

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?**

**APPLICATIONS**

* **Response Rate:** Pretty importat metric to measure the effectiveness of the different campaigns through different channels (X-sell). It should be monitored at least once a month or after a campaign is launched. RR% = Responses / Total communications sent
* **Approval Rate:** Once we receive a positive response from Clients it is more likeble that a Client turns into a Booking. It should be monitored once a month. AR% = Approvals / Total Responses
* **Booking Rate:** Booking of a Credit Card and the final metric to see how many CCs we are delivering. It should be monitored once a month. BR = AR% \* RR% or Bookings / Total Applications

**CLIENT BEHAVIOR**

* **Utilization % :** Percentage of utilization of the total credit line of each client. Ut % = Balance / Credit Line
* **Balance/Salary:** Also important to know how much percentage of the Client’s salary is compromised to pay their debts.
* **Paybal**: How much percentage of the balance a Client is paying, massive importance to track payments behavior. Paybal = Payments / Balance

**PORTFOLIO BEHAVIOR**

* **Revenue:** All money we receive such as commissions, interests, fees. It should be monitored once a month or if we are testing something to see potential gains/losses of this metric in a specific period.
* **GCL/NCL:** Gross/Net credit losses: All money we loose because of lack oy payment, special give aways, cashback, etc.It should be monitored once a month.
* **Write OFF :** All clients we lose because of lack of payment after 4th month, this brings a negative balance due to all the money we cannot recover. It should be monitored once a month.
* **EBITDA:** Earnings before interests, taxes, depreciation and amortization. All money we received discounting write off, ncm, expenses, salaries, etc and before taxes. Very important to see the health of the portfolio.
* **Average Weighted Line:** Important to see the average line we are offering to our Customers. An increase on this metric could bring more revenue. AWL% = Credit line at booking / Accounts

**COLLECTIONS**

Right clients through all channels with the right offers. To see the behaviour we need to answer the following questions:

1. How many of our clients are enrolled with Rappi?
2. How many present a delay in their payments ?
3. Did we check if changing the limit payment date increases the payments?
4. Do we offer special offers once they present a delay ?
5. Which best practices around the globe are we testing right now?

We need to segment our Clients based on delinquency (Prevent, pre-delinquent, front-end, back-end,etc), detect patterns and monitor if there is a change in the payment behavior of each of them. Recovery should be relevant for the metric as well.

* **Recovery:** Pretty important to recover part of the balance we lost due to Write Off. Should be monitored several times in a month.

[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?**

First of all we have to understand that it is very important for a bank to have lines of credit go to people who are going to make transactions, charges or at least maintain a certain level of credit utilization. Having an inactivity in the Credit Card (CC) represents a cost to Rappi of not having that reserve of money (Credit line) in utilization. Since the demand of the RappiCard is high (A lot of applicants in the waiting list) a quicker response to this situation must be done.

For the best strategy a deep analysis can be done.

1. See how many clients got transactions back after a 4 month ‘dormant’.
2. Then, how many of those clients did them between the 4th and 6th month after the last transaction?
3. If the number is not high, what would be the increase of bookings, revenue, credit utilization, etc in that period of 2 months (between 4th and 6th month) if we set the period of 4 months ‘dormant’.

Also different projects to retain those customers should be done like:

* Sending push notifications to motivate clients to use the CC
* Reminders of CC utilization
* Checking at ‘Buro de Crédito’ special reports to answer questions like.. the client has other CCs? the utilization of competitors is higher? Has the client acquired a big debt in the previous months?

[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?**

We should have all data in a centralized tool in the cloud like Snowflake which controls not only data access pretty well but also implements consistent policies for data access according to specific roles.

Some of the best practices for data governance are the following:

* A security tool must be there to safeguard all information. There is no option for this.
* Enable roles to users allowing them to access to specific tables according to the role within RappiCard (RRHH, Marketing, Ops, etc)
* Once created the databases/tables a firewall must be there to avoid any kind of intrusion.
* Data masking and encryption must be there as well especially for sensitive data like ccv, email, passwords, etc.
* Data in final tables should be organized, clean and with some metadata about the table.
* It is very important to have all this data redundant so in case of failure or drop of information data can be restored easily.
* Finally a consistent audit should be done to avoid any compliances or privacy violations.

[4] Download the attached .csv file. *Preferably upload it to a SQL db and query your way through the challenge.*

**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:

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.

This point is already on a jupyter notebook in my github. Please go to the following Link:

* https://github.com/Jorch39/Rappi\_Collections

The Rappi team wishes you the best of lucks.