**Milestone #4, Final Report, (Due date - 8th week) 200 points**

I want you to design ten more interesting questions for this milestone and write their corresponding queries. These questions and queries should be diverse and able to cover the contents we have covered in the class. You should run this new set of queries in Python. Here is an example code that you use if you need some help with that.

Query performance: choose three to five questions that you designed for this milestone and write two different queries for each of them to observe and compare their performances. After making your comparison, explain why you would choose one over the other. And ultimately, come up with an optimized query plan for them.

Please record a 5-10 minutes presentation of your work. You may include the following in your presentation:

* Intro: What is your topic? Why did you choose the topic? Why is your project interesting or important? What datasets have you gathered and used?
* Present some questions and queries and interesting outputs you have seen in your results
* What did you learn when comparing different queries?
* Any interesting or valuable things that you have learned in this project
* For grad students: include some of your visualization results

Only for CS 586 students (grad students)

Create at least three visualizations for your query results.

You can use Python libraries, any other software, or your preferred way of visualization. (Note: You may use excel, but it is not preferred.)

**What to turn in?**

Please add all the following into a zip file and submit it on Canvas.

* Attach your Python code file written for database connection and implementation and used to run various queries. Along with how to run your code and a sample output of your code.
* A well structured PDF report including:
  + Your new ten questions and corresponding queries, with screenshots of the first 5 rows of output and the total number of rows in the result
  + 3-5 query performance comparison results, and optimized query plan for those selected five queries. Please include screenshots of your query performances.
  + A cloud link to your presentation. (A link to the recording of your presentation showing your database from Postgres, code logic, and results. You may use zoom to record a meeting and share the cloud link in your PDF file.)
  + Include a final ER diagram and any changes you made and the reason for those changes. If you haven’t made any modifications, it is excellent.
  + For CS586 (graduate) students, include visualizations in the report and briefly explain why you chose this data for visualization.

**Grading** is slightly different for grad and undergrad students for milestone #4, and it is as following:

For **undergraduate (CS 486)** students:

* [50 points] Query results through python code
* [50 points] Query optimization analysis
* [50 points] Your written code and its effectiveness
* [50 points] Your presentation

For **graduate (CS 586)** students:

* [40 points] Query results through python code
* [40 points] Query optimization analysis
* [40 points] Your written code and its effectiveness
* [40 points] Your presentation
* [40 points] Your visualizations and your analytical approach to them

**Note for milestone #3 & #4:**

* This project will require you only to know or learn some basic Python. You may use the examples from the sample code provided in this project.
* Although the documents here are in Python, and I can assist you the best in Python, you are welcome to do so if you are interested in using other programming languages.