**Steps for Cleaning up data**

1. **Merging Documents** 
   1. Merge following documents into one document. The document should contain player statistics for all years in one file
      1. NBA Players Stats 201617.csv
      2. NBA Players Stats 201718.csv
      3. NBA Players Stats 201819.csv
      4. NBA STATs19-20.csv
   2. Merge following documents to one document. The document should contain player salaries for all years in one file
      1. nba\_stats\_salaries\_16-19.csv
      2. NBA\_salaries\_2019\_present.csv
2. **Cleaning Stats Documents for 2016 to 2019**
   1. Split Player column into player\_id and player\_name
   2. When a player gets traded, a player has multiple rows. Pick the row where Tm(team) is TOT.
      1. 2016-17: 486/595 players are unique
      2. 2017-18: 540/664 players are unique
      3. 2018-19: 530/708 players are unique
      4. 2019-20: 529/651 players are unique
   3. Filter and keep only the players available in all seasons.
   4. Check whether filtered players have salary info
   5. Clean up null values
      1. 3Point percentage was not there. We are replacing with 0
      2. Free Throw percentage
3. **Cleaning Stats Documents for 2019-20**
   1. Check whether player\_id matches with stats data for 2016-19
   2. When a player gets traded, a player has multiple rows. Clean this up
      1. 2019-20: 529/651 players are unique
   3. Clean up null values
4. **Cleaning up Salary for 2016-2019**
   1. Only 415/436 rows have salary values
   2. 2016-17: 436/436 unique playerids
   3. 2016-17: 498/498 unique playerids
   4. 2017-18: 474/474 unique playerids
5. **Cleaning up Salary for 2019-20**
   1. Split playerid and playername
   2. Extract salary for 2019-20