# Problem Set #1

EH6105 - Quantitative Methods

Steven V. Miller

This homework makes use of data available in `{stevedata}` and implies the use of `{tidyverse}` to answer the questions. `{tidyverse}` is not necessary to answer these questions though it will assuredly make the process easier. Load these two libraries to get started answering these questions.

```
library(tidyverse)
library(stevedata)
```

## Systemic Banking Crises Database II

This homework assignment will refer to the `SBCD` data set that is available in `{stevedata}`. This data set contains a (very) basic summary of banking, currency, debt, and debt-restructuring crises from 1970 to 2017 and are made available by Luc Laeven and Fabian Valencia (2020) in the *IMF Economic Review*. You can find out more information about the data by visiting this part of the package's website, or with the following command.

```
?SBCD
```

Here's a little preview of these data.

```
SBCD
```

```
## # A tibble: 547 x 4
##    country type               year month
##    <chr>   <chr>             <dbl> <int>
##  1 Albania banking            1994    NA
##  2 Albania currency           1997     1
##  3 Albania debt               1990    NA
##  4 Albania debtrestructuring  1992    NA
##  5 Algeria banking            1990    NA
##  6 Algeria currency           1988    11
##  7 Algeria currency           1994     4
##  8 Angola  currency           1991     3
##  9 Angola  debt               1988    NA
## 10 Angola  debtrestructuring  1992    NA
## # ... with 537 more rows
```

Answer these questions. A successful answer of these question must include the R code you used to help you answer the question. Each question is worth a point.

1. Which country appears the most in these data?
2. How many distinct countries are there in these data?
3. What is the most common type of crisis to occur in these data? What is the least common?
4. Which year had the most crises?
5. How many crises were there in 2008, the height of the Great Recession?
6. There are three occurrences for Sweden in these data. Reproduce these observations for me.
7. These data are helpfully already structured alphabetically by country (in English), and then chronologically within country (i.e. the earliest banking crises are listed first for each country). For each country in the data, show me how you could create a data/table that lists the first occurrence of a banking crisis.[1]
8. Subset the data to just the crises involving Denmark, Finland, Iceland, Norway, and Sweden. Assign this subset of the data to a new object called `Nordics` and reproduce these observations here.
9. The `month` variable is mostly missing data as the authors do not often record the month of the crisis. If I wanted to get rid of this column, how might I do that?
10. Suppose I wanted to create a `cold_war` dummy variable that equals 1 if the crisis occurred on or before 1990 and equals a 0 if it occurred afterward. How might I do that?

[1] Here's a hint: this condensed data/table would be 157 rows long. You don't necessarily have to print all 157 rows.