

June 25, 2023

The results below are generated from an R script.

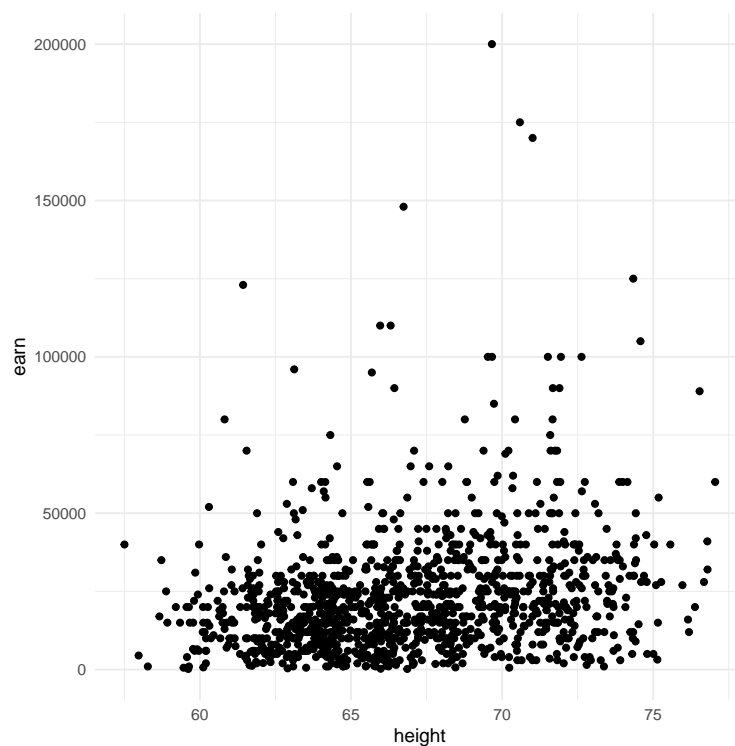
```
# Assignment: ASSIGNMENT 3
# Name: Smith, David
# Date: 2023-06-25

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

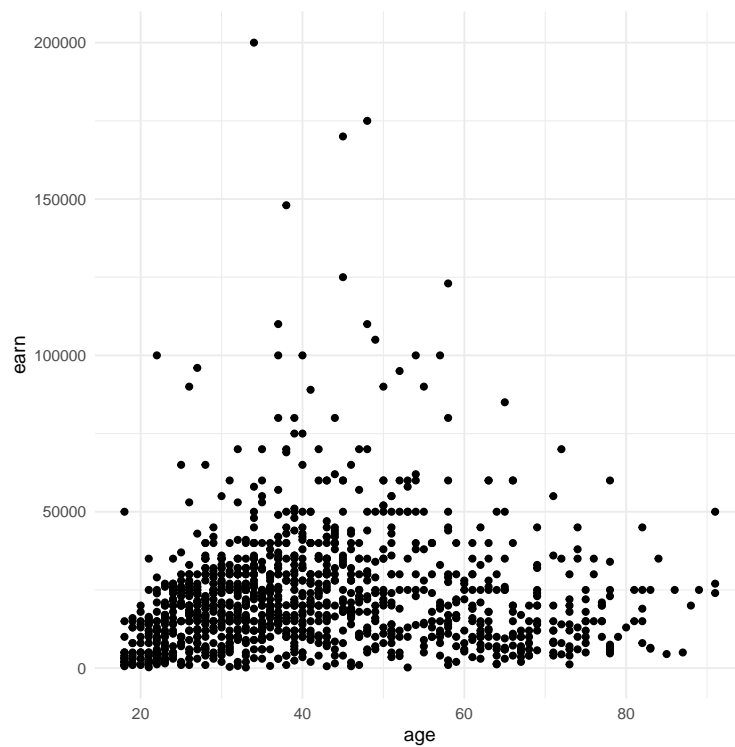
## Set the working directory to the root of your DSC 520 directory
setwd("F:\\GitLab-Projects\\Bellevue\\SMITH-DSC520")

## Load the 'data/r4ds/heights.csv' to
heights_df <- read.csv("data/r4ds/heights.csv")

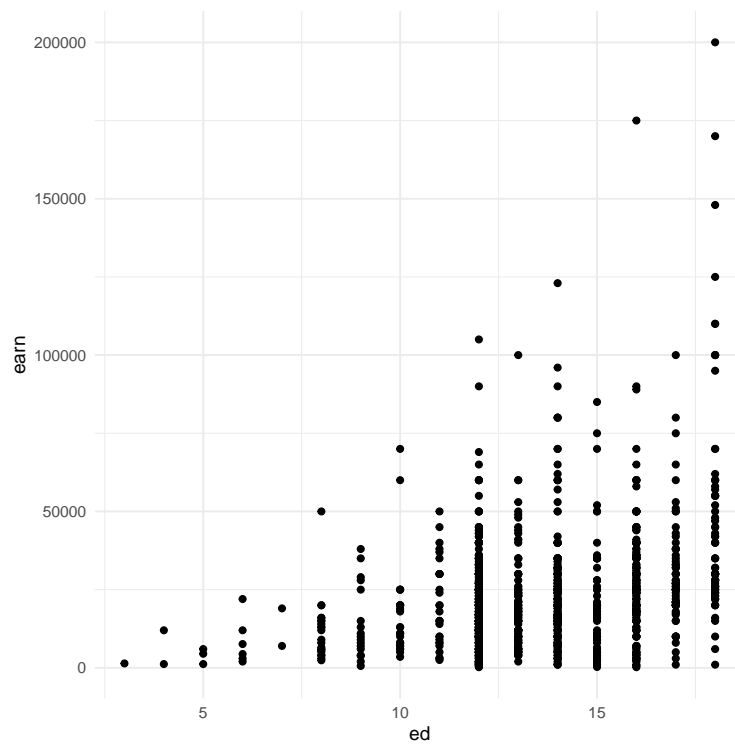
# https://ggplot2.tidyverse.org/reference/geom_point.html
## Using 'geom_point()' create three scatterplots for
## 'height' vs. 'earn'
ggplot(heights_df, aes(x=height, y=earn)) + geom_point()
```



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

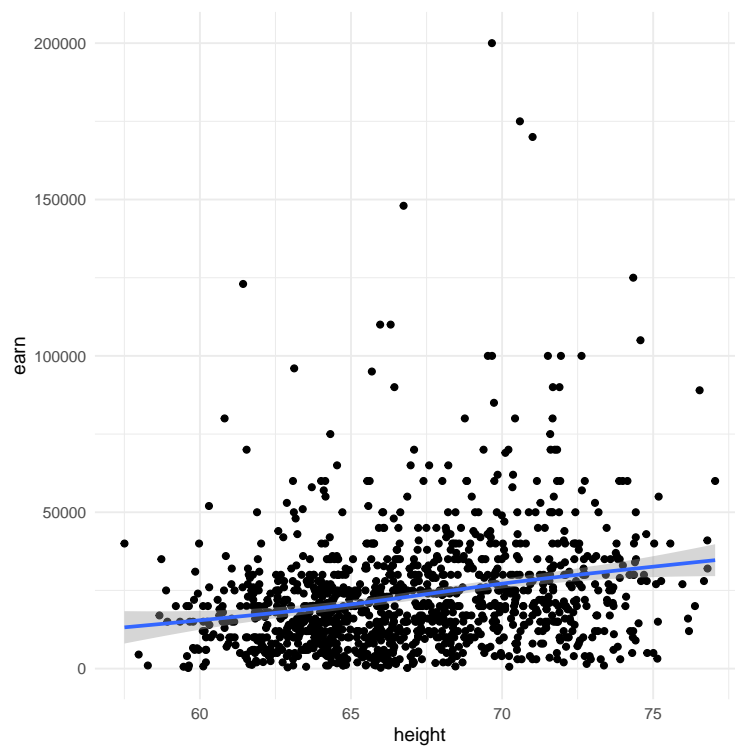


```
## 'ed' vs. 'earn'  
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point()
```



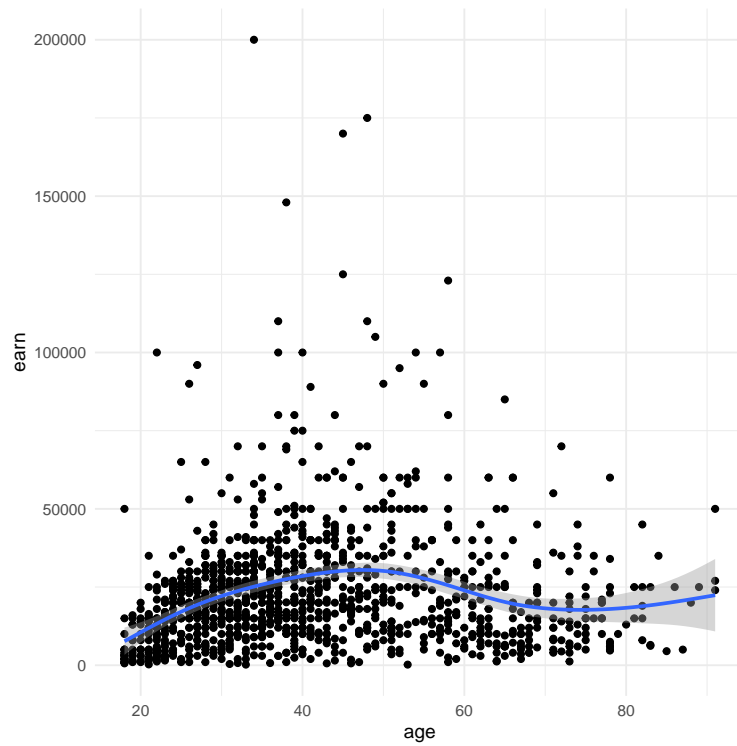
```
## Re-create the three scatterplots and add a regression trend line using
## the 'geom_smooth()' function
## 'height' vs. 'earn'
ggplot(heights_df, aes(x=height, y=earn)) + geom_point() + geom_smooth()

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



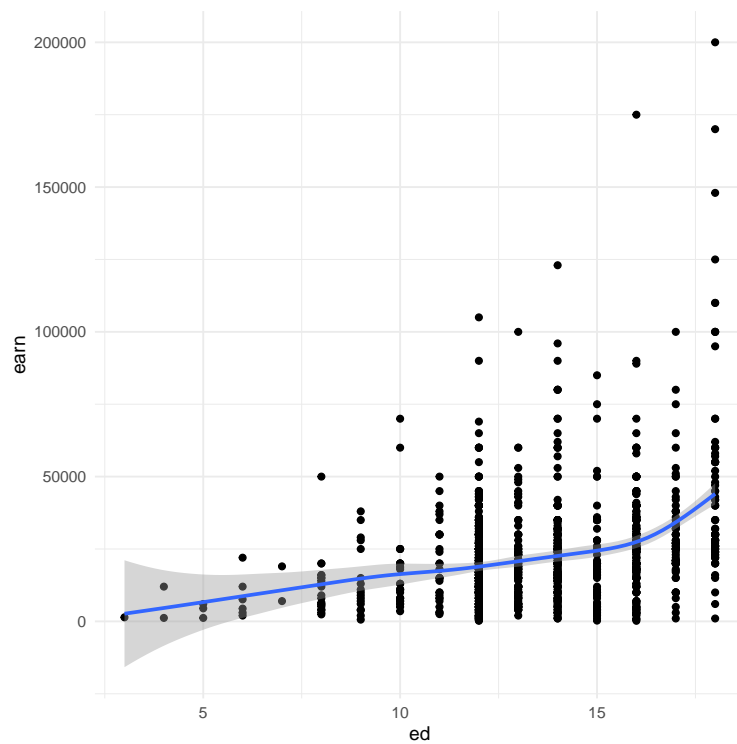
```
## 'age' vs. 'earn'
ggplot(heights_df, aes(x=age, y=earn)) + geom_point() + geom_smooth()

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

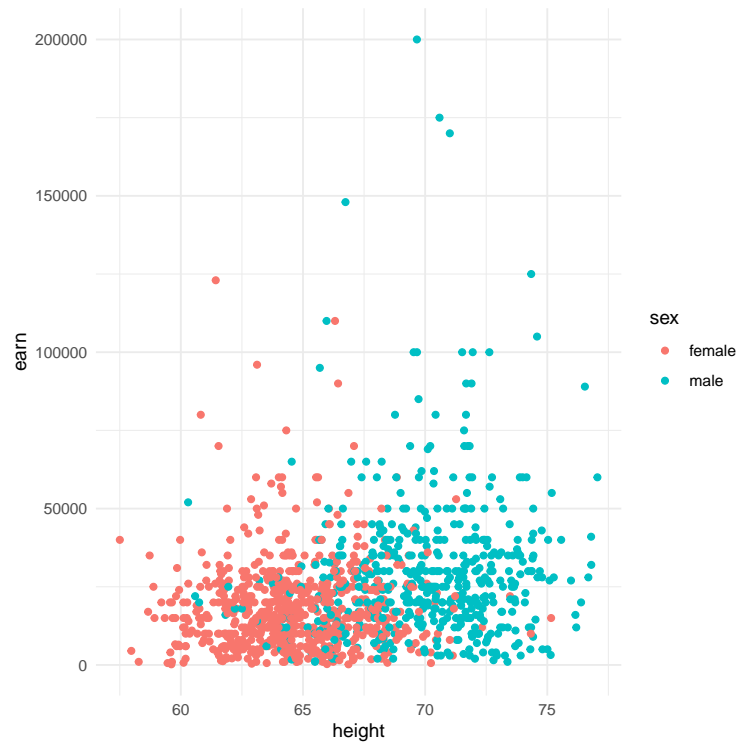


```
## 'ed' vs. 'earn'
ggplot(heights_df, aes(x=ed, y=earn)) + geom_point() + geom_smooth()

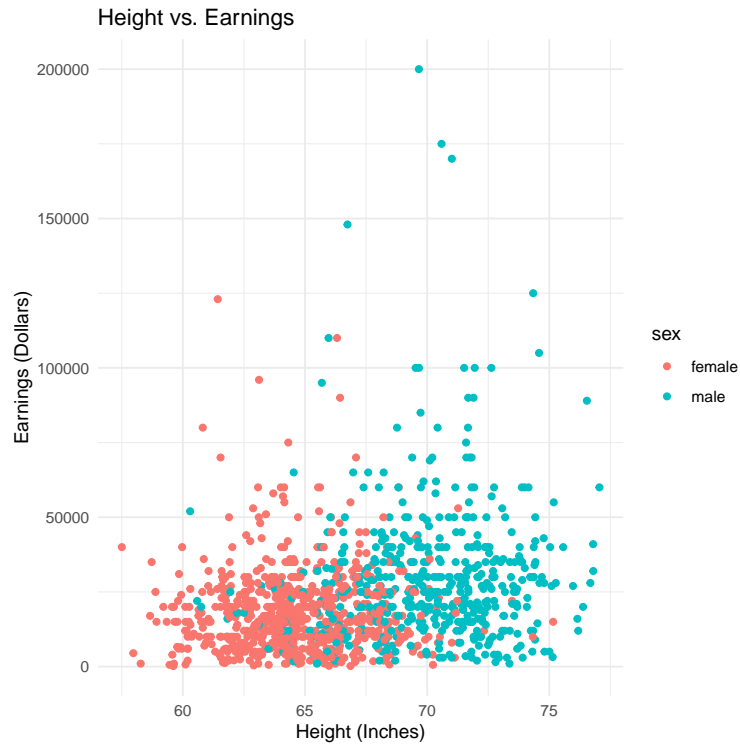
## 'geom_smooth()' using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```



```
## 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()
```

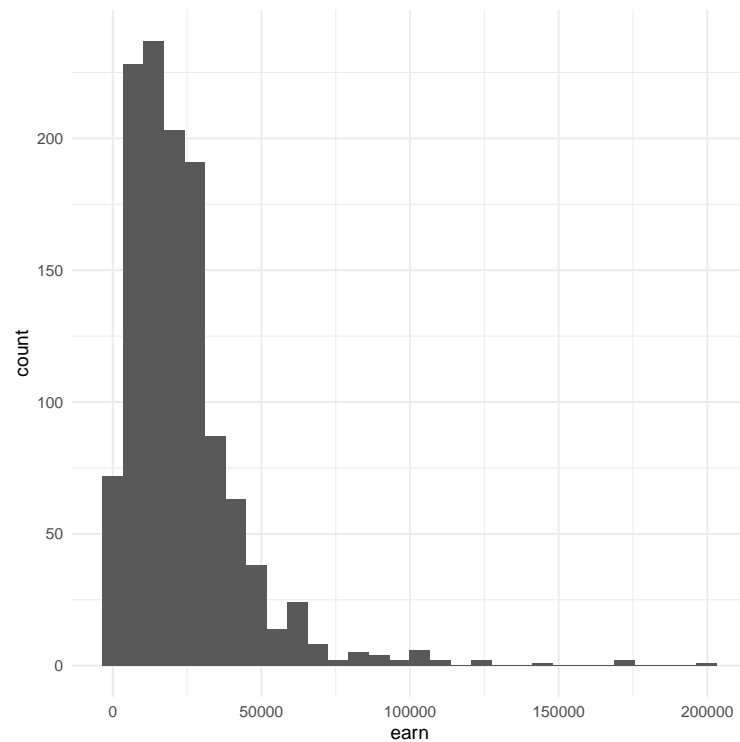


```
## Using 'ggtitle()', 'xlab()', and 'ylab()' to add a title, x label, and y label to the previous plot
## Title: Height vs. Earnings
## X label: Height (Inches)
## Y Label: Earnings (Dollars)
ggplot(heights_df, aes(x=height, y=earn, col=sex)) + geom_point() + ggtitle("Height vs. Earnings") + xlab("Height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

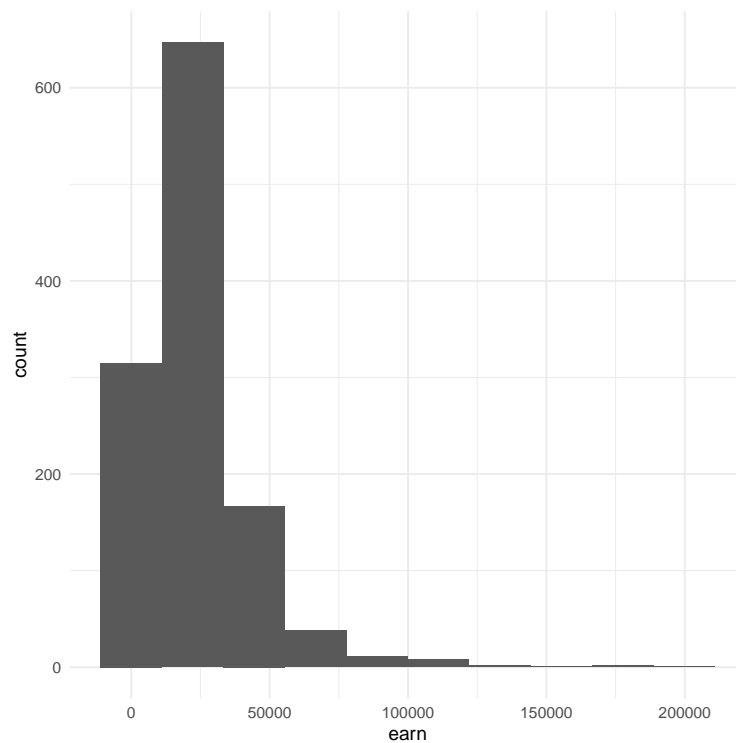


```
# https://ggplot2.tidyverse.org/reference/geom\_histogram.html
## Create a histogram of the 'earn' variable using 'geom_histogram()'
ggplot(heights_df, aes(earn)) + geom_histogram()

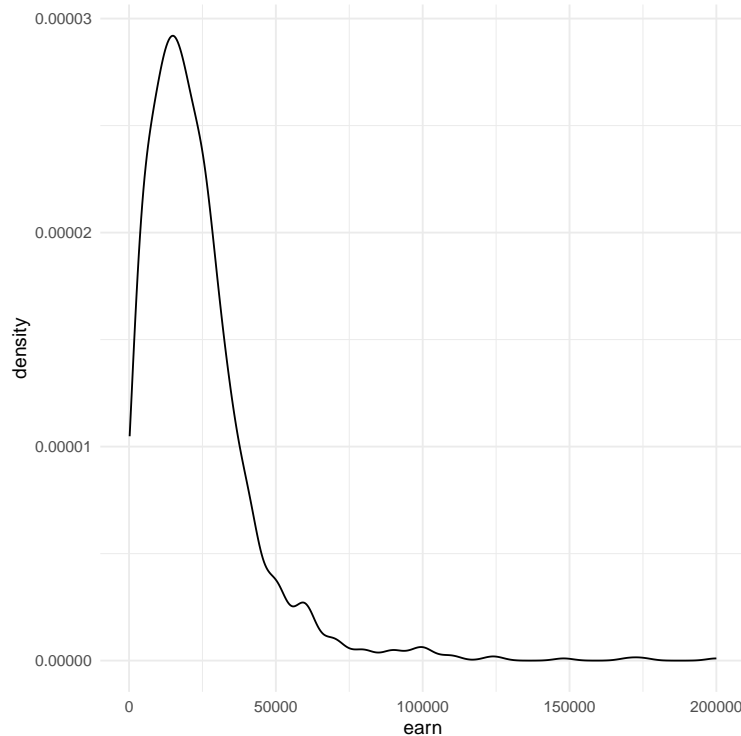
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
## 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()
```



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

```
sessionInfo()

## R version 4.3.1 (2023-06-16 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/New_York
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] RSQLite_2.3.1  moments_0.14.1  pastecs_1.3.21  ggplot2_3.4.2
##
## loaded via a namespace (and not attached):
## [1] Matrix_1.5-4.1  bit_4.0.5       gtable_0.3.3    dplyr_1.1.2     compiler_4.3.1
## [6] highr_0.10      tinytex_0.45    tidyselect_1.2.0 blob_1.2.4      splines_4.3.1
## [11] scales_1.2.1    boot_1.3-28.1   fastmap_1.1.1   lattice_0.21-8  R6_2.5.1
## [16] labeling_0.4.2  generics_0.1.3  knitr_1.43      tibble_3.2.1    munsell_0.5.0
## [21] DBI_1.1.3       pillar_1.9.0    rlang_1.1.1     utf8_1.2.3      cachem_1.0.8
```



```
## [26] xfun_0.39      bit64_4.0.5      memoise_2.0.1     cli_3.6.1         mgcv_1.8-42
## [31] withr_2.5.0     magrittr_2.0.3    grid_4.3.1        rstudioapi_0.14   nlme_3.1-162
## [36] lifecycle_1.0.3 vctrs_0.6.3       evaluate_0.21     glue_1.6.2        farver_2.1.1
## [41] fansi_1.0.4     colorspace_2.1-0 tools_4.3.1        pkgconfig_2.0.3

Sys.time()

## [1] "2023-06-25 21:48:28 EDT"
```