**Data Science Ex1**

1. Final headers: ['game.id', 'date', 'home\_team\_score', 'period', 'postseason', 'season', ‘status', 'time', 'visitor\_team\_score', 'home\_team.id','home\_team.abbreviation', 'home\_team.city', 'home\_team.conference', 'home\_team.division', 'home\_team.full\_name', 'home\_team.name', 'visitor\_team.id', 'visitor\_team.abbreviation', 'visitor\_team.city', 'visitor\_team.conference', 'visitor\_team.division', 'visitor\_team.full\_name', 'visitor\_team.name', 'id', 'assists', 'blocks', 'defensive\_rebounds', 'three\_point\_pct', 'three\_point\_attempts', 'three\_point\_made', 'field\_goal\_pct', 'field\_goal\_attempts', 'field\_goal\_made', 'free\_throw\_pct', 'free\_throw\_attempts', 'free\_throw\_made', 'minutes\_played', 'offensive\_rebounds', 'personal\_fouls', 'points', 'reb', 'steals', 'turnover', 'game.home\_team\_id', 'player.id', 'player.height\_feet', 'player.height\_inches', 'player.position', 'player.weight\_pounds', 'team\_won', 'abs\_score\_difference', 'player\_full\_name', 'player\_height\_cm', 'home\_player\_stats', 'same\_division', 'points\_per\_minute', 'winning\_team'] – We removed all feature that appeared both in the game and stats data (with different names).
2. We created 3 new features:
   1. Same\_division: A boolean feature that represent if both playing teams are of the same division.
   2. Points\_per\_minute: A feature that present that ration between the amount of points a player scored and the amount of minutes he played in the game.
   3. Winning\_team: the name of the winning team.
3. 1. According to our assessment the team to win the most games is LA Lakers.
   2. According to our assessment the team with the least wins Is Golden State Warriors.