# 21162101012_CBA_Yash

## Practical 2: NBA Data Exploration

The **National Basketball Association (NBA)** is a professional basketball league in North America. The league is composed of 30 teams (29 in the United States and 1 in Canada) and is one of the four major professional sports leagues in the United States and Canada. It is the premier men's professional basketball league in the world.

The dataset provided contains details of the NBA (National Basketball Association) including:

- Name of the player
- Team name
- Age of the player
- Number of games a player has played
- 3-Point Field Goal Attempts and 2-Point Field Goal Attempts by a player
- 2-Point Field Goal Percentage and 3-Point Field Goal Percentage
- Free throws and free throw attempts made by the players
- Offensive Rebounds and Defensive Rebounds by the players
- Assists, Steals, Blocks, Turnovers, Personal Fouls, and Points made by a player
- Start time of the game and the number of minutes played

Using the concept of Data exploration, try to read the given dataset using Python and find out the following information from the dataset.

```
!pip install pandas
import pandas as pd
```

```
# Read the CSV file
df = pd.read_csv('data_DMWp2.csv')
```

```
# Question 1: Average age of NBA players
avg_age = df['Age'].mean()
print(f"1. Average age of NBA players: {avg_age:.2f}")
```

```
1. Average age of NBA players: 26.05
```

```
# Question 2: Number of games played by each player
games_by_player = df.groupby('Player')['Games'].count().reset_index()
print("\n2. Number of games played by each player:")
print(games_by_player)
```

```
2. Number of games played by each player:
              Player  Games
0       Aaron Gordon      1
1      Aaron Holiday      1
2        Abdel Nader      1
3         Al Horford      1
4    Al-Farouq Aminu      1
..               ...    ...
525    Zaza Pachulia      1
526     Zhaire Smith      1
527          Zhou Qi      1
528     Álex Abrines      1
529    Ángel Delgado      1

[530 rows x 2 columns]
```

```
# Question 3: Total number of teams in NBA
total_teams = df['Team'].nunique()
print(f"\n3. Total number of teams in NBA: {total_teams}")
```

```
3. Total number of teams in NBA: 30
```

```
# Question 4: Minimum age of NBA players
min_age = df['Age'].min()
print(f"\n4. Minimum age of NBA players: {min_age}")
```

```
4. Minimum age of NBA players: 19
```

```
# Question 5: Maximum age of NBA players and player details
max_age = df['Age'].max()
max_age_player_details = df[df['Age'] == max_age]
print(f"\n5. Maximum age of NBA players: {max_age}")
print("   Details of the player with the maximum age:")
print(max_age_player_details[['Player', 'Team', 'Age']])
```

```
5. Maximum age of NBA players: 42
   Details of the player with the maximum age:
           Player           Team  Age
302  Vince Carter  Atlanta Hawks   42
```

```
# Question 6: Total games organized in the Eastern region
eastern_games = df[df['Conference'] == 'Eastern']['Games'].sum()
print(f"\n6. Total number of games organized in the Eastern region: {eastern_games}")
```

```
6. Total number of games organized in the Eastern region: 12976
```

```
# Question 7: Number of unique regions where games have been organized
unique_regions = df['Conference'].nunique()
print(f"\n7. Total number of unique regions where games have been organized: {unique_regions}")
```

```
7. Total number of unique regions where games have been organized: 2
```

```
# Question 8: List of players who played for the "Boston Celtics" team
boston_players = df[df['Team'] == 'Boston Celtics']['Player']
print("\n8. List of players who played for the 'Boston Celtics' team:")
print(boston_players)
```

```
8. List of players who played for the 'Boston Celtics' team:
0           Jayson Tatum
1           Kyrie Irving
2           Marcus Smart
3          Marcus Morris
4             Al Horford
5           Jaylen Brown
6         Gordon Hayward
7           Terry Rozier
8           Daniel Theis
9            Aron Baynes
10          Semi Ojeleye
11        Brad Wanamaker
12       Robert Williams
13    Guerschon Yabusele
14            PJ Dozier
15           R.J. Hunter
16           Greg Monroe
Name: Player, dtype: object
```

```
# Question 9: Total games organized in each division
games_by_division = df.groupby('Division')['Games'].sum().reset_index()
print("\n9. Total number of games organized in each division:")
print(games_by_division)
```

```
9. Total number of games organized in each division:
    Division  Games
0   Atlantic   4319
1    Central   4345
2  Northwest   4396
3    Pacific   4453
4  Southeast   4312
5  Southwest   4276
```

```
# Question 10: Player who scored the maximum number of points
max_points_player = df.loc[df['Points'].idxmax()]['Player']
print(f"\n10. Player who scored the maximum number of points: {max_points_player}")
```

```
10. Player who scored the maximum number of points: James Harden
```

```
# Question 11: Player with the lowest number of personal fouls
min_fouls_player = df.loc[df['Personal Fouls'].idxmin()]['Player']
print(f"\n11. Player with the lowest number of personal fouls: {min_fouls_player}")
```

```
11. Player with the lowest number of personal fouls: Greg Monroe
```

```
# Question 12: Player with the most 3-point attempts and success percentage
max_3pt_attempts_player = df.loc[df['3-Point Field Goal Attempts'].idxmax()]['Player']
success_3pt_percentage = (df['3-Point Field Goals'] / df['3-Point Field Goal Attempts']).max() * 100
print(f"\n12. Player with the most 3-point attempts: {max_3pt_attempts_player}")
print(f"    Success percentage of 3-point field goals: {success_3pt_percentage:.2f}%")
```

```
12. Player with the most 3-point attempts: James Harden
    Success percentage of 3-point field goals: 100.00%
```

```
# Question 13: Average points scored by all players
avg_points = df['Points'].mean()
print(f"\n13. Average points scored by all players: {avg_points:.2f}")
```

```
13. Average points scored by all players: 439.83
```

```
# Question 14: Average age of players division-wise
avg_age_divisionwise = df.groupby('Division')['Age'].mean()
print("\n14. Average age of players division-wise:")
print(avg_age_divisionwise)
```

```
14. Average age of players division-wise:
Division
Atlantic     25.728972
Central      25.736364
Northwest    25.709677
Pacific      26.104762
Southeast    25.979798
Southwest    26.981481
Name: Age, dtype: float64
```

```
# Question 15: Total number of fouls in each team
total_fouls_by_team = df.groupby('Team')['Personal Fouls'].sum()
print("\n15. Total number of fouls in each team:")
print(total_fouls_by_team)
```

```
15. Total number of fouls in each team:
Team
Atlanta Hawks             1932
Boston Celtics            1670
Brooklyn Nets             1763
Charlotte Hornets         1550
Chicago Bulls             1663
Cleveland Cavaliers       1642
Dallas Mavericks          1650
Denver Nuggets            1644
Detroit Pistons           1811
Golden State Warriors     1757
Houston Rockets           1803
Indiana Pacers            1594
Los Angeles Clippers      1913
Los Angeles Lakers        1701
Memphis Grizzlies         1801
Miami Heat                1712
Milwaukee Bucks           1608
Minnesota Timberwolves    1664
New Orleans Pelicans      1732
New York Knicks           1713
Oklahoma City Thunder     1839
Orlando Magic             1526
Philadelphia 76ers        1745
Phoenix Suns              1932
Portland Trail Blazers    1669
Sacramento Kings          1751
San Antonio Spurs         1487
Toronto Raptors           1724
Utah Jazz                 1728
Washington Wizards        1701
Name: Personal Fouls, dtype: int64
```