Cryptography Course Target Audience.(EDA)

Gathoni Njoroge

2022-05-26

Exploratory Data Analysis of an online Cryptography Course Advert.

1) Defining the question.

a) Specifying the question.

Who is the most likely target audience for an online cryptography course?

b) The metric of success.

Finding a specific group of people who are likely to click on the advertisments given the data available.

c) The context.

Determining which factors make a person one of the target audiences for future forecasting.

d) Experimental design.

i) Loading and reading the data

ii) Data cleaning.

iii) Exploratory analysis.

iv) Conclusion.

v) Recommendation.

e) Appropriateness of the Data Available.

The data available has been collected from previous targeted adverts by the same enterpreneur and is therefore apprpriate for this study.

2) Loading and reading the data.

```
# loading libraries
#
library(data.table)
library(tibble)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.6     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## v purrr   0.3.4

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::between()   masks data.table::between()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x purrr::transpose() masks data.table::transpose()
```

```
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
library(ggplot2)
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
# loading the dataset.
#
target <- fread('http://bit.ly/IPAdvertisingData')

# checking the top of the dataset.
#
head(target)
```

```
##    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
```

```
# checking the bottom of the dataset
#
tail(target)
```

```
##    Daily Time Spent on Site Age Area Income Daily Internet Usage
## 1:                    43.70  28    63126.96               173.01
## 2:                    72.97  30    71384.57               208.58
## 3:                    51.30  45    67782.17               134.42
## 4:                    51.63  51    42415.72               120.37
## 5:                    55.55  19    41920.79               187.95
## 6:                    45.01  26    29875.80               178.35
##                          Ad Topic Line          City Male
## 1:        Front-line bifurcated ability  Nicholasland    0
## 2:        Fundamental modular algorithm     Duffystad    1
## 3:      Grass-roots cohesive monitoring   New Darlene    1
## 4:         Expanded intangible solution South Jessica    1
## 5: Proactive bandwidth-monitored policy   West Steven    0
## 6:      Virtual 5thgeneration emulation   Ronniemouth    0
##                   Country           Timestamp Clicked on Ad
## 1:                Mayotte 2016-04-04 03:57:48             1
## 2:                Lebanon 2016-02-11 21:49:00             1
## 3: Bosnia and Herzegovina 2016-04-22 02:07:01             1
## 4:               Mongolia 2016-02-01 17:24:57             1
## 5:              Guatemala 2016-03-24 02:35:54             0
## 6:                 Brazil 2016-06-03 21:43:21             1
```

```
# Converting dataset into tibble
#
target1 <- tibble(target)
target1
```

```
## # A tibble: 1,000 x 10
##    `Daily Time Spen~`   Age `Area Income` `Daily Interne~` `Ad Topic Line` City
##                 <dbl> <int>         <dbl>            <dbl> <chr>           <chr>
##  1               69.0    35        61834.             256. Cloned 5thgene~ Wrig~
##  2               80.2    31        68442.             194. Monitored nati~ West~
##  3               69.5    26        59786.             236. Organic bottom~ Davi~
##  4               74.2    29        54806.             246. Triple-buffere~ West~
##  5               68.4    35        73890.             226. Robust logisti~ Sout~
```

```
##  6              60.0    23        59762.            227. Sharable clien~ Jami~
##  7              88.9    33        53853.            208. Enhanced dedic~ Bran~
##  8              66      48        24593.            132. Reactive local~ Port~
##  9              74.5    30        68862             222. Configurable c~ West~
## 10              69.9    20        55642.            184. Mandatory homo~ Rami~
## # ... with 990 more rows, and 4 more variables: Male <int>, Country <chr>,
## #   Timestamp <dttm>, `Clicked on Ad` <int>
```

3) Data cleaning.

```
# checking for missing values
#
colSums(is.na(target1))
```

```
## 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
```

```
# The dataset contains no missing values
```

```
# checking for duplicates
#
duplicates <- target1[duplicated(target1),]
duplicates
```

```
## # A tibble: 0 x 10
## # ... with 10 variables: Daily Time Spent on Site <dbl>, Age <int>,
## #   Area Income <dbl>, Daily Internet Usage <dbl>, Ad Topic Line <chr>,
## #   City <chr>, Male <int>, Country <chr>, Timestamp <dttm>,
## #   Clicked on Ad <int>
```

```
# handling duplicates
#
# The dataset contains no duplicated rows
```

```
# checking for and removing outliers using IQR
#

Q1 = quantile(target1$'Age')
Q3 = quantile(target1$'Age')
IQR = IQR(target1$'Age')
no_out = subset(target1, target1$'Age'> (Q1 - 1.5*IQR) & target1$'Age'< (Q3 + 1.5*IQR))
no_out
```

```
## # A tibble: 748 x 10
##    `Daily Time Spen~`   Age `Area Income` `Daily Interne~` `Ad Topic Line` City
##                 <dbl> <int>         <dbl>            <dbl> <chr>           <chr>
##  1               69.0    35        61834.             256. Cloned 5thgene~ Wrig~
##  2               80.2    31        68442.             194. Monitored nati~ West~
##  3               69.5    26        59786.             236. Organic bottom~ Davi~
##  4               74.2    29        54806.             246. Triple-buffere~ West~
##  5               60.0    23        59762.             227. Sharable clien~ Jami~
##  6               88.9    33        53853.             208. Enhanced dedic~ Bran~
##  7               66      48        24593.             132. Reactive local~ Port~
##  8               74.5    30        68862              222. Configurable c~ West~
##  9               83.1    37        62491.             231. Team-oriented ~ East~
## 10               69.6    48        51637.             113. Centralized co~ West~
## # ... with 738 more rows, and 4 more variables: Male <int>, Country <chr>,
## #   Timestamp <dttm>, `Clicked on Ad` <int>
```

```
# selecting rows where clicked on ad is 1
#
target2 <- no_out %>% filter(no_out$'Clicked on Ad' == 1)
target2
```

```
## # A tibble: 361 x 10
##    `Daily Time Spen~`   Age `Area Income` `Daily Interne~` `Ad Topic Line` City
##                 <dbl> <int>         <dbl>            <dbl> <chr>           <chr>
##  1               66      48        24593.             132. Reactive local~ Port~
##  2               69.6    48        51637.             113. Centralized co~ West~
##  3               63.4    23        52182.             141. Persistent dem~ New ~
##  4               55.4    37        23937.             129. Customizable m~ West~
##  5               54.7    36        31088.             118. Grass-roots so~ Jess~
##  6               41.5    52        32636.             165. Mandatory disi~ Sout~
##  7               48.5    28        38067.             134. Ameliorated cl~ West~
##  8               52.0    52        58296.             129. Monitored syst~ Sout~
##  9               70.2    34        32709.             119. Open-architect~ Palm~
## 10               55.6    23        30228.             213. Multi-layered ~ Port~
## # ... with 351 more rows, and 4 more variables: Male <int>, Country <chr>,
## #   Timestamp <dttm>, `Clicked on Ad` <int>
```

```
# dropping redundant variables
#
drop <- c("Timestamp", "Ad Topic Line")
target_ad = target2[,!(names(target2) %in% drop)]
target_ad
```

```
## # A tibble: 361 x 8
##    `Daily Time Spent ~`   Age `Area Income` `Daily Interne~` City   Male Country
##                   <dbl> <int>         <dbl>            <dbl> <chr> <int> <chr>
##  1                 66      48        24593.             132. Port~     1 Austra~
##  2                 69.6    48        51637.             113. West~     1 Egypt
##  3                 63.4    23        52182.             141. New ~     1 Spain
##  4                 55.4    37        23937.             129. West~     0 Palest~
##  5                 54.7    36        31088.             118. Jess~     1 Britis~
##  6                 41.5    52        32636.             165. Sout~     0 Burundi
##  7                 48.5    28        38067.             134. West~     1 Tuvalu
```

```
##  8                52.0    52        58296.             129. Sout~     0 Greece
##  9                70.2    34        32709.             119. Palm~     0 Britis~
## 10                55.6    23        30228.             213. Port~     0 Senegal
## # ... with 351 more rows, and 1 more variable: `Clicked on Ad` <int>
```

4) Exploratory Analysis.

Univariate Analysis.

```
#calculating mean
#
mean(target_ad$'Daily Time Spent on Site')
```

Measures of Central Tendency.

```
## [1] 52.9731
```

```
mean(target_ad$'Age')
```

```
## [1] 39.30194
```

```
mean(target_ad$'Area Income')
```

```
## [1] 49106.58
```

```
mean(target_ad$'Daily Internet Usage')
```

```
## [1] 144.7557
```

```
#Calculating median
#
median(target_ad$'Daily Time Spent on Site')
```

```
## [1] 51.58
```

```
median(target_ad$'Age')
```

```
## [1] 39
```

```
median(target_ad$'Area Income')
```

```
## [1] 50055.33
```

```r
median(target_ad$'Daily Internet Usage')
```

```
## [1] 137.43
```

```r
# Calculating mode
#
getmode <- function(v) {
   uniqv <- unique(v)
   uniqv[which.max(tabulate(match(v, uniqv)))]
}
getmode(target_ad$'Daily Time Spent on Site')
```

```
## [1] 55.6
```

```r
getmode(target_ad$'Age')
```

```
## [1] 45
```

```r
getmode(target_ad$'Area Income')
```

```
## [1] 24593.33
```

```r
getmode(target_ad$'Daily Internet Usage')
```

```
## [1] 120.06
```

```r
getmode(target_ad$'Male')
```

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

```r
# calculating minimum values
#
min(target_ad$'Daily Time Spent on Site')
```

Measures of Dispersion.

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

```r
min(target_ad$'Age')
```

```
## [1] 20
```

```r
min(target_ad$'Area Income')
```

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

```r
min(target_ad$'Daily Internet Usage')
```

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

```r
'---------------'
```

```
## [1] "---------------"
```

```r
#calculating maximum values
#
max(target_ad$'Daily Time Spent on Site')
```

```
## [1] 88.97
```

```r
max(target_ad$'Age')
```

```
## [1] 60
```

```r
max(target_ad$'Area Income')
```

```
## [1] 77143.61
```

```r
max(target_ad$'Daily Internet Usage')
```

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

```r
# Get range, 1st and 3rd quantiles
#
quantile(target_ad$'Daily Time Spent on Site')
```

```
##    0%   25%   50%   75%  100%
## 32.60 43.07 51.58 61.22 88.97
```

```r
quantile(target_ad$'Age')
```

```
##   0%  25%  50%  75% 100%
##   20   33   39   45   60
```

```r
quantile(target_ad$'Area Income')
```

```
##       0%      25%      50%      75%     100%
## 13996.50 39799.73 50055.33 59243.46 77143.61
```

```r
quantile(target_ad$'Daily Internet Usage')
```

```
##     0%    25%    50%    75%   100%
## 104.78 122.45 137.43 158.56 269.96
```

```r
# variance and standard deviation
#
var(target_ad$'Daily Time Spent on Site')
```

```
## [1] 158.9318
```

```r
var(target_ad$'Age')
```

```
## [1] 71.61136
```

```r
var(target_ad$'Area Income')
```

```
## [1] 190658357
```

```r
var(target_ad$'Daily Internet Usage')
```

```
## [1] 971.8155
```

```r
'------------'
```

```
## [1] "------------"
```

```r
sd(target_ad$'Daily Time Spent on Site')
```

```
## [1] 12.60682
```

```r
sd(target_ad$'Age')
```

```
## [1] 8.462349
```

```r
sd(target_ad$'Area Income')
```

```
## [1] 13807.91
```

```r
sd(target_ad$'Daily Internet Usage')
```

```
## [1] 31.17396
```

```r
# frequency of categorical variables
#
library(plyr)
```

```
## ------------------------------------------------------------------------------
```

```
## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)
```

```
## ------------------------------------------------------------------------

##
## Attaching package: 'plyr'

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

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

```
library(dplyr)
country_freq = count(target_ad,  'Country')
country_freq
```

```
##                                                Country freq
## 1                                          Afghanistan    4
## 2                                              Albania    3
## 3                                              Algeria    2
## 4                                       American Samoa    3
## 5                                              Andorra    2
## 6                                               Angola    1
## 7                                             Anguilla    2
## 8          Antarctica (the territory South of 60 deg S)    1
## 9                                   Antigua and Barbuda    4
## 10                                            Argentina    1
## 11                                              Armenia    1
## 12                                            Australia    5
## 13                                              Austria    1
## 14                                              Bahamas    3
## 15                                              Bahrain    2
## 16                                           Bangladesh    2
## 17                                              Belarus    2
## 18                                              Belgium    1
## 19                                               Belize    2
## 20                                                Benin    1
## 21                                               Bhutan    1
## 22                           Bosnia and Herzegovina    3
## 23                       Bouvet Island (Bouvetoya)    2
## 24                                               Brazil    2
## 25  British Indian Ocean Territory (Chagos Archipelago)    1
## 26                           British Virgin Islands    1
## 27                                Brunei Darussalam    2
## 28                                     Burkina Faso    1
## 29                                          Burundi    2
## 30                                         Cambodia    1
## 31                                           Canada    1
## 32                                   Cayman Islands    2
## 33                       Central African Republic    1
## 34                                             Chad    2
```

```
## 35                                    Chile    2
## 36                                    China    3
## 37                          Christmas Island    3
## 38                                 Colombia    1
## 39                                  Comoros    1
## 40                                    Congo    2
## 41                             Cook Islands    1
## 42                           Cote d'Ivoire    2
## 43                                     Cuba    3
## 44                                   Cyprus    2
## 45                           Czech Republic    2
## 46                                  Denmark    2
## 47                                 Djibouti    1
## 48                                 Dominica    2
## 49                       Dominican Republic    1
## 50                                  Ecuador    1
## 51                                    Egypt    3
## 52                              El Salvador    3
## 53                        Equatorial Guinea    3
## 54                                  Eritrea    1
## 55                                  Estonia    1
## 56                                 Ethiopia    3
## 57                             Faroe Islands    2
## 58                                     Fiji    3
## 59                                  Finland    1
## 60                                   France    3
## 61                             French Guiana    1
## 62                          French Polynesia    1
## 63                French Southern Territories    1
## 64                                   Gambia    1
## 65                                  Georgia    2
## 66                                  Germany    1
## 67                                   Greece    3
## 68                                Greenland    1
## 69                               Guadeloupe    1
## 70                                     Guam    1
## 71                                Guatemala    2
## 72                                 Guernsey    2
## 73                                   Guinea    2
## 74                            Guinea-Bissau    1
## 75                                   Guyana    3
## 76          Heard Island and McDonald Islands    2
## 77              Holy See (Vatican City State)    1
## 78                                 Honduras    2
## 79                                Hong Kong    4
## 80                                  Hungary    4
## 81                                  Iceland    1
## 82                                Indonesia    4
## 83                                     Iran    1
## 84                                  Ireland    1
## 85                                   Israel    2
## 86                                    Italy    1
## 87                                  Jamaica    1
## 88                                    Japan    2
```

```
## 89                                        Jersey    3
## 90                                    Kazakhstan    1
## 91                                         Kenya    4
## 92                                         Korea    3
## 93                                        Kuwait    1
## 94                  Lao People's Democratic Republic    2
## 95                                        Latvia    4
## 96                                       Lebanon    3
## 97                                       Liberia    3
## 98                         Libyan Arab Jamahiriya    1
## 99                                  Liechtenstein    4
## 100                                     Lithuania    2
## 101                                    Luxembourg    3
## 102                                         Macao    1
## 103                                     Macedonia    1
## 104                                        Malawi    1
## 105                                          Mali    1
## 106                                         Malta    1
## 107                               Marshall Islands    1
## 108                                     Martinique    2
## 109                                     Mauritania    1
## 110                                        Mayotte    3
## 111                                         Mexico    3
## 112                                     Micronesia    4
## 113                                        Moldova    2
## 114                                         Monaco    1
## 115                                       Mongolia    3
## 116                                     Montenegro    1
## 117                                     Montserrat    1
## 118                                        Morocco    1
## 119                                        Namibia    1
## 120                                    Netherlands    2
## 121                            Netherlands Antilles    1
## 122                                  New Caledonia    2
## 123                                    New Zealand    2
## 124                                          Niger    2
## 125                                 Norfolk Island    1
## 126                        Northern Mariana Islands    2
## 127                                       Pakistan    1
## 128                                          Palau    2
## 129                           Palestinian Territory    2
## 130                               Papua New Guinea    3
## 131                                       Paraguay    1
## 132                                           Peru    3
## 133                                    Philippines    3
## 134                                 Pitcairn Islands    1
## 135                                         Poland    3
## 136                                    Puerto Rico    1
## 137                                        Romania    1
## 138                                         Rwanda    2
## 139                                Saint Barthelemy    2
## 140                                    Saint Helena    1
## 141                            Saint Kitts and Nevis    1
## 142                                     Saint Lucia    1
```

```
## 143                                     Saint Martin    2
## 144                           Saint Pierre and Miquelon    3
## 145                       Saint Vincent and the Grenadines    3
## 146                                               Samoa    3
## 147                                        Saudi Arabia    1
## 148                                             Senegal    3
## 149                                              Serbia    2
## 150                                        Sierra Leone    2
## 151                                            Slovenia    1
## 152                                             Somalia    1
## 153                                        South Africa    6
## 154       South Georgia and the South Sandwich Islands    1
## 155                                               Spain    2
## 156                                            Suriname    1
## 157                          Svalbard & Jan Mayen Islands    3
## 158                                              Sweden    1
## 159                                         Switzerland    3
## 160                                Syrian Arab Republic    1
## 161                                              Taiwan    4
## 162                                          Tajikistan    2
## 163                                            Tanzania    1
## 164                                            Thailand    1
## 165                                          Timor-Leste    1
## 166                                              Tokelau    2
## 167                                                Tonga    1
## 168                                  Trinidad and Tobago    1
## 169                                               Turkey    6
## 170                                         Turkmenistan    1
## 171                             Turks and Caicos Islands    3
## 172                                               Tuvalu    3
## 173                                               Uganda    4
## 174                                              Ukraine    1
## 175                                United Arab Emirates    1
## 176                                      United Kingdom    1
## 177                  United States Minor Outlying Islands    1
## 178                               United States of America    3
## 179                          United States Virgin Islands    2
## 180                                           Uzbekistan    1
## 181                                              Vanuatu    1
## 182                                            Venezuela    2
## 183                                              Vietnam    2
## 184                                       Western Sahara    3
## 185                                               Zambia    3
## 186                                             Zimbabwe    2
```

```
sorted_by_freq <- country_freq[order(-country_freq$freq),]
sorted_by_freq
```

```
##                                         Country freq
## 153                                    South Africa    6
## 169                                          Turkey    6
## 12                                        Australia    5
## 1                                       Afghanistan    4
## 9                                 Antigua and Barbuda    4
```

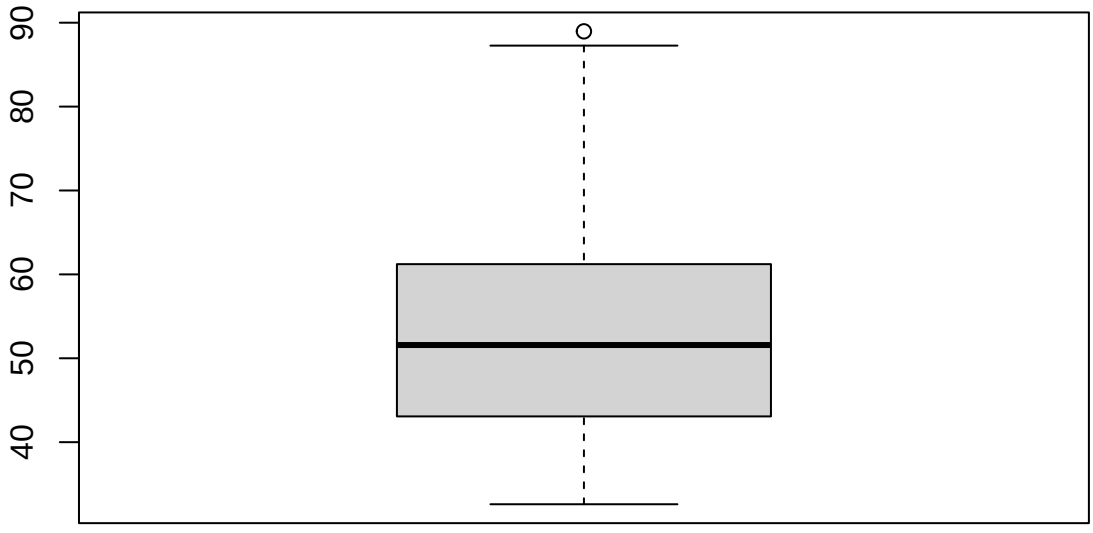
```
## 79                                  Hong Kong    4
## 80                                    Hungary    4
## 82                                  Indonesia    4
## 91                                      Kenya    4
## 95                                     Latvia    4
## 99                                Liechtenstein    4
## 112                                 Micronesia    4
## 161                                     Taiwan    4
## 173                                     Uganda    4
## 2                                       Albania    3
## 4                                American Samoa    3
## 14                                     Bahamas    3
## 22                   Bosnia and Herzegovina    3
## 36                                      China    3
## 37                            Christmas Island    3
## 43                                       Cuba    3
## 51                                      Egypt    3
## 52                                  El Salvador    3
## 53                            Equatorial Guinea    3
## 56                                   Ethiopia    3
## 58                                       Fiji    3
## 60                                     France    3
## 67                                     Greece    3
## 75                                     Guyana    3
## 89                                     Jersey    3
## 92                                      Korea    3
## 96                                    Lebanon    3
## 97                                    Liberia    3
## 101                                Luxembourg    3
## 110                                   Mayotte    3
## 111                                    Mexico    3
## 115                                  Mongolia    3
## 130                          Papua New Guinea    3
## 132                                      Peru    3
## 133                               Philippines    3
## 135                                    Poland    3
## 144                 Saint Pierre and Miquelon    3
## 145 Saint Vincent and the Grenadines    3
## 146                                     Samoa    3
## 148                                   Senegal    3
## 157      Svalbard & Jan Mayen Islands    3
## 159                               Switzerland    3
## 171          Turks and Caicos Islands    3
## 172                                    Tuvalu    3
## 178          United States of America    3
## 184                    Western Sahara    3
## 185                            Zambia    3
## 3                                      Algeria    2
## 5                                      Andorra    2
## 7                                     Anguilla    2
## 15                                     Bahrain    2
## 16                                  Bangladesh    2
## 17                                     Belarus    2
## 19                                      Belize    2
```

```
## 23                            Bouvet Island (Bouvetoya)    2
## 24                                               Brazil    2
## 27                                     Brunei Darussalam    2
## 29                                              Burundi    2
## 32                                        Cayman Islands    2
## 34                                                 Chad    2
## 35                                                Chile    2
## 40                                                Congo    2
## 42                                         Cote d'Ivoire    2
## 44                                               Cyprus    2
## 45                                       Czech Republic    2
## 46                                              Denmark    2
## 48                                             Dominica    2
## 57                                         Faroe Islands    2
## 65                                              Georgia    2
## 71                                            Guatemala    2
## 72                                             Guernsey    2
## 73                                               Guinea    2
## 76                     Heard Island and McDonald Islands    2
## 78                                             Honduras    2
## 85                                               Israel    2
## 88                                                Japan    2
## 94                      Lao People's Democratic Republic    2
## 100                                            Lithuania    2
## 108                                           Martinique    2
## 113                                              Moldova    2
## 120                                          Netherlands    2
## 122                                        New Caledonia    2
## 123                                          New Zealand    2
## 124                                                Niger    2
## 126                           Northern Mariana Islands    2
## 128                                                Palau    2
## 129                              Palestinian Territory    2
## 138                                               Rwanda    2
## 139                                    Saint Barthelemy    2
## 143                                        Saint Martin    2
## 149                                              Serbia    2
## 150                                        Sierra Leone    2
## 155                                               Spain    2
## 162                                          Tajikistan    2
## 166                                             Tokelau    2
## 179                          United States Virgin Islands    2
## 182                                           Venezuela    2
## 183                                             Vietnam    2
## 186                                            Zimbabwe    2
## 6                                                Angola    1
## 8          Antarctica (the territory South of 60 deg S)    1
## 10                                             Argentina    1
## 11                                               Armenia    1
## 13                                               Austria    1
## 18                                               Belgium    1
## 20                                                 Benin    1
## 21                                                Bhutan    1
## 25  British Indian Ocean Territory (Chagos Archipelago)    1
```

```
## 26                          British Virgin Islands    1
## 28                                    Burkina Faso    1
## 30                                        Cambodia    1
## 31                                          Canada    1
## 33                       Central African Republic    1
## 38                                        Colombia    1
## 39                                         Comoros    1
## 41                                    Cook Islands    1
## 47                                        Djibouti    1
## 49                              Dominican Republic    1
## 50                                         Ecuador    1
## 54                                         Eritrea    1
## 55                                         Estonia    1
## 59                                         Finland    1
## 61                                  French Guiana     1
## 62                                French Polynesia    1
## 63                  French Southern Territories     1
## 64                                          Gambia    1
## 66                                         Germany    1
## 68                                       Greenland    1
## 69                                      Guadeloupe    1
## 70                                            Guam    1
## 74                                   Guinea-Bissau    1
## 77                  Holy See (Vatican City State)     1
## 81                                         Iceland    1
## 83                                            Iran    1
## 84                                         Ireland    1
## 86                                           Italy    1
## 87                                         Jamaica    1
## 90                                      Kazakhstan    1
## 93                                          Kuwait    1
## 98                            Libyan Arab Jamahiriya    1
## 102                                           Macao    1
## 103                                        Macedonia    1
## 104                                           Malawi    1
## 105                                             Mali    1
## 106                                            Malta    1
## 107                                 Marshall Islands    1
## 109                                       Mauritania    1
## 114                                           Monaco    1
## 116                                       Montenegro    1
## 117                                       Montserrat    1
## 118                                          Morocco    1
## 119                                          Namibia    1
## 121                             Netherlands Antilles    1
## 125                                   Norfolk Island    1
## 127                                         Pakistan    1
## 131                                         Paraguay    1
## 134                                 Pitcairn Islands    1
## 136                                      Puerto Rico    1
## 137                                          Romania    1
## 140                                      Saint Helena    1
## 141                              Saint Kitts and Nevis    1
## 142                                       Saint Lucia    1
```

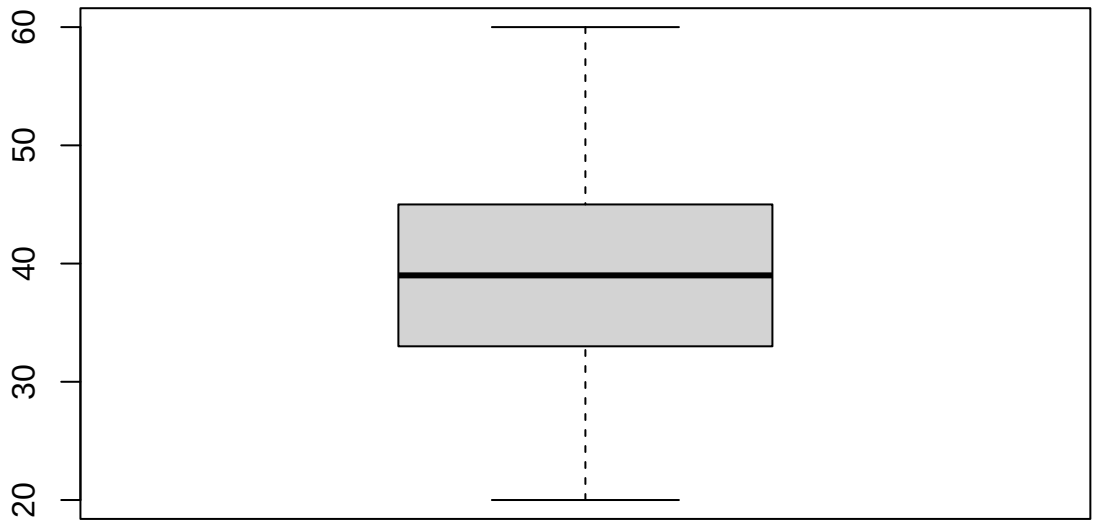
```
## 147                                      Saudi Arabia    1
## 151                                          Slovenia    1
## 152                                           Somalia    1
## 154       South Georgia and the South Sandwich Islands    1
## 156                                          Suriname    1
## 158                                            Sweden    1
## 160                               Syrian Arab Republic    1
## 163                                          Tanzania    1
## 164                                          Thailand    1
## 165                                         Timor-Leste    1
## 167                                             Tonga    1
## 168                               Trinidad and Tobago    1
## 170                                      Turkmenistan    1
## 174                                           Ukraine    1
## 175                               United Arab Emirates    1
## 176                                    United Kingdom    1
## 177            United States Minor Outlying Islands    1
## 180                                        Uzbekistan    1
## 181                                           Vanuatu    1
```

```
#summary of above using boxplots
#
boxplot(target_ad$'Daily Time Spent on Site')
```

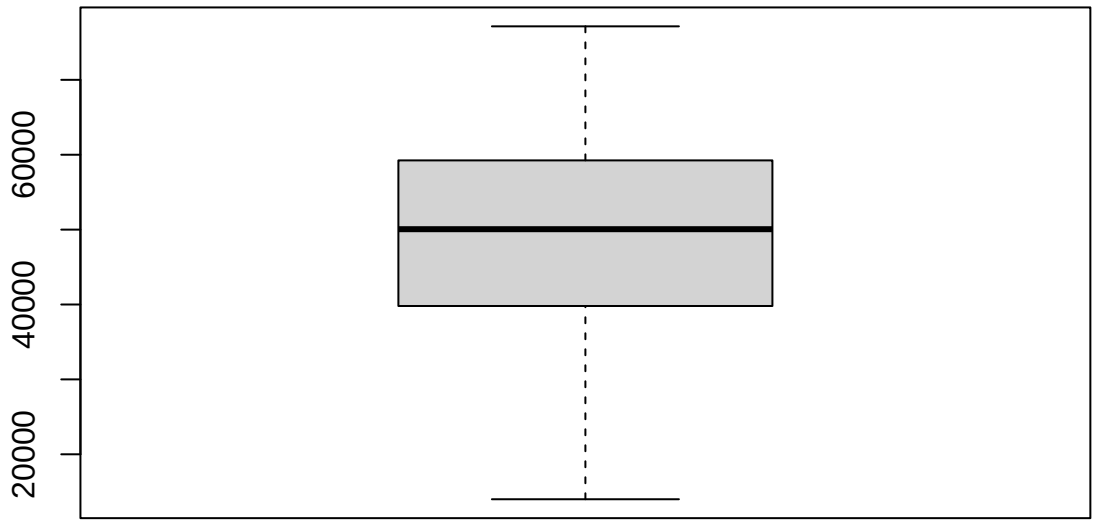


Visualization.

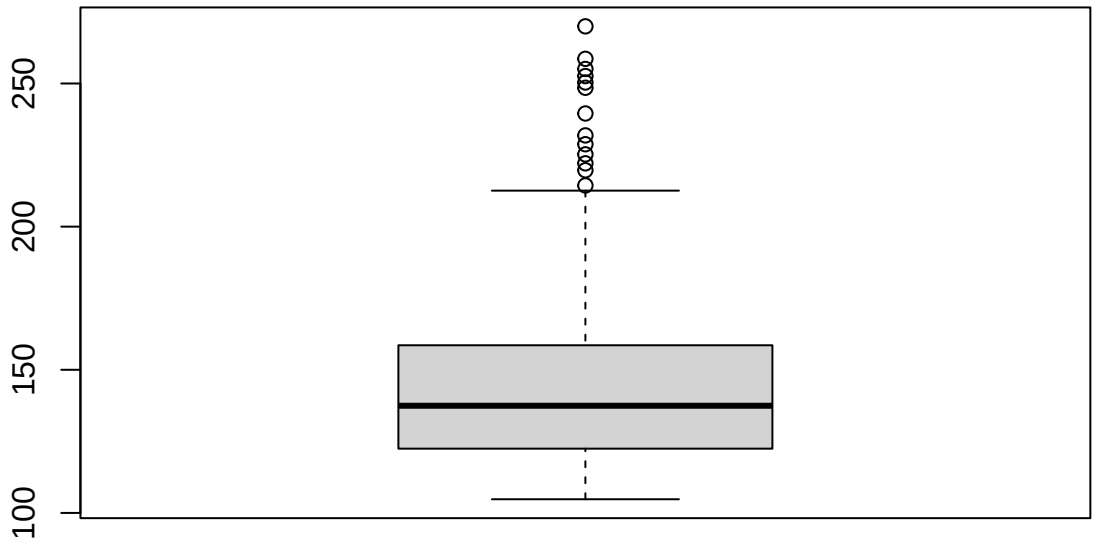
```
boxplot(target_ad$'Age')
```



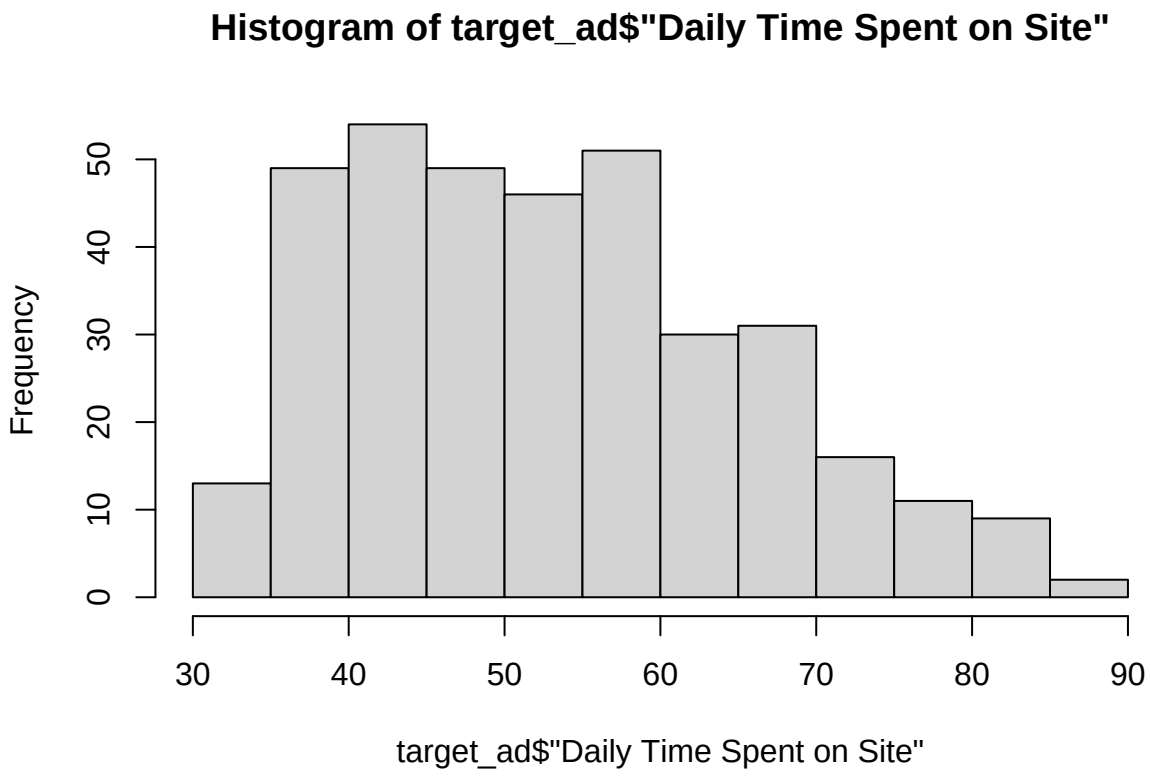

```
boxplot(target_ad$'Area Income')
```



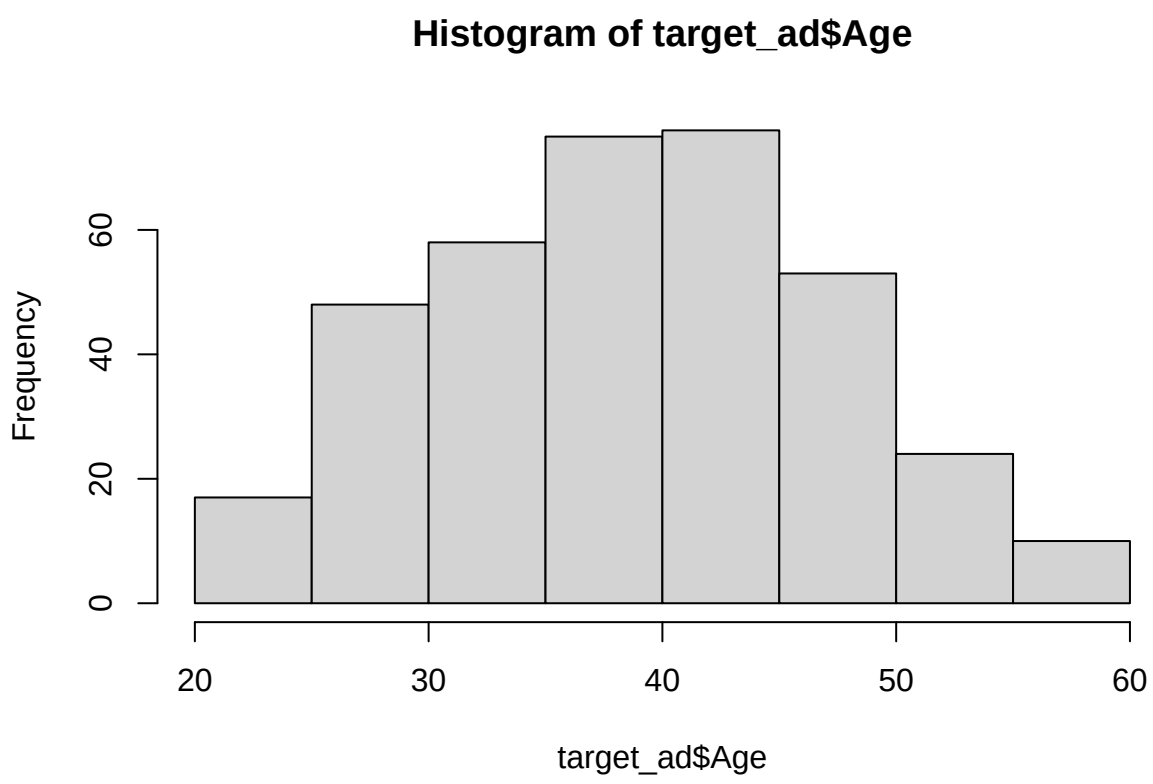

```
boxplot(target_ad$'Daily Internet Usage')
```



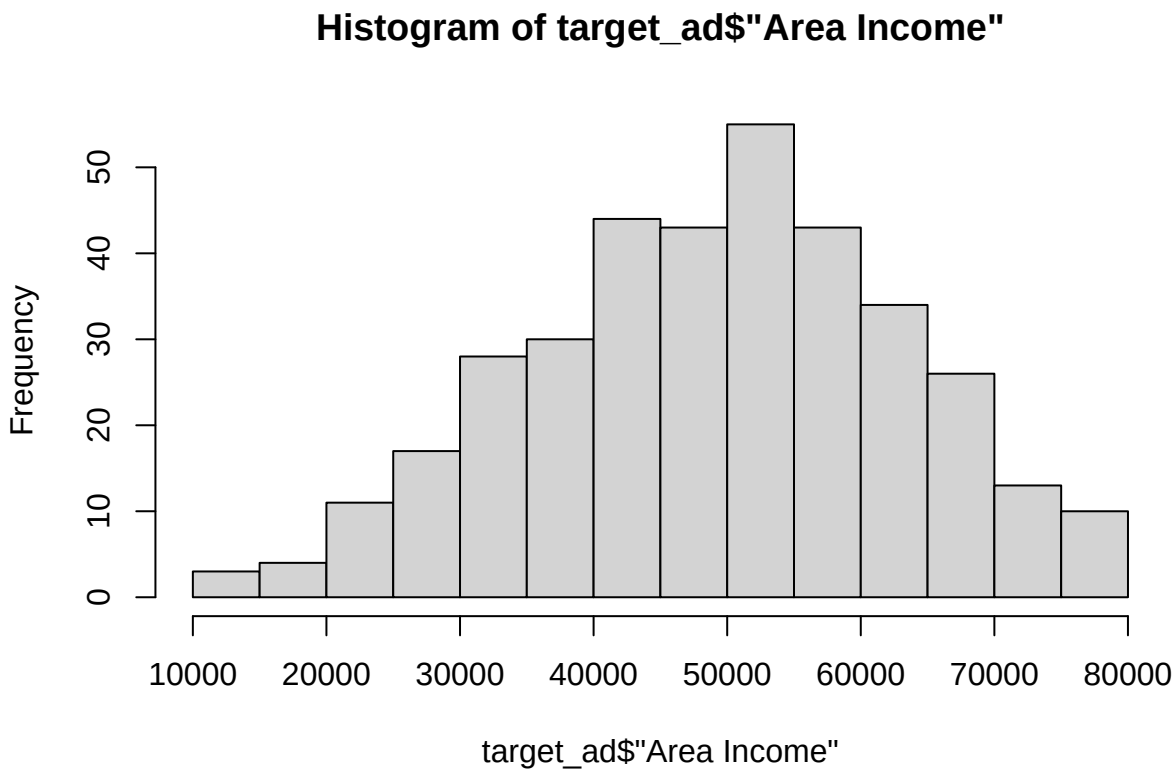

```
# visualisation of numeric variables using a histogam
#
hist(target_ad$'Daily Time Spent on Site')
```



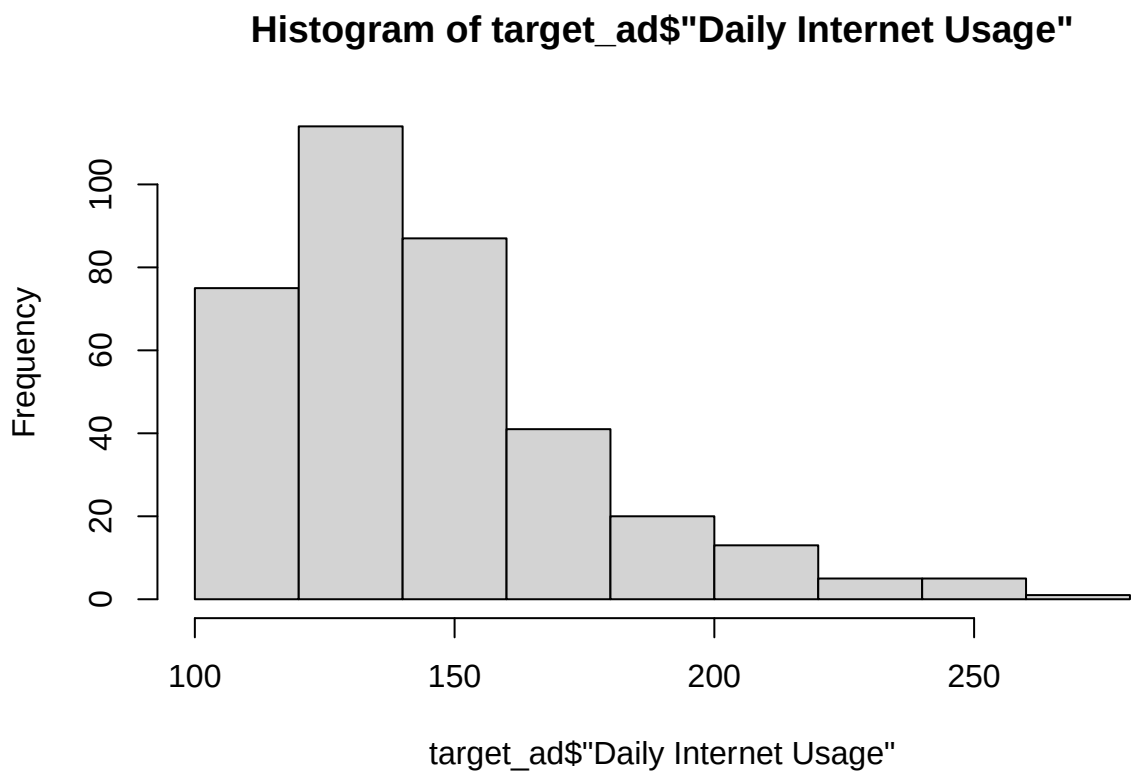

```
hist(target_ad$'Age')
```



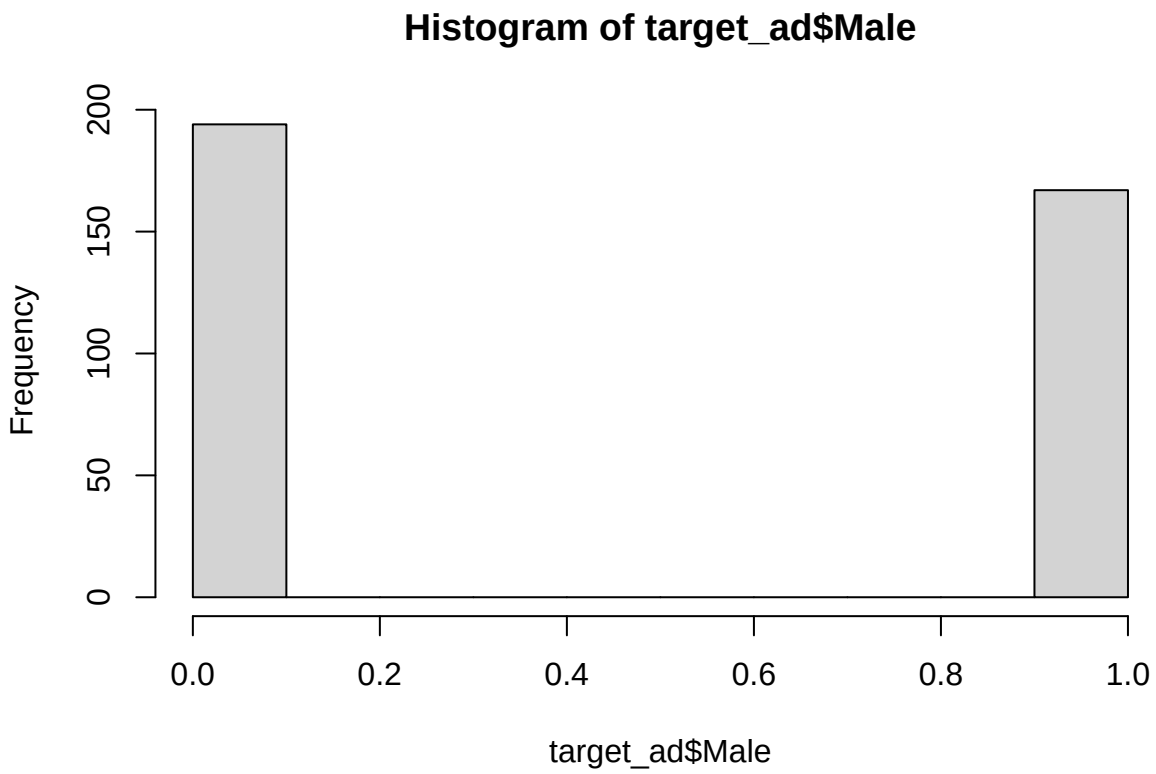

```
hist(target_ad$'Area Income')
```




```
hist(target_ad$'Daily Internet Usage')
```

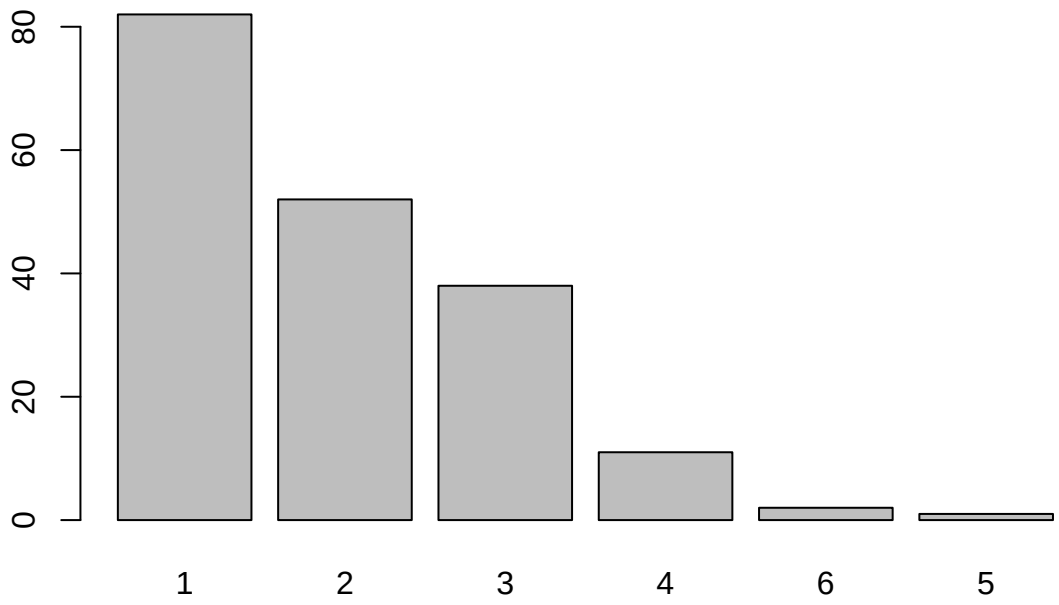



```
hist(target_ad$'Male')
```




```
# Barplot of countries.
#

x <- table(country_freq$freq)
barplot(x[order(x, decreasing=TRUE)])
```



Bivariate Analysis #### Covariance and Correlation.

```
# Covariance of numeric variables
#
drop <- c("Country", "City", "Location", "Ad Topic Line", "Clicked on Ad")
target4 = target_ad[,!(names(target_ad) %in% drop)]
target4
```

```
## # A tibble: 361 x 5
##    `Daily Time Spent on Site`   Age `Area Income` `Daily Internet Usage`  Male
##                         <dbl> <int>         <dbl>                  <dbl> <int>
##  1                       66      48        24593.                   132.     1
##  2                       69.6    48        51637.                   113.     1
##  3                       63.4    23        52182.                   141.     1
##  4                       55.4    37        23937.                   129.     0
##  5                       54.7    36        31088.                   118.     1
##  6                       41.5    52        32636.                   165.     0
##  7                       48.5    28        38067.                   134.     1
##  8                       52.0    52        58296.                   129.     0
##  9                       70.2    34        32709.                   119.     0
## 10                       55.6    23        30228.                   213.     0
## # ... with 351 more rows
```

```
cov(target4)
```

```
##                         Daily Time Spent on Site          Age   Area Income
```

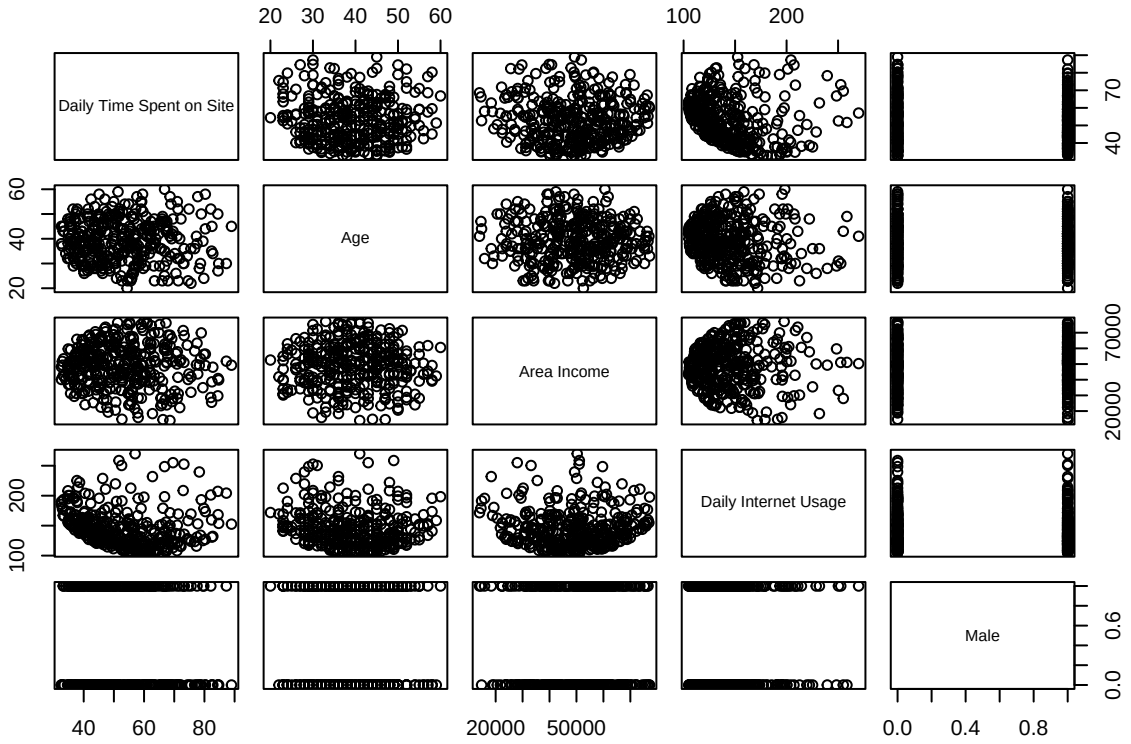
```
## Daily Time Spent on Site              158.9317865 -1.329912e+00 -7.350829e+03
## Age                                    -1.3299116  7.161136e+01 -2.411858e+03
## Area Income                         -7350.8286871 -2.411858e+03  1.906584e+08
## Daily Internet Usage                   -68.1159257 -3.085422e+01 -7.959702e+03
## Male                                    -0.7673003 -2.339951e-02  3.556316e+02
##                           Daily Internet Usage          Male
## Daily Time Spent on Site        -6.811593e+01  -0.767300323
## Age                             -3.085422e+01  -0.023399508
## Area Income                     -7.959702e+03 355.631570714
## Daily Internet Usage             9.718155e+02  -0.003606494
## Male                            -3.606494e-03   0.249292090
```

```
#correlation matrix
#
correlation = cor(target4)
correlation
```

```
##                           Daily Time Spent on Site          Age Area Income
## Daily Time Spent on Site               1.00000000 -0.012465980 -0.04222824
## Age                                   -0.01246598  1.000000000 -0.02064110
## Area Income                           -0.04222824 -0.020641098  1.00000000
## Daily Internet Usage                  -0.17332107 -0.116958473 -0.01849171
## Male                                  -0.12190057 -0.005538109  0.05158437
##                           Daily Internet Usage         Male
## Daily Time Spent on Site         -0.173321069 -0.121900574
## Age                              -0.116958473 -0.005538109
## Area Income                      -0.018491706  0.051584371
## Daily Internet Usage              1.000000000 -0.000231707
## Male                             -0.000231707  1.000000000
```

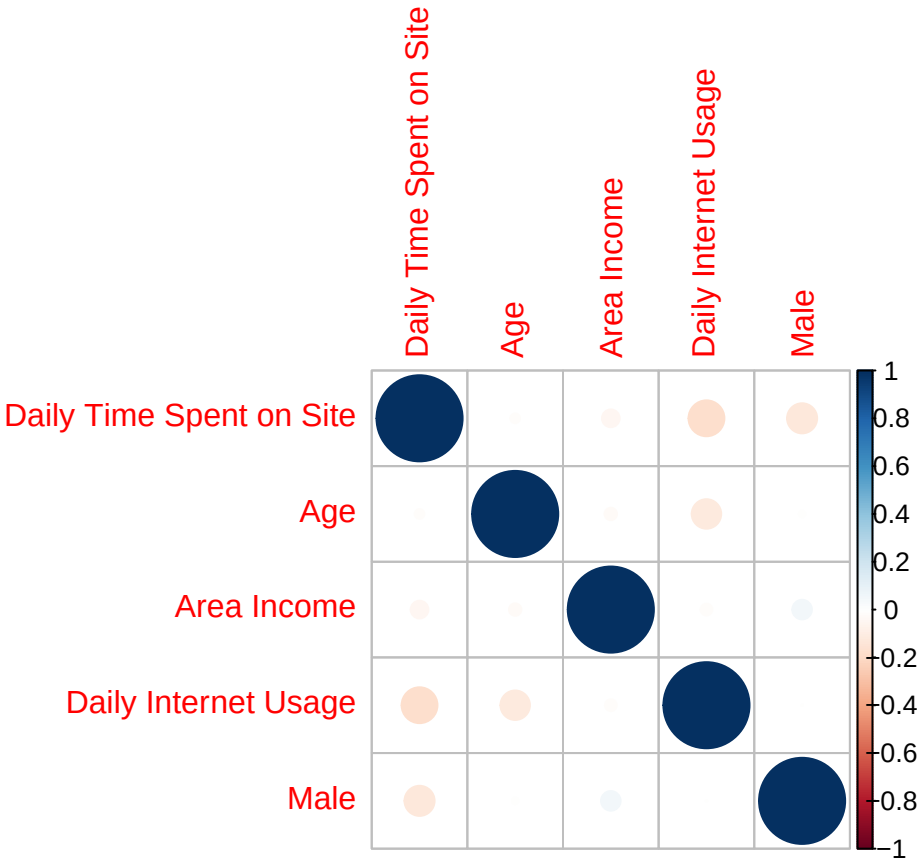
```
# With a scatterplot matrix
#

plot(target4)
```



Visualization.

```
# visualization of the correlation matrix
#
corrplot(correlation)
```



5) Conclusion

Give the results of this analysis the most likely group of people to click on the targeted ads are:

a)Those aged between 35-45

b) Who spend between 50 minutes to an hour per day on the site

c) From South Africa, Turkey, Australia. though not limited to this countries only.

6) Recommendation.

More data is needed for further analysis. However, what is currently analyzed could still be useful to launch a pilot advertising program and gauge how well it works and reaches the target audience.