Analysis

Total Observations in 2018: 17981

Total Observations in 2019: 18207

Duplicate records in  2018 dataset :52    (Refer DuplicateData\_2018.xlsx)

Join the data based on ID column           (Refer Final\_JoinedData\_UsingIDColumn.xlsx)

Player match 2018 and 2019 :12539

Player available in  2018 and not available in 2019 :5442

Join the data based on Name Column      (Refer JoinedData\_UsingNameColumn.xlsx)

Player match 2018 and 2019 : 15041

Player available in  2018 and not available in 2019 : 5187

Removed Column From 2018 Dataset

Column 0,Photo,Flag,Club Logo,Value,Wage

Challenges

>> There are multiple player with exact same name who playing for different clubs .while join data using Name column it give duplicate data

>> Found 52 Duplicate records in 2018 dataset