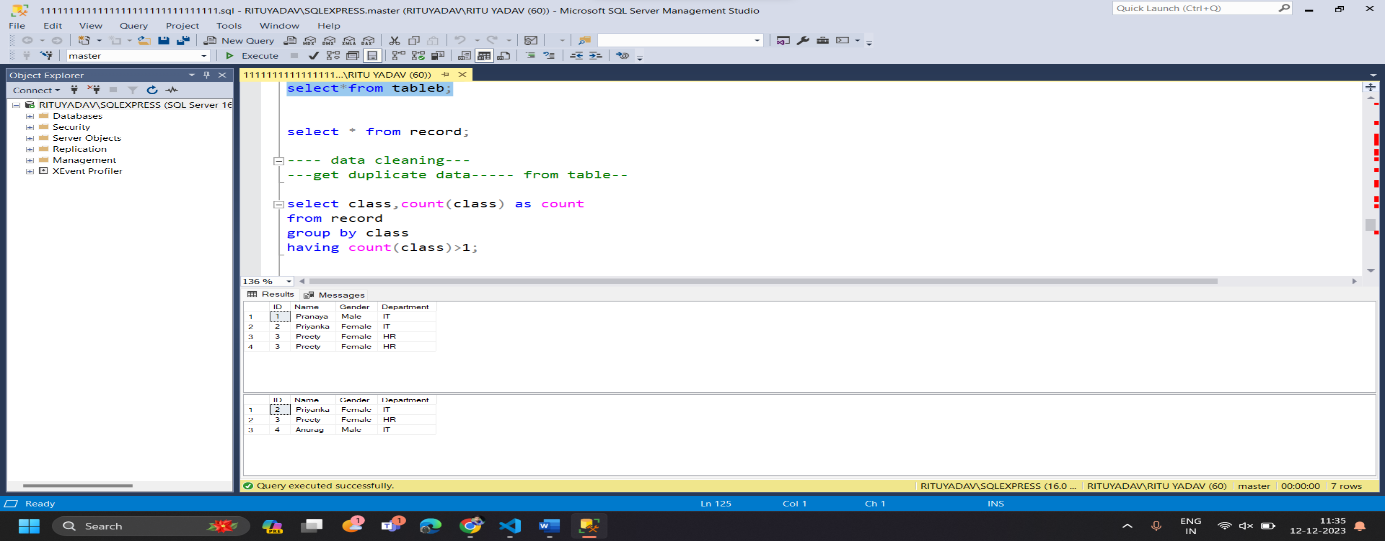
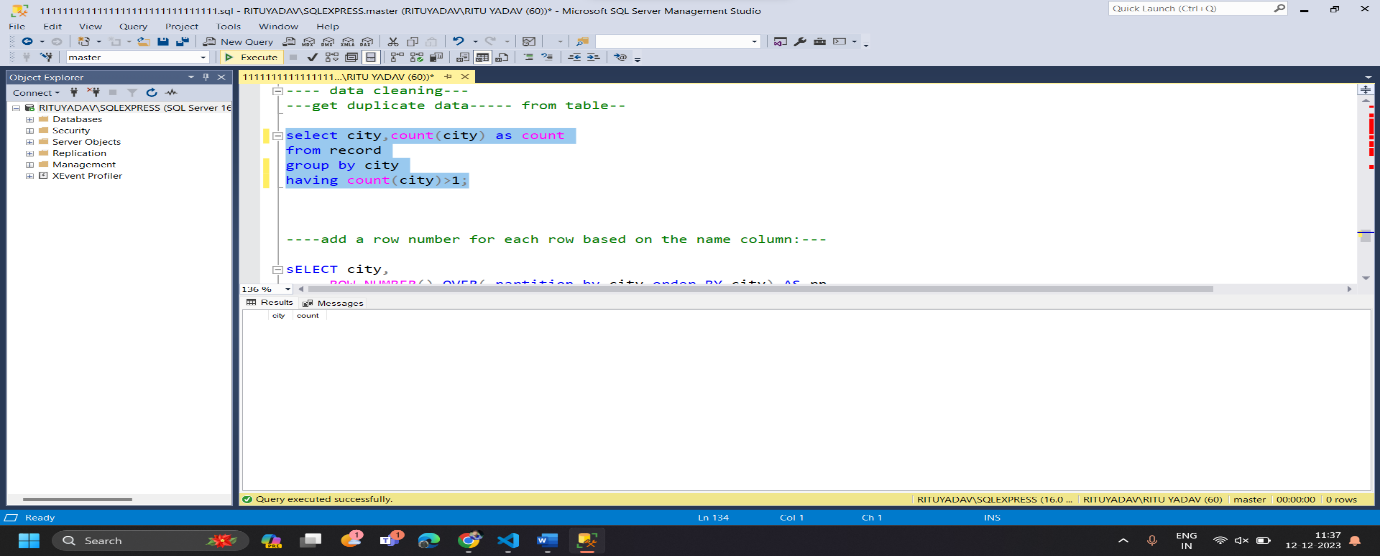
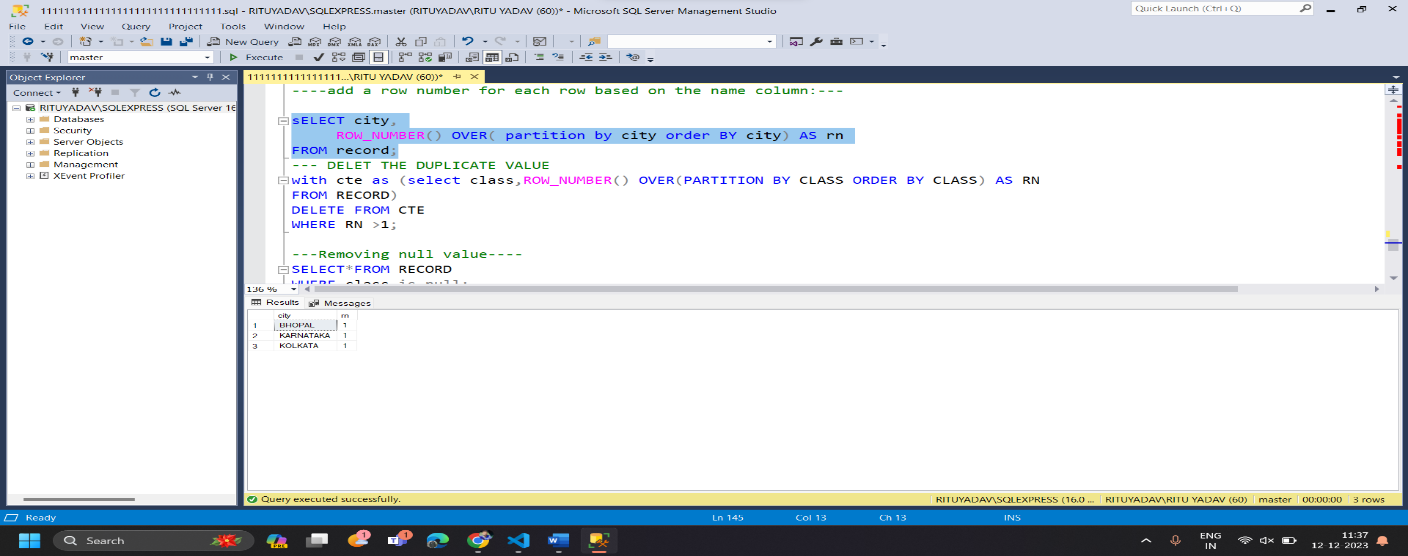
11/12/2023

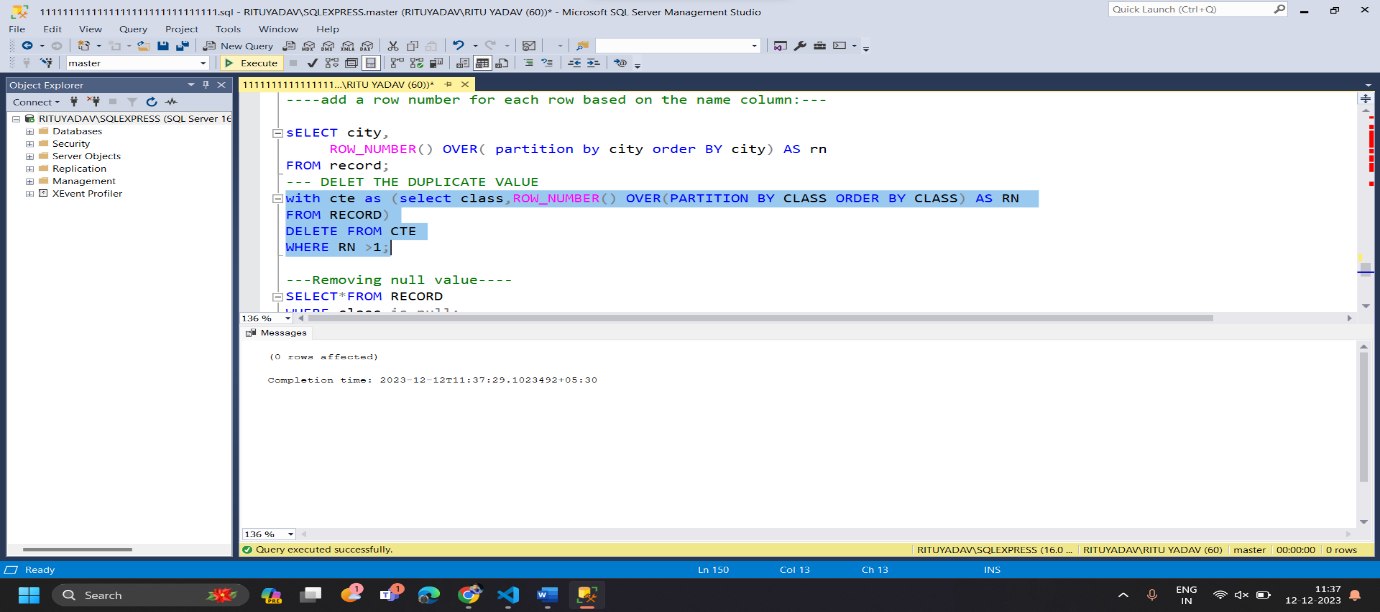
DATA CLEANING, SCHEMA,SNOWFLAKE

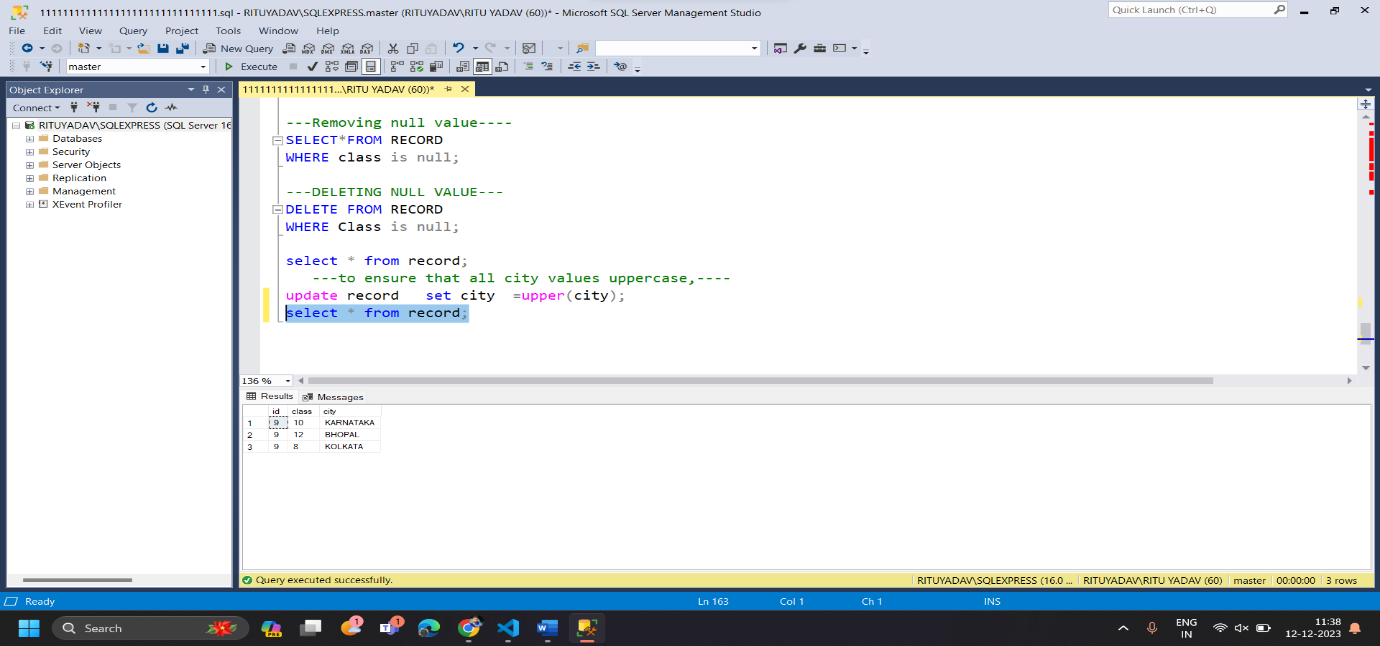
DATA CLEANING:

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabelled.









SCHEMA: **It represents the organization of data and provides information about the relationships between the tables in a given database. In this topic, we will understand more about database schema and its types. Before understanding database schema, lets first understand what a Database is**



TYPE OF SCHEMA

* Logical Schema
* Physical schema
* View Schema

Snowflake Schema: -

**The snowflake schema is an adaption of a star schema. There is a main "Fact" table in the star schema that contains the main data points and reference to its dimension tables. But in snowflake, dimension tables can have their own dimension tables**.