Assignment2_DecisionTree

Mukhtar

3/28/2022

```
library(ISLR)
library(rpart)
library(rattle)
```

```
## Warning: package 'rattle' was built under R version 4.1.2

## Loading required package: tibble

## Loading required package: bitops

## Rattle: A free graphical interface for data science with R.
## Version 5.4.0 Copyright (c) 2006-2020 Togaware Pty Ltd.
## Type 'rattle()' to shake, rattle, and roll your data.
```

```
library(rpart.plot)
```

```
## Warning: package 'rpart.plot' was built under R version 4.1.2
```

```
library(caret)
```

```
## Loading required package: ggplot2

## Loading required package: lattice
```

```
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
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

Problem Statement

Data Preparation

```
getwd()
```

```
## [1] "C:/Users/Mukht/OneDrive/Desktop/Kent State University/54050-Project-ADM/Assignment2"
```

```
setwd("C:\\Users\\Mukht\\OneDrive\\Desktop\\Kent State University\\54050-Project-ADM\\Assignment2")
```

```
ADM_Assignment2<-read.csv("carseats_ADM.csv")
str(ADM_Assignment2)
```

```
## 'data.frame':    400 obs. of  11 variables:
##  $ ï..Sales   : num  9.5 11.22 10.06 7.4 4.15 ...
##  $ Income     : int  73 48 35 100 64 113 105 81 110 113 ...
##  $ Advertising: int  11 16 10 4 3 13 0 15 0 0 ...
##  $ Population : int  276 260 269 466 340 501 45 425 108 131 ...
##  $ Price      : int  120 83 80 97 128 72 108 120 124 124 ...
##  $ Age        : int  42 65 59 55 38 78 71 67 76 76 ...
##  $ Education  : int  17 10 12 14 13 16 15 10 10 17 ...
##  $ Urban      : chr  "Yes" "Yes" "Yes" "Yes" ...
##  $ US         : chr  "Yes" "Yes" "Yes" "Yes" ...
##  $ CompPrice  : int  138 111 113 117 141 124 115 136 132 132 ...
##  $ ShelveLoc  : chr  "Bad" "Good" "Medium" "Medium" ...
```

```
head(ADM_Assignment2)
```

```
##   ï..Sales Income Advertising Population Price Age Education Urban  US
## 1     9.50     73          11        276   120  42        17   Yes Yes
## 2    11.22     48          16        260    83  65        10   Yes Yes
## 3    10.06     35          10        269    80  59        12   Yes Yes
## 4     7.40    100           4        466    97  55        14   Yes Yes
## 5     4.15     64           3        340   128  38        13   Yes  No
## 6    10.81    113          13        501    72  78        16    No Yes
##   CompPrice ShelveLoc
## 1       138       Bad
## 2       111      Good
## 3       113    Medium
## 4       117    Medium
## 5       141       Bad
## 6       124       Bad
```

#Three of the variables are factors, while the rest are numeric. Currently there are no missing observations.

```
Carseats_Filtered <- ADM_Assignment2 %>% select("ï..Sales", "Price", "Advertising","Population","Age","I
Carseats_Filtered
```

```
##     ï..Sales Price Advertising Population Age Income Education
## 1       9.50   120          11        276  42     73        17
## 2      11.22    83          16        260  65     48        10
## 3      10.06    80          10        269  59     35        12
## 4       7.40    97           4        466  55    100        14
## 5       4.15   128           3        340  38     64        13
## 6      10.81    72          13        501  78    113        16
## 7       6.63   108           0         45  71    105        15
## 8      11.85   120          15        425  67     81        10
## 9       6.54   124           0        108  76    110        10
## 10      4.69   124           0        131  76    113        17
## 11      9.01   100           9        150  26     78        10
## 12     11.96    94           4        503  50     94        13
## 13      3.98   136           2        393  62     35        18
## 14     10.96    86          11         29  53     28        18
## 15     11.17   118          11        148  52    117        18
## 16      8.71   144           5        400  76     95        18
## 17      7.58   110           0        284  63     32        13
## 18     12.29   131          13        251  52     74        10
## 19     13.91    68           0        408  46    110        17
## 20      8.73   121          16         58  69     76        12
## 21      6.41   131           2        367  35     90        18
## 22     12.13   109          12        239  62     29        18
## 23      5.08   138           6        497  42     46        13
## 24      5.87   109           0        292  79     31        10
## 25     10.14   113          16        294  42    119        12
## 26     14.90    82           0        176  54     32        11
## 27      8.33   131          11        496  50    115        11
## 28      5.27   107           0         19  64    118        17
## 29      2.99    97           0        359  55     74        11
## 30      7.81   102          15        226  58     99        17
## 31     13.55    89           0        447  30     94        12
## 32      8.25   131          16        241  44     58        18
## 33      6.20   137          12        236  64     32        10
## 34      8.77   128          13        317  50     38        16
## 35      2.67   128           0        406  42     54        17
## 36     11.07    96          11         29  44     84        17
## 37      8.89   100           0        270  60     76        18
## 38      4.95   110           5        412  54     41        10
## 39      6.59   102           0        454  65     73        15
## 40      3.24   138           0        144  38     60        10
## 41      2.07   126           0         18  73     98        17
## 42      7.96   124           0        403  58     53        16
## 43     10.43    24           0         25  50     69        18
## 44      4.12   134          11         16  59     42        13
## 45      4.16    95           6        325  69     79        13
## 46      4.56   135           0        168  44     63        12
## 47     12.44    70          14         16  48     90        15
## 48      4.38   108           0        173  55     98        16
```

```
## 49       3.91    98              0         349  69      52          18
## 50      10.61   149              0          51  32      93          17
## 51       1.42   108             18         341  80      32          16
## 52       4.42   108              0         150  75      90          16
## 53       7.91   129              3         112  39      40          18
## 54       6.92   119             13          39  61      64          17
## 55       4.90   144             13          25  76     103          17
## 56       6.85   154              5          60  61      81          18
## 57      11.91    84              0          54  50      82          17
## 58       0.91   117              0          22  75      91          11
## 59       5.42   103             15         188  74      93          16
## 60       5.21   114              4         148  80      71          13
## 61       8.32   123             19         469  29     102          13
## 62       7.32   107              0         358  26      32          13
## 63       1.82   133              0         146  77      45          17
## 64       8.47   101             10         170  61      88          13
## 65       7.80   104             12         184  32      67          16
## 66       4.90   128              0         197  55      26          13
## 67       8.85    91              0         508  56      92          18
## 68       9.01   115             14         152  47      61          16
## 69      13.39   134             20         366  60      69          13
## 70       7.99    99              0         339  65      59          12
## 71       9.46    99             15         237  74      81          12
## 72       6.50   150             16         148  58      51          17
## 73       5.52   116              0         432  25      45          15
## 74      12.61   104             10          54  31      90          11
## 75       6.20   136              5         125  64      68          13
## 76       8.55    92             23         480  36     111          16
## 77      10.64    70             10         346  64      87          15
## 78       7.70    89             12          44  67      71          18
## 79       4.43   145              1         139  65      48          12
## 80       9.14    90              0         286  41      67          13
## 81       8.01    79             16         353  68     100          11
## 82       7.52   128              0         237  70      72          13
## 83      11.62   139              4         325  28      83          17
## 84       4.42    94              7         468  56      36          11
## 85       2.23   121              0          52  43      25          18
## 86       8.47   112              0         304  49     103          13
## 87       8.70   134              9         432  64      84          15
## 88      11.70   126              7         272  54      67          16
## 89       6.56   111              7         144  62      42          10
## 90       7.95   119              3         493  45      66          16
## 91       5.33   103              0         491  64      22          11
## 92       4.81   107             11         267  80      46          15
## 93       4.53   125              0          97  29     113          12
## 94       8.86   104              0          67  55      30          17
## 95       8.39    84              5         134  55      97          11
## 96       5.58   148             10         237  59      25          13
## 97       9.48   132             10         407  73      42          16
## 98       7.45   129              5         287  33      82          16
## 99      12.49   127             24         382  36      77          16
## 100      4.88   107              3         220  56      47          16
## 101      4.11   106             11          94  76      69          12
## 102      6.20   118              0          89  34      93          18
```

```
## 103      5.30    97             0          57  65      22          16
## 104      5.07    96             0         334  78      91          17
## 105      4.62   138             0         472  51      96          12
## 106      5.55    97             8         398  61     100          11
## 107      0.16   139             0         217  70      33          18
## 108      8.55   108             0         104  60     107          12
## 109      3.47   103             2         488  65      79          16
## 110      8.98    90             0         217  60      65          17
## 111      9.00   116             7         125  43      62          14
## 112      6.62   151            12         272  43     118          14
## 113      6.67   125             5         298  62      99          12
## 114      6.01   127            11         335  33      29          12
## 115      9.31   106             9          17  65      87          13
## 116      8.54   129             0          95  42      35          13
## 117      5.08   128             0         202  80      75          10
## 118      8.80   119             0         507  41      53          12
## 119      7.57    99             2         243  62      88          11
## 120      7.37   128             8         137  64      94          12
## 121      6.87   131            11         249  63     105          13
## 122     11.67    87            10         380  28      89          10
## 123      6.88   108             5          45  75     100          10
## 124      8.19   155             0         125  29     103          15
## 125      8.87   120             0         181  63     113          14
## 126      9.34    49             0         181  43      78          15
## 127     11.27   133             2          60  59      68          16
## 128      6.52   116             3         192  51      48          14
## 129      4.96   126             3         350  55     100          13
## 130      4.47   147             7         279  40     120          10
## 131      8.41    77            13         497  51      84          12
## 132      6.50    94             3         208  77      69          16
## 133      9.54   136             9         232  72      87          10
## 134      7.62    97             2         265  62      98          12
## 135      3.67   131             0         327  76      31          16
## 136      6.44   120            14         384  36      94          18
## 137      5.17   120             0          10  31      75          18
## 138      6.52   118             0         436  80      42          11
## 139     10.27   109            12         371  44     103          10
## 140     12.30    94            10         310  30      62          13
## 141      6.03   129            10         277  45      60          18
## 142      6.53   131             0         331  28      42          15
## 143      7.44   104             0         300  77      84          15
## 144      0.53   159             7          36  28      88          17
## 145      9.09   123             0         264  34      68          11
## 146      8.77   117            11          27  47      63          17
## 147      3.90   131             0         412  39      83          14
## 148     10.51   119             9         402  41      54          16
## 149      7.56    97             0         384  72     119          14
## 150     11.48    87            13         140  56     120          11
## 151     10.49   114             8         176  57      84          10
## 152     10.77   103            17         407  75      58          17
## 153      7.64   128             0         341  45      78          13
## 154      5.93   150             7         488  25      36          17
## 155      6.89   110            10         289  50      69          16
## 156      7.71    69             0          59  65      72          16
```

```
## 157      7.49   157            0        220  51      34          16
## 158     10.21    90            8        249  48      58          13
## 159     12.53   112            1        189  39      90          10
## 160      9.32    70            0        372  30      60          18
## 161      4.67   111            0        486  29      28          12
## 162      2.93   160            5         81  67      21          12
## 163      3.63   149            0        424  51      74          13
## 164      5.68   106            0         40  39      64          17
## 165      8.22   141            0         58  27      64          13
## 166      0.37   191            7        100  27      58          15
## 167      6.71   137           17        151  55      67          11
## 168      6.71    93            0        216  60      73          13
## 169      7.30   117            0        425  45      89          10
## 170     11.48    77           15        492  73      41          18
## 171      8.01   118           12        356  71      39          10
## 172     12.49    55           12        416  75     106          15
## 173      9.03   110           13        123  35     102          16
## 174      6.38   128            5        207  66      91          18
## 175      0.00   185            0        358  79      24          15
## 176      7.54   122            0         38  25      89          12
## 177      5.61   154            9        480  47     107          11
## 178     10.48    94            0        148  27      72          17
## 179     10.66    81           14         89  25      71          14
## 180      7.78   116            3         70  77      25          18
## 181      4.94   149           15        434  66     112          13
## 182      7.43    91            0         79  68      83          11
## 183      4.74   140            4        230  25      60          13
## 184      5.32   102            6        426  80      74          18
## 185      9.95    97            7         35  60      33          11
## 186     10.07   107           11        449  64     100          10
## 187      8.68    86            0         93  46      51          17
## 188      6.03    96            0        142  62      32          17
## 189      8.07    90            0        426  76      37          15
## 190     12.11   104           18        509  26     117          15
## 191      8.79   101           13        297  37      37          13
## 192      6.67   173           13        170  74      42          14
## 193      7.56    93            0        408  56      26          14
## 194     13.28    96            7         71  61      70          10
## 195      7.23   128           18        481  45      98          11
## 196      4.19   112            4        420  66      93          11
## 197      4.10   133            6        410  72      28          16
## 198      2.52   138            0        333  76      61          16
## 199      3.62   128            5        500  69      80          10
## 200      6.42   126            5        335  64      88          14
## 201      5.56   146            0        349  62      92          12
## 202      5.94   134            0        139  54      83          18
## 203      4.10   130            4        413  46      78          10
## 204      2.05   157            0        132  25      82          14
## 205      8.74   124            0        237  37      80          14
## 206      5.68   132            1        317  28      22          12
## 207      4.97   160            0         27  77      67          17
## 208      8.19    97            0        466  61     105          10
## 209      7.78    64            0        497  33      54          12
## 210      3.02    90           11        326  76      21          11
```

```
## 211     4.36   123           2        357  47     41         14
## 212     9.39   120          14        445  32    118         15
## 213    12.04   105          19        501  45     69         11
## 214     8.23   139           5        220  33     84         10
## 215     4.83   107           3         48  73    115         18
## 216     2.34   144          15        170  71     83         11
## 217     5.73   144           0        243  34     33         17
## 218     4.34   111           0        481  70     44         14
## 219     9.70   120          12        156  25     61         14
## 220    10.62   116          19        359  58     79         17
## 221    10.59   124          15        262  30    120         10
## 222     6.43   107           0        125  80     44         11
## 223     7.49   145           6        178  35    119         13
## 224     3.45   125           9        276  62     45         14
## 225     4.10   141           0        464  48     82         13
## 226     6.68    82           0        412  36     25         14
## 227     7.80   122           0        245  56     33         14
## 228     8.69   101          10         68  57     64         16
## 229     5.40   163          13        381  26     73         11
## 230    11.19    72           0        404  27    104         18
## 231     5.16   114           0        119  38     60         14
## 232     8.09   122           0        123  27     69         11
## 233    13.14   105          10         24  61     80         15
## 234     8.65   120          18        218  29     76         14
## 235     9.43   129          11        289  56     62         16
## 236     5.53   132           8         95  50     32         17
## 237     9.32   108          16        361  69     34         10
## 238     9.62   135           8        499  48     28         10
## 239     7.36   133           0        200  73     24         13
## 240     3.89   118           0        149  62    105         16
## 241    10.31   121           0        362  26     80         18
## 242    12.01    94           0        160  38     63         12
## 243     4.68   135           0        199  52     46         14
## 244     7.82   110          13         87  57     25         10
## 245     8.78   100           0        391  26     30         18
## 246    10.00    88           0        199  57     43         10
## 247     6.90    90          20        266  78     56         18
## 248     5.04   151           0        298  34    114         16
## 249     5.36   101           0         12  61     52         11
## 250     5.05   117           0         86  65     67         11
## 251     9.16   156          10        435  72    105         14
## 252     3.72   132           5        310  62    111         13
## 253     8.31   117           0         70  32     97         16
## 254     5.64   122           5        288  57     24         12
## 255     9.58   129          23        353  37    104         17
## 256     7.71    81           8        198  80     81         15
## 257     4.20   144           0        277  73     40         10
## 258     8.67   112          14        477  80     62         13
## 259     3.47    81           0        251  72     38         14
## 260     5.12   100          10        467  74     36         11
## 261     7.67   101           8        400  36    117         10
## 262     5.71   118           4        188  54     42         15
## 263     6.37   132          15         86  48     77         18
## 264     7.77   115           6        434  25     26         17
```

```
## 265     6.95  159           5         324  31     29         15
## 266     5.31  129          10         402  39     35         17
## 267     9.10  112          12         343  73     93         17
## 268     5.83  112           7         473  51     82         12
## 269     6.53  105           0          66  39     57         11
## 270     5.01  166           0         438  46     69         17
## 271    11.99   89           0         284  26     26         10
## 272     4.55  110           0         504  62     56         16
## 273    12.98   63           0          14  38     33         12
## 274    10.04   86           8         244  58    106         12
## 275     7.22  119           2          67  34     93         11
## 276     6.67  132          11         210  53    119         11
## 277     6.93  130          14         296  73     69         15
## 278     7.80  125          12         326  36     48         16
## 279     7.22  151           2         129  40    113         15
## 280     3.42  158          13         376  64     57         18
## 281     2.86  145          10         496  51     86         10
## 282    11.19  105           7         303  45     69         16
## 283     7.74  154           0          80  61     96         11
## 284     5.36  117           0         112  80    110         16
## 285     6.97   96          11         414  79     46         17
## 286     7.60  131          11         261  39     26         10
## 287     7.53  113          11         429  67    118         18
## 288     6.88   72           4         208  44     44         17
## 289     6.98   97           0          74  76     40         15
## 290     8.75  156          25         448  43     77         17
## 291     9.49  103          14         400  41    111         11
## 292     6.64   89           0         106  39     70         17
## 293    11.82   74          16         322  76     66         15
## 294    11.28   89           0          74  59     84         10
## 295    12.66   99           3         126  60     76         11
## 296     4.21  137          14         502  79     35         10
## 297     8.21  123          13         160  63     44         18
## 298     3.07  104          13         276  75     83         10
## 299    10.98  130           0         312  63     63         15
## 300     9.40   96          17         497  54     40         17
## 301     8.57   99           1         158  45     78         11
## 302     7.41   87           0         198  57     93         16
## 303     5.28  110          13         388  74     77         14
## 304    10.01   99          16         290  43     52         11
## 305    11.93  134          12         408  29     98         10
## 306     8.03  132          26         394  33     29         13
## 307     4.78  133           1          85  48     32         12
## 308     5.90  120           0          13  61     92         12
## 309     9.24  126          19         436  52     80         10
## 310    11.18   80          13          33  68    111         18
## 311     9.53  166          29         419  53     65         12
## 312     6.15  132          12         328  51     68         14
## 313     6.80  135           5         337  38    117         10
## 314     9.33   54           3         491  66     81         13
## 315     7.72  129          10         333  71     33         14
## 316     6.39  171           8         220  29     21         14
## 317    15.63   72           5         369  35     36         10
## 318     6.41  136           0         472  80     30         15
```

```
## 319     10.08   130            10         456  41     72          14
## 320      6.97   129            19         459  57     45          11
## 321      5.86   152            12         171  44     70          18
## 322      7.52    98             5         499  34     39          15
## 323      9.16   139            10         300  60     50          15
## 324     10.36   103            18         428  34    105          12
## 325      2.66   150             4         133  53     65          13
## 326     11.70   104            11         131  47     69          11
## 327      4.69   122             0         152  53     30          17
## 328      6.23   104            17         316  80     38          16
## 329      3.15   111             1          65  55     66          11
## 330     11.27    89             9         433  45     54          12
## 331      4.99   112             0         501  32     59          14
## 332     10.10   134            15         213  32     63          10
## 333      5.74   104            20         354  61     33          12
## 334      5.87   147             7         303  41     60          10
## 335      7.63    83             9         489  42    117          13
## 336      6.18   110            15         464  72     70          15
## 337      5.17   143             6          60  28     35          18
## 338      8.61   102             0         283  80     38          15
## 339      5.97   101             0         164  45     24          11
## 340     11.54   126             4         219  44     44          15
## 341      7.50    91             0         105  43     29          16
## 342      7.38    93             0         268  72    120          10
## 343      7.81   118            13         422  71    102          10
## 344      5.99   121            10         371  26     42          14
## 345      8.43   126             0         108  70     80          13
## 346      4.81   149             0         279  79     68          12
## 347      8.97   125             0         144  33    107          13
## 348      6.88   112             0         161  27     39          14
## 349     12.57   107            20         459  49    102          11
## 350      9.32    96            18         467  49     27          14
## 351      8.64    91            17         266  63    101          17
## 352     10.44   105            16         458  62    115          16
## 353     13.44   122            14         288  61    103          17
## 354      9.45    92            12         430  35     67          12
## 355      5.30   145             1          80  42     31          18
## 356      7.02   146             0         306  42    100          11
## 357      3.58   164             0         111  72    109          12
## 358     13.36    72             3         276  34     73          15
## 359      4.17   118            10          71  69     96          11
## 360      3.13   130            11         396  66     62          14
## 361      8.77   114             7         265  52     86          15
## 362      8.68   104            10         183  56     25          15
## 363      5.25   110             0          26  79     55          12
## 364     10.26   108             1         377  25     75          12
## 365     10.50   131            16         488  30     21          14
## 366      6.53   162             0         122  57     30          17
## 367      5.98   134            11         447  53     56          12
## 368     14.37    53             0         256  52    106          17
## 369     10.71    79            10         348  74     22          14
## 370     10.26   122            22         463  36    100          14
## 371      7.68   119            22         403  42     41          12
## 372      9.08   126             0         191  54     81          16
```

```
## 373      7.80    98            0         508  65      50          11
## 374      5.58   116            0         402  78      71          17
## 375      9.44   118            7          90  47      47          12
## 376      7.90   124            4         206  73      46          11
## 377     16.27    92           19         319  44      60          11
## 378      6.81   125            0         263  41      61          12
## 379      6.11   119            3         105  79      88          12
## 380      5.81   107            0         404  54     111          15
## 381      9.64    89           10          17  68      64          17
## 382      3.90   151           21         496  77      65          13
## 383      4.95   121           19         315  66      28          14
## 384      9.35    68            0          76  63     117          10
## 385     12.85   112           15         348  28      37          12
## 386      5.87   132           13         455  62      73          17
## 387      5.32   160            0         170  39     116          16
## 388      8.67   115           14         238  73      73          14
## 389      8.14    78           11         245  79      89          16
## 390      8.44   107            8         328  35      42          12
## 391      5.47   111            9          61  67      75          12
## 392      6.10   124            0          49  56      63          16
## 393      4.53   130           13         315  34      42          13
## 394      5.57   120           10          26  30      51          17
## 395      5.35   139           19         366  33      58          16
## 396     12.57   128           17         203  33     108          14
## 397      6.14   120            3          37  55      23          11
## 398      7.41   159           12         368  40      26          18
## 399      5.94    95            7         284  50      79          12
## 400      9.71   120            0          27  49      37          16
```

#Plot correlation headmap

```
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 4.1.2
```

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

```
ggcorr(Carseats_Filtered, label = TRUE, palette = "RdBu", name = "Correlation", hjust = 0.75, label_size
```

#Check to explore Misssing Data

```
#We Look at the summary of the dataset and see if there are NA's present in variables/columns
NA_perct <- function(df, fmt = F) {
  return (df %>%
            is.na() %>%
            colMeans() %>%
            sapply(function(x) {
              if (fmt) {
                return(sprintf("%.5f%", x))
              }
              return (x)
            })
          )
}
NA_perct_df <- NA_perct(Carseats_Filtered) %>%
  data_frame(Columns = names(.), `NA %` = .) %>%
  mutate_at(
    vars(`NA %`),
    funs(round(. * 100, 2))
  ) %>%
  mutate(label = sprintf("%g%%", `NA %`)) %>%
  arrange(desc(`NA %`))
```

```
## Warning: `data_frame()` was deprecated in tibble 1.1.0.
## Please use `tibble()` instead.
```

```
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.

## Warning: `funs()` was deprecated in dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

```
NA_perct_df %>% select(-label)
```

```
## # A tibble: 7 x 2
##   Columns     `NA %`
##   <chr>        <dbl>
## 1 ï..Sales         0
## 2 Price            0
## 3 Advertising      0
## 4 Population       0
## 5 Age              0
## 6 Income           0
## 7 Education        0
```

This indicates that there are no missing data

```
summary(Carseats_Filtered)
```

```
##     ï..Sales          Price        Advertising       Population
##  Min.   : 0.000   Min.   : 24.0   Min.   : 0.000   Min.   : 10.0
##  1st Qu.: 5.390   1st Qu.:100.0   1st Qu.: 0.000   1st Qu.:139.0
##  Median : 7.490   Median :117.0   Median : 5.000   Median :272.0
##  Mean   : 7.496   Mean   :115.8   Mean   : 6.635   Mean   :264.8
##  3rd Qu.: 9.320   3rd Qu.:131.0   3rd Qu.:12.000   3rd Qu.:398.5
##  Max.   :16.270   Max.   :191.0   Max.   :29.000   Max.   :509.0
##       Age            Income         Education
##  Min.   :25.00   Min.   : 21.00   Min.   :10.0
##  1st Qu.:39.75   1st Qu.: 42.75   1st Qu.:12.0
##  Median :54.50   Median : 69.00   Median :14.0
##  Mean   :53.32   Mean   : 68.66   Mean   :13.9
##  3rd Qu.:66.00   3rd Qu.: 91.00   3rd Qu.:16.0
##  Max.   :80.00   Max.   :120.00   Max.   :18.0
```

```r
glimpse(Carseats_Filtered)
```

```
## Rows: 400
## Columns: 7
## $ ï..Sales    <dbl> 9.50, 11.22, 10.06, 7.40, 4.15, 10.81, 6.63, 11.85, 6.54, ~
## $ Price       <int> 120, 83, 80, 97, 128, 72, 108, 120, 124, 124, 100, 94, 136~
## $ Advertising <int> 11, 16, 10, 4, 3, 13, 0, 15, 0, 0, 9, 4, 2, 11, 11, 5, 0, ~
## $ Population  <int> 276, 260, 269, 466, 340, 501, 45, 425, 108, 131, 150, 503,~
## $ Age         <int> 42, 65, 59, 55, 38, 78, 71, 67, 76, 76, 26, 50, 62, 53, 52~
## $ Income      <int> 73, 48, 35, 100, 64, 113, 105, 81, 110, 113, 78, 94, 35, 2~
## $ Education   <int> 17, 10, 12, 14, 13, 16, 15, 10, 10, 17, 10, 13, 18, 18, 18~
```

Predicting Sales of Baby Car Seats

```r
MyDecisionTree <-Carseats_Filtered[, 1:7]
Model_2 <-rpart(ï..Sales~., data=Carseats_Filtered, method="anova")
summary(Model_2)
```

```
## Call:
## rpart(formula = ï..Sales ~ ., data = Carseats_Filtered, method = "anova")
##   n= 400
##
##            CP nsplit rel error    xerror       xstd
## 1  0.14251535      0 1.0000000 1.0025580 0.06924325
## 2  0.08034146      1 0.8574847 0.9049745 0.06512823
## 3  0.06251702      2 0.7771432 0.8715420 0.06434565
## 4  0.02925241      3 0.7146262 0.7820490 0.05668452
## 5  0.02537341      4 0.6853738 0.7731898 0.05343946
## 6  0.02127094      5 0.6600003 0.7645757 0.05231488
## 7  0.02059174      6 0.6387294 0.7539634 0.05251800
## 8  0.01632010      7 0.6181377 0.7362951 0.05084472
## 9  0.01521801      8 0.6018176 0.7583298 0.05538272
## 10 0.01042023      9 0.5865996 0.8027608 0.05777675
## 11 0.01000559     10 0.5761793 0.8120555 0.05809388
## 12 0.01000000     12 0.5561681 0.8171755 0.05885499
##
## Variable importance
##       Price Advertising         Age      Income  Population   Education
##          49          18          16           8           6           3
##
## Node number 1: 400 observations,    complexity param=0.1425153
##   mean=7.496325, MSE=7.955687
##   left son=2 (329 obs) right son=3 (71 obs)
##   Primary splits:
##       Price       < 94.5  to the right, improve=0.14251530, (0 missing)
##       Advertising < 7.5   to the left,  improve=0.07303226, (0 missing)
##       Age         < 61.5  to the right, improve=0.07120203, (0 missing)
##       Income      < 61.5  to the left,  improve=0.02840494, (0 missing)
##       Population  < 174.5 to the left,  improve=0.01077467, (0 missing)
##
## Node number 2: 329 observations,    complexity param=0.08034146
```

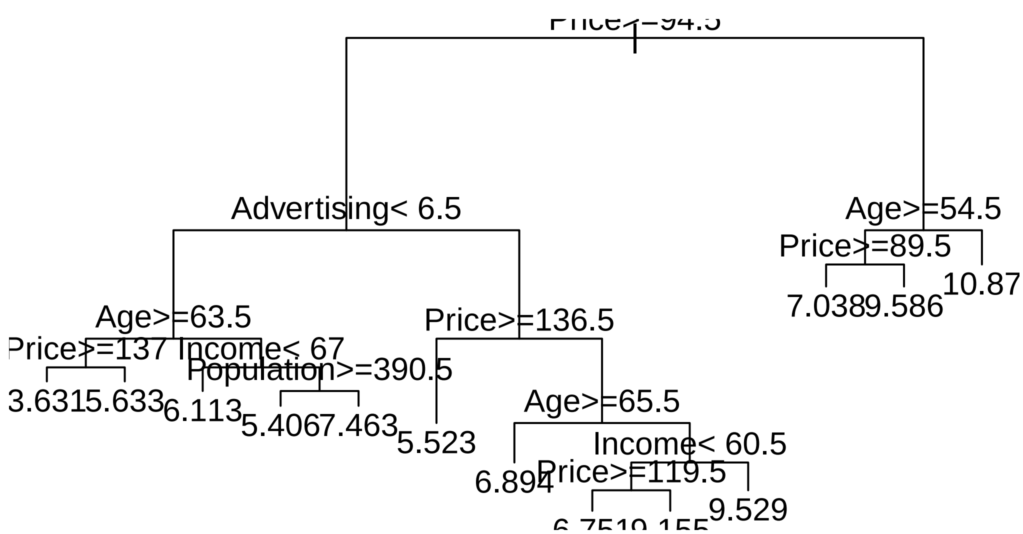
```
##   mean=7.001672, MSE=6.815199
##   left son=4 (174 obs) right son=5 (155 obs)
##   Primary splits:
##       Advertising < 6.5   to the left,  improve=0.11402580, (0 missing)
##       Price       < 136.5 to the right, improve=0.08411056, (0 missing)
##       Age         < 63.5  to the right, improve=0.08091745, (0 missing)
##       Income      < 60.5  to the left,  improve=0.03394126, (0 missing)
##       Population  < 23    to the left,  improve=0.01831455, (0 missing)
##   Surrogate splits:
##       Population < 223   to the left,  agree=0.599, adj=0.148, (0 split)
##       Education  < 10.5  to the right, agree=0.565, adj=0.077, (0 split)
##       Age        < 53.5  to the right, agree=0.547, adj=0.039, (0 split)
##       Income     < 114.5 to the left,  agree=0.547, adj=0.039, (0 split)
##       Price      < 106.5 to the right, agree=0.544, adj=0.032, (0 split)
##
## Node number 3: 71 observations,    complexity param=0.02537341
##   mean=9.788451, MSE=6.852836
##   left son=6 (36 obs) right son=7 (35 obs)
##   Primary splits:
##       Age        < 54.5  to the right, improve=0.16595410, (0 missing)
##       Price      < 75.5  to the right, improve=0.08365773, (0 missing)
##       Income     < 30.5  to the left,  improve=0.03322169, (0 missing)
##       Education  < 10.5  to the right, improve=0.03019634, (0 missing)
##       Population < 268.5 to the left,  improve=0.02383306, (0 missing)
##   Surrogate splits:
##       Advertising < 4.5   to the right, agree=0.606, adj=0.200, (0 split)
##       Price       < 73    to the right, agree=0.592, adj=0.171, (0 split)
##       Population  < 272.5 to the left,  agree=0.592, adj=0.171, (0 split)
##       Income      < 79.5  to the right, agree=0.592, adj=0.171, (0 split)
##       Education   < 11.5  to the left,  agree=0.577, adj=0.143, (0 split)
##
## Node number 4: 174 observations,    complexity param=0.02127094
##   mean=6.169655, MSE=4.942347
##   left son=8 (58 obs) right son=9 (116 obs)
##   Primary splits:
##       Age         < 63.5  to the right, improve=0.078712160, (0 missing)
##       Price       < 130.5 to the right, improve=0.048919280, (0 missing)
##       Population  < 26.5  to the left,  improve=0.030421540, (0 missing)
##       Income      < 67.5  to the left,  improve=0.027749670, (0 missing)
##       Advertising < 0.5   to the left,  improve=0.006795377, (0 missing)
##   Surrogate splits:
##       Income     < 22.5  to the left,  agree=0.678, adj=0.034, (0 split)
##       Price      < 96.5  to the left,  agree=0.672, adj=0.017, (0 split)
##       Population < 26.5  to the left,  agree=0.672, adj=0.017, (0 split)
##
## Node number 5: 155 observations,    complexity param=0.06251702
##   mean=7.935677, MSE=7.268151
##   left son=10 (28 obs) right son=11 (127 obs)
##   Primary splits:
##       Price       < 136.5 to the right, improve=0.17659580, (0 missing)
##       Age         < 73.5  to the right, improve=0.08000201, (0 missing)
##       Income      < 60.5  to the left,  improve=0.05360755, (0 missing)
##       Advertising < 13.5  to the left,  improve=0.03920507, (0 missing)
##       Population  < 399   to the left,  improve=0.01037956, (0 missing)
```

```
##   Surrogate splits:
##       Advertising < 24.5  to the right, agree=0.826, adj=0.036, (0 split)
##
## Node number 6: 36 observations,    complexity param=0.0163201
##   mean=8.736944, MSE=4.961043
##   left son=12 (12 obs) right son=13 (24 obs)
##   Primary splits:
##       Price       < 89.5  to the right, improve=0.29079360, (0 missing)
##       Income      < 39.5  to the left,  improve=0.19043350, (0 missing)
##       Advertising < 11.5  to the left,  improve=0.17891930, (0 missing)
##       Age         < 75.5  to the right, improve=0.04316067, (0 missing)
##       Education   < 14.5  to the left,  improve=0.03411396, (0 missing)
##   Surrogate splits:
##       Advertising < 16.5  to the right, agree=0.722, adj=0.167, (0 split)
##       Income      < 37.5  to the left,  agree=0.722, adj=0.167, (0 split)
##       Age         < 56.5  to the left,  agree=0.694, adj=0.083, (0 split)
##
## Node number 7: 35 observations
##   mean=10.87, MSE=6.491674
##
## Node number 8: 58 observations,    complexity param=0.01042023
##   mean=5.287586, MSE=3.93708
##   left son=16 (10 obs) right son=17 (48 obs)
##   Primary splits:
##       Price      < 137   to the right, improve=0.14521540, (0 missing)
##       Education  < 15.5  to the right, improve=0.07995394, (0 missing)
##       Income     < 35.5  to the left,  improve=0.04206708, (0 missing)
##       Age        < 79.5  to the left,  improve=0.02799057, (0 missing)
##       Population < 52.5  to the left,  improve=0.01914342, (0 missing)
##
## Node number 9: 116 observations,    complexity param=0.01000559
##   mean=6.61069, MSE=4.861446
##   left son=18 (58 obs) right son=19 (58 obs)
##   Primary splits:
##       Income     < 67    to the left,  improve=0.05085914, (0 missing)
##       Population < 392   to the right, improve=0.04476721, (0 missing)
##       Price      < 127   to the right, improve=0.04210762, (0 missing)
##       Age        < 37.5  to the right, improve=0.02858424, (0 missing)
##       Education  < 14.5  to the left,  improve=0.01187387, (0 missing)
##   Surrogate splits:
##       Education   < 12.5  to the right, agree=0.586, adj=0.172, (0 split)
##       Age         < 58.5  to the left,  agree=0.578, adj=0.155, (0 split)
##       Price       < 144.5 to the left,  agree=0.569, adj=0.138, (0 split)
##       Population  < 479   to the right, agree=0.560, adj=0.121, (0 split)
##       Advertising < 2.5   to the right, agree=0.543, adj=0.086, (0 split)
##
## Node number 10: 28 observations
##   mean=5.522857, MSE=5.084213
##
## Node number 11: 127 observations,    complexity param=0.02925241
##   mean=8.467638, MSE=6.183142
##   left son=22 (29 obs) right son=23 (98 obs)
##   Primary splits:
##       Age         < 65.5  to the right, improve=0.11854590, (0 missing)
```

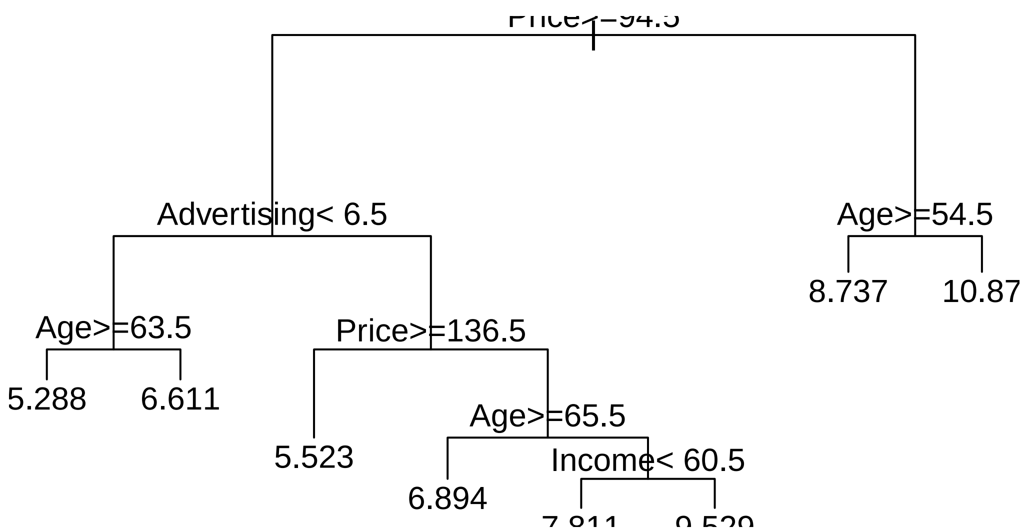
```
##       Income      < 51.5  to the left,  improve=0.08076060, (0 missing)
##       Advertising < 13.5  to the left,  improve=0.04801701, (0 missing)
##       Education   < 11.5  to the right, improve=0.02471512, (0 missing)
##       Population  < 479   to the left,  improve=0.01908657, (0 missing)
##
## Node number 12: 12 observations
##   mean=7.038333, MSE=2.886964
##
## Node number 13: 24 observations
##   mean=9.58625, MSE=3.834123
##
## Node number 16: 10 observations
##   mean=3.631, MSE=5.690169
##
## Node number 17: 48 observations
##   mean=5.632708, MSE=2.88102
##
## Node number 18: 58 observations
##   mean=6.113448, MSE=3.739109
##
## Node number 19: 58 observations,    complexity param=0.01000559
##   mean=7.107931, MSE=5.489285
##   left son=38 (10 obs) right son=39 (48 obs)
##   Primary splits:
##       Population  < 390.5 to the right, improve=0.10993270, (0 missing)
##       Price       < 124.5 to the right, improve=0.07534567, (0 missing)
##       Advertising < 0.5   to the left,  improve=0.07060488, (0 missing)
##       Age         < 45.5  to the right, improve=0.04611510, (0 missing)
##       Education   < 11.5  to the right, improve=0.03722944, (0 missing)
##
## Node number 22: 29 observations
##   mean=6.893793, MSE=6.08343
##
## Node number 23: 98 observations,    complexity param=0.02059174
##   mean=8.933367, MSE=5.262759
##   left son=46 (34 obs) right son=47 (64 obs)
##   Primary splits:
##       Income      < 60.5  to the left,  improve=0.12705480, (0 missing)
##       Advertising < 13.5  to the left,  improve=0.07114001, (0 missing)
##       Price       < 118.5 to the right, improve=0.06932216, (0 missing)
##       Education   < 11.5  to the right, improve=0.03377416, (0 missing)
##       Age         < 49.5  to the right, improve=0.02289004, (0 missing)
##   Surrogate splits:
##       Education < 17.5  to the right, agree=0.663, adj=0.029, (0 split)
##
## Node number 38: 10 observations
##   mean=5.406, MSE=2.508524
##
## Node number 39: 48 observations
##   mean=7.4625, MSE=5.381106
##
## Node number 46: 34 observations,    complexity param=0.01521801
##   mean=7.811471, MSE=4.756548
##   left son=92 (19 obs) right son=93 (15 obs)
```

```
##   Primary splits:
##       Price       < 119.5 to the right, improve=0.29945020, (0 missing)
##       Advertising < 11.5  to the left,  improve=0.14268440, (0 missing)
##       Income      < 40.5  to the right, improve=0.12781140, (0 missing)
##       Population  < 152   to the left,  improve=0.03601768, (0 missing)
##       Age         < 49.5  to the right, improve=0.02748814, (0 missing)
##   Surrogate splits:
##       Education   < 12.5  to the right, agree=0.676, adj=0.267, (0 split)
##       Advertising < 7.5   to the right, agree=0.647, adj=0.200, (0 split)
##       Age         < 53.5  to the left,  agree=0.647, adj=0.200, (0 split)
##       Population  < 240   to the right, agree=0.618, adj=0.133, (0 split)
##       Income      < 41.5  to the right, agree=0.618, adj=0.133, (0 split)
##
## Node number 47: 64 observations
##   mean=9.529375, MSE=4.5078
##
## Node number 92: 19 observations
##   mean=6.751053, MSE=3.378915
##
## Node number 93: 15 observations
##   mean=9.154667, MSE=3.273025
```

```
# Too Complex
plot(Model_2)
text(Model_2)
```



```
##Less Complex
Model_2 = rpart (ï..Sales~.,data=Carseats_Filtered, method="anova", control = rpart.control(minsplit =
plot(Model_2)
text(Model_2)
```

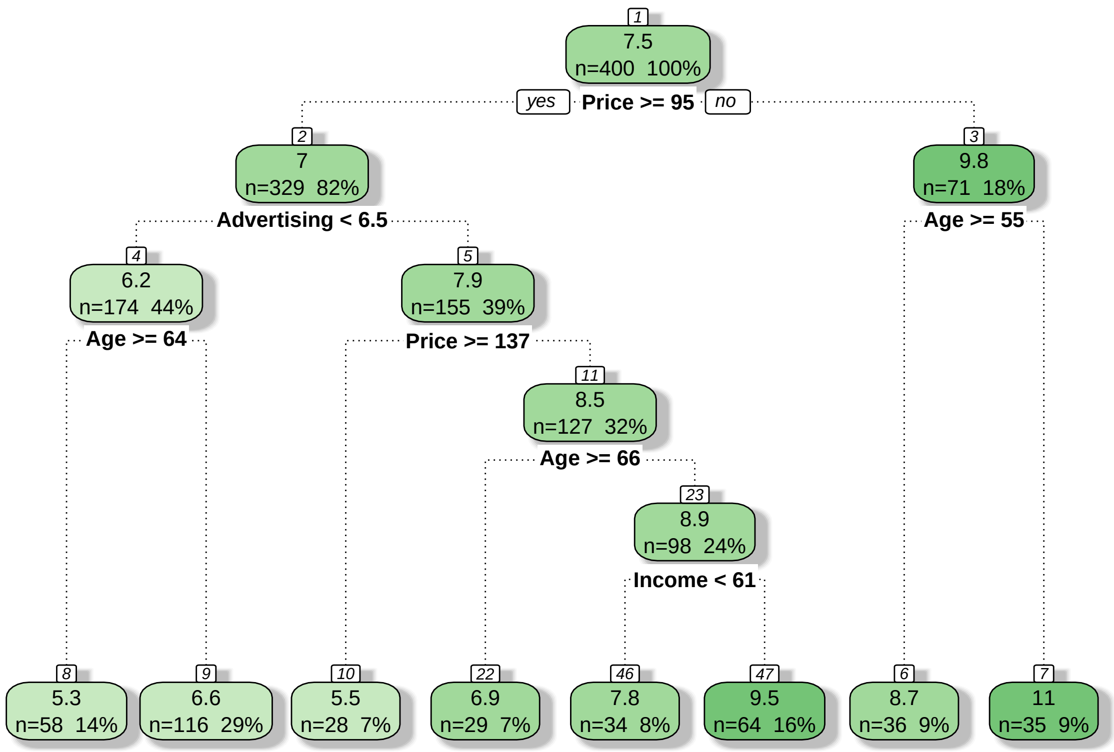



```
#Difficult to follow
print(Model_2)
```

```
## n= 400
##
## node), split, n, deviance, yval
##       * denotes terminal node
##
##  1) root 400 3182.2750  7.496325
##    2) Price>=94.5 329 2242.2000  7.001672
##      4) Advertising< 6.5 174  859.9684  6.169655
##        8) Age>=63.5 58  228.3507  5.287586 *
##        9) Age< 63.5 116  563.9277  6.610690 *
##      5) Advertising>=6.5 155 1126.5630  7.935677
##       10) Price>=136.5 28  142.3580  5.522857 *
##       11) Price< 136.5 127  785.2591  8.467638
##         22) Age>=65.5 29  176.4195  6.893793 *
##         23) Age< 65.5 98  515.7504  8.933367
##           46) Income< 60.5 34  161.7226  7.811471 *
##           47) Income>=60.5 64  288.4992  9.529375 *
```

```
##    3) Price< 94.5 71  486.5513  9.788451
##      6) Age>=54.5 36  178.5976  8.736944 *
##      7) Age< 54.5 35  227.2086 10.870000 *
```

```
library(rattle)
library(rpart.plot)
fancyRpartPlot(Model_2)
```



Rattle 2022−Apr−12 13:39:32 Mukht

##QB2. Consider the following input: Sales=9, Price=6.54, Population=124, Advertising=0, Age=76, Income= 110, Education=10 What will be the estimated Sales for this record using the decision tree model?

```
Model_3 <- data.frame(Price=6.54, Population=124, Advertising=0, Age=76, Income= 110, Education=10)
predict(Model_2, Model_3 , method = "anova")
```

```
##        1
## 8.736944
```