# HW2_DataMining

Yaniv Bronshtein

3/9/2021

### import libraries

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.5     v dplyr   1.0.3
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(glmnet)
```

```
## Loading required package: Matrix

##
## Attaching package: 'Matrix'

## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack

## Loaded glmnet 4.1
```

```
library(Matrix)
library(rsample)
library(ISLR)
library(pls)
```

```
##
## Attaching package: 'pls'

## The following object is masked from 'package:stats':
##
##     loadings
```

## Question 3

*Bootstrap with Least squares, Ridge and Lasso* Let $\beta = (\beta_1, \beta_2, ..., \beta_p)$ and let x,y be random variables such that the entries of x are i.i.d standard normal variables (i.e, with mean zero and variance one) and $y = \beta^T x + \epsilon$ where $\epsilon \sim N(0, 1)$. (a). Simulate a dataset $(x_1, y_1), ...(x_n, y_n)$ as n i.i.d copies of the random variables x,y defined above, with $n = 800$, $p = 200$, and $\beta_j = j^{-1}$.

```
#Define n and p
p <- 200
n <- 800
#Construct x by create p copies of rnorm() which will generate n numbers. Each column represents a rand
x <- replicate(p, rnorm(n = n, mean = 0, sd = 1))
#Create the j vector
j = 1:p
#Beta is j inverted
beta <- j^-1
#Transform beta
beta_t <- beta %>% as.matrix()
#epsilon is a column vector nX1
epsilon <- rnorm(n = n, mean = 0, sd = 1) %>% as.matrix()
#Construct matrix y according to the formula in the question
y = x %*% beta_t + epsilon

#Create a data frame from y and x
df <- data.frame("y"= y, x)
```

(b) The goal of this problem is to construct confidence intervals for $\beta_1$ using Bootstrap method.

**Specify a 95% confidence interval.** Use $\psi$ to avoid confusion with alpha in glmnet()

```
psi = 0.05
```

**Generate the bootstraps to be used for all the models**

```
set.seed(1)
samples <-df %>% bootstraps(1000)
```

(i). Construct confidence intervals for $\beta_1$ by boostrapping the data and applying Least Squares to the boostrapped data set. **Create a function to return the coefficients of the linear model**

```
get_beta1_lm <- function(data){
  x_mat <- model.matrix(y~.,data)[,-1]
  y <- as.matrix(data['y'])
  lm_model <- lm(y ~ x_mat)
  beta_1_estim <- coef(lm_model)[2]
  return(beta_1_estim)
}
```

**Create a vector of 1000 bootstrap estimates of beta 1 for least square**

```r
estim_lm <- samples$splits %>%
  map(.,~as.data.frame(.)) %>%
  map(.,~get_beta1_lm(.)) %>%
  simplify()

estim_lm
```

```
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9198947 0.9786692 1.0156690 1.0458242 1.0760794 1.0131668 0.8911744 0.9711342
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9678475 1.0213258 0.9937837 0.8931286 0.9403398 1.0231839 0.9306646 0.9438314
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9739355 1.0196414 0.9151550 1.0583078 0.9765379 0.9689989 1.0548590 1.0341454
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9720323 0.9241653 1.0570991 0.9490265 0.8600840 1.0898541 0.9056544 0.9745879
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9642519 1.0502443 1.1072987 1.0166539 1.1099852 1.0315581 1.0320642 1.0339638
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0196268 0.9681221 0.9987922 1.0298655 1.0013170 1.0558315 0.9612283 0.9519157
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9338356 0.9942892 0.9811803 1.0256725 1.0509844 0.9399390 0.9512246 1.0676239
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0113035 0.9648494 1.0283646 1.0375995 1.0034919 1.0458048 1.0193061 1.0733394
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9614978 1.0177641 0.9741181 0.9477524 1.0009987 1.0020269 1.0124911 0.9561293
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0279186 1.0365276 1.0425889 1.0496619 1.0320294 1.0543084 1.0518831 0.8887751
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.8798863 0.9982975 1.0729815 0.9989488 0.9386808 1.0490345 1.0022620 1.0385067
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9975377 0.9273037 0.9575338 0.9958526 0.8886100 0.9631446 1.0139194 0.9639169
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9832154 1.0150638 0.9464973 0.9762201 0.9759765 1.0052108 0.9635617 0.9473460
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0102224 1.0132266 1.0499850 1.0162747 0.9592215 0.9999271 1.0584122 0.9738106
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0572437 1.0672817 0.9111970 1.0496229 0.9032576 0.9526896 1.0284540 1.0276221
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.8893982 1.0239462 1.0089273 0.9548201 1.0511459 1.0074496 1.0212367 1.0864375
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0394440 0.9620582 0.9487107 1.0350930 1.0404489 0.9597783 0.9751477 0.9629864
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0145551 1.0020142 1.0492624 1.0140154 0.9905807 1.0253498 0.9878853 0.9332883
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0848748 0.9884597 0.9545988 0.9557274 0.9779127 0.9834163 1.0032455 1.1074624
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0855759 1.0076888 1.0414008 0.9698381 0.9968597 1.0263921 1.0546542 1.0105672
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9601930 0.9636028 1.0864052 1.0344336 1.0241582 1.0748722 0.9570658 1.0250436
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0083599 0.9744628 1.0020109 0.9251354 0.9094637 0.9248985 0.9951025 1.0494659
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0736551 0.9333748 0.9896772 1.0372710 0.9960045 1.0568946 0.9710753 0.9807220
```

```
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9784021 1.0122495 0.9585275 1.0795563 0.9813089 1.0249847 0.9724334 0.9651843
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9818387 1.0185284 1.0228870 0.9679080 1.0093621 0.9638427 0.9343023 1.0210490
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0242613 1.0334996 1.0540815 0.9632900 0.9479806 0.9334249 1.0875427 0.9684958
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0464284 0.9520619 0.9515980 0.9435644 1.0267710 0.9794059 0.9244876 0.9287340
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9497247 1.0037994 1.0425150 1.0487849 0.9828232 0.9626040 0.9537726 0.9798339
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9662414 1.0195409 0.9971579 0.9243153 1.0028242 0.9791771 0.9625319 0.8821420
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9232935 0.9943758 1.0691520 0.9910427 1.0229763 0.8951269 1.0761666 0.9934374
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0150844 0.8341990 1.0475569 1.0635977 0.9799951 1.0049584 0.9260274 1.0503732
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0097795 0.8904975 0.9342607 0.9912580 1.0127379 0.9958113 0.9202575 0.9624243
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9342328 1.0424202 1.0246135 1.0456124 1.0105457 1.0215931 0.9488350 0.9068206
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0076574 0.9133780 1.0287347 1.0764292 0.9985598 0.9785774 0.9339965 0.9154130
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9767432 1.0229045 1.0269412 1.0633537 1.0445864 0.9977654 0.9779324 0.9485023
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9764459 1.0145993 0.9086280 0.9905611 1.0384079 0.9653729 0.9710351 1.0177884
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9522261 0.9927985 0.9496147 0.9598676 1.0132632 0.9475689 1.0119673 1.0738431
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0011266 1.0503475 1.0304129 1.0226183 0.9528920 1.0378564 1.0129812 1.0192174
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9241491 0.9887877 0.9366591 1.0530886 0.9826103 0.9422098 0.9005180 1.0232565
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9081222 1.0302755 0.9152258 0.9524972 1.0489665 1.0319952 1.0333879 0.9662192
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9574447 1.0269621 1.0225189 0.9929998 1.0330632 0.9838714 1.0629970 1.0817117
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0405276 0.9861349 1.0039630 0.9940203 0.9861131 0.9441590 0.9867795 0.9373021
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0437247 0.9770879 1.0078018 0.9971368 1.1008363 1.0509842 0.9520789 1.0227216
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0301348 0.9477487 1.0620227 1.0349867 0.9063626 1.0237830 0.9972308 0.9823694
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0145841 0.9869689 1.0414538 1.0132268 1.0237223 0.9845674 1.0767852 0.9980984
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9456004 1.0424701 1.0268175 0.9900264 0.9954535 1.0693083 0.9884517 0.9705501
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9570614 0.9700292 0.9710956 0.9583065 0.9606046 0.9395316 1.0329075 1.0126627
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0691616 0.9539271 1.0257187 0.9869830 0.9666469 0.9327526 1.0030623 0.9998625
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0106294 0.9267039 0.8877596 1.0036918 0.9095863 1.0816621 1.0185402 1.0020580
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0001161 0.9581982 1.0246167 0.9412208 1.0177996 1.0029916 1.0310539 0.9220063
```

```
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9796296 1.0391503 1.0938147 1.0217783 0.9535952 0.9426859 0.9799765 0.9982980
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9955421 0.9948887 0.9776098 1.0405606 1.0156920 0.9917367 1.0272228 1.0716395
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.8766995 1.0556697 0.9852289 0.9694291 1.0516120 0.9806091 0.9628718 1.1053335
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0249155 1.0179028 0.9766770 0.8864963 0.9655074 1.0238217 0.9612055 1.0164119
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9781061 0.9954311 0.9925768 0.9972625 0.9755829 1.0045905 0.9880040 1.0900249
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9265634 0.9954194 0.9230058 1.0520386 1.0354745 1.0101128 1.0898633 1.0141580
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9265133 1.0204116 1.0500035 0.9630580 0.9911423 1.0406082 1.0155915 1.0294068
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0853703 0.9676933 0.9631054 1.0229940 0.9582459 1.0166378 1.0292647 0.9639793
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9954665 0.9663809 1.0946169 0.9075531 0.9658719 0.9151069 0.9625784 1.0061609
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0427209 0.9954236 0.9936099 1.0130083 1.0459145 1.0262712 0.9572908 1.0356357
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9779514 1.0226078 1.0594543 1.0300157 1.0703748 1.0170745 1.0455704 0.9947855
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9876327 0.8885033 0.9596444 1.0091180 1.0079374 1.0063500 1.0078268 1.0371378
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0021207 1.0372033 1.1149432 0.9865899 0.9905592 0.9630380 0.9194118 0.8833411
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0008034 1.0052205 0.9543905 1.0606851 1.0020223 0.9596290 0.9942491 1.0017267
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0869545 1.0023012 1.0360228 0.8917000 0.9541829 1.0154291 1.0332189 0.9743751
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0009020 1.0409807 0.9696967 0.9083316 0.9515143 0.9803010 0.9539481 0.9691387
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9886653 0.9861861 0.9491991 1.0279549 1.0420720 1.0521232 0.9962758 0.9997138
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9890617 0.9448175 1.0092044 0.9275262 0.9818877 1.0458290 1.0820439 1.0251156
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0232302 1.0314866 1.0316533 1.0245384 0.9331168 1.0569764 0.9419619 1.0615734
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0263318 0.9471860 1.0192736 1.0366561 0.9194088 1.0305972 0.9966190 1.0265806
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9742424 1.0521756 0.9844927 0.9805152 1.0286272 1.0134838 1.0320326 0.9734771
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0357675 1.0095465 1.0206727 1.0194645 1.0112973 1.0050470 1.0245748 1.0016489
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0265058 1.0251117 0.8872741 1.1165488 0.9928213 1.0238229 1.0117414 0.9734900
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0392448 0.9914664 0.9428528 0.9539791 1.0097614 0.9589860 1.0047060 0.9647388
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0632959 0.9299326 0.9508690 0.9741759 1.0133988 0.9583345 0.9764993 0.9780540
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9345610 1.0500132 0.9518282 1.0201744 0.9293902 1.0105311 0.9698785 1.0537628
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0124679 0.9761560 1.0004173 0.9839735 1.0264126 1.0781989 0.8995902 1.0026680
```

```
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9862048 0.9486174 1.0245011 1.0377305 0.9851592 0.9813818 0.9897931 0.9972467 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0290009 1.0287190 0.9975577 0.9917136 0.9377211 0.9895274 1.0429694 1.0060400 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.1071904 0.9668809 0.9460837 1.0730427 1.0069356 0.9296926 0.9933641 0.9801744 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0034317 0.9340553 0.9739997 0.9793247 1.1222979 1.0095372 0.9781989 0.9533980 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9890538 0.9882338 0.9860329 0.9434437 0.9277253 1.0262789 0.9139467 1.0379330 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0067856 1.0027502 1.0247727 0.9659297 0.9955808 0.9851412 1.0576272 1.0068194 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9679356 0.9739452 1.0427578 0.9966072 1.0408772 0.9767807 0.9884497 1.0294630 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9939010 0.9496163 0.9992926 1.0250294 0.9628703 1.0243225 0.9622338 1.0309822 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0336219 0.9959452 1.0510831 0.9999705 0.9366851 1.0242429 0.9961332 0.9583154 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9285697 0.9505965 1.0288405 0.9535355 0.9256597 1.0044224 0.9971810 0.9389225 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0468110 0.9648289 0.9381886 1.0060417 0.9811304 1.0818845 1.0441552 1.0247566 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9787934 1.0232679 0.9689720 0.9701929 1.0249431 1.0046446 0.9504088 0.9725746 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0270122 0.9676728 1.0553297 0.9755483 1.0584560 1.0020032 0.9708451 1.0613370 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0814580 0.9444859 0.9862882 0.9508051 0.9836599 0.9625821 1.0357592 1.0180692 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9623491 0.9804979 1.0162162 0.9654607 1.0381891 1.0656816 1.0125973 1.0593711 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0053229 1.0659910 0.9615151 1.0033519 1.0304715 0.9503345 0.9589905 0.9149634 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.8961051 0.9652091 0.9856552 1.0250545 0.9606064 1.0080408 1.0569807 0.9507197 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9842999 1.0565412 1.0098300 0.9638092 1.0038317 0.9119774 0.9589941 1.0346690 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0030086 0.9414487 1.0070078 1.0043016 1.0680164 1.0096902 1.0209148 0.9134212 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0153167 1.0296054 0.8664424 0.9647438 0.9858585 1.0243345 1.0062961 1.0099822 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9648660 1.1073495 0.9950626 0.9236892 0.9679534 0.9994973 1.0129287 0.9859358 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0534142 1.0474483 0.9913526 1.0161071 0.9655156 0.9660394 1.0647183 0.9796698 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9659897 0.9937891 1.0098357 1.0394305 0.9640827 1.0273612 0.9829232 1.0426499 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9990075 0.9147371 1.0693006 0.9905609 0.9909176 1.0562756 0.9834951 1.0407299 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 1.0699743 0.9227365 1.0031370 0.9052164 1.0402731 0.9735231 1.0719420 0.9819743 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9953251 0.9652544 0.9904909 1.0120996 0.9863178 0.9919137 0.9837995 0.9921276 
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1 
## 0.9802121 0.9521174 1.0319118 1.0105381 0.9489376 0.9780002 0.9265885 0.9794099 
```

```
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9668309 0.9581089 1.0155393 1.0066032 0.9368677 1.0068993 0.9944453 0.9614527
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9645243 0.9889811 1.0024713 0.9200306 0.9627901 0.9833683 0.9792891 0.9761990
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9074045 0.9282546 1.0485030 0.9460663 0.9450125 0.9754917 1.0114060 1.0678821
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9936269 0.9896551 1.0332804 1.0408454 0.9975460 0.8968997 0.9063927 1.0612425
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9341073 0.9467477 0.9376071 1.0104275 0.9803537 0.9878755 0.9973880 1.0135220
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0350569 1.0020078 0.9990109 0.9642229 0.9724353 0.9768769 1.0384327 1.0067318
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0015202 0.8844905 0.9279512 1.0246801 0.9731415 0.9515739 1.0370910 0.9844573
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9473984 0.9548493 1.0178420 1.0510811 0.9293358 1.0169315 0.9827815 1.0347438
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9587153 0.9712059 1.0267707 1.0267899 1.0516450 0.9621788 1.0996896 1.0481467
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9717446 0.9392353 0.9552818 0.9628091 1.0067333 1.0458453 0.9887961 1.0245239
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9150036 1.0026744 0.9805006 0.9505085 1.0252100 1.0012170 1.0335829 1.0531080
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0554487 1.0404811 1.0688584 1.0474394 1.0540501 1.0562356 0.8986464 0.9428480
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0399277 0.9586207 1.0006316 1.0506873 0.9673024 1.0401208 1.0079508 0.8996452
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9758642 1.0053933 0.9551616 0.9568309 0.9676261 1.0235171 0.9749138 1.0195205
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9376173 0.9633816 1.0284842 0.9758496 0.9509036 0.9682979 1.0188714 0.9346805
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0082968 1.0401229 0.9859845 1.0384567 0.9673877 0.9605891 1.0427459 0.9742667
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0472991 0.9656051 0.9838263 1.0746011 0.9594095 0.9709409 0.9436661 1.0278385
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0373878 1.0518851 0.9821714 0.9729706 0.9298226 1.0046148 0.9758360 1.0034662
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 1.0244752 0.9835934 0.9533373 0.9715194 0.9758628 1.0439805 0.9678169 0.9495575
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9845194 0.9431443 1.0377010 1.0352064 0.9351620 1.0193471 0.9819925 1.0630688
##   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1   x_matX1
## 0.9987477 0.9423180 1.0281257 1.0296351 0.9476046 1.1142084 0.9719002 1.0439778
```

**Display the confidence interval for the bootstrap estimates by applying the quantile() function to get the lower and upper bounds and median() to get the estimate**

```
lm_confint <- estim_lm %>%
  as.data.frame()
colnames(lm_confint)[1] = "estimate"
lm_confint <- lm_confint %>%
  summarise(conf.low = quantile(estimate, psi / 2),
            median = median(estimate),
            conf.high = quantile(estimate, 1 - psi / 2))
lm_confint
```

```
##    conf.low    median conf.high
## 1 0.9004962 0.9964415  1.081889
```

ii. Construct confidence intervals for $\beta_1$ by boostrapping the data and applying Ridge to the boostrapped data set. **Define x__mat__cv and y__cv to be used for cv.glmnet() for both ridge and lasso**

```
x_mat_cv = model.matrix(y~.,df)[,-1]
y_cv = as.matrix(df['y'])
```

**Use cv.glmnet() to perform cross validation for ridge to extract the optimal(minimal) lambda to be used in glmnet()**

```
best_lambda_ridge <- (cv.glmnet(x_mat_cv, y_cv, alpha = 0))$lambda.min
```

**Create a function to return the coefficients of the Ridge model**

```
get_beta1_ridge = function(data){
  x_mat = model.matrix(y~.,data)[,-1]
  y = as.matrix(data['y'])
  ridge_model = glmnet(x_mat, y, alpha = 0, lambda = best_lambda_ridge)
  coef_ridge <- coef(ridge_model)
  beta_1_estim <- coef_ridge[2, 1]
  return(beta_1_estim)
}
```

**Create a vector of 1000 bootstrap estimates of beta 1 for ridge**

```
estim_ridge <- samples$splits %>%
  map(.,~as.data.frame(.)) %>%
  map(.,~get_beta1_ridge(.)) %>%
  simplify()

estim_ridge
```

```
##    [1] 0.7097176 0.7445196 0.7714905 0.7787428 0.7923692 0.7782389 0.6957261
##    [8] 0.7594222 0.7358038 0.7792644 0.7573250 0.6939873 0.7479135 0.7827265
##   [15] 0.7213151 0.7386054 0.7369787 0.7724025 0.7011172 0.8167422 0.7465947
##   [22] 0.7604086 0.7958661 0.8050109 0.7317098 0.6951037 0.8238828 0.7287042
##   [29] 0.6599533 0.8284938 0.6838492 0.7464226 0.7135441 0.7978399 0.7980748
##   [36] 0.7664522 0.8258698 0.7680932 0.7519327 0.8009131 0.7585040 0.7386098
##   [43] 0.7772201 0.7452146 0.7508061 0.7990401 0.7473985 0.7578202 0.7311751
##   [50] 0.7297944 0.7180124 0.7937719 0.7807356 0.7270569 0.7241785 0.8020031
##   [57] 0.7762751 0.7288494 0.7819297 0.7861039 0.7891623 0.7911025 0.7693903
##   [64] 0.8069206 0.7645307 0.7727808 0.7802573 0.7218666 0.7601952 0.7730647
##   [71] 0.7651433 0.7300846 0.7776865 0.7751960 0.7902243 0.7969392 0.7571641
##   [78] 0.7973073 0.7938214 0.6889560 0.6896220 0.7616942 0.8148593 0.7433733
##   [85] 0.7248862 0.7985370 0.7734775 0.7946140 0.7434004 0.7174536 0.7267286
##   [92] 0.7727373 0.7063185 0.7502314 0.7769451 0.7218164 0.7460744 0.7824586
##   [99] 0.7265576 0.7338202 0.7587752 0.7889481 0.7674774 0.7185510 0.8053014
##  [106] 0.7708317 0.8091550 0.7889674 0.7395048 0.7574946 0.7979135 0.7481547
##  [113] 0.7838699 0.7911026 0.7046881 0.8006087 0.6830706 0.7640539 0.7798347
##  [120] 0.7967030 0.6695028 0.7819089 0.7789580 0.7281855 0.7953795 0.7808871
```

```
## [127] 0.7898206 0.8051974 0.7992139 0.7240514 0.7428429 0.7926884 0.7943105
## [134] 0.7413032 0.7474622 0.7376180 0.7889143 0.7602737 0.8171497 0.7808483
## [141] 0.7618910 0.7954293 0.7757222 0.7022213 0.8008767 0.7689037 0.7549544
## [148] 0.7364487 0.7535600 0.7372698 0.7549883 0.8601659 0.8177538 0.7510477
## [155] 0.7951937 0.7357499 0.7633083 0.7767061 0.8156750 0.7451277 0.7450715
## [162] 0.7527099 0.8150727 0.7855672 0.7718950 0.8149486 0.7251206 0.7934454
## [169] 0.7471314 0.7528225 0.7520146 0.7222922 0.6779255 0.7207259 0.7197297
## [176] 0.7691013 0.8449649 0.7320536 0.7554997 0.7985872 0.7795602 0.7930994
## [183] 0.7415957 0.7608608 0.7378618 0.7774668 0.7356658 0.7999217 0.7417687
## [190] 0.7759793 0.7452051 0.7327038 0.7480114 0.7852465 0.7776302 0.7377749
## [197] 0.7715552 0.7545905 0.7197125 0.7765360 0.7718110 0.7827341 0.8137529
## [204] 0.7421285 0.7143915 0.7106810 0.8425367 0.7441481 0.8117139 0.7437224
## [211] 0.7427385 0.7088623 0.7603480 0.7575295 0.6992575 0.6870935 0.7392695
## [218] 0.7770561 0.7811901 0.7904581 0.7573385 0.7406791 0.7078041 0.7533992
## [225] 0.7265085 0.7846565 0.7542615 0.7049146 0.7595367 0.7349457 0.7468969
## [232] 0.6863485 0.6960311 0.7409287 0.8092361 0.7318471 0.7878724 0.6775576
## [239] 0.8116021 0.7244805 0.7778419 0.6469124 0.8068277 0.8031002 0.7515035
## [246] 0.7799974 0.7050617 0.8095600 0.7768901 0.6939549 0.7362192 0.7548875
## [253] 0.7699461 0.7703982 0.7012211 0.7428939 0.7210711 0.7882435 0.7988014
## [260] 0.7781047 0.7691159 0.7809858 0.7361348 0.7086043 0.7635100 0.6959352
## [267] 0.7830189 0.8112004 0.7766608 0.7476133 0.7006352 0.6951825 0.7370439
## [274] 0.7565240 0.7703847 0.8108358 0.7900504 0.7616386 0.7735282 0.7383459
## [281] 0.7701028 0.7456687 0.7060310 0.7468749 0.7862308 0.7475704 0.7386600
## [288] 0.7924289 0.7306495 0.7557731 0.7381370 0.7339678 0.7843271 0.7367506
## [295] 0.7720199 0.8176789 0.7638613 0.7572934 0.7572932 0.7735786 0.7290114
## [302] 0.7619875 0.7730341 0.7690902 0.7135374 0.7364904 0.7146498 0.7892280
## [309] 0.7672202 0.7345109 0.6858267 0.7476967 0.6850172 0.7501022 0.7317778
## [316] 0.7549400 0.7821506 0.7753397 0.7795530 0.7588050 0.7238428 0.7709579
## [323] 0.7861103 0.7633756 0.7874541 0.7751759 0.8114995 0.8216292 0.7715249
## [330] 0.7363183 0.7560442 0.7510116 0.7547421 0.7231771 0.7544025 0.7418099
## [337] 0.7694342 0.7242777 0.7684125 0.7772745 0.8173440 0.8365975 0.7192787
## [344] 0.7757309 0.7752063 0.7032879 0.7950761 0.8021546 0.7107581 0.7631285
## [351] 0.7778836 0.7791999 0.7901445 0.7642313 0.7964979 0.7857416 0.7661779
## [358] 0.7549131 0.8145135 0.7612992 0.7194277 0.7887460 0.7864489 0.7610941
## [365] 0.7570200 0.8253470 0.7464123 0.7508252 0.7407733 0.7564472 0.7471730
## [372] 0.7368950 0.7348487 0.7215540 0.7783964 0.7828364 0.8024278 0.7095538
## [379] 0.7970126 0.7671227 0.7419712 0.7098767 0.7577754 0.7663527 0.7468424
## [386] 0.7168323 0.7005595 0.7424421 0.7074310 0.8047293 0.7679581 0.7595712
## [393] 0.7652986 0.7480108 0.7672823 0.7322590 0.7804416 0.7731963 0.7963475
## [400] 0.7106855 0.7683639 0.7853027 0.8248033 0.7367410 0.7408710 0.7381330
## [407] 0.7238854 0.7660995 0.7845773 0.7511353 0.7660659 0.7979913 0.7916460
## [414] 0.7812328 0.7717177 0.8019255 0.6697820 0.8130005 0.7549317 0.7572629
## [421] 0.7936881 0.7638536 0.7791103 0.8196490 0.7947665 0.7742734 0.7629127
## [428] 0.6777188 0.7540567 0.7728792 0.7273836 0.7541639 0.7327987 0.7604712
## [435] 0.7452571 0.7512687 0.7595014 0.7624224 0.7458503 0.8305900 0.7149168
## [442] 0.7547369 0.7231186 0.8079475 0.7832152 0.7639900 0.8102649 0.7805384
## [449] 0.7191559 0.7832159 0.7842970 0.7284506 0.7433402 0.7837665 0.7678254
## [456] 0.7855148 0.8195710 0.7437230 0.7380321 0.7811874 0.7595103 0.7595085
## [463] 0.7798696 0.7314806 0.7625771 0.7362997 0.8343001 0.7070793 0.7498922
## [470] 0.7107318 0.7508251 0.7649803 0.8025812 0.7741666 0.7808625 0.7606939
## [477] 0.7766998 0.7906052 0.7339892 0.7754406 0.7513138 0.7968455 0.7755948
## [484] 0.7439476 0.8141716 0.7912462 0.7592575 0.7578330 0.7574911 0.6790819
## [491] 0.7481432 0.7316896 0.7688730 0.7741761 0.7884836 0.7757666 0.7642496
## [498] 0.7809706 0.8198417 0.7593918 0.7719383 0.7322211 0.7346099 0.6930967
```

```
## [505] 0.7621863 0.7618844 0.7325877 0.7702570 0.7660062 0.7331280 0.7400238
## [512] 0.7527823 0.8097231 0.7643873 0.7827824 0.6728654 0.7383619 0.7646211
## [519] 0.7815219 0.7684025 0.7489344 0.7799688 0.7544311 0.7103718 0.7103173
## [526] 0.7575945 0.7324794 0.7275563 0.7370850 0.7507998 0.7380905 0.8034782
## [533] 0.7944511 0.7876981 0.7639128 0.7588500 0.7629149 0.7197824 0.7865061
## [540] 0.7167282 0.7433368 0.7581316 0.8045285 0.7604060 0.7679702 0.8055961
## [547] 0.7941208 0.7610057 0.7237212 0.7818208 0.7151214 0.7925691 0.7824176
## [554] 0.7310375 0.7877092 0.8019362 0.7017412 0.7658456 0.7640428 0.7867915
## [561] 0.7481731 0.8048671 0.7594008 0.7432810 0.7796324 0.7734923 0.7716872
## [568] 0.7552097 0.7938998 0.7332623 0.7738276 0.7662496 0.7654629 0.7515155
## [575] 0.7834108 0.7583199 0.7686469 0.7741013 0.6720262 0.8317202 0.7547630
## [582] 0.7908994 0.7578089 0.7539210 0.7804128 0.7429877 0.7173809 0.7003543
## [589] 0.7472377 0.7340865 0.7555046 0.7449298 0.7871542 0.7167109 0.7302139
## [596] 0.7478731 0.7704538 0.7373981 0.7434166 0.7379777 0.7138844 0.8197764
## [603] 0.7234499 0.7730975 0.7146831 0.7843648 0.7580857 0.7896640 0.7705150
## [610] 0.7640991 0.7651008 0.7683909 0.7608487 0.8188121 0.6859476 0.7663132
## [617] 0.7663199 0.7177644 0.7818596 0.8037115 0.7456440 0.7609530 0.7448636
## [624] 0.7765114 0.7756016 0.7995925 0.7693231 0.7382889 0.7249796 0.7663714
## [631] 0.7829138 0.7564638 0.8361472 0.7259572 0.7096493 0.8058101 0.7550947
## [638] 0.7108840 0.7506040 0.7865525 0.7655187 0.7121616 0.7314727 0.7534783
## [645] 0.8392904 0.7987876 0.7362875 0.7207819 0.7531336 0.7751723 0.7796229
## [652] 0.7061945 0.7051802 0.7710498 0.7160704 0.7871573 0.7761228 0.7734939
## [659] 0.8027999 0.7329060 0.7677719 0.7381065 0.7972752 0.7640089 0.7476565
## [666] 0.7567894 0.7922771 0.7638907 0.7762012 0.7694710 0.7393224 0.7966868
## [673] 0.7500533 0.7377243 0.7553378 0.7942636 0.7356740 0.7830284 0.7137869
## [680] 0.7760154 0.7804262 0.7689830 0.8085716 0.7836000 0.7011377 0.7810696
## [687] 0.7722431 0.7468239 0.7131480 0.7446722 0.7942309 0.7209722 0.7097414
## [694] 0.7580007 0.7565439 0.7295744 0.8001474 0.7222959 0.7019543 0.7421976
## [701] 0.7487925 0.8096339 0.7984880 0.7610823 0.7355186 0.7762745 0.7433679
## [708] 0.7503059 0.7988963 0.7270844 0.7431145 0.7526247 0.7929404 0.7503846
## [715] 0.7895025 0.7375688 0.7955349 0.7625262 0.7259245 0.7946503 0.8327328
## [722] 0.7034930 0.7506833 0.7176934 0.7526820 0.7406013 0.8127844 0.7638394
## [729] 0.7371841 0.7555121 0.7913393 0.7291205 0.7740573 0.7913644 0.7677698
## [736] 0.7865419 0.7677646 0.8215546 0.7415451 0.7699418 0.7984056 0.7181179
## [743] 0.7264955 0.7018749 0.6965446 0.7370281 0.7392357 0.7704490 0.7528796
## [750] 0.7640266 0.7879673 0.7433790 0.7321303 0.8216673 0.7499268 0.7096615
## [757] 0.7463737 0.6972008 0.7266166 0.8093551 0.7617726 0.7083978 0.7763628
## [764] 0.7714957 0.7998181 0.7679219 0.7835073 0.6979469 0.7876820 0.7876157
## [771] 0.6942559 0.7237274 0.7373187 0.7743651 0.7880079 0.7742051 0.7306328
## [778] 0.8493174 0.7752680 0.7283702 0.7464663 0.7459073 0.7442638 0.7523335
## [785] 0.7974775 0.7694005 0.7613752 0.7659485 0.7575711 0.7332705 0.7843000
## [792] 0.7693880 0.7582131 0.7522422 0.7635245 0.7937395 0.7579158 0.7932399
## [799] 0.7513301 0.7942216 0.7512858 0.7154570 0.7747557 0.7485355 0.7555953
## [806] 0.7848658 0.7490410 0.8148709 0.8093258 0.7350039 0.7610974 0.7257832
## [813] 0.8068983 0.7245355 0.8224420 0.7527587 0.7454078 0.7289901 0.7841109
## [820] 0.7638379 0.7516077 0.7476620 0.7535752 0.7720954 0.7590873 0.7121733
## [827] 0.7972156 0.7673343 0.7209848 0.7527451 0.7121147 0.7459590 0.7293050
## [834] 0.7356005 0.7570230 0.7854624 0.7106302 0.7778324 0.7795340 0.7447704
## [841] 0.7264073 0.7718415 0.7843854 0.7207865 0.7548043 0.7490838 0.7630751
## [848] 0.7573036 0.6943022 0.7183966 0.7938692 0.7369656 0.7350734 0.7161560
## [855] 0.7937013 0.8132343 0.7597536 0.7408224 0.7667556 0.8024384 0.7472141
## [862] 0.7108751 0.7064760 0.7870626 0.7180252 0.7304149 0.7094265 0.7610742
## [869] 0.7231457 0.7545931 0.7689343 0.7853897 0.7811566 0.7438201 0.7670147
## [876] 0.7380250 0.7287908 0.7473960 0.7693339 0.7647018 0.7736287 0.6992859
```

```
## [883] 0.6916131 0.7861045 0.7436295 0.7014249 0.7778297 0.7686290 0.7298030
## [890] 0.7277556 0.7663882 0.7841627 0.6807730 0.7546451 0.7401919 0.7884830
## [897] 0.7299611 0.7371765 0.7453238 0.7749179 0.7911399 0.7584592 0.8105124
## [904] 0.7987124 0.7537020 0.7273182 0.6899221 0.7432251 0.7729911 0.7882767
## [911] 0.7533967 0.7634334 0.6943599 0.7540386 0.7674488 0.7162514 0.7855322
## [918] 0.7808445 0.8012213 0.7984689 0.8047773 0.7620774 0.7971858 0.7877581
## [925] 0.8074601 0.7949674 0.7053898 0.7145091 0.7846863 0.7289324 0.7587493
## [932] 0.7870515 0.7547258 0.8029106 0.7556521 0.6989972 0.7476167 0.7574723
## [939] 0.7405180 0.7253315 0.7218436 0.7708241 0.7564524 0.7813669 0.7204354
## [946] 0.7367840 0.7626278 0.7539889 0.7426674 0.7391746 0.7770520 0.7039320
## [953] 0.7823384 0.7912773 0.7854495 0.7940831 0.7431181 0.7453755 0.7980380
## [960] 0.7663854 0.7944726 0.7378449 0.7742843 0.8121794 0.7385644 0.7351148
## [967] 0.7364993 0.7707877 0.8021556 0.8088461 0.7556215 0.7719632 0.7422482
## [974] 0.7687719 0.7366197 0.7583285 0.7792891 0.7508156 0.7257168 0.7504903
## [981] 0.7404687 0.7868759 0.7454054 0.7397834 0.7414770 0.7272082 0.7946580
## [988] 0.8125782 0.7194974 0.7723783 0.7628173 0.7962465 0.7639691 0.7253625
## [995] 0.7991080 0.7642114 0.7397696 0.8099999 0.7453567 0.7944218
```

**Display the confidence interval for the bootstrap estimates by applying the quantile() function to get the lower and upper bounds and median() to get the estimate**

```r
ridge_confint <- estim_ridge %>%
  as.data.frame()
colnames(ridge_confint)[1] = "estimate"
ridge_confint <- ridge_confint %>%
  summarise(conf.low = quantile(estimate, psi / 2),
            median = median(estimate),
            conf.high = quantile(estimate, 1 - psi / 2))

ridge_confint
```

```
##    conf.low    median conf.high
## 1 0.6942491 0.7595539 0.8177802
```

(iii). Construct confidence intervals for $\beta_1$ by boostrapping the data and applying Lasso to the bootstrapped data set.

**Use cv.glmnet() to perform cross validation for lasso to extract the optimal(minimal) lambda to be used in glmnet()**

```r
best_lambda_lasso <- (cv.glmnet(x_mat_cv, y_cv, alpha = 1))$lambda.min
```

**Create a function to return the coefficients of the Lasso model**

```r
get_beta1_lasso = function(data){
  x_mat = model.matrix(y~.,data)[,-1]
  y = as.matrix(data['y'])
  lasso_model = glmnet(x_mat, y, alpha = 1, lambda = best_lambda_lasso)
  coef_lasso <- coef(lasso_model)
  beta_1_estim <- coef_lasso[2, 1]
  return(beta_1_estim)
}
```

**Create a vector of 1000 bootstrap estimates of beta 1 for lasso**

```
estim_lasso <- samples$splits %>%
  map(.,~as.data.frame(.)) %>%
  map(.,~get_beta1_lasso(.)) %>%
  simplify()

estim_lasso
```

```
##    [1] 0.8866644 0.8985255 0.9242725 0.9276995 0.9444668 0.9588975 0.8451853
##    [8] 0.9685211 0.8828291 0.9695218 0.9179477 0.8466190 0.9453458 0.9591695
##   [15] 0.9070580 0.9405312 0.9109404 0.9459126 0.8728344 0.9832534 0.9481791
##   [22] 0.9221063 0.9767601 0.9782987 0.9139078 0.8230130 0.9839861 0.8906673
##   [29] 0.8222634 1.0125764 0.8500391 0.9036113 0.9040392 0.9362062 0.9641979
##   [36] 0.9397687 1.0315318 0.9499756 0.9458378 0.9884270 0.9271215 0.8784637
##   [43] 0.9294682 0.9095834 0.9086202 0.9986124 0.9135295 0.9273830 0.9089430
##   [50] 0.9104951 0.8653036 0.9758097 0.9529394 0.8828750 0.8982052 0.9616433
##   [57] 0.9667523 0.8925547 0.9533790 0.9416007 0.9736191 0.9516655 0.9467800
##   [64] 0.9602473 0.9771661 0.9415753 0.9713124 0.9155157 0.8914027 0.9580916
##   [71] 0.9517025 0.8936908 0.9302457 0.9513721 0.9829752 0.9733080 0.9259420
##   [78] 0.9617770 1.0102212 0.8696085 0.8574328 0.9398746 0.9870652 0.8998634
##   [85] 0.8803742 0.9753343 0.9514251 0.9511127 0.9267590 0.9324064 0.8633897
##   [92] 0.9520876 0.8629779 0.9260873 0.9420001 0.8877811 0.9212558 0.9318625
##   [99] 0.8902473 0.8719358 0.9470212 0.9681946 0.9395341 0.8557472 0.9669304
##  [106] 0.9253206 0.9867080 0.9515093 0.8979274 0.9184626 0.9686145 0.9114991
##  [113] 0.9690648 0.9838493 0.8648765 0.9783980 0.8598395 0.9491019 0.9294148
##  [120] 0.9875992 0.8568638 0.9476656 0.9546155 0.9026445 0.9783159 0.9572546
##  [127] 0.9510655 0.9849803 0.9580782 0.8967775 0.8938675 0.9846444 0.9700814
##  [134] 0.8978489 0.9219558 0.9122594 0.9490374 0.9509542 0.9990840 0.9677496
##  [141] 0.9065593 0.9718334 0.9428539 0.8886073 0.9499756 0.9492699 0.9298228
##  [148] 0.9165874 0.9392285 0.9054387 0.9341410 1.0421854 1.0169062 0.9316233
##  [155] 0.9708052 0.9136841 0.9307772 0.9403303 0.9885627 0.9165400 0.9209717
##  [162] 0.9215720 1.0148325 0.9538789 0.9620842 0.9917551 0.9045457 0.9710063
##  [169] 0.8899936 0.8990656 0.9254089 0.8985601 0.8311603 0.8959058 0.8865723
##  [176] 0.9141958 1.0216179 0.9082475 0.9321217 0.9830749 0.9509241 0.9434456
##  [183] 0.9193370 0.9437260 0.8969349 0.9681287 0.8833780 0.9733986 0.8911919
##  [190] 0.9618219 0.9243516 0.8838296 0.9361712 0.9737420 0.9439583 0.8917888
##  [197] 0.9318855 0.9621731 0.9067167 0.9640811 0.9307029 0.9488687 0.9739263
##  [204] 0.8900583 0.8761258 0.9042318 1.0263563 0.9370366 0.9896724 0.9185603
##  [211] 0.9361371 0.8570320 0.8956238 0.9165398 0.8515741 0.8517502 0.9026755
##  [218] 0.9429912 0.9592920 0.9510306 0.9393869 0.9222988 0.8868377 0.9299533
##  [225] 0.8845944 0.9521581 0.9259703 0.8615563 0.9515886 0.8910442 0.9453006
##  [232] 0.8799167 0.8812506 0.9374611 0.9758763 0.8881374 0.9631494 0.8569769
##  [239] 0.9895383 0.9055557 0.9276039 0.8200130 0.9594868 0.9690657 0.9519035
##  [246] 0.9586939 0.8982767 0.9925048 0.9102716 0.8876119 0.9242961 0.9001998
##  [253] 0.9483803 0.9454428 0.8621033 0.9172735 0.9026592 0.9867498 0.9535515
##  [260] 0.9098192 0.9426451 0.9468557 0.9082065 0.8692055 0.9170538 0.8461652
##  [267] 0.9530598 0.9792622 0.9563348 0.9137164 0.8691390 0.8877322 0.9379656
##  [274] 0.9530166 0.9417752 0.9640225 0.9629021 0.9345061 0.9944415 0.9600937
##  [281] 0.9610732 0.9202189 0.8805679 0.8781146 0.9263370 0.9111285 0.9206605
##  [288] 0.9624451 0.8991046 0.9338518 0.8972598 0.9097565 0.9925934 0.8999039
##  [295] 0.8982250 1.0262274 0.9444289 0.9044345 0.9224141 0.9595652 0.9087727
##  [302] 0.9130903 0.9461421 0.9271604 0.8736601 0.9124153 0.9233584 0.9519133
##  [309] 0.9618227 0.8916857 0.8798648 0.9088922 0.8739441 0.9173379 0.9237045
```

```
## [316] 0.9382833 0.9654414 0.9317185 0.9512035 0.9522639 0.8936655 0.9497259
## [323] 0.9694224 0.9635932 0.9783170 0.9876113 0.9806181 0.9937135 0.9158193
## [330] 0.8947092 0.9223355 0.9220797 0.9273547 0.8765571 0.9435266 0.9188782
## [337] 0.9377335 0.8922068 0.9298727 0.9676217 0.9780848 1.0435483 0.9069473
## [344] 0.9307395 0.9744894 0.8621080 0.9684136 0.9723867 0.8977981 0.9509037
## [351] 0.9216552 0.9688464 0.9798631 0.9360583 0.9666594 0.9550947 0.9188485
## [358] 0.9014406 0.9792581 0.9350031 0.9009371 0.9642318 0.9534072 0.9470833
## [365] 0.9485706 1.0004734 0.9020469 0.9067655 0.9026468 0.9446484 0.9070012
## [372] 0.8760680 0.8670577 0.8796168 0.9642046 0.9332871 1.0084645 0.8866684
## [379] 0.9759188 0.9378454 0.9125887 0.8723202 0.9372468 0.9506056 0.9003314
## [386] 0.8913659 0.8692945 0.9209380 0.8779876 0.9440423 0.9443209 0.9361099
## [393] 0.9526900 0.9174495 0.9402769 0.9011420 0.9705029 1.0042196 0.9750606
## [400] 0.8801731 0.9589637 0.9576255 0.9854596 0.9008702 0.9042897 0.9221151
## [407] 0.8907855 0.9416375 0.9591653 0.9130747 0.9651485 0.9519591 0.9870198
## [414] 0.9831067 0.9490494 0.9789145 0.8495590 0.9814351 0.9288278 0.9516465
## [421] 0.9769981 0.9170088 0.9788284 0.9852910 0.9713214 0.9304622 0.9133804
## [428] 0.8483685 0.9505148 0.9273458 0.9002192 0.9027701 0.8968946 0.9228471
## [435] 0.9614104 0.9246483 0.9340204 0.9558707 0.8949437 1.0065415 0.9025513
## [442] 0.9602958 0.9059914 0.9995737 0.9663029 0.9335824 0.9863118 0.9505091
## [449] 0.9115006 0.9690425 0.9787149 0.8905684 0.9349397 0.9672669 0.9454125
## [456] 0.9845726 1.0075539 0.9148866 0.9057964 0.9257944 0.9457807 0.9420082
## [463] 0.9241650 0.9089932 0.9620919 0.9005941 0.9995585 0.8712128 0.9369516
## [470] 0.8732645 0.9226638 0.9534720 1.0018015 0.9468630 0.9634924 0.9027596
## [477] 0.9647279 0.9576303 0.9284995 0.9311921 0.9324603 0.9823110 0.9580949
## [484] 0.8882159 0.9895329 0.9659656 0.9513654 0.9369393 0.9104691 0.8518139
## [491] 0.9078652 0.8811713 0.9197790 0.9345732 0.9770401 0.9535744 0.9248660
## [498] 0.9447733 0.9992565 0.9461078 0.9566259 0.9236143 0.9088534 0.8830453
## [505] 0.9783057 0.9308610 0.9095528 0.9411041 0.9404197 0.8842071 0.8951187
## [512] 0.8837429 0.9726079 0.9162542 0.9399184 0.7955815 0.8939660 0.9107874
## [519] 0.9435764 0.9884149 0.9087844 0.9393128 0.9333108 0.9107384 0.8581953
## [526] 0.9373549 0.9063806 0.8872081 0.8982553 0.9239925 0.8986625 0.9951444
## [533] 0.9623525 0.9641764 0.9492742 0.9376193 0.9251601 0.8911874 0.9856384
## [540] 0.8904433 0.9105990 0.9249443 0.9823329 0.9218468 0.9434050 0.9558703
## [547] 0.9743601 0.9447398 0.9157164 0.9620428 0.9057118 0.9592745 0.9718097
## [554] 0.8856001 0.9628420 0.9557460 0.8811211 0.9569168 0.9382922 1.0049189
## [561] 0.9217286 0.9879093 0.9212269 0.9308166 0.9622571 0.9452699 0.9666980
## [568] 0.9081755 0.9902459 0.8943485 0.9367541 0.9436589 0.9331485 0.9257579
## [575] 0.9356795 0.9190294 0.9281084 0.9580137 0.8464440 1.0092268 0.9189857
## [582] 0.9506337 0.9282524 0.9321918 0.9294667 0.9314024 0.9109988 0.8713535
## [589] 0.9275105 0.9359836 0.9486917 0.8998948 0.9325894 0.9115309 0.9171521
## [596] 0.9435213 0.9567113 0.9102142 0.9074259 0.9083162 0.8816563 1.0008049
## [603] 0.8773306 0.9369637 0.8665796 0.9686827 0.9143324 0.9515226 0.9380244
## [610] 0.9396287 0.9278115 0.9326262 0.9103974 0.9876507 0.8496464 0.9295589
## [617] 0.9299233 0.8693104 0.9445426 0.9495715 0.9153985 0.9388286 0.9092283
## [624] 0.9718196 0.9542364 0.9708383 0.9211981 0.8953992 0.8842211 0.9336912
## [631] 0.9457765 0.9487786 1.0447112 0.9001316 0.8803392 0.9854424 0.9040696
## [638] 0.8748309 0.9482077 0.9893531 0.9745078 0.8657095 0.9065751 0.9624518
## [645] 1.0290546 0.9712131 0.8943905 0.8677562 0.9424910 0.9323225 0.9592605
## [652] 0.8578565 0.8904409 0.9182194 0.9045339 0.9660344 0.9774270 0.9279601
## [659] 0.9833361 0.8998064 0.9448223 0.9048127 0.9533589 0.9355986 0.9215257
## [666] 0.9276309 0.9636859 0.9345018 0.9651786 0.9680817 0.9401477 1.0104849
## [673] 0.8523572 0.9057358 0.9324831 0.9727885 0.8900941 0.9448230 0.8829993
## [680] 0.9407559 0.9379378 0.9604633 0.9859476 0.9380876 0.8574344 0.9436663
## [687] 0.9572460 0.9238615 0.8753180 0.9283253 0.9820996 0.8914759 0.8937371
```

```
## [694] 0.9374554 0.9210139 0.8876993 0.9596984 0.8961654 0.8970571 0.9144826
## [701] 0.9071671 0.9789637 0.9686429 0.9540766 0.9228848 0.9414327 0.9253510
## [708] 0.9384899 0.9868870 0.8683542 0.9179274 0.9117385 0.9626102 0.9193853
## [715] 0.9793735 0.9406260 0.9751693 0.9297479 0.8569206 0.9577573 1.0286614
## [722] 0.8892711 0.9104528 0.8961753 0.9249526 0.9103764 1.0036500 0.9193341
## [729] 0.9164189 0.9549762 0.9758549 0.8939863 0.9546693 0.9807017 0.9056936
## [736] 0.9680070 0.9712677 1.0216583 0.9069216 0.9399154 0.9766730 0.8454296
## [743] 0.9110339 0.8750734 0.8963494 0.9242699 0.9194636 0.9148462 0.9221658
## [750] 0.9270006 0.9809607 0.9214238 0.8887594 0.9977908 0.9331726 0.8996824
## [757] 0.9270204 0.8496010 0.8763421 1.0044351 0.9311213 0.8517576 0.9371484
## [764] 0.9392324 0.9807213 0.9587610 0.9590912 0.8522558 0.9656227 0.9643888
## [771] 0.8597593 0.8990989 0.9151626 0.9514713 0.9773499 0.9358680 0.9217786
## [778] 1.0105848 0.9859097 0.8949263 0.9261717 0.9077551 0.9246400 0.9048629
## [785] 0.9611311 0.9539389 0.9386465 0.9489694 0.9256297 0.9092482 0.9600399
## [792] 0.9476788 0.9155018 0.9136217 0.9376709 0.9766098 0.9470553 0.9486325
## [799] 0.9296047 0.9677817 0.9367750 0.9136920 0.9563303 0.9552591 0.9164872
## [806] 0.9679249 0.9343103 1.0152978 0.9605367 0.9412048 0.9336704 0.9444096
## [813] 0.9965777 0.8731299 0.9901314 0.9257516 0.9116948 0.8931713 0.9694283
## [820] 0.9267795 0.9741094 0.8916320 0.9300911 0.9661125 0.9174317 0.9046186
## [827] 0.9753166 0.9054280 0.9031615 0.9206182 0.8683808 0.9230620 0.8960727
## [834] 0.9176075 0.8960054 0.9628376 0.8802459 0.9598903 0.9850366 0.8959338
## [841] 0.9661284 0.9821680 0.9345171 0.8992460 0.9338311 0.9412226 0.9409629
## [848] 0.9131058 0.8941911 0.9070923 0.9574283 0.9425502 0.9114247 0.8907186
## [855] 0.9656708 0.9742025 0.9597091 0.9196423 0.9454393 0.9716225 0.9236624
## [862] 0.8887601 0.8490858 0.9499420 0.8928048 0.9076606 0.8709874 0.9133311
## [869] 0.9020108 0.9090331 0.9350471 0.9646367 0.9922725 0.9148800 0.9318201
## [876] 0.8952693 0.9096098 0.9339558 0.9282400 0.9129124 0.9388362 0.8897677
## [883] 0.8785407 0.9922907 0.9034501 0.8656308 0.9409620 0.9758169 0.9283276
## [890] 0.8789611 0.9414521 0.9457248 0.8670255 0.9004152 0.9172234 0.9741263
## [897] 0.8833637 0.9090982 0.8852771 0.9673446 0.9583805 0.9102225 0.9378227
## [904] 0.9748980 0.9350438 0.9427443 0.8468758 0.9347754 0.9320504 0.9694287
## [911] 0.9534614 0.9578739 0.8367737 0.9186027 0.9615586 0.8818722 0.9804749
## [918] 0.9884936 0.9929520 0.9567397 0.9899265 0.8901431 1.0018319 0.9681582
## [925] 0.9801150 0.9562148 0.9027278 0.8708203 0.9616410 0.9154818 0.9339876
## [932] 0.9619435 0.9089991 0.9847698 0.9121370 0.8682660 0.9309232 0.9441279
## [939] 0.9101988 0.9084480 0.8812352 0.9479253 0.9495053 0.9362900 0.8917249
## [946] 0.9218943 0.9006198 0.9273403 0.9327713 0.8891859 0.9505138 0.8671086
## [953] 0.9498452 0.9614892 0.9488980 0.9849925 0.8665598 0.9568795 0.9467416
## [960] 0.9390410 0.9792772 0.9163278 0.9278371 0.9921725 0.9341025 0.9392010
## [967] 0.9173451 0.9698551 0.9648440 1.0186835 0.9347241 0.9448899 0.9377759
## [974] 0.9498827 0.9107805 0.9165861 0.9694012 0.9231000 0.9062803 0.9063814
## [981] 0.9256209 0.9492561 0.9343131 0.8778119 0.9242609 0.8966138 0.9995873
## [988] 1.0217461 0.8768014 0.9777821 0.9450588 0.9638950 0.9458049 0.8789258
## [995] 0.9858269 0.9485212 0.9412506 0.9843307 0.9111920 0.9766610
```

**Display the confidence interval for the bootstrap estimates by applying the quantile() function to get the lower and upper bounds and median() to get the estimate**

```
lasso_confint <- estim_lasso %>%
  as.data.frame()
colnames(lasso_confint)[1] = "estimate"
lasso_confint <- lasso_confint %>%
  summarise(conf.low = quantile(estimate, psi / 2),
            median = median(estimate),
```

```
           conf.high = quantile(estimate, 1 - psi / 2))
lasso_confint
```

```
##    conf.low    median conf.high
## 1 0.8568359 0.9338414  1.004225
```

(c). Comment on the obtained results **From the results, we assertain that lasso and least squares produce similar estimates for $\beta_1$ with lasso providing the tightest bound of all 3. Ridge produces a significantly lower estimate than both lasso and least squares with a bound that is not as tight as lasso but significantly tighter than least squares**

## Problem 4(Question 9 in ISL Pg. 263 )

9.In this exercise, we will predict the number of applications received
using the other variables in the College data set.

```
data(College)
```

(a). Split the data set into a training set and a test set. set.seed(1)

```
data(College)
set.seed(1)
smp_siz <- dim(College)[1] / 2

train <- sample(seq_len(nrow(College)),size = smp_siz)
test <- -train
data_train <- College[train,]
data_test <- College[test,]
```

(b). Fit a linear model using least squares on the training set, and report the test error obtained.

```
lm_model <- lm(Apps ~ ., data = data_train)
lm_predict <- predict(lm_model, data_test)
lm_predict
```

```
##                          Agnes Scott College
##                                  2054.898055
##                           Albertson College
##                                   673.657532
##                     Albertus Magnus College
##                                  -149.929987
##                              Albion College
##                                  2722.169420
##                            Albright College
##                                   847.596822
##                   Alderson-Broaddus College
##                                   564.067517
##                           Alfred University
##                                  1876.762135
##                           Allegheny College
```

```
##                                     3021.916028
##       Allentown Coll. of St. Francis de Sales
##                                     2103.586449
##                                    Alma College
##                                     2089.054859
##                                 Amherst College
##                                     4021.813888
##                                 Aquinas College
##                                      279.238899
##          Arizona State University Main campus
##                                    15388.046640
##                        Arkansas Tech University
##                                     1739.892865
##                              Assumption College
##                                     2200.024338
##                            Augustana College IL
##                                     2404.481987
##                                   Barat College
##                                      189.927754
##                                 Barnard College
##                                     3064.308673
##                              Bellarmine College
##                                     1582.064502
##                        Bemidji State University
##                                     1127.599072
##                             Benedictine College
##                                      643.902085
##                                   Berry College
##                                     2198.723251
##                               Bethel College KS
##                                       48.483671
##                                  Bethel College
##                                       39.979913
##                          Bethune Cookman College
##                                     1498.360522
##                     Birmingham-Southern College
##                                     2616.612399
##              Bloomsburg Univ. of Pennsylvania
##                                     4096.048123
##                                Bluffton College
##                                      173.140508
##                                 Bowdoin College
##                                     3906.766362
##                                Bradford College
##                                     -227.930183
##                              Bradley University
##                                     5625.391819
##                               Brenau University
##                                     -116.347870
##                           Brewton-Parker College
##                                     -166.425775
##                             Briar Cliff College
##                                      279.516939
##                             Bridgewater College
```

```
##                                    873.392683
##           Brigham Young University at Provo
##                                   6871.083493
##                           Buena Vista College
##                                    549.788656
##                             Butler University
##                                   3302.882740
##                               Cabrini College
##                                    602.669882
##              California Lutheran University
##                                    -61.762048
##      California State University at Fresno
##                                   3911.210406
##                           Campbell University
##                                   1198.215103
##                            Capital University
##                                   2088.526232
##                              Carleton College
##                                   4209.193690
##                  Carnegie Mellon University
##                                   9474.711129
##                        Carson-Newman College
##                                   1011.182444
##             Case Western Reserve University
##                                   6580.419320
##                               Catawba College
##                                    809.468312
##              Catholic University of America
##                                   2143.807971
##                             Cazenovia College
##                                   5902.705647
##                           Cedar Crest College
##                                    895.012915
##                            Cedarville College
##                                   1653.297911
##              Centenary College of Louisiana
##                                   1012.055042
##                 Center for Creative Studies
##                                   1247.389412
##                               Central College
##                                   1474.177755
##                                Centre College
##                                   2471.491235
##                               Chatham College
##                                    132.570559
##               Christian Brothers University
##                                      7.187550
##              Christopher Newport University
##                                    296.797058
##                    Claremont McKenna College
##                                   3187.364649
##                           Clarkson University
##                                   2935.944391
## Clinch Valley Coll. of  the Univ. of Virginia
```

```
##                                 845.718438
##                              Colby College
##                                3072.535064
##                      College Misericordia
##                                1962.389804
##              College of Mount St. Vincent
##                                1221.043263
##                 College of Saint Benedict
##                                 894.473493
##                College of Saint Catherine
##                                 485.387588
##                     College of Saint Rose
##                                1072.339197
##                     College of St. Joseph
##                                 314.233474
##                College of St. Scholastica
##                                 845.541600
##               College of William and Mary
##                                6506.186190
##                 Colorado State University
##                                9785.996455
##                          Columbia College
##                                 419.459649
##             Concordia College at St. Paul
##                                -315.803498
##                      Concordia University
##                                 436.799538
##                          Converse College
##                                 744.953424
##                           Cornell College
##                                1606.377657
##                      Creighton University
##                                4821.067906
##                   Culver-Stockton College
##                                1539.813629
##                        Cumberland College
##                                1524.315585
##                        D'Youville College
##                                 624.205205
##                              Dana College
##                                  15.243377
##                    Daniel Webster College
##                                 545.769532
##                         Dartmouth College
##                                6142.651321
##                          Davidson College
##                                3499.495954
##                         Dickinson College
##                                3525.354608
##                Dickinson State University
##                                 365.992449
##                             Dordt College
##                                 470.691569
##                           Dowling College
```

```
##                                      38.087731
##                               Drake University
##                                    3750.312779
##                                  Drury College
##                                     991.742535
##                                Duke University
##                                    8686.745839
##                       East Carolina University
##                                    8744.087198
##                East Tennessee State University
##                                    3663.271944
##                    Eastern Illinois University
##                                    6078.761481
##                       Eastern Nazarene College
##                                     344.188401
##                          Elizabethtown College
##                                    2913.551837
##                                   Elms College
##                                     240.239278
##                                   Elon College
##                                    3367.212879
##                          Emory & Henry College
##                                    1397.852166
##                                 Eureka College
##                                    1121.887686
##                        Evergreen State College
##                                    1788.516051
##                                 Ferrum College
##                                     898.727067
##                                Flagler College
##                                     834.507749
##               Florida International University
##                                    3772.520809
##                       Florida State University
##                                   14306.970264
##                             Fordham University
##                                    4740.359992
##                             Fort Lewis College
##                                    3660.644249
##                      Francis Marion University
##                                    1758.816865
##                               Franklin College
##                                     923.464024
##                        Franklin Pierce College
##                                    6655.980696
##                         Fresno Pacific College
##                                    1120.473048
##                              Gannon University
##                                    2542.881259
##                        Gardner Webb University
##                                    1004.118394
##                             Georgetown College
##                                    1180.335191
##                Georgia Institute of Technology
```

```
##                             9636.161174
##                Georgia State University
##                             2785.003681
##                   Georgian Court College
##                             -292.832076
##                      Gettysburg College
##                             3393.017864
##                    Goldey Beacom College
##                              575.519112
##                      Gonzaga University
##                             2096.601990
##                          Goshen College
##                              802.267907
##                         Goucher College
##                             2078.320275
##            Grace College and Seminary
##                              443.411485
##                      Greensboro College
##                              157.612168
##                      Greenville College
##                              351.385386
##                      Grove City College
##                             3156.209768
##                        Guilford College
##                             1092.634160
##                        Hamilton College
##                             2465.101862
##                      Hamline University
##                              781.747466
##                Hampden - Sydney College
##                              528.429036
##                     Hampton University
##                             4996.896686
##                     Harding University
##                             1633.674287
##                        Hartwick College
##                             1943.881807
##                     Harvard University
##                             6280.605273
##                        Hastings College
##                              589.512413
##                         Hendrix College
##                             1859.077906
##                           Hiram College
##                             1407.651272
##       Hobart and William Smith Colleges
##                             2520.582855
##                     Hofstra University
##                             9112.390541
##                         Hollins College
##                              747.509831
##                            Hope College
##                             2244.431479
##                        Houghton College
```

```
##                                  764.110944
##                            Huntingdon College
##                                  945.801856
##                            Huntington College
##                                  105.274392
##                              Huron University
##                                 -189.330820
##                                Husson College
##                                  837.223549
##                  Illinois Benedictine College
##                                  612.135829
##              Illinois Institute of Technology
##                                 2118.522576
##                            Immaculata College
##                                  256.966834
##                      Indiana State University
##                                 4473.360944
##             Indiana University at Bloomington
##                                17312.797489
##                   Indiana Wesleyan University
##                                  694.469286
##                         Iowa State University
##                                10259.734800
##                         John Brown University
##                                  588.226618
##                         Johnson State College
##                                  608.598608
##                               Juniata College
##                                 1374.466108
##                       Kansas State University
##                                 4753.934201
##                     Kentucky Wesleyan College
##                                  634.535842
##                                  King College
##                                  681.577317
##                                  Knox College
##                                 1643.936858
##                              La Roche College
##                                  -85.903902
##                           Lake Forest College
##                                  322.099903
##                              Lakeland College
##                                  704.795958
##                             Lander University
##                                 1112.916990
##                             Le Moyne College
##                                  887.076489
##                             Lehigh University
##                                 6407.165106
##                         Lenoir-Rhyne College
##                                  905.967781
##                                Lesley College
##                                  303.170639
##                        LeTourneau University
```

```
##                                     1145.423001
##                          Lewis and Clark College
##                                     3006.462513
##                               Lincoln University
##                                     1703.651271
##                                Lindenwood College
##                                      459.437196
##                                  Linfield College
##                                     2365.996377
##                                     Loras College
##                                     1514.046220
##     Louisiana State University at Baton Rouge
##                                     6204.643802
##                     Loyola Marymount University
##                                     4982.879359
##                                    Luther College
##                                     1937.689335
##                                  Lycoming College
##                                      724.854019
##                                 Lynchburg College
##                                     1566.295688
##                                Macalester College
##                                     3163.460193
##                                    Malone College
##                                      704.801359
##                            Manhattanville College
##                                      605.523658
##                         Mankato State University
##                                     3580.443374
##                      Marian College of Fond du Lac
##                                      234.517418
##                                    Marist College
##                                     4771.274014
##                               Mary Baldwin College
##                                     1482.906336
##                           Mary Washington College
##                                     4342.770107
##                       Marymount Manhattan College
##                                     1645.981908
##                              Marymount University
##                                      539.584739
##                                Maryville College
##                                     1117.461420
##                              Maryville University
##                                      816.756892
##                                 Mercer University
##                                     2439.346157
##                                  Meredith College
##                                      998.992783
##                                 Merrimack College
##                                     2330.138872
##                                Mesa State College
##                                     2116.792932
##                                   Messiah College
```

```
##                                  2185.691759
##                      Miami University at Oxford
##                                 11756.600121
##                    Michigan State University
##                                 20084.836678
##              Millersville University of Penn.
##                                  4668.576280
##                          Millikin University
##                                  1778.684436
##             Milwaukee School of Engineering
##                                  1900.186869
##                  Mississippi State University
##                                  3674.828707
##                     Missouri Valley College
##                                   819.687391
##                         Monmouth College IL
##                                   181.684941
##                     Montana State University
##                                  3247.375032
##                    Moorhead State University
##                                  2577.501430
##                            Moravian College
##                                   717.826170
##                           Morehouse College
##                                  3639.959396
##                         Morningside College
##                                   484.111603
##                       Mount Holyoke College
##                                  2499.552386
##                          Mount Mary College
##                                  -117.677318
##                         Mount Mercy College
##                                    64.392998
##                   Mount Saint Mary's College
##                                  1562.481707
##                    Mount Saint Mary College
##                                   715.718603
##                     Mount St. Mary's College
##                                  1277.736381
##                         Mount Union College
##                                   827.223804
##                Mount Vernon Nazarene College
##                                   572.699606
##                          Muhlenberg College
##                                  2456.788542
##                            Newberry College
##                                   469.654341
##                          Niagara University
##                                  4546.825442
##                    North Adams State College
##                                  1422.980349
##       North Carolina A. & T. State University
##                                  3486.103705
##                       North Central College
```

```
##                                   2003.026668
##                 North Dakota State University
##                                   1923.875657
##                             North Park College
##                                    185.404050
##                     Northern Arizona University
##                                   7302.293866
##                        Northwestern University
##                                  10466.636291
##                            Norwich University
##                                   1289.681926
##                            Notre Dame College
##                                     34.604359
##                            Occidental College
##                                   2827.852011
##                        Oglethorpe University
##                                   1823.638014
##                               Ohio University
##                                  11818.950859
##                             Otterbein College
##                                   1619.783352
##                        Pacific Union College
##                                    695.735628
##                    Pembroke State University
##                                    709.122502
##        Pennsylvania State Univ. Main Campus
##                                  17009.683053
##                        Pepperdine University
##                                   5689.277531
##                             Pfeiffer College
##                                    502.228103
##     Philadelphia Coll. of Textiles and Sci.
##                                   1919.870975
##                             Piedmont College
##                                    653.524947
##                               Pitzer College
##                                    746.576495
##                            Point Park College
##                                    847.234268
##             Prairie View A. and M. University
##                                   3429.537256
##                         Presbyterian College
##                                   1264.822112
##                          Providence College
##                                   4806.981192
##                               Queens College
##                                    834.608133
##                           Radford University
##                                   7126.060635
##                      Randolph-Macon College
##                                   1398.773253
##              Randolph-Macon Woman's College
##                                   1367.901344
##                               Rhodes College
```

```
##                                           4032.701634
##                                         Ripon College
##                                            172.936602
##                                        Rivier College
##                                             63.611233
##                                       Roanoke College
##                                           2377.364539
##                               Rocky Mountain College
##                                              2.544508
##                                        Rosary College
##                                            654.023346
##                   Rutgers State University at Newark
##                                           5061.277715
##                             Saint Ambrose University
##                                            224.024154
##                                 Saint Anselm College
##                                           1929.869642
##                             Saint Francis College IN
##                                             38.781269
##                            Saint Joseph's University
##                                           3371.248392
##                                 Saint Joseph College
##                                           -218.249751
##                               Salem-Teikyo University
##                                            577.502623
##                                    Samford University
##                                           1897.046050
##                                     Schreiner College
##                                            726.623181
##                                       Scripps College
##                                           2738.223533
##                           Seattle Pacific University
##                                           1769.951086
##                                    Seattle University
##                                           1837.257016
##                                Seton Hall University
##                                           5327.961524
##                     Shippensburg University of Penn.
##                                           4286.677710
##                                         Siena College
##                                           2780.943741
##                                Siena Heights College
##                                            245.004280
##                                  Sioux Falls College
##                                           -292.972074
##                                      Skidmore College
##                                           3669.871085
##                                         Smith College
##                                           3247.461319
##                          South Dakota State University
##                                           2806.536174
##                     Southeast Missouri State University
##                                           1991.875790
## Southern Illinois University at Edwardsville
```

```
##                                 2876.676285
##                   Southwest State University
##                                 1458.376895
##                         Southwestern College
##                                  193.404734
##                      Southwestern University
##                                 1938.780201
##                         Spring Arbor College
##                                  470.610059
##               St. Mary's College of Maryland
##                                 2329.908498
##         St. Mary's University of San Antonio
##                                 1817.943297
##                    St. Thomas Aquinas College
##                                  873.100832
##              Stevens Institute of Technology
##                                 2397.779348
##               Stockton College of New Jersey
##                                 2864.333555
##                               SUNY at Albany
##                                14539.680463
##                     SUNY College at Cortland
##                                 4764.509318
##                     SUNY College at Fredonia
##                                 4087.129733
##                      SUNY College at Geneseo
##                                 8757.857436
##                       Susquehanna University
##                                 1713.112191
##                                Tabor College
##                                   46.153986
##                            Taylor University
##                                 2074.804588
##            Texas A&M University at Galveston
##                                  572.855789
##                   Texas Christian University
##                                 4491.871376
##                       Texas Lutheran College
##                                  573.366543
##                    Texas Southern University
##                                 1015.378237
##                          Trinity College DC
##                                  319.052141
##                          Trinity College VT
##                                  292.335285
##                         Tuskegee University
##                                 2401.441948
##                Univ. of Wisconsin at OshKosh
##                                 3194.273054
##         University of California at Irvine
##                                20201.965743
##                     University of Charleston
##                                  819.402531
##                    University of Cincinnati
```

```
##                                                 7821.671230
##                       University of Connecticut at Storrs
##                                                10934.529378
##                                      University of Dallas
##                                                 1788.519062
##                              University of Detroit Mercy
##                                                  754.121188
##                                    University of Dubuque
##                                                 -144.898111
##                           University of Illinois - Urbana
##                                                16398.738475
##                         University of Illinois at Chicago
##                                                 7978.940150
##                              University of Indianapolis
##                                                 1797.820037
##                                University of Louisville
##                                                 3792.154325
##                          University of Maine at Machias
##                                                  746.091677
##                     University of Maine at Presque Isle
##                                                  -82.713679
##                  University of Maryland at College Park
##                                                15530.655642
##                 University of Massachusetts at Amherst
##                                                17832.246271
##               University of Massachusetts at Dartmouth
##                                                 3575.763968
##                                       University of Miami
##                                                 8629.660503
##                       University of Minnesota at Duluth
##                                                 3420.304221
##                     University of Minnesota Twin Cities
##                                                 8358.198297
##                    University of Missouri at Saint Louis
##                                                 1608.809258
##                                      University of Mobile
##                                                  -33.480505
##                       University of Nebraska at Lincoln
##                                                 7662.509796
##              University of North Carolina at Asheville
##                                                 1578.767924
##            University of North Carolina at Chapel Hill
##                                                10282.073664
##             University of North Carolina at Wilmington
##                                                 4619.806033
##                              University of North Dakota
##                                                 2125.203655
##                               University of North Texas
##                                                 3600.272200
##                            University of Northern Iowa
##                                                 4233.009902
##                                      University of Oregon
##                                                 9244.189110
##                                    University of Portland
```

```
##                              1863.686370
##                 University of Rhode Island
##                             11648.410396
##                    University of Scranton
##                              4695.592718
##         University of Southern California
##                             13552.994621
##           University of Southern Indiana
##                              2356.486269
##        University of Southern Mississippi
##                              3071.849835
##              University of St. Thomas TX
##                              1037.272234
##            University of Texas at Austin
##                             13454.168424
##        University of Texas at San Antonio
##                              4459.459346
##                   University of the South
##                              1840.520656
##                       University of Tulsa
##                              2472.922724
##                    University of Virginia
##                              9159.403622
##            University of Wisconsin-Stout
##                              2170.083543
##         University of Wisconsin-Superior
##                               520.797110
##        University of Wisconsin-Whitewater
##                              4758.448014
##      University of Wisconsin at Green Bay
##                              2379.650528
##      University of Wisconsin at Milwaukee
##                              4457.057911
##                     Upper Iowa University
##                              -101.923829
##                          Ursuline College
##                               600.524869
##              Valley City State University
##                               157.806288
##                     Valparaiso University
##                              3225.510183
##                            Vassar College
##                              3473.172153
##                 Virginia State University
##                              3179.644109
##                             Virginia Tech
##                             16612.020525
##                 Virginia Union University
##                              1784.693906
##                 Virginia Wesleyan College
##                              1177.889754
##                            Wabash College
##                              1275.490453
##                    Wake Forest University
```

```
##                                6760.832102
##                              Walsh University
##                                  522.118368
##                              Wartburg College
##                                 1617.162144
##                   Washington and Lee University
##                                 3600.513472
##                            Washington College
##                                 1327.508245
##                     Washington State University
##                                 8749.758530
##                           Wayne State College
##                                 1832.050800
##                            Waynesburg College
##                                 1058.023796
##                            Webster University
##                                  362.340633
##                             Wellesley College
##                                 4300.701940
##               Wentworth Institute of Technology
##                                 3051.807327
##                                Wesley College
##                                  928.679849
##                           Wesleyan University
##                                 4240.993141
##                West Chester University of Penn.
##                                 4361.885015
##                     Western Carolina University
##                                 3040.814028
##               Western State College of Colorado
##                                 2059.419051
##                   Western Washington University
##                                 5340.463086
##                        Westfield State College
##                                 2658.548865
##                         Westminster College MO
##                                  518.351648
##       Westminster College of Salt Lake City
##                                  600.338167
##                         Westminster College PA
##                                 1676.912512
##                             Whitworth College
##                                 1910.057624
##                            Widener University
##                                 1605.005506
##                             Wilkes University
##                                 1857.012219
##                         Willamette University
##                                 2385.468953
##                      William Woods University
##                                  535.184986
##                                Wilson College
##                                 -295.074534
##                       Winona State University
```

```
##                                    2765.598209
##                          Wittenberg University
##                                    2702.474799
##                                Wofford College
##                                    2026.872540
##                   York College of Pennsylvania
##                                    3035.206068
```

```
test_error_least_square = mean((data_test[, 'Apps'] - lm_predict)^2)
test_error_least_square
```

```
## [1] 1135758
```

**For the next series of questions, we need our feature data to be in matrix form because glmnet() requires it. Thus we create a train and test matrix containing only feature values from our respectful train and test data**

```
x_mat_train <- model.matrix(Apps ~ ., data = data_train)[, -1]
x_mat_test <- model.matrix(Apps ~ ., data = data_test)[, -1]
```

(c). Fit a ridge regression model on the training set, with $\lambda$ chosen by cross-validation. Report the test error obtained. **Extract the optimal lambda for lasso using cv.glmnet()**

```
lambda_min_ridge <- cv.glmnet(x_mat_train, data_train[, 'Apps'], alpha = 0)$lambda.min
lambda_min_ridge
```

```
## [1] 405.8404
```

**Fit a ridge_model using glmnet() with the optimal lambda Then, call predict() to get the prediction values to ultimately compute the test error**

```
ridge_model <- glmnet(x_mat_train, data_train[, 'Apps'], lambda = lambda_min_ridge, alpha = 0)
ridge_predict <- predict(ridge_model, newx = x_mat_test, s = lambda_min_ridge)

test_error_ridge = mean((data_test[, 'Apps'] - ridge_predict)^2)
test_error_ridge
```

```
## [1] 976897.6
```

(d). Fit a lasso model on the training set, with $\lambda$ chosen by cross-validation. Report the test error obtained, along with the number of non-zero coefficient estimates.

**Extract the optimal lambda for lasso using cv.glmnet()**

```
lambda_min_lasso <- cv.glmnet(x_mat_train, data_train[, 'Apps'], alpha = 1)$lambda.min
lambda_min_lasso
```

```
## [1] 1.97344
```

**Fit a ridge_model using glmnet() with the optimal lambda Then, call predict() to get the prediction values to ultimately compute the test error**

```r
lasso_model <- glmnet(x_mat_train, data_train[, 'Apps'], lambda = lambda_min_lasso, alpha = 1)
lasso_predict <- predict(lasso_model, newx = x_mat_test, s = lambda_min_lasso)

test_error_lasso = mean((data_test[, 'Apps'] - lasso_predict)^2)
test_error_lasso
```

```
## [1] 1116402
```

**Extract the coefficients from the lasso model**

```r
coef_lasso <- coef(lasso_model)
coef_lasso
```

```
## 18 x 1 sparse Matrix of class "dgCMatrix"
##                        s0
## (Intercept) -763.46736132
## PrivateYes  -313.76789130
## Accept         1.76295802
## Enroll        -1.32011619
## Top10perc     64.98040139
## Top25perc    -20.93406827
## F.Undergrad    0.07157355
## P.Undergrad    0.01197052
## Outstate      -0.10493165
## Room.Board     0.20915417
## Books          0.29265420
## Personal       0.00351509
## PhD          -14.48803558
## Terminal       5.33334391
## S.F.Ratio     21.67584609
## perc.alumni    0.51564913
## Expend         0.04812700
## Grad.Rate      7.01799292
```

**Filter out the coefficients equal to zero. Note this process removes the names**

```r
coef_lasso_vec <- coef_lasso %>% as.vector()
coef_lasso_vec
```

```
##  [1] -763.46736132 -313.76789130    1.76295802   -1.32011619   64.98040139
##  [6]  -20.93406827    0.07157355    0.01197052   -0.10493165    0.20915417
## [11]    0.29265420    0.00351509  -14.48803558    5.33334391   21.67584609
## [16]    0.51564913    0.04812700    7.01799292
```

```r
coef_lasso_non_zero <- subset(coef_lasso_vec, !(coef_lasso_vec %in% 0.0))
coef_lasso_non_zero
```

```
##  [1] -763.46736132 -313.76789130    1.76295802   -1.32011619   64.98040139
##  [6]  -20.93406827    0.07157355    0.01197052   -0.10493165    0.20915417
## [11]    0.29265420    0.00351509  -14.48803558    5.33334391   21.67584609
## [16]    0.51564913    0.04812700    7.01799292
```

(e). Fit a PCR model on the training set, with M chosen by cross validation. Report the test error obtained, along with the value of M selected by cross-validation.

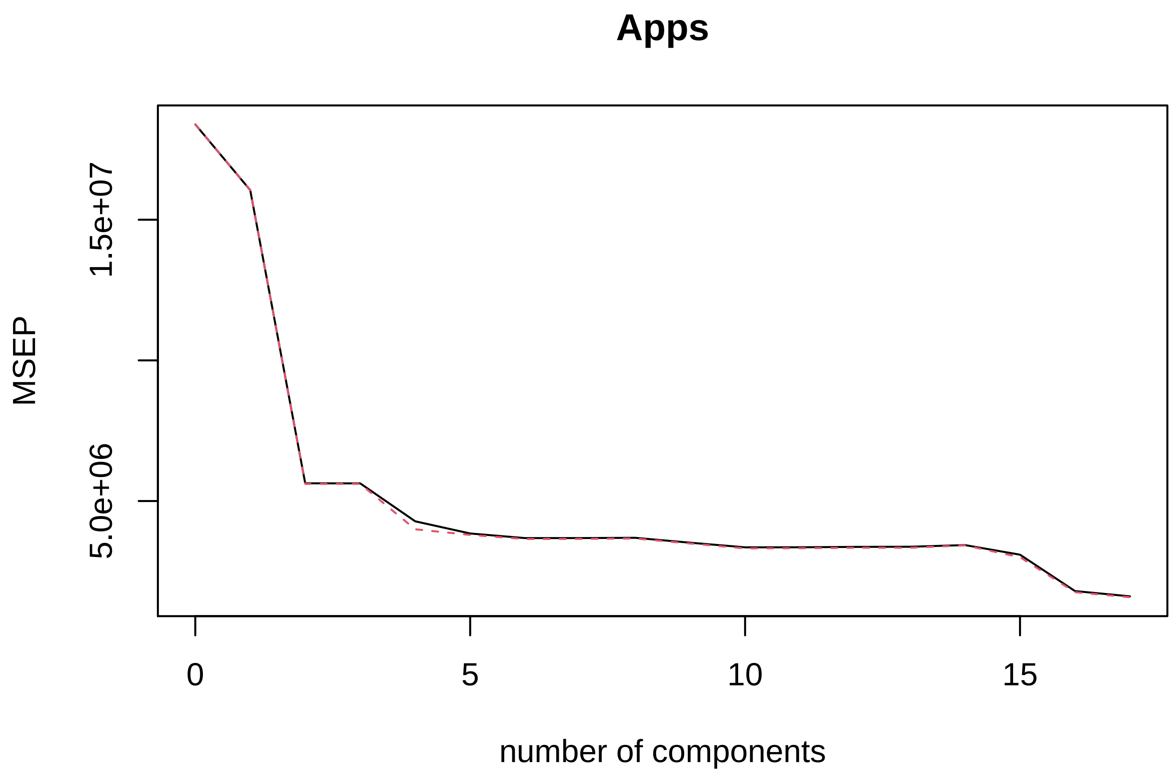
**Fit a PCR model using Cross Validation**

```
set.seed(1)
pcr_model <- pcr(Apps ~., data = data_train, scale = TRUE,
                 validation = "CV")

summary(pcr_model)
```

```
## Data:    X dimension: 388 17
##  Y dimension: 388 1
## Fit method: svdpc
## Number of components considered: 17
##
## VALIDATION: RMSEP
## Cross-validated using 10 random segments.
##        (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
## CV            4288     4006     2373     2372     2069     1961     1919
## adjCV         4288     4007     2368     2369     1999     1948     1911
##        7 comps  8 comps  9 comps  10 comps  11 comps  12 comps  13 comps
## CV        1919     1921     1876      1832      1832      1836      1837
## adjCV     1912     1915     1868      1821      1823      1827      1827
##        14 comps  15 comps  16 comps  17 comps
## CV         1853      1759      1341      1270
## adjCV      1850      1733      1326      1257
##
## TRAINING: % variance explained
##       1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps
## X       32.20    57.78    65.31    70.99    76.37    81.27     84.8    87.85
## Apps    13.44    70.93    71.07    79.87    81.15    82.25     82.3    82.33
##       9 comps  10 comps  11 comps  12 comps  13 comps  14 comps  15 comps
## X       90.62     92.91     94.98     96.74     97.79     98.72     99.42
## Apps    83.38     84.76     84.80     84.84     85.11     85.14     90.55
##       16 comps  17 comps
## X        99.88    100.00
## Apps     93.42     93.89
```

**Generate a validation plot to determine the optimal value of M for pcr Based on the plot, M is either 5 or 6. We shall go with 6**

```
validationplot(pcr_model, val.type = "MSEP")
```




**Report the test error obtained for pcr**

```
pcr_predict <- predict(pcr_model, x_mat_test, ncomp = 6)
test_error_pcr <- mean((data_test[, 'Apps'] - pcr_predict)^2)
test_error_pcr
```

```
## [1] 1966028
```

(f). Fit a PLS model on the training set, with M chosen by cross-validation. Report the test error obtained, along with the value of M selected by cross-validation.

**Fit a PLS model using Cross Validation**

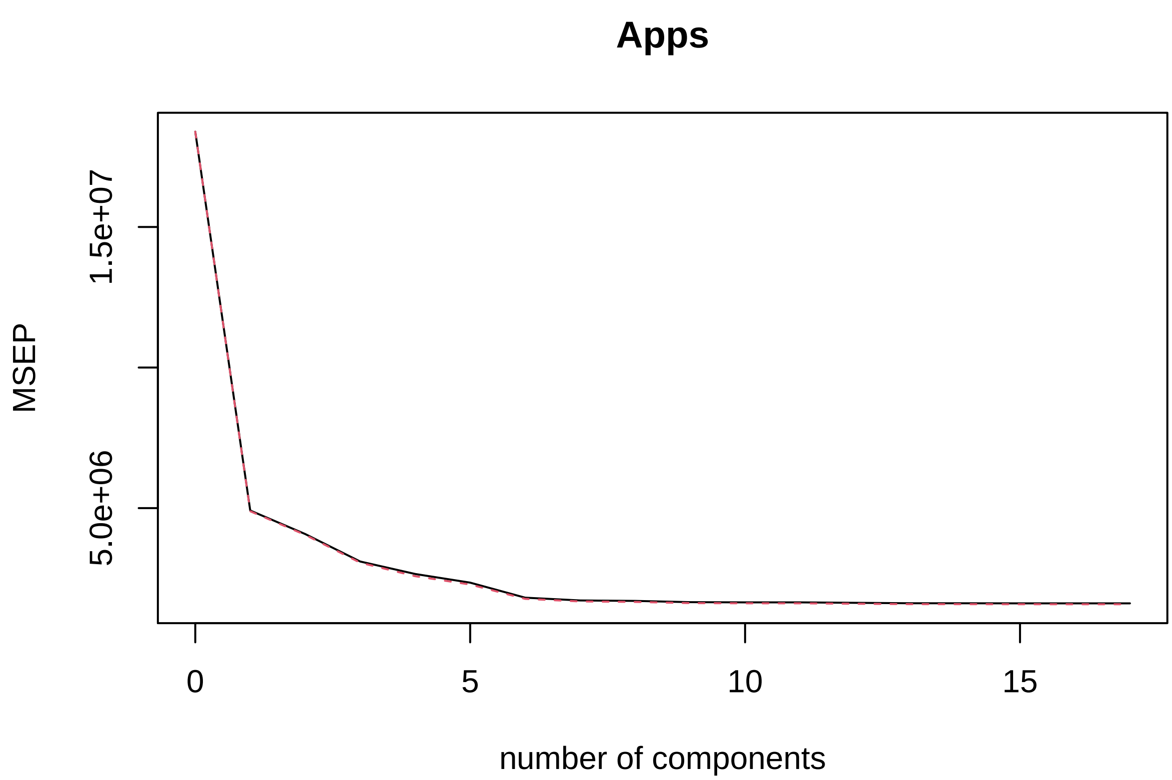
```
set.seed(1)
pls_model <- plsr(Apps ~ ., data = data_train, scale = TRUE, validation = "CV")

summary(pls_model)
```

```
## Data:    X dimension: 388 17
##  Y dimension: 388 1
## Fit method: kernelpls
## Number of components considered: 17
##
## VALIDATION: RMSEP
## Cross-validated using 10 random segments.
```

```
##        (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
## CV            4288     2217     2019     1761     1630     1533     1347
## adjCV         4288     2211     2012     1749     1605     1510     1331
##        7 comps  8 comps  9 comps  10 comps  11 comps  12 comps  13 comps
## CV        1309     1303     1286      1283      1283      1277      1271
## adjCV     1296     1289     1273      1270      1270      1264      1258
##        14 comps  15 comps  16 comps  17 comps
## CV         1270      1270      1270      1270
## adjCV      1258      1257      1257      1257
##
## TRAINING: % variance explained
##       1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps
## X       27.21    50.73    63.06    65.52    70.20    74.20    78.62    80.81
## Apps    75.39    81.24    86.97    91.14    92.62    93.43    93.56    93.68
##       9 comps  10 comps  11 comps  12 comps  13 comps  14 comps  15 comps
## X       83.29     87.17     89.15     91.37     92.58     94.42     96.98
## Apps    93.76     93.79     93.83     93.86     93.88     93.89     93.89
##       16 comps  17 comps
## X        98.78    100.00
## Apps     93.89     93.89
```

**Generate a validation plot to determine the optimal value of M for pls Based on the plot, M is either 6 or 7. We shall go with 6**

```
validationplot(pls_model, val.type = "MSEP")
```



**Report the test error obtained for pcr**

```
pls_predict <- predict(pls_model, x_mat_test, ncomp = 6)
test_error_pls <- mean((data_test[, 'Apps'] - pls_predict)^2)
test_error_pls
```

```
## [1] 1066991
```

(g). Comment on the results obtained. How accurately can we predict the number of college applications received? Is there much difference among the test errors resulting from these five approaches?

```
lm_test_r2 <- 1 - test_error_least_square / mean((data_test[, 'Apps'] - mean(data_test[, 'Apps']))^2)
ridge_test_r2 <- 1 - test_error_ridge / mean((data_test[, 'Apps'] - mean(data_test[, 'Apps']))^2)
lasso_test_r2 <- 1 - test_error_lasso / mean((data_test[, 'Apps'] - mean(data_test[, 'Apps']))^2)
pcr_test_r2 <- 1 - test_error_pcr / mean((data_test[, 'Apps'] - mean(data_test[, 'Apps']))^2)
pls_test_r2 <- 1 - test_error_pls / mean((data_test[, 'Apps'] - mean(data_test[, 'Apps']))^2)
```

**OLS**

```
cat("Test Error:", test_error_least_square, "\n")
```

```
## Test Error: 1135758
```

```
cat("R squared:", lm_test_r2)
```

```
## R squared: 0.9015413
```

**Ridge**

```
cat("Test Error:", test_error_ridge, "\n")
```

```
## Test Error: 976897.6
```

```
cat("R squared:", ridge_test_r2)
```

```
## R squared: 0.9153129
```

**Lasso**

```
cat("Test Error:", test_error_lasso, "\n")
```

```
## Test Error: 1116402
```

```
cat("R squared:", lasso_test_r2)
```

```
## R squared: 0.9032193
```

**PCR**

```r
cat("Test Error:", test_error_pcr, "\n")
```

```
## Test Error: 1966028
```

```r
cat("R squared:", pcr_test_r2)
```

```
## R squared: 0.8295654
```

**PLS**

```r
cat("Test Error:", test_error_pls, "\n")
```

```
## Test Error: 1066991
```

```r
cat("R squared:", pls_test_r2)
```

```
## R squared: 0.9075028
```

**Discussion: All the models have similar test error. With regards to R squared, all the models produced similar values close to 1. Surprisingly, PCR had significantly lower R_squared then the rest and unsurprisingly, Ridge had the highest All models could be used to accurately predict college applications**