DMP (Data Management Plan) is a document that describes a dataset you acquired and how you will store, manage and analyze it. It also contains details regarding what mechanisms you will use to share and preserve your data.

In the beginning of my data plan, I started off with a Microsoft Word document and wrote down how I found the dataset and which organization collected this original data. Then, I took notes of available metadata, which I thought might provide to be useful for interesting narratives later on.

Afterward, I considered what would be good repository options to not only protect the data to any potential security hazard, but also preserve the data after the project ends. Then, for each option, I considered how I could respond to unforeseen risks with the data, such as losing the raw file or file corruption. Lastly, I filtered the options by examining their strengths and limitations and whether they can match my needs.

For example, GitHub is a well-known repository service, which can fulfill my needs. However, it still has its limitation. You can only upload a strict file of or less than 100MB. Although this does not impose much of a serious problem when you are uploading codes, it could be an issue when you are uploading massive raw data. In my case, my dataset—crime dataset—is too big to be simply dragged to my GitHub repository. Therefore, I will need to break down the raw data into different smaller files and then upload them.

Lastly, for my data plan, I then added which software and analysis tools I plan to use to analyze my dataset. Then, I explained whether these tools are licensed by my institution and if not, whether they are free. Moreover, I mentioned how my analysis techniques would format my output data.

As I included in the previously mentioned elements onto my word document, I would also occasionally transfer my document’s content into appropriate sections of a data plan I made on DMPTool. Before making a plan on DMPTool, it’s important to note that I first needed to go over my choices and decide what was the best fit for me. In my case, for my crime dataset, I decided a template provided by National Endowment for the Humanities (NEH) was the right fit.

After completing my data plan, I re-examined my data plan thoroughly a few times and shared my data plan on DMPTool with my collaborator, Ms.Hannah Gunderman.