

# Moringa\_Data\_Science\_Prep\_W12\_Independent\_Project\_2020\_02\_Gideon\_Cheruiyot\_DataReport

Gideon Cheruiyot

28/02/2020

#Problem statement: I identify individuals who are more likely to click on the Ads posted on his/her blog

#Metrics of success: our success will be measured by our ability to identify attributes of the the users who click the adverts.#

#Data relevance

*The appropriateness of the data will be addressed by checking:*

- whether the data is accurate?
- whether the dataset is enough to sufficiently address the problem at hand?
- whether the dataset was biased or imbalanced? # Experimental design:
- loading and previewing the dataset
- Cleaning , checking for outliers
- Conducting exploratory data analysis to find out patterns and relationships
- Communicating the observations and findings
- Drawing conclusion

#Challenging the solution. More analysis need to be done to clearly predict the person likely to click on the Ad, this may include modeling for instance incorporating a logistic regression model.

##Loading the dataset

```
data <- read.csv('advertising.csv')
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
```

## Understanding the datatypes for our features

```
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",...: 92 465 567 904 767 806 223 724 108 455 ...
## $ City                    : Factor w/ 969 levels
"Adamsbury","Adamside",...: 962 904 112 940 806 283 47 672 885 713 ...
## $ 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 79 ...
## $ Timestamp                : Factor w/ 1000 levels "2016-01-01
02:52:10",...: 440 475 368 57 768 690 131 334 549 942 ...
## $ Clicked.on.Ad           : int    0 0 0 0 0 0 0 1 0 0 ...
```

## Finding any 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
```

## Exploratory Data Analysis

### Univariate analysis:

\*we will be analysis individual featurtes to find interesting facts about them. The features include;Daily.Time.Spent.on.Site, Age, Area.Income, Male and daily time spent on the ineternet.

```
summary(data)
```

```
## Daily.Time.Spent.on.Site      Age      Area.Income
Daily.Internet.Usage
## Min.      :32.60      Min.      :19.00      Min.      :13996      Min.      :104.8
## 1st Qu.:51.36      1st Qu.:29.00      1st Qu.:47032      1st Qu.:138.8
## Median :68.22      Median :35.00      Median :57012      Median :183.1
## Mean   :65.00      Mean   :36.01      Mean   :55000      Mean   :180.0
## 3rd Qu.:78.55      3rd Qu.:42.00      3rd Qu.:65471      3rd Qu.:218.8
## Max.    :91.43      Max.    :61.00      Max.    :79485      Max.    :270.0
##
##                               Ad.Topic.Line      City
## Adaptive 24hour Graphic Interface      : 1      Lisamouth      : 3
## Adaptive asynchronous attitude          : 1      Williamsport   : 3
## Adaptive context-sensitive application   : 1      Benjaminechester: 2
## Adaptive contextually-based methodology: 1      East John      : 2
## Adaptive demand-driven knowledgebase    : 1      East Timothy   : 2
## Adaptive uniform capability              : 1      Johnstad       : 2
## (Other)                                :994      (Other)        :986
##      Male      Country      Timestamp
Clicked.on.Ad
## Min.      :0.000      Czech Republic: 9      2016-01-01 02:52:10: 1      Min.
:0.0
## 1st Qu.:0.000      France      : 9      2016-01-01 03:35:35: 1      1st
Qu.:0.0
## Median :0.000      Afghanistan : 8      2016-01-01 05:31:22: 1      Median
:0.5
## Mean   :0.481      Australia  : 8      2016-01-01 08:27:06: 1      Mean
:0.5
## 3rd Qu.:1.000      Cyprus     : 8      2016-01-01 15:14:24: 1      3rd
Qu.:1.0
## Max.    :1.000      Greece     : 8      2016-01-01 20:17:49: 1      Max.
:1.0
##      (Other)      :950      (Other)      :994
```

### Finding outliers

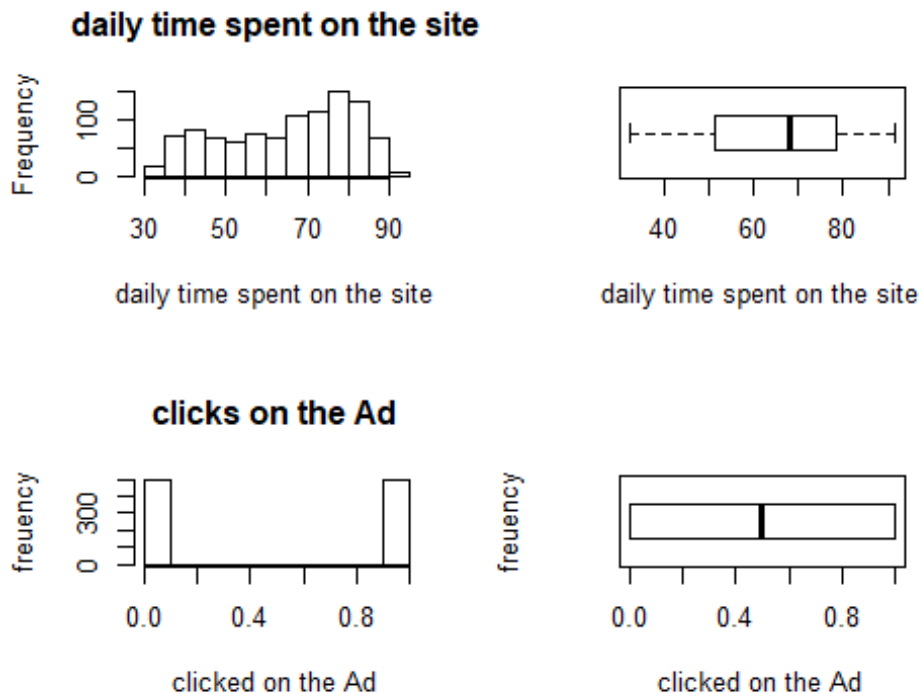
*could not find any outliers*

```
Dt_spent <- data$Daily.Time.Spent.on.Site
dt_clicked <- data$Clicked.on.Ad
```

```

par(mfrow=c(2,2))
hist(Dt_spent,xlab = 'daily time spent on the site', main = 'daily time spent on the site' )
boxplot(Dt_spent,xlab = 'daily time spent on the site', horizontal=TRUE)
hist(dt_clicked, xlab = 'clicked on the Ad', ylab = 'freuency',main = 'clicks on the Ad')
boxplot(dt_clicked, xlab = 'clicked on the Ad', ylab = 'freuency',
horizontal=TRUE)

```



on Average people spent an 65 minutes on the site on a daily basis and a maximum of 88 minutes. clicks on the the Ad are balanced, the number of people who clicked and those who did not click on the AD are equal.

```

library(e1071)
skewness(data$Clicked.on.Ad)

## [1] 0

skewness(data$Daily.Time.Spent.on.Site)

## [1] -0.370646

```

conclusion: 'clicked.on.Ad' and 'Daily.time.Spent.on.Site' features have no skewness.

```

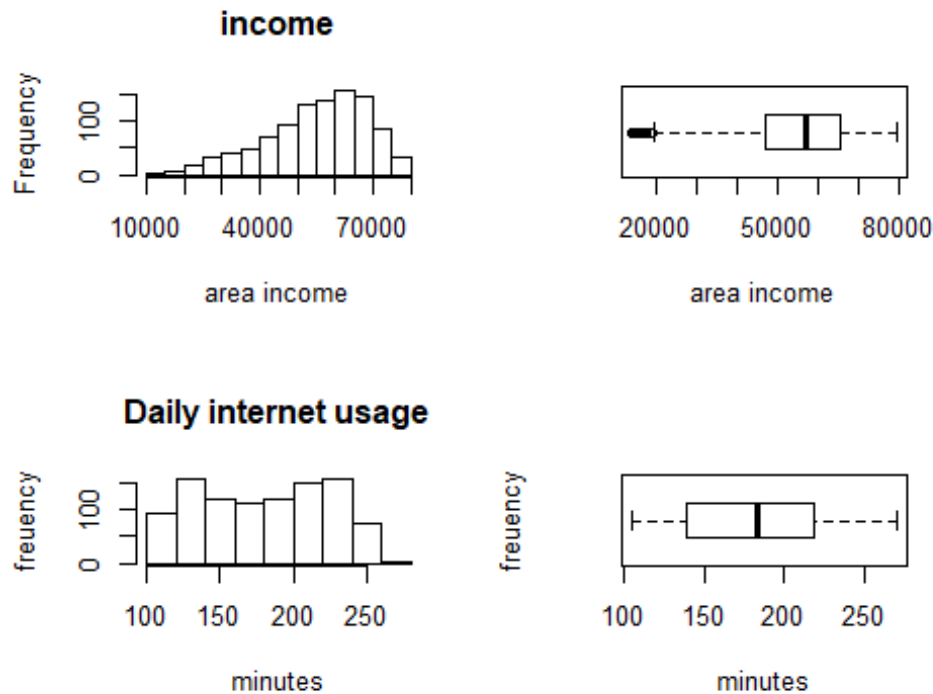
a_income <- data$Area.Income
d_internet_usage <- data$Daily.Internet.Usage
par(mfrow=c(2,2))
hist(a_income,xlab = 'area income', main = 'income' )

```

```

boxplot(a_income,xlab = 'area income', horizontal=TRUE)
hist(d_internet_usage, xlab = 'minutes', ylab = 'freuency',main = 'Daily
internet usage')
boxplot(d_internet_usage, xlab = 'minutes', ylab = 'freuency',
horizontal=TRUE)

```



```

boxplot.stats(data$Area.Income)$out

```

```

## [1] 17709.98 18819.34 15598.29 15879.10 14548.06 13996.50 14775.50
18368.57

```

*The values above are deemed to be outliers in the area income feature as listed above, however they won't be dropped due to the understanding that different countries have different income levels*

```

skewness(data$Daily.Internet.Usage)

```

```

## [1] -0.03343681

```

```

skewness(data$Area.Income)

```

```

## [1] -0.6484229

```

Daily.Internet.Usage and Area.Income features appear to have no skewness. \*The maximum area income was approximately 8,000 and an average of between 5,000 and 6,000. Additional area income appears to be skewed to the left

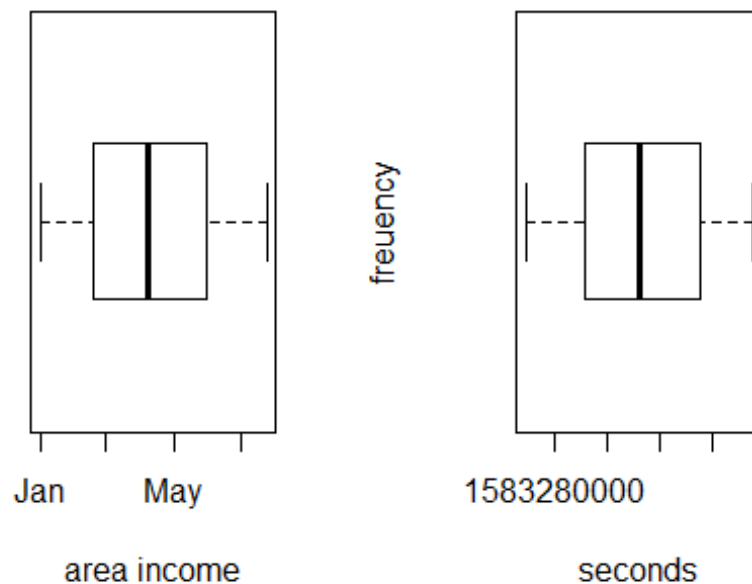
**found no outliers in the daily internet usage** \*The outliers were retained due to the fact that, different countries have different range of income and they appear to be normal, there are low and high income earning countries.

**Feature engineering to allow analysis based on time and date. first step was to split the Timestamp into 'Date' and 'Time' then converted them into Date and Time formats respectively\***

```
data$Date <- sapply(strsplit(as.character(data$Timestamp), " "), "[", 1)
data$Time <- sapply(strsplit(as.character(data$Timestamp), " "), "[", 2)

date <- as.Date(data$Date)
Time <- as.POSIXct(data$Time, format="%H:%M:%S")

par(mfrow=c(1,2))
#plot.ts(date,xlab = 'area income', main = 'income' )
boxplot(date,xlab = 'area income', horizontal=TRUE)
#plot.ts(Time, xlab = 'minutes', ylab = 'freuency',main = 'Daily internet
usage')
boxplot(Time, xlab = 'seconds', ylab = 'freuency', horizontal=TRUE)
```



\*There were no outliers, for Date and Time. # Bivariate analysis **here we will see the relationship between the our dependent variable(Clicked.on.Ad) and the independent variables(aily.Time.Spent.on.Site","Age","Area.Income","Daily.Internet.Usage", etc...)**

```
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"
## [11] "Date"                    "Time"

library(ggplot2)
library(dplyr)

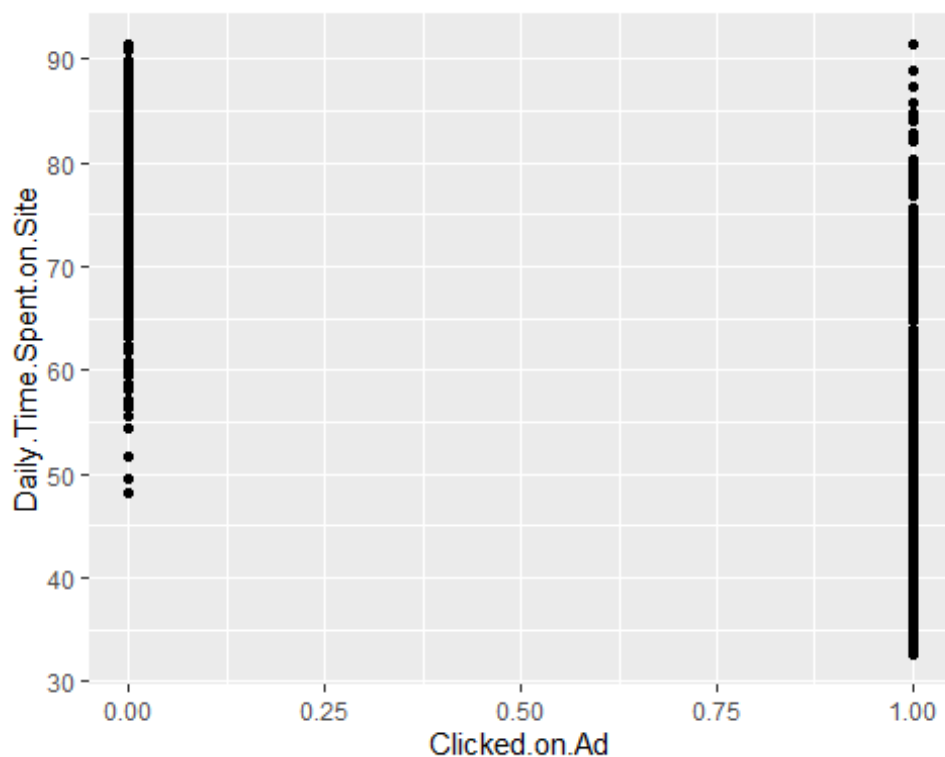
##
## Attaching package: 'dplyr'

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

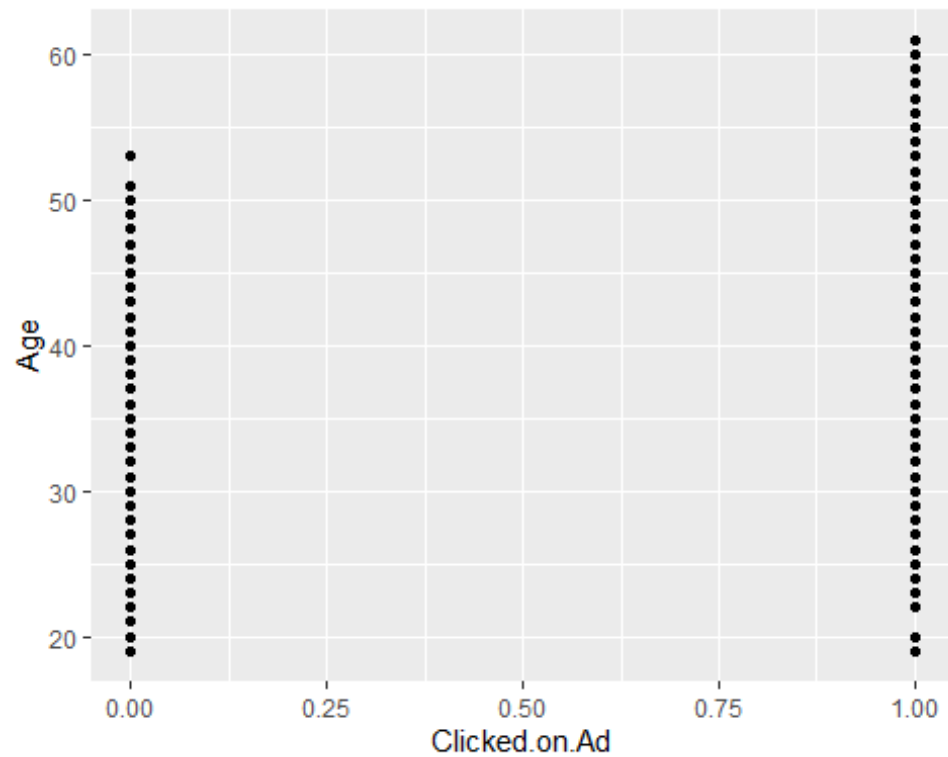
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

par(mfrow=c(2,2))

#ggplot(data = data, aes(x = Daily.Time.Spent.on.Site , fill =
Clicked.on.Ad))+ #geom_histogram(bins= 20, colours = 'cyan' )
ggplot(data = data) +
  geom_point(mapping = aes(x = Clicked.on.Ad , y= Daily.Time.Spent.on.Site))
```

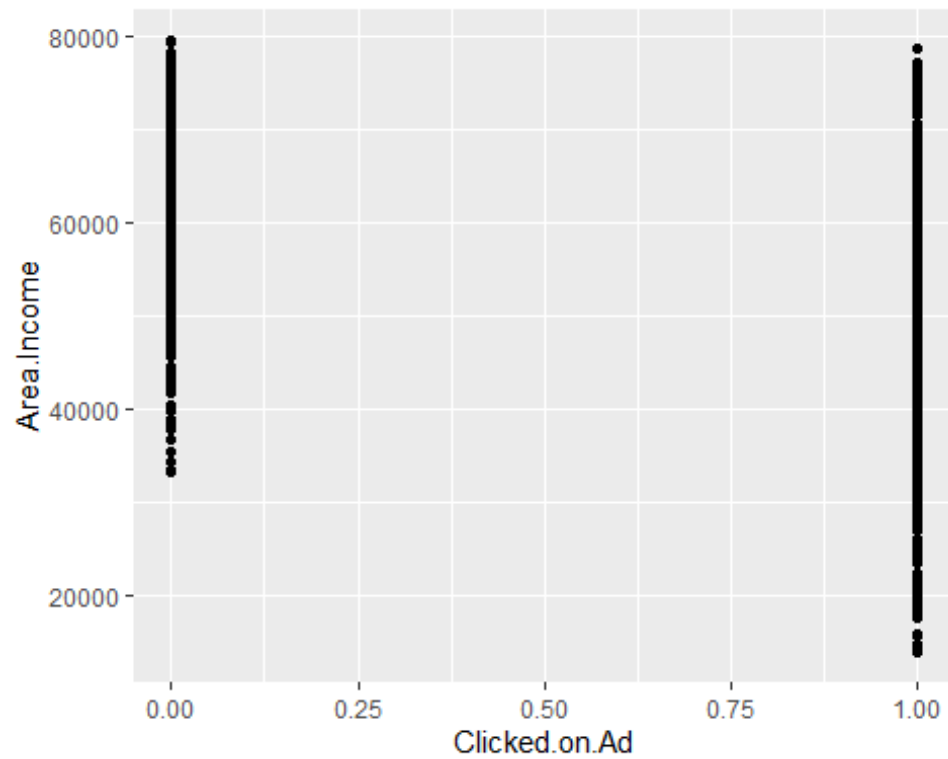


```
ggplot(data = data) +  
  geom_point(mapping = aes(x = Clicked.on.Ad , y= Age))
```

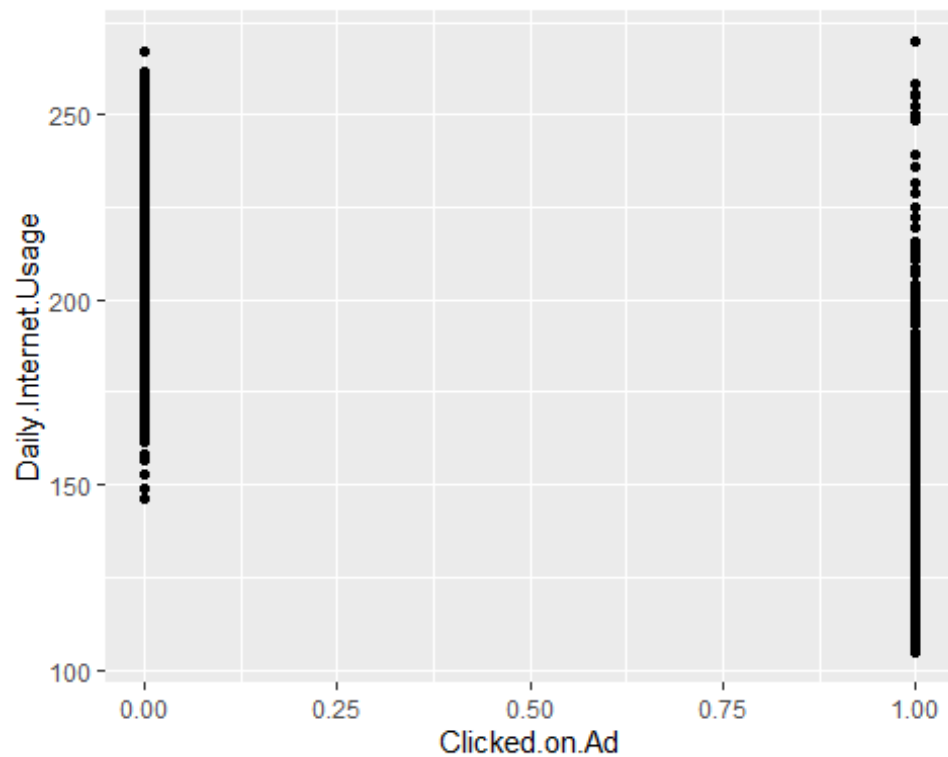


```
ggplot(data = data) +  
  geom_point(mapping = aes(x = Clicked.on.Ad , y= Area.Income))
```

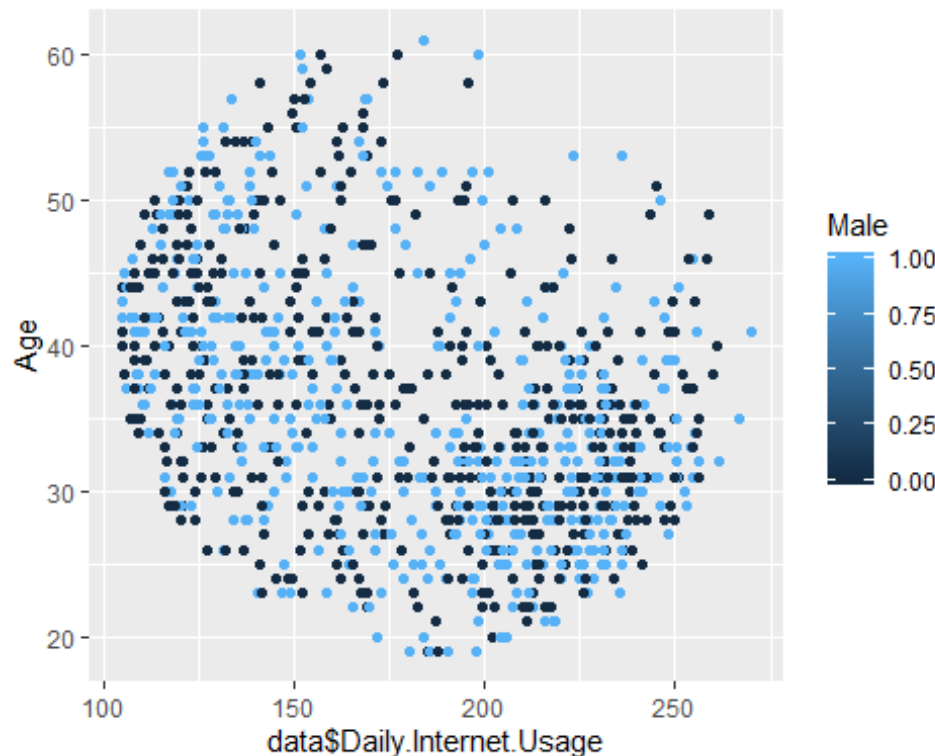




```
ggplot(data = data) +  
  geom_point(mapping = aes(x = Clicked.on.Ad , y= Daily.Internet.Usage))
```



```
ggplot(data = data) +
  geom_point(mapping = aes(x = data$Daily.Internet.Usage , y= Age, fill =
Male, colour = Male))
```



users who spend less than 50 minutes in the site on a daily basis are more likely to click on the Ads. While people with less than 150 minutes on the internet are more likely to click on the Ads. Users range between 25 and 60 year are from both genders. Age does not have much influence on the likelihood of one clicking the Ads, however, older people above 50 years are more likely to click on the Ads. people from low area income of below 3,000 are more likely to click on the Ads

```
factor <- factor(data$Country)
as.numeric(factor)
```

```
##      [1] 216 148 185 104  97 159 146  13  83  79 172  35  61  27  19 198 162
1
##     [19]  30 175  37  37  35 112 213 140 220  81  31  28 166  12 129 188  58
123
##     [37] 142 222 178 119 175 216 218 178 155 218 172 199 215 104  31 224  89
138
##     [55] 217  52 103 202  15 100  35 183  35  33  45  38 176 219 216 157  28
219
##     [73]  49 217  86  51  68 172 101 222 139 154 143 211  28 171  41 231  13
233
##     [91] 107 171 184  81  8  2  95 122  61  18 234 189 129  55  87 209  25
45
##    [109]  88 138 126 117  64  90 215 107 223 133 195  25  81  23 164 229 196
```

|     |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 61  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [127] | 96  | 67  | 58  | 107 | 122 | 181 | 53  | 225 | 22  | 22  | 8   | 183 | 113 | 210 | 80  | 93  | 112 |
| 178 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [145] | 219 | 55  | 150 | 20  | 58  | 196 | 153 | 212 | 110 | 161 | 211 | 36  | 22  | 53  | 50  | 121 | 112 |
| 222 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [163] | 6   | 148 | 63  | 141 | 202 | 211 | 29  | 42  | 170 | 127 | 172 | 192 | 88  | 109 | 113 | 176 | 44  |
| 28  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [181] | 232 | 86  | 166 | 136 | 184 | 192 | 105 | 16  | 38  | 3   | 69  | 110 | 10  | 28  | 167 | 188 | 201 |
| 119 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [199] | 85  | 223 | 9   | 10  | 77  | 108 | 187 | 196 | 52  | 69  | 13  | 186 | 69  | 54  | 114 | 160 | 190 |
| 184 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [217] | 33  | 134 | 55  | 43  | 169 | 65  | 218 | 116 | 69  | 217 | 109 | 17  | 46  | 32  | 207 | 189 | 182 |
| 13  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [235] | 42  | 159 | 219 | 70  | 196 | 133 | 1   | 138 | 74  | 167 | 3   | 185 | 87  | 191 | 208 | 121 | 60  |
| 205 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [253] | 139 | 70  | 71  | 231 | 53  | 166 | 217 | 2   | 74  | 164 | 121 | 210 | 128 | 135 | 3   | 45  | 106 |
| 82  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [271] | 186 | 188 | 84  | 21  | 103 | 94  | 65  | 165 | 114 | 134 | 72  | 158 | 117 | 182 | 4   | 14  | 214 |
| 214 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [289] | 74  | 189 | 152 | 207 | 226 | 144 | 201 | 125 | 233 | 43  | 75  | 75  | 93  | 190 | 136 | 221 | 36  |
| 9   |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [307] | 37  | 195 | 117 | 182 | 58  | 96  | 207 | 180 | 156 | 71  | 54  | 74  | 50  | 14  | 236 | 48  | 226 |
| 168 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [325] | 22  | 7   | 196 | 192 | 70  | 133 | 37  | 204 | 152 | 27  | 192 | 67  | 27  | 135 | 99  | 55  | 64  |
| 137 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [343] | 80  | 91  | 67  | 64  | 95  | 76  | 19  | 148 | 166 | 62  | 120 | 36  | 177 | 173 | 9   | 184 | 1   |
| 15  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [361] | 167 | 6   | 2   | 96  | 68  | 55  | 202 | 1   | 176 | 163 | 184 | 225 | 81  | 51  | 160 | 7   | 54  |
| 166 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [379] | 110 | 42  | 114 | 2   | 75  | 59  | 237 | 52  | 36  | 141 | 94  | 126 | 172 | 44  | 18  | 203 | 209 |
| 64  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [397] | 38  | 179 | 34  | 220 | 62  | 126 | 18  | 4   | 116 | 139 | 7   | 18  | 68  | 207 | 92  | 103 | 26  |
| 16  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [415] | 50  | 146 | 151 | 55  | 97  | 161 | 120 | 109 | 72  | 220 | 48  | 224 | 123 | 73  | 164 | 129 | 236 |
| 49  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [433] | 48  | 188 | 146 | 59  | 17  | 171 | 43  | 26  | 189 | 128 | 65  | 82  | 215 | 210 | 167 | 156 | 1   |
| 6   |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [451] | 61  | 69  | 170 | 14  | 78  | 163 | 226 | 45  | 63  | 138 | 131 | 60  | 200 | 115 | 183 | 205 | 198 |
| 219 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [469] | 99  | 49  | 13  | 70  | 160 | 101 | 64  | 71  | 14  | 92  | 234 | 119 | 59  | 214 | 115 | 226 | 21  |
| 99  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [487] | 52  | 32  | 4   | 151 | 210 | 81  | 73  | 87  | 102 | 93  | 62  | 44  | 146 | 124 | 13  | 227 | 137 |
| 57  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [505] | 51  | 130 | 105 | 174 | 40  | 76  | 3   | 171 | 157 | 217 | 88  | 139 | 81  | 4   | 94  | 141 | 66  |
| 66  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [523] | 199 | 144 | 223 | 234 | 234 | 36  | 153 | 13  | 33  | 120 | 19  | 73  | 228 | 228 | 29  | 231 | 146 |
| 131 |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [541] | 105 | 17  | 3   | 220 | 77  | 36  | 85  | 209 | 99  | 195 | 22  | 189 | 13  | 85  | 45  | 164 | 16  |
| 47  |       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ##  | [559] | 234 | 154 | 85  | 230 | 26  | 127 | 231 | 149 | 224 | 2   | 126 | 90  | 235 | 98  | 171 | 227 | 9   |

```

72
## [577] 9 218 94 190 54 182 169 207 51 138 119 187 149 79 100 153 120
199
## [595] 223 99 183 141 94 164 114 66 176 114 83 212 160 67 107 40 196
138
## [613] 208 26 37 60 236 89 138 16 39 73 187 71 35 116 144 231 161
102
## [631] 166 21 52 71 194 166 20 26 23 233 15 141 56 175 29 66 90
66
## [649] 135 57 206 181 151 81 126 188 34 55 115 151 172 5 121 44 232
208
## [667] 64 140 103 96 192 53 173 236 75 58 16 213 218 21 72 133 73
60
## [685] 171 223 34 123 105 8 44 234 117 95 230 230 86 82 206 178 117
131
## [703] 45 222 131 104 106 135 217 147 44 150 80 48 188 96 168 193 227
140
## [721] 170 217 221 157 156 222 230 225 11 204 211 74 70 183 188 35 16
204
## [739] 202 214 112 114 50 121 237 50 96 69 150 204 19 165 104 20 197
7
## [757] 191 181 221 187 82 231 119 130 27 32 197 55 62 213 71 75 33
34
## [775] 136 195 2 26 107 31 178 27 98 77 225 111 79 184 100 50 158
121
## [793] 83 169 110 100 21 147 54 106 237 5 123 54 217 95 150 227 132
234
## [811] 183 226 6 40 203 233 237 42 181 176 139 75 56 202 169 69 167
232
## [829] 107 99 162 116 131 1 14 138 137 43 53 20 127 1 123 196 149
198
## [847] 95 193 40 221 230 7 205 237 228 119 61 81 17 199 109 82 139
169
## [865] 7 41 137 212 11 154 64 38 52 205 235 213 11 63 19 4 180
3
## [883] 218 136 196 124 71 63 130 136 160 84 56 153 151 20 207 62 207
166
## [901] 119 35 124 231 123 104 185 126 157 230 216 165 125 92 66 62 155
211
## [919] 228 195 128 112 115 16 90 66 27 54 192 59 24 105 177 2 145
237
## [937] 77 29 206 162 83 79 32 122 129 118 55 97 167 40 91 46 123
223
## [955] 101 38 202 131 200 60 188 36 20 90 130 100 33 1 119 151 95
161
## [973] 127 228 54 137 155 71 106 157 33 229 137 32 71 235 158 169 17
182
## [991] 214 47 142 102 136 117 27 141 86 29

```

```
summary(data)
```

```

## Daily.Time.Spent.on.Site      Age      Area.Income
Daily.Internet.Usage
## Min.      :32.60      Min.      :19.00      Min.      :13996      Min.      :104.8
## 1st Qu.:51.36      1st Qu.:29.00      1st Qu.:47032      1st Qu.:138.8
## Median :68.22      Median :35.00      Median :57012      Median :183.1
## Mean   :65.00      Mean   :36.01      Mean   :55000      Mean   :180.0
## 3rd Qu.:78.55      3rd Qu.:42.00      3rd Qu.:65471      3rd Qu.:218.8
## Max.   :91.43      Max.   :61.00      Max.   :79485      Max.   :270.0
##
##                               Ad.Topic.Line      City
## Adaptive 24hour Graphic Interface      : 1      Lisamouth      : 3
## Adaptive asynchronous attitude      : 1      Williamsport      : 3
## Adaptive context-sensitive application : 1      Benjaminchester: 2
## Adaptive contextually-based methodology: 1      East John      : 2
## Adaptive demand-driven knowledgebase : 1      East Timothy      : 2
## Adaptive uniform capability      : 1      Johnstad      : 2
## (Other)      :994      (Other)      :986
##      Male      Country      Timestamp
Clicked.on.Ad
## Min.      :0.000      Czech Republic: 9      2016-01-01 02:52:10: 1      Min.
:0.0
## 1st Qu.:0.000      France      : 9      2016-01-01 03:35:35: 1      1st
Qu.:0.0
## Median :0.000      Afghanistan : 8      2016-01-01 05:31:22: 1      Median
:0.5
## Mean   :0.481      Australia  : 8      2016-01-01 08:27:06: 1      Mean
:0.5
## 3rd Qu.:1.000      Cyprus      : 8      2016-01-01 15:14:24: 1      3rd
Qu.:1.0
## Max.   :1.000      Greece      : 8      2016-01-01 20:17:49: 1      Max.
:1.0
##      (Other)      :950      (Other)      :994
##      Date      Time
## Length:1000      Length:1000
## Class :character      Class :character
## Mode :character      Mode :character
##
##
##
##

```

#multivariate Analysis here we explore simltenously explore how other (more than 1) independent variabel influence the likelihood of one clicking the Ad.

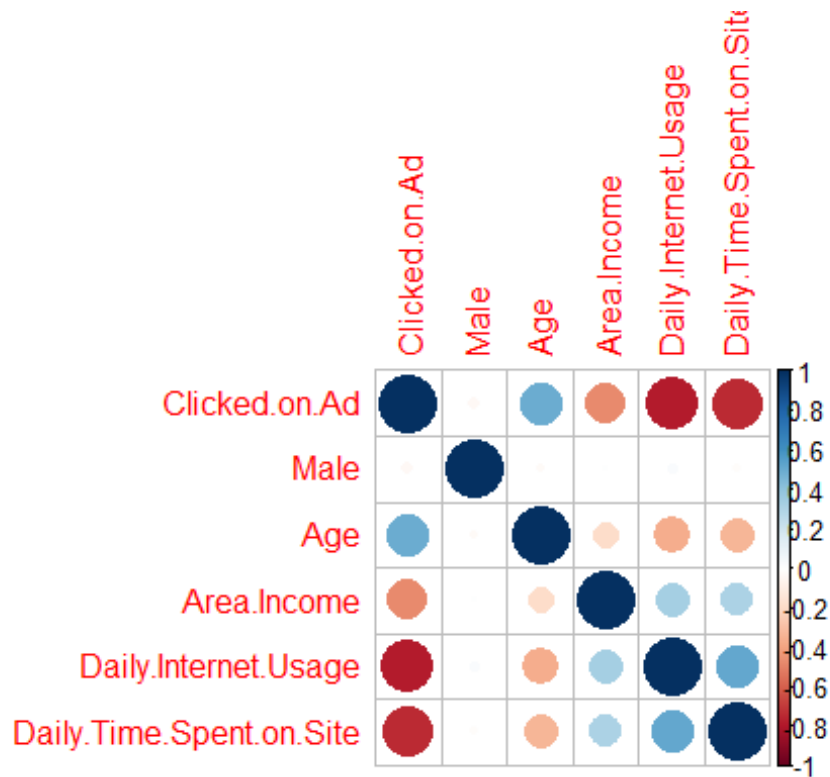
```

sub_data <- c('Clicked.on.Ad', 'Male', 'Age', 'Area.Income',
'Daily.Internet.Usage', 'Daily.Time.Spent.on.Site')
num <- data[,sub_data]
num_cor <- cor(num)
library(corrplot)

```

```
## corrplot 0.84 loaded
```

```
corrplot(num_cor)
```



- There is a strong positive correlation between Age and the those who clicked the AD,hence the older the user the more likely he/she will click the ad.\*
- Daily inter usage and Daily time spent o the site has a strong negative correlation to Clicking the Ad.Thus, the less time one spends on both the site and the internet he/she will be more likely to click on the Ads.\*
- Gender(Rep' by - 'Male') has no correlation to the likelihood of one clicking the Ad.\*
- Area income as weak positive corelation with the the likelihood of one clicking the Ad, suggesting that , the higher the area income the more likely that the person will clkick on the Ad.\*