

STREAMLIT-PLOTLY PROJECT DOCUMENTATION

(An Insight into Ten season of EPL Football 2010 – 2020)

Abubakar Muhammed Muktar & Amalai Rham Machunga

Technologies Utilized: Plotly, Streamlit, Pandas, and Numpy

Problem Statement: Use a dataset of your choice and make various plots explaining some insights into your dataset.

Datasets Used: *all_season_table.csv, All_tables.csv, epl-allseasons-matchstats.csv*

Image Used: *pl.png*

Plots: bar chart, scatter plot, sunburst chart, line chart, tree map, histogram chart, funnel chart, pie chart, box plot, strip chart, and donut chart.

PROCESS:

1. Downloaded the dataset.
2. Preprocessed our dataset so that is it fit for visualization without giving errors or incorrect visualizations.
3. Imported the packages (Plotly, Streamlit, Pandas, and Numpy), needed for the project to run without errors - *see app.py*
4. Loaded the .csv datasets - *see app.py*
5. Carried out some 'groupby' of the dataset – *see app.py*
NB: *load the app.py to streamlit using the command 'streamlit run app.py'*
6. Uploaded an 'English Premier League' image to our streamlit project.



7. Used and st.expander function to group our tabs into a single dropdown (Visualization 1, 2, 3 and 4 for easy access.

Visualization 1 ▾

Visualization 2 ▾

Visualization 3 ▾

Visualization 4 ▾

8. Used a selectbox for the user to choose between the premier league seasons of interest.

Select Season?

2010/11 ▾

2010/11

2011/12

2012/13

2013/14

2014/15

2015/16

2016/17

2017/18

9. Under *Visualization 1* dropdown, we used the **Bar Chart** to visualize the ‘Total points per Team for the season’, the **Scatter Plot** to visualize the ‘Total Goals scored per team per season’ and ‘Total goals Conceded per team per season’.
10. Under *Visualization 2* dropdown, we used the **Sunburst Chart** to visualize the ‘Position and games Lost Per Season’, the **Line Chart** to visualize ‘Total point across all seasons’ and the **Treemap** to visualize the ‘Win Percentage per Team per Season’.
11. Under *Visualization 3* dropdown, we used the **Histogram** to visualize the ‘Goals scored per team for Each Season’, the **Funnel Plot** to visualize the ‘Total Goals Scored and Conceded Per team’, and the **Pie Chart** to visualize the ‘Total Goals & Percentage across 10 Seasons’.
12. Under *Visualization 4* dropdown, we used the **Box Plot** to visualize the ‘Total Yellow Cards per club’, the **Strip Chart** to visualize the ‘Total Red Cards per club’, and the **Donut Chart** to visualize the ‘Total Cards Issued by Referees’.

SUMMARY:

In Summary, this Project gives an insight into 10 seasons of premier league action from 2010 – 2020 season using various charts and plots to visualize the dataset. We chose this dataset because of our love for the game of soccer and we both support the same premier league club (name withheld, lol).