

# Data Visualization Project Proposal

- **Basic Info.**
  - Project Title : Soccer Stats
  - Names : Sravan Kumar Neerati, Sreekanth Reddy Konda
  - Emails : [u0931101@utah.edu](mailto:u0931101@utah.edu) , [u1140489@utah.edu](mailto:u1140489@utah.edu)
  - UID : 0931101 , u1140489
  - Project Repository Link : <https://github.com/Meerkat3/SoccerStats>
  - **Background and Motivation:** Soccer Data set encapsulates good amount of information around the Player attributes , Matches and League statistics and Betting results around the games
  - **Project Objectives.** We aim to show the highly sought after mainstream Soccer stats in overview and details of the Player attributes and their performance over time , Insights of a club's performance , Transfer of players between clubs/leagues and how the betting fared on the matches(tentative)
  - **Data.** The data is presented in SQLite format and we intend to do some pre-processing before storing them as csv/json for our final usage. If time permits we intend to have a backend server, serving the data.
  - Link: <https://www.kaggle.com/hugomathien/soccer/data>
  - **Data Processing.** There is not much data cleanup but a good amount of data extraction that needs to be done from the SQLite data. We intend to use python code to do the pre-processing and store the data in csv.
  - **Visualization Design.** How will you display your data? Provide some general ideas that you have for the visualization design. Develop **three alternative prototype designs for your visualization**. Create **one final design that incorporates the best of your three designs**. Describe your designs and justify your choices of visual encodings. We recommend you use the [Five Design Sheet Methodology](#).
  - **Must-Have Features.**
- 1) Player :
- Attributes Summary :
    - Individual Player attributes(Top 50 players)
  - Player performance over the years
    - Evolution of the Best Players – change in overall rating over the years
    - Comparison among the players(Overall Rating or specific attributes) for a particular year.

## 2) League :

- Clubs' performance over years per league (attribute Overall goals scored)
- Share of influence by players(top 20) on betting, per club
- **Optional Features.**
- Betting patterns :  
investigate a few simple betting approaches with applying different financial strategies
  - Flat - fixed stake
  - Expected Profit
- Match stats : The incidents over a given match (Shots taken, corners etc ..)
- **Project Schedule.**
- Week1:
  - Individual Player Attributes
  - Evolution of Best players
- Week2:
  - Comparison of Players' performance in a year
  - Implement interaction
- Week3:
  - Performance of club over years in a league
  - Share of influence by player in betting.
- Week4:
  - Work on Optional Features and Enhancements over the Existing visualization.

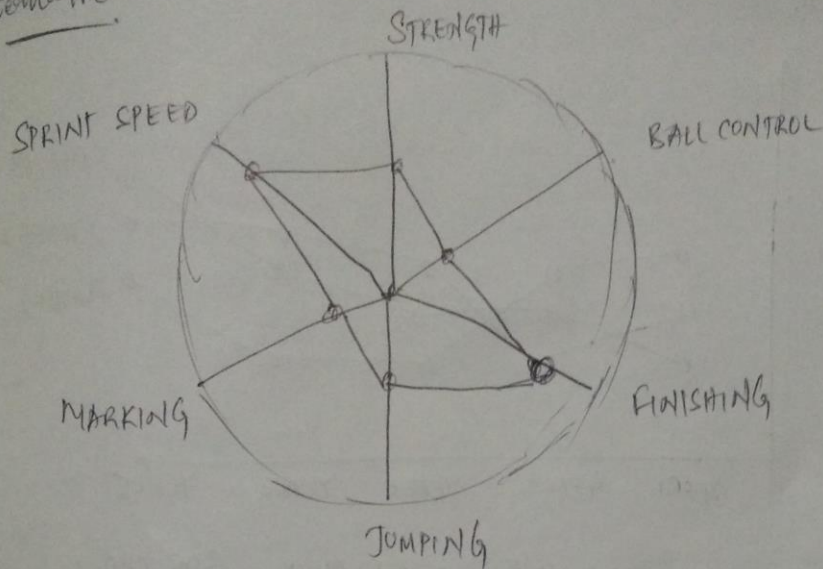
Comparison between alternative designs shown in pictures below:

- Individual Player attributes
  - Alternate 1 : Radar Chart :
    - Pros:
      - Summarizes all attributes and the main stand out attribute can be seen easily.
      - Juxtaposition of the attributes is straightforward.
  - Alternate 2 : Bar Chart :
    - Cons:
      - The Attribute progression on the x axis, unlike time is not linear.
      - Two different attributes can be important to compare but their distance on the x axis can make it difficult to compare
  - Alternate 3 : Pie Chart :
    - Cons:
      - The channel “Area” is often not the best choice, when it comes to comparison.
- Evolution of best Players
  - Alternate 1 : Line Chart :
    - Pros:
      - Time progresses linearly and the evolution of a player can be seen easily as ups and downs in the line chart.
      - It tells a story of how a player evolved over time.
  - Alternate 2 : Bar Chart : On click over year timeline
    - Cons:
      - Zooms in on a particular year and misses the bigger picture of the trend over years
  - Alternate 3 : combination of bar charts
    - Cons:
      - Too many bar charts , cluttering the visualization.
- Comparison of players per year
  - Alternate 1 : Bar Chart :
    - Pros:
      - Length is a good choice of channel when it comes to comparison
      - The standout players can be identified easily
  - Alternate 2 : Pie Chart :
    - Cons:
      - Comparison of channel, area is not easy
  - Alternate 3 : Radial lines :
    - Cons:
      - Comparison and standout identification of radial lines is difficult.
- Clubs performance over the years per league
  - Alternate 1 : Line Chart :
    - Pros:
      - Trend over time can be understood easily

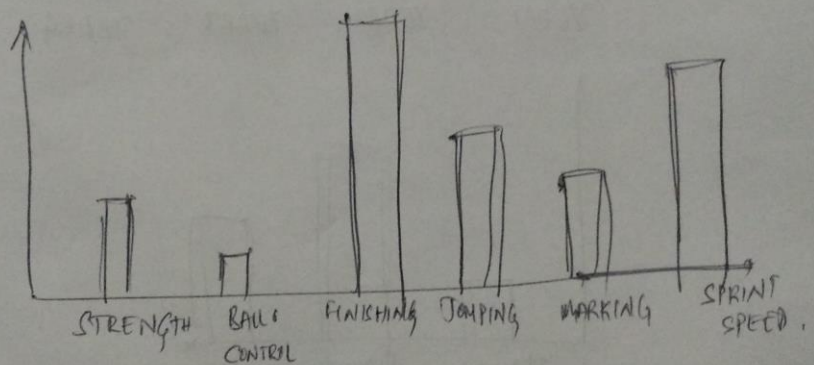
- Alternate 2 : Bar Chart :
  - Cons:
    - Connectivity amongst the bars is not evident as it is in Line Chart.
- Alternate 3 : Bubble Chart :
  - Cons:
    - Area / size of circles is not a good choice for comparison.
- Share of influence by player on betting
  - Alternate 1 : Bar Chart :
    - Pros:
      - Length is a good choice of channel when it comes to comparison
      - The standout players in influence can be identified easily
  - Alternate 2 : Stacked Bar Chart :
    - Cons:
      - Although the summation of all bar lengths comes to 1, individual standouts are difficult to identify.
  - Alternate 3 : Radar Chart :
    - Cons:
      - Difficult to compare and identify individual share.

# Individual Player Attributes:

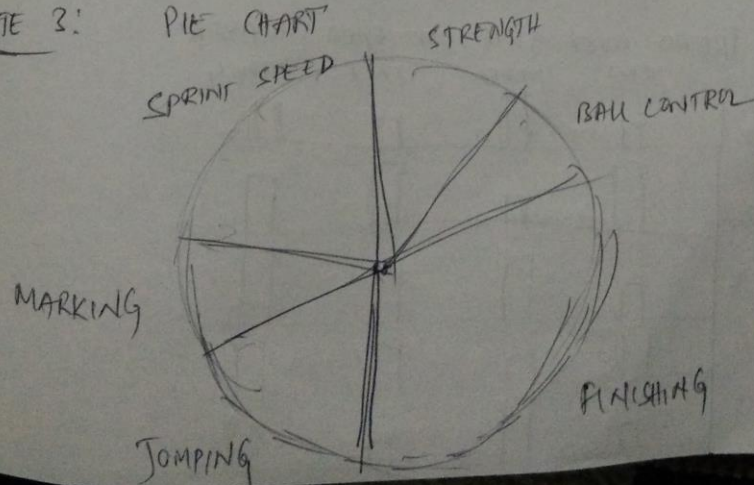
## Alternative 1: RADAR CHART



## ALTERNATE 2: BAR CHART



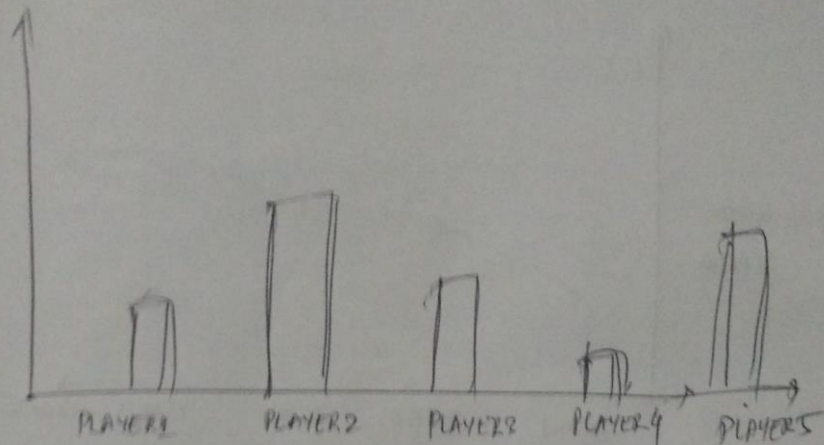
## ALTERNATE 3: PIE CHART



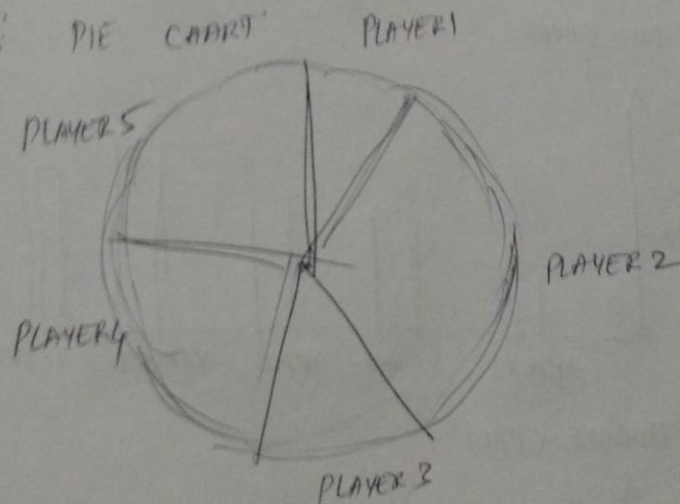
# COMPARISON OF PLAYERS PER YEAR (INTERACTIVE)

ON SELECTION OF YEAR.

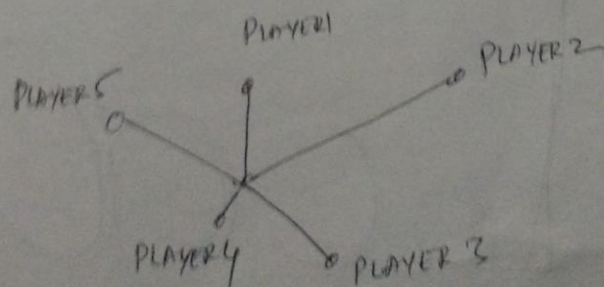
ALTERNATE 1: BAR CHART



ALTERNATE 2: PIE CHART



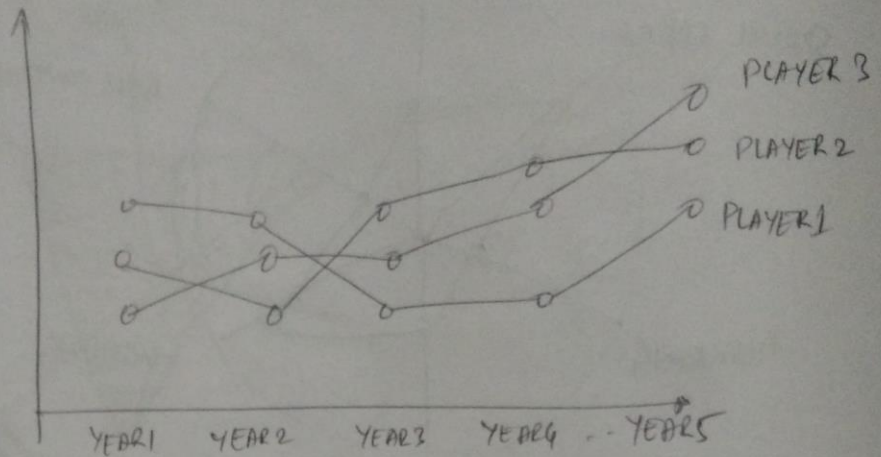
ALTERNATE 3: RADIAL LINES



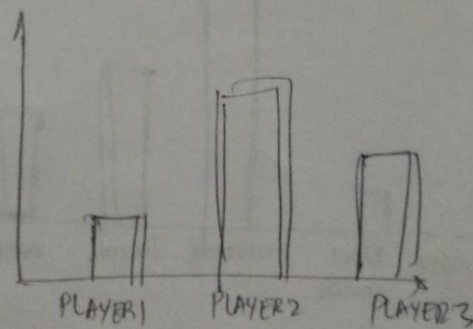
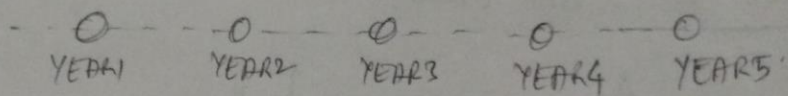


EVOLUTION OF BEST PLAYERS!

ALTERNATIVE 1: LINE CHART

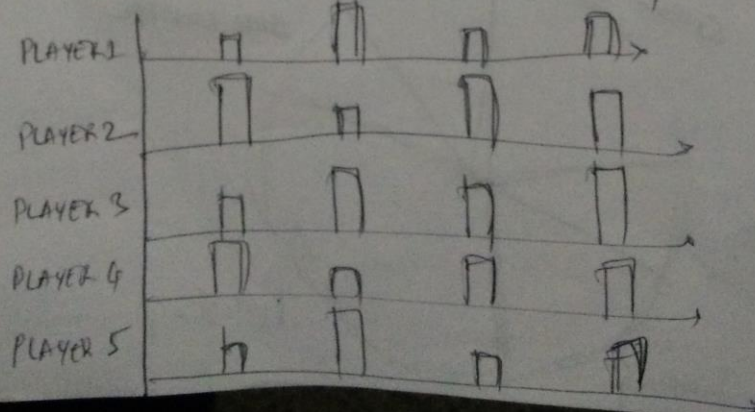


ALTERNATIVE 2: YEAR TIMELINE + ON CLICK BAR CHART



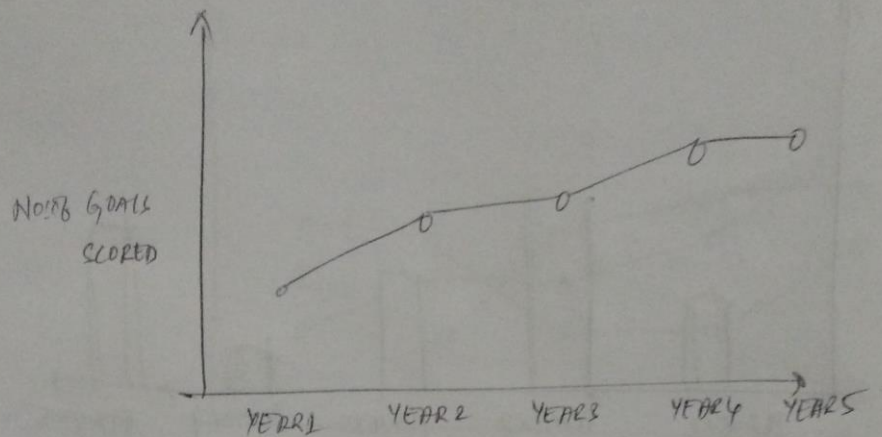
ALTERNATIVE 3:

TREND OVER TIME FOR EACH PLAYER!

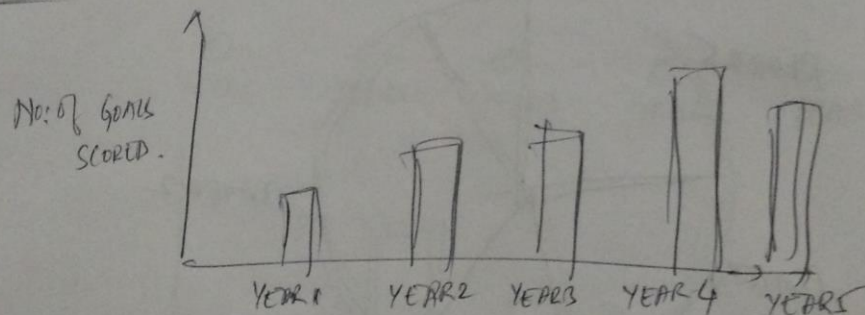


# CLUBS' PERFORMANCE OVER YEARS PER LEAGUE!

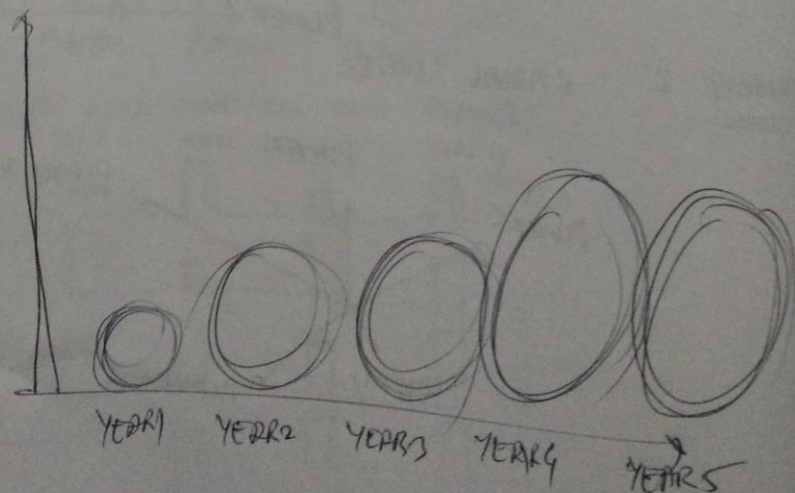
## ALTERNATE 1: LINE CHART!



## ALTERNATE 2: BAR CHART



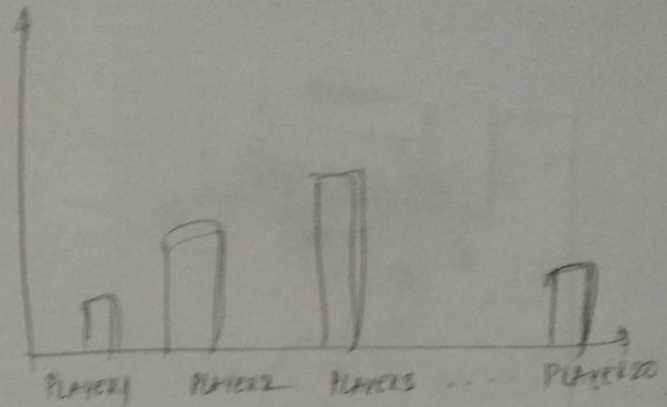
## ALTERNATE 3: BUBBLE CHART



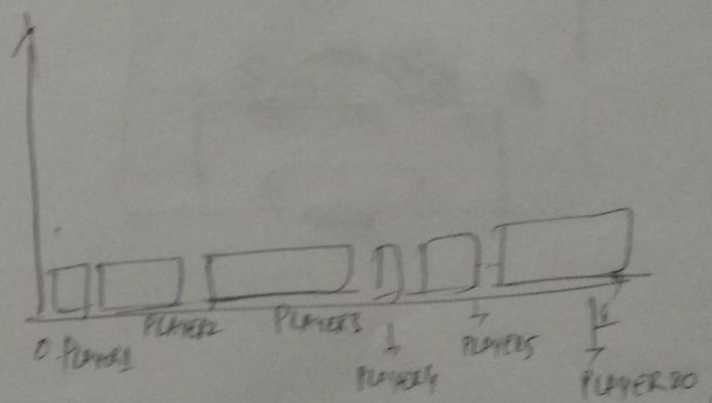


# SPARE OF INFLUENCE BY PLAYER ON GETTING

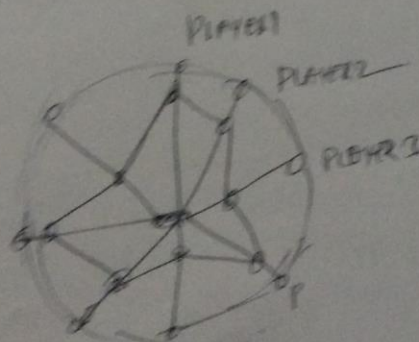
ALTERNATE 1: BAR CHART



ALTERNATE 2: STACKED BAR CHART



ALTERNATE 3: RADAR CHART



MATCH STATS : OPTIONAL :

INCIDENT DATA: EX: SHOTS TAKEN.

