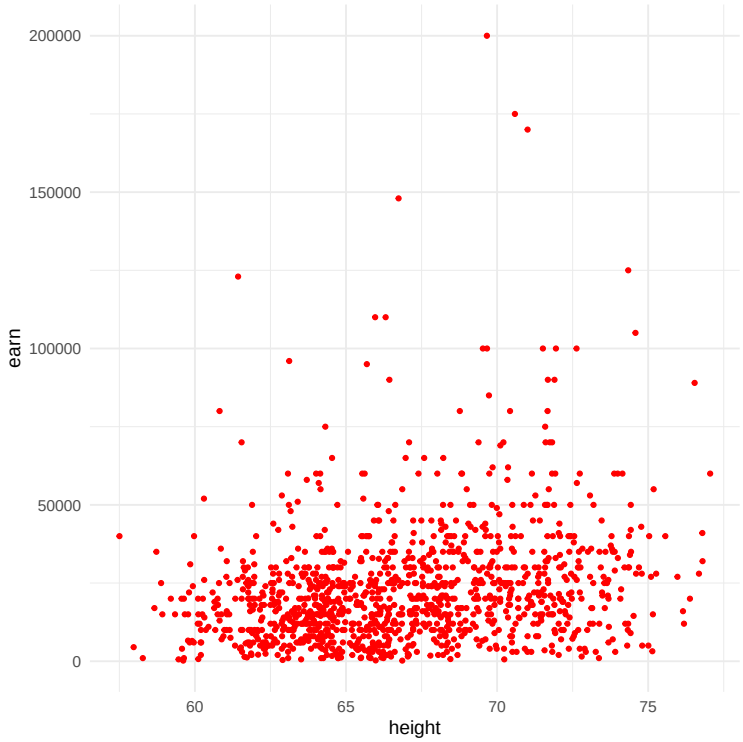
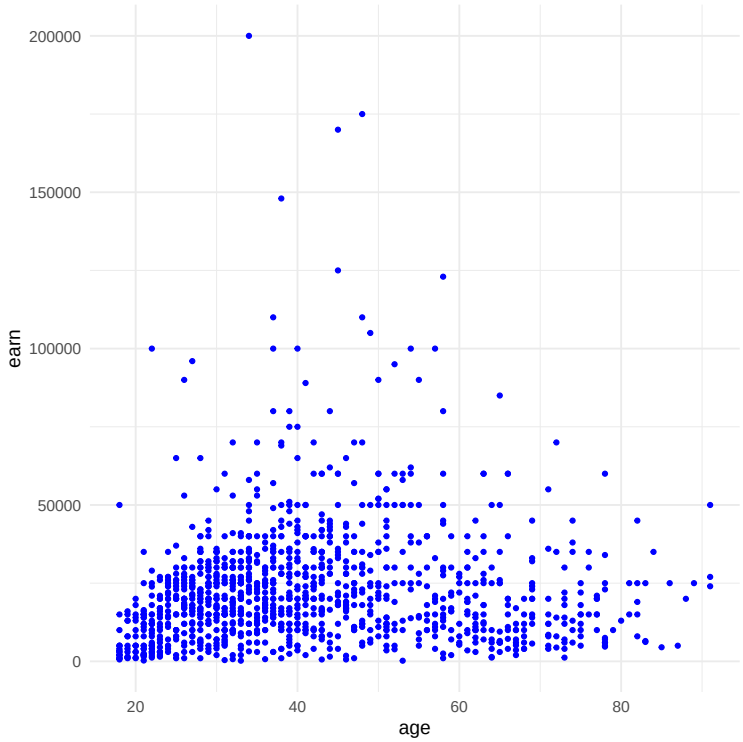
April 2, 2023

The results below are generated from an R script.

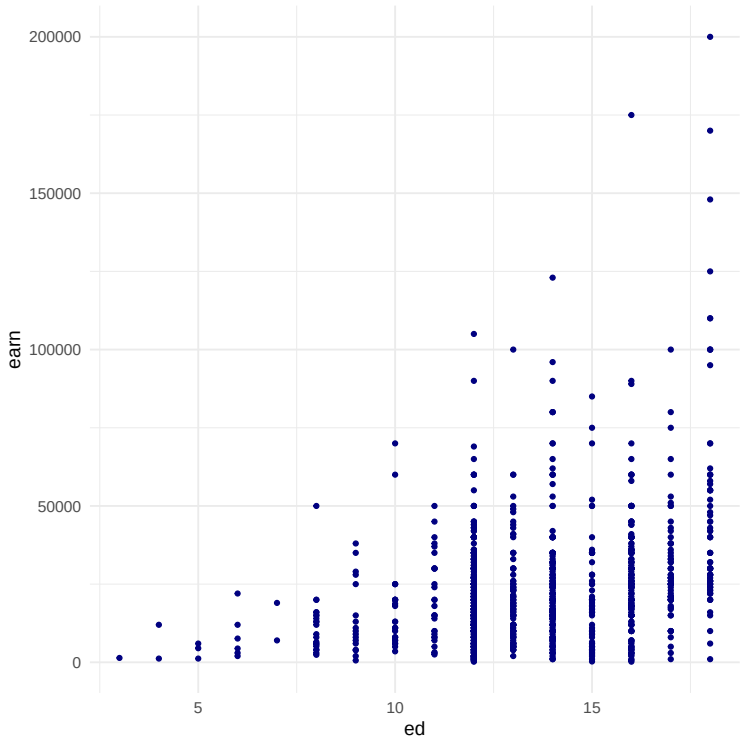
```
# Assignment: ASSIGNMENT 3
# Name: Syverson, Luke
# Date: 2023-04-02

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

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

## Error in setwd("/home/jdoe/Workspaces/dsc520"):  cannot change working directory

## 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`
x <- heights_df # After completing this assignment, I realize I shouldn't use 'x' as a table name shortl
ggplot(x, aes(x=height, y=earn)) + geom_point(colour = "red", size = 1)
```



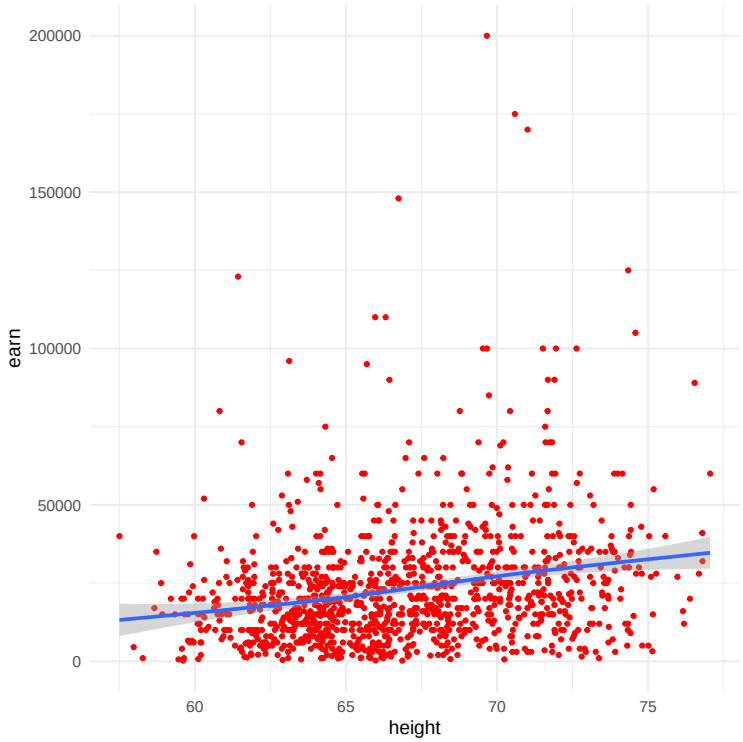

```
## `age` vs. `earn`
ggplot(x, aes(x=age, y=earn)) + geom_point(colour = "blue", size = 1)
```



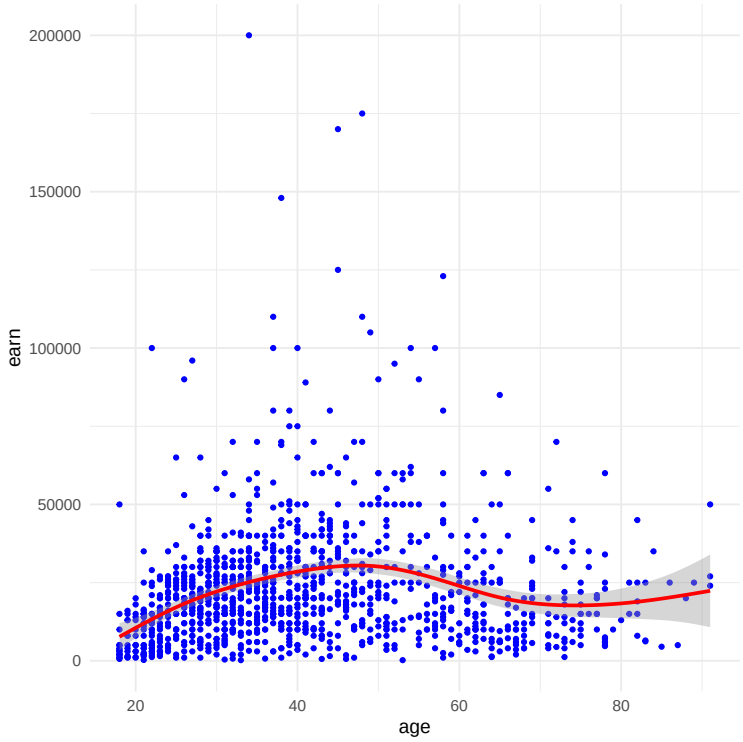
```
## `ed` vs. `earn`
ggplot(x, aes(x=ed, y=earn)) + geom_point(colour = "navy", size = 1)
```



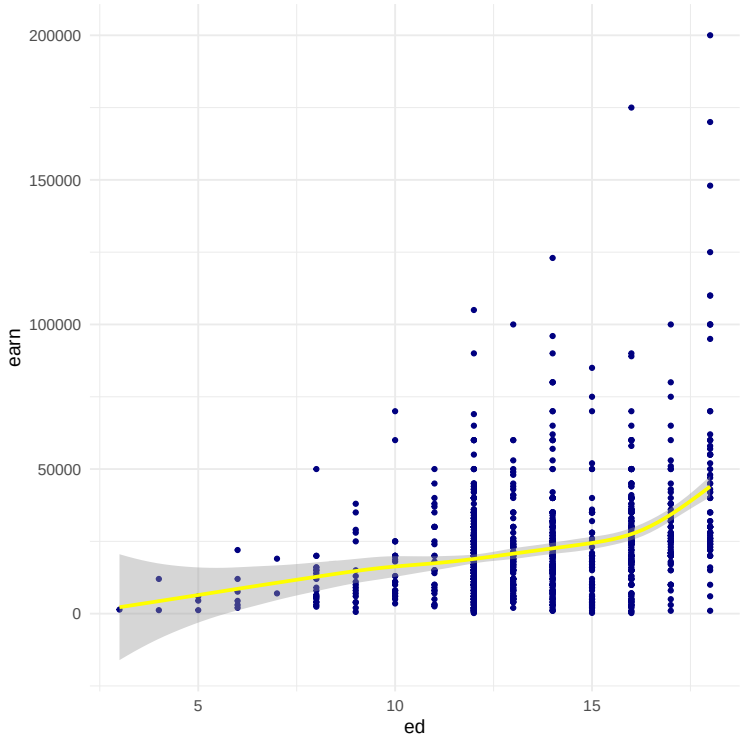

```
## Re-create the three scatterplots and add a regression trend line using
## the `geom_smooth()` function
## `height` vs. `earn`
ggplot(x, aes(x=height, y=earn)) + geom_point(colour = "red", size = 1) + geom_smooth()

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



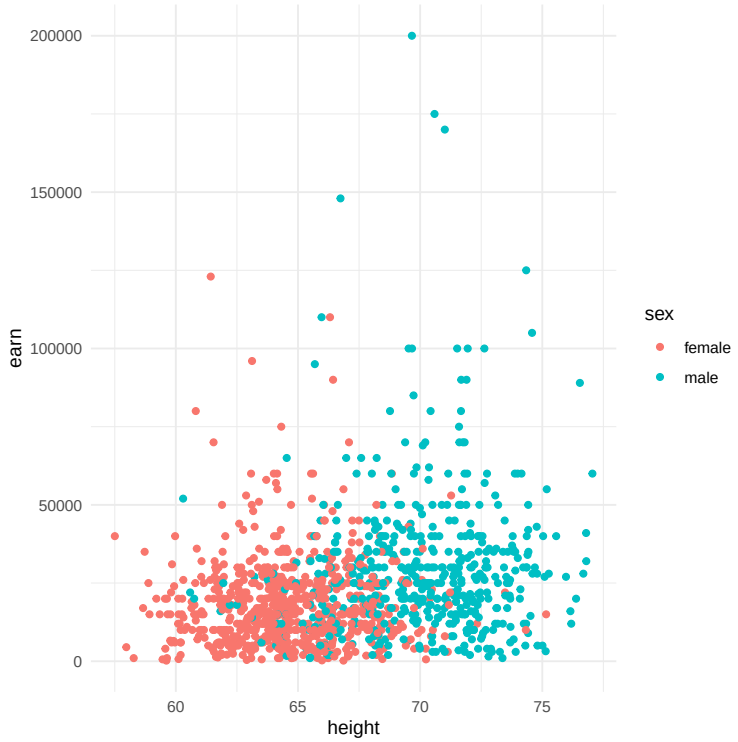

```
## `age` vs. `earn`
ggplot(x, aes(x=age, y=earn)) + geom_point(colour = "blue", size = 1) + geom_smooth(colour = "red")

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



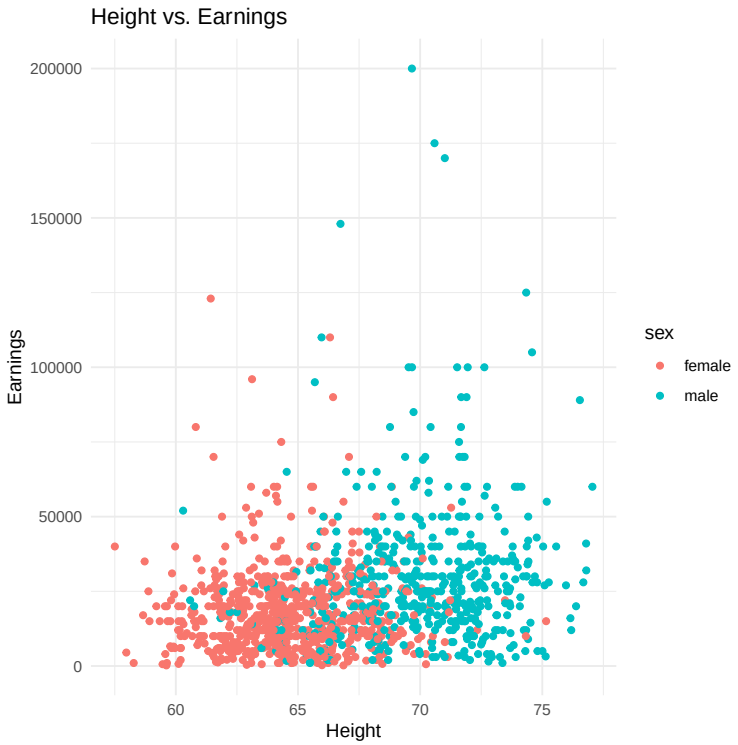
```
## `ed` vs. `earn`
ggplot(x, aes(x=ed, y=earn)) + geom_point(colour = "navy", size = 1) + geom_smooth(colour = "yellow")

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



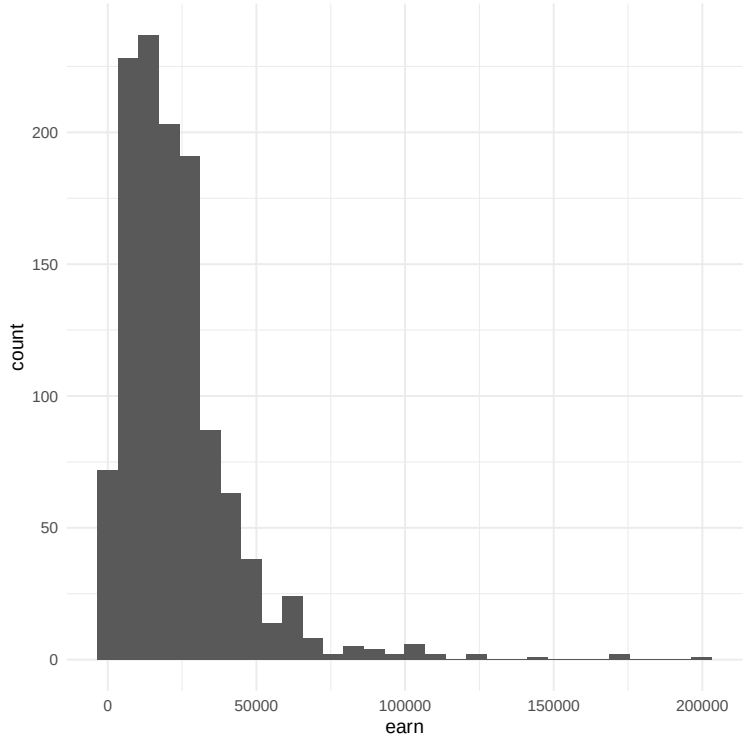

```
## Create a scatterplot of ``height`` vs. `earn`.  Use `sex` as the `col` (color) attribute
ggplot(x, aes(x=height, y=earn, col=sex)) + geom_point()
```



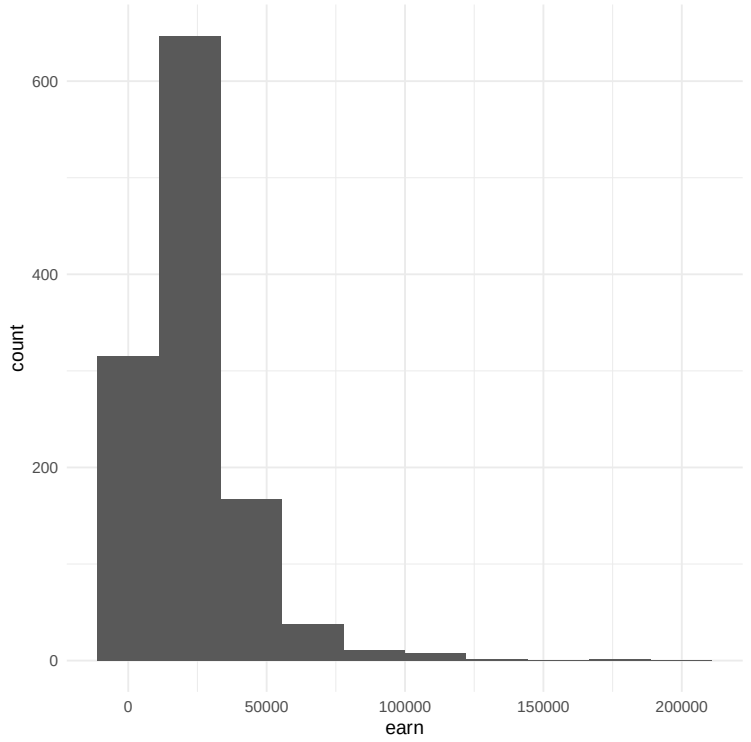

```
## Using 'ggtitle()', 'xlab()', and 'ylab()' to add a title, x label, and y label to the previous plot
## Title: Height vs. Earnings
## X label: Height (Inches)
## Y Label: Earnings (Dollars)
ggplot(x, aes(x=height, y=earn, col=sex)) + geom_point() + xlab("Height") + ylab("Earnings") + ggtitle('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



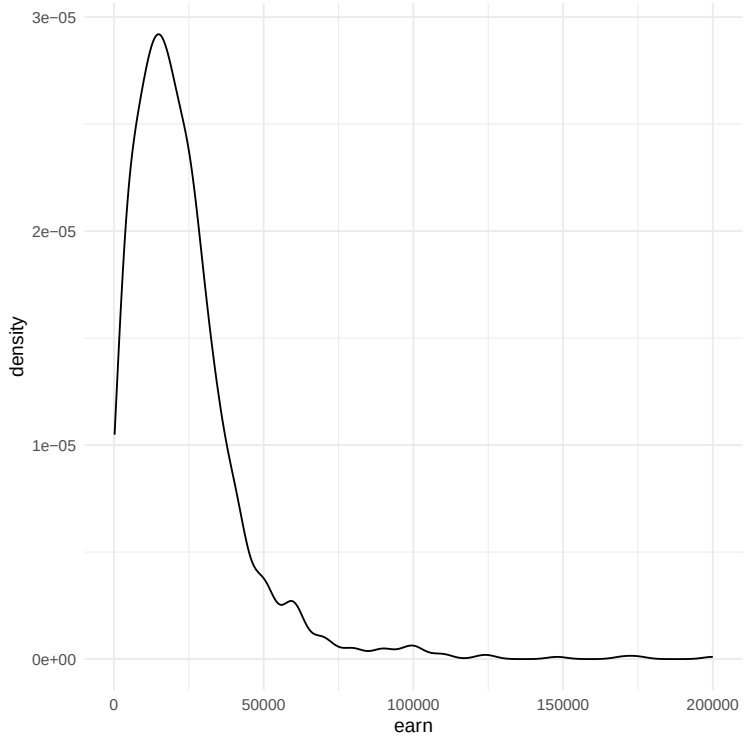
```
# https://ggplot2.tidyverse.org/reference/geom_histogram.html
## Create a histogram of the `earn` variable using `geom_histogram()`
ggplot(x, aes(earn)) + geom_histogram()

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




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



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

```r
sessionInfo()
```

```
## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 22000)
##
## 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
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] knitr_1.42        ggplot2_3.4.1     DescTools_0.99.48 pastecs_1.3.21
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.10       highr_0.10        cellranger_1.1.0  compiler_4.2.3
##  [5] pillar_1.9.0      class_7.3-21      tools_4.2.3       boot_1.3-28.1
##  [9] evaluate_0.20     nlme_3.1-162      lifecycle_1.0.3   tibble_3.2.1
## [13] gtable_0.3.3      rootSolve_1.8.2.3 lattice_0.20-45   mgcv_1.8-42
## [17] pkgconfig_2.0.3   rlang_1.1.0       Matrix_1.5-3      cli_3.6.1
## [21] rstudioapi_0.14   mvtnorm_1.1-3     expm_0.999-7      xfun_0.38
## [25] e1071_1.7-13      withr_2.5.0       dplyr_1.1.1       httr_1.4.5
## [29] generics_0.1.3    vctrs_0.6.1       tidyselect_1.2.0  gld_2.6.6
## [33] grid_4.2.3        glue_1.6.2        data.table_1.14.8 R6_2.5.1
## [37] fansi_1.0.4       readxl_1.4.2      lmom_2.9          farver_2.1.1
## [41] magrittr_2.0.3    splines_4.2.3     scales_1.2.1      MASS_7.3-58.2
## [45] Exact_3.2         colorspace_2.1-0  labeling_0.4.2    tinytex_0.44
## [49] utf8_1.2.3        proxy_0.4-27      munsell_0.5.0
```

```r
Sys.time()
```

```
## [1] "2023-04-02 18:07:02 CDT"
```