Soccer Team Management

Group 3
Jeemisaba Chudasama (jac786)
Ankitha Gavinolla(ag1527)
Ajinkya patankar(aap256)
Instructor: Prof. James Abello

Index:

Dataset

- The data is from the website https://sofifa.com by extracting the Player personal data, followed by Player IDs and their playing and style statistics.
- The dataset is taken from the Kaggle website. It contains information about 18000 players and each player has 88 attributes.
- For skill attribute, we have an integer from 0 to 100 that measures how good a player is at that attribute.
- Examples of attributes are: dribbling, aggression, vision, marking and ball control.

ID	Name	Age	Photo	Nationality	Flag	Overall	Potential	Club	Club Logo	Value	Wage	Special	Preferred F
158023	L. Messi	31	https://cdr	Argentina	https://cdr	94	94	FC Barcelo	https://cdr	€110.5M	€565K	2202	Left
20801	Cristiano R	33	https://cdr	Portugal	https://cdr	94	94	Juventus	https://cdr	€77M	€405K	2228	Right
190871	Neymar Jr	26	https://cdr	Brazil	https://cdr	92	93	Paris Saint-	https://cdr	€118.5M	€290K	2143	Right
193080	De Gea	27	https://cdr	Spain	https://cdr	91	93	Mancheste	https://cdr	€72M	€260K	1471	Right
192985	K. De Bruyı	27	https://cdr	Belgium	https://cdr	91	92	Mancheste	https://cdr	€102M	€355K	2281	Right
183277	E. Hazard	27	https://cdr	Belgium	https://cdr	91	91	Chelsea	https://cdr	€93M	€340K	2142	Right
177003	L. Modrić	32	https://cdr	Croatia	https://cdr	91	91	Real Madri	https://cdr	€67M	€420K	2280	Right
176580	L. Suárez	31	https://cdr	Uruguay	https://cdr	91	91	FC Barcelo	https://cdr	€80M	€455K	2346	Right
155862	Sergio Ram	32	https://cdr	Spain	https://cdr	91	91	Real Madri	https://cdr	€51M	€380K	2201	Right
200389	J. Oblak	25	https://cdr	Slovenia	https://cdr	90	93	Atlético Ma	https://cdr	€68M	€94K	1331	Right
188545	R. Lewandc	29	https://cdr	Poland	https://cdr	90	90	FC Bayern	https://cdr	€77M	€205K	2152	Right
182521	T. Kroos	28	https://cdr	Germany	https://cdr	90	90	Real Madri	https://cdr	€76.5M	€355K	2190	Right
182493	D. Godín	32	https://cdr	Uruguay	https://cdr	90	90	Atlético Ma	https://cdr	€44M	€125K	1946	Right
168542	David Silva	32	https://cdr	Spain	https://cdr	90	90	Mancheste	https://cdr	€60M	€285K	2115	Left
215914	N. Kanté	27	https://cdr	France	https://cdr	89	90	Chelsea	https://cdr	€63M	€225K	2189	Right
211110	P. Dybala	24	https://cdr	Argentina	https://cdr	89	94	Juventus	https://cdr	€89M	€205K	2092	Left
202126	H. Kane	24	https://cdr	England	https://cdr	89	91	Tottenham	https://cdr	€83.5M	€205K	2165	Right
194765	A. Griezma	27	https://cdr	France	https://cdr	89	90	Atlético Ma	https://cdr	€78M	€145K	2246	Left
192448	M. ter Steg	26	https://cdr	Germany	https://cdr	89	92	FC Barcelo	https://cdr	€58M	€240K	1328	Right
192119	T. Courtois	26	https://cdr	Belgium	https://cdr	89	90	Real Madri	https://cdr	€53.5M	€240K	1311	Left
189511	Sergio Busc	29	https://cdr	Spain	https://cdr	89	89	FC Barcelo	https://cdr	€51.5M	€315K	2065	Right
179813	E. Cavani	31	https://cdr	Uruguay	https://cdr	89	89	Paris Saint-	https://cdr	€60M	€200K	2161	Right
167495	M. Neuer	32	https://cdr	Germany	https://cdr	89	89	FC Bayern	https://cdr	€38M	€130K	1473	Right
153079	S. Agüero	30	https://cdr	Argentina	https://cdr	89	89	Mancheste	https://cdr	€64.5M	€300K	2107	Right

Dataset : Data Dictionary

- ID = unique id for every player
- Age = Age of player
- Photo =url to the player's photo
- Flag = url to player's country flag
- Overall = overall rating of a player out of 100
- Potential = potential rating of a player out of 100
- Value = current market value of a player
- Wage = current wage of a player
- Preferred Foot = Preferred foot of a player left/right
- Accelerations: is the increment of a player’s running speed.
- Agility: how agile the player is while moving or turning.
- Crossing: how accurately the player crosses the ball during both normal running and free kick set pieces.
- Dribbling: is the player’s ability to carry the ball and past an opponent.
- Marking is the ability to track and defend an opposing player
- RF: Right forward
- CF: Center forward
- LF: Left forward
- LWM: Left Wing Midfield

Data:

Player personal data like:

- Id, Nationality, Photo, Club, Age, Wage, Value, Position, Preferred Foot etc.

Skill attributes like:

- Accelerations, Crossing, Dribbling, Marking,Vision, Penalties, Finishing, Heading, Jumping, Stamina, Accuracy, Short Passing, Volleys,, Standing Tackle, GK Diving, GK Handling, GK Kicking, GK Positioning, GK Reflexes, and Release Clause, LS, ST, RS, LW, LF, CF, RF, RW, LAM, CAM, RAM, LM, LCM, CM etc.

Cleaning and Preprocessing data

- We need to do some preprocessing for variables like Height, Weight and monetary variables like Value, Wage so that they're converted to a numeric format.
- Missing values were not removed directly as we thought we might need it for the future, so what we did it for all the column that had missing values we created one more column were we replace the missing value with standard deviation.

Why We Need It?

- This will help the manager who manages the soccer team for country, clubs in order build best team based on the requirement.
- Players can use this to see how other player are performing.
- Game developer can use it to get more insights and improve the game.

Questions:

- Maximum players belong to which country?
- What is the average overall ratings of a player?
- Relationship between Potential and Overall power of a player.
- Do player's wages depend on the value?
- Is the players worth real?
- Market value of the player based on the finishing?
- What are the best skills of a player?

Types of Users:

Players (Naive):

- Analyze the performance of other teams and player.
- Improve his game by seeing the past performance of the player having similar skills and potentials as him.

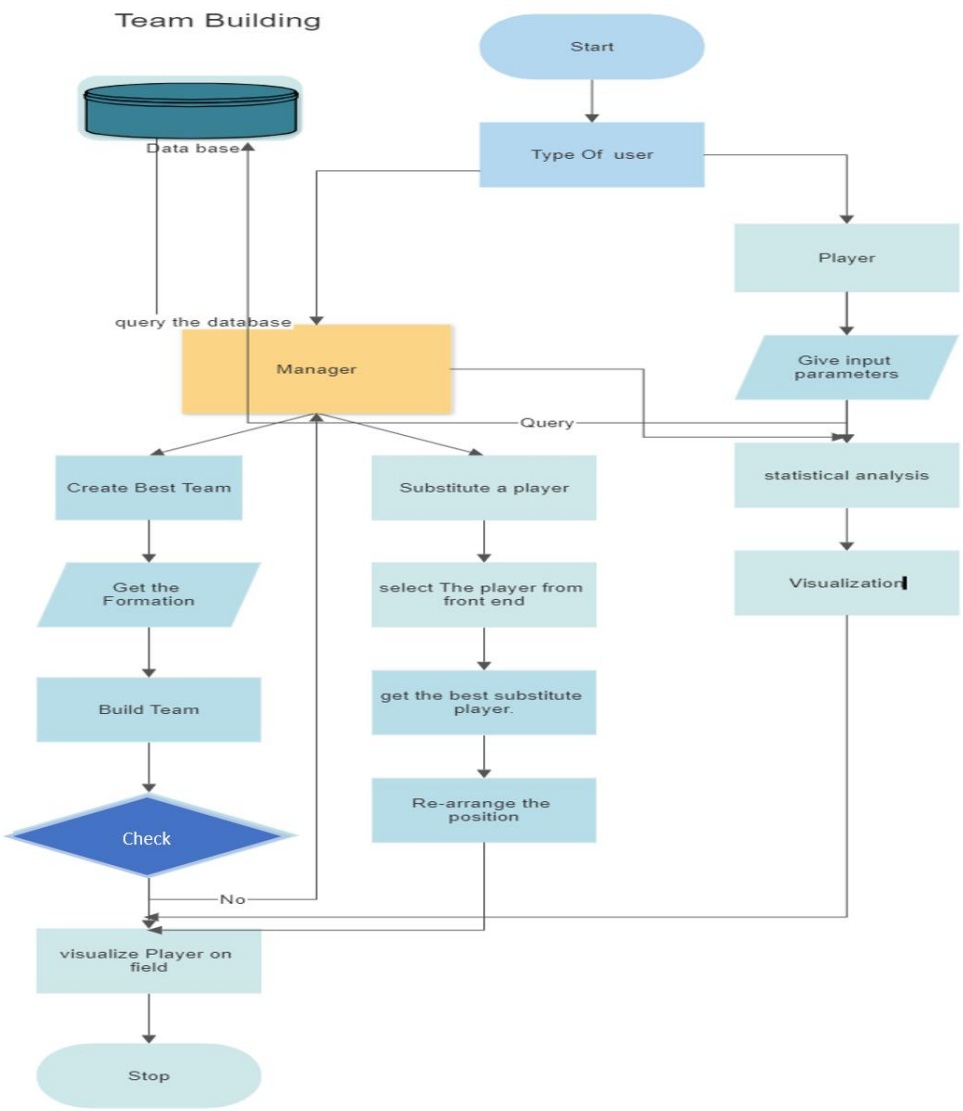
Club Manager and Investor(Intermediate):

- Create the best team based on different formation.
- Create the best team for a budget.
- Substitute a player.

Game Developer(Professional):

- Find the complex relation like between performance and potential of the player with respect to their value.

Flow Diagram:

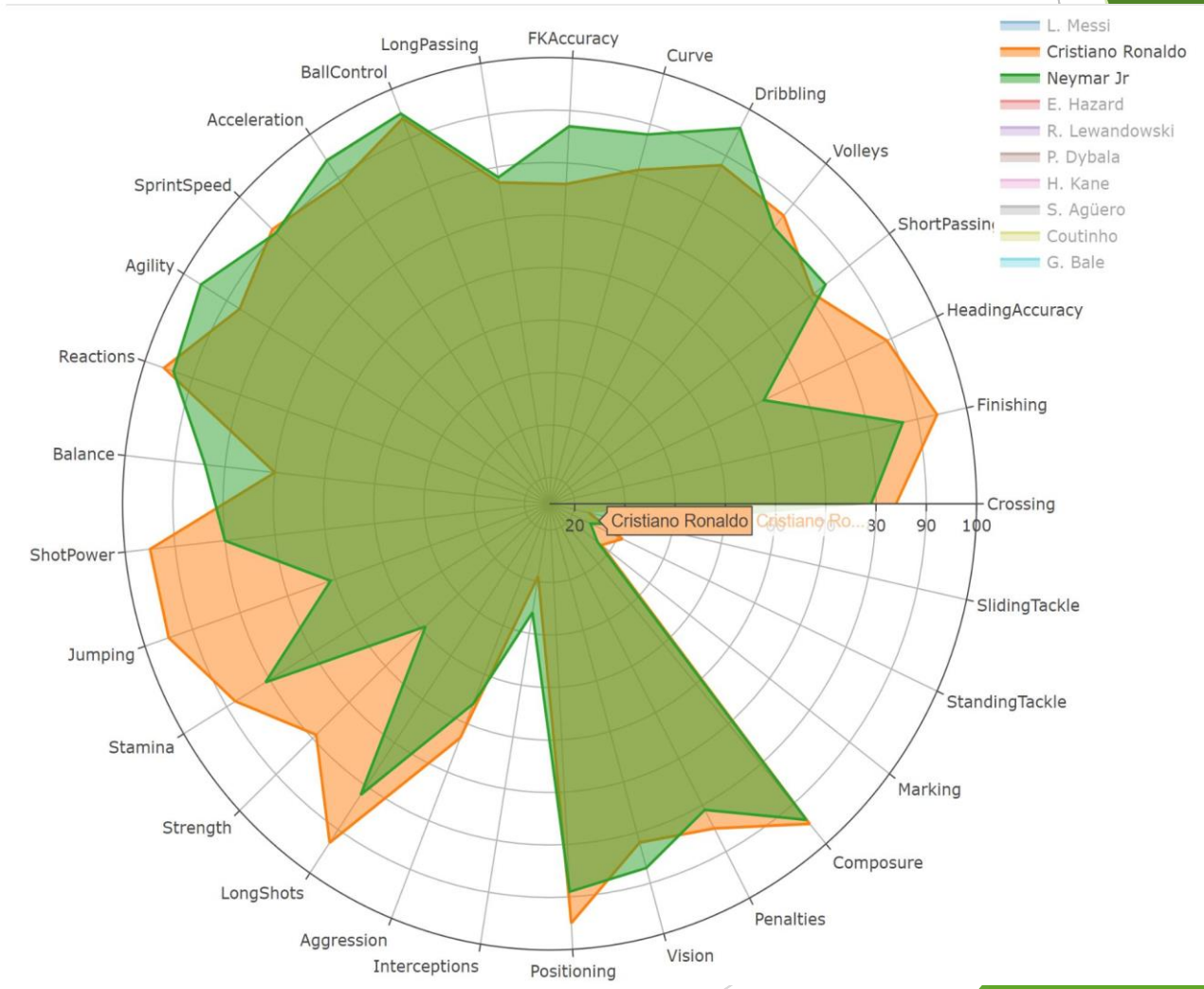


Technologies To be used:

- Database: SQLlight
- Back end: Python
- Front end: HTML, CSS, JAVA-Script, Django, Ajax, D3.js.
- Libraries : Plotly, Matplotlib, Pandas, numpy.

Compare the Skills Of Player:

- Radar plot to analyze and compare the performance of player.



Algorithm:

Improved

Select a Club:

Liverpool

Select a Nation:

Choose a Nation

Select a Formation:

4-2-3-1A

Build Team

Subsitute a Player:

Choose a Substitute

Substitute a Player

Team Score

905

Search:

Name	Position	Overall	Nationality
A. Robertson	LB	82	Scotland
Alisson	GK	85	Brazil
D. Lovren	RCB	81	Croatia
Fabinho	RB	84	Brazil
G. Wijnaldum	LCM	82	Netherlands
J. Henderson	RCM	82	England
M. Salah	RAM	88	Egypt
N. Keïta	CAM	83	Guinea

Showing 1 to 11 of 11 entries

Build The Best Team

hy greedy approach doesn't work.

y all possible combinations and pick up best one
me complexity is $({}^{n}C_{11}) * 11!$)

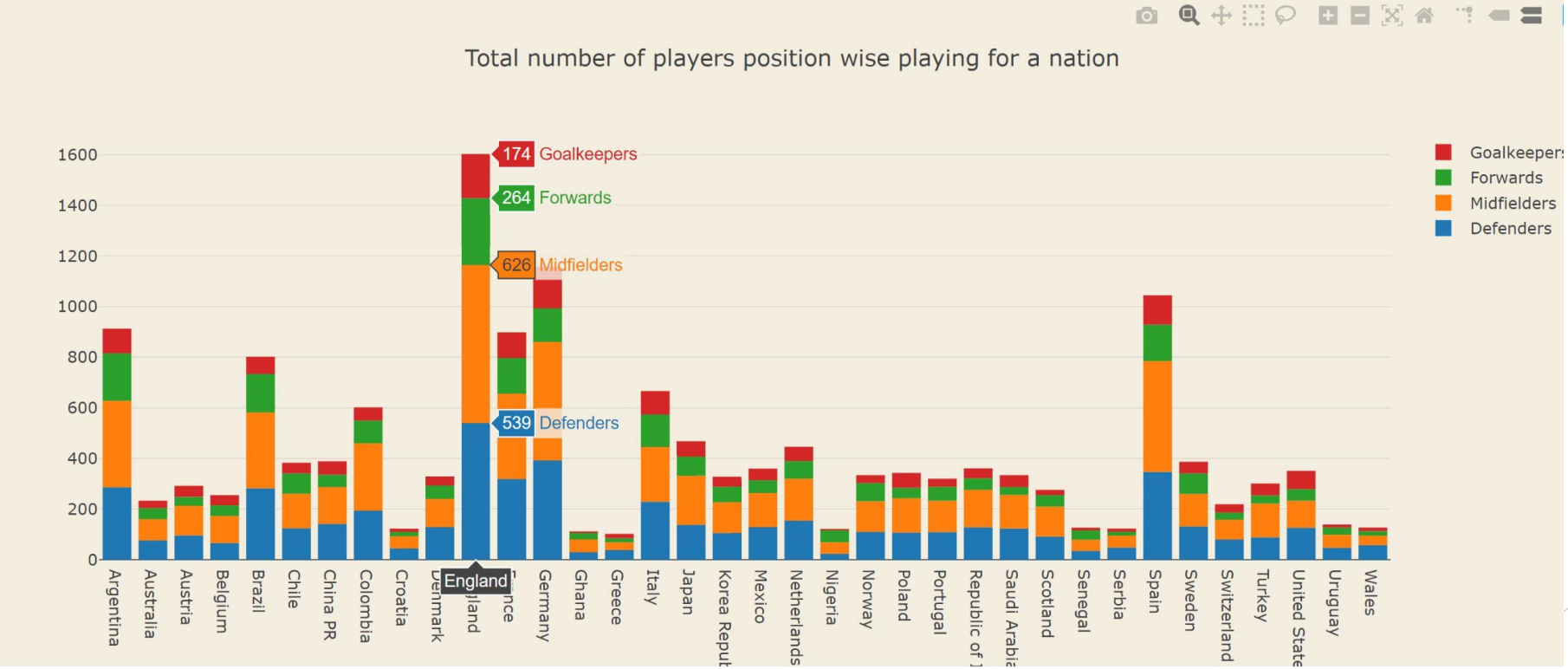
Algorithm used:

- User selects a Formation (like 4-3-3, 4-4-2)
- Once user selects a formation, we get the positions related to that formation (ST,LW,RW etc)
- Get list attacking, midfield, defender positions [[ST,RW,LW,CAM], [LCM,RCM],[RCB,LCB,LB,RB]] for 4-3-2-1A
- For each type of configuration like (Attackers, Defenders), we choose C(total_attackers,num_attackers_required) and try out all the possible permutations and choose the best one

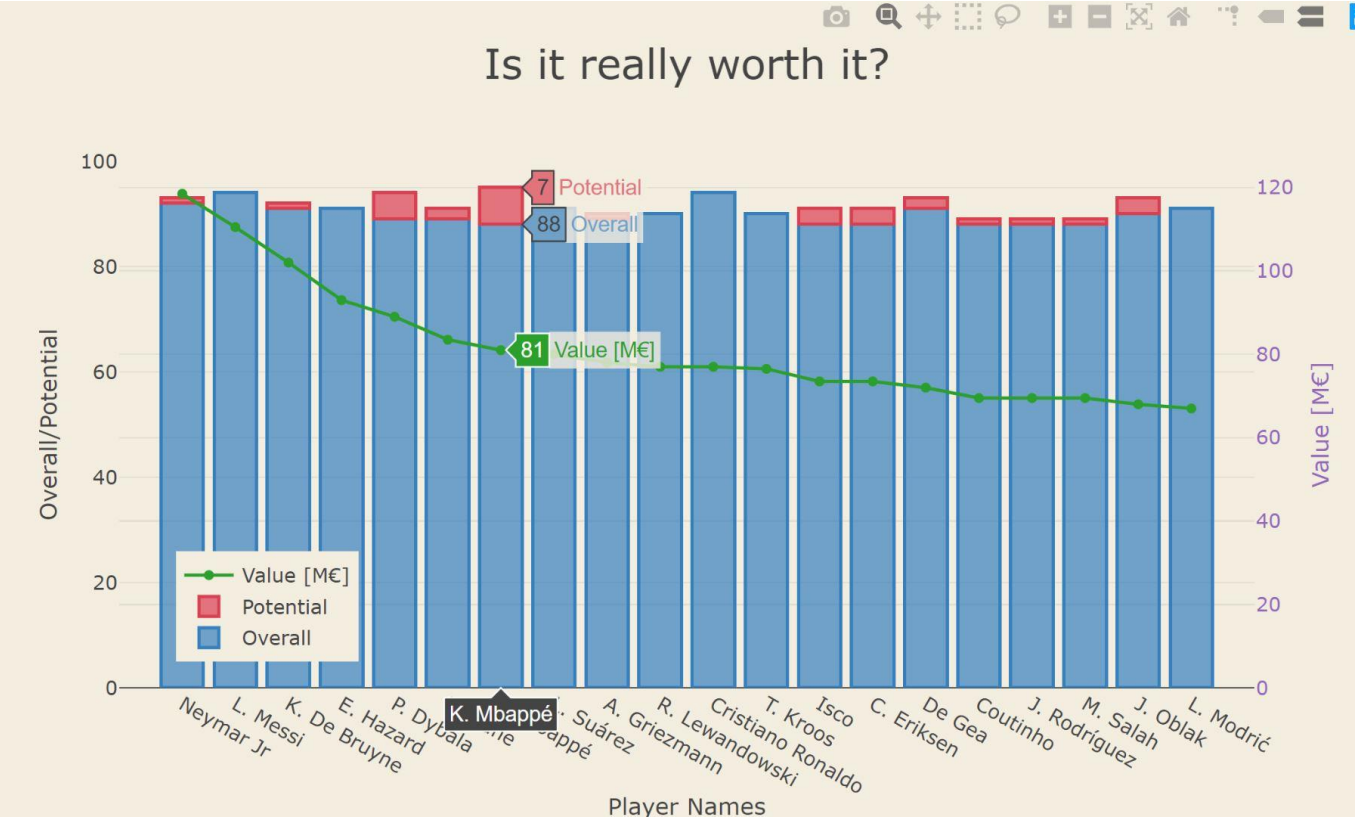
Substitute A Player

- We cannot substitute directly a similar player from bench because there is a possibility that there might be a player in current playing 11 who might perform better at the same position than from the bench.
- For example:
  - **(Player A, Player B)** are in the current playing 11
  - If we want to substitute **Player B**, There might be a similar **Player C** from the bench, But **Player A** plays better in the **Player B** position than **Player C**. So, we can instead choose a substitute for **Player A** to build a better team when we substitute **Player B**.
- Once the user selects a player to substitute, we add each player from the bench and build different teams accordingly. We choose the best team according to the approach discussed earlier.

Sample Data Analysis:

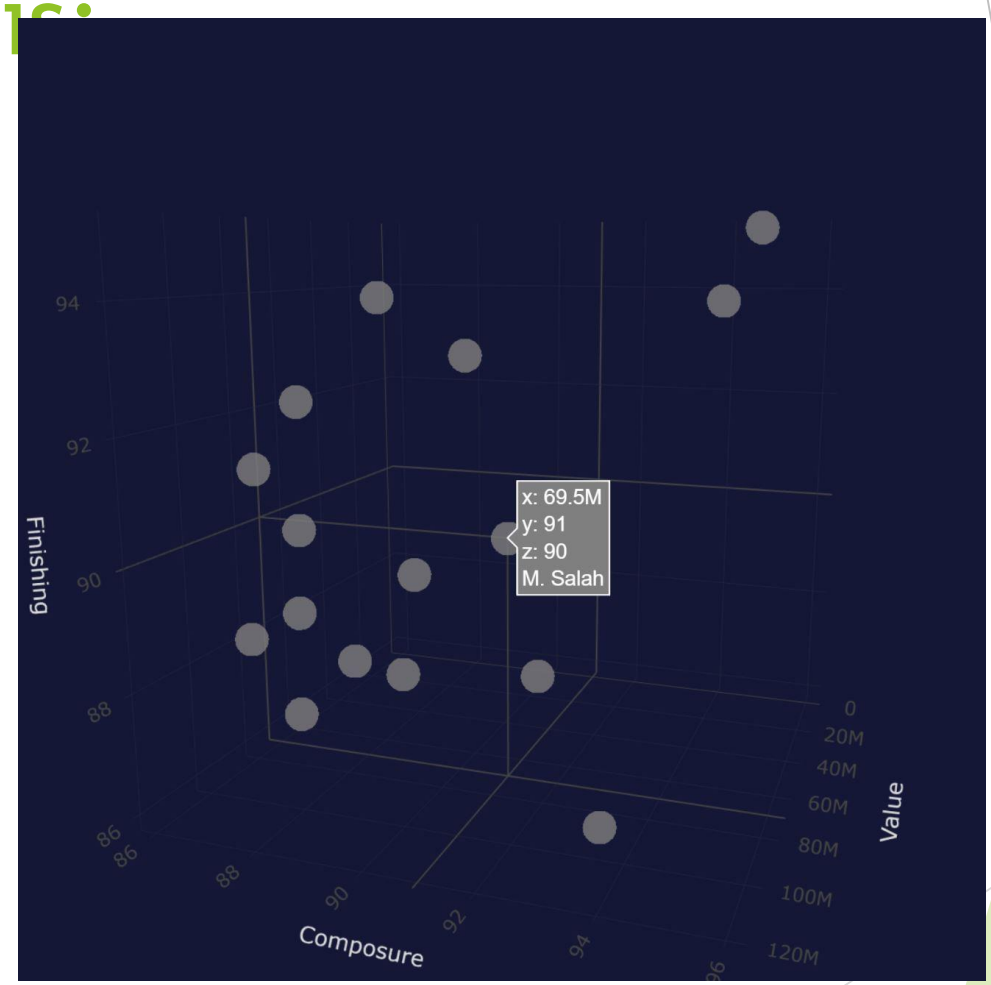


Sample Data Analysis:



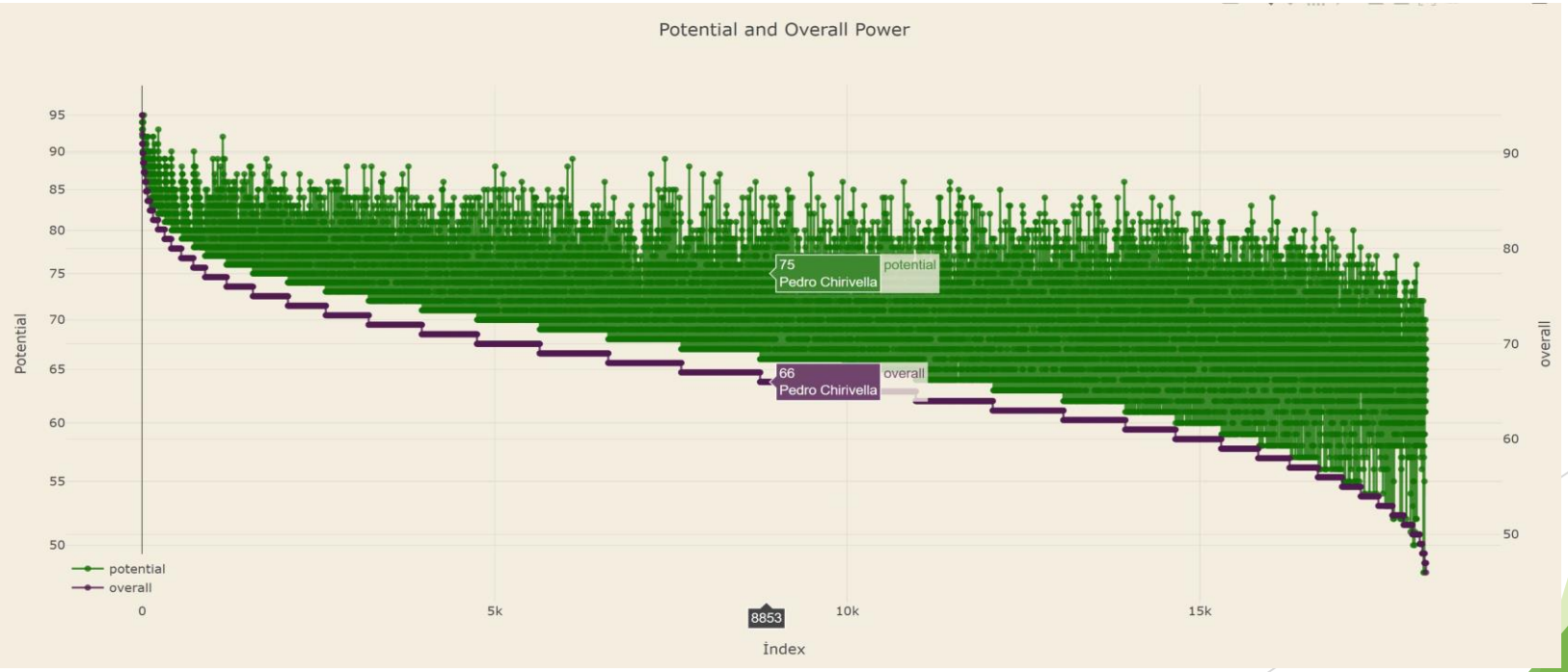
Sample Data Analysis:

- We try to find some if there are trends.
- Market value of the player is related to its finishing using the 3D graph. We can analyze the affect of good or bad finishing and composure on there value.

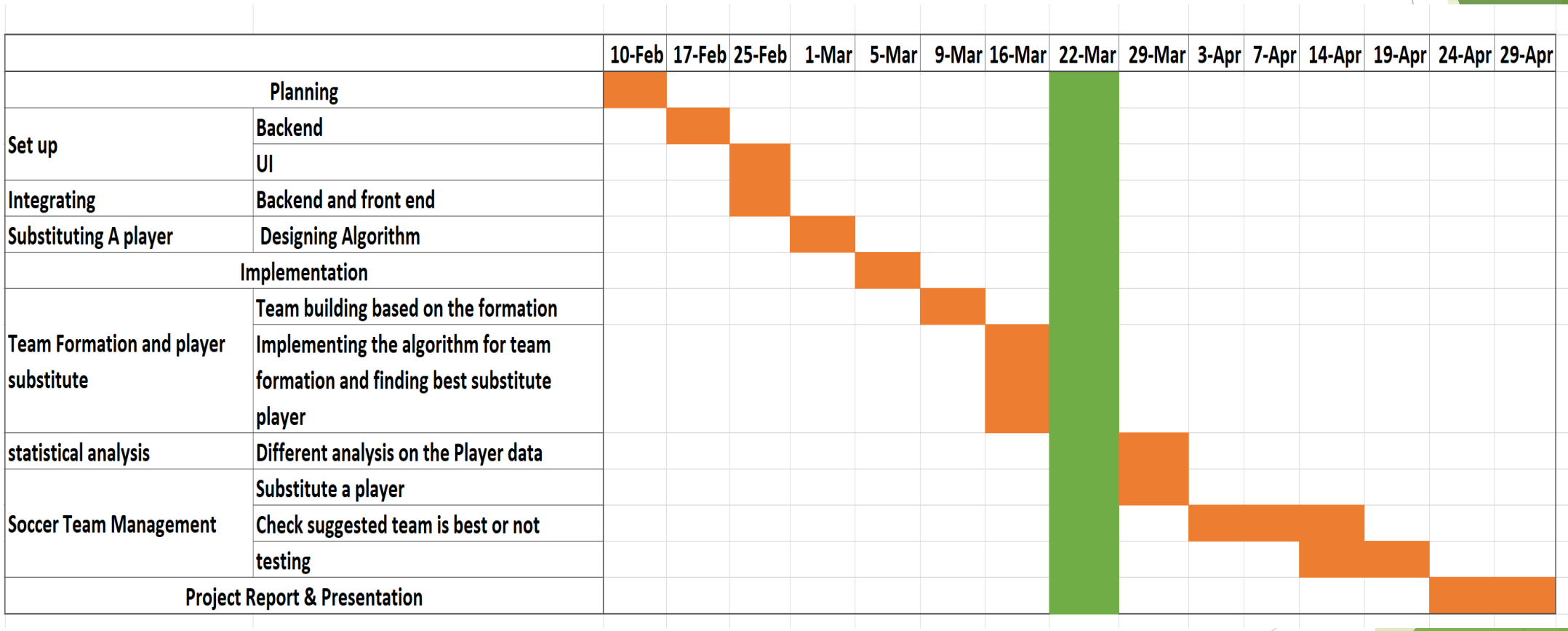


Sample Data Analysis:

- The fig on the top is the line scatter graph of potential and overall performance of a player.



Time Line


10-Feb
17-Feb
25-Feb
1-Mar
5-Mar
9-Mar
16-Mar
22-Mar
29-Mar
3-Apr
7-Apr
14-Apr
19-Apr
24-Apr
29-Apr
Planning
Set up
Backend
UI
Integrating
Backend and front end
Substituting A player
Designing Algorithm
Implementation
Team Formation and player substitute
Team building based on the formation
Implementing the algorithm for team formation and finding best substitute player
statistical analysis
Different analysis on the Player data
Soccer Team Management
Substitute a player
Check suggested team is best or not
testing
Project Report & Presentation

References

- https://www.degruyter.com/downloadpdf/j/hukin.2012.31.issue--1/v10078-012-0015-7/v10078-012-0015-7.pdf
- https://books.google.com/books?hl=en&lr=&id=ChpdantgTqQC&oi=fnd&pg=PA77&dq=fifa+best+teams&ots=GNWNfcXs0j&sig=xZmZd5CKrRfco7mu4RlhNnlmVys#v=onepage&q=fifa%20best%20teams&f=false
- https://arxiv.org/abs/1506.07768
- https://oaktrust.library.tamu.edu/handle/1969.1/173614
- https://link.springer.com/article/10.1007/s40279-017-0836-6
- https://www.kaggle.com/jsaguiar/complete-eda-time-series-with-plotly
- https://github.com/amanthedorkknight/fifa18-all-player-statistics

Thank YOU