**Final Project: Milestone 5**

DSC 540

Trish Girmus

3/6/2021

This semester, we enhanced our skill sets by learning data wrangling with Python and SQL. I must admit, I felt I have learned more in this class than previous ones. What I enjoy about these courses are these final projects, where we can get hands-on practice and really learn the applications and processes when working with data. This is how I learn the best as I am a kinesthetic learner.

For our project, I learned a lot! I appreciate that we had this broken into milestones every two weeks, and for each milestone, we needed to extract that data differently. My biggest takeaways from this project – time and repetition. I really enjoyed this week analyzing the data and wish I would have had more time to do an extensive analysis and show the results. I realize that was the scope of this project, but maybe that is why I enjoy working with data!

What I learned from each milestone:

**Milestone 1:** it took days and hours to find the datasets required for this project. Searching for these datasets had me looking everywhere for information. Allowing/budgeting the time to find this information is a must.

**Milestone 2:** this was a great first exercise in learning how to clean data from a .csv or flat file. It was good practice and refresher from our assignments.

**Milestone 3:** this assignment was more challenging as I learned the table I found was not going to work for the project and I had to change course with how I needed to complete the assignment. This is part of the learning process.

**Milestone 4:** this was the most difficult milestone for me as learning about API’s was not an easy subject for me. I know I need to practice more with this in the future so I can improve my skills.

**Milestone 5:** this was a great exercise of joining the three datasets together. I also realize I need to practice more with SQL as this is also important when working with relational databases in the future.

My biggest learning lessons from this assignment were:

1. Time – allow plenty of it as I have learned nothing ever goes as planned.
2. The data itself – understanding your dataset is so crucial. Also, having cleaned and structured data will ensure accurate results. I also learned when picking a dataset, less is more. My dataset was almost 5,000 rows and 41 columns which was almost overwhelming for time I had to work with for this assignment.
3. There is still so much to learn – but this was great practice. I know repetition with coding will only make us better as students. I also learned that ideas I had originally, I had to adjust. For example, I wanted to look at data from the fiscal year 2015 through 2020. However, when I joined my datasets together in Milestone 5, I lost the year 2015 and could only use 2016-2020.

Regarding the findings of the actual project itself, I learned that our country is in serious debt. On the five data visualizations I provided, they all show an increase in expenses and debt from 2016-2020. The most alarming was in the year 2020. I presume our current pandemic was part of the contribution. I looked at the mean of Intragovernmental Holdings by Year (2016-2020) and that stayed consistent through but did slightly increase each year. On the other hand, total public debt per fiscal year dramatically increased from 2019 to 2020. One other visualization I wanted to mention was looking at the fiscal YTD expense per fiscal year. That total expense increased the most in 2020 at 25.1%.

This was a great project, and I really enjoyed working on it. I do regret not having more time to spend on it as I would have enjoyed adding more visualizations. I would also be curious to re-review this information now that we are in a new year, with a new president, and a pandemic that will hopefully be ending soon.