



## SQL Case Study

### Business Context:

A retail store would like to understand customer behaviour using their point of sale data (POS).

### Data:

Customer Demographics: customer.csv

Transactions: Transaction.csv

Product Category: prod\_cat.csv

### BigQuery Access:

- Go to <https://cloud.google.com/bigquery> using your gmail id.
- Click on "Console" at the top right corner.
- Agree to the terms of service and click Continue.
- Search for "bigquery" on the search bar at the top.
- Click on "Create Project"
- Give the project a name you like.
- Click on the three dots on the left dataset panel's in the Explorer and click "Create Dataset".



## SQL Case Study

- Name the dataset "statfinity\_sql\_case". Select data location as US and click "Create Dataset".
- Click on the 3 dots in front of **statfinity\_sql\_case** on the left panel. Click "Create Table".
- In "Source", click "Upload". Browse and upload the CSV files and name the table accordingly.
- Write the queries in the editor and send them as separate txt files.

## Questions:

### DATA ANALYSIS

1. Which channel is most frequently used for transactions?
2. What is the count of Male and Female customers in the database?
3. From which city do we have the maximum number of customers and how many?
4. How many sub-categories are there under the Books category?
5. What is the maximum quantity of products ever ordered?
6. What is the net total revenue generated in categories Electronics and Books?
7. How many customers have >10 transactions with us, excluding returns?
8. What is the combined revenue earned from the "Electronics" & "Clothing" categories, from "Flagship stores"?
9. What is the total revenue generated from "Male" customers in "Electronics" category? Output should display total revenue by prod sub-cat.
10. What is percentage of sales and returns by product sub category; display only top 5 sub categories in terms of sales?
11. For all customers aged between 25 to 35 years find what is the net total revenue generated by these consumers in last 30 days of transactions from max transaction date available in the data?



12. Which product category has seen the max value of returns in the last 3 months of transactions?
13. Which store-type sells the maximum products; by value of sales amount and by quantity sold?
14. What are the categories for which average revenue is above the overall average.
15. Find the average and total revenue by each subcategory for the categories which are among top 5 categories in terms of quantity sold.