**# Assignment: ASSIGNMENT 3**

**# Name: Thangaraj, Arunkumar**

**# Date: 2020-12-16**

> ## Load the ggplot2 package

> library(ggplot2)

> theme\_set(theme\_minimal())

> ## Set the working directory to the root of your DSC 520 directory

> setwd("C:/Users/arun8/R/dsc520\_fork")

> ## Load the `data/r4ds/heights.csv` to

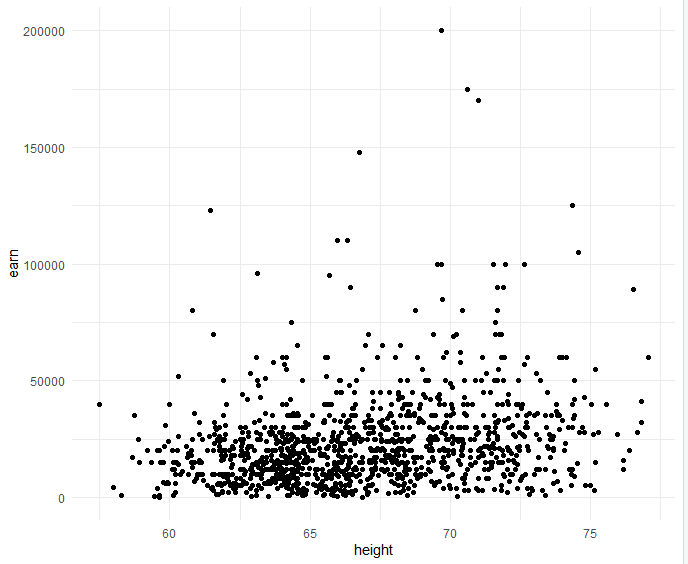
> heights\_df <- read.csv("data/r4ds/heights.csv")

# https://ggplot2.tidyverse.org/reference/geom\_point.html

## Using `geom\_point()` create three scatterplots for

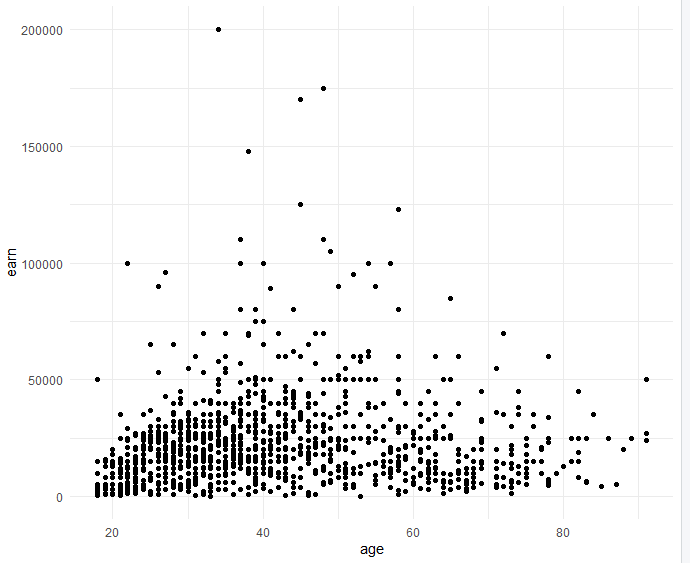
## `height` vs. `earn`

ggplot(heights\_df, aes(x=height, y=earn)) + geom\_point()



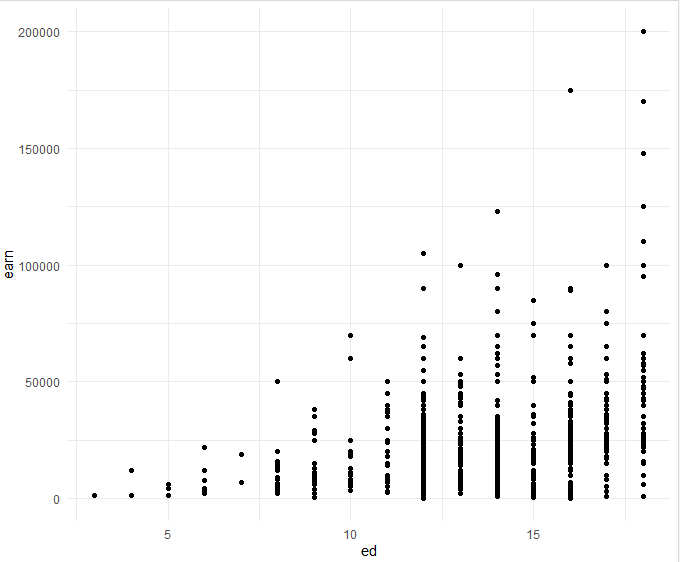
## `age` vs. `earn`

ggplot(heights\_df, aes(x=age, y=earn)) + geom\_point()



## `ed` vs. `earn`

ggplot(heights\_df, aes(x=ed, y=earn)) + geom\_point()



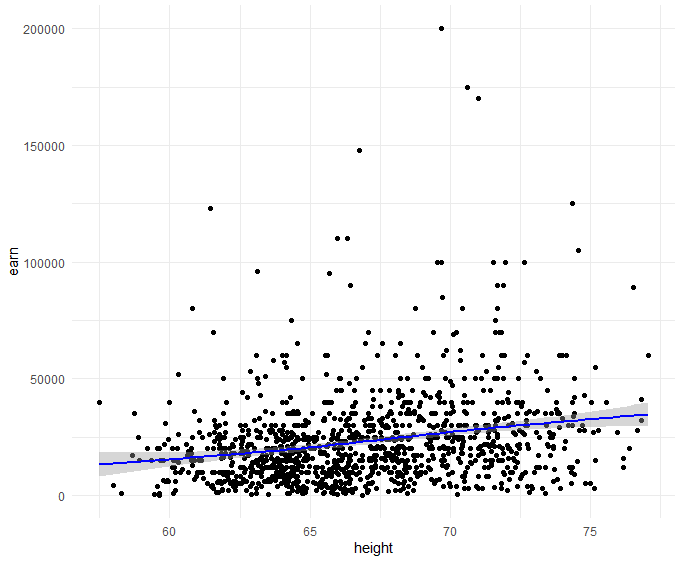
## Re-create the three scatterplots and add a regression trend line using

## the `geom\_smooth()` function

## `height` vs. `earn`

ggplot(heights\_df, aes(x=height, y=earn)) + geom\_point() +

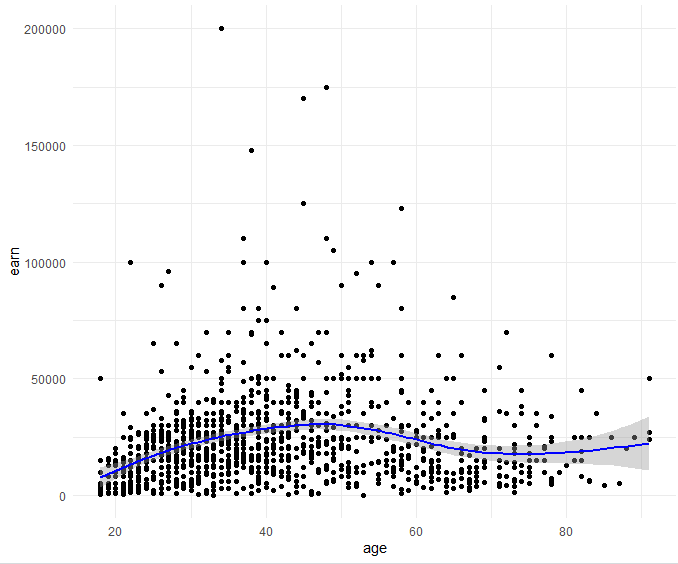
geom\_smooth(method = NULL, colour = "Blue")



## `age` vs. `earn`

ggplot(heights\_df, aes(x=age, y=earn)) + geom\_point() +

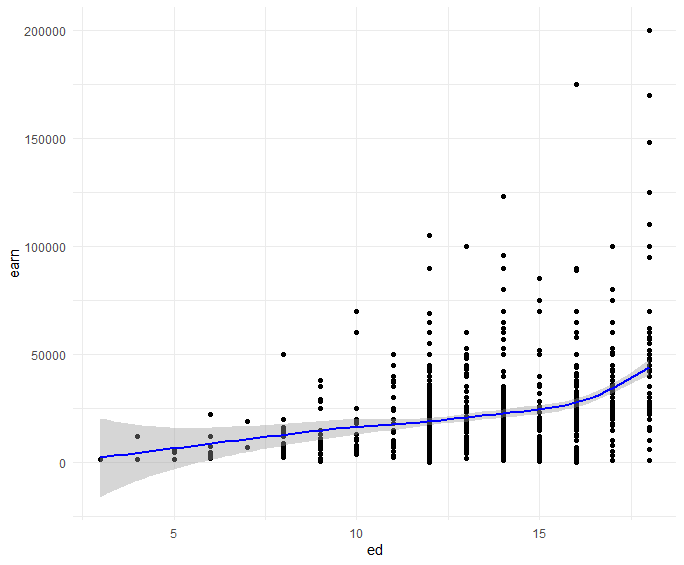
geom\_smooth(method = NULL, colour = "Blue")



## `ed` vs. `earn`

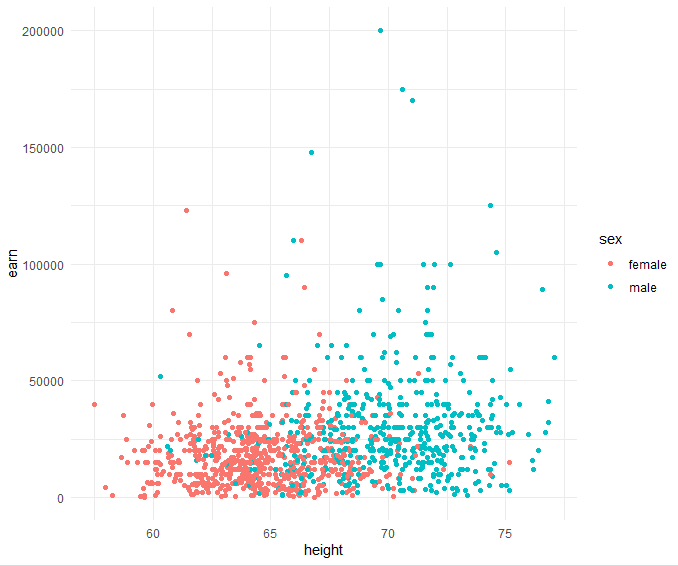
ggplot(heights\_df, aes(x=ed, y=earn)) + geom\_point() +

geom\_smooth(method = NULL, colour = "Blue")



## Create a scatterplot of `height`` vs. `earn`. Use `sex` as the `col` (color) attribute

ggplot(heights\_df, aes(x=height, y=earn, col=sex)) + geom\_point()



## Using `ggtitle()`, `xlab()`, and `ylab()` to add a title, x label, and y label to the previous plot

## Title: Height vs. Earnings

## X label: Height (Inches)

## Y Label: Earnings (Dollars)

ggplot(heights\_df, aes(x=height, y=earn, col=sex)) +

geom\_point() +

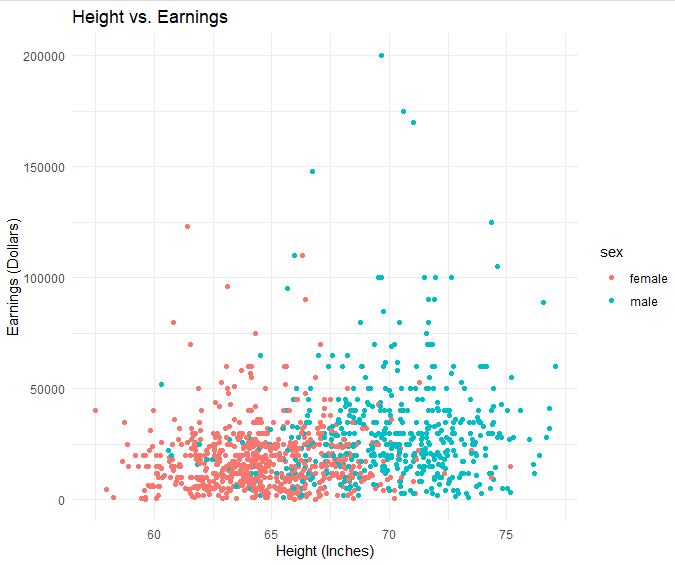
labs(x = "Height (Inches)", y = "Earnings (Dollars)", title = "Height vs. Earnings")

or

ggplot(heights\_df, aes(x=height, y=earn, col=sex)) +

geom\_point() +

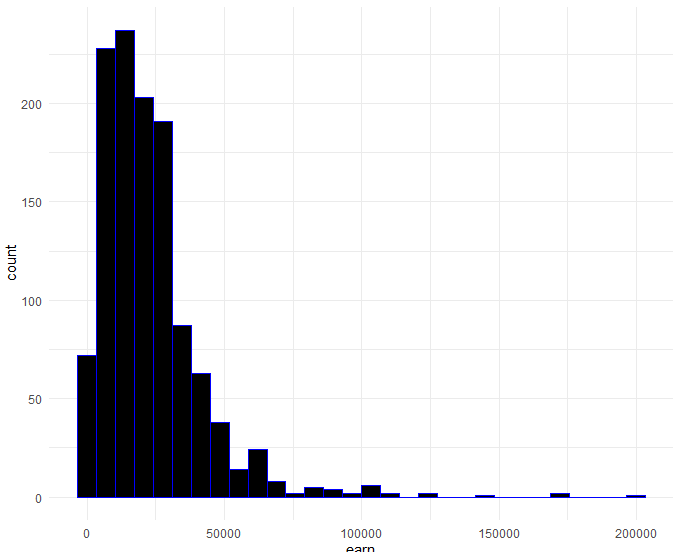
xlab("Height (Inches)") + ylab("Earnings (Dollars)")+ggtitle("Height vs. Earnings")



## https://ggplot2.tidyverse.org/reference/geom\_histogram.html

## Create a histogram of the `earn` variable using `geom\_histogram()`

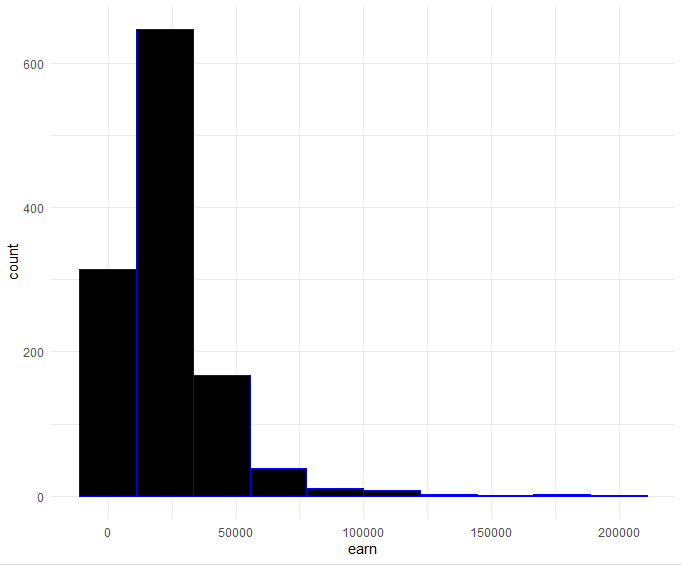
ggplot(heights\_df, aes(earn)) + geom\_histogram(colour="Blue", fill="black")



## Create a histogram of the `earn` variable using `geom\_histogram()`

## Use 10 bins

ggplot(heights\_df, aes(earn)) + geom\_histogram(colour="Blue", fill="black",bins=10)



# https://ggplot2.tidyverse.org/reference/geom\_density.html

## Create a kernel density plot of `earn` using `geom\_density()`

ggplot(heights\_df, aes(earn)) + geom\_density(color="darkblue", fill="lightblue")

