

# assignment\_05\_LeathermanKatie.R

ktlea

2022-04-29

```
# Assignment: ASSIGNMENT 5
# Name: Leatherman, Katie
# Date: 2022-04-28
```

```
## Set the working directory to the root of your DSC 520 directory
setwd("C:/Users/ktlea/OneDrive/Documents/Bellevue Local/520/dsc520")
```

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

```
## 'data.frame': 1192 obs. of 6 variables:
## $ earn : num 50000 60000 30000 50000 51000 9000 29000 32000 2000 27000 ...
## $ height: num 74.4 65.5 63.6 63.1 63.4 ...
## $ sex : chr "male" "female" "female" "female" ...
## $ ed : int 16 16 16 16 17 15 12 17 15 12 ...
## $ age : int 45 58 29 91 39 26 49 46 21 26 ...
## $ race : chr "white" "white" "white" "other" ...
```

```
## Using `cor()` compute correclation coefficients for
## height vs. earn
cor(heights_df$height, heights_df$earn)
```

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

```
### age vs. earn
cor(heights_df$age, heights_df$earn)
```

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

```
### ed vs. earn
cor(heights_df$ed, heights_df$earn)
```

```
## [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)
```

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