Mohammad Raahim Anil

18124

HBL PSL Project

Problem Statement:

As we all know cricket fans not only want to see the live match but also, they want to enjoy all the fun statistics and analysis of players, grounds and other aspects of cricket. Furthermore, the coaches and team selectors use this analysis to opt the best playing XI against different teams in different grounds. The award ceremony also takes place using the insights from the data analysis performed. This is the reason why we will be analyzing all the aspects of cricket by the given KPIs and dimensions.

Data Wrangling

The data was comparatively cleaner than regular data, but it did have some issues that needed to fix and those are the following:

- Column names were not specified; however, the column names were given in the first row, so I replaced all the column name with the first row.
- The figures were given in the wrong data type creating confusion therefore I corrected the datatype.
- I had to split many columns to retrieve only the useful data from each column such as I was given a column which had runs as well as the wickets which were of no use to me therefore, I split the column to extract useful data only.
- There were a lot of blanks spaces and there was a need to remove them so replaced them with null.
- There were a lot of unnecessary columns that needed to remove which was done.

There were several **dimensions** such as

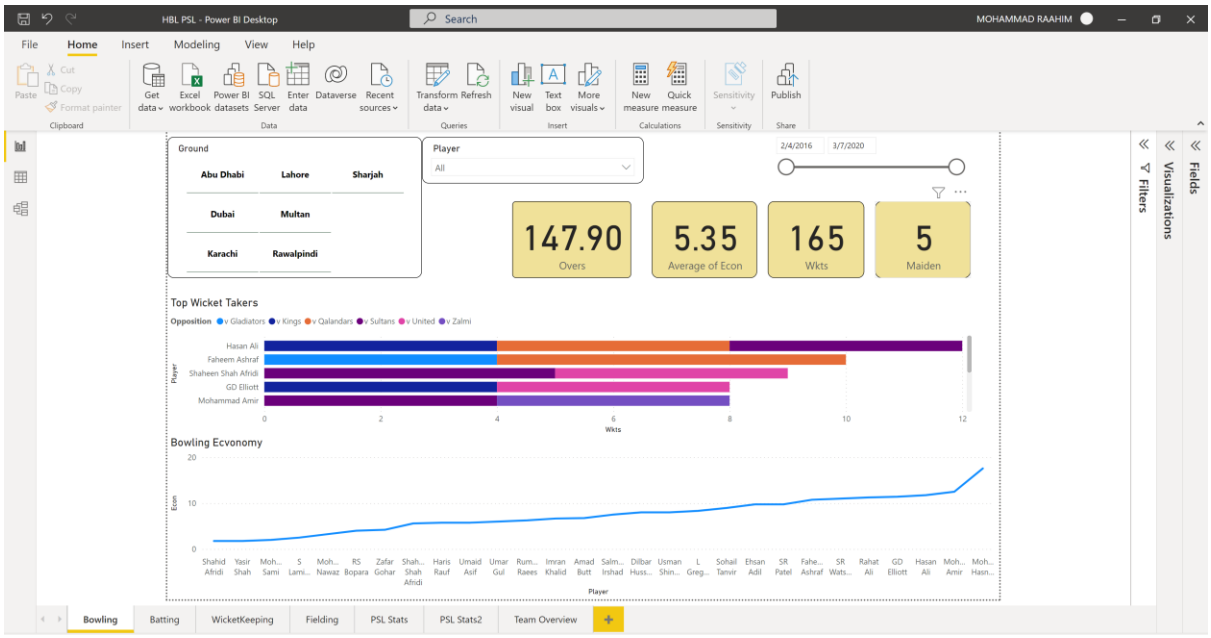
- Date
- Year
- Ground
- Teams
- Overs
- Player

The **KPIs** that were deduced were:

- Runs
- Dismissals
- Strike Rate
- Catches

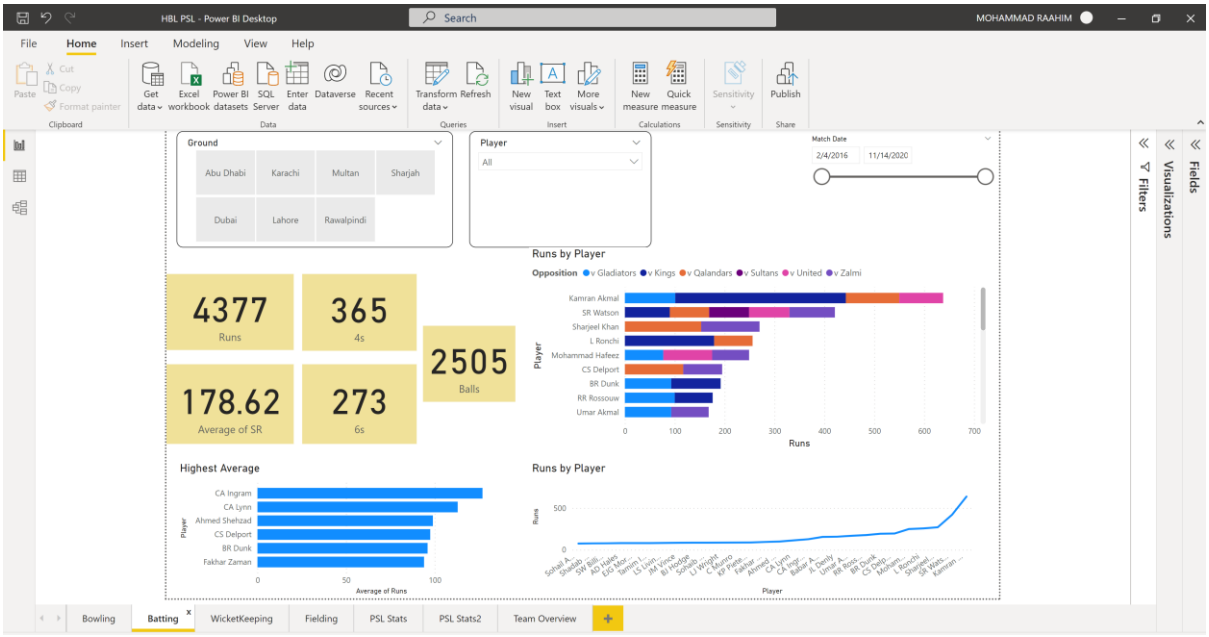
- Matches
- Average Runs

Dashboards and Answers to Queries



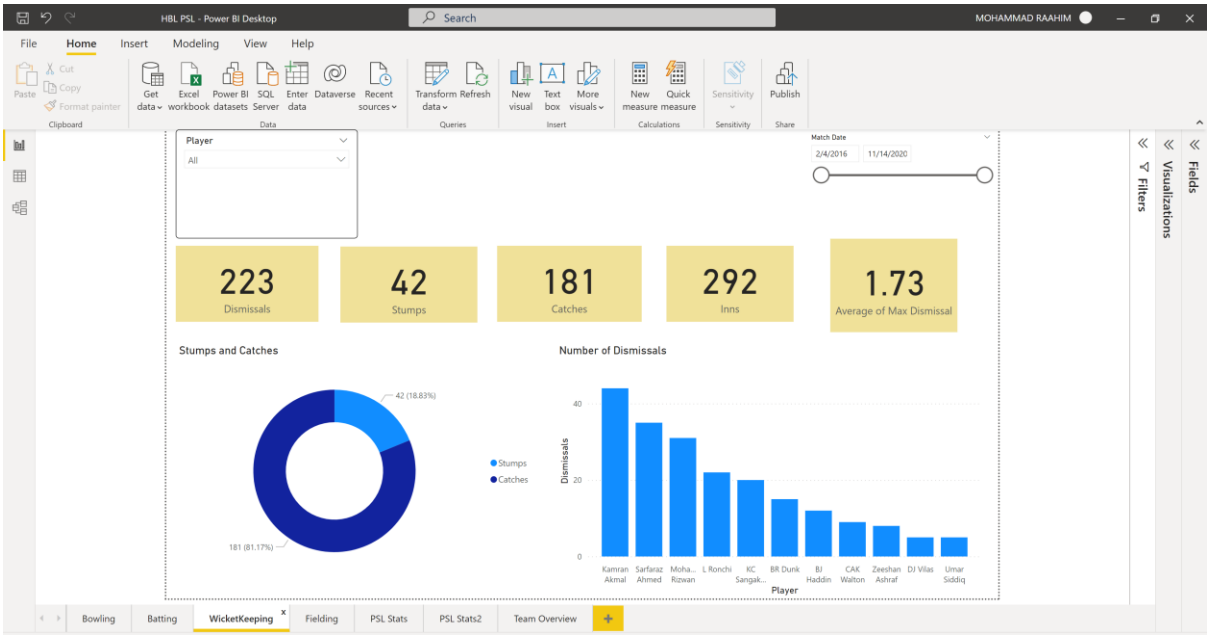
Who are the top wicket takers in the HBL PSL from 2016-2020 and at what economy against different teams in different grounds?

The above given screen shows all the data of the bowlers in the all the editions of PSL and it is also filtered by ground locations, and we can monitor and analyze each player individually on different pitches and against different opponents as shown in the stacked bar chart. The below given line graph also helps us in understanding the rankings of bowlers in terms of their economy and we can rank them accordingly. From these charts we can easily answer our query that Hasan Ali leads the top wicket taker chart followed by Faheem Ashraf. Slicers will help us cover the second part of the question which is filtering on the basis of opposition team and grounds.



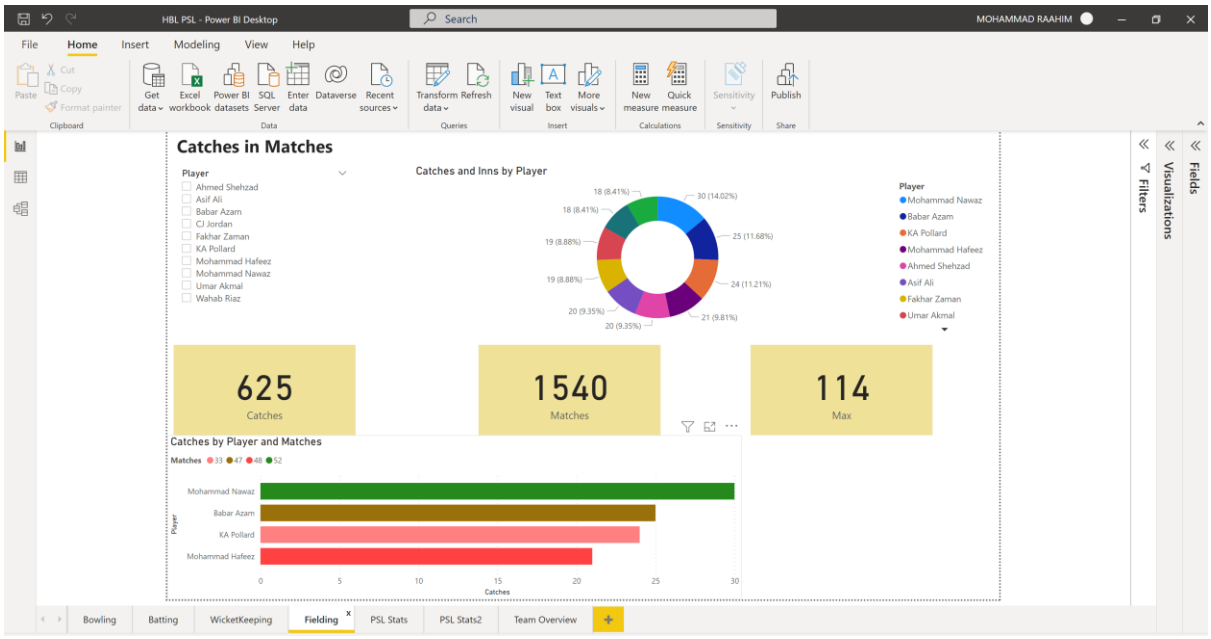
Who are the top run scorers in the HBL PSL from 2016-2020 and at what average, strike rate against different teams in different grounds?

The above given screen shows all the data of the batsmen in the all the editions of PSL and it can also be filtered by ground locations aswell as players themselves and we can monitor and analyze each player individually on different pitches and also against different opponents as shown in the stacked bar chart. The below given line graph also helps us in understanding the rankings of batsmen in terms of the number of runs they have made on various ground and we can rank them accordingly. On the bottom left is a stacked bar chart to show the top 5 batsmen of PSL in terms of highest average throughout.



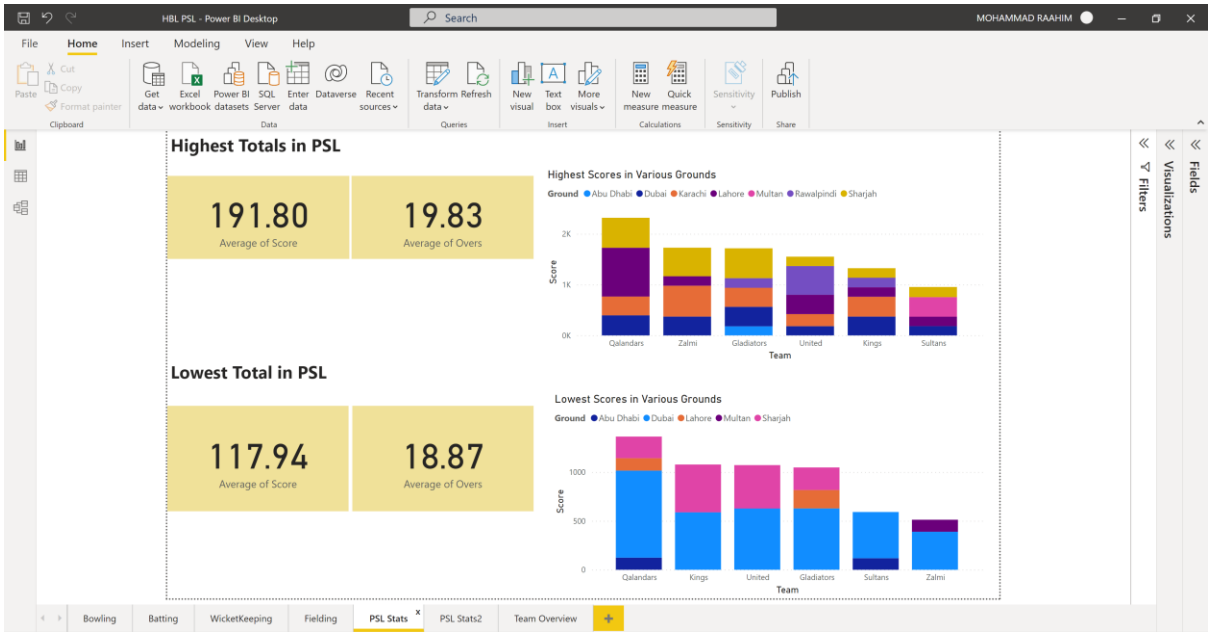
Who are the best wicketkeepers and fielders in the HBL PSL from 2016-2020 and at with what rate of stumps and catches against different teams in different grounds?

The above data is based upon dismissals that involve the wicketkeeper. The data show clear division of data into dismissals through stump and dismissals through caught behinds. This data varies from keeper to keeper and can also be analyzed through the date on which the match has been played. This data also shows the maximum number of dismissals done by each player individually as well. Overall we can compare wicket keeper's performances with each other to rank them as shown in the bar chart on the bottom right.



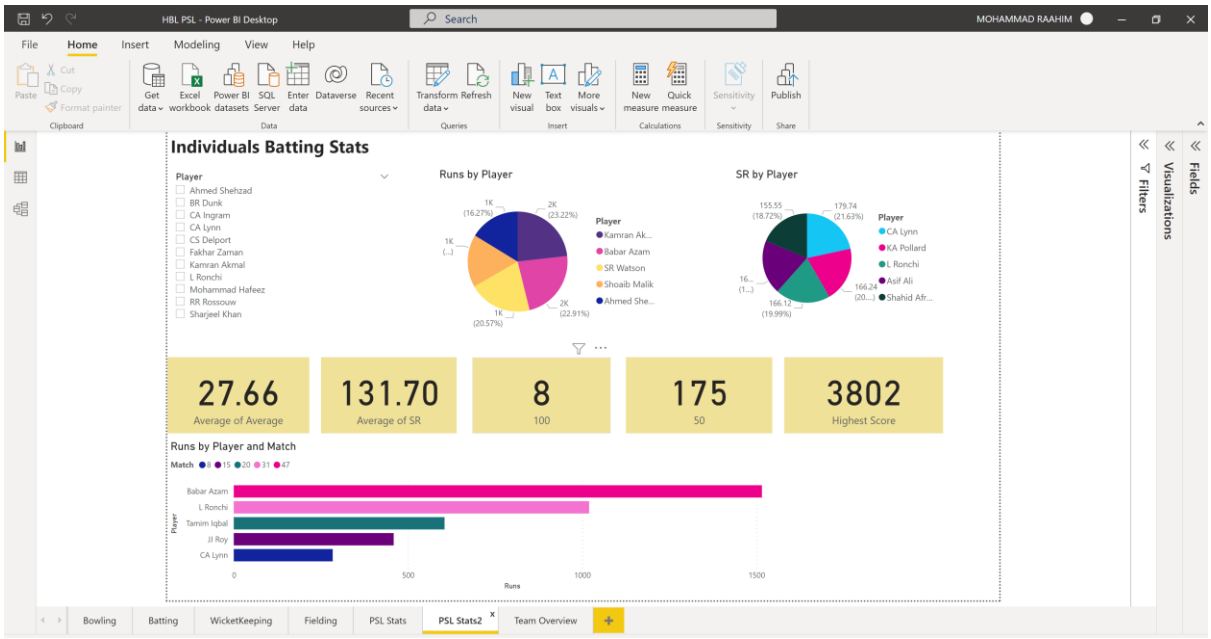
Who are the best wicketkeepers and fielders in the HBL PSL from 2016-2020 and at with what rate of stumps and catches against different teams in different grounds?

This data analyzes the fielding department of each team and individual. The above given filter helps us monitor data individually and we can monitor it collectively aswell. The datacards dsiplay the data accordingly and the donut chart shows top 9 fielders in the whole PSL tenure. When hover across each player's data we can also analyze the number of matches played by each individual to achieve these statistics. The below stacked bar chart shows the top 4 fielders in terms of most catches and their number of matches played are also specified.



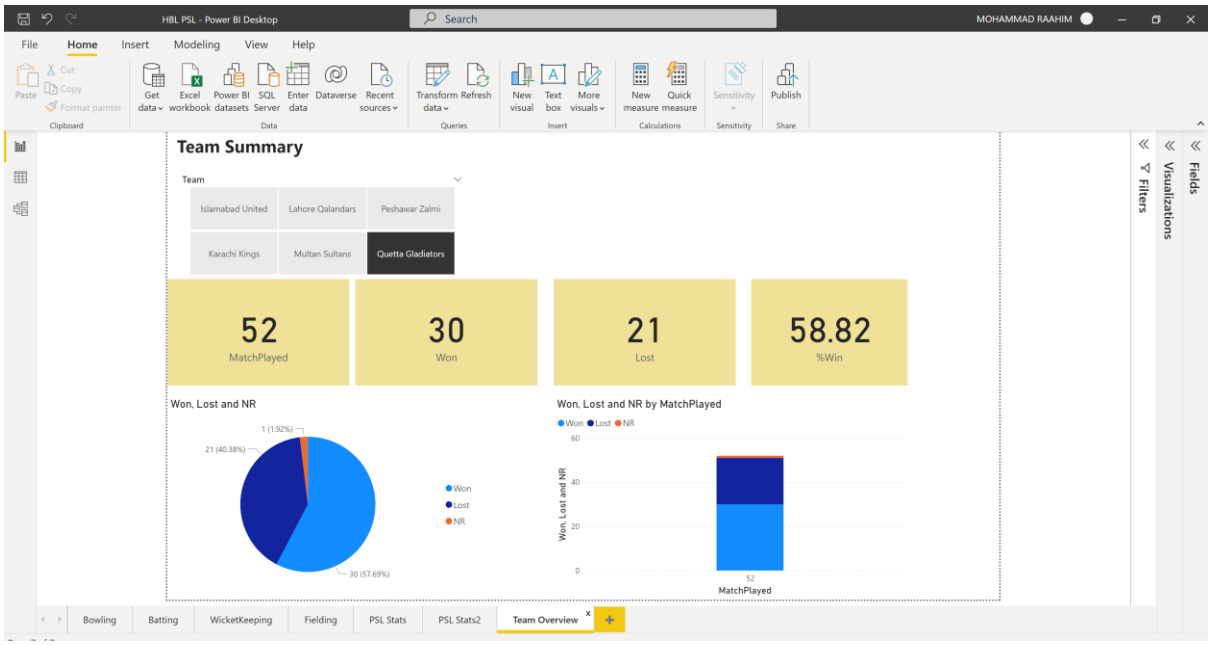
Which grounds have the highest average total and which ground have the lowest average totals by any team?

This is an overview of totals in PSL. This above stacked column chart shows which teams have scored a high total in which ground and it clearly shows that every team has scored high totals in Sharjah ground. This helps us analyze pitch conditions in several cricket grounds and how the teams are expected to perform. The similar thing happens in the below column chart which shows the lowest totals in PSL.



The individual battimg stats show various data and it also shows the rankings of top performers as a batsman and using the Top N filter we managed to select the highest scorers of PSL . The other pie chart also shows batsmen who might not be top scorers but have a very high strike rate. This shows different batsmen with different style of playing. The data cards show a lot of data on each

individual batter. The Data in the below bar chart shows runs made by players taking in consideration number of matches played aswell.



Which team has what win/loss % overall throughout the the PSL journey of 4 years?

The final report displays overall team performance in PSL. We can clearly see and tell which team has played how many matches and what were their results. The pie chart and column chart can clearly show the distribution of win, lost and No result of each team individually. The percentage win can easily help us compare between teams over a long period of time.

Story

According to the analysis, we can deduce that grounds and opposition teams have a very large impact on the performances of the teams and individual players. We can also deduce that not all the grounds have the same pitch conditions therefore in some ground venues, bowlers perform well and, in some grounds, batsmen are performing well. Using such analysis help us predict results of the match and also the stakeholders of each team franchise would try to make the Best Playing XI after looking at all the conditions and matching it from the analysis already performed from the past records. These analyses also help in identifying and finalizing awards in the award ceremony at the end of the tournament. The individuals that are performing best according to these analysis would be wanted by each and every team in the next season therefore the teams can plan which players to buy using the analysis performed above. Finally, we can clearly see so many benefits of the insights generated and how these can be useful for the teams.