Core Week 12 IP

1. INTRODUCTION

1.1 Defining the Question

My work is to identify which facators determine whether a user clicks on an ad or not.

1.2 Setting the Metric for Success

The project will be considered a success when I am able to identify what makes a user more likely to click on an ad.

1.3 Outlining the Context

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 would now like to employ your services as a Data Science Consultant to help her identify which individuals are most likely to click on her ads.

1.4 Drafting the Experimental Design

1. Define the question, set the metric for success, outline the context, drafting the experimental design, and determining the appropriateness of the data.
2. Load the dataset and previewing it.
3. Check for missing and duplicated values and deal with them where necessary.
4. Check for outliers and other anomalies and deal with them where necessary.
5. Perform univariate and bivariate analysis.
6. Create a baseline model and assess its accuracy score.
7. Challenge the solution.
8. Conclude and provide insights on how this project can be improved.

1.5 Determining the Appropriateness of the Data

2. Data Preparation and Cleaning

```
# importing and previewing the dataset
data <- read.csv("advertising.csv", header = TRUE)
head(data)
```

```
##   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
## 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
```

```
# finding the number of rows and columns
dim(data)
```

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

This shows us that we have 1000 rows and 10 columns.

```
# previewing basic information
str(data)
```

```
## '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           : Factor w/ 1000 levels "Adaptive 24hour Graphic Interface",..: 92 465 567
##  $ City                    : Factor w/ 969 levels "Adamsbury","Adamside",..: 962 904 112 940 806 283
##  $ Male                    : int  0 1 0 1 0 1 0 1 1 1 ...
##  $ Country                 : Factor w/ 237 levels "Afghanistan",..: 216 148 185 104 97 159 146 13 83
##  $ Timestamp               : Factor w/ 1000 levels "2016-01-01 02:52:10",..: 440 475 368 57 768 690 1
##  $ Clicked.on.Ad           : int  0 0 0 0 0 0 0 1 0 0 ...
```

Here we see that the following columns have the following data types:

- Daily.Time.Spent.on.Site : numerical
- Age : integer
- Area.Income : numerical
- Daily.Internet.Usage : numerical
- Ad.Topic.Line : Factor with 1000 levels

- City : Factor with 969 levels
- Male : integer
- Country : Factor with 237 levels
- Timestamp : Factor with 1000 levels
- Clicked.on.Ad : integer
- Numerical means it is a number which can be either a whole number or a decimal.
- Integer means it is a whole number only.
- Factor means it is a categorical (non-numeric) value. Factor with x levels means it has x unique values, e.g. Country is a Factor wit 237 levels meaning it has 237 unique categorical values.

```
# checking for duplicates
anyDuplicated(data)
```

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

There are no duplicated records so there is no need to remove any of them.

```
# looking for missing values
colSums(is.na(data))
```

```
## 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
##                        0
```

There are no missing values in each column so we don't need to carry out imputation or replacement.

We should modify the dataset so as to make it easier to work with. We will start by changing the column names and then change the "Male" and "Clicked on Ad" columns to be categorical variables (Factors) instead of numerical variables because it makes more logical sense that way.

```
# get column names
colnames(data)
```

```
##  [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
# rename them
names(data)[names(data) == "Daily.Time.Spent.on.Site"] <- "daily_time_spent"
names(data)[names(data) == "Age"] <- "age"
names(data)[names(data) == "Area.Income"] <- "area_income"
names(data)[names(data) == "Daily.Internet.Usage"] <- "daily_internet_usage"
names(data)[names(data) == "Ad.Topic.Line"] <- "ad_topic_line"
names(data)[names(data) == "City"] <- "city"
names(data)[names(data) == "Male"] <- "male"
names(data)[names(data) == "Country"] <- "country"
names(data)[names(data) == "Timestamp"] <- "timestamp"
names(data)[names(data) == "Clicked.on.Ad"] <- "clicked_on_ad"

# now previewing to confirm they've been changed
colnames(data)
```

```
##  [1] "daily_time_spent"     "age"                  "area_income"
##  [4] "daily_internet_usage" "ad_topic_line"        "city"
##  [7] "male"                 "country"              "timestamp"
## [10] "clicked_on_ad"
```

```r
# changing the data types of the "male" and "clicked_on_ad" columns from integer to factor
data$male <- as.factor(data$male)
data$clicked_on_ad <- as.factor(data$clicked_on_ad)

str(data$male)
```

```
##  Factor w/ 2 levels "0","1": 1 2 1 2 1 2 1 2 2 2 ...
```

```r
str(data$clicked_on_ad)
```

```
##  Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 2 1 1 ...
```

```r
# split timestamp column into year, month, day, and hour
# NB: minute and second are irrelevant to our analysis
data$year <- format(as.POSIXct(data$timestamp, format="%Y-%m-%d %H:%M:%S"), "%Y")
data$month <- format(as.POSIXct(data$timestamp, format="%Y-%m-%d %H:%M:%S"), "%m")
data$day <- format(as.POSIXct(data$timestamp, format="%Y-%m-%d %H:%M:%S"), "%d")
data$hour <- format(as.POSIXct(data$timestamp, format="%Y-%m-%d %H:%M:%S"), "%H")
head(data)
```

```
##   daily_time_spent 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
## 1 2016-03-27 00:53:11             0 2016    03  27   00
## 2 2016-04-04 01:39:02             0 2016    04  04   01
## 3 2016-03-13 20:35:42             0 2016    03  13   20
## 4 2016-01-10 02:31:19             0 2016    01  10   02
## 5 2016-06-03 03:36:18             0 2016    06  03   03
## 6 2016-05-19 14:30:17             0 2016    05  19   14
```

```
# drop the timestamp column since it is no longer useful
data$timestamp <- NULL
colnames(data)
```

```
##  [1] "daily_time_spent"     "age"                  "area_income"
##  [4] "daily_internet_usage" "ad_topic_line"        "city"
##  [7] "male"                 "country"              "clicked_on_ad"
## [10] "year"                 "month"                "day"
## [13] "hour"
```

```
# check the data types of the new columns
paste("Year:", class(data$year))
```

```
## [1] "Year: character"
```

```
paste("Month:", class(data$month))
```

```
## [1] "Month: character"
```

```
paste("Day:", class(data$day))
```

```
## [1] "Day: character"
```

```
paste("Hour:", class(data$hour))
```

```
## [1] "Hour: character"
```

```
# set the new columns to be of data type Factor
data$year <- as.factor(data$year)
data$month <- as.factor(data$month)
data$day <- as.factor(data$day)
data$hour <- as.factor(data$hour)
```

```
# move the 'clicked_on_ad' column to the end
data <- data[, c(1:8, 10:13, 9)]
head(data)
```

```
##   daily_time_spent 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 year
## 1    Cloned 5thgeneration orchestration    Wrightburgh    0    Tunisia 2016
## 2    Monitored national standardization      West Jodi    1      Nauru 2016
## 3      Organic bottom-line service-desk       Davidton    0 San Marino 2016
## 4 Triple-buffered reciprocal time-frame West Terrifurt    1      Italy 2016
## 5         Robust logistical utilization   South Manuel    0    Iceland 2016
## 6       Sharable client-driven software      Jamieberg    1     Norway 2016
##   month day hour clicked_on_ad
## 1    03  27   00             0
## 2    04  04   01             0
## 3    03  13   20             0
## 4    01  10   02             0
## 5    06  03   03             0
## 6    05  19   14             0
```

```
str(data)
```

```
## 'data.frame':    1000 obs. of  13 variables:
##  $ daily_time_spent    : 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       : Factor w/ 1000 levels "Adaptive 24hour Graphic Interface",..: 92 465 567 904
##  $ city                : Factor w/ 969 levels "Adamsbury","Adamside",..: 962 904 112 940 806 283 47 (
##  $ male                : Factor w/ 2 levels "0","1": 1 2 1 2 1 2 1 2 2 2 ...
##  $ country             : Factor w/ 237 levels "Afghanistan",..: 216 148 185 104 97 159 146 13 83 79
##  $ year                : Factor w/ 1 level "2016": 1 1 1 1 1 1 1 1 1 1 ...
##  $ month               : Factor w/ 7 levels "01","02","03",..: 3 4 3 1 6 5 1 3 4 7 ...
##  $ day                 : Factor w/ 31 levels "01","02","03",..: 27 4 13 10 3 19 28 7 18 11 ...
##  $ hour                : Factor w/ 24 levels "00","01","02",..: 1 2 21 3 4 15 21 2 10 2 ...
##  $ clicked_on_ad       : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 2 1 1 ...
```

From this, we see that there is only one value for year (2016), 7 for month (Jan to July), 31 for day and 24 for hour.

We can now proceed to carry out exploratory data analysis.

3. Exploratory Data Analysis

3.1 Univariate Analysis

3.1.1 Daily Time Spent

```
# calculate mean
mean(data$daily_time_spent)
```

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

```
# calculate median
median(data$daily_time_spent)
```

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

```
# create function to calculate mode since R doesn't have an in-built function to do that
getmode <- function(v) {
   uniqv <- unique(v)
   uniqv[which.max(tabulate(match(v, uniqv)))]
}

# now calling the mode function on our column
getmode(data$daily_time_spent)
```

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

```
# find variance
var(data$daily_time_spent)
```

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

```
# find standard deviation
sd(data$daily_time_spent)
```

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

```
# computing minimum value
min(data$daily_time_spent)
```

```
## [1] 32.6
```

```
# computing maximum value
max(data$daily_time_spent)
```

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

```
# calculate range
#range(data$daily_time_spent)
max(data$daily_time_spent) - min(data$daily_time_spent)
```

```
## [1] 58.83
```

```
# get first quantile
quantile(data$daily_time_spent, 0.25)
```

```
##   25%
## 51.36
```

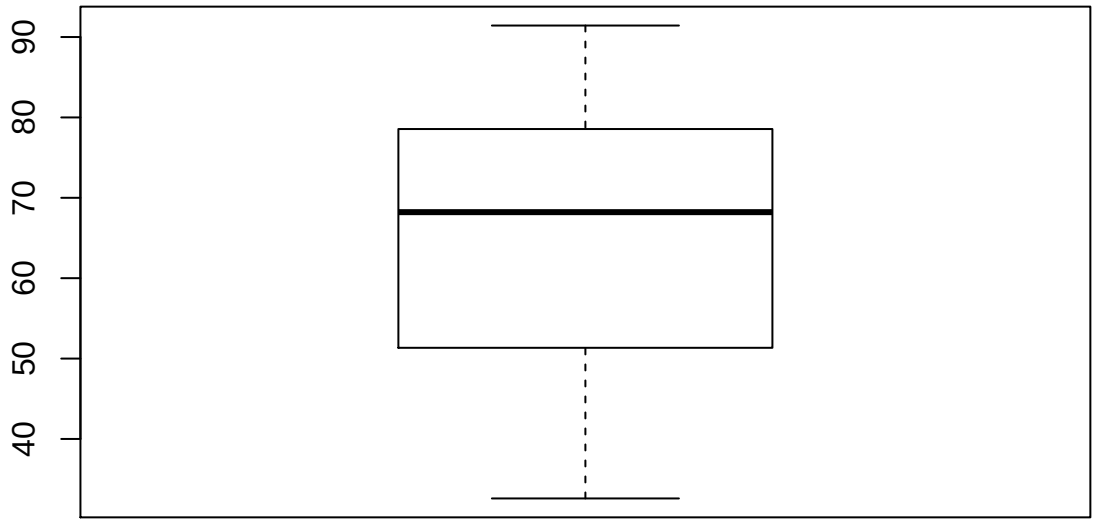
```r
# get third quantile
quantile(data$daily_time_spent, 0.75)
```

```
##     75%
## 78.5475
```

```r
# get interquantile range
quantile(data$daily_time_spent, 0.75) - quantile(data$daily_time_spent, 0.25)
```

```
##     75%
## 27.1875
```

```r
# graph boxplot
boxplot(data$daily_time_spent)
```



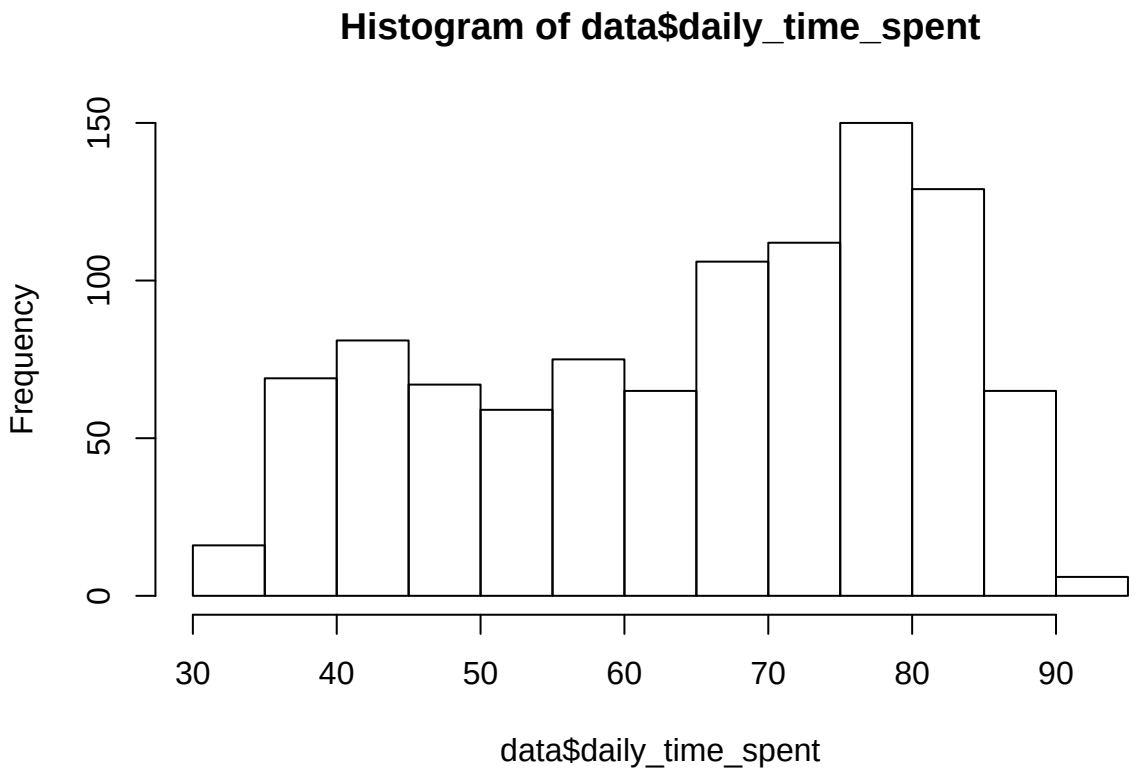
This variable does not have any outliers.

```r
# find the kurtosis of this variable
library(moments)
kurtosis(data$daily_time_spent)
```

```
## [1] 1.903942
```

This kurtosis value is less than 3 implying that the distribution of this variable is platykurtic. This means that there are few to no outliers.

```
# display histogram
hist(data$daily_time_spent)
```



We see that the distribution of the 'daily_time_spent' variable is not normally distributed. It looks to be negatively skewed. We can confirm it by getting the skewness value.

```
skewness(data$daily_time_spent)
```

```
## [1] -0.3712026
```

This proves that this variable is slightly negatively skewed.

3.1.2 Age

To save on time and space, I will use the functions that are like shortcuts to what I've done above.

```
# getting the minimum, maximum, mean, and quartiles
summary(data$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   19.00   29.00   35.00   36.01   42.00   61.00
```

```
# getting mode
getmode(data$age)
```

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

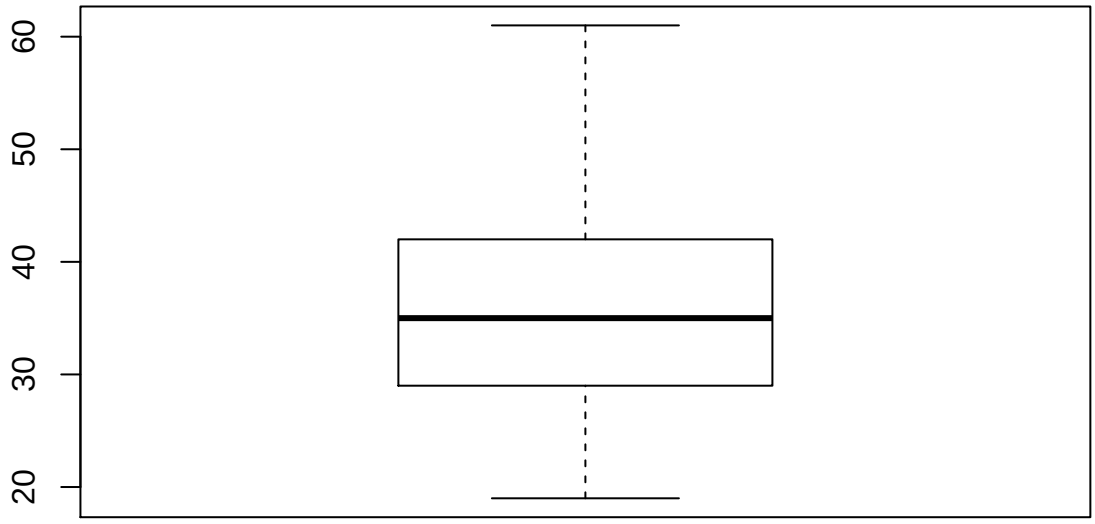
```
# standard deviation
sd(data$age)
```

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

```
# calculate IQR
IQR(data$age)
```

```
## [1] 13
```

```
# check for outliers
boxplot(data$age)
```



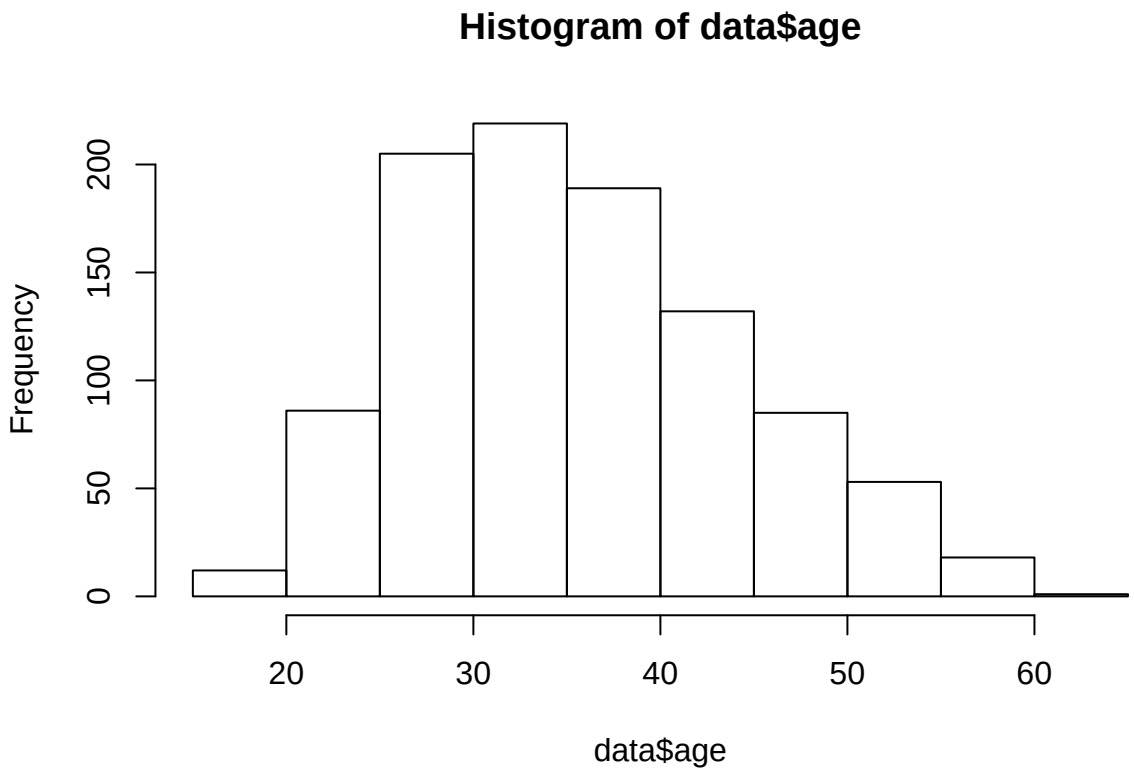
No outliers.

```
# check kurtosis
kurtosis(data$age)
```

```
## [1] 2.595482
```

The distribution is platykurtic implying the existence of few to no outliers.

```r
# check distribution
hist(data$age)
```



The distribution looks almost normal except for the fact that it appears slightly positively skewed. To confirm this, we will test for its skewness.

```r
skewness(data$age)
```

```
## [1] 0.4784227
```

This skewness value implies that the distribution is almost fairly symmetrical, so our initial assumption based on just looking at the visualization of the distribution is slightly wrong.

3.1.3 Area Income

```r
# getting the minimum, maximum, mean, and quartiles
summary(data$area_income)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13996   47032   57012   55000   65471   79485
```

```r
# getting mode
getmode(data$area_income)
```

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

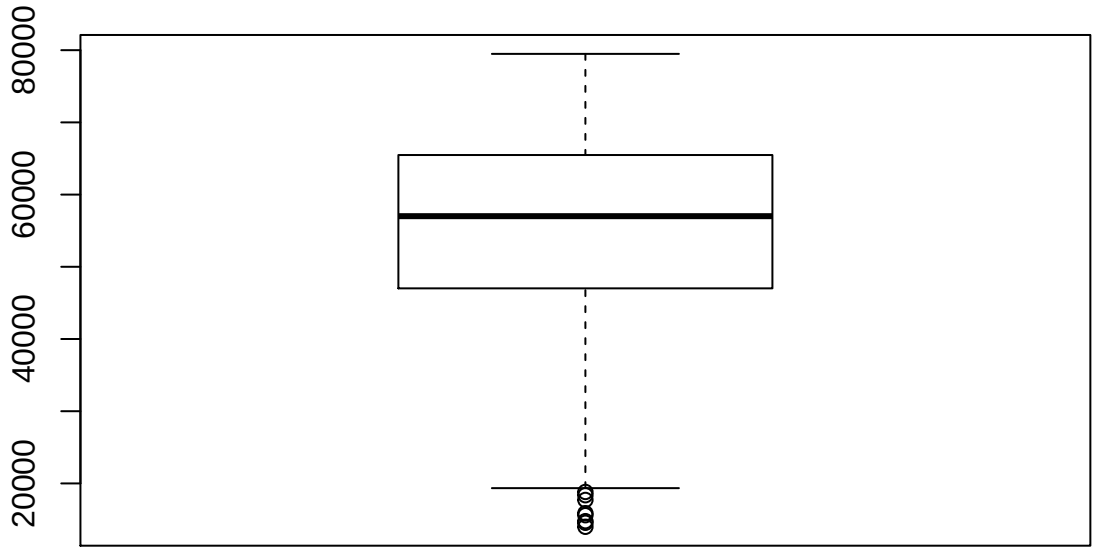
```
# standard deviation
sd(data$area_income)
```

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

```
# calculate IQR
IQR(data$area_income)
```

```
## [1] 18438.83
```

```
# check for outliers
boxplot(data$area_income)
```



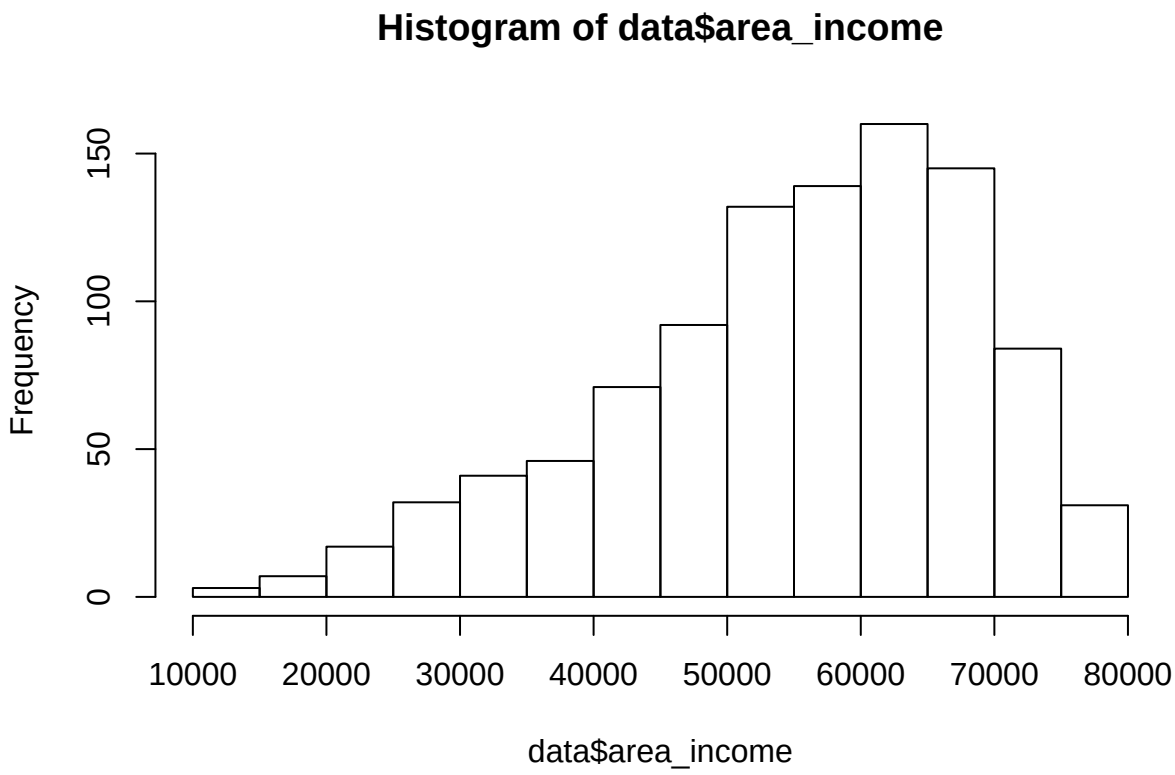
There are outliers below the 20,000 mark. This is to be expected since people's income varies depending on factors such as their employer/company, their position at work, etc.

```
# check kurtosis
kurtosis(data$area_income)
```

```
## [1] 2.894694
```

A kurtosis value of 2.89 indicates that the distribution is platykurtic although it is getting very close to being mesokurtic.

```r
# check distribution
hist(data$area_income)
```



The distribution is negatively skewed.

```r
# check skewness
skewness(data$area_income)
```

```
## [1] -0.6493967
```

3.1.4 Daily Internet Usage

```r
# getting the minimum, maximum, mean, and quartiles
summary(data$daily_internet_usage)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   104.8   138.8   183.1   180.0   218.8   270.0
```

```r
# getting mode
getmode(data$daily_internet_usage)
```

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

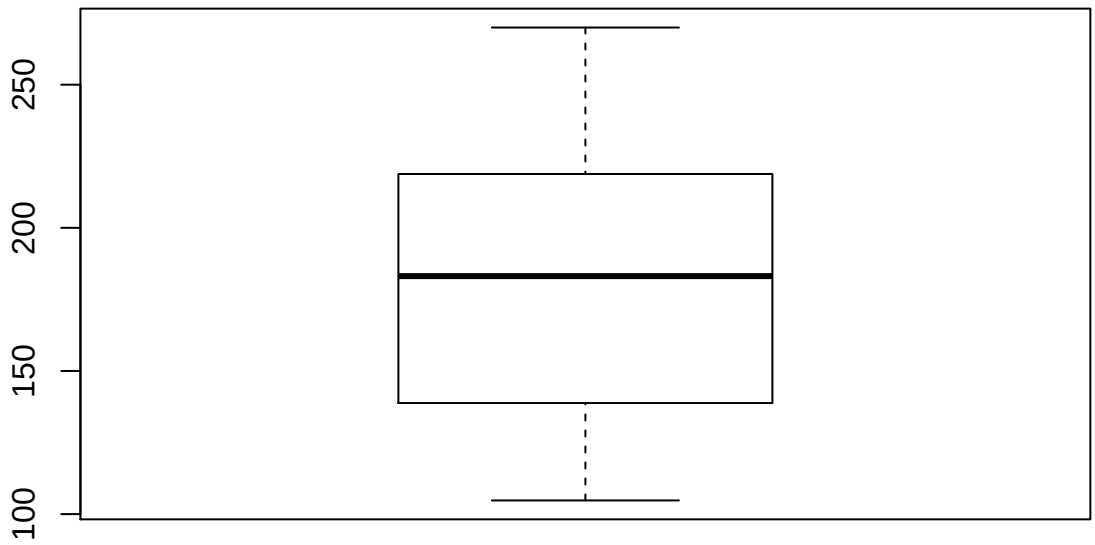
```r
# standard deviation
sd(data$daily_internet_usage)
```

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

```r
# calculate IQR
IQR(data$daily_internet_usage)
```

```
## [1] 79.9625
```

```r
# check for outliers
boxplot(data$daily_internet_usage)
```



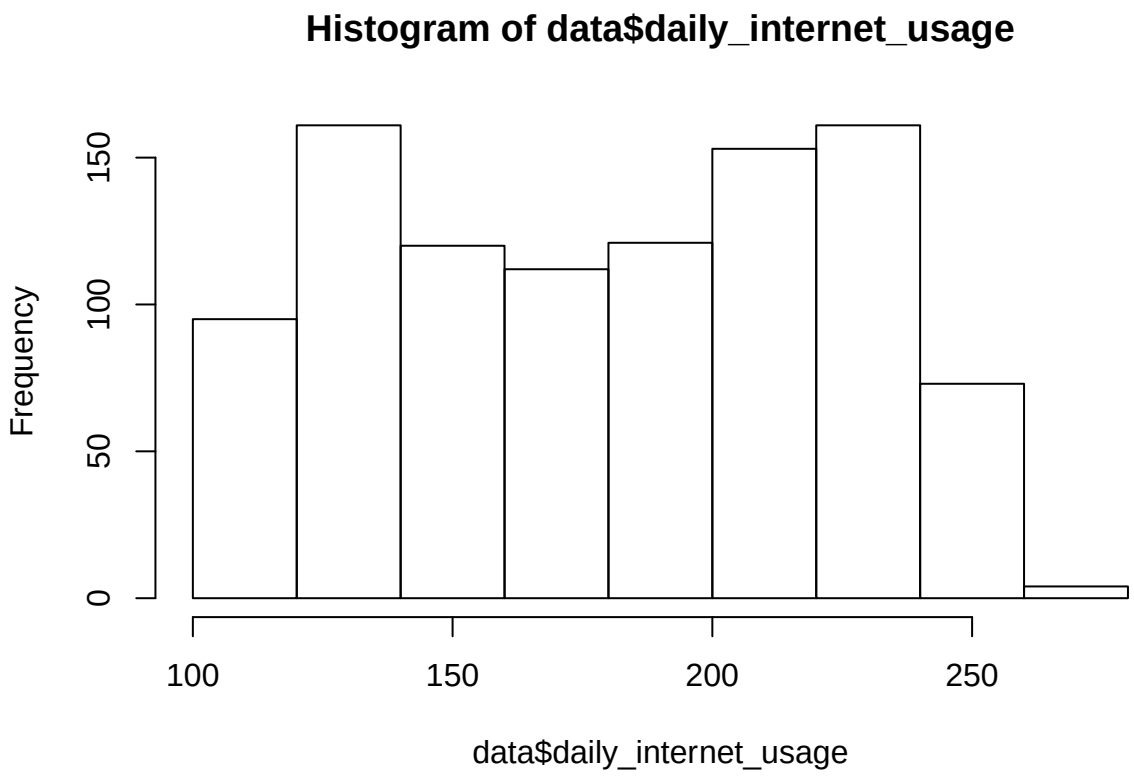
There are no outliers in this column.

```r
# check kurtosis
kurtosis(data$daily_internet_usage)
```

```
## [1] 1.727701
```

The distribution is platykurtic.

```r
# check distribution
hist(data$daily_internet_usage)
```



The distribution appears to be relatively uniform and bimodal.

```r
# check skewness
skewness(data$daily_internet_usage)
```

```
## [1] -0.03348703
```

3.1.5 city

```r
# displaying the first 6 frequently occurring cities
library(plyr)
count_city <- count(data$city)
count_city_head <- head(arrange(count_city, desc(freq)))
count_city_head
```

```
##                 x freq
## 1       Lisamouth    3
## 2    Williamsport    3
## 3 Benjaminchester    2
## 4       East John    2
## 5    East Timothy    2
## 6        Johnstad    2
```

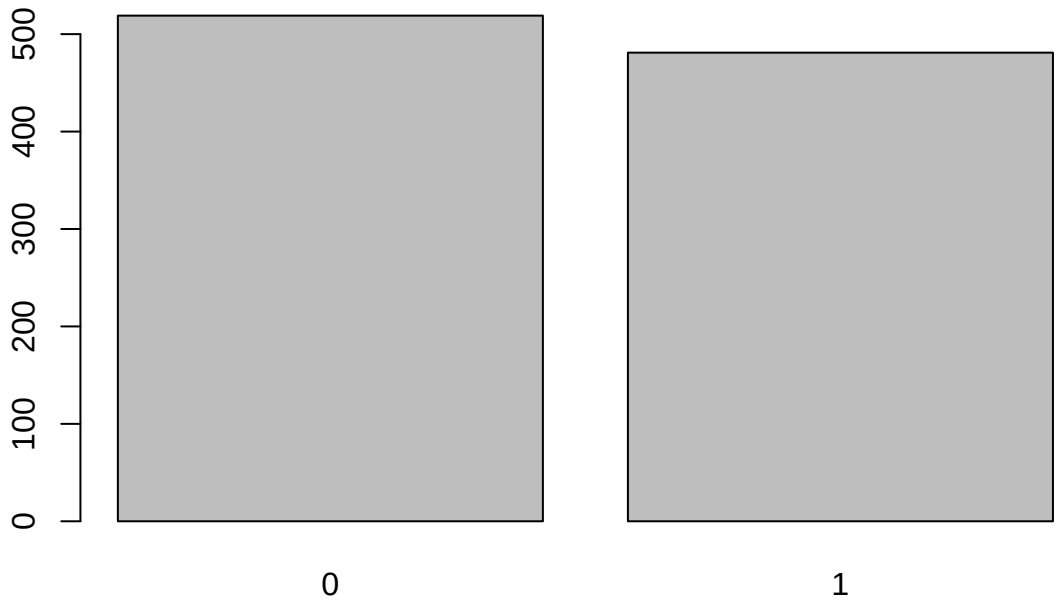
3.1.6 male

```
male_table <- table(data$male)
male_table
```

```
##
##   0   1
## 519 481
```

We see here that 591 are not male while 481 are. To easily visualize this:

```
barplot(male_table)
```



3.1.7 country

```
# displaying the first 10 frequently occuring countries
count_country <- count(data$country)
count_country_head <- head(arrange(count_country, desc(freq)), 10)
count_country_head
```

```
##                x freq
## 1 Czech Republic    9
## 2         France    9
```

```
## 3     Afghanistan    8
## 4       Australia    8
## 5          Cyprus    8
## 6          Greece    8
## 7         Liberia    8
## 8      Micronesia    8
## 9            Peru    8
## 10        Senegal    8
```

3.1.8 month

```
# displaying the months in order of most frequently occurring to least frequently occuring
count_months <- count(data$month)
arrange(count_months, desc(freq))
```

```
##    x freq
## 1 02  160
## 2 03  156
## 3 01  147
## 4 04  147
## 5 05  147
## 6 06  142
## 7 07  101
```

We see here that February is the most frequently occuring month with July being the least frequently occuring month. Could Valentine's Day have something to do with this? LOL.

3.1.9 day

```
# displaying top 5 frequently occuring days
count_days <- count(data$day)
head(arrange(count_days, desc(freq)), 5)
```

```
##    x freq
## 1 03   46
## 2 17   42
## 3 15   41
## 4 10   37
## 5 04   36
```

The 3rd day is the most frequently occuring day overall. However, to get a more accurate picture of this, we will look at which day occurs most frequently in which month. We will do this in bivariate analysis.

```
tail(arrange(count_days, desc(freq)), 1)
```

```
##     x freq
## 31 31   18
```

The 31st day seems to be the least occurring day. Is it because people are splurging since it's end month?

3.1.10 hour

```
# displaying the top 5 hours
count_hours <- count(data$hour)
head(arrange(count_hours, desc(freq)), 5)
```

```
##    x freq
## 1 07   54
## 2 20   50
## 3 09   49
## 4 21   48
## 5 00   45
```

Most frequently occurring time appears to be around 7 AM.

```
tail(arrange(count_hours, desc(freq)), 1)
```

```
##     x freq
## 24 10   31
```

Least frequently occurring time appears to be around 10 AM. This is probably because more people get engrossed in the day's work.

3.1.11 clicked on ad

```
ad_table <- table(data$clicked_on_ad)
print(ad_table)
```

```
##
##   0   1
## 500 500
```

Looks like the number of people who both clicked on the ad and didn't click on the ad is the same (500 each).

3.2 Bivariate Analysis

3.2.1 Research-specific Bivariate Analysis

We will start by looking at the relationship between our target variable (clicked_on_ad) and the other variables.

```
# how many males clicked on ads
ad_male.table <- table(data$clicked_on_ad, data$male)
names(dimnames(ad_male.table)) <- c("Clicked on Ad?", "Male?")
ad_male.table
```

```
##               Male?
## Clicked on Ad?   0   1
##              0 250 250
##              1 269 231
```

From this we see that of those who clicked on the ad, 269 were female while 231 were male. There was no difference in gender of those who did not click on the ad.

```
# ad clicked per month
ad_month.table <- table(data$month, data$clicked_on_ad)
names(dimnames(ad_month.table)) <- c("Month", "Clicked on Ad?")
ad_month.table
```

```
##      Clicked on Ad?
## Month  0  1
##    01 78 69
##    02 77 83
##    03 82 74
##    04 73 74
##    05 68 79
##    06 71 71
##    07 51 50
```

Looking at this table, we see that February reports the highest number of ads clicked and July the least.

```
# ad clicked per day
ad_day.table <- table(data$day, data$clicked_on_ad)
names(dimnames(ad_day.table)) <- c("Day", "Clicked on Ad?")
ad_day.table
```

```
##     Clicked on Ad?
## Day   0  1
##   01 14 19
##   02 15 10
##   03 20 26
##   04 22 14
##   05 17 18
##   06 11 14
##   07 18 14
##   08 20 15
##   09 14 20
##   10 18 19
##   11 17 15
##   12  9 20
##   13 13 17
##   14 12 21
##   15 21 20
##   16 21 14
##   17 24 18
##   18 18 17
##   19 17 12
##   20 22 11
##   21 17 15
##   22 14 10
##   23 13 22
##   24 15 18
##   25  8 15
##   26 21 15
```

```
##   27 19 16
##   28 13 17
##   29 14 15
##   30 14 14
##   31  9  9
```

Day 03 has the highest number of ads clicked. Day 31 has the least.

```
# ad clicked per hour
ad_hour.table <- table(data$hour, data$clicked_on_ad)
names(dimnames(ad_hour.table)) <- c("Hour", "Clicked on Ad?")
ad_hour.table
```

```
##     Clicked on Ad?
## Hour  0  1
##   00 19 26
##   01 16 16
##   02 19 17
##   03 19 23
##   04 21 21
##   05 23 21
##   06 16 23
##   07 28 26
##   08 22 21
##   09 21 28
##   10 17 14
##   11 16 24
##   12 22 16
##   13 21 21
##   14 22 21
##   15 16 19
##   16 23 16
##   17 18 23
##   18 16 25
##   19 20 19
##   20 26 24
##   21 29 19
##   22 24 19
##   23 26 18
```

Hour 09 (9 AM) returned the highest number of ads clicked, 28, whereas Hour 10 (10 AM) returned the lowest, 14.

Improving the solution: create a function that returns the highest and lowest values of a specific column so that you do not have to manually go through each individual record.

```
# ad clicked per country
ad_country.table <- table(data$country, data$clicked_on_ad)
names(dimnames(ad_country.table)) <- c("Country", "Clicked on Ad?")
ad_country.table
```

```
##                                                 Clicked on Ad?
## Country                                          0 1
```

```
##   Afghanistan                                          3 5
##   Albania                                              3 4
##   Algeria                                              3 3
##   American Samoa                                       2 3
##   Andorra                                              0 2
##   Angola                                               3 1
##   Anguilla                                             3 3
##   Antarctica (the territory South of 60 deg S)         1 2
##   Antigua and Barbuda                                  1 4
##   Argentina                                            1 1
##   Armenia                                              2 1
##   Aruba                                                1 0
##   Australia                                            1 7
##   Austria                                              4 1
##   Azerbaijan                                           2 1
##   Bahamas                                              3 4
##   Bahrain                                              3 2
##   Bangladesh                                           2 2
##   Barbados                                             3 2
##   Belarus                                              3 3
##   Belgium                                              3 2
##   Belize                                               2 3
##   Benin                                                1 1
##   Bermuda                                              1 0
##   Bhutan                                               1 1
##   Bolivia                                              6 0
##   Bosnia and Herzegovina                               4 3
##   Bouvet Island (Bouvetoya)                            3 2
##   Brazil                                               2 3
##   British Indian Ocean Territory (Chagos Archipelago)  0 1
##   British Virgin Islands                               2 1
##   Brunei Darussalam                                    3 2
##   Bulgaria                                             2 4
##   Burkina Faso                                         3 1
##   Burundi                                              5 2
##   Cambodia                                             5 2
##   Cameroon                                             5 0
##   Canada                                               2 3
##   Cape Verde                                           1 0
##   Cayman Islands                                       2 3
##   Central African Republic                             1 1
##   Chad                                                 2 2
##   Chile                                                1 3
##   China                                                2 4
##   Christmas Island                                     2 4
##   Colombia                                             1 1
##   Comoros                                              1 1
##   Congo                                                1 3
##   Cook Islands                                         2 1
##   Costa Rica                                           4 2
##   Cote d'Ivoire                                        1 3
##   Croatia                                              6 0
##   Cuba                                                 1 4
##   Cyprus                                               4 4
```

```
##   Czech Republic                                5 4
##   Denmark                                       1 2
##   Djibouti                                      1 1
##   Dominica                                      3 2
##   Dominican Republic                            2 2
##   Ecuador                                       3 2
##   Egypt                                         2 3
##   El Salvador                                   2 4
##   Equatorial Guinea                             1 3
##   Eritrea                                       4 3
##   Estonia                                       2 1
##   Ethiopia                                      0 7
##   Falkland Islands (Malvinas)                   2 2
##   Faroe Islands                                 1 2
##   Fiji                                          4 3
##   Finland                                       4 1
##   France                                        4 5
##   French Guiana                                 1 3
##   French Polynesia                              4 1
##   French Southern Territories                   4 1
##   Gabon                                         6 0
##   Gambia                                        1 1
##   Georgia                                       2 2
##   Germany                                       0 1
##   Ghana                                         2 2
##   Gibraltar                                     3 0
##   Greece                                        5 3
##   Greenland                                     4 1
##   Grenada                                       2 2
##   Guadeloupe                                    1 1
##   Guam                                          2 2
##   Guatemala                                     1 3
##   Guernsey                                      1 2
##   Guinea                                        1 2
##   Guinea-Bissau                                 1 1
##   Guyana                                        2 3
##   Haiti                                         1 1
##   Heard Island and McDonald Islands             1 2
##   Holy See (Vatican City State)                 2 1
##   Honduras                                      3 2
##   Hong Kong                                     2 4
##   Hungary                                       1 5
##   Iceland                                       2 1
##   India                                         2 0
##   Indonesia                                     2 4
##   Iran                                          2 3
##   Ireland                                       2 1
##   Isle of Man                                   2 1
##   Israel                                        2 2
##   Italy                                         4 1
##   Jamaica                                       3 2
##   Japan                                         2 2
##   Jersey                                        2 4
##   Jordan                                        1 0
```

```
##   Kazakhstan                                 2 2
##   Kenya                                      0 4
##   Kiribati                                   0 1
##   Korea                                      2 3
##   Kuwait                                     1 1
##   Kyrgyz Republic                            5 1
##   Lao People's Democratic Republic           2 2
##   Latvia                                     0 4
##   Lebanon                                    2 4
##   Lesotho                                    1 0
##   Liberia                                    2 6
##   Libyan Arab Jamahiriya                     2 2
##   Liechtenstein                              0 6
##   Lithuania                                  0 3
##   Luxembourg                                 4 3
##   Macao                                      0 3
##   Macedonia                                  1 1
##   Madagascar                                 4 2
##   Malawi                                     2 2
##   Malaysia                                   3 0
##   Maldives                                   2 2
##   Mali                                       3 1
##   Malta                                      3 3
##   Marshall Islands                           0 1
##   Martinique                                 1 3
##   Mauritania                                 1 1
##   Mauritius                                  3 1
##   Mayotte                                    1 5
##   Mexico                                     2 4
##   Micronesia                                 4 4
##   Moldova                                    4 2
##   Monaco                                     2 1
##   Mongolia                                   2 4
##   Montenegro                                 0 2
##   Montserrat                                 0 1
##   Morocco                                    2 1
##   Mozambique                                 1 0
##   Myanmar                                    4 1
##   Namibia                                    1 1
##   Nauru                                      2 1
##   Nepal                                      3 0
##   Netherlands                                1 3
##   Netherlands Antilles                       4 2
##   New Caledonia                              0 2
##   New Zealand                                2 2
##   Nicaragua                                  3 0
##   Niger                                      1 2
##   Niue                                       3 0
##   Norfolk Island                             3 2
##   Northern Mariana Islands                   1 2
##   Norway                                     1 1
##   Pakistan                                   4 1
##   Palau                                      2 2
##   Palestinian Territory                      1 2
```

```
##   Panama                                          2 0
##   Papua New Guinea                                2 3
##   Paraguay                                        2 1
##   Peru                                            3 5
##   Philippines                                     3 3
##   Pitcairn Islands                                1 1
##   Poland                                          3 3
##   Portugal                                        2 1
##   Puerto Rico                                     3 3
##   Qatar                                           4 2
##   Reunion                                         2 0
##   Romania                                         0 1
##   Russian Federation                              2 1
##   Rwanda                                          3 2
##   Saint Barthelemy                                0 2
##   Saint Helena                                    3 2
##   Saint Kitts and Nevis                           0 1
##   Saint Lucia                                     1 1
##   Saint Martin                                    2 2
##   Saint Pierre and Miquelon                       2 3
##   Saint Vincent and the Grenadines                3 3
##   Samoa                                           2 4
##   San Marino                                      2 1
##   Sao Tome and Principe                           0 2
##   Saudi Arabia                                    1 3
##   Senegal                                         3 5
##   Serbia                                          2 3
##   Seychelles                                      2 1
##   Sierra Leone                                    0 2
##   Singapore                                       5 1
##   Slovakia (Slovak Republic)                      2 0
##   Slovenia                                        0 1
##   Somalia                                         3 2
##   South Africa                                    2 6
##   South Georgia and the South Sandwich Islands    1 1
##   Spain                                           0 3
##   Sri Lanka                                       4 0
##   Sudan                                           2 0
##   Suriname                                        1 1
##   Svalbard & Jan Mayen Islands                    2 4
##   Swaziland                                       2 0
##   Sweden                                          3 1
##   Switzerland                                     1 3
##   Syrian Arab Republic                            2 1
##   Taiwan                                          3 4
##   Tajikistan                                      1 2
##   Tanzania                                        2 1
##   Thailand                                        2 2
##   Timor-Leste                                     4 1
##   Togo                                            2 1
##   Tokelau                                         1 3
##   Tonga                                           3 2
##   Trinidad and Tobago                             1 2
##   Tunisia                                         3 1
```

```
##   Turkey                                          1 7
##   Turkmenistan                                    4 2
##   Turks and Caicos Islands                        2 3
##   Tuvalu                                          1 3
##   Uganda                                          0 4
##   Ukraine                                         4 1
##   United Arab Emirates                            3 3
##   United Kingdom                                  1 2
##   United States Minor Outlying Islands            2 2
##   United States of America                        2 3
##   United States Virgin Islands                    2 2
##   Uruguay                                         4 1
##   Uzbekistan                                      1 1
##   Vanuatu                                         5 1
##   Venezuela                                       4 3
##   Vietnam                                         1 2
##   Wallis and Futuna                               3 1
##   Western Sahara                                  3 4
##   Yemen                                           1 2
##   Zambia                                          1 3
##   Zimbabwe                                        2 4
```

```r
# ad clicked per city
ad_city.table <- table(data$city, data$clicked_on_ad)
names(dimnames(ad_city.table)) <- c("City", "Clicked on Ad?")
ad_city.table
```

```
##                         Clicked on Ad?
## City                     0 1
##   Adamsbury              0 1
##   Adamside               0 1
##   Adamsstad              1 0
##   Alanview               1 0
##   Alexanderfurt          0 1
##   Alexanderview          0 1
##   Alexandrafort          1 0
##   Alexisland             1 0
##   Aliciatown             0 1
##   Alvaradoport           0 1
##   Alvarezland            0 1
##   Amandafort             0 1
##   Amandahaven            0 1
##   Amandaland             1 0
##   Amyfurt                1 0
##   Amyhaven               1 0
##   Andersonchester        0 1
##   Andersonfurt           0 1
##   Andersonton            1 0
##   Andrewborough          0 1
##   Andrewmouth            1 0
##   Angelhaven             1 0
##   Anthonyfurt            1 0
##   Ashleychester          1 0
##   Ashleymouth            1 0
```

```
##   Austinborough          1 0
##   Austinland             1 0
##   Bakerhaven             1 0
##   Barbershire            1 0
##   Beckton                1 0
##   Benjaminchester        2 0
##   Bernardton             0 1
##   Bethburgh              0 1
##   Birdshire              1 0
##   Blairborough           0 1
##   Blairville             1 0
##   Blevinstown            0 1
##   Bowenview              1 0
##   Boyerberg              0 1
##   Bradleyborough         1 0
##   Bradleyburgh           0 1
##   Bradleyside            0 1
##   Bradshawborough        1 0
##   Bradyfurt              0 1
##   Brandiland             0 1
##   Brandonbury            0 1
##   Brandonstad            1 0
##   Brandymouth            0 1
##   Brendaburgh            1 0
##   Brendachester          0 1
##   Brianabury             1 0
##   Brianfurt              0 1
##   Brianland              0 1
##   Brittanyborough        0 1
##   Brownbury              1 0
##   Brownport              0 1
##   Brownton               0 1
##   Browntown              0 1
##   Brownview              1 0
##   Bruceburgh             1 0
##   Burgessside            0 1
##   Butlerfort             0 1
##   Calebberg              1 0
##   Cameronberg            0 1
##   Campbellstad           1 0
##   Cannonbury             1 0
##   Carsonshire            1 0
##   Carterburgh            1 0
##   Carterland             0 1
##   Carterport             1 0
##   Carterton              1 0
##   Cassandratown          1 0
##   Catherinefort          0 1
##   Cervantesshire         0 1
##   Chapmanland            1 0
##   Chapmanmouth           0 1
##   Charlenetown           0 1
##   Charlesbury            1 0
##   Charlesport            0 1
```

```
##   Charlottefort          0 1
##   Chaseshire             0 1
##   Chrismouth             0 1
##   Christinehaven         0 1
##   Christinetown          0 1
##   Christopherchester     1 0
##   Christopherport        0 1
##   Christopherville       1 0
##   Clarkborough           0 1
##   Claytonside            1 0
##   Clineshire             1 0
##   Codyburgh              0 1
##   Coffeytown             1 0
##   Colebury               0 1
##   Colemanshire           1 0
##   Collinsburgh           1 0
##   Combsstad              0 1
##   Contrerasshire         1 0
##   Costaburgh             0 1
##   Courtneyfort           0 1
##   Coxhaven               1 0
##   Cranemouth             1 0
##   Crawfordfurt           0 1
##   Cunninghamhaven        0 1
##   Curtisport             0 1
##   Curtisview             1 0
##   Cynthiaside            1 0
##   Daisymouth             1 0
##   Danielview             0 1
##   Davidmouth             0 1
##   Davidside              0 1
##   Davidstad              0 1
##   Davidton               1 0
##   Davidview              0 1
##   Daviesborough          1 0
##   Davieshaven            1 0
##   Davilachester          0 1
##   Davisfurt              0 1
##   Dayton                 1 0
##   Deannaville            1 0
##   Debraburgh             0 1
##   Derrickhaven           1 0
##   Destinyfurt            0 1
##   Dianashire             1 0
##   Dianaville             0 1
##   Donaldshire            1 0
##   Douglasview            1 0
##   Duffystad              0 1
##   Dustinborough          1 0
##   Dustinchester          1 0
##   Dustinmouth            0 1
##   East Aaron             1 0
##   East Anthony           0 1
##   East Barbara           0 1
```

```
##   East Benjaminville      1 0
##   East Breannafurt        0 1
##   East Brettton           0 1
##   East Brianberg          1 0
##   East Brittanyville      0 1
##   East Carlos             1 0
##   East Christopher        1 0
##   East Christopherbury    1 0
##   East Connie             1 0
##   East Dana               0 1
##   East Deborahhaven       1 0
##   East Debraborough       1 0
##   East Donna              0 1
##   East Donnatown          1 0
##   East Eric               0 1
##   East Ericport           0 1
##   East Georgeside         0 1
##   East Graceland          1 0
##   East Heatherside        0 1
##   East Heidi              0 1
##   East Henry              1 0
##   East Jason              0 1
##   East Jennifer           1 0
##   East Jessefort          0 1
##   East John               1 1
##   East Johnport           1 0
##   East Kevinbury          0 1
##   East Lindsey            0 1
##   East Maureen            0 1
##   East Michaelland        1 0
##   East Michaelmouth       0 1
##   East Michaeltown        1 0
##   East Michele            0 1
##   East Michelleberg       0 1
##   East Mike               0 1
##   East Paul               1 0
##   East Rachaelfurt        0 1
##   East Rachelview         0 1
##   East Ronald             0 1
##   East Samanthashire      0 1
##   East Sharon             0 1
##   East Shawn              0 1
##   East Shawnchester       1 0
##   East Sheriville         1 0
##   East Stephen            0 1
##   East Susanland          1 0
##   East Tammie             0 1
##   East Theresashire       1 0
##   East Tiffanyport        1 0
##   East Timothy            2 0
##   East Timothyport        1 0
##   East Toddfort           1 0
##   East Troyhaven          1 0
##   East Tylershire         0 1
```

```
##   East Valerie           1 0
##   East Vincentstad       0 1
##   East Yvonnechester     0 1
##   Edwardmouth            1 0
##   Edwardsmouth           1 0
##   Edwardsport            0 1
##   Elizabethbury          0 1
##   Elizabethmouth         1 0
##   Elizabethport          0 1
##   Elizabethstad          0 1
##   Emilyfurt              1 0
##   Ericksonmouth          0 1
##   Erikville              1 0
##   Erinmouth              1 0
##   Erinton                0 1
##   Estesfurt              0 1
##   Estradafurt            1 0
##   Estradashire           0 1
##   Evansfurt              1 0
##   Evansville             0 1
##   Faithview              1 0
##   Florestown             0 1
##   Fosterside             0 1
##   Frankbury              0 1
##   Frankchester           1 0
##   Frankport              0 1
##   Fraziershire           0 1
##   Garciamouth            0 1
##   Garciaside             0 1
##   Garciatown             1 0
##   Garciaview             0 1
##   Garnerberg             1 0
##   Garrettborough         1 0
##   Garychester            1 0
##   Gilbertville           1 0
##   Gomezport              1 0
##   Gonzalezburgh          1 0
##   Grahamberg             0 1
##   Gravesport             1 0
##   Greenechester          1 0
##   Greentown              1 0
##   Greerport              0 1
##   Greerton               1 0
##   Greghaven              1 0
##   Guzmanland             0 1
##   Haleberg               1 0
##   Haleview               1 0
##   Hallfort               1 0
##   Hamiltonfort           0 1
##   Hammondport            1 0
##   Hannahside             1 0
##   Hannaport              0 1
##   Hansenland             0 1
##   Hansenmouth            0 1
```

```
##   Harmonhaven            1 0
##   Harperborough          0 1
##   Harrishaven            1 0
##   Harrisonmouth          1 0
##   Hartmanchester         0 1
##   Hartport               1 0
##   Harveyport             0 1
##   Hatfieldshire          1 0
##   Hawkinsbury            0 1
##   Hayesmouth             1 0
##   Heatherberg            0 1
##   Helenborough           0 1
##   Hendrixmouth           0 1
##   Henryfort              1 0
##   Henryland              0 1
##   Hernandezchester       1 0
##   Hernandezfort          1 0
##   Hernandezside          0 1
##   Hernandezville         0 1
##   Hessstad               1 0
##   Hintonport             0 1
##   Hobbsbury              0 1
##   Holderville            0 1
##   Hollandberg            1 0
##   Hollyfurt              1 0
##   Hubbardmouth           0 1
##   Huffmanchester         0 1
##   Hughesport             0 1
##   Hurleyborough          1 0
##   Ianmouth               1 0
##   Ingramberg             1 0
##   Isaacborough           0 1
##   Jacksonburgh           0 1
##   Jacksonmouth           1 0
##   Jacksonstad            0 1
##   Jacobstad              0 1
##   Jacquelineshire        0 1
##   Jamesberg              1 0
##   Jamesfurt              0 1
##   Jamesmouth             0 1
##   Jamesville             1 0
##   Jamieberg              1 0
##   Jamiefort              1 0
##   Janiceview             1 0
##   Jasminefort            1 0
##   Jayville               1 0
##   Jeffreyburgh           0 1
##   Jeffreymouth           0 1
##   Jeffreyshire           1 0
##   Jenniferhaven          0 1
##   Jenniferstad           1 0
##   Jensenborough          0 1
##   Jensenton              0 1
##   Jeremybury             0 1
```

```
##   Jeremyshire            1 0
##   Jessicahaven           0 1
##   Jessicashire           0 1
##   Jessicastad            0 1
##   Joanntown              1 0
##   Joechester             0 1
##   Johnport               1 0
##   Johnsonfort            1 0
##   Johnsontown            0 1
##   Johnsonview            0 1
##   Johnsport              1 0
##   Johnstad               2 0
##   Johnstonmouth          0 1
##   Johnstonshire          1 0
##   Jonathanland           0 1
##   Jonathantown           0 1
##   Jonesland              1 0
##   Jonesmouth             1 0
##   Jonesshire             0 1
##   Joneston               1 1
##   Jordanmouth            1 0
##   Jordanshire            0 1
##   Jordantown             0 1
##   Josephberg             0 1
##   Josephmouth            0 1
##   Josephstad             0 1
##   Joshuaburgh            1 0
##   Joshuamouth            1 0
##   Juanport               1 0
##   Juliaport              1 0
##   Julietown              0 1
##   Karenmouth             1 0
##   Karenton               1 0
##   Katieport              0 1
##   Kaylashire             1 0
##   Keithtown              0 1
##   Kellytown              1 0
##   Kennedyfurt            1 0
##   Kennethview            1 0
##   Kentmouth              0 1
##   Kevinberg              0 1
##   Kevinchester           1 0
##   Kimberlyhaven          1 0
##   Kimberlymouth          0 1
##   Kimberlytown           1 0
##   Kingchester            0 1
##   Kingshire              1 0
##   Klineside              0 1
##   Knappburgh             1 0
##   Kristineberg           1 0
##   Kristinfurt            0 1
##   Kristintown            0 1
##   Kyleborough            0 1
##   Kylieview              1 0
```

```
##   Lake Adrian             1 0
##   Lake Allenville         0 1
##   Lake Amanda             0 1
##   Lake Amy                1 0
##   Lake Angela             1 0
##   Lake Annashire          1 0
##   Lake Beckyburgh         0 1
##   Lake Brandonview        0 1
##   Lake Brian              1 0
##   Lake Cassandraport      0 1
##   Lake Charlottestad      0 1
##   Lake Christopherfurt    0 1
##   Lake Conniefurt         0 1
##   Lake Courtney           1 0
##   Lake Craigview          0 1
##   Lake Cynthia            1 0
##   Lake Danielle           1 0
##   Lake David              0 2
##   Lake Deannaborough      1 0
##   Lake Deborahburgh       1 0
##   Lake Dustin             0 1
##   Lake Edward             0 1
##   Lake Elizabethside      1 0
##   Lake Evantown           0 1
##   Lake Faith              0 1
##   Lake Gerald             0 1
##   Lake Hailey             1 0
##   Lake Ian                0 1
##   Lake Jacob              1 0
##   Lake Jacqueline         1 0
##   Lake James              0 2
##   Lake Jasonchester       1 0
##   Lake Jennifer           0 1
##   Lake Jenniferton        1 0
##   Lake Jessica            0 1
##   Lake Jessicaville       0 1
##   Lake Jesus              0 1
##   Lake Jillville          1 0
##   Lake John               0 1
##   Lake Johnbury           0 1
##   Lake Jonathanview       1 0
##   Lake Jose               1 1
##   Lake Joseph             1 0
##   Lake Josetown           1 0
##   Lake Joshuafurt         0 1
##   Lake Kevin              1 0
##   Lake Kurtmouth          1 0
##   Lake Lisa               1 0
##   Lake Matthew            0 1
##   Lake Matthewland        1 0
##   Lake Melindamouth       1 0
##   Lake Michael            1 0
##   Lake Michaelport        1 0
##   Lake Michelle           0 1
```

```
##   Lake Michellebury       0 1
##   Lake Nicole             1 0
##   Lake Patrick            2 0
##   Lake Rhondaburgh        0 1
##   Lake Stephenborough     0 1
##   Lake Susan              1 1
##   Lake Timothy            1 0
##   Lake Tracy              0 1
##   Lake Vanessa            0 1
##   Lake Zacharyfurt        1 0
##   Lauraburgh              1 0
##   Laurieside              1 0
##   Lawrenceborough         1 0
##   Lawsonshire             0 1
##   Leahside                0 1
##   Leonchester             1 0
##   Lesliebury              0 1
##   Lesliefort              1 0
##   Lewismouth              0 1
##   Lindaside               1 0
##   Lindsaymouth            1 0
##   Lisaberg                1 0
##   Lisafort                1 0
##   Lisamouth               1 2
##   Lopezberg               0 1
##   Lopezmouth              1 0
##   Loriville               0 1
##   Lovemouth               0 1
##   Luischester             1 0
##   Luisfurt                1 0
##   Lukeport                1 0
##   Mackenziemouth          1 0
##   Marcushaven             1 0
##   Mariahview              0 1
##   Mariebury               1 0
##   Mariemouth              1 0
##   Markhaven               0 1
##   Masonhaven              1 0
##   Masseyshire             0 1
##   Mataberg                1 0
##   Matthewtown             0 1
##   Mauricefurt             0 1
##   Mauriceshire            1 0
##   Mcdonaldfort            1 0
##   Mclaughlinbury          1 0
##   Meaganfort              1 0
##   Meghanchester           0 1
##   Melanieton              0 1
##   Melissachester          0 1
##   Melissafurt             1 0
##   Melissastad             1 0
##   Meyerchester            1 0
##   Meyersstad              0 1
##   Mezaton                 0 1
```

```
##   Michaelland            1 0
##   Michaelmouth           1 0
##   Michaelshire           0 1
##   Micheletown            0 1
##   Michellefort           0 1
##   Michelleside           0 2
##   Millerbury             0 2
##   Millerchester          0 1
##   Millerfort             1 0
##   Millerland             1 0
##   Millerside             0 1
##   Millertown             1 1
##   Millerview             1 0
##   Mollyport              1 0
##   Monicaview             0 1
##   Morganfort             1 0
##   Morganport             0 1
##   Morrismouth            0 1
##   Mosleyburgh            1 0
##   Mullenside             1 0
##   Munozberg              1 0
##   Murphymouth            1 0
##   Nelsonfurt             0 1
##   New Amanda             0 1
##   New Angelview          0 1
##   New Brandy             1 0
##   New Brendafurt         0 1
##   New Charleschester     0 1
##   New Christinatown      0 1
##   New Cynthia            1 0
##   New Daniellefort       0 1
##   New Darlene            0 1
##   New Dawnland           1 0
##   New Debbiestad         0 1
##   New Denisebury         0 1
##   New Frankshire         1 0
##   New Gabriel            1 0
##   New Henry              0 1
##   New Hollyberg          0 1
##   New James              0 1
##   New Jamestown          1 0
##   New Jasmine            1 0
##   New Jay                0 1
##   New Jeffreychester     1 0
##   New Jessicaport        2 0
##   New Johnberg           1 0
##   New Joshuaport         0 1
##   New Juan               1 0
##   New Julianberg         0 1
##   New Julie              1 0
##   New Karenberg          0 1
##   New Kayla              1 0
##   New Keithburgh         0 1
##   New Lindaberg          0 1
```

```
##   New Lucasburgh          0 1
##   New Marcusbury          0 1
##   New Maria               1 0
##   New Matthew             0 1
##   New Michael             0 1
##   New Michaeltown         1 0
##   New Nancy               0 1
##   New Nathan              1 0
##   New Patriciashire       1 0
##   New Patrick             0 1
##   New Paul                1 0
##   New Rachel              0 1
##   New Rebecca             0 1
##   New Sabrina             0 1
##   New Sean                1 0
##   New Shane               1 0
##   New Sharon              1 0
##   New Sheila              2 0
##   New Sonialand           1 0
##   New Steve               1 0
##   New Tammy               0 1
##   New Taylorburgh         1 0
##   New Teresa              0 1
##   New Theresa             0 1
##   New Thomas              0 1
##   New Timothy             0 1
##   New Tina                0 1
##   New Tinamouth           1 0
##   New Traceystad          1 0
##   New Travis              1 0
##   New Travistown          0 1
##   New Tyler               1 0
##   New Wanda               1 0
##   New Williammouth        0 1
##   New Williamville        0 1
##   Newmanberg              1 0
##   Nicholasland            0 1
##   Nicholasport            1 0
##   North Aaronburgh        0 1
##   North Aaronchester      0 1
##   North Alexandra         1 0
##   North Anaport           1 0
##   North Andrew            0 1
##   North Andrewstad        0 1
##   North Angelastad        0 1
##   North Angelatown        0 1
##   North Anna              1 0
##   North April             0 1
##   North Brandon           1 0
##   North Brittanyburgh     0 1
##   North Cassie            0 1
##   North Charlesbury       0 1
##   North Christopher       1 0
##   North Daniel            1 1
```

```
##   North Debra             1 0
##   North Debrashire        0 1
##   North Derekville        0 1
##   North Destiny           0 1
##   North Elizabeth         1 0
##   North Frankstad         1 0
##   North Garyhaven         1 0
##   North Isabellaville     1 0
##   North Jenniferburgh     0 1
##   North Jeremyport        1 0
##   North Jessicaville      0 1
##   North Johnside          1 0
##   North Johntown          0 1
##   North Jonathan          0 1
##   North Joshua            1 0
##   North Katie             0 1
##   North Kennethside       1 0
##   North Kevinside         0 1
##   North Kimberly          0 1
##   North Kristine          1 0
##   North Lauraland         0 1
##   North Laurenview        1 0
##   North Leonmouth         1 0
##   North Lisachester       1 0
##   North Loriburgh         1 0
##   North Mark              0 1
##   North Maryland          0 1
##   North Mercedes          0 1
##   North Michael           0 1
##   North Monicaville       1 0
##   North Randy             1 0
##   North Raymond           1 0
##   North Regina            0 1
##   North Ricardotown       0 1
##   North Richardburgh      0 1
##   North Ronaldshire       1 0
##   North Russellborough    0 1
##   North Samantha          0 1
##   North Sarashire         0 1
##   North Shannon           1 0
##   North Stephanieberg     1 0
##   North Tara              1 0
##   North Tiffany           1 0
##   North Tracyport         1 0
##   North Tylerland         1 0
##   North Virginia          0 1
##   North Wesleychester     1 0
##   Novaktown               1 0
##   Odomville               1 0
##   Olsonside               0 1
##   Olsonstad               0 1
##   Palmerside              0 1
##   Pamelamouth             2 0
##   Parkerhaven             1 0
```

```
##   Patriciahaven           1 0
##   Patrickmouth            1 0
##   Pattymouth              0 1
##   Paulhaven               1 0
##   Paulport                1 0
##   Paulshire               1 0
##   Pearsonfort             1 0
##   Penatown                0 1
##   Perezland               1 0
##   Perryburgh              0 1
##   Petersonfurt            0 1
##   Phelpschester           1 0
##   Philipberg              0 1
##   Phillipsbury            0 1
##   Port Aliciabury         1 0
##   Port Angelamouth        0 1
##   Port Anthony            1 0
##   Port Aprilville         0 1
##   Port Beth               0 1
##   Port Blake              0 1
##   Port Brenda             0 1
##   Port Brian              0 1
##   Port Brianfort          1 0
##   Port Brittanyville      1 0
##   Port Brookeland         0 1
##   Port Calvintown         1 0
##   Port Cassie             0 1
##   Port Chasemouth         1 0
##   Port Christina          0 1
##   Port Christinemouth     1 0
##   Port Christopher        0 1
##   Port Christopherborough 0 1
##   Port Crystal            0 1
##   Port Daniel             1 0
##   Port Danielleberg       1 0
##   Port Davidland          1 0
##   Port Dennis             0 1
##   Port Derekberg          0 1
##   Port Destiny            1 0
##   Port Douglasborough     0 1
##   Port Elijah             1 0
##   Port Eric               0 1
##   Port Erikhaven          0 1
##   Port Erinberg           0 1
##   Port Eugeneport         1 0
##   Port Georgebury         0 1
##   Port Gregory            1 0
##   Port Jacqueline         1 0
##   Port Jacquelinestad     1 0
##   Port James              1 0
##   Port Jasmine            1 0
##   Port Jason              1 1
##   Port Jefferybury        0 1
##   Port Jeffrey            1 0
```

```
##   Port Jennifer           0 1
##   Port Jessica            0 1
##   Port Jessicamouth       1 0
##   Port Jodi               1 0
##   Port Joshuafort         0 1
##   Port Juan               1 1
##   Port Julie              1 1
##   Port Karenfurt          1 0
##   Port Katelynview        0 1
##   Port Kathleenfort       0 1
##   Port Kevinborough       1 0
##   Port Lawrence           0 1
##   Port Maria              1 0
##   Port Mathew             1 0
##   Port Melissaberg        0 1
##   Port Melissastad        1 0
##   Port Michaelmouth       0 1
##   Port Michealburgh       0 1
##   Port Mitchell           0 1
##   Port Patrickton         0 1
##   Port Paultown           0 1
##   Port Rachel             0 1
##   Port Raymondfort        1 0
##   Port Robin              1 0
##   Port Sarahhaven         0 1
##   Port Sarahshire         0 1
##   Port Sherrystad         0 1
##   Port Stacey             1 0
##   Port Stacy              1 0
##   Port Susan              1 0
##   Port Whitneyhaven       1 0
##   Portermouth             1 0
##   Pottermouth             0 1
##   Princebury              1 0
##   Pruittmouth             1 0
##   Rachelhaven             1 0
##   Ramirezhaven            0 1
##   Ramirezland             1 0
##   Ramirezside             0 1
##   Ramirezton              1 0
##   Ramosstad               1 0
##   Randolphport            1 0
##   Randyshire              1 0
##   Rebeccamouth            0 1
##   Reginamouth             0 1
##   Reneechester            0 1
##   Reyesfurt               1 0
##   Reyesland               1 0
##   Rhondaborough           1 0
##   Richardshire            0 1
##   Richardsland            1 0
##   Richardsonland          0 1
##   Richardsonmouth         1 0
##   Richardsonshire         0 1
```

```
##   Richardsontown          1 0
##   Rickymouth              1 0
##   Riggsstad               1 0
##   Rivasland               0 1
##   Robertbury              1 0
##   Robertfurt              0 2
##   Robertmouth             1 0
##   Robertside              0 1
##   Robertsonburgh          0 1
##   Robertstown             0 1
##   Roberttown              0 1
##   Robinsonland            1 0
##   Robinsontown            0 1
##   Rochabury               0 1
##   Rogerburgh              0 1
##   Rogerland               1 0
##   Ronaldport              0 1
##   Ronniemouth             0 1
##   Russellville            0 1
##   Ryanhaven               0 1
##   Sabrinaview             1 0
##   Salazarbury             0 1
##   Samanthaland            0 1
##   Samuelborough           1 0
##   Sanchezland             1 0
##   Sanchezmouth            1 0
##   Sandersland             1 0
##   Sanderstown             0 1
##   Sandraland              1 0
##   Sandrashire             0 1
##   Sandraville             1 0
##   Sarafurt                1 0
##   Sarahland               0 1
##   Sarahton                1 0
##   Sellerstown             1 0
##   Shaneland               1 0
##   Sharpberg               1 0
##   Shawnside               1 0
##   Shawstad                1 0
##   Shelbyport              1 1
##   Sherrishire             1 0
##   Shirleyfort             1 0
##   Silvaton                0 1
##   Smithburgh              1 0
##   Smithside               0 1
##   Smithtown               1 0
##   South Aaron             0 1
##   South Adam              0 1
##   South Adamhaven         1 0
##   South Alexisborough     0 1
##   South Blakestad         1 0
##   South Brian             1 0
##   South Cathyfurt         0 1
##   South Christopher       1 0
```

```
##   South Corey             1 0
##   South Cynthiashire      0 1
##   South Daniel            0 1
##   South Daniellefort      1 0
##   South Davidhaven        0 1
##   South Davidmouth        0 1
##   South Denise            1 0
##   South Denisefurt        1 0
##   South Dianeshire        1 0
##   South George            0 1
##   South Henry             0 1
##   South Jackieberg        0 1
##   South Jade              0 1
##   South Jaimeview         1 0
##   South Jasminebury       0 1
##   South Jeanneport        0 1
##   South Jennifer          1 0
##   South Jessica           0 1
##   South John              0 1
##   South Johnnymouth       0 1
##   South Kyle              0 1
##   South Lauraton          0 1
##   South Lauratown         0 1
##   South Lisa              0 2
##   South Manuel            1 0
##   South Margaret          0 1
##   South Mark              0 1
##   South Meghan            0 1
##   South Meredithmouth     1 0
##   South Pamela            1 0
##   South Patrickfort       1 0
##   South Peter             0 1
##   South Rebecca           0 1
##   South Renee             1 0
##   South Robert            1 0
##   South Ronald            1 0
##   South Stephanieport     1 0
##   South Tiffanyton        0 1
##   South Tomside           1 0
##   South Troy              1 0
##   South Vincentchester    0 1
##   South Walter            0 1
##   Staceyfort              0 1
##   Stephenborough          1 0
##   Stewartbury             1 0
##   Suzannetown             0 1
##   Sylviaview              1 0
##   Tammymouth              0 1
##   Tammyshire              0 1
##   Taylorberg              1 0
##   Taylorhaven             0 1
##   Taylormouth             0 1
##   Taylorport              1 0
##   Teresahaven             1 0
```

```
##   Thomasstad             1 0
##   Thomasview             1 0
##   Timothyfurt            0 1
##   Timothymouth           0 1
##   Timothyport            0 1
##   Timothytown            1 0
##   Tinachester            1 0
##   Tinaton                0 1
##   Townsendfurt           1 0
##   Tracyhaven             0 1
##   Tranland               1 0
##   Troyville              1 0
##   Turnerchester          0 1
##   Turnerview             1 0
##   Turnerville            1 0
##   Tylerport              0 1
##   Valerieland            1 0
##   Vanessastad            0 1
##   Vanessaview            0 1
##   Villanuevastad         1 0
##   Villanuevaton          1 0
##   Wademouth              1 0
##   Wadestad               1 0
##   Wagnerchester          1 0
##   Wallacechester         1 0
##   Walshhaven             1 0
##   Waltertown             0 1
##   Watsonfort             1 0
##   Welchshire             0 1
##   Wendyton               1 0
##   Wendyville             0 1
##   West Alice             1 0
##   West Alyssa            1 0
##   West Amanda            0 2
##   West Andrew            1 0
##   West Angela            1 0
##   West Angelabury        1 0
##   West Annefort          0 1
##   West Aprilport         0 1
##   West Arielstad         1 0
##   West Barbara           1 0
##   West Benjamin          1 0
##   West Brad              0 1
##   West Brandonton        0 1
##   West Brenda            1 0
##   West Carmenfurt        1 0
##   West Casey             0 1
##   West Chloeborough      0 1
##   West Christopher       0 1
##   West Colin             1 0
##   West Connor            0 1
##   West Courtney          1 0
##   West Daleborough       1 0
##   West Dannyberg         1 0
```

```
##   West David             0 1
##   West Dennis            1 0
##   West Derekmouth        0 1
##   West Dylanberg         0 1
##   West Eduardotown       0 1
##   West Ericaport         0 1
##   West Ericfurt          0 1
##   West Gabriellamouth    0 1
##   West Gregburgh         1 0
##   West Guybury           1 0
##   West James             0 1
##   West Jane              0 1
##   West Jeremyside        0 1
##   West Jessicahaven      0 1
##   West Jodi              1 0
##   West Joseph            1 0
##   West Julia             0 1
##   West Justin            0 1
##   West Katiefurt         0 1
##   West Kevinfurt         0 1
##   West Lacey             1 0
##   West Leahton           0 1
##   West Lindseybury       0 1
##   West Lisa              1 0
##   West Lucas             1 0
##   West Mariafort         1 0
##   West Melaniefurt       0 1
##   West Melissashire      0 1
##   West Michaelhaven      1 0
##   West Michaelport       1 0
##   West Michaelshire      1 0
##   West Michaelstad       1 0
##   West Pamela            0 1
##   West Randy             0 1
##   West Raymondmouth      0 1
##   West Rhondamouth       1 0
##   West Ricardo           0 1
##   West Richard           0 1
##   West Robertside        1 0
##   West Roytown           1 0
##   West Russell           1 0
##   West Ryan              0 1
##   West Samantha          1 0
##   West Shannon           0 2
##   West Sharon            1 0
##   West Shaun             1 0
##   West Steven            2 0
##   West Sydney            1 0
##   West Tanner            1 0
##   West Tanya             0 1
##   West Terrifurt         1 0
##   West Thomas            1 0
##   West Tinashire         0 1
##   West Travismouth       0 1
```

```
##   West Wendyland          1 0
##   West William            0 1
##   West Zacharyborough     1 0
##   Westshire               0 1
##   Whiteport               0 1
##   Whitneyfort             1 0
##   Wilcoxport              0 1
##   Williammouth            0 1
##   Williamport             1 0
##   Williamsborough         0 1
##   Williamsfort            0 1
##   Williamsmouth           0 1
##   Williamsport            1 2
##   Williamsside            1 0
##   Williamstad             0 1
##   Wilsonburgh             1 0
##   Wintersfort             1 0
##   Wongland                1 0
##   Wrightburgh             2 0
##   Wrightview              0 1
##   Yangside                0 1
##   Youngburgh              1 0
##   Youngfort               0 1
##   Yuton                   0 1
##   Zacharystad             1 0
##   Zacharyton              0 1
```

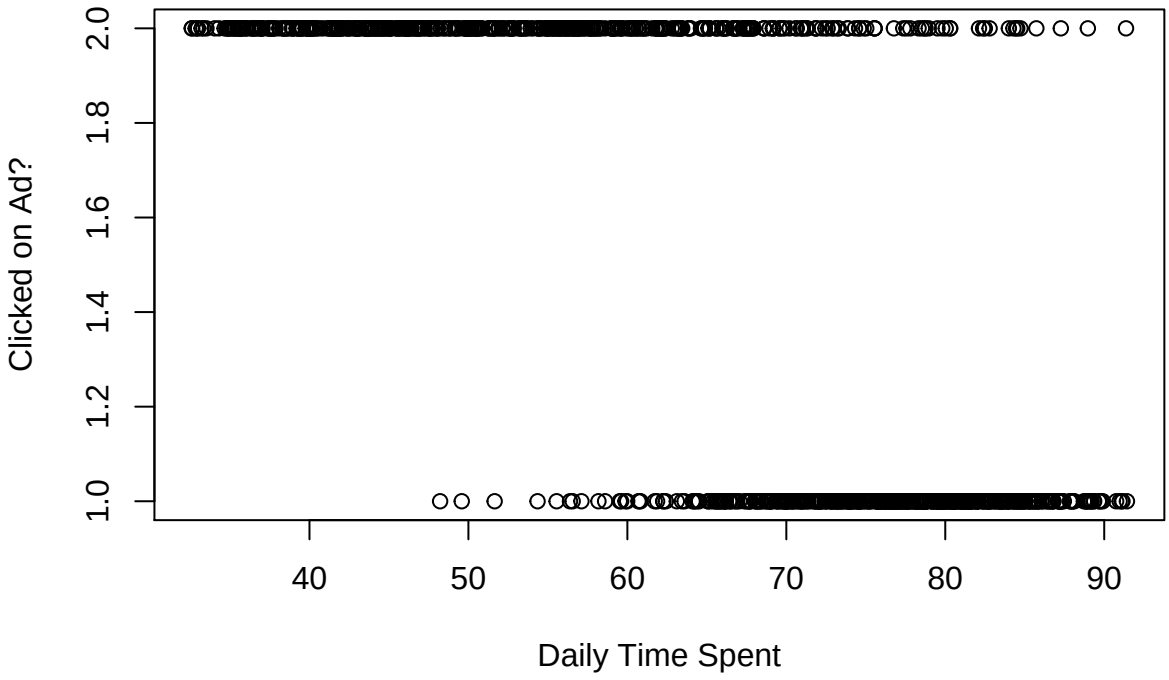
```
head(data$clicked_on_ad, 10)
```

```
##  [1] 0 0 0 0 0 0 0 1 0 0
## Levels: 0 1
```

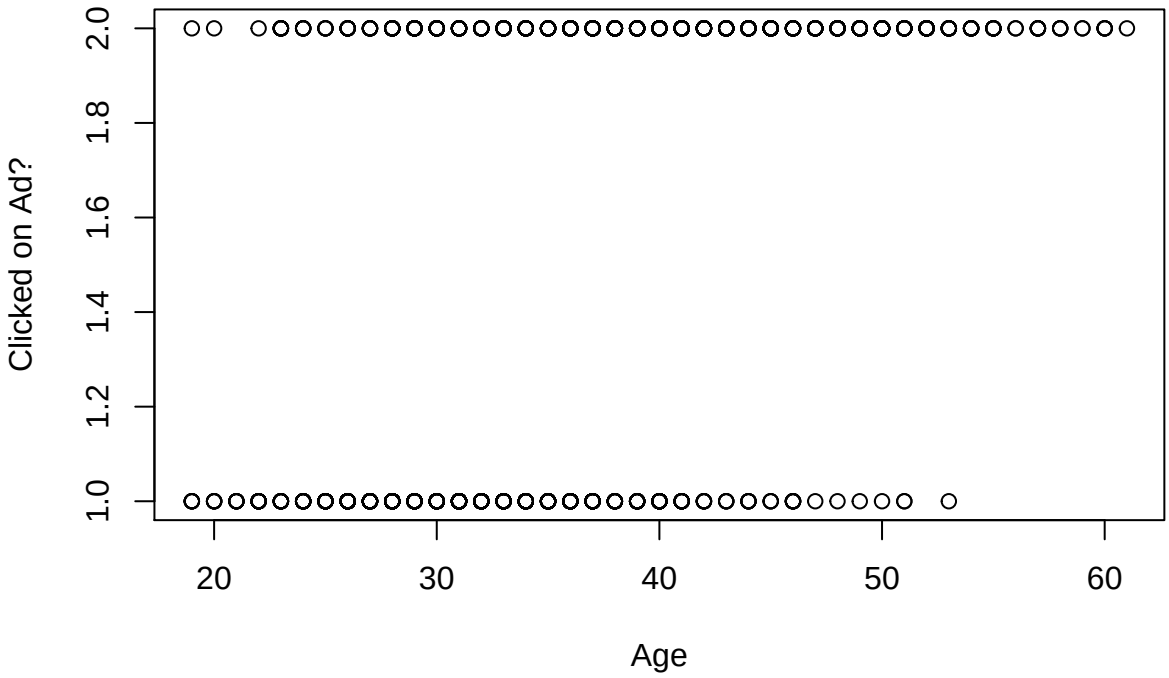
```
# to work with the other variables which are numeric in nature, we will convert the clicked_on_ad varia
ad_int <- as.integer(data$clicked_on_ad)
head(ad_int, 10)
```

```
##  [1] 1 1 1 1 1 1 1 2 1 1
```

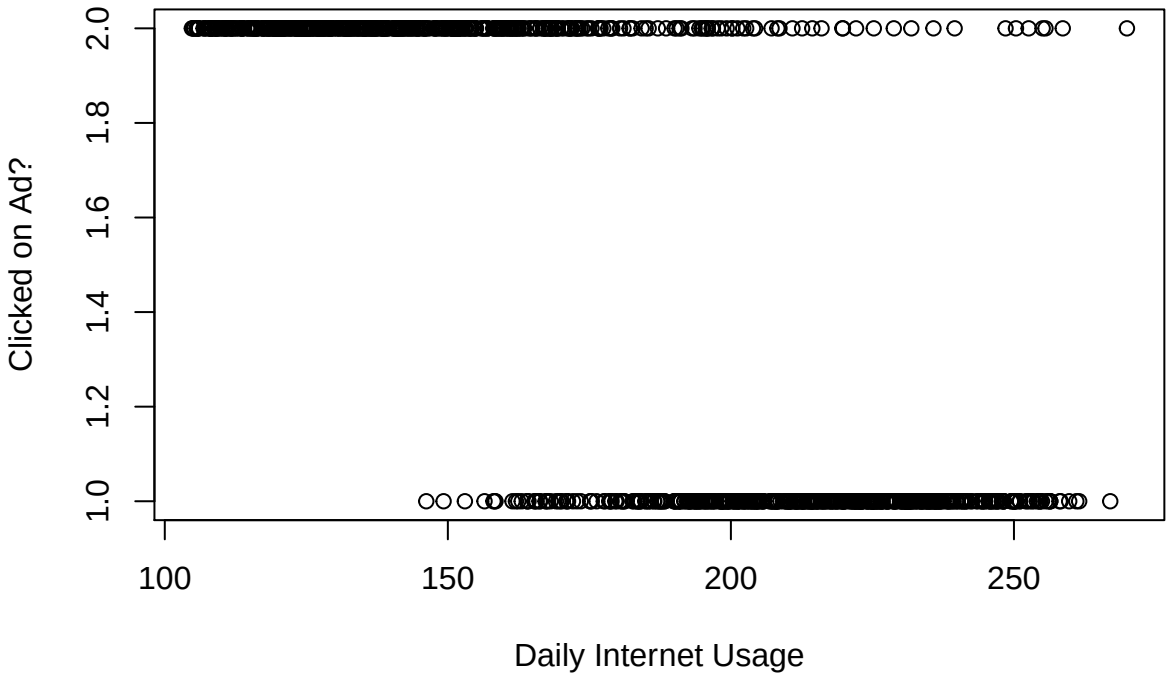
```
# scatter plot of how daily time spent impacts ad being clicked
plot(data$daily_time_spent, ad_int, ylab = "Clicked on Ad?", xlab = "Daily Time Spent")
```



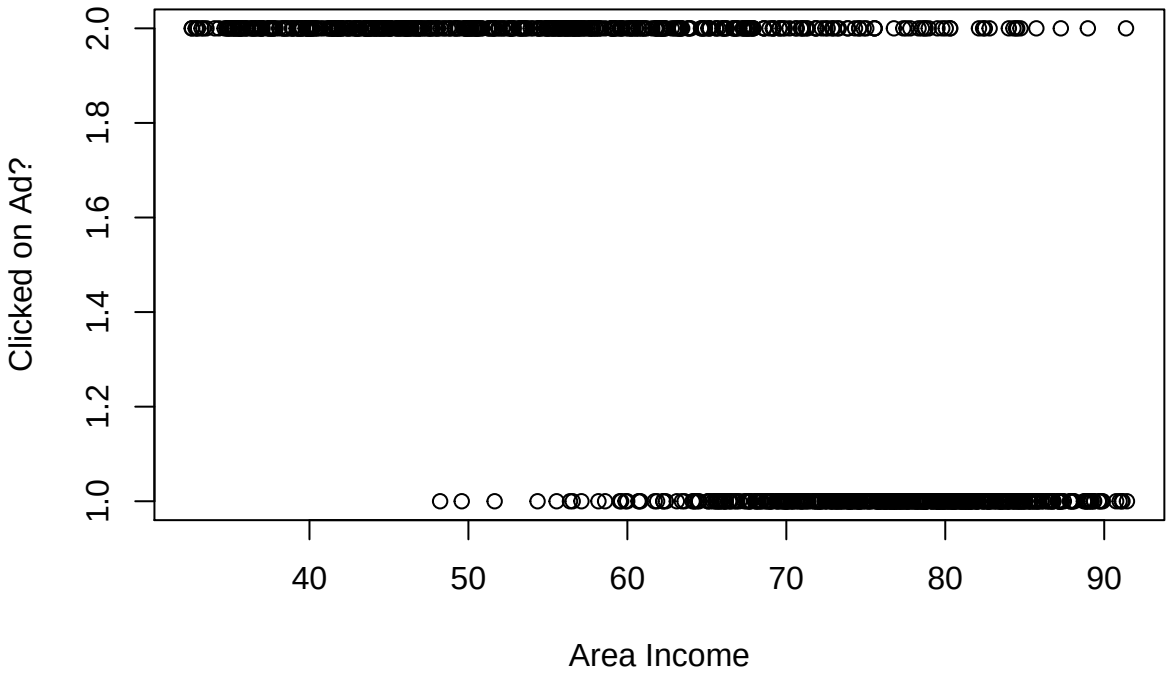

```r
# scatter plot of how age impacts ad being clicked
plot(data$age, ad_int, ylab = "Clicked on Ad?", xlab = "Age")
```




```
# scatter plot of how daily internet usage impacts ad being clicked
plot(data$daily_internet_usage, ad_int, ylab = "Clicked on Ad?", xlab = "Daily Internet Usage")
```




```r
# scatter plot of how area income impacts ad being clicked
plot(data$daily_time_spent, ad_int, ylab = "Clicked on Ad?", xlab = "Area Income")
```




```
colnames(data)
```

```
##  [1] "daily_time_spent"     "age"                  "area_income"
##  [4] "daily_internet_usage" "ad_topic_line"        "city"
##  [7] "male"                 "country"              "year"
## [10] "month"                "day"                  "hour"
## [13] "clicked_on_ad"
```

4. Conclusion

We will use the results we have obtained from our exploratory data analysis to make conclusions.

To begin with, we see that the dataset was already slightly biased by having slightly more females than males. Because of this, more females than males clicked on the ad.

People with lower area incomes clicked more on the ad than people with higher area incomes.

People who spent less time online were more likely to click on the ad than people who spent more time online.

The month of February and the 3rd days were prime times for ad clicking. The 31st days and the month of July, not so much.