MANOJ KUMAR - 2048015

CAC-2

19/03/2021

suppressMessages(library(tidyverse)) # Data manipulation and plots  
suppressMessages(library(funModeling)) # Overview stats  
library(magrittr) # To use pipes

##   
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':  
##   
## set\_names

## The following object is masked from 'package:tidyr':  
##   
## extract

library(skimr) # To get a quick summary table  
library(caret) # To create the partition for training/test datasets

##   
## Attaching package: 'caret'

## The following object is masked from 'package:survival':  
##   
## cluster

## The following object is masked from 'package:purrr':  
##   
## lift

options(scipen = 999) # Turn off scientific notation for numbers  
options(repr.plot.width=12,   
 repr.plot.height=8) # Set universal plot size

# Reading the dataset  
df <- read.csv('insurance.csv')  
  
# Denote factor variables  
df$sex <- factor(df$sex)  
df$smoker <- factor(df$smoker)  
df$region <- factor(df$region)  
df$children <- factor(df$children)

# check for missing values  
df %>%  
 is.na() %>%  
 sum()

## [1] 0

# check data types  
df %>%  
 str()

## 'data.frame': 1338 obs. of 7 variables:  
## $ age : int 19 18 28 33 32 31 46 37 37 60 ...  
## $ sex : Factor w/ 2 levels "female","male": 1 2 2 2 2 1 1 1 2 1 ...  
## $ bmi : num 27.9 33.8 33 22.7 28.9 ...  
## $ children: Factor w/ 6 levels "0","1","2","3",..: 1 2 4 1 1 1 2 4 3 1 ...  
## $ smoker : Factor w/ 2 levels "no","yes": 2 1 1 1 1 1 1 1 1 1 ...  
## $ region : Factor w/ 4 levels "northeast","northwest",..: 4 3 3 2 2 3 3 2 1 2 ...  
## $ charges : num 16885 1726 4449 21984 3867 ...

skim(df)

Data summary

|  |  |
| --- | --- |
| Name | df |
| Number of rows | 1338 |
| Number of columns | 7 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Column type frequency: |  |
| factor | 4 |
| numeric | 3 |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |  |
| Group variables | None |

**Variable type: factor**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| skim\_variable | n\_missing | complete\_rate | ordered | n\_unique | top\_counts |
| sex | 0 | 1 | FALSE | 2 | mal: 676, fem: 662 |
| children | 0 | 1 | FALSE | 6 | 0: 574, 1: 324, 2: 240, 3: 157 |
| smoker | 0 | 1 | FALSE | 2 | no: 1064, yes: 274 |
| region | 0 | 1 | FALSE | 4 | sou: 364, nor: 325, sou: 325, nor: 324 |

**Variable type: numeric**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| skim\_variable | n\_missing | complete\_rate | mean | sd | p0 | p25 | p50 | p75 | p100 | hist |
| age | 0 | 1 | 39.21 | 14.05 | 18.00 | 27.00 | 39.00 | 51.00 | 64.00 | ▇▅▅▆▆ |
| bmi | 0 | 1 | 30.66 | 6.10 | 15.96 | 26.30 | 30.40 | 34.69 | 53.13 | ▂▇▇▂▁ |
| charges | 0 | 1 | 13270.42 | 12110.01 | 1121.87 | 4740.29 | 9382.03 | 16639.91 | 63770.43 | ▇▂▁▁▁ |

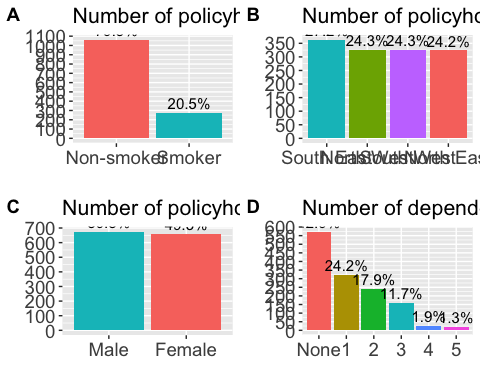
figsize <- options(repr.plot.width=12, repr.plot.height=12) # set plot size for this plot   
  
# Smoker count plot  
smoker <- df %>%  
 ggplot(aes(x=smoker, fill=smoker)) +  
 geom\_bar(show.legend = FALSE) +  
 geom\_text( stat='count',  
 aes(label=paste0(round(after\_stat(prop\*100), digits=1), "%"),group=1),  
 vjust=-0.4, size=4 ) +  
 labs( x = "", y = "", title = "Number of policyholders by smoking" ) +  
 scale\_x\_discrete( labels = c("no" = "Non-smoker", "yes" = "Smoker") ) +  
 scale\_y\_continuous( breaks=seq(0,2000,100) ) +  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text.x = element\_text(size=14),  
 axis.text.y = element\_text(size=14)  
 )

# Region count plot  
region <- df %>%  
 ggplot(aes(x=forcats::fct\_infreq(region), fill=region)) +  
 geom\_bar(show.legend = FALSE) +  
 geom\_text( stat='count',   
 aes(label = paste0(round(after\_stat(prop\*100), digits=1), "%"), group=1),   
 vjust=-0.4, size=4 ) +  
 labs( x = "", y = "", title = "Number of policyholders by region" ) +  
 scale\_x\_discrete( labels = c("northeast" = "North East", "northwest" = "North West",  
 "southeast" = "South East", "southwest" = "South West") ) +  
 scale\_y\_continuous( breaks=seq(0,350,50) ) +  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text.x = element\_text(size=14),  
 axis.text.y = element\_text(size=14)  
 )

# Sex count plot  
sex <- df %>%  
 ggplot(aes(x=forcats::fct\_infreq(sex), fill=sex)) +  
 geom\_bar(show.legend = FALSE) +  
 geom\_text(stat='count',  
 aes( label=paste0(round(after\_stat(prop\*100), digits=1), "%"), group=1),  
 vjust=-0.4, size=4 ) +  
 labs( x = "", y = "", title = "Number of policyholders by sex", fill = "Sex" ) +  
 scale\_x\_discrete( labels = c("male" = "Male", "female" = "Female") ) +  
 scale\_y\_continuous( breaks=seq(0,700,100) ) +  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text.x = element\_text(size=14),  
 axis.text.y = element\_text(size=14)  
 )

# Children count plot  
children <- df %>%  
 ggplot(aes(x=forcats::fct\_infreq(children), fill=children)) +  
 geom\_bar(show.legend = FALSE) +  
 geom\_text(stat='count',  
 aes(label=paste0(round(after\_stat(prop\*100), digits=1), "%"), group=1),  
 vjust=-0.4, size=4 ) +  
 labs( x = "", y = "", title = "Number of dependents per policy" ) +  
 scale\_x\_discrete( labels = c("0" = "None") ) +  
 scale\_y\_continuous( breaks=seq(0,600,50) ) +  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text.x = element\_text(size=14),  
 axis.text.y = element\_text(size=14)  
 )

# Plot grid  
cowplot::plot\_grid( smoker, region, sex, children, labels="AUTO", ncol = 2, nrow = 2 )



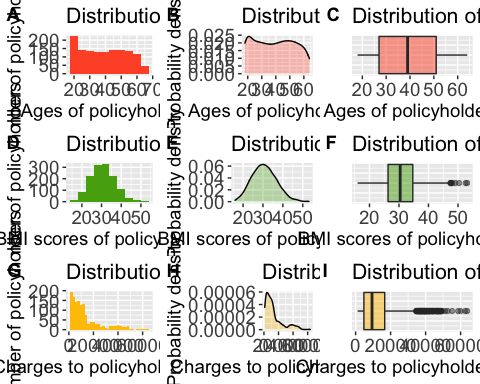
options(figsize)

figsize <- options(repr.plot.width=20, repr.plot.height=16)  
  
# Age distribution  
age\_hist <- df %>%  
 ggplot(aes(x=age))+  
 geom\_histogram( binwidth = 5, show.legend = FALSE, fill="#ff5733" )+  
 labs( x = "Ages of policyholders", y = "Number of policyholders",  
 title = "Distribution of ages" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
age\_dens <- df %>%  
 ggplot(aes(x=age)) + geom\_density( alpha=.3, fill="#ff5733" )+  
 labs( x = "Ages of policyholders", y = "Probability density",  
 title = "Distribution of ages" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
age\_box <- df %>%  
 ggplot(aes(y=age)) +  
 geom\_boxplot( alpha=.5, fill="#ff5733" )+  
 coord\_flip() +  
 theme(  
 axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank()  
 )+  
 labs( y = "Ages of policyholders", x = "", title = "Distribution of ages" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

# BMI distribution  
bmi\_hist <- df %>%  
 ggplot(aes(x=bmi))+  
 geom\_histogram(binwidth = 4, show.legend = FALSE, fill = "#55ab11" )+  
 labs( x = "BMI scores of policyholders", y = "Number of policyholders",  
 title = "Distribution of BMI scores" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_dens <- df %>%  
 ggplot(aes(x=bmi)) + geom\_density( alpha=.3, fill="#55ab11" )+  
 labs(  
 x = "BMI scores of policyholders",  
 y = "Probability density",  
 title = "Distribution of BMI scores" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_box <- df %>%  
 ggplot(aes(y=bmi)) +  
 geom\_boxplot( alpha=.5, fill="#55ab11" )+  
 coord\_flip() +  
 theme( axis.text.y = element\_blank(),  
 axis.ticks.y = element\_blank() )+  
 labs( y = "BMI scores of policyholders", x = "",  
 title = "Distribution of BMI scores" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

# Charges distribution  
charges\_hist <- df %>%  
 ggplot(aes(x=charges)) +  
 geom\_histogram( binwidth = 2000, show.legend = FALSE, fill = "#FFC300" )+  
 labs( x = "Charges to policyholders ($)", y = "Number of policyholders",  
 title = "Distribution of medical charges" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
charges\_dens <- df %>%  
 ggplot( aes(x=charges) ) +  
 geom\_density( alpha=.3, fill="#FFC300" ) +  
 labs( x = "Charges to policyholders ($)", y = "Probability density",  
 title = "Distribution of medical charges" ) +  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
charges\_box <- df %>%  
 ggplot(aes(y=charges))+  
 geom\_boxplot( alpha=.5, fill="#FFC300" )+ coord\_flip()+  
 theme( axis.text.y = element\_blank(), axis.ticks.y = element\_blank() )+  
 labs( y = "Charges to policyholders ($)", x = "",  
 title = "Distribution of medical charges" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

cowplot::plot\_grid( age\_hist, age\_dens, age\_box, bmi\_hist, bmi\_dens, bmi\_box,  
 charges\_hist, charges\_dens, charges\_box, labels="AUTO", ncol = 3, nrow = 3 )

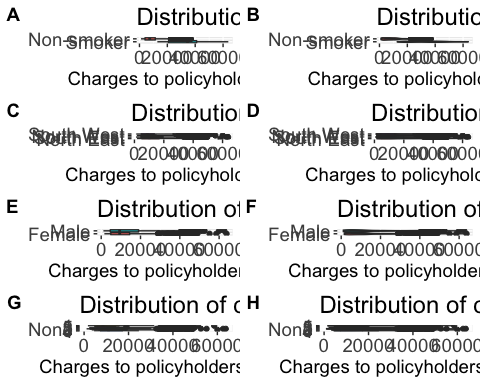


options(figsize)

figsize <- options(repr.plot.width=20, repr.plot.height=26)  
  
# Boxplots  
chargesBysmoker <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(smoker, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=smoker ) ) +  
 geom\_boxplot(show.legend = FALSE) +  
 coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges by smoking" )+  
 scale\_x\_discrete(  
 labels = c("no" = "Non-smoker", "yes" = "Smoker") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
chargesByregion <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(region, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=region ) ) +  
 geom\_boxplot(show.legend = FALSE) +  
 coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges by region" )+  
 scale\_x\_discrete( labels = c("northeast" = "North East", "northwest" = "North West",  
 "southeast" = "South East", "southwest" = "South West") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
chargesBychildren <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(children, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=children ) ) +  
 geom\_boxplot(show.legend = FALSE) + coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges by dependents" )+  
 scale\_x\_discrete( labels = c("0" = "None") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
chargesBysex <- df %>%  
 ggplot(  
 aes( x=forcats::fct\_reorder(sex, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=sex ) ) +  
 geom\_boxplot(show.legend = FALSE) +  
 coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",   
 title = "Distribution of charges by sex" )+  
 scale\_x\_discrete( labels = c("male" = "Male", "female" = "Female") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

# Density plots with medians  
densityBySmoker <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(smoker, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=smoker ) ) +  
 geom\_violin(show.legend = FALSE) +  
 geom\_boxplot( width=0.1, show.legend = FALSE )+ coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges with density by smoking"  
 )+ scale\_x\_discrete( labels = c("no" = "Non-smoker", "yes" = "Smoker") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
densityByRegion <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(region, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=region ) ) +  
 geom\_violin(show.legend = FALSE) +  
 geom\_boxplot(  
 width=0.1,  
 show.legend = FALSE  
 )+ coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges with density by region" )+  
 scale\_x\_discrete(labels = c("northeast" = "North East", "northwest" = "North West",  
 "southeast" = "South East", "southwest" = "South West") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
densityBySex <- df %>%  
 ggplot(  
 aes( x=forcats::fct\_reorder(sex, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=sex ) ) +  
 geom\_violin(show.legend = FALSE) +  
 geom\_boxplot( width=0.1, show.legend = FALSE )+ coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges with density by sex" )+  
 scale\_x\_discrete( labels = c("male" = "Male", "female" = "Female") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
densityByChildren <- df %>%  
 ggplot( aes( x=forcats::fct\_reorder(children, charges, .fun=median, .desc=TRUE),  
 y=charges, fill=children ) ) +  
 geom\_violin(show.legend = FALSE) +  
 geom\_boxplot( width=0.1, show.legend = FALSE )+ coord\_flip() +  
 labs( x = "", y = "Charges to policyholders ($)",  
 title = "Distribution of charges with density by dependents")+  
 scale\_x\_discrete( labels = c("0" = "None") )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

# Plot grid of all plots  
cowplot::plot\_grid(  
 chargesBysmoker, densityBySmoker,  
 chargesByregion, densityByRegion,  
 chargesBysex, densityBySex,  
 chargesBychildren, densityByChildren,  
 labels="AUTO", ncol = 2, nrow = 4  
)

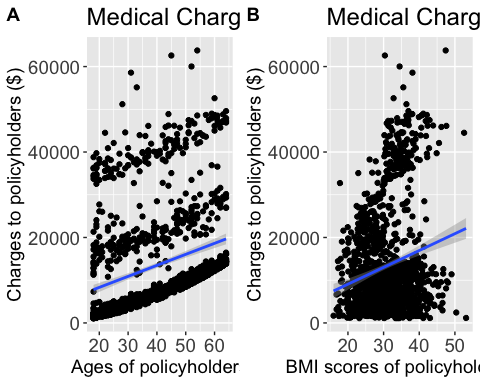


options(figsize)

figsize <- options(repr.plot.width=12, repr.plot.height=8)  
  
age\_scatter <- df %>%  
 ggplot(aes(x=age, y=charges)) + geom\_point()+ geom\_smooth(method='lm')+  
 labs( x = "Ages of policyholders",y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder Age" )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_scatter <- df %>%  
 ggplot(aes(x=bmi, y=charges)) +  
 geom\_point()+  
 geom\_smooth(method='lm')+  
 labs( x = "BMI scores of policyholders", y = "Charges to policyholders ($)",   
 title = "Medical Charges vs Policyholder BMI" )+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

cowplot::plot\_grid( age\_scatter, bmi\_scatter, labels="AUTO", ncol = 2, nrow = 1 )

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



options(figsize)

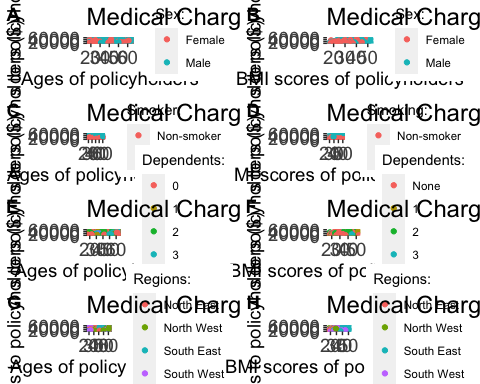
figsize <- options(repr.plot.width=20, repr.plot.height=22)  
  
# by sex  
age\_scatter\_sex <- df %>%  
 ggplot(aes(x=age, y=charges, color=sex)) +  
 geom\_point()+  
 labs( x = "Ages of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder Age by Sex", color = "Sex:" )+  
 scale\_color\_hue(labels = c("male" = "Male", "female" = "Female"))+ guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_scatter\_sex <- df %>%  
 ggplot(aes(x=bmi, y=charges, color=sex)) + geom\_point()+  
 labs( x = "BMI scores of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder BMI by Sex", color = "Sex:" )+  
 scale\_color\_hue(labels = c("male" = "Male", "female" = "Female"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

#by smoker  
age\_scatter\_smoker <- df %>%  
 ggplot(aes(x=age, y=charges, color=smoker)) +  
 geom\_point()+  
 labs( x = "Ages of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder Age by Smoking", color = "Smoker:" )+  
 scale\_color\_hue(labels = c("no" = "Non-smoker", "yes" = "Smoker"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_scatter\_smoker <- df %>%  
 ggplot(aes(x=bmi, y=charges, color=smoker)) +  
 geom\_point()+  
 labs( x = "BMI scores of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder BMI by Smoking",  
 color = "Smoking:" )+  
 scale\_color\_hue(labels = c("no" = "Non-smoker", "yes" = "Smoker"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

#by children  
age\_scatter\_kids <- df %>%  
 ggplot(aes(x=age, y=charges, color=children)) +  
 geom\_point()+  
 labs( x = "Ages of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder Age by Dependents",  
 color = "Dependents:" )+  
 scale\_color\_hue(labels = c("no" = "Non-smoker", "yes" = "Smoker"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
bmi\_scatter\_kids <- df %>%  
 ggplot(aes(x=bmi, y=charges, color=children)) +  
 geom\_point()+  
 labs( x = "BMI scores of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder BMI by Dependents",  
 color = "Dependents:" )+  
 scale\_color\_hue(labels = c("0" = "None"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

#by region  
age\_scatter\_region <- df %>%  
 ggplot(aes(x=age, y=charges, color=region)) +  
 geom\_point()+  
 labs( x = "Ages of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder Age by Region", color = "Regions:" )+  
 scale\_color\_hue(labels = c("northeast" = "North East", "northwest" = "North West",  
 "southeast" = "South East", "southwest" = "South West"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
bmi\_scatter\_region <- df %>%  
 ggplot(aes(x=bmi, y=charges, color=region)) +  
 geom\_point()+  
 labs( x = "BMI scores of policyholders", y = "Charges to policyholders ($)",  
 title = "Medical Charges vs Policyholder BMI by Regions", color = "Regions:" )+  
 scale\_color\_hue(labels = c("northeast" = "North East", "northwest" = "North West",  
 "southeast" = "South East", "southwest" = "South West"))+  
 guides(fill=FALSE)+  
 theme(  
 plot.title = element\_text(size=18),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

# make a grid  
cowplot::plot\_grid(  
 age\_scatter\_sex, bmi\_scatter\_sex,  
 age\_scatter\_smoker, bmi\_scatter\_smoker,  
 age\_scatter\_kids, bmi\_scatter\_kids,  
 age\_scatter\_region, bmi\_scatter\_region,  
 labels="AUTO", ncol = 2, nrow = 4 )



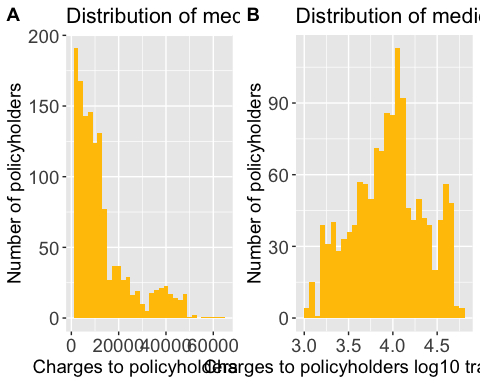
options(figsize)

## Multiple linear regression

charges\_hist <- df %>%  
 ggplot( aes(x=charges) ) +  
 geom\_histogram( binwidth = 2000, show.legend = FALSE, fill = "#FFC300" )+  
 labs( x = "Charges to policyholders ($)", y = "Number of policyholders",  
 title = "Distribution of medical charges" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )  
  
charges\_hist\_log10 <- df %>%  
 ggplot( aes(x=log10(charges)) ) +  
 geom\_histogram( show.legend = FALSE, fill = "#FFC300" )+  
 labs( x = "Charges to policyholders log10 transformed",  
 y = "Number of policyholders", title = "Distribution of medical charges after log10 transform" )+  
 theme(  
 plot.title = element\_text(size=16),  
 axis.text = element\_text(size=14),  
 axis.title = element\_text(size=14)  
 )

cowplot::plot\_grid( charges\_hist, charges\_hist\_log10, labels="AUTO", ncol = 2, nrow = 1 )

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



## Split the dataset and train the model

# log10 transform of response variable   
df$logCharges<- log10(df$charges)  
  
drops <- c("charges")  
df[ , !(names(df) %in% drops)]

## age sex bmi children smoker region logCharges  
## 1 19 female 27.900 0 yes southwest 4.227499  
## 2 18 male 33.770 1 no southeast 3.236928  
## 3 28 male 33.000 3 no southeast 3.648308  
## 4 33 male 22.705 0 no northwest 4.342116  
## 5 32 male 28.880 0 no northwest 3.587358  
## 6 31 female 25.740 0 no southeast 3.574797  
## 7 46 female 33.440 1 no southeast 3.915958  
## 8 37 female 27.740 3 no northwest 3.862221  
## 9 37 male 29.830 2 no northeast 3.806615  
## 10 60 female 25.840 0 no northwest 4.461245  
## 11 25 male 26.220 0 no northeast 3.434780  
## 12 62 female 26.290 0 yes southeast 4.444181  
## 13 23 male 34.400 0 no southwest 3.261701  
## 14 56 female 39.820 0 no southeast 4.044960  
## 15 27 male 42.130 0 yes southeast 4.597824  
## 16 19 male 24.600 1 no southwest 3.264165  
## 17 52 female 30.780 1 no northeast 4.033317  
## 18 23 male 23.845 0 no northeast 3.379337  
## 19 56 male 40.300 0 no southwest 4.025404  
## 20 30 male 35.300 0 yes southwest 4.566290  
## 21 60 female 36.005 0 no northeast 4.121522  
## 22 30 female 32.400 1 no southwest 3.618020  
## 23 18 male 34.100 0 no southeast 3.055765  
## 24 34 female 31.920 1 yes northeast 4.576363  
## 25 37 male 28.025 2 no northwest 3.792665  
## 26 59 female 27.720 3 no southeast 4.146163  
## 27 63 female 23.085 0 no northeast 4.159923  
## 28 55 female 32.775 2 no northwest 4.088796  
## 29 23 male 17.385 1 no northwest 3.443293  
## 30 31 male 36.300 2 yes southwest 4.587834  
## 31 22 male 35.600 0 yes southwest 4.551274  
## 32 18 female 26.315 0 no northeast 3.342065  
## 33 19 female 28.600 5 no southwest 3.670969  
## 34 63 male 28.310 0 no northwest 4.138937  
## 35 28 male 36.400 1 yes southwest 4.709224  
## 36 19 male 20.425 0 no northwest 3.210969  
## 37 62 female 32.965 3 no northwest 4.193464  
## 38 26 male 20.800 0 no southwest 3.362162  
## 39 35 male 36.670 1 yes northeast 4.599602  
## 40 60 male 39.900 0 yes southwest 4.682807  
## 41 24 female 26.600 0 no northeast 3.483739  
## 42 31 female 36.630 2 no southeast 3.694584  
## 43 41 male 21.780 1 no southeast 3.797439  
## 44 37 female 30.800 2 no southeast 3.800288  
## 45 38 male 37.050 1 no northeast 3.783880  
## 46 55 male 37.300 0 no southwest 4.314505  
## 47 18 female 38.665 2 no northeast 3.530629  
## 48 28 female 34.770 0 no northwest 3.551074  
## 49 60 female 24.530 0 no southeast 4.101400  
## 50 36 male 35.200 1 yes southeast 4.587814  
## 51 18 female 35.625 0 no northeast 3.344614  
## 52 21 female 33.630 2 no northwest 3.553862  
## 53 48 male 28.000 1 yes southwest 4.372328  
## 54 36 male 34.430 0 yes southeast 4.576832  
## 55 40 female 28.690 3 no northwest 3.906318  
## 56 58 male 36.955 2 yes northwest 4.676662  
## 57 58 female 31.825 2 no northeast 4.133774  
## 58 18 male 31.680 2 yes southeast 4.535334  
## 59 53 female 22.880 1 yes southeast 4.366326  
## 60 34 female 37.335 2 no northwest 3.777392  
## 61 43 male 27.360 3 no northeast 3.934812  
## 62 25 male 33.660 4 no southeast 3.653662  
## 63 64 male 24.700 1 no northwest 4.479527  
## 64 28 female 25.935 1 no northwest 3.616333  
## 65 20 female 22.420 0 yes northwest 4.167664  
## 66 19 female 28.900 0 no southwest 3.241351  
## 67 61 female 39.100 2 no southwest 4.153360  
## 68 40 male 26.315 1 no northwest 3.805459  
## 69 40 female 36.190 0 no southeast 3.772329  
## 70 28 male 23.980 3 yes southeast 4.247068  
## 71 27 female 24.750 0 yes southeast 4.219526  
## 72 31 male 28.500 5 no northeast 3.832474  
## 73 53 female 28.100 3 no southwest 4.069732  
## 74 58 male 32.010 1 no southeast 4.077245  
## 75 44 male 27.400 2 no southwest 3.888003  
## 76 57 male 34.010 0 no northwest 4.055251  
## 77 29 female 29.590 1 no southeast 3.596313  
## 78 21 male 35.530 0 no southeast 3.185392  
## 79 22 female 39.805 0 no northeast 3.440125  
## 80 41 female 32.965 0 no northwest 3.817633  
## 81 31 male 26.885 1 no northeast 3.647502  
## 82 45 female 38.285 0 no northeast 3.899563  
## 83 22 male 37.620 1 yes southeast 4.570136  
## 84 48 female 41.230 4 no northwest 4.042720  
## 85 37 female 34.800 2 yes southwest 4.600281  
## 86 45 male 22.895 2 yes northwest 4.324253  
## 87 57 female 31.160 0 yes northwest 4.639277  
## 88 56 female 27.200 0 no southwest 4.044272  
## 89 46 female 27.740 0 no northwest 3.904535  
## 90 55 female 26.980 0 no northwest 4.044641  
## 91 21 female 39.490 0 no southeast 3.306848  
## 92 53 female 24.795 1 no northwest 4.039102  
## 93 59 male 29.830 3 yes northeast 4.479790  
## 94 35 male 34.770 2 no northwest 3.758079  
## 95 64 female 31.300 2 yes southwest 4.674779  
## 96 28 female 37.620 1 no southeast 3.575982  
## 97 54 female 30.800 3 no southwest 4.082976  
## 98 55 male 38.280 0 no southeast 4.009718  
## 99 56 male 19.950 0 yes northeast 4.350493  
## 100 38 male 19.300 0 yes southwest 4.199226  
## 101 41 female 31.600 0 no southwest 3.791419  
## 102 30 male 25.460 0 no northeast 3.561708  
## 103 18 female 30.115 0 no northeast 4.329293  
## 104 61 female 29.920 3 yes southeast 4.490551  
## 105 34 female 27.500 1 no southwest 3.699305  
## 106 20 male 28.025 1 yes northwest 4.244534  
## 107 19 female 28.400 1 no southwest 3.367639  
## 108 26 male 30.875 2 no northwest 3.588530  
## 109 29 male 27.940 0 no southeast 3.457446  
## 110 63 male 35.090 0 yes southeast 4.672611  
## 111 54 male 33.630 1 no northwest 4.034438  
## 112 55 female 29.700 2 no southwest 4.074866  
## 113 37 male 30.800 0 no southwest 3.667150  
## 114 21 female 35.720 0 no northwest 3.381067  
## 115 52 male 32.205 3 no northeast 4.060256  
## 116 60 male 28.595 0 no northeast 4.480869  
## 117 58 male 49.060 0 no southeast 4.056193  
## 118 29 female 27.940 1 yes southeast 4.281210  
## 119 49 female 27.170 0 no southeast 3.934566  
## 120 37 female 23.370 2 no northwest 3.825194  
## 121 44 male 37.100 2 no southwest 3.888760  
## 122 18 male 23.750 0 no northeast 3.231883  
## 123 20 female 28.975 0 no northwest 3.353623  
## 124 44 male 31.350 1 yes northeast 4.597218  
## 125 47 female 33.915 3 no northwest 4.004966  
## 126 26 female 28.785 0 no northeast 3.529610  
## 127 19 female 28.300 0 yes southwest 4.232515  
## 128 52 female 37.400 0 no southwest 3.983831  
## 129 32 female 17.765 2 yes northwest 4.515002  
## 130 38 male 34.700 2 no southwest 3.784075  
## 131 59 female 26.505 0 no northeast 4.107734  
## 132 61 female 22.040 0 no northeast 4.134061  
## 133 53 female 35.900 2 no southwest 4.047803  
## 134 19 male 25.555 0 no northwest 3.212870  
## 135 20 female 28.785 0 no northeast 3.390442  
## 136 22 female 28.050 0 no southeast 3.333585  
## 137 19 male 34.100 0 no southwest 3.100867  
## 138 22 male 25.175 0 no northwest 3.310839  
## 139 54 female 31.900 3 no southeast 4.436524  
## 140 22 female 36.000 0 no southwest 3.335805  
## 141 34 male 22.420 2 no northeast 4.437368  
## 142 26 male 32.490 1 no northeast 3.542894  
## 143 34 male 25.300 2 yes southeast 4.278124  
## 144 29 male 29.735 2 no northwest 4.259065  
## 145 30 male 28.690 3 yes northwest 4.316934  
## 146 29 female 38.830 3 no southeast 3.710816  
## 147 46 male 30.495 3 yes northwest 4.609814  
## 148 51 female 37.730 1 no southeast 3.994652  
## 149 53 female 37.430 1 no northwest 4.039798  
## 150 19 male 28.400 1 no southwest 3.265412  
## 151 35 male 24.130 1 no northwest 3.709712  
## 152 48 male 29.700 0 no southeast 3.891517  
## 153 32 female 37.145 3 no northeast 3.801702  
## 154 42 female 23.370 0 yes northeast 4.300264  
## 155 40 female 25.460 1 no northeast 3.849861  
## 156 44 male 39.520 0 no northwest 3.841904  
## 157 48 male 24.420 0 yes southeast 4.326821  
## 158 18 male 25.175 0 yes northeast 4.190841  
## 159 30 male 35.530 0 yes southeast 4.567617  
## 160 50 female 27.830 3 no southeast 4.295554  
## 161 42 female 26.600 0 yes northwest 4.329372  
## 162 18 female 36.850 0 yes southeast 4.558102  
## 163 54 male 39.600 1 no southwest 4.019139  
## 164 32 female 29.800 2 no southwest 3.711987  
## 165 37 male 29.640 0 no northwest 3.701408  
## 166 47 male 28.215 4 no northeast 4.017329  
## 167 20 female 37.000 5 no southwest 3.684004  
## 168 32 female 33.155 3 no northwest 3.787375  
## 169 19 female 31.825 1 no northwest 3.434454  
## 170 27 male 18.905 3 no northeast 3.683759  
## 171 63 male 41.470 0 no southeast 4.127279  
## 172 49 male 30.300 0 no southwest 3.909378  
## 173 18 male 15.960 0 no northeast 3.229118  
## 174 35 female 34.800 1 no southwest 3.719832  
## 175 24 female 33.345 0 no northwest 3.455673  
## 176 63 female 37.700 0 yes southwest 4.688637  
## 177 38 male 27.835 2 no northwest 3.809954  
## 178 54 male 29.200 1 no southwest 4.018538  
## 179 46 female 28.900 2 no southwest 3.945630  
## 180 41 female 33.155 3 no northeast 3.931371  
## 181 58 male 28.595 0 no northwest 4.069516  
## 182 18 female 38.280 0 no southeast 3.212673  
## 183 22 male 19.950 3 no northeast 3.602648  
## 184 44 female 26.410 0 no northwest 3.870373  
## 185 44 male 30.690 2 no southeast 3.888260  
## 186 36 male 41.895 3 yes northeast 4.641011  
## 187 26 female 29.920 2 no southeast 3.600099  
## 188 30 female 30.900 3 no southwest 3.726373  
## 189 41 female 32.200 1 no southwest 3.830971  
## 190 29 female 32.110 2 no northwest 3.692222  
## 191 61 male 31.570 0 no southeast 4.098907  
## 192 36 female 26.200 0 no southwest 3.688764  
## 193 25 male 25.740 0 no southeast 3.329937  
## 194 56 female 26.600 1 no northwest 4.080783  
## 195 18 male 34.430 0 no southeast 3.055940  
## 196 19 male 30.590 0 no northwest 3.214728  
## 197 39 female 32.800 0 no southwest 3.752027  
## 198 45 female 28.600 2 no southeast 3.930278  
## 199 51 female 18.050 0 no northwest 3.984269  
## 200 64 female 39.330 0 no northeast 4.173230  
## 201 19 female 32.110 0 no northwest 3.328517  
## 202 48 female 32.230 1 no southeast 3.947980  
## 203 60 female 24.035 0 no northwest 4.114351  
## 204 27 female 36.080 0 yes southeast 4.569771  
## 205 46 male 22.300 0 no southwest 3.854130  
## 206 28 female 28.880 1 no northeast 3.637263  
## 207 59 male 26.400 0 no southeast 4.069790  
## 208 35 male 27.740 2 yes northeast 4.321890  
## 209 63 female 31.800 0 no southwest 4.142419  
## 210 40 male 41.230 1 no northeast 3.820209  
## 211 20 male 33.000 1 no southwest 3.296681  
## 212 40 male 30.875 4 no northwest 3.911835  
## 213 24 male 28.500 2 no northwest 3.548721  
## 214 34 female 26.730 1 no southeast 3.699212  
## 215 45 female 30.900 2 no southwest 3.930441  
## 216 41 female 37.100 2 no southwest 3.867572  
## 217 53 female 26.600 0 no northwest 4.015177  
## 218 27 male 23.100 0 no southeast 3.395105  
## 219 26 female 29.920 1 no southeast 3.530581  
## 220 24 female 23.210 0 no southeast 4.399358  
## 221 34 female 33.700 1 no southwest 3.700052  
## 222 53 female 33.250 0 no northeast 4.023865  
## 223 32 male 30.800 3 no southwest 3.720451  
## 224 19 male 34.800 0 yes southwest 4.541325  
## 225 42 male 24.640 0 yes southeast 4.290381  
## 226 55 male 33.880 3 no southeast 4.078717  
## 227 28 male 38.060 0 no southeast 3.429671  
## 228 58 female 41.910 0 no southeast 4.384306  
## 229 41 female 31.635 1 no northeast 3.866770  
## 230 47 male 25.460 2 no northeast 3.964978  
## 231 42 female 36.195 1 no northwest 3.871786  
## 232 59 female 27.830 3 no southeast 4.146168  
## 233 19 female 17.800 0 no southwest 3.237490  
## 234 59 male 27.500 1 no southwest 4.091098  
## 235 39 male 24.510 2 no northwest 3.826735  
## 236 40 female 22.220 2 yes southeast 4.288792  
## 237 18 female 26.730 0 no southeast 3.208379  
## 238 31 male 38.390 2 no southeast 3.649647  
## 239 19 male 29.070 0 yes northwest 4.239367  
## 240 44 male 38.060 1 no southeast 3.854468  
## 241 23 female 36.670 2 yes northeast 4.585592  
## 242 33 female 22.135 1 no northeast 3.728684  
## 243 55 female 26.800 1 no southwest 4.546051  
## 244 40 male 35.300 3 no southwest 3.857143  
## 245 63 female 27.740 0 yes northeast 4.470163  
## 246 54 male 30.020 0 no northwest 4.388749  
## 247 60 female 38.060 0 no southeast 4.102046  
## 248 24 male 35.860 0 no southeast 3.298183  
## 249 19 male 20.900 1 no southwest 3.262948  
## 250 29 male 28.975 1 no northeast 3.606441  
## 251 18 male 17.290 2 yes northeast 4.108208  
## 252 63 female 32.200 2 yes southwest 4.674910  
## 253 54 male 34.210 2 yes southeast 4.646019  
## 254 27 male 30.300 3 no southwest 3.629485  
## 255 50 male 31.825 0 yes northeast 4.613812  
## 256 55 female 25.365 3 no northeast 4.115522  
## 257 56 male 33.630 0 yes northwest 4.642674  
## 258 38 female 40.150 0 no southeast 3.732473  
## 259 51 male 24.415 4 no northwest 4.061456  
## 260 19 male 31.920 0 yes northwest 4.528278  
## 261 58 female 25.200 0 no southwest 4.073248  
## 262 20 female 26.840 1 yes southeast 4.232622  
## 263 52 male 24.320 3 yes northeast 4.395673  
## 264 19 male 36.955 0 yes northwest 4.558941  
## 265 53 female 38.060 3 no southeast 4.310969  
## 266 46 male 42.350 3 yes southeast 4.664182  
## 267 40 male 19.800 1 yes southeast 4.235011  
## 268 59 female 32.395 3 no northeast 4.164074  
## 269 45 male 30.200 1 no southwest 3.871634  
## 270 49 male 25.840 1 no northeast 3.967664  
## 271 18 male 29.370 1 no southeast 3.235386  
## 272 50 male 34.200 2 yes southwest 4.632020  
## 273 41 male 37.050 2 no northwest 3.861278  
## 274 50 male 27.455 1 no northeast 3.983070  
## 275 25 male 27.550 0 no northwest 3.401946  
## 276 47 female 26.600 2 no northeast 3.987480  
## 277 19 male 20.615 2 no northwest 3.447731  
## 278 22 female 24.300 0 no southwest 3.332533  
## 279 59 male 31.790 2 no southeast 4.111558  
## 280 51 female 21.560 1 no southeast 3.993662  
## 281 40 female 28.120 1 yes northeast 4.348919  
## 282 54 male 40.565 3 yes northeast 4.686182  
## 283 30 male 27.645 1 no northeast 3.627071  
## 284 55 female 32.395 1 no northeast 4.074784  
## 285 52 female 31.200 0 no southwest 3.983442  
## 286 46 male 26.620 1 no southeast 3.888859  
## 287 46 female 48.070 2 no northeast 3.974646  
## 288 63 female 26.220 0 no northwest 4.154004  
## 289 59 female 36.765 1 yes northeast 4.680306  
## 290 52 male 26.400 3 no southeast 4.414853  
## 291 28 female 33.400 0 no southwest 3.501336  
## 292 29 male 29.640 1 no northeast 4.307021  
## 293 25 male 45.540 2 yes southeast 4.624408  
## 294 22 female 28.820 0 no southeast 3.333800  
## 295 25 male 26.800 3 no southwest 3.591746  
## 296 18 male 22.990 0 no northeast 3.231614  
## 297 19 male 27.700 0 yes southwest 4.212130  
## 298 47 male 25.410 1 yes southeast 4.342002  
## 299 31 male 34.390 3 yes northwest 4.588231  
## 300 48 female 28.880 1 no northwest 3.966118  
## 301 36 male 27.550 3 no northeast 3.829094  
## 302 53 female 22.610 3 yes northeast 4.395735  
## 303 56 female 37.510 2 no southeast 4.088686  
## 304 28 female 33.000 2 no southeast 3.638436  
## 305 57 female 38.000 2 no southwest 4.101960  
## 306 29 male 33.345 2 no northwest 4.288749  
## 307 28 female 27.500 2 no southwest 4.304871  
## 308 30 female 33.330 1 no southeast 3.618156  
## 309 58 male 34.865 0 no northeast 4.077171  
## 310 41 female 33.060 2 no northwest 3.889254  
## 311 50 male 26.600 0 no southwest 3.926573  
## 312 19 female 24.700 0 no southwest 3.239894  
## 313 43 male 35.970 3 yes southeast 4.624535  
## 314 49 male 35.860 0 no southeast 3.909792  
## 315 27 female 31.400 0 yes southwest 4.542064  
## 316 52 male 33.250 0 no northeast 3.987790  
## 317 50 male 32.205 0 no northwest 3.946220  
## 318 54 male 32.775 0 no northeast 4.018495  
## 319 44 female 27.645 0 no northwest 3.870474  
## 320 32 male 37.335 1 no northeast 3.669094  
## 321 34 male 25.270 1 no northwest 3.689731  
## 322 26 female 29.640 4 no northeast 4.392198  
## 323 34 male 30.800 0 yes southwest 4.550126  
## 324 57 male 40.945 0 no northeast 4.063194  
## 325 29 male 27.200 0 no southwest 3.457290  
## 326 40 male 34.105 1 no northeast 3.819557  
## 327 27 female 23.210 1 no southeast 3.551680  
## 328 45 male 36.480 2 yes northwest 4.631043  
## 329 64 female 33.800 1 yes southwest 4.680590  
## 330 52 male 36.700 0 no southwest 3.961163  
## 331 61 female 36.385 1 yes northeast 4.685899  
## 332 52 male 27.360 0 yes northwest 4.387276  
## 333 61 female 31.160 0 no northwest 4.128045  
## 334 56 female 28.785 0 no northeast 4.066638  
## 335 43 female 35.720 2 no northeast 4.282046  
## 336 64 male 34.500 0 no southwest 4.140596  
## 337 60 male 25.740 0 no southeast 4.084311  
## 338 62 male 27.550 1 no northwest 4.144190  
## 339 50 male 32.300 1 yes northeast 4.622412  
## 340 46 female 27.720 1 no southeast 3.915539  
## 341 24 female 27.600 0 no southwest 4.277729  
## 342 62 male 30.020 0 no northwest 4.125550  
## 343 60 female 27.550 0 no northeast 4.121136  
## 344 63 male 36.765 0 no northeast 4.145565  
## 345 49 female 41.470 4 no southeast 4.040492  
## 346 34 female 29.260 3 no southeast 3.791291  
## 347 33 male 35.750 2 no southeast 3.689309  
## 348 46 male 33.345 1 no northeast 3.920877  
## 349 36 female 29.920 1 no southeast 3.738625  
## 350 19 male 27.835 0 no northwest 3.213713  
## 351 57 female 23.180 0 no northwest 4.073007  
## 352 50 female 25.600 0 no southwest 3.950953  
## 353 30 female 27.700 0 no southwest 3.550742  
## 354 33 male 35.245 0 no northeast 4.093593  
## 355 18 female 38.280 0 no southeast 4.150236  
## 356 46 male 27.600 0 no southwest 4.390989  
## 357 46 male 43.890 3 no southeast 3.951537  
## 358 47 male 29.830 3 no northwest 3.983190  
## 359 23 male 41.910 0 no southeast 3.264176  
## 360 18 female 20.790 0 no southeast 3.206154  
## 361 48 female 32.300 2 no northeast 4.001874  
## 362 35 male 30.500 1 no southwest 3.676791  
## 363 19 female 21.700 0 yes southwest 4.141277  
## 364 21 female 26.400 1 no southwest 3.414602  
## 365 21 female 21.890 2 no southeast 3.502497  
## 366 49 female 30.780 1 no northeast 3.990265  
## 367 56 female 32.300 3 no northeast 4.128085  
## 368 42 female 24.985 2 no northwest 3.904015  
## 369 44 male 32.015 2 no northwest 3.909356  
## 370 18 male 30.400 3 no northeast 3.541812  
## 371 61 female 21.090 0 no northwest 4.127592  
## 372 57 female 22.230 0 no northeast 4.080240  
## 373 42 female 33.155 1 no northeast 3.883060  
## 374 26 male 32.900 2 yes southwest 4.557329  
## 375 20 male 33.330 0 no southeast 3.143492  
## 376 23 female 28.310 0 yes northwest 4.256091  
## 377 39 female 24.890 3 yes northeast 4.335657  
## 378 24 male 40.150 0 yes southeast 4.581224  
## 379 64 female 30.115 3 no northwest 4.216317  
## 380 62 male 31.460 1 no southeast 4.431380  
## 381 27 female 17.955 2 yes northeast 4.176282  
## 382 55 male 30.685 0 yes northeast 4.626378  
## 383 55 male 33.000 0 no southeast 4.317677  
## 384 35 female 43.340 2 no southeast 3.766927  
## 385 44 male 22.135 2 no northeast 3.919211  
## 386 19 male 34.400 0 no southwest 3.101011  
## 387 58 female 39.050 0 no southeast 4.073953  
## 388 50 male 25.365 2 no northwest 4.481222  
## 389 26 female 22.610 0 no northwest 3.501992  
## 390 24 female 30.210 3 no northwest 3.664461  
## 391 48 male 35.625 4 no northeast 4.030878  
## 392 19 female 37.430 0 no northwest 3.330022  
## 393 48 male 31.445 1 no northeast 3.952505  
## 394 49 male 31.350 1 no northeast 3.968022  
## 395 46 female 32.300 2 no northeast 3.973636  
## 396 46 male 19.855 0 no northwest 3.876605  
## 397 43 female 34.400 3 no southwest 3.930542  
## 398 21 male 31.020 0 no southeast 4.219755  
## 399 64 male 25.600 2 no southwest 4.175756  
## 400 18 female 38.170 0 no southeast 3.212632  
## 401 51 female 20.600 0 no southwest 3.966836  
## 402 47 male 47.520 1 no southeast 3.907622  
## 403 64 female 32.965 0 no northwest 4.167101  
## 404 49 male 32.300 3 no northwest 4.011548  
## 405 31 male 20.400 0 no southwest 3.513244  
## 406 52 female 38.380 2 no northeast 4.056787  
## 407 33 female 24.310 0 no southeast 3.621706  
## 408 47 female 23.600 1 no southwest 3.931441  
## 409 38 male 21.120 3 no southeast 3.822987  
## 410 32 male 30.030 1 no southeast 3.610069  
## 411 19 male 17.480 0 no northwest 3.209874  
## 412 44 female 20.235 1 yes northeast 4.292141  
## 413 26 female 17.195 2 yes northeast 4.160037  
## 414 25 male 23.900 5 no southwest 3.705872  
## 415 19 female 35.150 0 no northwest 3.329378  
## 416 43 female 35.640 1 no southeast 3.866035  
## 417 52 male 34.100 0 no southeast 3.960991  
## 418 36 female 22.600 2 yes southwest 4.269706  
## 419 64 male 39.160 1 no southeast 4.158913  
## 420 63 female 26.980 0 yes northwest 4.461656  
## 421 64 male 33.880 0 yes southeast 4.671073  
## 422 61 male 35.860 0 yes southeast 4.668378  
## 423 40 male 32.775 1 yes northeast 4.592458  
## 424 25 male 30.590 0 no northeast 3.435748  
## 425 48 male 30.200 2 no southwest 3.952712  
## 426 45 male 24.310 5 no southeast 3.990732  
## 427 38 female 27.265 1 no northeast 3.816577  
## 428 18 female 29.165 0 no northeast 3.864733  
## 429 21 female 16.815 1 no northeast 3.500711  
## 430 27 female 30.400 3 no northwest 4.274268  
## 431 19 male 33.100 0 no southwest 4.363291  
## 432 29 female 20.235 2 no northwest 3.690764  
## 433 42 male 26.900 0 no southwest 3.775954  
## 434 60 female 30.500 0 no southwest 4.101685  
## 435 31 male 28.595 1 no northwest 3.627733  
## 436 60 male 33.110 3 no southeast 4.143634  
## 437 22 male 31.730 0 no northeast 3.353107  
## 438 35 male 28.900 3 no southwest 3.772824  
## 439 52 female 46.750 5 no southeast 4.100113  
## 440 26 male 29.450 0 no northeast 3.461997  
## 441 31 female 32.680 1 no northwest 3.675620  
## 442 33 female 33.500 0 yes southwest 4.569132  
## 443 18 male 43.010 0 no southeast 3.060470  
## 444 59 female 36.520 1 no southeast 4.451601  
## 445 56 male 26.695 1 yes northwest 4.416796  
## 446 45 female 33.100 0 no southwest 3.865997  
## 447 60 male 29.640 0 no northeast 4.104863  
## 448 56 female 25.650 0 no northwest 4.058958  
## 449 40 female 29.600 0 no southwest 3.771657  
## 450 35 male 38.600 1 no southwest 3.677819  
## 451 39 male 29.600 4 no southwest 3.875771  
## 452 30 male 24.130 1 no northwest 3.605546  
## 453 24 male 23.400 0 no southwest 3.294381  
## 454 20 male 29.735 0 no northwest 3.247858  
## 455 32 male 46.530 2 no southeast 3.670838  
## 456 59 male 37.400 0 no southwest 4.338397  
## 457 55 female 30.140 2 no southeast 4.074888  
## 458 57 female 30.495 0 no northwest 4.073380  
## 459 56 male 39.600 0 no southwest 4.025364  
## 460 40 female 33.000 3 no southeast 3.885512  
## 461 49 female 36.630 3 no southeast 4.016259  
## 462 42 male 30.000 0 yes southwest 4.345257  
## 463 62 female 38.095 2 no northeast 4.182709  
## 464 56 male 25.935 0 no northeast 4.047875  
## 465 19 male 25.175 0 no northwest 3.212730  
## 466 30 female 28.380 1 yes southeast 4.290524  
## 467 60 female 28.700 1 no southwest 4.121386  
## 468 56 female 33.820 2 no northwest 4.101863  
## 469 28 female 24.320 1 no northeast 4.367150  
## 470 18 female 24.090 1 no southeast 3.342639  
## 471 27 male 32.670 0 no southeast 3.397425  
## 472 18 female 30.115 0 no northeast 3.343108  
## 473 19 female 29.800 0 no southwest 3.241662  
## 474 47 female 33.345 0 no northeast 4.319705  
## 475 54 male 25.100 3 yes southwest 4.404531  
## 476 61 male 28.310 1 yes northwest 4.460427  
## 477 24 male 28.500 0 yes northeast 4.545895  
## 478 25 male 35.625 0 no northwest 3.403874  
## 479 21 male 36.850 0 no southeast 3.185912  
## 480 23 male 32.560 0 no southeast 3.261093  
## 481 63 male 41.325 3 no northwest 4.191875  
## 482 49 male 37.510 2 no southeast 3.968702  
## 483 18 female 31.350 0 no southeast 3.210101  
## 484 51 female 39.500 1 no southwest 3.994760  
## 485 48 male 34.300 3 no southwest 3.980595  
## 486 31 female 31.065 0 no northeast 3.638192  
## 487 54 female 21.470 3 no northwest 4.096053  
## 488 19 male 28.700 0 no southwest 3.098275  
## 489 44 female 38.060 0 yes southeast 4.689177  
## 490 53 male 31.160 1 no northwest 4.019614  
## 491 19 female 32.900 0 no southwest 3.242734  
## 492 61 female 25.080 0 no southeast 4.389398  
## 493 18 female 25.080 0 no northeast 3.341726  
## 494 61 male 43.400 0 no southwest 4.099475  
## 495 21 male 25.700 4 yes southwest 4.253873  
## 496 20 male 27.930 0 no northeast 3.293809  
## 497 31 female 23.600 2 no southwest 3.692992  
## 498 45 male 28.700 2 no southwest 3.904606  
## 499 44 female 23.980 2 no southeast 3.914401  
## 500 62 female 39.200 0 no southwest 4.129395  
## 501 29 male 34.400 0 yes southwest 4.558681  
## 502 43 male 26.030 0 no northeast 3.834889  
## 503 51 male 23.210 1 yes southeast 4.346707  
## 504 19 male 30.250 0 yes southeast 4.512529  
## 505 38 female 28.930 1 no southeast 3.776293  
## 506 37 male 30.875 3 no northwest 3.832309  
## 507 22 male 31.350 1 no northwest 3.422141  
## 508 21 male 23.750 2 no northwest 3.488141  
## 509 24 female 25.270 0 no northeast 3.483475  
## 510 57 female 28.700 0 no southwest 4.059006  
## 511 56 male 32.110 1 no northeast 4.070518  
## 512 27 male 33.660 0 no southeast 3.397664  
## 513 51 male 22.420 0 no northeast 3.971337  
## 514 19 male 30.400 0 no southwest 3.099093  
## 515 39 male 28.300 1 yes southwest 4.323915  
## 516 58 male 35.700 0 no southwest 4.055484  
## 517 20 male 35.310 1 no southeast 4.442860  
## 518 45 male 30.495 2 no northwest 3.924975  
## 519 35 female 31.000 1 no southwest 3.719395  
## 520 31 male 30.875 0 no northeast 3.586335  
## 521 50 female 27.360 0 no northeast 4.409199  
## 522 32 female 44.220 0 no southeast 3.601427  
## 523 51 female 33.915 0 no northeast 3.994155  
## 524 38 female 37.730 0 no southeast 3.732202  
## 525 42 male 26.070 1 yes southeast 4.582581  
## 526 18 female 33.880 0 no southeast 4.060042  
## 527 19 female 30.590 2 no northwest 4.381290  
## 528 51 female 25.800 1 no southwest 3.993922  
## 529 46 male 39.425 1 no northeast 3.921317  
## 530 18 male 25.460 0 no northeast 3.232488  
## 531 57 male 42.130 1 yes southeast 4.687311  
## 532 62 female 31.730 0 no northeast 4.147475  
## 533 59 male 29.700 2 no southeast 4.111460  
## 534 37 male 36.190 0 no southeast 4.283634  
## 535 64 male 40.480 0 no southeast 4.140857  
## 536 38 male 28.025 1 no northeast 3.782983  
## 537 33 female 38.900 3 no southwest 3.776147  
## 538 46 female 30.200 2 no southwest 3.945719  
## 539 46 female 28.050 1 no southeast 3.915563  
## 540 53 male 31.350 0 no southeast 4.436894  
## 541 34 female 38.000 3 no southwest 3.792143  
## 542 20 female 31.790 2 no southeast 3.485209  
## 543 63 female 36.300 0 no southeast 4.142615  
## 544 54 female 47.410 0 yes southeast 4.804619  
## 545 54 male 30.210 0 no northwest 4.009939  
## 546 49 male 25.840 2 yes northwest 4.376709  
## 547 28 male 35.435 0 no northeast 3.514395  
## 548 54 female 46.700 2 no southwest 4.062146  
## 549 25 female 28.595 0 no northeast 3.506995  
## 550 43 female 46.200 0 yes southeast 4.661464  
## 551 63 male 30.800 0 no southwest 4.126799  
## 552 32 female 28.930 0 no southeast 3.599110  
## 553 62 male 21.400 0 no southwest 4.112508  
## 554 52 female 31.730 2 no northwest 4.048739  
## 555 25 female 41.325 0 no northeast 4.252341  
## 556 28 male 23.800 2 no southwest 3.585198  
## 557 46 male 33.440 1 no northeast 3.920884  
## 558 34 male 34.210 0 no southeast 3.594965  
## 559 35 female 34.105 3 yes northwest 4.601880  
## 560 19 male 35.530 0 no northwest 3.216543  
## 561 46 female 19.950 2 no northwest 3.963497  
## 562 54 female 32.680 0 no northeast 4.038379  
## 563 27 male 30.500 0 no southwest 3.396900  
## 564 50 male 44.770 1 no southeast 3.957067  
## 565 18 female 32.120 2 no southeast 3.447353  
## 566 19 female 30.495 0 no northwest 3.328060  
## 567 38 female 40.565 1 no northwest 3.804382  
## 568 41 male 30.590 2 no northwest 3.860741  
## 569 49 female 31.900 5 no southwest 4.062691  
## 570 48 male 40.565 2 yes northwest 4.659935  
## 571 31 female 29.100 0 no southwest 3.575337  
## 572 18 female 37.290 1 no southeast 3.346244  
## 573 30 female 43.120 2 no southeast 3.677026  
## 574 62 female 36.860 1 no northeast 4.499962  
## 575 57 female 34.295 2 no northeast 4.121365  
## 576 58 female 27.170 0 no northwest 4.087174  
## 577 22 male 26.840 0 no southeast 3.221414  
## 578 31 female 38.095 1 yes northeast 4.767683  
## 579 52 male 30.200 1 no southwest 3.987869  
## 580 25 female 23.465 0 no northeast 3.506030  
## 581 59 male 25.460 1 no northeast 4.111061  
## 582 19 male 30.590 0 no northwest 3.214728  
## 583 39 male 45.430 2 no southeast 3.803202  
## 584 32 female 23.650 1 no southeast 4.246160  
## 585 19 male 20.700 0 no southwest 3.094407  
## 586 33 female 28.270 1 no southeast 3.679392  
## 587 21 male 20.235 3 no northeast 3.586723  
## 588 34 female 30.210 1 yes northwest 4.642898  
## 589 61 female 35.910 0 no northeast 4.134675  
## 590 38 female 30.690 1 no southeast 3.776471  
## 591 58 female 29.000 0 no southwest 4.073441  
## 592 47 male 19.570 1 no northwest 3.925728  
## 593 20 male 31.130 2 no southeast 3.409336  
## 594 21 female 21.850 1 yes northeast 4.186366  
## 595 41 male 40.260 0 no southeast 3.756573  
## 596 46 female 33.725 1 no northeast 3.945665  
## 597 42 female 29.480 2 no southeast 3.883111  
## 598 34 female 33.250 1 no northeast 3.747788  
## 599 43 male 32.600 2 no southwest 3.871661  
## 600 52 female 37.525 2 no northwest 4.524681  
## 601 18 female 39.160 0 no southeast 3.212998  
## 602 51 male 31.635 0 no northwest 3.962565  
## 603 56 female 25.300 0 no southwest 4.044169  
## 604 64 female 39.050 3 no southeast 4.206425  
## 605 19 female 28.310 0 yes northwest 4.242268  
## 606 51 female 34.100 0 no southeast 3.967715  
## 607 27 female 25.175 0 no northeast 3.551282  
## 608 59 female 23.655 0 yes northwest 4.409574  
## 609 28 male 26.980 2 no northeast 3.646903  
## 610 30 male 37.800 2 yes southwest 4.593745  
## 611 47 female 29.370 1 no southeast 3.931849  
## 612 38 female 34.800 2 no southwest 3.817667  
## 613 18 female 33.155 0 no northeast 3.343940  
## 614 34 female 19.000 3 no northeast 3.829499  
## 615 20 female 33.000 0 no southeast 3.274174  
## 616 47 female 36.630 1 yes southeast 4.633164  
## 617 56 female 28.595 0 no northeast 4.066628  
## 618 49 male 25.600 2 yes southwest 4.367478  
## 619 19 female 33.110 0 yes southeast 4.537061  
## 620 55 female 37.100 0 no southwest 4.029937  
## 621 30 male 31.400 1 no southwest 3.563403  
## 622 37 male 34.100 4 yes southwest 4.604034  
## 623 49 female 21.300 1 no southwest 3.962945  
## 624 18 male 33.535 0 yes northeast 4.539300  
## 625 59 male 28.785 0 no northwest 4.083847  
## 626 29 female 26.030 0 no northwest 3.572461  
## 627 36 male 28.880 3 no northeast 3.829213  
## 628 33 male 42.460 1 no southeast 4.054104  
## 629 58 male 38.000 0 no southwest 4.055606  
## 630 44 female 38.950 0 yes northwest 4.633301  
## 631 53 male 36.100 1 no southwest 4.003712  
## 632 24 male 29.300 0 no southwest 3.296186  
## 633 29 female 35.530 0 no southeast 3.527201  
## 634 40 male 22.705 2 no northeast 3.855723  
## 635 51 male 39.700 1 no southwest 3.972728  
## 636 64 male 38.190 0 no northeast 4.158692  
## 637 19 female 24.510 1 no northwest 3.432827  
## 638 35 female 38.095 2 no northeast 4.396462  
## 639 39 male 26.410 0 yes northeast 4.304260  
## 640 56 male 33.660 4 no southeast 4.112241  
## 641 33 male 42.400 5 no southwest 3.823881  
## 642 42 male 28.310 3 yes northwest 4.515708  
## 643 61 male 33.915 0 no northeast 4.118723  
## 644 23 female 34.960 3 no northwest 3.649979  
## 645 43 male 35.310 2 no southeast 4.274300  
## 646 48 male 30.780 3 no northeast 4.006087  
## 647 39 male 26.220 1 no northwest 3.787005  
## 648 40 female 23.370 3 no northeast 3.916574  
## 649 18 male 28.500 0 no northeast 3.233561  
## 650 58 female 32.965 0 no northeast 4.094504  
## 651 49 female 42.680 2 no southeast 3.991265  
## 652 53 female 39.600 1 no southeast 4.024474  
## 653 48 female 31.130 0 no southeast 3.918063  
## 654 45 female 36.300 2 no southeast 3.930823  
## 655 59 female 35.200 0 no southeast 4.087942  
## 656 52 female 25.300 2 yes southeast 4.392124  
## 657 26 female 42.400 1 no southwest 3.532796  
## 658 27 male 33.155 2 no northwest 3.608388  
## 659 48 female 35.910 1 no northeast 4.421477  
## 660 57 female 28.785 4 no northeast 4.158194  
## 661 37 male 46.530 3 no southeast 3.808591  
## 662 57 female 23.980 1 no southeast 4.346205  
## 663 32 female 31.540 1 no northeast 3.711685  
## 664 18 male 33.660 0 no southeast 3.055531  
## 665 64 female 22.990 0 yes southeast 4.431973  
## 666 43 male 38.060 2 yes southeast 4.629006  
## 667 49 male 28.700 1 no southwest 3.939692  
## 668 40 female 32.775 2 yes northwest 4.602096  
## 669 62 male 32.015 0 yes northeast 4.660013  
## 670 40 female 29.810 1 no southeast 3.812929  
## 671 30 male 31.570 3 no southeast 3.684628  
## 672 29 female 31.160 0 no northeast 3.595892  
## 673 36 male 29.700 0 no southeast 3.643426  
## 674 41 female 31.020 0 no southeast 3.791362  
## 675 44 female 43.890 2 yes southeast 4.664651  
## 676 45 male 21.375 0 no northwest 3.858705  
## 677 55 female 40.810 3 no southeast 4.096416  
## 678 60 male 31.350 3 yes northwest 4.663988  
## 679 56 male 36.100 3 no southwest 4.092143  
## 680 49 female 23.180 2 no northwest 4.006756  
## 681 21 female 17.400 1 no southwest 3.412506  
## 682 19 male 20.300 0 no southwest 3.094213  
## 683 39 male 35.300 2 yes southwest 4.603187  
## 684 53 male 24.320 0 no northwest 3.994030  
## 685 33 female 18.500 1 no southwest 3.678156  
## 686 53 male 26.410 2 no northeast 4.050935  
## 687 42 male 26.125 2 no northeast 3.888160  
## 688 40 male 41.690 0 no southeast 3.735499  
## 689 47 female 24.100 1 no southwest 4.418907  
## 690 27 male 31.130 1 yes southeast 4.541660  
## 691 21 male 27.360 0 no northeast 3.323069  
## 692 47 male 36.200 1 no southwest 3.906776  
## 693 20 male 32.395 1 no northwest 3.373322  
## 694 24 male 23.655 0 no northwest 3.371616  
## 695 27 female 34.800 1 no southwest 3.553640  
## 696 26 female 40.185 0 no northwest 3.505319  
## 697 53 female 32.300 2 no northeast 4.465182  
## 698 41 male 35.750 1 yes southeast 4.605021  
## 699 56 male 33.725 0 no northwest 4.040454  
## 700 23 female 39.270 2 no southeast 3.544144  
## 701 21 female 34.870 0 no southeast 3.305470  
## 702 50 female 44.745 0 no northeast 3.979626  
## 703 53 male 41.470 0 no southeast 3.977921  
## 704 34 female 26.410 1 no northwest 3.731213  
## 705 47 female 29.545 1 no northwest 3.950897  
## 706 33 female 32.900 2 no southwest 3.730382  
## 707 51 female 38.060 0 yes southeast 4.647387  
## 708 49 male 28.690 3 no northwest 4.011335  
## 709 31 female 30.495 3 no northeast 3.786271  
## 710 36 female 27.740 0 no northeast 3.737908  
## 711 18 male 35.200 1 no southeast 3.237428  
## 712 50 female 23.540 2 no southeast 4.004632  
## 713 43 female 30.685 2 no northwest 3.919645  
## 714 20 male 40.470 0 no northeast 3.297641  
## 715 24 female 22.600 0 no southwest 3.390494  
## 716 60 male 28.900 0 no southwest 4.084468  
## 717 49 female 22.610 1 no northwest 3.980775  
## 718 60 male 24.320 1 no northwest 4.117689  
## 719 51 female 36.670 2 no northwest 4.035355  
## 720 58 female 33.440 0 no northwest 4.087484  
## 721 51 female 40.660 0 no northeast 3.994567  
## 722 53 male 36.600 3 no southwest 4.051714  
## 723 62 male 37.400 0 no southwest 4.113253  
## 724 19 male 35.400 0 no southwest 3.101489  
## 725 50 female 27.075 1 no northeast 4.004585  
## 726 30 female 39.050 3 yes southeast 4.612068  
## 727 41 male 28.405 1 no northwest 3.823780  
## 728 29 female 21.755 1 yes northeast 4.221615  
## 729 18 female 40.280 0 no northeast 3.345883  
## 730 41 female 36.080 1 no southeast 3.831316  
## 731 35 male 24.420 3 yes southeast 4.286950  
## 732 53 male 21.400 1 no southwest 4.002832  
## 733 24 female 30.100 3 no southwest 3.626846  
## 734 48 female 27.265 1 no northeast 3.975305  
## 735 59 female 32.100 3 no southwest 4.146352  
## 736 49 female 34.770 1 no northwest 3.981542  
## 737 37 female 38.390 0 yes southeast 4.606586  
## 738 26 male 23.700 2 no southwest 3.542119  
## 739 23 male 31.730 3 yes northeast 4.558578  
## 740 29 male 35.500 2 yes southwest 4.649193  
## 741 45 male 24.035 2 no northeast 3.934725  
## 742 27 male 29.150 0 yes southeast 4.261179  
## 743 53 male 34.105 0 yes northeast 4.636030  
## 744 31 female 26.620 0 no southeast 3.574939  
## 745 50 male 26.410 0 no northwest 3.945823  
## 746 50 female 30.115 1 no northwest 3.996089  
## 747 34 male 27.000 2 no southwest 4.069589  
## 748 19 male 21.755 0 no northwest 3.211463  
## 749 47 female 36.000 1 no southwest 3.932317  
## 750 28 male 30.875 0 no northwest 3.486077  
## 751 37 female 26.400 0 yes southeast 4.290908  
## 752 21 male 28.975 0 no northwest 3.280205  
## 753 64 male 37.905 0 no northwest 4.152610  
## 754 58 female 22.770 0 no southeast 4.073124  
## 755 24 male 33.630 4 no northeast 4.233717  
## 756 31 male 27.645 2 no northeast 3.701678  
## 757 39 female 22.800 3 no northeast 3.902319  
## 758 47 female 27.830 0 yes southeast 4.362961  
## 759 30 male 37.430 3 no northeast 3.734698  
## 760 18 male 38.170 0 yes southeast 4.560000  
## 761 22 female 34.580 2 no northeast 3.593924  
## 762 23 male 35.200 1 no southwest 3.383269  
## 763 33 male 27.100 1 yes southwest 4.279687  
## 764 27 male 26.030 0 no northeast 3.487253  
## 765 45 female 25.175 2 no northeast 3.958806  
## 766 57 female 31.825 0 no northwest 4.073448  
## 767 47 male 32.300 1 no southwest 3.906484  
## 768 42 female 29.000 1 no southwest 3.848229  
## 769 64 female 39.700 0 no southwest 4.155914  
## 770 38 female 19.475 2 no northwest 3.840936  
## 771 61 male 36.100 3 no southwest 4.446246  
## 772 53 female 26.700 2 no southwest 4.047305  
## 773 44 female 36.480 0 no northeast 4.107115  
## 774 19 female 28.880 0 yes northwest 4.249162  
## 775 41 male 34.200 2 no northwest 3.861041  
## 776 51 male 33.330 3 no southeast 4.023684  
## 777 40 male 32.300 2 no northwest 3.844272  
## 778 45 male 39.805 0 no northeast 3.872063  
## 779 35 male 34.320 3 no southeast 3.773375  
## 780 53 male 28.880 0 no northwest 3.994309  
## 781 30 male 24.400 3 yes southwest 4.261482  
## 782 18 male 41.140 0 no southeast 3.059486  
## 783 51 male 35.970 1 no southeast 3.972488  
## 784 50 female 27.600 1 yes southwest 4.389525  
## 785 31 female 29.260 1 no southeast 3.638541  
## 786 35 female 27.700 3 no southwest 3.807141  
## 787 60 male 36.955 0 no northeast 4.105209  
## 788 21 male 36.860 0 no northwest 3.282694  
## 789 29 male 22.515 3 no northeast 3.716803  
## 790 62 female 29.920 0 no southeast 4.128979  
## 791 39 female 41.800 0 no southeast 3.752987  
## 792 19 male 27.600 0 no southwest 3.097745  
## 793 22 female 23.180 0 no northeast 3.436467  
## 794 53 male 20.900 0 yes southeast 4.326250  
## 795 39 female 31.920 2 no northwest 3.857905  
## 796 27 male 28.500 0 yes northwest 4.262706  
## 797 30 male 44.220 2 no southeast 3.630038  
## 798 30 female 22.895 1 no northeast 3.673898  
## 799 58 female 33.100 0 no southwest 4.073650  
## 800 33 male 24.795 0 yes northeast 4.252963  
## 801 42 female 26.180 1 no southeast 3.847987  
## 802 64 female 35.970 0 no southeast 4.155756  
## 803 21 male 22.300 1 no southwest 3.322856  
## 804 18 female 42.240 0 yes southeast 4.588750  
## 805 23 male 26.510 0 no southeast 3.259086  
## 806 45 female 35.815 0 no northwest 3.888284  
## 807 40 female 41.420 1 no northwest 4.454490  
## 808 19 female 36.575 0 no northwest 3.329781  
## 809 18 male 30.140 0 no southeast 3.053657  
## 810 25 male 25.840 1 no northeast 3.519801  
## 811 46 female 30.800 3 no southwest 3.973817  
## 812 33 female 42.940 3 no northwest 3.803525  
## 813 54 male 21.010 2 no southeast 4.041934  
## 814 28 male 22.515 2 no northeast 3.646295  
## 815 36 male 34.430 2 no southeast 3.746969  
## 816 20 female 31.460 0 no southeast 3.273679  
## 817 24 female 24.225 0 no northwest 3.453740  
## 818 23 male 37.100 3 no southwest 3.556012  
## 819 47 female 26.125 1 yes northeast 4.369240  
## 820 33 female 35.530 0 yes northwest 4.741431  
## 821 45 male 33.700 1 no southwest 3.871918  
## 822 26 male 17.670 0 no northwest 3.428289  
## 823 18 female 31.130 0 no southeast 3.210019  
## 824 44 female 29.810 2 no southeast 3.914830  
## 825 60 male 24.320 0 no northwest 4.097729  
## 826 64 female 31.825 2 no northeast 4.205991  
## 827 56 male 31.790 2 yes southeast 4.641612  
## 828 36 male 28.025 1 yes northeast 4.317512  
## 829 41 male 30.780 3 yes northeast 4.597667  
## 830 39 male 21.850 1 no northwest 3.786574  
## 831 63 male 33.100 0 no southwest 4.126902  
## 832 36 female 25.840 0 no northwest 3.721511  
## 833 28 female 23.845 2 no northwest 3.673918  
## 834 58 male 34.390 0 no northwest 4.069814  
## 835 36 male 33.820 1 no northwest 3.730577  
## 836 42 male 35.970 2 no southeast 3.854933  
## 837 36 male 31.500 0 no southwest 3.643673  
## 838 56 female 28.310 0 no northeast 4.066614  
## 839 35 female 23.465 2 no northeast 3.806335  
## 840 59 female 31.350 0 no northwest 4.101134  
## 841 21 male 31.100 0 no southwest 3.183643  
## 842 59 male 24.700 0 no northeast 4.090749  
## 843 23 female 32.780 2 yes southeast 4.556556  
## 844 57 female 29.810 0 yes southeast 4.439868  
## 845 53 male 30.495 0 no northeast 4.003118  
## 846 60 female 32.450 0 yes southeast 4.653299  
## 847 51 female 34.200 1 no southwest 3.994436  
## 848 23 male 50.380 1 no southeast 3.387044  
## 849 27 female 24.100 0 no southwest 3.473359  
## 850 55 male 32.775 0 no northwest 4.025373  
## 851 37 female 30.780 0 yes northeast 4.571361  
## 852 61 male 32.300 2 no northwest 4.149823  
## 853 46 female 35.530 0 yes northeast 4.624402  
## 854 53 female 23.750 2 no northeast 4.069286  
## 855 49 female 23.845 3 yes northeast 4.382142  
## 856 20 female 29.600 0 no southwest 3.273081  
## 857 48 female 33.110 0 yes southeast 4.612510  
## 858 25 male 24.130 0 yes northwest 4.199151  
## 859 25 female 32.230 1 no southeast 4.260505  
## 860 57 male 28.100 0 no southwest 4.040026  
## 861 37 female 47.600 2 yes southwest 4.663828  
## 862 38 female 28.000 3 no southwest 3.854372  
## 863 55 female 33.535 2 no northwest 4.088834  
## 864 36 female 19.855 0 no northeast 3.737037  
## 865 51 male 25.400 0 no southwest 3.943617  
## 866 40 male 29.900 2 no southwest 3.819568  
## 867 18 male 37.290 0 no southeast 3.057455  
## 868 57 male 43.700 1 no southwest 4.063563  
## 869 61 male 23.655 0 no northeast 4.118252  
## 870 25 female 24.300 3 no southwest 3.642628  
## 871 50 male 36.200 0 no southwest 3.927258  
## 872 26 female 29.480 1 no southeast 3.530503  
## 873 42 male 24.860 0 no southeast 3.775748  
## 874 43 male 30.100 1 no southwest 3.835629  
## 875 44 male 21.850 3 no northeast 3.948957  
## 876 23 female 28.120 0 no northwest 3.429771  
## 877 49 female 27.100 1 no southwest 4.417312  
## 878 33 male 33.440 5 no southeast 3.823069  
## 879 41 male 28.800 1 no southwest 3.798114  
## 880 37 female 29.500 2 no southwest 3.800164  
## 881 22 male 34.800 3 no southwest 3.536945  
## 882 23 male 27.360 1 no northwest 3.445457  
## 883 21 female 22.135 0 no northeast 3.412603  
## 884 51 female 37.050 3 yes northeast 4.665160  
## 885 25 male 26.695 4 no northwest 3.688240  
## 886 32 male 28.930 1 yes southeast 4.294900  
## 887 57 male 28.975 0 yes northeast 4.434863  
## 888 36 female 30.020 0 no northwest 3.721990  
## 889 22 male 39.500 0 no southwest 3.225980  
## 890 57 male 33.630 1 no northwest 4.077191  
## 891 64 female 26.885 0 yes northwest 4.467327  
## 892 36 female 29.040 4 no southeast 3.859967  
## 893 54 male 24.035 0 no northeast 4.017989  
## 894 47 male 38.940 2 yes southeast 4.645448  
## 895 62 male 32.110 0 no northeast 4.132100  
## 896 61 female 44.000 0 no southwest 4.116072  
## 897 43 female 20.045 2 yes northeast 4.296623  
## 898 19 male 25.555 1 no northwest 3.346659  
## 899 18 female 40.260 0 no southeast 3.213404  
## 900 19 female 22.515 0 no northwest 3.325790  
## 901 49 male 22.515 0 no northeast 3.938963  
## 902 60 male 40.920 0 yes southeast 4.687293  
## 903 26 male 27.265 3 no northeast 3.668506  
## 904 49 male 36.850 0 no southeast 3.909865  
## 905 60 female 35.100 0 no southwest 4.101905  
## 906 26 female 29.355 2 no northeast 3.659364  
## 907 27 male 32.585 3 no northeast 3.685466  
## 908 44 female 32.340 1 no southeast 3.882736  
## 909 63 male 39.800 3 no southwest 4.180988  
## 910 32 female 24.600 0 yes southwest 4.242946  
## 911 22 male 28.310 1 no northwest 3.421446  
## 912 18 male 31.730 0 yes northeast 4.528051  
## 913 59 female 26.695 3 no northwest 4.157841  
## 914 44 female 27.500 1 no southwest 3.882353  
## 915 33 male 24.605 2 no northwest 3.720780  
## 916 24 female 33.990 0 no southeast 3.393283  
## 917 43 female 26.885 0 yes northwest 4.337945  
## 918 45 male 22.895 0 yes northeast 4.544928  
## 919 61 female 28.200 0 no southwest 4.115342  
## 920 35 female 34.210 1 no southeast 3.719764  
## 921 62 female 25.000 0 no southwest 4.128759  
## 922 62 female 33.200 0 no southwest 4.129126  
## 923 38 male 31.000 1 no southwest 3.739435  
## 924 34 male 35.815 0 no northwest 3.635525  
## 925 43 male 23.200 0 no southwest 3.795910  
## 926 50 male 32.110 2 no northeast 4.403692  
## 927 19 female 23.400 2 no southwest 3.464425  
## 928 57 female 20.100 1 no southwest 4.080350  
## 929 62 female 39.160 0 no southeast 4.129394  
## 930 41 male 34.210 1 no southeast 3.798634  
## 931 26 male 46.530 1 no southeast 3.466432  
## 932 39 female 32.500 1 no southwest 3.795066  
## 933 46 male 25.800 5 no southwest 4.004191  
## 934 45 female 35.300 0 no southwest 3.866178  
## 935 32 male 37.180 2 no southeast 3.669632  
## 936 59 female 27.500 0 no southwest 4.087562  
## 937 44 male 29.735 2 no northeast 4.506622  
## 938 39 female 24.225 5 no northwest 3.952589  
## 939 18 male 26.180 2 no southeast 3.362483  
## 940 53 male 29.480 0 no southeast 3.977158  
## 941 18 male 23.210 0 no southeast 3.049944  
## 942 50 female 46.090 1 no southeast 3.979984  
## 943 18 female 40.185 0 no northeast 3.345858  
## 944 19 male 22.610 0 no northwest 3.211780  
## 945 62 male 39.930 0 no southeast 4.113371  
## 946 56 female 35.800 1 no southwest 4.067225  
## 947 42 male 35.800 2 no southwest 3.854919  
## 948 37 male 34.200 1 yes northeast 4.591591  
## 949 42 male 31.255 0 no northwest 3.803374  
## 950 25 male 29.700 3 yes southwest 4.299583  
## 951 57 male 18.335 0 no northeast 4.062013  
## 952 51 male 42.900 2 yes southeast 4.676354  
## 953 30 female 28.405 1 no northwest 3.655828  
## 954 44 male 30.200 2 yes southwest 4.591048  
## 955 34 male 27.835 1 yes northwest 4.301239  
## 956 31 male 39.490 1 no southeast 3.588354  
## 957 54 male 30.800 1 yes southeast 4.623244  
## 958 24 male 26.790 1 no northwest 4.100711  
## 959 43 male 34.960 1 yes northeast 4.613146  
## 960 48 male 36.670 1 no northwest 4.454371  
## 961 19 female 39.615 1 no northwest 3.436180  
## 962 29 female 25.900 0 no southwest 3.525470  
## 963 63 female 35.200 1 no southeast 4.160609  
## 964 46 male 24.795 3 no northeast 3.977750  
## 965 52 male 36.765 2 no northwest 4.422706  
## 966 35 male 27.100 1 no southwest 3.676359  
## 967 51 male 24.795 2 yes northwest 4.379621  
## 968 44 male 25.365 1 no northwest 3.876104  
## 969 21 male 25.745 2 no northeast 3.515856  
## 970 39 female 34.320 5 no southeast 3.934338  
## 971 50 female 28.160 3 no southeast 4.029491  
## 972 34 female 23.560 0 no northeast 3.698307  
## 973 22 female 20.235 0 no northwest 3.402746  
## 974 19 female 40.500 0 no southwest 3.245349  
## 975 26 male 35.420 0 no southeast 3.365978  
## 976 29 male 22.895 0 yes northeast 4.207870  
## 977 48 male 40.150 0 no southeast 3.892326  
## 978 26 male 29.150 1 no southeast 3.462833  
## 979 45 female 39.995 3 no northeast 3.986981  
## 980 36 female 29.920 0 no southeast 3.689223  
## 981 54 male 25.460 1 no northeast 4.406832  
## 982 34 male 21.375 0 no northeast 3.653245  
## 983 31 male 25.900 3 yes southwest 4.283300  
## 984 27 female 30.590 1 no northeast 4.225217  
## 985 20 male 30.115 5 no northeast 3.691529  
## 986 44 female 25.800 1 no southwest 3.882219  
## 987 43 male 30.115 3 no northwest 3.924798  
## 988 45 female 27.645 1 no northwest 4.452403  
## 989 34 male 34.675 0 no northeast 3.655026  
## 990 24 female 20.520 0 yes northeast 4.163516  
## 991 26 female 19.800 1 no southwest 3.528777  
## 992 38 female 27.835 2 no northeast 3.853994  
## 993 50 female 31.600 2 no southwest 4.005113  
## 994 38 male 28.270 1 no southeast 3.739134  
## 995 27 female 20.045 3 yes northwest 4.215386  
## 996 39 female 23.275 3 no northeast 3.902355  
## 997 39 female 34.100 3 no southwest 3.870317  
## 998 63 female 36.850 0 no southeast 4.142639  
## 999 33 female 36.290 3 no northeast 3.816357  
## 1000 36 female 26.885 0 no northwest 3.721631  
## 1001 30 male 22.990 2 yes northwest 4.239594  
## 1002 24 male 32.700 0 yes southwest 4.537477  
## 1003 24 male 25.800 0 no southwest 3.295116  
## 1004 48 male 29.600 0 no southwest 4.326995  
## 1005 47 male 19.190 1 no northeast 3.935887  
## 1006 29 male 31.730 2 no northwest 3.646736  
## 1007 28 male 29.260 2 no northeast 3.647213  
## 1008 47 male 28.215 3 yes northwest 4.396465  
## 1009 25 male 24.985 2 no northeast 4.366264  
## 1010 51 male 27.740 1 no northeast 3.998160  
## 1011 48 female 22.800 0 no southwest 3.917455  
## 1012 43 male 20.130 2 yes southeast 4.273412  
## 1013 61 female 33.330 4 no southeast 4.563247  
## 1014 48 male 32.300 1 no northwest 3.942764  
## 1015 38 female 27.600 0 no southwest 3.731068  
## 1016 59 male 25.460 0 no northwest 4.083681  
## 1017 19 female 24.605 1 no northwest 3.432848  
## 1018 26 female 34.200 2 no southwest 3.600747  
## 1019 54 female 35.815 3 no northwest 4.096746  
## 1020 21 female 32.680 2 no northwest 4.415290  
## 1021 51 male 37.000 0 no southwest 3.944413  
## 1022 22 female 31.020 3 yes southeast 4.551396  
## 1023 47 male 36.080 1 yes southeast 4.625427  
## 1024 18 male 23.320 1 no southeast 3.233257  
## 1025 47 female 45.320 1 no southeast 3.932974  
## 1026 21 female 34.600 0 no southwest 3.305389  
## 1027 19 male 26.030 1 yes northwest 4.216190  
## 1028 23 male 18.715 0 no northwest 4.334361  
## 1029 54 male 31.600 0 no southwest 3.993455  
## 1030 37 female 17.290 2 no northeast 3.837461  
## 1031 46 female 23.655 1 yes northwest 4.336005  
## 1032 55 female 35.200 0 yes southeast 4.647616  
## 1033 30 female 27.930 0 no northeast 3.616740  
## 1034 18 male 21.565 0 yes northeast 4.138235  
## 1035 61 male 38.380 0 no northwest 4.112272  
## 1036 54 female 23.000 3 no southwest 4.082587  
## 1037 22 male 37.070 2 yes southeast 4.573851  
## 1038 45 female 30.495 1 yes northwest 4.599070  
## 1039 22 male 28.880 0 no northeast 3.352344  
## 1040 19 male 27.265 2 no northwest 4.352060  
## 1041 35 female 28.025 0 yes northwest 4.306100  
## 1042 18 male 23.085 0 no northeast 3.231648  
## 1043 20 male 30.685 0 yes northeast 4.524731  
## 1044 28 female 25.800 0 no southwest 3.499887  
## 1045 55 male 35.245 1 no northeast 4.056679  
## 1046 43 female 24.700 2 yes northwest 4.340064  
## 1047 43 female 25.080 0 no northeast 3.864810  
## 1048 22 male 52.580 1 yes southeast 4.648374  
## 1049 25 female 22.515 1 no northwest 3.555599  
## 1050 49 male 30.900 0 yes southwest 4.599092  
## 1051 44 female 36.955 1 no northwest 3.904344  
## 1052 64 male 26.410 0 no northeast 4.158198  
## 1053 49 male 29.830 1 no northeast 3.967923  
## 1054 47 male 29.800 3 yes southwest 4.403283  
## 1055 27 female 21.470 0 no northwest 3.525494  
## 1056 55 male 27.645 0 no northwest 4.025081  
## 1057 48 female 28.900 0 no southwest 3.917900  
## 1058 45 female 31.790 0 no southeast 4.253563  
## 1059 24 female 39.490 0 no southeast 3.394623  
## 1060 32 male 33.820 1 no northwest 3.649600  
## 1061 24 male 32.010 0 no southeast 3.297012  
## 1062 57 male 27.940 1 no southeast 4.062741  
## 1063 59 male 41.140 1 yes southeast 4.689932  
## 1064 36 male 28.595 3 no northwest 3.816122  
## 1065 29 female 25.600 4 no southwest 3.756550  
## 1066 42 female 25.300 1 no southwest 3.847912  
## 1067 48 male 37.290 2 no southeast 3.953189  
## 1068 39 male 42.655 0 no northeast 3.760227  
## 1069 63 male 21.660 1 no northwest 4.156847  
## 1070 54 female 31.900 1 no southeast 4.038574  
## 1071 37 male 37.070 1 yes southeast 4.600665  
## 1072 63 male 31.445 0 no northeast 4.145335  
## 1073 21 male 31.255 0 no northwest 3.280926  
## 1074 54 female 28.880 2 no northeast 4.082665  
## 1075 60 female 18.335 0 no northeast 4.120715  
## 1076 32 female 29.590 1 no southeast 3.659235  
## 1077 47 female 32.000 1 no southwest 3.932035  
## 1078 21 male 26.030 0 no northeast 3.322687  
## 1079 28 male 31.680 0 yes southeast 4.539981  
## 1080 63 male 33.660 3 no southeast 4.180743  
## 1081 18 male 21.780 2 no southeast 4.074964  
## 1082 32 male 27.835 1 no northwest 3.648789  
## 1083 38 male 19.950 1 no northwest 3.767594  
## 1084 32 male 31.500 1 no southwest 3.610287  
## 1085 62 female 30.495 2 no northwest 4.176663  
## 1086 39 female 18.300 5 yes southwest 4.279285  
## 1087 55 male 28.975 0 no northeast 4.033277  
## 1088 57 male 31.540 0 no northwest 4.055119  
## 1089 52 male 47.740 1 no southeast 3.988956  
## 1090 56 male 22.100 0 no southwest 4.024366  
## 1091 47 male 36.190 0 yes southeast 4.619887  
## 1092 55 female 29.830 0 no northeast 4.052561  
## 1093 23 male 32.700 3 no southwest 3.555273  
## 1094 22 female 30.400 0 yes northwest 4.530296  
## 1095 50 female 33.700 4 no southwest 4.053053  
## 1096 18 female 31.350 4 no northeast 3.659078  
## 1097 51 female 34.960 2 yes northeast 4.649736  
## 1098 22 male 33.770 0 no southeast 3.223919  
## 1099 52 female 30.875 0 no northeast 4.362587  
## 1100 25 female 33.990 1 no southeast 3.508815  
## 1101 33 female 19.095 2 yes northeast 4.224696  
## 1102 53 male 28.600 3 no southwest 4.051285  
## 1103 29 male 38.940 1 no southeast 3.540506  
## 1104 58 male 36.080 0 no southeast 4.055504  
## 1105 37 male 29.800 0 no southwest 4.310069  
## 1106 54 female 31.240 0 no southeast 4.014476  
## 1107 49 female 29.925 0 no northwest 3.953671  
## 1108 50 female 26.220 2 no northwest 4.020939  
## 1109 26 male 30.000 1 no southwest 3.463010  
## 1110 45 male 20.350 3 no southeast 3.934769  
## 1111 54 female 32.300 1 no northeast 4.061166  
## 1112 38 male 38.390 3 yes southeast 4.622724  
## 1113 48 female 25.850 3 yes southeast 4.383473  
## 1114 28 female 26.315 3 no northwest 3.725272  
## 1115 23 male 24.510 0 no northeast 3.379504  
## 1116 55 male 32.670 1 no southeast 4.033725  
## 1117 41 male 29.640 5 no northeast 3.964844  
## 1118 25 male 33.330 2 yes southeast 4.557803  
## 1119 33 male 35.750 1 yes southeast 4.583003  
## 1120 30 female 19.950 3 no northwest 3.755374  
## 1121 23 female 31.400 0 yes southwest 4.533598  
## 1122 46 male 38.170 2 no southeast 3.921539  
## 1123 53 female 36.860 3 yes northwest 4.668958  
## 1124 27 female 32.395 1 no northeast 4.276542  
## 1125 23 female 42.750 1 yes northeast 4.611768  
## 1126 63 female 25.080 0 no northwest 4.153955  
## 1127 55 male 29.900 0 no southwest 4.009223  
## 1128 35 female 35.860 2 no southeast 3.766154  
## 1129 34 male 32.800 1 no southwest 4.157105  
## 1130 19 female 18.600 0 no southwest 3.237769  
## 1131 39 female 23.870 5 no southeast 3.933604  
## 1132 27 male 45.900 2 no southwest 3.567430  
## 1133 57 male 40.280 0 no northeast 4.316160  
## 1134 52 female 18.335 0 no northwest 3.999611  
## 1135 28 male 33.820 0 no northwest 4.293878  
## 1136 50 female 28.120 3 no northwest 4.044759  
## 1137 44 female 25.000 1 no southwest 3.882155  
## 1138 26 female 22.230 0 no northwest 3.501920  
## 1139 33 male 30.250 0 no southeast 3.568713  
## 1140 19 female 32.490 0 yes northwest 4.567011  
## 1141 50 male 37.070 1 no southeast 3.956554  
## 1142 41 female 32.600 3 no southwest 3.900614  
## 1143 52 female 24.860 0 no southeast 4.433258  
## 1144 39 male 32.340 2 no southeast 3.801957  
## 1145 50 male 32.300 2 no southwest 3.983644  
## 1146 52 male 32.775 3 no northwest 4.052660  
## 1147 60 male 32.800 0 yes southwest 4.720910  
## 1148 20 female 31.920 0 no northwest 3.354410  
## 1149 55 male 21.500 1 no southwest 4.033100  
## 1150 42 male 34.100 0 no southwest 3.776682  
## 1151 18 female 30.305 0 no northeast 3.343160  
## 1152 58 female 36.480 0 no northwest 4.087634  
## 1153 43 female 32.560 3 yes southeast 4.612161  
## 1154 35 female 35.815 1 no northwest 3.750544  
## 1155 48 female 27.930 4 no northwest 4.041991  
## 1156 36 female 22.135 3 no northeast 3.859031  
## 1157 19 male 44.880 0 yes southeast 4.599039  
## 1158 23 female 23.180 2 no northwest 4.159148  
## 1159 20 female 30.590 0 no northeast 3.390886  
## 1160 32 female 41.100 0 no southwest 3.600956  
## 1161 43 female 34.580 1 no northwest 3.888025  
## 1162 34 male 42.130 2 no southeast 3.709625  
## 1163 30 male 38.830 1 no southeast 4.277911  
## 1164 18 female 28.215 0 no northeast 3.342587  
## 1165 41 female 28.310 1 no northwest 3.854522  
## 1166 35 female 26.125 0 no northeast 3.718335  
## 1167 57 male 40.370 0 no southeast 4.040701  
## 1168 29 female 24.600 2 no southwest 3.656048  
## 1169 32 male 35.200 2 no southwest 3.669376  
## 1170 37 female 34.105 1 no northwest 3.786208  
## 1171 18 male 27.360 1 yes northeast 4.234990  
## 1172 43 female 26.700 2 yes southwest 4.351769  
## 1173 56 female 41.910 0 no southeast 4.045073  
## 1174 38 male 29.260 2 no northwest 3.810088  
## 1175 29 male 32.110 2 no northwest 3.646787  
## 1176 22 female 27.100 0 no southwest 3.333318  
## 1177 52 female 24.130 1 yes northwest 4.378174  
## 1178 40 female 27.400 1 no southwest 3.812705  
## 1179 23 female 34.865 0 no northeast 3.462322  
## 1180 31 male 29.810 0 yes southeast 4.286689  
## 1181 42 female 41.325 1 no northeast 3.883705  
## 1182 24 female 29.925 0 no northwest 3.454949  
## 1183 25 female 30.300 0 no southwest 3.420450  
## 1184 48 female 27.360 1 no northeast 3.975311  
## 1185 23 female 28.490 1 yes southeast 4.263121  
## 1186 45 male 23.560 2 no northeast 3.934691  
## 1187 20 male 35.625 3 yes northwest 4.573630  
## 1188 62 female 32.680 0 no northwest 4.141287  
## 1189 43 female 25.270 1 yes northeast 4.337885  
## 1190 23 female 28.000 0 no southwest 4.118155  
## 1191 31 female 32.775 2 no northwest 3.726515  
## 1192 41 female 21.755 1 no northeast 4.137527  
## 1193 58 female 32.395 1 no northeast 4.114583  
## 1194 48 female 36.575 0 no northwest 3.938079  
## 1195 31 female 21.755 0 no northwest 3.616379  
## 1196 19 female 27.930 3 no northwest 4.275051  
## 1197 19 female 30.020 0 yes northwest 4.522543  
## 1198 41 male 33.550 0 no southeast 3.755862  
## 1199 40 male 29.355 1 no northwest 3.805746  
## 1200 31 female 25.800 2 no southwest 3.693261  
## 1201 37 male 24.320 2 no northwest 3.792304  
## 1202 46 male 40.375 2 no northwest 3.941175  
## 1203 22 male 32.110 0 no northwest 3.312880  
## 1204 51 male 32.300 1 no northeast 3.998436  
## 1205 18 female 27.280 3 yes southeast 4.260631  
## 1206 35 male 17.860 1 no northwest 3.708973  
## 1207 59 female 34.800 2 no southwest 4.567151  
## 1208 36 male 33.400 2 yes southwest 4.584506  
## 1209 37 female 25.555 1 yes northeast 4.307429  
## 1210 59 male 37.100 1 no southwest 4.091567  
## 1211 36 male 30.875 1 no northwest 3.730246  
## 1212 39 male 34.100 2 no southeast 4.372231  
## 1213 18 male 21.470 0 no northeast 3.231076  
## 1214 52 female 33.300 2 no southwest 4.033699  
## 1215 27 female 31.255 1 no northwest 3.597264  
## 1216 18 male 39.140 0 no northeast 4.110255  
## 1217 40 male 25.080 0 no southeast 3.733651  
## 1218 29 male 37.290 2 no southeast 3.608324  
## 1219 46 female 34.600 1 yes southwest 4.619736  
## 1220 38 female 30.210 3 no northwest 3.877208  
## 1221 30 female 21.945 1 no northeast 3.673777  
## 1222 40 male 24.970 2 no southeast 3.819117  
## 1223 50 male 25.300 0 no southeast 3.926480  
## 1224 20 female 24.420 0 yes southeast 4.417068  
## 1225 41 male 23.940 1 no northeast 3.836228  
## 1226 33 female 39.820 1 no southeast 3.680848  
## 1227 38 male 16.815 2 no northeast 3.822204  
## 1228 42 male 37.180 2 no southeast 3.855035  
## 1229 56 male 34.430 0 no southeast 4.025069  
## 1230 58 male 30.305 0 no northeast 4.076941  
## 1231 52 male 34.485 3 yes northwest 4.778306  
## 1232 20 female 21.800 0 yes southwest 4.304649  
## 1233 54 female 24.605 3 no northwest 4.096204  
## 1234 58 male 23.300 0 no southwest 4.054824  
## 1235 45 female 27.830 2 no southeast 3.930223  
## 1236 26 male 31.065 0 no northwest 3.431294  
## 1237 63 female 21.660 0 no northeast 4.159863  
## 1238 58 female 28.215 0 no northwest 4.087226  
## 1239 37 male 22.705 3 no northeast 3.844198  
## 1240 25 female 42.130 1 no southeast 3.510335  
## 1241 52 male 41.800 2 yes southeast 4.674584  
## 1242 64 male 36.960 2 yes southeast 4.695286  
## 1243 22 female 21.280 3 no northwest 3.633092  
## 1244 28 female 33.110 0 no southeast 3.501280  
## 1245 18 male 33.330 0 no southeast 3.055356  
## 1246 28 male 24.300 5 no southwest 3.749378  
## 1247 45 female 25.700 3 no southwest 3.959127  
## 1248 33 male 29.400 4 no southwest 3.782413  
## 1249 18 female 39.820 0 no southeast 3.213242  
## 1250 32 male 33.630 1 yes northeast 4.575275  
## 1251 24 male 29.830 0 yes northeast 4.270642  
## 1252 19 male 19.800 0 no southwest 3.093969  
## 1253 20 male 27.300 0 yes southwest 4.210395  
## 1254 40 female 29.300 4 no southwest 4.199449  
## 1255 34 female 27.720 0 no southeast 3.644946  
## 1256 42 female 37.900 0 no southwest 3.811174  
## 1257 51 female 36.385 3 no northwest 4.058302  
## 1258 54 female 27.645 1 no northwest 4.053306  
## 1259 55 male 37.715 3 no northwest 4.478041  
## 1260 52 female 23.180 0 no northeast 4.008505  
## 1261 32 female 20.520 0 no northeast 3.657461  
## 1262 28 male 37.100 1 no southwest 3.515498  
## 1263 41 female 28.050 1 no southeast 3.830601  
## 1264 43 female 29.900 1 no southwest 3.865563  
## 1265 49 female 33.345 2 no northeast 4.015817  
## 1266 64 male 23.760 0 yes southeast 4.430180  
## 1267 55 female 30.500 0 no southwest 4.029565  
## 1268 24 male 31.065 0 yes northeast 4.534712  
## 1269 20 female 33.300 0 no southwest 3.274270  
## 1270 45 male 27.500 3 no southwest 3.935270  
## 1271 26 male 33.915 1 no northwest 3.517530  
## 1272 25 female 34.485 0 no northwest 3.480267  
## 1273 43 male 25.520 5 no southeast 4.160718  
## 1274 35 male 27.610 1 no southeast 3.676424  
## 1275 26 male 27.060 0 yes southeast 4.231555  
## 1276 57 male 23.700 0 no southwest 4.039784  
## 1277 22 female 30.400 0 no northeast 3.438059  
## 1278 32 female 29.735 0 no northwest 3.639192  
## 1279 39 male 29.925 1 yes northeast 4.351449  
## 1280 25 female 26.790 2 no northwest 3.622122  
## 1281 48 female 33.330 0 no southeast 3.918223  
## 1282 47 female 27.645 2 yes northwest 4.389798  
## 1283 18 female 21.660 0 yes northeast 4.154833  
## 1284 18 male 30.030 1 no southeast 3.235618  
## 1285 61 male 36.300 1 yes southwest 4.675814  
## 1286 47 female 24.320 0 no northeast 3.931187  
## 1287 28 female 17.290 0 no northeast 3.572014  
## 1288 36 female 25.900 1 no southwest 3.738182  
## 1289 20 male 39.400 2 yes southwest 4.583704  
## 1290 44 male 34.320 1 no southeast 3.854153  
## 1291 38 female 19.950 2 no northeast 3.853327  
## 1292 19 male 34.900 0 yes southwest 4.541937  
## 1293 21 male 23.210 0 no southeast 3.180511  
## 1294 46 male 25.745 3 no northwest 3.968571  
## 1295 58 male 25.175 0 no northeast 4.076681  
## 1296 20 male 22.000 1 no southwest 3.293314  
## 1297 18 male 26.125 0 no northeast 3.232723  
## 1298 28 female 26.510 2 no southeast 3.637534  
## 1299 33 male 27.455 2 no northwest 3.721107  
## 1300 19 female 25.745 1 no northwest 3.433102  
## 1301 45 male 30.360 0 yes southeast 4.796525  
## 1302 62 male 30.875 3 yes northwest 4.669486  
## 1303 25 female 20.800 1 no southwest 3.506341  
## 1304 43 male 27.800 0 yes southwest 4.577833  
## 1305 42 male 24.605 2 yes northeast 4.327551  
## 1306 24 female 27.720 0 no southeast 3.391750  
## 1307 29 female 21.850 0 yes northeast 4.207239  
## 1308 32 male 28.120 4 yes northwest 4.331882  
## 1309 25 female 30.200 0 yes southwest 4.530208  
## 1310 41 male 32.200 2 no southwest 3.837333  
## 1311 42 male 26.315 1 no northwest 3.841416  
## 1312 33 female 26.695 0 no northwest 3.660050  
## 1313 34 male 42.900 1 no southwest 3.656698  
## 1314 19 female 34.700 2 yes southwest 4.561072  
## 1315 30 female 23.655 3 yes northwest 4.273369  
## 1316 18 male 28.310 1 no northeast 4.052014  
## 1317 19 female 20.600 0 no southwest 3.238467  
## 1318 18 male 53.130 0 no southeast 3.065752  
## 1319 35 male 39.710 4 no northeast 4.289962  
## 1320 39 female 26.315 2 no northwest 3.857435  
## 1321 31 male 31.065 3 no northwest 3.734402  
## 1322 62 male 26.695 0 yes northeast 4.448727  
## 1323 62 male 38.830 0 no southeast 4.113320  
## 1324 42 female 40.370 2 yes southeast 4.642429  
## 1325 31 male 25.935 1 no northwest 3.627355  
## 1326 61 male 33.535 0 no northeast 4.118706  
## 1327 42 female 32.870 0 no northeast 3.848190  
## 1328 51 male 30.030 1 no southeast 3.972106  
## 1329 23 female 24.225 2 no northeast 4.350165  
## 1330 52 male 38.600 2 no southwest 4.013899  
## 1331 57 female 25.740 2 no southeast 4.101375  
## 1332 23 female 33.400 0 no southwest 4.033260  
## 1333 52 female 44.700 3 no southwest 4.057350  
## 1334 50 male 30.970 3 no northwest 4.025328  
## 1335 18 female 31.920 0 no northeast 3.343602  
## 1336 18 female 36.850 0 no southeast 3.212143  
## 1337 21 female 25.800 0 no southwest 3.302752  
## 1338 61 female 29.070 0 yes northwest 4.464510

# Split the data into training and test sets  
set.seed(122) # Set the seed to make the partition reproducible  
training.samples <- df$logCharges %>%  
 createDataPartition(p = 0.8, list = FALSE)  
  
train <- df[training.samples, ]  
test <- df[-training.samples, ]

### Modelling

library(ridge)   
library(e1071)

##   
## Attaching package: 'e1071'

## The following object is masked from 'package:Hmisc':  
##   
## impute

library(ggplot2)  
library(rpart)  
library(rpart.plot)

## Multi Linear Regression

# Fitting Multiple Linear Regression to the Training set  
formula <- as.formula("logCharges ~ smoker + bmi + age + children + sex + region")  
model <- lm(formula, data = train)  
  
summary(model)

##   
## Call:  
## lm(formula = formula, data = train)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.40628 -0.09013 -0.02321 0.03314 0.93626   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.0308795 0.0342260 88.555 < 0.0000000000000002 \*\*\*  
## smokeryes 0.6760329 0.0144515 46.779 < 0.0000000000000002 \*\*\*  
## bmi 0.0058070 0.0009931 5.848 0.00000000663898698 \*\*\*  
## age 0.0153611 0.0004142 37.090 < 0.0000000000000002 \*\*\*  
## children1 0.0538452 0.0146927 3.665 0.000260 \*\*\*  
## children2 0.1286999 0.0161328 7.978 0.00000000000000385 \*\*\*  
## children3 0.1086741 0.0189414 5.737 0.00000001254227630 \*\*\*  
## children4 0.2109837 0.0411729 5.124 0.00000035470555674 \*\*\*  
## children5 0.1835554 0.0552900 3.320 0.000931 \*\*\*  
## sexmale -0.0304837 0.0115905 -2.630 0.008661 \*\*   
## regionnorthwest -0.0305449 0.0164321 -1.859 0.063325 .   
## regionsoutheast -0.0599089 0.0168307 -3.559 0.000388 \*\*\*  
## regionsouthwest -0.0562769 0.0165515 -3.400 0.000699 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1884 on 1059 degrees of freedom  
## Multiple R-squared: 0.7789, Adjusted R-squared: 0.7764   
## F-statistic: 310.9 on 12 and 1059 DF, p-value: < 0.00000000000000022

### Evaluating the model

# Make predictions on the training dataset  
predictions <- model %>% predict(train)  
# Calculating the residuals  
residuals <- train$logCharges - predictions  
# Calculating Root Mean Squared Error  
rmse <- sqrt(mean(residuals^2))  
  
rmse %>%  
 round(digits=3)

## [1] 0.187

predictions <- model %>% predict(test)  
residuals <- test$logCharges - predictions  
rmse <- sqrt(mean(residuals^2))  
  
rmse %>%  
 round(digits=3)

## [1] 0.208

# Calculating RMSE for training data with backtransformed data  
predictions <- model %>% predict(train)  
residuals <- 10^train$logCharges - 10^predictions # backtransform measured and predicted values  
rmse <- sqrt(mean(residuals^2))  
round(rmse)

## [1] 8334

# Calculating RMSE for testing data with backtransformed data  
predictions <- model %>% predict(test)  
residuals <- 10^test$logCharges - 10^predictions # backtransform measured and predicted values  
rmse <- sqrt(mean(residuals^2))  
round(rmse)

## [1] 9000

### Linear Ridge

linRidgeMod = linearRidge(logCharges ~ ., data = train)  
predicted = predict(linRidgeMod, test)   
compare1 = cbind (actual=test$logCharges, predicted)  
  
mean (apply(compare1, 1, min)/apply(compare1, 1, max))

## [1] 0.9769878

RMSE = sqrt(mean((test$logCharges-predicted)^2))  
RMSE

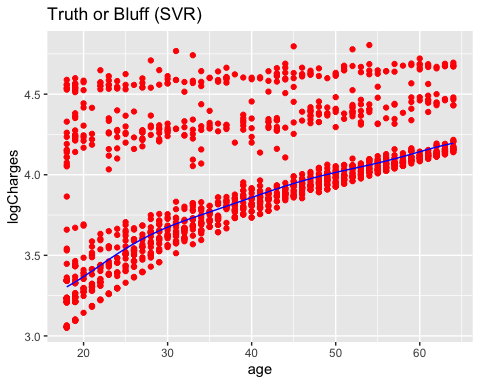
## [1] 0.1335739

# summary  
summary(linRidgeMod)

##   
## Call:  
## linearRidge(formula = logCharges ~ ., data = train)  
##   
##   
## Coefficients:  
## Estimate Scaled estimate Std. Error (scaled)  
## (Intercept) 3.30297677 NA NA  
## age 0.00929640 4.27152666 0.15559432  
## sexmale -0.02915808 -0.47730851 0.13198861  
## bmi -0.00145019 -0.29048685 0.14480523  
## children1 0.05049776 0.70773365 0.14311672  
## children2 0.08768795 1.10974346 0.14271249  
## children3 0.08344239 0.88589104 0.14001409  
## children4 0.15305531 0.71048842 0.13317015  
## children5 0.15881510 0.54706375 0.13250258  
## smokeryes 0.14952019 1.97042375 0.24340272  
## regionnorthwest -0.01979536 -0.28064772 0.16164160  
## regionsoutheast -0.04597714 -0.66128119 0.16792405  
## regionsouthwest -0.04039165 -0.56977292 0.16204902  
## charges 0.00002236 8.80624906 0.26059639  
## t value (scaled) Pr(>|t|)   
## (Intercept) NA NA   
## age 27.453 < 0.0000000000000002 \*\*\*  
## sexmale 3.616 0.000299 \*\*\*  
## bmi 2.006 0.044851 \*   
## children1 4.945 0.000000760853017479 \*\*\*  
## children2 7.776 0.000000000000007550 \*\*\*  
## children3 6.327 0.000000000249720467 \*\*\*  
## children4 5.335 0.000000095443154713 \*\*\*  
## children5 4.129 0.000036481571797919 \*\*\*  
## smokeryes 8.095 0.000000000000000666 \*\*\*  
## regionnorthwest 1.736 0.082522 .   
## regionsoutheast 3.938 0.000082171165242118 \*\*\*  
## regionsouthwest 3.516 0.000438 \*\*\*  
## charges 33.793 < 0.0000000000000002 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Ridge parameter: 0.002308026, chosen automatically, computed using 10 PCs  
##   
## Degrees of freedom: model 12.95 , variance 12.9 , residual 13

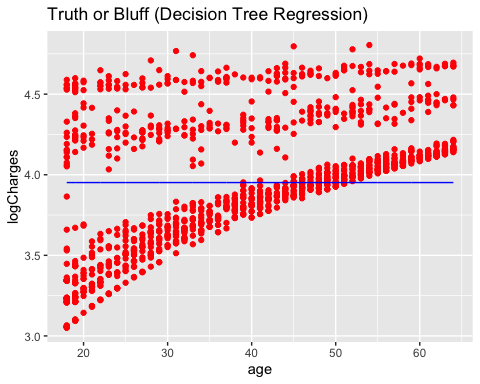
### SVR

regressor = svm(formula = logCharges ~ age, data = df, type = 'eps-regression', kernel = 'radial')  
y\_pred = predict(regressor, data.frame(age = 37))  
  
# Visualizing the SVR results  
ggplot() + geom\_point(aes(x = df$age, y = df$logCharges), colour = 'red') +  
 geom\_line(aes(x = df$age, y = predict(regressor, newdata = df)), colour = 'blue') +  
 ggtitle('Truth or Bluff (SVR)') + xlab('age') + ylab('logCharges')



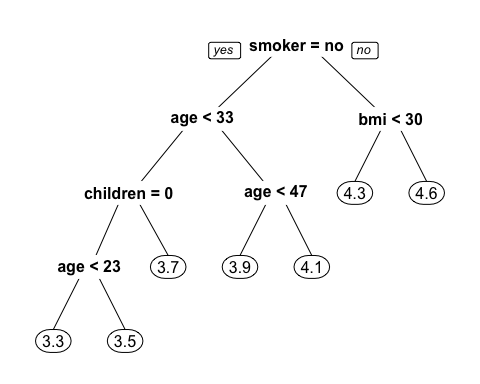
### Decision Tree Regression

regressor = rpart(formula = logCharges ~ age, data = df, control = rpart.control(minsplit = 1))  
y\_pred = predict(regressor, data.frame(age = 37))  
  
ggplot() + geom\_point(aes(x = df$age, y = df$logCharges), colour = 'red') +  
 geom\_line(aes(x = df$age, y = predict(regressor, newdata = df)), colour = 'blue') +  
 ggtitle('Truth or Bluff (Decision Tree Regression)') + xlab('age') + ylab('logCharges')



### Tree plot

tree = rpart(logCharges ~ smoker + bmi + age + children + sex + region, data=df)  
prp(tree)



## Random Forest

myControl = trainControl(method = "cv", number = 5, verboseIter = FALSE)  
model\_rf = train(logCharges ~ smoker + bmi + age + children + sex + region, data = train,  
 tuneLength = 1, method = "ranger", importance = 'impurity',  
 trControl = myControl)  
model\_rf

## Random Forest   
##   
## 1072 samples  
## 6 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (5 fold)   
## Summary of sample sizes: 856, 857, 858, 858, 859   
## Resampling results across tuning parameters:  
##   
## splitrule RMSE Rsquared MAE   
## variance 0.1727875 0.8297888 0.1130003  
## extratrees 0.1796149 0.8260375 0.1245521  
##   
## Tuning parameter 'mtry' was held constant at a value of 3  
## Tuning  
## parameter 'min.node.size' was held constant at a value of 5  
## RMSE was used to select the optimal model using the smallest value.  
## The final values used for the model were mtry = 3, splitrule = variance  
## and min.node.size = 5.