Assignment1

library(GGally)

## Loading required package: ggplot2

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

#install.packages("data.table")  
library(data.table)  
library(car)

## Loading required package: carData

library(rpart)  
library(chemometrics)  
#install.packages("mvoutlier")  
library(mvoutlier)

## Loading required package: sgeostat

library(sgeostat)  
library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

Preparing the data in the environment

# Clear plots  
if(!is.null(dev.list())) dev.off()

## null device   
## 1

# Clean workspace  
rm(list=ls())  
#load data  
df <- read.csv("insurance.csv")

### Data cleaning

#### Data format

is.null(df) #no nulls in the data

## [1] FALSE

replace(df,which(df %like% " "), '') #close all blank space

## age sex bmi children smoker region charges  
## 1   
## 2   
## 3   
## 4   
## 5   
## 6   
## 7   
## 8   
## 9   
## 10   
## 11   
## 12   
## 13   
## 14   
## 15   
## 16   
## 17   
## 18   
## 19   
## 20   
## 21   
## 22   
## 23   
## 24   
## 25   
## 26   
## 27   
## 28   
## 29   
## 30   
## 31   
## 32   
## 33   
## 34   
## 35   
## 36   
## 37   
## 38   
## 39   
## 40   
## 41   
## 42   
## 43   
## 44   
## 45   
## 46   
## 47   
## 48   
## 49   
## 50   
## 51   
## 52   
## 53   
## 54   
## 55   
## 56   
## 57   
## 58   
## 59   
## 60   
## 61   
## 62   
## 63   
## 64   
## 65   
## 66   
## 67   
## 68   
## 69   
## 70   
## 71   
## 72   
## 73   
## 74   
## 75   
## 76   
## 77   
## 78   
## 79   
## 80   
## 81   
## 82   
## 83   
## 84   
## 85   
## 86   
## 87   
## 88   
## 89   
## 90   
## 91   
## 92   
## 93   
## 94   
## 95   
## 96   
## 97   
## 98   
## 99   
## 100   
## 101   
## 102   
## 103   
## 104   
## 105   
## 106   
## 107   
## 108   
## 109   
## 110   
## 111   
## 112   
## 113   
## 114   
## 115   
## 116   
## 117   
## 118   
## 119   
## 120   
## 121   
## 122   
## 123   
## 124   
## 125   
## 126   
## 127   
## 128   
## 129   
## 130   
## 131   
## 132   
## 133   
## 134   
## 135   
## 136   
## 137   
## 138   
## 139   
## 140   
## 141   
## 142   
## 143   
## 144   
## 145   
## 146   
## 147   
## 148   
## 149   
## 150   
## 151   
## 152   
## 153   
## 154   
## 155   
## 156   
## 157   
## 158   
## 159   
## 160   
## 161   
## 162   
## 163   
## 164   
## 165   
## 166   
## 167   
## 168   
## 169   
## 170   
## 171   
## 172   
## 173   
## 174   
## 175   
## 176   
## 177   
## 178   
## 179   
## 180   
## 181   
## 182   
## 183   
## 184   
## 185   
## 186   
## 187   
## 188   
## 189   
## 190   
## 191   
## 192   
## 193   
## 194   
## 195   
## 196   
## 197   
## 198   
## 199   
## 200   
## 201   
## 202   
## 203   
## 204   
## 205   
## 206   
## 207   
## 208   
## 209   
## 210   
## 211   
## 212   
## 213   
## 214   
## 215   
## 216   
## 217   
## 218   
## 219   
## 220   
## 221   
## 222   
## 223   
## 224   
## 225   
## 226   
## 227   
## 228   
## 229   
## 230   
## 231   
## 232   
## 233   
## 234   
## 235   
## 236   
## 237   
## 238   
## 239   
## 240   
## 241   
## 242   
## 243   
## 244   
## 245   
## 246   
## 247   
## 248   
## 249   
## 250   
## 251   
## 252   
## 253   
## 254   
## 255   
## 256   
## 257   
## 258   
## 259   
## 260   
## 261   
## 262   
## 263   
## 264   
## 265   
## 266   
## 267   
## 268   
## 269   
## 270   
## 271   
## 272   
## 273   
## 274   
## 275   
## 276   
## 277   
## 278   
## 279   
## 280   
## 281   
## 282   
## 283   
## 284   
## 285   
## 286   
## 287   
## 288   
## 289   
## 290   
## 291   
## 292   
## 293   
## 294   
## 295   
## 296   
## 297   
## 298   
## 299   
## 300   
## 301   
## 302   
## 303   
## 304   
## 305   
## 306   
## 307   
## 308   
## 309   
## 310   
## 311   
## 312   
## 313   
## 314   
## 315   
## 316   
## 317   
## 318   
## 319   
## 320   
## 321   
## 322   
## 323   
## 324   
## 325   
## 326   
## 327   
## 328   
## 329   
## 330   
## 331   
## 332   
## 333   
## 334   
## 335   
## 336   
## 337   
## 338   
## 339   
## 340   
## 341   
## 342   
## 343   
## 344   
## 345   
## 346   
## 347   
## 348   
## 349   
## 350   
## 351   
## 352   
## 353   
## 354   
## 355   
## 356   
## 357   
## 358   
## 359   
## 360   
## 361   
## 362   
## 363   
## 364   
## 365   
## 366   
## 367   
## 368   
## 369   
## 370   
## 371   
## 372   
## 373   
## 374   
## 375   
## 376   
## 377   
## 378   
## 379   
## 380   
## 381   
## 382   
## 383   
## 384   
## 385   
## 386   
## 387   
## 388   
## 389   
## 390   
## 391   
## 392   
## 393   
## 394   
## 395   
## 396   
## 397   
## 398   
## 399   
## 400   
## 401   
## 402   
## 403   
## 404   
## 405   
## 406   
## 407   
## 408   
## 409   
## 410   
## 411   
## 412   
## 413   
## 414   
## 415   
## 416   
## 417   
## 418   
## 419   
## 420   
## 421   
## 422   
## 423   
## 424   
## 425   
## 426   
## 427   
## 428   
## 429   
## 430   
## 431   
## 432   
## 433   
## 434   
## 435   
## 436   
## 437   
## 438   
## 439   
## 440   
## 441   
## 442   
## 443   
## 444   
## 445   
## 446   
## 447   
## 448   
## 449   
## 450   
## 451   
## 452   
## 453   
## 454   
## 455   
## 456   
## 457   
## 458   
## 459   
## 460   
## 461   
## 462   
## 463   
## 464   
## 465   
## 466   
## 467   
## 468   
## 469   
## 470   
## 471   
## 472   
## 473   
## 474   
## 475   
## 476   
## 477   
## 478   
## 479   
## 480   
## 481   
## 482   
## 483   
## 484   
## 485   
## 486   
## 487   
## 488   
## 489   
## 490   
## 491   
## 492   
## 493   
## 494   
## 495   
## 496   
## 497   
## 498   
## 499   
## 500   
## 501   
## 502   
## 503   
## 504   
## 505   
## 506   
## 507   
## 508   
## 509   
## 510   
## 511   
## 512   
## 513   
## 514   
## 515   
## 516   
## 517   
## 518   
## 519   
## 520   
## 521   
## 522   
## 523   
## 524   
## 525   
## 526   
## 527   
## 528   
## 529   
## 530   
## 531   
## 532   
## 533   
## 534   
## 535   
## 536   
## 537   
## 538   
## 539   
## 540   
## 541   
## 542   
## 543   
## 544   
## 545   
## 546   
## 547   
## 548   
## 549   
## 550   
## 551   
## 552   
## 553   
## 554   
## 555   
## 556   
## 557   
## 558   
## 559   
## 560   
## 561   
## 562   
## 563   
## 564   
## 565   
## 566   
## 567   
## 568   
## 569   
## 570   
## 571   
## 572   
## 573   
## 574   
## 575   
## 576   
## 577   
## 578   
## 579   
## 580   
## 581   
## 582   
## 583   
## 584   
## 585   
## 586   
## 587   
## 588   
## 589   
## 590   
## 591   
## 592   
## 593   
## 594   
## 595   
## 596   
## 597   
## 598   
## 599   
## 600   
## 601   
## 602   
## 603   
## 604   
## 605   
## 606   
## 607   
## 608   
## 609   
## 610   
## 611   
## 612   
## 613   
## 614   
## 615   
## 616   
## 617   
## 618   
## 619   
## 620   
## 621   
## 622   
## 623   
## 624   
## 625   
## 626   
## 627   
## 628   
## 629   
## 630   
## 631   
## 632   
## 633   
## 634   
## 635   
## 636   
## 637   
## 638   
## 639   
## 640   
## 641   
## 642   
## 643   
## 644   
## 645   
## 646   
## 647   
## 648   
## 649   
## 650   
## 651   
## 652   
## 653   
## 654   
## 655   
## 656   
## 657   
## 658   
## 659   
## 660   
## 661   
## 662   
## 663   
## 664   
## 665   
## 666   
## 667   
## 668   
## 669   
## 670   
## 671   
## 672   
## 673   
## 674   
## 675   
## 676   
## 677   
## 678   
## 679   
## 680   
## 681   
## 682   
## 683   
## 684   
## 685   
## 686   
## 687   
## 688   
## 689   
## 690   
## 691   
## 692   
## 693   
## 694   
## 695   
## 696   
## 697   
## 698   
## 699   
## 700   
## 701   
## 702   
## 703   
## 704   
## 705   
## 706   
## 707   
## 708   
## 709   
## 710   
## 711   
## 712   
## 713   
## 714   
## 715   
## 716   
## 717   
## 718   
## 719   
## 720   
## 721   
## 722   
## 723   
## 724   
## 725   
## 726   
## 727   
## 728   
## 729   
## 730   
## 731   
## 732   
## 733   
## 734   
## 735   
## 736   
## 737   
## 738   
## 739   
## 740   
## 741   
## 742   
## 743   
## 744   
## 745   
## 746   
## 747   
## 748   
## 749   
## 750   
## 751   
## 752   
## 753   
## 754   
## 755   
## 756   
## 757   
## 758   
## 759   
## 760   
## 761   
## 762   
## 763   
## 764   
## 765   
## 766   
## 767   
## 768   
## 769   
## 770   
## 771   
## 772   
## 773   
## 774   
## 775   
## 776   
## 777   
## 778   
## 779   
## 780   
## 781   
## 782   
## 783   
## 784   
## 785   
## 786   
## 787   
## 788   
## 789   
## 790   
## 791   
## 792   
## 793   
## 794   
## 795   
## 796   
## 797   
## 798   
## 799   
## 800   
## 801   
## 802   
## 803   
## 804   
## 805   
## 806   
## 807   
## 808   
## 809   
## 810   
## 811   
## 812   
## 813   
## 814   
## 815   
## 816   
## 817   
## 818   
## 819   
## 820   
## 821   
## 822   
## 823   
## 824   
## 825   
## 826   
## 827   
## 828   
## 829   
## 830   
## 831   
## 832   
## 833   
## 834   
## 835   
## 836   
## 837   
## 838   
## 839   
## 840   
## 841   
## 842   
## 843   
## 844   
## 845   
## 846   
## 847   
## 848   
## 849   
## 850   
## 851   
## 852   
## 853   
## 854   
## 855   
## 856   
## 857   
## 858   
## 859   
## 860   
## 861   
## 862   
## 863   
## 864   
## 865   
## 866   
## 867   
## 868   
## 869   
## 870   
## 871   
## 872   
## 873   
## 874   
## 875   
## 876   
## 877   
## 878   
## 879   
## 880   
## 881   
## 882   
## 883   
## 884   
## 885   
## 886   
## 887   
## 888   
## 889   
## 890   
## 891   
## 892   
## 893   
## 894   
## 895   
## 896   
## 897   
## 898   
## 899   
## 900   
## 901   
## 902   
## 903   
## 904   
## 905   
## 906   
## 907   
## 908   
## 909   
## 910   
## 911   
## 912   
## 913   
## 914   
## 915   
## 916   
## 917   
## 918   
## 919   
## 920   
## 921   
## 922   
## 923   
## 924   
## 925   
## 926   
## 927   
## 928   
## 929   
## 930   
## 931   
## 932   
## 933   
## 934   
## 935   
## 936   
## 937   
## 938   
## 939   
## 940   
## 941   
## 942   
## 943   
## 944   
## 945   
## 946   
## 947   
## 948   
## 949   
## 950   
## 951   
## 952   
## 953   
## 954   
## 955   
## 956   
## 957   
## 958   
## 959   
## 960   
## 961   
## 962   
## 963   
## 964   
## 965   
## 966   
## 967   
## 968   
## 969   
## 970   
## 971   
## 972   
## 973   
## 974   
## 975   
## 976   
## 977   
## 978   
## 979   
## 980   
## 981   
## 982   
## 983   
## 984   
## 985   
## 986   
## 987   
## 988   
## 989   
## 990   
## 991   
## 992   
## 993   
## 994   
## 995   
## 996   
## 997   
## 998   
## 999   
## 1000   
## 1001   
## 1002   
## 1003   
## 1004   
## 1005   
## 1006   
## 1007   
## 1008   
## 1009   
## 1010   
## 1011   
## 1012   
## 1013   
## 1014   
## 1015   
## 1016   
## 1017   
## 1018   
## 1019   
## 1020   
## 1021   
## 1022   
## 1023   
## 1024   
## 1025   
## 1026   
## 1027   
## 1028   
## 1029   
## 1030   
## 1031   
## 1032   
## 1033   
## 1034   
## 1035   
## 1036   
## 1037   
## 1038   
## 1039   
## 1040   
## 1041   
## 1042   
## 1043   
## 1044   
## 1045   
## 1046   
## 1047   
## 1048   
## 1049   
## 1050   
## 1051   
## 1052   
## 1053   
## 1054   
## 1055   
## 1056   
## 1057   
## 1058   
## 1059   
## 1060   
## 1061   
## 1062   
## 1063   
## 1064   
## 1065   
## 1066   
## 1067   
## 1068   
## 1069   
## 1070   
## 1071   
## 1072   
## 1073   
## 1074   
## 1075   
## 1076   
## 1077   
## 1078   
## 1079   
## 1080   
## 1081   
## 1082   
## 1083   
## 1084   
## 1085   
## 1086   
## 1087   
## 1088   
## 1089   
## 1090   
## 1091   
## 1092   
## 1093   
## 1094   
## 1095   
## 1096   
## 1097   
## 1098   
## 1099   
## 1100   
## 1101   
## 1102   
## 1103   
## 1104   
## 1105   
## 1106   
## 1107   
## 1108   
## 1109   
## 1110   
## 1111   
## 1112   
## 1113   
## 1114   
## 1115   
## 1116   
## 1117   
## 1118   
## 1119   
## 1120   
## 1121   
## 1122   
## 1123   
## 1124   
## 1125   
## 1126   
## 1127   
## 1128   
## 1129   
## 1130   
## 1131   
## 1132   
## 1133   
## 1134   
## 1135   
## 1136   
## 1137   
## 1138   
## 1139   
## 1140   
## 1141   
## 1142   
## 1143   
## 1144   
## 1145   
## 1146   
## 1147   
## 1148   
## 1149   
## 1150   
## 1151   
## 1152   
## 1153   
## 1154   
## 1155   
## 1156   
## 1157   
## 1158   
## 1159   
## 1160   
## 1161   
## 1162   
## 1163   
## 1164   
## 1165   
## 1166   
## 1167   
## 1168   
## 1169   
## 1170   
## 1171   
## 1172   
## 1173   
## 1174   
## 1175   
## 1176   
## 1177   
## 1178   
## 1179   
## 1180   
## 1181   
## 1182   
## 1183   
## 1184   
## 1185   
## 1186   
## 1187   
## 1188   
## 1189   
## 1190   
## 1191   
## 1192   
## 1193   
## 1194   
## 1195   
## 1196   
## 1197   
## 1198   
## 1199   
## 1200   
## 1201   
## 1202   
## 1203   
## 1204   
## 1205   
## 1206   
## 1207   
## 1208   
## 1209   
## 1210   
## 1211   
## 1212   
## 1213   
## 1214   
## 1215   
## 1216   
## 1217   
## 1218   
## 1219   
## 1220   
## 1221   
## 1222   
## 1223   
## 1224   
## 1225   
## 1226   
## 1227   
## 1228   
## 1229   
## 1230   
## 1231   
## 1232   
## 1233   
## 1234   
## 1235   
## 1236   
## 1237   
## 1238   
## 1239   
## 1240   
## 1241   
## 1242   
## 1243   
## 1244   
## 1245   
## 1246   
## 1247   
## 1248   
## 1249   
## 1250   
## 1251   
## 1252   
## 1253   
## 1254   
## 1255   
## 1256   
## 1257   
## 1258   
## 1259   
## 1260   
## 1261   
## 1262   
## 1263   
## 1264   
## 1265   
## 1266   
## 1267   
## 1268   
## 1269   
## 1270   
## 1271   
## 1272   
## 1273   
## 1274   
## 1275   
## 1276   
## 1277   
## 1278   
## 1279   
## 1280   
## 1281   
## 1282   
## 1283   
## 1284   
## 1285   
## 1286   
## 1287   
## 1288   
## 1289   
## 1290   
## 1291   
## 1292   
## 1293   
## 1294   
## 1295   
## 1296   
## 1297   
## 1298   
## 1299   
## 1300   
## 1301   
## 1302   
## 1303   
## 1304   
## 1305   
## 1306   
## 1307   
## 1308   
## 1309   
## 1310   
## 1311   
## 1312   
## 1313   
## 1314   
## 1315   
## 1316   
## 1317   
## 1318   
## 1319   
## 1320   
## 1321   
## 1322   
## 1323   
## 1324   
## 1325   
## 1326   
## 1327   
## 1328   
## 1329   
## 1330   
## 1331   
## 1332   
## 1333   
## 1334   
## 1335   
## 1336   
## 1337   
## 1338

which(df=="") #no blanks found in the data

## integer(0)

#check for distinct values and whether there are differences in them  
unique(df$sex) #expecting 2 values

## [1] "female" "male"

unique(df$smoker) #expecting 2 values

## [1] "yes" "no"

unique(df$region) #expecting 4 values

## [1] "southwest" "southeast" "northwest" "northeast"

#we can see that data is consistent for categorical variables  
df$f.sex <- factor(df$sex,labels = c("female","male"));  
df$f.smoker <- factor(df$smoker,labels = c("no","yes"))  
df$f.region <- factor(df$region,labels = c("northeast","northwest","southeast","southwest"))  
summary(df) #from the summary we can see the factor values, it seems that sex and region are distributed equally and not much smokers compare to the non smokers.

## age sex bmi children   
## Min. :18.00 Length:1338 Min. :15.96 Min. :0.000   
## 1st Qu.:27.00 Class :character 1st Qu.:26.30 1st Qu.:0.000   
## Median :39.00 Mode :character Median :30.40 Median :1.000   
## Mean :39.21 Mean :30.66 Mean :1.095   
## 3rd Qu.:51.00 3rd Qu.:34.69 3rd Qu.:2.000   
## Max. :64.00 Max. :53.13 Max. :5.000   
## smoker region charges f.sex f.smoker   
## Length:1338 Length:1338 Min. : 1122 female:662 no :1064   
## Class :character Class :character 1st Qu.: 4740 male :676 yes: 274   
## Mode :character Mode :character Median : 9382   
## Mean :13270   
## 3rd Qu.:16640   
## Max. :63770   
## f.region   
## northeast:324   
## northwest:325   
## southeast:364   
## southwest:325   
##   
##

dim(df)

## [1] 1338 10

unique(df)

## age sex bmi children smoker region charges f.sex f.smoker  
## 1 19 female 27.900 0 yes southwest 16884.924 female yes  
## 2 18 male 33.770 1 no southeast 1725.552 male no  
## 3 28 male 33.000 3 no southeast 4449.462 male no  
## 4 33 male 22.705 0 no northwest 21984.471 male no  
## 5 32 male 28.880 0 no northwest 3866.855 male no  
## 6 31 female 25.740 0 no southeast 3756.622 female no  
## 7 46 female 33.440 1 no southeast 8240.590 female no  
## 8 37 female 27.740 3 no northwest 7281.506 female no  
## 9 37 male 29.830 2 no northeast 6406.411 male no  
## 10 60 female 25.840 0 no northwest 28923.137 female no  
## 11 25 male 26.220 0 no northeast 2721.321 male no  
## 12 62 female 26.290 0 yes southeast 27808.725 female yes  
## 13 23 male 34.400 0 no southwest 1826.843 male no  
## 14 56 female 39.820 0 no southeast 11090.718 female no  
## 15 27 male 42.130 0 yes southeast 39611.758 male yes  
## 16 19 male 24.600 1 no southwest 1837.237 male no  
## 17 52 female 30.780 1 no northeast 10797.336 female no  
## 18 23 male 23.845 0 no northeast 2395.172 male no  
## 19 56 male 40.300 0 no southwest 10602.385 male no  
## 20 30 male 35.300 0 yes southwest 36837.467 male yes  
## 21 60 female 36.005 0 no northeast 13228.847 female no  
## 22 30 female 32.400 1 no southwest 4149.736 female no  
## 23 18 male 34.100 0 no southeast 1137.011 male no  
## 24 34 female 31.920 1 yes northeast 37701.877 female yes  
## 25 37 male 28.025 2 no northwest 6203.902 male no  
## 26 59 female 27.720 3 no southeast 14001.134 female no  
## 27 63 female 23.085 0 no northeast 14451.835 female no  
## 28 55 female 32.775 2 no northwest 12268.632 female no  
## 29 23 male 17.385 1 no northwest 2775.192 male no  
## 30 31 male 36.300 2 yes southwest 38711.000 male yes  
## 31 22 male 35.600 0 yes southwest 35585.576 male yes  
## 32 18 female 26.315 0 no northeast 2198.190 female no  
## 33 19 female 28.600 5 no southwest 4687.797 female no  
## 34 63 male 28.310 0 no northwest 13770.098 male no  
## 35 28 male 36.400 1 yes southwest 51194.559 male yes  
## 36 19 male 20.425 0 no northwest 1625.434 male no  
## 37 62 female 32.965 3 no northwest 15612.193 female no  
## 38 26 male 20.800 0 no southwest 2302.300 male no  
## 39 35 male 36.670 1 yes northeast 39774.276 male yes  
## 40 60 male 39.900 0 yes southwest 48173.361 male yes  
## 41 24 female 26.600 0 no northeast 3046.062 female no  
## 42 31 female 36.630 2 no southeast 4949.759 female no  
## 43 41 male 21.780 1 no southeast 6272.477 male no  
## 44 37 female 30.800 2 no southeast 6313.759 female no  
## 45 38 male 37.050 1 no northeast 6079.672 male no  
## 46 55 male 37.300 0 no southwest 20630.284 male no  
## 47 18 female 38.665 2 no northeast 3393.356 female no  
## 48 28 female 34.770 0 no northwest 3556.922 female no  
## 49 60 female 24.530 0 no southeast 12629.897 female no  
## 50 36 male 35.200 1 yes southeast 38709.176 male yes  
## 51 18 female 35.625 0 no northeast 2211.131 female no  
## 52 21 female 33.630 2 no northwest 3579.829 female no  
## 53 48 male 28.000 1 yes southwest 23568.272 male yes  
## 54 36 male 34.430 0 yes southeast 37742.576 male yes  
## 55 40 female 28.690 3 no northwest 8059.679 female no  
## 56 58 male 36.955 2 yes northwest 47496.494 male yes  
## 57 58 female 31.825 2 no northeast 13607.369 female no  
## 58 18 male 31.680 2 yes southeast 34303.167 male yes  
## 59 53 female 22.880 1 yes southeast 23244.790 female yes  
## 60 34 female 37.335 2 no northwest 5989.524 female no  
## 61 43 male 27.360 3 no northeast 8606.217 male no  
## 62 25 male 33.660 4 no southeast 4504.662 male no  
## 63 64 male 24.700 1 no northwest 30166.618 male no  
## 64 28 female 25.935 1 no northwest 4133.642 female no  
## 65 20 female 22.420 0 yes northwest 14711.744 female yes  
## 66 19 female 28.900 0 no southwest 1743.214 female no  
## 67 61 female 39.100 2 no southwest 14235.072 female no  
## 68 40 male 26.315 1 no northwest 6389.378 male no  
## 69 40 female 36.190 0 no southeast 5920.104 female no  
## 70 28 male 23.980 3 yes southeast 17663.144 male yes  
## 71 27 female 24.750 0 yes southeast 16577.780 female yes  
## 72 31 male 28.500 5 no northeast 6799.458 male no  
## 73 53 female 28.100 3 no southwest 11741.726 female no  
## 74 58 male 32.010 1 no southeast 11946.626 male no  
## 75 44 male 27.400 2 no southwest 7726.854 male no  
## 76 57 male 34.010 0 no northwest 11356.661 male no  
## 77 29 female 29.590 1 no southeast 3947.413 female no  
## 78 21 male 35.530 0 no southeast 1532.470 male no  
## 79 22 female 39.805 0 no northeast 2755.021 female no  
## 80 41 female 32.965 0 no northwest 6571.024 female no  
## 81 31 male 26.885 1 no northeast 4441.213 male no  
## 82 45 female 38.285 0 no northeast 7935.291 female no  
## 83 22 male 37.620 1 yes southeast 37165.164 male yes  
## 84 48 female 41.230 4 no northwest 11033.662 female no  
## 85 37 female 34.800 2 yes southwest 39836.519 female yes  
## 86 45 male 22.895 2 yes northwest 21098.554 male yes  
## 87 57 female 31.160 0 yes northwest 43578.939 female yes  
## 88 56 female 27.200 0 no southwest 11073.176 female no  
## 89 46 female 27.740 0 no northwest 8026.667 female no  
## 90 55 female 26.980 0 no northwest 11082.577 female no  
## 91 21 female 39.490 0 no southeast 2026.974 female no  
## 92 53 female 24.795 1 no northwest 10942.132 female no  
## 93 59 male 29.830 3 yes northeast 30184.937 male yes  
## 94 35 male 34.770 2 no northwest 5729.005 male no  
## 95 64 female 31.300 2 yes southwest 47291.055 female yes  
## 96 28 female 37.620 1 no southeast 3766.884 female no  
## 97 54 female 30.800 3 no southwest 12105.320 female no  
## 98 55 male 38.280 0 no southeast 10226.284 male no  
## 99 56 male 19.950 0 yes northeast 22412.648 male yes  
## 100 38 male 19.300 0 yes southwest 15820.699 male yes  
## 101 41 female 31.600 0 no southwest 6186.127 female no  
## 102 30 male 25.460 0 no northeast 3645.089 male no  
## 103 18 female 30.115 0 no northeast 21344.847 female no  
## 104 61 female 29.920 3 yes southeast 30942.192 female yes  
## 105 34 female 27.500 1 no southwest 5003.853 female no  
## 106 20 male 28.025 1 yes northwest 17560.380 male yes  
## 107 19 female 28.400 1 no southwest 2331.519 female no  
## 108 26 male 30.875 2 no northwest 3877.304 male no  
## 109 29 male 27.940 0 no southeast 2867.120 male no  
## 110 63 male 35.090 0 yes southeast 47055.532 male yes  
## 111 54 male 33.630 1 no northwest 10825.254 male no  
## 112 55 female 29.700 2 no southwest 11881.358 female no  
## 113 37 male 30.800 0 no southwest 4646.759 male no  
## 114 21 female 35.720 0 no northwest 2404.734 female no  
## 115 52 male 32.205 3 no northeast 11488.317 male no  
## 116 60 male 28.595 0 no northeast 30259.996 male no  
## 117 58 male 49.060 0 no southeast 11381.325 male no  
## 118 29 female 27.940 1 yes southeast 19107.780 female yes  
## 119 49 female 27.170 0 no southeast 8601.329 female no  
## 120 37 female 23.370 2 no northwest 6686.431 female no  
## 121 44 male 37.100 2 no southwest 7740.337 male no  
## 122 18 male 23.750 0 no northeast 1705.624 male no  
## 123 20 female 28.975 0 no northwest 2257.475 female no  
## 124 44 male 31.350 1 yes northeast 39556.495 male yes  
## 125 47 female 33.915 3 no northwest 10115.009 female no  
## 126 26 female 28.785 0 no northeast 3385.399 female no  
## 127 19 female 28.300 0 yes southwest 17081.080 female yes  
## 128 52 female 37.400 0 no southwest 9634.538 female no  
## 129 32 female 17.765 2 yes northwest 32734.186 female yes  
## 130 38 male 34.700 2 no southwest 6082.405 male no  
## 131 59 female 26.505 0 no northeast 12815.445 female no  
## 132 61 female 22.040 0 no northeast 13616.359 female no  
## 133 53 female 35.900 2 no southwest 11163.568 female no  
## 134 19 male 25.555 0 no northwest 1632.564 male no  
## 135 20 female 28.785 0 no northeast 2457.211 female no  
## 136 22 female 28.050 0 no southeast 2155.682 female no  
## 137 19 male 34.100 0 no southwest 1261.442 male no  
## 138 22 male 25.175 0 no northwest 2045.685 male no  
## 139 54 female 31.900 3 no southeast 27322.734 female no  
## 140 22 female 36.000 0 no southwest 2166.732 female no  
## 141 34 male 22.420 2 no northeast 27375.905 male no  
## 142 26 male 32.490 1 no northeast 3490.549 male no  
## 143 34 male 25.300 2 yes southeast 18972.495 male yes  
## 144 29 male 29.735 2 no northwest 18157.876 male no  
## 145 30 male 28.690 3 yes northwest 20745.989 male yes  
## 146 29 female 38.830 3 no southeast 5138.257 female no  
## 147 46 male 30.495 3 yes northwest 40720.551 male yes  
## 148 51 female 37.730 1 no southeast 9877.608 female no  
## 149 53 female 37.430 1 no northwest 10959.695 female no  
## 150 19 male 28.400 1 no southwest 1842.519 male no  
## 151 35 male 24.130 1 no northwest 5125.216 male no  
## 152 48 male 29.700 0 no southeast 7789.635 male no  
## 153 32 female 37.145 3 no northeast 6334.344 female no  
## 154 42 female 23.370 0 yes northeast 19964.746 female yes  
## 155 40 female 25.460 1 no northeast 7077.189 female no  
## 156 44 male 39.520 0 no northwest 6948.701 male no  
## 157 48 male 24.420 0 yes southeast 21223.676 male yes  
## 158 18 male 25.175 0 yes northeast 15518.180 male yes  
## 159 30 male 35.530 0 yes southeast 36950.257 male yes  
## 160 50 female 27.830 3 no southeast 19749.383 female no  
## 161 42 female 26.600 0 yes northwest 21348.706 female yes  
## 162 18 female 36.850 0 yes southeast 36149.484 female yes  
## 163 54 male 39.600 1 no southwest 10450.552 male no  
## 164 32 female 29.800 2 no southwest 5152.134 female no  
## 165 37 male 29.640 0 no northwest 5028.147 male no  
## 166 47 male 28.215 4 no northeast 10407.086 male no  
## 167 20 female 37.000 5 no southwest 4830.630 female no  
## 168 32 female 33.155 3 no northwest 6128.797 female no  
## 169 19 female 31.825 1 no northwest 2719.280 female no  
## 170 27 male 18.905 3 no northeast 4827.905 male no  
## 171 63 male 41.470 0 no southeast 13405.390 male no  
## 172 49 male 30.300 0 no southwest 8116.680 male no  
## 173 18 male 15.960 0 no northeast 1694.796 male no  
## 174 35 female 34.800 1 no southwest 5246.047 female no  
## 175 24 female 33.345 0 no northwest 2855.438 female no  
## 176 63 female 37.700 0 yes southwest 48824.450 female yes  
## 177 38 male 27.835 2 no northwest 6455.863 male no  
## 178 54 male 29.200 1 no southwest 10436.096 male no  
## 179 46 female 28.900 2 no southwest 8823.279 female no  
## 180 41 female 33.155 3 no northeast 8538.288 female no  
## 181 58 male 28.595 0 no northwest 11735.879 male no  
## 182 18 female 38.280 0 no southeast 1631.821 female no  
## 183 22 male 19.950 3 no northeast 4005.423 male no  
## 184 44 female 26.410 0 no northwest 7419.478 female no  
## 185 44 male 30.690 2 no southeast 7731.427 male no  
## 186 36 male 41.895 3 yes northeast 43753.337 male yes  
## 187 26 female 29.920 2 no southeast 3981.977 female no  
## 188 30 female 30.900 3 no southwest 5325.651 female no  
## 189 41 female 32.200 1 no southwest 6775.961 female no  
## 190 29 female 32.110 2 no northwest 4922.916 female no  
## 191 61 male 31.570 0 no southeast 12557.605 male no  
## 192 36 female 26.200 0 no southwest 4883.866 female no  
## 193 25 male 25.740 0 no southeast 2137.654 male no  
## 194 56 female 26.600 1 no northwest 12044.342 female no  
## 195 18 male 34.430 0 no southeast 1137.470 male no  
## 196 19 male 30.590 0 no northwest 1639.563 male no  
## 197 39 female 32.800 0 no southwest 5649.715 female no  
## 198 45 female 28.600 2 no southeast 8516.829 female no  
## 199 51 female 18.050 0 no northwest 9644.253 female no  
## 200 64 female 39.330 0 no northeast 14901.517 female no  
## 201 19 female 32.110 0 no northwest 2130.676 female no  
## 202 48 female 32.230 1 no southeast 8871.152 female no  
## 203 60 female 24.035 0 no northwest 13012.209 female no  
## 204 27 female 36.080 0 yes southeast 37133.898 female yes  
## 205 46 male 22.300 0 no southwest 7147.105 male no  
## 206 28 female 28.880 1 no northeast 4337.735 female no  
## 207 59 male 26.400 0 no southeast 11743.299 male no  
## 208 35 male 27.740 2 yes northeast 20984.094 male yes  
## 209 63 female 31.800 0 no southwest 13880.949 female no  
## 210 40 male 41.230 1 no northeast 6610.110 male no  
## 211 20 male 33.000 1 no southwest 1980.070 male no  
## 212 40 male 30.875 4 no northwest 8162.716 male no  
## 213 24 male 28.500 2 no northwest 3537.703 male no  
## 214 34 female 26.730 1 no southeast 5002.783 female no  
## 215 45 female 30.900 2 no southwest 8520.026 female no  
## 216 41 female 37.100 2 no southwest 7371.772 female no  
## 217 53 female 26.600 0 no northwest 10355.641 female no  
## 218 27 male 23.100 0 no southeast 2483.736 male no  
## 219 26 female 29.920 1 no southeast 3392.977 female no  
## 220 24 female 23.210 0 no southeast 25081.768 female no  
## 221 34 female 33.700 1 no southwest 5012.471 female no  
## 222 53 female 33.250 0 no northeast 10564.885 female no  
## 223 32 male 30.800 3 no southwest 5253.524 male no  
## 224 19 male 34.800 0 yes southwest 34779.615 male yes  
## 225 42 male 24.640 0 yes southeast 19515.542 male yes  
## 226 55 male 33.880 3 no southeast 11987.168 male no  
## 227 28 male 38.060 0 no southeast 2689.495 male no  
## 228 58 female 41.910 0 no southeast 24227.337 female no  
## 229 41 female 31.635 1 no northeast 7358.176 female no  
## 230 47 male 25.460 2 no northeast 9225.256 male no  
## 231 42 female 36.195 1 no northwest 7443.643 female no  
## 232 59 female 27.830 3 no southeast 14001.287 female no  
## 233 19 female 17.800 0 no southwest 1727.785 female no  
## 234 59 male 27.500 1 no southwest 12333.828 male no  
## 235 39 male 24.510 2 no northwest 6710.192 male no  
## 236 40 female 22.220 2 yes southeast 19444.266 female yes  
## 237 18 female 26.730 0 no southeast 1615.767 female no  
## 238 31 male 38.390 2 no southeast 4463.205 male no  
## 239 19 male 29.070 0 yes northwest 17352.680 male yes  
## 240 44 male 38.060 1 no southeast 7152.671 male no  
## 241 23 female 36.670 2 yes northeast 38511.628 female yes  
## 242 33 female 22.135 1 no northeast 5354.075 female no  
## 243 55 female 26.800 1 no southwest 35160.135 female no  
## 244 40 male 35.300 3 no southwest 7196.867 male no  
## 245 63 female 27.740 0 yes northeast 29523.166 female yes  
## 246 54 male 30.020 0 no northwest 24476.479 male no  
## 247 60 female 38.060 0 no southeast 12648.703 female no  
## 248 24 male 35.860 0 no southeast 1986.933 male no  
## 249 19 male 20.900 1 no southwest 1832.094 male no  
## 250 29 male 28.975 1 no northeast 4040.558 male no  
## 251 18 male 17.290 2 yes northeast 12829.455 male yes  
## 252 63 female 32.200 2 yes southwest 47305.305 female yes  
## 253 54 male 34.210 2 yes southeast 44260.750 male yes  
## 254 27 male 30.300 3 no southwest 4260.744 male no  
## 255 50 male 31.825 0 yes northeast 41097.162 male yes  
## 256 55 female 25.365 3 no northeast 13047.332 female no  
## 257 56 male 33.630 0 yes northwest 43921.184 male yes  
## 258 38 female 40.150 0 no southeast 5400.980 female no  
## 259 51 male 24.415 4 no northwest 11520.100 male no  
## 260 19 male 31.920 0 yes northwest 33750.292 male yes  
## 261 58 female 25.200 0 no southwest 11837.160 female no  
## 262 20 female 26.840 1 yes southeast 17085.268 female yes  
## 263 52 male 24.320 3 yes northeast 24869.837 male yes  
## 264 19 male 36.955 0 yes northwest 36219.405 male yes  
## 265 53 female 38.060 3 no southeast 20462.998 female no  
## 266 46 male 42.350 3 yes southeast 46151.124 male yes  
## 267 40 male 19.800 1 yes southeast 17179.522 male yes  
## 268 59 female 32.395 3 no northeast 14590.632 female no  
## 269 45 male 30.200 1 no southwest 7441.053 male no  
## 270 49 male 25.840 1 no northeast 9282.481 male no  
## 271 18 male 29.370 1 no southeast 1719.436 male no  
## 272 50 male 34.200 2 yes southwest 42856.838 male yes  
## 273 41 male 37.050 2 no northwest 7265.703 male no  
## 274 50 male 27.455 1 no northeast 9617.662 male no  
## 275 25 male 27.550 0 no northwest 2523.169 male no  
## 276 47 female 26.600 2 no northeast 9715.841 female no  
## 277 19 male 20.615 2 no northwest 2803.698 male no  
## 278 22 female 24.300 0 no southwest 2150.469 female no  
## 279 59 male 31.790 2 no southeast 12928.791 male no  
## 280 51 female 21.560 1 no southeast 9855.131 female no  
## 281 40 female 28.120 1 yes northeast 22331.567 female yes  
## 282 54 male 40.565 3 yes northeast 48549.178 male yes  
## 283 30 male 27.645 1 no northeast 4237.127 male no  
## 284 55 female 32.395 1 no northeast 11879.104 female no  
## 285 52 female 31.200 0 no southwest 9625.920 female no  
## 286 46 male 26.620 1 no southeast 7742.110 male no  
## 287 46 female 48.070 2 no northeast 9432.925 female no  
## 288 63 female 26.220 0 no northwest 14256.193 female no  
## 289 59 female 36.765 1 yes northeast 47896.791 female yes  
## 290 52 male 26.400 3 no southeast 25992.821 male no  
## 291 28 female 33.400 0 no southwest 3172.018 female no  
## 292 29 male 29.640 1 no northeast 20277.808 male no  
## 293 25 male 45.540 2 yes southeast 42112.236 male yes  
## 294 22 female 28.820 0 no southeast 2156.752 female no  
## 295 25 male 26.800 3 no southwest 3906.127 male no  
## 296 18 male 22.990 0 no northeast 1704.568 male no  
## 297 19 male 27.700 0 yes southwest 16297.846 male yes  
## 298 47 male 25.410 1 yes southeast 21978.677 male yes  
## 299 31 male 34.390 3 yes northwest 38746.355 male yes  
## 300 48 female 28.880 1 no northwest 9249.495 female no  
## 301 36 male 27.550 3 no northeast 6746.743 male no  
## 302 53 female 22.610 3 yes northeast 24873.385 female yes  
## 303 56 female 37.510 2 no southeast 12265.507 female no  
## 304 28 female 33.000 2 no southeast 4349.462 female no  
## 305 57 female 38.000 2 no southwest 12646.207 female no  
## 306 29 male 33.345 2 no northwest 19442.354 male no  
## 307 28 female 27.500 2 no southwest 20177.671 female no  
## 308 30 female 33.330 1 no southeast 4151.029 female no  
## 309 58 male 34.865 0 no northeast 11944.594 male no  
## 310 41 female 33.060 2 no northwest 7749.156 female no  
## 311 50 male 26.600 0 no southwest 8444.474 male no  
## 312 19 female 24.700 0 no southwest 1737.376 female no  
## 313 43 male 35.970 3 yes southeast 42124.515 male yes  
## 314 49 male 35.860 0 no southeast 8124.408 male no  
## 315 27 female 31.400 0 yes southwest 34838.873 female yes  
## 316 52 male 33.250 0 no northeast 9722.770 male no  
## 317 50 male 32.205 0 no northwest 8835.265 male no  
## 318 54 male 32.775 0 no northeast 10435.065 male no  
## 319 44 female 27.645 0 no northwest 7421.195 female no  
## 320 32 male 37.335 1 no northeast 4667.608 male no  
## 321 34 male 25.270 1 no northwest 4894.753 male no  
## 322 26 female 29.640 4 no northeast 24671.663 female no  
## 323 34 male 30.800 0 yes southwest 35491.640 male yes  
## 324 57 male 40.945 0 no northeast 11566.301 male no  
## 325 29 male 27.200 0 no southwest 2866.091 male no  
## 326 40 male 34.105 1 no northeast 6600.206 male no  
## 327 27 female 23.210 1 no southeast 3561.889 female no  
## 328 45 male 36.480 2 yes northwest 42760.502 male yes  
## 329 64 female 33.800 1 yes southwest 47928.030 female yes  
## 330 52 male 36.700 0 no southwest 9144.565 male no  
## 331 61 female 36.385 1 yes northeast 48517.563 female yes  
## 332 52 male 27.360 0 yes northwest 24393.622 male yes  
## 333 61 female 31.160 0 no northwest 13429.035 female no  
## 334 56 female 28.785 0 no northeast 11658.379 female no  
## 335 43 female 35.720 2 no northeast 19144.577 female no  
## 336 64 male 34.500 0 no southwest 13822.803 male no  
## 337 60 male 25.740 0 no southeast 12142.579 male no  
## 338 62 male 27.550 1 no northwest 13937.666 male no  
## 339 50 male 32.300 1 yes northeast 41919.097 male yes  
## 340 46 female 27.720 1 no southeast 8232.639 female no  
## 341 24 female 27.600 0 no southwest 18955.220 female no  
## 342 62 male 30.020 0 no northwest 13352.100 male no  
## 343 60 female 27.550 0 no northeast 13217.094 female no  
## 344 63 male 36.765 0 no northeast 13981.850 male no  
## 345 49 female 41.470 4 no southeast 10977.206 female no  
## 346 34 female 29.260 3 no southeast 6184.299 female no  
## 347 33 male 35.750 2 no southeast 4889.999 male no  
## 348 46 male 33.345 1 no northeast 8334.458 male no  
## 349 36 female 29.920 1 no southeast 5478.037 female no  
## 350 19 male 27.835 0 no northwest 1635.734 male no  
## 351 57 female 23.180 0 no northwest 11830.607 female no  
## 352 50 female 25.600 0 no southwest 8932.084 female no  
## 353 30 female 27.700 0 no southwest 3554.203 female no  
## 354 33 male 35.245 0 no northeast 12404.879 male no  
## 355 18 female 38.280 0 no southeast 14133.038 female no  
## 356 46 male 27.600 0 no southwest 24603.048 male no  
## 357 46 male 43.890 3 no southeast 8944.115 male no  
## 358 47 male 29.830 3 no northwest 9620.331 male no  
## 359 23 male 41.910 0 no southeast 1837.282 male no  
## 360 18 female 20.790 0 no southeast 1607.510 female no  
## 361 48 female 32.300 2 no northeast 10043.249 female no  
## 362 35 male 30.500 1 no southwest 4751.070 male no  
## 363 19 female 21.700 0 yes southwest 13844.506 female yes  
## 364 21 female 26.400 1 no southwest 2597.779 female no  
## 365 21 female 21.890 2 no southeast 3180.510 female no  
## 366 49 female 30.780 1 no northeast 9778.347 female no  
## 367 56 female 32.300 3 no northeast 13430.265 female no  
## 368 42 female 24.985 2 no northwest 8017.061 female no  
## 369 44 male 32.015 2 no northwest 8116.269 male no  
## 370 18 male 30.400 3 no northeast 3481.868 male no  
## 371 61 female 21.090 0 no northwest 13415.038 female no  
## 372 57 female 22.230 0 no northeast 12029.287 female no  
## 373 42 female 33.155 1 no northeast 7639.417 female no  
## 374 26 male 32.900 2 yes southwest 36085.219 male yes  
## 375 20 male 33.330 0 no southeast 1391.529 male no  
## 376 23 female 28.310 0 yes northwest 18033.968 female yes  
## 377 39 female 24.890 3 yes northeast 21659.930 female yes  
## 378 24 male 40.150 0 yes southeast 38126.247 male yes  
## 379 64 female 30.115 3 no northwest 16455.708 female no  
## 380 62 male 31.460 1 no southeast 27000.985 male no  
## 381 27 female 17.955 2 yes northeast 15006.579 female yes  
## 382 55 male 30.685 0 yes northeast 42303.692 male yes  
## 383 55 male 33.000 0 no southeast 20781.489 male no  
## 384 35 female 43.340 2 no southeast 5846.918 female no  
## 385 44 male 22.135 2 no northeast 8302.536 male no  
## 386 19 male 34.400 0 no southwest 1261.859 male no  
## 387 58 female 39.050 0 no southeast 11856.412 female no  
## 388 50 male 25.365 2 no northwest 30284.643 male no  
## 389 26 female 22.610 0 no northwest 3176.816 female no  
## 390 24 female 30.210 3 no northwest 4618.080 female no  
## 391 48 male 35.625 4 no northeast 10736.871 male no  
## 392 19 female 37.430 0 no northwest 2138.071 female no  
## 393 48 male 31.445 1 no northeast 8964.061 male no  
## 394 49 male 31.350 1 no northeast 9290.139 male no  
## 395 46 female 32.300 2 no northeast 9411.005 female no  
## 396 46 male 19.855 0 no northwest 7526.706 male no  
## 397 43 female 34.400 3 no southwest 8522.003 female no  
## 398 21 male 31.020 0 no southeast 16586.498 male no  
## 399 64 male 25.600 2 no southwest 14988.432 male no  
## 400 18 female 38.170 0 no southeast 1631.668 female no  
## 401 51 female 20.600 0 no southwest 9264.797 female no  
## 402 47 male 47.520 1 no southeast 8083.920 male no  
## 403 64 female 32.965 0 no northwest 14692.669 female no  
## 404 49 male 32.300 3 no northwest 10269.460 male no  
## 405 31 male 20.400 0 no southwest 3260.199 male no  
## 406 52 female 38.380 2 no northeast 11396.900 female no  
## 407 33 female 24.310 0 no southeast 4185.098 female no  
## 408 47 female 23.600 1 no southwest 8539.671 female no  
## 409 38 male 21.120 3 no southeast 6652.529 male no  
## 410 32 male 30.030 1 no southeast 4074.454 male no  
## 411 19 male 17.480 0 no northwest 1621.340 male no  
## 412 44 female 20.235 1 yes northeast 19594.810 female yes  
## 413 26 female 17.195 2 yes northeast 14455.644 female yes  
## 414 25 male 23.900 5 no southwest 5080.096 male no  
## 415 19 female 35.150 0 no northwest 2134.901 female no  
## 416 43 female 35.640 1 no southeast 7345.727 female no  
## 417 52 male 34.100 0 no southeast 9140.951 male no  
## 418 36 female 22.600 2 yes southwest 18608.262 female yes  
## 419 64 male 39.160 1 no southeast 14418.280 male no  
## 420 63 female 26.980 0 yes northwest 28950.469 female yes  
## 421 64 male 33.880 0 yes southeast 46889.261 male yes  
## 422 61 male 35.860 0 yes southeast 46599.108 male yes  
## 423 40 male 32.775 1 yes northeast 39125.332 male yes  
## 424 25 male 30.590 0 no northeast 2727.395 male no  
## 425 48 male 30.200 2 no southwest 8968.330 male no  
## 426 45 male 24.310 5 no southeast 9788.866 male no  
## 427 38 female 27.265 1 no northeast 6555.070 female no  
## 428 18 female 29.165 0 no northeast 7323.735 female no  
## 429 21 female 16.815 1 no northeast 3167.456 female no  
## 430 27 female 30.400 3 no northwest 18804.752 female no  
## 431 19 male 33.100 0 no southwest 23082.955 male no  
## 432 29 female 20.235 2 no northwest 4906.410 female no  
## 433 42 male 26.900 0 no southwest 5969.723 male no  
## 434 60 female 30.500 0 no southwest 12638.195 female no  
## 435 31 male 28.595 1 no northwest 4243.590 male no  
## 436 60 male 33.110 3 no southeast 13919.823 male no  
## 437 22 male 31.730 0 no northeast 2254.797 male no  
## 438 35 male 28.900 3 no southwest 5926.846 male no  
## 439 52 female 46.750 5 no southeast 12592.534 female no  
## 440 26 male 29.450 0 no northeast 2897.323 male no  
## 441 31 female 32.680 1 no northwest 4738.268 female no  
## 442 33 female 33.500 0 yes southwest 37079.372 female yes  
## 443 18 male 43.010 0 no southeast 1149.396 male no  
## 444 59 female 36.520 1 no southeast 28287.898 female no  
## 445 56 male 26.695 1 yes northwest 26109.329 male yes  
## 446 45 female 33.100 0 no southwest 7345.084 female no  
## 447 60 male 29.640 0 no northeast 12731.000 male no  
## 448 56 female 25.650 0 no northwest 11454.022 female no  
## 449 40 female 29.600 0 no southwest 5910.944 female no  
## 450 35 male 38.600 1 no southwest 4762.329 male no  
## 451 39 male 29.600 4 no southwest 7512.267 male no  
## 452 30 male 24.130 1 no northwest 4032.241 male no  
## 453 24 male 23.400 0 no southwest 1969.614 male no  
## 454 20 male 29.735 0 no northwest 1769.532 male no  
## 455 32 male 46.530 2 no southeast 4686.389 male no  
## 456 59 male 37.400 0 no southwest 21797.000 male no  
## 457 55 female 30.140 2 no southeast 11881.970 female no  
## 458 57 female 30.495 0 no northwest 11840.775 female no  
## 459 56 male 39.600 0 no southwest 10601.412 male no  
## 460 40 female 33.000 3 no southeast 7682.670 female no  
## 461 49 female 36.630 3 no southeast 10381.479 female no  
## 462 42 male 30.000 0 yes southwest 22144.032 male yes  
## 463 62 female 38.095 2 no northeast 15230.324 female no  
## 464 56 male 25.935 0 no northeast 11165.418 male no  
## 465 19 male 25.175 0 no northwest 1632.036 male no  
## 466 30 female 28.380 1 yes southeast 19521.968 female yes  
## 467 60 female 28.700 1 no southwest 13224.693 female no  
## 468 56 female 33.820 2 no northwest 12643.378 female no  
## 469 28 female 24.320 1 no northeast 23288.928 female no  
## 470 18 female 24.090 1 no southeast 2201.097 female no  
## 471 27 male 32.670 0 no southeast 2497.038 male no  
## 472 18 female 30.115 0 no northeast 2203.472 female no  
## 473 19 female 29.800 0 no southwest 1744.465 female no  
## 474 47 female 33.345 0 no northeast 20878.784 female no  
## 475 54 male 25.100 3 yes southwest 25382.297 male yes  
## 476 61 male 28.310 1 yes northwest 28868.664 male yes  
## 477 24 male 28.500 0 yes northeast 35147.528 male yes  
## 478 25 male 35.625 0 no northwest 2534.394 male no  
## 479 21 male 36.850 0 no southeast 1534.304 male no  
## 480 23 male 32.560 0 no southeast 1824.285 male no  
## 481 63 male 41.325 3 no northwest 15555.189 male no  
## 482 49 male 37.510 2 no southeast 9304.702 male no  
## 483 18 female 31.350 0 no southeast 1622.188 female no  
## 484 51 female 39.500 1 no southwest 9880.068 female no  
## 485 48 male 34.300 3 no southwest 9563.029 male no  
## 486 31 female 31.065 0 no northeast 4347.023 female no  
## 487 54 female 21.470 3 no northwest 12475.351 female no  
## 488 19 male 28.700 0 no southwest 1253.936 male no  
## 489 44 female 38.060 0 yes southeast 48885.136 female yes  
## 490 53 male 31.160 1 no northwest 10461.979 male no  
## 491 19 female 32.900 0 no southwest 1748.774 female no  
## 492 61 female 25.080 0 no southeast 24513.091 female no  
## 493 18 female 25.080 0 no northeast 2196.473 female no  
## 494 61 male 43.400 0 no southwest 12574.049 male no  
## 495 21 male 25.700 4 yes southwest 17942.106 male yes  
## 496 20 male 27.930 0 no northeast 1967.023 male no  
## 497 31 female 23.600 2 no southwest 4931.647 female no  
## 498 45 male 28.700 2 no southwest 8027.968 male no  
## 499 44 female 23.980 2 no southeast 8211.100 female no  
## 500 62 female 39.200 0 no southwest 13470.860 female no  
## 501 29 male 34.400 0 yes southwest 36197.699 male yes  
## 502 43 male 26.030 0 no northeast 6837.369 male no  
## 503 51 male 23.210 1 yes southeast 22218.115 male yes  
## 504 19 male 30.250 0 yes southeast 32548.340 male yes  
## 505 38 female 28.930 1 no southeast 5974.385 female no  
## 506 37 male 30.875 3 no northwest 6796.863 male no  
## 507 22 male 31.350 1 no northwest 2643.269 male no  
## 508 21 male 23.750 2 no northwest 3077.095 male no  
## 509 24 female 25.270 0 no northeast 3044.213 female no  
## 510 57 female 28.700 0 no southwest 11455.280 female no  
## 511 56 male 32.110 1 no northeast 11763.001 male no  
## 512 27 male 33.660 0 no southeast 2498.414 male no  
## 513 51 male 22.420 0 no northeast 9361.327 male no  
## 514 19 male 30.400 0 no southwest 1256.299 male no  
## 515 39 male 28.300 1 yes southwest 21082.160 male yes  
## 516 58 male 35.700 0 no southwest 11362.755 male no  
## 517 20 male 35.310 1 no southeast 27724.289 male no  
## 518 45 male 30.495 2 no northwest 8413.463 male no  
## 519 35 female 31.000 1 no southwest 5240.765 female no  
## 520 31 male 30.875 0 no northeast 3857.759 male no  
## 521 50 female 27.360 0 no northeast 25656.575 female no  
## 522 32 female 44.220 0 no southeast 3994.178 female no  
## 523 51 female 33.915 0 no northeast 9866.305 female no  
## 524 38 female 37.730 0 no southeast 5397.617 female no  
## 525 42 male 26.070 1 yes southeast 38245.593 male yes  
## 526 18 female 33.880 0 no southeast 11482.635 female no  
## 527 19 female 30.590 2 no northwest 24059.680 female no  
## 528 51 female 25.800 1 no southwest 9861.025 female no  
## 529 46 male 39.425 1 no northeast 8342.909 male no  
## 530 18 male 25.460 0 no northeast 1708.001 male no  
## 531 57 male 42.130 1 yes southeast 48675.518 male yes  
## 532 62 female 31.730 0 no northeast 14043.477 female no  
## 533 59 male 29.700 2 no southeast 12925.886 male no  
## 534 37 male 36.190 0 no southeast 19214.706 male no  
## 535 64 male 40.480 0 no southeast 13831.115 male no  
## 536 38 male 28.025 1 no northeast 6067.127 male no  
## 537 33 female 38.900 3 no southwest 5972.378 female no  
## 538 46 female 30.200 2 no southwest 8825.086 female no  
## 539 46 female 28.050 1 no southeast 8233.097 female no  
## 540 53 male 31.350 0 no southeast 27346.042 male no  
## 541 34 female 38.000 3 no southwest 6196.448 female no  
## 542 20 female 31.790 2 no southeast 3056.388 female no  
## 543 63 female 36.300 0 no southeast 13887.204 female no  
## 544 54 female 47.410 0 yes southeast 63770.428 female yes  
## 545 54 male 30.210 0 no northwest 10231.500 male no  
## 546 49 male 25.840 2 yes northwest 23807.241 male yes  
## 547 28 male 35.435 0 no northeast 3268.847 male no  
## 548 54 female 46.700 2 no southwest 11538.421 female no  
## 549 25 female 28.595 0 no northeast 3213.622 female no  
## 550 43 female 46.200 0 yes southeast 45863.205 female yes  
## 551 63 male 30.800 0 no southwest 13390.559 male no  
## 552 32 female 28.930 0 no southeast 3972.925 female no  
## 553 62 male 21.400 0 no southwest 12957.118 male no  
## 554 52 female 31.730 2 no northwest 11187.657 female no  
## 555 25 female 41.325 0 no northeast 17878.901 female no  
## 556 28 male 23.800 2 no southwest 3847.674 male no  
## 557 46 male 33.440 1 no northeast 8334.590 male no  
## 558 34 male 34.210 0 no southeast 3935.180 male no  
## 559 35 female 34.105 3 yes northwest 39983.426 female yes  
## 560 19 male 35.530 0 no northwest 1646.430 male no  
## 561 46 female 19.950 2 no northwest 9193.838 female no  
## 562 54 female 32.680 0 no northeast 10923.933 female no  
## 563 27 male 30.500 0 no southwest 2494.022 male no  
## 564 50 male 44.770 1 no southeast 9058.730 male no  
## 565 18 female 32.120 2 no southeast 2801.259 female no  
## 566 19 female 30.495 0 no northwest 2128.431 female no  
## 567 38 female 40.565 1 no northwest 6373.557 female no  
## 568 41 male 30.590 2 no northwest 7256.723 male no  
## 569 49 female 31.900 5 no southwest 11552.904 female no  
## 570 48 male 40.565 2 yes northwest 45702.022 male yes  
## 571 31 female 29.100 0 no southwest 3761.292 female no  
## 572 18 female 37.290 1 no southeast 2219.445 female no  
## 573 30 female 43.120 2 no southeast 4753.637 female no  
## 574 62 female 36.860 1 no northeast 31620.001 female no  
## 575 57 female 34.295 2 no northeast 13224.057 female no  
## 576 58 female 27.170 0 no northwest 12222.898 female no  
## 577 22 male 26.840 0 no southeast 1665.000 male no  
## 578 31 female 38.095 1 yes northeast 58571.074 female yes  
## 579 52 male 30.200 1 no southwest 9724.530 male no  
## 580 25 female 23.465 0 no northeast 3206.491 female no  
## 581 59 male 25.460 1 no northeast 12913.992 male no  
## 583 39 male 45.430 2 no southeast 6356.271 male no  
## 584 32 female 23.650 1 no southeast 17626.240 female no  
## 585 19 male 20.700 0 no southwest 1242.816 male no  
## 586 33 female 28.270 1 no southeast 4779.602 female no  
## 587 21 male 20.235 3 no northeast 3861.210 male no  
## 588 34 female 30.210 1 yes northwest 43943.876 female yes  
## 589 61 female 35.910 0 no northeast 13635.638 female no  
## 590 38 female 30.690 1 no southeast 5976.831 female no  
## 591 58 female 29.000 0 no southwest 11842.442 female no  
## 592 47 male 19.570 1 no northwest 8428.069 male no  
## 593 20 male 31.130 2 no southeast 2566.471 male no  
## 594 21 female 21.850 1 yes northeast 15359.104 female yes  
## 595 41 male 40.260 0 no southeast 5709.164 male no  
## 596 46 female 33.725 1 no northeast 8823.986 female no  
## 597 42 female 29.480 2 no southeast 7640.309 female no  
## 598 34 female 33.250 1 no northeast 5594.846 female no  
## 599 43 male 32.600 2 no southwest 7441.501 male no  
## 600 52 female 37.525 2 no northwest 33471.972 female no  
## 601 18 female 39.160 0 no southeast 1633.044 female no  
## 602 51 male 31.635 0 no northwest 9174.136 male no  
## 603 56 female 25.300 0 no southwest 11070.535 female no  
## 604 64 female 39.050 3 no southeast 16085.128 female no  
## 605 19 female 28.310 0 yes northwest 17468.984 female yes  
## 606 51 female 34.100 0 no southeast 9283.562 female no  
## 607 27 female 25.175 0 no northeast 3558.620 female no  
## 608 59 female 23.655 0 yes northwest 25678.778 female yes  
## 609 28 male 26.980 2 no northeast 4435.094 male no  
## 610 30 male 37.800 2 yes southwest 39241.442 male yes  
## 611 47 female 29.370 1 no southeast 8547.691 female no  
## 612 38 female 34.800 2 no southwest 6571.544 female no  
## 613 18 female 33.155 0 no northeast 2207.697 female no  
## 614 34 female 19.000 3 no northeast 6753.038 female no  
## 615 20 female 33.000 0 no southeast 1880.070 female no  
## 616 47 female 36.630 1 yes southeast 42969.853 female yes  
## 617 56 female 28.595 0 no northeast 11658.115 female no  
## 618 49 male 25.600 2 yes southwest 23306.547 male yes  
## 619 19 female 33.110 0 yes southeast 34439.856 female yes  
## 620 55 female 37.100 0 no southwest 10713.644 female no  
## 621 30 male 31.400 1 no southwest 3659.346 male no  
## 622 37 male 34.100 4 yes southwest 40182.246 male yes  
## 623 49 female 21.300 1 no southwest 9182.170 female no  
## 624 18 male 33.535 0 yes northeast 34617.841 male yes  
## 625 59 male 28.785 0 no northwest 12129.614 male no  
## 626 29 female 26.030 0 no northwest 3736.465 female no  
## 627 36 male 28.880 3 no northeast 6748.591 male no  
## 628 33 male 42.460 1 no southeast 11326.715 male no  
## 629 58 male 38.000 0 no southwest 11365.952 male no  
## 630 44 female 38.950 0 yes northwest 42983.459 female yes  
## 631 53 male 36.100 1 no southwest 10085.846 male no  
## 632 24 male 29.300 0 no southwest 1977.815 male no  
## 633 29 female 35.530 0 no southeast 3366.670 female no  
## 634 40 male 22.705 2 no northeast 7173.360 male no  
## 635 51 male 39.700 1 no southwest 9391.346 male no  
## 636 64 male 38.190 0 no northeast 14410.932 male no  
## 637 19 female 24.510 1 no northwest 2709.112 female no  
## 638 35 female 38.095 2 no northeast 24915.046 female no  
## 639 39 male 26.410 0 yes northeast 20149.323 male yes  
## 640 56 male 33.660 4 no southeast 12949.155 male no  
## 641 33 male 42.400 5 no southwest 6666.243 male no  
## 642 42 male 28.310 3 yes northwest 32787.459 male yes  
## 643 61 male 33.915 0 no northeast 13143.865 male no  
## 644 23 female 34.960 3 no northwest 4466.621 female no  
## 645 43 male 35.310 2 no southeast 18806.145 male no  
## 646 48 male 30.780 3 no northeast 10141.136 male no  
## 647 39 male 26.220 1 no northwest 6123.569 male no  
## 648 40 female 23.370 3 no northeast 8252.284 female no  
## 649 18 male 28.500 0 no northeast 1712.227 male no  
## 650 58 female 32.965 0 no northeast 12430.953 female no  
## 651 49 female 42.680 2 no southeast 9800.888 female no  
## 652 53 female 39.600 1 no southeast 10579.711 female no  
## 653 48 female 31.130 0 no southeast 8280.623 female no  
## 654 45 female 36.300 2 no southeast 8527.532 female no  
## 655 59 female 35.200 0 no southeast 12244.531 female no  
## 656 52 female 25.300 2 yes southeast 24667.419 female yes  
## 657 26 female 42.400 1 no southwest 3410.324 female no  
## 658 27 male 33.155 2 no northwest 4058.712 male no  
## 659 48 female 35.910 1 no northeast 26392.260 female no  
## 660 57 female 28.785 4 no northeast 14394.398 female no  
## 661 37 male 46.530 3 no southeast 6435.624 male no  
## 662 57 female 23.980 1 no southeast 22192.437 female no  
## 663 32 female 31.540 1 no northeast 5148.553 female no  
## 664 18 male 33.660 0 no southeast 1136.399 male no  
## 665 64 female 22.990 0 yes southeast 27037.914 female yes  
## 666 43 male 38.060 2 yes southeast 42560.430 male yes  
## 667 49 male 28.700 1 no southwest 8703.456 male no  
## 668 40 female 32.775 2 yes northwest 40003.332 female yes  
## 669 62 male 32.015 0 yes northeast 45710.208 male yes  
## 670 40 female 29.810 1 no southeast 6500.236 female no  
## 671 30 male 31.570 3 no southeast 4837.582 male no  
## 672 29 female 31.160 0 no northeast 3943.595 female no  
## 673 36 male 29.700 0 no southeast 4399.731 male no  
## 674 41 female 31.020 0 no southeast 6185.321 female no  
## 675 44 female 43.890 2 yes southeast 46200.985 female yes  
## 676 45 male 21.375 0 no northwest 7222.786 male no  
## 677 55 female 40.810 3 no southeast 12485.801 female no  
## 678 60 male 31.350 3 yes northwest 46130.526 male yes  
## 679 56 male 36.100 3 no southwest 12363.547 male no  
## 680 49 female 23.180 2 no northwest 10156.783 female no  
## 681 21 female 17.400 1 no southwest 2585.269 female no  
## 682 19 male 20.300 0 no southwest 1242.260 male no  
## 683 39 male 35.300 2 yes southwest 40103.890 male yes  
## 684 53 male 24.320 0 no northwest 9863.472 male no  
## 685 33 female 18.500 1 no southwest 4766.022 female no  
## 686 53 male 26.410 2 no northeast 11244.377 male no  
## 687 42 male 26.125 2 no northeast 7729.646 male no  
## 688 40 male 41.690 0 no southeast 5438.749 male no  
## 689 47 female 24.100 1 no southwest 26236.580 female no  
## 690 27 male 31.130 1 yes southeast 34806.468 male yes  
## 691 21 male 27.360 0 no northeast 2104.113 male no  
## 692 47 male 36.200 1 no southwest 8068.185 male no  
## 693 20 male 32.395 1 no northwest 2362.229 male no  
## 694 24 male 23.655 0 no northwest 2352.968 male no  
## 695 27 female 34.800 1 no southwest 3577.999 female no  
## 696 26 female 40.185 0 no northwest 3201.245 female no  
## 697 53 female 32.300 2 no northeast 29186.482 female no  
## 698 41 male 35.750 1 yes southeast 40273.645 male yes  
## 699 56 male 33.725 0 no northwest 10976.246 male no  
## 700 23 female 39.270 2 no southeast 3500.612 female no  
## 701 21 female 34.870 0 no southeast 2020.552 female no  
## 702 50 female 44.745 0 no northeast 9541.696 female no  
## 703 53 male 41.470 0 no southeast 9504.310 male no  
## 704 34 female 26.410 1 no northwest 5385.338 female no  
## 705 47 female 29.545 1 no northwest 8930.935 female no  
## 706 33 female 32.900 2 no southwest 5375.038 female no  
## 707 51 female 38.060 0 yes southeast 44400.406 female yes  
## 708 49 male 28.690 3 no northwest 10264.442 male no  
## 709 31 female 30.495 3 no northeast 6113.231 female no  
## 710 36 female 27.740 0 no northeast 5469.007 female no  
## 711 18 male 35.200 1 no southeast 1727.540 male no  
## 712 50 female 23.540 2 no southeast 10107.221 female no  
## 713 43 female 30.685 2 no northwest 8310.839 female no  
## 714 20 male 40.470 0 no northeast 1984.453 male no  
## 715 24 female 22.600 0 no southwest 2457.502 female no  
## 716 60 male 28.900 0 no southwest 12146.971 male no  
## 717 49 female 22.610 1 no northwest 9566.991 female no  
## 718 60 male 24.320 1 no northwest 13112.605 male no  
## 719 51 female 36.670 2 no northwest 10848.134 female no  
## 720 58 female 33.440 0 no northwest 12231.614 female no  
## 721 51 female 40.660 0 no northeast 9875.680 female no  
## 722 53 male 36.600 3 no southwest 11264.541 male no  
## 723 62 male 37.400 0 no southwest 12979.358 male no  
## 724 19 male 35.400 0 no southwest 1263.249 male no  
## 725 50 female 27.075 1 no northeast 10106.134 female no  
## 726 30 female 39.050 3 yes southeast 40932.429 female yes  
## 727 41 male 28.405 1 no northwest 6664.686 male no  
## 728 29 female 21.755 1 yes northeast 16657.717 female yes  
## 729 18 female 40.280 0 no northeast 2217.601 female no  
## 730 41 female 36.080 1 no southeast 6781.354 female no  
## 731 35 male 24.420 3 yes southeast 19361.999 male yes  
## 732 53 male 21.400 1 no southwest 10065.413 male no  
## 733 24 female 30.100 3 no southwest 4234.927 female no  
## 734 48 female 27.265 1 no northeast 9447.250 female no  
## 735 59 female 32.100 3 no southwest 14007.222 female no  
## 736 49 female 34.770 1 no northwest 9583.893 female no  
## 737 37 female 38.390 0 yes southeast 40419.019 female yes  
## 738 26 male 23.700 2 no southwest 3484.331 male no  
## 739 23 male 31.730 3 yes northeast 36189.102 male yes  
## 740 29 male 35.500 2 yes southwest 44585.456 male yes  
## 741 45 male 24.035 2 no northeast 8604.484 male no  
## 742 27 male 29.150 0 yes southeast 18246.496 male yes  
## 743 53 male 34.105 0 yes northeast 43254.418 male yes  
## 744 31 female 26.620 0 no southeast 3757.845 female no  
## 745 50 male 26.410 0 no northwest 8827.210 male no  
## 746 50 female 30.115 1 no northwest 9910.360 female no  
## 747 34 male 27.000 2 no southwest 11737.849 male no  
## 748 19 male 21.755 0 no northwest 1627.282 male no  
## 749 47 female 36.000 1 no southwest 8556.907 female no  
## 750 28 male 30.875 0 no northwest 3062.508 male no  
## 751 37 female 26.400 0 yes southeast 19539.243 female yes  
## 752 21 male 28.975 0 no northwest 1906.358 male no  
## 753 64 male 37.905 0 no northwest 14210.536 male no  
## 754 58 female 22.770 0 no southeast 11833.782 female no  
## 755 24 male 33.630 4 no northeast 17128.426 male no  
## 756 31 male 27.645 2 no northeast 5031.270 male no  
## 757 39 female 22.800 3 no northeast 7985.815 female no  
## 758 47 female 27.830 0 yes southeast 23065.421 female yes  
## 759 30 male 37.430 3 no northeast 5428.728 male no  
## 760 18 male 38.170 0 yes southeast 36307.798 male yes  
## 761 22 female 34.580 2 no northeast 3925.758 female no  
## 762 23 male 35.200 1 no southwest 2416.955 male no  
## 763 33 male 27.100 1 yes southwest 19040.876 male yes  
## 764 27 male 26.030 0 no northeast 3070.809 male no  
## 765 45 female 25.175 2 no northeast 9095.068 female no  
## 766 57 female 31.825 0 no northwest 11842.624 female no  
## 767 47 male 32.300 1 no southwest 8062.764 male no  
## 768 42 female 29.000 1 no southwest 7050.642 female no  
## 769 64 female 39.700 0 no southwest 14319.031 female no  
## 770 38 female 19.475 2 no northwest 6933.242 female no  
## 771 61 male 36.100 3 no southwest 27941.288 male no  
## 772 53 female 26.700 2 no southwest 11150.780 female no  
## 773 44 female 36.480 0 no northeast 12797.210 female no  
## 774 19 female 28.880 0 yes northwest 17748.506 female yes  
## 775 41 male 34.200 2 no northwest 7261.741 male no  
## 776 51 male 33.330 3 no southeast 10560.492 male no  
## 777 40 male 32.300 2 no northwest 6986.697 male no  
## 778 45 male 39.805 0 no northeast 7448.404 male no  
## 779 35 male 34.320 3 no southeast 5934.380 male no  
## 780 53 male 28.880 0 no northwest 9869.810 male no  
## 781 30 male 24.400 3 yes southwest 18259.216 male yes  
## 782 18 male 41.140 0 no southeast 1146.797 male no  
## 783 51 male 35.970 1 no southeast 9386.161 male no  
## 784 50 female 27.600 1 yes southwest 24520.264 female yes  
## 785 31 female 29.260 1 no southeast 4350.514 female no  
## 786 35 female 27.700 3 no southwest 6414.178 female no  
## 787 60 male 36.955 0 no northeast 12741.167 male no  
## 788 21 male 36.860 0 no northwest 1917.318 male no  
## 789 29 male 22.515 3 no northeast 5209.579 male no  
## 790 62 female 29.920 0 no southeast 13457.961 female no  
## 791 39 female 41.800 0 no southeast 5662.225 female no  
## 792 19 male 27.600 0 no southwest 1252.407 male no  
## 793 22 female 23.180 0 no northeast 2731.912 female no  
## 794 53 male 20.900 0 yes southeast 21195.818 male yes  
## 795 39 female 31.920 2 no northwest 7209.492 female no  
## 796 27 male 28.500 0 yes northwest 18310.742 male yes  
## 797 30 male 44.220 2 no southeast 4266.166 male no  
## 798 30 female 22.895 1 no northeast 4719.524 female no  
## 799 58 female 33.100 0 no southwest 11848.141 female no  
## 800 33 male 24.795 0 yes northeast 17904.527 male yes  
## 801 42 female 26.180 1 no southeast 7046.722 female no  
## 802 64 female 35.970 0 no southeast 14313.846 female no  
## 803 21 male 22.300 1 no southwest 2103.080 male no  
## 804 18 female 42.240 0 yes southeast 38792.686 female yes  
## 805 23 male 26.510 0 no southeast 1815.876 male no  
## 806 45 female 35.815 0 no northwest 7731.858 female no  
## 807 40 female 41.420 1 no northwest 28476.735 female no  
## 808 19 female 36.575 0 no northwest 2136.882 female no  
## 809 18 male 30.140 0 no southeast 1131.507 male no  
## 810 25 male 25.840 1 no northeast 3309.793 male no  
## 811 46 female 30.800 3 no southwest 9414.920 female no  
## 812 33 female 42.940 3 no northwest 6360.994 female no  
## 813 54 male 21.010 2 no southeast 11013.712 male no  
## 814 28 male 22.515 2 no northeast 4428.888 male no  
## 815 36 male 34.430 2 no southeast 5584.306 male no  
## 816 20 female 31.460 0 no southeast 1877.929 female no  
## 817 24 female 24.225 0 no northwest 2842.761 female no  
## 818 23 male 37.100 3 no southwest 3597.596 male no  
## 819 47 female 26.125 1 yes northeast 23401.306 female yes  
## 820 33 female 35.530 0 yes northwest 55135.402 female yes  
## 821 45 male 33.700 1 no southwest 7445.918 male no  
## 822 26 male 17.670 0 no northwest 2680.949 male no  
## 823 18 female 31.130 0 no southeast 1621.883 female no  
## 824 44 female 29.810 2 no southeast 8219.204 female no  
## 825 60 male 24.320 0 no northwest 12523.605 male no  
## 826 64 female 31.825 2 no northeast 16069.085 female no  
## 827 56 male 31.790 2 yes southeast 43813.866 male yes  
## 828 36 male 28.025 1 yes northeast 20773.628 male yes  
## 829 41 male 30.780 3 yes northeast 39597.407 male yes  
## 830 39 male 21.850 1 no northwest 6117.494 male no  
## 831 63 male 33.100 0 no southwest 13393.756 male no  
## 832 36 female 25.840 0 no northwest 5266.366 female no  
## 833 28 female 23.845 2 no northwest 4719.737 female no  
## 834 58 male 34.390 0 no northwest 11743.934 male no  
## 835 36 male 33.820 1 no northwest 5377.458 male no  
## 836 42 male 35.970 2 no southeast 7160.330 male no  
## 837 36 male 31.500 0 no southwest 4402.233 male no  
## 838 56 female 28.310 0 no northeast 11657.719 female no  
## 839 35 female 23.465 2 no northeast 6402.291 female no  
## 840 59 female 31.350 0 no northwest 12622.180 female no  
## 841 21 male 31.100 0 no southwest 1526.312 male no  
## 842 59 male 24.700 0 no northeast 12323.936 male no  
## 843 23 female 32.780 2 yes southeast 36021.011 female yes  
## 844 57 female 29.810 0 yes southeast 27533.913 female yes  
## 845 53 male 30.495 0 no northeast 10072.055 male no  
## 846 60 female 32.450 0 yes southeast 45008.955 female yes  
## 847 51 female 34.200 1 no southwest 9872.701 female no  
## 848 23 male 50.380 1 no southeast 2438.055 male no  
## 849 27 female 24.100 0 no southwest 2974.126 female no  
## 850 55 male 32.775 0 no northwest 10601.632 male no  
## 851 37 female 30.780 0 yes northeast 37270.151 female yes  
## 852 61 male 32.300 2 no northwest 14119.620 male no  
## 853 46 female 35.530 0 yes northeast 42111.665 female yes  
## 854 53 female 23.750 2 no northeast 11729.680 female no  
## 855 49 female 23.845 3 yes northeast 24106.913 female yes  
## 856 20 female 29.600 0 no southwest 1875.344 female no  
## 857 48 female 33.110 0 yes southeast 40974.165 female yes  
## 858 25 male 24.130 0 yes northwest 15817.986 male yes  
## 859 25 female 32.230 1 no southeast 18218.161 female no  
## 860 57 male 28.100 0 no southwest 10965.446 male no  
## 861 37 female 47.600 2 yes southwest 46113.511 female yes  
## 862 38 female 28.000 3 no southwest 7151.092 female no  
## 863 55 female 33.535 2 no northwest 12269.689 female no  
## 864 36 female 19.855 0 no northeast 5458.046 female no  
## 865 51 male 25.400 0 no southwest 8782.469 male no  
## 866 40 male 29.900 2 no southwest 6600.361 male no  
## 867 18 male 37.290 0 no southeast 1141.445 male no  
## 868 57 male 43.700 1 no southwest 11576.130 male no  
## 869 61 male 23.655 0 no northeast 13129.603 male no  
## 870 25 female 24.300 3 no southwest 4391.652 female no  
## 871 50 male 36.200 0 no southwest 8457.818 male no  
## 872 26 female 29.480 1 no southeast 3392.365 female no  
## 873 42 male 24.860 0 no southeast 5966.887 male no  
## 874 43 male 30.100 1 no southwest 6849.026 male no  
## 875 44 male 21.850 3 no northeast 8891.139 male no  
## 876 23 female 28.120 0 no northwest 2690.114 female no  
## 877 49 female 27.100 1 no southwest 26140.360 female no  
## 878 33 male 33.440 5 no southeast 6653.789 male no  
## 879 41 male 28.800 1 no southwest 6282.235 male no  
## 880 37 female 29.500 2 no southwest 6311.952 female no  
## 881 22 male 34.800 3 no southwest 3443.064 male no  
## 882 23 male 27.360 1 no northwest 2789.057 male no  
## 883 21 female 22.135 0 no northeast 2585.851 female no  
## 884 51 female 37.050 3 yes northeast 46255.113 female yes  
## 885 25 male 26.695 4 no northwest 4877.981 male no  
## 886 32 male 28.930 1 yes southeast 19719.695 male yes  
## 887 57 male 28.975 0 yes northeast 27218.437 male yes  
## 888 36 female 30.020 0 no northwest 5272.176 female no  
## 889 22 male 39.500 0 no southwest 1682.597 male no  
## 890 57 male 33.630 1 no northwest 11945.133 male no  
## 891 64 female 26.885 0 yes northwest 29330.983 female yes  
## 892 36 female 29.040 4 no southeast 7243.814 female no  
## 893 54 male 24.035 0 no northeast 10422.917 male no  
## 894 47 male 38.940 2 yes southeast 44202.654 male yes  
## 895 62 male 32.110 0 no northeast 13555.005 male no  
## 896 61 female 44.000 0 no southwest 13063.883 female no  
## 897 43 female 20.045 2 yes northeast 19798.055 female yes  
## 898 19 male 25.555 1 no northwest 2221.564 male no  
## 899 18 female 40.260 0 no southeast 1634.573 female no  
## 900 19 female 22.515 0 no northwest 2117.339 female no  
## 901 49 male 22.515 0 no northeast 8688.859 male no  
## 902 60 male 40.920 0 yes southeast 48673.559 male yes  
## 903 26 male 27.265 3 no northeast 4661.286 male no  
## 904 49 male 36.850 0 no southeast 8125.784 male no  
## 905 60 female 35.100 0 no southwest 12644.589 female no  
## 906 26 female 29.355 2 no northeast 4564.191 female no  
## 907 27 male 32.585 3 no northeast 4846.920 male no  
## 908 44 female 32.340 1 no southeast 7633.721 female no  
## 909 63 male 39.800 3 no southwest 15170.069 male no  
## 910 32 female 24.600 0 yes southwest 17496.306 female yes  
## 911 22 male 28.310 1 no northwest 2639.043 male no  
## 912 18 male 31.730 0 yes northeast 33732.687 male yes  
## 913 59 female 26.695 3 no northwest 14382.709 female no  
## 914 44 female 27.500 1 no southwest 7626.993 female no  
## 915 33 male 24.605 2 no northwest 5257.508 male no  
## 916 24 female 33.990 0 no southeast 2473.334 female no  
## 917 43 female 26.885 0 yes northwest 21774.322 female yes  
## 918 45 male 22.895 0 yes northeast 35069.375 male yes  
## 919 61 female 28.200 0 no southwest 13041.921 female no  
## 920 35 female 34.210 1 no southeast 5245.227 female no  
## 921 62 female 25.000 0 no southwest 13451.122 female no  
## 922 62 female 33.200 0 no southwest 13462.520 female no  
## 923 38 male 31.000 1 no southwest 5488.262 male no  
## 924 34 male 35.815 0 no northwest 4320.411 male no  
## 925 43 male 23.200 0 no southwest 6250.435 male no  
## 926 50 male 32.110 2 no northeast 25333.333 male no  
## 927 19 female 23.400 2 no southwest 2913.569 female no  
## 928 57 female 20.100 1 no southwest 12032.326 female no  
## 929 62 female 39.160 0 no southeast 13470.804 female no  
## 930 41 male 34.210 1 no southeast 6289.755 male no  
## 931 26 male 46.530 1 no southeast 2927.065 male no  
## 932 39 female 32.500 1 no southwest 6238.298 female no  
## 933 46 male 25.800 5 no southwest 10096.970 male no  
## 934 45 female 35.300 0 no southwest 7348.142 female no  
## 935 32 male 37.180 2 no southeast 4673.392 male no  
## 936 59 female 27.500 0 no southwest 12233.828 female no  
## 937 44 male 29.735 2 no northeast 32108.663 male no  
## 938 39 female 24.225 5 no northwest 8965.796 female no  
## 939 18 male 26.180 2 no southeast 2304.002 male no  
## 940 53 male 29.480 0 no southeast 9487.644 male no  
## 941 18 male 23.210 0 no southeast 1121.874 male no  
## 942 50 female 46.090 1 no southeast 9549.565 female no  
## 943 18 female 40.185 0 no northeast 2217.469 female no  
## 944 19 male 22.610 0 no northwest 1628.471 male no  
## 945 62 male 39.930 0 no southeast 12982.875 male no  
## 946 56 female 35.800 1 no southwest 11674.130 female no  
## 947 42 male 35.800 2 no southwest 7160.094 male no  
## 948 37 male 34.200 1 yes northeast 39047.285 male yes  
## 949 42 male 31.255 0 no northwest 6358.776 male no  
## 950 25 male 29.700 3 yes southwest 19933.458 male yes  
## 951 57 male 18.335 0 no northeast 11534.873 male no  
## 952 51 male 42.900 2 yes southeast 47462.894 male yes  
## 953 30 female 28.405 1 no northwest 4527.183 female no  
## 954 44 male 30.200 2 yes southwest 38998.546 male yes  
## 955 34 male 27.835 1 yes northwest 20009.634 male yes  
## 956 31 male 39.490 1 no southeast 3875.734 male no  
## 957 54 male 30.800 1 yes southeast 41999.520 male yes  
## 958 24 male 26.790 1 no northwest 12609.887 male no  
## 959 43 male 34.960 1 yes northeast 41034.221 male yes  
## 960 48 male 36.670 1 no northwest 28468.919 male no  
## 961 19 female 39.615 1 no northwest 2730.108 female no  
## 962 29 female 25.900 0 no southwest 3353.284 female no  
## 963 63 female 35.200 1 no southeast 14474.675 female no  
## 964 46 male 24.795 3 no northeast 9500.573 male no  
## 965 52 male 36.765 2 no northwest 26467.097 male no  
## 966 35 male 27.100 1 no southwest 4746.344 male no  
## 967 51 male 24.795 2 yes northwest 23967.383 male yes  
## 968 44 male 25.365 1 no northwest 7518.025 male no  
## 969 21 male 25.745 2 no northeast 3279.869 male no  
## 970 39 female 34.320 5 no southeast 8596.828 female no  
## 971 50 female 28.160 3 no southeast 10702.642 female no  
## 972 34 female 23.560 0 no northeast 4992.376 female no  
## 973 22 female 20.235 0 no northwest 2527.819 female no  
## 974 19 female 40.500 0 no southwest 1759.338 female no  
## 975 26 male 35.420 0 no southeast 2322.622 male no  
## 976 29 male 22.895 0 yes northeast 16138.762 male yes  
## 977 48 male 40.150 0 no southeast 7804.160 male no  
## 978 26 male 29.150 1 no southeast 2902.907 male no  
## 979 45 female 39.995 3 no northeast 9704.668 female no  
## 980 36 female 29.920 0 no southeast 4889.037 female no  
## 981 54 male 25.460 1 no northeast 25517.114 male no  
## 982 34 male 21.375 0 no northeast 4500.339 male no  
## 983 31 male 25.900 3 yes southwest 19199.944 male yes  
## 984 27 female 30.590 1 no northeast 16796.412 female no  
## 985 20 male 30.115 5 no northeast 4915.060 male no  
## 986 44 female 25.800 1 no southwest 7624.630 female no  
## 987 43 male 30.115 3 no northwest 8410.047 male no  
## 988 45 female 27.645 1 no northwest 28340.189 female no  
## 989 34 male 34.675 0 no northeast 4518.826 male no  
## 990 24 female 20.520 0 yes northeast 14571.891 female yes  
## 991 26 female 19.800 1 no southwest 3378.910 female no  
## 992 38 female 27.835 2 no northeast 7144.863 female no  
## 993 50 female 31.600 2 no southwest 10118.424 female no  
## 994 38 male 28.270 1 no southeast 5484.467 male no  
## 995 27 female 20.045 3 yes northwest 16420.495 female yes  
## 996 39 female 23.275 3 no northeast 7986.475 female no  
## 997 39 female 34.100 3 no southwest 7418.522 female no  
## 998 63 female 36.850 0 no southeast 13887.969 female no  
## 999 33 female 36.290 3 no northeast 6551.750 female no  
## 1000 36 female 26.885 0 no northwest 5267.818 female no  
## 1001 30 male 22.990 2 yes northwest 17361.766 male yes  
## 1002 24 male 32.700 0 yes southwest 34472.841 male yes  
## 1003 24 male 25.800 0 no southwest 1972.950 male no  
## 1004 48 male 29.600 0 no southwest 21232.182 male no  
## 1005 47 male 19.190 1 no northeast 8627.541 male no  
## 1006 29 male 31.730 2 no northwest 4433.388 male no  
## 1007 28 male 29.260 2 no northeast 4438.263 male no  
## 1008 47 male 28.215 3 yes northwest 24915.221 male yes  
## 1009 25 male 24.985 2 no northeast 23241.475 male no  
## 1010 51 male 27.740 1 no northeast 9957.722 male no  
## 1011 48 female 22.800 0 no southwest 8269.044 female no  
## 1012 43 male 20.130 2 yes southeast 18767.738 male yes  
## 1013 61 female 33.330 4 no southeast 36580.282 female no  
## 1014 48 male 32.300 1 no northwest 8765.249 male no  
## 1015 38 female 27.600 0 no southwest 5383.536 female no  
## 1016 59 male 25.460 0 no northwest 12124.992 male no  
## 1017 19 female 24.605 1 no northwest 2709.244 female no  
## 1018 26 female 34.200 2 no southwest 3987.926 female no  
## 1019 54 female 35.815 3 no northwest 12495.291 female no  
## 1020 21 female 32.680 2 no northwest 26018.951 female no  
## 1021 51 male 37.000 0 no southwest 8798.593 male no  
## 1022 22 female 31.020 3 yes southeast 35595.590 female yes  
## 1023 47 male 36.080 1 yes southeast 42211.138 male yes  
## 1024 18 male 23.320 1 no southeast 1711.027 male no  
## 1025 47 female 45.320 1 no southeast 8569.862 female no  
## 1026 21 female 34.600 0 no southwest 2020.177 female no  
## 1027 19 male 26.030 1 yes northwest 16450.895 male yes  
## 1028 23 male 18.715 0 no northwest 21595.382 male no  
## 1029 54 male 31.600 0 no southwest 9850.432 male no  
## 1030 37 female 17.290 2 no northeast 6877.980 female no  
## 1031 46 female 23.655 1 yes northwest 21677.283 female yes  
## 1032 55 female 35.200 0 yes southeast 44423.803 female yes  
## 1033 30 female 27.930 0 no northeast 4137.523 female no  
## 1034 18 male 21.565 0 yes northeast 13747.872 male yes  
## 1035 61 male 38.380 0 no northwest 12950.071 male no  
## 1036 54 female 23.000 3 no southwest 12094.478 female no  
## 1037 22 male 37.070 2 yes southeast 37484.449 male yes  
## 1038 45 female 30.495 1 yes northwest 39725.518 female yes  
## 1039 22 male 28.880 0 no northeast 2250.835 male no  
## 1040 19 male 27.265 2 no northwest 22493.660 male no  
## 1041 35 female 28.025 0 yes northwest 20234.855 female yes  
## 1042 18 male 23.085 0 no northeast 1704.700 male no  
## 1043 20 male 30.685 0 yes northeast 33475.817 male yes  
## 1044 28 female 25.800 0 no southwest 3161.454 female no  
## 1045 55 male 35.245 1 no northeast 11394.066 male no  
## 1046 43 female 24.700 2 yes northwest 21880.820 female yes  
## 1047 43 female 25.080 0 no northeast 7325.048 female no  
## 1048 22 male 52.580 1 yes southeast 44501.398 male yes  
## 1049 25 female 22.515 1 no northwest 3594.171 female no  
## 1050 49 male 30.900 0 yes southwest 39727.614 male yes  
## 1051 44 female 36.955 1 no northwest 8023.135 female no  
## 1052 64 male 26.410 0 no northeast 14394.558 male no  
## 1053 49 male 29.830 1 no northeast 9288.027 male no  
## 1054 47 male 29.800 3 yes southwest 25309.489 male yes  
## 1055 27 female 21.470 0 no northwest 3353.470 female no  
## 1056 55 male 27.645 0 no northwest 10594.502 male no  
## 1057 48 female 28.900 0 no southwest 8277.523 female no  
## 1058 45 female 31.790 0 no southeast 17929.303 female no  
## 1059 24 female 39.490 0 no southeast 2480.979 female no  
## 1060 32 male 33.820 1 no northwest 4462.722 male no  
## 1061 24 male 32.010 0 no southeast 1981.582 male no  
## 1062 57 male 27.940 1 no southeast 11554.224 male no  
## 1063 59 male 41.140 1 yes southeast 48970.248 male yes  
## 1064 36 male 28.595 3 no northwest 6548.195 male no  
## 1065 29 female 25.600 4 no southwest 5708.867 female no  
## 1066 42 female 25.300 1 no southwest 7045.499 female no  
## 1067 48 male 37.290 2 no southeast 8978.185 male no  
## 1068 39 male 42.655 0 no northeast 5757.413 male no  
## 1069 63 male 21.660 1 no northwest 14349.854 male no  
## 1070 54 female 31.900 1 no southeast 10928.849 female no  
## 1071 37 male 37.070 1 yes southeast 39871.704 male yes  
## 1072 63 male 31.445 0 no northeast 13974.456 male no  
## 1073 21 male 31.255 0 no northwest 1909.527 male no  
## 1074 54 female 28.880 2 no northeast 12096.651 female no  
## 1075 60 female 18.335 0 no northeast 13204.286 female no  
## 1076 32 female 29.590 1 no southeast 4562.842 female no  
## 1077 47 female 32.000 1 no southwest 8551.347 female no  
## 1078 21 male 26.030 0 no northeast 2102.265 male no  
## 1079 28 male 31.680 0 yes southeast 34672.147 male yes  
## 1080 63 male 33.660 3 no southeast 15161.534 male no  
## 1081 18 male 21.780 2 no southeast 11884.049 male no  
## 1082 32 male 27.835 1 no northwest 4454.403 male no  
## 1083 38 male 19.950 1 no northwest 5855.903 male no  
## 1084 32 male 31.500 1 no southwest 4076.497 male no  
## 1085 62 female 30.495 2 no northwest 15019.760 female no  
## 1086 39 female 18.300 5 yes southwest 19023.260 female yes  
## 1087 55 male 28.975 0 no northeast 10796.350 male no  
## 1088 57 male 31.540 0 no northwest 11353.228 male no  
## 1089 52 male 47.740 1 no southeast 9748.911 male no  
## 1090 56 male 22.100 0 no southwest 10577.087 male no  
## 1091 47 male 36.190 0 yes southeast 41676.081 male yes  
## 1092 55 female 29.830 0 no northeast 11286.539 female no  
## 1093 23 male 32.700 3 no southwest 3591.480 male no  
## 1094 22 female 30.400 0 yes northwest 33907.548 female yes  
## 1095 50 female 33.700 4 no southwest 11299.343 female no  
## 1096 18 female 31.350 4 no northeast 4561.189 female no  
## 1097 51 female 34.960 2 yes northeast 44641.197 female yes  
## 1098 22 male 33.770 0 no southeast 1674.632 male no  
## 1099 52 female 30.875 0 no northeast 23045.566 female no  
## 1100 25 female 33.990 1 no southeast 3227.121 female no  
## 1101 33 female 19.095 2 yes northeast 16776.304 female yes  
## 1102 53 male 28.600 3 no southwest 11253.421 male no  
## 1103 29 male 38.940 1 no southeast 3471.410 male no  
## 1104 58 male 36.080 0 no southeast 11363.283 male no  
## 1105 37 male 29.800 0 no southwest 20420.605 male no  
## 1106 54 female 31.240 0 no southeast 10338.932 female no  
## 1107 49 female 29.925 0 no northwest 8988.159 female no  
## 1108 50 female 26.220 2 no northwest 10493.946 female no  
## 1109 26 male 30.000 1 no southwest 2904.088 male no  
## 1110 45 male 20.350 3 no southeast 8605.362 male no  
## 1111 54 female 32.300 1 no northeast 11512.405 female no  
## 1112 38 male 38.390 3 yes southeast 41949.244 male yes  
## 1113 48 female 25.850 3 yes southeast 24180.933 female yes  
## 1114 28 female 26.315 3 no northwest 5312.170 female no  
## 1115 23 male 24.510 0 no northeast 2396.096 male no  
## 1116 55 male 32.670 1 no southeast 10807.486 male no  
## 1117 41 male 29.640 5 no northeast 9222.403 male no  
## 1118 25 male 33.330 2 yes southeast 36124.574 male yes  
## 1119 33 male 35.750 1 yes southeast 38282.749 male yes  
## 1120 30 female 19.950 3 no northwest 5693.431 female no  
## 1121 23 female 31.400 0 yes southwest 34166.273 female yes  
## 1122 46 male 38.170 2 no southeast 8347.164 male no  
## 1123 53 female 36.860 3 yes northwest 46661.442 female yes  
## 1124 27 female 32.395 1 no northeast 18903.491 female no  
## 1125 23 female 42.750 1 yes northeast 40904.200 female yes  
## 1126 63 female 25.080 0 no northwest 14254.608 female no  
## 1127 55 male 29.900 0 no southwest 10214.636 male no  
## 1128 35 female 35.860 2 no southeast 5836.520 female no  
## 1129 34 male 32.800 1 no southwest 14358.364 male no  
## 1130 19 female 18.600 0 no southwest 1728.897 female no  
## 1131 39 female 23.870 5 no southeast 8582.302 female no  
## 1132 27 male 45.900 2 no southwest 3693.428 male no  
## 1133 57 male 40.280 0 no northeast 20709.020 male no  
## 1134 52 female 18.335 0 no northwest 9991.038 female no  
## 1135 28 male 33.820 0 no northwest 19673.336 male no  
## 1136 50 female 28.120 3 no northwest 11085.587 female no  
## 1137 44 female 25.000 1 no southwest 7623.518 female no  
## 1138 26 female 22.230 0 no northwest 3176.288 female no  
## 1139 33 male 30.250 0 no southeast 3704.354 male no  
## 1140 19 female 32.490 0 yes northwest 36898.733 female yes  
## 1141 50 male 37.070 1 no southeast 9048.027 male no  
## 1142 41 female 32.600 3 no southwest 7954.517 female no  
## 1143 52 female 24.860 0 no southeast 27117.994 female no  
## 1144 39 male 32.340 2 no southeast 6338.076 male no  
## 1145 50 male 32.300 2 no southwest 9630.397 male no  
## 1146 52 male 32.775 3 no northwest 11289.109 male no  
## 1147 60 male 32.800 0 yes southwest 52590.829 male yes  
## 1148 20 female 31.920 0 no northwest 2261.569 female no  
## 1149 55 male 21.500 1 no southwest 10791.960 male no  
## 1150 42 male 34.100 0 no southwest 5979.731 male no  
## 1151 18 female 30.305 0 no northeast 2203.736 female no  
## 1152 58 female 36.480 0 no northwest 12235.839 female no  
## 1153 43 female 32.560 3 yes southeast 40941.285 female yes  
## 1154 35 female 35.815 1 no northwest 5630.458 female no  
## 1155 48 female 27.930 4 no northwest 11015.175 female no  
## 1156 36 female 22.135 3 no northeast 7228.216 female no  
## 1157 19 male 44.880 0 yes southeast 39722.746 male yes  
## 1158 23 female 23.180 2 no northwest 14426.074 female no  
## 1159 20 female 30.590 0 no northeast 2459.720 female no  
## 1160 32 female 41.100 0 no southwest 3989.841 female no  
## 1161 43 female 34.580 1 no northwest 7727.253 female no  
## 1162 34 male 42.130 2 no southeast 5124.189 male no  
## 1163 30 male 38.830 1 no southeast 18963.172 male no  
## 1164 18 female 28.215 0 no northeast 2200.831 female no  
## 1165 41 female 28.310 1 no northwest 7153.554 female no  
## 1166 35 female 26.125 0 no northeast 5227.989 female no  
## 1167 57 male 40.370 0 no southeast 10982.501 male no  
## 1168 29 female 24.600 2 no southwest 4529.477 female no  
## 1169 32 male 35.200 2 no southwest 4670.640 male no  
## 1170 37 female 34.105 1 no northwest 6112.353 female no  
## 1171 18 male 27.360 1 yes northeast 17178.682 male yes  
## 1172 43 female 26.700 2 yes southwest 22478.600 female yes  
## 1173 56 female 41.910 0 no southeast 11093.623 female no  
## 1174 38 male 29.260 2 no northwest 6457.843 male no  
## 1175 29 male 32.110 2 no northwest 4433.916 male no  
## 1176 22 female 27.100 0 no southwest 2154.361 female no  
## 1177 52 female 24.130 1 yes northwest 23887.663 female yes  
## 1178 40 female 27.400 1 no southwest 6496.886 female no  
## 1179 23 female 34.865 0 no northeast 2899.489 female no  
## 1180 31 male 29.810 0 yes southeast 19350.369 male yes  
## 1181 42 female 41.325 1 no northeast 7650.774 female no  
## 1182 24 female 29.925 0 no northwest 2850.684 female no  
## 1183 25 female 30.300 0 no southwest 2632.992 female no  
## 1184 48 female 27.360 1 no northeast 9447.382 female no  
## 1185 23 female 28.490 1 yes southeast 18328.238 female yes  
## 1186 45 male 23.560 2 no northeast 8603.823 male no  
## 1187 20 male 35.625 3 yes northwest 37465.344 male yes  
## 1188 62 female 32.680 0 no northwest 13844.797 female no  
## 1189 43 female 25.270 1 yes northeast 21771.342 female yes  
## 1190 23 female 28.000 0 no southwest 13126.677 female no  
## 1191 31 female 32.775 2 no northwest 5327.400 female no  
## 1192 41 female 21.755 1 no northeast 13725.472 female no  
## 1193 58 female 32.395 1 no northeast 13019.161 female no  
## 1194 48 female 36.575 0 no northwest 8671.191 female no  
## 1195 31 female 21.755 0 no northwest 4134.082 female no  
## 1196 19 female 27.930 3 no northwest 18838.704 female no  
## 1197 19 female 30.020 0 yes northwest 33307.551 female yes  
## 1198 41 male 33.550 0 no southeast 5699.837 male no  
## 1199 40 male 29.355 1 no northwest 6393.603 male no  
## 1200 31 female 25.800 2 no southwest 4934.705 female no  
## 1201 37 male 24.320 2 no northwest 6198.752 male no  
## 1202 46 male 40.375 2 no northwest 8733.229 male no  
## 1203 22 male 32.110 0 no northwest 2055.325 male no  
## 1204 51 male 32.300 1 no northeast 9964.060 male no  
## 1205 18 female 27.280 3 yes southeast 18223.451 female yes  
## 1206 35 male 17.860 1 no northwest 5116.500 male no  
## 1207 59 female 34.800 2 no southwest 36910.608 female no  
## 1208 36 male 33.400 2 yes southwest 38415.474 male yes  
## 1209 37 female 25.555 1 yes northeast 20296.863 female yes  
## 1210 59 male 37.100 1 no southwest 12347.172 male no  
## 1211 36 male 30.875 1 no northwest 5373.364 male no  
## 1212 39 male 34.100 2 no southeast 23563.016 male no  
## 1213 18 male 21.470 0 no northeast 1702.455 male no  
## 1214 52 female 33.300 2 no southwest 10806.839 female no  
## 1215 27 female 31.255 1 no northwest 3956.071 female no  
## 1216 18 male 39.140 0 no northeast 12890.058 male no  
## 1217 40 male 25.080 0 no southeast 5415.661 male no  
## 1218 29 male 37.290 2 no southeast 4058.116 male no  
## 1219 46 female 34.600 1 yes southwest 41661.602 female yes  
## 1220 38 female 30.210 3 no northwest 7537.164 female no  
## 1221 30 female 21.945 1 no northeast 4718.204 female no  
## 1222 40 male 24.970 2 no southeast 6593.508 male no  
## 1223 50 male 25.300 0 no southeast 8442.667 male no  
## 1224 20 female 24.420 0 yes southeast 26125.675 female yes  
## 1225 41 male 23.940 1 no northeast 6858.480 male no  
## 1226 33 female 39.820 1 no southeast 4795.657 female no  
## 1227 38 male 16.815 2 no northeast 6640.545 male no  
## 1228 42 male 37.180 2 no southeast 7162.012 male no  
## 1229 56 male 34.430 0 no southeast 10594.226 male no  
## 1230 58 male 30.305 0 no northeast 11938.256 male no  
## 1231 52 male 34.485 3 yes northwest 60021.399 male yes  
## 1232 20 female 21.800 0 yes southwest 20167.336 female yes  
## 1233 54 female 24.605 3 no northwest 12479.709 female no  
## 1234 58 male 23.300 0 no southwest 11345.519 male no  
## 1235 45 female 27.830 2 no southeast 8515.759 female no  
## 1236 26 male 31.065 0 no northwest 2699.568 male no  
## 1237 63 female 21.660 0 no northeast 14449.854 female no  
## 1238 58 female 28.215 0 no northwest 12224.351 female no  
## 1239 37 male 22.705 3 no northeast 6985.507 male no  
## 1240 25 female 42.130 1 no southeast 3238.436 female no  
## 1241 52 male 41.800 2 yes southeast 47269.854 male yes  
## 1242 64 male 36.960 2 yes southeast 49577.662 male yes  
## 1243 22 female 21.280 3 no northwest 4296.271 female no  
## 1244 28 female 33.110 0 no southeast 3171.615 female no  
## 1245 18 male 33.330 0 no southeast 1135.941 male no  
## 1246 28 male 24.300 5 no southwest 5615.369 male no  
## 1247 45 female 25.700 3 no southwest 9101.798 female no  
## 1248 33 male 29.400 4 no southwest 6059.173 male no  
## 1249 18 female 39.820 0 no southeast 1633.962 female no  
## 1250 32 male 33.630 1 yes northeast 37607.528 male yes  
## 1251 24 male 29.830 0 yes northeast 18648.422 male yes  
## 1252 19 male 19.800 0 no southwest 1241.565 male no  
## 1253 20 male 27.300 0 yes southwest 16232.847 male yes  
## 1254 40 female 29.300 4 no southwest 15828.822 female no  
## 1255 34 female 27.720 0 no southeast 4415.159 female no  
## 1256 42 female 37.900 0 no southwest 6474.013 female no  
## 1257 51 female 36.385 3 no northwest 11436.738 female no  
## 1258 54 female 27.645 1 no northwest 11305.935 female no  
## 1259 55 male 37.715 3 no northwest 30063.581 male no  
## 1260 52 female 23.180 0 no northeast 10197.772 female no  
## 1261 32 female 20.520 0 no northeast 4544.235 female no  
## 1262 28 male 37.100 1 no southwest 3277.161 male no  
## 1263 41 female 28.050 1 no southeast 6770.193 female no  
## 1264 43 female 29.900 1 no southwest 7337.748 female no  
## 1265 49 female 33.345 2 no northeast 10370.913 female no  
## 1266 64 male 23.760 0 yes southeast 26926.514 male yes  
## 1267 55 female 30.500 0 no southwest 10704.470 female no  
## 1268 24 male 31.065 0 yes northeast 34254.053 male yes  
## 1269 20 female 33.300 0 no southwest 1880.487 female no  
## 1270 45 male 27.500 3 no southwest 8615.300 male no  
## 1271 26 male 33.915 1 no northwest 3292.530 male no  
## 1272 25 female 34.485 0 no northwest 3021.809 female no  
## 1273 43 male 25.520 5 no southeast 14478.330 male no  
## 1274 35 male 27.610 1 no southeast 4747.053 male no  
## 1275 26 male 27.060 0 yes southeast 17043.341 male yes  
## 1276 57 male 23.700 0 no southwest 10959.330 male no  
## 1277 22 female 30.400 0 no northeast 2741.948 female no  
## 1278 32 female 29.735 0 no northwest 4357.044 female no  
## 1279 39 male 29.925 1 yes northeast 22462.044 male yes  
## 1280 25 female 26.790 2 no northwest 4189.113 female no  
## 1281 48 female 33.330 0 no southeast 8283.681 female no  
## 1282 47 female 27.645 2 yes northwest 24535.699 female yes  
## 1283 18 female 21.660 0 yes northeast 14283.459 female yes  
## 1284 18 male 30.030 1 no southeast 1720.354 male no  
## 1285 61 male 36.300 1 yes southwest 47403.880 male yes  
## 1286 47 female 24.320 0 no northeast 8534.672 female no  
## 1287 28 female 17.290 0 no northeast 3732.625 female no  
## 1288 36 female 25.900 1 no southwest 5472.449 female no  
## 1289 20 male 39.400 2 yes southwest 38344.566 male yes  
## 1290 44 male 34.320 1 no southeast 7147.473 male no  
## 1291 38 female 19.950 2 no northeast 7133.903 female no  
## 1292 19 male 34.900 0 yes southwest 34828.654 male yes  
## 1293 21 male 23.210 0 no southeast 1515.345 male no  
## 1294 46 male 25.745 3 no northwest 9301.894 male no  
## 1295 58 male 25.175 0 no northeast 11931.125 male no  
## 1296 20 male 22.000 1 no southwest 1964.780 male no  
## 1297 18 male 26.125 0 no northeast 1708.926 male no  
## 1298 28 female 26.510 2 no southeast 4340.441 female no  
## 1299 33 male 27.455 2 no northwest 5261.469 male no  
## 1300 19 female 25.745 1 no northwest 2710.829 female no  
## 1301 45 male 30.360 0 yes southeast 62592.873 male yes  
## 1302 62 male 30.875 3 yes northwest 46718.163 male yes  
## 1303 25 female 20.800 1 no southwest 3208.787 female no  
## 1304 43 male 27.800 0 yes southwest 37829.724 male yes  
## 1305 42 male 24.605 2 yes northeast 21259.378 male yes  
## 1306 24 female 27.720 0 no southeast 2464.619 female no  
## 1307 29 female 21.850 0 yes northeast 16115.305 female yes  
## 1308 32 male 28.120 4 yes northwest 21472.479 male yes  
## 1309 25 female 30.200 0 yes southwest 33900.653 female yes  
## 1310 41 male 32.200 2 no southwest 6875.961 male no  
## 1311 42 male 26.315 1 no northwest 6940.910 male no  
## 1312 33 female 26.695 0 no northwest 4571.413 female no  
## 1313 34 male 42.900 1 no southwest 4536.259 male no  
## 1314 19 female 34.700 2 yes southwest 36397.576 female yes  
## 1315 30 female 23.655 3 yes northwest 18765.875 female yes  
## 1316 18 male 28.310 1 no northeast 11272.331 male no  
## 1317 19 female 20.600 0 no southwest 1731.677 female no  
## 1318 18 male 53.130 0 no southeast 1163.463 male no  
## 1319 35 male 39.710 4 no northeast 19496.719 male no  
## 1320 39 female 26.315 2 no northwest 7201.701 female no  
## 1321 31 male 31.065 3 no northwest 5425.023 male no  
## 1322 62 male 26.695 0 yes northeast 28101.333 male yes  
## 1323 62 male 38.830 0 no southeast 12981.346 male no  
## 1324 42 female 40.370 2 yes southeast 43896.376 female yes  
## 1325 31 male 25.935 1 no northwest 4239.893 male no  
## 1326 61 male 33.535 0 no northeast 13143.337 male no  
## 1327 42 female 32.870 0 no northeast 7050.021 female no  
## 1328 51 male 30.030 1 no southeast 9377.905 male no  
## 1329 23 female 24.225 2 no northeast 22395.744 female no  
## 1330 52 male 38.600 2 no southwest 10325.206 male no  
## 1331 57 female 25.740 2 no southeast 12629.166 female no  
## 1332 23 female 33.400 0 no southwest 10795.937 female no  
## 1333 52 female 44.700 3 no southwest 11411.685 female no  
## 1334 50 male 30.970 3 no northwest 10600.548 male no  
## 1335 18 female 31.920 0 no northeast 2205.981 female no  
## 1336 18 female 36.850 0 no southeast 1629.833 female no  
## 1337 21 female 25.800 0 no southwest 2007.945 female no  
## 1338 61 female 29.070 0 yes northwest 29141.360 female yes  
## f.region  
## 1 southwest  
## 2 southeast  
## 3 southeast  
## 4 northwest  
## 5 northwest  
## 6 southeast  
## 7 southeast  
## 8 northwest  
## 9 northeast  
## 10 northwest  
## 11 northeast  
## 12 southeast  
## 13 southwest  
## 14 southeast  
## 15 southeast  
## 16 southwest  
## 17 northeast  
## 18 northeast  
## 19 southwest  
## 20 southwest  
## 21 northeast  
## 22 southwest  
## 23 southeast  
## 24 northeast  
## 25 northwest  
## 26 southeast  
## 27 northeast  
## 28 northwest  
## 29 northwest  
## 30 southwest  
## 31 southwest  
## 32 northeast  
## 33 southwest  
## 34 northwest  
## 35 southwest  
## 36 northwest  
## 37 northwest  
## 38 southwest  
## 39 northeast  
## 40 southwest  
## 41 northeast  
## 42 southeast  
## 43 southeast  
## 44 southeast  
## 45 northeast  
## 46 southwest  
## 47 northeast  
## 48 northwest  
## 49 southeast  
## 50 southeast  
## 51 northeast  
## 52 northwest  
## 53 southwest  
## 54 southeast  
## 55 northwest  
## 56 northwest  
## 57 northeast  
## 58 southeast  
## 59 southeast  
## 60 northwest  
## 61 northeast  
## 62 southeast  
## 63 northwest  
## 64 northwest  
## 65 northwest  
## 66 southwest  
## 67 southwest  
## 68 northwest  
## 69 southeast  
## 70 southeast  
## 71 southeast  
## 72 northeast  
## 73 southwest  
## 74 southeast  
## 75 southwest  
## 76 northwest  
## 77 southeast  
## 78 southeast  
## 79 northeast  
## 80 northwest  
## 81 northeast  
## 82 northeast  
## 83 southeast  
## 84 northwest  
## 85 southwest  
## 86 northwest  
## 87 northwest  
## 88 southwest  
## 89 northwest  
## 90 northwest  
## 91 southeast  
## 92 northwest  
## 93 northeast  
## 94 northwest  
## 95 southwest  
## 96 southeast  
## 97 southwest  
## 98 southeast  
## 99 northeast  
## 100 southwest  
## 101 southwest  
## 102 northeast  
## 103 northeast  
## 104 southeast  
## 105 southwest  
## 106 northwest  
## 107 southwest  
## 108 northwest  
## 109 southeast  
## 110 southeast  
## 111 northwest  
## 112 southwest  
## 113 southwest  
## 114 northwest  
## 115 northeast  
## 116 northeast  
## 117 southeast  
## 118 southeast  
## 119 southeast  
## 120 northwest  
## 121 southwest  
## 122 northeast  
## 123 northwest  
## 124 northeast  
## 125 northwest  
## 126 northeast  
## 127 southwest  
## 128 southwest  
## 129 northwest  
## 130 southwest  
## 131 northeast  
## 132 northeast  
## 133 southwest  
## 134 northwest  
## 135 northeast  
## 136 southeast  
## 137 southwest  
## 138 northwest  
## 139 southeast  
## 140 southwest  
## 141 northeast  
## 142 northeast  
## 143 southeast  
## 144 northwest  
## 145 northwest  
## 146 southeast  
## 147 northwest  
## 148 southeast  
## 149 northwest  
## 150 southwest  
## 151 northwest  
## 152 southeast  
## 153 northeast  
## 154 northeast  
## 155 northeast  
## 156 northwest  
## 157 southeast  
## 158 northeast  
## 159 southeast  
## 160 southeast  
## 161 northwest  
## 162 southeast  
## 163 southwest  
## 164 southwest  
## 165 northwest  
## 166 northeast  
## 167 southwest  
## 168 northwest  
## 169 northwest  
## 170 northeast  
## 171 southeast  
## 172 southwest  
## 173 northeast  
## 174 southwest  
## 175 northwest  
## 176 southwest  
## 177 northwest  
## 178 southwest  
## 179 southwest  
## 180 northeast  
## 181 northwest  
## 182 southeast  
## 183 northeast  
## 184 northwest  
## 185 southeast  
## 186 northeast  
## 187 southeast  
## 188 southwest  
## 189 southwest  
## 190 northwest  
## 191 southeast  
## 192 southwest  
## 193 southeast  
## 194 northwest  
## 195 southeast  
## 196 northwest  
## 197 southwest  
## 198 southeast  
## 199 northwest  
## 200 northeast  
## 201 northwest  
## 202 southeast  
## 203 northwest  
## 204 southeast  
## 205 southwest  
## 206 northeast  
## 207 southeast  
## 208 northeast  
## 209 southwest  
## 210 northeast  
## 211 southwest  
## 212 northwest  
## 213 northwest  
## 214 southeast  
## 215 southwest  
## 216 southwest  
## 217 northwest  
## 218 southeast  
## 219 southeast  
## 220 southeast  
## 221 southwest  
## 222 northeast  
## 223 southwest  
## 224 southwest  
## 225 southeast  
## 226 southeast  
## 227 southeast  
## 228 southeast  
## 229 northeast  
## 230 northeast  
## 231 northwest  
## 232 southeast  
## 233 southwest  
## 234 southwest  
## 235 northwest  
## 236 southeast  
## 237 southeast  
## 238 southeast  
## 239 northwest  
## 240 southeast  
## 241 northeast  
## 242 northeast  
## 243 southwest  
## 244 southwest  
## 245 northeast  
## 246 northwest  
## 247 southeast  
## 248 southeast  
## 249 southwest  
## 250 northeast  
## 251 northeast  
## 252 southwest  
## 253 southeast  
## 254 southwest  
## 255 northeast  
## 256 northeast  
## 257 northwest  
## 258 southeast  
## 259 northwest  
## 260 northwest  
## 261 southwest  
## 262 southeast  
## 263 northeast  
## 264 northwest  
## 265 southeast  
## 266 southeast  
## 267 southeast  
## 268 northeast  
## 269 southwest  
## 270 northeast  
## 271 southeast  
## 272 southwest  
## 273 northwest  
## 274 northeast  
## 275 northwest  
## 276 northeast  
## 277 northwest  
## 278 southwest  
## 279 southeast  
## 280 southeast  
## 281 northeast  
## 282 northeast  
## 283 northeast  
## 284 northeast  
## 285 southwest  
## 286 southeast  
## 287 northeast  
## 288 northwest  
## 289 northeast  
## 290 southeast  
## 291 southwest  
## 292 northeast  
## 293 southeast  
## 294 southeast  
## 295 southwest  
## 296 northeast  
## 297 southwest  
## 298 southeast  
## 299 northwest  
## 300 northwest  
## 301 northeast  
## 302 northeast  
## 303 southeast  
## 304 southeast  
## 305 southwest  
## 306 northwest  
## 307 southwest  
## 308 southeast  
## 309 northeast  
## 310 northwest  
## 311 southwest  
## 312 southwest  
## 313 southeast  
## 314 southeast  
## 315 southwest  
## 316 northeast  
## 317 northwest  
## 318 northeast  
## 319 northwest  
## 320 northeast  
## 321 northwest  
## 322 northeast  
## 323 southwest  
## 324 northeast  
## 325 southwest  
## 326 northeast  
## 327 southeast  
## 328 northwest  
## 329 southwest  
## 330 southwest  
## 331 northeast  
## 332 northwest  
## 333 northwest  
## 334 northeast  
## 335 northeast  
## 336 southwest  
## 337 southeast  
## 338 northwest  
## 339 northeast  
## 340 southeast  
## 341 southwest  
## 342 northwest  
## 343 northeast  
## 344 northeast  
## 345 southeast  
## 346 southeast  
## 347 southeast  
## 348 northeast  
## 349 southeast  
## 350 northwest  
## 351 northwest  
## 352 southwest  
## 353 southwest  
## 354 northeast  
## 355 southeast  
## 356 southwest  
## 357 southeast  
## 358 northwest  
## 359 southeast  
## 360 southeast  
## 361 northeast  
## 362 southwest  
## 363 southwest  
## 364 southwest  
## 365 southeast  
## 366 northeast  
## 367 northeast  
## 368 northwest  
## 369 northwest  
## 370 northeast  
## 371 northwest  
## 372 northeast  
## 373 northeast  
## 374 southwest  
## 375 southeast  
## 376 northwest  
## 377 northeast  
## 378 southeast  
## 379 northwest  
## 380 southeast  
## 381 northeast  
## 382 northeast  
## 383 southeast  
## 384 southeast  
## 385 northeast  
## 386 southwest  
## 387 southeast  
## 388 northwest  
## 389 northwest  
## 390 northwest  
## 391 northeast  
## 392 northwest  
## 393 northeast  
## 394 northeast  
## 395 northeast  
## 396 northwest  
## 397 southwest  
## 398 southeast  
## 399 southwest  
## 400 southeast  
## 401 southwest  
## 402 southeast  
## 403 northwest  
## 404 northwest  
## 405 southwest  
## 406 northeast  
## 407 southeast  
## 408 southwest  
## 409 southeast  
## 410 southeast  
## 411 northwest  
## 412 northeast  
## 413 northeast  
## 414 southwest  
## 415 northwest  
## 416 southeast  
## 417 southeast  
## 418 southwest  
## 419 southeast  
## 420 northwest  
## 421 southeast  
## 422 southeast  
## 423 northeast  
## 424 northeast  
## 425 southwest  
## 426 southeast  
## 427 northeast  
## 428 northeast  
## 429 northeast  
## 430 northwest  
## 431 southwest  
## 432 northwest  
## 433 southwest  
## 434 southwest  
## 435 northwest  
## 436 southeast  
## 437 northeast  
## 438 southwest  
## 439 southeast  
## 440 northeast  
## 441 northwest  
## 442 southwest  
## 443 southeast  
## 444 southeast  
## 445 northwest  
## 446 southwest  
## 447 northeast  
## 448 northwest  
## 449 southwest  
## 450 southwest  
## 451 southwest  
## 452 northwest  
## 453 southwest  
## 454 northwest  
## 455 southeast  
## 456 southwest  
## 457 southeast  
## 458 northwest  
## 459 southwest  
## 460 southeast  
## 461 southeast  
## 462 southwest  
## 463 northeast  
## 464 northeast  
## 465 northwest  
## 466 southeast  
## 467 southwest  
## 468 northwest  
## 469 northeast  
## 470 southeast  
## 471 southeast  
## 472 northeast  
## 473 southwest  
## 474 northeast  
## 475 southwest  
## 476 northwest  
## 477 northeast  
## 478 northwest  
## 479 southeast  
## 480 southeast  
## 481 northwest  
## 482 southeast  
## 483 southeast  
## 484 southwest  
## 485 southwest  
## 486 northeast  
## 487 northwest  
## 488 southwest  
## 489 southeast  
## 490 northwest  
## 491 southwest  
## 492 southeast  
## 493 northeast  
## 494 southwest  
## 495 southwest  
## 496 northeast  
## 497 southwest  
## 498 southwest  
## 499 southeast  
## 500 southwest  
## 501 southwest  
## 502 northeast  
## 503 southeast  
## 504 southeast  
## 505 southeast  
## 506 northwest  
## 507 northwest  
## 508 northwest  
## 509 northeast  
## 510 southwest  
## 511 northeast  
## 512 southeast  
## 513 northeast  
## 514 southwest  
## 515 southwest  
## 516 southwest  
## 517 southeast  
## 518 northwest  
## 519 southwest  
## 520 northeast  
## 521 northeast  
## 522 southeast  
## 523 northeast  
## 524 southeast  
## 525 southeast  
## 526 southeast  
## 527 northwest  
## 528 southwest  
## 529 northeast  
## 530 northeast  
## 531 southeast  
## 532 northeast  
## 533 southeast  
## 534 southeast  
## 535 southeast  
## 536 northeast  
## 537 southwest  
## 538 southwest  
## 539 southeast  
## 540 southeast  
## 541 southwest  
## 542 southeast  
## 543 southeast  
## 544 southeast  
## 545 northwest  
## 546 northwest  
## 547 northeast  
## 548 southwest  
## 549 northeast  
## 550 southeast  
## 551 southwest  
## 552 southeast  
## 553 southwest  
## 554 northwest  
## 555 northeast  
## 556 southwest  
## 557 northeast  
## 558 southeast  
## 559 northwest  
## 560 northwest  
## 561 northwest  
## 562 northeast  
## 563 southwest  
## 564 southeast  
## 565 southeast  
## 566 northwest  
## 567 northwest  
## 568 northwest  
## 569 southwest  
## 570 northwest  
## 571 southwest  
## 572 southeast  
## 573 southeast  
## 574 northeast  
## 575 northeast  
## 576 northwest  
## 577 southeast  
## 578 northeast  
## 579 southwest  
## 580 northeast  
## 581 northeast  
## 583 southeast  
## 584 southeast  
## 585 southwest  
## 586 southeast  
## 587 northeast  
## 588 northwest  
## 589 northeast  
## 590 southeast  
## 591 southwest  
## 592 northwest  
## 593 southeast  
## 594 northeast  
## 595 southeast  
## 596 northeast  
## 597 southeast  
## 598 northeast  
## 599 southwest  
## 600 northwest  
## 601 southeast  
## 602 northwest  
## 603 southwest  
## 604 southeast  
## 605 northwest  
## 606 southeast  
## 607 northeast  
## 608 northwest  
## 609 northeast  
## 610 southwest  
## 611 southeast  
## 612 southwest  
## 613 northeast  
## 614 northeast  
## 615 southeast  
## 616 southeast  
## 617 northeast  
## 618 southwest  
## 619 southeast  
## 620 southwest  
## 621 southwest  
## 622 southwest  
## 623 southwest  
## 624 northeast  
## 625 northwest  
## 626 northwest  
## 627 northeast  
## 628 southeast  
## 629 southwest  
## 630 northwest  
## 631 southwest  
## 632 southwest  
## 633 southeast  
## 634 northeast  
## 635 southwest  
## 636 northeast  
## 637 northwest  
## 638 northeast  
## 639 northeast  
## 640 southeast  
## 641 southwest  
## 642 northwest  
## 643 northeast  
## 644 northwest  
## 645 southeast  
## 646 northeast  
## 647 northwest  
## 648 northeast  
## 649 northeast  
## 650 northeast  
## 651 southeast  
## 652 southeast  
## 653 southeast  
## 654 southeast  
## 655 southeast  
## 656 southeast  
## 657 southwest  
## 658 northwest  
## 659 northeast  
## 660 northeast  
## 661 southeast  
## 662 southeast  
## 663 northeast  
## 664 southeast  
## 665 southeast  
## 666 southeast  
## 667 southwest  
## 668 northwest  
## 669 northeast  
## 670 southeast  
## 671 southeast  
## 672 northeast  
## 673 southeast  
## 674 southeast  
## 675 southeast  
## 676 northwest  
## 677 southeast  
## 678 northwest  
## 679 southwest  
## 680 northwest  
## 681 southwest  
## 682 southwest  
## 683 southwest  
## 684 northwest  
## 685 southwest  
## 686 northeast  
## 687 northeast  
## 688 southeast  
## 689 southwest  
## 690 southeast  
## 691 northeast  
## 692 southwest  
## 693 northwest  
## 694 northwest  
## 695 southwest  
## 696 northwest  
## 697 northeast  
## 698 southeast  
## 699 northwest  
## 700 southeast  
## 701 southeast  
## 702 northeast  
## 703 southeast  
## 704 northwest  
## 705 northwest  
## 706 southwest  
## 707 southeast  
## 708 northwest  
## 709 northeast  
## 710 northeast  
## 711 southeast  
## 712 southeast  
## 713 northwest  
## 714 northeast  
## 715 southwest  
## 716 southwest  
## 717 northwest  
## 718 northwest  
## 719 northwest  
## 720 northwest  
## 721 northeast  
## 722 southwest  
## 723 southwest  
## 724 southwest  
## 725 northeast  
## 726 southeast  
## 727 northwest  
## 728 northeast  
## 729 northeast  
## 730 southeast  
## 731 southeast  
## 732 southwest  
## 733 southwest  
## 734 northeast  
## 735 southwest  
## 736 northwest  
## 737 southeast  
## 738 southwest  
## 739 northeast  
## 740 southwest  
## 741 northeast  
## 742 southeast  
## 743 northeast  
## 744 southeast  
## 745 northwest  
## 746 northwest  
## 747 southwest  
## 748 northwest  
## 749 southwest  
## 750 northwest  
## 751 southeast  
## 752 northwest  
## 753 northwest  
## 754 southeast  
## 755 northeast  
## 756 northeast  
## 757 northeast  
## 758 southeast  
## 759 northeast  
## 760 southeast  
## 761 northeast  
## 762 southwest  
## 763 southwest  
## 764 northeast  
## 765 northeast  
## 766 northwest  
## 767 southwest  
## 768 southwest  
## 769 southwest  
## 770 northwest  
## 771 southwest  
## 772 southwest  
## 773 northeast  
## 774 northwest  
## 775 northwest  
## 776 southeast  
## 777 northwest  
## 778 northeast  
## 779 southeast  
## 780 northwest  
## 781 southwest  
## 782 southeast  
## 783 southeast  
## 784 southwest  
## 785 southeast  
## 786 southwest  
## 787 northeast  
## 788 northwest  
## 789 northeast  
## 790 southeast  
## 791 southeast  
## 792 southwest  
## 793 northeast  
## 794 southeast  
## 795 northwest  
## 796 northwest  
## 797 southeast  
## 798 northeast  
## 799 southwest  
## 800 northeast  
## 801 southeast  
## 802 southeast  
## 803 southwest  
## 804 southeast  
## 805 southeast  
## 806 northwest  
## 807 northwest  
## 808 northwest  
## 809 southeast  
## 810 northeast  
## 811 southwest  
## 812 northwest  
## 813 southeast  
## 814 northeast  
## 815 southeast  
## 816 southeast  
## 817 northwest  
## 818 southwest  
## 819 northeast  
## 820 northwest  
## 821 southwest  
## 822 northwest  
## 823 southeast  
## 824 southeast  
## 825 northwest  
## 826 northeast  
## 827 southeast  
## 828 northeast  
## 829 northeast  
## 830 northwest  
## 831 southwest  
## 832 northwest  
## 833 northwest  
## 834 northwest  
## 835 northwest  
## 836 southeast  
## 837 southwest  
## 838 northeast  
## 839 northeast  
## 840 northwest  
## 841 southwest  
## 842 northeast  
## 843 southeast  
## 844 southeast  
## 845 northeast  
## 846 southeast  
## 847 southwest  
## 848 southeast  
## 849 southwest  
## 850 northwest  
## 851 northeast  
## 852 northwest  
## 853 northeast  
## 854 northeast  
## 855 northeast  
## 856 southwest  
## 857 southeast  
## 858 northwest  
## 859 southeast  
## 860 southwest  
## 861 southwest  
## 862 southwest  
## 863 northwest  
## 864 northeast  
## 865 southwest  
## 866 southwest  
## 867 southeast  
## 868 southwest  
## 869 northeast  
## 870 southwest  
## 871 southwest  
## 872 southeast  
## 873 southeast  
## 874 southwest  
## 875 northeast  
## 876 northwest  
## 877 southwest  
## 878 southeast  
## 879 southwest  
## 880 southwest  
## 881 southwest  
## 882 northwest  
## 883 northeast  
## 884 northeast  
## 885 northwest  
## 886 southeast  
## 887 northeast  
## 888 northwest  
## 889 southwest  
## 890 northwest  
## 891 northwest  
## 892 southeast  
## 893 northeast  
## 894 southeast  
## 895 northeast  
## 896 southwest  
## 897 northeast  
## 898 northwest  
## 899 southeast  
## 900 northwest  
## 901 northeast  
## 902 southeast  
## 903 northeast  
## 904 southeast  
## 905 southwest  
## 906 northeast  
## 907 northeast  
## 908 southeast  
## 909 southwest  
## 910 southwest  
## 911 northwest  
## 912 northeast  
## 913 northwest  
## 914 southwest  
## 915 northwest  
## 916 southeast  
## 917 northwest  
## 918 northeast  
## 919 southwest  
## 920 southeast  
## 921 southwest  
## 922 southwest  
## 923 southwest  
## 924 northwest  
## 925 southwest  
## 926 northeast  
## 927 southwest  
## 928 southwest  
## 929 southeast  
## 930 southeast  
## 931 southeast  
## 932 southwest  
## 933 southwest  
## 934 southwest  
## 935 southeast  
## 936 southwest  
## 937 northeast  
## 938 northwest  
## 939 southeast  
## 940 southeast  
## 941 southeast  
## 942 southeast  
## 943 northeast  
## 944 northwest  
## 945 southeast  
## 946 southwest  
## 947 southwest  
## 948 northeast  
## 949 northwest  
## 950 southwest  
## 951 northeast  
## 952 southeast  
## 953 northwest  
## 954 southwest  
## 955 northwest  
## 956 southeast  
## 957 southeast  
## 958 northwest  
## 959 northeast  
## 960 northwest  
## 961 northwest  
## 962 southwest  
## 963 southeast  
## 964 northeast  
## 965 northwest  
## 966 southwest  
## 967 northwest  
## 968 northwest  
## 969 northeast  
## 970 southeast  
## 971 southeast  
## 972 northeast  
## 973 northwest  
## 974 southwest  
## 975 southeast  
## 976 northeast  
## 977 southeast  
## 978 southeast  
## 979 northeast  
## 980 southeast  
## 981 northeast  
## 982 northeast  
## 983 southwest  
## 984 northeast  
## 985 northeast  
## 986 southwest  
## 987 northwest  
## 988 northwest  
## 989 northeast  
## 990 northeast  
## 991 southwest  
## 992 northeast  
## 993 southwest  
## 994 southeast  
## 995 northwest  
## 996 northeast  
## 997 southwest  
## 998 southeast  
## 999 northeast  
## 1000 northwest  
## 1001 northwest  
## 1002 southwest  
## 1003 southwest  
## 1004 southwest  
## 1005 northeast  
## 1006 northwest  
## 1007 northeast  
## 1008 northwest  
## 1009 northeast  
## 1010 northeast  
## 1011 southwest  
## 1012 southeast  
## 1013 southeast  
## 1014 northwest  
## 1015 southwest  
## 1016 northwest  
## 1017 northwest  
## 1018 southwest  
## 1019 northwest  
## 1020 northwest  
## 1021 southwest  
## 1022 southeast  
## 1023 southeast  
## 1024 southeast  
## 1025 southeast  
## 1026 southwest  
## 1027 northwest  
## 1028 northwest  
## 1029 southwest  
## 1030 northeast  
## 1031 northwest  
## 1032 southeast  
## 1033 northeast  
## 1034 northeast  
## 1035 northwest  
## 1036 southwest  
## 1037 southeast  
## 1038 northwest  
## 1039 northeast  
## 1040 northwest  
## 1041 northwest  
## 1042 northeast  
## 1043 northeast  
## 1044 southwest  
## 1045 northeast  
## 1046 northwest  
## 1047 northeast  
## 1048 southeast  
## 1049 northwest  
## 1050 southwest  
## 1051 northwest  
## 1052 northeast  
## 1053 northeast  
## 1054 southwest  
## 1055 northwest  
## 1056 northwest  
## 1057 southwest  
## 1058 southeast  
## 1059 southeast  
## 1060 northwest  
## 1061 southeast  
## 1062 southeast  
## 1063 southeast  
## 1064 northwest  
## 1065 southwest  
## 1066 southwest  
## 1067 southeast  
## 1068 northeast  
## 1069 northwest  
## 1070 southeast  
## 1071 southeast  
## 1072 northeast  
## 1073 northwest  
## 1074 northeast  
## 1075 northeast  
## 1076 southeast  
## 1077 southwest  
## 1078 northeast  
## 1079 southeast  
## 1080 southeast  
## 1081 southeast  
## 1082 northwest  
## 1083 northwest  
## 1084 southwest  
## 1085 northwest  
## 1086 southwest  
## 1087 northeast  
## 1088 northwest  
## 1089 southeast  
## 1090 southwest  
## 1091 southeast  
## 1092 northeast  
## 1093 southwest  
## 1094 northwest  
## 1095 southwest  
## 1096 northeast  
## 1097 northeast  
## 1098 southeast  
## 1099 northeast  
## 1100 southeast  
## 1101 northeast  
## 1102 southwest  
## 1103 southeast  
## 1104 southeast  
## 1105 southwest  
## 1106 southeast  
## 1107 northwest  
## 1108 northwest  
## 1109 southwest  
## 1110 southeast  
## 1111 northeast  
## 1112 southeast  
## 1113 southeast  
## 1114 northwest  
## 1115 northeast  
## 1116 southeast  
## 1117 northeast  
## 1118 southeast  
## 1119 southeast  
## 1120 northwest  
## 1121 southwest  
## 1122 southeast  
## 1123 northwest  
## 1124 northeast  
## 1125 northeast  
## 1126 northwest  
## 1127 southwest  
## 1128 southeast  
## 1129 southwest  
## 1130 southwest  
## 1131 southeast  
## 1132 southwest  
## 1133 northeast  
## 1134 northwest  
## 1135 northwest  
## 1136 northwest  
## 1137 southwest  
## 1138 northwest  
## 1139 southeast  
## 1140 northwest  
## 1141 southeast  
## 1142 southwest  
## 1143 southeast  
## 1144 southeast  
## 1145 southwest  
## 1146 northwest  
## 1147 southwest  
## 1148 northwest  
## 1149 southwest  
## 1150 southwest  
## 1151 northeast  
## 1152 northwest  
## 1153 southeast  
## 1154 northwest  
## 1155 northwest  
## 1156 northeast  
## 1157 southeast  
## 1158 northwest  
## 1159 northeast  
## 1160 southwest  
## 1161 northwest  
## 1162 southeast  
## 1163 southeast  
## 1164 northeast  
## 1165 northwest  
## 1166 northeast  
## 1167 southeast  
## 1168 southwest  
## 1169 southwest  
## 1170 northwest  
## 1171 northeast  
## 1172 southwest  
## 1173 southeast  
## 1174 northwest  
## 1175 northwest  
## 1176 southwest  
## 1177 northwest  
## 1178 southwest  
## 1179 northeast  
## 1180 southeast  
## 1181 northeast  
## 1182 northwest  
## 1183 southwest  
## 1184 northeast  
## 1185 southeast  
## 1186 northeast  
## 1187 northwest  
## 1188 northwest  
## 1189 northeast  
## 1190 southwest  
## 1191 northwest  
## 1192 northeast  
## 1193 northeast  
## 1194 northwest  
## 1195 northwest  
## 1196 northwest  
## 1197 northwest  
## 1198 southeast  
## 1199 northwest  
## 1200 southwest  
## 1201 northwest  
## 1202 northwest  
## 1203 northwest  
## 1204 northeast  
## 1205 southeast  
## 1206 northwest  
## 1207 southwest  
## 1208 southwest  
## 1209 northeast  
## 1210 southwest  
## 1211 northwest  
## 1212 southeast  
## 1213 northeast  
## 1214 southwest  
## 1215 northwest  
## 1216 northeast  
## 1217 southeast  
## 1218 southeast  
## 1219 southwest  
## 1220 northwest  
## 1221 northeast  
## 1222 southeast  
## 1223 southeast  
## 1224 southeast  
## 1225 northeast  
## 1226 southeast  
## 1227 northeast  
## 1228 southeast  
## 1229 southeast  
## 1230 northeast  
## 1231 northwest  
## 1232 southwest  
## 1233 northwest  
## 1234 southwest  
## 1235 southeast  
## 1236 northwest  
## 1237 northeast  
## 1238 northwest  
## 1239 northeast  
## 1240 southeast  
## 1241 southeast  
## 1242 southeast  
## 1243 northwest  
## 1244 southeast  
## 1245 southeast  
## 1246 southwest  
## 1247 southwest  
## 1248 southwest  
## 1249 southeast  
## 1250 northeast  
## 1251 northeast  
## 1252 southwest  
## 1253 southwest  
## 1254 southwest  
## 1255 southeast  
## 1256 southwest  
## 1257 northwest  
## 1258 northwest  
## 1259 northwest  
## 1260 northeast  
## 1261 northeast  
## 1262 southwest  
## 1263 southeast  
## 1264 southwest  
## 1265 northeast  
## 1266 southeast  
## 1267 southwest  
## 1268 northeast  
## 1269 southwest  
## 1270 southwest  
## 1271 northwest  
## 1272 northwest  
## 1273 southeast  
## 1274 southeast  
## 1275 southeast  
## 1276 southwest  
## 1277 northeast  
## 1278 northwest  
## 1279 northeast  
## 1280 northwest  
## 1281 southeast  
## 1282 northwest  
## 1283 northeast  
## 1284 southeast  
## 1285 southwest  
## 1286 northeast  
## 1287 northeast  
## 1288 southwest  
## 1289 southwest  
## 1290 southeast  
## 1291 northeast  
## 1292 southwest  
## 1293 southeast  
## 1294 northwest  
## 1295 northeast  
## 1296 southwest  
## 1297 northeast  
## 1298 southeast  
## 1299 northwest  
## 1300 northwest  
## 1301 southeast  
## 1302 northwest  
## 1303 southwest  
## 1304 southwest  
## 1305 northeast  
## 1306 southeast  
## 1307 northeast  
## 1308 northwest  
## 1309 southwest  
## 1310 southwest  
## 1311 northwest  
## 1312 northwest  
## 1313 southwest  
## 1314 southwest  
## 1315 northwest  
## 1316 northeast  
## 1317 southwest  
## 1318 southeast  
## 1319 northeast  
## 1320 northwest  
## 1321 northwest  
## 1322 northeast  
## 1323 southeast  
## 1324 southeast  
## 1325 northwest  
## 1326 northeast  
## 1327 northeast  
## 1328 southeast  
## 1329 northeast  
## 1330 southwest  
## 1331 southeast  
## 1332 southwest  
## 1333 southwest  
## 1334 northwest  
## 1335 northeast  
## 1336 southeast  
## 1337 southwest  
## 1338 northwest

#There is only one observation which repeat twice, it makes sense that a person with the same properties will have the same charge and since it's only one we decide to leave it there.  
#outliers

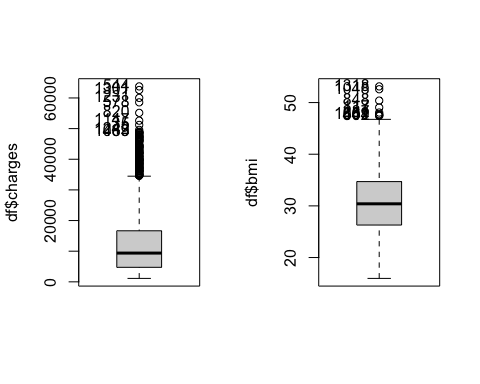
#### Outlier detection

##### Univariate

par(mfrow=c(1,2))  
Boxplot(df$charges)

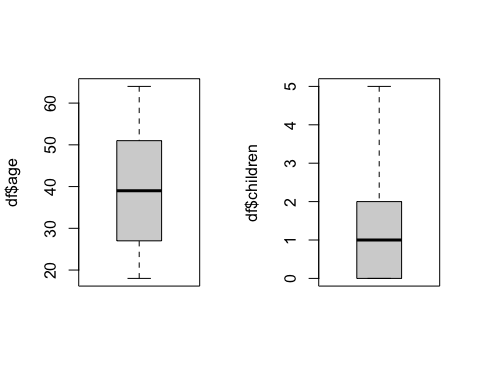
## [1] 544 1301 1231 578 820 1147 35 1242 1063 489

Boxplot(df$bmi)



## [1] 117 287 402 544 848 861 1048 1089 1318

Boxplot(df$age)  
Boxplot(df$children)



#  
  
# treat outliers for charges variable  
sevout<-quantile(df$charges,0.75,na.rm=TRUE)+3\*(quantile(df$charges,0.75,na.rm=TRUE)-quantile(df$charges,0.25,na.rm=TRUE))  
sevout

## 75%   
## 52338.79

sev\_out\_lower <- quantile(df$charges,0.25,na.rm=TRUE)-3\*(quantile(df$charges,0.75,na.rm=TRUE)-quantile(df$charges,0.25,na.rm=TRUE))  
  
mist<-quantile(df$charges,0.75,na.rm=TRUE)+1.5\*(quantile(df$charges,0.75,na.rm=TRUE)-quantile(df$charges,0.25,na.rm=TRUE))  
mist

## 75%   
## 34489.35

mist\_out\_lower <- quantile(df$charges,0.25,na.rm=TRUE)-1.5\*(quantile(df$charges,0.75,na.rm=TRUE)-quantile(df$charges,0.25,na.rm=TRUE))  
  
# get list of outliers  
loutse<-which(df$charges>sevout);length(loutse)

## [1] 6

loutmist <-which(df$charges>mist);length(loutmist)

## [1] 139

low\_out\_sever <- which(df$charges<sev\_out\_lower);low\_out\_sever

## integer(0)

low\_out\_mild <- which(df$charges<mist\_out\_lower);low\_out\_mild

## integer(0)

#table(loutse)  
#table(loutmist)  
  
# see outliers  
Boxplot(df$charges)

## [1] 544 1301 1231 578 820 1147 35 1242 1063 489

abline(h=sevout,col="red")  
abline(h=mist,col="yellow")  
  
# Since there are only 6 severe outliers, we will remove them from the dataset,   
df <- df[-which(df$charges >= sevout),]  
  
# check severe outliers for bmi atrribute  
sevout\_bmi<-quantile(df$bmi,0.75,na.rm=TRUE)+3\*(quantile(df$bmi,0.75,na.rm=TRUE)-quantile(df$bmi,0.25,na.rm=TRUE));sevout\_bmi

## 75%   
## 59.815

mist\_bmi <- quantile(df$bmi,0.75,na.rm=TRUE)+1.5\*(quantile(df$bmi,0.75,na.rm=TRUE)-quantile(df$bmi,0.25,na.rm=TRUE))  
loutse\_bmi<-which(df$bmi>sevout\_bmi);length(loutse\_bmi) # no severe outliers for bmi

## [1] 0

colSums(is.na(df))

## age sex bmi children smoker region charges f.sex   
## 0 0 0 0 0 0 0 0   
## f.smoker f.region   
## 0 0

serout\_lower\_bmi <- quantile(df$bmi,0.25,na.rm=TRUE)-3\*(quantile(df$bmi,0.75,na.rm=TRUE)-quantile(df$bmi,0.25,na.rm=TRUE));serout\_lower\_bmi

## 25%   
## 1.02375

mist\_lower\_bmi <- quantile(df$bmi,0.25,na.rm=TRUE)-1.5\*(quantile(df$bmi,0.75,na.rm=TRUE)-quantile(df$bmi,0.25,na.rm=TRUE));mist\_lower\_bmi

## 25%   
## 13.62187

up\_sever\_bmi <- which(df$bmi > sevout\_bmi); up\_sever\_bmi

## integer(0)

up\_mild\_bmi <- which(df$bmi > mist\_bmi); up\_mild\_bmi

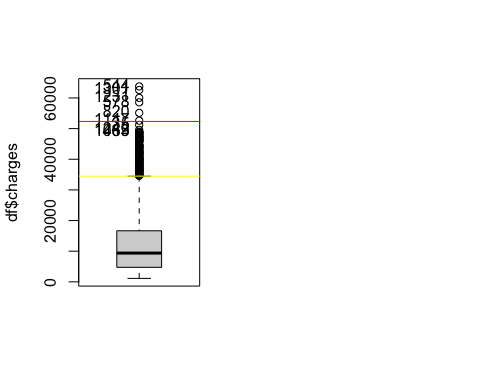
## [1] 117 287 402 845 858 1045 1086 1312

low\_sever\_bmi <- which(df$bmi < serout\_lower\_bmi); low\_sever\_bmi

## integer(0)

low\_mild\_bmi <- which(df$bmi < mist\_lower\_bmi); low\_mild\_bmi

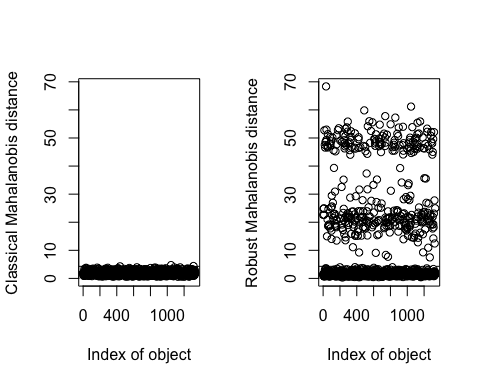
## integer(0)



We can see extreme outliers for both charges and bmi, since it’s just several observation it might be the case that for a certain bmi, age or smokers the charge value is raising by a lot compare to the rest. from looking at the high value of column charges it can be seen that all are smokers and mid-high bmi, also some of the ages I see are relatively high. For the target variable we can see there is no lower bound for extreme and mild outliers, it’s also can be seen on the Boxplot(). For variable bmi, mild outliers on the upper bound and no sever upper bound outliers and not lower bound outliers. We decided to delete the 6 univariate outliers since the charges are very high, even though all 6 observation are smokers, there are 274 smokers in the dataset and their charges values are not as high as the extreme outliers observations

##### Multivariate

res.out<-Moutlier(df[,c(7,3,1,4)],quantile=0.999)



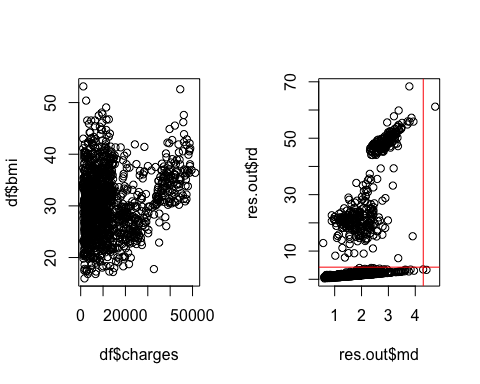
#str(res.out)  
plot(df$charges,df$bmi)  
res.out$cutoff

## [1] 4.297305

#quantile(res.out$md,seq(0,1,0.001))  
which((res.out$md > res.out$cutoff) & (res.out$rd > res.out$cutoff))

## 1048   
## 1045

plot( res.out$md, res.out$rd )  
#text(res.out$md, res.out$rd, labels=rownames(df),adj=1, cex=0.5)  
abline(h=res.out$cutoff, col="red")  
abline(v=res.out$cutoff, col="red")



df <- df[-which(res.out$md > res.out$cutoff & res.out$rd > res.out$cutoff),]

For the multivariate outliers, we have chosen the quantile to be a very high value so outliers we get are very extreme compare to our values in the dataset. Observation number 1048 is the multivatiate outlier we have got and it’s indeed a very high value of charge and bmi. Since this observation is so extreme we will remove it from the dataset. We see from the plot classical Mahalanobis distance vs robust Mahalanobis distance that there is one observation (1048) that is behind the cutoff value, in addition we can indicate 3 clusters and number of observations that a bit far from the clusters, it can be suspected as influential data. We also plot charges vs bmi and we can see on the top right corner of the graph there is one observation which has high charge and bmi.

#### Missing data

# check missing data  
  
colSums(is.na(df))

## age sex bmi children smoker region charges f.sex   
## 0 0 0 0 0 0 0 0   
## f.smoker f.region   
## 0 0

There is no missing data in the dataframe so no further imputation is needed

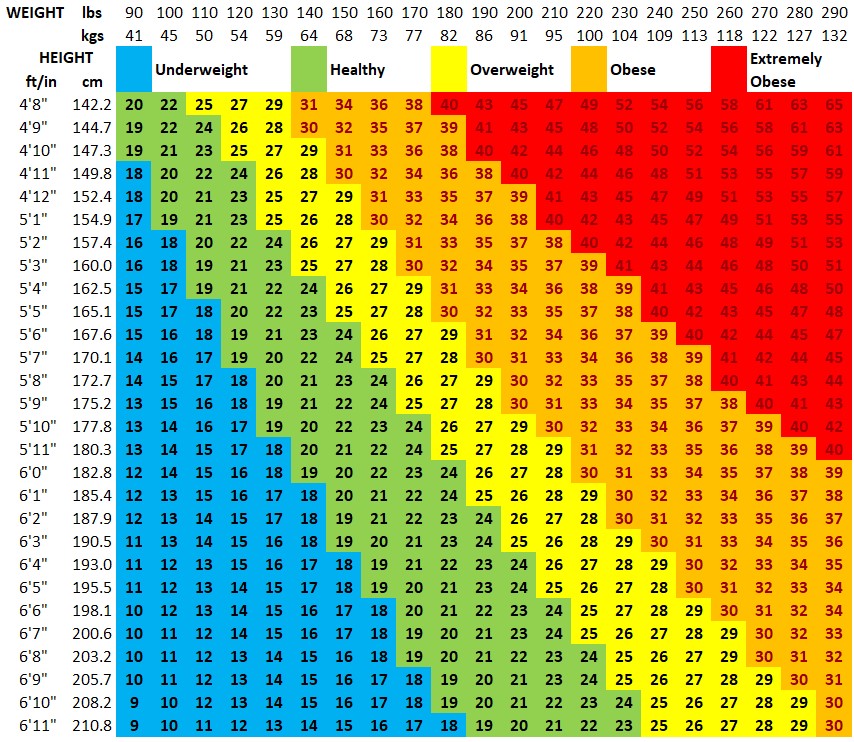
#### Data Validation

After doing the pre processing steps where we detected and removed outliers, we will check if data makes sense using common sense and domain knowledge.

summary(df)

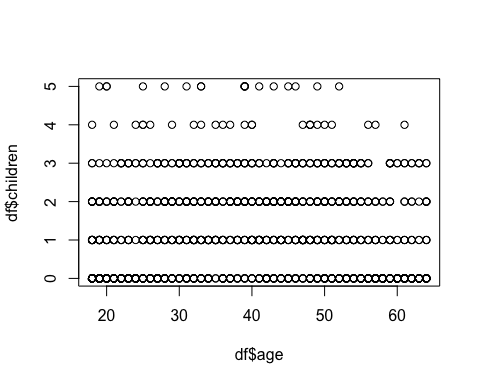
## age sex bmi children   
## Min. :18.00 Length:1331 Min. :15.96 Min. :0.000   
## 1st Qu.:26.50 Class :character 1st Qu.:26.22 1st Qu.:0.000   
## Median :39.00 Mode :character Median :30.30 Median :1.000   
## Mean :39.19 Mean :30.62 Mean :1.097   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.000   
## Max. :64.00 Max. :53.13 Max. :5.000   
## smoker region charges f.sex f.smoker   
## Length:1331 Length:1331 Min. : 1122 female:659 no :1064   
## Class :character Class :character 1st Qu.: 4720 male :672 yes: 267   
## Mode :character Mode :character Median : 9302   
## Mean :13042   
## 3rd Qu.:16359   
## Max. :51195   
## f.region   
## northeast:323   
## northwest:323   
## southeast:361   
## southwest:324   
##   
##

We have ages ranging from 18 to 64, and which bmi ranging from 16 to 53 which are values that are in the following table. The balance between factor variable is really good. However, only 20% of the sample are smokers.



Let’s see how the relationship between children per age.

plot(df$children~df$age)



As we can see in the plot, there are individuals with age 20 that have from 3 to 5 children which is really strange.

thr2five\_children <- which(df$age <= 20 & df$children>2)  
  
thr2five\_children

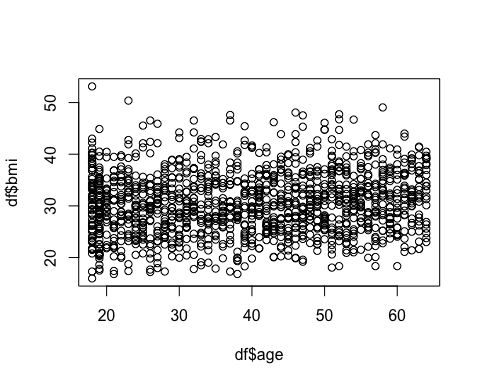
## [1] 33 167 370 982 1092 1182 1191 1200

This observations will be removed since it’s something very unlikely.

df <- df[-thr2five\_children,]

Let’s check now the bmi values per age to see if there is any weird case:

plot(df$bmi~df$age)



In this case the plot shows there are young people who have a really high bmi. Since data is from EEUU, and there are a lof of obesity problems, we decide that these observations are not going to be removed.

summary(df)

## age sex bmi children   
## Min. :18.00 Length:1323 Min. :15.96 Min. :0.00   
## 1st Qu.:27.00 Class :character 1st Qu.:26.22 1st Qu.:0.00   
## Median :39.00 Mode :character Median :30.30 Median :1.00   
## Mean :39.31 Mean :30.62 Mean :1.08   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.00   
## Max. :64.00 Max. :53.13 Max. :5.00   
## smoker region charges f.sex f.smoker   
## Length:1323 Length:1323 Min. : 1122 female:654 no :1058   
## Class :character Class :character 1st Qu.: 4729 male :669 yes: 265   
## Mode :character Mode :character Median : 9305   
## Mean :13047   
## 3rd Qu.:16265   
## Max. :51195   
## f.region   
## northeast:320   
## northwest:321   
## southeast:360   
## southwest:322   
##   
##

### Explanatory data analysis

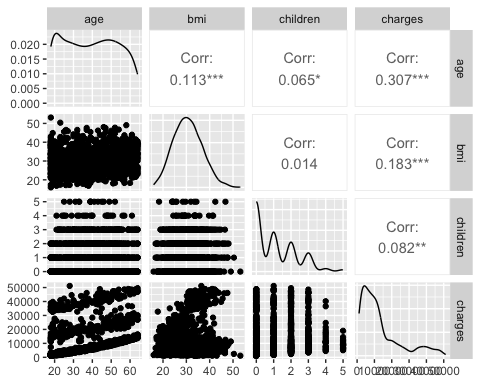
summary(df)

## age sex bmi children   
## Min. :18.00 Length:1323 Min. :15.96 Min. :0.00   
## 1st Qu.:27.00 Class :character 1st Qu.:26.22 1st Qu.:0.00   
## Median :39.00 Mode :character Median :30.30 Median :1.00   
## Mean :39.31 Mean :30.62 Mean :1.08   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.00   
## Max. :64.00 Max. :53.13 Max. :5.00   
## smoker region charges f.sex f.smoker   
## Length:1323 Length:1323 Min. : 1122 female:654 no :1058   
## Class :character Class :character 1st Qu.: 4729 male :669 yes: 265   
## Mode :character Mode :character Median : 9305   
## Mean :13047   
## 3rd Qu.:16265   
## Max. :51195   
## f.region   
## northeast:320   
## northwest:321   
## southeast:360   
## southwest:322   
##   
##

#numeric variables  
summary(df[,c(1,3,4,7)])

## age bmi children charges   
## Min. :18.00 Min. :15.96 Min. :0.00 Min. : 1122   
## 1st Qu.:27.00 1st Qu.:26.22 1st Qu.:0.00 1st Qu.: 4729   
## Median :39.00 Median :30.30 Median :1.00 Median : 9305   
## Mean :39.31 Mean :30.62 Mean :1.08 Mean :13047   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.00 3rd Qu.:16265   
## Max. :64.00 Max. :53.13 Max. :5.00 Max. :51195

#plot(df[,c(1,3,4,7)])  
ggpairs(df[,c(1,3,4,7)])



#categorical variables  
summary(df[,c(1,4,8:10)])

## age children f.sex f.smoker f.region   
## Min. :18.00 Min. :0.00 female:654 no :1058 northeast:320   
## 1st Qu.:27.00 1st Qu.:0.00 male :669 yes: 265 northwest:321   
## Median :39.00 Median :1.00 southeast:360   
## Mean :39.31 Mean :1.08 southwest:322   
## 3rd Qu.:51.00 3rd Qu.:2.00   
## Max. :64.00 Max. :5.00

From the summary we can see the factor values, it seems that sex and region are distributed equally and not much smokers compare to the non smokers. age and number of children looks about right and there is values in a range that makes sense. In addition, we see low correlation (0.198) between the target variable and the other numeric explantory variable bmi. We don’t see any pattern in the relation between the two variables. We see number of extreme values with high bmi and/or charges.

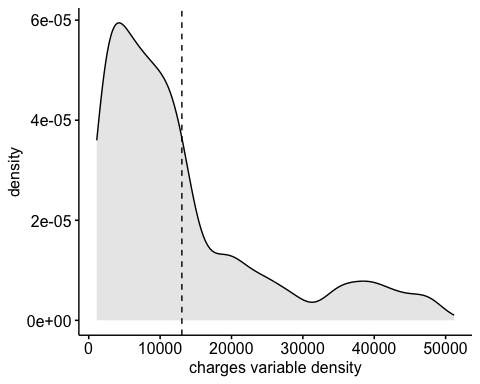
* Determine if the response variable (charges) has an acceptably normal distribution.

# Density plot to check the distribution  
ggpubr::ggdensity(df$charges, fill = "lightgray", add = "mean", xlab = "charges variable density")

## Warning: `geom\_vline()`: Ignoring `mapping` because `xintercept` was provided.

## Warning: `geom\_vline()`: Ignoring `data` because `xintercept` was provided.

## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.  
## ℹ Please use `after\_stat(density)` instead.  
## ℹ The deprecated feature was likely used in the ggpubr package.  
## Please report the issue at <]8;;https://github.com/kassambara/ggpubr/issueshttps://github.com/kassambara/ggpubr/issues]8;;>.



# Shapiro Test to asses that data on response variable is normaly distribution  
# H0 = Data is normally distributed  
# H1 = Data is not normally distributed  
# alfa = 0.05  
shapiro.test(df$charges)

##   
## Shapiro-Wilk normality test  
##   
## data: df$charges  
## W = 0.81754, p-value < 2.2e-16

As we can see, the density plot shows that data is not normally distributed. To asses that, we can use one of many statistical tests that check normality on data. In this case, we use Shapiro test.

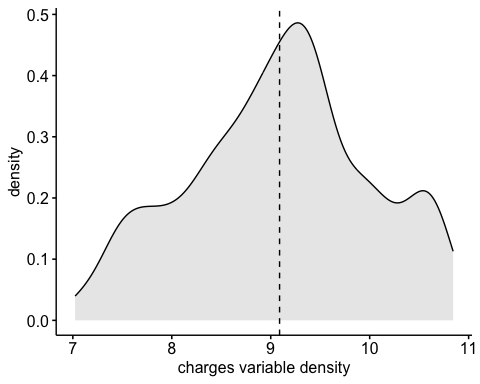
The result of the Shapiro test shows that data in variable **charges** is not normally distributed since *p-value* is less than the significance level (0.05) so we reject the null hypothesis (data is normally distributed) and we conclude that data is not normally distributed (alternative hypothesis)

Let’s try to apply the log transformation

# Density plot to check the distribution  
ggpubr::ggdensity(log(df$charges), fill = "lightgray", add = "mean", xlab = "charges variable density")

## Warning: `geom\_vline()`: Ignoring `mapping` because `xintercept` was provided.

## Warning: `geom\_vline()`: Ignoring `data` because `xintercept` was provided.

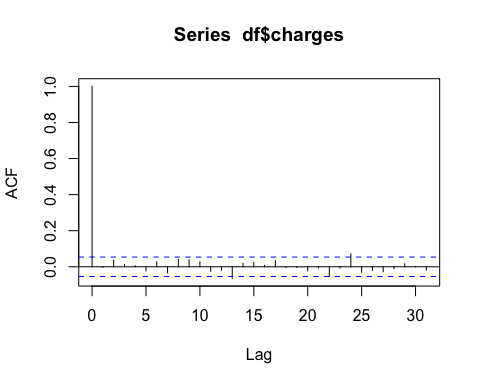


# Shapiro Test to asses that data on response variable is normaly distribution  
# H0 = Data is normally distributed  
# H1 = Data is not normally distributed  
# alfa = 0.05  
shapiro.test(log(df$charges))

##   
## Shapiro-Wilk normality test  
##   
## data: log(df$charges)  
## W = 0.98152, p-value = 5.679e-12

The null hypothesis can be still rejected so data still not being not normally distributed.

par(mfrow=c(1,1))  
acf(df$charges)



dwtest(df$charges~1)

##   
## Durbin-Watson test  
##   
## data: df$charges ~ 1  
## DW = 2.0054, p-value = 0.5394  
## alternative hypothesis: true autocorrelation is greater than 0

Address tests to discard serial correlation: In the acf (auto correlation function) we can see from the graph that the data is not correlated where we have the blue threshold and all lines are within the threshold, we do see that there is one or two lines that crosses the threshold but just in a little bit so we leave it as it is without random the order of the observations. In addition we address Durbin-Watson test to check whether true autocorrelation is greater or not than 0. We see p-value 0.5183, thus we don’t reject the null hypothesis and say that true autocorrelation is not greater than 0.

#library(DataExplorer)  
#create\_report(df, y= "charges")  
  
library(FactoMineR)  
res.con <- condes(df[,c(1,3,4,7,8:10)], num.var = 4 , proba = 0.01 )  
res.con$quanti

## correlation p.value  
## age 0.30679657 3.128392e-30  
## bmi 0.18280602 2.091908e-11  
## children 0.08239851 2.705520e-03

res.con$quali

## R2 p.value  
## f.smoker 0.6169962 1.418037e-277

res.con$category

## Estimate p.value  
## f.smoker=yes 11493.36 1.418037e-277  
## f.smoker=no -11493.36 1.418037e-277

Association to the target variable, we see the numeric variable age 0.301 which is the most associated but the number is quite low and it is not strong association. Following age, we have bmi and then children.

For categorical variables we see that f.smoker is globally associated to charges, in particular, f.smoker=yes is very remarkable. Let’s check the case of smoker category.

res.cat <- catdes(df[,c(9,3,4,7,8, 1, 10)], num.var = 1 , proba = 0.01 )  
  
res.cat$quanti

## $no  
## v.test Mean in category Overall mean sd in category Overall sd  
## charges -28.55992 8443.049 13047.35 5993.179 11712.29  
## p.value  
## charges 2.115342e-179  
##   
## $yes  
## v.test Mean in category Overall mean sd in category Overall sd  
## charges 28.55992 31429.78 13047.35 10904.12 11712.29  
## p.value  
## charges 2.115342e-179

res.cat$category

## $no  
## Cla/Mod Mod/Cla Global p.value v.test  
## f.sex=female 83.02752 51.32325 49.43311 0.006035893 2.745825  
## f.sex=male 76.98057 48.67675 50.56689 0.006035893 -2.745825  
##   
## $yes  
## Cla/Mod Mod/Cla Global p.value v.test  
## f.sex=male 23.01943 58.11321 50.56689 0.006035893 2.745825  
## f.sex=female 16.97248 41.88679 49.43311 0.006035893 -2.745825

res.cat$quanti.var

## Eta2 P-value  
## charges 0.6169962 1.418037e-277

We can see that the mean of charges for smokers is much more higher than people who don’t smoke. Smoking seems a very important influence in the price for having high insurance charges.

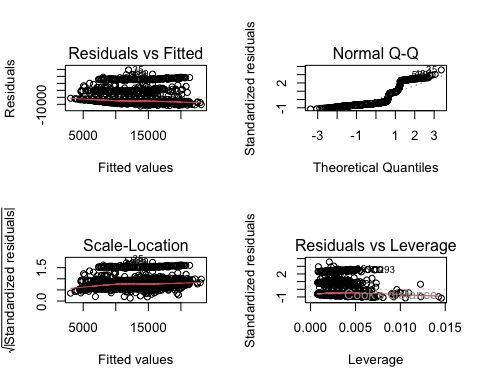
### Building the model

#### First model

m1<-lm(charges~bmi+age+children, data = df)  
summary(m1)

##   
## Call:  
## lm(formula = charges ~ bmi + age + children, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12628 -6735 -5057 5894 39232   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5853.05 1710.99 -3.421 0.000643 \*\*\*  
## bmi 288.72 50.14 5.758 1.06e-08 \*\*\*  
## age 239.14 21.79 10.977 < 2e-16 \*\*\*  
## children 610.04 255.28 2.390 0.017003 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 11000 on 1319 degrees of freedom  
## Multiple R-squared: 0.1202, Adjusted R-squared: 0.1182   
## F-statistic: 60.05 on 3 and 1319 DF, p-value: < 2.2e-16

par(mfrow=c(2,2))  
plot(m1)



par(mfrow=c(1,1))

Looking at the summary of the model, the RSquared is very low and there is a lot of residual standard error.

If we study the residual error looking at the plots we can see that the data is not following a normal distribution since there are deviations of the line (Normal Q-Q plot). Also there are a lot of sparsity in the variance (Scale-Location plot).

#### Asses multicollinearity

Maybe there is multicollinearity that is causing bad results

car::vif(m1)

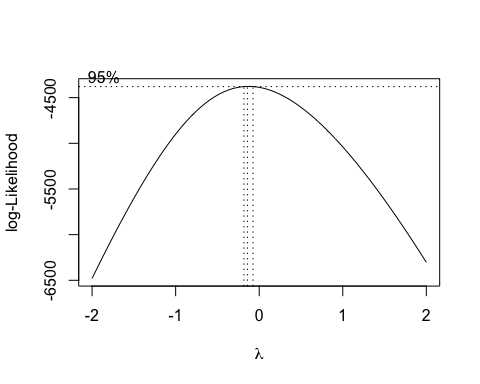
## bmi age children   
## 1.012957 1.017013 1.004239

The vif values are low (less than 5) so there aren’t problems of multidisciplinary.

Let’s try to do some transformations to the data.

#### Transformation

library(MASS)  
  
boxcox(charges~bmi+age+children, data = df)



The boxplots shows that the lambda values are close to 0 so a logarithmic transformation to the target variable should help to improve the results

# (only for numerical variables)  
  
boxTidwell(log(charges) ~ bmi + age + I(children+0.5), data=df)

## MLE of lambda Score Statistic (z) Pr(>|z|)   
## bmi -1.07828 -1.4110 0.15824   
## age 0.42692 -1.7687 0.07694 .  
## I(children + 0.5) 0.25004 -1.7969 0.07235 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## iterations = 16

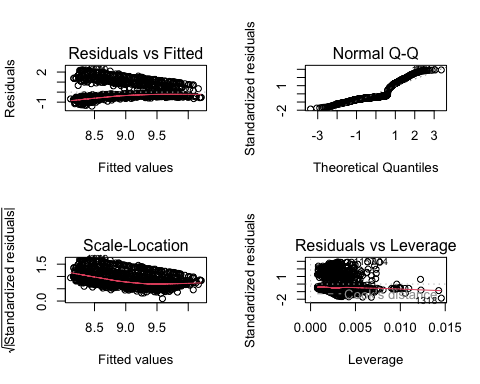
# poly(age,3) for adding ortogonal polynomial

The transformations of the explanatory variables are not performed since all p-values are above 0.05 significance level.

m2 <- lm(log(charges)~bmi+ age+children, data = df)  
  
summary(m2)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.3991 -0.4339 -0.3051 0.4777 2.2134   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.360187 0.117968 62.392 < 2e-16 \*\*\*  
## bmi 0.009188 0.003457 2.657 0.00797 \*\*   
## age 0.033885 0.001502 22.558 < 2e-16 \*\*\*  
## children 0.107190 0.017601 6.090 1.48e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7586 on 1319 degrees of freedom  
## Multiple R-squared: 0.3105, Adjusted R-squared: 0.309   
## F-statistic: 198 on 3 and 1319 DF, p-value: < 2.2e-16

par(mfrow=c(2,2))  
plot(m2)

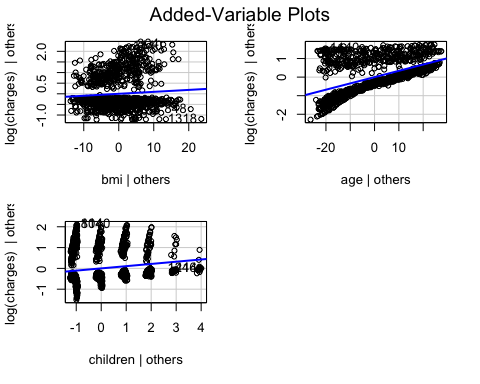


par(mfrow=c(1,1))

The model is still not performing very well. However if we check the study of residuals we can see that it results in an improvement.

The normal Q-Q plot still have a deviation but is that big as the m1 and if we check the Scale-Location of the standard residuals the variance is better.

avPlots(m2)



The partial regressions plots shows that all regresors have two big clusters of data.

AIC(m1,m2)

## df AIC  
## m1 5 28383.903  
## m2 5 3029.469

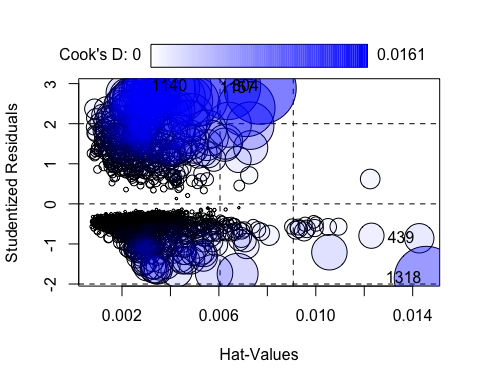
The AIC test shows that model 2 is performing much better than model 1 so we will continue with it.

#### Inlfuential data

Maybe, removing influential data the results can be improved.

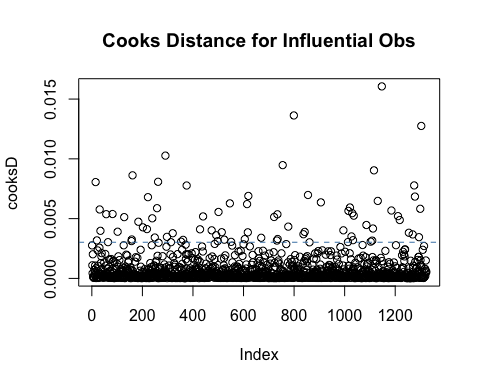
* Residual outliers
* Influential values

library(car)  
  
influencePlot(m2)



## StudRes Hat CookD  
## 439 -0.8586974 0.014270736 0.002669288  
## 804 2.9280088 0.006355073 0.013629744  
## 1140 2.9305989 0.003026175 0.006479950  
## 1157 2.8840078 0.007706514 0.016060091  
## 1318 -1.8596140 0.014562223 0.012751918

# there are a lot of influential data  
  
  
# With cooks distance  
cooksD <- cooks.distance(m2)  
n <- nrow(df)  
plot(cooksD, main = "Cooks Distance for Influential Obs")  
abline(h = 4/n, lty = 2, col = "steelblue") # add cutoff line



influential\_obs <- as.numeric(names(cooksD)[(cooksD > (4/n))])  
influential\_obs

## [1] 15 20 31 35 58 65 83 103 129 158 159 162 186 204 220  
## [16] 224 241 251 260 264 293 299 315 322 355 363 378 431 443 477  
## [31] 495 501 504 517 527 550 555 610 619 622 624 675 726 737 739  
## [46] 740 760 782 804 843 848 861 912 1002 1020 1022 1028 1034 1037 1040  
## [61] 1043 1094 1112 1118 1121 1125 1140 1157 1197 1224 1232 1268 1283 1289 1292  
## [76] 1309 1314 1318

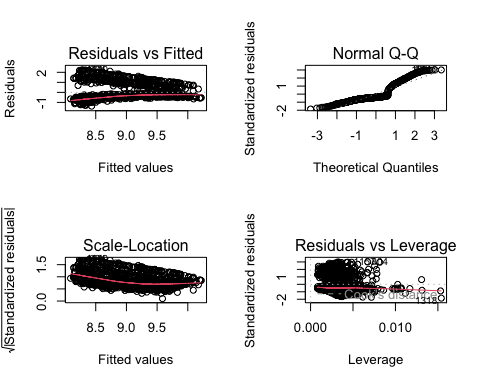
length(influential\_obs)

## [1] 78

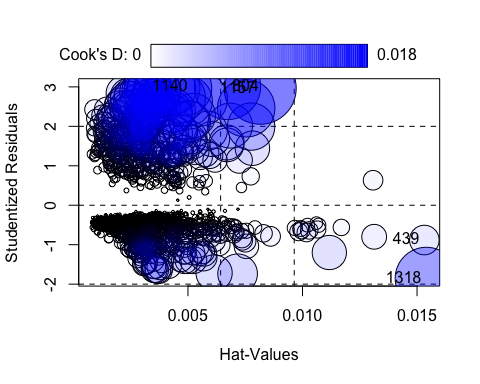
m3 <- lm(log(charges)~bmi+age+children, data=df[-influential\_obs,])  
summary(m3)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children, data = df[-influential\_obs,   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.3732 -0.4262 -0.3001 0.4490 2.2392   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.320468 0.120584 60.708 < 2e-16 \*\*\*  
## bmi 0.009217 0.003512 2.624 0.00879 \*\*   
## age 0.034569 0.001524 22.691 < 2e-16 \*\*\*  
## children 0.109175 0.017948 6.083 1.57e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.747 on 1241 degrees of freedom  
## Multiple R-squared: 0.3241, Adjusted R-squared: 0.3225   
## F-statistic: 198.4 on 3 and 1241 DF, p-value: < 2.2e-16

par(mfrow=c(2,2))  
plot(m3)



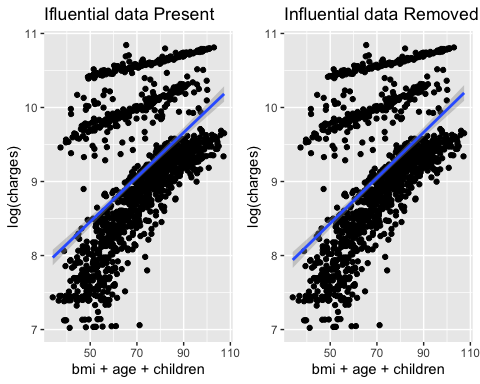
par(mfrow=c(1,1))  
  
influencePlot(m3)



## StudRes Hat CookD  
## 439 -0.8822072 0.015318123 0.003027391  
## 804 3.0105579 0.006725594 0.015243428  
## 1140 3.0123385 0.003233415 0.007311376  
## 1157 2.9649173 0.008153223 0.017952793  
## 1318 -1.8545683 0.015404544 0.013426531

#create scatterplot with influential data present  
outliers\_present <- ggplot(data = df, aes(x = bmi + age + children, y = log(charges))) +  
 geom\_point() +  
 geom\_smooth(method = lm) +  
# ylim(0, 200) +  
 ggtitle("Ifluential data Present")  
  
#create scatterplot with influential data removed  
outliers\_removed <- ggplot(data = df[-influential\_obs,], aes(x = bmi + age + children, y = log(charges))) +  
 geom\_point() +  
 geom\_smooth(method = lm) +  
# ylim(0, 200) +  
 ggtitle("Influential data Removed")  
  
#plot both scatterplots side by side  
gridExtra::grid.arrange(outliers\_present, outliers\_removed, ncol = 2)

## `geom\_smooth()` using formula = 'y ~ x'  
## `geom\_smooth()` using formula = 'y ~ x'



#### Adding factors

* Check that meaning of a factor could not be related to the numerical variables so one should be used.
* AIC test to compare

summary(df)

## age sex bmi children   
## Min. :18.00 Length:1323 Min. :15.96 Min. :0.00   
## 1st Qu.:27.00 Class :character 1st Qu.:26.22 1st Qu.:0.00   
## Median :39.00 Mode :character Median :30.30 Median :1.00   
## Mean :39.31 Mean :30.62 Mean :1.08   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.00   
## Max. :64.00 Max. :53.13 Max. :5.00   
## smoker region charges f.sex f.smoker   
## Length:1323 Length:1323 Min. : 1122 female:654 no :1058   
## Class :character Class :character 1st Qu.: 4729 male :669 yes: 265   
## Mode :character Mode :character Median : 9305   
## Mean :13047   
## 3rd Qu.:16265   
## Max. :51195   
## f.region   
## northeast:320   
## northwest:321   
## southeast:360   
## southwest:322   
##   
##

m4 <- lm(log(charges)~bmi+age+children+f.sex+f.smoker+f.region, data=df[-influential\_obs,])  
summary(m4)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children + f.sex + f.smoker +   
## f.region, data = df[-influential\_obs, ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.05304 -0.19908 -0.05181 0.05959 2.16809   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.0366548 0.0752441 93.518 < 2e-16 \*\*\*  
## bmi 0.0124863 0.0021630 5.773 9.86e-09 \*\*\*  
## age 0.0349941 0.0008989 38.929 < 2e-16 \*\*\*  
## children 0.1023007 0.0105952 9.655 < 2e-16 \*\*\*  
## f.sexmale -0.0746327 0.0250596 -2.978 0.002956 \*\*   
## f.smokeryes 1.5244571 0.0316950 48.098 < 2e-16 \*\*\*  
## f.regionnorthwest -0.0613902 0.0357542 -1.717 0.086229 .   
## f.regionsoutheast -0.1431207 0.0359860 -3.977 7.38e-05 \*\*\*  
## f.regionsouthwest -0.1267455 0.0360563 -3.515 0.000455 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4402 on 1236 degrees of freedom  
## Multiple R-squared: 0.7662, Adjusted R-squared: 0.7647   
## F-statistic: 506.4 on 8 and 1236 DF, p-value: < 2.2e-16

Let’s try to check if there are factors that could be removed

Anova(m4)

## Anova Table (Type II tests)  
##   
## Response: log(charges)  
## Sum Sq Df F value Pr(>F)   
## bmi 6.46 1 33.3225 9.863e-09 \*\*\*  
## age 293.68 1 1515.5052 < 2.2e-16 \*\*\*  
## children 18.07 1 93.2266 < 2.2e-16 \*\*\*  
## f.sex 1.72 1 8.8698 0.0029559 \*\*   
## f.smoker 448.30 1 2313.3955 < 2.2e-16 \*\*\*  
## f.region 3.80 3 6.5320 0.0002206 \*\*\*  
## Residuals 239.52 1236   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

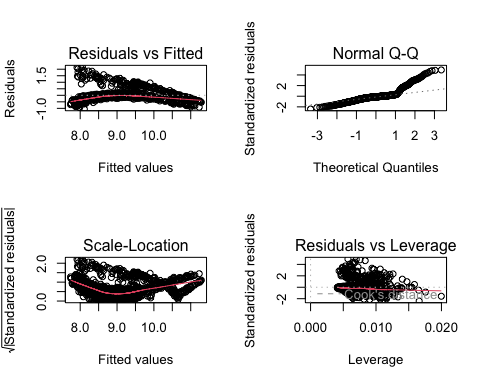
m5 <- step( m4 )

## Start: AIC=-2034.08  
## log(charges) ~ bmi + age + children + f.sex + f.smoker + f.region  
##   
## Df Sum of Sq RSS AIC  
## <none> 239.52 -2034.08  
## - f.sex 1 1.72 241.24 -2027.17  
## - f.region 3 3.80 243.32 -2020.49  
## - bmi 1 6.46 245.98 -2002.96  
## - children 1 18.07 257.59 -1945.54  
## - age 1 293.68 533.20 -1039.74  
## - f.smoker 1 448.30 687.82 -722.73

summary(m5)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children + f.sex + f.smoker +   
## f.region, data = df[-influential\_obs, ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.05304 -0.19908 -0.05181 0.05959 2.16809   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.0366548 0.0752441 93.518 < 2e-16 \*\*\*  
## bmi 0.0124863 0.0021630 5.773 9.86e-09 \*\*\*  
## age 0.0349941 0.0008989 38.929 < 2e-16 \*\*\*  
## children 0.1023007 0.0105952 9.655 < 2e-16 \*\*\*  
## f.sexmale -0.0746327 0.0250596 -2.978 0.002956 \*\*   
## f.smokeryes 1.5244571 0.0316950 48.098 < 2e-16 \*\*\*  
## f.regionnorthwest -0.0613902 0.0357542 -1.717 0.086229 .   
## f.regionsoutheast -0.1431207 0.0359860 -3.977 7.38e-05 \*\*\*  
## f.regionsouthwest -0.1267455 0.0360563 -3.515 0.000455 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4402 on 1236 degrees of freedom  
## Multiple R-squared: 0.7662, Adjusted R-squared: 0.7647   
## F-statistic: 506.4 on 8 and 1236 DF, p-value: < 2.2e-16

par( mfrow = c(2,2))  
plot( m5, id.n=0 )



par( mfrow = c(1,1))

Le’s try to transform age into a factor

df$age\_range <- cut(df$age, breaks = quantile(df$age,probs = c(0,0.5,1)), include.lowest = T)  
summary(df)

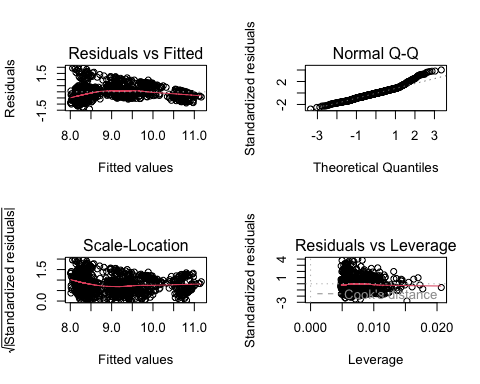
## age sex bmi children   
## Min. :18.00 Length:1323 Min. :15.96 Min. :0.00   
## 1st Qu.:27.00 Class :character 1st Qu.:26.22 1st Qu.:0.00   
## Median :39.00 Mode :character Median :30.30 Median :1.00   
## Mean :39.31 Mean :30.62 Mean :1.08   
## 3rd Qu.:51.00 3rd Qu.:34.60 3rd Qu.:2.00   
## Max. :64.00 Max. :53.13 Max. :5.00   
## smoker region charges f.sex f.smoker   
## Length:1323 Length:1323 Min. : 1122 female:654 no :1058   
## Class :character Class :character 1st Qu.: 4729 male :669 yes: 265   
## Mode :character Mode :character Median : 9305   
## Mean :13047   
## 3rd Qu.:16265   
## Max. :51195   
## f.region age\_range   
## northeast:320 [18,39]:663   
## northwest:321 (39,64]:660   
## southeast:360   
## southwest:322   
##   
##

We have created a new variable called age\_range where we divide the ages into 4 groups according to the 4 quantiles. From the summary (and the new column in the data set) we see 4 groups of ages and how many observations were fit into each age group. The results does not changed a lot with 4 quantiles and we tried with 2 groups and this got a more interesting result.

m6 <- lm(log(charges)~bmi+children+f.sex+f.smoker+f.region+age\_range, data=df[-influential\_obs,])  
summary(m6)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + children + f.sex + f.smoker +   
## f.region + age\_range, data = df[-influential\_obs, ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.40598 -0.32131 -0.02742 0.25982 2.02237   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.923742 0.080374 98.586 < 2e-16 \*\*\*  
## bmi 0.014752 0.002484 5.939 3.73e-09 \*\*\*  
## children 0.113214 0.012172 9.301 < 2e-16 \*\*\*  
## f.sexmale -0.083762 0.028814 -2.907 0.003715 \*\*   
## f.smokeryes 1.520719 0.036448 41.722 < 2e-16 \*\*\*  
## f.regionnorthwest -0.060609 0.041120 -1.474 0.140747   
## f.regionsoutheast -0.158606 0.041374 -3.833 0.000133 \*\*\*  
## f.regionsouthwest -0.125455 0.041467 -3.025 0.002534 \*\*   
## age\_range(39,64] 0.838466 0.028852 29.061 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5062 on 1236 degrees of freedom  
## Multiple R-squared: 0.6908, Adjusted R-squared: 0.6888   
## F-statistic: 345.2 on 8 and 1236 DF, p-value: < 2.2e-16

par( mfrow = c(2,2))  
plot( m6, id.n=0 )



par( mfrow = c(1,1))

AIC(m5,m6)

## df AIC  
## m5 10 1501.080  
## m6 10 1849.079

Removing age as a numerical exploratory variable and adding it as a factor does not improve things in general. However we can see that the normal Q-Q plot from the residuals is better since we reduced the impact of age variable. We will continue with model 5.

#### Adding interactions

With the model added factors will try to check adding double interactions between all numerical and factors.

m7 <- lm(log(charges)~bmi+age  
 +children \* (f.sex+f.smoker+f.region), data=df[-influential\_obs,])  
  
summary(m7)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children \* (f.sex + f.smoker +   
## f.region), data = df[-influential\_obs, ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.99576 -0.21193 -0.05748 0.06408 2.17375   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.0207859 0.0783886 89.564 < 2e-16 \*\*\*  
## bmi 0.0125351 0.0021465 5.840 6.68e-09 \*\*\*  
## age 0.0350238 0.0008924 39.246 < 2e-16 \*\*\*  
## children 0.1122317 0.0247948 4.526 6.58e-06 \*\*\*  
## f.sexmale -0.0968541 0.0336195 -2.881 0.004034 \*\*   
## f.smokeryes 1.6711212 0.0435590 38.365 < 2e-16 \*\*\*  
## f.regionnorthwest -0.0724783 0.0484149 -1.497 0.134643   
## f.regionsoutheast -0.1654342 0.0470327 -3.517 0.000452 \*\*\*  
## f.regionsouthwest -0.0973501 0.0479806 -2.029 0.042679 \*   
## children:f.sexmale 0.0250942 0.0210920 1.190 0.234374   
## children:f.smokeryes -0.1319506 0.0272271 -4.846 1.42e-06 \*\*\*  
## children:f.regionnorthwest 0.0101619 0.0304850 0.333 0.738934   
## children:f.regionsoutheast 0.0174022 0.0296315 0.587 0.557118   
## children:f.regionsouthwest -0.0230416 0.0297476 -0.775 0.438741   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4364 on 1231 degrees of freedom  
## Multiple R-squared: 0.7711, Adjusted R-squared: 0.7687   
## F-statistic: 319.1 on 13 and 1231 DF, p-value: < 2.2e-16

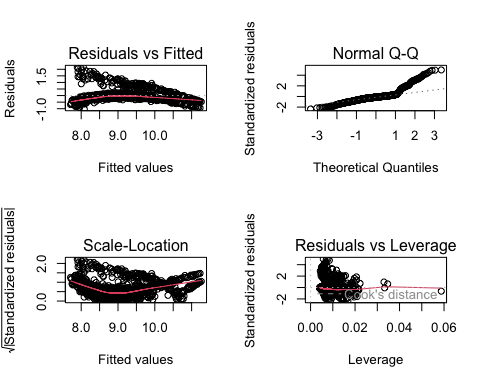
m8 <- step(m7) # see which is the best combination

## Start: AIC=-2050.6  
## log(charges) ~ bmi + age + children \* (f.sex + f.smoker + f.region)  
##   
## Df Sum of Sq RSS AIC  
## - children:f.region 3 0.418 234.89 -2054.4  
## - children:f.sex 1 0.270 234.74 -2051.2  
## <none> 234.47 -2050.6  
## - children:f.smoker 1 4.474 238.94 -2029.1  
## - bmi 1 6.496 240.97 -2018.6  
## - age 1 293.369 527.84 -1042.3  
##   
## Step: AIC=-2054.38  
## log(charges) ~ bmi + age + children + f.sex + f.smoker + f.region +   
## children:f.sex + children:f.smoker  
##   
## Df Sum of Sq RSS AIC  
## - children:f.sex 1 0.247 235.14 -2055.1  
## <none> 234.89 -2054.4  
## - f.region 3 3.864 238.75 -2040.1  
## - children:f.smoker 1 4.514 239.40 -2032.7  
## - bmi 1 6.407 241.30 -2022.9  
## - age 1 294.200 529.09 -1045.4  
##   
## Step: AIC=-2055.08  
## log(charges) ~ bmi + age + children + f.sex + f.smoker + f.region +   
## children:f.smoker  
##   
## Df Sum of Sq RSS AIC  
## <none> 235.14 -2055.1  
## - f.sex 1 1.503 236.64 -2049.1  
## - f.region 3 3.887 239.02 -2040.7  
## - children:f.smoker 1 4.384 239.52 -2034.1  
## - bmi 1 6.435 241.57 -2023.5  
## - age 1 294.017 529.15 -1047.2

summary(m8)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children + f.sex + f.smoker +   
## f.region + children:f.smoker, data = df[-influential\_obs,   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.02519 -0.20790 -0.05672 0.06422 2.17187   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.009093 0.074803 93.700 < 2e-16 \*\*\*  
## bmi 0.012465 0.002144 5.814 7.76e-09 \*\*\*  
## age 0.035014 0.000891 39.297 < 2e-16 \*\*\*  
## children 0.125946 0.011601 10.857 < 2e-16 \*\*\*  
## f.sexmale -0.069849 0.024859 -2.810 0.005036 \*\*   
## f.smokeryes 1.669037 0.043529 38.343 < 2e-16 \*\*\*  
## f.regionnorthwest -0.060998 0.035440 -1.721 0.085472 .   
## f.regionsoutheast -0.147420 0.035681 -4.132 3.84e-05 \*\*\*  
## f.regionsouthwest -0.123750 0.035745 -3.462 0.000554 \*\*\*  
## children:f.smokeryes -0.130210 0.027135 -4.799 1.79e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4363 on 1235 degrees of freedom  
## Multiple R-squared: 0.7705, Adjusted R-squared: 0.7688   
## F-statistic: 460.7 on 9 and 1235 DF, p-value: < 2.2e-16

par( mfrow = c(2,2))  
plot( m8, id.n=0 )



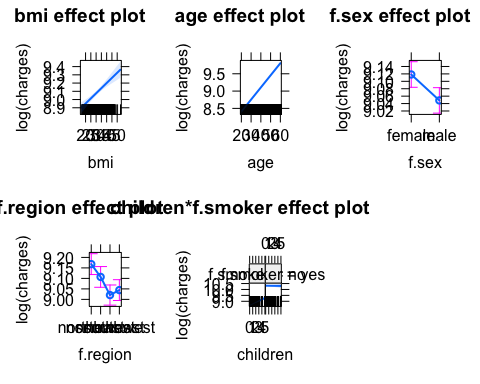
par( mfrow = c(1,1))

This will be our final model after several iterations.

library(effects)

## lattice theme set by effectsTheme()  
## See ?effectsTheme for details.

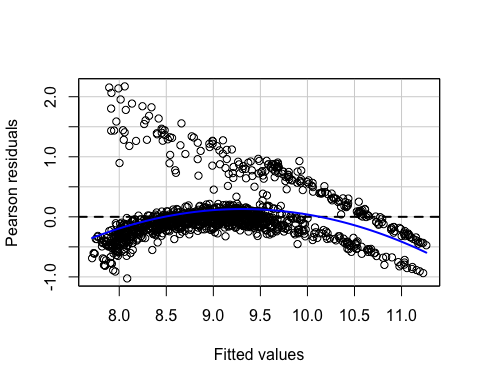
plot(allEffects(m8))



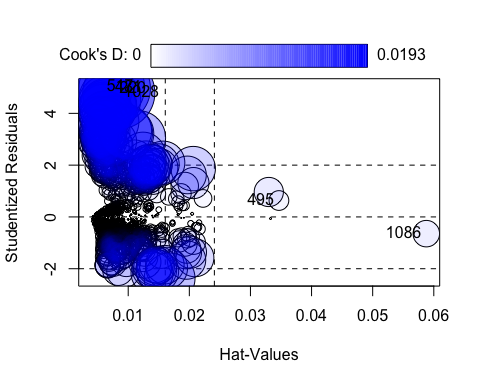
The allEffects plot shows that being a female have an effect of increasing the charges. In addition, we can see that having more children has an effect of increasing the charges on no smokers. On the other hand, smokers seem to have to pay much more regardless to the number of children .

#### Validation of the model

library(car)  
  
residualPlot(m8)



influencePlot(m8)



## StudRes Hat CookD  
## 220 4.9672541 0.007782677 0.018989292  
## 431 5.0007994 0.007074228 0.017477513  
## 495 0.6337794 0.034659237 0.001442863  
## 517 5.0408291 0.005715592 0.014323673  
## 1028 4.7909844 0.008465434 0.019254782  
## 1086 -0.6454875 0.058775305 0.002603050

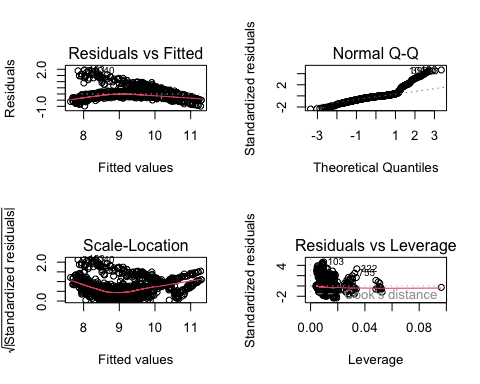
# there are a lot of influential data  
  
influential\_after\_iteractions <- which(rownames(df) %in% c("517","1028", "220", "431"))  
   
influential\_after\_iteractions

## [1] 218 428 514 1021

influential\_obs <- c(influential\_obs, influential\_after\_iteractions)  
  
m9 <- lm(log(charges)~bmi+age+children+children \* (f.sex+f.smoker+f.region), data=df[-influential\_obs,])  
summary(m9)

##   
## Call:  
## lm(formula = log(charges) ~ bmi + age + children + children \*   
## (f.sex + f.smoker + f.region), data = df[-influential\_obs,   
## ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.98467 -0.20570 -0.05223 0.06808 1.94903   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.9808247 0.0755475 92.403 < 2e-16 \*\*\*  
## bmi 0.0132065 0.0020684 6.385 2.43e-10 \*\*\*  
## age 0.0356108 0.0008598 41.416 < 2e-16 \*\*\*  
## children 0.1109512 0.0238384 4.654 3.60e-06 \*\*\*  
## f.sexmale -0.1081588 0.0324006 -3.338 0.000869 \*\*\*  
## f.smokeryes 1.6869411 0.0418988 40.262 < 2e-16 \*\*\*  
## f.regionnorthwest -0.0851715 0.0466150 -1.827 0.067924 .   
## f.regionsoutheast -0.1863705 0.0453029 -4.114 4.15e-05 \*\*\*  
## f.regionsouthwest -0.1109497 0.0461916 -2.402 0.016456 \*   
## children:f.sexmale 0.0287765 0.0202966 1.418 0.156503   
## children:f.smokeryes -0.1374206 0.0261779 -5.249 1.80e-07 \*\*\*  
## children:f.regionnorthwest 0.0159379 0.0293268 0.543 0.586913   
## children:f.regionsoutheast 0.0225097 0.0284977 0.790 0.429752   
## children:f.regionsouthwest -0.0177469 0.0286148 -0.620 0.535240   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4195 on 1227 degrees of freedom  
## Multiple R-squared: 0.7884, Adjusted R-squared: 0.7861   
## F-statistic: 351.6 on 13 and 1227 DF, p-value: < 2.2e-16

par(mfrow=c(2,2))  
plot(m9)



par(mfrow=c(1,1))

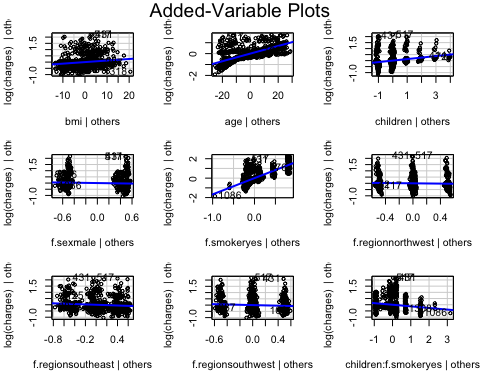
We addressed again influential data after adding interactions and we removed some observations.

The final model created have an adjusted R-squared score of 0.78 which is good. However, studying the residual plots there are patterns that are producing a deviation in the normal Q-Q.

This patters are mainly introduced by the **age** variable which we tried to reduce the impact transforming it into a factor with an age-range variable and removing it from the numerical explanatory variables. This transformation helped us to have a better normal Q-Q plot but reduced significantly the R-squared score.

Our decision is that we keep the age variable since we consider it an important variable and it has a positive impact in the r-squared score so the model will be better explained.

avPlots(m8)



We can see in the partial regression plots that people who are older are paying more charges and also people who smoke are significantly paying more. Also having more childrens and having a high bmi is affecting to paying more.

We managed to reach a good R-Squared which explains a lot of the variable charges and could help to make an prediction of what would a person be paying.