July 16, 2023

The results below are generated from an R script.

```
# Assignment: ASSIGNMENT 5
# Name: Lastname, Firstname
# Date: 2010-02-14

## 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")
#View(heights_df)

## Using `cor()` compute correclation coefficients for
## height vs. earn
height_LM <- lm(height ~ earn, data = heights_df)
height_LM
```

```
##
## Call:
## lm(formula = height ~ earn, data = heights_df)
##
## Coefficients:
## (Intercept)         earn
##   6.581e+01    4.787e-05
```

```
cor(heights_df$height, heights_df$earn, use = "everything", method = c("pearson", "kendall", "spearman")
```

```
## [1] 0.2418481
```

```
### age vs. earn
cor(heights_df$age, heights_df$earn, use = "everything", method = c("pearson", "kendall", "spearman"))
```

```
## [1] 0.08100297
```

```
### ed vs. earn
cor(heights_df$ed, heights_df$earn, use = "everything", method = c("pearson", "kendall", "spearman"))
```

```
## [1] 0.3399765
```

```
## Spurious correlation
## The following is data on US spending on science, space, and technology in millions of today's dollars
## and Suicides by hanging strangulation and suffocation for the years 1999 to 2009
## Compute the correlation between these variables
tech_spending <- c(18079, 18594, 19753, 20734, 20831, 23029, 23597, 23584, 25525, 27731, 29449)
suicides <- c(5427, 5688, 6198, 6462, 6635, 7336, 7248, 7491, 8161, 8578, 9000)

cor(tech_spending, suicides, use = "everything", method = c("pearson", "kendall", "spearman"))
```

```
## [1] 0.9920817
```

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/Chicago
## tzcode source: internal
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] dbplyr_2.3.2
##
## loaded via a namespace (and not attached):
##  [1] utf8_1.2.3       R6_2.5.1         xfun_0.39        tidyselect_1.2.0 magrittr_2.0.3
##  [6] glue_1.6.2       tibble_3.2.1     knitr_1.43       pkgconfig_2.0.3  dplyr_1.1.2
## [11] generics_0.1.3   tinytex_0.45     lifecycle_1.0.3  cli_3.6.1        fansi_1.0.4
## [16] vctrs_0.6.2      DBI_1.1.3        compiler_4.3.1   highr_0.10       rstudioapi_0.14
## [21] tools_4.3.1      evaluate_0.21    pillar_1.9.0     rlang_1.1.1

Sys.time()

## [1] "2023-07-16 16:23:24 CDT"
```