

```
# Importing necessary libraries

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

# Load the dataset (replace the path with the actual file location)
url = "your_dataset_url_or_file_path.csv" # Put the dataset URL or file path here
df = pd.read_csv(url)

# Display the first few rows of the dataset to understand its structure
print(df.head())

# Check for missing values and data types
print(df.info())

# Data preprocessing: handle missing data or fill missing values
df.fillna(method='ffill', inplace=True) # Forward fill missing data (example)

# Check for duplicates
df.drop_duplicates(inplace=True)

# Basic statistics on unemployment rate column (assuming it's named 'unemployment_rate')
print(df['unemployment_rate'].describe())

# Data visualization: Line plot of unemployment rate over time
# Assume the data has a 'date' column and 'unemployment_rate' column
df['date'] = pd.to_datetime(df['date'])
plt.figure(figsize=(12, 6))
plt.plot(df['date'], df['unemployment_rate'], color='blue', label='Unemployment Rate')
plt.title("Unemployment Rate Over Time")
```

```
plt.xlabel("Date")
plt.ylabel("Unemployment Rate (%)")
plt.legend()
plt.grid(True)
plt.show()
```

# If there's a specific time period (e.g., COVID), filter the data and analyze trends

```
covid_start = '2020-03-01'
```

```
covid_end = '2022-12-31'
```

```
df_covid = df[(df['date'] >= covid_start) & (df['date'] <= covid_end)]
```

```
plt.figure(figsize=(12, 6))
```

```
plt.plot(df_covid['date'], df_covid['unemployment_rate'], color='red', label='COVID Unemployment Rate')
```

```
plt.title("Unemployment Rate During COVID-19")
```

```
plt.xlabel("Date")
```

```
plt.ylabel("Unemployment Rate (%)")
```

```
plt.legend()
```

```
plt.grid(True)
```

```
plt.show()
```

# You can also analyze trends before and after COVID by creating subsets:

```
df_pre_covid = df[df['date'] < covid_start]
```

```
df_post_covid = df[df['date'] > covid_end]
```

# You can perform additional statistical analysis or even build forecasting models using ARIMA or other methods if needed.