APPLIED STATISTICAL ANALYSIS LAB

NAME: ADITI KULKARNI

ROLL NO: 55

YEAR: SY

DIVISION: E (E3)

SRN NO: 202201893

ASSIGNMENT 7

STATEMENT: *To use One-Way ANOVA with post hoc tests to explore the relationship between several variables Using the preexisting data file.*

THEORY:

1. Data Import and Inspection:

- *data <- read.csv(file.choose(): This line reads your data from a user-selected CSV file and stores it in the 'data' variable.*
- *View(data): This command opens the data in a data viewer for manual inspection.*
- *dim(data): It displays the dimensions of the data (number of rows and columns).*
- *str(data): Provides the structure of the data, including variable types.*

2. One-Way ANOVA:

- *anova_result <- aov(SLpM ~ stance, data = data): This line performs a One-Way ANOVA. It analyzes the relationship between the 'SLpM' (Strikes Landed per Minute) variable and the 'stance' variable. 'SLpM' is treated as the dependent variable, while 'stance' is the independent variable.*

3. ANOVA Summary:

- *anova_summary <- summary(anova_result): This command summarizes the results of the ANOVA analysis. It includes statistics such as the F-statistic, p-value, and group means.*

4. Post Hoc Test:

- *posthoc <- TukeyHSD(anova_result): This line performs Tukey's HSD (Honestly Significant Difference) post hoc test. It's used to compare the means of different 'stance' groups and identify significant differences.*

5. Visualization:

- *library(ggplot2): Loads the 'ggplot2' library for data visualization.*
- *p <- ggplot(data, aes(x = stance, y = SLpM)) + ...: This code creates a box plot using 'ggplot2' to visualize the distribution of 'SLpM' values across different 'stance' categories.*
- *geom_boxplot(fill = "lightsteelblue1"): This part specifies that the data should be represented as box plots with a light blue fill.*
- *labs(...): It adds labels and titles to the plot for clarity.*

6. Display:

- *print(p): This command displays the box plot, allowing you to visualize how 'SLpM' varies among different stances.*

SOURCE CODE:

```
data <- read.csv(file.choose())

View(data)

dim(data)

str(data)

# Perform One-Way ANOVA to analyze the relationship between 'stance' and 'SLpM'

anova_result <- aov(SLpM ~ stance, data = data)


# Check ANOVA summary

anova_summary <- summary(anova_result)

print(anova_summary)



library(dplyr)

library(stats)

library(multcomp)


# Perform Tukey's HSD post hoc test

posthoc <- TukeyHSD(anova_result)

print(posthoc)


# Visualize the results with box plots

library(ggplot2)

p <- ggplot(data, aes(x = stance, y = SLpM)) +

  geom_boxplot(fill = "lightsteelblue1") +
```

```
  labs(
    title = "Box Plot of SLpM by Stance",
    x = "Stance",
    y = "SLpM"
  )

# Display the box plots
print(p)
```

OUTPUT:

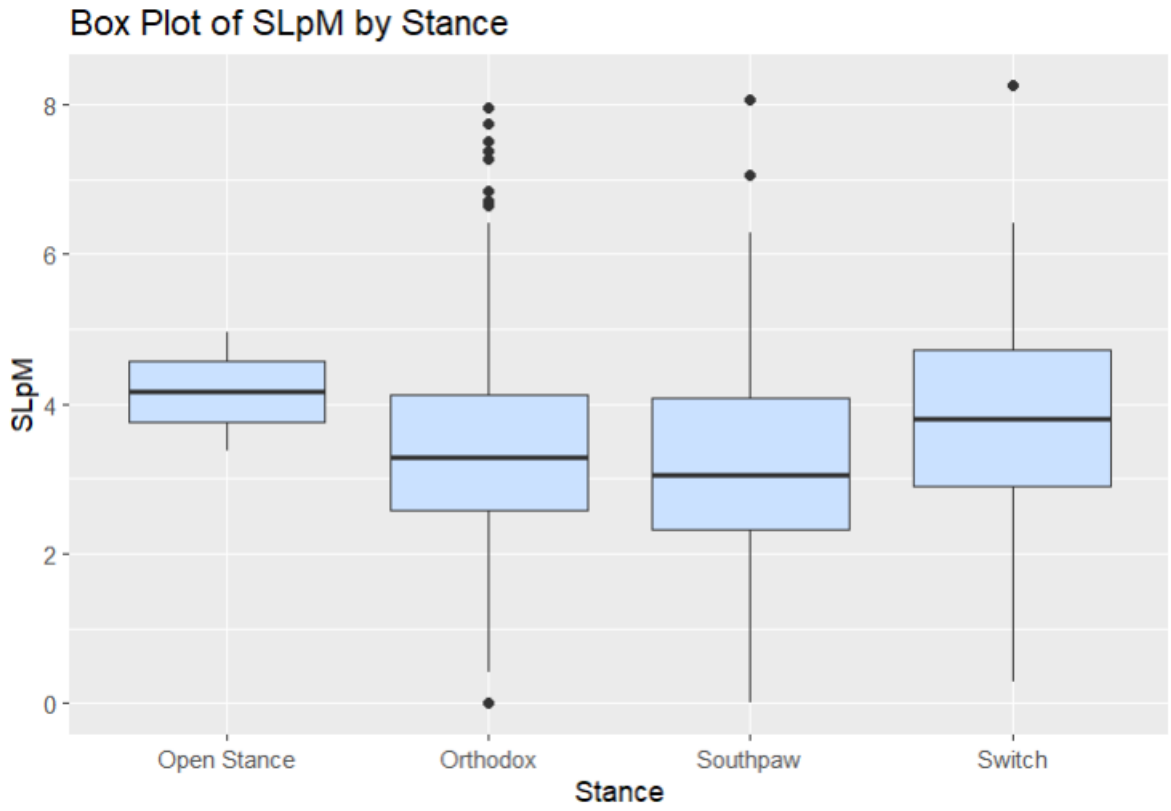
```
> dim(data)
[1] 632   9
> str(data)
'data.frame':   632 obs. of  9 variables:
 $ X            : int  2 3 14 26 27 42 46 49 50 52 ...
 $ name         : chr  "David" "Shamil" "Israel" "Omari" ...
 $ last_name    : chr  "Abbott" "Abdurakhimov" "Adesanya" "Akhmedov" ...
 $ stance       : chr  "Switch" "Orthodox" "Switch" "Orthodox" ...
 $ wins         : int  10 20 22 21 14 35 13 31 10 35 ...
 $ losses       : int  15 6 1 6 6 10 6 7 4 17 ...
 $ SLpM         : num  1.35 2.53 3.84 2.97 3.1 2.72 5.61 3.61 4.05 2.1 ...
 $ fights       : int  25 26 23 27 20 45 19 38 14 52 ...
 $ tko_win_ratio: num  0.62 0.4 0.36 0.22 0.36 0.36 0.33 0.52 0 0 ...
> # Perform One-Way ANOVA to analyze the relationship between 'stance' and 'SLpM'
> anova_result <- aov(SLpM ~ stance, data = data)
>
> # Check ANOVA summary
> anova_summary <- summary(anova_result)
> print(anova_summary)
             Df Sum Sq Mean Sq F value Pr(>F)
stance        3      9   3.003    1.84  0.139
Residuals   628   1025   1.632
```

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = SLpM ~ stance, data = data)

$stance
                           diff        lwr       upr     p adj
Orthodox-Open Stance -0.7694885 -3.1013035 1.5623265 0.8304027
Southpaw-Open Stance -0.9268110 -3.2720241 1.4184020 0.7389709
Switch-Open Stance   -0.3233333 -2.7453123 2.0986456 0.9860104
Southpaw-Orthodox    -0.1573225 -0.4857738 0.1711288 0.6055346
Switch-Orthodox       0.4461552 -0.2422037 1.1345141 0.3407280
Switch-Southpaw       0.6034777 -0.1289847 1.3359400 0.1471112

>
```



CONCLUSION:

The provided code conducts a One-Way Analysis of Variance (ANOVA) and post hoc tests to explore the relationship between the 'SLpM' (Strikes Landed per Minute) variable and the 'stance' variable in a dataset. It then visualizes the results using box plots. The ANOVA assesses whether there are significant differences in 'SLpM' across different stances. If significant differences are found, the post hoc tests identify which stance groups differ from each other. The box plots provide a clear visual representation of these differences. This code is a useful tool for analyzing and visualizing the impact of fighter stances on striking performance in the context of mixed martial arts or a related field.