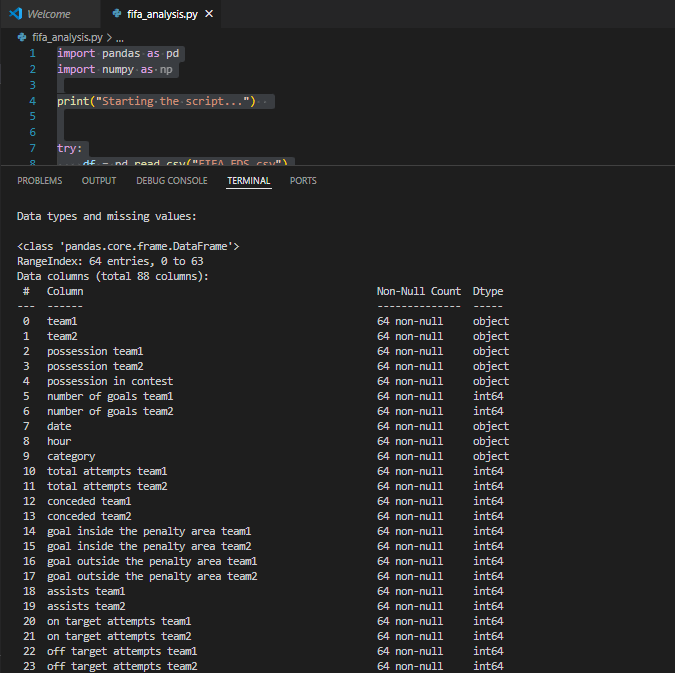
EDS ACTIVITY-1

Name: Sanika Jaind

Roll no.: CS7-48

PRN: 202401110032

Dataset: FIFA DATASETS



#1. total no of matches played

print("1. Total matches:", df.shape[0])

OUTPUT : 

#2. which team appeared most of the time

print("2. Most frequent team1:", df['team1'].value\_counts().idxmax())

OUTPUT: 

#3. Which match had the highest number of total goals?

df['total\_goals'] = df['number of goals team1'] + df['number of goals team2']

print(df.loc[df['total\_goals'].idxmax(), ['team1', 'team2', 'total\_goals']])

OUTPUT: 

#4. average possession for team1

df['possession team1'] = df['possession team1'].str.replace('%', '').astype(float)

print("4. Avg possession team1:", df['possession team1'].mean())

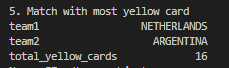
OUTPUT: 

#5. which match had the most yellow card

print("5. Match with most yellow card")

df['total\_yellow\_cards'] = df['yellow cards team1'] + df['yellow cards team2']

print(df.loc[df['total\_yellow\_cards'].idxmax(), ['team1', 'team2', 'total\_yellow\_cards']])

OUTPUT: 

#6. total red cards in the tournament

print("6. Total red cards:", df['red cards team1'].sum() + df['red cards team2'].sum())

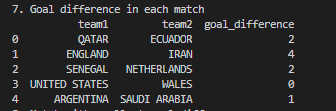
OUTPUT: 

#7. Goal difference in each match

print("7. Goal difference in each match")

df['goal\_difference'] = abs(df['number of goals team1'] - df['number of goals team2'])

print(df[['team1', 'team2', 'goal\_difference']].head())

OUTPUT: 

#8. match with smallest goal difference

print("8. Match with smallest goal difference: ")

print(df[df['goal\_difference'] == df['goal\_difference'].min()][['team1', 'team2', 'goal\_difference']])

OUTPUT: 

#9. average number of fouls per match

df['fouls\_total'] = df['fouls against team1'] + df['fouls against team2']

print("9. Avg fouls per match:", df['fouls\_total'].mean())

OUTPUT: 

#10. find number  of matches with more than 5 goals total

print("10. Matches with >5 total goals:", df[df['total\_goals'] > 5].shape[0])

OUTPUT: 

#11. total number of own goals

print("11. Total own goals:", df['own goals team1'].sum() + df['own goals team2'].sum())

OUTPUT: 

#12. team with most penalties scored

print("12. Team with most penalties scored:")

pen1 = df.groupby('team1')['penalties scored team1'].sum()

pen2 = df.groupby('team2')['penalties scored team2'].sum()

pen\_total = pen1.add(pen2, fill\_value=0)

print(pen\_total.idxmax())

OUTPUT: 

#13. team with a highest number of fouls

print("13. Team with the highest number of fouls committed:")

fouls\_by\_team1 = df.groupby('team2')['fouls against team1'].sum()

fouls\_by\_team2 = df.groupby('team1')['fouls against team2'].sum()

total\_fouls\_committed = fouls\_by\_team1.add(fouls\_by\_team2, fill\_value=0)

print(total\_fouls\_committed.idxmax())

OUTPUT: 

#14. team that applied most defensive pressure

print("14. Team that applied the most defensive pressure:")

d1 = df.groupby('team1')['defensive pressures applied team1'].sum()

d2 = df.groupby('team2')['defensive pressures applied team2'].sum()

total\_def\_pressures = d1.add(d2, fill\_value=0)

print(total\_def\_pressures.idxmax())

OUTPUT: 

#15. average  shots on target per team per match

print("15. Average shots on target per team per match:")

df['avg\_on\_target'] = (df['on target attempts team1'] + df['on target attempts team2']) / 2

print(df['avg\_on\_target'].mean())

OUTPUT: 

#16. team that never scored a penalty

print("16. Teams that never scored a penalty:")

all\_teams = pd.unique(df[['team1', 'team2']].values.ravel())

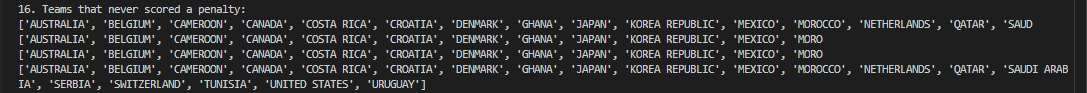
pen\_team1 = df.groupby('team1')['penalties scored team1'].sum()

pen\_team2 = df.groupby('team2')['penalties scored team2'].sum()

pen\_total = pen\_team1.add(pen\_team2, fill\_value=0)

no\_penalty\_teams = pen\_total[pen\_total == 0].index.tolist()

print(no\_penalty\_teams)

OUTPUT: 

#17. team with the lowest yellow and red card

yellow\_cards\_team1 = df.groupby('team1')['yellow cards team1'].sum()

yellow\_cards\_team2 = df.groupby('team2')['yellow cards team2'].sum()

red\_cards\_team1 = df.groupby('team1')['red cards team1'].sum()

red\_cards\_team2 = df.groupby('team2')['red cards team2'].sum()

total\_yellow\_cards = yellow\_cards\_team1.add(yellow\_cards\_team2, fill\_value=0)

total\_red\_cards = red\_cards\_team1.add(red\_cards\_team2, fill\_value=0)

fair\_play\_score = total\_yellow\_cards.add(total\_red\_cards, fill\_value=0)

fair\_play\_team = fair\_play\_score.idxmin()

print(f"17. Team with the lowest yellow and red cards (Fair Play Match): {fair\_play\_team}")

OUTPUT: 

#18. match with the most total attempts

df['total\_attempts'] = df['total attempts team1'] + df['total attempts team2']

max\_attempts\_match = df.loc[df['total\_attempts'].idxmax()]

print(f"18. Match with the most total attempts: {max\_attempts\_match['team1']} vs {max\_attempts\_match['team2']} - {max\_attempts\_match['total\_attempts']} attempts")

OUTPUT: 

#19. team with the most successful passes

total\_passes\_completed\_team1 = df['passes completed team1'].sum()

total\_passes\_completed\_team2 = df['passes completed team2'].sum()

if total\_passes\_completed\_team1 > total\_passes\_completed\_team2:

    print(f"19. Team with the most successful passes: Team1 with {total\_passes\_completed\_team1} successful passes")

else:

    print(f"19. Team with the most successful passes: Team2 with {total\_passes\_completed\_team2} successful passes")

OUTPUT: 

#20. match with the highest number of free kicks

df['total\_free\_kicks'] = df['free kicks team1'] + df['free kicks team2']

max\_free\_kicks\_match = df.loc[df['total\_free\_kicks'].idxmax()]

print(f"20. Match with the highest number of free kicks: {max\_free\_kicks\_match['team1']} vs {max\_free\_kicks\_match['team2']} on {max\_free\_kicks\_match['date']} with {max\_free\_kicks\_match['total\_free\_kicks']} free kicks.")

OUTPUT: 