

Analysis Document: Top Scoring Minutes Visualization

Objective:

The goal of the code is to visualize the top 4 scoring minutes of all time (from 1872–2024) in football, using a pie chart and a horizontal bar chart. This analysis provides insights into the minutes during which goals were most frequently scored in recorded matches.

Data Description:

- **Dataset:** `'Goalscorers.csv'` (CSV file)
- **Column Used:** `'minute'` – Contains the minutes during which goals were scored in various matches.
- **Null Value Handling:**
 - Rows with missing values in the `'minute'` column are dropped.
 - Remaining NaN values are replaced with the placeholder 'No Goals.'

Process Breakdown:

1. **Data Preprocessing:**

- Data is read from the CSV file using `'pandas'`.
- The `'minute'` column is extracted, null values are dropped, and remaining empty values are filled.

2. **Counting Goal Occurrences:**

- The `'Counter'` class from `'collections'` is used to count occurrences of each scoring minute.
- The `'most_common(4)'` method identifies the top 4 scoring minutes.

3. **Preparing Data for Visualization:**

- The scoring minutes and their respective goal counts are extracted into separate lists for easy visualization.
- Lists are reversed to ensure data is in ascending order for better readability in the bar chart.

4. **Visualization:**

- **Pie Chart:**

- Displays the percentage distribution of goals scored in the top 4 minutes.
- Uses `'explode'` to slightly emphasize the minute with the highest number of goals.

- **Horizontal Bar Chart:**

- Illustrates the absolute number of goals scored for each of the top 4 scoring minutes.
- Includes labeled axes for clarity.

Insights:

The visualization highlights the most goal-rich minutes in football history, offering valuable insights for analysts, coaches, and enthusiasts. The combination of a pie chart and bar chart ensures both percentage-based and count-based perspectives are provided.

Code Style and Design:

- **Visualization Style:**

- `matplotlib` is used with the `fivethirtyeight` style for a clean and professional appearance.

- **Titles and Labels:**

- Both charts are appropriately titled and labeled to ensure interpretability.

- **Color Scheme:**

- The pie chart uses an explosion effect for emphasis, while the bar chart uses a consistent 'sky blue' theme for clarity.

Recommendations for Improvement:

1. Validate the dataset to ensure consistency in the `minute` column (e.g., filtering out invalid minute values like negative numbers).
2. Extend the analysis to include more than the top 4 scoring minutes for broader insights.
3. Consider using interactive visualizations with tools like Plotly for enhanced user engagement.