Name: Vinayak Soni
Roll No: 67

Practical No. 6

Aim: To perform ETL process for building a data warehouse using power query.

Theory: The ETL process (Extract, Transform, Load) is critical in building a data warehouse. Its main aim is to extract data from various sources, transform it into a suitable format for analysis, and load it into the data warehouse. Power Query is a data transformation and cleansing tool that can perform the ETL process in Microsoft Excel and Power BI. It allows users to connect to various data sources, filter, and clean data, and transform it into a format that can be easily loaded into a data warehouse. The ETL process using Power Query typically involves the following steps:

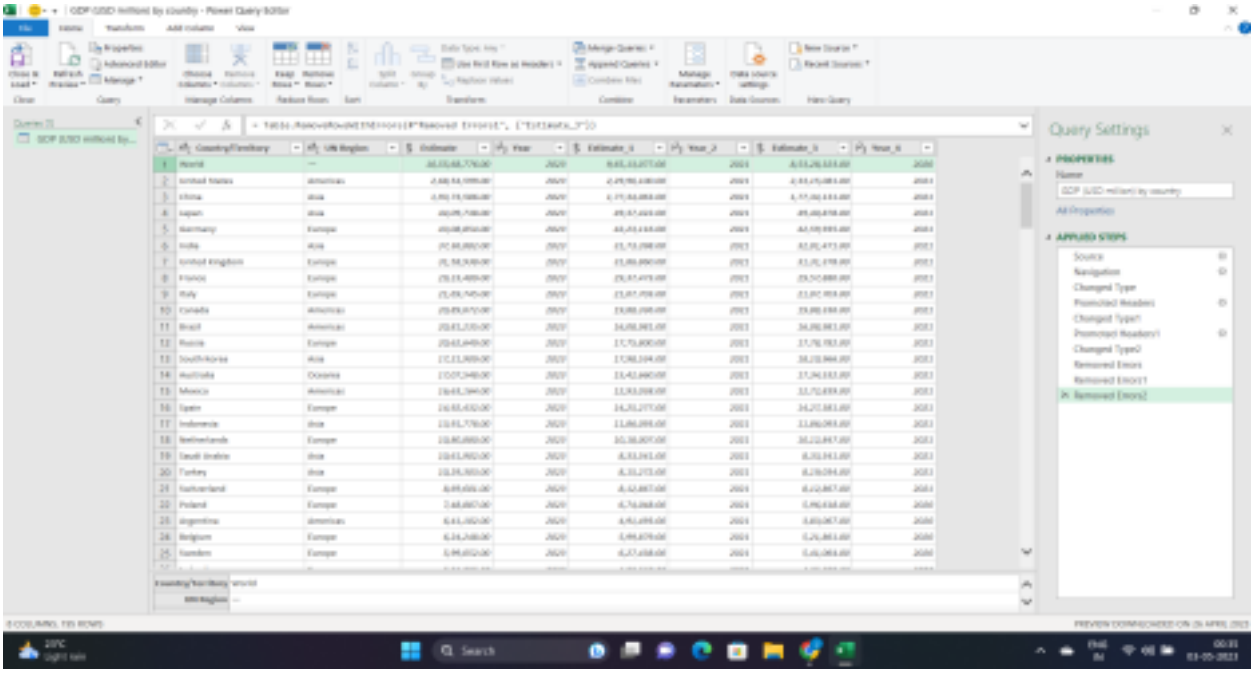
1. Extraction: Data is extracted from various sources such as databases, spreadsheets, or flat files.

2. Transformation: The extracted data is cleaned, filtered, and transformed into a format that is suitable for analysis. This includes tasks such as removing duplicates, renaming columns, converting data types, and creating calculated fields.

3. Loading: The transformed data is loaded into the data warehouse and stored for future analysis and reporting.

Screenshots:

Here we use a URL: https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal) and load it to power query
We perform transform and do preprocessing- Putting correct headers, changing data types, removing errors, etc.



Finally loading the data in excel sheet: