**DSCI 551 project questions**

1. What does it mean when you say datasets must be “relevant to each other”?
2. Can we use the MapPartitions() function of Spark to implement Map/Reduce or do we have to create our own map reduce functionality using the Partition function of MySQL?
3. Being a 2-member team, we have to do implementation on Jupyter Notebook. I am trying to understand Part 3 of the project which states creating an app. How would we implement Part 3 of the project on Jupyter Notebook.
4. Do the search and analytical queries be dynamic that is take an input from the user or we can predefine a few queries just to show how our EDFS works and the results of the queries?