**Data Storage And Soloutions**

**Individual Contribution Report**

**Sales Analysis For Automobile Company**

**Name: Adit Deshpande**

**Roll Number:10505324**

This report is regarding my contribution in the Data Storage and Solution group assessment. The assessment included various activities such as finding the dataset to perform analysis, then using that data set to create a data warehouse. Further, once the data warehouse was ready, we had to perform Extraction, Transformation and Loading (ETL) of data from our data set to the data warehouse. I helped my group in finding the data set. Once, the data set was final I started preparing the schema for our data warehouse. I discussed with my group, and then created the dimension tables and fact tables, used Microsoft SQL Server Management Studio (SSMS) to load the data in those dimension and fact tables. Further, for the analysis purpose we needed to generate SSRS reports and tableau visualization which was done by group members. In this process, I was also involved in the discussion part where we suggested each other how can we show our insights and can get better visualizations.

Another part of our assessment was to explain how Relational Database Management System (RDBMS) differs from the Graph Databases. For that we were asked to download ‘AdventureWorks’ database and to execute seven queries in both CQL and SQL. In this task we distributed these seven queries among ourselves. Two out of seven queries were done by me using Neo4j for CQL and SSMS for SQL. I was quiet familiar with the syntax of SQL, so that part was easy for me to do. But I learned for the first time CQL using videos uploaded on moodle. Adapting to new syntax and then using it for the assessment was bit tricky.

There were many difficulties faced during the entire process of the assignment. I used to get stuck on some parts or processes. One such difficulty faced was during loading the data in one of the fact tables where duplicate values were getting extracted from dimension table. So, it gave me error while debugging. For this, I again had a look at the queries generated in query builder of source, then I realised that there was problem in that query and gave me duplicated values. To overcome this, I wrote the entire query in SSMS and copied that at the place of ‘SQL command text’ of ADO.NET Source and got the error solved. Also, during the executing the CQL queries we were getting error, so we watched moodle videos again and again until we understand what the error was. There was also one measure issue we all faced due to the lockdown because of COVID-19. There was not personal interaction between me and my group members. It made communication bit difficult, however we used Microsoft Teams and Zoom in order to discuss all the difficulties faced and got them solved. Not only among the group but we attended doubt solving sessions arranged by our faculties which made us easy to work on assessments.

From this assessment I gained the practical knowledge of ETL, SSRS Reports, Tableau visualizations and Neo4j. I think that having a practical knowledge is as important as knowing the theory behind it. I felt like preparing a presentation and reports for a real company, which gave me experience of real-world process. Also, involving in a group discussion, putting my point forward, convincing group members how we can improve our project and getting suggestions and learning their point view gave me a different experience. I feel like this is the better way of learning than the theory exams and the knowledge gained by me in this would surely help me in my theory exams.

