

There is a lot of data that has been given to me, and it's definitely been more complex than the previous data set I was using. I sadly had to change the topic of my data analysis as the files were not compatible with my laptop which I do most of my work on. It is organized alphabetically by state per year from 1997 to 2016. There was other data such as national averages of state revenue and math or reading scores, but due to the overwhelming amount of data I decided to only focus my efforts on one part of the dataset as a whole. I'm not quite sure if I'm going to utilize as many variables as I currently have, due to the nature of having to change my entire data set, as I'm working on a new hypothesis.

All of the values in this data set, aside from the state name, are numerical. This will make it easy to determine trends among the varying states and use line or scatter plots to show the data.

The only coding schemes I've found in my data set are the difference in Primary Key as it has the state name, and then the year, and the difference between how the different grades scored.