**Final Project Instruction**

This project should illustrate your understanding and competency in the following area.

* Relevant topic selection
* Data gathering (cleansing if necessary)
* ERD implementation (Table, Columns, Data Types)
* SQL queries
* Performance tuning
* Other advanced skills in DBMS

You should upload your PPT slide deck the day before your presentation starts. (by 11:59pm). Both of your team members should take turns in the presentation.

Complete the following tasks.

1. Implement an ERD to satisfy the requirement of 3 NF. Provide it with a data dictionary.
2. Provide tables with data and SQL queries to obtain meaningful results from the dataset. The number of queries should be more than four. The first query should include a **SELECT** statement involving **more than two** tables, if possible. Also, make sure to use **table aliases**. In the second query, use **VIEW**. In the third query, use **CTE**. You can add more queries to illustrate more skills up to six in total.
3. Use **Materialized View** or **Recursive CTE** and mention which one you have used.
4. Graphical methods with analysis should be employed either by using Python, pgAdmin or other methods. (If you want to use other methods, check with me beforehand.)
5. Choose one query and perform the analysis using “index” and find the optimal number of the index(indexes) in the query. Justify your conclusion in your own words based on your analysis.

Your final report should include the following and have APA format.

* Cover page
* Introduction – should include why you select this topic and the data.
* ERD
* SQL queries & Results of the queries
* Conclusion – should include how you applied advanced skills and improved your design.
* References – should include the source of your raw data
* Roles and responsibilities of each team member.

In your main body, you can compare two or more scenarios and compare them, discussing advantages and disadvantages, and depending on your situation, you can suggest one over the other. (e.g index comes with a cost)

Choose the topic of your interest and get data. (Recommendation: Finance, Health, Ocean, Nova Scotia)

Data: Obtaining real-world data would be ideal. If not, try to get significantly sizable data.(Reference site: kaggle.com)

Other considerations:

\* You should also privately email me about what and how other member has done within 24 hours of your final presentation.

\*The scope of the DBMS and the size of data should be significantly more comprehensive than individual projects in the previous courses. The result should be something you can proudly present on your portfolio website. You can first contact people you know, who are working in the field and ask them if there’s any data publicly available.

\* Your naming in your projects should be consistent and well-organized, including but not limited to consistent in your table names, column names, etc.

**Important deadlines**

One of your team members needs to upload PPT slide decks (or whatever means you choose) on Brightspace the day before your presentation.

One of your team members should upload the Final Project Report on BrightSpace: Deadline: April 16th (Tuesday), 11:59 pm.

Rubric for presentation & final report

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 5 | 4 | 3 | 2 | 1 | 0 |
| ERD | - Includes all the necessary entities  - Includes all the relevant relationships  - Consistent and meaningful naming |  | - Includes part of necessary entities  - Includes part of the relevant relationships  - Some consistent and meaningful naming |  | Minimal effort has been made in some areas. | No effort has been made. |
| SQL | - More than four meaningful queries  - Complete Use of required Keywords | Four… | - Three meaningful queries  - Some use of required Keywords | Two… | - One meaningful query  - Minimal Analysis | No meaningful queries |
| Graphical Tool | Comprehensive analysis |  | Partial… |  |  | No effort has been made. |
| Recursive CTE/ Materialized View | Comprehensive analysis |  | Partial… |  |  | No effort has been made. |
| Performance tuning skills | Comprehensive analysis using performance tuning skills including indexes, cost analysis |  | Partial analysis using performance tuning skills including indexes, cost analysis |  | A minimal analysis is present. | No effort has been made. |
| Introduction & Conclusion & Reference | - Comprehensive contents are present & followed APA format. |  | Some relevant contents are present. |  | Minimal relevant contents are present | No relevant contents are present. |