R Bridge Week 2 Assignment

Donald Butler

7/30/2021

Select a random dataset

US Macroeconomic Data (1957–2005, Stock & Watson) https://vincentarelbundock.github.io/Rdatasets/doc/AER/USMacroSW.html https://vincentarelbundock.github.io/Rdatasets/csv/AER/USMacroSW.csv

```
# github csv location
csvfile <- 'https://raw.githubusercontent.com/dab31415/SPS-Bridge-R/main/USMacroSW.csv'

df <- read.csv(csvfile)
```

Exercise #1.

Use the summary function to gain an overview of the data set. Then display the mean and median for at least two attributes.

```
summary(df)
```

```
##        X           unemp             cpi             ffrate
##  Min.   :  1   Min.   : 3.400   Min.   : 27.78   Min.   : 0.930
##  1st Qu.: 49   1st Qu.: 5.000   1st Qu.: 35.87   1st Qu.: 3.480
##  Median : 97   Median : 5.700   Median : 87.93   Median : 5.400
##  Mean   : 97   Mean   : 5.891   Mean   : 91.73   Mean   : 5.953
##  3rd Qu.:145   3rd Qu.: 6.833   3rd Qu.:143.07   3rd Qu.: 7.760
##  Max.   :193   Max.   :10.667   Max.   :192.17   Max.   :19.100
##      tbill            tbond           gbpusd          gdpjp
##  Min.   : 0.830   Min.   : 1.01   Min.   :112.5   Min.   : 10149
##  1st Qu.: 3.500   1st Qu.: 3.91   1st Qu.:159.6   1st Qu.: 57632
##  Median : 5.080   Median : 5.62   Median :185.5   Median :254560
##  Mean   : 5.435   Mean   : 6.04   Mean   :204.9   Mean   :259306
##  3rd Qu.: 6.740   3rd Qu.: 7.55   3rd Qu.:246.9   3rd Qu.:482328
##  Max.   :15.490   Max.   :16.52   Max.   :281.5   Max.   :523638
```

```
sprintf('3-month treasury bill: mean = %.3f; median = %.3f',mean(df$tbill),median(df$tbill))
```

```
## [1] "3-month treasury bill: mean = 5.435; median = 5.080"
```

```
sprintf('1-year treasury bond: mean = %.3f; median = %.3f',mean(df$tbond),median(df$tbond))
```

```
## [1] "1-year treasury bond: mean = 6.040; median = 5.620"
```

Exercise #2.

Create a new data frame with a subset of the columns and rows. Make sure to rename it.

```
# select the first 50 rows, with columns x, unemp, tbill, and tbond
my_df <- df[1:50,c(1:2,5:6)]
```

Exercise #3.

Create new column names for the new data frame.

```
names(my_df) <- c('index','unemployement_rate','3-month tbill rate','1-year bond rate')
```

Exercise #4.

Use the summary function to create an overview of your new data frame. Then print the mean and median for the same two attributes. Please compare.

```
summary(my_df)
```

```
##      index       unemployement_rate 3-month tbill rate 1-year bond rate
##  Min.   : 1.00   Min.   :3.400      Min.   :0.830      Min.   :1.230
##  1st Qu.:13.25   1st Qu.:3.875      1st Qu.:2.772      1st Qu.:3.098
##  Median :25.50   Median :5.117      Median :3.500      Median :3.875
##  Mean   :25.50   Mean   :5.008      Mean   :3.603      Mean   :4.054
##  3rd Qu.:37.75   3rd Qu.:5.625      3rd Qu.:4.410      3rd Qu.:4.970
##  Max.   :50.00   Max.   :7.367      Max.   :6.440      Max.   :7.040
```

```
sprintf('3-month treasury bill: mean = %.3f; median = %.3f',mean(my_df$`3-month tbill rate`),median(my_
```

```
## [1] "3-month treasury bill: mean = 3.603; median = 3.500"
```

```
sprintf('1-year treasury bond: mean = %.3f; median = %.3f',mean(my_df$`1-year bond rate`),median(my_df$
```

```
## [1] "1-year treasury bond: mean = 4.054; median = 3.875"
```

Exercise #5.

For at least 3 values in a column please rename so that every value in that column is renamed. For example, suppose I have 20 values of the letter “e” in one column. Rename those values so that all 20 would show as “excellent”.

```
my_df$year <- 1957 + floor((my_df$index-1)/4)
my_df$qtr[(my_df$index-1) %% 4 == 0] <- 'First'
my_df$qtr[(my_df$index-1) %% 4 == 1] <- 'Second'
my_df$qtr[(my_df$index-1) %% 4 == 2] <- 'Third'
my_df$qtr[(my_df$index-1) %% 4 == 3] <- 'Fourth'
```

Exercise #6.

Display enough rows to see examples of all of steps 1-5 above.

```
head(my_df,15)
```

```
##    index unemployement_rate 3-month tbill rate 1-year bond rate year    qtr
## 1      1           3.933333               3.08             3.42 1957  First
## 2      2           4.100000               3.29             3.65 1957 Second
## 3      3           4.233333               3.53             4.07 1957  Third
## 4      4           4.933333               3.04             3.18 1957 Fourth
## 5      5           6.300000               1.30             1.84 1958  First
## 6      6           7.366667               0.83             1.23 1958 Second
## 7      7           7.333333               2.44             3.05 1958  Third
## 8      8           6.366667               2.77             3.29 1958 Fourth
## 9      9           5.833333               2.80             3.61 1959  First
## 10    10           5.100000               3.21             4.07 1959 Second
## 11    11           5.266667               4.04             5.00 1959  Third
## 12    12           5.600000               4.49             5.14 1959 Fourth
## 13    13           5.133333               3.31             4.02 1960  First
## 14    14           5.233333               2.46             3.36 1960 Second
## 15    15           5.533333               2.48             3.07 1960  Third
```