**Introduction to Data Analysis using Tableau PROJECT REPORT**

**Project Report on 2022 T20 WC Batting Card**

Submitted by **Varkuti Sathwik Reddy**

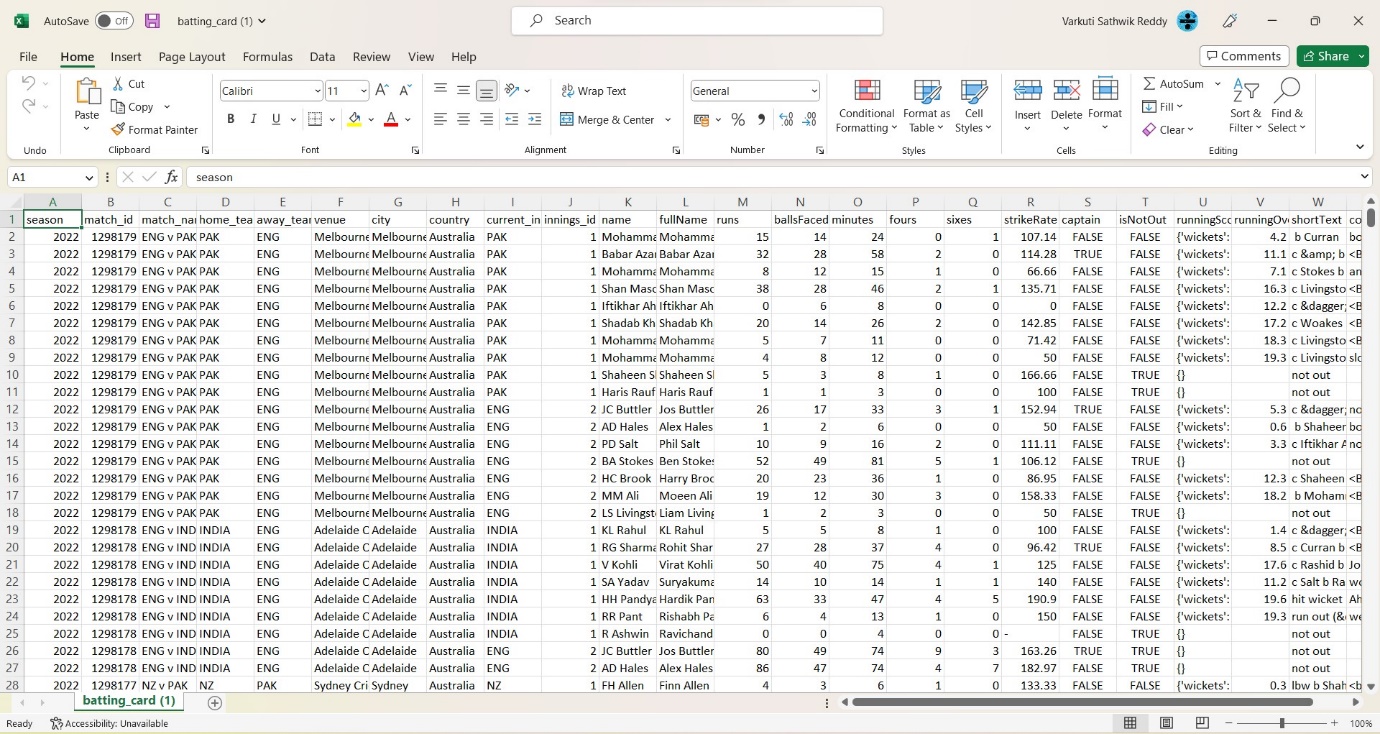
**Discipline of CSE – DS**

Guru Nanak Institutions of Technical Campus , Hyderabad -501506

INTRODUCTION :

This data set contains ICC Men’s T20 WC 2022 that runs from mid OCT to mid NOV in Australia .

We are using dataset for analyzing using Tableau visualization tool .



# This Dataset consists of following column’s :

* Match\_name : It describes the match between two teams.
* Venue : It tells about the where the match is happening.
* City : The place of the city where the match is taking place.
* Country : The Name of Country where the matches are taking place.
* Current\_Innings : The team which is batting currently.
* Name : It shows the Name of the player.
* Full\_name : It shows the full name of the player.
* Runs : Total number of runs hit by the batsman.
* Balls faced : The number of balls faced by the batsman.
* Fours : Number of fours hit by the batsman.
* Sixes : Number of sixes hit by the batsman.
* Strike Rate : the number of runs of a batter divided by the number of balls faced.
* Balls faced : The number of balls faced by the batsman.
* Fours : Number of fours hit by the batsman.
* Sixes : Number of sixes hit by the batsman.
* Strike Rate : the number of runs of a batter divided by the number of balls faced.

**OBJECTIVE :**

The main aim of this project is to visualize batting performance which consists of player’s name , balls faced , strike rate using Tableau tool.

After analysis of the dataset, the aim of this project is to give answer of given objectives in easy way:

* Player’s and number of runs scored in the tournament.
* Player’s having more than 200 runs in the tournament.
* Player’s having more than 200 runs with the strike rate above 120.
* Match name and total number of boundaries takes place in match.

* Player’s having more than 100 runs at a single venue.
* Team and their boundary count.

**SOURCE OF DATASET :**

[ICC Men's T20 World Cup 2022 | Kaggle](https://www.kaggle.com/datasets/rajsengo/icc-mens-t20-world-cup?select=batting_card.csv)

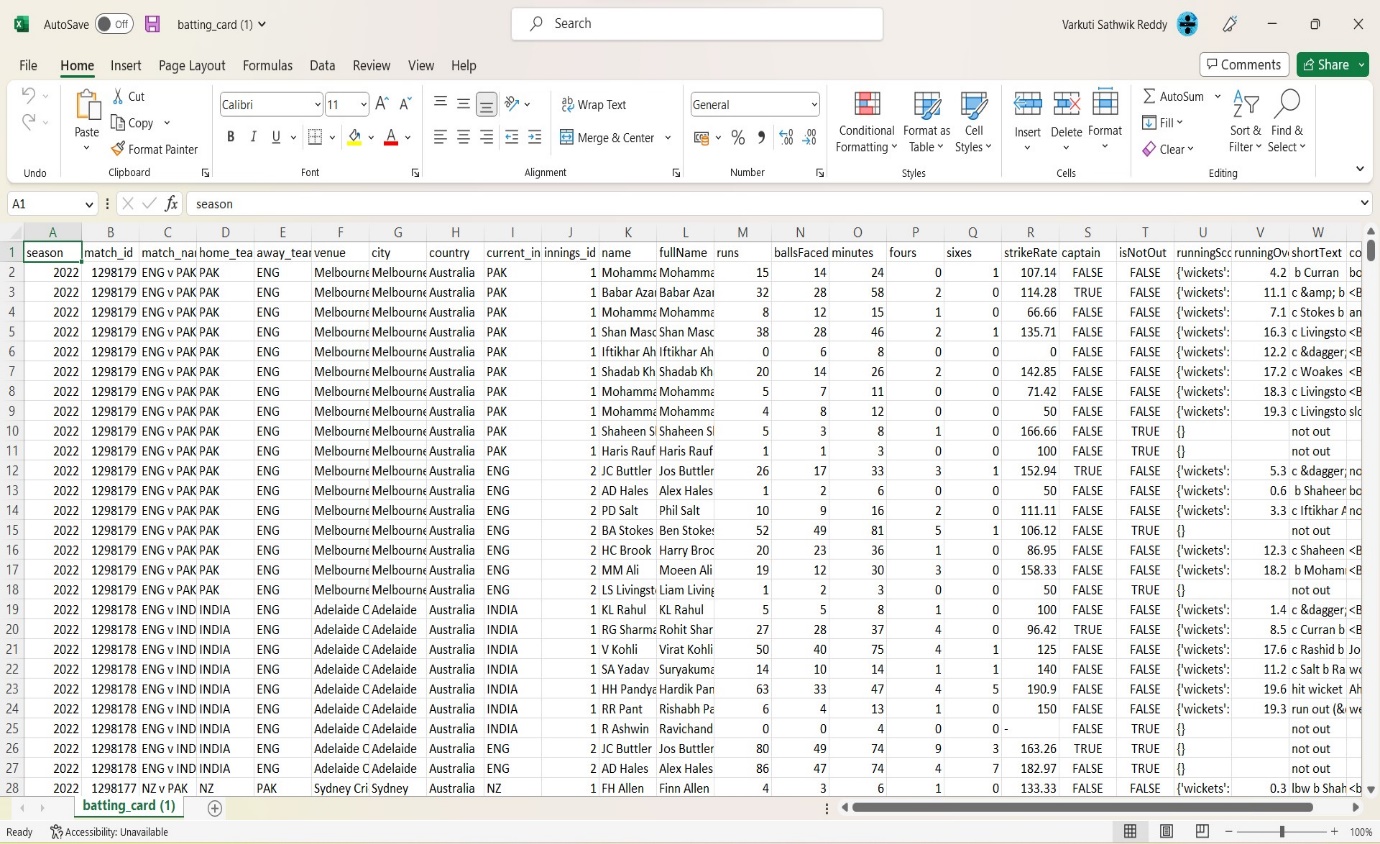
**KAGGLE :**

Kaggle is an online community for data scientists and machine learners, developed by Google. Kaggle allows users to find and publish data sets, explore, and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges. Kaggle got its start by offering machine learning competitions and now also offers a public data platform, a cloud - based workbench for data science, and short form Al education. On 8 March 2017, Google announced that they were acquiring Kaggle.

This data science project analyzes the Batting cards of 2022 T20 WC dataset. It was created for the (B. Tech CSE 6th semester Introduction to Data Science course) project.

**Sample of Dataset with Data fields :**

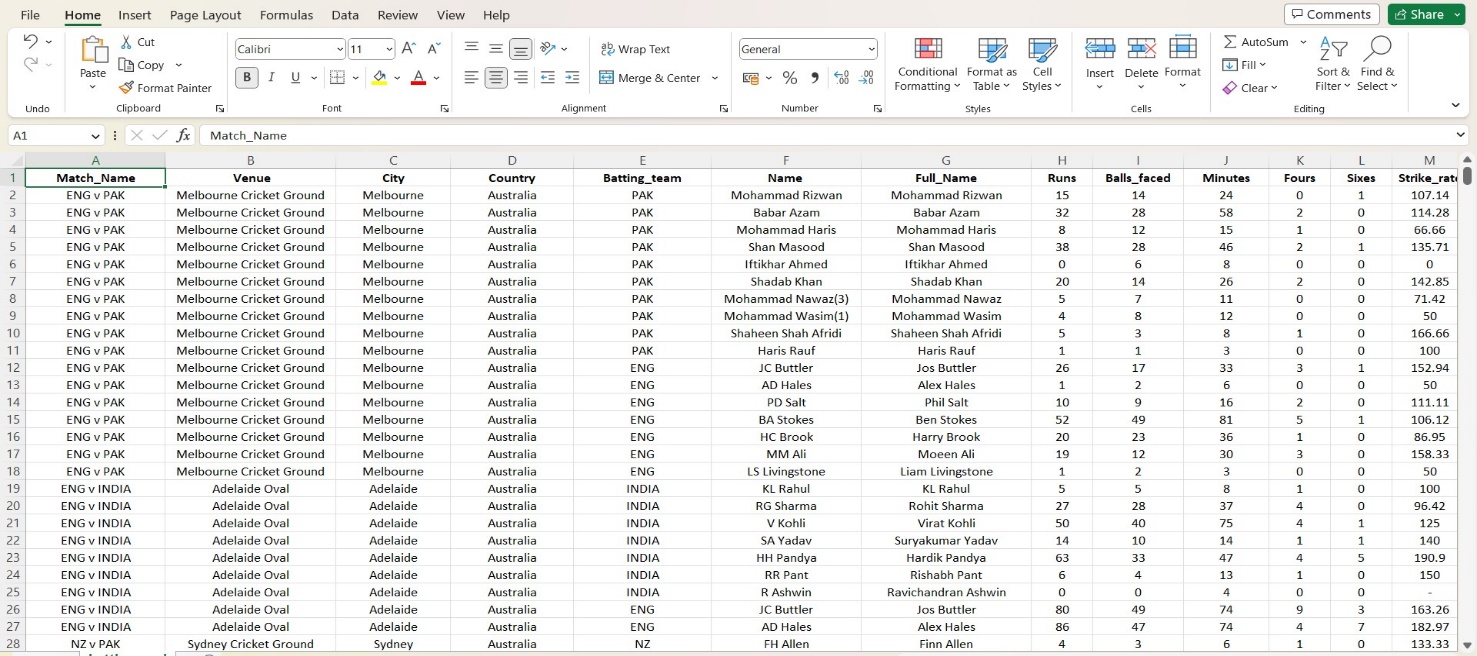
The dataset contains the following columns like Match name , Venue , City , Player’s name , Runs and so on.

****

**DATA PRE-PROCESSING :**

Data pre-processing is an important step in the data mining process. It refers to the cleaning, transforming, and integrating of data in order to make it ready for analysis. The goal of data pre-processing is to improve the quality of the data and to make it more suitable for the specific data mining task.

After removing the unnecessary columns from the dataset which are not beneficial for Visualization.

**Analysis of Dataset :**

* Player’s and their run’s:

* Introduction

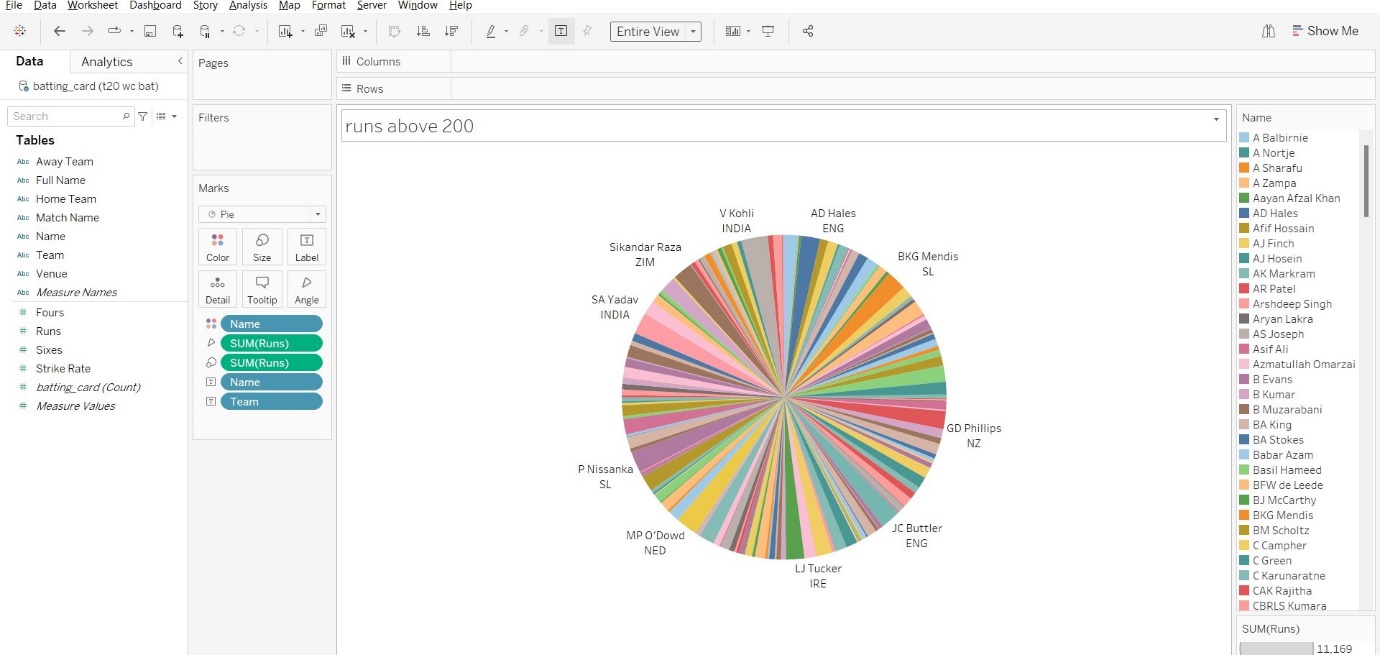
By performing this analysis, we will get number of Sales with respect to Product type

* Analysis Results

We are using Pie chart for visualizing the runs hit by the player’s in the 2022 T20WC.

We are visualizing player’s and runs using Pie Chart.

We labeled the Player’s name and Team using Labels in the Marks.

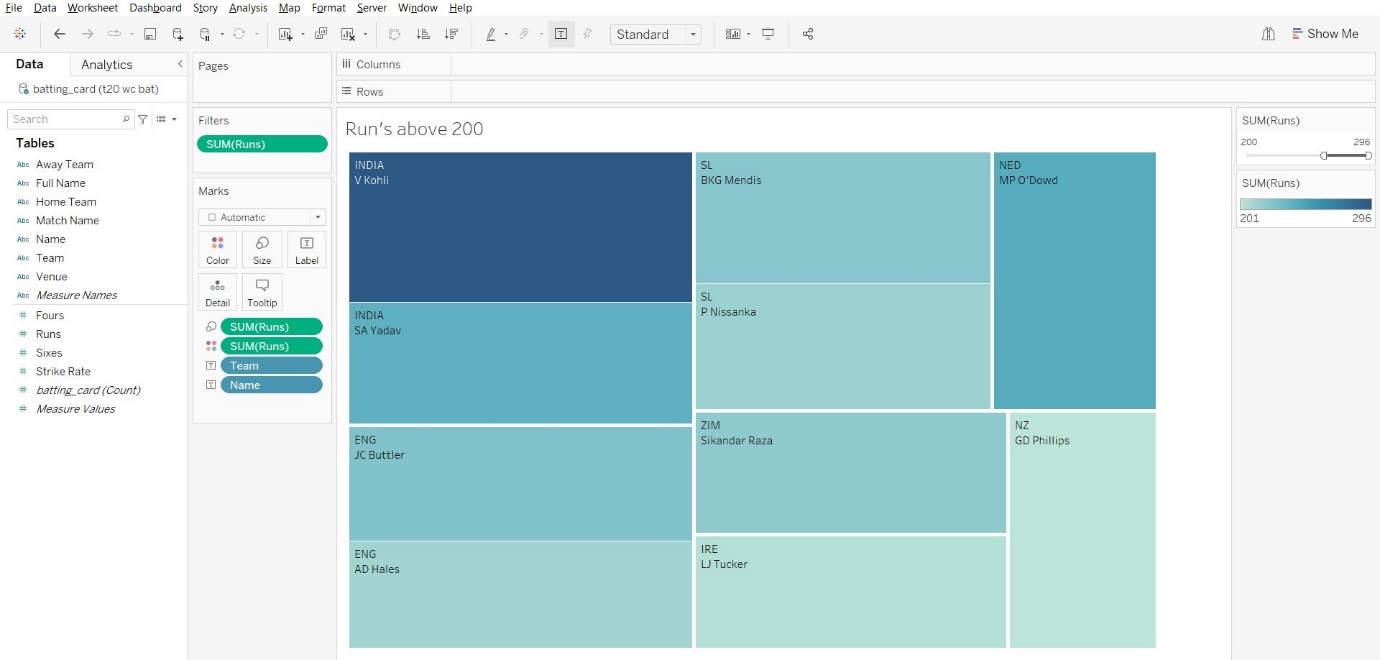


[player's runs.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0pnzMbuFCQOAuwyw?e=Tamb9S)

* Player’s having more than 200 Runs :

From the above Pie chart we are filtering Player’s who are having runs above 200 using Treemap Visualization.

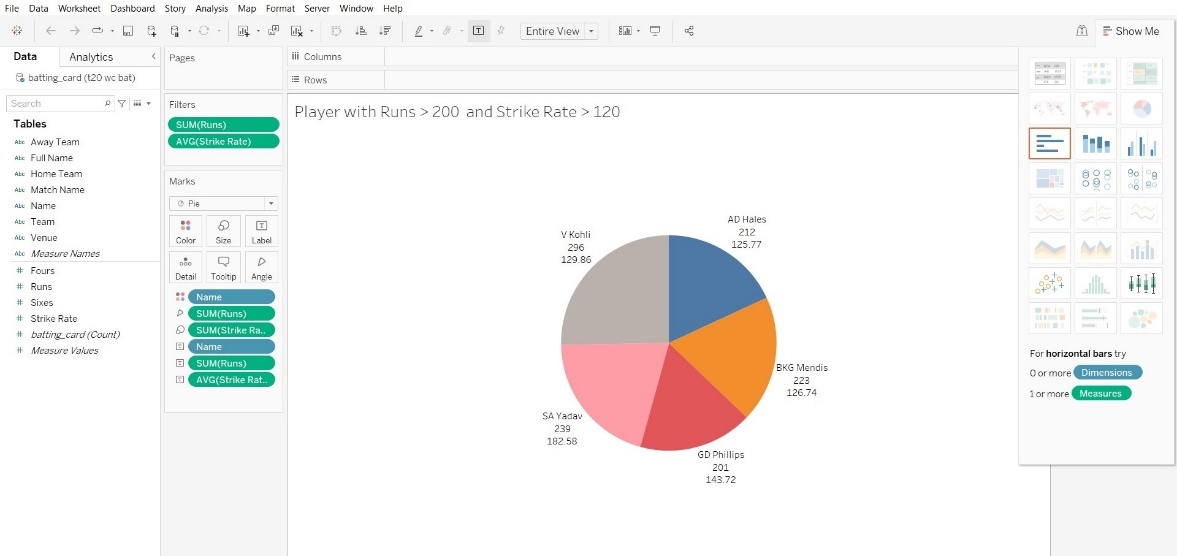
And we labeled player’s name and team using Marks.

 [Runs above 200.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0XkjYaKWMGqJyaPw?e=UQdH57)

* Player’s having Runs more than 200 and Strike Rate more than 120.

From the above Two sheets now are again filtering player’s who are having strike rate above 120.0

Now it shows Player’s Name having runs above 200 and Strike rate above 120.0 using Pie Chart



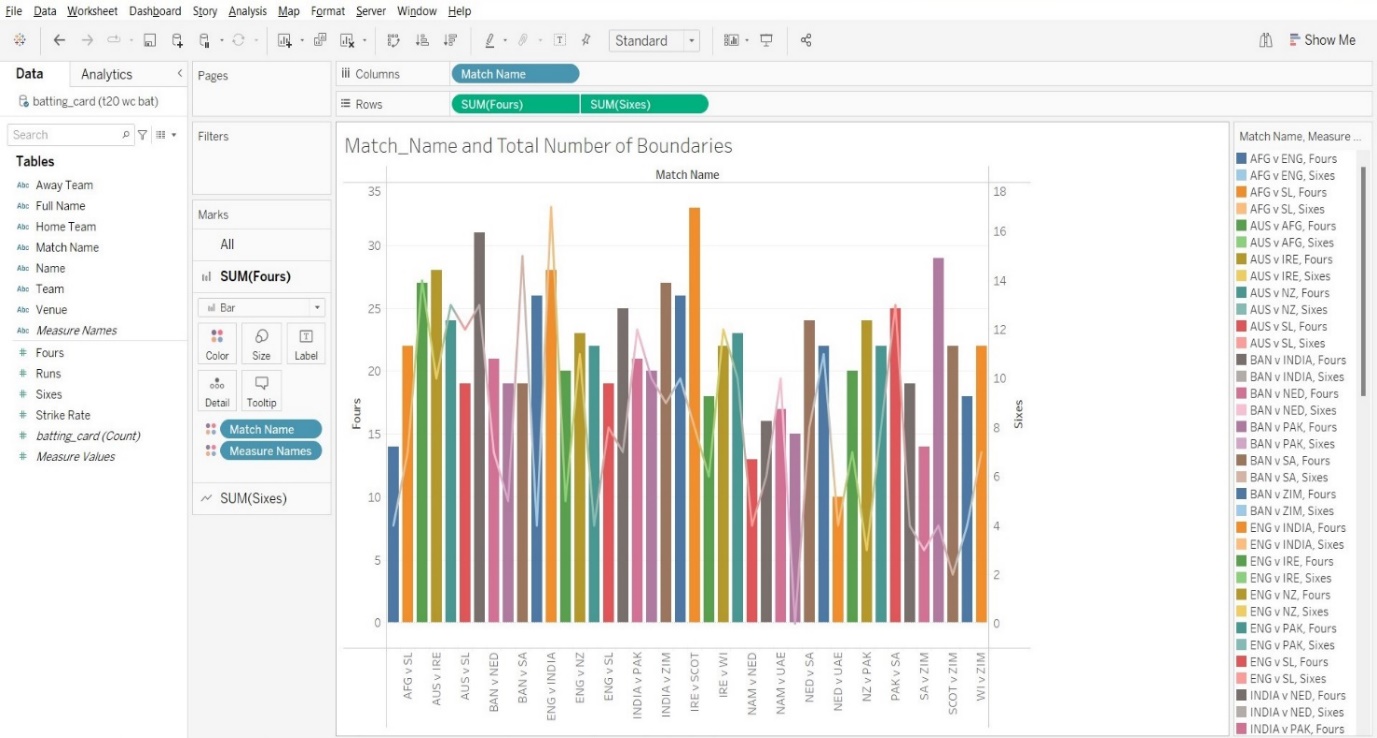
[player\_strike\_rate\_runs.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0T9WWQuSpd0-9PYg?e=Wek8Ka)

* Match name and total number of boundaries in the match

In this Sheet we are visualizing no.of boundaries (four’s and sixes) occurred in a match.

In the columns we are displaying match\_name and in the rows we are arranged fours and sixes.

In the visualization we are showing fours using bar chart and sixes using line chart.



[match\_name vs boundaries.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0Sd4onjeFZbCcxKA?e=9NfKmK)

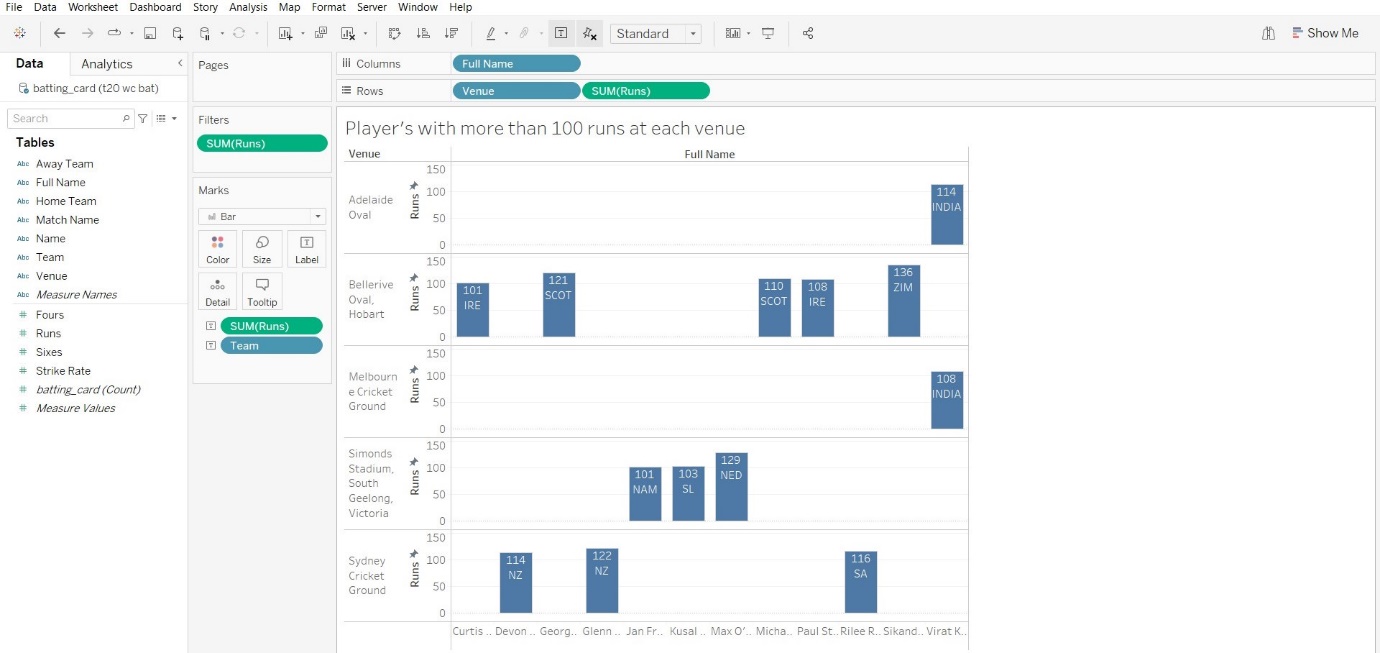
* Player's with more than 100 runs at each venue :

In these sheet we are showing Player’s having More than 100 runs at a venue .

In the rows we are arranged venue and runs.

In the columns we placed Name of the player.

We are visualizing using stacked bars tool in the Tableau.



[Player's with more than 100 runs at each venue.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0fTR8hz-rvh7P7dw?e=zjOfob)

* Team and their Boundary count :

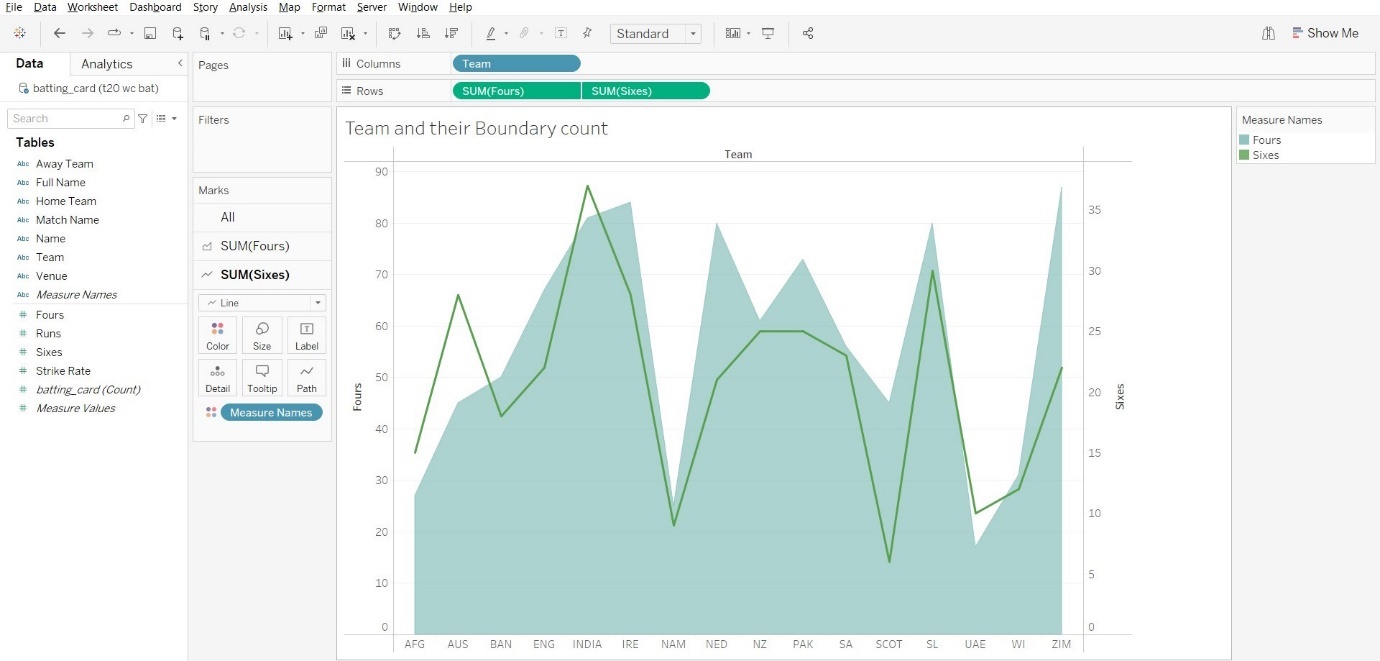
In the sheet we are visualizing team and their boundary count in the entire tournament .

In the rows we arranged fours and sixes .

In the columns we placed Team.

In the sheet we showing fours in the form of Area sheet and sixes in the line chart.

[Team and their Boundary count.tbm](https://1drv.ms/u/s!Aj4KRQDZtjQpiP0ddnq2PkSAkz_ueg?e=bYNalo)



**DASHBOARD :**

