Unveiling Insights through Data-driven Exploration of Cricket World Cup 2023

Team Members - Anshuman Mohanty (USC ID - 4257570790), Kaustubh Sharma (USC ID - 1765749035)

Problem Statement

With the Cricket World cup 2023 in full swing, there is a wealth of untapped data which holds the key to unlocking strategic insights. Our project aims to uncover hidden patterns, correlations of individual and team performances from the tournament data, venue trends etc. thereby enhancing our understanding of the dynamics of the tournament. The objective of our project is to empower cricket stakeholders with data-driven intelligence for better decision-making, ultimately contributing to the evolution of cricket strategies.

Data Sources

Cricbuzz - https://www.cricbuzz.com/cricket-series/6732/icc-cricket-world-cup-2023
ICC - https://www.cricketworldcup.com/tournament-stats
Espn -https://www.espncricinfo.com/series/icc-cricket-world-cup-2023-24-1367856/stats

Analysis and Visualizations

- **Web Scraping** - utilize BeautifulSoup and requests module to scrape data on match results, player statistics, team compositions, venue details etc. from official websites.

- **Data cleaning and preprocessing** - visualizing missing values using MissingNo library and handling missing values if any.

- **Data Transformation and outlier analysis -** convert data types and perform outlier analysis using boxplots.

- **Statistical analysis -** perform statistical analysis of the dataset to summarize different features.

- **Visualizations -** use Seaborn, Matplotlib and Plotly to create bar charts, pie charts, histograms and scatter plots to visualize trends like venue specific patterns, player statistics, team strategies etc.