

DS ASSIGNMENT-2



**SUBMITTED TO:
PROF. MANDAR KARYAKARTE**

**VISHWAKARMA INSTITUTE OF INFORMATION
TECHNOLOGY, PUNE**

COMPUTER ENGINEERING DEPARTMENT

**BY:
NAME: ABHISHEK MORE**

G.R No.: 21810033

ROLL NO.: 323036

CLASS: T.Y COMP

BATCH: COMP-C2

Description:

This Assignment includes understanding of Statistics, Data Visualization (using histograms plotting), Data Integration, Data Cleaning and Data Transformation. Statistics -> Mean, mode, median, variance, standard deviation, range, percentiles, maximum and minimum value.

Interpretation:

I performed analysis of the data to understand the relation between the features(columns) of the dataset. Also I analysed how one feature affects the other. I did analysis on the toss winning team and match winning team(combined with data integration) datasets and interpreted how the factors like toss winning team, match winning team, win by runs, win by wickets etc. are related to the match winners.

I reached to the following analysis:

- Every team playing the tournament has won at least 5 matches
- Maximum number of wins are for SRH and MI with 19 wins
- Most number of matches were played in Bangalore with 17 matches and least in Raipur with only 2 matches
- Most matches were won with a margin of 1,7 and 15 runs
- Most matches were with a margin of 6 to 7 wickets
- Mumbai Indians were most successful with their toss decisions as it was the team winning most of the matches after winning the tosses.
- Average winning runs margin was 12.3 runs
- Average winning wickets margin was 3.7 wickets

Outcomes:

Most of the toss winning teams decided to field first and won the matches.