IR PROJECT REPORT

NAME: G SAIMANI

ROLL.NO: S20180010057

INTRODUCTION:

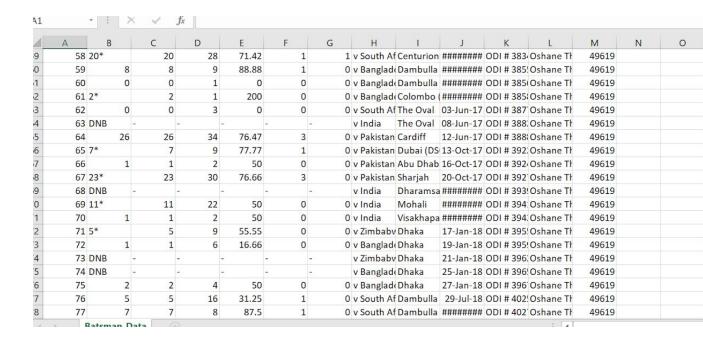
This project is mainly about searching about the Cricket Player and Stats about them. For implementing this, Dataset about them is directly taken from the web. Based on the query given by the user the Related documents were retrieved using a ranked retrieval algorithm and ranked retrieval algorithm is Tf-idf.

TASKS:

- 1) Search engine for the info and it gives us the details of the Particular Player in the way user wants
- 2) If a user wants to compare an individual player stats with other team then it retrieves all the related documents

TASK-1:

1) Dataset for this retrieval system has 15000 documents in the csy file.



2) Data was extracted from the given dataset (Batsman_Data.csv). The Data contains the total info of the players. It gives the player stats when the user searched for a player and their records. And I have Stored the inverted index using Trie data structure

Because of faster retrieval of the docs corresponding to terms And I have created max-tf for all documents.

TASK 2:

1) If a user wants to compare an individual player stats with another team then it retrieves all the related documents.

```
The highest number of score for this batsman is:
127.0

The highest Strike Rate for this batsman is:
380.0

The most number of 4s for this batsman is:
14.0

The most number of 6s score for this batsman is:
7.0

The highest number of score against this team is:
189.0

The highest Strike Rate against this team is:
400.0

The most number of 4s against this team is:
19.0

The most number of 6s against this team is:
14.0

PS C:\Users\saimani\Desktop\IR-end>
```

2) Users can compare any player's runs ,strike rate,sixes,fours,highest score with another player or with another team.