# Assignment: ASSIGNMENT 3

# Name: Kalaikkovan, Vasanthakumar

# Date: 2010-03-30

## Load the ggplot2 package

library(ggplot2)

theme\_set(theme\_minimal())

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

setwd("E:/Repos/StatisticsR/DSC520-Statistics/")

## 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 diamonds,aes(x=carat,y=price))+geom\_point()

## `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="gam",formula=y~s(x,bs="cs"))

## `age` vs. `earn`

ggplot(heights\_df, aes(x=age, y=earn)) + geom\_point() + geom\_smooth(method = "gam",formula = y~s(x,bs="cs"))

## `ed` vs. `earn`

ggplot(heights\_df, aes(x=ed, y=earn)) + geom\_point() + geom\_smooth(method = "gam",formula = y~s(x,bs="cs"))

## 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() + 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()

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

## Use 10 bins

ggplot(heights\_df, aes(earn)) + geom\_histogram(binwidth=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()