Find a dataset that you find interesting. You can look at a variety of sources for data including:

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?

* + <https://lib.utsa.edu> (log in with UTSA credentials)
  + Main page of UTSA library> Databases> Demography> RAND State Statistics>Population and Demographics>Deaths (Mortality)> Deaths and Death Rates by Cause (50 states)>Choose state or county(Texas)>Choose cause of death (Intentional self-harm) >submit

**In R: Import the data. Provide a list of variables in the dataset**

**[1] "Area" "State" "Chapter Name"**

**[4] "Sub-chapter Name" "Category" "2012"**

**[7] "2013" "2014" "2015"**

**[10] "2016"**

**What is the structure of the data? Which variables are character and which are numeric?**

Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 83 obs. of 10 variables:

$ Area : chr "Texas (state totals)" "Anderson County" "Angelina County" "Atascosa County" ...

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

$ Chapter Name : chr "External causes of morbidity and mortality" "External causes of morbidity and mortality" "External causes of morbidity and mortality" "External causes of morbidity and mortality" ...

$ Sub-chapter Name: chr "Intentional self-harm" "Intentional self-harm" "Intentional self-harm" "Intentional self-harm" ...

$ Category : chr "Population" "Population" "Population" "Population" ...

$ 2012 : num 26060796 58080 87488 NA NA ...

$ 2013 : num 26448193 NA NA NA 75825 ...

$ 2014 : num 26956958 NA 87750 47774 78069 ...

$ 2015 : num 27469114 57580 88255 NA 80527 ...

$ 2016 : num 27862596 57734 NA NA 82733 ...

**Describe the data through Central Tendencies:**

Give me the mean, median, and mode of **five variables**.

2012: Mean(836996), Median (136928), Mode(Mode does not exist)

2013: Mean(866834 ), Median(149378), Mode(Mode does not exist)

2014: Mean(866480), Median(141196), Mode (Mode does not exist)

2015: Mean(822645), Median(131705), Mode(Mode does not exist)

2016: Mean(872946), Median(145900), Mode(Mode does not exist)

**Describe the data through variation:**

2016: Variance(1.310377e+13), Range (27828565), Standard Deviation(3619857)

**Histogram**



**Boxplot**

****