BUSINESS INTELLIGENCE / BUSINESS ANALYST, LEAD

BUSINESS CASE

Welcome to the RappiPay challenge for **Business Intelligence / Business Analyst.** From the interview we had, we can already tell that you're a very knowledgeable professional that we would love to call teammate. In this challenge, we are looking for ownership: ownership of the decisions, ownership of the data, and ownership of the business, hence imagine you're already a crucial part of the team and that many business decisions rely on your analysis. Let's get started…

[1] We just launched the credit card to market. As you might be aware, everyone was extremely busy planning and developing the product, but no one thought of coming up nor monitoring the key performance indicators of the business. **What would be the key performance indicators you would come up as the most important to monitor a credit card business? How often would you suggest such indicators must be monitored?**

[2] Dealing with diverse stakeholders is difficult. Where one might interpret a concept in a way, another one might differ from such interpretation. Let's take for example the concept 'dormant': some stakeholders might interpret the dormant customer as one that has not done any transactions in 6 months, where another one might say it takes only 4 months to reach this state. **Propose a problem resolution strategy with the stakeholders. How would you deal with this issue? Which facts would you present?**

[3] It is a common practice to have many systems scattered all over: where one might be hosting the app, others might be hosting models needed for daily operations. This usually benefits usability over scalability. Nevertheless, data centralization is crucial for data exploitation. For simplicity, imagine there are 4 systems:

- The first system hosts the app. It generates data that is stored in an internal database (ignore the database's architecture for now). Every time the user interacts with a screen, clicks a button, or opens the app, this is stored as an event.
- The second system hosts the risk model. Every time a customer asks for a credit, the system retrieves the risk data from the credit bureau and evaluates whether the customer is prone to be a defaulter.
- The third system hosts the customers information. Here, unrestricted information is hosted. This database contains the name, address, email, etc…
- Finally, the fourth and last system hosts all the payments information, this means, all the information related to the usage of the credit card: swipes, payments, recurrent payments, credit line, etc…

All systems share a unique identifier for all of our customers. This is the key that allows data to be joined on other databases.

What should we do to centralize the data in order to display it in charts for KPI monitoring? What would you propose the data governance strategy should be?

[4] Download the attached .csv file.

YOUR TASK IS TO exploit the information contained in the aforementioned file **as you find fit**. Some things to take into consideration:

[1] This database contains credit card information and transactions from multiple customers. Use your favorite data visualization tool / programming language to explore the data and present the results [R, Python, PowerBI, Spotfire, etc…]. The database has the following architecture:

ID	UPDATE	STATUS	MOTIVE	INTEREST_RATE	AMOUNT	CAT	TXN	CP	DELIVERY_SCORE

Where:

ID = This is the user's unique identifier.
UPDATE = Date when the event happened.
STATUS = The event, which can take the following values:

EMPTY - The user did not respond to the communication OR there was a transaction (this is reflected in the TXN column).

RESPONSE – The customer responded to the MKT campaign.

RISK – The customer was checked on the risk model whether the customer is fit to get a credit or not.

REJECTED – The risk model determined the customer is not fit to get a credit.

APPROVED – The risk model determined the customer as fit to get a credit. The customer is granted a credit. Here some of the other columns are populated.

DELIVERED – The customer received a physical credit card.

MOTIVE = The reason of rejection OR the type of card.
INTEREST_RATE = The interest rate of the customer's credit card.
AMOUNT = The amount of the credit granted to the customer.
CAT = The annual cost of the credit granted to the customer.
TXN = The amount of each transaction for each customer.
CP = Zip Code where the physical credit card was delivered to.
DELIVERY_SCORE = A score the customer gives to the delivery company for the delivery service.

Usually, the sign-up process starts when the customer responded to the communication, and ends-up with an approval, either with physical or digital card.

[2] Display and plot the information you consider to be the most relevant for a Credit card business. You could consider the following departments: Operations, Growth (Marketing), Finance, Customer Service, and Product.

[3] Use your imagination to best describe the data with charts and tables. Select those key performance indicators you consider that drive the business. **Present recommendations on those indicators that, to the best of your knowledge, might be low or could be boosted.**

[4] Think outside the box. If you feel that, extra information might be needed to support your arguments, include it in the folder: Power Point presentations, word documents, etc…

[5] Uploading your results to a git repo is desired but not mandatory.

The Rappi team and I wish you the best of lucks.

Fernando González Paulin
Head of Data Science, MX.
fernando.gonzalez@rappi.com