WIRECARD BI CHALLENGE

One of the challenges we face at Wirecard is to provide an accurate financial overview in order to help our decision makers to drive the company to its highest potential.

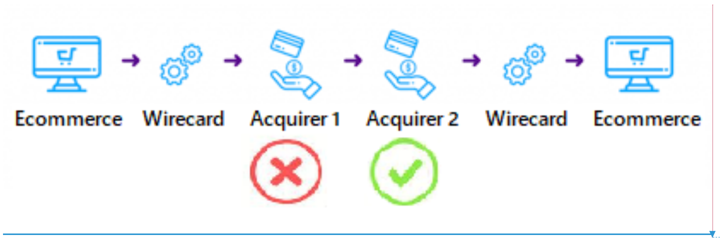
- **Business Case**

The main service provided by Wirecard Brazil is online payments. As a payment facilitator, we operate using four Brazilian acquirers in order to deliver the highest approval rate through requisition retrials to our clients.

The picture below presents the overall process used to authorize a purchase. Firstly, after navigating and selecting their desired products, buyers finishes they processes by selecting preferable payments methods and inputting required data (ie.: Credit Card Number).

After that, the eCommerce store sends data to Wirecard, who validates it and begins the authorization process. As a third step, the authorization process consists in individually and sequentially sending payment data to an Acquirer waiting for a “success” or “failure” status return.

If the transaction is approved by the first Acquirer, Wirecard communicates the eCommerce provider and the payment process is finished. If it is denied, Wirecard sends the authorization request to the second Acquirer. As a denial is returned, a new request is sent to the next Acquirer in line until the final one is reached. If all Acquirers reject the transaction, Wirecard then communicates the eCommerce and the payments process is finished with a negative status.



Given the explanation above, there are three tasks you must provide a solution:

1- **Dimensional Modelling**
2- **ETL**
3- **Business Visualizations**

Deleted: today

Deleted: ¶
Below follows your challenge description.¶

Deleted: Scenario

Deleted: You can picture it as a subacquirer service that

Deleted: operates

Deleted: b

Deleted: achieve

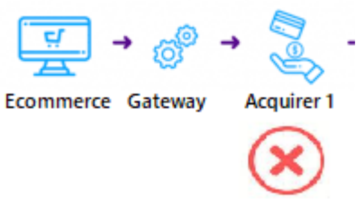
Deleted: Ex.: In t

Deleted: s

Deleted: are

Deleted: s

Deleted: , the payment was denied by Acquirer 1, then it was sent to Acquirer 2 which approved the transaction.¶
* If all four acquirers deny the transaction, the payment is cancelled.¶
¶



Deleted:

Deleted: ¶

Deleted: this scenario, we will split the challenge into three steps

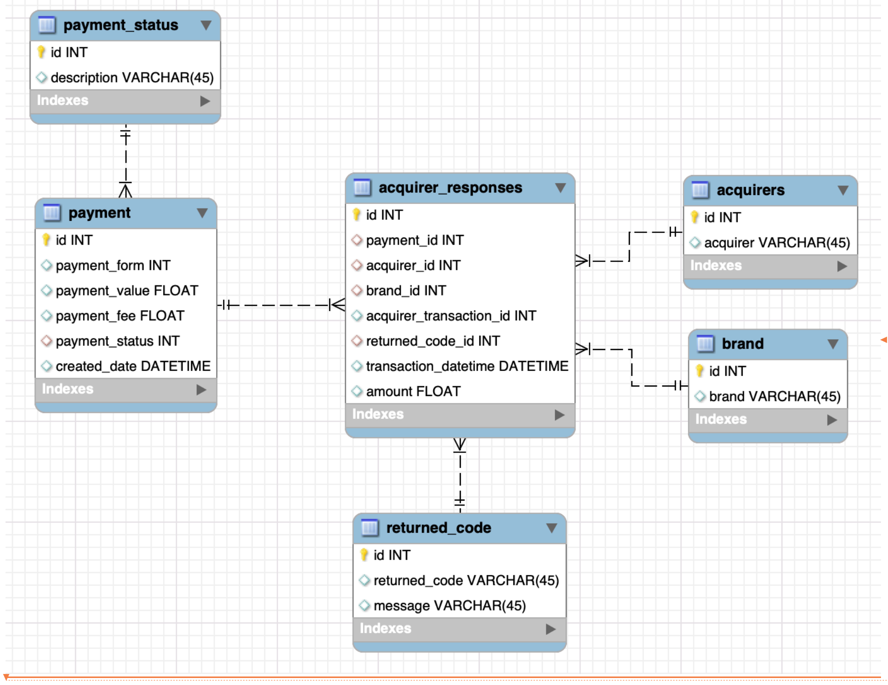
Task 1 - Dimensional Modelling

Your first step will be to build a dimensional model aiming to answer the following business questions: Total Payment Volume (TPV), Overall conversion rate, Acquirer efficiency (TPV per Acquirer and conversion rate), Retry rate and Retry Approval Rate.

The final deliverable must be a dimensional model which can be constructed using MySQL Workbench, or any other tool you are more comfortable with.

The final model, must be converted into one of the following formats: pdf, png or jpeg.

The original relational model can be found here.



Task 2 – ETL

Once you defined the dimensional model, let's implement your ETL process! We need to extract and transform the raw data and then load the model, to be able to work with it.

You must provide the architecture (and explain it) and implement it. You're free to use any technology you prefer (HDFS, sqoop, kafka, flume, hive, spark, shell script, python, et al.).

Deleted: STEP

Deleted: ¶

Deleted: such as

Deleted: acquirer

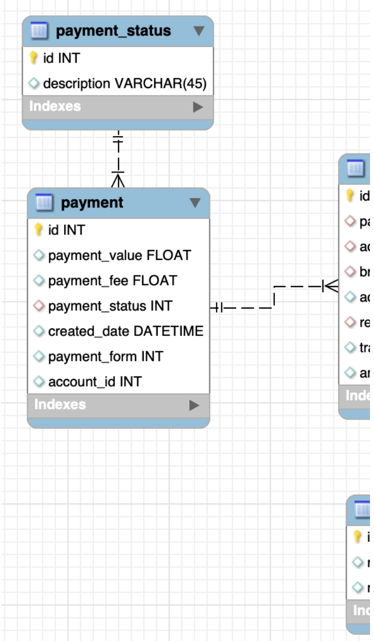
Deleted: , among others.

Deleted: As result, you must send your proposed dimensional model, which can be developed using opensource tools, such as

Deleted: like to work with.

Deleted: xls,

Deleted: (adicionar o link).



Deleted:

Formatted: Centered

Deleted: STEP

Deleted: ¶
As you may know, w

Deleted: In this step, y

Deleted: opensource

Deleted: etc.

Please, send us all scripts developed for this process. You must provide an explanation of all resources and thought process you used to build the solution, which will be used by is to test it.

Task 3 - Business Visualizations

Now that you have the dimensional model defined, the raw data extracted, transformed and loaded, it's time to have fun and show us how would you present your results to business people!

Here you can use all your creative resources to build business visualizations easy to understand. Feel free to use any visualization tool you want (Looker, Tableau, Excel, etc.)

As mentioned above, the basic required visualizations are:

- Total Payment Volume (TPV);
- Overall conversion rate (orders approved divided by total order amount);
- Acquirer efficiency (TPV per Acquirer and conversion rate);
- Retry rate and Retry Approval Rate.

You're encouraged to provide more insights and develop more visualizations you deem interesting for business management.

The final visualizations, must be converted one of the following formats: pdf, png or jpeg.

- **Raw Data**

Each acquirer has its own table sets, and we have a table (acquirer_responses) where you can find keys to track information from all table sets. Below follows the description of the tables that you'll use.

payment: it has data about the payments, such as the amount paid by the customer and the transaction fee. Columns description:

id: Is the table`s primary key. Wirecard's payment id. It can also be found in 'acquirer_responses' table in the column 'payment_id'.

payment_form: it identifies if the payment was made by credit card (1) or debit card (2). You can assume that all payments were made without installments.

payment_value: is the total amount paid by the customer.

payment_fee: is the transaction fee paid by our client. Is our transaction revenue.

payment_status: is the final payment status. Its description are in the table 'payment_status'.

created_date: is the payment creation date.

payment_status: it has the payment status description. Columns description:

id: Is the table`s primary key, also found in table 'payment' in the column 'payment_status'.

description: is the status`s description.

Deleted: you used while developing your process

Deleted: us and

Deleted: explain

Deleted: the

Deleted: knowledge

Deleted: c

Deleted: then we'll be able to validate and test it

Deleted: STEP

Deleted: ¶

Formatted: List Paragraph, Bulleted + Level: 1 + Aligned at: 0,63 cm + Indent at: 1,27 cm

Deleted: Hereby follows some visualization suggestions, but you'll have bonus points if you present other different concepts you can extract from data. (we highly recommend you to do so)¶

Deleted: ¶
Total credit card TPV¶
Credit card TPV by acquirer¶
Acquirer conversion rate¶

Deleted: xls,

Formatted: Font: Bold

Deleted: ¶
→account_id: each account id represents a client of ours.

Formatted: Font color: Text 1

Formatted: Font: Not Bold

acquirer_responses: it has the transaction data from all acquirers, including canceled transactions. Columns description:

id: Is the table`s primary key.

payment_id: Wirecard's payment id. It can also be found in 'payment' table in the column 'id'.

acquirer_id: is the acquirers`s id, also found in table 'acquirers' in the column 'id'.

brand_id: is the brand id, also found in table 'brand', in the column 'id'.

acquirer_transaction_id: is the acquirers`s transaction id.

returned_code_id: is the returned_code_id, also found in 'returned_code' table, in column 'id'.

transaction_datetime: is the datetime which the transaction was captured by the acquirer.

amount: is the total transaction amount captured by the acquirer.

returned_code: it has the acquirers`s returned codes and messages. Columns description:

id: Is the table`s primary key.

returned_code: acquirers`s returned codes.

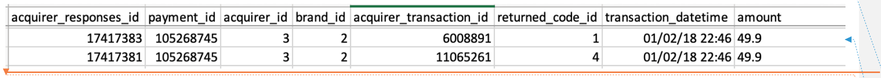
message: it has the correspondent messages from acquirers`s returned codes.

acquirers: it has the acquirers`s ids and names.

brand: it has the brands`s ids and names.

Observations about the data:

- You may find a payment_id more than once in acquirer_responses table. It can occur if the payment was denied by some acquirer and analized by others until its finally approved or denied.
- An easy way to identify an approved payment is by the returned_code_id = 1 (returned_code 0).
- An easy way to identify an cancelled payment is by the returned_code_id = 4 (returned_code 75).
- You may find a same payment approved by an acquirer and canceled moments after. This occurs because we have our own fraud prevention system, that can cancel suspicious transactions even if its approved by the acquirers. See an example below:



acquirer_responses_id	payment_id	acquirer_id	brand_id	acquirer_transaction_id	returned_code_id	transaction_datetime	amount
17417383	105268745	3	2	6008891	1	01/02/18 22:46	49.9
17417381	105268745	3	2	11065261	4	01/02/18 22:46	49.9

Deleted: the data you need to identify a payment into the acquirers' table sets. Here we have the field 'brand', which represents five different credit card brands.

Formatted: Font: Bold

Deleted: acquirer1: it has the keys to link to acquirer_responses table and the other tables from the acquirer 1 table set.¶

Formatted: Font: Bold

Formatted: Font: Bold

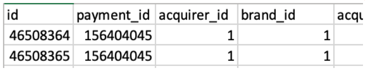
Deleted: acquirer1_authorizations: it shows information about the acquirer 1 authorization process. ¶
acquirer1_cancels: it shows information about the payments cancelled by acquirer 1.¶
acquirer1_captures: it shows all the transaction data from the payments captured by acquirer1.¶
¶
acquirer2: it has the keys to link to acquirer_responses table and the other tables from the acquirer2 table set.¶
acquirer2_authorizations: it shows information about the acquirer 2 authorization process.¶
acquirer2_cancels: it shows information about the payments cancelled by acquirer 2.¶
acquirer2_captures: it shows all the transaction data from the payments captured by acquirer2.¶
¶
acquirer3: it has the keys to link to acquirer_responses table and the other tables from the acquirer 3 table set.¶
acquirer3_authorizations: it shows information about the acquirer 3 authorization process.¶
acquirer3_cancels: it shows information about the payments cancelled by acquirer 3.¶
acquirer3_captures: it shows all the transaction data from the payments captured by acquirer3.¶
¶
acquirer4: it has the keys to link to acquirer_responses table and the other tables from the acquirer 4 table set.¶
acquirer4_authorizations: it shows information about the acquirer 4 authorization process.¶
acquirer4_cancels: it shows information about the payments cancelled by acquirer 4.¶
acquirer4_captures: it shows all the transaction data from the payments captured by acquirer4.¶

Deleted: <#>Note that the fields 'amount', 'total', 'total_amount' represents the value paid by the client, and in some tables, this fields are 'int' type, and other are 'float' type. When in 'int' type, the value is represented in cents. Ex.: R$10,00 is represented in some tables as 1000 ('int' type) or 10.00 ('float' type). We would like to see how you can handle that.¶ ... [1]

Deleted: Note that tables received from each acquirer has its own data model, and some acquirers send more data than others. We also would like to see how you can handle that.

Formatted: Indent: Left: 1,27 cm, No bullets or numbering

Deleted:



id	payment_id	acquirer_id	brand_id	acqu
46508364	156404045	1	1	
46508365	156404045	1	1	

Formatted: Normal, Centered, Indent: Left: 0,63 cm, No bullets or numbering

- Important Guidelines

- You need to use all table files provided.
- During the process, you can use any opensource tool you want (MySQL, pentaho, python, etc.).
- We highly recommend you to document your scripts. Remember that we'll evaluate all your work by this, so a well written documentation can avoid misunderstandings.
- Remember to check if you're sending all files you used in this challenge, we'll need it to evaluate your work.

- What do We Expect From You

As you'll develop the data pipeline from top to bottom, since the raw data extraction to business clients visualizations, we expect you to be very curious (don't be shy to ask us questions), creative, and able to communicate to developers, and business people. You're totally free to suggest any improvement to the process and build the solution in the best way you can.
We also recommend you to contact us if you have any question about the challenge, the business or the raw data. We'll be happy to help you!

We wish you can do a nice job!

Page 4: [1] Deleted	Microsoft Office User	9/27/19 5:59:00 PM

-