Analista de dados Jr.

Introduction

You are part of a data team working at a *fintech* called **LCGBR Bank**.

The data department's mission is to provide **BI tools**, **information**, and **insights** that enable the company to excel in a fiercely competitive market.

Data Architecture Overview

Here at LCGBR Bank, we have three different environments: Production, Data Warehouse and Reporting. (Figure 1).

The Production Environment is where LCGBR Bank's services live. In this environment, services are responsible for getting customers' information and storing in their respective databases.

Then, the datasets generated by the services are maintained in the Data Warehouse Environment, where Business Analysts and Analytics Engineers create new tables, improve data models, and work to turn raw data into easily consumable tables for analysis.

Finally, these tables feed into the Reporting Environment, where it is possible to create all sorts of dashboards and visualisations for monitoring and decision-making.

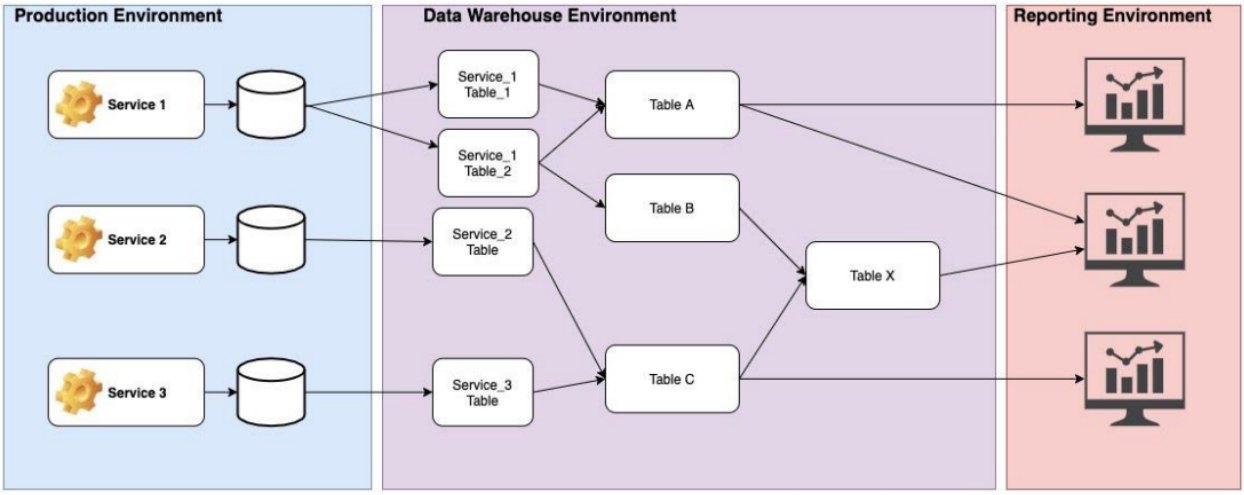

Figure 1. LCGBR Bank's environments: Production, Data Warehouse and Reporting.

A slice of the table structure from the Data Warehouse Environment is depicted in Figure 2.

Apart from time (d_time, d_year, d_month, d_week, d_weekday), location (city, state, country), accounts, and customers tables, three tables store the financial movements of the accounts:

- **transfer_ins**: non PIX transfers made to an account (money arriving)
- **transfer_outs**: non PIX transfers made from an account (money leaving)
- **pix_movements**: transfers that are either received by or sent from an account using PIX

Figure 2. A slice of the table structure from the DW Environment (diagram/table_diagram.png). You can check the code used to generate these diagrams using dbdiagram.io on file (diagram/table_diagram.txt).

Business Context 1

To solve this case you need to be familiar with the concept of "Account Monthly Balance". Account Monthly Balance is the amount of money a customer had in their account at the end of a given month. This information can be calculated by adding all the transfers in and subtracting all the transfers out from the previous

account balance. Remember you should also consider **transfers made by PIX**, the newest Brazilian transfer method. You can see an example below:

Month	Account	R$ Total Transfer In	R$ Total Transfer Out	R$ Account Monthly Balance
2020-12	1	R$ 1.000,00	R$ 200,00	R$ 800,00
2020-12	2	R$ 2.000,00	R$ -	R$ 2.000,00
2020-11	1	R$ -	R$ 200,00	R$ 600,00
2020-11	2	R$ 100,00	R$ 500,00	R$ 1.600,00
2020-11	3	R$ 500,00	R$ 100,00	R$ 400,00

Table 1. An example of account monthly balance data.

Business Context 2

To solve this case, you need to understand the concept of "Sliding Window Calculations." These calculations allow you to aggregate data row by row across your dataset, providing **running totals, sums, counts, and averages in any order you need**.

In this scenario, the company aims to assess the performance of **transfers out** made by Pix over time. This assessment includes analyzing the number of accounts, quantity and value of transfers, as well as comparing these metrics to the previous month (L1M) and the average of the last three months (L3M).

Month	# Number of Accounts	# Number of transfers	R$ Amount of transfers	% L1M Number of transfers	% L3M AVG Number of transfers	% L1M Number of accounts	% L3M AVG Number of accounts
2020-12	380	2.003	R$ 2.003.862,90	3,84%	2,74%	15,50%	21,15%
2020-11	329	1.929	R$ 1.815.864,60	-3,79%	-2,44%	9,67%	5,90%
2020-10	300	2.005	R$ 2.092.486,00	4,70%	1,48%	-3,85%	-1,32%
2020-09	312	1.915	R$ 1.911.123,30	-4,82%	-4,54%	-2,50%	4,00%
2020-08	320	2.012	R$ 1.904.263,80	0,60%	0,60%	14,29%	14,29%
2020-07	280	2.000	R$ 1.929.831,30				

Table 2. An example of summary with comparisons

Problem Statement

Your colleague Renato Silva, the manager responsible for analyzing customer behavior who directly uses data from the Data Warehouse Environment, needs to obtain all account monthly balances. Renato also needs to summarize the performance of all **transfers out** made via Pix to assess growth over time, comparing it to the previous month and the average of the last three months. Unable to complete this task alone, she has requested your assistance.

Add to your resolution the SQL query used to retrieve the data needed (the necessary tables in csv format were sent along with this pdf, on folder tables/).

Renato's friend, João Alves, wants to know how well PIX is doing inside LCGBR Bank. For that, he wants your help to come up with indicators that can be used to track the performance of the product. Which metrics would you suggest to track it and why?

Finally, propose a dashboard featuring metrics and insights relevant to the company.

Here is a summary of what we expect from you:

1. Create a SQL query to help Renato retrieve the monthly balance of all accounts.
2. Create a SQL query to help Renato retrieve the company's monthly performance using sliding window calculations.
3. Propose metrics to track PIX performance and its impact on LCGBR Bank. Feel free to suggest any metrics you consider relevant.
4. Propose a dashboard to demonstrate useful information and suggested metrics (a sketch is sufficient).
5. Create a dashboard using the tool of your preference (Bônus)

Here are some tips that might help you create your resolution:

- Feel free to send additional files if you wish.
- Although you were introduced to the entire data architecture, for this case, you can focus specifically on the data warehouse environment.