R programming project

Name: H D Sathvik

SRN: PES1UG21CS212

Sem: 6th (completed)

DATASET USED:

https://data.world/datatouille/stephen-curry-stats/workspace/file?filename=Stephen+Curry+Stats.xlsx

Packages installed:

```
# Install and load necessary packages
install.packages(c("readxl", "tidyverse", "ggplot2", "dplyr"))
library(readxl)
library(tidyverse)
library(ggplot2)
library(dplyr)
```

Loading the dataset:

```
# Load the dataset
file_path <- "C:\\Users\\sathv\\Desktop\\R project\\Stephen Curry Stats dataset.xlsx"
nba_data <- read_excel(file_path)
```

DATA ANALYSIS:

How does Stephen Curry's points scored vary over time?

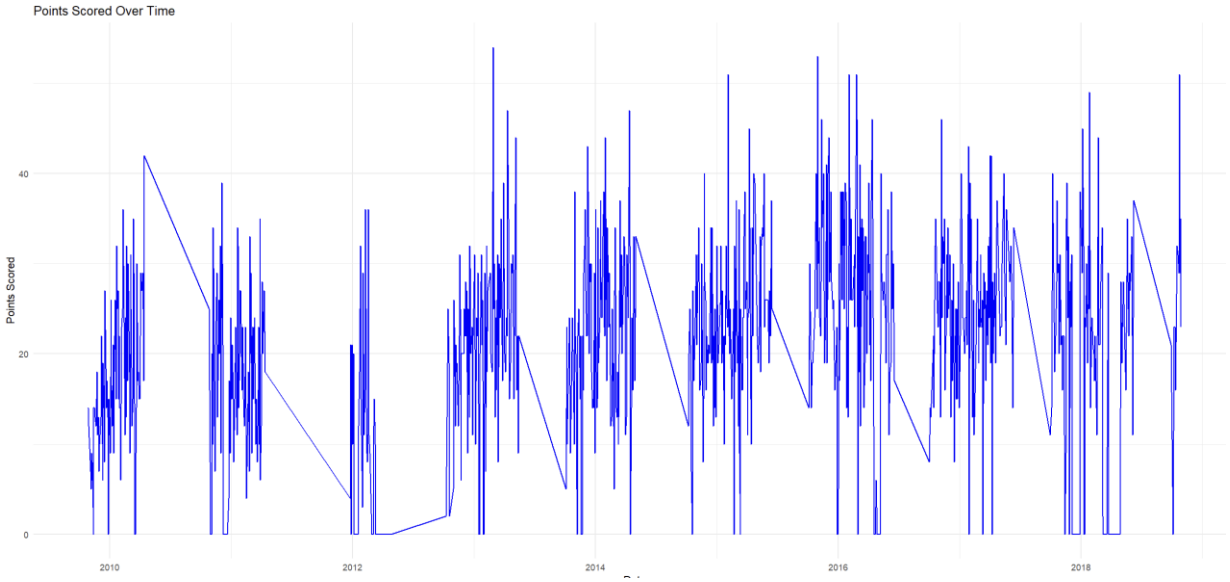
This analysis is mainly done to check the trends in Curry's performance over time, for example improvements or decline throughout his career. This can be helpful in understanding his overall career performance and patterns related to different seasons.

The analysis is done with the help of a line plot.

CODE:

```
## 1. Points scored over time
ggplot(data, aes(x=Dates, y=PTS)) +
  geom_line(color="blue") +
  labs(title="Points Scored Over Time", x="Date", y="Points Scored") +
  theme_minimal()
```

OUTPUT:



ANALYSIS:

From around 2012 onwards, Curry's performance has increased noticeably, this shows his development as a player. The sharp drops to 0 indicates missed games due to injuries, rest days etc. Despite the fluctuations, Curry's overall performance is strong, consistent mainly after the year 2012.

Against which opponent does Stephen Curry score the most points?

This analysis identifies against which opponent teams Curry performs better by scoring more and the teams against which he has tough competition. The analysis is valuable for game planning and understanding the players strong and weak points.

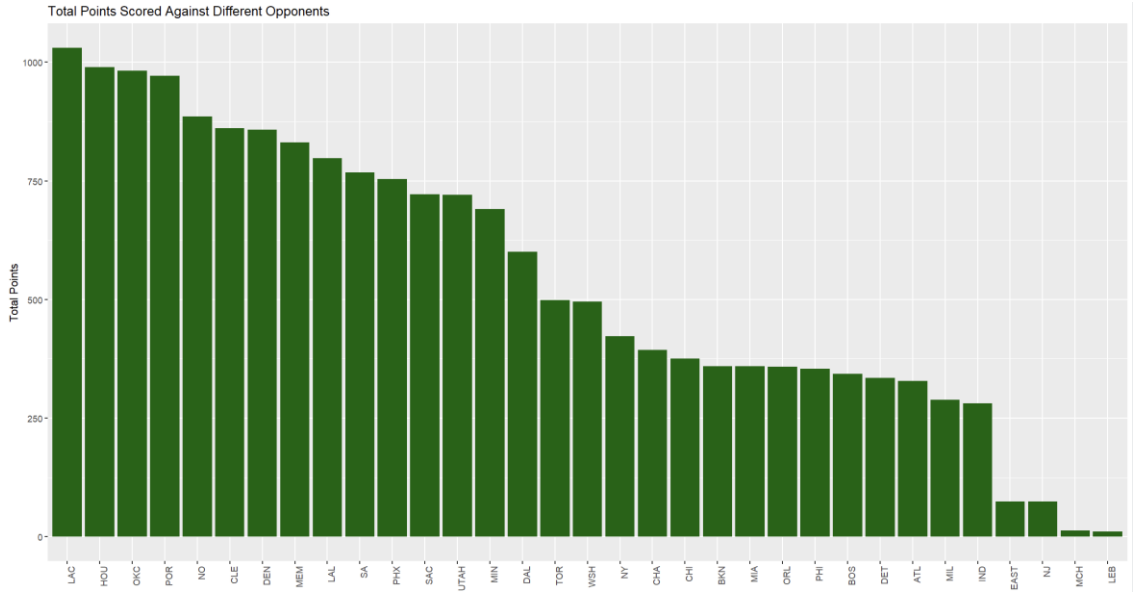
This analysis is done using the bar plot.

CODE:

```
## 2. Total points scored against different opponents
total_points_by_opponent <- aggregate(PTS ~ Opponent, data=data, sum)

ggplot(total_points_by_opponent, aes(x=reorder(Opponent, -PTS), y=PTS)) +
  geom_bar(stat="identity", fill="darkgreen") +
  labs(title="Total Points Scored Against Different Opponents", x="Opponent", y="Total Points")
  theme(axis.text.x = element_text(angle=90, hjust=1))
```

OUTPUT:



ANALYSIS:

As seen in the graph, Curry has scored the most points against LAC (Los Angeles Clippers) and then on HOU (Houston Rockets) and so on. If we can take a closer look we notice that these teams are few of the top teams in the NBA and hence Curry's performance is great under high pressure. We also notice that the points scored against the other teams like the EAST, NJ and LEB are lesser as Curry has not played too many games against these teams.

Is there a relationship between the minutes played and points scored by Stephen Curry?

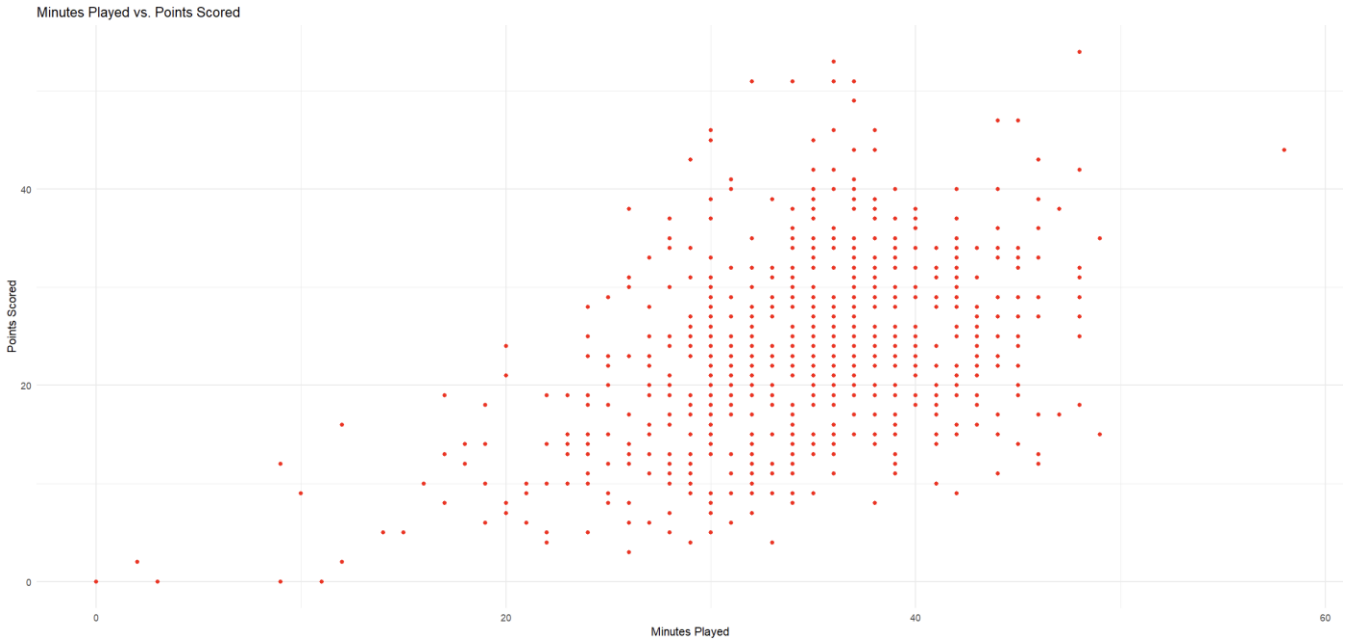
This analysis is done to check how efficient the player is on the court by correlating playing time and scoring. This can help in strategizing player rotation in game and managing his playing time and hence making the game more optimal.

The analysis here is done using the scatter plot.

CODE:

```
## 3. Relationship between minutes played and points scored
ggplot(nba_data, aes(x=Minutes, y=PTS)) +
  geom_point(color="red") +
  labs(title="Minutes Played vs. Points Scored", x="Minutes Played", y="Points Scored") +
  theme_minimal()
```

OUTPUT:



ANALYSIS:

Generally playing more minutes lead to the player scoring more points, in this case, curry scores up to 40 points if he has an in-game time of 30-40 minutes and scores a little lesser if he plays 0-20 minutes. This score is above average for an average NBA player and hence Curry is a great player and the outcome of most of the games depends on his playing time in it.

What is the distribution of points scored by Stephen Curry?

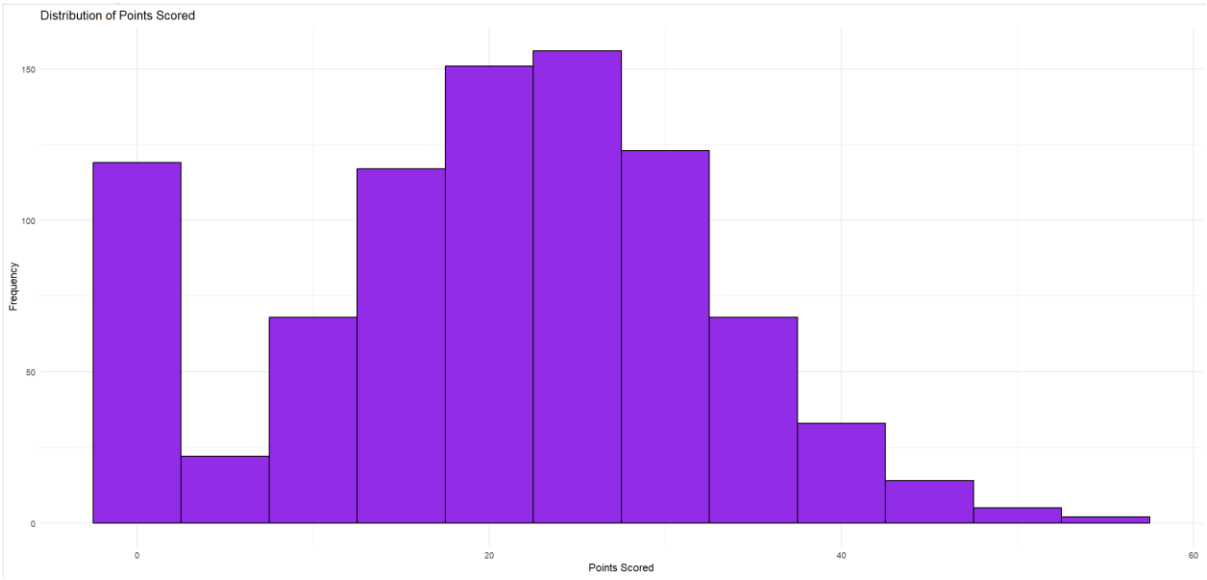
This is to mainly check how consistent curry is across his games and his scoring ability throughout his career.

The analysis is done using a Histogram.

CODE:

```
## 4. Histogram: Distribution of points scored
ggplot(data, aes(x=PTS)) +
  geom_histogram(binwidth=5, fill="purple", color="black") +
  labs(title="Distribution of Points Scored", x="Points Scored", y="Frequency") +
  theme_minimal()
```

OUTPUT:



ANALYSIS:

Curry averages 20 -30 points per game showing that he is very consistent with his scoring ability, although there are games where he is scored lower than 10 points this is mainly due to reasons like opponents defensive strategies, travel fatigue, physical conditions and health etc.

How does Stephen Curry's scoring differ between regular season and playoff games?

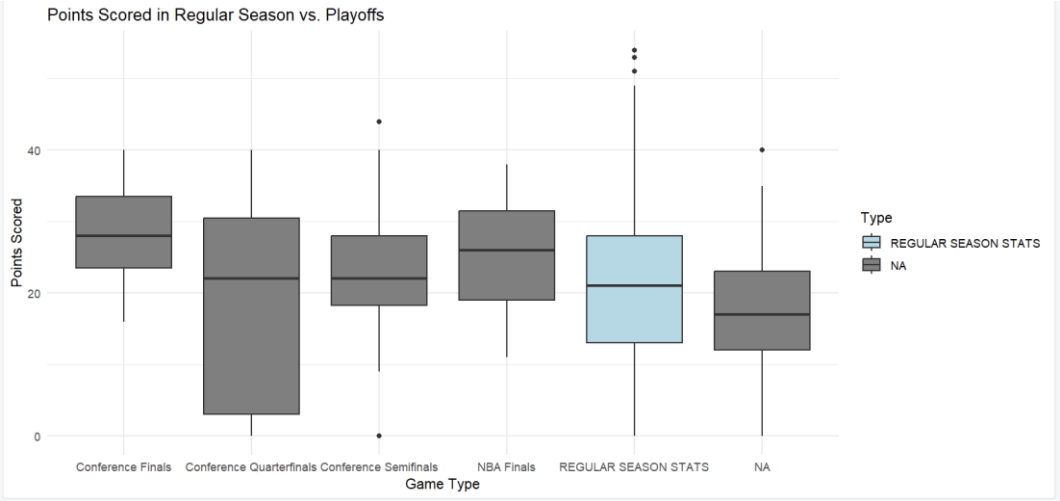
This is done mainly to check if Curry can elevate his game even during high pressure situations as this can be important in checking his clutch performance and reliability in crucial games

The analysis here is done using the box plot.

CODE:

```
## 5. Boxplot for points scored in regular season vs. playoffs
ggplot(data, aes(x=Type, y=PTS, fill=Type)) +
  geom_boxplot() +
  scale_fill_manual(values=c("REGULAR SEASON STATS"="lightblue", "PLAYOFFS STATS"="lightcoral")) +
  labs(title="Points Scored in Regular Season vs. Playoffs", x="Game Type", y="Points Scored") +
  theme_minimal()
```

OUTPUT:



ANALYSIS:

Curry's scoring can be different in high-stake games whereas in playoffs there is a lot of variability and we can see both higher and lower performances based on the condition of the game and opponent strength.

Curry scores less in conference quarter finals or semi-finals as the opponent team form their strategies around Curry's gameplay to limit his effectiveness on the court.

INSIGHTS GAINED FROM THE ANALYSIS –

- **Impact of Game Context**: Stephen Curry's scoring varies with game context, especially in high-stakes playoff matches. While he excels under pressure, his performance fluctuates with game intensity. This insight can guide coaches to develop strategies to support peak performance in critical moments.
- **Opponent Strength and Defensive Strategies**: Curry's scoring is significantly impacted by stronger opponents' defensive strategies. Teams focusing on limiting his effectiveness reduce his output, highlighting the need for adaptive game plans to counteract specific defenses.
- **Physical and Mental Factors**: Injuries, fatigue, and overall health affect Curry's performance, with observed scoring drops during certain periods. This underscores the importance of maintaining optimal physical and mental conditions for consistent high-level performance throughout the season.