

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_\(nominal\)](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.

Screenshot of Power Query Editor showing the 'GDP (USD) million by country' query. The table contains 29 rows of data with columns: Country/Region, Life_Region, Estimate_1, Year_1, Estimate_2, Year_2, Estimate_3, Year_3, and Estimate_4, Year_4. The 'APPLIED STEPS' pane shows the 'Removed Errors' step.

Country/Region	Life_Region	Estimate_1	Year_1	Estimate_2	Year_2	Estimate_3	Year_3	Estimate_4	Year_4
1. Namibia	—	101,048,776.00	2021	94,11,207.00	2021	4,51,323.00	2021	4,51,323.00	2021
2. United States	Americas	2,611,54,000.00	2020	2,47,95,400.00	2021	4,71,42,000.00	2021	4,71,42,000.00	2021
3. China	Asia	1,811,15,000.00	2020	1,71,15,000.00	2021	3,71,15,000.00	2021	3,71,15,000.00	2021
4. Japan	Asia	462,09,700.00	2020	45,37,000.00	2021	45,37,000.00	2021	45,37,000.00	2021
5. Germany	Europe	401,08,000.00	2020	44,13,000.00	2021	44,13,000.00	2021	44,13,000.00	2021
6. India	Asia	17,44,000.00	2020	17,73,000.00	2021	31,05,473.00	2021	31,05,473.00	2021
7. United Kingdom	Europe	17,18,000.00	2020	17,08,000.00	2021	31,12,198.00	2021	31,12,198.00	2021
8. France	Europe	17,13,000.00	2020	17,08,000.00	2021	31,12,198.00	2021	31,12,198.00	2021
9. Italy	Europe	17,07,700.00	2020	17,07,700.00	2021	31,07,700.00	2021	31,07,700.00	2021
10. Canada	Americas	17,05,000.00	2020	17,05,000.00	2021	31,05,000.00	2021	31,05,000.00	2021
11. Brazil	Americas	17,01,300.00	2020	14,00,000.00	2021	34,00,000.00	2021	34,00,000.00	2021
12. Russia	Europe	17,01,000.00	2020	17,01,000.00	2021	31,75,000.00	2021	31,75,000.00	2021
13. South Korea	Asia	17,11,000.00	2020	17,08,100.00	2021	38,10,344.00	2021	38,10,344.00	2021
14. Australia	Oceania	17,07,000.00	2020	17,42,000.00	2021	37,04,000.00	2021	37,04,000.00	2021
15. Mexico	Americas	16,63,394.00	2020	14,03,000.00	2021	33,70,000.00	2021	33,70,000.00	2021
16. Spain	Europe	16,61,000.00	2020	14,73,217.00	2021	34,27,343.00	2021	34,27,343.00	2021
17. Indonesia	Asia	17,01,700.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
18. Netherlands	Europe	16,95,000.00	2020	17,08,000.00	2021	31,12,047.00	2021	31,12,047.00	2021
19. Saudi Arabia	Asia	17,01,000.00	2020	17,01,000.00	2021	31,01,313.00	2021	31,01,313.00	2021
20. Turkey	Asia	17,01,000.00	2020	17,01,000.00	2021	31,01,000.00	2021	31,01,000.00	2021
21. Netherlands	Europe	16,95,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
22. Poland	Europe	16,88,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
23. Argentina	Americas	16,11,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
24. Belgium	Europe	16,11,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
25. Vietnam	Europe	16,04,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021

Finally loading the data in excel sheet:

Screenshot of Microsoft Excel showing the 'GDP (USD) million by country' table loaded into a new sheet named 'Sheet1'. The table has 29 rows and 10 columns, matching the data from the Power Query Editor.

Country/Region	Life_Region	Estimate_1	Year_1	Estimate_2	Year_2	Estimate_3	Year_3	Estimate_4	Year_4
1. Namibia	—	101,048,776.00	2021	94,11,207.00	2021	4,51,323.00	2021	4,51,323.00	2021
2. United States	Americas	2,611,54,000.00	2020	2,47,95,400.00	2021	4,71,42,000.00	2021	4,71,42,000.00	2021
3. China	Asia	1,811,15,000.00	2020	1,71,15,000.00	2021	3,71,15,000.00	2021	3,71,15,000.00	2021
4. Japan	Asia	462,09,700.00	2020	45,37,000.00	2021	45,37,000.00	2021	45,37,000.00	2021
5. Germany	Europe	401,08,000.00	2020	44,13,000.00	2021	44,13,000.00	2021	44,13,000.00	2021
6. India	Asia	17,44,000.00	2020	17,73,000.00	2021	31,05,473.00	2021	31,05,473.00	2021
7. United Kingdom	Europe	17,18,000.00	2020	17,08,000.00	2021	31,12,198.00	2021	31,12,198.00	2021
8. France	Europe	17,13,000.00	2020	17,08,000.00	2021	31,12,198.00	2021	31,12,198.00	2021
9. Italy	Europe	17,07,700.00	2020	17,07,700.00	2021	31,07,700.00	2021	31,07,700.00	2021
10. Canada	Americas	17,05,000.00	2020	17,05,000.00	2021	31,05,000.00	2021	31,05,000.00	2021
11. Brazil	Americas	17,01,300.00	2020	14,00,000.00	2021	34,00,000.00	2021	34,00,000.00	2021
12. Russia	Europe	17,01,000.00	2020	17,01,000.00	2021	31,75,000.00	2021	31,75,000.00	2021
13. South Korea	Asia	17,11,000.00	2020	17,08,100.00	2021	38,10,344.00	2021	38,10,344.00	2021
14. Australia	Oceania	17,07,000.00	2020	17,42,000.00	2021	37,04,000.00	2021	37,04,000.00	2021
15. Mexico	Americas	16,63,394.00	2020	14,03,000.00	2021	33,70,000.00	2021	33,70,000.00	2021
16. Spain	Europe	16,61,000.00	2020	14,73,217.00	2021	34,27,343.00	2021	34,27,343.00	2021
17. Indonesia	Asia	17,01,700.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
18. Netherlands	Europe	16,95,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
19. Turkey	Asia	17,01,000.00	2020	17,01,000.00	2021	31,01,313.00	2021	31,01,313.00	2021
20. Switzerland	Europe	16,88,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
21. Poland	Europe	16,88,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
22. Argentina	Americas	16,11,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
23. Belgium	Europe	16,11,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021
24. Vietnam	Europe	16,04,000.00	2020	17,08,000.00	2021	31,08,000.00	2021	31,08,000.00	2021