

Transforming and Analyzing Data with SQL

by
Abi Afolabi
12th July 2023



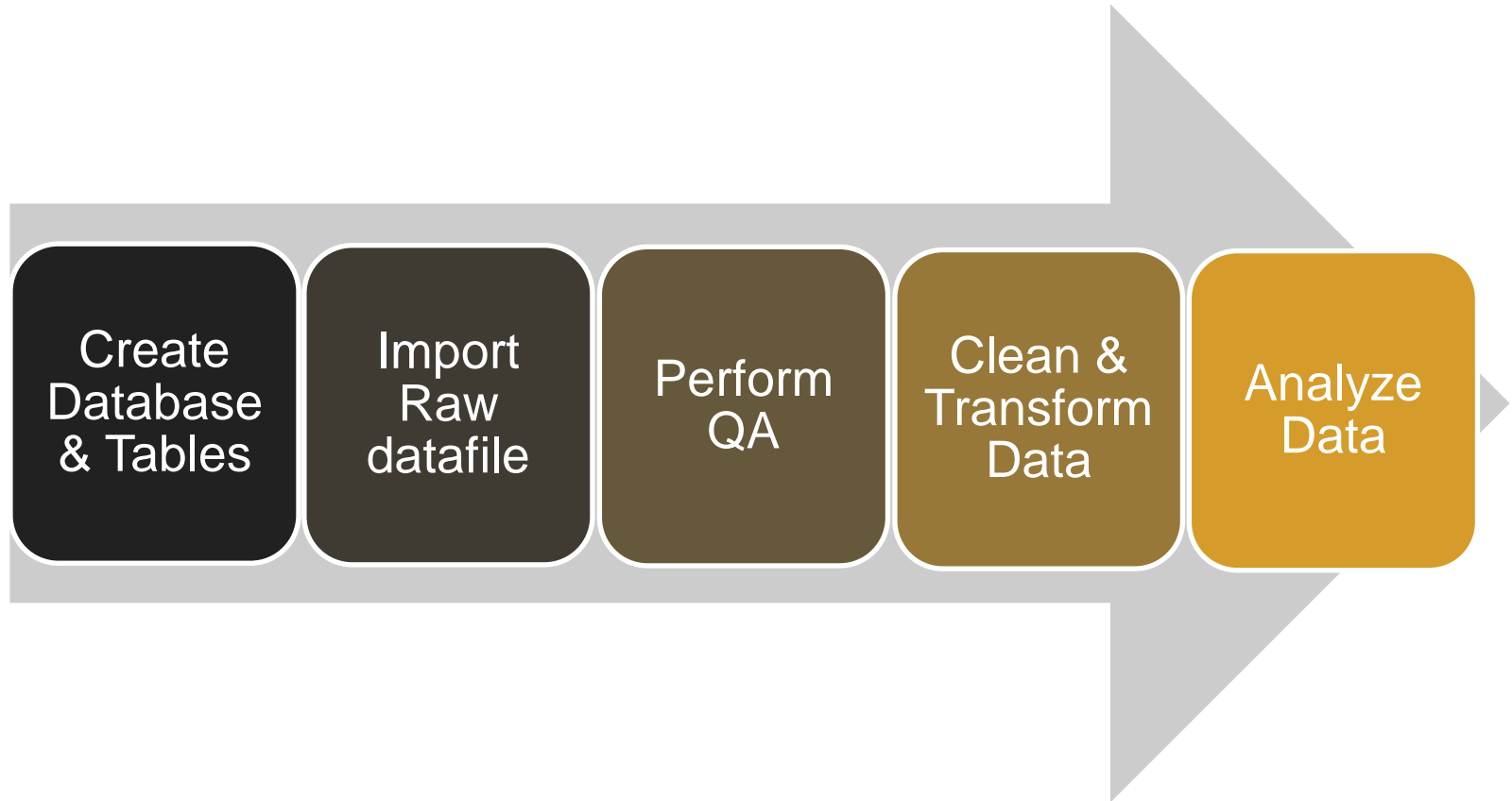
Project Overview & Goals

- Reinforce learning with hands-on experience
 - Database & Table creation
 - Data Import & QA
 - Data Cleaning & Transformation
 - Data Analysis
- Deductive reasoning
 - Extract hidden information
 - Gain insight



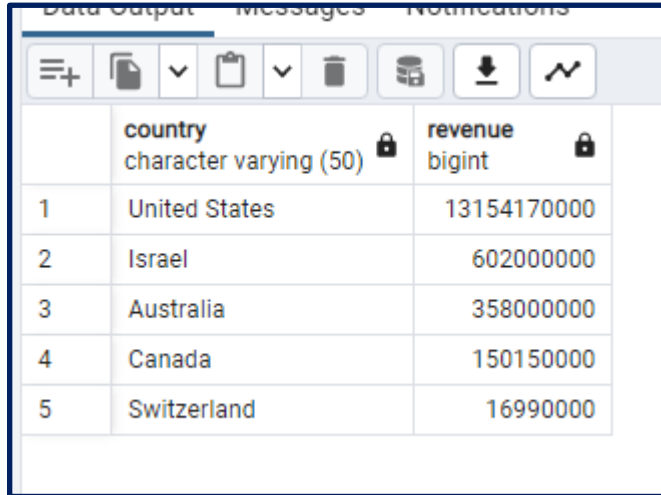
- Data from online store
- Customer interactions
- Customers across the globe
- Records of sales

Project Execution Steps



Discoveries

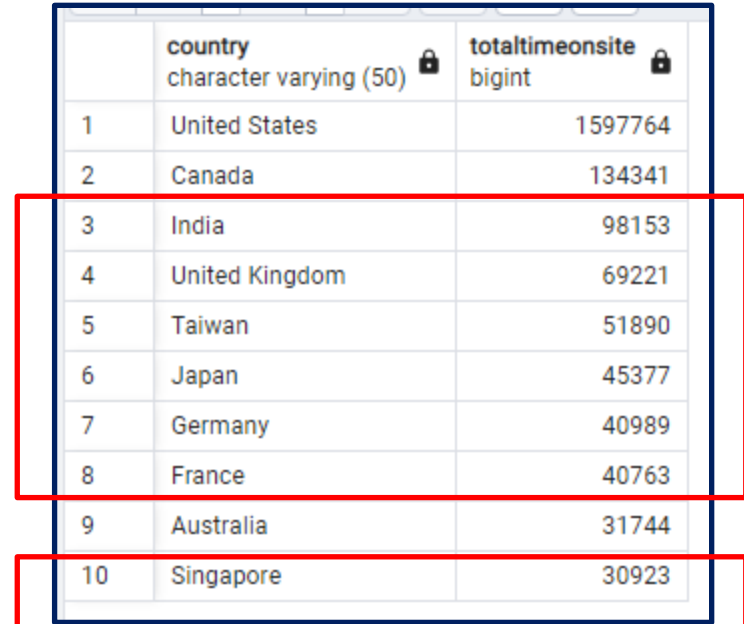
- Current top ranked customers
 - 5 out of 136 (~4%)



A screenshot of a data table interface. At the top, there are tabs for 'Data Output', 'Messages', and 'Notifications'. Below the tabs is a toolbar with icons for adding, deleting, and other actions. The table has two columns: 'country' (character varying (50)) and 'revenue' (bigint). The data is sorted by revenue in descending order, showing the top 5 customers.

	country character varying (50)	revenue bigint
1	United States	13154170000
2	Israel	602000000
3	Australia	358000000
4	Canada	150150000
5	Switzerland	16990000

- Potential customers (Limit 10)



A screenshot of a data table interface showing potential customers. The table has two columns: 'country' (character varying (50)) and 'totaltimeonsite' (bigint). The data is sorted by total time on site in descending order, showing the top 10 customers. A red box highlights the rows for India, United Kingdom, Taiwan, Japan, Germany, France, Australia, and Singapore.

	country character varying (50)	totaltimeonsite bigint
1	United States	1597764
2	Canada	134341
3	India	98153
4	United Kingdom	69221
5	Taiwan	51890
6	Japan	45377
7	Germany	40989
8	France	40763
9	Australia	31744
10	Singapore	30923

QA Process

- Incomplete data (visual & SQL)
- Inconsistent / unreasonable values
 - Quantityordered < 0
- Validate column names in different tables for JOINS
 - Fullvisitorid
- Referential Integrity Check failed

With more time I will....

- Investigate why Referential Integrity Check between tables failed.
- Explore best opportunity to update null values with valid data.
- Generate insight to guide on ways to get customers from more countries.

Conclusion

- Questionable dataset ??
- Investigate opportunities to increase revenue
 - Promotional offers
 - Payment currency
 - Discount for customers in select countries

A dark green, irregularly shaped graphic resembling a watercolor splash or a rock. The words "Summary" and "Notes" are written in white, bold, sans-serif font, stacked vertically.

**Summary
Notes**



https://github.com/AbiAfolabi/Abis_New_Project1