

Samran Mirza

## Data Engineer Work Sample Readme

I completed this project using a Jupyter Notebook using python 3.8 as the programming language. I also used libraries such as pandas, NumPy, and ast. In order to run this program the machine must have these capabilities to run these libraries and applications. The output csv file was saved as Final\_Output.csv. The input csv file has to be stored in the same file folder as the program. The program is saved as GuildWorkSample.ipynb. I removed the word GATORADE from the csv file. So GATORADE has to not be in the csv file to run. I also created a public repo on my github. The URL is [https://github.com/smirza10/Movies\\_DataPipeline](https://github.com/smirza10/Movies_DataPipeline).