```
#Web Data Analysis

getwd()

setwd("~/Web Data Analysis")

install.packages("readxl","gdata")

library(readxl)

library(gdata)

internet_data<- read_xlsx("internet_dataset.xlsx")

View(internet_data)

#Analysis Tasks:
# 1. The team wants to analyze each variable of the data collected
through data
#   summarization to get a basic understanding of the dataset and to
prepare for
#   further analysis.

str(internet_data)

summary(internet_data)


#2. As mentioned earlier, a unique page view represents the number of
sessions during
#   which that page was viewed one or more times. A visit counts all
instances, no matter
#   how many times the same visitor may have been to your site. So the
team needs to
#   know whether the unique page view value depends on visits.

cor(internet_data$Uniquepageviews,internet_data$Visits)
ano<- aov(Uniquepageviews~Visits,data= internet_data)
ano
summary(ano)


#3. Find out the probable factors from the dataset, which could affect
the exits.
# Exit Page Analysis is usually required to get an idea about why a
user leaves
# the website for a session and moves on to another one. Please keep in
mind that
# exits should not be confused with bounces.
```

```
ano_exits<- aov(Exits~.,data = internet_data)
ano_exits
summary(ano_exits)


#4. Every site wants to increase the time on page for a visitor.
# This increases the chances of the visitor understanding the site
content
# better and hence there are more chances of a transaction taking
place.
# Find the variables which possibly have an effect on the time on page

ano_timeinpage<- aov(Timeinpage~.,data = internet_data)
ano_timeinpage
summary(ano_timeinpage)


#5. A high bounce rate is a cause of alarm for websites which depend on
visitor
# engagement. Help the team in determining the factors that are
impacting the bounce.

rmm<-
glm(BouncesNew~Timeinpage+Continent+Exits+Sourcegroup+Uniquepageviews+Vi
 = internet_data,family = "binomial")
rmm
summary(rmm)
```