CA - 1 – Data Storage Solutions

Individual Contribution report submitted in partial fulfilment of the degree of

MSc in Data Analytics

Module: Data Storage Solutions

Author: Bharath Shakthivel Balaih Raveendran
Student ID: 20057027

1. Dataset Selection and Initial Implementation

Our team met at the beginning of the project to choose a suitable dataset, and we ultimately decided on the one that my teammate Daniel had proposed. I was able to put the ideas I learned in lectures into practice by importing and getting this dataset ready for use. I gained a solid understanding of SQL Server and completed the first setup by looking up information online and paying attention to in-class demonstrations. In addition to strengthening the module's theoretical components, this practical experience gave me SQL skills that will be useful in future data-related positions.

2. Schema Design and Data Warehouse Development

After discussing roles and responsibilities, our team started working on the data warehouse design. At first, our schema looked like a snowflake. I refined our model into a proper star schema by working with our lecturer to analyze the database and find dimensions that weren't needed for the fact table. My knowledge of fact tables, dimension tables, and their importance in creating a successful data warehouse for practical analytics applications has increased as a result of this exercise.

3. ETL Process and Issue Resolution

I helped my teammate Mrunal resolve a significant SSRS integration issue during the development phase. I learned a lot about the ETL process by investigating possible fixes and resolving the issue that prevented data from moving from the operational database to the warehouse. This experience exposed me to real-world problem-solving techniques in database management workflows and improved my ability to work with SSIS for data population.

4. Data Visualization and Tableau Integration

Making Tableau visualizations based on the sales fact table was another way I helped. To do this, Tableau and SQL Server Management Studio had to be integrated in order to produce insightful data. Working with both platforms enhanced my ability to effectively retrieve data using SQL and convert it into interactive, business-focused visual dashboards in Tableau are directly applicable to tasks involving professional data analytics.

5. Neo4j Graph Database Implementation

My implementation of Neo4j as a graph database was another important aspect of my contribution. I constructed nodes and relationships, imported data from the warehouse, and composed seven Cypher queries to compare with corresponding SQL queries. My technical

skill set was further expanded by this comparison, which gave me useful insight into graph database concepts and their benefits when working with connected data.

6. Conclusion

Overall, this project improved my teamwork, time management, and report-writing abilities in addition to my technical proficiency in data warehousing, ETL, visualization, and graph databases. Effective peer collaboration reaffirmed the value of teamwork in accomplishing a shared objective, and hands-on application helped close the knowledge gap between classroom instruction and real-world application.