

```

# Assignment: ASSIGNMENT 3
# Name: Koby-Hercsky, Theodore
# Date: 2021-03-31

## Load the ggplot2 package
library(ggplot2)
theme_set(theme_minimal())

## Set the working directory to the root of your DSC 520 directory
setwd("/home/downloads/DSC520/dsc520")

## Load the `data/r4ds/heights.csv` to
heights_df <- read_csv("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(aes(color=sex))
## `age` vs. `earn`
ggplot(heights_df, aes(x=age, y=earn)) + geom_point(aes(color=race))
## `ed` vs. `earn`
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point(aes(alpha=age))

## 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(aes(color=sex)) + geom_smooth(orientation = "x")
## `age` vs. `earn`
ggplot(heights_df, aes(x=age, y=earn)) + geom_point(aes(color=race)) + geom_smooth(span = 0.5)
## `ed` vs. `earn`
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point(aes(alpha=age)) + geom_smooth(se = FALSE, method = lm)

## 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(aes(alpha=age))

## 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(aes(alpha=age)) + ggtitle("Height vs. Earnings") +
xlab("Height (Inches)") + ylab("Earnings (Dollars)")

# 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(binwidth = 50000)

## Create a histogram of the `earn` variable using `geom_histogram()`
## Use 10 bins
ggplot(heights_df, aes(earn)) + geom_histogram(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(kernel = "gaussian")

```