Complete the following tasks:

* Find a dataset that you find interesting.
* Once you have identified a dataset, explain how you accessed it. What links did you click? Was there any registration required? Did you download directly or was there an online system you navigated?

Logged into UTSA Library > Databases > Demography > RAND State Statistics > Logged in again using UTSA log in > Population & Demographics > Border Crossings > Selected Categories, Border Port Names, & Timeframe > Download Document

* In R:
  + Import the data
  + Provide a list of variables in the dataset
    - Copy and paste this list in your word document

[1] "State" "Crossing" "Category" "2013" "2014" "2015" "2016" "2017"

* + What is the structure of the data? Which variables are character and which are numeric?
    - Copy and paste this list in your word document

Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 36 obs. of 8 variables:

$ State : chr "Texas" "Texas" "Texas" "Texas" ...

$ Crossing: chr "Boquillas" "Boquillas" "Boquillas" "Brownsville" ...

$ Category: chr "Incoming Passengers in Personal Vehicles" "Incoming Passengers on Buses" "Incoming Pedestrian Crossings" "Incoming Passengers in Personal Vehicles" ...

$ 2013 : num NA NA NA 8468635 8443 ...

$ 2014 : num NA NA NA 8527359 53493 ...

$ 2015 : num 0 0 7007 8576187 48583 ...

$ 2016 : num 0 0 14099 9428627 46395 ...

$ 2017 : num 0 0 10965 10047891 43733 ...

* + Describe the data through Central Tendencies:
    - Give me the mean, median, and mode of **five variables**.

> summary(TxBorderCrossing$`2013`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

0 21595 600347 2272189 2459890 17545433 3

> summary(TxBorderCrossing$`2014`)

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

0 56644 748772 2434183 2683861 19134740 3

> summary(TxBorderCrossing$`2015`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0 20424 430044 2297787 2440429 19982407

> summary(TxBorderCrossing$`2016`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0 22705 471336 2380660 2599926 20767737

> summary(TxBorderCrossing$`2017`)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0 28718 456142 2394135 2725038 22046772

> mode(TxBorderCrossing$`2013`)

[1] 0

> mode(TxBorderCrossing$`2014`)

[1] 0

> mode(TxBorderCrossing$`2015`)

[1] 0

> mode(TxBorderCrossing$`2016`)

[1] 0

> mode(TxBorderCrossing$`2017`)

[1] 0

* + Describe the data through variation:
    - Choose one variable and give me:
    - The variance, range, and standard deviation.

> describe(TxBorderCrossing$`2017`)

vars n mean **sd** median trimmed mad min max **range** skew kurtosis se

X1 1 36 2394135 **4472522** 456142 1453515 676276.1 0 22046772 **22046772** 2.72 8.11 745420.3

> var((TxBorderCrossing$`2017`))

[1] 2.000345e+13

* + - A histogram and either box-plot or stem-leaf plot
* > stem(TxBorderCrossing$`2017`, scale = 1, width = 100, atom = 2e+08)
* The decimal point is 7 digit(s) to the right of the |
* 0 | 00000000000000000011111122333367900
* 2 | 2

**What to turn in**

* Publish your homework to GitHub under your user account, the appropriate format (check syllabus for instructions). This means turn in a separate document from your code that answers the questions above.
* Upload the code you used to GitHub.
* Provide me the links of these two files via e-mail no later than Thursday, September 20th at 6:00pm.