**Assignment 01**

The chosen dataset is *Youtube Videos Dataset (~3400 videos)*. I decided to use this dataset for the assignment because it is one of the dataset I want to use and explore in my final project. I want to test the water with this dataset to see the potential of my project and if I’ll need to adjust and do some changes accordingly in this artistic process. I faced a couple of issues during the package installation and data importation process. It was a bit difficult to follow through all the steps of the installation. There were a lot of precise steps to follow, and I was a bit nervous while doing it for fear of messing things up on my computer. I had to re-install Chocolatey because I forgot a step, but in the end, I could install everything correctly. The second challenge was importing the data, it took me a while to understand the mongoimport function, the handshake was not happening so it took me awhile to figure it out. In the end I understood how to write the path for importing my csv file and that it required my connection string at the beginning of the code line to make the handshake happen. The third struggle was to connect my mongodb profile with my code. It took me a while to realize I forgot to add my URI to the env file.

I discovered that the used dataset for this assignment might not be a good fit for my final project, or at least very limited in its options. I tried various db functions and only a few of them were able to pull some data from this dataset. I think it is in the way things were categorized, there are only few categories in this dataset, which limits the possibilities of pulling out data. Plus, I look at the csv file and I noticed that some elements in the categories could have been categorized. For example, in the description category, many elements written in it could have been divided such as the content creators’ names and the total of subscribers.