

```
import pandas as pd

# Load the dataset
data = pd.read_csv('/content/Unemployment_Rate_upto_11_2020.csv')

# Display basic information about the dataset
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 267 entries, 0 to 266
Data columns (total 9 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   Region                                267 non-null    object
 1   Date                                  267 non-null    object
 2   Frequency                             267 non-null    object
 3   Estimated Unemployment Rate (%)       267 non-null    float64
 4   Estimated Employed                    267 non-null    int64
 5   Estimated Labour Participation Rate (%) 267 non-null    float64
 6   Region.1                              267 non-null    object
 7   longitude                             267 non-null    float64
 8   latitude                              267 non-null    float64
dtypes: float64(4), int64(1), object(4)
memory usage: 18.9+ KB
```

```
# Identify numerical columns
numerical_columns = data.select_dtypes(include=['float64', 'int64']).columns

# Calculate mean, median, and standard deviation for each numerical column
descriptive_stats = data[numerical_columns].agg(['mean', 'median', 'std'])
descriptive_stats
```

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

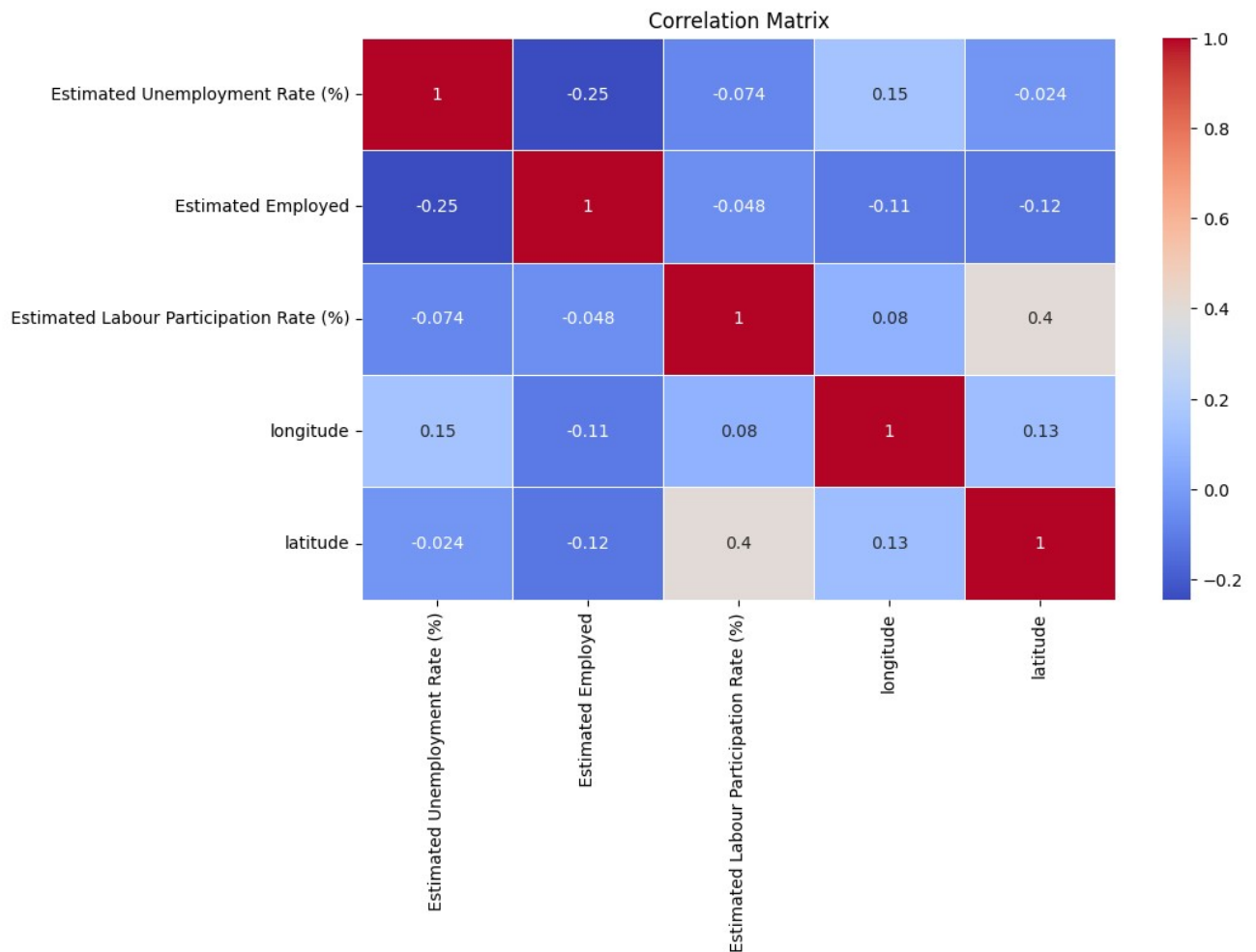
	Estimated Unemployment Rate (%)	Estimated Employed	Estimated Labour Participation Rate (%)	longitude	latitude
mean	12.236929	1.396211e+07	41.681573	22.826048	80.532425
median	9.650000	9.732417e+06	40.390000	23.610200	79.019300

```
import seaborn as sns
import matplotlib.pyplot as plt

# Compute the correlation matrix
correlation_matrix = data[numerical_columns].corr()

# Plot the heatmap
```

```
# Plot the heatmap
plt.figure(figsize=(10, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Correlation Matrix')
plt.show()
```



```
# Clean up column names
```

```
data.columns = data.columns.str.strip()

# Scatter plot for Estimated Employed vs. Estimated Unemployment Rate (%)
plt.figure(figsize=(10, 6))
sns.scatterplot(x=data['Estimated Employed'], y=data['Estimated Unemployment Rate (%)'])
plt.title('Estimated Employed vs. Estimated Unemployment Rate (%)')
plt.xlabel('Estimated Employed')
plt.ylabel('Estimated Unemployment Rate (%)')
plt.show()

# Scatter plot for Estimated Labour Participation Rate (%) vs. Estimated Employed
plt.figure(figsize=(10, 6))
sns.scatterplot(x=data['Estimated Labour Participation Rate (%)'], y=data['Estimated Empl
plt.title('Estimated Labour Participation Rate (%) vs. Estimated Employed')
plt.xlabel('Estimated Labour Participation Rate (%)')
plt.ylabel('Estimated Employed')
plt.show()
```

