
Title: Cryptograph Ads Prediction Author: Joseph Njuguna Date: 27/5/22

#1.Defining the question

#a) Specifying the question Identify individuals most likely to click on ads.

#b) Defining the metric for success Ability to identify individuals that click on ads.

#c) Understanding 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.

#d) Recording the experimental design -Define the question, the metric for success, the context, experimental design taken and the appropriateness of the available data to answer the given question. -Find and deal with outliers, anomalies, and missing data within the dataset. -Perform univariate and bivariate analysis. -From your insights provide a conclusion and recommendation.

#2. Reading the data

```
# choosing working directory that has uploaded file
getwd()
```

```
## [1] "C:/Users/jojo/Desktop/R"
```

```
#setwd("C:/Users/jojo/Downloads/R basics")
# using .csv to read dataset
# must have utils package installed.
adsop <- read.csv("advertising.csv", header= TRUE, sep= ",")
# view dataset
View(adsop)
```

#3. Checking the data

viewing first 5 rows of our dataset

```
head(adsop)
```

```
##   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
```

viewing last 5 rows of our dataset

```
tail(adsop)
```

```
##      Daily.Time.Spent.on.Site Age Area.Income Daily.Internet.Usage
## 995                     43.70  28    63126.96               173.01
## 996                     72.97  30    71384.57               208.58
## 997                     51.30  45    67782.17               134.42
## 998                     51.63  51    42415.72               120.37
## 999                     55.55  19    41920.79               187.95
## 1000                    45.01  26    29875.80               178.35
##                            Ad.Topic.Line          City Male
## 995         Front-line bifurcated ability  Nicholasland    0
## 996         Fundamental modular algorithm     Duffystad    1
## 997       Grass-roots cohesive monitoring   New Darlene    1
## 998          Expanded intangible solution South Jessica    1
## 999  Proactive bandwidth-monitored policy   West Steven    0
## 1000      Virtual 5thgeneration emulation   Ronniemouth    0
##                      Country           Timestamp Clicked.on.Ad
## 995                 Mayotte 2016-04-04 03:57:48             1
## 996                 Lebanon 2016-02-11 21:49:00             1
## 997  Bosnia and Herzegovina 2016-04-22 02:07:01             1
## 998                Mongolia 2016-02-01 17:24:57             1
## 999               Guatemala 2016-03-24 02:35:54             0
## 1000                 Brazil 2016-06-03 21:43:21             1
```

checking data types

```
str(adsop)
```

```
## 'data.frame':    1000 obs. of  10 variables:
##  $ Daily.Time.Spent.on.Site: num  69 80.2 69.5 74.2 68.4 ...
##  $ Age                     : int  35 31 26 29 35 23 33 48 30 20 ...
##  $ Area.Income             : num  61834 68442 59786 54806 73890 ...
##  $ Daily.Internet.Usage    : num  256 194 236 246 226 ...
##  $ Ad.Topic.Line           : chr  "Cloned 5thgeneration orchestration" "Monitored national standardi
##  $ City                    : chr  "Wrightburgh" "West Jodi" "Davidton" "West Terrifurt" ...
```

```
##  $ Male                    : int  0 1 0 1 0 1 0 1 1 1 ...
##  $ Country                 : chr  "Tunisia" "Nauru" "San Marino" "Italy" ...
##  $ Timestamp               : chr  "2016-03-27 00:53:11" "2016-04-04 01:39:02" "2016-03-13 20:35:42" "
##  $ Clicked.on.Ad           : int  0 0 0 0 0 0 0 1 0 0 ...

# Our data types are numeric, integer, character.
```

shape of data

```
dim(adsop)
```

```
## [1] 1000   10

# Our dataset has 1000rows, 10 columns
```

descriptive statistical summary of our dataset

```
summary(adsop)
```

```
##  Daily.Time.Spent.on.Site      Age          Area.Income    Daily.Internet.Usage
##  Min.   :32.60            Min.   :19.00   Min.   :13996   Min.   :104.8
##  1st Qu.:51.36            1st Qu.:29.00   1st Qu.:47032   1st Qu.:138.8
##  Median :68.22            Median :35.00   Median :57012   Median :183.1
##  Mean   :65.00            Mean   :36.01   Mean   :55000   Mean   :180.0
##  3rd Qu.:78.55            3rd Qu.:42.00   3rd Qu.:65471   3rd Qu.:218.8
##  Max.   :91.43            Max.   :61.00   Max.   :79485   Max.   :270.0
##  Ad.Topic.Line          City                Male         Country
##  Length:1000        Length:1000        Min.   :0.000   Length:1000
##  Class :character   Class :character   1st Qu.:0.000   Class :character
##  Mode  :character   Mode  :character   Median :0.000   Mode  :character
##                                        Mean   :0.481
##                                        3rd Qu.:1.000
##                                        Max.   :1.000
##   Timestamp         Clicked.on.Ad
##  Length:1000        Min.   :0.0
##  Class :character   1st Qu.:0.0
##  Mode  :character   Median :0.5
##                     Mean   :0.5
##                     3rd Qu.:1.0
##                     Max.   :1.0
```

#4. Tidying the data

checking for duplicate records in our df

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

```
##  [1] Daily.Time.Spent.on.Site Age                      Area.Income
##  [4] Daily.Internet.Usage     Ad.Topic.Line            City
##  [7] Male                     Country                  Timestamp
## [10] Clicked.on.Ad
## <0 rows> (or 0-length row.names)

# No duplicate records in our dataset
```

missing values

list of columns and mising values

```
colSums(is.na(adsop))
```

```
## 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

# No missing values in our dataset.
```

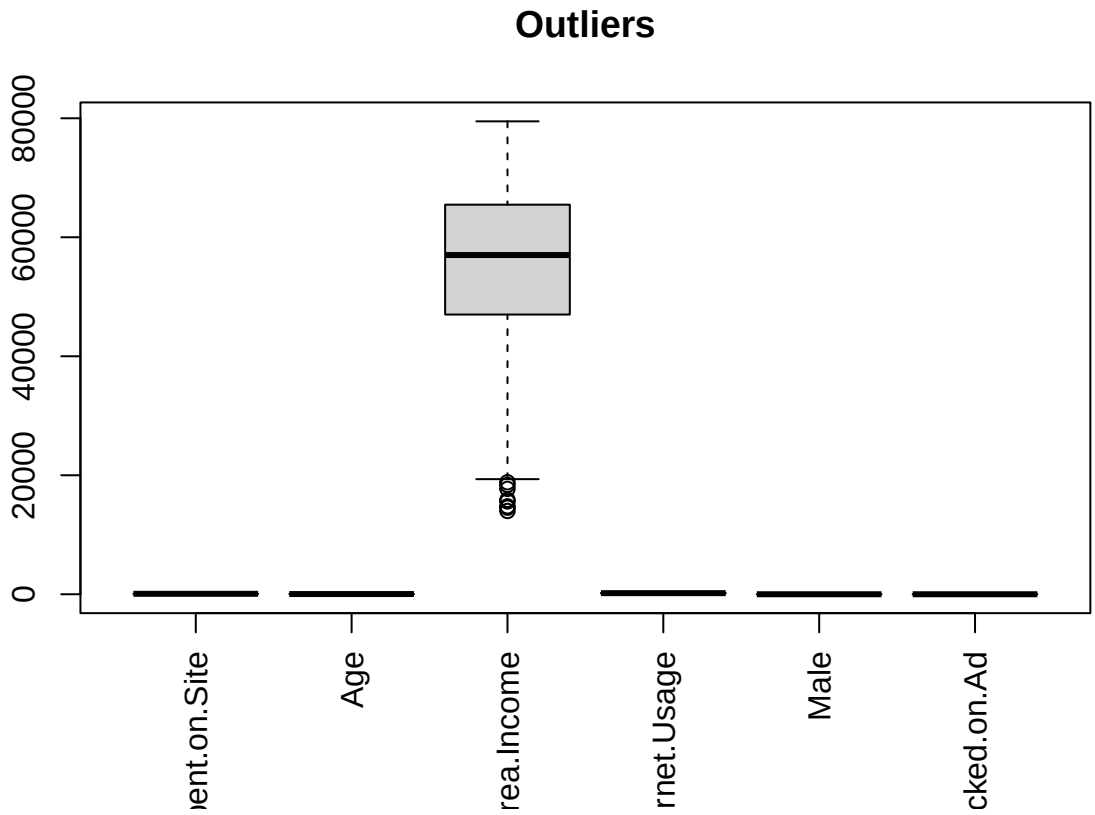
checking for outliers

listing numerical columns as we can only get outliers for numerical columns

```
numerical <- list(adsop$Daily.Time.Spent.on.Site,adsop$Age,
adsop$Area.Income,adsop$Daily.Internet.Usage,adsop$Male, adsop$Clicked.on.Ad)
```

creating boxplots

```
boxplot(numerical, names=c('Daily.Time.Spent.on.Site', 'Age', 'Area.Income', 'Daily.Internet.Usage', 'Ma
```



Outliers only exist in our area.income column # It is not necessary to remove them.

#5. Data Anlaysis # Univariate Analysis # Measures of central tendecy

Mean

```
colMeans(adsop[sapply(adsop, is.numeric)])
```

```
## Daily.Time.Spent.on.Site                      Age              Area.Income
##                  65.0002                 36.0090                 55000.0001
##     Daily.Internet.Usage                     Male            Clicked.on.Ad
##                 180.0001                  0.4810                  0.5000
```

```
# The mean age of respondents is 36 years.
# Mean area income is $55,000.
# Mean time spent on site daily is 65 minutes.
```

Median

daily time spent on site

```
median(adsop$Daily.Time.Spent.on.Site)
```

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

age

```
median(adsop$Age)
```

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

area income

```
median(adsop$Area.Income)
```

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

daily internet usage

```
median(adsop$Daily.Internet.Usage)
```

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

Measures of dispersion

Variance

daily time spent on site

```
var(adsop$Daily.Time.Spent.on.Site)
```

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

age

```
var(adsop$Age)
```

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

area income

```
var(adsop$Area.Income)
```

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

daily internet usage

```
var(adsop$Daily.Internet.Usage)
```

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

Standard deviation

daily time spent on site, age, area income, internet usage, male, clicked on ad

```
sd(adsop$Daily.Time.Spent.on.Site)
```

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

```
sd(adsop$Age)
```

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

```
sd(adsop$Area.Income)
```

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

```
sd(adsop$Daily.Internet.Usage)
```

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

```
sd(adsop$Male)
```

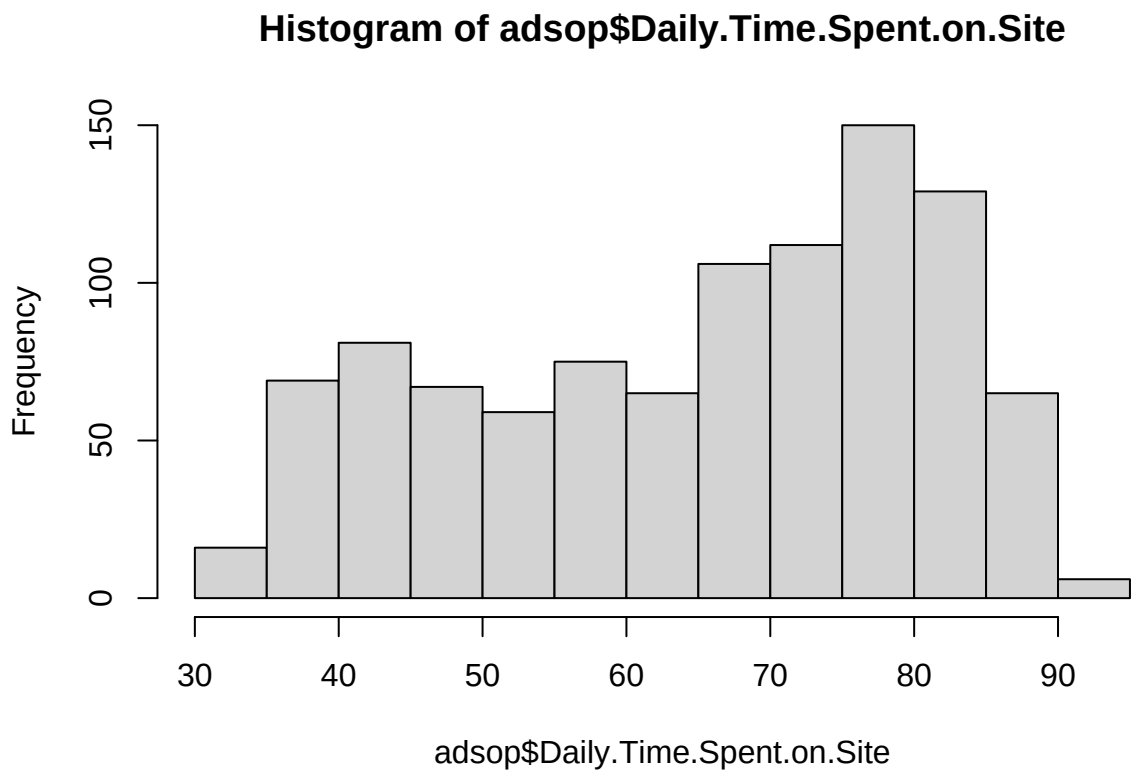
```
## [1] 0.4998889
```

```
sd(adsop$Clicked.on.Ad)
```

```
## [1] 0.5002502
```

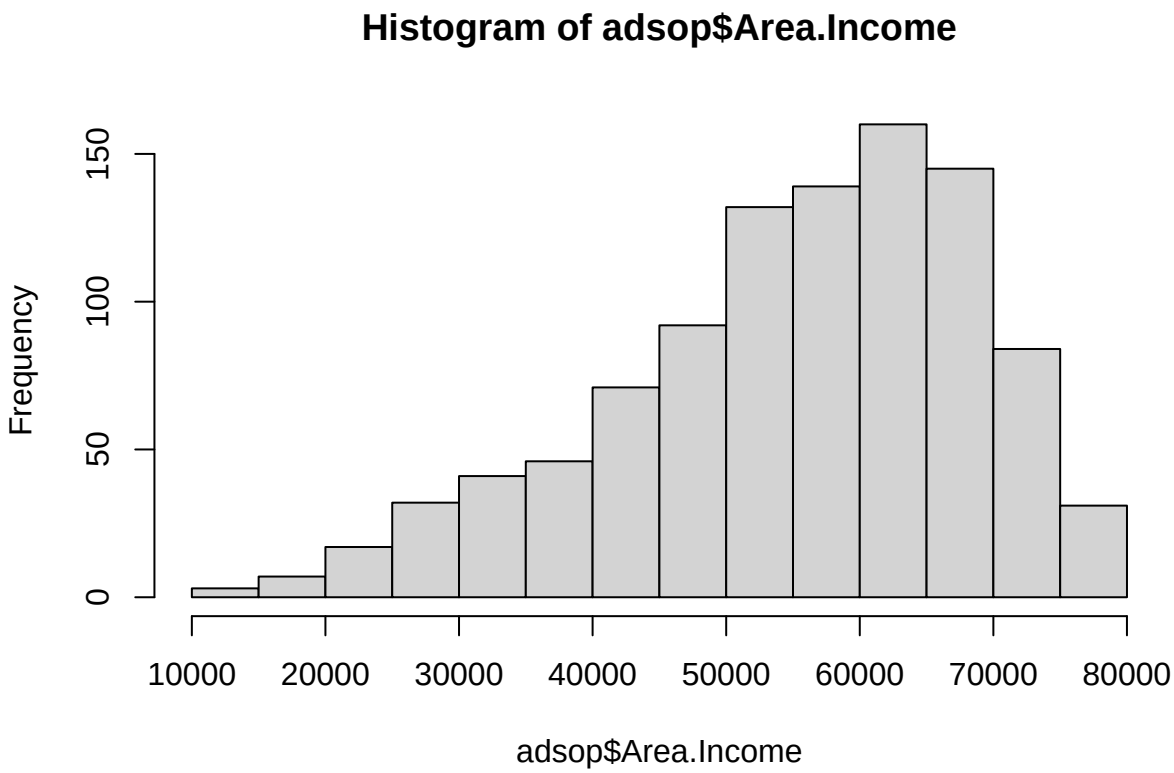
histogram - time spent on site

```
hist(adsop$Daily.Time.Spent.on.Site)
```



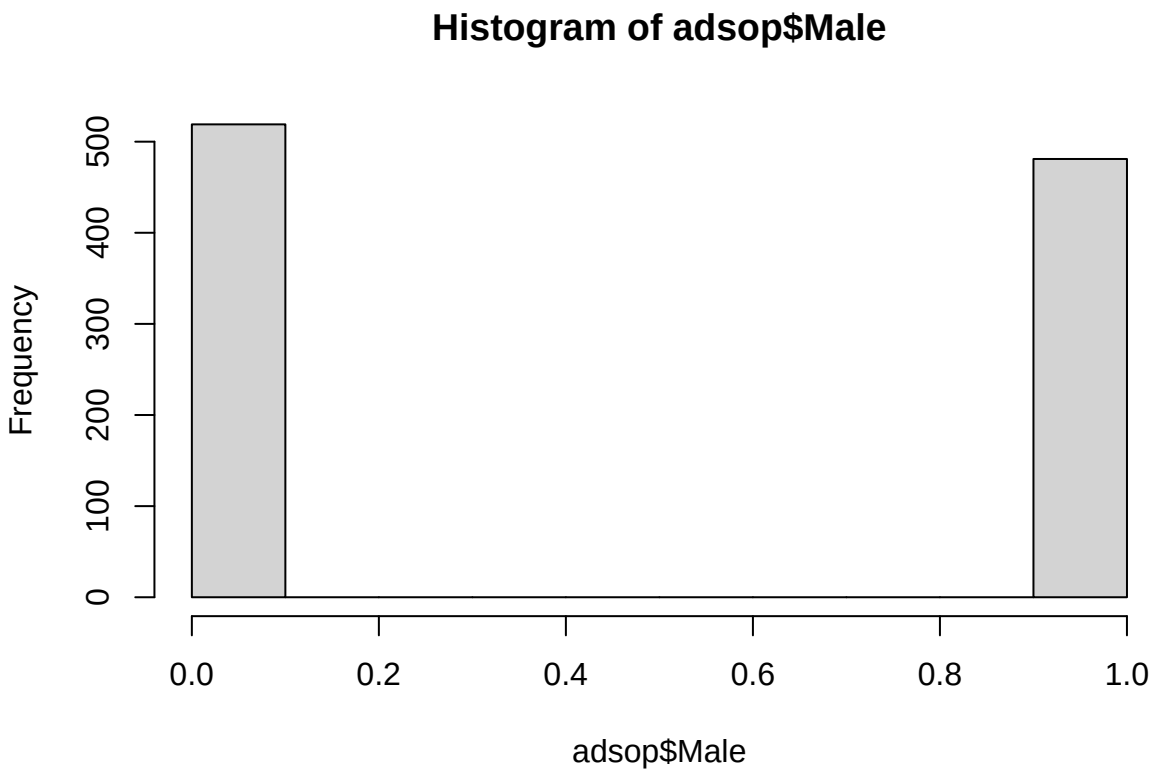
80 minutes is the most frequent time spent by users on site # hist - Area income

```
hist(adsop$Area.Income)
```



Highest area income revenue is $60,000 # histogram on gender distribution

```
hist(adsop$Male)
```

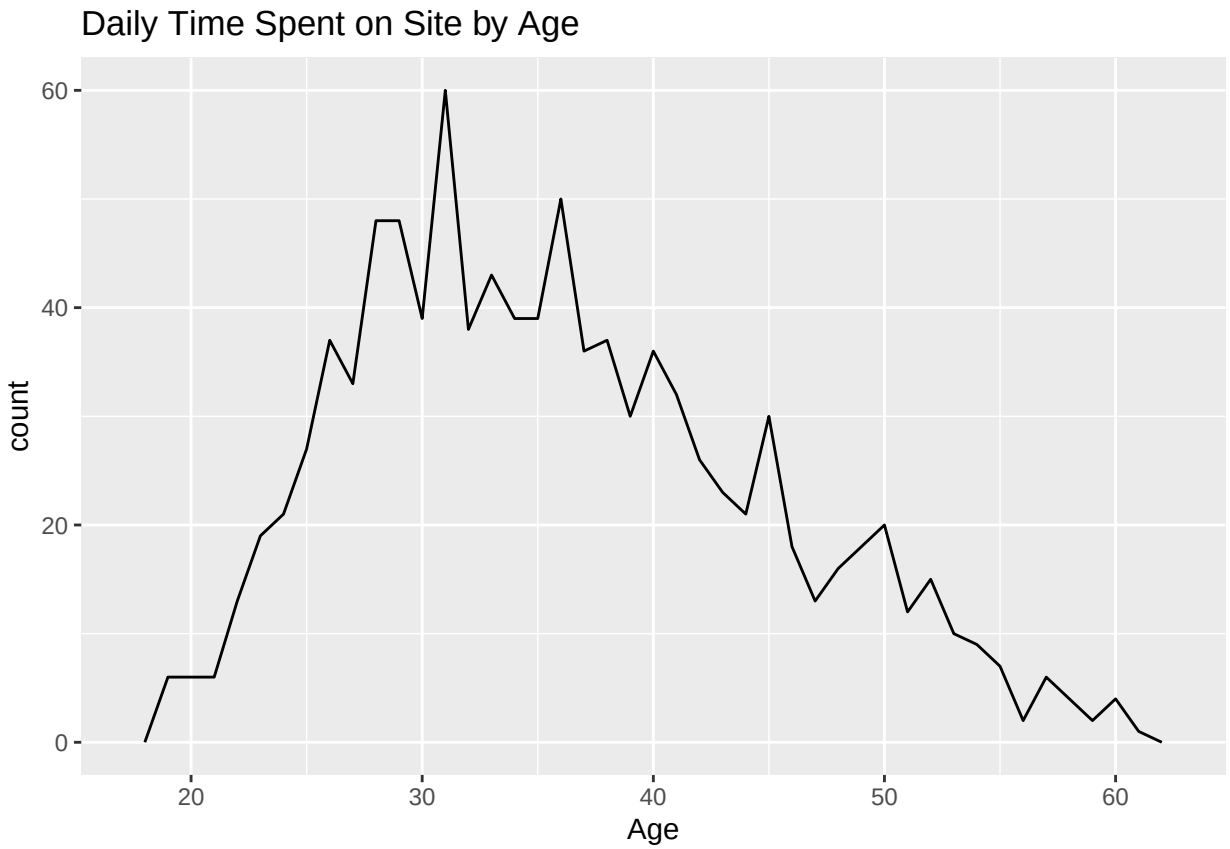


Number of female respondents were slightly higher than male respondents.

Bivariate Analysis

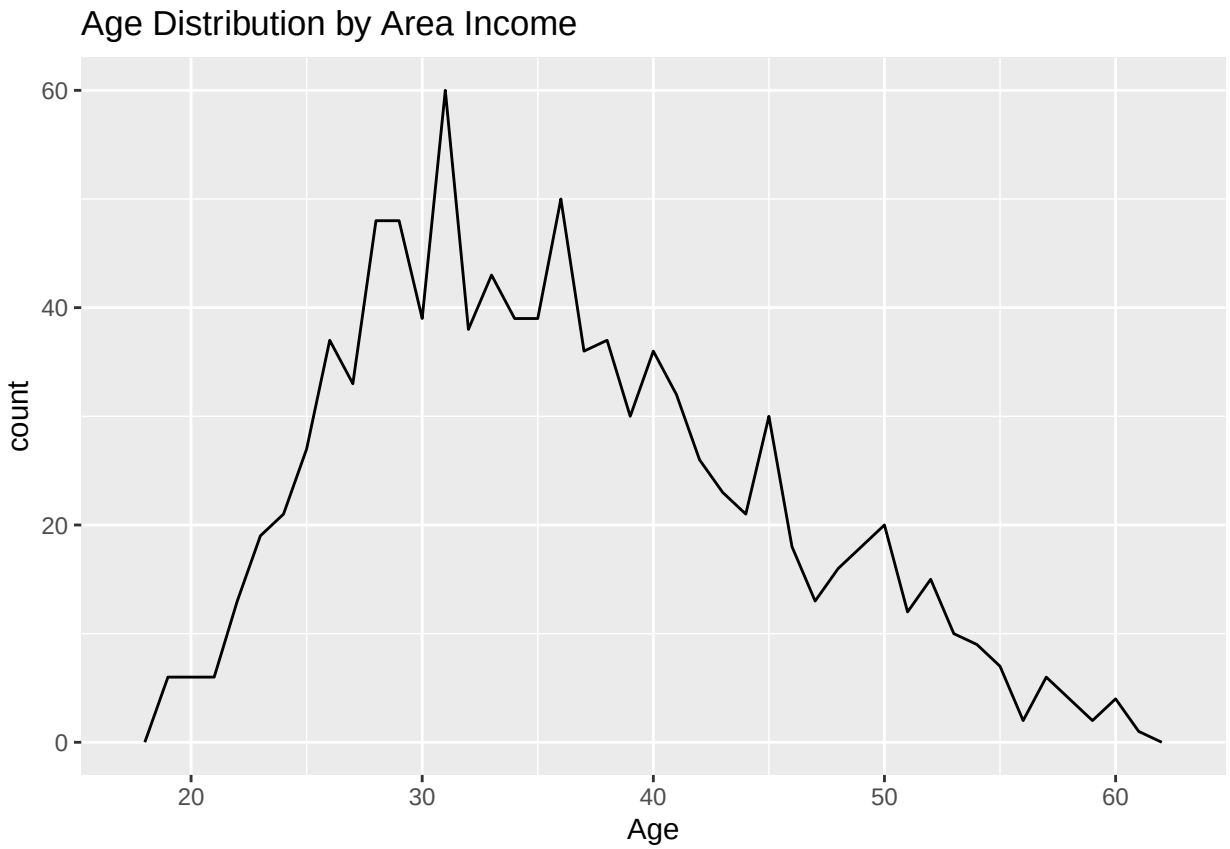
Numerical-Numerical variables

```
#ViZ of daily time spent on site by age
# ---
#
library(ggplot2)
ggplot(adsop, aes(Age, colour = Daily.Time.Spent.on.Site)) +
geom_freqpoly(binwidth = 1) + labs(title="Daily Time Spent on Site by Age")
```



We can observe that the highest amount of daily time spent on site increases steadily from age group 23-33 and decreases after that.

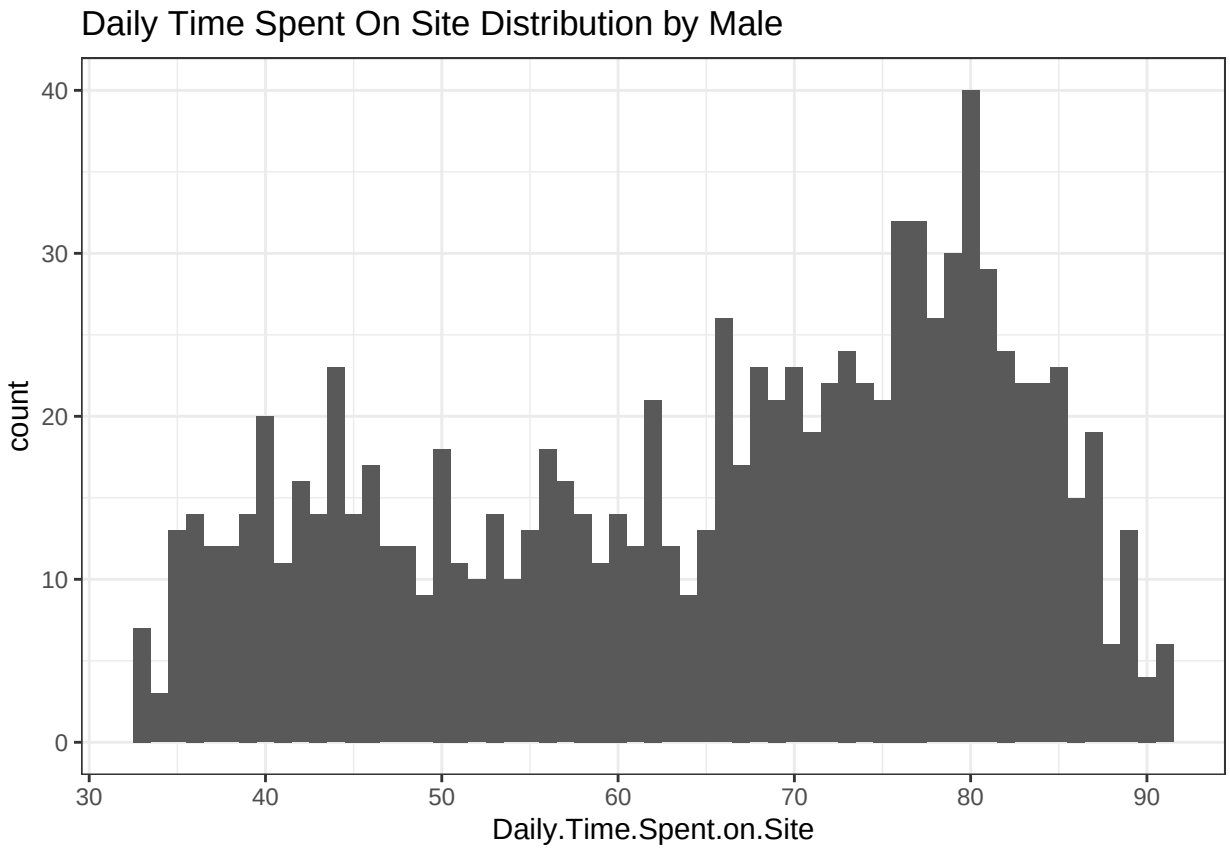
```
ggplot(adsop, aes(Age, colour = Area.Income)) +
geom_freqpoly(binwidth = 1) + labs(title="Age Distribution by Area Income")
```



We observe that area income was highest around the 30-33 years age group.

Numerical-Categorical

```
c <- ggplot(adsop, aes(x=Daily.Time.Spent.on.Site, fill=Male, color=Male)) +
geom_histogram(binwidth = 1) + labs(title="Daily Time Spent On Site Distribution by Male")
c + theme_bw()
```



#6. Modelling

Decision trees

```
#data splicing
set.seed(12345)
train <- sample(1:nrow(adsop),size = ceiling(0.80*nrow(adsop)),replace = FALSE)
# training set
adsop_train <- adsop[train,]
# test set
adsop_test <- adsop[-train,]

adsop_test[-10]
```

```
##      Daily.Time.Spent.on.Site Age Area.Income Daily.Internet.Usage
## 20                      74.58  40    23821.72               135.51
## 22                      84.59  35    60015.57               226.54
## 29                      70.20  34    32708.94               119.20
## 35                      57.64  57    45580.92               133.81
## 39                      50.43  46    57425.87               119.32
## 58                      44.98  49    52336.64               129.31
## 78                      82.73  33    54541.56               238.99
## 87                      77.51  36    73600.28               200.55
```

```
## 97                  45.72  36    22473.08             154.02
## 98                  39.94  41    64927.19             156.30
## 100                 79.71  34    69456.83             211.65
## 101                 41.49  53    31947.65             169.18
## 102                 63.60  23    51864.77             235.28
## 105                 66.49  20    56884.74             202.16
## 116                 76.99  31    56729.78             244.34
## 126                 79.57  31    61227.59             230.93
## 134                 75.70  29    67633.44             215.44
## 141                 83.86  31    64122.36             190.25
## 143                 55.60  44    65953.76             124.38
## 147                 56.04  49    65120.86             128.95
## 151                 58.18  25    69112.84             176.28
## 155                 79.52  27    64287.78             183.48
## 160                 75.55  36    73234.87             159.24
## 163                 72.01  31    57756.89             251.00
## 168                 86.06  31    53767.12             219.72
## 178                 72.04  22    60309.58             199.43
## 179                 44.57  31    38349.78             133.17
## 183                 62.95  60    36752.24             157.04
## 186                 46.88  54    43444.86             136.64
## 191                 50.08  30    41629.86             123.91
## 194                 43.67  53    46004.31             143.79
## 206                 58.22  29    37345.24             120.90
## 209                 59.51  58    39132.64             140.83
## 213                 76.87  27    54774.77             235.35
## 219                 34.87  40    59621.02             200.23
## 221                 77.88  37    55353.41             254.57
## 223                 49.95  39    68737.75             136.59
## 224                 60.94  41    76893.84             154.97
## 228                 34.86  38    49942.66             154.75
## 235                 37.32  50    56735.14             199.25
## 236                 40.42  45    40183.75             133.90
## 237                 76.77  36    58348.41             123.51
## 241                 80.03  44    24030.06             150.84
## 254                 46.08  30    66784.81             164.63
## 273                 76.28  33    52686.47             254.34
## 277                 87.30  35    58337.18             216.87
## 284                 71.76  35    69285.69             237.39
## 288                 82.87  37    67938.77             213.36
## 293                 55.20  39    76560.59             159.46
## 295                 71.55  36    75687.46             163.99
## 298                 72.07  31    69710.51             226.45
## 299                 80.39  31    66269.49             214.74
## 301                 69.97  28    55041.60             250.00
## 305                 33.52  43    42191.61             165.56
## 306                 79.81  24    56194.56             178.85
## 308                 82.70  35    61383.79             231.07
## 310                 54.92  54    23975.35             161.16
## 317                 75.84  21    48537.18             186.98
## 321                 51.50  34    67050.16             135.31
## 327                 34.78  48    42861.42             208.21
## 328                 76.84  32    65883.39             231.59
## 330                 41.47  31    60953.93             219.79
```

```
## 331  80.71  26  58476.57  200.58
## 334  79.36  34  57260.41  245.78
## 344  83.97  28  56986.73  205.50
## 350  60.91  19  53309.61  184.94
## 353  78.19  30  62475.99  228.81
## 362  62.14  41  59397.89  110.93
## 365  69.11  42  73608.99  231.48
## 366  71.90  54  61228.96  140.15
## 384  74.63  26  51975.41  235.99
## 386  80.59  37  67744.56  224.23
## 397  42.05  51  28357.27  174.55
## 398  50.52  28  66929.03  219.69
## 402  84.95  34  56379.30  230.36
## 408  48.73  27  34191.23  142.04
## 416  41.18  43  41866.55  129.25
## 426  60.07  42  65963.37  120.75
## 427  54.43  37  75180.20  154.74
## 431  88.72  32  43870.51  211.87
## 434  85.23  36  64238.71  212.92
## 437  56.39  27  38817.40  248.12
## 443  36.44  39  52400.88  147.64
## 447  38.10  34  60641.09  214.38
## 460  79.82  26  51512.66  223.28
## 461  48.03  40  25598.75  134.60
## 462  32.99  45  49282.87  177.46
## 466  45.70  33  67384.31  151.12
## 477  78.76  24  46422.76  219.98
## 482  83.26  40  70225.60  187.76
## 489  35.49  48  43974.49  159.77
## 492  59.12  33  28210.03  124.54
## 501  51.56  46  63102.19  124.85
## 505  66.12  49  39031.89  113.80
## 507  77.05  31  62161.26  236.64
## 523  79.16  32  69562.46  202.90
## 524  67.94  43  68447.17  128.16
## 530  89.37  34  42760.22  162.03
## 542  75.65  39  64021.55  247.90
## 551  75.80  36  71222.40  224.90
## 556  72.53  37  73474.82  223.93
## 557  56.20  49  53549.94  114.85
## 560  78.64  31  60283.47  235.28
## 563  59.52  44  67511.86  251.08
## 569  64.51  42  79484.80  190.71
## 573  72.80  35  63551.67  249.54
## 575  46.66  45  49101.67  118.16
## 581  69.15  46  36424.94  112.72
## 593  73.15  25  64631.22  211.12
## 594  82.07  24  50337.93  193.97
## 596  43.97  36  68863.95  156.97
## 599  83.53  36  67686.16  204.56
## 602  52.13  50  40926.93  118.27
## 609  62.06  40  63493.60  116.27
## 616  40.01  53  51463.17  161.77
## 626  67.91  23  55002.05  146.80
```

```
## 627                85.77  27    52261.73            191.78
## 630                73.94  26    55411.06            236.15
## 631                63.43  29    66504.16            236.75
## 632                84.59  36    47169.14            241.80
## 636                35.76  51    45522.44            195.07
## 641                64.63  45    15598.29            158.80
## 642                81.29  28    33239.20            219.72
## 646                46.14  28    39723.97            137.97
## 650                84.31  29    47139.21            225.87
## 666                58.05  32    15879.10            195.54
## 671                77.89  26    64188.50            201.54
## 673                89.21  33    44078.24            210.53
## 679                87.85  34    51816.27            153.01
## 683                43.57  36    50971.73            125.20
## 685                42.06  34    43241.19            131.55
## 688                73.27  28    67301.39            216.24
## 693                66.26  47    14548.06            179.04
## 695                68.25  33    76480.16            198.86
## 696                73.49  38    67132.46            244.23
## 703                87.27  30    51824.01            204.27
## 705                76.02  40    73174.19            219.55
## 707                71.33  23    31072.44            169.40
## 711                45.44  43    48453.55            119.27
## 712                69.96  31    73413.87            214.06
## 722                43.07  36    60583.02            137.63
## 731                65.53  19    51049.47            190.17
## 732                61.87  35    66629.61            250.20
## 734                44.11  41    43111.41            121.24
## 738                71.23  52    41521.28            122.59
## 743                71.90  29    72203.96            193.29
## 750                54.35  42    76984.21            164.02
## 755                76.87  36    72948.76            212.59
## 756                65.07  34    73941.91            227.53
## 758                48.86  35    62463.70            128.37
## 761                75.81  40    71157.05            229.19
## 762                72.94  31    74159.69            190.84
## 774                66.79  60    60575.99            198.30
## 782                75.55  36    31998.72            123.71
## 789                85.62  29    57032.36            195.68
## 792                35.49  47    36884.23            170.04
## 793                56.56  26    68783.45            204.47
## 797                79.22  27    63429.18            198.79
## 801                56.64  29    55984.89            123.24
## 805                53.92  41    25739.09            125.46
## 808                53.22  44    44307.18            108.85
## 811                43.16  29    50666.50            143.04
## 814                74.18  28    69874.18            203.87
## 828                43.49  45    47968.32            124.67
## 832                42.83  34    54324.73            132.38
## 839                44.13  40    33601.84            128.48
## 842                61.88  45    38260.89            108.18
## 844                74.61  38    71055.22            231.28
## 848                88.82  36    58638.75            169.10
## 850                59.05  52    50086.17            118.45
```

```
## 860                   83.17  43    54251.78               244.40
## 865                   74.06  50    60938.73               246.29
## 870                   82.41  36    65882.81               222.08
## 875                   77.35  34    72330.57               167.26
## 879                   81.75  43    52520.75               249.45
## 886                   46.84  45    34903.67               123.22
## 892                   54.39  20    52563.22               171.90
## 894                   69.86  25    50506.44               241.36
## 897                   78.84  32    62430.55               235.29
## 904                   84.76  30    61039.13               178.69
## 907                   37.01  50    48826.14               216.01
## 910                   75.24  24    52252.91               226.49
## 915                   41.86  39    53041.77               128.62
## 916                   34.04  34    40182.84               174.88
## 921                   81.05  34    54045.39               245.50
## 926                   35.98  45    43241.88               150.79
## 936                   83.98  36    68305.91               194.62
## 939                   32.60  38    40159.20               190.05
## 940                   60.83  19    40478.83               185.46
## 947                   72.82  34    75769.82               191.82
## 954                   45.53  29    56129.89               141.58
## 959                   79.83  29    55764.43               234.23
## 961                   56.78  28    39939.39               124.32
## 975                   41.88  40    44217.68               126.11
## 977                   61.84  45    46024.29               105.63
## 989                   89.71  48    51501.38               204.40
## 990                   70.96  31    55187.85               256.40
## 992                   38.96  38    36497.22               140.67
## 997                   51.30  45    67782.17               134.42
## 1000                  45.01  26    29875.80               178.35
##                                                 Ad.Topic.Line
## 20                                 Advanced 24/7 productivity
## 22                            Streamlined non-volatile analyzer
## 29                  Open-architected impactful productivity
## 35                      Synchronized dedicated service-desk
## 39                Persevering needs-based open architecture
## 58                                   Organized global model
## 78                     Grass-roots impactful system engine
## 87                           Total 5thgeneration encoding
## 97                         Versatile homogeneous capacity
## 98                       Function-based optimizing protocol
## 100                            Seamless holistic time-frame
## 101                         Persevering reciprocal firmware
## 102                      Centralized logistical secured line
## 105                  Polarized 6thgeneration info-mediaries
## 116                           Switchable mobile framework
## 126                User-friendly asymmetric info-mediaries
## 134            Profound zero administration instruction set
## 141                          Up-sized executive moderator
## 143                 Fully-configurable neutral open system
## 147                        Decentralized real-time circuit
## 151                            Polarized analyzing concept
## 155                         Automated coherent flexibility
## 160                    Managed grid-enabled standardization
```

```
## 163                         Horizontal multi-state interface
## 168                               Extended local methodology
## 178                            Realigned systematic function
## 179                   Adaptive context-sensitive application
## 183                         Polarized tangible collaboration
## 186                            Virtual scalable secured line
## 191                             Sharable grid-enabled matrix
## 194                 Cloned analyzing artificial intelligence
## 206              User-friendly client-server instruction set
## 209                            Assimilated fault-tolerant hub
## 213                  Configurable fault-tolerant monitoring
## 219         Configurable logistical Graphical User Interface
## 221                    Multi-tiered human-resource structure
## 223                         Cloned object-oriented benchmark
## 224                    Fundamental fault-tolerant neural-net
## 228                Pre-emptive cohesive budgetary management
## 235                             Versatile dedicated software
## 236                           Stand-alone reciprocal synergy
## 237                           Universal even-keeled analyzer
## 241                                 Automated static concept
## 254                         Synergistic value-added extranet
## 273                         Right-sized asynchronous website
## 277                         Triple-buffered regional toolset
## 284                          Customizable zero-defect matrix
## 288                    Front-line fresh-thinking open system
## 293                        Exclusive zero tolerance alliance
## 295                      Inverse next generation moratorium
## 298                   Grass-roots eco-centric instruction set
## 299  Fully-configurable incremental Graphical User Interface
## 301                         Mandatory 3rdgeneration moderator
## 305                    Vision-oriented optimizing middleware
## 306                      Proactive context-sensitive project
## 308                         Visionary multi-tasking alliance
## 310                               Extended interactive model
## 317                         Reactive bi-directional workforce
## 321                              Configurable dynamic adapter
## 327                         Exclusive even-keeled moratorium
## 328                         Reduced incremental productivity
## 330                Secured scalable Graphical User Interface
## 331             Team-oriented context-sensitive installation
## 334                               Profound maximized workforce
## 344                     Reactive needs-based instruction set
## 350                     Upgradable directional system engine
## 353           Triple-buffered needs-based Local Area Network
## 362                      Upgradable multi-tasking initiative
## 365                  Face-to-face mission-critical definition
## 366                   Inverse zero tolerance customer loyalty
## 384                         Enhanced asymmetric installation
## 386                          Robust responsive collaboration
## 397                                    Configurable 24/7 hub
## 398                        Versatile responsive knowledge user
## 402                         Mandatory dedicated data-warehouse
## 408                           Versatile content-based protocol
## 416             Enterprise-wide incremental Internet solution
```

```
## 426                            Total user-facing hierarchy
## 427           Balanced contextually-based pricing structure
## 431           Customer-focused system-worthy superstructure
## 434                    Visionary client-driven installation
## 437                        Configurable tertiary capability
## 443                  Configurable disintermediate throughput
## 447                               Optional secondary access
## 460                  Function-based directional productivity
## 461                                  Networked stable array
## 462                                Phased full-range hardware
## 466                  Self-enabling multimedia system engine
## 477                            Reactive interactive protocol
## 482                  Decentralized foreground infrastructure
## 489                   Fully-configurable holistic throughput
## 492                        Synergistic stable infrastructure
## 501                  Business-focused client-driven forecast
## 505                        Stand-alone motivating moratorium
## 507                               Upgradable local migration
## 523                          Compatible intermediate concept
## 524                     Assimilated next generation firmware
## 530                        Compatible dedicated productivity
## 542                    Organized upward-trending contingency
## 551                  Multi-layered fresh-thinking neural-net
## 556                     Synchronized grid-enabled moratorium
## 557                               Adaptive uniform capability
## 560                      Integrated client-server definition
## 563                         Managed 5thgeneration time-frame
## 569                         Virtual 5thgeneration neural-net
## 573                Public-key solution-oriented focus group
## 575                  Grass-roots mission-critical emulation
## 581                            Automated multi-state toolset
## 593                         Face-to-face executive encryption
## 594                         Monitored local Internet solution
## 596                       User-friendly grid-enabled analyzer
## 599         Customizable methodical Graphical User Interface
## 602                         Open-source even-keeled database
## 609                                  Digitized radical array
## 616                       Sharable multimedia conglomeration
## 626                    Re-engineered non-volatile neural-net
## 627                     Implemented disintermediate attitude
## 630                  Front-line non-volatile implementation
## 631                                  Ergonomic 24/7 solution
## 632             Integrated grid-enabled budgetary management
## 636                          Versatile optimizing projection
## 641              Triple-buffered high-level Internet solution
## 642       Synergized well-modulated Graphical User Interface
## 646              Seamless bandwidth-monitored knowledge user
## 650                            Object-based optimal solution
## 666                          Total asynchronous architecture
## 671                           Switchable multi-state success
## 673                  Fundamental clear-thinking knowledgebase
## 679                          Digitized interactive initiative
## 683                      Front-line upward-trending groupware
## 685                          Multi-layered secondary software
```

```
## 688                            Front-line multi-state hub
## 693                          Optional full-range projection
## 695                         Organic well-modulated database
## 696                        Organic 3rdgeneration encryption
## 703                  Fully-configurable 5thgeneration circuit
## 705                     Distributed leadingedge orchestration
## 707                   Optimized attitude-oriented initiative
## 711                    Enhanced intermediate standardization
## 712                         Realigned tangible collaboration
## 722               Monitored zero administration collaboration
## 731                    Pre-emptive client-server installation
## 732                 Multi-channeled attitude-oriented toolset
## 734                          Organic next generation matrix
## 738              Self-enabling zero administration neural-net
## 743                   De-engineered intangible flexibility
## 750                     Proactive interactive service-desk
## 755                    Down-sized bandwidth-monitored core
## 756                           Monitored explicit hierarchy
## 758                           Universal empowering adapter
## 761                         Inverse zero-defect capability
## 762                    Multi-tiered real-time implementation
## 774                Decentralized client-driven data-warehouse
## 782                Decentralized attitude-oriented interface
## 789                Triple-buffered systematic info-mediaries
## 792                     Front-line actuating functionalities
## 793                           Compatible composite project
## 797                        Intuitive modular system engine
## 801                     Versatile 6thgeneration parallelism
## 805                  Business-focused transitional solution
## 808                       Horizontal client-server database
## 811                             Robust holistic application
## 814                         Front-line even-keeled website
## 828                         Synchronized full-range portal
## 832                       Vision-oriented methodical support
## 839                            Ameliorated discrete extranet
## 842                  Balanced disintermediate conglomeration
## 844                            Networked impactful framework
## 848                            Assimilated discrete strategy
## 850                       Upgradable logistical flexibility
## 860                               Managed national hardware
## 865                       Horizontal intermediate monitoring
## 870                      Stand-alone logistical service-desk
## 875                             Total bi-directional success
## 879                              Progressive 24/7 definition
## 886                  Virtual homogeneous budgetary management
## 892                           Digitized content-based circuit
## 894                     Triple-buffered foreground encryption
## 897                        Optimized 5thgeneration moratorium
## 904                                Monitored 24/7 moratorium
## 907              Secured encompassing Graphical User Interface
## 910                   Vision-oriented next generation solution
## 915                              Open-source stable paradigm
## 916             Reverse-engineered context-sensitive emulation
## 921                       Configurable asynchronous application
```

```
## 926                  Front-line fresh-thinking installation
## 936                        Profit-focused secondary portal
## 939                    Customizable homogeneous contingency
## 940              Versatile next generation pricing structure
## 947                           User-centric discrete success
## 954                             Front-line neutral alliance
## 959                           Innovative regional groupware
## 961                        Mandatory empowering focus group
## 975                             Streamlined exuding adapter
## 977                  Reactive bi-directional standardization
## 989                               Versatile local forecast
## 990                       Ameliorated user-facing help-desk
## 992                  Versatile mission-critical application
## 997                         Grass-roots cohesive monitoring
## 1000                        Virtual 5thgeneration emulation
##                      City Male                                        Country
## 20              Millertown    1                         Russian Federation
## 22             Lake Nicole    1                                   Cameroon
## 29              Palmerside    0                British Virgin Islands
## 35              New Thomas    1                                   Dominica
## 39            Charlesport    1                                Saint Helena
## 58         Port Sarahshire    0             Svalbard & Jan Mayen Islands
## 78              Cranemouth    1                                      Qatar
## 87          North Kristine    0                Central African Republic
## 97            Williammouth    1                                   Hong Kong
## 98         Williamsborough    0                                   Lithuania
## 100        Benjaminchester    1                                  Bangladesh
## 101         Hernandezville    0                              Western Sahara
## 102             Youngburgh    1                                      Serbia
## 105        Bradshawborough    0                                   Guernsey
## 116        West Angelabury    1                                      Jersey
## 126               Lisaberg    0                                       Egypt
## 134           Port Jeffrey    0    United States Minor Outlying Islands
## 141          Lake Josetown    0                                   Gibraltar
## 143          New Debbiestad    1                                       Korea
## 147            West Ricardo    1                                 Netherlands
## 151             Tinachester    1                                 New Zealand
## 155       East Timothyport    1                                 Timor-Leste
## 160            Ramirezside    0                               Liechtenstein
## 163      West Michaelhaven    0                                      Angola
## 168             Timothytown    1                                 Timor-Leste
## 178          Hernandezfort    0                                      Rwanda
## 179               Joneston    1                                       China
## 183              Youngfort    0                                        Peru
## 186       Port Melissaberg    0                                   Singapore
## 191          West Ericfurt    0                                        Fiji
## 194       South Tiffanyton    1                 Bouvet Island (Bouvetoya)
## 206  North Russellborough    0                                South Africa
## 209               Penatown    0                                   Australia
## 213           Mcdonaldfort    1                             Kyrgyz Republic
## 219          Davilachester    0                             Czech Republic
## 221            Melissafurt    0                                      Poland
## 223            Millerbury    0                                Turkmenistan
## 224            Garciaview    0                                       Latvia
```

```
## 228           West Justin    0                                  Bahrain
## 235             New Nancy    1                                    Chad
## 236             Lisamouth    1                                   Norway
## 237            Harveyport    0                Turks and Caicos Islands
## 241         Christinetown    0                              Afghanistan
## 254             Brownport    1                                  Finland
## 273       Port Aliciabury    0                               Guadeloupe
## 277       North Loriburgh    0                                  Estonia
## 284         South Tomside    0             Saint Pierre and Miquelon
## 288           West Brenda    0                                    Tonga
## 293           Jordantown     1             United States of America
## 295           South Troy     1                                 Suriname
## 298     Port Jessicamouth    0                                    Chile
## 299              Paulport    0                                    Gabon
## 301           Cynthiaside    0      Holy See (Vatican City State)
## 305          Jessicahaven    0                                 Cambodia
## 306          North Daniel    1                      Antigua and Barbuda
## 308               Amyfurt    0                                  Somalia
## 310           Roberttown     0             Saint Pierre and Miquelon
## 317           Hayesmouth     0                                  Cyprus
## 321      East Vincentstad    0                                  Zambia
## 327          Jeffreyburgh    1                            South Africa
## 328             Faithview    0                               Singapore
## 330       Port Brookeland    0                              Martinique
## 331  East Christopherbury    0                                Cameroon
## 334              Hessstad    1              Bosnia and Herzegovina
## 344           Garychester    1                                   Haiti
## 350     Port Kathleenfort    0                                   Nauru
## 353             Mollyport    0              Libyan Arab Jamahiriya
## 362      North Andrewstad    1                                  Angola
## 365             Novaktown    1                           Faroe Islands
## 366          Timothymouth    1                          Czech Republic
## 384        New Traceystad    1                      Dominican Republic
## 386          Lake Courtney    0                                 Croatia
## 397      West Eduardotown     1                                  Canada
## 398             Davisfurt    0                  Saint Kitts and Nevis
## 402           Lake Brian     0                              Madagascar
## 408           Monicaview     0                              Bangladesh
## 416        New Denisebury    1                                 Myanmar
## 426  South Vincentchester    1                          United Kingdom
## 427         Williamsmouth    1                              Luxembourg
## 431             East John    1                                  Zambia
## 434            Paulhaven     1                                 Senegal
## 437       North Tylerland    1                                 Bahrain
## 443       Port Patrickton    1                                 Estonia
## 447    North Aaronchester    1                             Philippines
## 460              Jayville    1                              Micronesia
## 461          East Lindsey    1                                   Malta
## 462          Masseyshire     0                                 Ecuador
## 466      New Williammouth    1                             Switzerland
## 477             Joanntown    1                                 Austria
## 482             Alanview     1                                   Tonga
## 489          Timothyport     0                          American Samoa
## 492              Silvaton    0                                  Greece
```

```
## 501          Helenborough    0                          Australia
## 505   South Alexisborough    0                      Cote d'Ivoire
## 507             Hallfort    0                            Jamaica
## 523         East Timothy    1                          Sri Lanka
## 524       North Kimberly    0                            Morocco
## 530          Port Stacey    1                          Australia
## 542          Brianabury    1                            Bahrain
## 551     East Michaelland    0                             Belize
## 556    South Patrickfort    0                  Papua New Guinea
## 557      East Georgeside    1                            Bahamas
## 560          South Renee    1                          Nicaragua
## 563          Amandaland    1                            Bolivia
## 569          Edwardmouth    1                         Madagascar
## 573          Austinland    0                        Puerto Rico
## 575          Wrightview    0          Antigua and Barbuda
## 581      Richardsonshire    0                             Cyprus
## 593         Lake Patrick    1          Libyan Arab Jamahiriya
## 594      Richardsonmouth    0                          Sri Lanka
## 596            Boyerberg    1                          Indonesia
## 599          New Dawnland    0                           Honduras
## 602    West Raymondmouth    1                           Ethiopia
## 609     East Rachaelfurt    1                             Jersey
## 616        East Brettton    0                            Ecuador
## 626         Hubbardmouth    1                             Latvia
## 627          South Brian    1                            Morocco
## 630             Lukeport    1                        Isle of Man
## 631            New Shane    1                               Peru
## 632       Lake Jillville    1                            Belgium
## 636     North Angelastad    0                               Peru
## 641         Isaacborough    1                         Azerbaijan
## 642         Lake Michael    0                           Mongolia
## 646            Markhaven    1                           Ethiopia
## 650       West Dannyberg    0                           Djibouti
## 666          Sanderstown    1                         Tajikistan
## 671          West Sharon    0                          Singapore
## 673         Lake Cynthia    0                            Reunion
## 679         Meyerchester    0                       Turkmenistan
## 683      Lake Joshuafurt    1                    French Polynesia
## 685           Jamesmouth    0                        Puerto Rico
## 688          West Angela    1                         Luxembourg
## 693          Matthewtown    1                            Lebanon
## 695       New Jamestown    1                            Vanuatu
## 696          Mosleyburgh    0                            Vanuatu
## 703    West Melissashire    1                  Christmas Island
## 705          Lesliefort    0                              Malta
## 707          Josephmouth    0                              Japan
## 711              Mezaton    0                              China
## 712           New Kayla    1                        Netherlands
## 722       East Ericport    1                             Turkey
## 731          Coffeytown    1                         Timor-Leste
## 732       North Johnside    1      French Southern Territories
## 734          Lake David    1  Saint Vincent and the Grenadines
## 738       Ericksonmouth    0                             Sweden
## 743         Williamsside    1                         Costa Rica
```

```
## 750           Lake Hailey    0                                      Sweden
## 755            Mullenside    0 South Georgia and the South Sandwich Islands
## 756            Princebury    1                                    Anguilla
## 758         Elizabethbury    1                                 Saint Martin
## 761           Sanchezland    0                                   Greenland
## 762             Rogerland    0                                   Venezuela
## 774            Shelbyport    1                                Burkina Faso
## 782      South Davidmouth    1                  Bosnia and Herzegovina
## 789      South Dianeshire    0                                        Iran
## 792   North Brittanyburgh    0                               Liechtenstein
## 793          Port Jasmine    1                                     Grenada
## 797            North Debra    1                                     Belgium
## 801              Lake John    1                                    Zimbabwe
## 805            Lake Amanda    1                                      Turkey
## 808       Port Aprilville    0          United States Virgin Islands
## 811            Wendyville    1       Saint Vincent and the Grenadines
## 814              Lake Lisa    0                              Cayman Islands
## 828              Smithside    0                                     Vietnam
## 832       Cunninghamhaven    1                                      Latvia
## 839       West Derekmouth    1                                        Cuba
## 842      North Debrashire    0                                 Afghanistan
## 844      East Christopher    1                                South Africa
## 848         East Johnport    0            Slovakia (Slovak Republic)
## 850           Kristinfurt    1                                      Uganda
## 860     Port Kevinborough    1                                   Sri Lanka
## 865   Port Jacquelinestad    1                                    Anguilla
## 870          North Brandon    0                                   Nicaragua
## 875        East Troyhaven    0                                       Yemen
## 879             Lake Jose    0                                    Barbados
## 886            Jonesshire    0                                       Macao
## 892          Richardshire    1                                  Guadeloupe
## 894         Austinborough    0                                 New Zealand
## 897         Gonzalezburgh    1                                      Taiwan
## 904           South Pamela    0                                   Venezuela
## 907        Port Derekberg    0                                  San Marino
## 910     South Christopher    0                                     Vanuatu
## 915           Hawkinsbury    1                                    Ethiopia
## 916         Elizabethport    1                                 El Salvador
## 921              Adamsstad    0                                    Malaysia
## 926        Huffmanchester    0                                    Ethiopia
## 936        West Gregburgh    0                                    Zimbabwe
## 939             Tylerport    0                  Syrian Arab Republic
## 940            West Lacey    1                 Palestinian Territory
## 947            Jonesmouth    1                        Czech Republic
## 954         North Destiny    0                  United Arab Emirates
## 959       North Frankstad    1                                   Sudan
## 961            Melanieton    0                                 Senegal
## 975           Port Rachel    1                                  Cyprus
## 977       Port Joshuafort    1                                   Niger
## 989              Sarafurt    1                                 Bahrain
## 990          South Denise    0           Saint Pierre and Miquelon
## 992           Mauricefurt    1                                 Comoros
## 997           New Darlene    1                  Bosnia and Herzegovina
## 1000          Ronniemouth    0                                  Brazil
```

```
##                Timestamp
## 20   2016-02-27 04:43:07
## 22   2016-03-18 13:22:35
## 29   2016-04-14 05:08:35
## 35   2016-03-15 03:12:25
## 39   2016-05-07 17:11:49
## 58   2016-04-10 00:13:47
## 78   2016-05-20 21:31:24
## 87   2016-07-01 04:41:57
## 97   2016-04-19 15:14:58
## 98   2016-01-08 22:47:10
## 100  2016-07-02 14:57:53
## 101  2016-07-03 09:22:30
## 102  2016-06-01 09:27:34
## 105  2016-06-10 11:31:33
## 116  2016-06-22 07:33:21
## 126  2016-05-17 04:27:31
## 134  2016-01-14 00:23:10
## 141  2016-02-05 19:06:01
## 143  2016-06-14 11:59:58
## 147  2016-02-07 08:02:31
## 151  2016-02-10 06:37:56
## 155  2016-06-10 10:11:00
## 160  2016-05-09 02:58:58
## 163  2016-02-29 12:31:57
## 168  2016-04-28 18:34:56
## 178  2016-02-10 06:52:07
## 179  2016-03-27 09:11:10
## 183  2016-05-24 13:30:38
## 186  2016-02-04 08:53:37
## 191  2016-01-03 05:34:33
## 194  2016-07-21 21:16:35
## 206  2016-02-11 02:40:02
## 209  2016-06-16 02:01:24
## 213  2016-05-29 21:17:10
## 219  2016-06-12 17:52:43
## 221  2016-07-02 00:24:22
## 223  2016-03-24 09:12:52
## 224  2016-02-14 07:30:24
## 228  2016-06-18 17:56:32
## 235  2016-03-30 23:40:52
## 236  2016-03-16 07:59:37
## 237  2016-05-04 00:01:33
## 241  2016-07-23 14:47:23
## 254  2016-07-21 10:54:35
## 273  2016-02-15 07:55:10
## 277  2016-01-09 07:28:16
## 284  2016-05-26 13:18:30
## 288  2016-06-28 20:13:41
## 293  2016-05-30 07:36:31
## 295  2016-07-20 23:08:28
## 298  2016-04-16 16:38:35
## 299  2016-05-03 08:21:23
## 301  2016-05-22 00:01:58
```

```
## 305  2016-05-09 21:54:38
## 306  2016-05-31 11:44:45
## 308  2016-01-09 15:49:28
## 310  2016-06-13 13:59:51
## 317  2016-02-07 07:41:06
## 321  2016-03-15 11:25:48
## 327  2016-03-24 02:01:55
## 328  2016-03-03 22:31:16
## 330  2016-07-06 15:56:39
## 331  2016-06-24 05:50:22
## 334  2016-04-28 22:54:37
## 344  2016-06-16 02:33:22
## 350  2016-06-09 14:24:06
## 353  2016-01-31 06:14:10
## 362  2016-01-23 18:59:21
## 365  2016-04-21 12:34:28
## 366  2016-04-22 20:32:17
## 384  2016-07-05 20:16:13
## 386  2016-07-21 23:14:35
## 397  2016-06-20 14:20:52
## 398  2016-04-03 06:17:22
## 402  2016-02-26 17:01:01
## 408  2016-07-09 16:23:33
## 416  2016-07-05 00:54:11
## 426  2016-04-15 10:16:49
## 427  2016-05-31 09:06:29
## 431  2016-01-03 17:10:05
## 434  2016-02-27 12:34:19
## 437  2016-03-23 21:06:51
## 443  2016-05-31 06:21:02
## 447  2016-05-30 18:08:19
## 460  2016-07-17 22:04:54
## 461  2016-06-02 22:16:08
## 462  2016-04-30 19:42:04
## 466  2016-05-26 10:33:00
## 477  2016-01-08 08:08:47
## 482  2016-01-02 14:36:03
## 489  2016-04-08 14:35:44
## 492  2016-05-12 21:32:06
## 501  2016-01-07 13:58:51
## 505  2016-02-15 07:27:41
## 507  2016-03-20 22:27:25
## 523  2016-03-30 20:23:48
## 524  2016-05-01 00:23:13
## 530  2016-04-16 12:26:31
## 542  2016-05-30 02:34:25
## 551  2016-01-21 04:30:43
## 556  2016-07-21 10:01:50
## 557  2016-04-21 18:31:27
## 560  2016-01-16 17:56:05
## 563  2016-05-27 18:45:35
## 569  2016-04-29 07:49:01
## 573  2016-01-03 23:21:26
## 575  2016-03-09 02:07:17
```

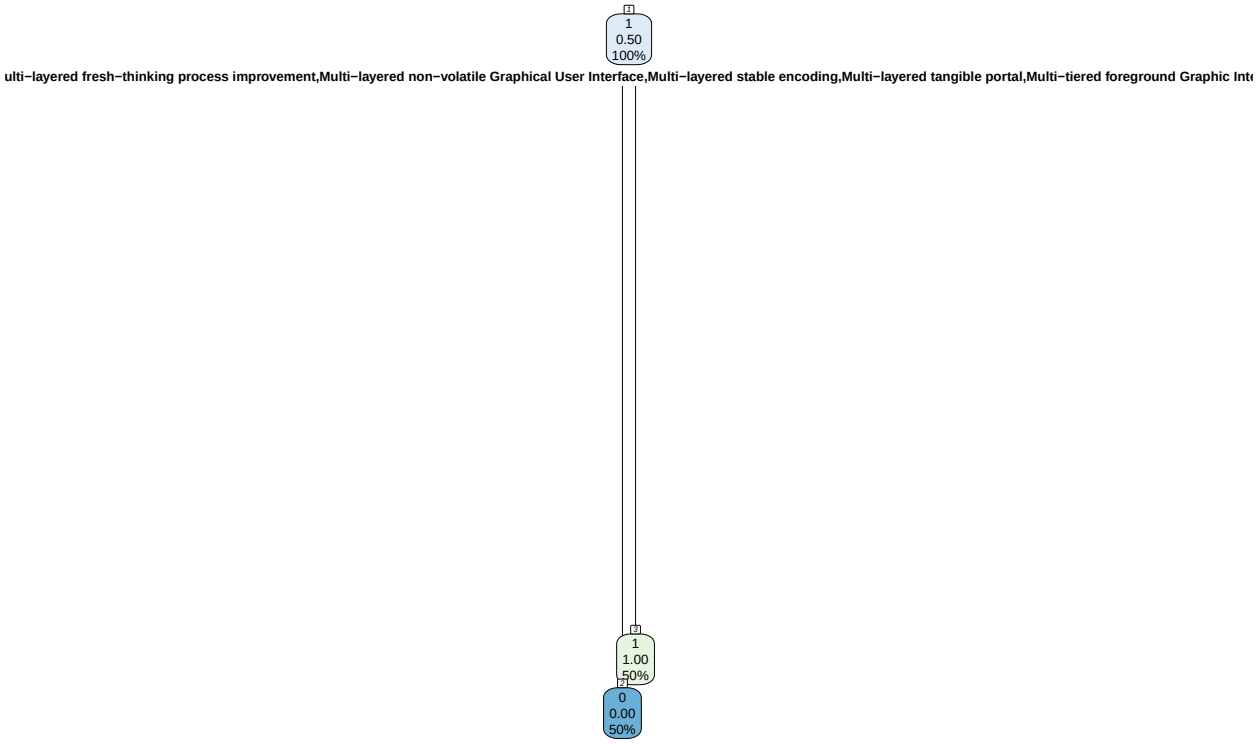
```
## 581  2016-07-24 00:22:16
## 593  2016-01-29 07:14:04
## 594  2016-06-14 07:02:09
## 596  2016-01-30 09:54:03
## 599  2016-07-06 05:34:52
## 602  2016-01-28 17:03:54
## 609  2016-03-16 20:10:53
## 616  2016-03-01 22:06:37
## 626  2016-01-12 03:28:31
## 627  2016-04-09 23:26:42
## 630  2016-01-24 01:53:14
## 631  2016-04-15 10:18:55
## 632  2016-04-26 13:13:20
## 636  2016-07-18 04:53:22
## 641  2016-06-12 03:11:04
## 642  2016-02-15 20:41:05
## 646  2016-04-04 13:56:14
## 650  2016-01-06 21:43:22
## 666  2016-02-12 10:39:10
## 671  2016-05-15 03:10:50
## 673  2016-07-19 12:05:58
## 679  2016-03-19 08:00:58
## 683  2016-06-25 17:33:35
## 685  2016-06-29 02:48:44
## 688  2016-05-22 15:17:25
## 693  2016-04-25 19:31:39
## 695  2016-05-30 08:02:35
## 696  2016-02-14 11:36:08
## 703  2016-03-06 06:51:23
## 705  2016-07-13 14:30:14
## 707  2016-03-15 06:54:21
## 711  2016-05-04 05:01:37
## 712  2016-05-20 12:17:28
## 722  2016-02-11 20:45:46
## 731  2016-06-20 02:25:12
## 732  2016-07-10 19:15:52
## 734  2016-04-20 16:49:15
## 738  2016-01-05 11:53:17
## 743  2016-04-24 07:20:16
## 750  2016-06-20 08:22:09
## 755  2016-07-08 22:30:10
## 756  2016-03-11 00:05:48
## 758  2016-01-04 00:44:57
## 761  2016-03-27 19:50:11
## 762  2016-04-29 13:38:19
## 774  2016-04-07 20:34:42
## 782  2016-03-12 06:05:12
## 789  2016-04-14 21:37:49
## 792  2016-04-13 07:07:36
## 793  2016-02-03 22:11:13
## 797  2016-03-11 12:39:19
## 801  2016-02-27 08:52:50
## 805  2016-06-12 11:17:25
## 808  2016-05-26 06:03:57
```

```
## 811  2016-03-05 20:53:19
## 814  2016-01-20 02:31:36
## 828  2016-06-14 12:08:10
## 832  2016-04-23 09:42:08
## 839  2016-03-03 20:20:32
## 842  2016-05-30 08:02:27
## 844  2016-04-06 23:10:40
## 848  2016-01-30 00:05:37
## 850  2016-04-23 03:46:34
## 860  2016-04-05 18:02:49
## 865  2016-06-17 09:38:22
## 870  2016-04-12 04:22:42
## 875  2016-03-09 12:10:08
## 879  2016-06-07 05:41:16
## 886  2016-05-15 18:44:50
## 892  2016-02-29 18:06:21
## 894  2016-01-12 21:17:15
## 897  2016-04-09 09:26:39
## 904  2016-05-14 23:08:14
## 907  2016-03-20 02:44:13
## 910  2016-02-25 16:33:24
## 915  2016-03-26 15:28:07
## 916  2016-02-16 07:37:28
## 921  2016-05-18 00:07:43
## 926  2016-06-27 21:51:47
## 936  2016-02-24 00:44:44
## 939  2016-02-12 03:39:09
## 940  2016-02-19 20:49:27
## 947  2016-01-27 16:06:05
## 954  2016-04-22 22:01:21
## 959  2016-05-18 19:33:51
## 961  2016-06-19 23:04:45
## 975  2016-02-28 23:54:44
## 977  2016-05-05 11:07:13
## 989  2016-02-17 07:00:38
## 990  2016-06-26 07:01:47
## 992  2016-07-21 16:02:40
## 997  2016-04-22 02:07:01
## 1000 2016-06-03 21:43:21
```

```
# penalty matrix
penalty.matrix <- matrix(c(0,1,10,0), byrow=TRUE, nrow=2)
```

```
# building the classification tree with rpart
library(rpart)
library(rpart.plot)
tree <- rpart(Clicked.on.Ad~.,
data=adsop_train,
parms = list(loss = penalty.matrix),
method = "class")
```

```
# Visualize the decision tree with rpart.plot
rpart.plot(tree, nn=TRUE)
```



#7. Challenging the solution

```
adsop_train[["Clicked.on.Ad"]] = factor(adsop_train[["Clicked.on.Ad"]])
```

```
#library(caret)
#trctrl <- trainControl(method = "repeatedcv", number = 10, repeats = 3)

#svm_Linear <- train(Clicked.on.Ad ~., data = adsop_train, method = "svmLinear",
#trControl=trctrl,
#preProcess = c("center", "scale"),
#tuneLength = 10)
```

#8. Conclusion - Average of 60 min is spent on the site per day.

#9. Recommendations - Use a computer with high computational power to run this model.

#10. Follow up questions

#a) Did we have right data? Yes. #b) Do we need other data to answer our question? No. #c) Did we have the right question? Yes.