Week 12 IP

Ted Askoye Samuel

1. Business Understanding

1 a.) Defining the Question

A Kenyan entrepreneur has created an online cryptography course and would want to advertise it on her blog. She currently targets audiences originating from various countries. In the past, she ran ads to advertise a related course on the same blog and collected data in the process. She has employed the services of Skoko Limited, a Data Science Consultancy to help her identify which individuals are most likely to click on her ads.

2. Defining the Metrics of Success

The success of this analysis will occur when the target audience is known as per the adverts.

3. Context

Advertising is everywhere online, but we've gotten pretty good at ignoring it. To win back our attention, advertisers have adapted to our digital viewing habits by remembering what we read and buy online, then using this information to sell us things they think we might like. Part of this strategy is Targeted advertising. Targeted Advertising is a form of online advertising that focuses on the specific traits, interests, and preferences of a consumer. Advertisers discover this information by tracking your activity on the Internet.

4. Experimental Design

We will define the question, the metric of success, context and experimental design taken. This will be followed by reading and exploring the dataset and its appropriateness of the available data to answer the given question. This will be followed by cleaning the data off outliers, anomalies and null values from missing data, perfom an exploratory data analysis after which we will record our observations and provide a conclusion and reccomendation.

5. Data Relevance

Our data is very relevant to our research question. The more you know about your audience, the better you'll be able to sell advertising to them. The dataset provided has relevant information about the blog's audience.

6. Loading relevant Libraries and Reading the Data

```r
# Importing the required packages

library("data.table")
```

```r
library("plyr")
library("dplyr")
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

## The following objects are masked from 'package:data.table':
##
##     between, first, last

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------------------------------- tidyverse 1.3.0

## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v stringr 1.4.0
## v tidyr   1.1.2     v forcats 0.5.0
## v readr   1.3.1

## -- Conflicts ------------------------------------------------------------------ tidyverse_conflicts()
## x dplyr::arrange()   masks plyr::arrange()
## x dplyr::between()   masks data.table::between()
## x purrr::compact()   masks plyr::compact()
## x dplyr::count()     masks plyr::count()
## x dplyr::failwith()  masks plyr::failwith()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::id()        masks plyr::id()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x dplyr::mutate()    masks plyr::mutate()
## x dplyr::rename()    masks plyr::rename()
## x dplyr::summarise() masks plyr::summarise()
## x dplyr::summarize() masks plyr::summarize()
## x purrr::transpose() masks data.table::transpose()
```

```r
library("tidyr")
library("lubridate")
```

```
##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:data.table':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday, week,
##     yday, year

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
library("ggcorrplot")
library("ggplot2")
library("corrplot")
```

```
## corrplot 0.84 loaded
```

```r
library("moments")
library('psych')
```

```
##
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```r
library('countrycode')
library('class')
library("rpart")
library("rpart.plot")
library("mlbench")
library('e1071')
```

```
##
## Attaching package: 'e1071'

## The following objects are masked from 'package:moments':
##
##     kurtosis, moment, skewness
```

```r
library('rpart')
library('caret')
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
##     lift
```

```r
library('ranger')
library('kernlab')
```

```
##
## Attaching package: 'kernlab'

## The following object is masked from 'package:psych':
##
##     alpha

## The following object is masked from 'package:purrr':
##
##     cross

## The following object is masked from 'package:ggplot2':
##
##     alpha
```

```r
library('pdp')
```

```
##
## Attaching package: 'pdp'

## The following object is masked from 'package:purrr':
##
##     partial
```

```r
library('vip')
```

```
##
## Attaching package: 'vip'

## The following object is masked from 'package:utils':
##
##     vi
```

```r
# Loading the Dataset

ad_df <- read.csv(url("http://bit.ly/IPAdvertisingData"))
```

Previewing the data

```r
# Previewing The First Seven records in the Dataset

head(ad_df, n=7)
```

```
##   Daily.Time.Spent.on.Site Age Area.Income Daily.Internet.Usage
## 1                    68.95  35    61833.90               256.09
## 2                    80.23  31    68441.85               193.77
## 3                    69.47  26    59785.94               236.50
## 4                    74.15  29    54806.18               245.89
## 5                    68.37  35    73889.99               225.58
## 6                    59.99  23    59761.56               226.74
## 7                    88.91  33    53852.85               208.36
##                          Ad.Topic.Line           City Male    Country
## 1    Cloned 5thgeneration orchestration    Wrightburgh    0    Tunisia
## 2    Monitored national standardization      West Jodi    1      Nauru
## 3      Organic bottom-line service-desk       Davidton    0 San Marino
## 4 Triple-buffered reciprocal time-frame West Terrifurt    1      Italy
## 5         Robust logistical utilization   South Manuel    0    Iceland
## 6       Sharable client-driven software      Jamieberg    1     Norway
## 7            Enhanced dedicated support    Brandonstad    0    Myanmar
##             Timestamp Clicked.on.Ad
## 1 2016-03-27 00:53:11             0
## 2 2016-04-04 01:39:02             0
## 3 2016-03-13 20:35:42             0
## 4 2016-01-10 02:31:19             0
## 5 2016-06-03 03:36:18             0
## 6 2016-05-19 14:30:17             0
## 7 2016-01-28 20:59:32             0
```

```r
# Previewing The Last Seven records in the Dataset

tail(ad_df, n=7)
```

```
##      Daily.Time.Spent.on.Site Age Area.Income Daily.Internet.Usage
## 994                     64.20  27    66200.96               227.63
## 995                     43.70  28    63126.96               173.01
## 996                     72.97  30    71384.57               208.58
## 997                     51.30  45    67782.17               134.42
## 998                     51.63  51    42415.72               120.37
## 999                     55.55  19    41920.79               187.95
## 1000                    45.01  26    29875.80               178.35
##                              Ad.Topic.Line           City Male
## 994        Phased zero tolerance extranet  Edwardsmouth    1
## 995         Front-line bifurcated ability  Nicholasland    0
## 996         Fundamental modular algorithm     Duffystad    1
## 997       Grass-roots cohesive monitoring   New Darlene    1
## 998          Expanded intangible solution South Jessica    1
## 999 Proactive bandwidth-monitored policy    West Steven    0
## 1000      Virtual 5thgeneration emulation    Ronniemouth    0
##                   Country           Timestamp Clicked.on.Ad
## 994           Isle of Man 2016-02-11 23:45:01             0
## 995               Mayotte 2016-04-04 03:57:48             1
```

```
## 996                 Lebanon 2016-02-11 21:49:00             1
## 997  Bosnia and Herzegovina 2016-04-22 02:07:01             1
## 998                Mongolia 2016-02-01 17:24:57             1
## 999               Guatemala 2016-03-24 02:35:54             0
## 1000                 Brazil 2016-06-03 21:43:21             1
```

```r
# Checking the Data Dimensions

dim(ad_df)
```

```
## [1] 1000   10
```

The dataset has 1000 records and 10 columns

```r
# Checking the Structure of the Dataset

str(ad_df)
```

```
## 'data.frame':    1000 obs. of  10 variables:
##  $ Daily.Time.Spent.on.Site: num  69 80.2 69.5 74.2 68.4 ...
##  $ Age                     : int  35 31 26 29 35 23 33 48 30 20 ...
##  $ Area.Income             : num  61834 68442 59786 54806 73890 ...
##  $ Daily.Internet.Usage    : num  256 194 236 246 226 ...
##  $ Ad.Topic.Line           : chr  "Cloned 5thgeneration orchestration" "Monitored national standardi
##  $ City                    : chr  "Wrightburgh" "West Jodi" "Davidton" "West Terrifurt" ...
##  $ Male                    : int  0 1 0 1 0 1 0 1 1 1 ...
##  $ Country                 : chr  "Tunisia" "Nauru" "San Marino" "Italy" ...
##  $ Timestamp               : chr  "2016-03-27 00:53:11" "2016-04-04 01:39:02" "2016-03-13 20:35:42"
##  $ Clicked.on.Ad           : int  0 0 0 0 0 0 0 1 0 0 ...
```

```r
# Checking The Data present in each column

glimpse(ad_df)
```

```
## Rows: 1,000
## Columns: 10
## $ Daily.Time.Spent.on.Site <dbl> 68.95, 80.23, 69.47, 74.15, 68.37, 59.99, ...
## $ Age                      <int> 35, 31, 26, 29, 35, 23, 33, 48, 30, 20, 49...
## $ Area.Income              <dbl> 61833.90, 68441.85, 59785.94, 54806.18, 73...
## $ Daily.Internet.Usage     <dbl> 256.09, 193.77, 236.50, 245.89, 225.58, 22...
## $ Ad.Topic.Line            <chr> "Cloned 5thgeneration orchestration", "Mon...
## $ City                     <chr> "Wrightburgh", "West Jodi", "Davidton", "W...
## $ Male                     <int> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ Country                  <chr> "Tunisia", "Nauru", "San Marino", "Italy",...
## $ Timestamp                <chr> "2016-03-27 00:53:11", "2016-04-04 01:39:0...
## $ Clicked.on.Ad            <int> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, ...
```

7. Data Preparation

Uniformity

```r
# Check column names

colnames(ad_df)
```

```
##  [1] "Daily.Time.Spent.on.Site" "Age"
##  [3] "Area.Income"              "Daily.Internet.Usage"
##  [5] "Ad.Topic.Line"            "City"
##  [7] "Male"                     "Country"
##  [9] "Timestamp"                "Clicked.on.Ad"
```

```r
# Renaming column names

names(ad_df)[1] <- "daily_time_spent_on_site"
names(ad_df)[2] <- "age"
names(ad_df)[3] <- "area_income"
names(ad_df)[4] <- "daily_internet_usage"
names(ad_df)[5] <- "ad_topic_line"
names(ad_df)[6] <- "city"
names(ad_df)[7] <- "male"
names(ad_df)[8] <- "country"
names(ad_df)[9] <- "timestamp"
names(ad_df)[10] <- "clicked_on_ad"

# Checking whether the column names have been changed

colnames(ad_df)
```

We'll rename the column names for Uniformity purposes

```
##  [1] "daily_time_spent_on_site" "age"
##  [3] "area_income"              "daily_internet_usage"
##  [5] "ad_topic_line"            "city"
##  [7] "male"                     "country"
##  [9] "timestamp"                "clicked_on_ad"
```

```r
# Checking for the length of unique values in each column

lapply(ad_df, function (x) {length(unique(x))})
```

```
## $daily_time_spent_on_site
## [1] 900
##
## $age
## [1] 43
```

```
##
## $area_income
## [1] 1000
##
## $daily_internet_usage
## [1] 966
##
## $ad_topic_line
## [1] 1000
##
## $city
## [1] 969
##
## $male
## [1] 2
##
## $country
## [1] 237
##
## $timestamp
## [1] 1000
##
## $clicked_on_ad
## [1] 2
```

We can observe that the 'Male'and 'Clicked_on_ad'columns are categorical since they only have 2 factor variables

Appropriateness

```r
# Converting timestamp column to datetime datatype

ad_df[["timestamp"]] <- as.POSIXct(ad_df$timestamp, tz=Sys.timezone())
str(ad_df)
```

```
## 'data.frame':    1000 obs. of  10 variables:
##  $ daily_time_spent_on_site: num  69 80.2 69.5 74.2 68.4 ...
##  $ age                     : int  35 31 26 29 35 23 33 48 30 20 ...
##  $ area_income             : num  61834 68442 59786 54806 73890 ...
##  $ daily_internet_usage    : num  256 194 236 246 226 ...
##  $ ad_topic_line           : chr  "Cloned 5thgeneration orchestration" "Monitored national standardi
##  $ city                    : chr  "Wrightburgh" "West Jodi" "Davidton" "West Terrifurt" ...
##  $ male                    : int  0 1 0 1 0 1 0 1 1 1 ...
##  $ country                 : chr  "Tunisia" "Nauru" "San Marino" "Italy" ...
##  $ timestamp               : POSIXct, format: "2016-03-27 00:53:11" "2016-04-04 01:39:02" ...
##  $ clicked_on_ad           : int  0 0 0 0 0 0 0 1 0 0 ...
```

```r
glimpse(ad_df)
```

```
## Rows: 1,000
```

```
## Columns: 10
## $ daily_time_spent_on_site <dbl> 68.95, 80.23, 69.47, 74.15, 68.37, 59.99, ...
## $ age                      <int> 35, 31, 26, 29, 35, 23, 33, 48, 30, 20, 49...
## $ area_income              <dbl> 61833.90, 68441.85, 59785.94, 54806.18, 73...
## $ daily_internet_usage     <dbl> 256.09, 193.77, 236.50, 245.89, 225.58, 22...
## $ ad_topic_line            <chr> "Cloned 5thgeneration orchestration", "Mon...
## $ city                     <chr> "Wrightburgh", "West Jodi", "Davidton", "W...
## $ male                     <int> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ country                  <chr> "Tunisia", "Nauru", "San Marino", "Italy",...
## $ timestamp                <dttm> 2016-03-27 00:53:11, 2016-04-04 01:39:02,...
## $ clicked_on_ad            <int> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, ...
```

We can observe that the change has taken shape successfully. We now want to split the column to date and time

```r
# Splitting datetime into date and time

Time <- format(as.POSIXct(strptime(ad_df$timestamp,"%Y-%m-%d %H:%M:%S",tz="")) ,format = "%H:%M:%S")
head(Time)
```

```
## [1] "00:53:11" "01:39:02" "20:35:42" "02:31:19" "03:36:18" "14:30:17"
```

```r
Dates <- format(as.POSIXct(strptime(ad_df$timestamp,"%Y-%m-%d %H:%M:%S",tz="")) ,format = "%Y-%m-%d")
head(Dates)
```

```
## [1] "2016-03-27" "2016-04-04" "2016-03-13" "2016-01-10" "2016-06-03"
## [6] "2016-05-19"
```

```r
ad_df$Dates <- Dates
ad_df$Time <- Time

str(ad_df)
```

```
## 'data.frame':    1000 obs. of  12 variables:
##  $ daily_time_spent_on_site: num  69 80.2 69.5 74.2 68.4 ...
##  $ age                     : int  35 31 26 29 35 23 33 48 30 20 ...
##  $ area_income             : num  61834 68442 59786 54806 73890 ...
##  $ daily_internet_usage    : num  256 194 236 246 226 ...
##  $ ad_topic_line           : chr  "Cloned 5thgeneration orchestration" "Monitored national standardi
##  $ city                    : chr  "Wrightburgh" "West Jodi" "Davidton" "West Terrifurt" ...
##  $ male                    : int  0 1 0 1 0 1 0 1 1 1 ...
##  $ country                 : chr  "Tunisia" "Nauru" "San Marino" "Italy" ...
##  $ timestamp               : POSIXct, format: "2016-03-27 00:53:11" "2016-04-04 01:39:02" ...
##  $ clicked_on_ad           : int  0 0 0 0 0 0 0 1 0 0 ...
##  $ Dates                   : chr  "2016-03-27" "2016-04-04" "2016-03-13" "2016-01-10" ...
##  $ Time                    : chr  "00:53:11" "01:39:02" "20:35:42" "02:31:19" ...
```

```r
# Separating dates to hours minutes and days and dropping the timestamp column

ad_df <- separate(ad_df, "Dates", c("year", "month", "day"), sep = "-")
```

```r
ad_df <- separate(ad_df, "Time", c("hour", "minutes", "seconds"), sep = ":")


colnames(ad_df)
```

```
##  [1] "daily_time_spent_on_site" "age"
##  [3] "area_income"              "daily_internet_usage"
##  [5] "ad_topic_line"            "city"
##  [7] "male"                     "country"
##  [9] "timestamp"                "clicked_on_ad"
## [11] "year"                     "month"
## [13] "day"                      "hour"
## [15] "minutes"                  "seconds"
```

```r
# Changing the new derived columns to factors for ease of analysis


ad_df$Male = factor(ad_df$male)
ad_df$Year = factor(ad_df$year)
ad_df$Month = factor(ad_df$month)
ad_df$Day = factor(ad_df$day)
ad_df$Hour = factor(ad_df$hour)
ad_df$Minutes = factor(ad_df$minutes)
ad_df$Seconds = factor(ad_df$seconds)
```

We can see that the date and time have their respective columns

#——- ## Completeness

```r
# Checking for missing values

colSums(is.na(ad_df))
```

```
## daily_time_spent_on_site                      age              area_income
##                        0                        0                        0
##     daily_internet_usage            ad_topic_line                     city
##                        0                        0                        0
##                     male                  country                timestamp
##                        0                        0                        0
##            clicked_on_ad                     year                    month
##                        0                        0                        0
##                      day                     hour                  minutes
##                        0                        0                        0
##                  seconds                     Male                     Year
##                        0                        0                        0
##                    Month                      Day                     Hour
##                        0                        0                        0
##                  Minutes                  Seconds
##                        0                        0
```

Our data is complete hence no missing values

#——- ## Consistency

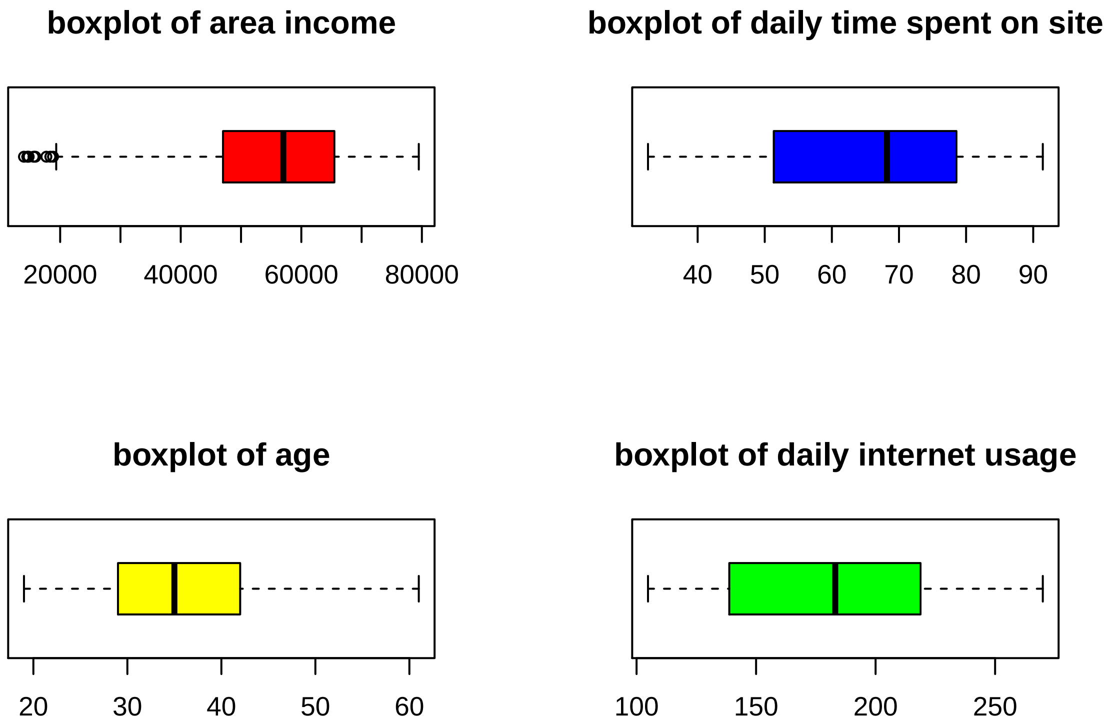
```
# Checking for duplicate values

duplicates <- ad_df[duplicated(ad_df),]
duplicates
```

```
##  [1] daily_time_spent_on_site age                      area_income
##  [4] daily_internet_usage     ad_topic_line            city
##  [7] male                     country                  timestamp
## [10] clicked_on_ad            year                     month
## [13] day                      hour                     minutes
## [16] seconds                  Male                     Year
## [19] Month                    Day                      Hour
## [22] Minutes                  Seconds
## <0 rows> (or 0-length row.names)
```

Our data is consistent due to no duplicate values present #——- ### Anomaly Detection #### # Checking for anomalies in our numerical variables i.e daily_time_spent_on_site, area income, age, and daily_internet usage

Boxplots

```
# Plotting boxplots for all the numerical variables

par(mfrow=c(2,2))
boxplot((ad_df$`area_income`), horizontal = TRUE, col = 'red', main = "boxplot of area income")
boxplot((ad_df$`daily_time_spent_on_site`), horizontal = TRUE, col = 'blue', main = "boxplot of daily ti
boxplot((ad_df$`age`), horizontal = TRUE, col = 'yellow', main = "boxplot of age")
boxplot((ad_df$`daily_internet_usage`), horizontal = TRUE, col = 'green', main = "boxplot of daily inte
```



1. **Area income variable has values ranging from below 0 to 80,000. We have a few values below 20,000 which are outliers but we'll keep them because they represent crucial data for analysis**

2. **Daily time spent on site has values from around 20 to 90 with the mode between 50 to 80**

3. **Age variable has observations from the age of 20 to 60 with the mode between 30 to 40**

4. **Daily internet usage has values from 100 to slightly above 250 with the mode between 150 to 200**

---

8. Exploratory Data Analysis

Univariate Analysis

```
# Checking the statistical summary of the data

summary(ad_df)
```

```
##  daily_time_spent_on_site      age          area_income    daily_internet_usage
##  Min.   :32.60              Min.   :19.00   Min.   :13996   Min.   :104.8
##  1st Qu.:51.36              1st Qu.:29.00   1st Qu.:47032   1st Qu.:138.8
```

```
##  Median :68.22           Median :35.00   Median :57012   Median :183.1
##  Mean   :65.00           Mean   :36.01   Mean   :55000   Mean   :180.0
##  3rd Qu.:78.55           3rd Qu.:42.00   3rd Qu.:65471   3rd Qu.:218.8
##  Max.   :91.43           Max.   :61.00   Max.   :79485   Max.   :270.0
##
##  ad_topic_line          city                male         country
##  Length:1000        Length:1000        Min.   :0.000   Length:1000
##  Class :character   Class :character   1st Qu.:0.000   Class :character
##  Mode  :character   Mode  :character   Median :0.000   Mode  :character
##                                        Mean   :0.481
##                                        3rd Qu.:1.000
##                                        Max.   :1.000
##
##    timestamp                    clicked_on_ad     year
##  Min.   :2016-01-01 02:52:10   Min.   :0.0   Length:1000
##  1st Qu.:2016-02-18 02:55:42   1st Qu.:0.0   Class :character
##  Median :2016-04-07 17:27:29   Median :0.5   Mode  :character
##  Mean   :2016-04-10 10:34:06   Mean   :0.5
##  3rd Qu.:2016-05-31 03:18:14   3rd Qu.:1.0
##  Max.   :2016-07-24 00:22:16   Max.   :1.0
##
##     month               day                hour             minutes
##  Length:1000        Length:1000        Length:1000        Length:1000
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    seconds           Male      Year      Month         Day           Hour
##  Length:1000        0:519   2016:1000   01:147   03     : 46   07     : 54
##  Class :character   1:481               02:160   17     : 42   20     : 50
##  Mode  :character                       03:156   15     : 41   09     : 49
##                                         04:147   10     : 37   21     : 48
##                                         05:147   04     : 36   00     : 45
##                                         06:142   26     : 36   05     : 44
##                                         07:101   (Other):762   (Other):710
##     Minutes       Seconds
##  02     : 26   22     : 28
##  07     : 24   10     : 27
##  13     : 24   35     : 27
##  10     : 22   37     : 27
##  21     : 21   38     : 24
##  33     : 21   15     : 23
##  (Other):862   (Other):844
```

The timestamp has a conflicting datatype compared to what its normal date/time format as well as gender and and clicked on ad datatypes which should be categorical instead of integers

The daily time spent on the site seems to be in minutes and seconds ranging from 32.60 to 91.43. The values are likely to be close to normally distributed as the median is 68.22 and the mean is 65.

The area income are not likely to be close to normally distributed due to a large difference in ranges i.e from 13996 to 79485, with a median of 57012 and a mean of 55000.

The daily internet usage ranges from 104.8 to 270.0, with a median of 183.1 and a mean of 180.0. The values are likely to be close to normally distributed.

The ad topic line, City, male, Country are categrocial features, with a different value for each record.

The feature male is categorical (binary) with a mean of 0.481, which means there are more records from individuals that are female.

The clicked on ad variable is categorical (binary) with a mean of 0.5, which means that the variable of interest is balanced in this dataset.

#——

Measures of Central Tendancy and Dispersion - Summary

Central Tendancy - Mode, Mean and Median

```
# First, a function for mode will be created since R does not have a built in function.

getmode <- function(v) {
    uniqv <- unique(v)
    uniqv[which.max(tabulate(match(v, uniqv)))]
}
```

```
# City
# This column represents the city where the most users are from
mode.city <- getmode(ad_df$city)
mode.city
```

```
## [1] "Lisamouth"
```

```
# Country
# This column represents the country where the most users are from
mode.country <- getmode(ad_df$country)
mode.country
```

```
## [1] "Czech Republic"
```

```r
# Age
# This column represents the Age That most users are, its mean and median
mode.age <- getmode(ad_df$age)
mode.age
```

```
## [1] 31
```

```r
mean(ad_df$age)
```

```
## [1] 36.009
```

```r
median(ad_df$age)
```

```
## [1] 35
```

```r
# Daily Internet Usage
# This column represents the daily internet usage for most users, its mean and median
mode.usage <- getmode(ad_df$daily_internet_usage)
mode.usage
```

```
## [1] 167.22
```

```r
mean(ad_df$daily_internet_usage)
```

```
## [1] 180.0001
```

```r
median(ad_df$daily_internet_usage)
```

```
## [1] 183.13
```

```r
# Area Income
# This column represents most of the Area Income
mode.income <- getmode(ad_df$area_income)
mode.income
```

```
## [1] 61833.9
```

```r
mean(ad_df$area_income)
```

```
## [1] 55000
```

```r
median(ad_df$area_income)
```

```
## [1] 57012.3
```

```r
# Male
# This column represents gender with the most users
mode.male <- getmode(ad_df$male)
mode.male
```

```
## [1] 0
```

```r
# Ad_Topic_line
# This column represents most advertisement topic line

mode.adline<-getmode(ad_df$ad_topic_line)
mode.adline
```

```
## [1] "Cloned 5thgeneration orchestration"
```

```r
# Daily_Time_Spent
# This column represents most frequent daily time spent on site

mode.time <- getmode(ad_df$daily_time_spent_on_site)
mode.time
```

```
## [1] 62.26
```

```r
mean(ad_df$daily_time_spent_on_site)
```

```
## [1] 65.0002
```

```r
median(ad_df$daily_time_spent_on_site)
```

```
## [1] 68.215
```

```r
# Month
# This column represents most frequent months during usage

mode.month <- getmode(ad_df$month)
mode.month
```

```
## [1] "02"
```

```r
# Day
# This column represents most frequent day during usage

mode.day <- getmode(ad_df$day)
mode.day
```

```
## [1] "03"
```

```r
# Hour
# This column represents most frequent hour during usage

mode.hour <- getmode(ad_df$hour)
mode.hour
```

```
## [1] "07"
```

```r
# Minute
# This column represents most frequent Minutes during usage

mode.minutes <- getmode(ad_df$minutes)
mode.minutes
```

```
## [1] "02"
```

```r
# Seconds
# This column represents most frequent months during usage

mode.seconds <- getmode(ad_df$seconds)
mode.seconds
```

```
## [1] "22"
```

```r
# Age

sd.age <- sd(ad_df$age)
sd.age
```

Measure of Dispersion - Standard Deviation, Variance, Skewness, Kurtosis and Range

```
## [1] 8.785562
```

```r
var.age <- var(ad_df$age)
var.age
```

```
## [1] 77.18611
```

```r
range.age <- range(ad_df$age)
range.age
```

```
## [1] 19 61
```

```r
skew.age <- skewness(ad_df$age)
skew.age
```

```
## [1] 0.4777052
```

```r
kurt.age <- kurtosis(ad_df$age)
kurt.age
```

```
## [1] -0.4097066
```

```r
# Daily Internet Usage

sd.daily_internet_usage <- sd(ad_df$daily_internet_usage)
sd.daily_internet_usage
```

```
## [1] 43.90234
```

```r
var.daily_internet_usage <- var(ad_df$daily_internet_usage)
var.daily_internet_usage
```

```
## [1] 1927.415
```

```r
range.daily_internet_usage <- range(ad_df$daily_internet_usage)
range.daily_internet_usage
```

```
## [1] 104.78 269.96
```

```r
skew.daily_internet_usage <- skewness(ad_df$daily_internet_usage)
skew.daily_internet_usage
```

```
## [1] -0.03343681
```

```r
kurt.daily_internet_usage <- kurtosis(ad_df$daily_internet_usage)
kurt.daily_internet_usage
```

```
## [1] -1.275752
```

```r
# Daily time spent on site

sd.daily_time_spent_on_site <- sd(ad_df$daily_time_spent_on_site)
sd.daily_time_spent_on_site
```

```
## [1] 15.85361
```

```r
var.daily_time_spent_on_site <- var(ad_df$daily_time_spent_on_site)
var.daily_time_spent_on_site
```

```
## [1] 251.3371
```

```r
range.daily_time_spent_on_site <- range(ad_df$daily_time_spent_on_site)
range.daily_time_spent_on_site
```

```
## [1] 32.60 91.43
```

```r
skew.daily_time_spent_on_site <- skewness(ad_df$daily_time_spent_on_site)
skew.daily_time_spent_on_site
```

```
## [1] -0.370646
```

```r
kurt.daily_time_spent_on_site <- kurtosis(ad_df$daily_time_spent_on_site)
kurt.daily_time_spent_on_site
```

```
## [1] -1.099864
```

```r
# Area Income

sd.area_income <- sd(ad_df$area_income)
sd.area_income
```

```
## [1] 13414.63
```

```r
var.area_income <- var(ad_df$area_income)
var.area_income
```

```
## [1] 179952406
```

```r
range.area_income <- range(ad_df$area_income)
range.area_income
```

```
## [1] 13996.5 79484.8
```

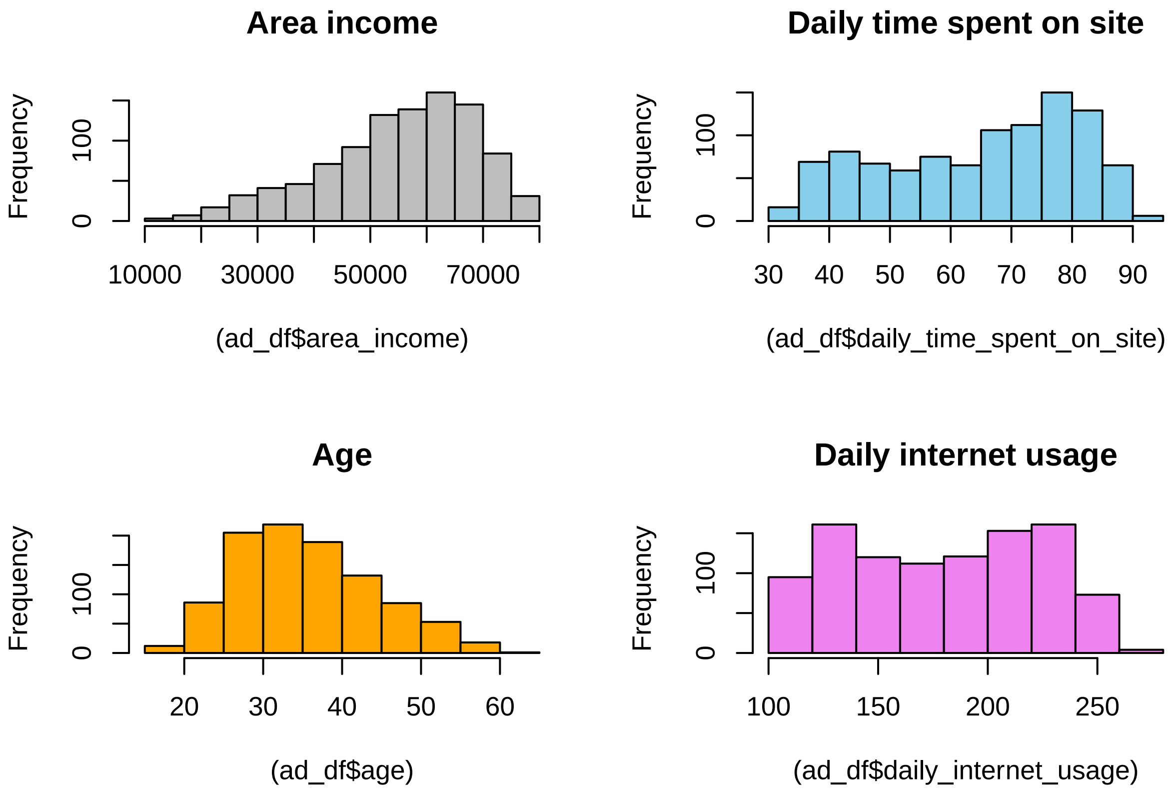
```r
skew.area_income <- skewness(ad_df$area_income)
skew.area_income
```

```
## [1] -0.6484229
```

```r
kurt.area_income <- kurtosis(ad_df$area_income)
kurt.area_income
```

```
## [1] -0.1110924
```

```r


# Plotting multiple histograms for Area income, Age, Daily time spent on site and Daily Internet Usage

par(mfrow=c(2,2))

hist((ad_df$`area_income`), col = 'grey', main = "Area income")
hist((ad_df$`daily_time_spent_on_site`), col = 'skyblue', main = "Daily time spent on site")
hist((ad_df$`age`), col = 'orange', main = "Age")
hist((ad_df$`daily_internet_usage`), col = 'violet', main = "Daily internet usage")
```



Histograms

Observations: ##### 1. Area income variable is negatively skewed as most of the observations recorded are lower compared to the high area income ##### 2. Age variable is positively skewed as most of the ages recorded are younger ##### 3. Daily internet usage and daily time spent on site are bimodal as they have an almost normal distribution

Bivariate Analysis

```
# Correlation Matrix
# Calling all the numerical data present

age<- ad_df$age
income<-ad_df$area_income
time<-ad_df$daily_time_spent_on_site
usage<-ad_df$daily_internet_usage

# Creating a new dataframe num with numerical data variables

num_data <- data.frame(age, income, time, usage)
head(num_data)
```

Correlation

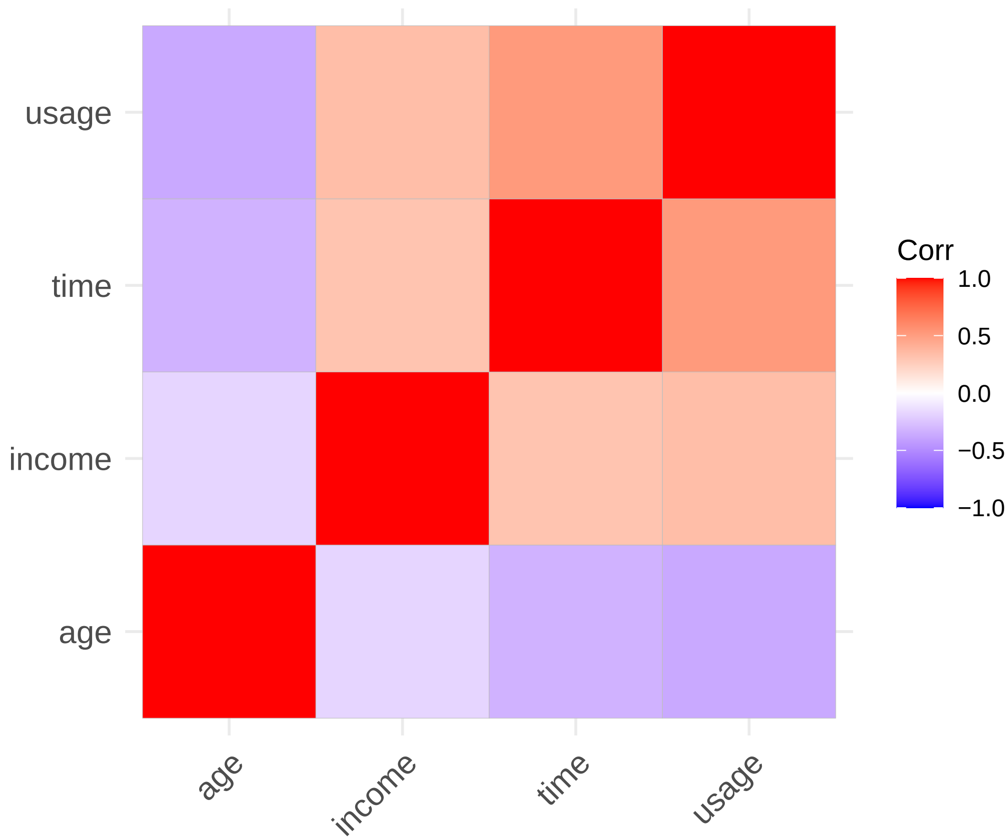
```
##   age   income  time  usage
## 1  35 61833.90 68.95 256.09
## 2  31 68441.85 80.23 193.77
## 3  26 59785.94 69.47 236.50
```

```
## 4  29 54806.18 74.15 245.89
## 5  35 73889.99 68.37 225.58
## 6  23 59761.56 59.99 226.74
```

```
# Correlation is a statistical technique that can show whether and how strongly pairs of variables are

# Calculating the correlation matrix

corr <- cor(num_data)
head(corr)
```

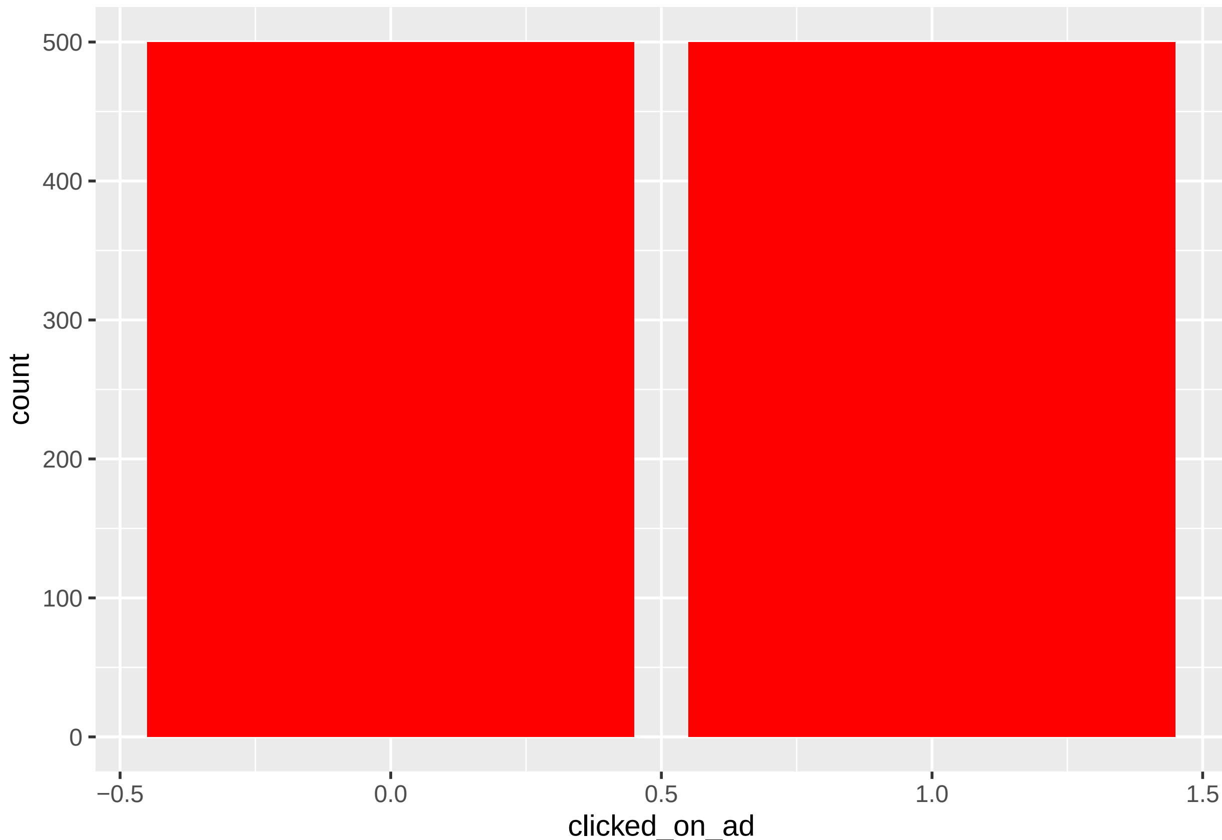
```
##               age     income       time      usage
## age     1.0000000 -0.1826050 -0.3315133 -0.3672086
## income -0.1826050  1.0000000  0.3109544  0.3374955
## time   -0.3315133  0.3109544  1.0000000  0.5186585
## usage  -0.3672086  0.3374955  0.5186585  1.0000000
```

```
# Plotting the correlation matrix

ggcorrplot(corr,hc.order = TRUE)
```



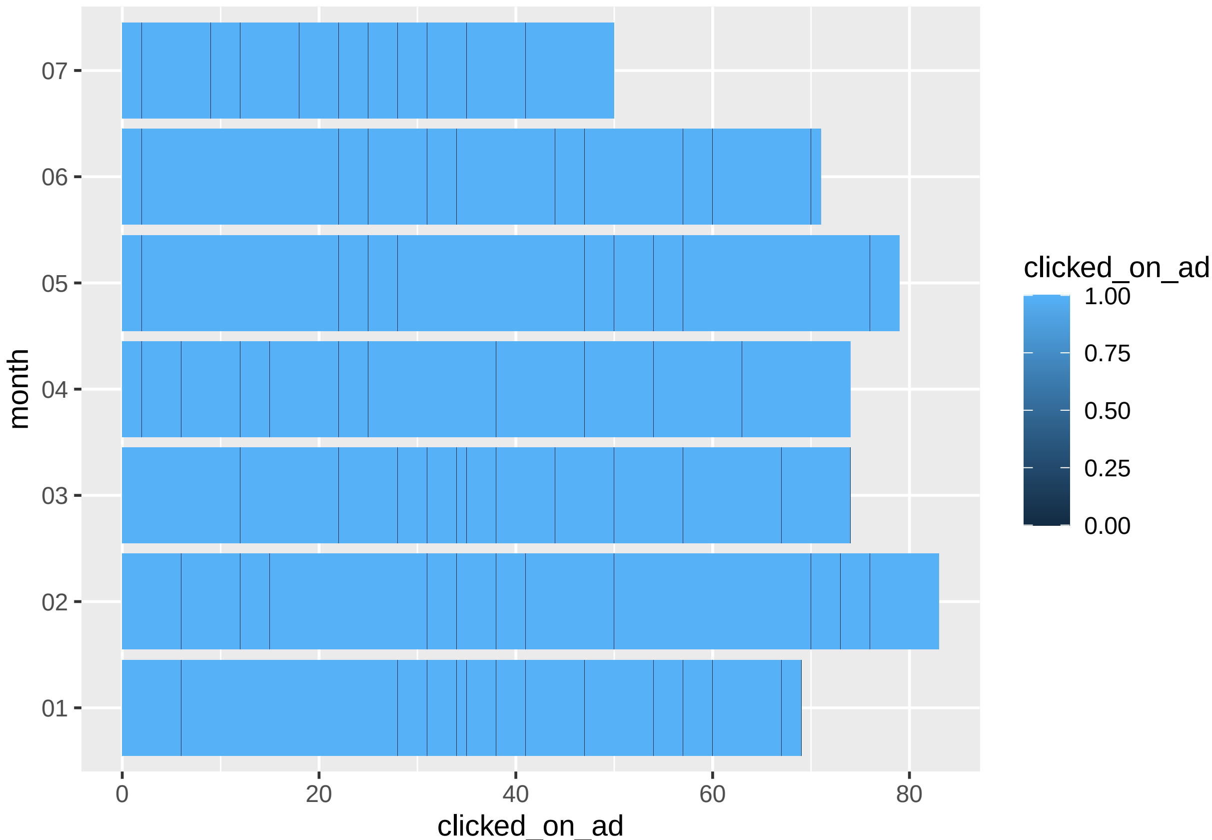
Observations ###### 1. Daily_internet_usage and Daily_time_spent_on_site seem to have a moderate positive correlation ###### 2. Daily_internet_usage and Age seem to have a negative correlation ###### 3. Area Income and Age are weakly correlated

```r
# Finding out and previewing the Number of clicked and no clicked ads

ggplot(ad_df, aes(clicked_on_ad)) + geom_bar(fill = "red")
```



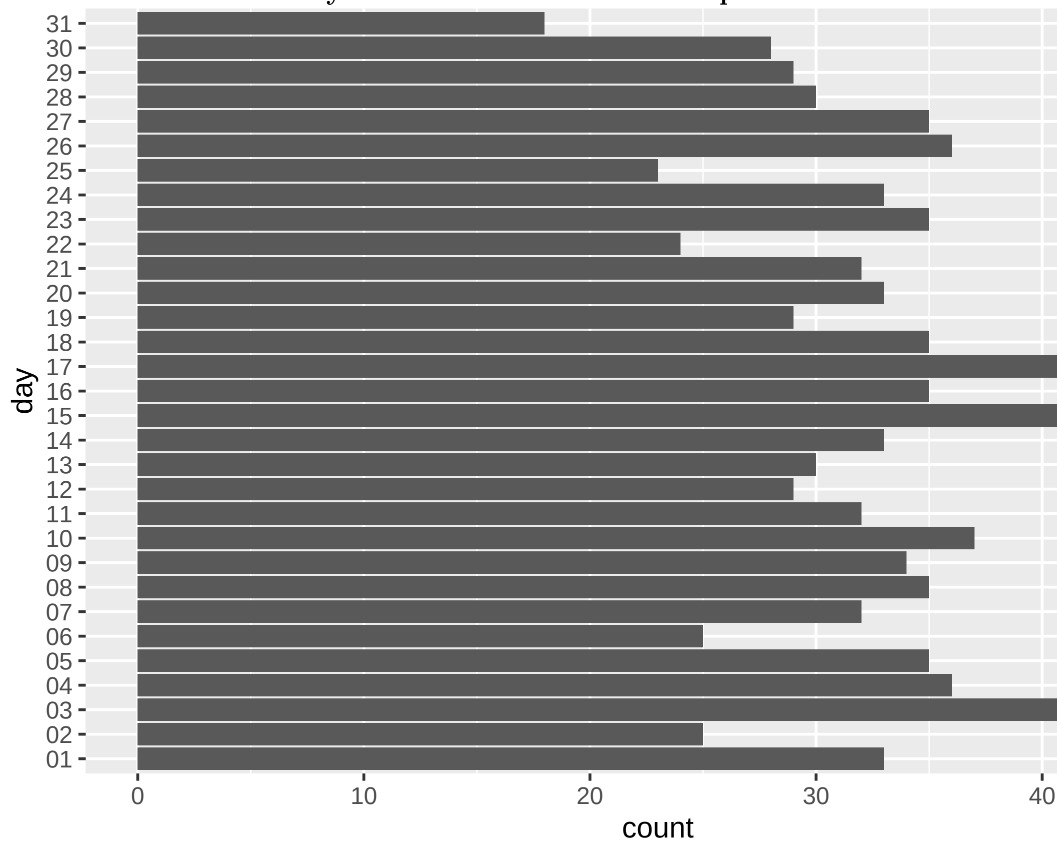
Barplots
The clicked ads and no clicked ads in our dataset were equal

```r
# Finding out and previewing the month with the most clicked ads

ggplot(ad_df, aes(x = `clicked_on_ad`, y = `month`)) + geom_col(aes(fill = `clicked_on_ad`))
```



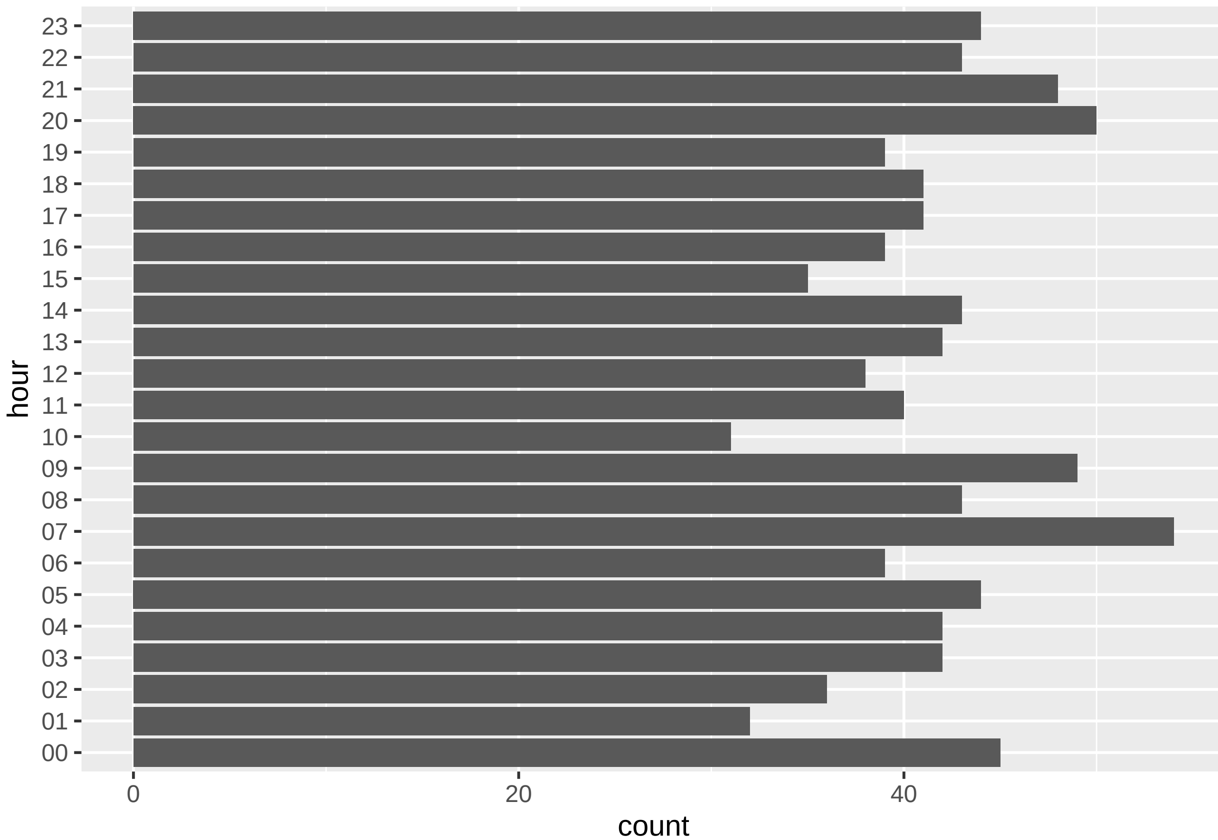

```r
# Finding out and previewing the day with the most clicked ads

ggplot(data = ad_df) +
  geom_bar(mapping = aes(y = day, fill = clicked_on_ad), position = "dodge")
```

February and May had the most clicked ads while July had the least. March and April had an equal number of clicked ads.



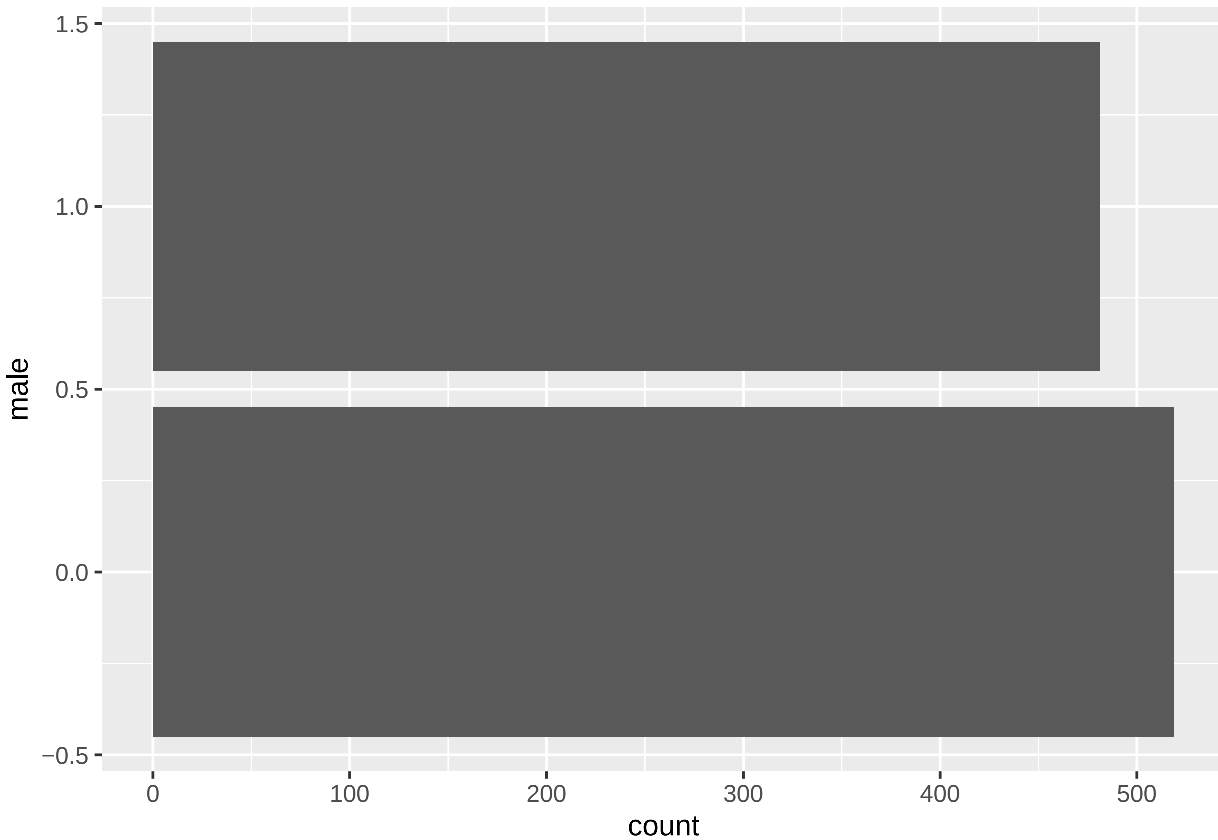
The most activity recorded is in the first 3 months, from both who clicked the ads and those who didn't. ###### January (1), March (3) and July (7) had more activity from those who did not click on the ads as compared to those who clicked on the ads. ###### Months February (2), April (4) and May (5) had more people who clicked on the ads as compared to those who did not click on the ads ###### June (6) had an equal number of people who clicked on the ads and those who did Not ###### We observe that at around mid month we had more people who were not clicking on the ads as compared to the beginning and the end of the month

```r
# Finding out and previewing the hours with the most clicked ads

ggplot(data = ad_df) +
  geom_bar(mapping = aes(y = hour, fill = clicked_on_ad), position = "dodge")
```



Observations - From around 8 pm to 11 pm, we have more people not clicking on ads as compared to those who clicked on the ads before 8 pm and a little after Midnight. 3, 6, 9 and 11 am are the morning hours with the most clicked ads while 3,5 and 6 pm are the hours with the most clicks on the ads in the evening.

```r
# Finding out and previewing the gender with the most clicked ads


ggplot(data = ad_df) +
  geom_bar(mapping = aes(y = male, fill = clicked_on_ad), position = "dodge")
```

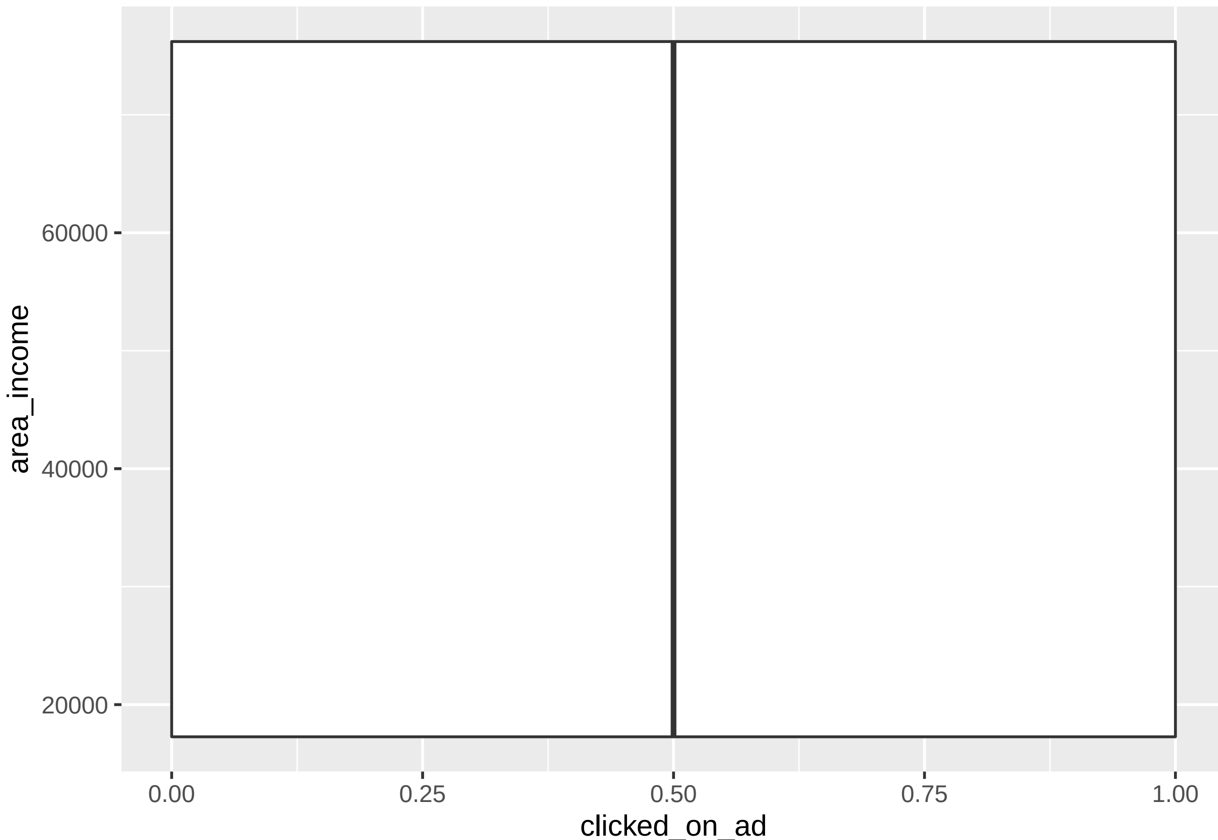


Observations - We have more number of females who clicked on the ads as compared to those who did not. Most males did not click on the ads.

Boxplots

```
# Area Income vs Number of ad clicks
# Finding out and previewing boxplots to show how the area income relates with the number of clicks

ggplot(data = ad_df, mapping = aes( x = area_income, y = clicked_on_ad, fill = clicked_on_ad)) +
  geom_boxplot() +
  coord_flip()
```

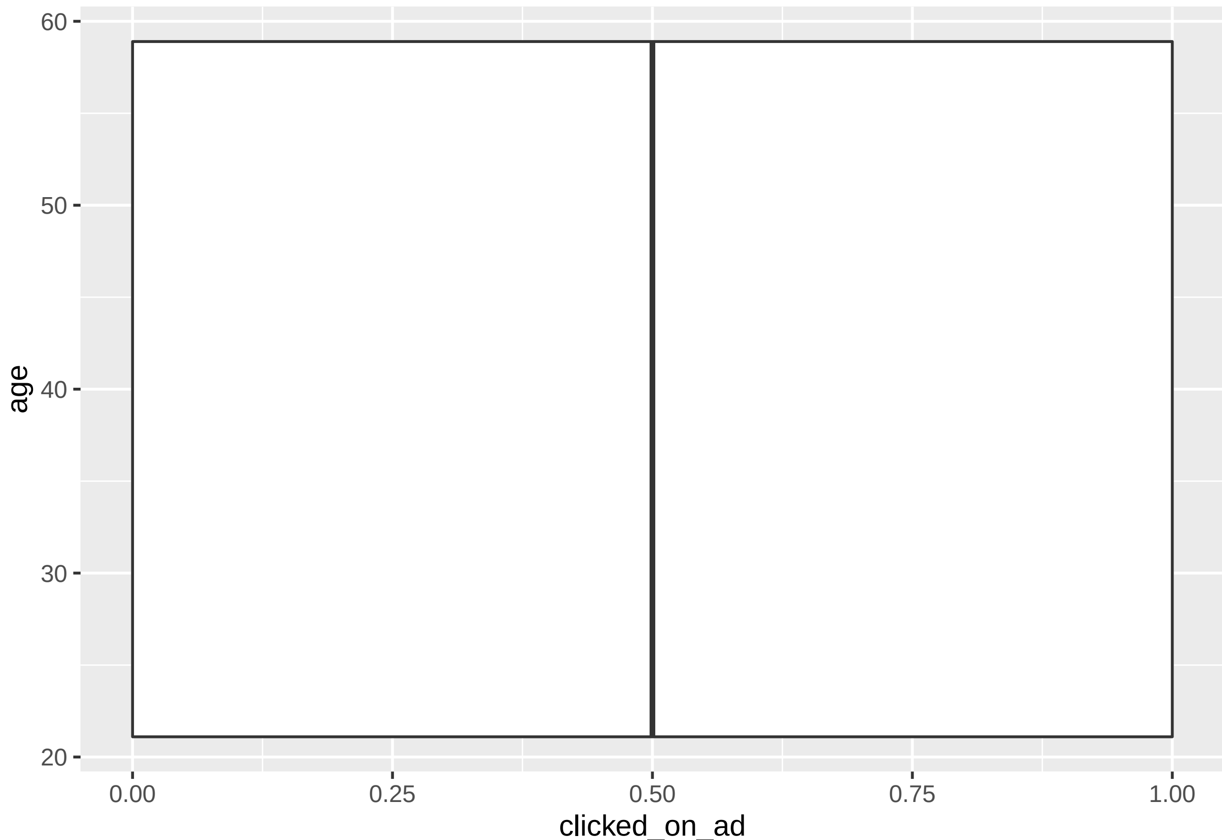
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



Most people who clicked on the ads have a lower income as compared to those who did Not click on the ads

```r
# Age vs Number of ad clicks
# Finding out and previewing boxplots to show how the age relates with the number of clicks

ggplot(data = ad_df, mapping = aes( x = age, y = clicked_on_ad, fill = clicked_on_ad)) +
  geom_boxplot() +
  coord_flip()
```

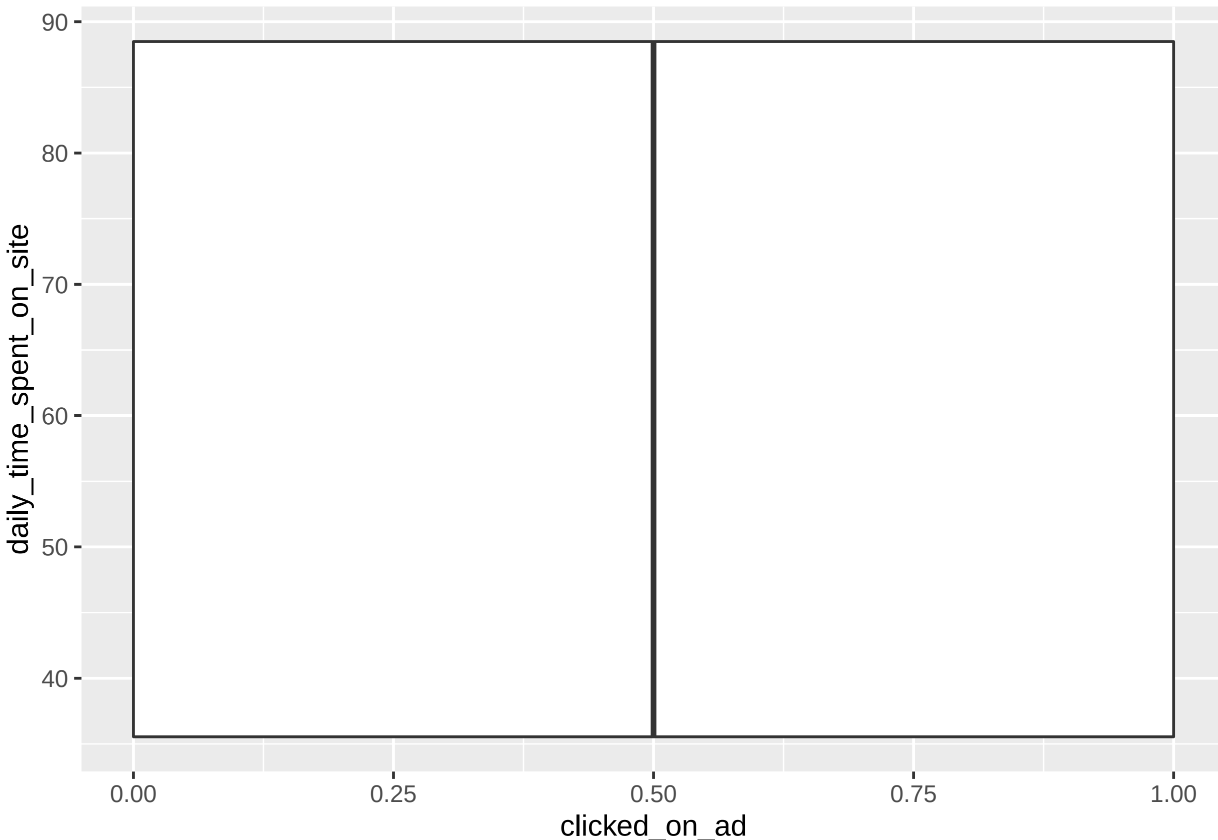
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



Most people who clicked on the ads were older than those who did NOt click on the ads

```
# Daily Time spent on site vs Number of ad clicks
# Finding out and previewing boxplots to show how the daily time spent on site relates with the number

ggplot(data = ad_df, mapping = aes( x = daily_time_spent_on_site, y = clicked_on_ad, fill = clicked_on_a
  geom_boxplot() +
  coord_flip()
```

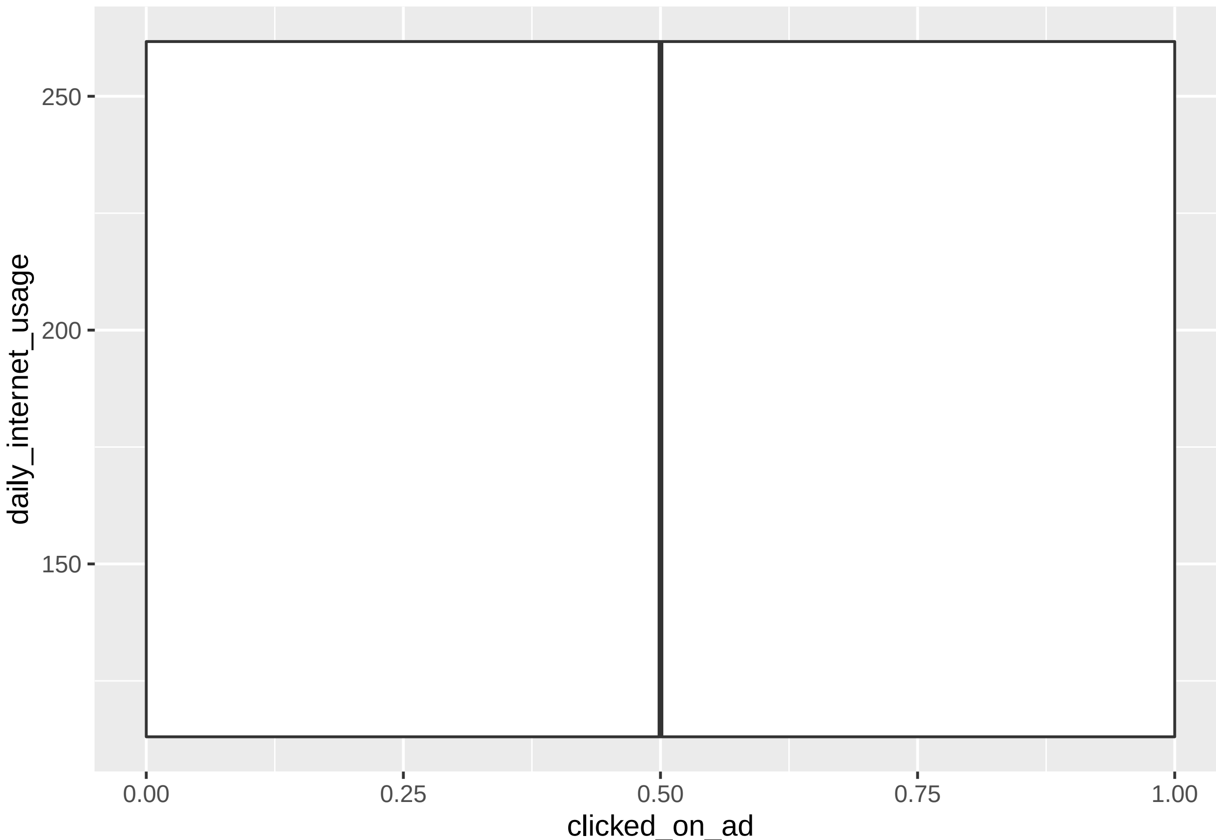
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



Most people who clicked on the ads spent way less time on the site as compared to thos who did not click on the ads

```
# Daily internet usage vs Number of ad clicks
# Finding out and previewing boxplots to show how the daily internet usage relates with the number of c

ggplot(data = ad_df, mapping = aes( x = daily_internet_usage, y = clicked_on_ad, fill = clicked_on_ad))
  geom_boxplot() +
  coord_flip()
```

```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```

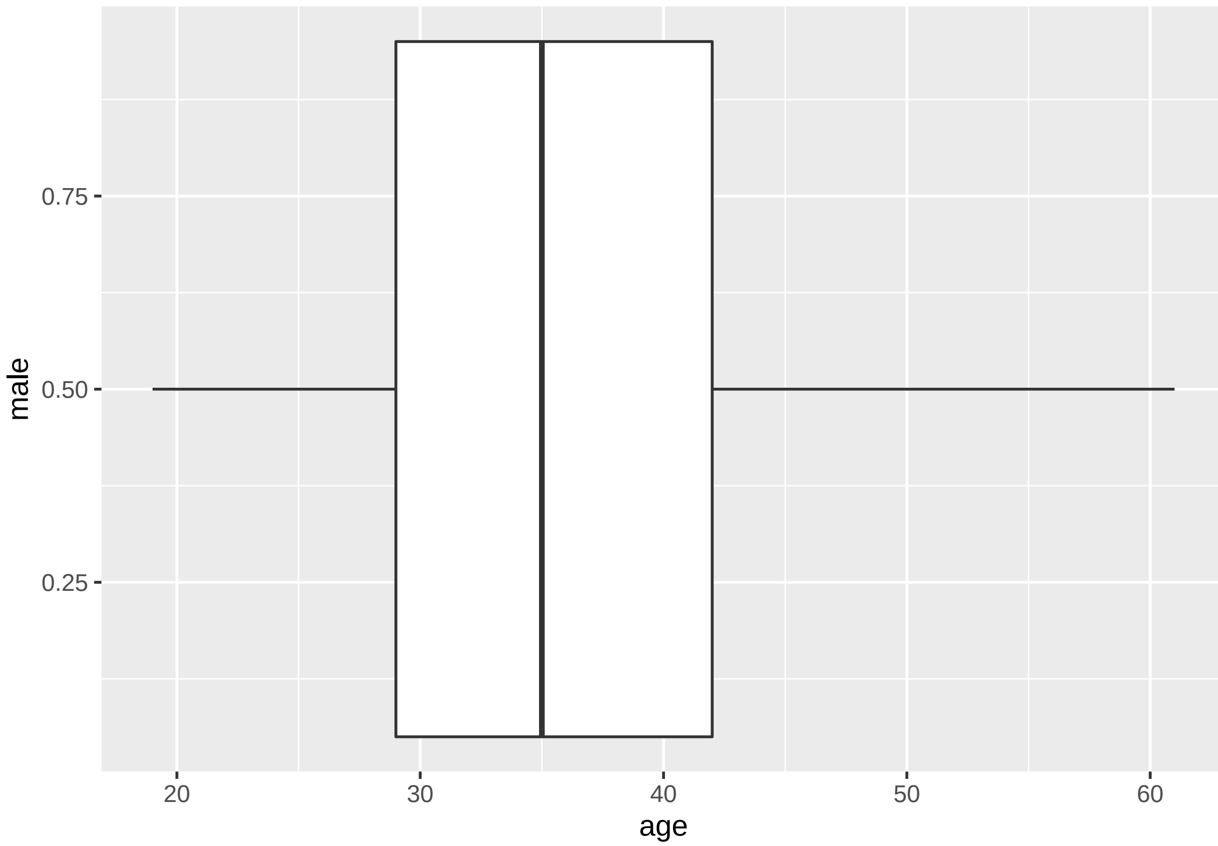


The daily internet usage of most people who clicked on the ads is way less than those who did NOt click on the ads

```r
# Age vs Gender
# Finding out and previewing boxplots to show how the Age relates with the gender

ggplot(data = ad_df, mapping = aes( x = male, y = age, fill = clicked_on_ad)) +
  geom_boxplot() +
  coord_flip()
```

Conclusion - The entepreneur should target people with lower area income levels, older and those who spend less time on the site.

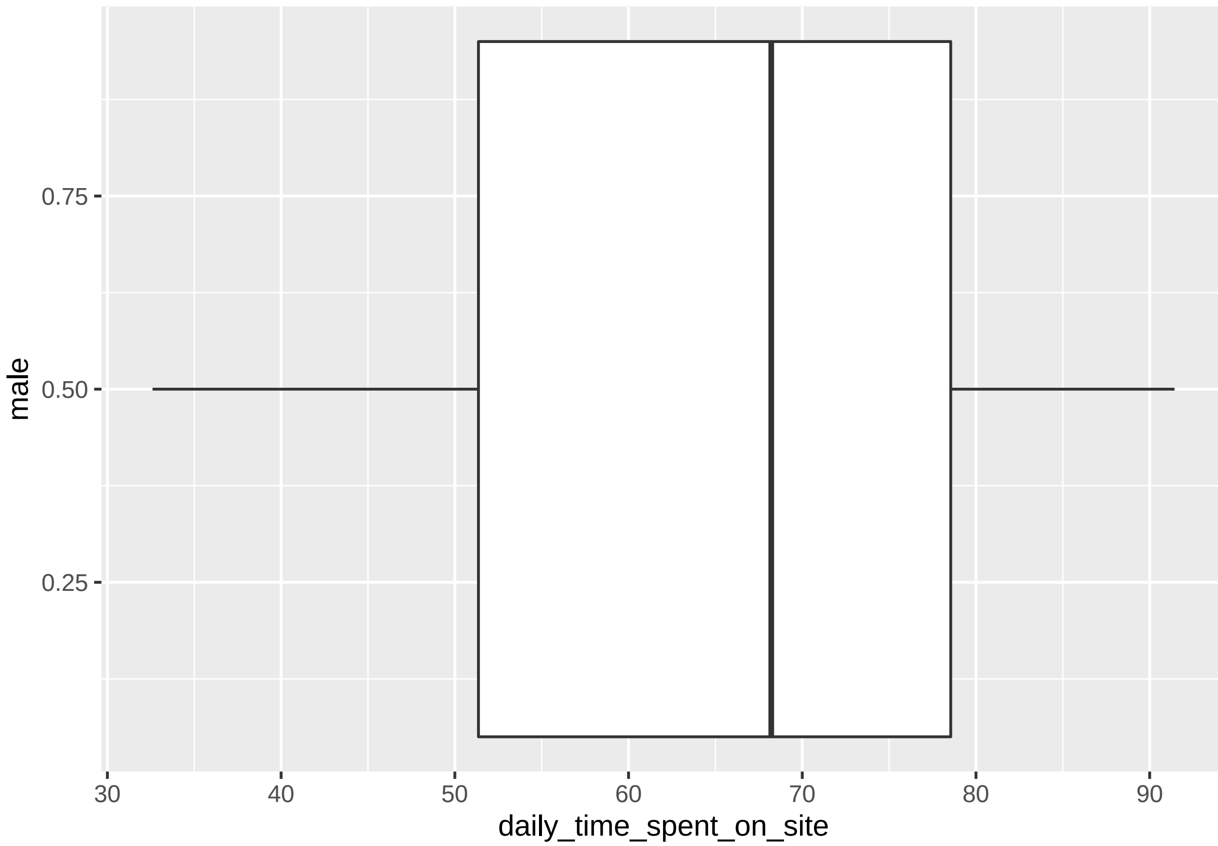
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



Generally, those who clicked on the ads were older, but the males were slightly older than the females

```
# Daily time spent on site vs Gender
# Finding out and previewing boxplots to show how the Age relates with the gender

ggplot(data = ad_df, mapping = aes( x = male , y = daily_time_spent_on_site, fill = clicked_on_ad)) +
  geom_boxplot() +
  coord_flip()
```

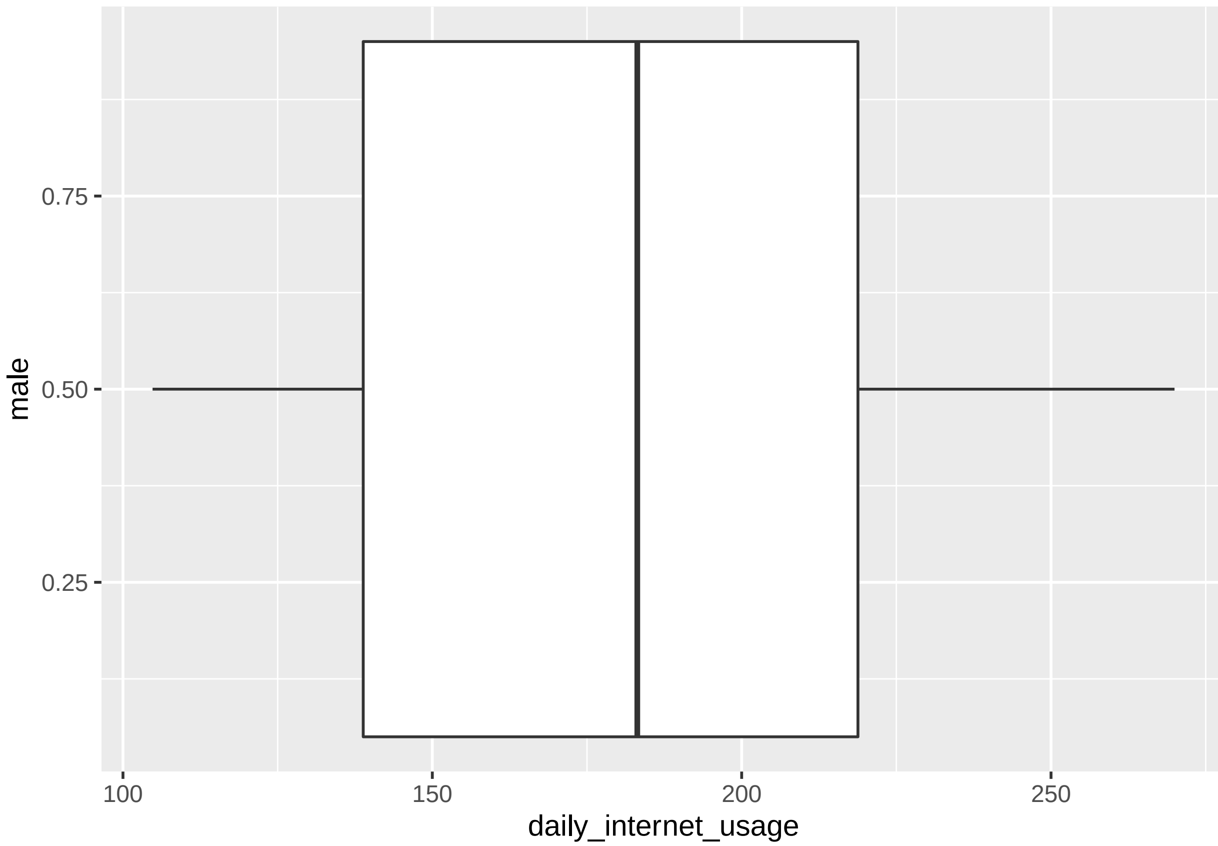
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



More of those who click on the ads spend less time on the site. Of those who click on the ads, the females generally spend more time on the site as compared to the males

```
# Daily internet usage vs Gender
# Finding out and previewing boxplots to show how the Age relates with the gender

ggplot(data = ad_df, mapping = aes( x = male , y = daily_internet_usage, fill = clicked_on_ad)) +
  geom_boxplot() +
  coord_flip()
```

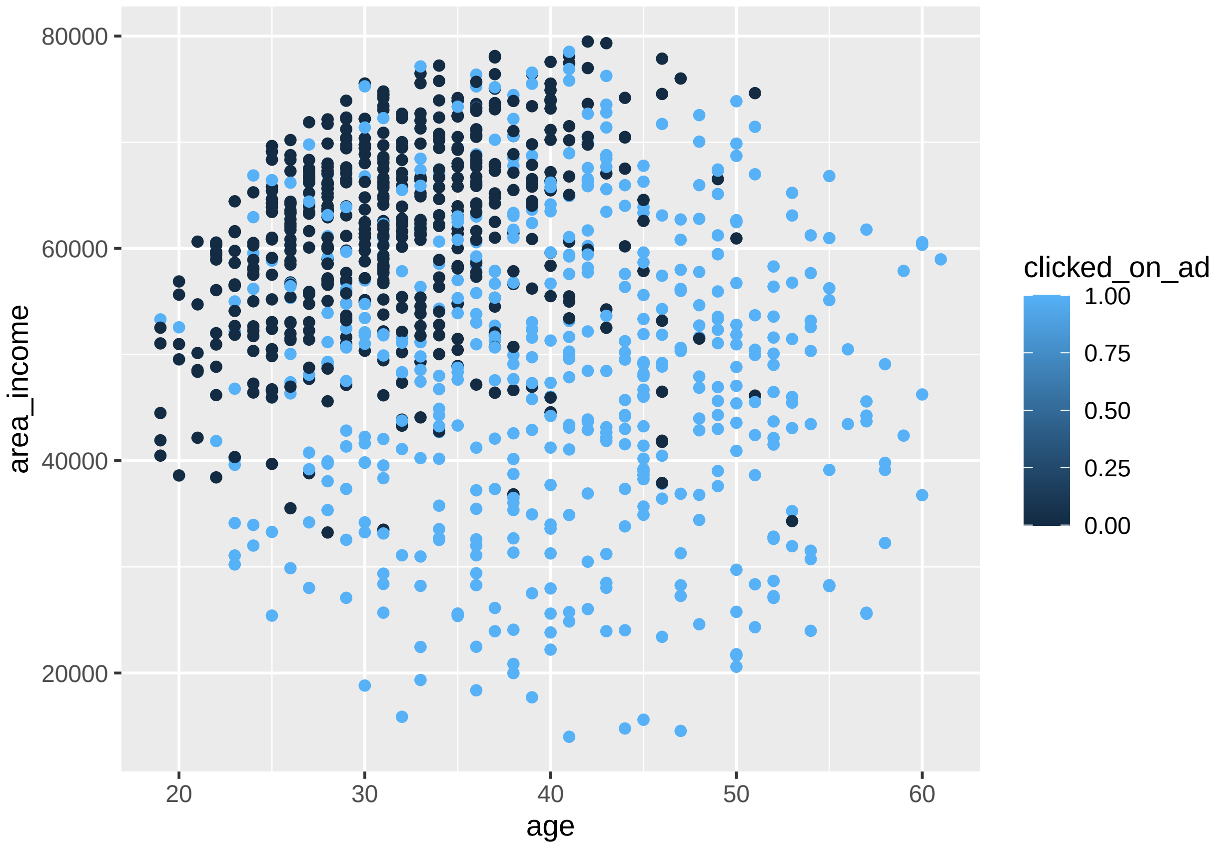
```
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?
```



In general, those who click on the ads have a lower daily internet usage, with a few observations as outlier values with the males were slightly more than the females

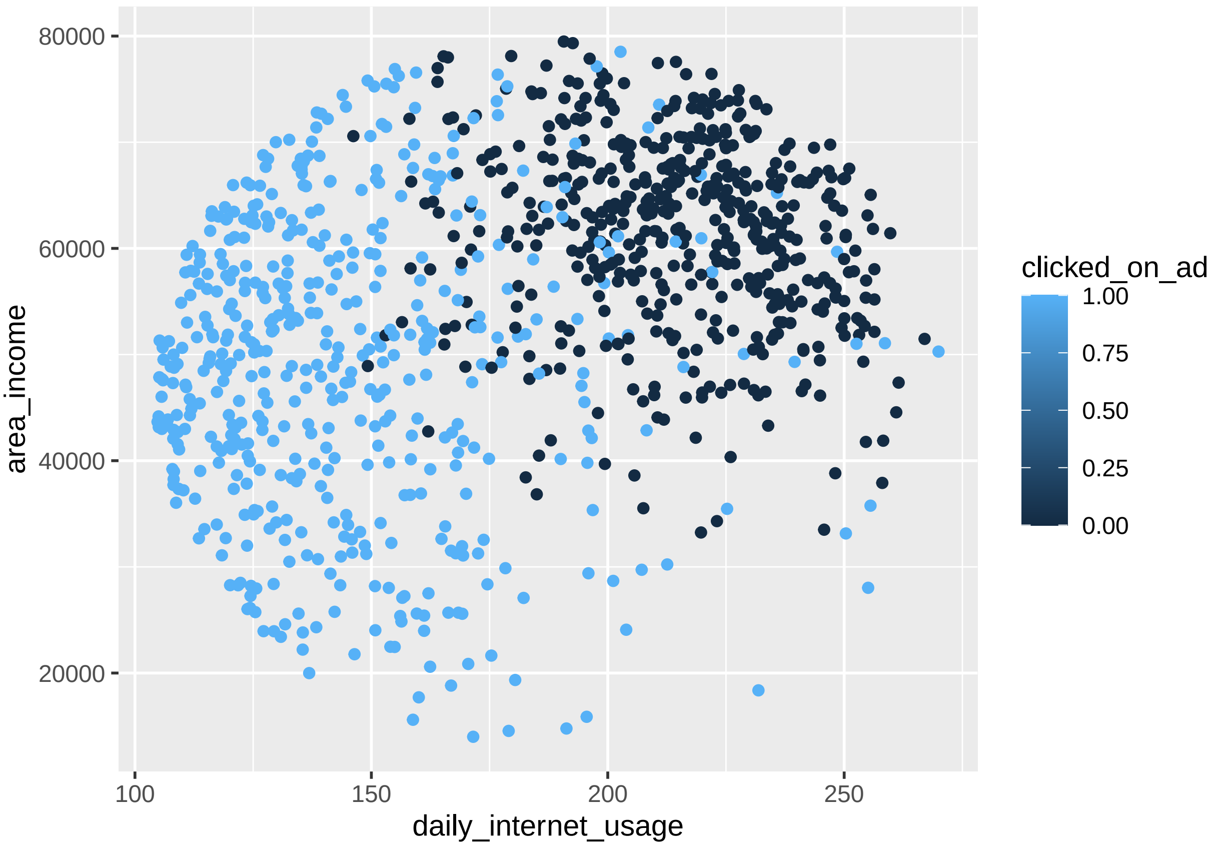
Scatterplots

```
# Age vs Area Income
# Finding out and previewing scatterplots showing how the Age relates with the Area Income

ggplot(data = ad_df) +
  geom_point(mapping = aes(x = age , y = area_income, color = clicked_on_ad))
```



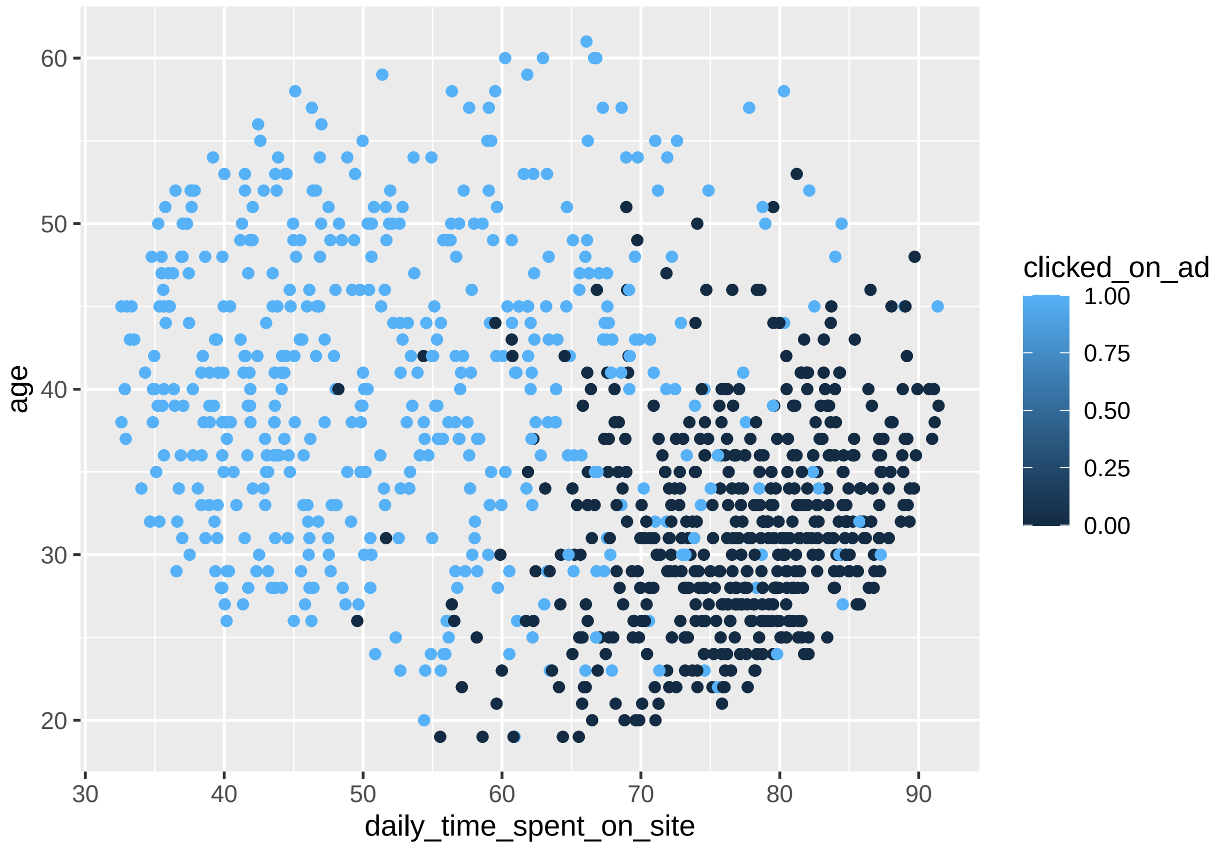
We observe that the number of people who clicked on the ads are more evenly distributed while most of the people who did not click on the ads have a higher area income and a bit younger

```
# Daily Internet usage vs Area Income
# Finding out and previewing scatterplots showing how the Daily internet usage relates with the Area In

ggplot(data = ad_df) +
  geom_point(mapping = aes(x = daily_internet_usage , y = area_income, color = clicked_on_ad))
```



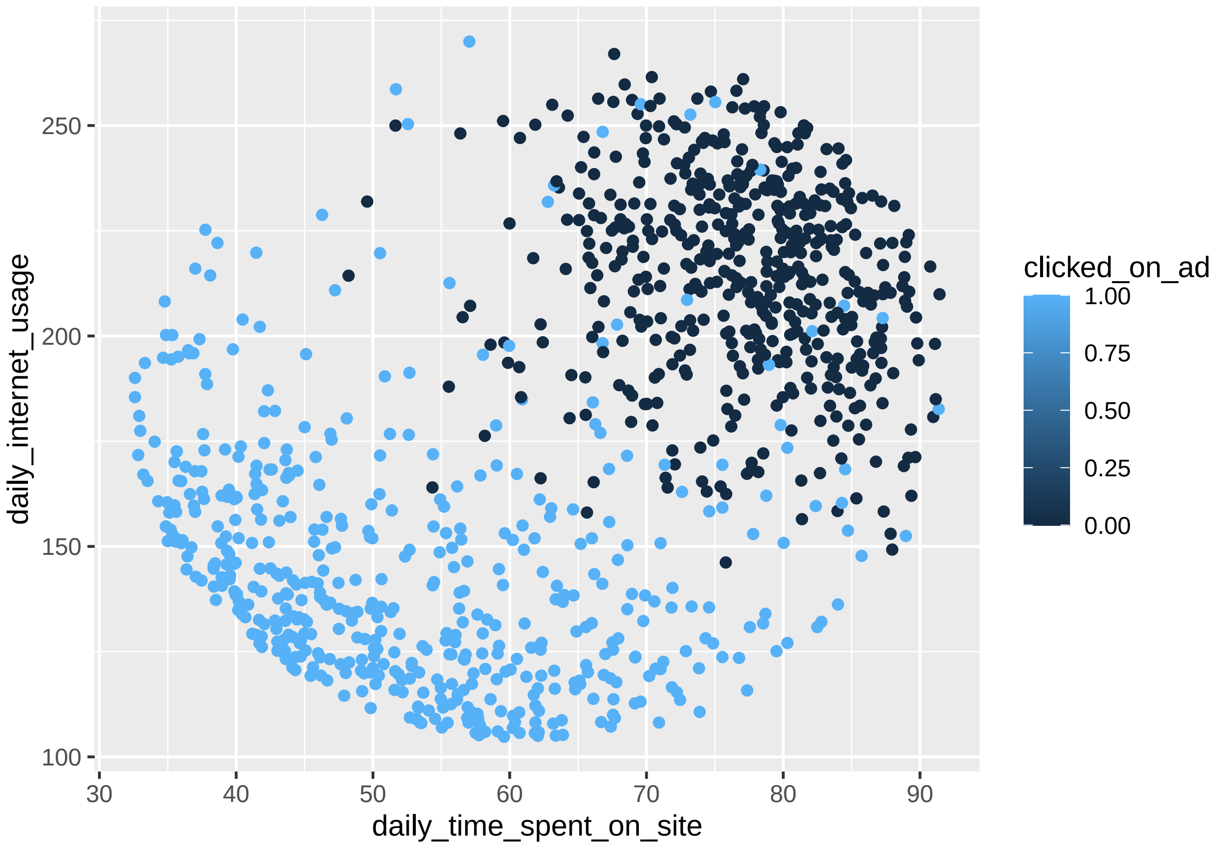
A great number of clicks comes from people who's daily internet usage is quite low and area income is also lower as compared to those who do Not click on the ads whose daily internet usage is significantly higher

```
# Age vs Daily time spent on site
# Finding out and previewing scatterplots showing how the Daily time spent on sites relates with the Ag

ggplot(data = ad_df) +
  geom_point(mapping = aes(x = daily_time_spent_on_site , y = age, color = clicked_on_ad))
```



A huge chuck of clicks come from people who spend significantly little time on the site as compared to those who spend more time on the site regardless of age

```
# Daily Internet Usage vs Daily time spent on site
# Finding out and previewing scatterplots showing how the Daily time spent on sites relates with the Da

ggplot(data = ad_df) +
  geom_point(mapping = aes(x = daily_time_spent_on_site , y = daily_internet_usage, color = clicked_on_a
```



Clearly, more clicks come from people who spend less time on the site and people whose daily internet usage is significantly lower as compared to those who spend more time on the site and have a high daily internet usage

The ads are getting more clicks from people who spend less time on the site and those whose daily internet usage is low.

9. Challenging the solution

Conclusion

Older people were more likely to be interested in cryptography than young users. The mean age of a person who clciked the ad was 40 years of age.

Females were more likely to click the cryptography ad than males however more analysis can be carried out in this particular area to determine the cause of this action.

The individuals from Lisamouth city were more likely to click the ad

People from the middle income areas clicked the ads more than the ones from a higher income area.

The lower daily internet usage users clicked the ads more than the ones who had a higher internet usage

Recommendations

We have observed that the users who were mostly interested in the ads were females who were older, had a lower area income and spent less time on the ads as they had less daily internet usage

To generate more intakes in the course, the company is better off increasing the number of ads towards the end and the beginning of the month and year as compared to the middle of the month and year.

Overall, we can say the study was successful based on the metrics of success.

Follow up Questions

Given that we had access to more data, would we be able to obtain better results?

10 Implementing the solution

```
# Getting a glimpse of our columns and datatypes

glimpse(ad_df)
```

```
## Rows: 1,000
## Columns: 23
## $ daily_time_spent_on_site <dbl> 68.95, 80.23, 69.47, 74.15, 68.37, 59.99, ...
## $ age                      <int> 35, 31, 26, 29, 35, 23, 33, 48, 30, 20, 49...
## $ area_income              <dbl> 61833.90, 68441.85, 59785.94, 54806.18, 73...
## $ daily_internet_usage     <dbl> 256.09, 193.77, 236.50, 245.89, 225.58, 22...
## $ ad_topic_line            <chr> "Cloned 5thgeneration orchestration", "Mon...
## $ city                     <chr> "Wrightburgh", "West Jodi", "Davidton", "W...
## $ male                     <int> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ country                  <chr> "Tunisia", "Nauru", "San Marino", "Italy",...
## $ timestamp                <dttm> 2016-03-27 00:53:11, 2016-04-04 01:39:02,...
## $ clicked_on_ad            <int> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, ...
## $ year                     <chr> "2016", "2016", "2016", "2016", "2016", "2...
## $ month                    <chr> "03", "04", "03", "01", "06", "05", "01", ...
## $ day                      <chr> "27", "04", "13", "10", "03", "19", "28", ...
## $ hour                     <chr> "00", "01", "20", "02", "03", "14", "20", ...
## $ minutes                  <chr> "53", "39", "35", "31", "36", "30", "59", ...
## $ seconds                  <chr> "11", "02", "42", "19", "18", "17", "32", ...
## $ Male                     <fct> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ Year                     <fct> 2016, 2016, 2016, 2016, 2016, 2016, 2016, ...
## $ Month                    <fct> 03, 04, 03, 01, 06, 05, 01, 03, 04, 07, 03...
## $ Day                      <fct> 27, 04, 13, 10, 03, 19, 28, 07, 18, 11, 16...
## $ Hour                     <fct> 00, 01, 20, 02, 03, 14, 20, 01, 09, 01, 20...
## $ Minutes                  <fct> 53, 39, 35, 31, 36, 30, 59, 40, 33, 42, 19...
## $ Seconds                  <fct> 11, 02, 42, 19, 18, 17, 32, 15, 42, 51, 01...
```

```
# Conversion of City and Country to Continent
ad_df$continent <- countrycode(sourcevar = ad_df[, "country"],
                            origin = "country.name",
                            destination = "continent")
```

The cities and countries have high cardinality hence we will convert the countries to continents using the countrycode package and use the continents to perform the modelling

```
## Warning in countrycode(sourcevar = ad_df[, "country"], origin = "country.name", : Some values were no
```

```r
# Finding out and Previewing if our columns have changed

head(ad_df)
```

```
##   daily_time_spent_on_site age area_income daily_internet_usage
## 1                    68.95  35    61833.90               256.09
## 2                    80.23  31    68441.85               193.77
## 3                    69.47  26    59785.94               236.50
## 4                    74.15  29    54806.18               245.89
## 5                    68.37  35    73889.99               225.58
## 6                    59.99  23    59761.56               226.74
##                          ad_topic_line           city male    country
## 1    Cloned 5thgeneration orchestration    Wrightburgh    0    Tunisia
## 2    Monitored national standardization      West Jodi    1      Nauru
## 3      Organic bottom-line service-desk       Davidton    0 San Marino
## 4 Triple-buffered reciprocal time-frame West Terrifurt    1      Italy
## 5         Robust logistical utilization   South Manuel    0    Iceland
## 6       Sharable client-driven software      Jamieberg    1     Norway
##             timestamp clicked_on_ad year month day hour minutes seconds Male
## 1 2016-03-27 00:53:11             0 2016    03  27   00      53      11    0
## 2 2016-04-04 01:39:02             0 2016    04  04   01      39      02    1
## 3 2016-03-13 20:35:42             0 2016    03  13   20      35      42    0
## 4 2016-01-10 02:31:19             0 2016    01  10   02      31      19    1
## 5 2016-06-03 03:36:18             0 2016    06  03   03      36      18    0
## 6 2016-05-19 14:30:17             0 2016    05  19   14      30      17    1
##   Year Month Day Hour Minutes Seconds continent
## 1 2016    03  27   00      53      11    Africa
## 2 2016    04  04   01      39      02   Oceania
## 3 2016    03  13   20      35      42    Europe
## 4 2016    01  10   02      31      19    Europe
## 5 2016    06  03   03      36      18    Europe
## 6 2016    05  19   14      30      17    Europe
```

```r
# Encompasing the continents into factors

ad_df$continent <- factor(ad_df$continent, order = TRUE, levels =c('Africa', 'Americas', 'Asia', 'Europ
```

```r
# Exploring the Continent Column

table(ad_df$continent)
```

```
##
##   Africa Americas     Asia   Europe  Oceania
##      214      219      218      214      100
```

```r
# Pre-processing:
# Converting variables to appropriate data types for modeling

# Converting factor variables to integers
ad_df$male = as.numeric(ad_df$male)
ad_df$month = as.numeric(ad_df$month)
ad_df$day = as.numeric(ad_df$day)
ad_df$hour = as.numeric(ad_df$hour)
ad_df$minutes = as.numeric(ad_df$minutes)
ad_df$seconds = as.numeric(ad_df$seconds)
```

```r
# Converting the clicked on ad variable as a factor

ad_df$clicked_on_ad = as.factor(ad_df$clicked_on_ad)
```

```r
# Previewing the dataset to see if changes were effected

glimpse(ad_df)
```

All the continents seem to be equally distributed except Oceania

```
## Rows: 1,000
## Columns: 24
## $ daily_time_spent_on_site <dbl> 68.95, 80.23, 69.47, 74.15, 68.37, 59.99, ...
## $ age                      <int> 35, 31, 26, 29, 35, 23, 33, 48, 30, 20, 49...
## $ area_income              <dbl> 61833.90, 68441.85, 59785.94, 54806.18, 73...
## $ daily_internet_usage     <dbl> 256.09, 193.77, 236.50, 245.89, 225.58, 22...
## $ ad_topic_line            <chr> "Cloned 5thgeneration orchestration", "Mon...
## $ city                     <chr> "Wrightburgh", "West Jodi", "Davidton", "W...
## $ male                     <dbl> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ country                  <chr> "Tunisia", "Nauru", "San Marino", "Italy",...
## $ timestamp                <dttm> 2016-03-27 00:53:11, 2016-04-04 01:39:02,...
## $ clicked_on_ad            <fct> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, ...
## $ year                     <chr> "2016", "2016", "2016", "2016", "2016", "2...
## $ month                    <dbl> 3, 4, 3, 1, 6, 5, 1, 3, 4, 7, 3, 5, 6, 4, ...
## $ day                      <dbl> 27, 4, 13, 10, 3, 19, 28, 7, 18, 11, 16, 8...
## $ hour                     <dbl> 0, 1, 20, 2, 3, 14, 20, 1, 9, 1, 20, 8, 1,...
## $ minutes                  <dbl> 53, 39, 35, 31, 36, 30, 59, 40, 33, 42, 19...
## $ seconds                  <dbl> 11, 2, 42, 19, 18, 17, 32, 15, 42, 51, 1, ...
## $ Male                     <fct> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ Year                     <fct> 2016, 2016, 2016, 2016, 2016, 2016, 2016, ...
## $ Month                    <fct> 03, 04, 03, 01, 06, 05, 01, 03, 04, 07, 03...
## $ Day                      <fct> 27, 04, 13, 10, 03, 19, 28, 07, 18, 11, 16...
## $ Hour                     <fct> 00, 01, 20, 02, 03, 14, 20, 01, 09, 01, 20...
## $ Minutes                  <fct> 53, 39, 35, 31, 36, 30, 59, 40, 33, 42, 19...
## $ Seconds                  <fct> 11, 02, 42, 19, 18, 17, 32, 15, 42, 51, 01...
## $ continent                <ord> Africa, Oceania, Europe, Europe, Europe, E...
```

```r
# Encoding the continent character variable

ad_df$continent_Numeric <-mapvalues(ad_df$continent, from = c('Africa', 'Americas', 'Asia', 'Europe', '(

glimpse(ad_df)
```

```
## Rows: 1,000
## Columns: 25
## $ daily_time_spent_on_site <dbl> 68.95, 80.23, 69.47, 74.15, 68.37, 59.99, ...
## $ age                      <int> 35, 31, 26, 29, 35, 23, 33, 48, 30, 20, 49...
## $ area_income              <dbl> 61833.90, 68441.85, 59785.94, 54806.18, 73...
## $ daily_internet_usage     <dbl> 256.09, 193.77, 236.50, 245.89, 225.58, 22...
## $ ad_topic_line            <chr> "Cloned 5thgeneration orchestration", "Mon...
## $ city                     <chr> "Wrightburgh", "West Jodi", "Davidton", "W...
## $ male                     <dbl> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ country                  <chr> "Tunisia", "Nauru", "San Marino", "Italy",...
## $ timestamp                <dttm> 2016-03-27 00:53:11, 2016-04-04 01:39:02,...
## $ clicked_on_ad            <fct> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, ...
## $ year                     <chr> "2016", "2016", "2016", "2016", "2016", "2...
## $ month                    <dbl> 3, 4, 3, 1, 6, 5, 1, 3, 4, 7, 3, 5, 6, 4, ...
## $ day                      <dbl> 27, 4, 13, 10, 3, 19, 28, 7, 18, 11, 16, 8...
## $ hour                     <dbl> 0, 1, 20, 2, 3, 14, 20, 1, 9, 1, 20, 8, 1,...
## $ minutes                  <dbl> 53, 39, 35, 31, 36, 30, 59, 40, 33, 42, 19...
## $ seconds                  <dbl> 11, 2, 42, 19, 18, 17, 32, 15, 42, 51, 1, ...
## $ Male                     <fct> 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, ...
## $ Year                     <fct> 2016, 2016, 2016, 2016, 2016, 2016, 2016, ...
## $ Month                    <fct> 03, 04, 03, 01, 06, 05, 01, 03, 04, 07, 03...
## $ Day                      <fct> 27, 04, 13, 10, 03, 19, 28, 07, 18, 11, 16...
## $ Hour                     <fct> 00, 01, 20, 02, 03, 14, 20, 01, 09, 01, 20...
## $ Minutes                  <fct> 53, 39, 35, 31, 36, 30, 59, 40, 33, 42, 19...
## $ Seconds                  <fct> 11, 02, 42, 19, 18, 17, 32, 15, 42, 51, 01...
## $ continent                <ord> Africa, Oceania, Europe, Europe, Europe, E...
## $ continent_Numeric        <ord> 1, 5, 4, 4, 4, 4, 3, 5, 2, 1, 3, 1, 1, 4, ...
```

Some countries could not be recognized by the function countrycode hence were not assigned a continent, this gave rise to some null values, let's look into this

```r
# Checking for null values brought about by unassigned countries and drop them

colSums(is.na(ad_df))
```

```
## daily_time_spent_on_site                      age              area_income
##                        0                        0                        0
##     daily_internet_usage            ad_topic_line                     city
##                        0                        0                        0
##                     male                  country                timestamp
##                        0                        0                        0
##            clicked_on_ad                     year                    month
##                        0                        0                        0
##                      day                     hour                  minutes
##                        0                        0                        0
##                  seconds                     Male                     Year
```

```
##                     0                        0                        0
##                 Month                      Day                     Hour
##                     0                        0                        0
##               Minutes                  Seconds                continent
##                     0                        0                       35
##     continent_Numeric
##                    35
```

```r
ad_df <- na.omit(ad_df)

colSums(is.na(ad_df))
```

```
## daily_time_spent_on_site                      age              area_income
##                        0                        0                        0
##     daily_internet_usage            ad_topic_line                     city
##                        0                        0                        0
##                     male                  country                timestamp
##                        0                        0                        0
##             clicked_on_ad                     year                    month
##                        0                        0                        0
##                      day                     hour                  minutes
##                        0                        0                        0
##                  seconds                     Male                     Year
##                        0                        0                        0
##                    Month                      Day                     Hour
##                        0                        0                        0
##                  Minutes                  Seconds                continent
##                        0                        0                        0
##        continent_Numeric
##                        0
```

```r
# We will delete the timestamp, year and continent columns as they are irrelevant for modeling as well 

# We will preview the results to see if the changes have been effected

ad_df$timestamp <- NULL
ad_df$year <- NULL
ad_df$ad_topic_line <- NULL
ad_df$city <- NULL
ad_df$country <- NULL
ad_df$continent <- NULL
ad_df$Male <- NULL
ad_df$Day <- NULL
ad_df$Minutes <- NULL
ad_df$Seconds <- NULL
ad_df$Hour <- NULL
ad_df$Month <- NULL
ad_df$Year <- NULL

colnames(ad_df)
```

```
##  [1] "daily_time_spent_on_site" "age"
##  [3] "area_income"              "daily_internet_usage"
```

```
##  [5] "male"                     "clicked_on_ad"
##  [7] "month"                    "day"
##  [9] "hour"                     "minutes"
## [11] "seconds"                  "continent_Numeric"
```

```r
describe(ad_df)
```

```
##                          vars   n     mean       sd   median  trimmed      mad
## daily_time_spent_on_site    1 965    65.15    15.76    68.25    65.91    17.76
## age                         2 965    36.04     8.83    35.00    35.54     8.90
## area_income                 3 965 54972.55 13433.69 56986.73 55990.94 13370.69
## daily_internet_usage        4 965   179.86    43.96   182.65   179.85    58.73
## male                        5 965     0.48     0.50     0.00     0.47     0.00
## clicked_on_ad*              6 965     1.50     0.50     2.00     1.50     0.00
## month                       7 965     3.81     1.92     4.00     3.76     2.97
## day                         8 965    15.54     8.76    15.00    15.48    11.86
## hour                        9 965    11.68     6.97    12.00    11.71     8.90
## minutes                    10 965    29.13    17.22    30.00    29.09    22.24
## seconds                    11 965    29.71    16.88    30.00    29.80    20.76
## continent_Numeric*         12 965     2.76     1.30     3.00     2.70     1.48
##                               min      max    range  skew kurtosis     se
## daily_time_spent_on_site    32.60    91.43    58.83 -0.38    -1.07   0.51
## age                         19.00    61.00    42.00  0.47    -0.43   0.28
## area_income              13996.50 79484.80 65488.30 -0.64    -0.13 432.45
## daily_internet_usage       104.78   269.96   165.18 -0.03    -1.27   1.41
## male                         0.00     1.00     1.00  0.09    -1.99   0.02
## clicked_on_ad*               1.00     2.00     1.00  0.00    -2.00   0.02
## month                        1.00     7.00     6.00  0.09    -1.18   0.06
## day                          1.00    31.00    30.00  0.04    -1.18   0.28
## hour                         0.00    23.00    23.00 -0.01    -1.23   0.22
## minutes                      0.00    59.00    59.00  0.02    -1.18   0.55
## seconds                      0.00    59.00    59.00 -0.03    -1.15   0.54
## continent_Numeric*           1.00     5.00     4.00  0.13    -1.14   0.04
```

```r
# Normalizing the continous variables
normalize <- function(x) (
  return( ((x - min(x)) /(max(x)-min(x))) )
)

ad_df$daily_time_spent_on_site <- normalize(ad_df$daily_time_spent_on_site)
ad_df$daily_internet_usage <- normalize(ad_df$daily_internet_usage)
ad_df$area_income <- normalize(ad_df$area_income)
ad_df$age<- normalize(ad_df$age)

head(ad_df)
```

We observe that all the variables have varying range values both continous and factor variables. Hence we will normalize the continous variables

```
##   daily_time_spent_on_site       age area_income daily_internet_usage male
```

```
## 1                0.6178820 0.3809524   0.7304725            0.9160310    0
## 2                0.8096209 0.2857143   0.8313752            0.5387456    1
## 3                0.6267211 0.1666667   0.6992003            0.7974331    0
## 4                0.7062723 0.2380952   0.6231599            0.8542802    1
## 5                0.6080231 0.3809524   0.9145678            0.7313234    0
## 6                0.4655788 0.0952381   0.6988280            0.7383460    1
##   clicked_on_ad month day hour minutes seconds continent_Numeric
## 1             0     3  27    0      53      11                 1
## 2             0     4   4    1      39       2                 5
## 3             0     3  13   20      35      42                 4
## 4             0     1  10    2      31      19                 4
## 5             0     6   3    3      36      18                 4
## 6             0     5  19   14      30      17                 4
```

Supervised Machine Learning Models

Baseline Model

K-Nearest Neighbours

```r
set.seed(123)
# Creating a random number equal 70% of total number of rows
ran <- sample(1:nrow(ad_df),0.7 * nrow(ad_df))

# The training dataset extracted
ad_train <- ad_df[ran,]
#head(ad_train)

# The test dataset extracted
ad_test <- ad_df[-ran,]
#ad_test

# Extracting the target variable from the target variable
ad_target <- (ad_df[ran,6])
#ad_target

# Extracting the target variable from the test dataset
test_target <- (ad_df[-ran,6])
#test_target

# Calculating the square root of the length of the target variable to get an optimal k
print(sqrt(length(ad_test)))
```

```
## [1] 3.464102
```

```r
#3

# Running the knn function, with k = 3 as from the above calculation
library(class)
k <- knn(ad_train,ad_test,cl=ad_target,k=3)
```

```r
# Creating the confucion matrix
matrix <- table(k,test_target)
print(matrix)
```

```
##    test_target
## k    0  1
##   0 75 59
##   1 76 80
```

```r
# Checking the accuracy
#Calculating the accuracy score of our model
accuracy <- function(x){sum(diag(x)/(sum(rowSums(x)))) * 100}
accuracy(matrix)
```

```
## [1] 53.44828
```

Our baseline model has an accuracy score of 53%, which is poor hence we will use random forests, svm and naive bayes to try and achieve a better accuracy

```r
# Building Naive Bayes model on our data
# setting the metod as cross validation with 10 iterations
model = train(ad_train, ad_target,'nb',trControl=trainControl(method='cv',number=10))
```

Naive Bayes Algorithm

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 21
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 21

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 68

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 68
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 21

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 21

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 21

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 68

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 23

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 49

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 61

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 68
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 69

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 10

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 11

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 14
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 36

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 37

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 40

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 42

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 52

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 59

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 63

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 68

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 1

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 2

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 3

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 4

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 5

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 6

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 7

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 12
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 13

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 16

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 17

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 18

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 19

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 20

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 22

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 26

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 27

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 28

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 29

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 30

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 31

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 32
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 33

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 34

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 38

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 39

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 41

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 43

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 44

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 47

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 48

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 50

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 51

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 54

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 57

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 60

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 62

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 67
```

```r
# Model Evalution
# ---
# Predicting our testing set
#
predict <- predict(model,newdata = ad_test )
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 8

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 9

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 15

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 24

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 25

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 35

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 45

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 46

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 53
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 55

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 56

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 58

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 64

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 65

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 66

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 72

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 73

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 76

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 79

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 81

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 83

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 85

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 88

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 91

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 92

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 93
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 98

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 108

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 111

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 112

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 115

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 117

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 120

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 122

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 124

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 125

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 127

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 130

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 144

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 147

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 148

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 150

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 152
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 158

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 161

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 166

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 167

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 170

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 172

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 177

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 178

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 182

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 184

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 188

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 190

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 193

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 194

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 196

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 198

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 203
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 211

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 218

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 219

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 221

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 228

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 237

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 238

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 240

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 250

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 254

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 256

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 258

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 260

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 265

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 269

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 272

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 276
```

```
## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 277

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 280

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 285

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 286

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 288

## Warning in FUN(X[[i]], ...): Numerical 0 probability for all classes with
## observation 290
```

```r
# Getting the confusion matrix to see accuracy value and other parameter values
confusionMatrix(predict, test_target )
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0 151   0
##          1   0 139
##
##                Accuracy : 1
##                  95% CI : (0.9874, 1)
##     No Information Rate : 0.5207
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##
##  Mcnemar's Test P-Value : NA
##
##             Sensitivity : 1.0000
##             Specificity : 1.0000
##          Pos Pred Value : 1.0000
##          Neg Pred Value : 1.0000
##              Prevalence : 0.5207
##          Detection Rate : 0.5207
##    Detection Prevalence : 0.5207
##       Balanced Accuracy : 1.0000
##
##        'Positive' Class : 0
##
```

From the Naive Bayes Model, we have an accuracy score of 100%, with all the observations classified correclty. This is an overfit which means that our model would not do well with new data

SVM model

```r
#setting the method to repeated cv and 10 number of iterations

trctrl <- trainControl(method = "repeatedcv", number = 10, repeats = 3)

#fitting svm linear
svm_Linear <- train(clicked_on_ad~., data = ad_train, method = "svmLinear",
trControl= trctrl,
preProcess = c("center", "scale"),
tuneLength = 10)

#result of our train model
svm_Linear
```

```
## Support Vector Machines with Linear Kernel
##
## 675 samples
##  11 predictor
##   2 classes: '0', '1'
##
## Pre-processing: centered (14), scaled (14)
## Resampling: Cross-Validated (10 fold, repeated 3 times)
## Summary of sample sizes: 607, 607, 607, 608, 607, 608, ...
## Resampling results:
##
##   Accuracy   Kappa
##   0.9619179  0.923818
##
## Tuning parameter 'C' was held constant at a value of 1
```

The svm model has an accuracy of approximately 98% on the data, with a total of 5 incorrect classifications. This is so far the best model for the data as it has a reasonable accuracy score.

```r
# Hyperparameter Tuning for SVM

# Control params for SVM

ctrl <- trainControl(
  method = "cv",
  number = 10,
)

# Tune an SVM
set.seed(7000)
svm <- train(
  clicked_on_ad ~ .,
  data = ad_train,
  method = "svmRadial",
  preProcess = c("center", "scale"),
  trControl = ctrl,
  tuneLength = 10
```

```r
)

svm
```

```
## Support Vector Machines with Radial Basis Function Kernel
##
## 675 samples
##  11 predictor
##   2 classes: '0', '1'
##
## Pre-processing: centered (14), scaled (14)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 608, 608, 607, 608, 607, 607, ...
## Resampling results across tuning parameters:
##
##   C       Accuracy   Kappa
##     0.25  0.9659570  0.9319573
##     0.50  0.9629719  0.9259837
##     1.00  0.9614794  0.9229966
##     2.00  0.9614574  0.9229199
##     4.00  0.9614794  0.9229470
##     8.00  0.9511414  0.9022643
##    16.00  0.9540825  0.9081575
##    32.00  0.9466637  0.8933170
##    64.00  0.9437006  0.8873733
##   128.00  0.9303556  0.8606424
##
## Tuning parameter 'sigma' was held constant at a value of 0.04223037
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were sigma = 0.04223037 and C = 0.25.
```

```r
# Looking at the SVM predictions
svm_pred <- predict(svm, newdata = ad_test)
```

```r
# Plotting confusion matrix
confusionMatrix(table(svm_pred, ad_test$clicked_on_ad))
```

```
## Confusion Matrix and Statistics
##
##
## svm_pred   0   1
##        0 151   8
##        1   0 131
##
##                Accuracy : 0.9724
##                  95% CI : (0.9464, 0.988)
##     No Information Rate : 0.5207
##     P-Value [Acc > NIR] : < 2e-16
##
##                   Kappa : 0.9446
##
##  Mcnemar's Test P-Value : 0.01333
```

```
##
##             Sensitivity : 1.0000
##             Specificity : 0.9424
##          Pos Pred Value : 0.9497
##          Neg Pred Value : 1.0000
##              Prevalence : 0.5207
##          Detection Rate : 0.5207
##    Detection Prevalence : 0.5483
##       Balanced Accuracy : 0.9712
##
##        'Positive' Class : 0
##
```

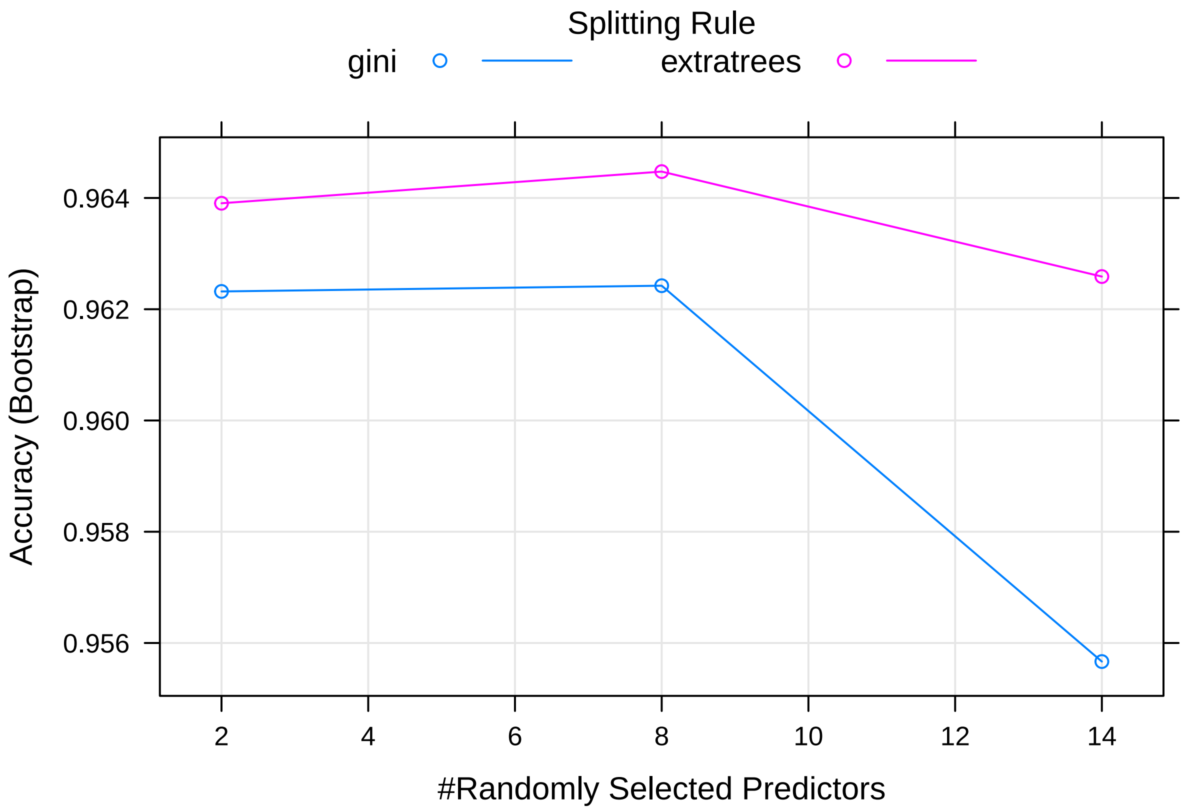
The svmRadial has actually performed worse than the svmLinear hence the svmLinear is the better svm model for this dataset

Random Forests

```
#fitting a single tree to the data
set.seed(12)
rforests <- train(clicked_on_ad ~ .,
               data = ad_df,
               method = "ranger")
rforests
```

```
## Random Forest
##
## 965 samples
##  11 predictor
##   2 classes: '0', '1'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 965, 965, 965, 965, 965, 965, ...
## Resampling results across tuning parameters:
##
##   mtry  splitrule   Accuracy   Kappa
##    2    gini        0.9623200  0.9245405
##    2    extratrees  0.9639053  0.9277429
##    8    gini        0.9624231  0.9247303
##    8    extratrees  0.9644744  0.9288505
##   14    gini        0.9556662  0.9112095
##   14    extratrees  0.9625875  0.9250781
##
## Tuning parameter 'min.node.size' was held constant at a value of 1
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were mtry = 8, splitrule = extratrees
##  and min.node.size = 1.
```

```
# plotting the model
plot(rforests)
```



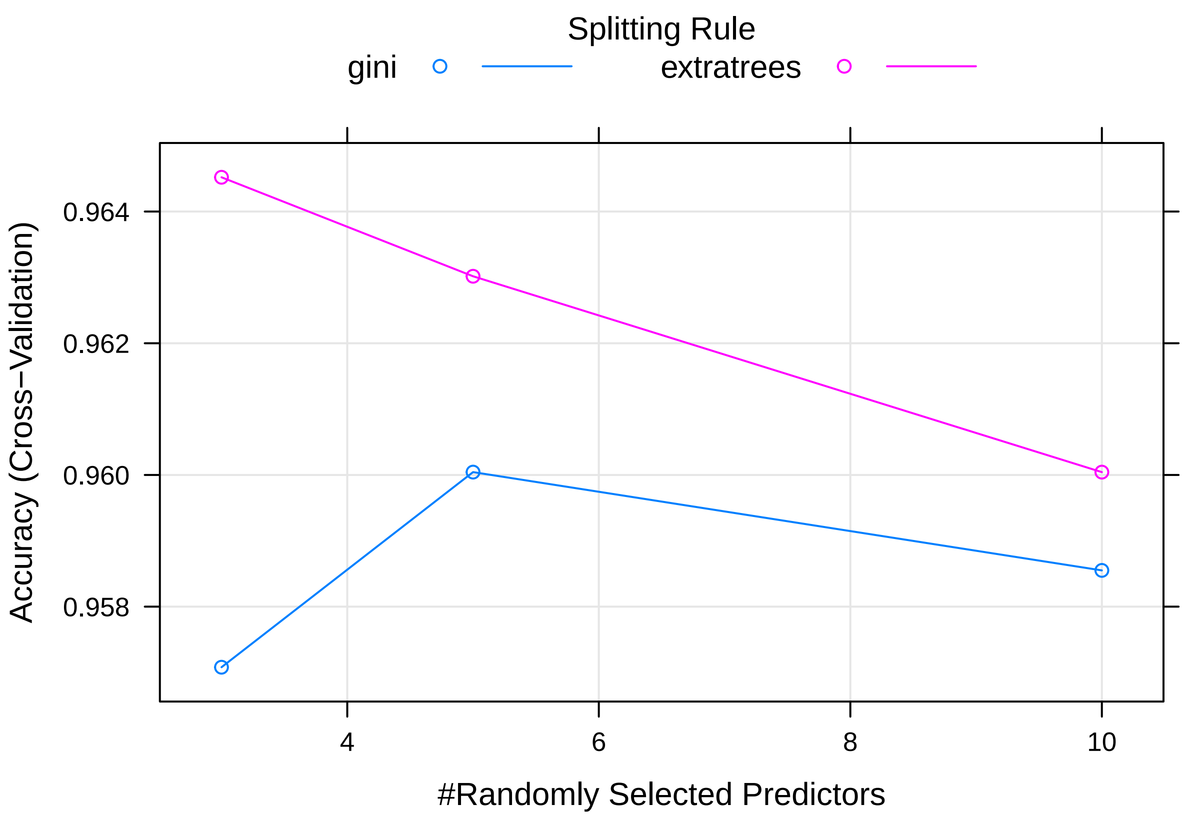
The Gini index reduces steadily from 2 to 8 and then gradually from 8 to 14, with an aim to reduce gini impurity. This means that the splitting rule is selecting the best gin but towards the end, it is overfitting

```
# Improving model performance
# Training the model
rforests_model <- train(clicked_on_ad ~ .,
                        data = ad_train,
                        method = "ranger",
                        tuneLength = 5)

# Setting grid search
set.seed(42)
myGrid <- expand.grid(mtry = c(3,5,10),
                      splitrule = c("gini", "extratrees"),
                      min.node.size = 10)
rforests_model <- train(clicked_on_ad ~ .,
                        data = ad_train,
                        method = "ranger",
                        tuneGrid = myGrid,
                        trControl = trainControl(method = "cv",
                                                 number = 5,
                                                 verboseIter = FALSE))

# Printing the model
rforests_model
```

```
## Random Forest
##
```

```
## 675 samples
##  11 predictor
##   2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (5 fold)
## Summary of sample sizes: 540, 541, 540, 540, 539
## Resampling results across tuning parameters:
##
##   mtry  splitrule   Accuracy   Kappa
##    3    gini        0.9570800  0.9141342
##    3    extratrees  0.9645204  0.9290764
##    5    gini        0.9600431  0.9200708
##    5    extratrees  0.9630169  0.9260458
##   10    gini        0.9585505  0.9170999
##   10    extratrees  0.9600429  0.9201043
##
## Tuning parameter 'min.node.size' was held constant at a value of 10
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were mtry = 3, splitrule = extratrees
##  and min.node.size = 10.
```

```
plot(rforests_model)
```



The Gini is a straight line with a descending slope while the splitting rule increases steadily at first then stagnates after sometime after selecting the split that minimizes the gini impurity

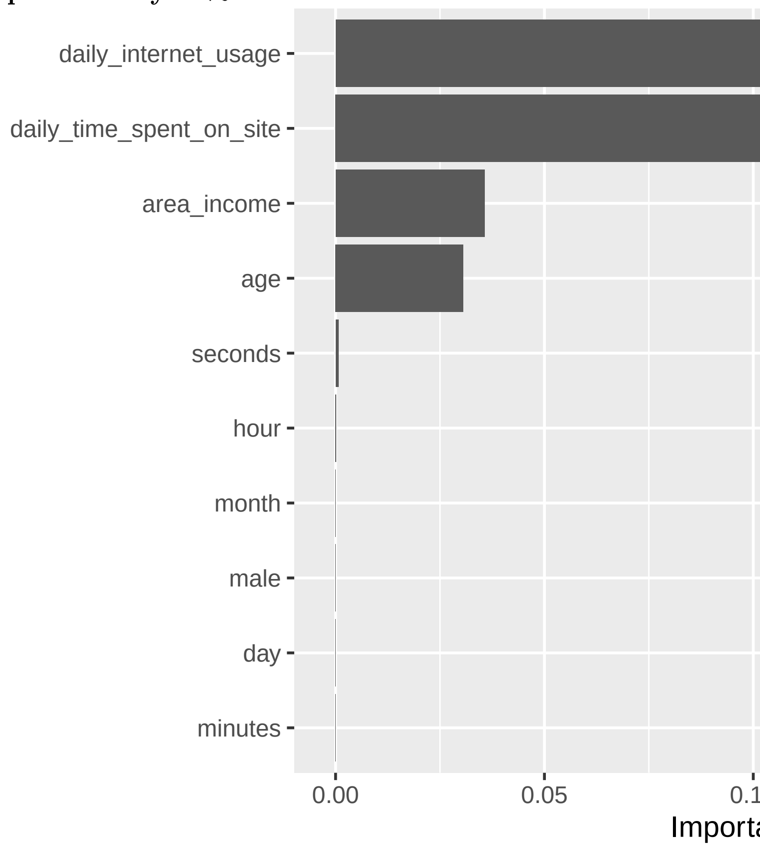
```r
# Evaluating model performance
# Predictions
rforests_pred <- predict(rforests_model, newdata = ad_test)

#confusion matrix
confusionMatrix(table(rforests_pred, ad_test$clicked_on_ad))
```

```
## Confusion Matrix and Statistics
##
##
## rforests_pred   0   1
##             0 148   7
##             1   3 132
##
##                Accuracy : 0.9655
##                  95% CI : (0.9375, 0.9833)
##     No Information Rate : 0.5207
##     P-Value [Acc > NIR] : <2e-16
##
##                   Kappa : 0.9308
##
##  Mcnemar's Test P-Value : 0.3428
##
##             Sensitivity : 0.9801
##             Specificity : 0.9496
##          Pos Pred Value : 0.9548
##          Neg Pred Value : 0.9778
##              Prevalence : 0.5207
##          Detection Rate : 0.5103
##    Detection Prevalence : 0.5345
##       Balanced Accuracy : 0.9649
##
##        'Positive' Class : 0
##
```

```r


# Feature Importance
# Re-run model with permutation-based variable importance
rforests_permutation <- ranger(
  formula = clicked_on_ad ~ .,
  data = ad_train,
  #num.trees = 2000,
  mtry = 3,
  min.node.size = 10,
  #sample.fraction = .80,
  replace = FALSE,
  importance = "permutation",
  respect.unordered.factors = "order",
  verbose = FALSE,
  seed  = 123
)
```

```r
#plotting a graph of the feature importance

vip::vip(rforests_permutation, num_features = 10)
```

The Random Forests has obtained an accuracy of approximately 98% with 6 incorrect observa-



tions. This is the second best model we have so far.

From the graph above, it is evident that the variables that are most important are;

Daily internet usage

Daily time spent on site

Age

Area income

Challenging the solution

From the solutions above, the svm model has performed the best followed by the Decision Trees with a total of 5 and 6 incorrect observations respectively. They were both highly accurate with approximately 98% accuracy hence we can say that this study has been successful.

Follow up questions

How would our models perform using different metrics of success and if we stacked the four models would we have used the combined prediction power to get better results?