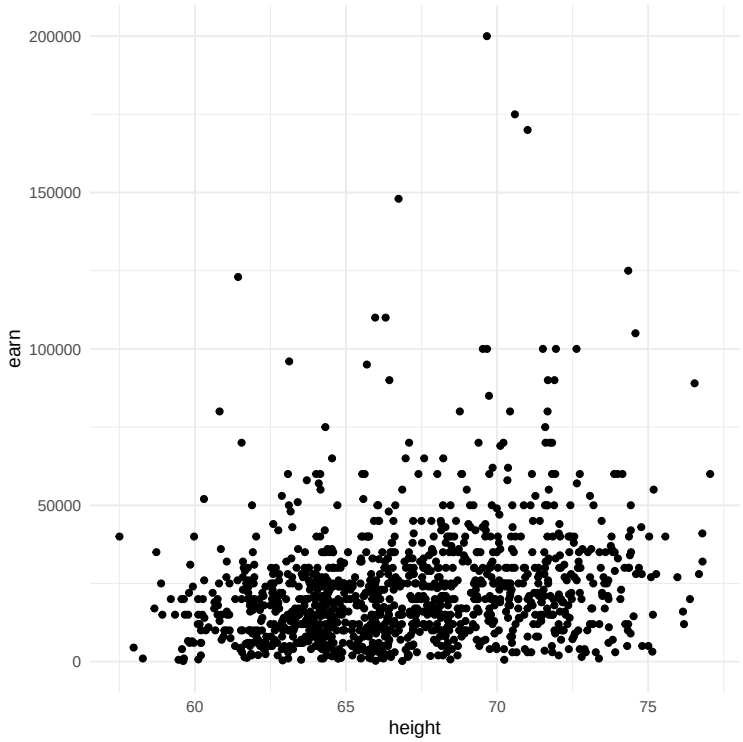
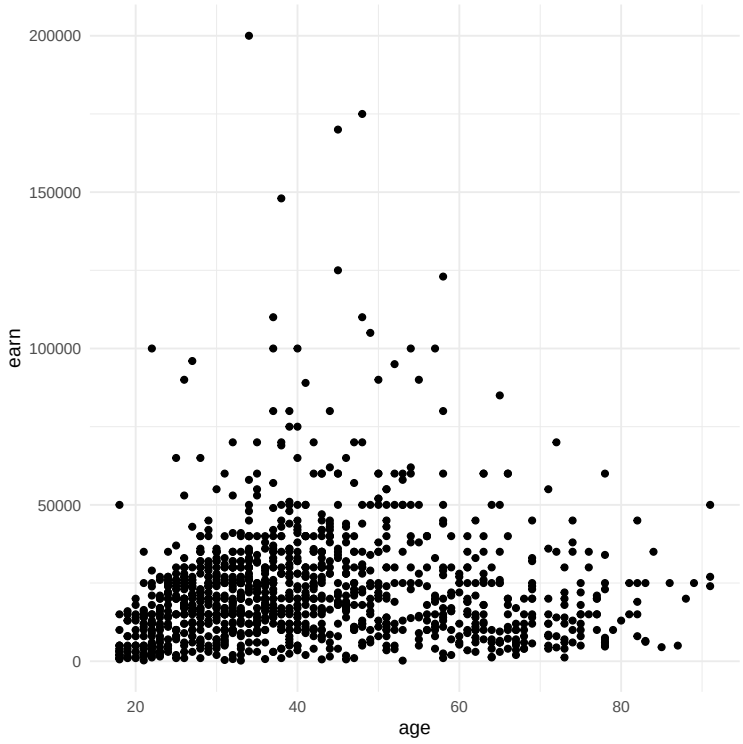
The results below are generated from an R script.

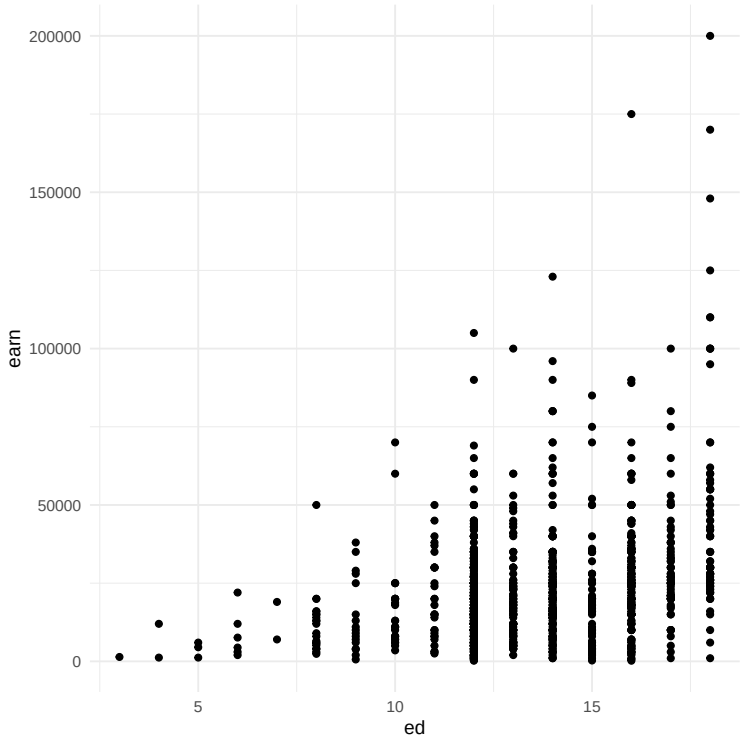
```
# Assignment: ASSIGNMENT 3
# Name: Lastname, Firstname
# Date: 2010-02-14

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

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

## 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()
```



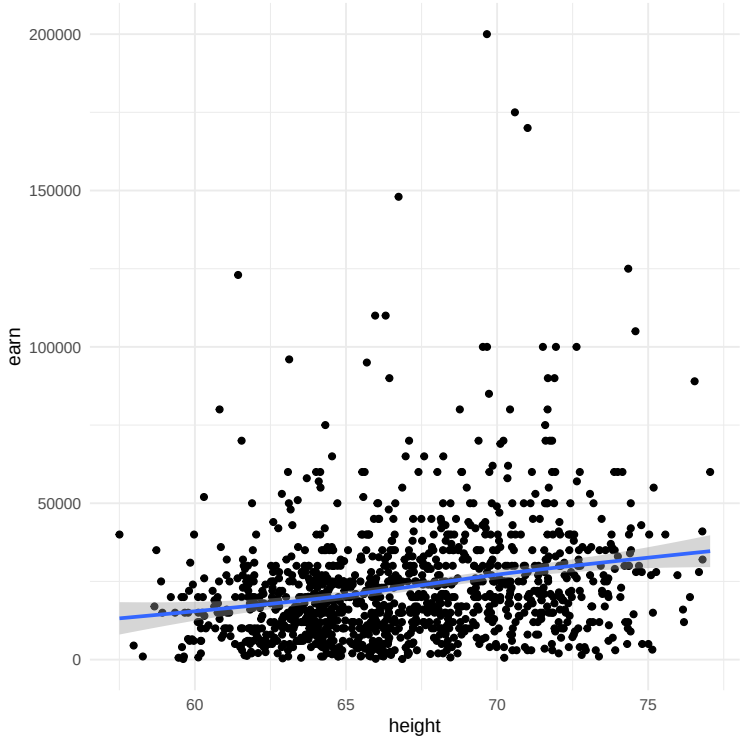

```
## `age` vs. `earn`
ggplot(heights_df, aes(x=age, y=earn)) + geom_point()
```



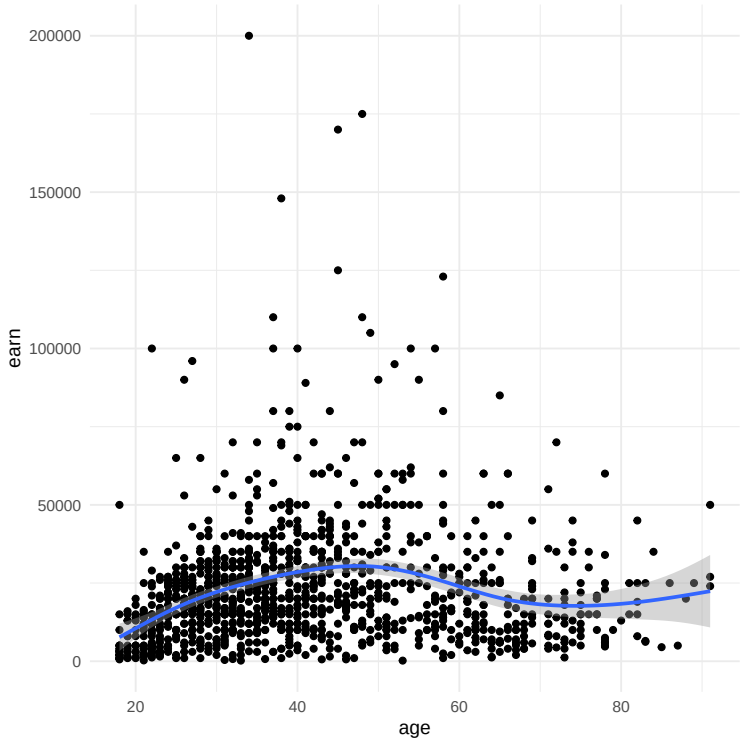

```
## `ed` vs. `earn`
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point()
```



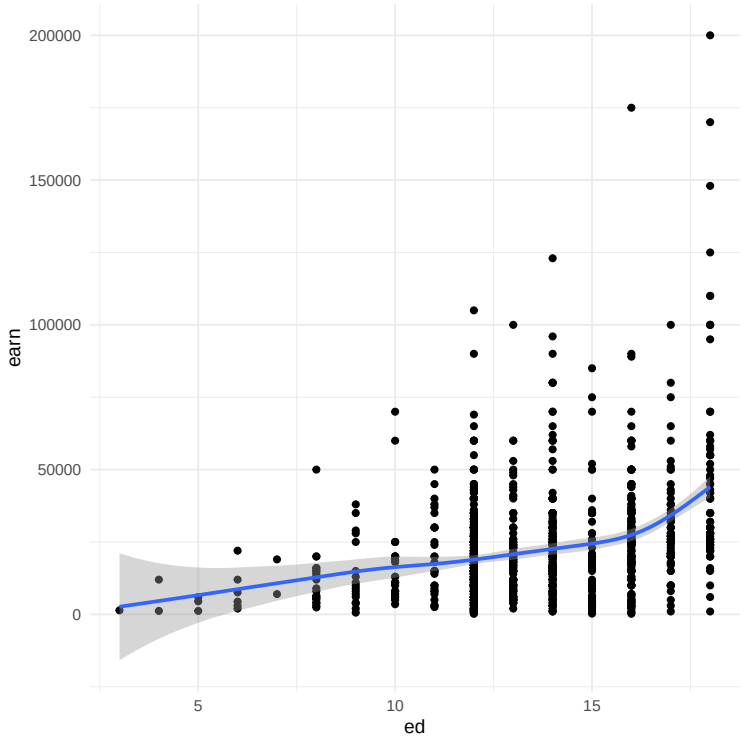
```
## Re-create the three scatterplots and add a regression trend line using
## the `geom_smooth()` function
## `height` vs. `earn`
ggplot(heights_df, aes(x=height, y=earn)) + geom_point() + geom_smooth()

## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



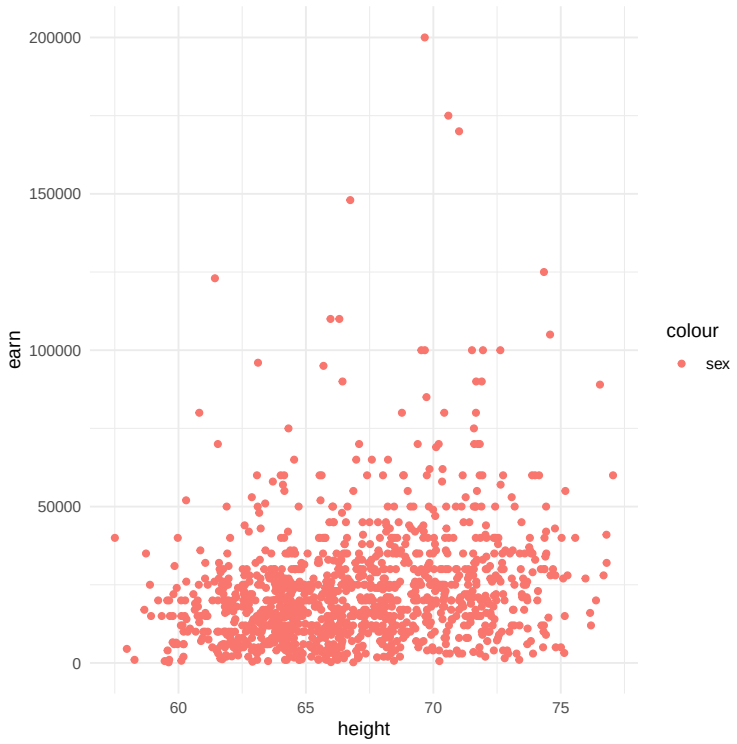

```
## `age` vs. `earn`
ggplot(heights_df, aes(x=age, y=earn)) + geom_point() + geom_smooth()

## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



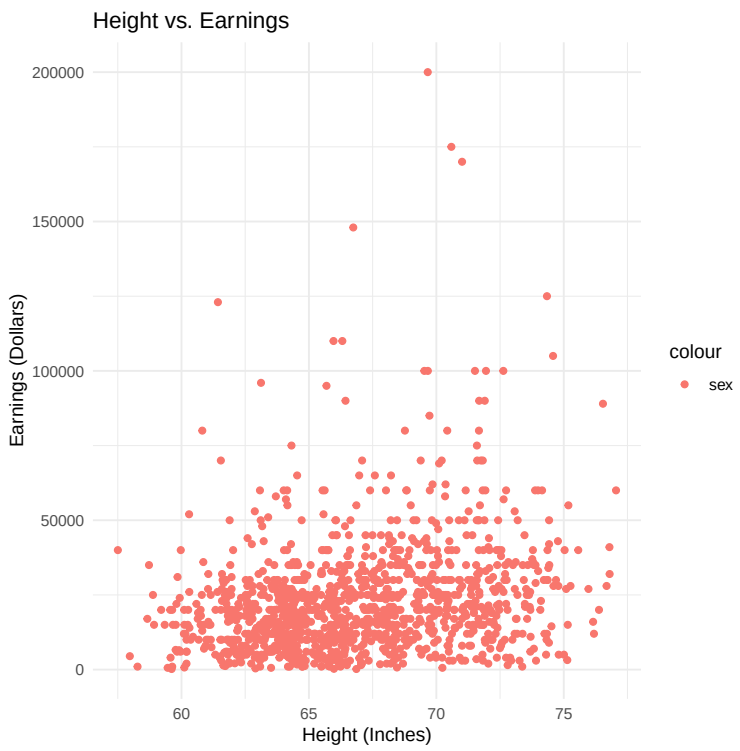

```
## `ed` vs. `earn`
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point() + geom_smooth()

## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



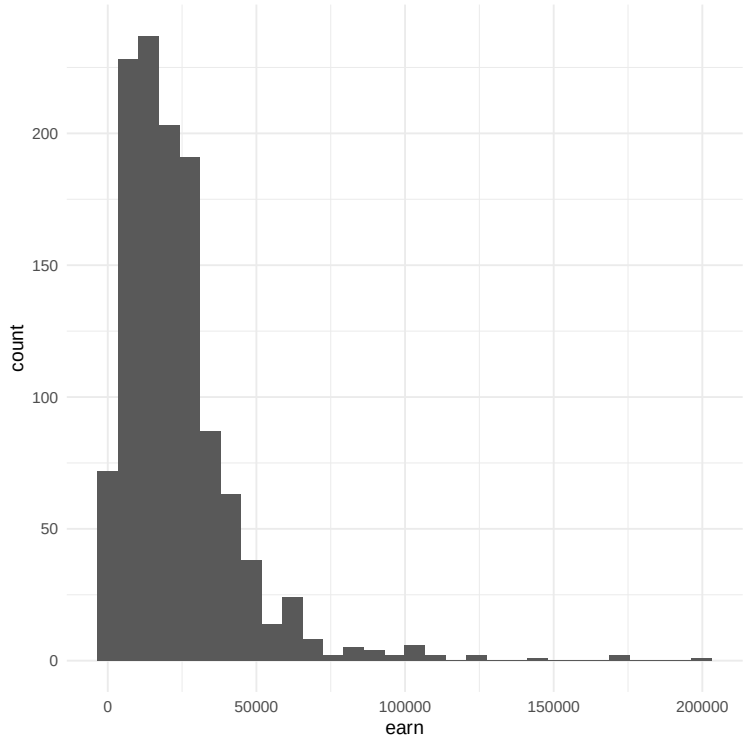
```
## Create a scatterplot of `height`` vs. `earn`.  Use `sex` as the `col` (color) attribute
ggplot(heights_df, aes(x=height, y=earn, col='sex')) + geom_point()
```



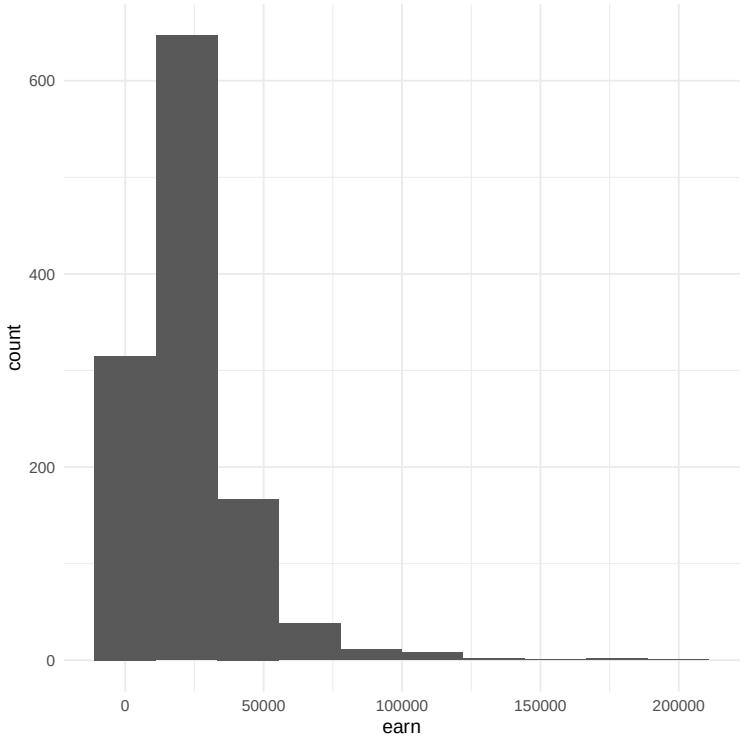

```
## Using `ggtitle()`, `xlab()`, and `ylab()` to add a title, x label, and y label to the previous plot
## Title: Height vs. Earnings
## X label: Height (Inches)
## Y Label: Earnings (Dollars)
ggplot(heights_df, aes(x=height, y=earn, col='sex')) + geom_point() + ggtitle("Height vs. Earnings") + x
```



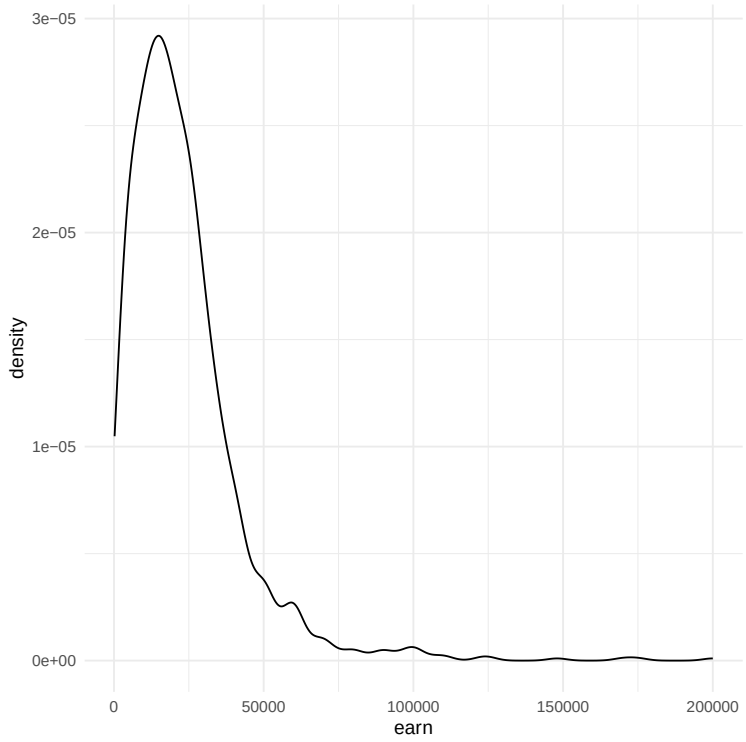

```
# https://ggplot2.tidyverse.org/reference/geom_histogram.html
## Create a histogram of the 'earn' variable using 'geom_histogram()'
ggplot(heights_df, aes(earn)) + geom_histogram()

## 'stat_bin()' using 'bins = 30'.  Pick better value with 'binwidth'.
```



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




```r
# 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()
```



The R session information (including the OS info, R version and all packages used):

```
sessionInfo()
```

```
## R version 4.3.0 (2023-04-21 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##
##
## locale:
## [1] LC_COLLATE=English_United States.utf8  LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8 LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: America/Chicago
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] tinytex_0.45   knitr_1.43     lattice_0.21-8 ggplot2_3.4.2
##
## loaded via a namespace (and not attached):
##  [1] vctrs_0.6.2      nlme_3.1-162     cli_3.6.1        rlang_1.1.1      xfun_0.39
##  [6] highr_0.10       glue_1.6.2       labeling_0.4.2   colorspace_2.1-0 scales_1.2.1
## [11] fansi_1.0.4      grid_4.3.0       evaluate_0.21    munsell_0.5.0    tibble_3.2.1
## [16] lifecycle_1.0.3  compiler_4.3.0   pkgconfig_2.0.3  mgcv_1.8-42      farver_2.1.1
## [21] R6_2.5.1         utf8_1.2.3       pillar_1.9.0     splines_4.3.0    magrittr_2.0.3
```

```
## [26] Matrix_1.5-4     tools_4.3.0      withr_2.5.0      gtable_0.3.3
```

```
Sys.time()
```

```
## [1] "2023-06-16 11:51:55 CDT"
```