Football

<https://www.kaggle.com/datasets/martj42/international-football-results-from-1872-to-2017>

import pandas as pd

df = pd.read\_csv('abfss://aamna@youexcel.dfs.core.windows.net/football/goalscorers.csv',

storage\_options = {'account\_key':'/dhYW+OiOeruQ2q2lRlOVQHX+VYV50A1MDZYASkOS6BP6NqNMeL12c2uqhhUzDp92mNhSurH6uI++AStKYyqpQ=='})

display(df)

1. Total Goals Scored by Each Team
2. import pandas as pd  
   import matplotlib.pyplot as plt  
     
   # Load the data  
   goalscorers = pd.read\_csv('goalscorers.csv')  
   results = pd.read\_csv('results.csv')  
     
   # Aggregate goals scored by each team  
   home\_goals = results.groupby('home\_team')['home\_score'].sum()  
   away\_goals = results.groupby('away\_team')['away\_score'].sum()  
   total\_goals = pd.concat([home\_goals, away\_goals], axis=1).fillna(0)  
   total\_goals.columns = ['HomeGoals', 'AwayGoals']  
   total\_goals['TotalGoals'] = total\_goals['HomeGoals'] + total\_goals['AwayGoals']  
   total\_goals = total\_goals.sort\_values(by='TotalGoals', ascending=False)  
     
   plt.figure(figsize=(14, 8))  
   total\_goals['TotalGoals'].plot(kind='bar')  
   plt.title('Total Goals Scored by Each Team')  
   plt.xlabel('Team')  
   plt.ylabel('Total Goals')  
   plt.xticks(rotation=90)  
   plt.grid(True)  
   plt.show()
3. Goals Scored Per Year

# Convert 'date' to datetime  
results['date'] = pd.to\_datetime(results['date'])  
results['Year'] = results['date'].dt.year  
  
# Calculate total goals per year  
results['TotalGoals'] = results['home\_score'] + results['away\_score']  
goals\_per\_year = results.groupby('Year')['TotalGoals'].sum()  
  
plt.figure(figsize=(12, 6))  
goals\_per\_year.plot(kind='line')  
plt.title('Total Goals Scored Per Year')  
plt.xlabel('Year')  
plt.ylabel('Total Goals')  
plt.grid(True)  
plt.show()

1. Number of Penalties Per Team
2. # Aggregate penalties by team  
   penalties = goalscorers[goalscorers['penalty'] == 1]  
   penalties\_home = penalties.groupby('home\_team').size()  
   penalties\_away = penalties.groupby('away\_team').size()  
   total\_penalties = pd.concat([penalties\_home, penalties\_away], axis=1).fillna(0)  
   total\_penalties.columns = ['HomePenalties', 'AwayPenalties']  
   total\_penalties['TotalPenalties'] = total\_penalties['HomePenalties'] + total\_penalties['AwayPenalties']  
   total\_penalties = total\_penalties.sort\_values(by='TotalPenalties', ascending=False)  
     
   plt.figure(figsize=(14, 8))  
   total\_penalties['TotalPenalties'].plot(kind='bar')  
   plt.title('Total Penalties by Each Team')  
   plt.xlabel('Team')  
   plt.ylabel('Total Penalties')  
   plt.xticks(rotation=90)  
   plt.grid(True)  
   plt.show()
3. Number of Own Goals by Team

# Filter own goals  
own\_goals = goalscorers[goalscorers['own\_goal'] == 1]  
own\_goals\_home = own\_goals.groupby('home\_team').size()  
own\_goals\_away = own\_goals.groupby('away\_team').size()  
total\_own\_goals = pd.concat([own\_goals\_home, own\_goals\_away], axis=1).fillna(0)  
total\_own\_goals.columns = ['HomeOwnGoals', 'AwayOwnGoals']  
total\_own\_goals['TotalOwnGoals'] = total\_own\_goals['HomeOwnGoals'] + total\_own\_goals['AwayOwnGoals']  
total\_own\_goals = total\_own\_goals.sort\_values(by='TotalOwnGoals', ascending=False)  
  
plt.figure(figsize=(14, 8))  
total\_own\_goals['TotalOwnGoals'].plot(kind='bar')  
plt.title('Total Own Goals by Each Team')  
plt.xlabel('Team')  
plt.ylabel('Total Own Goals')  
plt.xticks(rotation=90)  
plt.grid(True)  
plt.show()

1. Number of Matches by Tournament
2. tournament\_matches = results.groupby('tournament').size()  
     
   plt.figure(figsize=(12, 6))  
   tournament\_matches.plot(kind='bar')  
   plt.title('Number of Matches by Tournament')  
   plt.xlabel('Tournament')  
   plt.ylabel('Number of Matches')  
   plt.xticks(rotation=90)  
   plt.grid(True)  
   plt.show()
3. Number of Shoutouts by Team

# Aggregate shoutouts by team  
shoutouts = pd.read\_csv('shoutouts.csv')  
home\_shoutouts = shoutouts.groupby('home\_team').size()  
away\_shoutouts = shoutouts.groupby('away\_team').size()  
total\_shoutouts = pd.concat([home\_shoutouts, away\_shoutouts], axis=1).fillna(0)  
total\_shoutouts.columns = ['HomeShoutouts', 'AwayShoutouts']  
total\_shoutouts['TotalShoutouts'] = total\_shoutouts['HomeShoutouts'] + total\_shoutouts['AwayShoutouts']  
total\_shoutouts = total\_shoutouts.sort\_values(by='TotalShoutouts', ascending=False)  
  
plt.figure(figsize=(14, 8))  
total\_shoutouts['TotalShoutouts'].plot(kind='bar')  
plt.title('Total Shoutouts by Each Team')  
plt.xlabel('Team')  
plt.ylabel('Total Shoutouts')  
plt.xticks(rotation=90)  
plt.grid(True)  
plt.show()

1. Number of Goals per Match Over Time
2. # Calculate goals per match  
   results['GoalsPerMatch'] = results['home\_score'] + results['away\_score']  
   matches\_per\_year = results.groupby('Year')['GoalsPerMatch'].mean()  
     
   plt.figure(figsize=(12, 6))  
   matches\_per\_year.plot(kind='line')  
   plt.title('Average Goals Per Match by Year')  
   plt.xlabel('Year')  
   plt.ylabel('Average Goals Per Match')  
   plt.grid(True)  
   plt.show()