**Start Date: - 10-09-2021**

# Day-1

* Kick Off Meeting. Introduction to Trello Boards.
* Three tasks – System Requirements, Initial Data Analysis, Solution Candidates
* Spent a lot of time (4+ hrs) trying to learn how to create conda environments and integrate with Jupyter.
* Created a document for the same.
* Satisfied all the software requirements and created a report.
* Researched and learned the difference between IDA and EDA

# Day-2 (Saturday)

* Finding various ways to run the huge csv files. Explored options like Dask.
* Had a really hard time understanding Dask and yet the functional computations were as fast. 800 mb file took 16 mins to execute a function.
* System got hotter than Lana Rhodes.
* Tried to find more ways. Failed. Was told to use cloud services. But did not want to pay.
* Later found out that 13,300 rupees credit is provided for a new user. Tried signing up. Needed credit card details. [Create Your Azure Free Account Today | Microsoft Azure](https://azure.microsoft.com/en-in/free/)
* Explored options without a credit card. Found this: [Azure for Students – Free Account Credit | Microsoft Azure](https://azure.microsoft.com/en-us/free/students/) No need of a credit card. Free credit worth 6500. Enough for the project.
* Need to choose the best Virtual Machine to further work which would cost around 26 rupees an hour. Done that. Enabled SSH.
* Exploring the Azure Dashboard.
* Have to connect with the machine and find ways to do the second task IDA.
* Idea would be to chunk the dataset and work with it on the local machine. Once confident with the solution, I can then use that to the entire dataset.
* Tried many different ways to establish a connection but was not familiar with the idea of setting up an SSH.
* Awake till 4 am, tried a bunch of options and created a draft for the Initial Data Analysis report.
* Realized Pycharm was consuming a lot of RAM and was slowing down the whole system so thought of exploring Spyder next morning.

# Day-3 (Sunday)

* Need to submit the report today.
* It’s 1:00 pm and yet I have not begun with the IDA. Currently still exploring Azure services. If only I had started doing this in the morning of the previous day.
* Realized now at 2:30 pm with the help from mentor that I don’t need to upload data and run Jupyter notebooks on the cloud. Just do the development on my local machine. Connect to the Azure services using SSH and run the Python Script containing correct code through the cloud machine.
* Started exploring Spyder after I confirmed with the mentor that Pycharm wasn’t mandatory for this project.
* Set up a new environment and installed a bunch of libraries.
* Took help from sir and he explained via AnyDesk how to setup an SSH connection and an SCP connection.
* Was very satisfied to see the @azure on my local terminal.
* Uploaded all the csv files.
* Reconnected in the evening after having documented the steps and created a Python file for test execution.
* 6:15 pm, the first file ran on the cloud machine and I could see the output on my local machine. Was extremely satisfied with the result.
* Went for a walk. Tomorrow interview. Not worried.
* Came home, corrected the System setup report and created the IDA report. Uploaded by 8:50.
* Had Dinner. 1 pm and I am now writing my Solution Candidates report.
* Will upload this before sleeping.

# Day-4 (Monday)

# Day-5

# Day-6

# Day-7