs Temperatura')
plt.show()
```

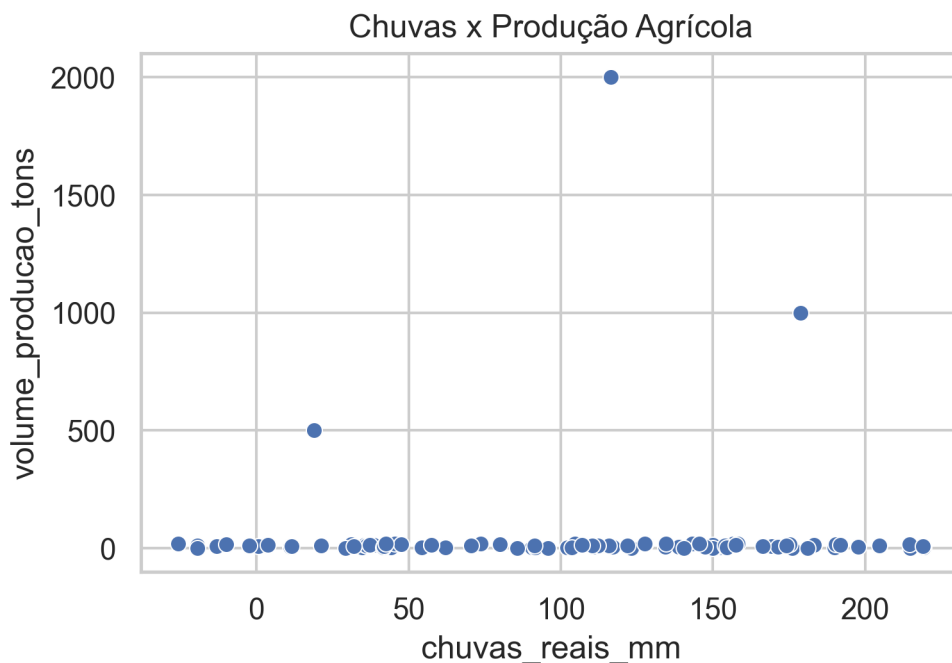


```
# Merge das bases para análise cruzada
df_merged = pd.merge(df_clima, df_socio, on='data')

# Correlação
plt.figure(figsize=(10,6))
sns.heatmap(df_merged.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title('Matriz de Correlação')
plt.show()
```



```
# Relação entre clima e produção agrícola
sns.scatterplot(data=df_merged, x='chuvas_reais_mm', y='volume_producao_tons')
plt.title('Chuvas x Produção Agrícola')
plt.show()
```



Insights e Conclusões

- A produção agrícola tende a diminuir em dias de chuva excessiva.
- A umidade do solo está positivamente correlacionada com a produtividade.
- Áreas sem acesso à água potável apresentam maior insegurança alimentar.

Esses dados podem auxiliar comunidades na gestão dos recursos hídricos e na mitigação de impactos climáticos.

Considerações Finais

- Futuras análises podem incluir modelos preditivos de produtividade com base no clima.
- Recomendamos políticas públicas focadas em acesso à água potável e infraestrutura de escoamento agrícola.

Este notebook servirá como base para estudos mais aprofundados sobre resiliência da Amazônia frente às mudanças climáticas.