**What I accomplished past week:**

For the past two weeks I was struggling to understand the columns in my dataset. Particularly ‘om’ which was Tornado number. Even after reading the data description, I couldn’t grasp what it meant. Last week I discussed with my group and Thomas gave some clarity on how that data is captured. It helped me to understand it.

Apparently, the column ‘om’ contains an ID for a tornado for each year. The first tornado in a year is given a tornado number 1 and second tornado is given 2 and so on. Each tornado might have affected multiple states, so I need to be mindful about the rows created for each state affected by that tornado.

If we need to find the total tornados in each year, we need to find the unique tornado numbers for that year. Below code does that.

A graph with blue lines

Description automatically generated

And below I have shared how I calculated the number of tornadoes each year and its plot.

A graph with lines and numbers

Description automatically generated

Looking at the graph I can’t find any difference, so trying to find the difference between them.



If you see the last column, there is some minor difference. This is not that drastic; hence the change isn’t visible in the plot.

**What I want to accomplish next week.**

1. I want to plot a heat map of the loss vs state. Something like below but with my dataset.

A map of the united states

Description automatically generated

1. Find Which state is affected by big tornadoes and visualize it.
2. I would like to find which months are more likely to get tornadoes.

**Roadblocks:**

I didn’t face any while working past week. Once I start working on answering the above questions, I might face some difficulties.