Name: Sakar Adhikari
Andrew Id: sakara
Subject Code: 12780-A2

Submission Date: Friday, Dec 09, 2023

Project: Interactive Soccer League Dashboard and Predictive Analysis

1. Introduction:

- Soccer has been a significant part of my life since a memorable trip to Manchester in 2004, where, alongside my dad, I experienced the thrill of a Manchester United match. That day, amidst the roaring

crowd and the intense game, I fell in love with the energy and spirit of the English Premier League. This wasn't just a sport; it was a world of excitement and passion that I instantly became a part of. From playing soccer at school to being an active member of my university team, the game has always been my constant companion. Every weekend, I'm glued to the TV, watching matches, analyzing plays, and cheering for my favorite teams. This project, 'Interactive Soccer League Dashboard and Predictive Analysis,' is more than an academic choice for me. It's a personal journey. It combines my technical skills with my deep passion for soccer. My aim is to create a platform that brings to life the rich statistics and dynamic nature of soccer, making it easy for others to understand and appreciate the beauty of this game, just like I do.

2. Background:

2.1 **Engineering Background:** The project is centered around the management and interpretation of large volume of soccer related statistical data. The engineering challenge here is to effectively process this data which includes various metrics like team performance, match outcomes and present it in more digestible format. The task also involves developing a user-friendly interface that allows users to explore this data interactively.

2.2 **Computational Methods:** The computational method for this project involves two primary component that is data processing (data analysis) and web development. In this case data processing involves structuring the large data sets for efficient analysis. The web development aspect involves creating an interactive interface where users can explore various data visualizations.

2.3 **Data Source:** The data required for this project is sourced from a comprehensive dataset available on Kaggle, specially focusing on the English Premier League. The dataset is publicly accessible at the following URL: https://www.kaggle.com/datasets/saife245/english-premier-league . The dataset is in the CSV format which encompasses a wide range of soccer statistics.

3. Learning Goal:

- To successfully execute this project, I need to enhance my skills in data processing particularly in handling and analyzing datasets. I will possibly learn about time series analysis for season wise data interrelation. I am thinking about creating various charts with plotly where I will use advanced visualization customizing graphs and making them interactive based on user input. I also need to learn web application development that focus on creating interactive user interfaces and integrating data visualization.

4. Project Plan:

- I have outlined a plan comprising various tasks along with an estimated time for each. My final report will encapsulate the outcomes of these tasks and the learning acquired during the project execution.

4.1 Data Collection and Initial Processing (~ 2 hours)

- My first task involves gathering the dataset from Kaggle which I will preliminary clean and do the structuring. My objective is to prepare the data set for initial analysis and ensuring making the dataset easier to understand for further processing.

4.2 In Depth Data Analysis and Learning Advanced Processing Techniques (~ 4 hours)

- Next, I will focus on advanced data analysis. In this phase I am planning to master more sophisticated data processing methods and thoroughly analyzing the dataset for key insights.

4.3 Web Development and Flask Framework Implementation (~ 6 hours)

- Based on class and my personal learning I will dedicate 5 hours for web development and flask framework. The goal during this time would be to set up basic web application structure using flask.

4.4 Integration of Data Visualization Tools (~6 hours)

- Integrating plotly for data visualization would be the focus for this task. I aim to create and embed interactive charts and graphs into the web application and enhancing its data presentation capabilities.

4.5 Enhancing User Interface and Interactivity (~3 hours)

- I will spend 3 hours to improve the web applications user interface and interactivity. I will also refine the front-end design and add dynamic user input features for data exploration.

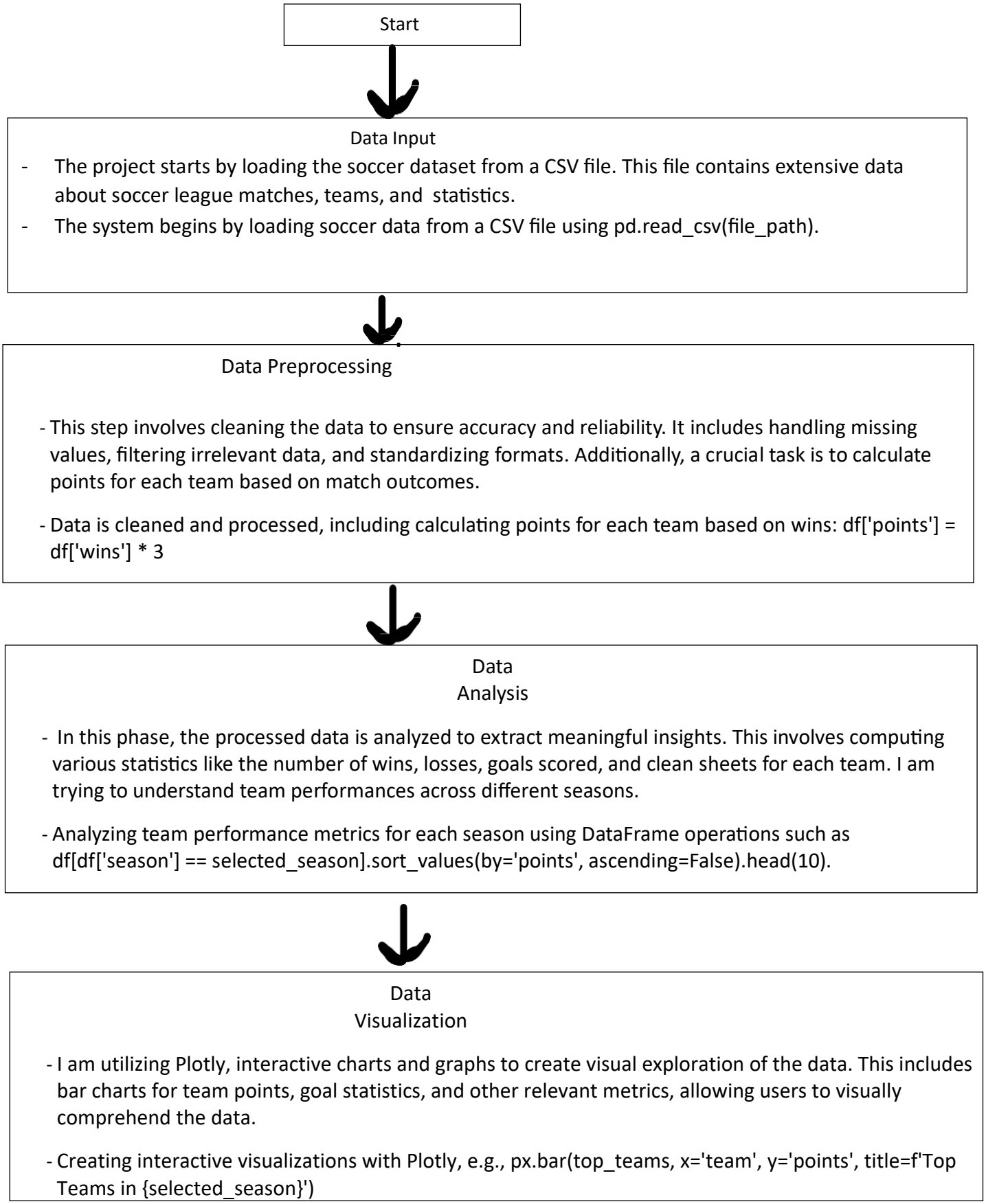
4.6 Final Testing and Documentation (~2 hours)

- The final phase of the project would involve testing the application, fixing the bug, and preparing comprehensive documentation. The outcome will be a fully functional and well documented web application.

5. Deliverables:

- The final deliverables for this project will include:
  1. A fully functional web application for soccer data visualization
  2. Interactive elements allowing users to filter and explore data based on teams and seasons.
  3. Various types of data visualization to represents different statistical aspects of the soccer data.

6. System Diagram



Web Interface

- I am developing user-friendly web interface using Flask. This interface incorporates interactive elements such as dropdown menus and buttons, enabling users to make selections (like choosing a team or a season) for data visualization.
- Developing a Flask web application and defining routes, as in app = Flask(__name__) and @app.route('/', methods=['GET', 'POST']).

User Interaction

- Users will interact with the application by selecting options like specific teams or seasons. Their interactions prompt the application to display corresponding data visualizations and statistics.
- Handling user input from web forms, such as selected_season = request.form.get('season').

Result

- Finally, the application presents the results of user interactions in an intuitive format. This includes displaying the data visualizations and insights derived from the users' selections, providing a clear and engaging presentation of soccer statistics.
- Displaying results on the web interface using render_template('index.html', seasons=seasons)

Stop

7. Preliminary code and Documentation

- **Implementation overview :** The Flask web application is structured around several key functions, each handling a specific aspect of the application:

- Process_data(): Handles the preprocessing of the Kaggle dataset, preparing it for analysis and visualization.

- Generate_visualization(): Uses Plotly to create interactive charts and graphs based on the processed data.

- Create_app: Uses Plotly to create interactive charts and graphs based on the processed data.

Code Snippet:

```python
from flask import Flask, render_template, request
import pandas as pd
import plotly.express as px

app = Flask(__name__)

# Loading the dataset and calculating the points(3 point to winner 0 to losser)
file_path = '/Users/sakaradhikari/Desktop/stats.csv'
df = pd.read_csv(file_path)
df['points'] = df['wins'] * 3  # 3 points for a win

@app.route('/', methods=['GET', 'POST'])
def index():
    seasons = df['season'].unique()
    selected_season = request.form.get('season') if request.method == 'POST' else seasons[0]

@app.route('/team_analysis', methods=['GET', 'POST'])
def team_analysis():
    teams = df['team'].unique()
    seasons = df['season'].unique()
    selected_team = request.form.get('team') if request.method == 'POST' else teams[0]
    selected_season = request.form.get('season') if request.method == 'POST' else seasons[0]

if __name__ == '__main__':
    app.run(debug=True)

```