

# R project: SLR and Bootstrap

Your Name

Date

## Contents

<b>Data</b>	<b>2</b>
<b>Theory used</b>	<b>3</b>
<b>Application of SLR to the mtcars data set</b>	<b>3</b>
<b>Now make your function:</b>	<b>3</b>
Decide on the output you want . . . . .	3
<b>myslr</b>	<b>4</b>
<b>Now invoke your function</b>	<b>4</b>
<b>Bootstrap</b>	<b>4</b>
<b>Make the function</b>	<b>4</b>
Plots . . . . .	4
Commandline . . . . .	5
File . . . . .	5
<b>Invoke your function for the mtcars dataset</b>	<b>5</b>
<b>Make a &lt;5 min recording of both functions running on the data sets with your explanation</b>	<b>5</b>
#This paragraph will explain what is needed to complete the project and will need to be deleted before submission	

- 1) Use this document to build your project
- 2) The headings can be kept but the descriptions under will need to be changed by you. They let you know what is needed
- 3) All R must go into code R Chunks
- 4) If you need to place a picture use `{width=60%}`
- 5) Mathematical annotation needs to be placed into Latex
- 6) R RStudio, Latex should already be installed.

- 7) You will need at least the following packages in R - s20x, ggplot2
- 8) The project is due at the end of semester - check with CANVAS.
- 9) Use subheadings as needed
- 10) Once finished with the supplied R chunks take off the `eval=FALSE` to see your function run.

#Introduction: What problem do you wish to solve with `myslr()`? – you will need to make a function that creates estimates for the betas in a slr. The function will take an x vector and a y vector.

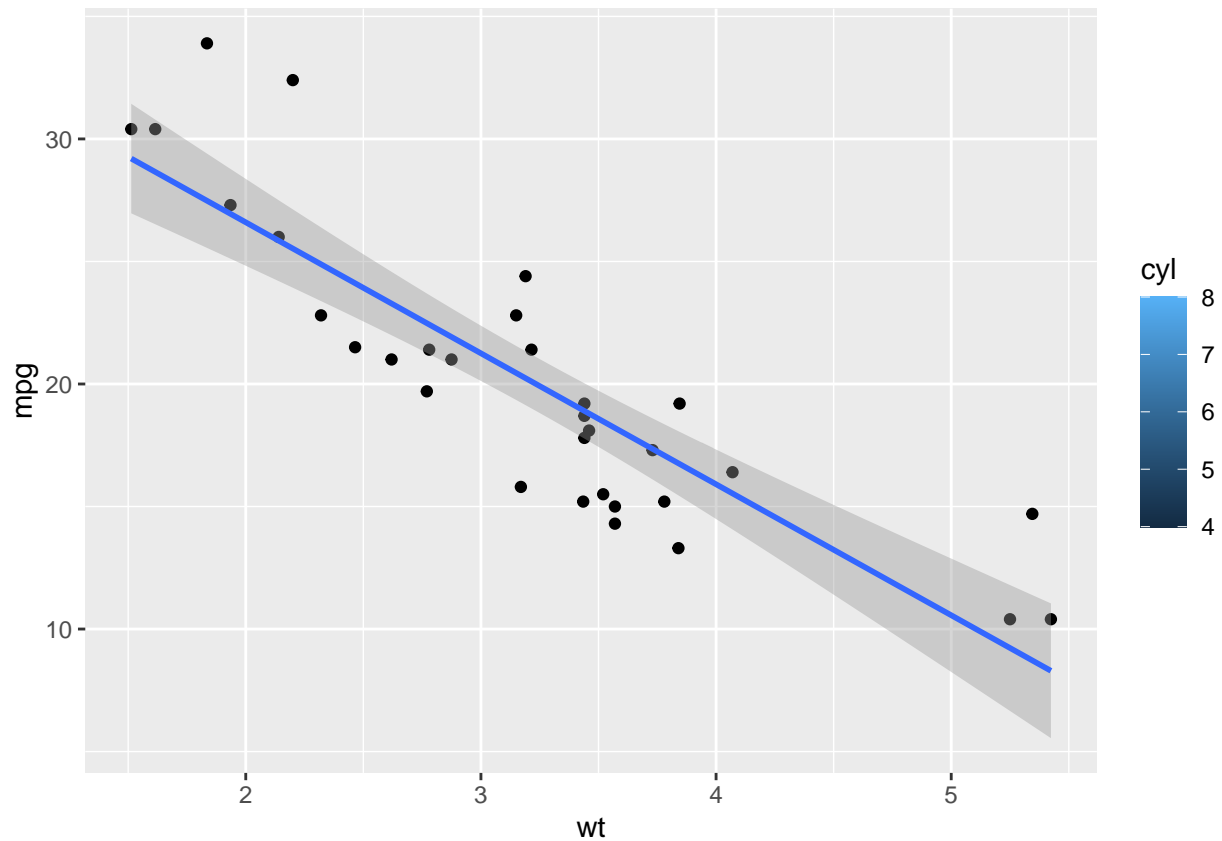
## Data

We will all apply our function to the mtcars data set and make a linear model `mpg ~ wt` - but remember the function will be completely general to SLR. Describe the data ?mtcars Describe the nature and type of the variables. Plot the data – be very creative and interpret the plots

```
data(mtcars)
head(mtcars)
```

```
##           mpg cyl  disp  hp  drat    wt  qsec vs am gear carb
## Mazda RX4      21.0   6  160 110 3.90 2.620 16.46 0  1   4    4
## Mazda RX4 Wag  21.0   6  160 110 3.90 2.875 17.02 0  1   4    4
## Datsun 710      22.8   4  108  93 3.85 2.320 18.61 1  1   4    1
## Hornet 4 Drive  21.4   6  258 110 3.08 3.215 19.44 1  0   3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02 0  0   3    2
## Valiant        18.1   6  225 105 2.76 3.460 20.22 1  0   3    1
```

```
library(ggplot2)
g = ggplot(mtcars, aes(x = wt, y=mpg, fill = cyl)) + geom_point() + geom_smooth(method="lm", formula=y~x)
g
```



## Theory used

Give the mathematical formulae in Latex. Interpret the meaning of the symbols.

## Application of SLR to the mtcars data set

Use R to analyze the data `y.lm=lm(y~x,data=mtcars)` Check assumptions `normcheck()` `plot(y.lm, which =1)`

Interpret the summary output. `summary(y.lm)` `ciReg(y.lm)` What are the point and interval estimates? Give full description of validity of model and what can be said from the summary information

## Now make your function:

This is where you get to be creative after seeing the sorts of things needed in the example above

## Decide on the output you want

Again – be creative and useful.

## Plots

- 1) Plot the data with the fitted line – you can use any package you want – ggplot would work well.
- 2) The plot can be made to appear when the function is used.
- 3) The plot should be saved to the working directory as well (automatic).

## Files created

You may wish to write summary information to a file using `write.csv` – what sort of information? Perhaps new predictions.

## Command line

Create a list of objects you think would be useful – like the beta estimates.

## myslr

```
myslr = function(x,y)
{
  y.lm = lm(y~x)
}
```

## Now invoke your function

```
obj1 = myslr(x=,y=,df=)
obj1
```

## Bootstrap

Make a function `myboot` that will create bootstrap estimates from a sample. The parameter estimated will be the population variance  $\sigma^2$  of mpg for 4 cylinder cars. Though the function will be entirely general.

## Make the function

`x` is a sample vector `alpha` is the error rate, the  $100(1-\alpha)\%$  confidence interval is `quantile(stat, c(alpha/2, 1-alpha/2))`

## Plots

histogram of the simulated statistic. What statistic will you use to estimate  $\sigma^2$

## Commandline

list of interesting objects like  $\hat{\sigma}^2$

## File

write a file to the commandline that contains the simulation.

```
myboot = function(x, alpha)
{
  }
}
```

## Invoke your function for the mtcars dataset

Must get x by using `subset()` or something else

```
obj = myboot(x=, alpha = 0.05)
obj
```

## Make a <5 min recording of both functions running on the data sets with your explanation

You can use quicktime .mov for a mac or use BBFLASH free edition for PC – make as a small file (keep as .fbr)

<https://www.flashbackrecorder.com/express/download/>