

Ebury What borders?®

Case Study

Analytics Engineer

How to Access

To access the Study Case, open the GCP project called "analytics-case-study" using your personal email account.

You can access the project by clicking this link:

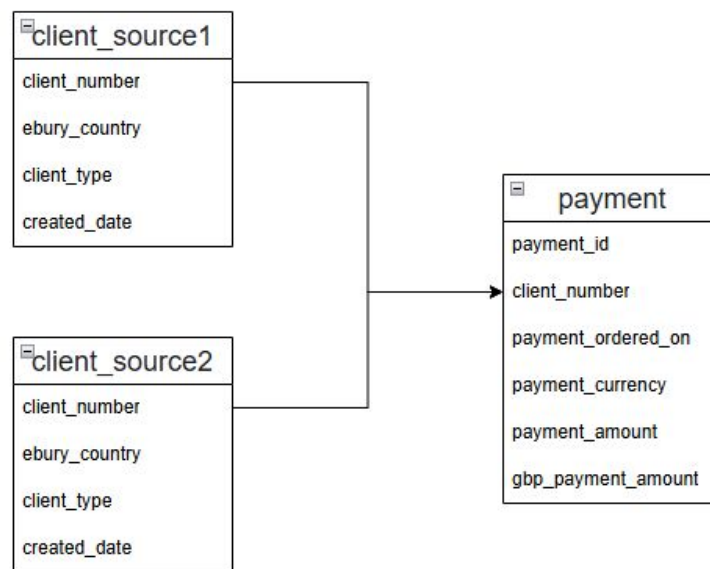
<https://console.cloud.google.com/welcome?inv=Ab6F2A&project=analytics-case-study&supportedpview=project>

If this link doesn't work, follow these steps:

1. Open the GCP console page: <https://console.cloud.google.com/>
2. On the top left corner, select the project drop-down box.
3. Select the "All" tab.
4. Click on the "analytics-case-study" project.
5. In the BigQuery find the dataset called: "ae_case_study_brazil_" + your name initials

If neither option works, please let us know so we can regularize your access.

Part 1: Modelling



You have been tasked with building a data model to provide business areas with payment-related information. However, the company has two distinct customer data sources, requiring you to consolidate this information in a consistent and non-redundant way.

The business users who will consume this model must have complete trust in the quality of the information. Therefore, perform all necessary exploratory analyses and, after development, run rigorous validation tests to ensure the data is accurate, complete, and ready for use.

Modeling Questions

Please use the workspace to create the tables and send the query to us.

1. We have a relational model with Payment and Client objects that needs to be modeled in a dimensional model, but we have 2 clients tables that the origin are different sources systems. **Draw your model proposal considering that we need to maintain the history.**
2. How would you do in your dimensional model if we needed to **filter only working days (excluding Brazilian, Spanish and English holidays)?**

Part 2: Analysis

Tables

| trade | | | | | | |
|----------|---------------------|------------|------------|------------|---------------------|--------|
| trade_id | order_date | account_id | product_id | trade_type | paid_date | amount |
| 111 | 2021-01-21T00:00:00 | 0000123 | 1 | internal | 2021-01-21T00:00:00 | 100 |
| 222 | 2020-11-11T00:00:00 | 0000123 | 2 | external | 2021-01-21T00:00:00 | 500 |
| 333 | NULL | 0000234 | 2 | internal | 2021-01-21T00:00:00 | 250 |
| 444 | 2021-01-21T00:00:00 | 0000345 | 2 | imported | 2021-01-21T00:00:00 | 250 |
| 555 | 2021-01-21T00:00:00 | 0000345 | 3 | NULL | 2021-01-21T00:00:00 | 150 |

| account | | | |
|------------|--------------|-----------------|---------------------|
| account_id | account_name | account_country | created_on |
| 0000123 | Name A | Spain | 2021-01-21T00:00:00 |
| 0000234 | Name B | Spain | 2021-01-03T00:00:00 |
| 0000345 | Name C | UK | 2020-01-18T00:00:00 |
| 0000456 | Name D | UK | 2021-01-03T00:00:00 |

| product | |
|------------|--------------|
| product_id | product_name |
| 1 | Payment |
| 2 | Invoice |
| 3 | Trade |

Analysis Questions

Please use the workspace to create the queries and send them to us.

1. Trades MTD. When order_date is NULL use paid_date. **Result trade_id, order_date and amount**
2. Return only spanish accounts that the total amount is more than 300. **Result account_id, account_country, total_amount**
3. Trades that are not external where a thread was bought. **Result trade_id, trade_type, order_date, amount**
4. All trades except the ones above. **Result trade_id, trade_type, order_date, amount**
5. Find if there are any duplicates in the trades table.
6. All accounts with the number of trades done, even if they don't have any trades. **Result account_id, number of trades**

Ebury

info@ebury.com | ebury.com

Ebury Partners Belgium NV/SA is authorised and regulated by the National Bank of Belgium as a Payment Institution under the Act of 11 March 2018, registered with the Crossroads Bank for Enterprises under number 0681.746.187. EBURY and EBURY What Borders? are trademarks.

© Copyright 2009-2022