Statistics, Modeling, and Machine Learning

Reproducibility

Steps of data analysis: -Define the question

-Define the ideal data set

-Determine what data you can access

-Obtain the data

-Clean the Data

-Exploratory data analysis

-Statistical prediction/Modeling

-Interpret results

-Challenge results

-Synthesize/Write up results

-Create reproducible code

Probability

P(A u B) = P(A) + P(B) - The probability of the union between scenario A and B is simply the sum of each. Random variable = is the numerical outcome of an experiment (discrete or continuous) - e.g. the flip of a coin (discrete random variable) - e.g. the roll of a dice (also discrete random variable because it's limited to 1-6) - e.g. BMI of a person is a continuous random variable

```
  For random variables, the probabilities must add  up to 1, and components are larger or equal to 0.
    Probability mass function - for discrete random variables
    Probability density function - for continuous random variables
        in R, probability of some instance occurring within a continuous distribution can be calculated
        Conditional probability- think lightning strike on a storming vs. sunny day.
            P(A|B) = P(A u B)/P(B) is the conditional probability. Written as the probability of A given
```

mean = center of a distribution variance and standard deviations = how spread out the distribution is

```
 ## Beta distribution
pbeta(0.75, 2, 1) ##here 0.75 is 75% probability of some density.
```

```
## [1] 0.5625
```

```
pbeta(c(.4, .5, .6), 2, 1) ##40-60% probability example of the same density.
```

```
## [1] 0.16 0.25 0.36
```

```
pnorm(70, mean = 80, sd = 10) ## probability to have value 70 given a set mean and standard dev. Data i
```

```
## [1] 0.1586553
```

```
qnorm(0.95, mean=1100, sd = 75) ## will return a value of a normal distribution given a specificed perc
```

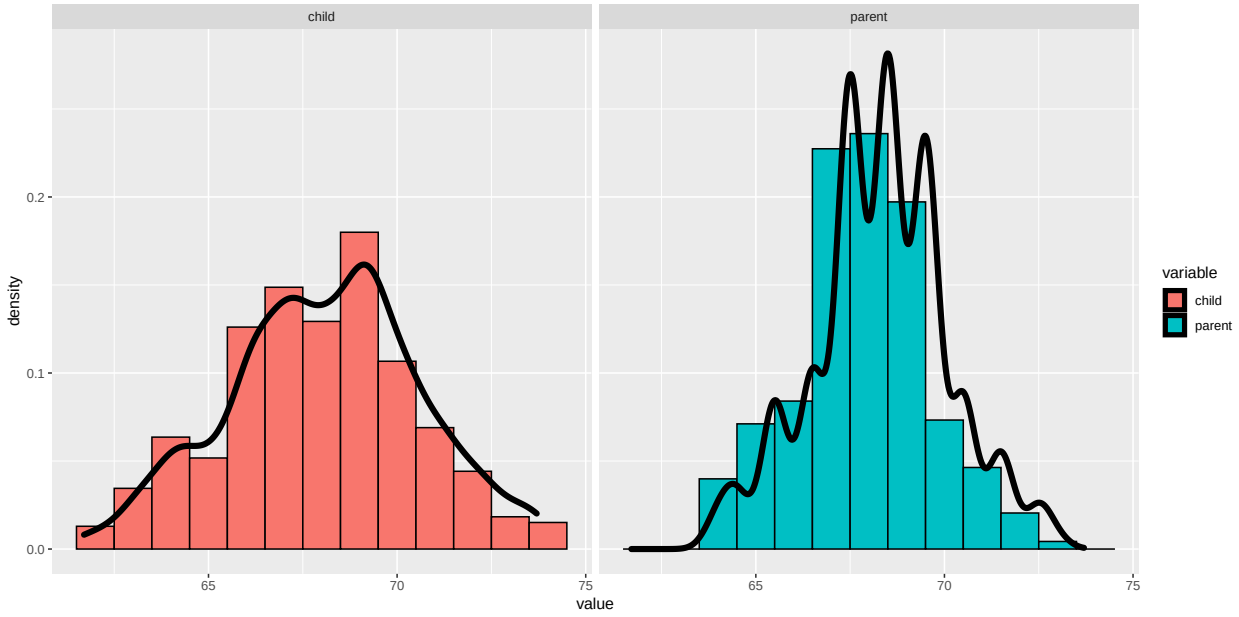
```
## [1] 1223.364
```

```
pnorm(16, mean = 15, sd = 1) - pnorm(14, mean = 15, sd = 1) ##probaility that a an output is between tw
```

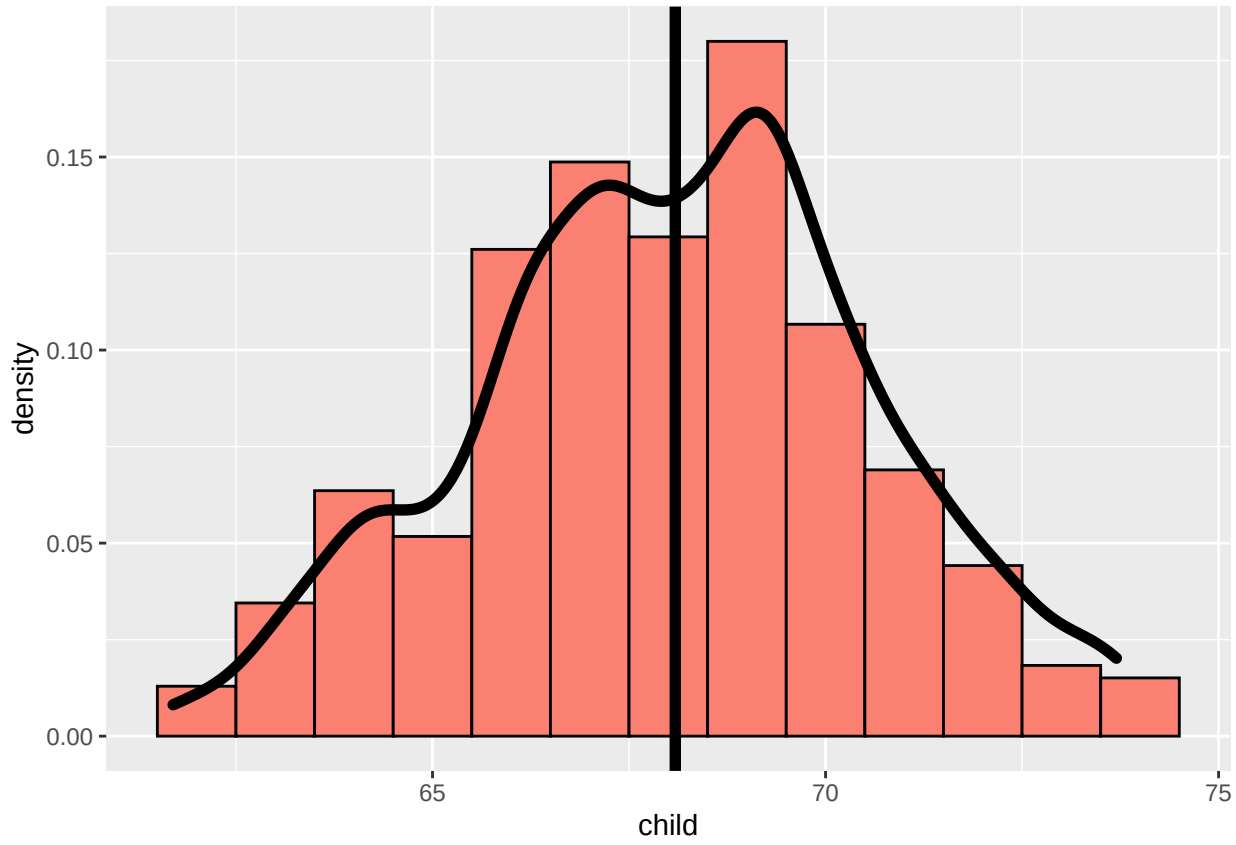
```
## [1] 0.6826895
```

```
ppois(10, lambda = 15) ##poisson distributed data. Probability of seeing 10 or less of an instance, whe
```

```
## [1] 0.1184644
```




```
g <- ggplot(galton, aes(x = child))
g <- g + geom_histogram(fill = "salmon",
  binwidth=1, aes(y = ..density..), colour = "black")
g <- g + geom_density(size = 2)
g <- g + geom_vline(xintercept = mean(galton$child), size = 2)
g
```



Variance = Total sum of squares

```
lambda <- 0.2
nsim <- 1:1000 # Number of Simulations/rows
n <- 40

Ematrix <- data.frame(x = sapply(nsim, function(x) {mean(rexp(n, lambda))})) ##creating a matrix of val
head(Ematrix)
```

```
##          x
## 1 3.349103
## 2 5.582835
## 3 4.038226
## 4 5.064299
## 5 5.987589
## 6 5.290295
```

```
Smean <- apply(Ematrix, 2, mean) ## calculating the simulated mean from above matrix.
Smean
```

```
##        x
## 4.971585
```

```r
Tmean <- 1/lambda  ##calculating the theoretical mean given lambda
Tmean
```

```
## [1] 5
```

```r
SSD <- sd(Ematrix$x)  ##The simulated standard deviation
SSD
```

```
## [1] 0.7920842
```

```r
SVar <- var(Ematrix$x)  ##The simulated variance from the above matrix.
SVar
```

```
## [1] 0.6273973
```

```r
TSD <- (1/lambda)/sqrt(n)  ##The theoretical standard deviation.
TSD
```

```
## [1] 0.7905694
```

```r
TVar <- TSD^2   ##The theoretical calculation for the Variance
TVar
```

```
## [1] 0.625
```

t-test

(X - X*)/(s/sqrt(n)) in other words, sample mean minus the hypothesized/test mean or value divided by standard error of the mean divided by square root of sample size, n. This follows a t-distribution with n-1 degrees of freedom. You can calculate the T-distribution using qt(.95, 15), where .95 is the percentile or an alpha of 0.05, and 15 is n-1 in this example.

```r
library(UsingR); data(father.son)
t.test(father.son$sheight - father.son$fheight)  ##testing whether there are significant differences be
```

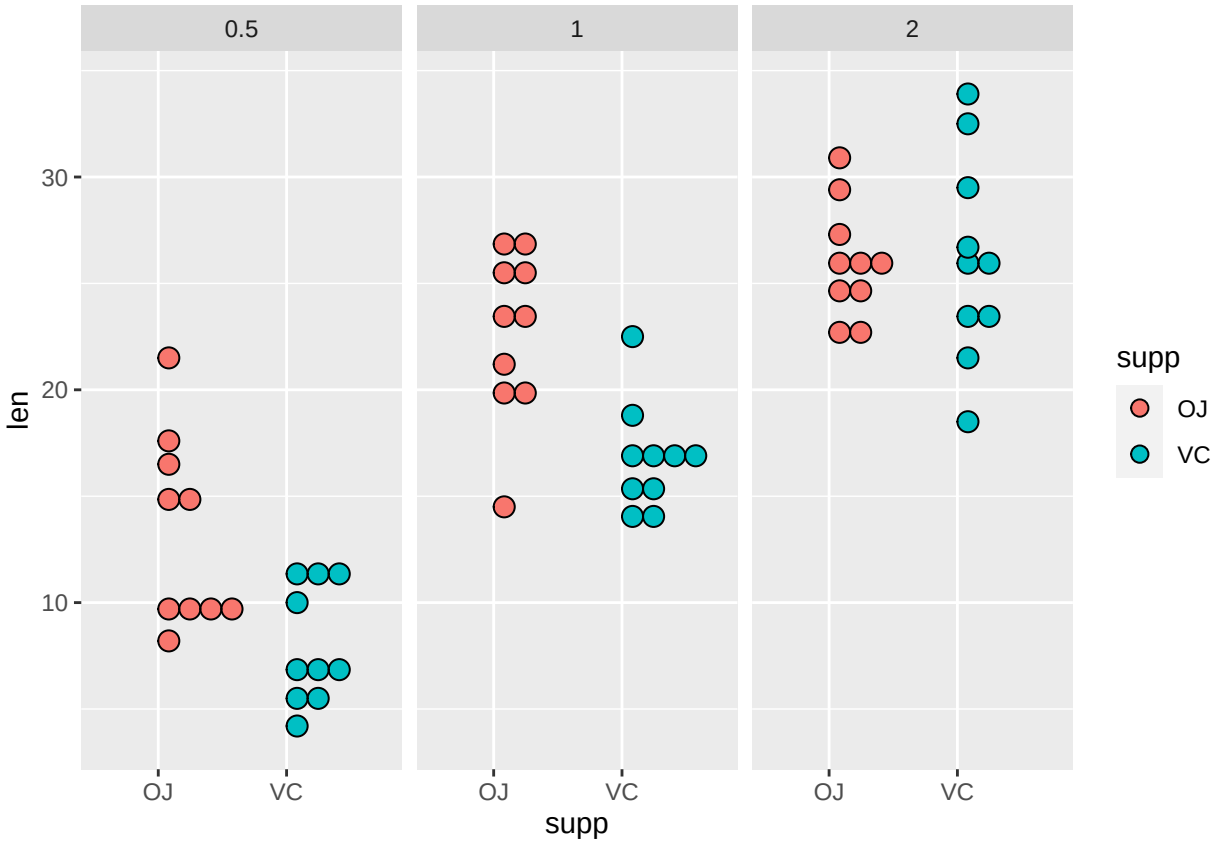
```
##
##  One Sample t-test
##
## data:  father.son$sheight - father.son$fheight
## t = 11.789, df = 1077, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.8310296 1.1629160
## sample estimates:
## mean of x
## 0.9969728
```

```
library(datasets)
data(ToothGrowth)
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
library(ggplot2)

plot0 <- ggplot(ToothGrowth, aes(supp, len, fill = supp))
plot1 <- plot0 + geom_dotplot(binaxis ="y") +
      facet_grid(.~ dose)
      plot1
```

```
## Bin width defaults to 1/30 of the range of the data. Pick better value with `binwidth`.
```




```
t.test(len ~ supp, data = ToothGrowth) ##testing sig difference between supp types and length
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means between group OJ and group VC is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

```
test1 <- subset(ToothGrowth, dose %in% c(.5,1)) ##subsetting doses to make pairwise comparisons
t.test(len ~ dose, data = test1) ##testing sig differences between discrete doses and length
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means between group 0.5 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

```
test2 <- subset(ToothGrowth, dose %in% c(1,2)) ##subsetting doses to make pairwise comparisons
t.test(len ~ dose, data = test2)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

t-test Confidence Interval

```
## Prompt for below: a sample of 9 men yielded a sample average brain volume of 1,100cc and a standard
mn = 1100
s = 30
n = 9
round(1100 + c(-1,1)*qt(.975, df = 8)*s/sqrt(n))
```

```
## [1] 1077 1123
```

```r
## diet pill is given to 9 subjects over six weeks. The average difference in weight (follow up - basel
n = 9
mn_dif = 2
t = .95
(y_d <- round(mn_dif*sqrt(n) / qt(.975, df = 8), 2))
```

```
## [1] 2.6
```

Two sample t-test

```r
##Running a two-sample t-test
data(ChickWeight); library(reshape2)
wideCW <- dcast(ChickWeight, Diet + Chick ~ Time, value.var = "weight") ##dcasting the long-form of the
names(wideCW)[-(1:2)] <- paste( "time", names(wideCW)[-(1:2)], sep= "") ##renames columns after the fir
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:Hmisc':
##
##     src, summarize

## The following object is masked from 'package:MASS':
##
##     select

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
wideCW <- mutate(wideCW, gain = time21-time0)  ##adds a column titled gain, where it subtracts or takes

wideCW_14 <- subset(wideCW, Diet %in% c(1,4))
t.test(gain~Diet, paired = FALSE, var.equal = TRUE, data = wideCW_14)
```

```
##
##  Two Sample t-test
##
## data:  gain by Diet
## t = -2.7252, df = 23, p-value = 0.01207
## alternative hypothesis: true difference in means between group 1 and group 4 is not equal to 0
## 95 percent confidence interval:
##  -108.14679  -14.81154
## sample estimates:
## mean in group 1 mean in group 4
##        136.1875        197.6667
```

P-Values

```
pt(2.5, 15, lower.tail = FALSE)  ##the probability of getting a T-statistic of 2.5 (first element of th
```

```
## [1] 0.0122529
```

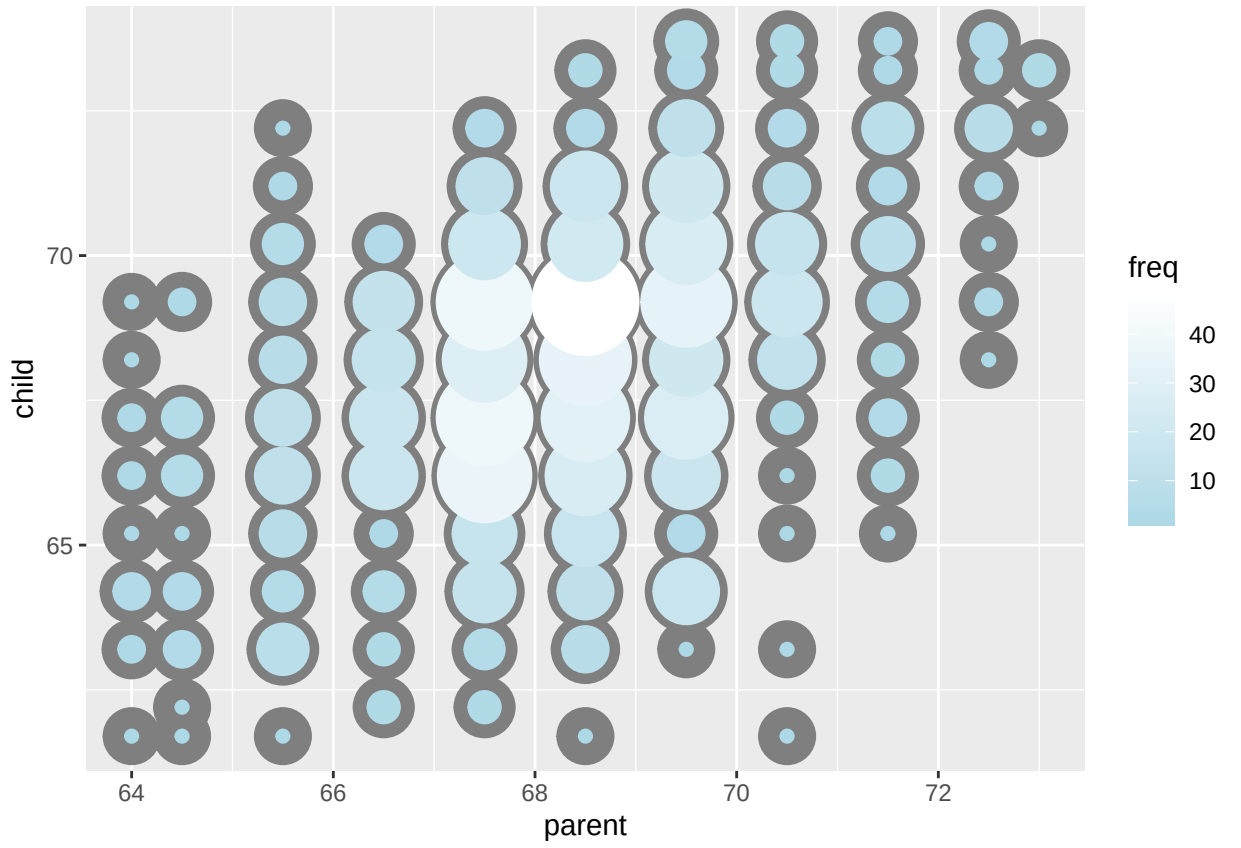
Least Squares (Regression)

Sum[i=1 to n] of [Yi - (Bo +B1Xi)]^2 Least squares best fit is the sum of i to the n of Yi which is the i'th sample data y-coordinate minus the intercept plus the slope times the i'th sample data x-coordinate, squarred. Bo = intercept of y-axis B1 = slope Yi = y of sample i Xi = x of sample i Think of it as subtracting the distance between sample's y-coordinate from the best fit, which is y - (ax + b) which is a linear regression of the best fit, or Yi - Yfit. Then square it.

```
library(UsingR)
library(ggplot2)

freqData <- as.data.frame(table(galton$child, galton$parent))
names(freqData) <- c("child", "parent", "freq")
freqData$child <- as.numeric(as.character(freqData$child))
freqData$parent <- as.numeric(as.character(freqData$parent))
g <- ggplot(filter(freqData, freq > 0), aes(x = parent, y = child))
g <- g  + scale_size(range = c(2, 20), guide = "none" )
g <- g + geom_point(colour="grey50", aes(size = freq+5, show_guide = FALSE))
```

```
## Warning: Ignoring unknown aesthetics: show_guide
```

```
g <- g + geom_point(aes(colour=freq, size = freq-5))
g <- g + scale_colour_gradient(low = "lightblue", high="white")
g
```




```
###_____________________Linear Least Squares calculated
y <- galton$child
x <- galton$parent
beta1 <- cor(y, x) *  sd(y) / sd(x)  ##The slope, which is the corelation between y and x times the st.
beta0 <- mean(y) - beta1 * mean(x) ##calculates the y-intercept of the best fit.
rbind(c(beta0, beta1), coef(lm(y ~ x))) ##lm stands for linear model. coef takes the output and gives u
```

```
##      (Intercept)         x
## [1,]    23.94153 0.6462906
## [2,]    23.94153 0.6462906
```

```
##This example gives you the same output, because one is the manual calculation, while the other is don

beta1 <- cor(y, x) *  sd(x) / sd(y)  ##if we swapped the predictor
beta0 <- mean(x) - beta1 * mean(y)
rbind(c(beta0, beta1), coef(lm(x ~ y)))  ##shows us the slope is the same
```

```
##      (Intercept)         y
## [1,]    46.13535 0.3256475
## [2,]    46.13535 0.3256475
```

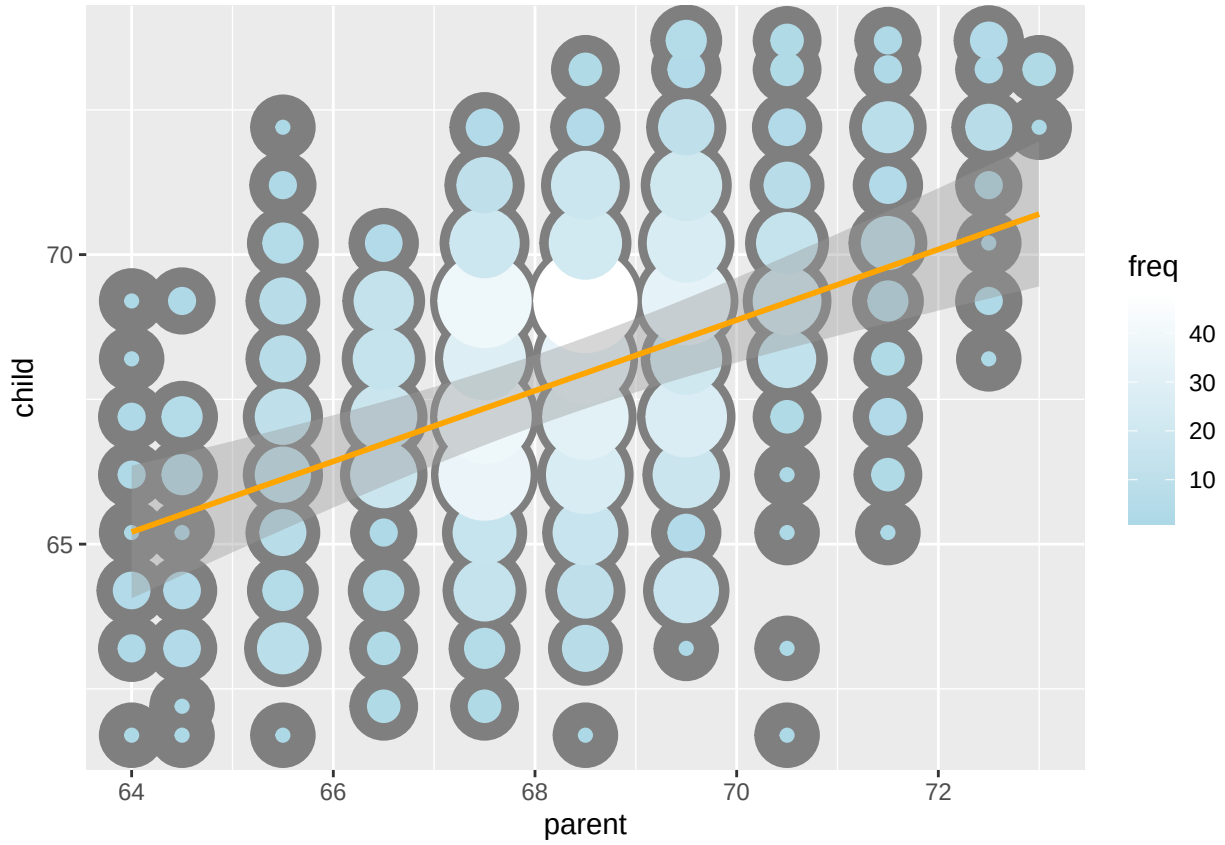
```
yc <- y - mean(y)
xc <- x - mean(x)
beta1 <- sum(yc * xc) / sum(xc ^ 2)
c(beta1, coef(lm(y ~ x))[2]) ##returns the same slope
```

```
##                   x
## 0.6462906 0.6462906
```

```
###simplified
g <- ggplot(filter(freqData, freq > 0), aes(x = parent, y = child))
g <- g  + scale_size(range = c(2, 20), guide = "none" )
g <- g + geom_point(colour="grey50", aes(size = freq+5, show_guide = FALSE))
```

```
## Warning: Ignoring unknown aesthetics: show_guide
```

```
g <- g + geom_point(aes(colour=freq, size = freq-10))
g <- g + scale_colour_gradient(low = "lightblue", high="white")
g <- g + geom_smooth(method="lm", col ="orange", formula=y~x) ##for fitting the linear best fit to a pl
g
```

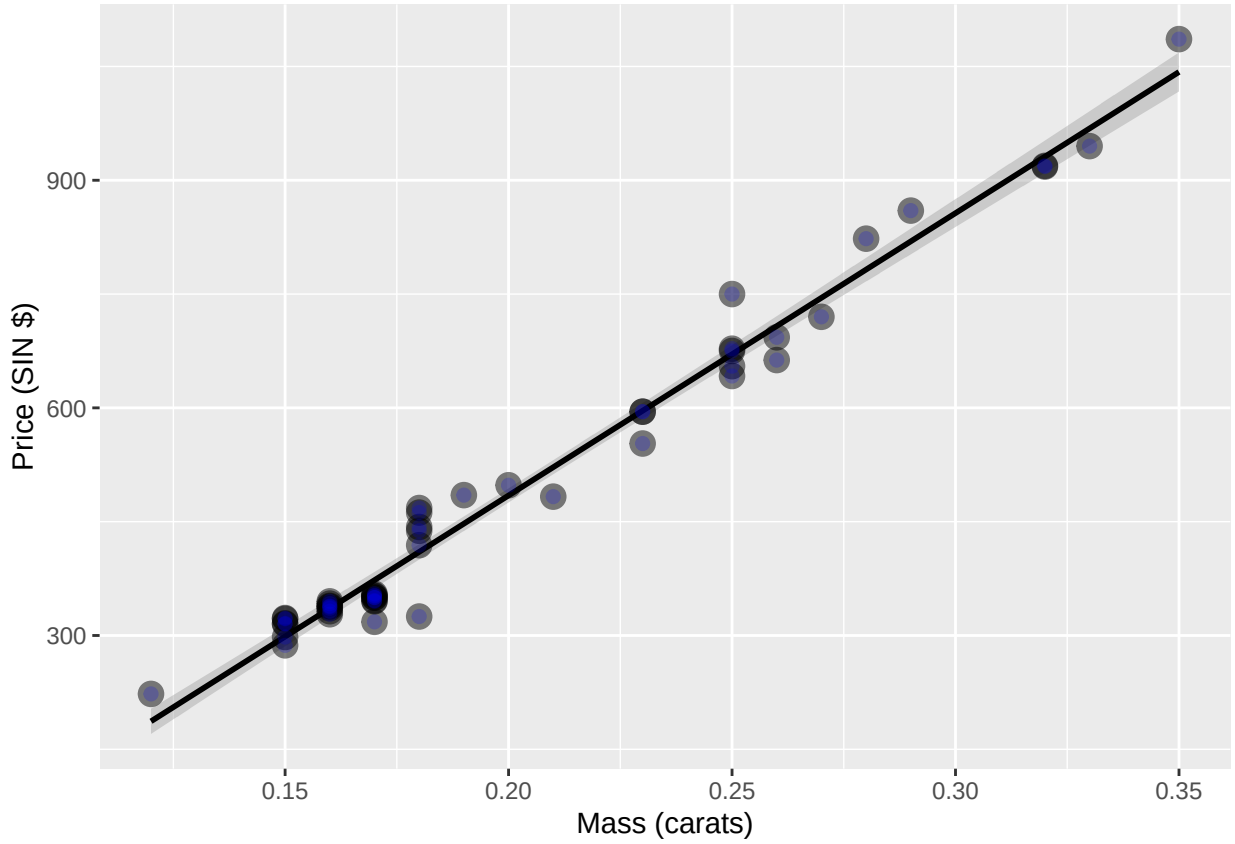


Linear Regression for Prediction

```
library(UsingR)
data(diamond)
library(ggplot2)
g = ggplot(diamond, aes(x = carat, y = price))
g = g + xlab("Mass (carats)")
```

```
g = g + ylab("Price (SIN $)")
g = g + geom_smooth(method = "lm", colour = "black")
g = g + geom_point(size = 4, colour = "black", alpha=0.5)
g = g + geom_point(size = 2, colour = "blue", alpha=0.2)
g
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
fit <- lm(price~ carat, data = diamond)  ## prints coefficient Bo and B1, slope and intercept
coef(fit)
```

```
## (Intercept)       carat
##   -259.6259   3721.0249
```

```
summary(fit) ##gives you the complete statistic
```

```
##
## Call:
## lm(formula = price ~ carat, data = diamond)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -85.159 -21.448  -0.869  18.972  79.370
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -259.63      17.32  -14.99   <2e-16 ***
## carat        3721.02      81.79   45.50   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 31.84 on 46 degrees of freedom
## Multiple R-squared:  0.9783, Adjusted R-squared:  0.9778
## F-statistic:  2070 on 1 and 46 DF,  p-value: < 2.2e-16
```

```
fit2 <- lm(price ~ I(carat - mean(carat)), data=diamond) ##how you mean center your predictor, carat. I
coef(fit2)
```

```
##            (Intercept) I(carat - mean(carat))
##              500.0833             3721.0249
```

```
fit3 <- lm(price ~ I(carat * 10), data=diamond) ##shortcut where fit3 produces the price change for 1/1
coef(fit3)
```

```
##   (Intercept) I(carat * 10)
##     -259.6259      372.1025
```

```
newx <- c(0.16, 0.27, 0.34)  ## an example of a list of carat sizes for diamonds, and we want to predic
predict(fit, newdata = data.frame(carat = newx))  ##taking original fit data, and calling to the newx c
```

```
##         1         2         3
##  335.7381  745.0508 1005.5225
```

Residuals

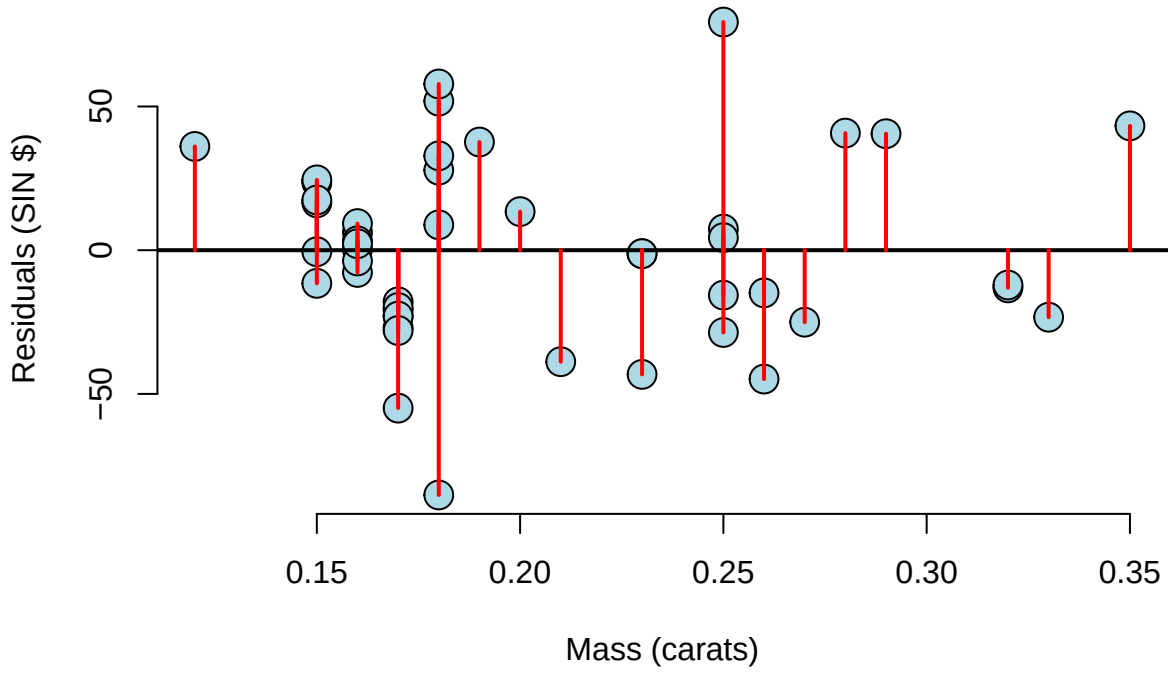
```
data(diamond)
y <- diamond$price; x <- diamond$carat; n <- length(y)
fit <- lm(y ~ x)
e <- resid(fit)  ##easiest way to calculate residuals
yhat <- predict(fit)
max(abs(e -(y - yhat)))
```

```
## [1] 9.485746e-13
```

```
max(abs(e - (y - coef(fit)[1] - coef(fit)[2] * x)))  ##will return the same as resid()
```

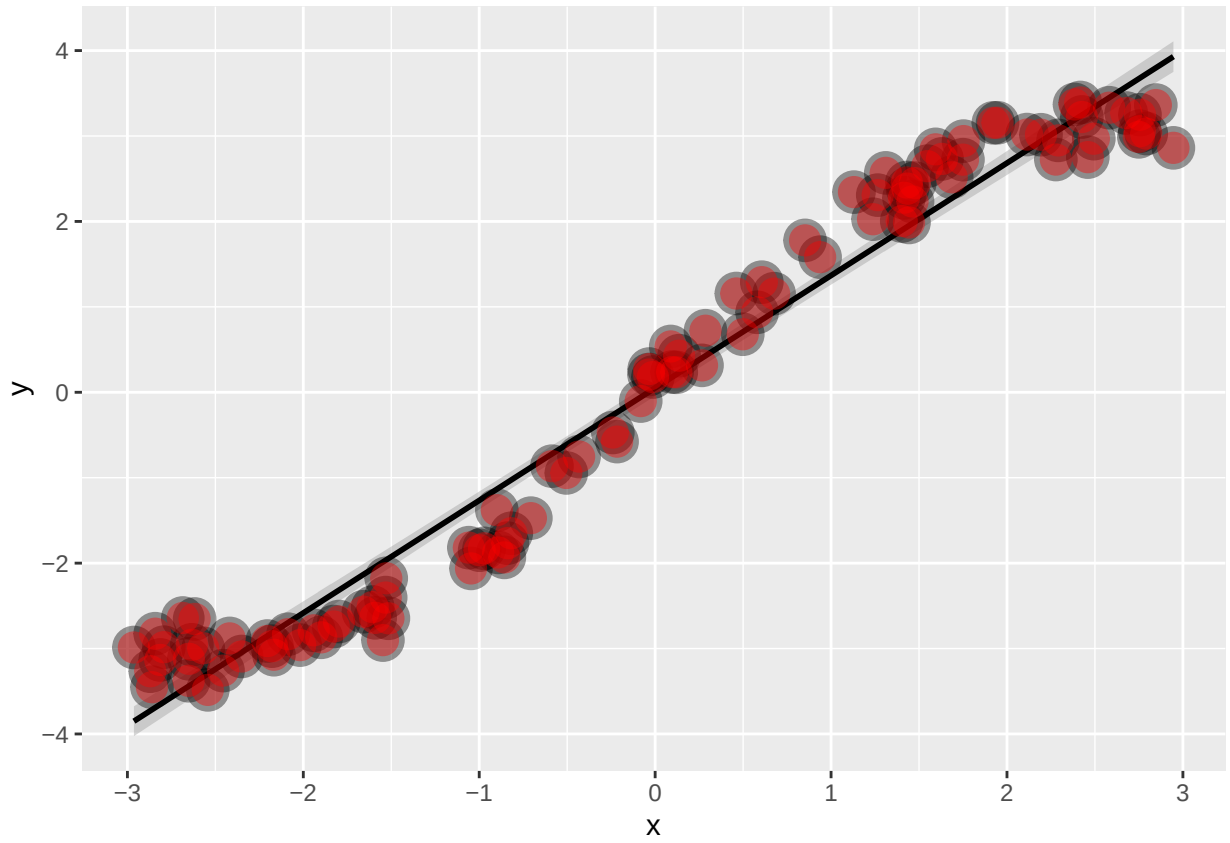
```
## [1] 9.485746e-13
```

```
plot(x, e,
     xlab = "Mass (carats)",
     ylab = "Residuals (SIN $)",
     bg = "lightblue",
     col = "black", cex = 2, pch = 21,frame = FALSE)
abline(h = 0, lwd = 2)
for (i in 1 : n)
  lines(c(x[i], x[i]), c(e[i], 0), col = "red" , lwd = 2)
```

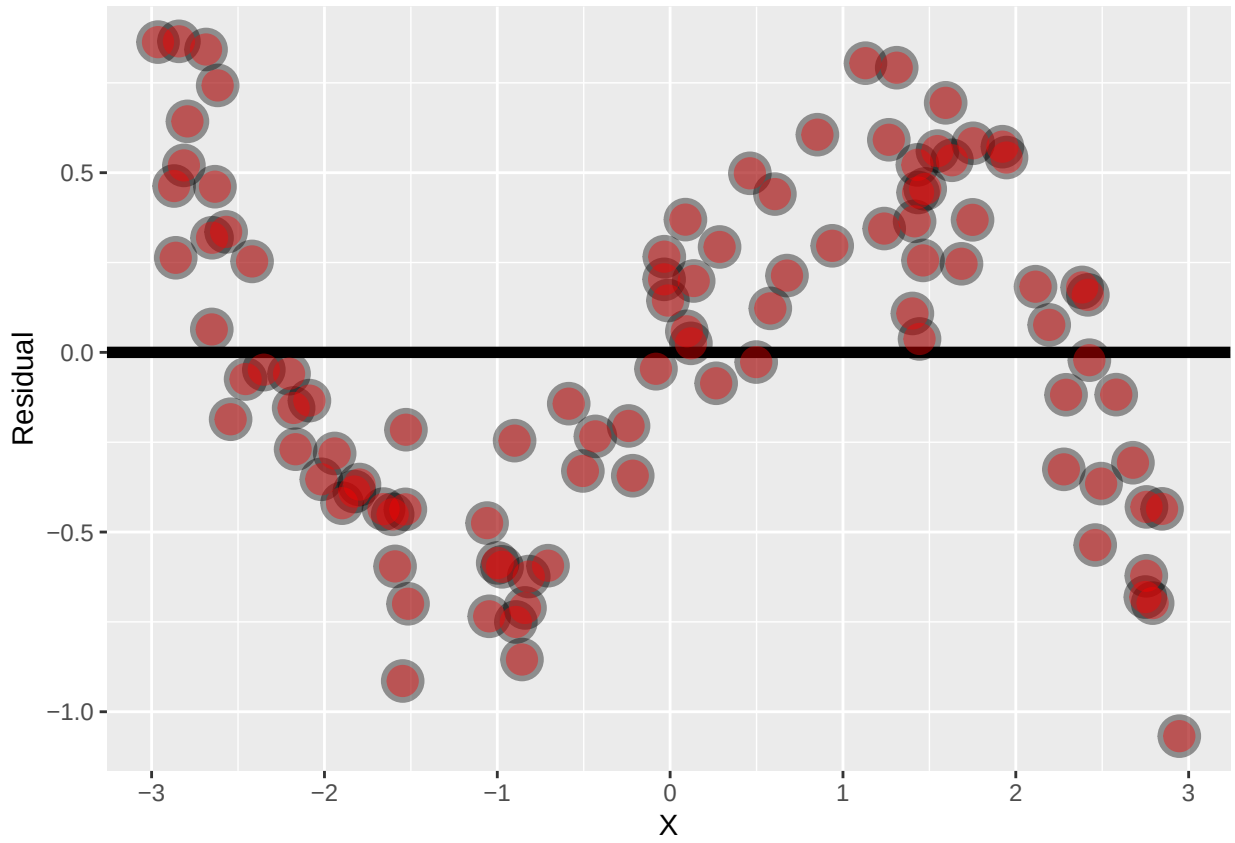



```
##_____________Example of non-linear data and residuals
x = runif(100, -3, 3); y = x + sin(x) + rnorm(100, sd = .2);
library(ggplot2)
library(RColorBrewer)
g = ggplot(data.frame(x = x, y = y), aes(x = x, y = y))
g = g + geom_smooth(method = "lm", colour = "black")
g = g + geom_point(size = 7, colour = "black", alpha = 0.4)
g = g + geom_point(size = 5, colour = "red", alpha = 0.4)
g
```

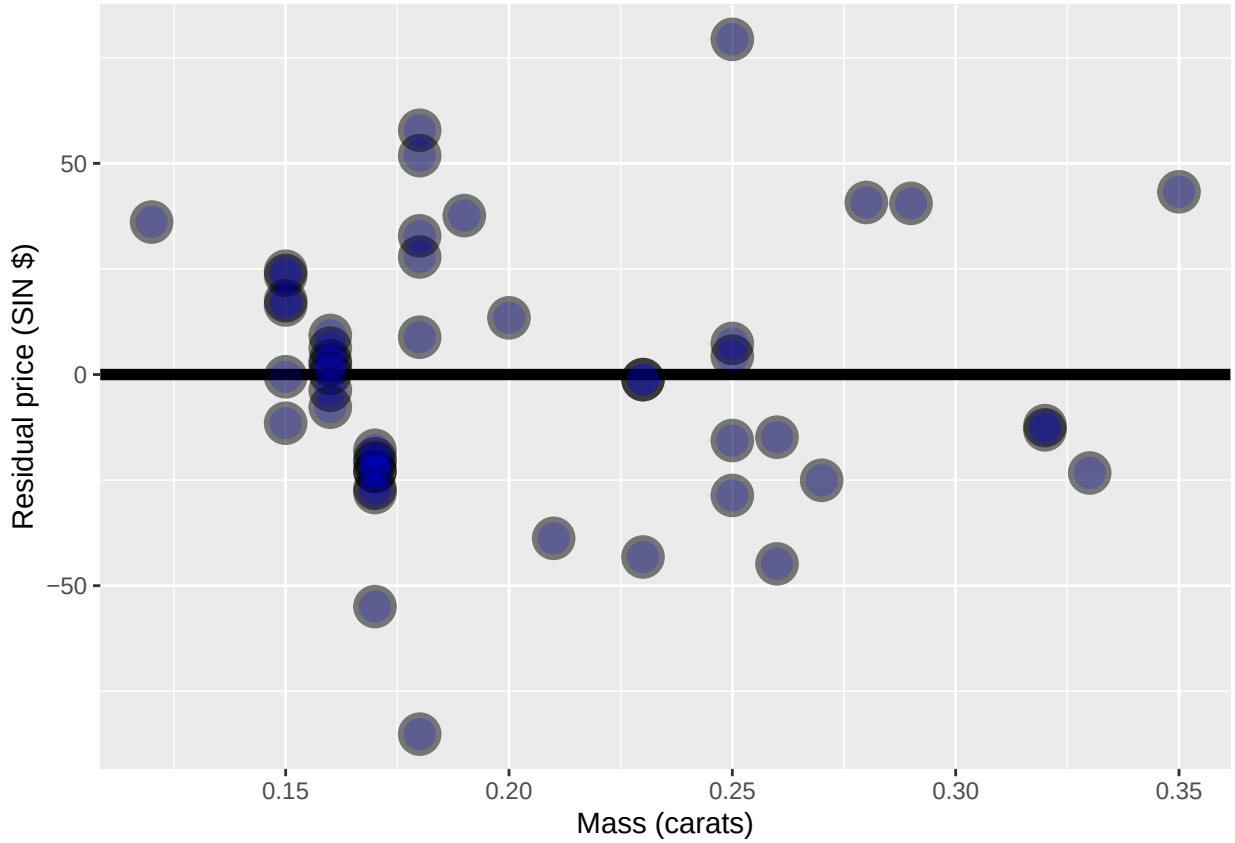
```
## `geom_smooth()` using formula 'y ~ x'
```




```
##_______________
g = ggplot(data.frame(x = x, y = resid(lm(y ~ x))), ##How to plot residuals
           aes(x = x, y = y))
g = g + geom_hline(yintercept = 0, size = 2);
g = g + geom_point(size = 7, colour = "black", alpha = 0.4)
g = g + geom_point(size = 5, colour = "red", alpha = 0.4)
g = g + xlab("X") + ylab("Residual")
g
```




```
##____________ Calculating the residual and plotting for diamond data
diamond$e <- resid(lm(price ~ carat, data = diamond))  ##adding a new column into data.frame for residu
g = ggplot(diamond, aes(x = carat, y = e))
g = g + xlab("Mass (carats)")
g = g + ylab("Residual price (SIN $)")
g = g + geom_hline(yintercept = 0, size = 2)
g = g + geom_point(size = 7, colour = "black", alpha=0.5)
g = g + geom_point(size = 5, colour = "blue", alpha=0.2)
g
```



Predictions

```
fit <- lm(mpg~wt, data= mtcars) ## creating a linear model of mtcars dataset, where weight is the predi
predict(fit, wt=3000) ##using the model above, and predicting the mpg for all cars weighing 3000lbs.
```

```
##          Mazda RX4      Mazda RX4 Wag         Datsun 710     Hornet 4 Drive
##          23.282611          21.919770          24.885952          20.102650
##  Hornet Sportabout            Valiant         Duster 360           Merc 240D
##          18.900144          18.793255          18.205363          20.236262
##           Merc 230           Merc 280          Merc 280C          Merc 450SE
##          20.450041          18.900144          18.900144          15.533127
##         Merc 450SL         Merc 450SLC  Cadillac Fleetwood Lincoln Continental
##          17.350247          17.083024           9.226650           8.296712
##  Chrysler Imperial           Fiat 128        Honda Civic     Toyota Corolla
##           8.718926          25.527289          28.653805          27.478021
##       Toyota Corona    Dodge Challenger         AMC Javelin          Camaro Z28
##          24.111004          18.472586          18.926866          16.762355
##    Pontiac Firebird           Fiat X1-9       Porsche 914-2        Lotus Europa
##          16.735633          26.943574          25.847957          29.198941
##      Ford Pantera L        Ferrari Dino       Maserati Bora          Volvo 142E
##          20.343151          22.480940          18.205363          22.427495
```

Multivariate Regressions

Shapiro Wilks test for Normal distribution

```
data <- rnorm(100)

#perform Shapiro-Wilk test for normality
shapiro.test(data)  ## where we reject normality when p < 0.05, otherwise we accept that the test data
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data
## W = 0.9854, p-value = 0.3395
```

Machine Learning

question - input data - features - algorithm - parameters - evaluation

Prediction has accuracy tradeoffs

In-sample error: the error rate you get on the same data set you used to build your predictor. Sometimes called re-substitution error. Out of sample error: the error rate you get on a new data set, sometimes called generalization error. Out of sample error is what we should care about In sample error is always less than out of sample error The reason is over-fitting. True positive vs. false positive, true negative vs. false negative Sensitivity - probability that the test positively predicted, and that the prediction was right Specificity- probability that the test was negative, and so too was the outcome Receiver Operating Characteristic (ROC) - Area under the curve is the measure of goodness of fit (model/prediction to data/outcome). AUC of 0.5 is random guessing. AUC > 0.8 is a good model. Cross-validation. 1.) use the training set 2.) split the training set into a test set and a smaller training set 3.) Build a model on the training set 4.) Evaluate the model on the test set 5.) repeat and average the estimated errors k-fold cross-validation is popular technique.

```
library(kernlab)
```

```
##
## Attaching package: 'kernlab'
```

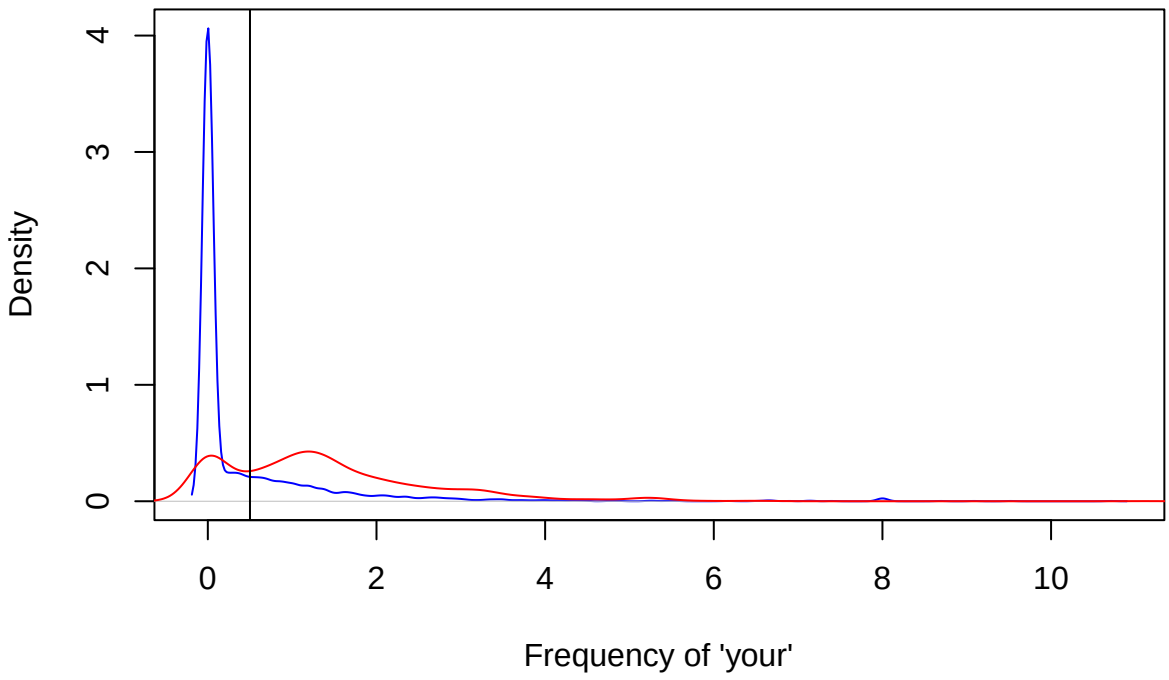
```
## The following object is masked from 'package:ggplot2':
##
##     alpha
```

```
data(spam)
head(spam)
```

```
##   make address  all num3d  our over remove internet order mail receive will
## 1 0.00    0.64 0.64     0 0.32 0.00   0.00     0.00  0.00 0.00    0.00 0.64
## 2 0.21    0.28 0.50     0 0.14 0.28   0.21     0.07  0.00 0.94    0.21 0.79
## 3 0.06    0.00 0.71     0 1.23 0.19   0.19     0.12  0.64 0.25    0.38 0.45
## 4 0.00    0.00 0.00     0 0.63 0.00   0.31     0.63  0.31 0.63    0.31 0.31
## 5 0.00    0.00 0.00     0 0.63 0.00   0.31     0.63  0.31 0.63    0.31 0.31
```

```
## 6 0.00    0.00 0.00     0 1.85 0.00   0.00     1.85  0.00 0.00    0.00 0.00
##   people report addresses free business email  you credit your font num000
## 1   0.00   0.00      0.00 0.32     0.00  1.29 1.93   0.00 0.96    0   0.00
## 2   0.65   0.21      0.14 0.14     0.07  0.28 3.47   0.00 1.59    0   0.43
## 3   0.12   0.00      1.75 0.06     0.06  1.03 1.36   0.32 0.51    0   1.16
## 4   0.31   0.00      0.00 0.31     0.00  0.00 3.18   0.00 0.31    0   0.00
## 5   0.31   0.00      0.00 0.31     0.00  0.00 3.18   0.00 0.31    0   0.00
## 6   0.00   0.00      0.00 0.00     0.00  0.00 0.00   0.00 0.00    0   0.00
##   money hp hpl george num650 lab labs telnet num857 data num415 num85
## 1  0.00  0   0      0      0   0    0      0      0    0      0     0
## 2  0.43  0   0      0      0   0    0      0      0    0      0     0
## 3  0.06  0   0      0      0   0    0      0      0    0      0     0
## 4  0.00  0   0      0      0   0    0      0      0    0      0     0
## 5  0.00  0   0      0      0   0    0      0      0    0      0     0
## 6  0.00  0   0      0      0   0    0      0      0    0      0     0
##   technology num1999 parts pm direct cs meeting original project   re  edu
## 1          0    0.00     0  0   0.00  0       0     0.00       0 0.00 0.00
## 2          0    0.07     0  0   0.00  0       0     0.00       0 0.00 0.00
## 3          0    0.00     0  0   0.06  0       0     0.12       0 0.06 0.06
## 4          0    0.00     0  0   0.00  0       0     0.00       0 0.00 0.00
## 5          0    0.00     0  0   0.00  0       0     0.00       0 0.00 0.00
## 6          0    0.00     0  0   0.00  0       0     0.00       0 0.00 0.00
##   table conference charSemicolon charRoundbracket charSquarebracket
## 1     0          0          0.00            0.000                 0
## 2     0          0          0.00            0.132                 0
## 3     0          0          0.01            0.143                 0
## 4     0          0          0.00            0.137                 0
## 5     0          0          0.00            0.135                 0
## 6     0          0          0.00            0.223                 0
##   charExclamation charDollar charHash capitalAve capitalLong capitalTotal type
## 1           0.778      0.000    0.000      3.756          61          278 spam
## 2           0.372      0.180    0.048      5.114         101         1028 spam
## 3           0.276      0.184    0.010      9.821         485         2259 spam
## 4           0.137      0.000    0.000      3.537          40          191 spam
## 5           0.135      0.000    0.000      3.537          40          191 spam
## 6           0.000      0.000    0.000      3.000          15           54 spam
```

```
plot(density(spam$your[spam$type=="nonspam"]),
     col="blue",main="",xlab="Frequency of 'your'")
lines(density(spam$your[spam$type=="spam"]),col="red")
abline(v=0.5,col="black")
```




```
##We want to find a value C, which is the cutoff point between spam frequencies, and non-spam frequenci

prediction <- ifelse(spam$your > 0.5,"spam","nonspam")
table(prediction,spam$type)/length(spam$type)
```

```
##
## prediction   nonspam      spam
##    nonspam 0.4590306 0.1017170
##    spam    0.1469246 0.2923278
```

###Caret for machine learning package

obj Class Package Predict Function Syntax lda MASS Predict(obj) (no options needed) glm stats predict(obj, type = "response") bgm gbm predict(obj, type = "response", n.trees) mda mda predict(obj, typoe ="posterior") rpart rpart predict(obj, type=- "prob") Weka RWeka predict(obj, type= "probability") LogitBoost caTools predict(obj, type= "raw", nIter)

** Different types of predictors have different object class syntax requirements.

```
library(caret)
```

```
##
## Attaching package: 'caret'

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

```r
library(kernlab) ## to get spam dataset

data(spam)
inTrain <- createDataPartition(y=spam$type, p= 0.75, list = FALSE) ## data is partitioned by spam type,
training <- spam[inTrain,]  ##subsetting the output of the partition function by spam
testing <- spam[-inTrain,]  ##subsetting the output of the partition function by not spam
dim(training)  ##shows the dimensions of the dataframe
```

```
## [1] 3451   58
```

```r
set.seed(11111)
modelFit <- train(type ~., data= training, method = "glm")  ##training a model, specifically a glm mode
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
modelFit
```

```
## Generalized Linear Model
##
## 3451 samples
##   57 predictor
##    2 classes: 'nonspam', 'spam'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 3451, 3451, 3451, 3451, 3451, 3451, ...
## Resampling results:
##
##   Accuracy  Kappa
##   0.919168  0.8302427
```

```
modelFit$finalModel  ##will return the fitted values for all of the other columns of data. The higher t
```

```
##
## Call:  NULL
##
## Coefficients:
##       (Intercept)               make            address                all
##         -1.671e+00         -2.453e-01         -1.538e-01          8.485e-02
##             num3d                our               over             remove
##          1.661e+00          5.668e-01          6.307e-01          2.443e+00
##          internet              order               mail            receive
##          5.826e-01          2.844e-01          5.552e-02         -5.869e-01
##              will             people             report          addresses
##         -1.568e-01         -1.696e-01          9.178e-02          9.142e-01
##              free           business              email                you
##          9.522e-01          9.009e-01          6.647e-02          1.119e-01
##            credit               your               font             num000
##          9.564e-01          2.165e-01          1.969e-01          2.149e+00
##             money                 hp                hpl             george
##          7.201e-01         -1.595e+00         -1.439e+00         -4.871e+00
##            num650                lab               labs             telnet
##          5.969e-01         -4.523e+00         -9.483e-03         -6.249e+00
##            num857               data             num415              num85
##          2.468e+00         -7.011e-01         -7.162e+00         -2.197e+00
```

```
##        technology            num1999              parts                 pm
##         7.988e-01          4.972e-02         -6.271e-01         -9.098e-01
##            direct                 cs            meeting           original
##         4.704e-01         -3.570e+02         -2.181e+00         -2.297e+00
##           project                 re                edu              table
##        -1.505e+00         -8.806e-01         -1.683e+00         -2.575e+00
##        conference      charSemicolon   charRoundbracket  charSquarebracket
##        -3.942e+00         -1.278e+00         -1.816e-01         -8.319e-01
##   charExclamation         charDollar           charHash         capitalAve
##         2.443e-01          6.219e+00          1.576e+00          8.079e-02
##       capitalLong       capitalTotal
##         1.132e-02          5.696e-04
##
## Degrees of Freedom: 3450 Total (i.e. Null);  3393 Residual
## Null Deviance:       4628
## Residual Deviance: 1369  AIC: 1485
```

```
predictions <- predict(modelFit, newdata = testing)  ##how we can predict on new samples, here using th
predictions
```

```
##    [1] spam    spam    spam    spam    spam    spam    nonspam nonspam nonspam
##   [10] spam    nonspam spam    spam    spam    spam    spam    spam    spam
##   [19] spam    spam    spam    spam    spam    nonspam spam    spam    spam
##   [28] spam    spam    nonspam spam    spam    nonspam spam    spam    nonspam
##   [37] spam    spam    spam    spam    spam    spam    spam    spam    spam
##   [46] spam    nonspam spam    spam    spam    spam    nonspam spam    spam
##   [55] spam    spam    nonspam spam    spam    spam    spam    spam    spam
##   [64] spam    spam    spam    spam    spam    nonspam spam    spam    spam
##   [73] spam    nonspam spam    spam    spam    spam    spam    spam    spam
##   [82] nonspam spam    nonspam nonspam spam    spam    spam    spam    spam
##   [91] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [100] nonspam spam    nonspam spam    spam    spam    spam    spam    spam
##  [109] spam    nonspam spam    spam    spam    spam    spam    spam    nonspam
##  [118] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [127] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [136] spam    nonspam spam    spam    spam    spam    nonspam spam    spam
##  [145] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [154] nonspam spam    spam    spam    nonspam spam    spam    spam    spam
##  [163] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [172] spam    spam    spam    nonspam spam    spam    nonspam spam    spam
##  [181] spam    spam    spam    spam    spam    spam    spam    nonspam spam
##  [190] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [199] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [208] spam    spam    nonspam spam    spam    spam    spam    spam    spam
##  [217] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [226] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [235] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [244] spam    nonspam spam    spam    spam    spam    spam    spam    spam
##  [253] spam    spam    spam    spam    spam    spam    nonspam spam    spam
##  [262] spam    spam    spam    spam    nonspam spam    spam    spam    spam
##  [271] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [280] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [289] spam    spam    spam    spam    spam    nonspam spam    spam    spam
##  [298] spam    spam    spam    spam    spam    nonspam spam    spam    spam
```

```
##  [307] nonspam spam    spam    spam    spam    spam    spam    spam    spam
##  [316] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [325] spam    spam    spam    nonspam spam    spam    spam    spam    spam
##  [334] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [343] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [352] nonspam spam    spam    spam    spam    spam    spam    spam    spam
##  [361] spam    nonspam spam    spam    spam    spam    spam    spam    spam
##  [370] spam    spam    spam    nonspam spam    spam    spam    spam    spam
##  [379] spam    spam    nonspam spam    spam    spam    spam    spam    spam
##  [388] spam    spam    nonspam spam    spam    spam    spam    spam    spam
##  [397] spam    spam    spam    spam    spam    spam    spam    spam    nonspam
##  [406] spam    nonspam spam    nonspam spam    spam    spam    spam    spam
##  [415] nonspam nonspam spam    spam    spam    nonspam nonspam spam    spam
##  [424] nonspam nonspam spam    spam    nonspam nonspam spam    spam    spam
##  [433] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [442] spam    spam    spam    spam    spam    spam    spam    spam    spam
##  [451] spam    spam    nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [460] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [469] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [478] nonspam nonspam nonspam nonspam nonspam nonspam spam    nonspam nonspam
##  [487] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [496] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [505] nonspam nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam
##  [514] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [523] spam    nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [532] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [541] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [550] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [559] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [568] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [577] nonspam spam    nonspam spam    nonspam nonspam nonspam nonspam nonspam
##  [586] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [595] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [604] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [613] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [622] nonspam nonspam nonspam nonspam nonspam spam    nonspam nonspam nonspam
##  [631] nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam    nonspam
##  [640] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [649] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [658] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [667] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [676] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [685] nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [694] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [703] nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam
##  [712] nonspam nonspam nonspam nonspam spam    nonspam nonspam nonspam nonspam
##  [721] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [730] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [739] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [748] nonspam nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam
##  [757] nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam    nonspam
##  [766] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [775] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [784] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
```

```
##  [793] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [802] nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [811] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [820] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [829] nonspam spam    spam    nonspam spam    nonspam nonspam nonspam nonspam
##  [838] nonspam nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam
##  [847] nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam
##  [856] nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [865] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam
##  [874] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [883] nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam
##  [892] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [901] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [910] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [919] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [928] nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam    nonspam
##  [937] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam
##  [946] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [955] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [964] nonspam nonspam spam    spam    nonspam nonspam nonspam nonspam nonspam
##  [973] nonspam spam    nonspam nonspam nonspam nonspam nonspam spam    nonspam
##  [982] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
##  [991] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1000] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1009] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1018] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1027] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1036] nonspam nonspam spam    nonspam nonspam nonspam nonspam nonspam nonspam
## [1045] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1054] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1063] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1072] nonspam nonspam nonspam nonspam nonspam nonspam nonspam spam    nonspam
## [1081] nonspam nonspam nonspam nonspam spam    nonspam nonspam spam    nonspam
## [1090] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1099] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1108] nonspam nonspam spam    nonspam spam    nonspam nonspam nonspam nonspam
## [1117] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1126] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1135] nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## [1144] nonspam nonspam nonspam nonspam nonspam nonspam nonspam
## Levels: nonspam spam
```

```
confusionMatrix(predictions, testing$type)  ##using the confusingMatrix argument to determine how well
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction nonspam spam
##    nonspam     664   52
##    spam         33  401
##
##                Accuracy : 0.9261
##                  95% CI : (0.9094, 0.9405)
##     No Information Rate : 0.6061
```

```
##     P-Value [Acc > NIR] : < 2e-16
##
##                   Kappa : 0.8441
##
##  Mcnemar's Test P-Value : 0.05089
##
##             Sensitivity : 0.9527
##             Specificity : 0.8852
##          Pos Pred Value : 0.9274
##          Neg Pred Value : 0.9240
##              Prevalence : 0.6061
##          Detection Rate : 0.5774
##    Detection Prevalence : 0.6226
##       Balanced Accuracy : 0.9189
##
##        'Positive' Class : nonspam
##
```

Data Slicing with caret

```
library(caret)
library(kernlab) ## to get spam dataset

data(spam)
inTrain <- createDataPartition(y=spam$type, p= 0.75, list = FALSE) ## data is partitioned by spam type,
training <- spam[inTrain,]  ##subsetting the output of the partition function by spam
testing <- spam[-inTrain,]  ##subsetting the output of the partition function by not spam
dim(training)
```

```
## [1] 3451   58
```

```
set.seed(11111)
folds <- createFolds(y= spam$type, k=10, list= TRUE, returnTrain=TRUE)  ##k is the number of folds, sma
sapply(folds, length)
```

```
## Fold01 Fold02 Fold03 Fold04 Fold05 Fold06 Fold07 Fold08 Fold09 Fold10
##   4141   4140   4141   4142   4142   4141   4140   4140   4141   4141
```

Plotting Predictors + Hmisc for cutting data.frames

```
library(ISLR)
library(ggplot2)
library(caret)

data(wage)
```

```
## Warning in data(wage): data set 'wage' not found
```

```
summary(Wage)
```

```
##       year           age                      maritl           race
##  Min.   :2003   Min.   :18.00   1. Never Married: 648   1. White:2480
##  1st Qu.:2004   1st Qu.:33.75   2. Married      :2074   2. Black: 293
##  Median :2006   Median :42.00   3. Widowed      :  19   3. Asian: 190
##  Mean   :2006   Mean   :42.41   4. Divorced     : 204   4. Other:  37
##  3rd Qu.:2008   3rd Qu.:51.00   5. Separated    :  55
##  Max.   :2009   Max.   :80.00
##
##                education                     region               jobclass
##  1. < HS Grad      :268   2. Middle Atlantic   :3000   1. Industrial :1544
##  2. HS Grad        :971   1. New England       :   0   2. Information:1456
##  3. Some College   :650   3. East North Central:   0
##  4. College Grad   :685   4. West North Central:   0
##  5. Advanced Degree:426   5. South Atlantic    :   0
##                           6. East South Central:   0
##                           (Other)              :   0
##             health      health_ins      logwage           wage
##  1. <=Good     : 858   1. Yes:2083   Min.   :3.000   Min.   : 20.09
##  2. >=Very Good:2142   2. No : 917   1st Qu.:4.447   1st Qu.: 85.38
##                                      Median :4.653   Median :104.92
##                                      Mean   :4.654   Mean   :111.70
##                                      3rd Qu.:4.857   3rd Qu.:128.68
##                                      Max.   :5.763   Max.   :318.34
##
```

```
head(Wage, n=20)
```

```
##        year age           maritl     race          education             region
## 231655 2006  18 1. Never Married 1. White       1. < HS Grad 2. Middle Atlantic
## 86582  2004  24 1. Never Married 1. White    4. College Grad 2. Middle Atlantic
## 161300 2003  45       2. Married 1. White    3. Some College 2. Middle Atlantic
## 155159 2003  43       2. Married 3. Asian    4. College Grad 2. Middle Atlantic
## 11443  2005  50      4. Divorced 1. White         2. HS Grad 2. Middle Atlantic
## 376662 2008  54       2. Married 1. White    4. College Grad 2. Middle Atlantic
## 450601 2009  44       2. Married 4. Other    3. Some College 2. Middle Atlantic
## 377954 2008  30 1. Never Married 3. Asian    3. Some College 2. Middle Atlantic
## 228963 2006  41 1. Never Married 2. Black    3. Some College 2. Middle Atlantic
## 81404  2004  52       2. Married 1. White         2. HS Grad 2. Middle Atlantic
## 302778 2007  45      4. Divorced 1. White    3. Some College 2. Middle Atlantic
## 305706 2007  34       2. Married 1. White         2. HS Grad 2. Middle Atlantic
## 8690   2005  35 1. Never Married 1. White         2. HS Grad 2. Middle Atlantic
## 153561 2003  39       2. Married 1. White    4. College Grad 2. Middle Atlantic
## 449654 2009  54       2. Married 1. White         2. HS Grad 2. Middle Atlantic
## 447660 2009  51       2. Married 1. White    3. Some College 2. Middle Atlantic
## 160191 2003  37 1. Never Married 3. Asian    4. College Grad 2. Middle Atlantic
## 230312 2006  50       2. Married 1. White 5. Advanced Degree 2. Middle Atlantic
## 301585 2007  56       2. Married 1. White    4. College Grad 2. Middle Atlantic
## 153682 2003  37 1. Never Married 1. White    3. Some College 2. Middle Atlantic
##              jobclass         health health_ins  logwage      wage
## 231655  1. Industrial      1. <=Good      2. No 4.318063  75.04315
```

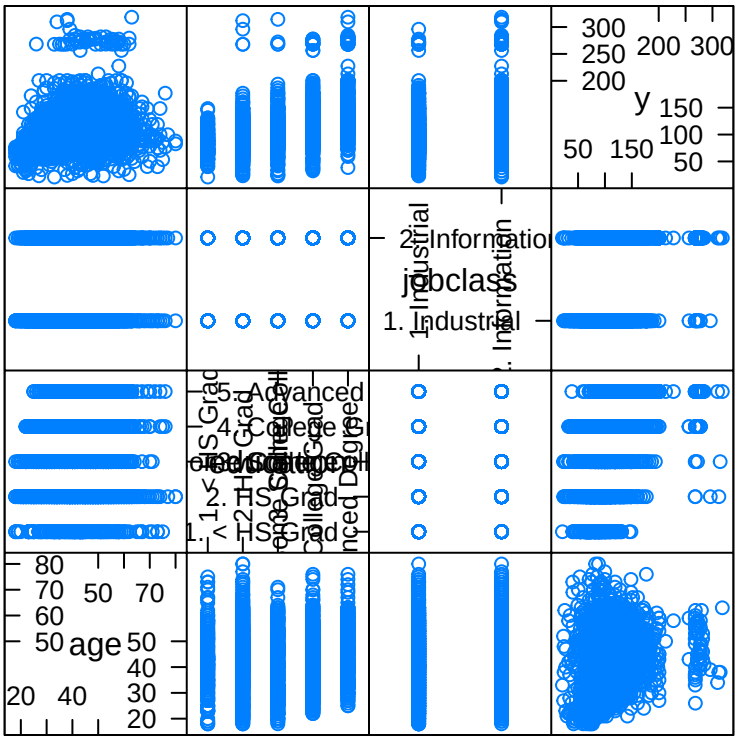
```
## 86582  2. Information 2. >=Very Good      2. No 4.255273  70.47602
## 161300  1. Industrial      1. <=Good     1. Yes 4.875061 130.98218
## 155159 2. Information 2. >=Very Good     1. Yes 5.041393 154.68529
## 11443  2. Information      1. <=Good     1. Yes 4.318063  75.04315
## 376662 2. Information 2. >=Very Good     1. Yes 4.845098 127.11574
## 450601  1. Industrial 2. >=Very Good     1. Yes 5.133021 169.52854
## 377954 2. Information      1. <=Good     1. Yes 4.716003 111.72085
## 228963 2. Information 2. >=Very Good     1. Yes 4.778151 118.88436
## 81404  2. Information 2. >=Very Good     1. Yes 4.857332 128.68049
## 302778 2. Information      1. <=Good     1. Yes 4.763428 117.14682
## 305706  1. Industrial 2. >=Very Good      2. No 4.397940  81.28325
## 8690   2. Information 2. >=Very Good     1. Yes 4.494155  89.49248
## 153561  1. Industrial 2. >=Very Good     1. Yes 4.903090 134.70538
## 449654 2. Information 2. >=Very Good     1. Yes 4.903090 134.70538
## 447660  1. Industrial 2. >=Very Good     1. Yes 4.505150  90.48191
## 160191  1. Industrial 2. >=Very Good      2. No 4.414973  82.67964
## 230312 2. Information 2. >=Very Good      2. No 5.360552 212.84235
## 301585  1. Industrial      1. <=Good     1. Yes 4.861026 129.15669
## 153682  1. Industrial 2. >=Very Good     1. Yes 4.591065  98.59934
```

```
inTrain <- createDataPartition(y=Wage$wage, p=0.7, list=FALSE)
training <- Wage[inTrain,]
testing <- Wage[-inTrain,]
dim(training); dim(testing)  ##returns the number of rows and columns
```

```
## [1] 2102   11
```

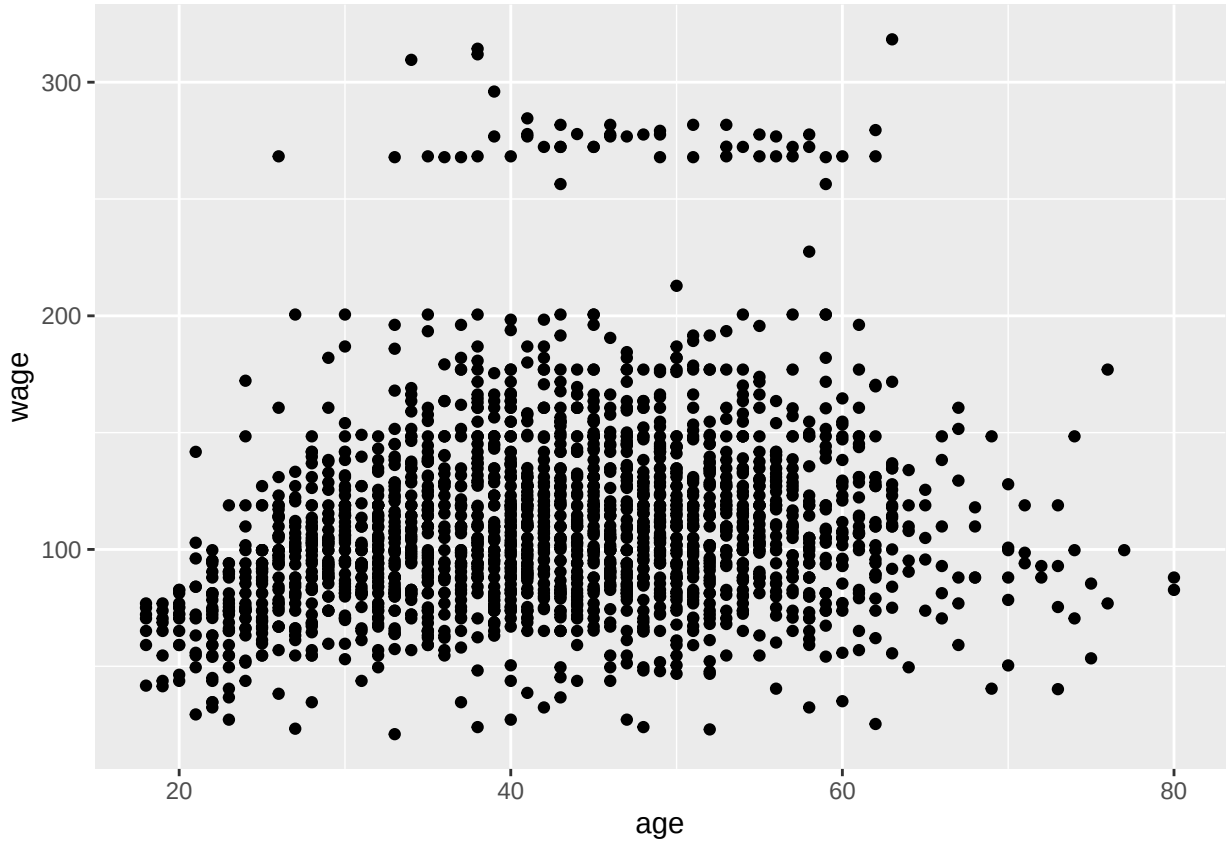
```
## [1] 898  11
```

```
featurePlot(x=training[, c("age", "education", "jobclass")], ##a plotting function that comes with care
            y= training$wage,  ##this is your output of interest
            plot="pairs")
```

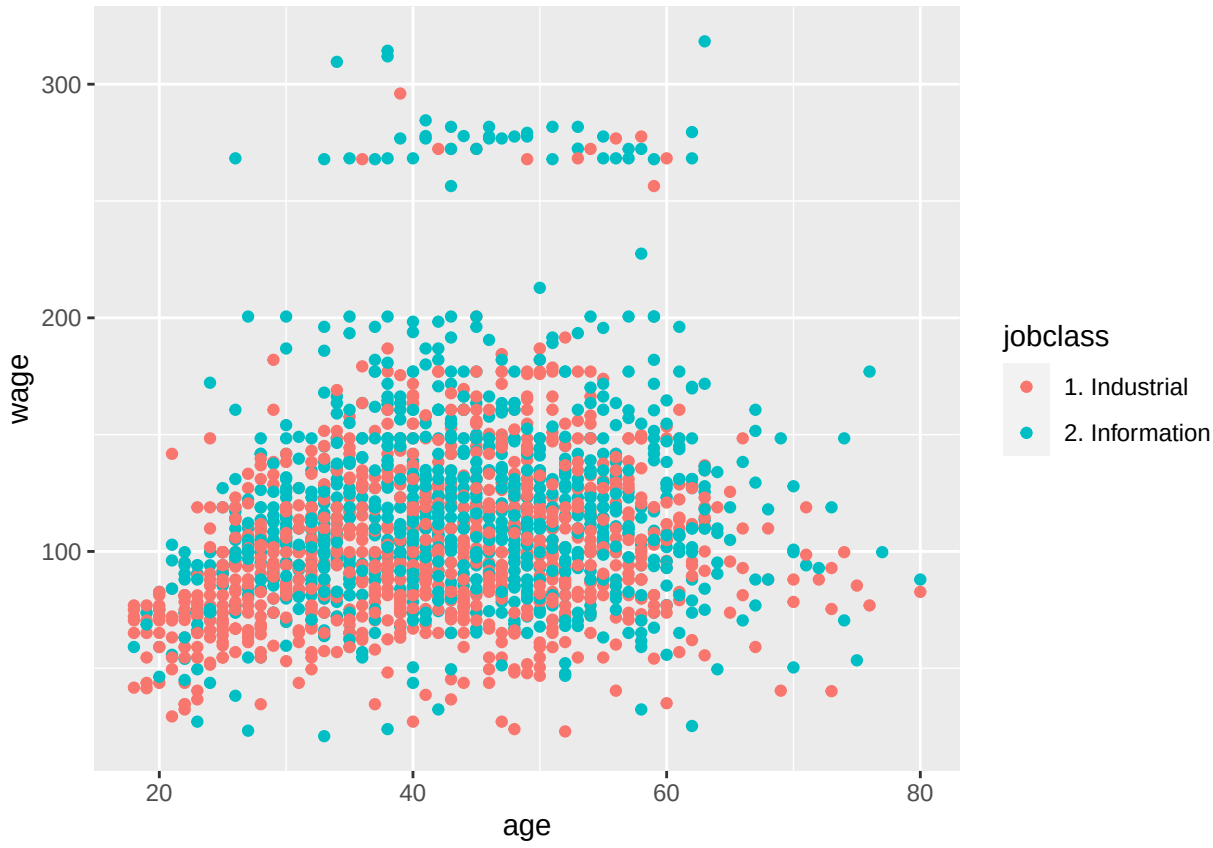


Scatter Plot Matrix

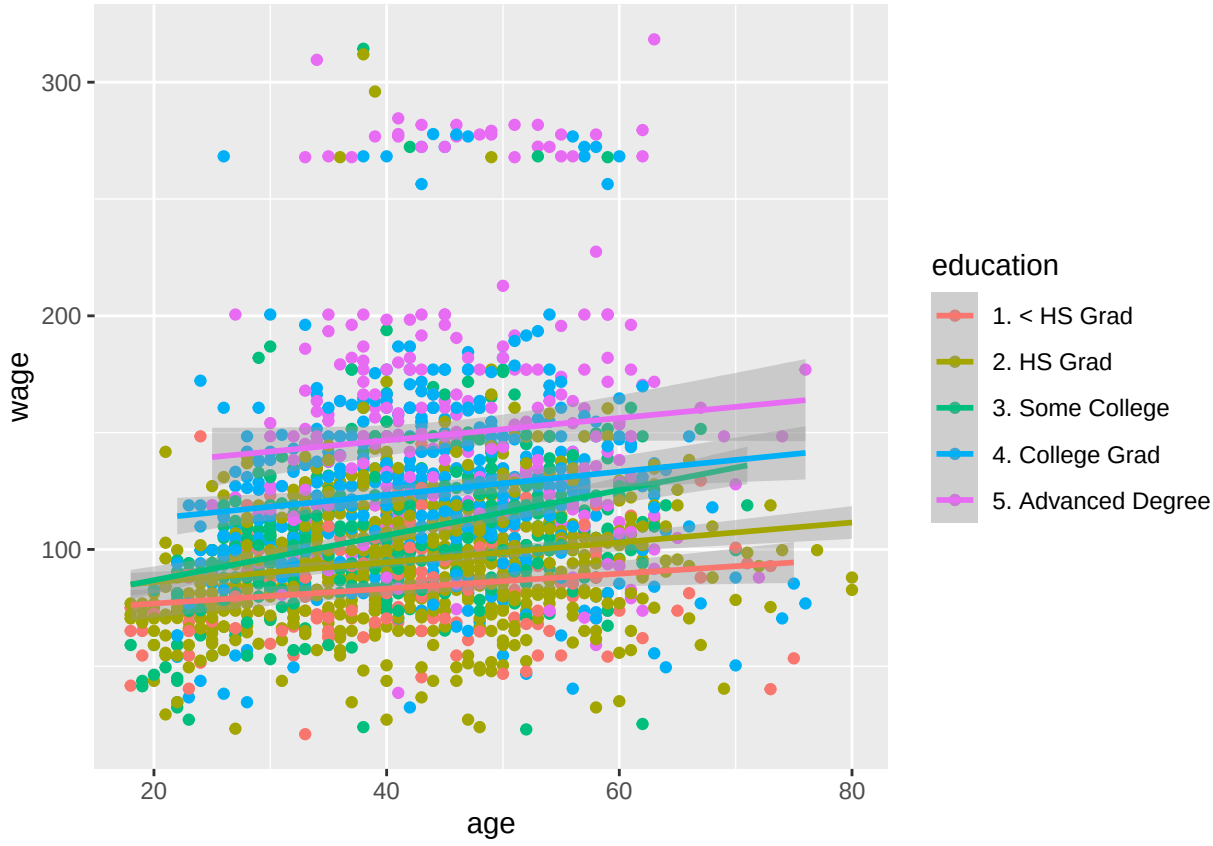
```
qplot(age, wage, data=training) ##here we see two distinct groups
```




```
qplot(age, wage, colour= jobclass, data= training)
```



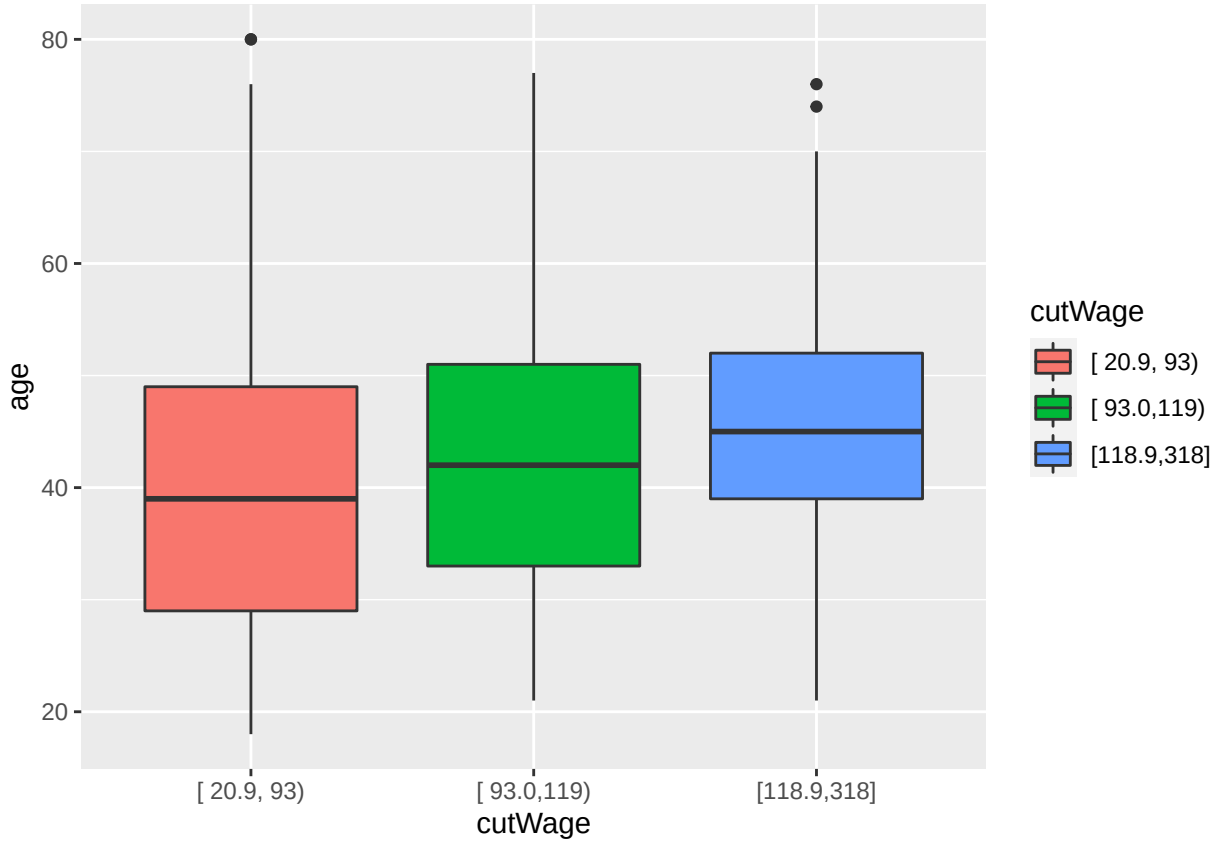

```
qp <- qplot(age, wage, color= education, data=training)
qp + geom_smooth(method = 'lm', formula= y~x)
```




```
library(Hmisc)  ##good library for cutting data.frames into pieces.
cutWage <- cut2(training$wage, g=3) ##cut2 argument performs the cut, by $wage column, into g=3 pieces.
table(cutWage)
```

```
## cutWage
## [ 20.9, 93) [ 93.0,119) [118.9,318]
##         707         718         677
```

```
p1 <- qplot(cutWage, age, data=training, fill=cutWage, geom=c("boxplot"))
p1
```



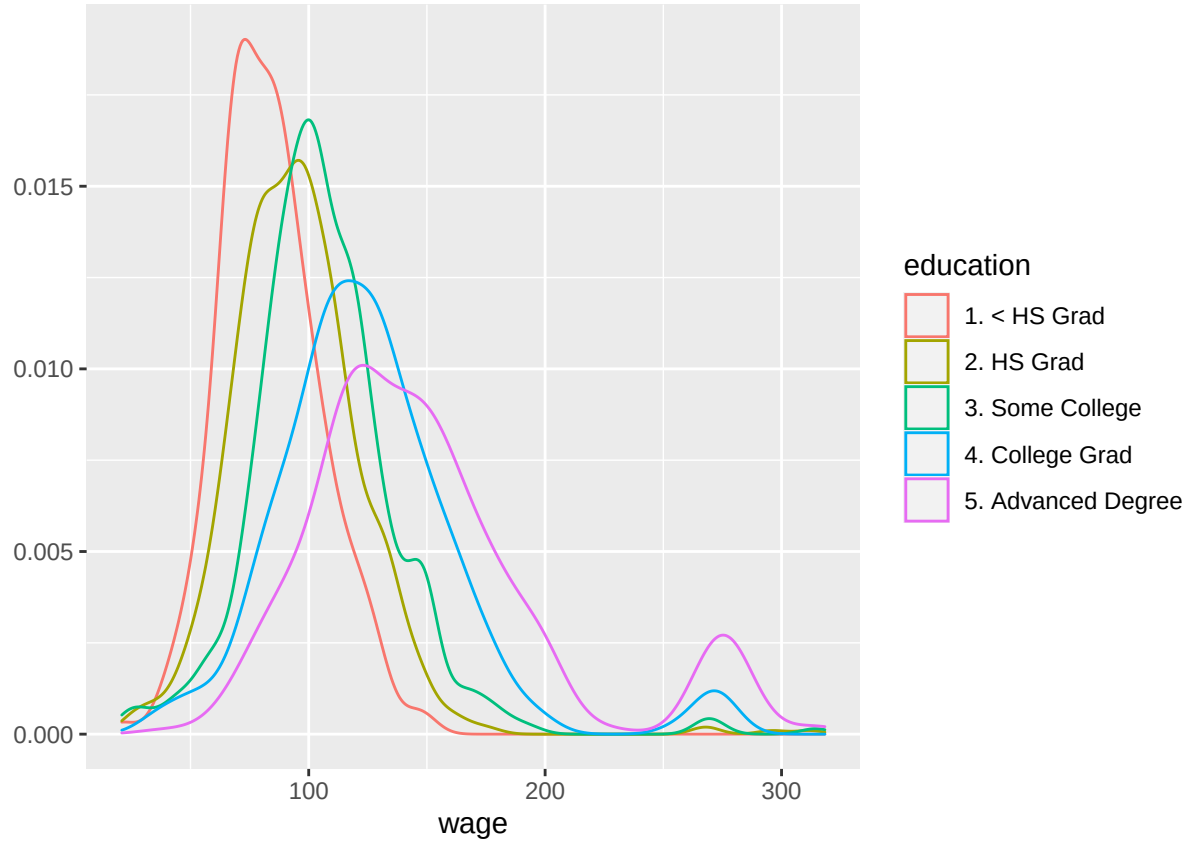

```
t1 <- table(cutWage, training$jobclass)  ##great way to make quick tables
t1
```

```
##
## cutWage       1. Industrial 2. Information
##   [ 20.9, 93)           455            252
##   [ 93.0,119)           362            356
##   [118.9,318]           276            401
```

```
prop.table(t1,1)  ##proportion table, where the 1 represents proportion in each row. a 2 would be each
```

```
##
## cutWage       1. Industrial 2. Information
##   [ 20.9, 93)     0.6435644      0.3564356
##   [ 93.0,119)     0.5041783      0.4958217
##   [118.9,318]     0.4076809      0.5923191
```

```
qplot(wage, color=education, data= training, geom="density")
```

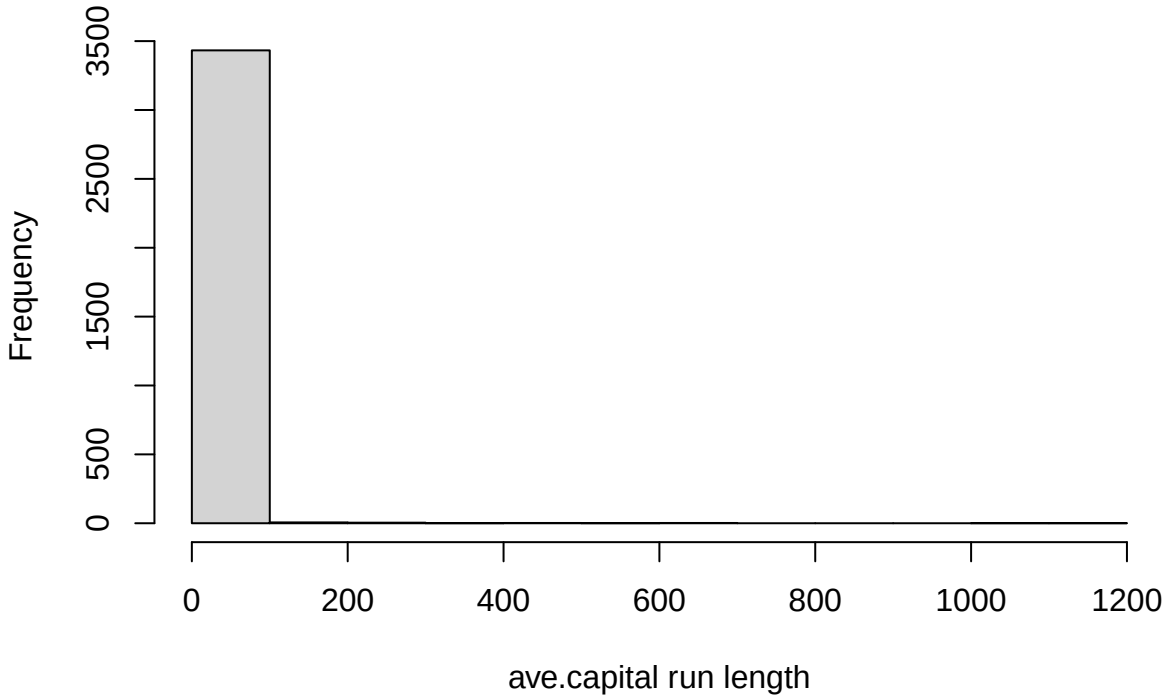


Pre-processing predictor variables

```
library(caret)
library(kernlab) ## to get spam dataset
library(RANN)

data(spam)
inTrain <- createDataPartition(y=spam$type, p= 0.75, list = FALSE) ## data is partitioned by spam type,
training <- spam[inTrain,]  ##subsetting the output of the partition function by spam
testing <- spam[-inTrain,]  ##subsetting the output of the partition function by not spam
dim(training)
```

```
## [1] 3451   58
```

```
hist(training$capitalAve, main="", xlab="ave.capital run length")
```




```
mean(training$capitalAve)
```

```
## [1] 5.577148
```

```
sd(training$capitalAve) ## what we see here is that the standard deviation is significanly higher than
```

```
## [1] 35.92989
```

```
trainCapAve <- training$capitalAve
trainCapAveS <- (trainCapAve - mean(trainCapAve))/sd(trainCapAve) ## A way of standardizing the data
mean(trainCapAveS)
```

```
## [1] -1.162129e-17
```

```
sd(trainCapAveS)
```

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

```
##If we want to then standardize the test set, we must use the mean from the training set, and the st.d
testCapAve <- testing$capitalAve
testCapAveS <- (testCapAve - mean(trainCapAve))/sd(trainCapAve)
mean(testCapAveS)
```

```
## [1] -0.04294102
```

```r
sd(testCapAveS)
```

```
## [1] 0.3437084
```

```r
##Alternative is passing preProcess into the train() argument
modelFit <- train(type~., data= training, preProcess= c("center", "scale"), method="glm")  ##This proce
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: algorithm did not converge

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
modelFit
```

```
## Generalized Linear Model
##
## 3451 samples
##   57 predictor
##    2 classes: 'nonspam', 'spam'
##
## Pre-processing: centered (57), scaled (57)
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 3451, 3451, 3451, 3451, 3451, 3451, ...
## Resampling results:
##
##   Accuracy   Kappa
##   0.9167869  0.8257754
```

```
## Prediction models tend to fail with NA. Well use K nearest imputation statistical function.
# Make some values NA
training$capAve <- training$capitalAve
selectNA <- rbinom(dim(training)[1],size=1,prob=0.05)==1  ## we are adding NA's into the dataset here.
training$capAve[selectNA] <- NA
# Impute and standardize
preObj <- preProcess(training[,-58],method="knnImpute")  ##removing the 58th column here
capAve <- predict(preObj,training[,-58])$capAve
# Standardize true values
capAveTruth <- training$capitalAve
capAveTruth <- (capAveTruth-mean(capAveTruth))/sd(capAveTruth)
```

Covariate (predictor) creation

```
 ##think about creating or calculating variables/features to summarize the sample data, in this instanc

library(ISLR); library(caret); data(Wage);
inTrain <- createDataPartition(y=Wage$wage,
                               p=0.7, list=FALSE)
training <- Wage[inTrain,]
```

```
testing <- Wage[-inTrain,]

##turning qualitative variables into quantitative ones
table(training$jobclass)
```

```
##
##  1. Industrial 2. Information
##           1067           1035
```

```
dummies <- dummyVars(wage ~ jobclass,data=training)  ##using the dummyVars argument to turn categorical
head(predict(dummies,newdata=training))
```

```
##        jobclass.1. Industrial jobclass.2. Information
## 161300                      1                       0
## 155159                      0                       1
## 376662                      0                       1
## 450601                      1                       0
## 377954                      0                       1
## 228963                      0                       1
```

```
##__________________
##A way to throw out less-meaningful predictors, i.e. near-zero variations in the data.
nsv <- nearZeroVar(training,saveMetrics=TRUE)
nsv  ##in this example we can throw out sex and region
```

```
##            freqRatio percentUnique zeroVar   nzv
## year        1.008721    0.33301618   FALSE FALSE
## age         1.013699    2.85442436   FALSE FALSE
## maritl      3.152505    0.23786870   FALSE FALSE
## race        8.964103    0.19029496   FALSE FALSE
## education   1.432150    0.23786870   FALSE FALSE
## region      0.000000    0.04757374    TRUE  TRUE
## jobclass    1.030918    0.09514748   FALSE FALSE
## health      2.434641    0.09514748   FALSE FALSE
## health_ins  2.269051    0.09514748   FALSE FALSE
## logwage     1.000000   19.36251189   FALSE FALSE
## wage        1.000000   19.36251189   FALSE FALSE
```

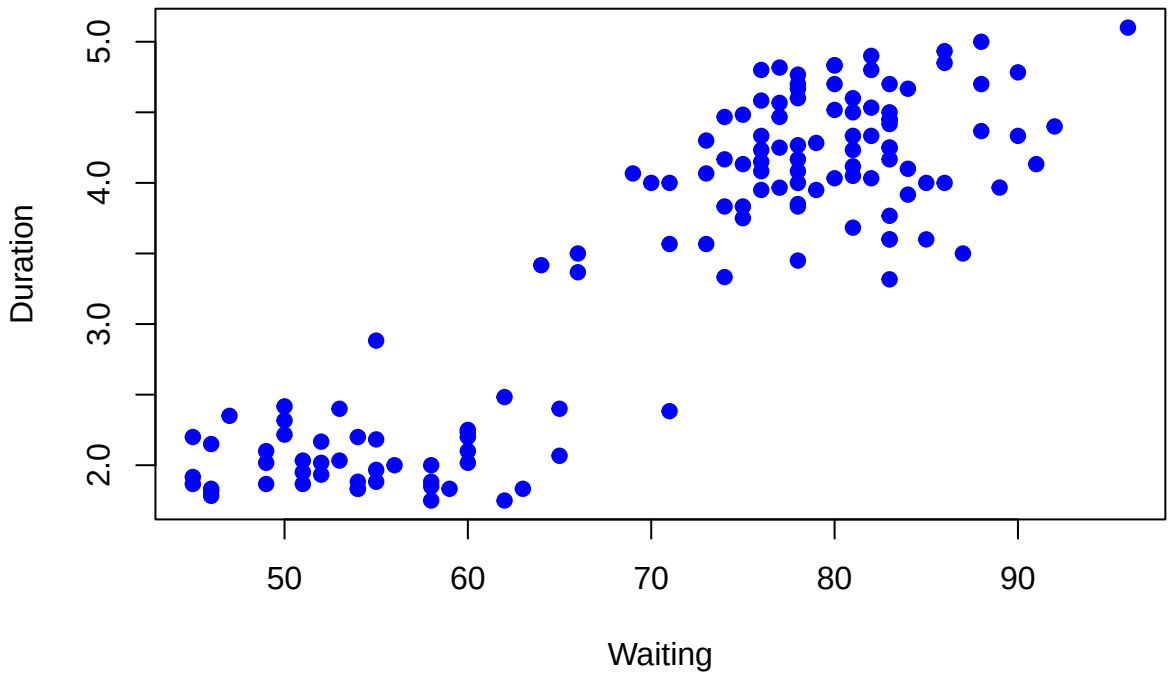
Predicting with Regression

```
##Using the old faithful dataset
library(caret);data(faithful); set.seed(333)
inTrain <- createDataPartition(y=faithful$waiting,
                               p=0.5, list=FALSE)
trainFaith <- faithful[inTrain,]; testFaith <- faithful[-inTrain,]
head(trainFaith)
```

```
##    eruptions waiting
## 3      3.333      74
```

```
## 6      2.883      55
## 7      4.700      88
## 8      3.600      85
## 9      1.950      51
## 11     1.833      54
```

```
plot(trainFaith$waiting,trainFaith$eruptions,pch=19,col="blue",xlab="Waiting",ylab="Duration") ##plotti
```




```
lm1 <- lm(eruptions ~ waiting, data=trainFaith)
summary(lm1)
```

```
##
## Call:
## lm(formula = eruptions ~ waiting, data = trainFaith)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13375 -0.36778  0.06064  0.36578  0.96057
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.648629   0.226603  -7.275 2.55e-11 ***
## waiting      0.072211   0.003136  23.026  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.4941 on 135 degrees of freedom
## Multiple R-squared:  0.7971, Adjusted R-squared:  0.7956
## F-statistic: 530.2 on 1 and 135 DF,  p-value: < 2.2e-16
```

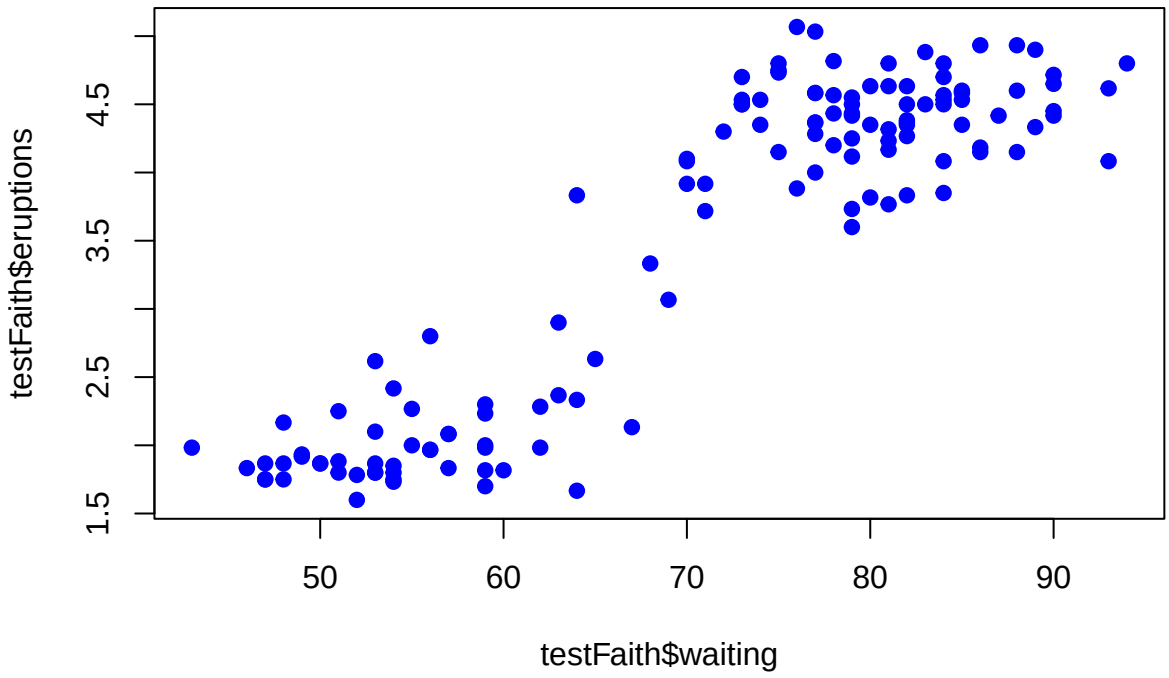
```
##output is y(eruption duration) = 0.073 * (waiting time) - 1.792

## To predict a new value, we can automate this by:
coef(lm1)[1] + coef(lm1)[2]*80  ## coef(lm1)[1] returns the intercept, and [2] returns the slope. We ar
```

```
## (Intercept)
##    4.128276
```

```
newdata <- data.frame(waiting=80)
predict(lm1,newdata)  ## a shortcut so we don't have to continuously calculate the coeficients.
```

```
##        1
## 4.128276
```

```
##We can not use the predictions from the training set on the TEST set
pred1 <- predict(lm1,newdata=testFaith,interval="prediction")
ord <- order(testFaith$waiting)
plot(testFaith$waiting,testFaith$eruptions,pch=19,col="blue")
```



```
##_________________
##We can use CARET to do the same, and much faster
modFit <- train(eruptions ~ waiting,data=trainFaith,method="lm")  ##eruptions is outcome, waiting is pr
summary(modFit$finalModel)  ##How we get final model output
```

```
##
## Call:
## lm(formula = .outcome ~ ., data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13375 -0.36778  0.06064  0.36578  0.96057
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.648629   0.226603  -7.275 2.55e-11 ***
## waiting      0.072211   0.003136  23.026  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4941 on 135 degrees of freedom
## Multiple R-squared:  0.7971, Adjusted R-squared:  0.7956
## F-statistic: 530.2 on 1 and 135 DF,  p-value: < 2.2e-16
```

Predicting Multiple Covariate Regression

```
library(ISLR); library(ggplot2); library(caret);
data(Wage); Wage <- subset(Wage,select=-c(logwage))  ##here we are subsetting and removing the variable
summary(Wage)
```

```
##       year           age                     maritl           race
##  Min.   :2003   Min.   :18.00   1. Never Married: 648   1. White:2480
##  1st Qu.:2004   1st Qu.:33.75   2. Married      :2074   2. Black: 293
##  Median :2006   Median :42.00   3. Widowed      :  19   3. Asian: 190
##  Mean   :2006   Mean   :42.41   4. Divorced     : 204   4. Other:  37
##  3rd Qu.:2008   3rd Qu.:51.00   5. Separated    :  55
##  Max.   :2009   Max.   :80.00
##
##                education                      region               jobclass
##  1. < HS Grad      :268   2. Middle Atlantic   :3000   1. Industrial :1544
##  2. HS Grad        :971   1. New England       :   0   2. Information:1456
##  3. Some College   :650   3. East North Central:   0
##  4. College Grad   :685   4. West North Central:   0
##  5. Advanced Degree:426   5. South Atlantic    :   0
##                           6. East South Central:   0
##                           (Other)              :   0
##             health      health_ins        wage
##  1. <=Good     : 858   1. Yes:2083   Min.   : 20.09
##  2. >=Very Good:2142   2. No : 917   1st Qu.: 85.38
##                                      Median :104.92
##                                      Mean   :111.70
```

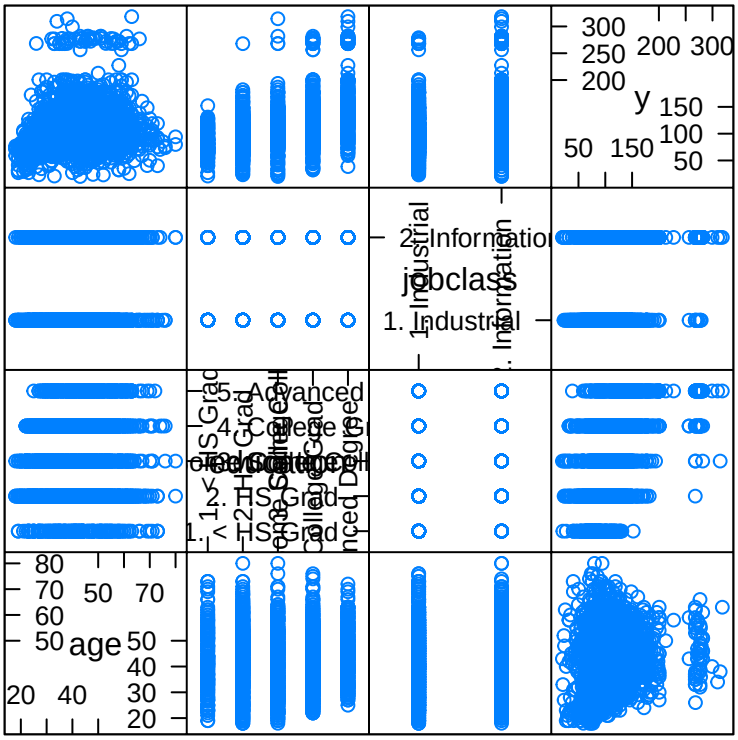
```
##                                       3rd Qu.:128.68
##                                       Max.   :318.34
##
```

```
inTrain <- createDataPartition(y=Wage$wage,##subsetting into test and train
                               p=0.7, list=FALSE)
training <- Wage[inTrain,]
testing <- Wage[-inTrain,]
dim(training); dim(testing)
```

```
## [1] 2102   10
```

```
## [1] 898  10
```

```
featurePlot(x=training[,c("age","education","jobclass")],
            y = training$wage,
            plot="pairs")
```

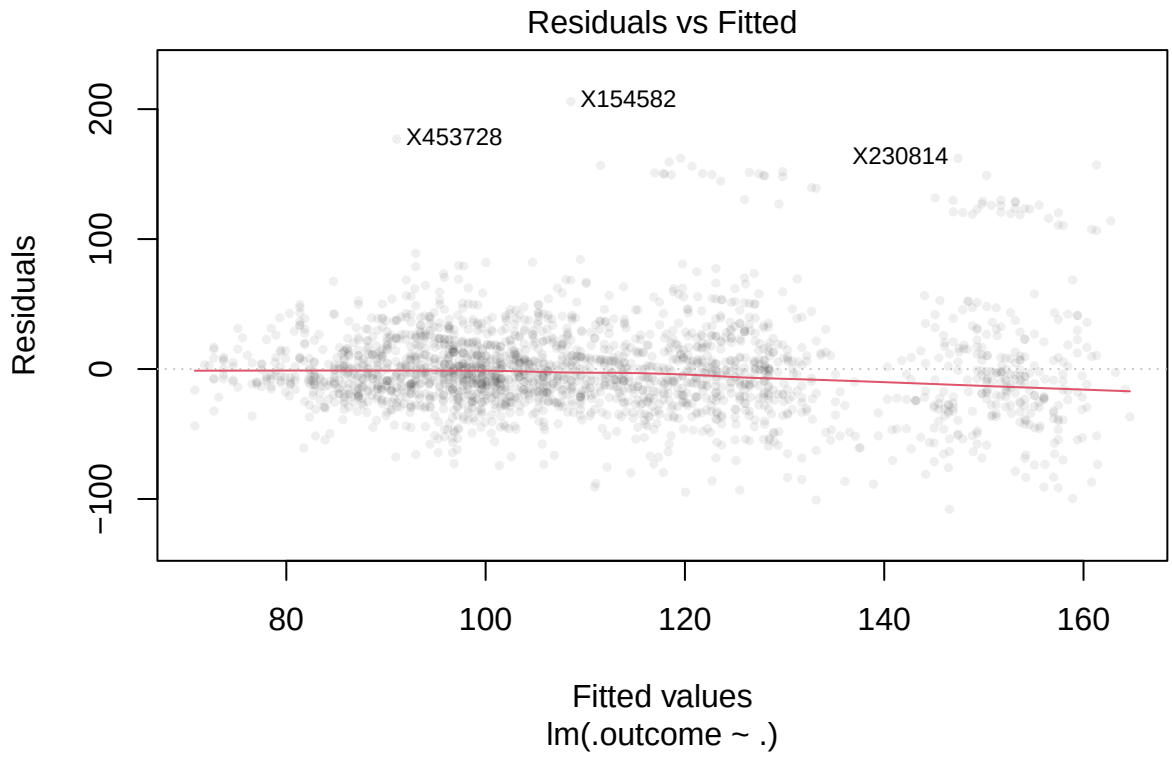


Scatter Plot Matrix

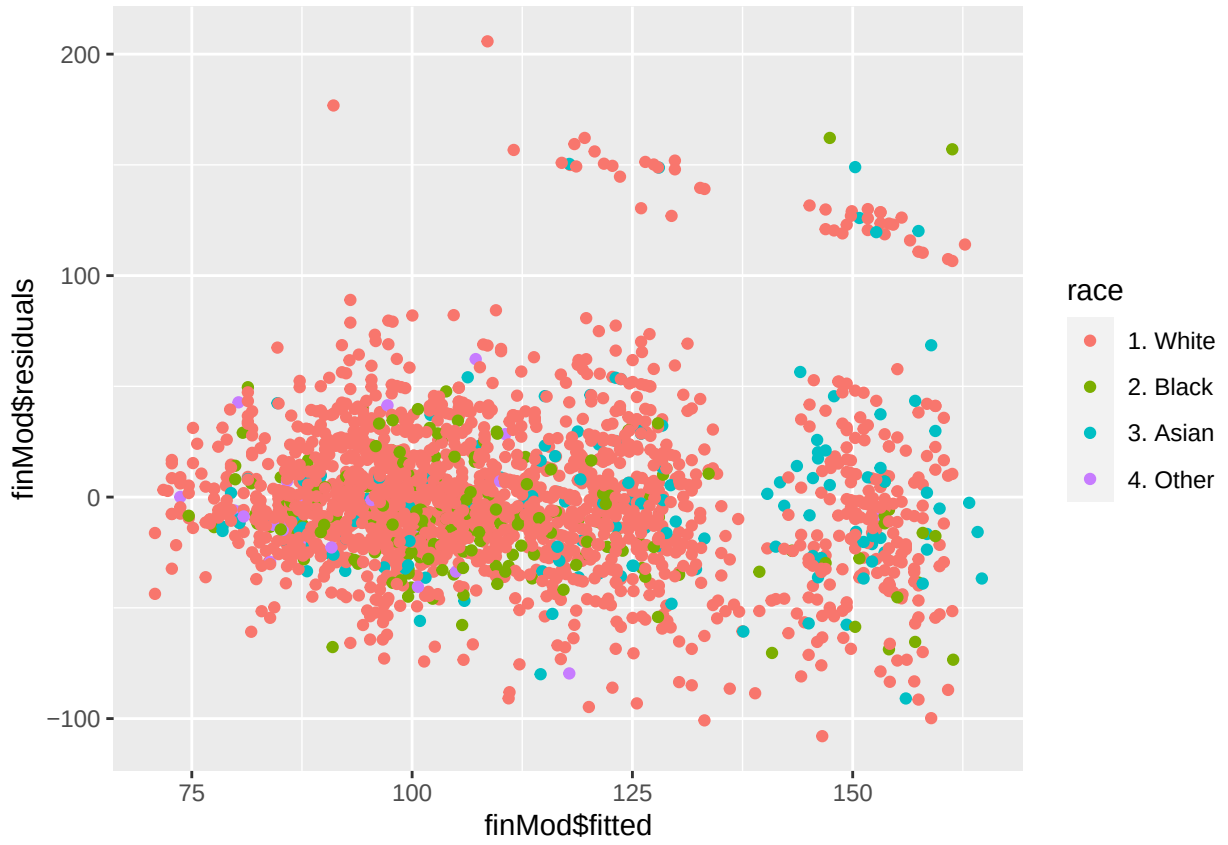
```
##_________________

modFit<- train(wage ~ age + jobclass + education,
               method = "lm",data=training)
finMod <- modFit$finalModel
print(modFit)
```

```
## Linear Regression
##
## 2102 samples
##    3 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 2102, 2102, 2102, 2102, 2102, 2102, ...
## Resampling results:
##
##   RMSE      Rsquared   MAE
##   35.56759  0.2589245  24.87554
##
## Tuning parameter 'intercept' was held constant at a value of TRUE
```

```
plot(finMod,1,pch=19,cex=0.5,col="#00000010")  ##residuals vs. fitted. We want to see a straight line
```




```
qplot(finMod$fitted,finMod$residuals,colour=race,data=training)  ##also fitted vs. residuals. Trying to
```



Predicting with Trees

Better performance in non-linear settings Can cause over-fitting, so be careful Considered non-linear models, so they use interactions between variables

1.) start with all variables in one group 2.) Find the variable/split that best separates the outcomes 3.) Divide the data into two groups "leaves" on that node 4.) Within each split, find the best variable/split that separates the outcomes 5.) Continue until the groups are too small or sufficiently "pure"

```
data(iris); library(ggplot2)
names(iris)
```

```
## [1] "Sepal.Length" "Sepal.Width"  "Petal.Length" "Petal.Width"  "Species"
```

```
table(iris$Species)
```

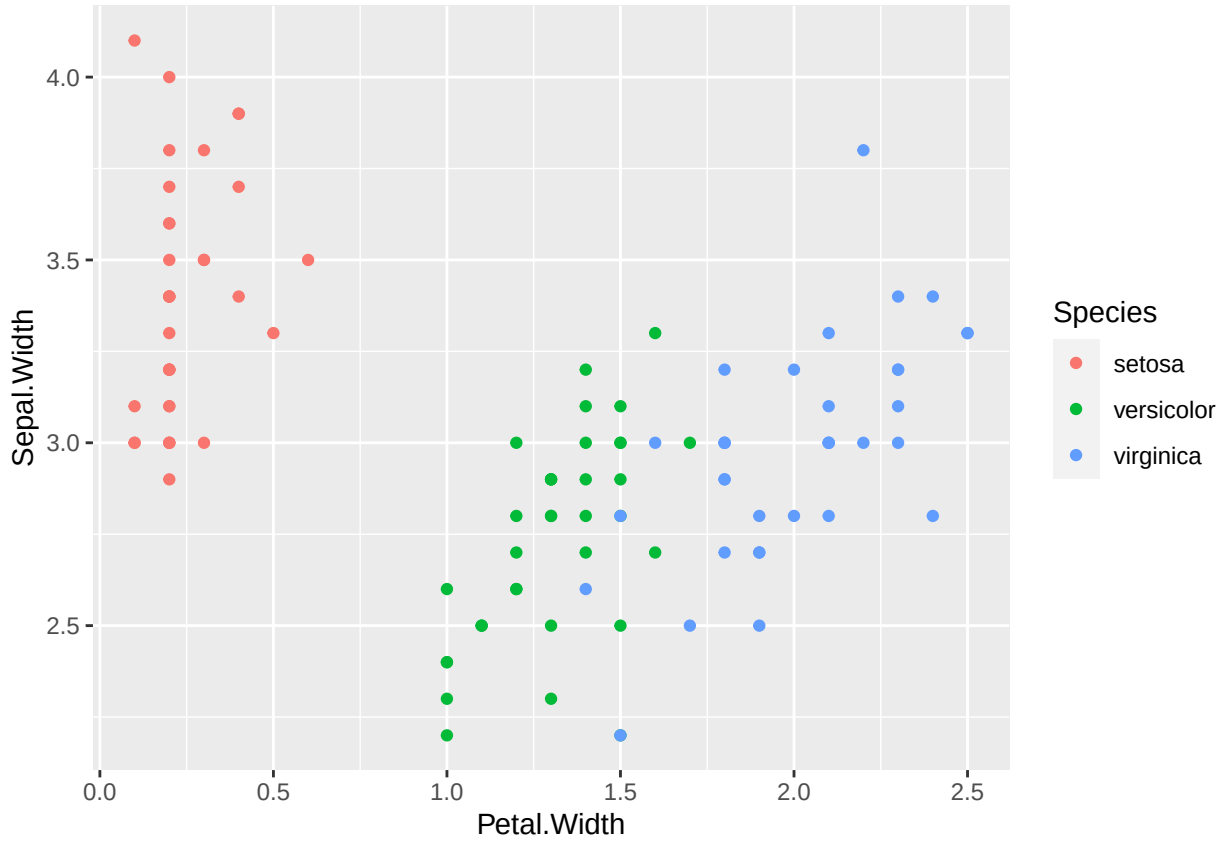
```
##
##     setosa versicolor  virginica
##         50         50         50
```

```
inTrain <- createDataPartition(y=iris$Species,
                               p=0.7, list=FALSE)
training <- iris[inTrain,]
testing <- iris[-inTrain,]
dim(training); dim(testing)
```

```
## [1] 105   5
```

```
## [1] 45  5
```

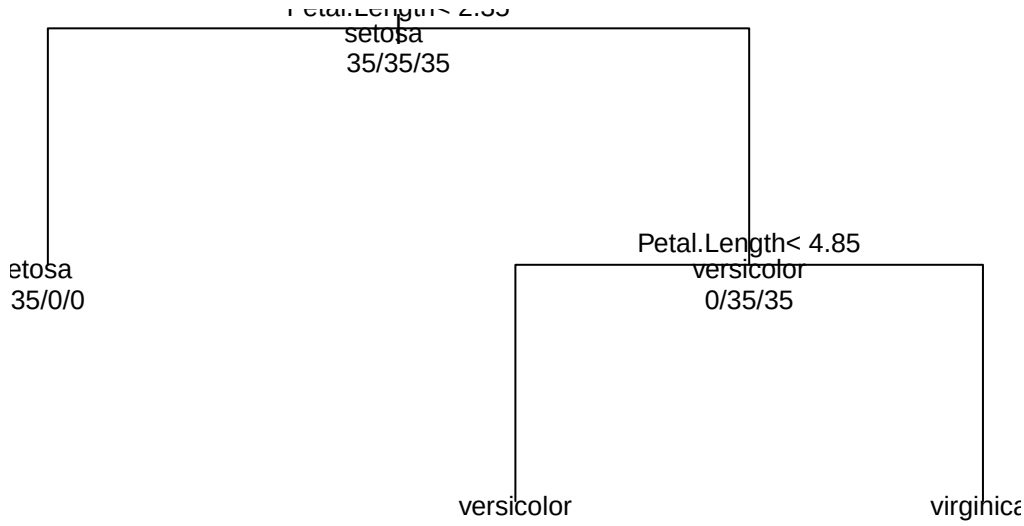
```
qplot(Petal.Width,Sepal.Width,colour=Species,data=training)
```




```
library(caret)
modFit <- train(Species ~ .,method="rpart",data=training) ##train function fits the model here. looking
print(modFit$finalModel)
```

```
## n= 105
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
## 1) root 105 70 setosa (0.33333333 0.33333333 0.33333333)
##   2) Petal.Length< 2.35 35  0 setosa (1.00000000 0.00000000 0.00000000) *
##   3) Petal.Length>=2.35 70 35 versicolor (0.00000000 0.50000000 0.50000000)
##     6) Petal.Length< 4.85 33  1 versicolor (0.00000000 0.96969697 0.03030303) *
##     7) Petal.Length>=4.85 37  3 virginica (0.00000000 0.08108108 0.91891892) *
```

```
plot(modFit$finalModel, uniform=TRUE,
      main="Classification Tree")
text(modFit$finalModel, use.n=TRUE, all=TRUE, cex=.8)  ##plotting the dendogram tree
```



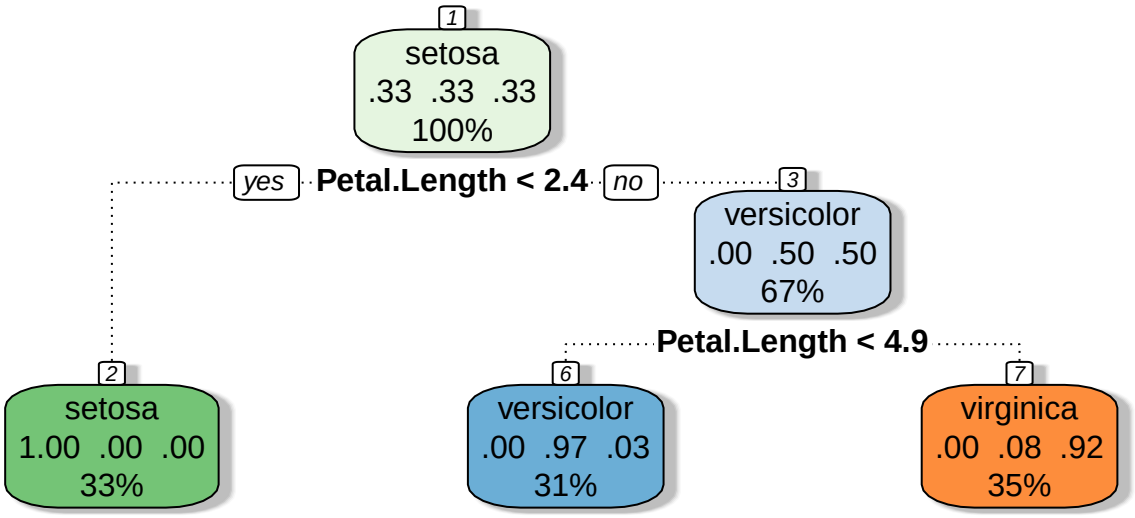

```
library(rattle)
```

```
## Loading required package: tibble
```

```
## Loading required package: bitops
```

```
## Rattle: A free graphical interface for data science with R.
## Version 5.5.1 Copyright (c) 2006-2021 Togaware Pty Ltd.
## Type 'rattle()' to shake, rattle, and roll your data.
```

```
fancyRpartPlot(modFit$finalModel)  ##makes a fancier version of the same dendogram
```



Bootstrap-aggregating "Bagging"

Re-sample the data, and recalculate predictions, then average or majority vote. Useful for non-linear functions When using bagging in caret, consider using "bagEarth", "treebag", "bagFDA" inside of the train function.

Below is the syntax/code, but ozone dataset not available predictors = data.frame(ozone=ozone$ozone$)temperature = $ozone$temperature treebag <- bag(predictors, temperature, B = 10, bagControl = bagControl(fit = ctreeBag$fit, predict = ctreeBag$pred, aggregate = ctreeBag$aggregate))

Random Forests

1.) Bootstrap samples 2.) at each split, bootstrap the variables again 3.) Grow multiple trees and vote Cons- prone to overfitting, but it is very accurate

```
data(iris); library(ggplot2); library(randomForest); library(caret)
```

```
## randomForest 4.7-1.1
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:rattle':
##
##     importance

## The following object is masked from 'package:dplyr':
##
##     combine

## The following object is masked from 'package:ggplot2':
##
##     margin
```

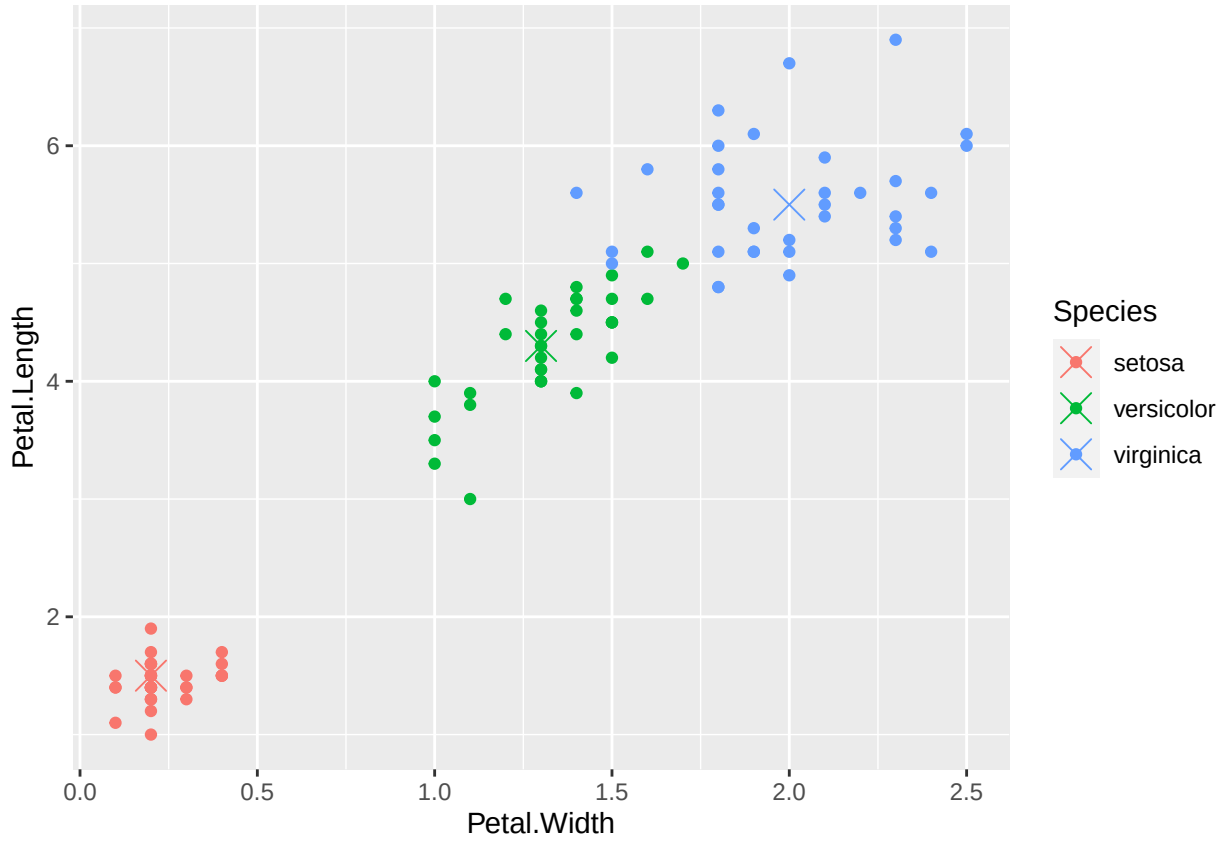
```
inTrain <- createDataPartition(y=iris$Species,
                               p=0.7, list=FALSE)
training <- iris[inTrain,]
testing <- iris[-inTrain,]


modFit <- train(Species~ .,data=training,method="rf",prox=TRUE)  ##using train, here method is "rf", wh
modFit
```

```
## Random Forest
##
## 105 samples
##   4 predictor
##   3 classes: 'setosa', 'versicolor', 'virginica'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 105, 105, 105, 105, 105, 105, ...
## Resampling results across tuning parameters:
##
##   mtry  Accuracy   Kappa
##   2     0.9497883  0.9240399
##   3     0.9508409  0.9256164
##   4     0.9508994  0.9256779
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was mtry = 4.
```

```
getTree(modFit$finalModel,k=2) ##to view the individual trees, you can use this function, which is look
```

```
##   left daughter right daughter split var split point status prediction
## 1             2              3         3        2.70      1          0
## 2             0              0         0        0.00     -1          1
## 3             4              5         4        1.75      1          0
## 4             6              7         3        5.35      1          0
## 5             0              0         0        0.00     -1          3
## 6             0              0         0        0.00     -1          2
## 7             0              0         0        0.00     -1          3
```

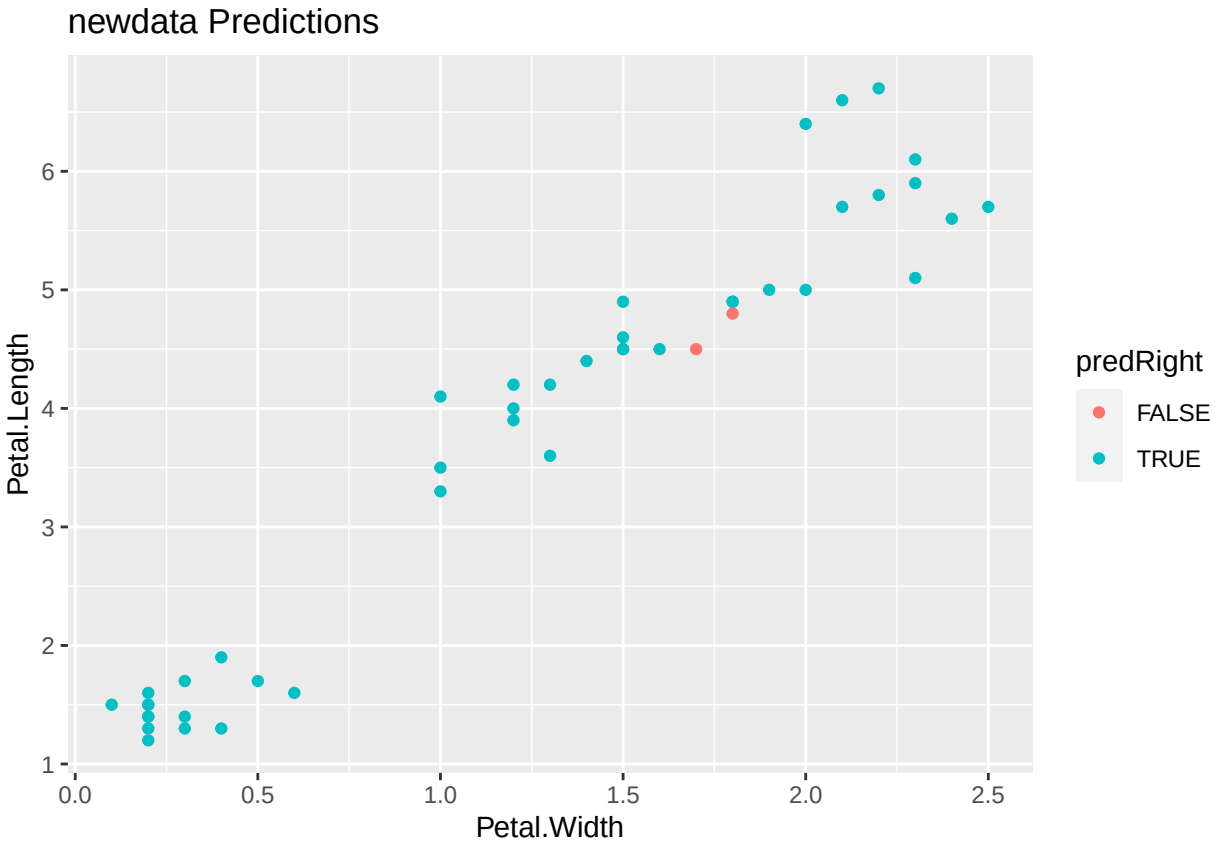
```
irisP <- classCenter(training[,c(3,4)], training$Species, modFit$finalModel$prox)  ##This centers your
irisP <- as.data.frame(irisP); irisP$Species <- rownames(irisP)
p <- qplot(Petal.Width, Petal.Length, col=Species,data=training)
p + geom_point(aes(x=Petal.Width,y=Petal.Length,col=Species),size=5,shape=4,data=irisP)
```




```
pred <- predict(modFit,testing); testing$predRight <- pred==testing$Species
table(pred,testing$Species)
```

```
##
## pred         setosa versicolor virginica
##   setosa         15          0         0
##   versicolor      0         14         1
##   virginica       0          1        14
```

```
qplot(Petal.Width,Petal.Length,colour=predRight,data=testing,main="newdata Predictions")
```



Boosting

1.) Takes a lot of weak predictors 2.) Weight them and add them up 3.) Get a stronger predictor

```
library(ISLR); data(Wage); library(ggplot2); library(caret);
Wage <- subset(Wage,select=-c(logwage))  ##removes the variable were interested in
inTrain <- createDataPartition(y=Wage$wage,
                               p=0.7, list=FALSE)
training <- Wage[inTrain,]; testing <- Wage[-inTrain,]

modFit <- train(wage ~ ., method="gbm",data=training,verbose=FALSE)  ## "gbm" is boosting with trees us
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.
```

```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.
```

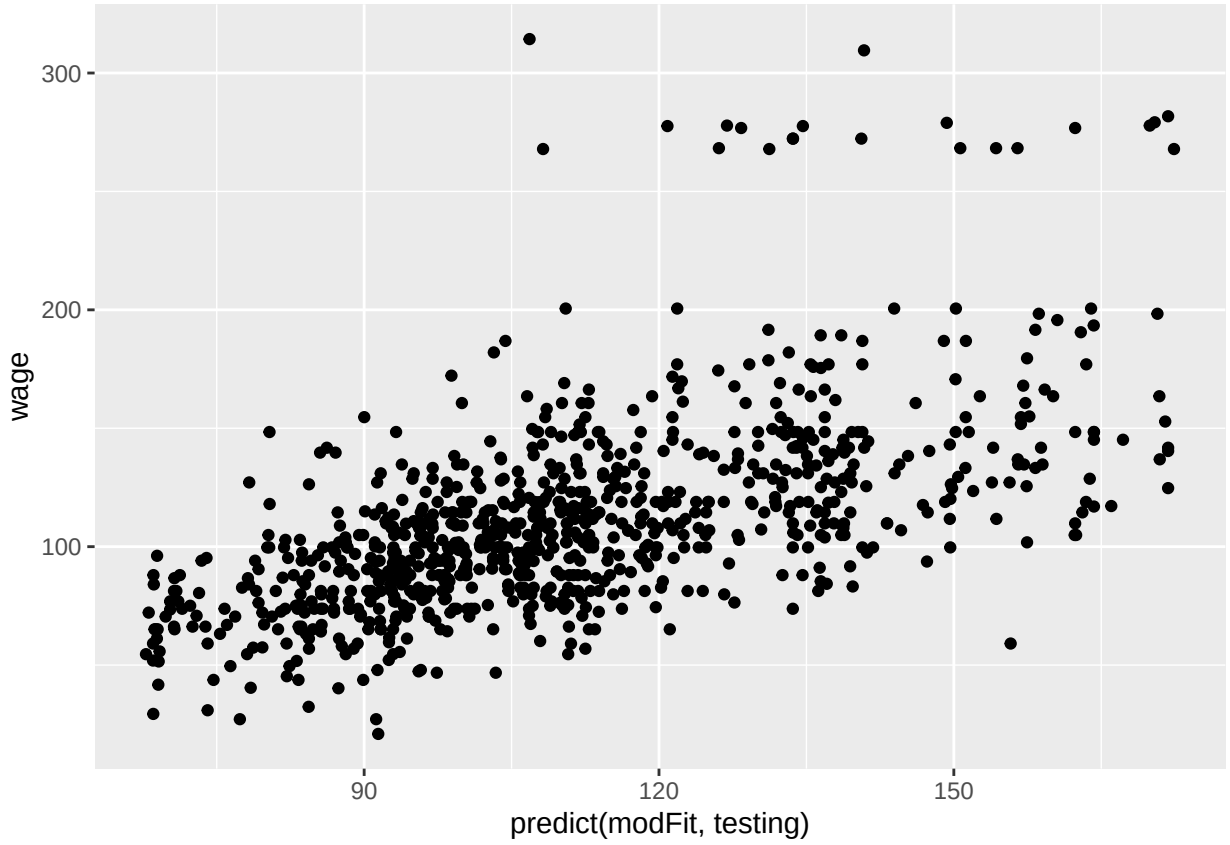
```
## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 14: region2. Middle Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 15: region3. East North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 16: region4. West North Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 17: region5. South Atlantic has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 18: region6. East South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 19: region7. West South Central has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 20: region8. Mountain has no variation.

## Warning in (function (x, y, offset = NULL, misc = NULL, distribution =
## "bernoulli", : variable 21: region9. Pacific has no variation.
```

```
print(modFit)
```

```
## Stochastic Gradient Boosting
##
## 2102 samples
##    9 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 2102, 2102, 2102, 2102, 2102, 2102, ...
## Resampling results across tuning parameters:
##
##   interaction.depth  n.trees  RMSE      Rsquared   MAE
##   1                   50      35.65470  0.3067370  24.18107
##   1                  100      35.08875  0.3164737  23.77081
##   1                  150      35.00625  0.3181686  23.78886
##   2                   50      34.99228  0.3216033  23.70188
##   2                  100      34.86318  0.3238252  23.70463
##   2                  150      34.91481  0.3222133  23.80882
##   3                   50      34.84892  0.3253748  23.62091
##   3                  100      34.95760  0.3212939  23.80812
##   3                  150      35.11354  0.3165658  23.99817
##
## Tuning parameter 'shrinkage' was held constant at a value of 0.1
```

```
##
## Tuning parameter 'n.minobsinnode' was held constant at a value of 10
## RMSE was used to select the optimal model using the smallest value.
## The final values used for the model were n.trees = 50, interaction.depth =
##  3, shrinkage = 0.1 and n.minobsinnode = 10.
```

```
qplot(predict(modFit,testing),wage,data=testing)
```



Practical Example 1

1. Subset the data to a training set and testing set based on the Case variable in the data set.
2. Set the seed to 125 and fit a caret model with the rpart method using all predictor variables and default caret settings.
3. In the final model what would be the final model prediction for cases with the following variable values:

a. TotalIntench2 = 23,000; FiberWidthCh1 = 10; PerimStatusCh1 = 2.

b. TotalIntench2 = 50,000; FiberWidthCh1 = 10; VarIntenCh4 = 100.

c. TotalIntench2 = 57,000; FiberWidthCh1 = 8; VarIntenCh4 = 100.

d. FiberWidthCh1 = 8; VarIntenCh4 = 100; PerimStatusCh1 = 2.

```
library(rattle)
library(AppliedPredictiveModeling)
data(segmentationOriginal)
seg1 <- segmentationOriginal
head(seg1, n=20)
```

```
##         Cell  Case Class   AngleCh1 AngleStatusCh1 AreaCh1 AreaStatusCh1
## 1  207827637  Test    PS 143.247705              1     185             0
## 2  207932307 Train    PS 133.752037              0     819             1
## 3  207932463 Train    WS 106.646387              0     431             0
## 4  207932470 Train    PS  69.150325              0     298             0
## 5  207932455  Test    PS   2.887837              2     285             0
## 6  207827656  Test    WS  40.748298              2     172             0
## 7  207827659  Test    WS 173.957833              1     177             0
## 8  207827661  Test    PS 179.800467              1     251             0
## 9  207932479  Test    WS  18.936420              2     495             0
## 10 207932480  Test    WS 153.252427              1     384             0
## 11 207827711  Test    WS  13.666671              2     424             0
## 12 207932484 Train    WS 109.416426              0     256             0
## 13 207827674  Test    PS  52.205447              2     236             0
## 14 207932302  Test    PS  92.901464              0     187             0
## 15 207932459 Train    PS 104.278654              0     258             0
## 16 207827779 Train    PS  77.991943              0     358             0
## 17 207827784 Train    PS  13.659972              2     158             0
## 18 207827790  Test    WS  23.071066              2     313             0
## 19 207827645 Train    WS 106.844369              0     315             0
## 20 207932306  Test    PS  93.591140              0     762             1
##    AvgIntenCh1 AvgIntenCh2 AvgIntenCh3 AvgIntenCh4 AvgIntenStatusCh1
## 1     15.71186    3.954802    9.548023    2.214689                 0
## 2     31.92327  205.878517   69.916880  164.153453                 0
## 3     28.03883  115.315534   63.941748  106.696602                 0
## 4     19.45614  101.294737   28.217544   31.028070                 0
## 5     24.27574  111.415441   20.474265   40.577206                 0
## 6    325.93333  652.606061  128.690909  346.854545                 1
## 7    260.05917  595.420118  123.840237  273.295858                 1
## 8     18.33750    4.729167   17.233333    1.554167                 0
## 9     16.05285   88.482030   13.716702   51.427061                 0
## 10    17.74387   88.931880   20.444142   63.111717                 0
## 11   174.35556  388.427160   38.762963  126.728395                 1
## 12    18.82857  125.938776   13.600000   46.800000                 0
## 13    18.19027    5.327434   17.088496    1.911504                 0
## 14    40.24022  213.189944   44.195531  136.251397                 0
## 15    17.57085  124.368421   22.461538   71.206478                 0
## 16    42.28363  217.131579   42.321637   67.540936                 0
## 17    31.41060  102.211921   41.490066  116.119205                 0
## 18   215.41137  568.913043  150.652174  378.612040                 1
## 19   294.76744  491.734219  193.435216  400.451827                 1
## 20    35.64099  163.719395   77.522696  213.110041                 0
##    AvgIntenStatusCh2 AvgIntenStatusCh3 AvgIntenStatusCh4 ConvexHullAreaRatioCh1
## 1                  2                 2                 2               1.124509
## 2                  0                 0                 0               1.263158
## 3                  0                 0                 0               1.053310
```

```
## 4                   0                   0                   2                1.202625
## 5                   0                   0                   0                1.109333
## 6                   1                   1                   1                1.005831
## 7                   1                   1                   0                1.014164
## 8                   2                   0                   2                1.200389
## 9                   0                   0                   0                1.192308
## 10                  0                   0                   0                1.160365
## 11                  1                   0                   0                1.038462
## 12                  0                   0                   0                1.081871
## 13                  2                   0                   2                1.292658
## 14                  0                   0                   0                1.035230
## 15                  0                   0                   0                1.075728
## 16                  0                   0                   0                1.037830
## 17                  0                   0                   0                1.379310
## 18                  1                   1                   1                1.040258
## 19                  1                   1                   1                1.014516
## 20                  0                   0                   0                1.202662
##    ConvexHullAreaRatioStatusCh1 ConvexHullPerimRatioCh1
## 1                             0                0.9196827
## 2                             1                0.7970801
## 3                             0                0.9354750
## 4                             0                0.8658291
## 5                             0                0.9568124
## 6                             0                0.9926363
## 7                             0                0.9840211
## 8                             0                0.8310554
## 9                             0                0.8219601
## 10                            0                0.8649792
## 11                            0                0.9725156
## 12                            0                0.9204595
## 13                            1                0.8013622
## 14                            0                0.9796029
## 15                            0                0.9313002
## 16                            0                0.9612698
## 17                            1                0.8960387
## 18                            0                0.9576618
## 19                            0                0.9775437
## 20                            0                0.8487132
##    ConvexHullPerimRatioStatusCh1 DiffIntenDensityCh1 DiffIntenDensityCh3
## 1                              0            29.51923            13.77564
## 2                              2            31.87500            43.12228
## 3                              0            32.48771            35.98577
## 4                              2            26.73228            22.91732
## 5                              0            31.58065            21.70968
## 6                              1            92.56291            61.92715
## 7                              1            89.22654            47.81553
## 8                              2            28.61792            19.22170
## 9                              2            26.98404            15.06613
## 10                             2            25.76036            18.23077
## 11                             0            67.25131            27.18848
## 12                             0            28.02004            14.87082
## 13                             2            28.52041            16.91327
## 14                             0            32.11585            25.60976
## 15                             0            27.87611            16.11947
```

```
## 16                          0           43.30938           43.36563
## 17                          0           38.45038           37.64885
## 18                          0           73.87410           78.32014
## 19                          0          154.57801          100.31560
## 20                          2           35.00584           54.41065
##    DiffIntenDensityCh4 DiffIntenDensityStatusCh1 DiffIntenDensityStatusCh3
## 1             6.826923                         2                         2
## 2            79.308424                         0                         0
## 3            51.357050                         0                         0
## 4            26.393701                         2                         0
## 5            25.032258                         0                         0
## 6           145.708609                         1                         1
## 7           139.291262                         1                         0
## 8             5.382075                         2                         0
## 9            36.117446                         2                         2
## 10           38.059172                         2                         0
## 11           84.141361                         0                         0
## 12           32.697105                         2                         2
## 13            5.918367                         2                         0
## 14           90.274390                         0                         0
## 15           36.185841                         2                         0
## 16           56.043750                         0                         0
## 17           74.923664                         0                         0
## 18          191.553957                         1                         1
## 19          237.407801                         1                         1
## 20          121.671043                         0                         0
##    DiffIntenDensityStatusCh4 EntropyIntenCh1 EntropyIntenCh3 EntropyIntenCh4
## 1                          2        4.969781        4.371017        2.718884
## 2                          0        6.087592        6.642761        7.880155
## 3                          0        5.883557        6.683000        7.144601
## 4                          2        5.420065        5.436732        5.778329
## 5                          2        5.658248        5.285671        5.235857
## 6                          1        6.998111        6.805538        7.119323
## 7                          1        6.896239        6.660887        7.032449
## 8                          2        5.318070        5.085643        2.295804
## 9                          0        5.189203        4.493851        6.281189
## 10                         0        5.334973        5.320425        6.740386
## 11                         0        7.383121        5.247167        5.997840
## 12                         2        5.383272        4.151183        6.194035
## 13                         2        5.293391        4.993778        2.423698
## 14                         0        5.929336        6.055303        6.821641
## 15                         0        5.184278        5.486225        6.624639
## 16                         0        6.136583        5.870027        5.995699
## 17                         0        5.756261        5.814347        6.121377
## 18                         1        7.468859        6.813520        7.540135
## 19                         1        7.402635        6.761025        7.560942
## 20                         0        6.249204        7.027566        7.240787
##    EntropyIntenStatusCh1 EntropyIntenStatusCh3 EntropyIntenStatusCh4
## 1                      2                     0                     2
## 2                      0                     1                     1
## 3                      0                     1                     0
## 4                      2                     0                     0
## 5                      2                     0                     0
## 6                      0                     1                     0
```

```
## 7                    0                    1                    0
## 8                    2                    0                    2
## 9                    2                    0                    0
## 10                   2                    0                    0
## 11                   1                    0                    0
## 12                   2                    0                    0
## 13                   2                    0                    2
## 14                   0                    0                    0
## 15                   2                    0                    0
## 16                   0                    0                    0
## 17                   2                    0                    0
## 18                   1                    1                    1
## 19                   1                    1                    1
## 20                   0                    1                    0
##    EqCircDiamCh1 EqCircDiamStatusCh1 EqEllipseLWRCh1 EqEllipseLWRStatusCh1
## 1       15.36954                   0        3.060676                     1
## 2       32.30558                   1        1.558394                     0
## 3       23.44892                   0        1.375386                     0
## 4       19.50279                   0        3.391220                     1
## 5       19.05279                   0        2.741380                     0
## 6       14.83940                   2        1.039726                     0
## 7       15.01819                   2        1.088583                     0
## 8       17.89698                   0        2.866415                     0
## 9       25.12493                   0        1.506684                     0
## 10      22.13132                   0        1.755000                     0
## 11      23.24886                   0        1.465172                     0
## 12      18.08245                   0        1.378748                     0
## 13      17.36715                   0        3.378319                     1
## 14      15.45613                   0        1.535489                     0
## 15      18.15611                   0        1.616053                     0
## 16      21.36423                   0        1.468210                     0
## 17      14.19589                   2        1.477694                     0
## 18      19.97606                   0        1.762458                     0
## 19      20.04276                   0        1.016647                     0
## 20      31.14881                   1        1.871580                     0
##    EqEllipseOblateVolCh1 EqEllipseOblateVolStatusCh1 EqEllipseProlateVolCh1
## 1               336.9691                           0                110.0963
## 2              2232.9055                           1               1432.8246
## 3               802.1945                           0                583.2504
## 4               724.7143                           0                213.7031
## 5               607.5185                           0                221.6105
## 6               176.7693                           0                170.0153
## 7               187.4918                           0                172.2348
## 8               514.8823                           0                179.6259
## 9              1032.8193                           0                685.4917
## 10              761.8320                           0                434.0923
## 11              806.9533                           0                550.7567
## 12              368.3097                           0                267.1334
## 13              510.7808                           0                151.1938
## 14              242.7306                           0                158.0803
## 15              403.6405                           0                249.7694
## 16              626.8373                           0                426.9397
## 17              184.4927                           0                124.8518
## 18              561.4198                           0                318.5436
```

```
## 19              430.6817                          0               423.6295
## 20             2193.4483                          1              1171.9766
##    EqEllipseProlateVolStatusCh1 EqSphereAreaCh1 EqSphereAreaStatusCh1
## 1                             0        742.1156                     0
## 2                             1       3278.7256                     1
## 3                             0       1727.4104                     0
## 4                             0       1194.9320                     0
## 5                             0       1140.4263                     0
## 6                             0        691.8027                     0
## 7                             0        708.5737                     0
## 8                             0       1006.2585                     0
## 9                             0       1983.1678                     0
## 10                            0       1538.7370                     0
## 11                            0       1698.0612                     0
## 12                            0       1027.2222                     0
## 13                            0        947.5601                     0
## 14                            0        750.5011                     0
## 15                            0       1035.6077                     0
## 16                            0       1433.9184                     0
## 17                            0        633.1043                     0
## 18                            0       1253.6304                     0
## 19                            0       1262.0159                     0
## 20                            1       3048.1247                     1
##    EqSphereVolCh1 EqSphereVolStatusCh1 FiberAlign2Ch3 FiberAlign2Ch4
## 1        1900.996                    0      0.0000000     0.00000000
## 2       17653.525                    1      0.4879354     0.35237425
## 3        6750.985                    0      0.3005220     0.52231582
## 4        3884.084                    0      0.2204239     0.73325044
## 5        3621.385                    0      0.4913869     0.38454715
## 6        1710.989                    0      0.3095039     0.36219126
## 7        1773.582                    0      0.2449735     0.54522523
## 8        3001.499                    0      0.0000000     0.00000000
## 9        8304.491                    0      0.3552463     0.45627167
## 10       5675.713                    0      0.7624199     0.65956639
## 11       6579.666                    0      0.5247012     0.53992350
## 12       3095.782                    0      0.3642360     0.48086857
## 13       2742.736                    0      0.0000000     0.00000000
## 14       1933.307                    0      0.6259856     0.73274592
## 15       3133.767                    0      0.3594184     0.24430781
## 16       5105.761                    0      0.4789638     0.46657686
## 17       1497.913                    0      0.7245099     0.59720458
## 18       4173.766                    0      0.4876284     0.46176931
## 19       4215.713                    0      0.2735832     0.01897489
## 20      15824.242                    1      0.4549632     0.41503726
##    FiberAlign2StatusCh3 FiberAlign2StatusCh4 FiberLengthCh1
## 1                     2                    2       26.98132
## 2                     0                    0       64.28230
## 3                     0                    0       21.14115
## 4                     0                    1       43.14112
## 5                     0                    0       34.74977
## 6                     0                    0       12.38420
## 7                     0                    0       12.59624
## 8                     2                    2       37.84863
## 9                     0                    0       44.14773
```

```
## 10                    1                    1       37.24447
## 11                    0                    0       20.24908
## 12                    0                    0       22.29743
## 13                    2                    2       43.64276
## 14                    1                    1       13.70788
## 15                    0                    0       26.52856
## 16                    0                    0       25.12696
## 17                    1                    1       27.63249
## 18                    0                    0       18.68085
## 19                    0                    2       16.75358
## 20                    0                    0       55.03983
##    FiberLengthStatusCh1 FiberWidthCh1 FiberWidthStatusCh1 IntenCoocASMCh3
## 1                     0      7.410365                   2     0.011183899
## 2                     1     13.167079                   0     0.028051061
## 3                     0     21.141150                   1     0.006862315
## 4                     0      7.404412                   2     0.030962071
## 5                     0      8.483574                   2     0.022768411
## 6                     0     12.384199                   0     0.008914083
## 7                     0     12.596237                   0     0.008166438
## 8                     0      7.117398                   2     0.006116300
## 9                     0     11.728889                   0     0.089600535
## 10                    0     10.792994                   0     0.012277917
## 11                    0     20.249084                   1     0.095391688
## 12                    0     12.057876                   0     0.108161298
## 13                    0      5.806209                   2     0.005475966
## 14                    0     13.707875                   0     0.009545651
## 15                    0     10.174239                   0     0.013034881
## 16                    0     14.610886                   0     0.024982639
## 17                    0      6.050330                   2     0.012319717
## 18                    0     17.422216                   1     0.022057270
## 19                    0     16.753579                   1     0.033156657
## 20                    1     14.306907                   0     0.013880529
##    IntenCoocASMCh4 IntenCoocASMStatusCh3 IntenCoocASMStatusCh4
## 1      0.050448005                     0                     0
## 2      0.012594975                     0                     0
## 3      0.006141691                     0                     0
## 4      0.011033195                     0                     0
## 5      0.079690842                     0                     0
## 6      0.009155300                     0                     0
## 7      0.008311984                     0                     0
## 8      0.127326686                     0                     0
## 9      0.019221940                     0                     0
## 10     0.008028741                     0                     0
## 11     0.093600421                     0                     0
## 12     0.009951830                     0                     0
## 13     0.103488064                     0                     0
## 14     0.008644145                     0                     0
## 15     0.008962990                     0                     0
## 16     0.044225099                     0                     0
## 17     0.020887482                     0                     0
## 18     0.010350158                     0                     0
## 19     0.009490846                     0                     0
## 20     0.038698122                     0                     0
##    IntenCoocContrastCh3 IntenCoocContrastCh4 IntenCoocContrastStatusCh3
```

```
## 1             40.751777           13.895439                        1
## 2              8.227953            6.984046                        0
## 3             14.446074           16.700843                        0
## 4              7.299457           13.390884                        0
## 5             15.854691            3.538895                        0
## 6              8.248344            9.867550                        0
## 7             12.881127           17.409342                        0
## 8             32.702178            8.833734                        1
## 9              9.943647            9.568663                        0
## 10            18.139723           12.330637                        1
## 11             4.302937            3.941435                        0
## 12             6.160585           10.590685                        0
## 13            38.771563           11.704375                        1
## 14            16.421096           16.470915                        1
## 15             9.402056           10.301528                        0
## 16            11.841077            8.084776                        0
## 17            11.913578            6.154895                        0
## 18             4.818421            5.870657                        0
## 19             5.696809            8.225177                        0
## 20             7.776731            3.789387                        0
##    IntenCoocContrastStatusCh4 IntenCoocEntropyCh3 IntenCoocEntropyCh4
## 1                           1           7.199458           5.249744
## 2                           0           6.822138           7.098988
## 3                           1           7.580100           7.671478
## 4                           1           6.312641           7.197026
## 5                           2           6.778093           5.502166
## 6                           0           7.016333           6.967390
## 7                           1           7.172120           7.142348
## 8                           0           7.557944           4.177876
## 9                           0           5.711782           6.960814
## 10                          1           7.309512           7.406901
## 11                          2           5.518710           5.550419
## 12                          0           5.044067           7.126900
## 13                          1           7.739265           4.512367
## 14                          1           7.236606           7.142452
## 15                          0           6.957487           7.144784
## 16                          0           6.794864           6.361635
## 17                          0           6.796920           6.485465
## 18                          0           6.546452           7.087970
## 19                          0           6.390147           7.161842
## 20                          2           7.096212           6.233320
##    IntenCoocEntropyStatusCh3 IntenCoocEntropyStatusCh4 IntenCoocMaxCh3
## 1                          0                         0      0.07741935
## 2                          0                         0      0.15321477
## 3                          1                         1      0.02835052
## 4                          0                         0      0.16279070
## 5                          0                         0      0.12741313
## 6                          0                         0      0.01986755
## 7                          0                         0      0.01948052
## 8                          1                         2      0.01643192
## 9                          0                         0      0.29519451
## 10                         1                         1      0.08630952
## 11                         0                         0      0.30079156
## 12                         0                         0      0.31531532
```

```
## 13                        1                        2      0.02083333
## 14                        0                        0      0.06586826
## 15                        0                        0      0.07391304
## 16                        0                        0      0.14012739
## 17                        0                        0      0.06818182
## 18                        0                        0      0.12727273
## 19                        0                        0      0.16312057
## 20                        0                        0      0.08333333
##    IntenCoocMaxCh4 IntenCoocMaxStatusCh3 IntenCoocMaxStatusCh4 KurtIntenCh1
## 1       0.17197452                     0                     0 -0.656744087
## 2       0.07387141                     0                     0 -0.248769067
## 3       0.02319588                     2                     2 -0.293484630
## 4       0.07751938                     0                     0  0.625856124
## 5       0.27848101                     0                     0  0.042106843
## 6       0.01986755                     2                     2 -0.196841369
## 7       0.01948052                     2                     2  0.250897694
## 8       0.34123223                     2                     1  0.112969507
## 9       0.11212815                     0                     0 -0.219208563
## 10      0.04705882                     0                     0  0.144554913
## 11      0.30079156                     0                     0 -0.625392634
## 12      0.05855856                     0                     0 -0.364691438
## 13      0.29500000                     2                     0 -0.001236013
## 14      0.03592814                     0                     0 -0.930684432
## 15      0.03478261                     0                     0 -0.556261231
## 16      0.19938650                     0                     0 -0.261787171
## 17      0.11363636                     0                     0 -0.365908061
## 18      0.05454545                     0                     0 -0.370863990
## 19      0.04609929                     0                     0 38.974284980
## 20      0.17903930                     0                     0  0.230696287
##    KurtIntenCh3 KurtIntenCh4 KurtIntenStatusCh1 KurtIntenStatusCh3
## 1   -0.60805827    0.7258145                  0                  0
## 2   -0.33078390   -0.2652638                  0                  0
## 3    1.05128134    0.1506140                  0                  0
## 4    0.12774061   -0.3472936                  0                  0
## 5    0.95229003   -0.1954188                  0                  0
## 6   -0.86079937   -0.4050669                  0                  0
## 7   -0.59831464   -0.8256696                  0                  0
## 8   -0.74863969    2.0335593                  0                  0
## 9    1.88482515   -0.7610117                  0                  0
## 10  -0.56811998    0.1918431                  0                  0
## 11  -0.35835018   -0.5315540                  0                  0
## 12   1.08333286   -0.6262704                  0                  0
## 13  -0.09820689    0.2876581                  0                  0
## 14  -0.58857201    6.7741213                  0                  0
## 15  -0.51231861   -0.6471575                  0                  0
## 16   0.80950143    0.7286793                  0                  0
## 17  -0.27882467   -0.7073985                  0                  0
## 18  -0.42094795   -0.4443304                  0                  0
## 19  -1.19024918   -1.3328752                  1                  0
## 20   5.19308674    2.7868190                  0                  1
##    KurtIntenStatusCh4 LengthCh1 LengthStatusCh1 MemberAvgAvgIntenStatusCh2
## 1                   0  26.20779                0                         0
## 2                   0  47.21855                1                         0
## 3                   0  28.14303                0                         0
```

```
## 4                   0  37.85957               0                           0
## 5                   0  35.99415               0                           0
## 6                   0  16.20494               2                           0
## 7                   0  15.91843               2                           0
## 8                   0  29.70812               0                           0
## 9                   0  32.37446               0                           0
## 10                  0  31.13589               0                           0
## 11                  0  28.28436               0                           0
## 12                  0  23.05547               0                           0
## 13                  0  34.75543               0                           0
## 14                  1  18.71589               0                           0
## 15                  0  26.31960               0                           0
## 16                  0  27.80421               0                           0
## 17                  0  20.56691               0                           0
## 18                  0  26.61822               0                           0
## 19                  0  21.28784               0                           0
## 20                  0  43.93737               1                           0
##    MemberAvgTotalIntenStatusCh2 NeighborAvgDistCh1 NeighborAvgDistStatusCh1
## 1                             0           370.4543                         1
## 2                             0           174.4442                         2
## 3                             0           158.4774                         2
## 4                             0           206.3344                         0
## 5                             0           204.9213                         0
## 6                             0           231.9933                         0
## 7                             0           230.2935                         0
## 8                             0           218.2644                         0
## 9                             0           235.0062                         0
## 10                            0           220.1542                         0
## 11                            0           334.7284                         1
## 12                            0           263.6345                         0
## 13                            0           204.2607                         0
## 14                            0           190.9553                         2
## 15                            0           231.1529                         0
## 16                            0           213.9487                         0
## 17                            0           228.9701                         0
## 18                            0           314.1311                         1
## 19                            0           288.1973                         0
## 20                            0           213.4731                         0
##    NeighborMinDistCh1 NeighborMinDistStatusCh1 NeighborVarDistCh1
## 1            99.10349                        1          127.96080
## 2            30.11114                        0           81.38063
## 3            34.94477                        0           90.43768
## 4            33.08030                        0          116.89276
## 5            27.02935                        0          110.96662
## 6            15.92674                        0           95.96952
## 7            15.92674                        0           94.81307
## 8            29.31744                        0          105.76881
## 9            38.43038                        0           79.58161
## 10           36.18265                        0           73.85064
## 11           28.66667                        0          130.46044
## 12           38.43038                        0           88.50411
## 13           29.31744                        0           78.31229
## 14           23.90107                        0           75.44312
## 15           29.84892                        0          103.45812
```

```
## 16           11.44654                       2          105.15059
## 17           17.40476                       0          112.56123
## 18           22.73226                       0          126.52233
## 19           49.92581                       1          119.63999
## 20           46.97267                       1          127.92813
##    NeighborVarDistStatusCh1  PerimCh1 PerimStatusCh1 ShapeBFRCh1
## 1                         0  68.78338              0   0.6651480
## 2                         2 154.89876              1   0.5397584
## 3                         2  84.56460              0   0.7243116
## 4                         0 101.09107              0   0.5891625
## 5                         0  86.46670              0   0.6001347
## 6                         0  49.53679              0   0.6869375
## 7                         0  50.38495              0   0.7468435
## 8                         0  89.93206              0   0.6858659
## 9                         2 111.75324              0   0.6318210
## 10                        2  96.07492              0   0.6738027
## 11                        0  80.99633              0   0.7574574
## 12                        2  68.71062              0   0.6347914
## 13                        2  98.89793              0   0.5453331
## 14                        2  54.83150              0   0.7968360
## 15                        0  73.40559              0   0.5570022
## 16                        0  79.47569              0   0.6635424
## 17                        0  67.36563              0   0.4423808
## 18                        0  72.20612              0   0.7609989
## 19                        0  67.01432              0   0.7161740
## 20                        0 138.69347              1   0.6034935
##    ShapeBFRStatusCh1 ShapeLWRCh1 ShapeLWRStatusCh1 ShapeP2ACh1
## 1                  0    2.462450                 0    1.883006
## 2                  2    1.468181                 0    2.255810
## 3                  1    1.328408                 0    1.272193
## 4                  0    2.826854                 1    2.545840
## 5                  0    2.727127                 1    2.018155
## 6                  0    1.043013                 0    1.086278
## 7                  1    1.068329                 0    1.091958
## 8                  0    2.406246                 0    2.389160
## 9                  0    1.335671                 0    1.919299
## 10                 0    1.698051                 0    1.827275
## 11                 1    1.427439                 0    1.197255
## 12                 0    1.313937                 0    1.397364
## 13                 2    2.780739                 1    3.071547
## 14                 1    1.487642                 0    1.237124
## 15                 2    1.490323                 0    1.588658
## 16                 0    1.430953                 0    1.369114
## 17                 2    1.182277                 0    2.160058
## 18                 1    1.720413                 0    1.274781
## 19                 0    1.028672                 0    1.099823
## 20                 0    1.528861                 0    1.943913
##    ShapeP2AStatusCh1 SkewIntenCh1 SkewIntenCh3 SkewIntenCh4 SkewIntenStatusCh1
## 1                  0   0.45450484   0.46039340    1.2327736                  0
## 2                  0   0.39870467   0.61973079    0.5272631                  0
## 3                  0   0.47248709   0.97137879    0.3247065                  0
## 4                  1   0.88167138   0.99989280    0.6044399                  1
## 5                  0   0.51702723   1.17675337    0.9258123                  0
## 6                  0  -0.70259526   0.15320233    0.6511667                  2
```

```
## 7                  0  -0.81606076  -0.22448186    0.1023062                  2
## 8                  1   0.66550394   0.01142938    1.6120078                  0
## 9                  0   0.59674364   1.51738349    0.4218934                  0
## 10                 0   0.58681762   0.40825422    0.6756981                  0
## 11                 0  -0.56913821   0.76536204    0.8186378                  2
## 12                 0   0.54673722   1.43236317    0.7038007                  0
## 13                 1   0.63116945   0.18603538    1.2101592                  0
## 14                 0  -0.08283974   0.08641559    2.0018820                  0
## 15                 0   0.44309238   0.55567549    0.1372252                  0
## 16                 0   0.25651828   1.07109399    1.0454462                  0
## 17                 0   0.41807947   0.88133481    0.7258839                  0
## 18                 0   0.27550348   0.77003268    0.6839946                  0
## 19                 0   5.27421649   0.17712659    0.2239078                  1
## 20                 0   0.66893477   1.93614592    1.6299375                  0
##    SkewIntenStatusCh3 SkewIntenStatusCh4 SpotFiberCountCh3 SpotFiberCountCh4
## 1                   0                  0                1                  4
## 2                   0                  0                4                 11
## 3                   0                  0                2                  6
## 4                   0                  0                4                  7
## 5                   0                  0                1                  7
## 6                   2                  0                1                  4
## 7                   2                  0                0                  4
## 8                   2                  1                2                  7
## 9                   0                  0                1                 11
## 10                  0                  0                1                  7
## 11                  0                  0                1                  4
## 12                  0                  0                0                  5
## 13                  2                  0                0                  6
## 14                  2                  1                2                  4
## 15                  0                  0                1                  4
## 16                  0                  0                1                  5
## 17                  0                  0                4                  4
## 18                  0                  0                1                  2
## 19                  2                  0                2                  2
## 20                  1                  1                2                 12
##    SpotFiberCountStatusCh3 SpotFiberCountStatusCh4 TotalIntenCh1 TotalIntenCh2
## 1                        0                       0          2781           700
## 2                        1                       1         24964        160997
## 3                        0                       0         11552         47510
## 4                        1                       0          5545         28869
## 5                        0                       0          6603         30305
## 6                        0                       0         53779        107680
## 7                        2                       0         43950        100626
## 8                        0                       0          4401          1135
## 9                        0                       1          7593         41852
## 10                       0                       0          6512         32638
## 11                       0                       0         70614        157313
## 12                       2                       0          4613         30855
## 13                       2                       0          4111          1204
## 14                       0                       0          7203         38161
## 15                       0                       0          4340         30719
## 16                       0                       0         14461         74259
## 17                       1                       0          4743         15434
## 18                       0                       0         64408        170105
```

```
## 19                     0                     0         88725        148012
## 20                     0                     1         25911        119024
##    TotalIntenCh3 TotalIntenCh4 TotalIntenStatusCh1 TotalIntenStatusCh2
## 1           1690           392                   0                   2
## 2          54675        128368                   0                   1
## 3          26344         43959                   0                   0
## 4           8042          8843                   0                   0
## 5           5569         11037                   0                   0
## 6          21234         57231                   1                   1
## 7          20929         46187                   0                   1
## 8           4136           373                   0                   2
## 9           6488         24325                   0                   0
## 10          7503         23162                   0                   0
## 11         15699         51325                   1                   1
## 12          3332         11466                   0                   0
## 13          3862           432                   0                   2
## 14          7911         24389                   0                   0
## 15          5548         17588                   0                   0
## 16         14474         23099                   0                   0
## 17          6265         17534                   0                   2
## 18         45045        113205                   1                   1
## 19         58224        120536                   1                   1
## 20         56359        154931                   0                   1
##    TotalIntenStatusCh3 TotalIntenStatusCh4 VarIntenCh1 VarIntenCh3 VarIntenCh4
## 1                    0                   2    12.47468    7.609035    2.714100
## 2                    1                   1    18.80923   56.715352  118.388139
## 3                    0                   0    17.29564   37.671053   49.470524
## 4                    0                   2    13.81897   30.005643   24.749537
## 5                    0                   0    15.40797   20.504288   45.450457
## 6                    0                   0   115.00462   70.918444  184.692907
## 7                    0                   0    93.38827   47.293662  124.475306
## 8                    0                   2    12.41264    9.872296    2.301536
## 9                    0                   0    11.79258   17.923023   42.837856
## 10                   0                   0    11.46656   14.866582   42.296424
## 11                   0                   0    67.84403   40.929727  139.875060
## 12                   0                   0    13.92294   18.643027   40.331747
## 13                   0                   2    12.26925    9.362472    2.578185
## 14                   0                   0    21.24513   23.862052  101.395285
## 15                   0                   0    12.32497   17.747143   41.928533
## 16                   0                   0    20.95648   42.316360   66.965709
## 17                   0                   0    20.58588   40.653090  109.643609
## 18                   1                   1   100.07556  138.203909  304.338787
## 19                   1                   1   201.36869  152.312103  305.582829
## 20                   1                   1    21.19878   81.758387  245.122043
##    VarIntenStatusCh1 VarIntenStatusCh3 VarIntenStatusCh4 WidthCh1
## 1                  0                 2                 2 10.64297
## 2                  0                 0                 0 32.16126
## 3                  0                 0                 0 21.18553
## 4                  0                 0                 2 13.39283
## 5                  0                 0                 0 13.19856
## 6                  1                 0                 0 15.53666
## 7                  1                 0                 0 14.90031
## 8                  0                 2                 2 12.34625
## 9                  0                 0                 0 24.23834
```

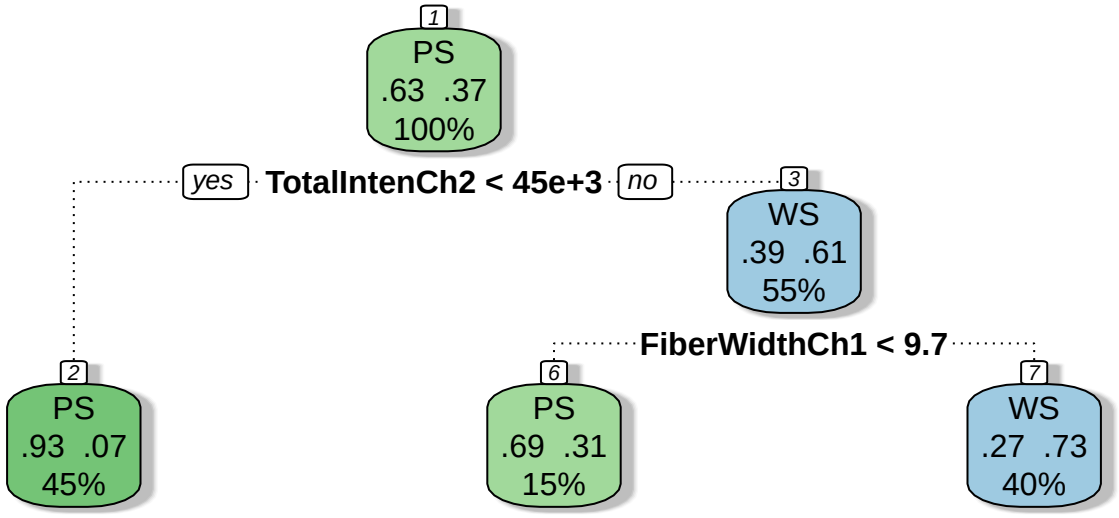
```
## 10                  0                  2                  2 18.33625
## 11                  0                  0                  0 19.81476
## 12                  0                  0                  2 17.54686
## 13                  0                  2                  2 12.49863
## 14                  0                  0                  0 12.58091
## 15                  0                  0                  2 17.66034
## 16                  0                  0                  0 19.43055
## 17                  0                  0                  0 17.39602
## 18                  1                  1                  1 15.47200
## 19                  1                  1                  1 20.69449
## 20                  0                  0                  1 28.73863
##    WidthStatusCh1 XCentroid YCentroid
## 1               2        42        14
## 2               1       215       347
## 3               0       371       252
## 4               0       487       295
## 5               0       283       159
## 6               0       191       127
## 7               0       180       138
## 8               2       373       181
## 9               1       236       467
## 10              0       303       468
## 11              0       473        15
## 12              0       211       495
## 13              2       347       252
## 14              2       229       277
## 15              0       172       207
## 16              0       276       385
## 17              0       239       404
## 18              0       466       449
## 19              0        95        95
## 20              1        77       344
```

```
set.seed(125)

# split train and test
library(dplyr)
training <- dplyr::filter(seg1, Case =="Train")
testing <- dplyr::filter(seg1, Case =="Test")

# fit CART model with rpart
fit <- train(Class ~ ., method="rpart", data=training)

fit$finalModel
```

```
## n= 1009
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
## 1) root 1009 373 PS (0.63032706 0.36967294)
##   2) TotalIntenCh2< 45323.5 454  34 PS (0.92511013 0.07488987) *
##   3) TotalIntenCh2>=45323.5 555 216 WS (0.38918919 0.61081081)
##     6) FiberWidthCh1< 9.673245 154  47 PS (0.69480519 0.30519481) *
```

```
##      7) FiberWidthCh1>=9.673245 401 109 WS (0.27182045 0.72817955) *
```

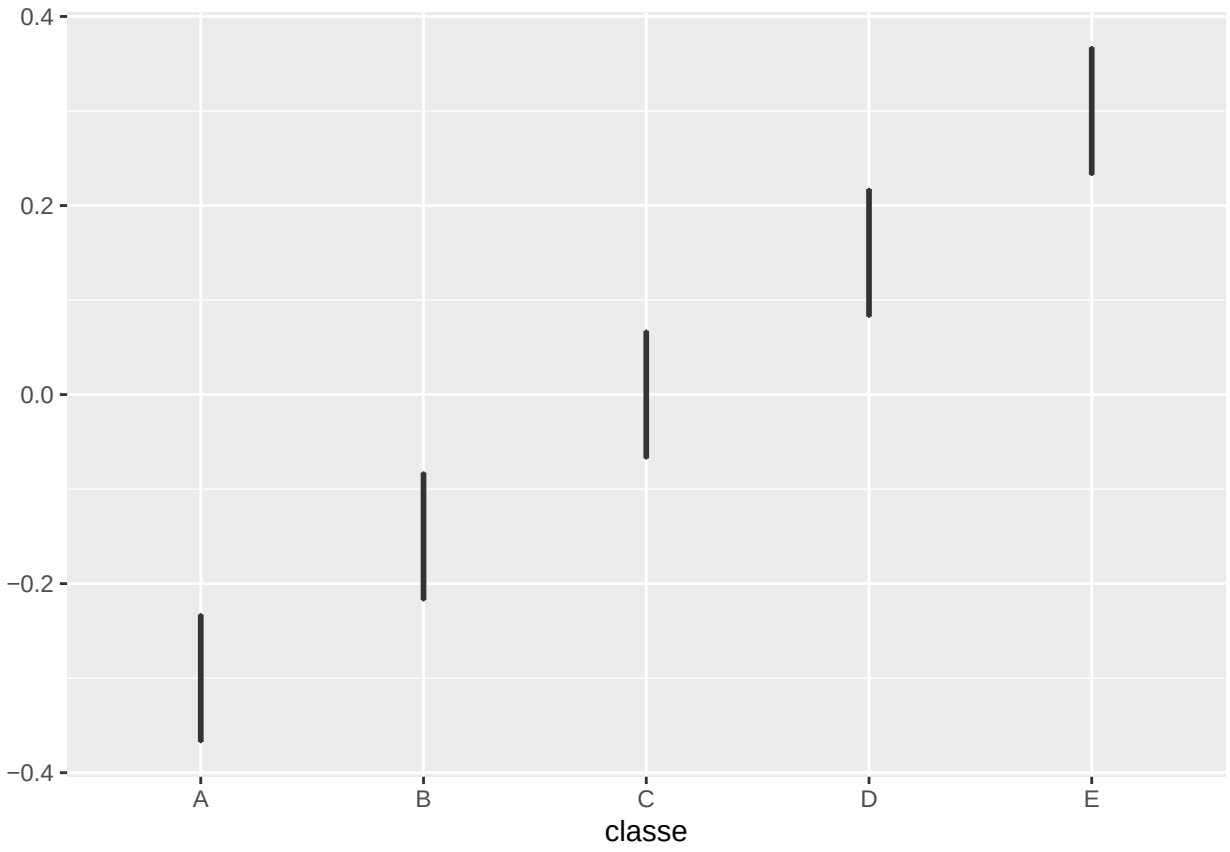
```
fancyRpartPlot(fit$finalModel)
```



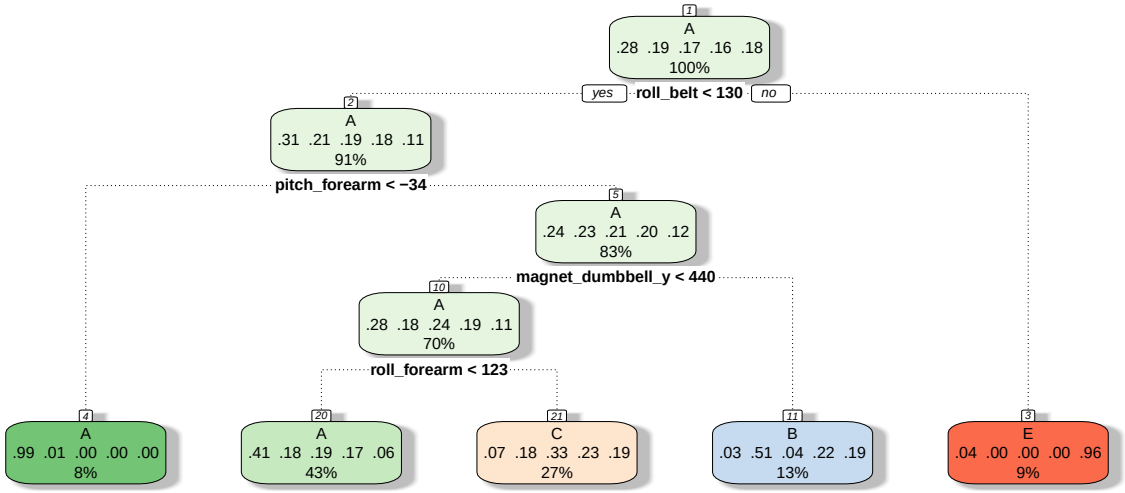
Practical example 2

```
# Set seed for reproducibility
set.seed(1111)

## Data processing- In this section the data is downloaded and manipulated. Some basic transformations

training   <-read.csv("./data/pml-training.csv", na.strings=c("NA","#DIV/0!", "")) ##Cleaning
testing <-read.csv("./data/pml-testing.csv", na.strings=c("NA", "#DIV/0!", ""))
training<-training[,colSums(is.na(training)) == 0]
testing <-testing[,colSums(is.na(testing)) == 0]

training   <-training[,-c(1:7)]  ##Subset for just the columns of interest.
testing <-testing[,-c(1:7)]


# Cross-validation: In this section cross-validation will be performed by splitting the training data i
library(caret)
subSamples <- createDataPartition(y=training$classe, p=0.75, list=FALSE)
subTraining <- training[subSamples, ]
subTesting <- training[-subSamples, ]
```

```
library(ggplot2) ##some exploration
p1 <- ggplot(subTraining, aes(classe))
plot <- p1 + geom_boxplot()
plot
```




```
## Prediction models: In this section a decision tree and random forest will be used on the data.
library(rattle)
# Fit model
modFitDT <- train(classe ~ ., method="rpart", data=subTraining)
# Perform prediction
predictDT <- predict(modFitDT, subTesting, type = "prob")
# Plot result

fancyRpartPlot(modFitDT$finalModel)
```



Rattle 2022−Nov−07 14:56:11 payashome

```
##This Random forest takes 1.5hrs to run, be careful!!!!!!!!!
# Fit model
library(caret)
library(randomForest)
modFit <- train(classe~ .,data=subTraining,method="rf")
modFit
```

```
## Random Forest
##
## 14718 samples
##    52 predictor
##     5 classes: 'A', 'B', 'C', 'D', 'E'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 14718, 14718, 14718, 14718, 14718, 14718, ...
## Resampling results across tuning parameters:
##
##   mtry  Accuracy   Kappa
##    2    0.9883892  0.9853054
##   27    0.9881484  0.9850017
##   52    0.9802377  0.9749918
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was mtry = 2.
```

```
gTree <- getTree(modFit$finalModel,k=3) ##to view the individual trees, you can use this function, whic
head(gTree, n=30)
```

```
##    left daughter right daughter split var split point status prediction
## 1              2              3        21  -270.50000      1          0
## 2              4              5        51  -247.50000      1          0
## 3              6              7        41   -32.95000      1          0
## 4              8              9        28    24.55794      1          0
## 5             10             11         3     1.78500      1          0
## 6             12             13        13  -312.50000      1          0
## 7             14             15        50  -434.50000      1          0
## 8             16             17        36    -2.50000      1          0
## 9             18             19        27    65.70858      1          0
## 10            20             21        48   235.50000      1          0
## 11            22             23         4    19.50000      1          0
## 12             0              0         0     0.00000     -1          1
## 13            24             25        52   787.50000      1          0
## 14            26             27         4    20.50000      1          0
## 15            28             29        36    26.50000      1          0
## 16            30             31        20     0.66500      1          0
## 17            32             33         9    32.00000      1          0
## 18             0              0         0     0.00000     -1          5
## 19             0              0         0     0.00000     -1          2
## 20            34             35        43    31.50000      1          0
## 21            36             37        35    -8.50000      1          0
## 22            38             39        45     0.98500      1          0
## 23             0              0         0     0.00000     -1          5
## 24            40             41        10    29.00000      1          0
## 25             0              0         0     0.00000     -1          2
## 26            42             43        52  -720.50000      1          0
## 27            44             45        36    61.00000      1          0
## 28            46             47        27    62.03453      1          0
## 29            48             49        14     2.19000      1          0
## 30            50             51        36  -144.50000      1          0
```

```
predictRF <- predict(modFit, subTesting, type = "prob")

# Perform prediction
predicts <- predict(modFit, testing, type="prob")
predicts
```

```
##        A     B     C     D     E
## 1  0.094 0.738 0.116 0.032 0.020
## 2  0.870 0.058 0.060 0.004 0.008
## 3  0.144 0.612 0.180 0.010 0.054
## 4  0.912 0.008 0.056 0.018 0.006
## 5  0.938 0.014 0.040 0.002 0.006
## 6  0.014 0.110 0.142 0.052 0.682
## 7  0.036 0.010 0.116 0.804 0.034
## 8  0.090 0.676 0.104 0.098 0.032
## 9  0.998 0.002 0.000 0.000 0.000
## 10 0.994 0.004 0.000 0.000 0.002
```

```
## 11 0.076 0.736 0.076 0.078 0.034
## 12 0.070 0.042 0.822 0.020 0.046
## 13 0.032 0.892 0.008 0.020 0.048
## 14 1.000 0.000 0.000 0.000 0.000
## 15 0.018 0.074 0.050 0.038 0.820
## 16 0.034 0.084 0.014 0.026 0.842
## 17 0.968 0.000 0.002 0.002 0.028
## 18 0.038 0.898 0.014 0.034 0.016
## 19 0.142 0.760 0.040 0.044 0.014
## 20 0.016 0.980 0.002 0.000 0.002
```

Forecasting

Traditionally independent variable over TIME