A new version of the app was released in August 2016. Anand wants to understand which month in the given time frame after the launch of the new version; the mean usage pattern would start to show a statistically significant shift.

Solution:-
We need to find, Is there any impact in the app after launch of new version?

```
# Lets read the data with Clipboard function
data5 <- read.csv("clipboard", sep = "\t", header = T)

#check the structure of data (suggested to do on every time)
str(data5)

# Install dplyr package and run library command
install.packages("dplyr")
library(dplyr)

#Install rlang package and run the library command
install.packages("rlang")
library(rlang)

#Install ggplot2 package and run library command
install.packages("ggplot2")
library(ggplot2)

# In order to proceed further, we have to chage the given date format to Year – Month - Date
data5$Month.Year <- as.Date(data5$Month.Year, format = "%Y-%m-%d")
str(data5$Date)

#Lets take a copy data and filter Month, Year and Usage
data6 <- data5 %>% filter(Month.Year >= "2016-08-01") %>%
  select(Month.Year, Usage)


# Group the Data to proceed with analysis
grouped_data <- data5 %>% group_by(Month.Year) %>%
summarize(Usage = sum(Usage))

# Lets filter the data after grouping using ifelse function.
grouped_data$group <- factor(ifelse(grouped_data$Month.Year >= "2016-08-01", "After",
"Before"))

#Lets do a Two Sample T Test to get the p value
t.test(Usage~group, data = grouped_data)
Welch Two Sample t-test
```

data:  Usage by group
t = 3.7775,
df = 25.647,
p-value = 0.0008473

alternative hypothesis: true difference in means between group After and group Before is not equal to 0
95 percent confidence interval:
  716.9608 2431.1144
sample estimates:
 mean in group After mean in group Before
        2838.895          1264.857

With this Welch test we got the result that,  before Aug 2016 and After Aug 2016 there is difference in the mean. It shows, there is a spike after new launch of App.  To find the in which on wards the spike happened, we have to test with line graph and have to use ggplot2 package.

```
###########
# we have to plot a line graph to see from which month is the shift happening.
rdate <- as.Date(data5$Month.Year, "%m-%d-%Y")
rdate

plot(data5$Usage~data5$Month.Year, type = "l", col = "red")

# Install the ggplot2 package and run library command
install.packages("ggplot2")
library(ggplot2)

ggplot(data5, aes(x = Month.Year, y = Usage)) + geom_line() +
  labs(title = "Time Series Chart - Weekly Data", x = "Week", y = "Usage")

#Creating a subset from October 2016 onwards

data7 <- data5[1:61,]
data8 <- data5[62:123,]

data7$dummy <- "before"
data8$dummy <- "after"

newdata<- rbind(data7, data8)

t.test(newdata$Usage~newdata$dummy, conf.level = 0.95, paired = FALSE, var.eq = TRUE)
```
***The Result:-***

Two Sample t-test

data:  newdata$Usage by newdata$dummy
t = 6.1002,
df = 121,
p-value = 1.311e-08

alternative hypothesis: true difference in means between group after and group before is not equal to 0
95 percent confidence interval:
 315.8774 619.4203
sample estimates:
mean in group after
        814.4194
mean in group before
        346.7705

## *The Conclusion:-*
-- The calculated t-statistic is more than the critical value of t; and thus, we reject the null hypothesis and conclude that the app usage is statistically significantly different before and after October 2016. So, after the new release in August 2016, the usage of JAT's app increased, starting from October 2016.