



**BIG DATA ANALYSIS**  
FIFA WORLD CUP



# 01. Data (Overview)

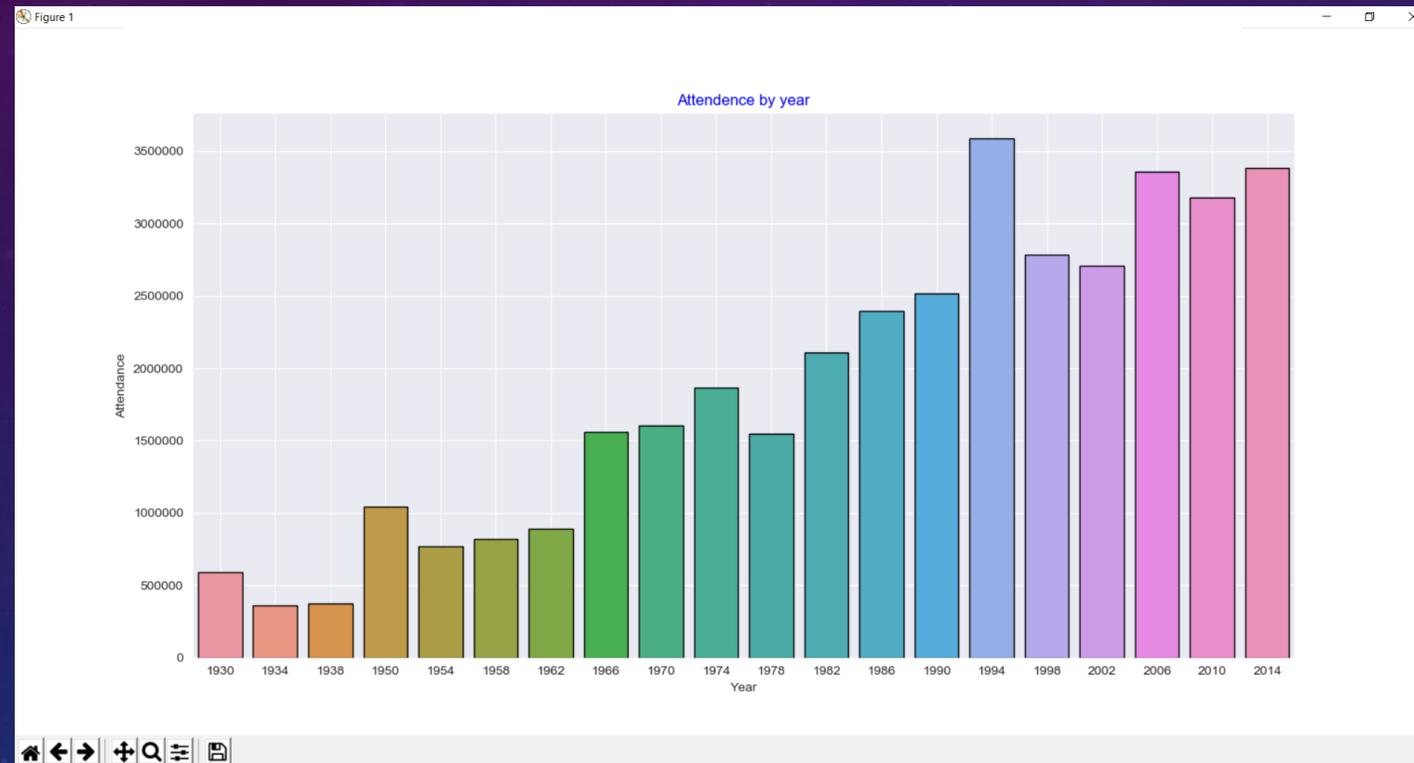
```
C:\WINDOWS\system32\cmd.exe
C:\CloudStation\University\2019 경기대학교\nodejs\FIFA\code>python 01_data-overview.py
MATCHES - DATA
  Year      Datetime      Stage  ... Home Team Initials  Away Team Initials  outcome
0  1930.0    13 Jul 1930 - 15:00  Group 1  ...          FRA          MEX      NaN
1  1930.0    13 Jul 1930 - 15:00  Group 4  ...          USA          BEL      NaN
2  1930.0    14 Jul 1930 - 12:45  Group 2  ...          YUG          BRA      NaN

[3 rows x 21 columns]
PLAYERS - DATA
  RoundID  MatchID  Team  Initials  ...  Player Name  Position  Event
0      201    1096   FRA  ...  Alex THEPOT    GK      NaN
1      201    1096   MEX  ...  Oscar BONFIGLIO  GK      NaN
2      201    1096   FRA  ...  Marcel LANGILLER  NaN    G40'

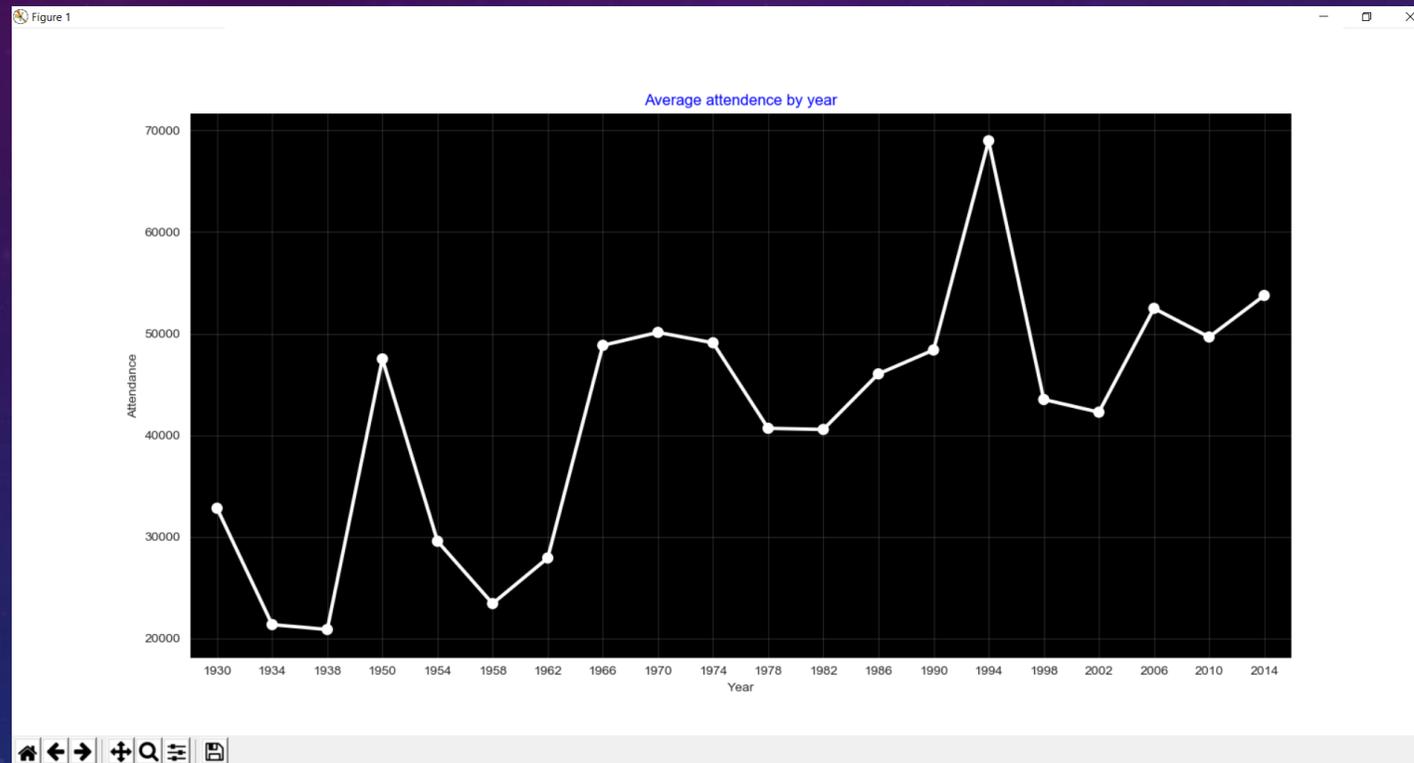
[3 rows x 9 columns]
WORLD CUPS - DATA
  Year  Country  Winner  Runners-Up  ...  QualifiedTeams  MatchesPlayed  Attendance  Year1
0  1930  Uruguay  Uruguay  Argentina  ...           13           18    590.549  NaN
1  1934   Italy   Italy  Czechoslovakia  ...           16           17    363.000  NaN
2  1938   France   Italy    Hungary  ...           15           18    375.700  NaN

[3 rows x 11 columns]
```

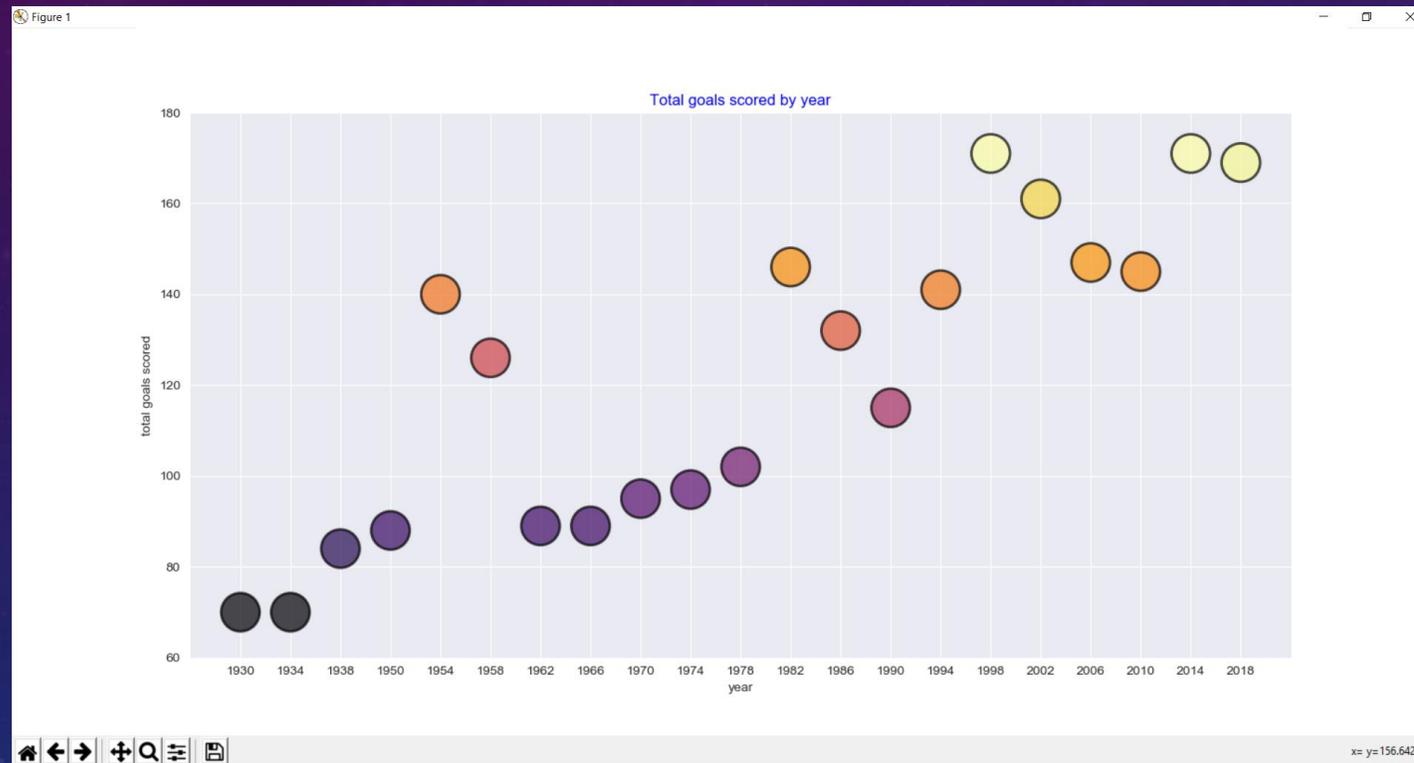
## 02. 연도별 관중 수



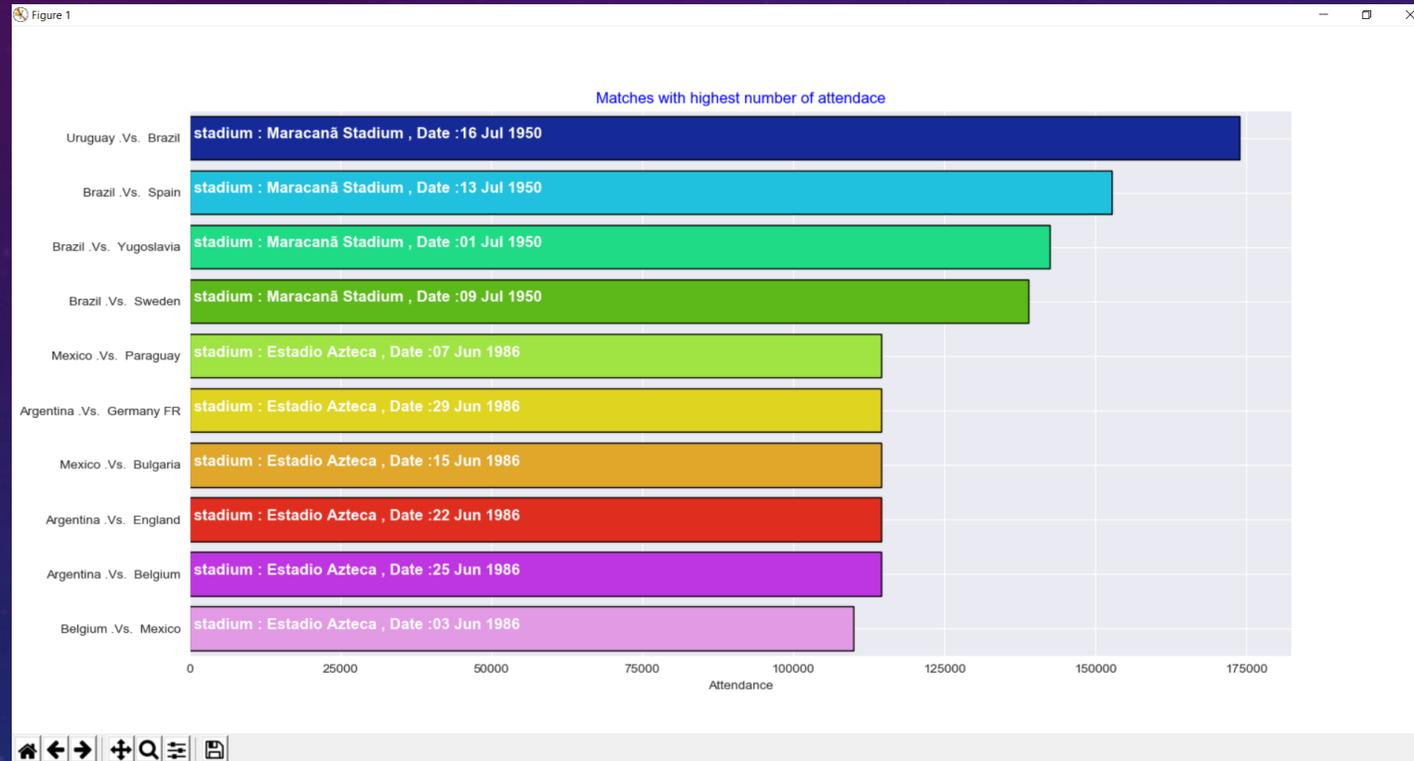
# 03. 연도별 평균 관중 수



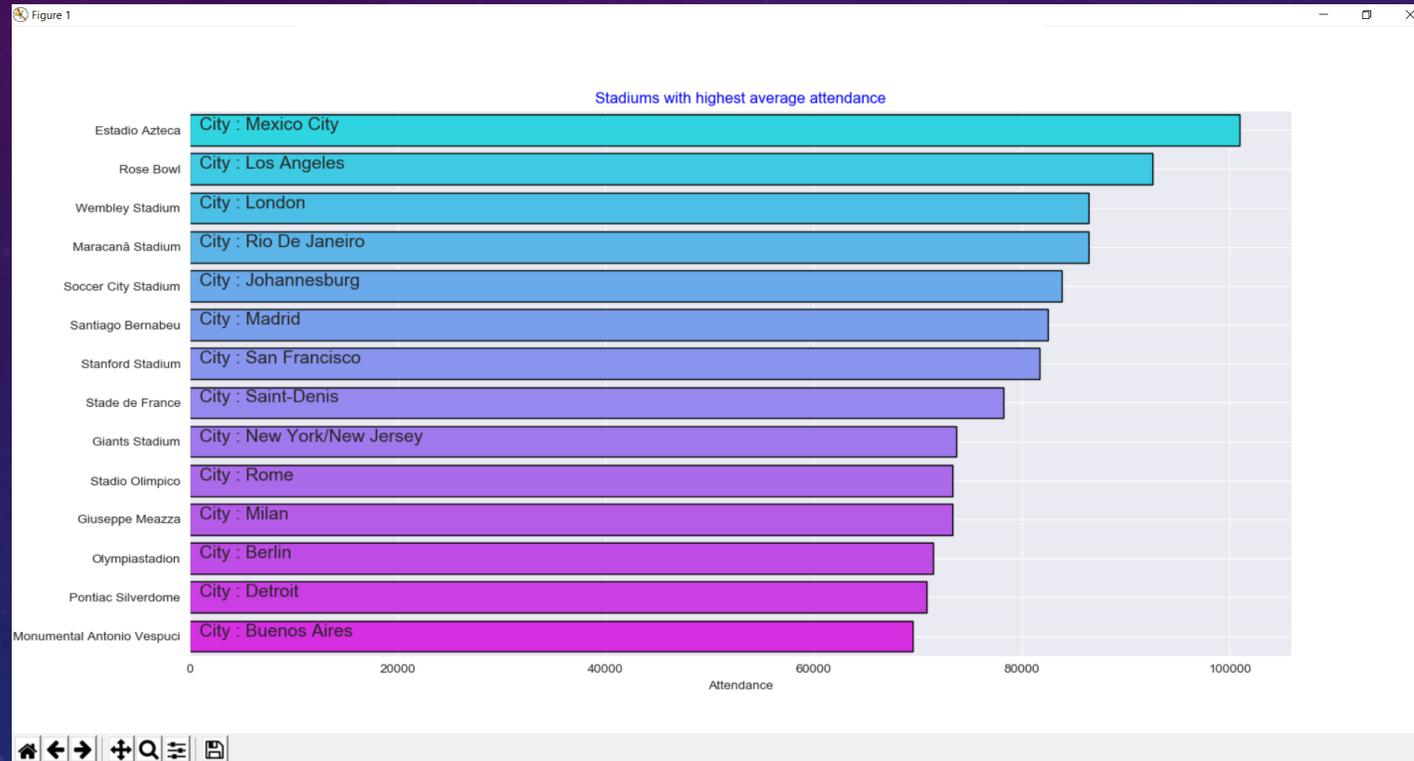
# 04. 연도별 총 득점 수



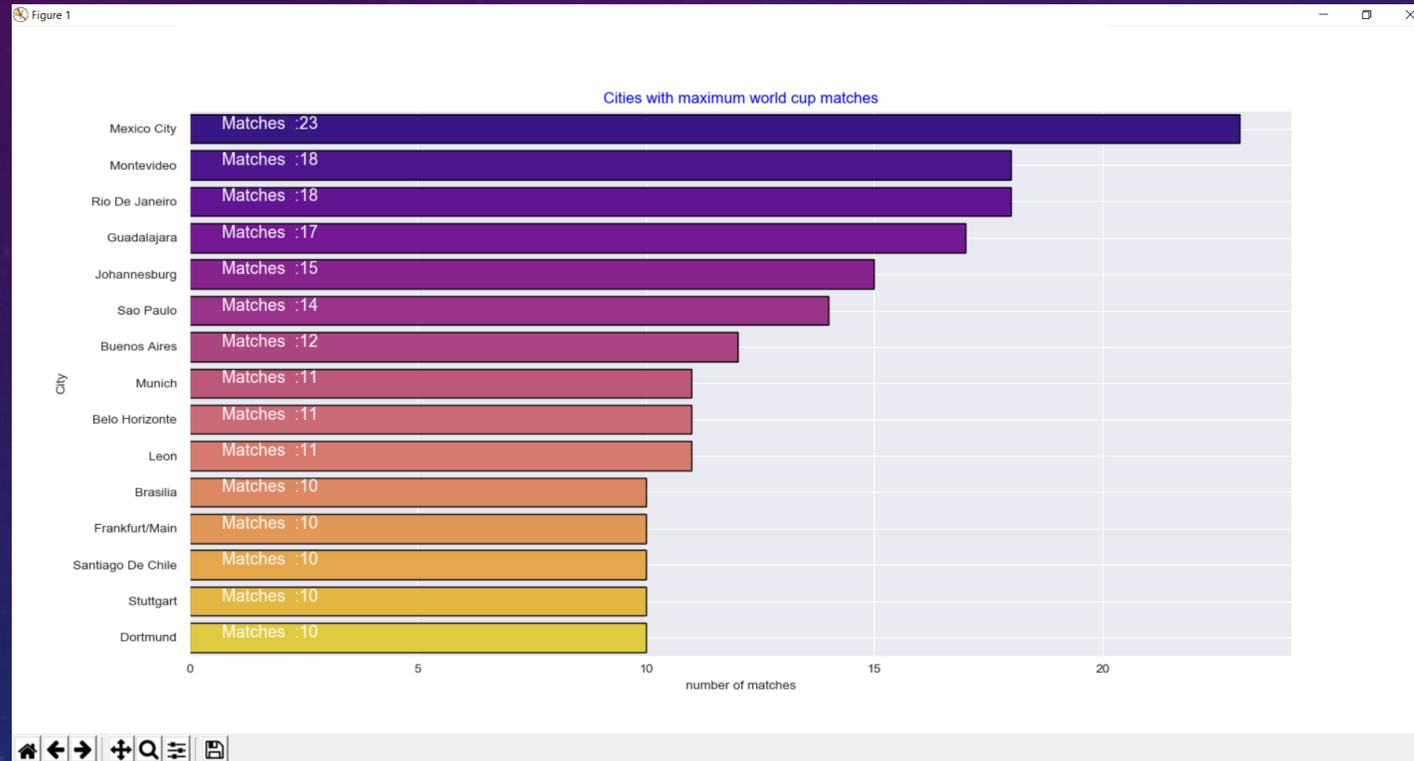
# 05. 경기당 관중 수 (내림차순)



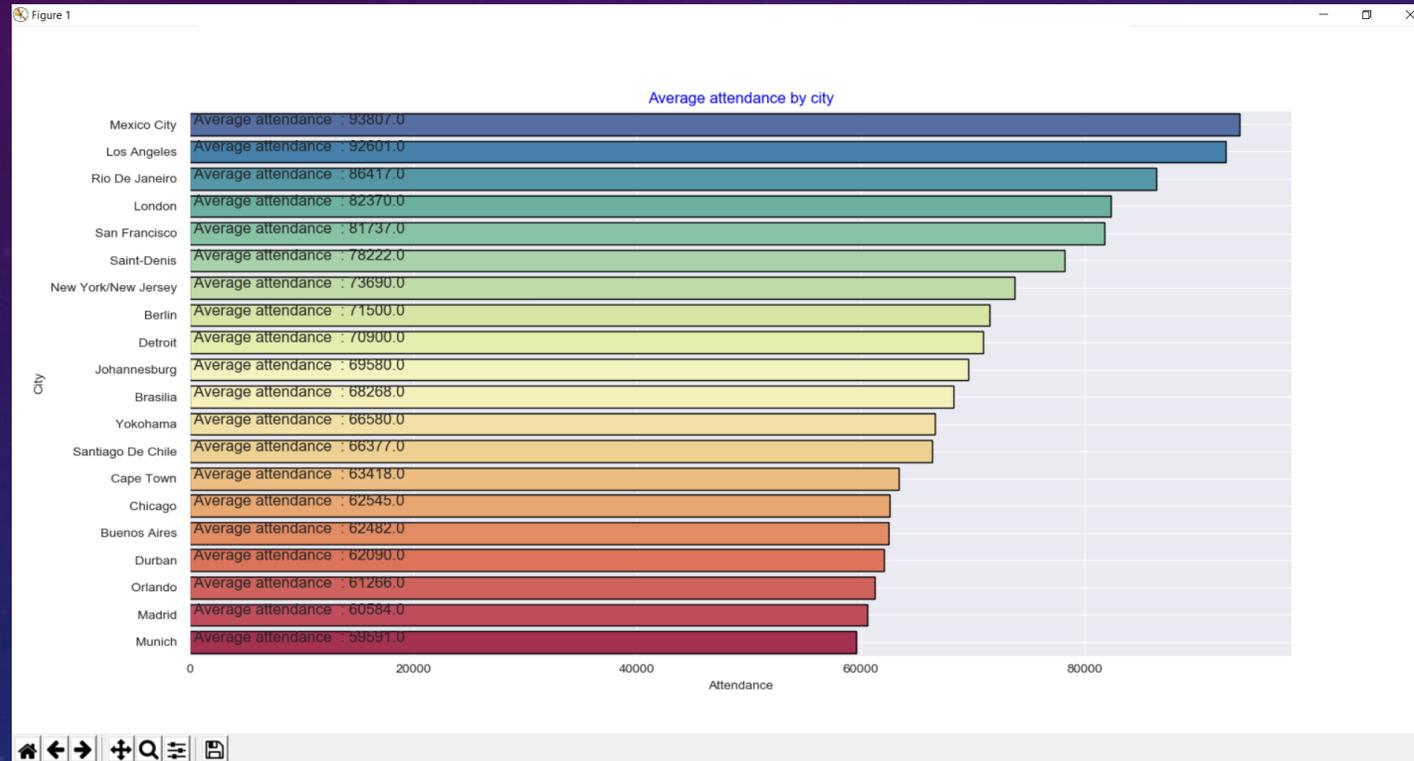
## 06. 경기장 별 평균 관중 수 (내림차순)



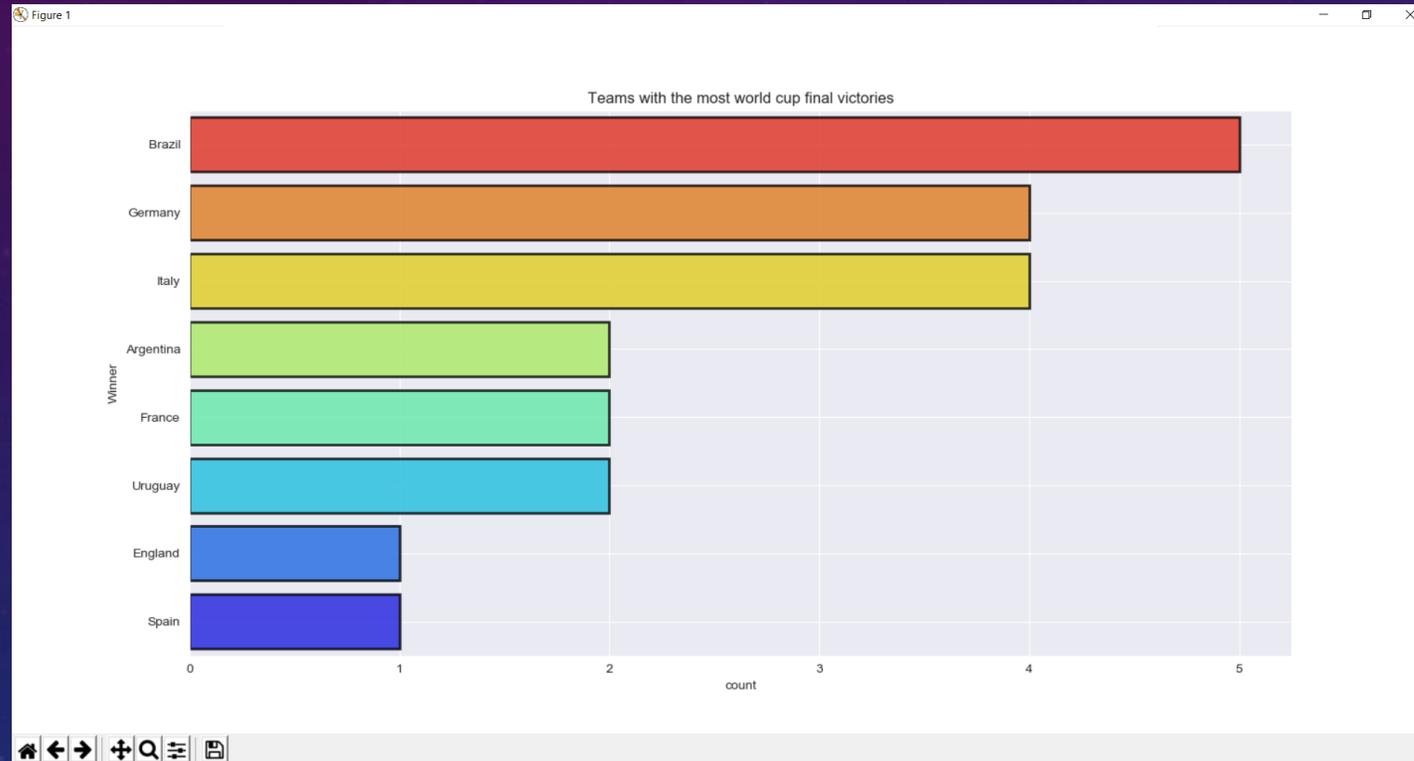
# 07. 경기 개최 도시 (내림차순)



# 08. 도시별 평균 관중 수 (내림차순)



# 09. 최종 우승 국가



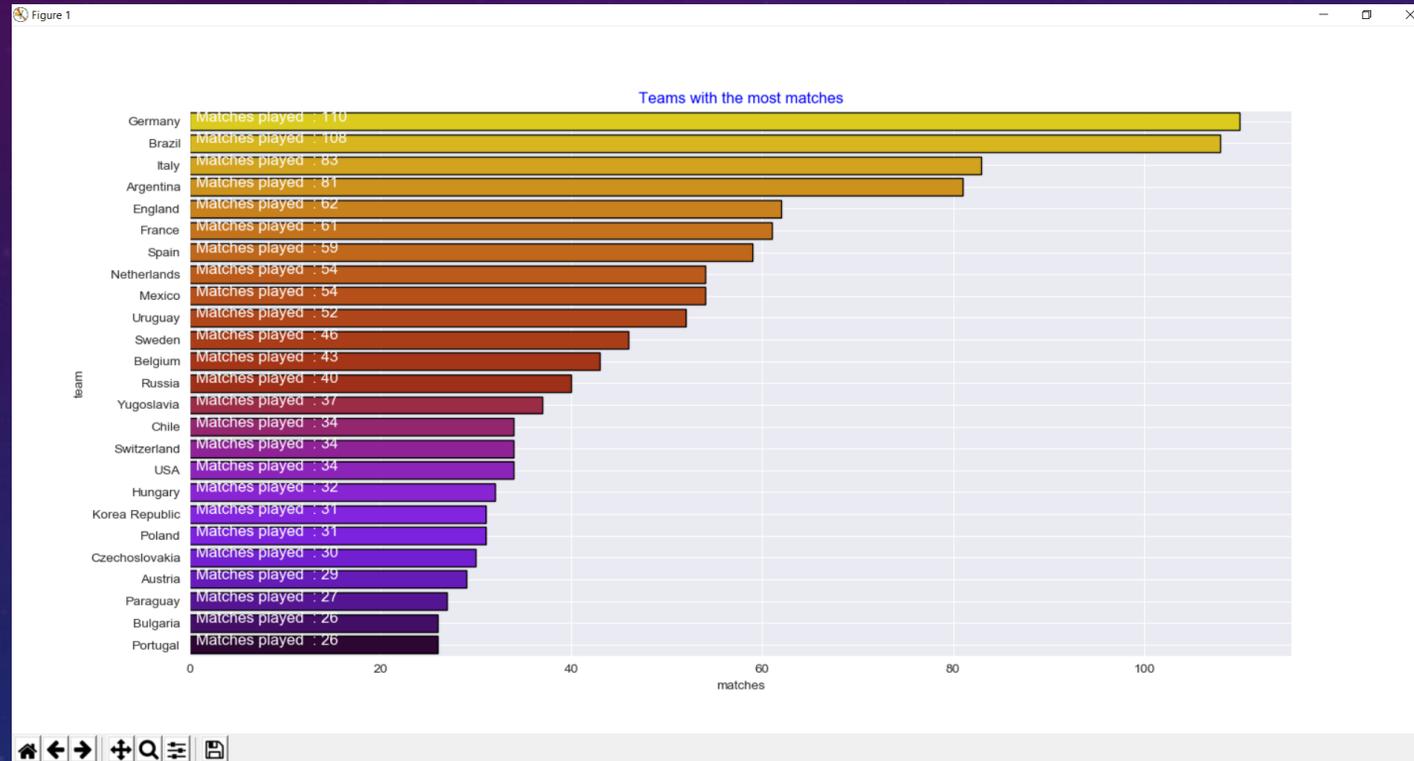
# 10. 최종 1,2위



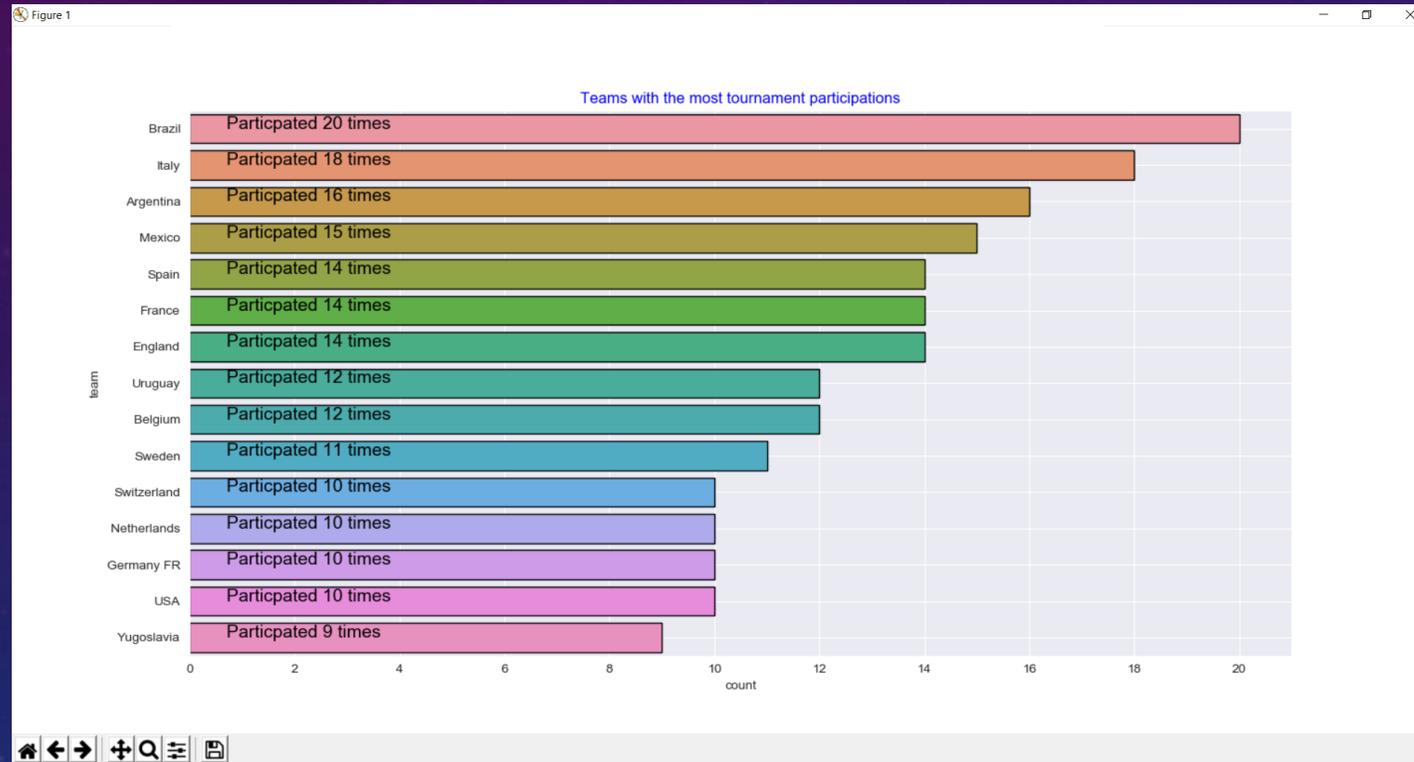
# 11. 최종 3,4위



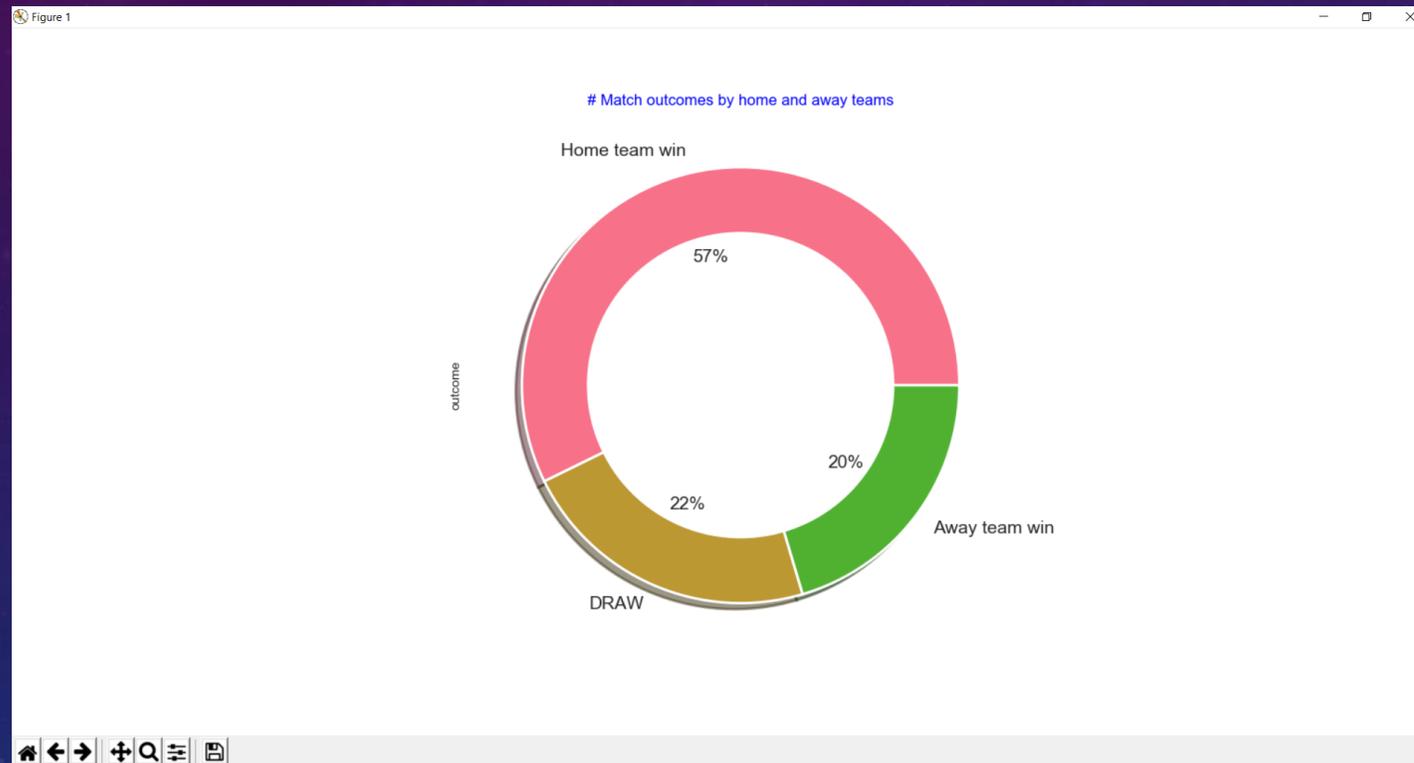
## 12. 팀(국가)별 경기 수 (내림차순)



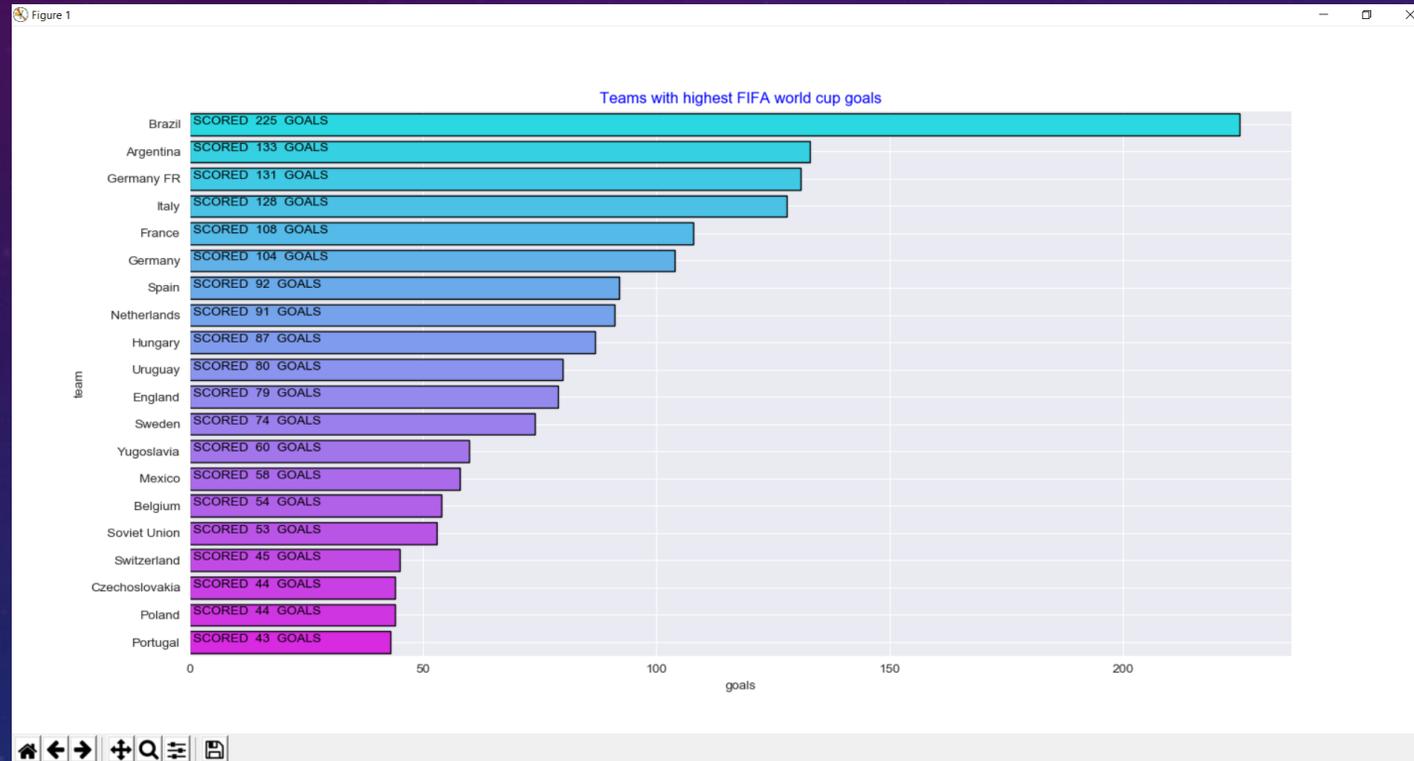
# 13. 팀(국가)별 토너먼트 경기 수 (내림차순)



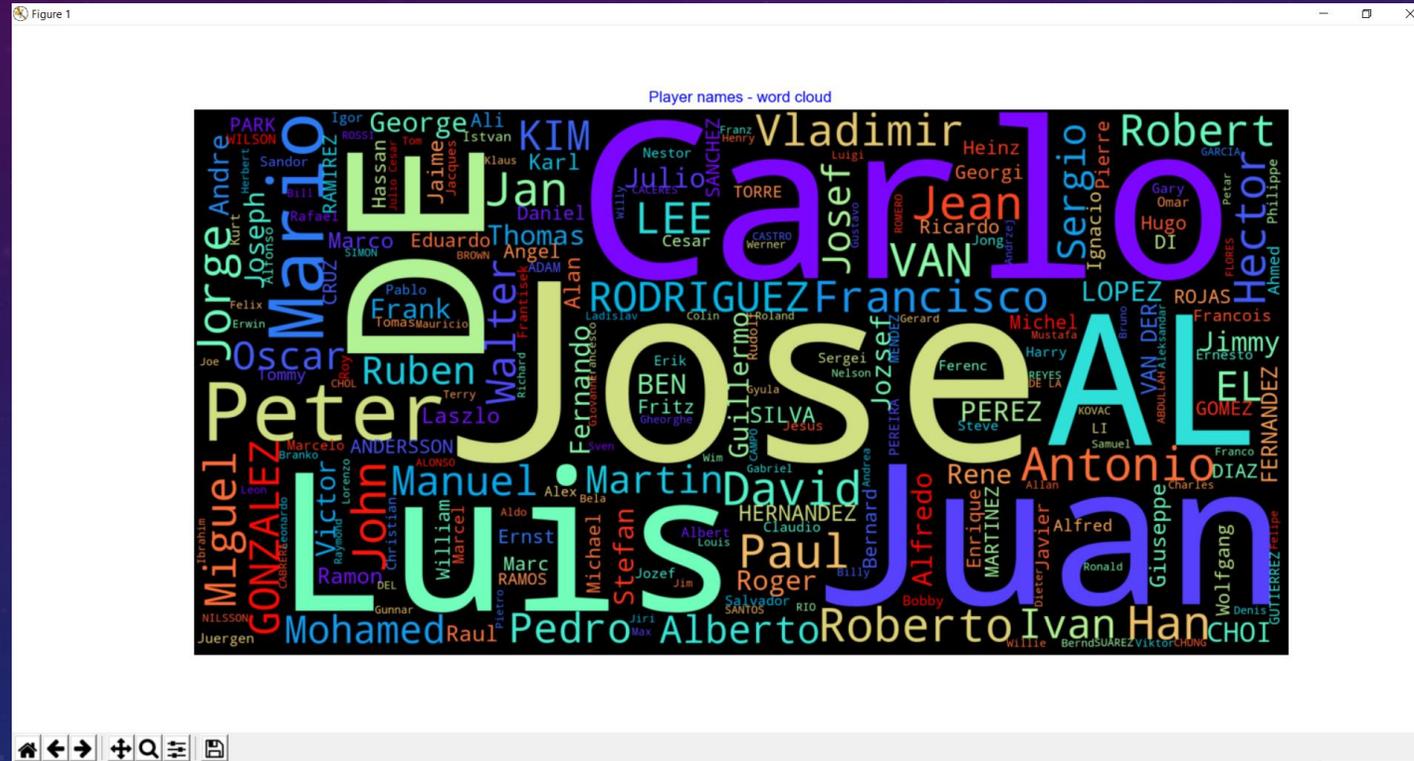
# 14. Home & Away 경기 결과



# 15. 팀(국가)별 총 득점 수 (내림차순)



# 16. 출전 선수 WordCloud



# 결론

1. 우승 국가 예측을 통한 수익 산출
2. 대량 정보의 분석과 확률을 통한 객관적인 국가대표 전력분석
3. 월드컵 진출 라운드에 따른 각 국가별 인프라 제공을 사전에 조절하여 비용 절감