**Insurance factors identification.**

Q1) The committee is interested to know each field of the data collected through descriptive analysis to gain basic insights into the data set and to prepare for further analysis.

Ans) For descriptive analysis to gain insights into the data set,the summary function on the insurance dataset will be helpful for further analysis. Below attached is the code in R which gives an overview of the descriptive analysis

print("Insurance Factors Identification")

insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")

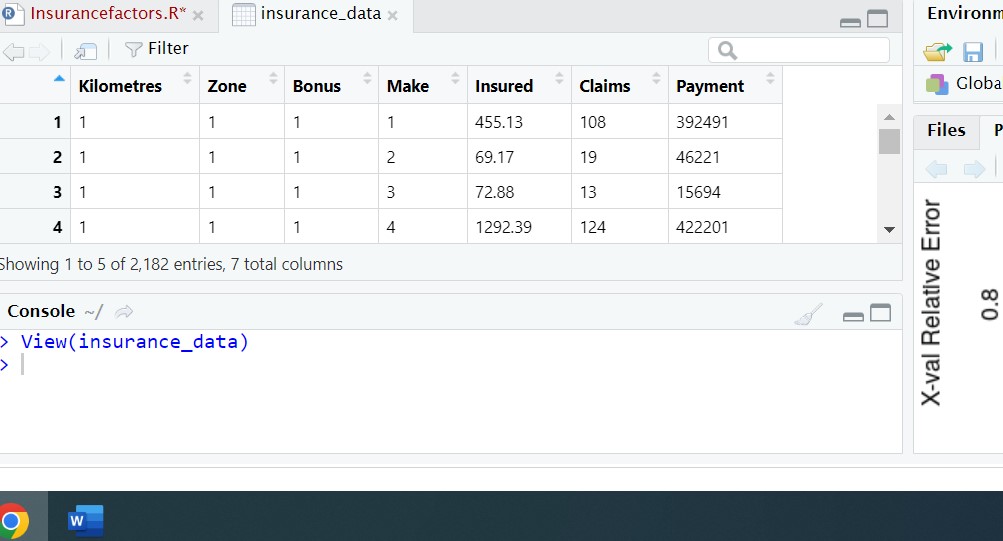
print(insurance\_data)

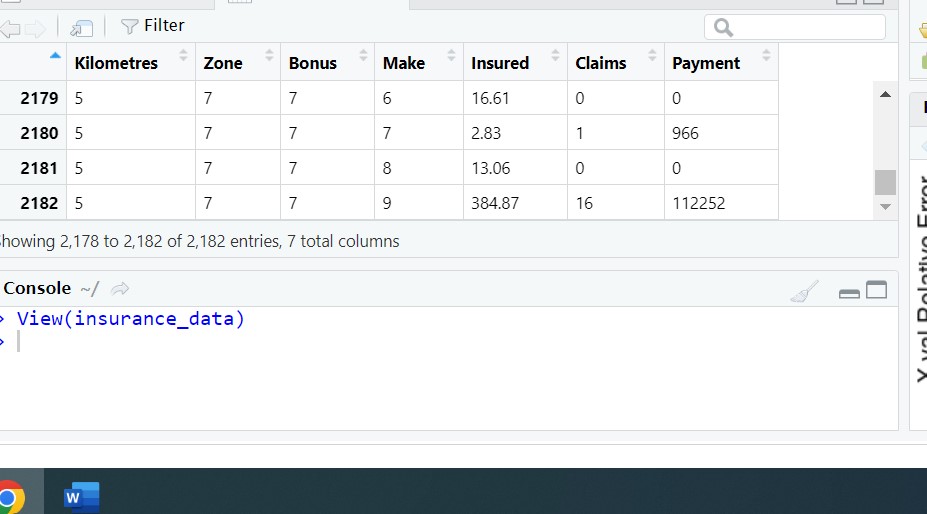
View(insurance\_data)

str(insurance\_data)

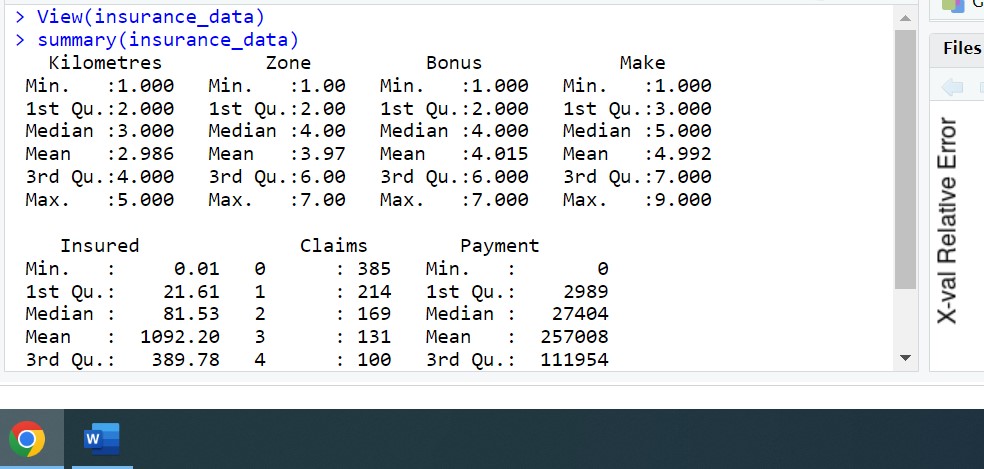
summary(insurance\_data)

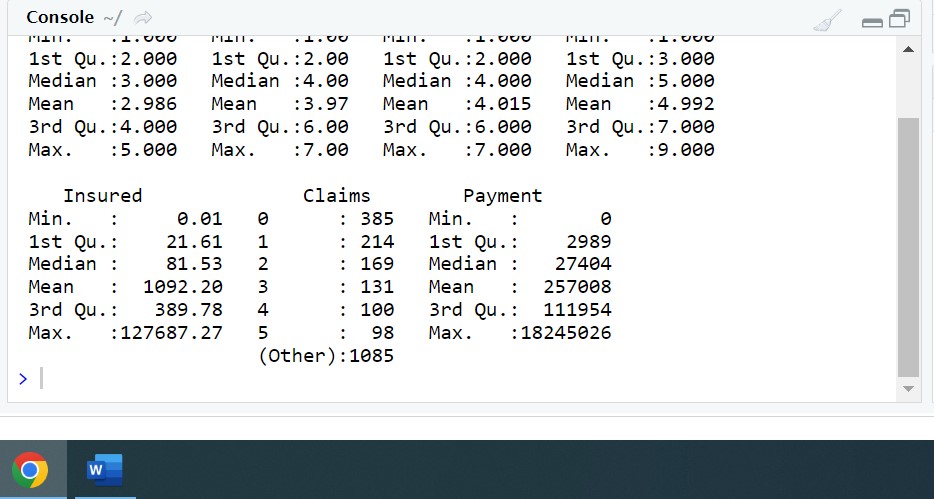
Here,are the screenshots of the descriptive analysis of data





 The table of insurance factors identification has 2,182 entries of 7 columns





The above attached are the summary of insurance dataset which has minimum, maximum,1st Quantile,Median and Mean values of all 7 columns of insurancedataset.

Q2) The total value of payment by an insurance company is an important factor to be monitored. So the committee has decided to find whether this payment is related to the number of claims and the number of insured policy years. They also want to visualize the results for better understanding.

Ans) For the analysis of payment wrt number of claims and the number of insured policy years,the concept of Simple Linear Regression with multiple variables can be used for determining the relationship. Below mention is the code in R

print("Insurance Factors Identification")

insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")

print(insurance\_data)

payment\_data<-lm(formula=Payment~Insured+Claims,data=insurance\_data)

print(payment\_data)

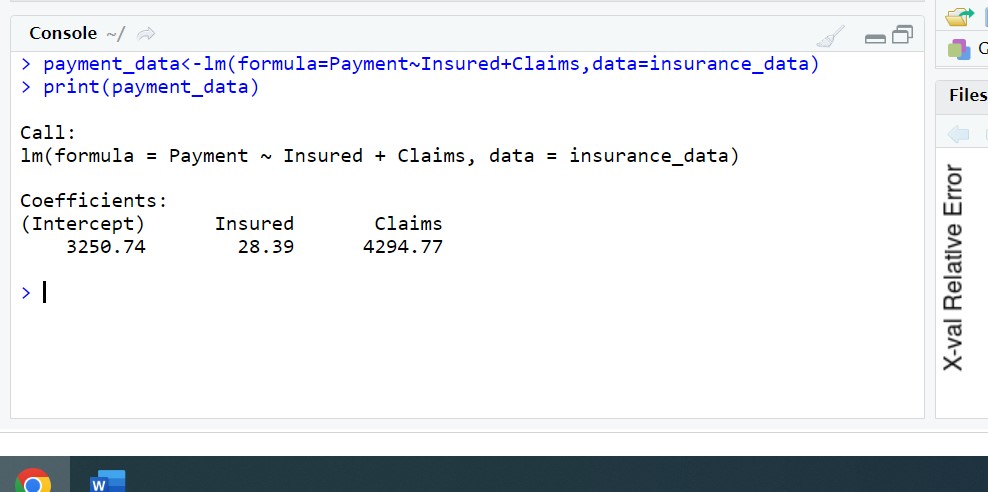
myinsureddata<-ggplot(data=insurance\_data,mapping=aes(x=Insured+Claims,y=Payment)) +geom\_point(alpha=0.1,color="blue")

myinsureddata

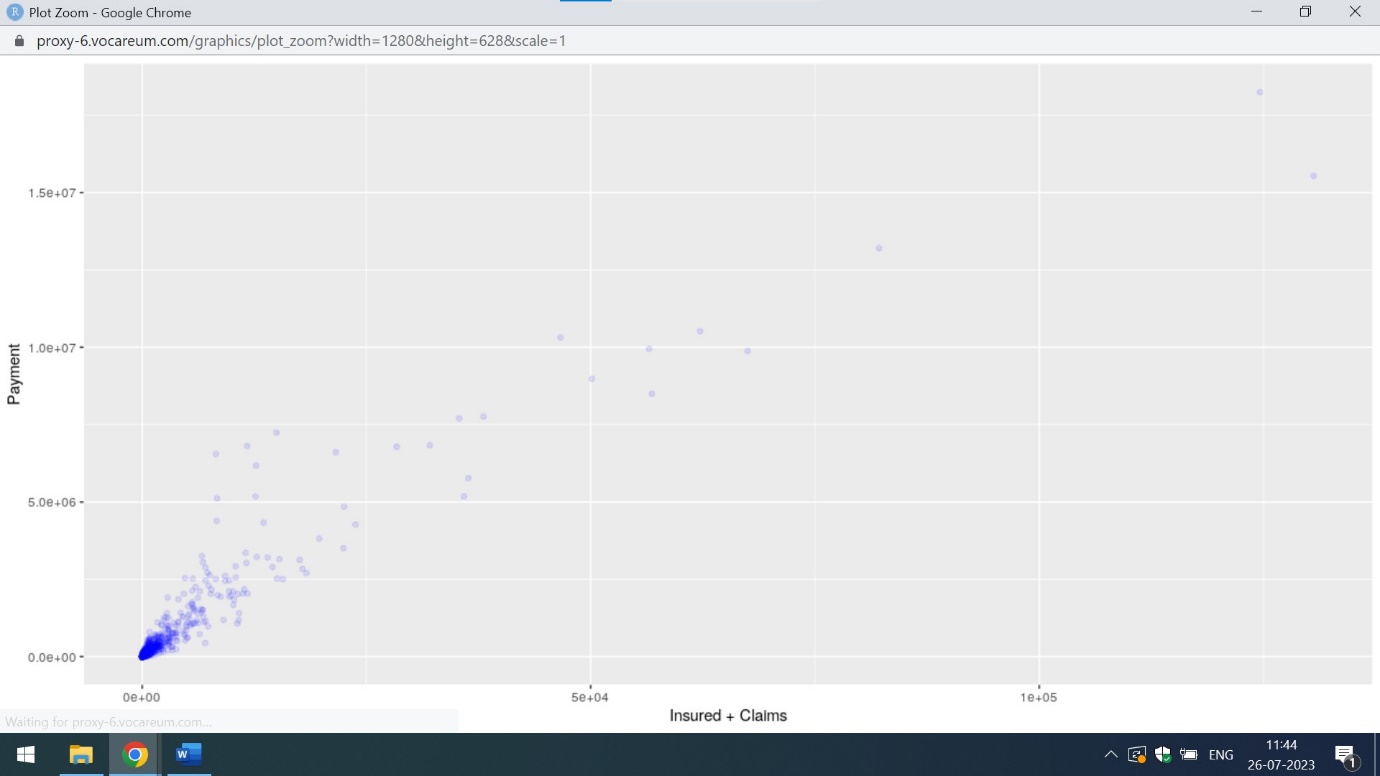
png(file="Insured.png")

boxplot(Payment~Insured,data=insurance\_data,xlab="Number of Insured in Policy Years",ylab="Total Value of Payment in SKR")

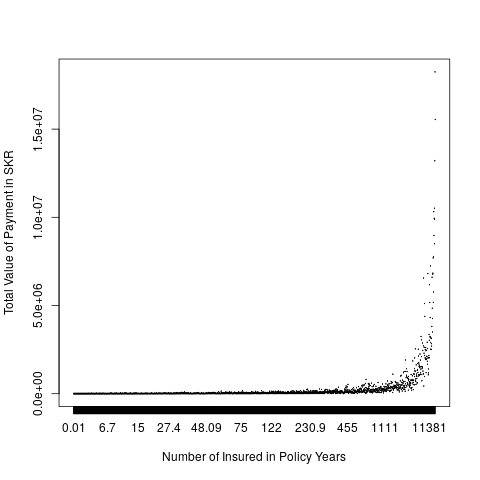
dev.off()



By applying the formula of Insured+Claims on Payment, it is clear that there is a positive Linear Regression between Payment wrt Insured and Claims as the value of both y and x intercepts are positive. Hence, payment is related to the number of claims and the number of insured policy years.



It is even cleared from the above attached ggplot view that there is a positive Linear Regression between Insured+Claims and Payment.



The above attached is the boxplot view of Number of Insured in Policy Years wrt Total Value in Payment. The total value of payment is increasing after 455 number of Insured in Policy Years

Q3) The committee wants to figure out the reasons for insurance payment increase and decrease. So they have decided to find whether distance, location, bonus, make, and insured amount or claims are affecting the payment or all or some of these are affecting it.

Ans) For analysing the reasons for insurance payment increase and decrease,I am using the concept of Simple Linear Regression with multiple variables. Below attached is the code in R

print("Insurance Factors Identification")

insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")

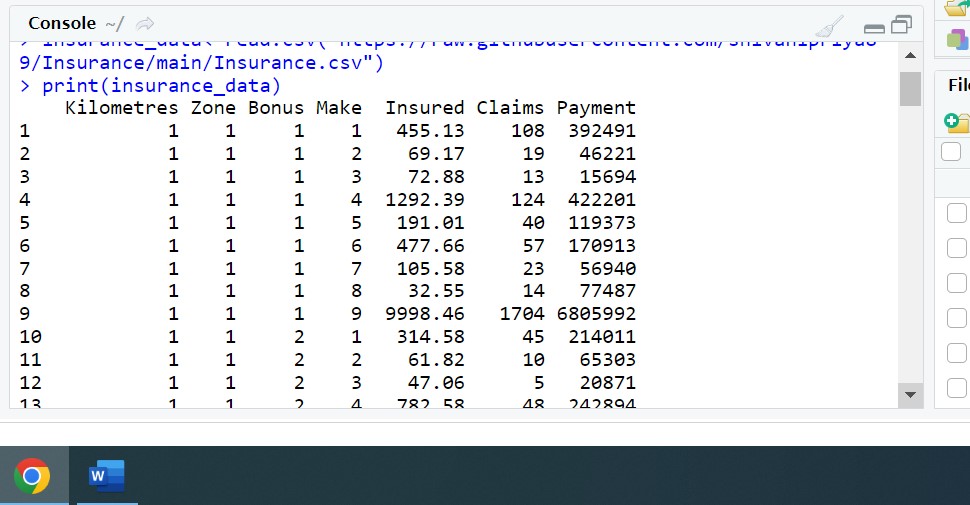
print(insurance\_data)

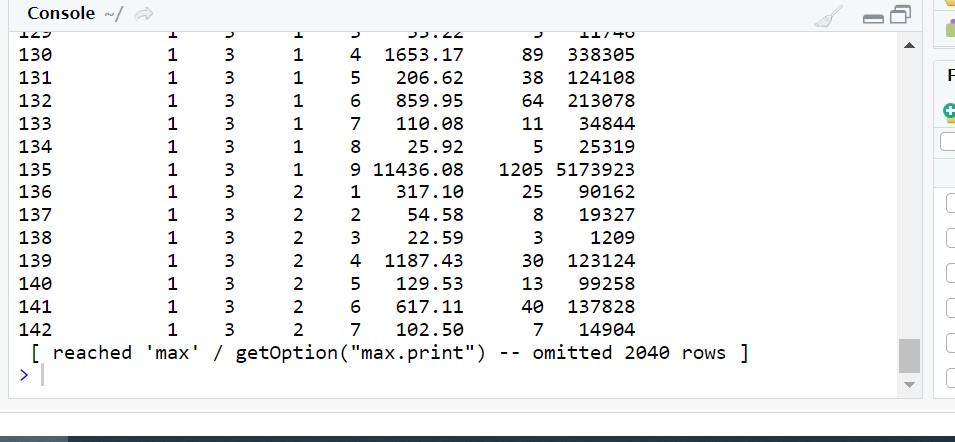
insurance\_results<-lm(formula=Payment~Kilometres+Zone+Bonus+Make+Insured+Claims,data=insurance\_data)

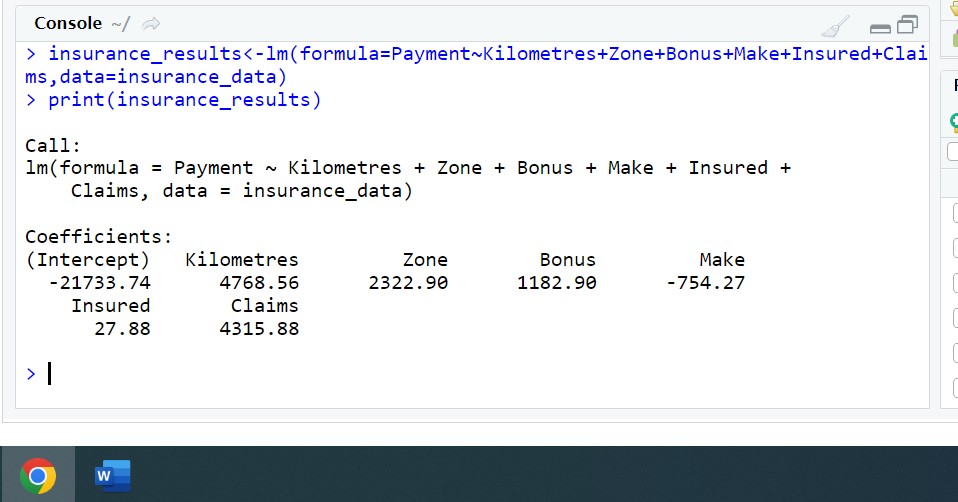
print(insurance\_results)

summary(insurance\_results)

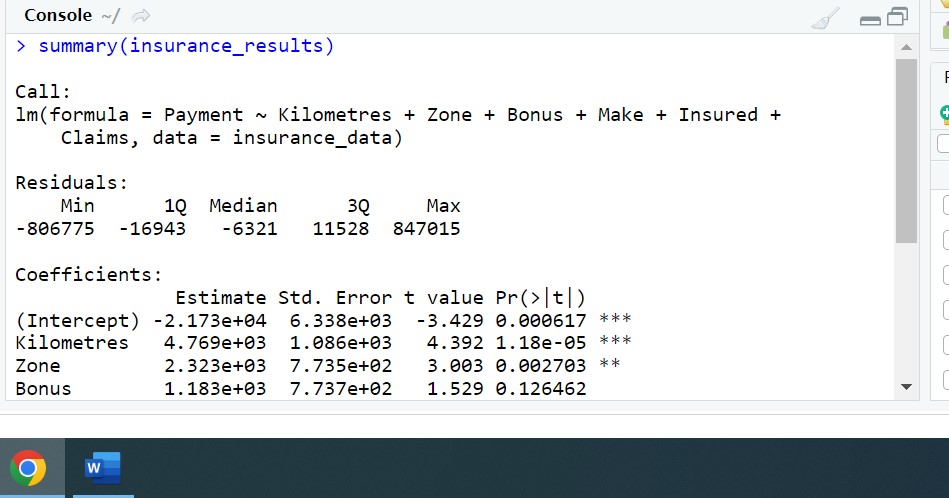
Below attached are the screenshots

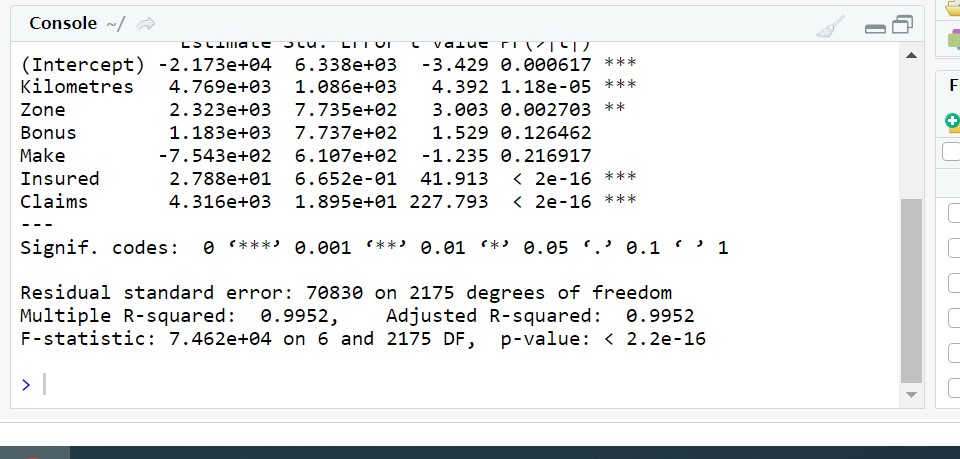






From the above mention screenshot,it is clear that Kilometres which is distance,Zone ie Location,Bonus ie No claims bonus; equal to the number of years, plus one, since the last claim,Insured ie The number of insured in policy-years and Claims are the major factors affecting the payment. Make ie 8 different common car models have no impact on payment





|  |
| --- |
|  |
|  |

It is even clear from the summary of the insurance factors identification dataset ie make has no impact on Payment while zone,bonus,insured and claims have positive impact

Q4) The insurance company is planning to establish a new branch office, so they are interested to find at what location, kilometre, and bonus level their insured amount, claims, and payment gets increased. (Hint: Aggregate Dataset)

Ans) By using the concept of Decision Tree Algorithm, one can analyze the various factors affecting the payment. Below attached is the code in R

print("Insurance Factors Identification")

insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")

print(insurance\_data)

# Convert zone to factor

insurance\_data$Zone<-sapply(insurance\_data$Zone,factor)

str(insurance\_data)

# Building the model

insurance\_analysis<-rpart(Zone~.,data=insurance\_data,method="class")

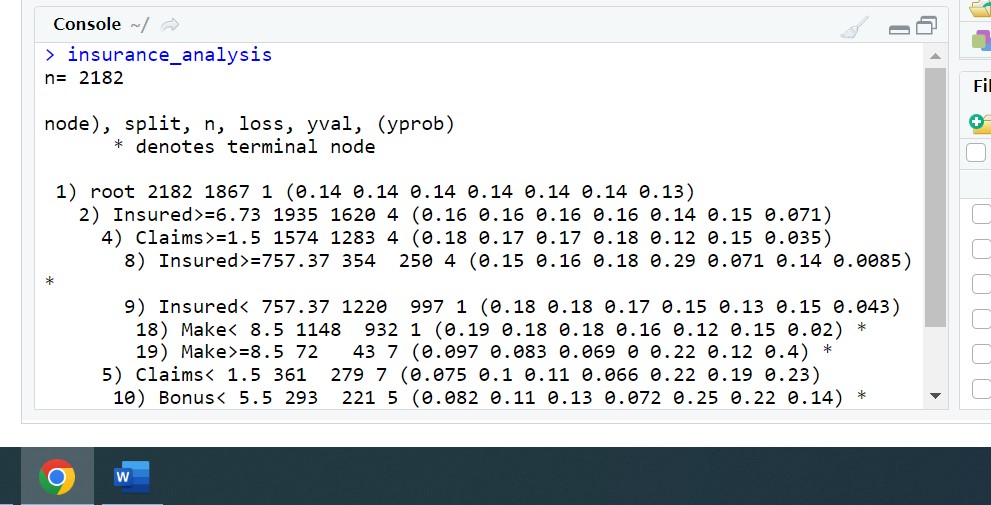
insurance\_analysis

summary(insurance\_analysis)

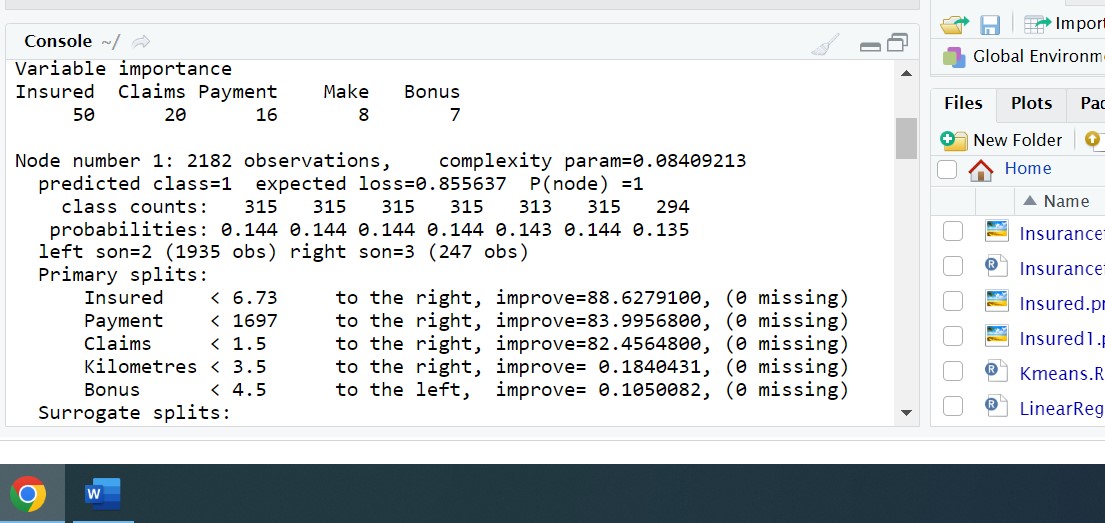
printcp(insurance\_analysis)

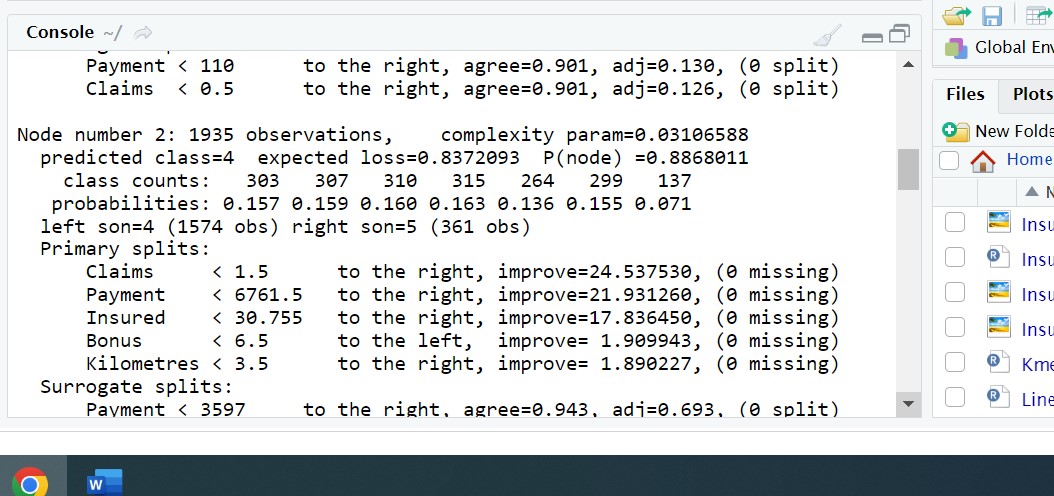
plotcp(insurance\_analysis)

Below attached are the screenshots



From the above mention screenshot,it is clear that there are total 2182 variables of the dataset which has splitends ,loss ,yval abd prob of the datset.





From the above attached screenshots,it is clear that Insured and claims are the most important factors in decision analysis

The actual output of the insurance factors identification are mention below:

> summary(insurance\_analysis)

Call:

rpart(formula = Zone ~ ., data = insurance\_data, method = "class")

n= 2182

CP nsplit rel error xerror xstd

1 0.08409213 0 1.0000000 1.0374933 0.007899074

2 0.03106588 1 0.9159079 0.9335833 0.010030192

3 0.01928227 2 0.8848420 0.8976968 0.010559432

4 0.01660418 3 0.8655597 0.8891269 0.010673764

5 0.01178361 4 0.8489555 0.8714515 0.010896037

6 0.01000000 5 0.8371719 0.8510980 0.011130575

Variable importance

Insured Claims Payment Make Bonus

50 20 16 8 7

Node number 1: 2182 observations, complexity param=0.08409213

predicted class=1 expected loss=0.855637 P(node) =1

class counts: 315 315 315 315 313 315 294

probabilities: 0.144 0.144 0.144 0.144 0.143 0.144 0.135

left son=2 (1935 obs) right son=3 (247 obs)

Primary splits:

Insured < 6.73 to the right, improve=88.6279100, (0 missing)

Payment < 1697 to the right, improve=83.9956800, (0 missing)

Claims < 1.5 to the right, improve=82.4564800, (0 missing)

Kilometres < 3.5 to the right, improve= 0.1840431, (0 missing)

Bonus < 4.5 to the left, improve= 0.1050082, (0 missing)

Surrogate splits:

Payment < 110 to the right, agree=0.901, adj=0.130, (0 split)

Claims < 0.5 to the right, agree=0.901, adj=0.126, (0 split)

Node number 2: 1935 observations, complexity param=0.03106588

predicted class=4 expected loss=0.8372093 P(node) =0.8868011

class counts: 303 307 310 315 264 299 137

probabilities: 0.157 0.159 0.160 0.163 0.136 0.155 0.071

left son=4 (1574 obs) right son=5 (361 obs)

Primary splits:

Claims < 1.5 to the right, improve=24.537530, (0 missing)

Payment < 6761.5 to the right, improve=21.931260, (0 missing)

Insured < 30.755 to the right, improve=17.836450, (0 missing)

Bonus < 6.5 to the left, improve= 1.909943, (0 missing)

Kilometres < 3.5 to the right, improve= 1.890227, (0 missing)

Surrogate splits:

Payment < 3597 to the right, agree=0.943, adj=0.693, (0 split)

Insured < 22.925 to the right, agree=0.882, adj=0.368, (0 split)

Node number 3: 247 observations

predicted class=7 expected loss=0.3643725 P(node) =0.1131989

class counts: 12 8 5 0 49 16 157

probabilities: 0.049 0.032 0.020 0.000 0.198 0.065 0.636

Node number 4: 1574 observations, complexity param=0.01928227

predicted class=4 expected loss=0.8151207 P(node) =0.7213566

class counts: 276 270 270 291 183 229 55

probabilities: 0.175 0.172 0.172 0.185 0.116 0.145 0.035

left son=8 (354 obs) right son=9 (1220 obs)

Primary splits:

Insured < 757.37 to the right, improve=7.096855, (0 missing)

Payment < 294105.5 to the right, improve=5.264421, (0 missing)

Claims < 4.5 to the right, improve=4.858938, (0 missing)

Make < 8.5 to the left, improve=4.084324, (0 missing)

Bonus < 5.5 to the left, improve=1.564487, (0 missing)

Surrogate splits:

Claims < 44.5 to the right, agree=0.935, adj=0.709, (0 split)

Payment < 272516 to the right, agree=0.930, adj=0.689, (0 split)

Make < 8.5 to the right, agree=0.837, adj=0.277, (0 split)

Bonus < 6.5 to the right, agree=0.801, adj=0.113, (0 split)

Node number 5: 361 observations, complexity param=0.01660418

predicted class=7 expected loss=0.7728532 P(node) =0.1654445

class counts: 27 37 40 24 81 70 82

probabilities: 0.075 0.102 0.111 0.066 0.224 0.194 0.227

left son=10 (293 obs) right son=11 (68 obs)

Primary splits:

Bonus < 5.5 to the left, improve=14.312300, (0 missing)

Kilometres < 3.5 to the right, improve= 9.077050, (0 missing)

Make < 1.5 to the right, improve= 7.278342, (0 missing)

Payment < 4228.5 to the right, improve= 2.401977, (0 missing)

Insured < 21.39 to the left, improve= 2.209768, (0 missing)

Surrogate splits:

Insured < 58.62 to the left, agree=0.82, adj=0.044, (0 split)

Node number 8: 354 observations

predicted class=4 expected loss=0.7062147 P(node) =0.1622365

class counts: 53 56 63 104 25 50 3

probabilities: 0.150 0.158 0.178 0.294 0.071 0.141 0.008

Node number 9: 1220 observations, complexity param=0.01178361

predicted class=1 expected loss=0.8172131 P(node) =0.5591201

class counts: 223 214 207 187 158 179 52

probabilities: 0.183 0.175 0.170 0.153 0.130 0.147 0.043

left son=18 (1148 obs) right son=19 (72 obs)

Primary splits:

Make < 8.5 to the left, improve=14.394020, (0 missing)

Bonus < 6.5 to the left, improve= 5.007216, (0 missing)

Payment < 375304.5 to the right, improve= 4.656929, (0 missing)

Claims < 85.5 to the right, improve= 4.149427, (0 missing)

Insured < 78.635 to the left, improve= 2.495010, (0 missing)

Surrogate splits:

Claims < 61 to the left, agree=0.944, adj=0.056, (0 split)

Payment < 421713 to the left, agree=0.943, adj=0.028, (0 split)

Node number 10: 293 observations

predicted class=5 expected loss=0.7542662 P(node) =0.1342805

class counts: 24 33 38 21 72 64 41

probabilities: 0.082 0.113 0.130 0.072 0.246 0.218 0.140

Node number 11: 68 observations

predicted class=7 expected loss=0.3970588 P(node) =0.03116407

class counts: 3 4 2 3 9 6 41

probabilities: 0.044 0.059 0.029 0.044 0.132 0.088 0.603

Node number 18: 1148 observations

predicted class=1 expected loss=0.8118467 P(node) =0.5261228

class counts: 216 208 202 187 142 170 23

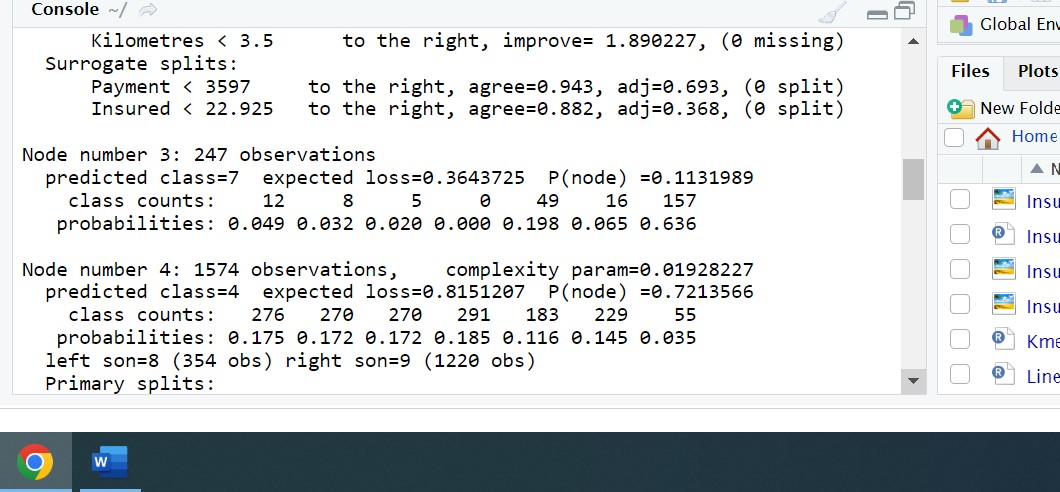
probabilities: 0.188 0.181 0.176 0.163 0.124 0.148 0.020

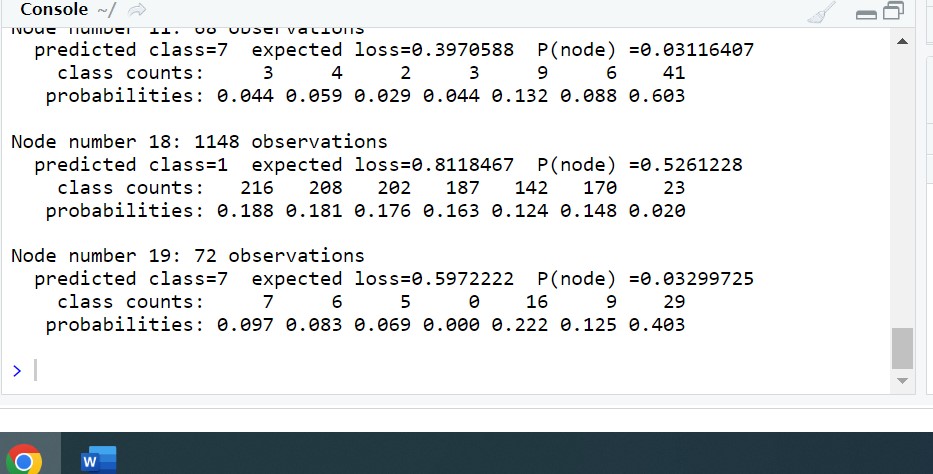
Node number 19: 72 observations

predicted class=7 expected loss=0.5972222 P(node) =0.03299725

class counts: 7 6 5 0 16 9 29

probabilities: 0.097 0.083 0.069 0.000 0.222 0.125 0.403

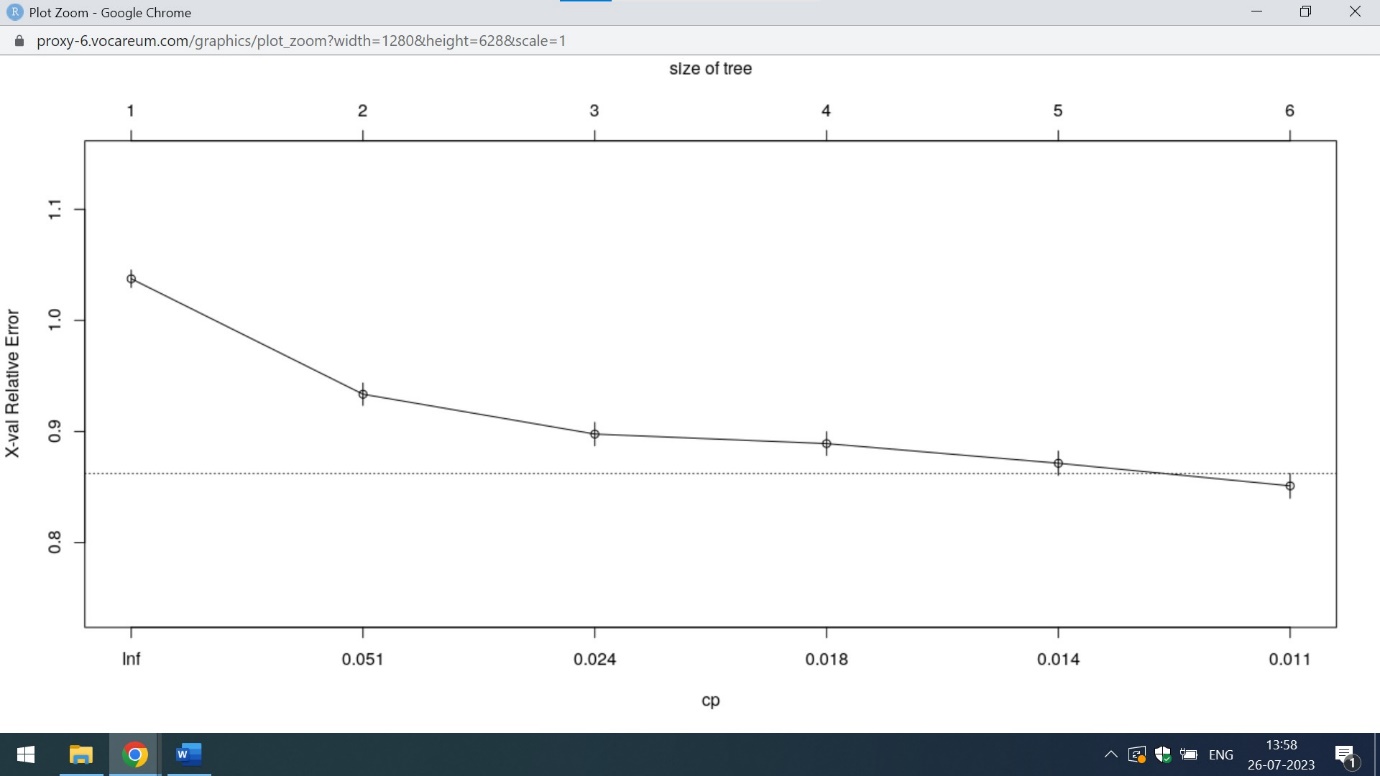




From the above attached screenshot,one can analyse that at nodenumber1,node number 2,node number 4 , 5 and 9 are some nodes reflecting the insured amount,claims and payment in a positive format

|  |
| --- |
| > printcp(insurance\_analysis)  Classification tree:  rpart(formula = Zone ~ ., data = insurance\_data, method = "class")  Variables actually used in tree construction:  [1] Bonus Claims Insured Make  Root node error: 1867/2182 = 0.85564  n= 2182  CP nsplit rel error xerror xstd  1 0.084092 0 1.00000 1.03749 0.0078991  2 0.031066 1 0.91591 0.93358 0.0100302  3 0.019282 2 0.88484 0.89770 0.0105594  4 0.016604 3 0.86556 0.88913 0.0106738  5 0.011784 4 0.84896 0.87145 0.0108960  6 0.010000 5 0.83717 0.85110 0.0111306 |
|  |
| |  | | --- | | > | |

The above attached screenshot shows the root node error which is 0.85



This tree gives an idea when the size of tree increases the relative error decreases

Q5) The committee wants to understand what affects their claim rates so as to decide the right premiums for a certain set of situations. Hence, they need to find whether the insured amount, zone, kilometre, bonus, or make affects the claim rates and to what extent.

Ans) Decision Tree Algorithm is helpful in analysing the factors for the claim rates.Below attached is the code in R

print("Insurance Factors Identification")

insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")

print(insurance\_data)

# Convert claim to factor

insurance\_data$Claims<-sapply(insurance\_data$Claims,factor)

str(insurance\_data)

# Build the model

insurance\_results<-rpart(Claims~.,data=insurance\_data,method="class")

insurance\_results

summary(insurance\_results)

printcp(insurance\_results)

plotcp(insurance\_results)

Below attached are the screenshots:

|  |
| --- |
| > print("Insurance Factors Identification")  [1] "Insurance Factors Identification"  > insurance\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Insurance/main/Insurance.csv")  > print(insurance\_data)  Kilometres Zone Bonus Make Insured Claims Payment  1 1 1 1 1 455.13 108 392491  2 1 1 1 2 69.17 19 46221  3 1 1 1 3 72.88 13 15694  4 1 1 1 4 1292.39 124 422201  5 1 1 1 5 191.01 40 119373  6 1 1 1 6 477.66 57 170913  7 1 1 1 7 105.58 23 56940  8 1 1 1 8 32.55 14 77487  9 1 1 1 9 9998.46 1704 6805992  10 1 1 2 1 314.58 45 214011  11 1 1 2 2 61.82 10 65303  12 1 1 2 3 47.06 5 20871  13 1 1 2 4 782.58 48 242894  14 1 1 2 5 115.43 11 23545  15 1 1 2 6 338.06 23 39598  16 1 1 2 7 70.44 7 48767  17 1 1 2 8 15.25 2 6560  18 1 1 2 9 6416.19 638 2873487  19 1 1 3 1 309.98 24 134931  20 1 1 3 2 49.18 6 50908  21 1 1 3 3 32.02 3 4399  22 1 1 3 4 497.20 23 112992  23 1 1 3 5 73.48 6 14788  24 1 1 3 6 278.01 9 48713  25 1 1 3 7 66.36 9 52076  26 1 1 3 8 17.86 3 13161  27 1 1 3 9 5063.15 408 1707680  28 1 1 4 1 318.48 29 103866  29 1 1 4 2 57.21 7 77588  30 1 1 4 3 35.33 4 11839  31 1 1 4 4 374.28 20 98140  32 1 1 4 5 85.18 7 27919  33 1 1 4 6 199.70 7 103910  34 1 1 4 7 60.46 4 38065  35 1 1 4 8 12.74 0 0  36 1 1 4 9 4263.09 300 1267678  37 1 1 5 1 444.37 25 69203  38 1 1 5 2 86.65 6 14620  39 1 1 5 3 53.81 5 40258  40 1 1 5 4 361.62 22 161455  41 1 1 5 5 117.91 3 20011  42 1 1 5 6 232.55 11 57214  43 1 1 5 7 81.27 3 4496  44 1 1 5 8 18.21 0 0  45 1 1 5 9 4761.37 301 1116208  46 1 1 6 1 1016.67 61 217617  47 1 1 6 2 150.56 12 58099  48 1 1 6 3 126.69 4 12268  49 1 1 6 4 517.31 16 59634  50 1 1 6 5 246.62 13 84966  51 1 1 6 6 482.96 19 137005  52 1 1 6 7 203.60 12 33767  53 1 1 6 8 25.88 3 6279  54 1 1 6 9 9197.99 522 1939894  55 1 1 7 1 5430.48 214 1048698  56 1 1 7 2 659.54 24 143915  57 1 1 7 3 657.34 22 153830  58 1 1 7 4 2795.72 60 202413  59 1 1 7 5 1119.12 41 180345  60 1 1 7 6 2861.69 92 484604  61 1 1 7 7 1111.00 37 152801  62 1 1 7 8 166.61 6 14084  63 1 1 7 9 48264.64 1875 8977527  64 1 2 1 1 458.89 98 532092  65 1 2 1 2 72.78 5 9006  66 1 2 1 3 33.23 7 45498  67 1 2 1 4 1544.55 101 337480  68 1 2 1 5 200.90 43 191982  69 1 2 1 6 663.98 65 300632  70 1 2 1 7 124.73 10 23349  71 1 2 1 8 29.24 4 13581  72 1 2 1 9 11381.00 1326 6173598  73 1 2 2 1 364.78 40 211494  74 1 2 2 2 51.89 5 10811  75 1 2 2 3 29.39 4 36204  76 1 2 2 4 1053.01 33 135007  77 1 2 2 5 110.10 16 49061  78 1 2 2 6 470.62 30 64287  79 1 2 2 7 93.29 8 51080  80 1 2 2 8 17.88 1 600  81 1 2 2 9 7607.66 591 2510207  82 1 2 3 1 315.14 17 106975  83 1 2 3 2 64.53 4 16922  84 1 2 3 3 27.24 2 8255  85 1 2 3 4 726.13 29 93656  86 1 2 3 5 96.49 4 44966  87 1 2 3 6 365.81 16 43426  88 1 2 3 7 80.80 5 48691  89 1 2 3 8 13.30 1 1325  90 1 2 3 9 5898.98 320 1392652  91 1 2 4 1 320.47 16 136143  92 1 2 4 2 69.55 4 34137  93 1 2 4 3 33.37 1 2702  94 1 2 4 4 507.57 9 22292  95 1 2 4 5 72.40 7 20295  96 1 2 4 6 316.14 9 57404  97 1 2 4 7 72.05 3 8538  98 1 2 4 8 18.35 0 0  99 1 2 4 9 4957.56 269 1375988  100 1 2 5 1 473.63 27 136376  101 1 2 5 2 88.09 8 19038  102 1 2 5 3 46.27 2 3604  103 1 2 5 4 467.96 9 10597  104 1 2 5 5 126.88 10 26433  105 1 2 5 6 316.15 11 52950  106 1 2 5 7 101.11 7 21620  107 1 2 5 8 23.37 1 2680  108 1 2 5 9 5481.31 282 1079230  109 1 2 6 1 996.27 61 236220  110 1 2 6 2 175.14 10 25036  111 1 2 6 3 111.97 5 22261  112 1 2 6 4 601.61 16 88961  113 1 2 6 5 260.69 14 64368  114 1 2 6 6 593.18 17 65578  115 1 2 6 7 229.72 7 46244  116 1 2 6 8 46.66 7 14385  117 1 2 6 9 9830.72 413 1840742  118 1 2 7 1 6021.43 233 1086534  119 1 2 7 2 852.80 33 165960  120 1 2 7 3 751.59 24 100564  121 1 2 7 4 3293.99 60 201401  122 1 2 7 5 1289.09 53 272610  123 1 2 7 6 3665.27 97 524316  124 1 2 7 7 1369.91 35 159658  125 1 2 7 8 183.98 5 18603  126 1 2 7 9 55084.54 1744 8500391  127 1 3 1 1 453.06 72 329632  128 1 3 1 2 67.13 9 79565  129 1 3 1 3 35.22 5 11746  130 1 3 1 4 1653.17 89 338305  131 1 3 1 5 206.62 38 124108  132 1 3 1 6 859.95 64 213078  133 1 3 1 7 110.08 11 34844  134 1 3 1 8 25.92 5 25319  135 1 3 1 9 11436.08 1205 5173923  136 1 3 2 1 317.10 25 90162  137 1 3 2 2 54.58 8 19327  138 1 3 2 3 22.59 3 1209  139 1 3 2 4 1187.43 30 123124  140 1 3 2 5 129.53 13 99258  141 1 3 2 6 617.11 40 137828  142 1 3 2 7 102.50 7 14904  [ reached 'max' / getOption("max.print") -- omitted 2040 rows ] |
|  |
| |  | | --- | | >  The above mentioned screenshot represents the tabular view of insurance data set with entries in all 7 columns  > summary(insurance\_results)  Call:  rpart(formula = Claims ~ ., data = insurance\_data, method = "class")  n= 2182  CP nsplit rel error xerror xstd  1 0.11908737 0 1.0000000 1.0000000 0.00990898  2 0.03171953 1 0.8809126 0.8814691 0.01159452  3 0.02003339 2 0.8491931 0.8619922 0.01179647  4 0.01001669 3 0.8291597 0.8352810 0.01204449  5 0.01000000 4 0.8191430 0.8358375 0.01203965  Variable importance  Payment Insured Zone  67 24 9  Node number 1: 2182 observations, complexity param=0.1190874  predicted class=0 expected loss=0.8235564 P(node) =1  class counts: 1 16 34 1 5 2 15 28 1 6 38 98 4 44 65 169 1 13 66 131 46 1 17 100 14 385 1 10 14 1 2 29 28 1 1 4 7 2 2 1 4 4 6 8 1 9 12 51 214 1 24 1 1 8 1 1 2 6 2 10 1 4 2 7 3 1 1 13 1 19 1 2 1 1 9 5 6 2 3 1 4 3 1 1 1 1 2 5 2 9 1 1 1 1 6 1 29 3 2 1 1 2 1 8 6 1 1 1 6 4 2 1 4 1 2 1 8 1 1 1 2 1 4 1 1 3 4 2 1 2 3 1 1 1 6 2 1 1 2 4 1 2 5 1 2 1 1 2 1 1 2 1 1 1 1 2 1 1 1 1 3 1 4 1 1 2 1 1 1 1 1 1 6 2 1 1 1 3 2 1 1 1 1 1 1 1 2 1 6 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1  probabilities: 0.000 0.007 0.016 0.000 0.002 0.001 0.007 0.013 0.000 0.003 0.017 0.045 0.002 0.020 0.030 0.077 0.000 0.006 0.030 0.060 0.021 0.000 0.008 0.046 0.006 0.176 0.000 0.005 0.006 0.000 0.001 0.013 0.013 0.000 0.000 0.002 0.003 0.001 0.001 0.000 0.002 0.002 0.003 0.004 0.000 0.004 0.005 0.023 0.098 0.000 0.011 0.000 0.000 0.004 0.000 0.000 0.001 0.003 0.001 0.005 0.000 0.002 0.001 0.003 0.001 0.000 0.000 0.006 0.000 0.009 0.000 0.001 0.000 0.000 0.004 0.002 0.003 0.001 0.001 0.000 0.002 0.001 0.000 0.000 0.000 0.000 0.001 0.002 0.001 0.004 0.000 0.000 0.000 0.000 0.003 0.000 0.013 0.001 0.001 0.000 0.000 0.001 0.000 0.004 0.003 0.000 0.000 0.000 0.003 0.002 0.001 0.000 0.002 0.000 0.001 0.000 0.004 0.000 0.000 0.000 0.001 0.000 0.002 0.000 0.000 0.001 0.002 0.001 0.000 0.001 0.001 0.000 0.000 0.000 0.003 0.001 0.000 0.000 0.001 0.002 0.000 0.001 0.002 0.000 0.001 0.000 0.000 0.001 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.001 0.000 0.002 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.003 0.001 0.000 0.000 0.000 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.003 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.001 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  left son=2 (385 obs) right son=3 (1797 obs)  Primary splits:  Payment < 36 to the left, improve=330.48330, (0 missing)  Insured < 18.405 to the left, improve=157.76770, (0 missing)  Zone < 6.5 to the right, improve= 74.63989, (0 missing)  Make < 8.5 to the left, improve= 14.34613, (0 missing)  Kilometres < 3.5 to the right, improve= 14.28534, (0 missing)  Surrogate splits:  Insured < 10.875 to the left, agree=0.904, adj=0.457, (0 split)  Zone < 6.5 to the right, agree=0.859, adj=0.203, (0 split)  Node number 2: 385 observations  predicted class=0 expected loss=0 P(node) =0.1764436  class counts: 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 385 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  probabilities: 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 1.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  Node number 3: 1797 observations, complexity param=0.03171953  predicted class=1 expected loss=0.8809126 P(node) =0.8235564  class counts: 1 16 34 1 5 2 15 28 1 6 38 98 4 44 65 169 1 13 66 131 46 1 17 100 14 0 1 10 14 1 2 29 28 1 1 4 7 2 2 1 4 4 6 8 1 9 12 51 214 1 24 1 1 8 1 1 2 6 2 10 1 4 2 7 3 1 1 13 1 19 1 2 1 1 9 5 6 2 3 1 4 3 1 1 1 1 2 5 2 9 1 1 1 1 6 1 29 3 2 1 1 2 1 8 6 1 1 1 6 4 2 1 4 1 2 1 8 1 1 1 2 1 4 1 1 3 4 2 1 2 3 1 1 1 6 2 1 1 2 4 1 2 5 1 2 1 1 2 1 1 2 1 1 1 1 2 1 1 1 1 3 1 4 1 1 2 1 1 1 1 1 1 6 2 1 1 1 3 2 1 1 1 1 1 1 1 2 1 6 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1  probabilities: 0.001 0.009 0.019 0.001 0.003 0.001 0.008 0.016 0.001 0.003 0.021 0.055 0.002 0.024 0.036 0.094 0.001 0.007 0.037 0.073 0.026 0.001 0.009 0.056 0.008 0.000 0.001 0.006 0.008 0.001 0.001 0.016 0.016 0.001 0.001 0.002 0.004 0.001 0.001 0.001 0.002 0.002 0.003 0.004 0.001 0.005 0.007 0.028 0.119 0.001 0.013 0.001 0.001 0.004 0.001 0.001 0.001 0.003 0.001 0.006 0.001 0.002 0.001 0.004 0.002 0.001 0.001 0.007 0.001 0.011 0.001 0.001 0.001 0.001 0.005 0.003 0.003 0.001 0.002 0.001 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.003 0.001 0.005 0.001 0.001 0.001 0.001 0.003 0.001 0.016 0.002 0.001 0.001 0.001 0.001 0.001 0.004 0.003 0.001 0.001 0.001 0.003 0.002 0.001 0.001 0.002 0.001 0.001 0.001 0.004 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.002 0.002 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.003 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.003 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.003 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.003 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001  left son=6 (225 obs) right son=7 (1572 obs)  Primary splits:  Payment < 4379 to the left, improve=89.935620, (0 missing)  Insured < 47.035 to the left, improve=74.469590, (0 missing)  Zone < 6.5 to the right, improve=14.227280, (0 missing)  Make < 8.5 to the left, improve= 9.340987, (0 missing)  Kilometres < 3.5 to the right, improve= 8.781400, (0 missing)  Surrogate splits:  Insured < 4.385 to the left, agree=0.881, adj=0.049, (0 split)  Node number 6: 225 observations  predicted class=1 expected loss=0.32 P(node) =0.1031164  class counts: 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 55 0 0 0 13 0 0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 153 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  probabilities: 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.004 0.000 0.000 0.000 0.244 0.000 0.000 0.000 0.058 0.000 0.000 0.000 0.013 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.680 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  Node number 7: 1572 observations, complexity param=0.02003339  predicted class=3 expected loss=0.9249364 P(node) =0.72044  class counts: 1 16 34 1 5 2 15 28 1 6 38 97 4 44 65 114 1 13 66 118 46 1 17 97 14 0 1 10 14 1 2 29 28 1 1 4 7 2 2 1 4 4 6 8 1 9 12 51 61 1 24 1 1 8 1 1 2 6 2 10 1 4 2 7 3 1 1 13 1 19 1 2 1 1 9 5 6 2 3 1 4 3 1 1 1 1 2 5 2 9 1 1 1 1 6 1 29 3 2 1 1 2 1 8 6 1 1 1 6 4 2 1 4 1 2 1 8 1 1 1 2 1 4 1 1 3 4 2 1 2 3 1 1 1 6 2 1 1 2 4 1 2 5 1 2 1 1 2 1 1 2 1 1 1 1 2 1 1 1 1 3 1 4 1 1 2 1 1 1 1 1 1 6 2 1 1 1 3 2 1 1 1 1 1 1 1 2 1 6 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1  probabilities: 0.001 0.010 0.022 0.001 0.003 0.001 0.010 0.018 0.001 0.004 0.024 0.062 0.003 0.028 0.041 0.073 0.001 0.008 0.042 0.075 0.029 0.001 0.011 0.062 0.009 0.000 0.001 0.006 0.009 0.001 0.001 0.018 0.018 0.001 0.001 0.003 0.004 0.001 0.001 0.001 0.003 0.003 0.004 0.005 0.001 0.006 0.008 0.032 0.039 0.001 0.015 0.001 0.001 0.005 0.001 0.001 0.001 0.004 0.001 0.006 0.001 0.003 0.001 0.004 0.002 0.001 0.001 0.008 0.001 0.012 0.001 0.001 0.001 0.001 0.006 0.003 0.004 0.001 0.002 0.001 0.003 0.002 0.001 0.001 0.001 0.001 0.001 0.003 0.001 0.006 0.001 0.001 0.001 0.001 0.004 0.001 0.018 0.002 0.001 0.001 0.001 0.001 0.001 0.005 0.004 0.001 0.001 0.001 0.004 0.003 0.001 0.001 0.003 0.001 0.001 0.001 0.005 0.001 0.001 0.001 0.001 0.001 0.003 0.001 0.001 0.002 0.003 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.004 0.001 0.001 0.001 0.001 0.003 0.001 0.001 0.003 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.003 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.004 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.004 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001  left son=14 (306 obs) right son=15 (1266 obs)  Primary splits:  Payment < 13583.5 to the left, improve=39.637770, (0 missing)  Insured < 103.93 to the left, improve=37.498660, (0 missing)  Make < 8.5 to the left, improve= 6.629778, (0 missing)  Bonus < 6.5 to the left, improve= 4.774978, (0 missing)  Kilometres < 3.5 to the right, improve= 4.735376, (0 missing)  Surrogate splits:  Insured < 39.155 to the left, agree=0.84, adj=0.176, (0 split)  Node number 14: 306 observations  predicted class=2 expected loss=0.7124183 P(node) =0.1402383  class counts: 0 0 0 0 0 0 0 0 0 0 0 31 0 0 8 88 0 0 12 82 2 0 0 46 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 34 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  probabilities: 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.101 0.000 0.000 0.026 0.288 0.000 0.000 0.039 0.268 0.007 0.000 0.000 0.150 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.010 0.111 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  Node number 15: 1266 observations, complexity param=0.01001669  predicted class=5 expected loss=0.9478673 P(node) =0.5802016  class counts: 1 16 34 1 5 2 15 28 1 6 38 66 4 44 57 26 1 13 54 36 44 1 17 51 14 0 1 10 14 1 2 29 28 1 1 4 7 2 2 1 4 4 6 8 1 9 12 48 27 1 24 1 1 8 1 1 2 6 2 10 1 4 2 7 3 1 1 13 1 19 1 2 1 1 9 5 6 2 3 1 4 3 1 1 1 1 2 5 2 9 1 1 1 1 6 1 29 3 2 1 1 2 1 8 6 1 1 1 6 4 2 1 4 1 2 1 8 1 1 1 2 1 4 1 1 3 4 2 1 2 3 1 1 1 6 2 1 1 2 4 1 2 5 1 2 1 1 2 1 1 2 1 1 1 1 2 1 1 1 1 3 1 4 1 1 2 1 1 1 1 1 1 6 2 1 1 1 3 2 1 1 1 1 1 1 1 2 1 6 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1  probabilities: 0.001 0.013 0.027 0.001 0.004 0.002 0.012 0.022 0.001 0.005 0.030 0.052 0.003 0.035 0.045 0.021 0.001 0.010 0.043 0.028 0.035 0.001 0.013 0.040 0.011 0.000 0.001 0.008 0.011 0.001 0.002 0.023 0.022 0.001 0.001 0.003 0.006 0.002 0.002 0.001 0.003 0.003 0.005 0.006 0.001 0.007 0.009 0.038 0.021 0.001 0.019 0.001 0.001 0.006 0.001 0.001 0.002 0.005 0.002 0.008 0.001 0.003 0.002 0.006 0.002 0.001 0.001 0.010 0.001 0.015 0.001 0.002 0.001 0.001 0.007 0.004 0.005 0.002 0.002 0.001 0.003 0.002 0.001 0.001 0.001 0.001 0.002 0.004 0.002 0.007 0.001 0.001 0.001 0.001 0.005 0.001 0.023 0.002 0.002 0.001 0.001 0.002 0.001 0.006 0.005 0.001 0.001 0.001 0.005 0.003 0.002 0.001 0.003 0.001 0.002 0.001 0.006 0.001 0.001 0.001 0.002 0.001 0.003 0.001 0.001 0.002 0.003 0.002 0.001 0.002 0.002 0.001 0.001 0.001 0.005 0.002 0.001 0.001 0.002 0.003 0.001 0.002 0.004 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.002 0.001 0.003 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.005 0.002 0.001 0.001 0.001 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.005 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001  left son=30 (457 obs) right son=31 (809 obs)  Primary splits:  Payment < 54528 to the left, improve=19.471720, (0 missing)  Insured < 229.725 to the left, improve=17.794150, (0 missing)  Make < 8.5 to the left, improve= 4.644546, (0 missing)  Bonus < 6.5 to the left, improve= 2.915686, (0 missing)  Kilometres < 3.5 to the left, improve= 1.787639, (0 missing)  Surrogate splits:  Insured < 177.345 to the left, agree=0.827, adj=0.521, (0 split)  Zone < 6.5 to the right, agree=0.648, adj=0.024, (0 split)  Node number 30: 457 observations  predicted class=5 expected loss=0.8708972 P(node) =0.2094409  class counts: 0 4 11 0 0 0 2 6 0 0 26 59 0 23 50 25 0 0 44 31 22 0 0 49 3 0 0 0 0 0 0 13 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 39 27 0 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  probabilities: 0.000 0.009 0.024 0.000 0.000 0.000 0.004 0.013 0.000 0.000 0.057 0.129 0.000 0.050 0.109 0.055 0.000 0.000 0.096 0.068 0.048 0.000 0.000 0.107 0.007 0.000 0.000 0.000 0.000 0.000 0.000 0.028 0.007 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.085 0.059 0.000 0.013 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.007 0.000 0.004 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.020 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000  Node number 31: 809 observations  predicted class=16 expected loss=0.9690977 P(node) =0.3707608  class counts: 1 12 23 1 5 2 13 22 1 6 12 7 4 21 7 1 1 13 10 5 22 1 17 2 11 0 1 10 14 1 2 16 25 1 1 4 7 2 2 1 4 4 6 8 1 9 12 9 0 1 18 1 1 8 1 1 2 6 2 10 1 4 2 7 3 1 1 10 1 17 1 2 1 1 9 5 6 2 3 1 4 3 1 1 1 1 2 5 2 9 1 1 1 1 6 1 20 3 2 1 1 2 1 8 6 1 1 1 6 4 2 1 4 1 2 1 8 1 1 1 2 1 4 1 1 3 4 2 1 2 3 1 1 1 6 2 1 1 2 4 1 2 5 1 2 1 1 2 1 1 2 1 1 1 1 2 1 1 1 1 3 1 4 1 1 2 1 1 1 1 1 1 6 2 1 1 1 3 2 1 1 1 1 1 1 1 2 1 6 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 2 2 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1  probabilities: 0.001 0.015 0.028 0.001 0.006 0.002 0.016 0.027 0.001 0.007 0.015 0.009 0.005 0.026 0.009 0.001 0.001 0.016 0.012 0.006 0.027 0.001 0.021 0.002 0.014 0.000 0.001 0.012 0.017 0.001 0.002 0.020 0.031 0.001 0.001 0.005 0.009 0.002 0.002 0.001 0.005 0.005 0.007 0.010 0.001 0.011 0.015 0.011 0.000 0.001 0.022 0.001 0.001 0.010 0.001 0.001 0.002 0.007 0.002 0.012 0.001 0.005 0.002 0.009 0.004 0.001 0.001 0.012 0.001 0.021 0.001 0.002 0.001 0.001 0.011 0.006 0.007 0.002 0.004 0.001 0.005 0.004 0.001 0.001 0.001 0.001 0.002 0.006 0.002 0.011 0.001 0.001 0.001 0.001 0.007 0.001 0.025 0.004 0.002 0.001 0.001 0.002 0.001 0.010 0.007 0.001 0.001 0.001 0.007 0.005 0.002 0.001 0.005 0.001 0.002 0.001 0.010 0.001 0.001 0.001 0.002 0.001 0.005 0.001 0.001 0.004 0.005 0.002 0.001 0.002 0.004 0.001 0.001 0.001 0.007 0.002 0.001 0.001 0.002 0.005 0.001 0.002 0.006 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.004 0.001 0.005 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.007 0.002 0.001 0.001 0.001 0.004 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.007 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.002 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.004 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.002 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001  The above screenshot represents the summary of the insurance factors identification. From the above attached screenshot,it is clear that factors affecting the claims are Payment,Insured and Zones. There are total 2182 variables of the dataset and 31 nodes. The values of the Payment,Insured,Zone,Make and Kms for each node are listed in the above mentioned screenshots | |