**Step 1: Data Manipulation**

First, let's load and inspect the data:

# Import necessary libraries

import pandas as pd

# Load the dataset

file\_path = 'path\_to\_your\_csv\_file.csv'

df = pd.read\_csv(file\_path)

# Display the first few rows of the dataframe

df.head()

**Step 2: Data Visualization**

Now, let's create some visualizations based on the combinations we discussed:

**1. Match Overview with Timeline**

import matplotlib.pyplot as plt

import seaborn as sns

# Convert 'Datetime' to datetime type

df['Datetime'] = pd.to\_datetime(df['Datetime'])

# Plot a timeline of matches

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

sns.scatterplot(data=df, x='Datetime', y='Home Team Name', hue='Stage', style='City', s=100)

plt.title('Match Overview with Timeline')

plt.xlabel('Datetime')

plt.ylabel('Home Team Name')

plt.legend(title='Stage', bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**2. Goals Comparison**

# Plot goals comparison

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

df\_goals = df[['Home Team Name', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Home Team Name',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.barplot(data=df\_goals, x='Home Team Name', y='Goals', hue='Team Type')

plt.title('Goals Comparison')

plt.xlabel('Team Name')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**3. Attendance by Location**

# Plot attendance by location

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

sns.barplot(data=df, x='City', y='Attendance', palette='viridis')

plt.title('Attendance by Location')

plt.xlabel('City')

plt.ylabel('Attendance')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**4. Team Performance**

# Plot team performance

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

df\_performance = df[['Home Team Name', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Home Team Name',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.barplot(data=df\_performance, x='Home Team Name', y='Goals', hue='Team Type')

plt.title('Team Performance')

plt.xlabel('Home Team Name')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**5. Referee Influence**

# Plot referee influence

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

df\_referee = df[['Referee', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Referee',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.scatterplot(data=df\_referee, x='Referee', y='Goals', hue='Team Type', style='Team Type', s=100)

plt.title('Referee Influence')

plt.xlabel('Referee')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

**Putting It All Together**

Here is the full Jupyter Notebook code with all the steps combined:

# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

file\_path = 'path\_to\_your\_csv\_file.csv'

df = pd.read\_csv(file\_path)

# Convert 'Datetime' to datetime type

df['Datetime'] = pd.to\_datetime(df['Datetime'])

# Display the first few rows of the dataframe

df.head()

# Plot a timeline of matches

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

sns.scatterplot(data=df, x='Datetime', y='Home Team Name', hue='Stage', style='City', s=100)

plt.title('Match Overview with Timeline')

plt.xlabel('Datetime')

plt.ylabel('Home Team Name')

plt.legend(title='Stage', bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

# Plot goals comparison

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

df\_goals = df[['Home Team Name', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Home Team Name',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.barplot(data=df\_goals, x='Home Team Name', y='Goals', hue='Team Type')

plt.title('Goals Comparison')

plt.xlabel('Team Name')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

# Plot attendance by location

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

sns.barplot(data=df, x='City', y='Attendance', palette='viridis')

plt.title('Attendance by Location')

plt.xlabel('City')

plt.ylabel('Attendance')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

# Plot team performance

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

df\_performance = df[['Home Team Name', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Home Team Name',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.barplot(data=df\_performance, x='Home Team Name', y='Goals', hue='Team Type')

plt.title('Team Performance')

plt.xlabel('Home Team Name')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

# Plot referee influence

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

df\_referee = df[['Referee', 'Home Team Goals', 'Away Team Goals']].melt(id\_vars='Referee',

value\_vars=['Home Team Goals', 'Away Team Goals'],

var\_name='Team Type',

value\_name='Goals')

sns.scatterplot(data=df\_referee, x='Referee', y='Goals', hue='Team Type', style='Team Type', s=100)

plt.title('Referee Influence')

plt.xlabel('Referee')

plt.ylabel('Goals')

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()