J-LEAGUE DATA ANALYSIS PROJECT REPORT

The purpose of the project was to derive insights from the 2024 J1 League table. Below are the steps taken and analysis I conducted.

1. Data Scraping

I first decided to scrape the 2024 J1 League table from the official website. I used the Beautiful Soup library to help me parse HTML, extract the table data and create the data frame.

2. Data Preparation

I removed special characters to avoid discrepancies. I then converted columns into their appropriate data types. Numeric columns were converted to ints and text columns were converted to string.

3. Goals Scored Analysis

I then plotted the top 10 teams with the most goals scored using Matplotlib

- **Insight:** Sanfrecce Hiroshima scored the most goals (72) in 2024.

I then plotted the top 10 teams with the least goals scored using Matplotlib

- **Insight:** Avispa Fukuoka scored the least goals (33) in 2024.

4. Goals Scored Grouping

I then grouped by the number of total goals scored, counted how many teams scored that number of goals, and listed the team names next to the count

- **Insight:** The modal number of goals scored was 43. 3 teams (Cerezo Osaka, Kyoto Sanga, and Hokkaido Consadole Sapporo) had scored that many goals.

5. Goals Conceded Analysis

I then plotted the top 10 teams with the least goals conceded using Matplotlib

- **Insight:** FC Machida Zelvia conceded the least goals (34) in 2024.

I then plotted the top 10 teams with the most goals conceded using Matplotlib

- **Insight:** Sagan Tosu and Jubilo Iwata conceded the most goals (68) in 2024.

6. Goals Conceded Grouping

I then grouped by the number of total goals conceded, counted how many teams conceded that number of goals, and listed the team names next to the count

- **Insight:** The modal number of goals conceded was 51. 3 teams (Tokyo Verdy, FC Tokyo, and Kashiwa Reysol) had conceded that many goals.

7. Recent Form Analysis

I then displayed the number of points each team earned in their last 5 matches in descending order alongside its position. I then plotted the number of points each team earned in their last 5 matches (in descending order) using Matplotlib

- **Insight:** Kashima Antlers had earned the most points in their last 5 matches, earning 11.

8. Goal Difference Analysis

I then plotted the goal difference by team (in descending order) using Matplotlib

- **Insight:** Sanfrecce Hiroshima had the best goal difference at +29. Hokkaido Consadole Sapporo had the worst goal difference at -23.

I then grouped by the goal difference (in descending order), counted how many teams had that goal difference, and listed the team names next to the count

- **Insight:** The modal goal difference was -5. Cerezo Osaka, Avispa Fukuoka, and Shonan Bellmare ended with a goal difference of -5.

9. Points Grouping

I then grouped by the number of points (in descending order), counted how many teams scored earned that many points, and listed the team names next to the count

- **Insight:** The modal number of points a team finished with is 52, with 3 teams (Kawasaki Frontale, Yokohama F. Marinos, and Cerezo Osaka) each finishing with that many points.

10. Points Per Game Analysis

I then displayed the position alongside the number of points each team earned per game (in descending order)

- **Insight:** Vissel Kobe earned the most points per game, with 1.8947 points/game. Sagan Tosu earned the least number of points per game, with 0.921 points/game.