DATA BUSINESS ANALYST TEST

ANA DELGADO GARCIA

OCTUBRE 2023

1. SQL

You have the customer_courier_chat_messages table that stores data about individual messages exchanged between customers and couriers via the in-app chat.

You also have access to the orders table where you have an order_id and city_code field. An example of the tables is in excel attached in the email.

Your task is to build queries for the following questions (Please add screenshots of the results obtained):

1. Show all the columns order by order_id

```
SELECT * FROM customer_courier_chat_messages
ORDER BY order_id
```

0%   18:2

Result Grid   Filter Rows: Search   Export:

sender_app_type	customer_id	from_id	to_id	chat_started_by_messa...	order_id	order_stage	courier_id	message_sent_ti...	city_cod
Courier Android	122	87	122	TRUE	38	ADDRESS_DELIVERY	87	9/8/22 7:55	BCN
Customer IOS	23	23	21	FALSE	134	PICKING_UP	21	7/8/22 10:02	OPO
Courier IOS	23	21	23	FALSE	134	ARRIVING	21	7/8/22 10:01	OPO
Customer IOS	23	23	21	FALSE	134	PICKING_UP	21	7/8/22 10:00	OPO
Customer IOS	99	99	21	FALSE	555	PICKING_UP	21	9/8/22 8:02	BCN
Courier IOS	99	21	99	FALSE	555	ARRIVING	21	9/8/22 8:01	BCN
Customer IOS	99	99	21	FALSE	555	PICKING_UP	21	9/8/22 8:00	BCN
Customer Android	43	43	75	FALSE	875	PICKING_UP	75	7/8/22 14:55	VAL
Courier Android	43	75	43	FALSE	875	ARRIVING	75	7/8/22 14:53	VAL
Customer Android	43	43	75	FALSE	875	PICKING_UP	75	7/8/22 14:51	VAL
Courier Android	43	75	43	TRUE	875	ADDRESS_DELIVERY	75	7/8/22 14:50	VAL

2. Show all the columns order by city_code

```
SELECT * FROM customer_courier_chat_messages
ORDER BY city_code
```

00%   19:2

Result Grid   Filter Rows: Search   Export:

sender_app_type	customer_id	from_id	to_id	chat_started_by_messa...	order_id	order_stage	courier_id	message_sent_ti...	city_cod
Customer IOS	99	99	21	FALSE	555	PICKING_UP	21	9/8/22 8:02	BCN
Courier IOS	99	21	99	FALSE	555	ARRIVING	21	9/8/22 8:01	BCN
Customer IOS	99	99	21	FALSE	555	PICKING_UP	21	9/8/22 8:00	BCN
Courier Android	122	87	122	TRUE	38	ADDRESS_DELIVERY	87	9/8/22 7:55	BCN
Customer IOS	23	23	21	FALSE	134	PICKING_UP	21	7/8/22 10:02	OPO
Courier IOS	23	21	23	FALSE	134	ARRIVING	21	7/8/22 10:01	OPO
Customer IOS	23	23	21	FALSE	134	PICKING_UP	21	7/8/22 10:00	OPO
Customer Android	43	43	75	FALSE	875	PICKING_UP	75	7/8/22 14:55	VAL
Courier Android	43	75	43	FALSE	875	ARRIVING	75	7/8/22 14:53	VAL
Customer Android	43	43	75	FALSE	875	PICKING_UP	75	7/8/22 14:51	VAL
Courier Android	43	75	43	TRUE	875	ADDRESS_DELIVERY	75	7/8/22 14:50	VAL

3. <u>Show the first message (row) sender (courier or customer)</u>

```sql
SELECT *
FROM customer_courier_chat_messages
ORDER BY message_sent_time
LIMIT 1;
```

sender_app_ty...	customer_id	from_id	to_id	chat_started_by_messa...	order_id	order_stage	courier_id	message_sent_ti...	city_code
Customer IOS	23	23	21	FALSE	134	PICKING_UP	21	7/8/22 10:00	OPO

4. <u>Show the number of messages sent by the customer and order_id</u>

```sql
SELECT customer_id, order_id, COUNT(*) as num_messages_sent
FROM customer_courier_chat_messages
WHERE sender_app_type LIKE 'Customer%'
GROUP BY customer_id, order_id;
```

customer_id	order_id	num_messages_se...
23	134	2
43	875	2
99	555	2

5. <u>Show the first message (row) in the conversation by order_id</u>

```sql
SELECT
    order_id,
    MIN(message_sent_time) AS first_message_time
FROM
    customer_courier_chat_messages
GROUP BY
    order_id;
```

order_id	first_message_ti...
38	9/8/22 7:55
134	7/8/22 10:00
555	9/8/22 8:00
875	7/8/22 14:50

6. <u>Show the last message (row) in the conversation by order_id</u>

```sql
SELECT
    order_id,
    max(message_sent_time) AS last_message_time
FROM
    customer_courier_chat_messages
GROUP BY
    order_id;
```

esult Grid Filter Rows: Search Export:

order_id	last_message_ti...
38	9/8/22 7:55
134	7/8/22 10:02
555	9/8/22 8:02
875	7/8/22 14:55

7. <u>Show the time (in secs) elapsed until the first message was responded by order_id</u>

```sql
SELECT
    order_id,
    TIMESTAMPDIFF(SECOND, MIN(CASE WHEN sender_app_type LIKE 'Customer%' THEN message_sent_time END),
    MIN(CASE WHEN sender_app_type LIKE 'Courier%' THEN message_sent_time END)) AS time_until_response_in_seconds
FROM
    customer_courier_chat_messages
GROUP BY
    order_id;
```

esult Grid Filter Rows: Search Export:

order_id	time_until_response_in_seco...
38	NULL
134	60
555	60
875	-60

8. <u>Build a query that aggregates individual messages into conversations. The query result should be used to create a table customer courier conversations (take into consideration that a conversation is unique per order).</u>

The required fields are the following:

• order_id

• city_code

• first_courier_message: Timestamp of the first courier message

- first_customer_message: Timestamp of the first customer message
- num_messages_courier: Number of messages sent by courier
- num_messages_customer: Number of messages sent by customer
- first_message_by: The first message sender (courier or customer)
- conversation_started_at: Timestamp of the first message in the conversation
- first_responsetime_delay_seconds: Time (in secs) elapsed until the first message was responded
- last_message_time: Timestamp of the last message sent
- last_message_order_stage: The stage of the order when the last message was sent

```
CREATE TABLE customer_courier_conversations AS
SELECT
    order_id,
    city_code,
    MIN(CASE WHEN sender_app_type LIKE 'Courier%' THEN message_sent_time END) AS first_courier_message,
    MIN(CASE WHEN sender_app_type LIKE 'Customer%' THEN message_sent_time END) AS first_customer_message,
    SUM(CASE WHEN sender_app_type LIKE 'Courier%' THEN 1 ELSE 0 END) AS num_messages_courier,
    SUM(CASE WHEN sender_app_type LIKE 'Customer%' THEN 1 ELSE 0 END) AS num_messages_customer,
    CASE WHEN MIN(message_sent_time) = MIN(CASE WHEN sender_app_type LIKE 'Courier%' THEN message_sent_time END) THEN 'Courier' ELSE 'Customer' END AS first_message_by,
    MIN(message_sent_time) AS conversation_started_at,
    TIMESTAMPDIFF(SECOND, MIN(CASE WHEN sender_app_type LIKE 'Customer%' THEN message_sent_time END),
    MIN(CASE WHEN sender_app_type LIKE 'Courier%' THEN message_sent_time END)) AS first_responsetime_delay_seconds,
    MAX(message_sent_time) AS last_message_time,
    MAX(CASE WHEN sender_app_type LIKE 'Courier%' THEN order_stage END) AS last_message_order_stage
FROM
    customer_courier_chat_messages
GROUP BY
    order_id, city_code;
```

```
SELECT * FROM customer_courier_conversations;
```

% 45:1

sult Grid Filter Rows: Search Export:

order_id	city_code	first_courier_messa...	first_customer_messa...	num_messages_courier	num_me...	first_message_...	conversation_started...	first_resp...	last_message_ti...	last_message_order_st
38	BCN	9/8/22 7:55	NULL	1	0	Courier	9/8/22 7:55	NULL	9/8/22 7:55	ADDRESS_DELIVERY
134	OPO	7/8/22 10:01	7/8/22 10:00	1	2	Customer	7/8/22 10:00	60	7/8/22 10:02	ARRIVING
555	BCN	9/8/22 8:01	9/8/22 8:00	1	2	Customer	9/8/22 8:00	60	9/8/22 8:02	ARRIVING
875	VAL	7/8/22 14:50	7/8/22 14:51	2	2	Courier	7/8/22 14:50	-60	7/8/22 14:55	ARRIVING

2. Email Request Simulation

You received the following email from a client:

Hello,

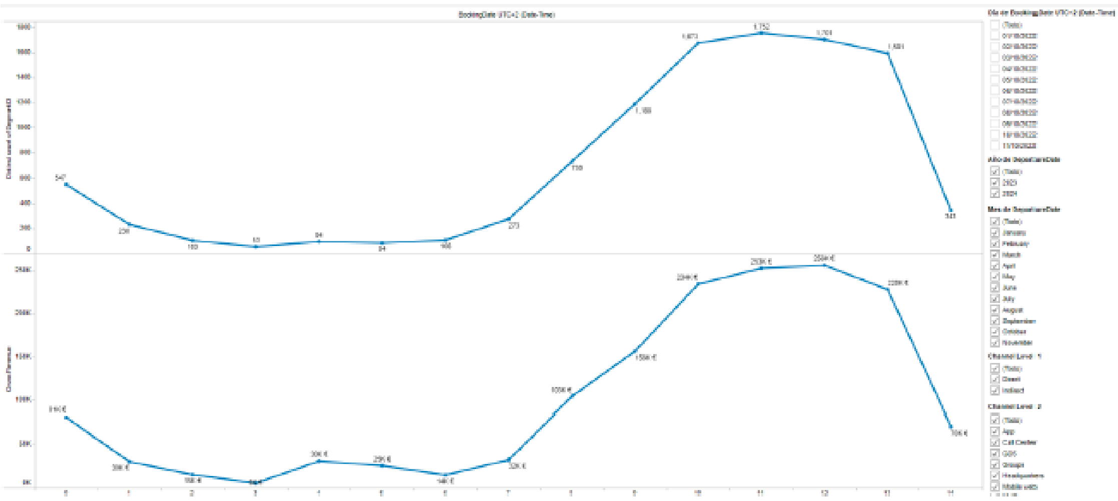
We would like to have a dashboard of “Hourly Payments” similar to our “Hourly Sales” (attached).

It’s not mandatory to be in Tableau, we can discuss about other platforms.

With this, we want to see similar graphics, but payment related. Maybe we could add an additional filter to see data by payments type. Please suggest any enhancements you believe will bring added value to business,

Thank you,

And this is the dashboard they are referring to:



After getting in touch with the technical team, you receive this feedback:

> Hi,
>
> This development is quite simple, it should take half a day to implement.
>
> Nevertheless, considering current ongoing projects, we can only do it in 3 weeks. Let me know if this is ok or if the client wants to change priorities.

With all the previous information, please elaborate:

1. A brief description of what you would transmit to the technical team so they can start developing what is being asked.
2. The answer you would provide to the client, considering the feedback from the tech team.

WHAT I WOULD TRANSMIT TO THE TECHNICAL TEAM

In this practical scenario, where the client mentions their desire for a "Hourly Payments" dashboard similar to the "Hourly Sales" dashboard, reflecting payment data on an hourly basis and resembling the style and functionality of the existing dashboard, the primary goal conveyed to the technical team is to present payment-related data and enable the visualization of graphics similar to those found in the hourly sales dashboard. Additionally, there is a request to add a filter to view data by payment type.

Furthermore, the technical team is informed that the client does not have specific platform preferences, allowing for discussions on alternative options beyond Tableau. It is suggested that alternatives such as Microsoft's Power BI be considered, as it integrates well with other Microsoft tools, which can be advantageous if the organization already utilizes Microsoft products. Additionally, options like QlikView and QlikSense are recommended due to their user-friendliness and flexibility in creating dashboards and visualizations.

However, whether for the use of Tableau or the suggestion of other tools, more information is needed, such as existing data infrastructure, technical team preferences, budget, and specific project requirements. It is essential to discuss these options with both the technical team and the client to make an informed decision regarding the platform that best suits the case.

In the event that the information to be provided to the technical team needs to be conveyed with the client present, it may be worthwhile to review the various steps required for implementing the changes. This approach adds value to the process by demonstrating clarity and transparency. Emphasis should be placed on the following parts of the process:

Extracting and Cleaning the Data: It should be noted that the extraction may involve one or several data sources. Subsequently, data cleaning and preparation are necessary.

1. Designing the Dashboard: In this case, it is essential to consider that, regardless of the tool used, the interface should effectively and attractively display the graphics while maintaining a style and functionalities similar to the existing one.

2. Implementing New Filters: Since the email mentions the need for a payment type filter, this functionality should be added to the dashboard.
3. Testing and Validation: Once the aforementioned steps are completed, testing will be conducted to ensure the dashboard functions correctly, and that the data is displayed as expected.
4. Documentation: Detailed documentation of all steps should be prepared to have information on the implementation and operation should it be needed for future reference.

MAIL TO THE CLIENT

Dear client

Thank you for your message and for trusting us to develop the new 'Hourly Payments' dashboard.

We have consulted with our technical team, and they are ready to begin the development of this project.

Considering our current workload and ongoing commitments, we estimate that the project implementation will take approximately 3 weeks. This extended timeframe will allow us to ensure a proper execution that meets your expectations.

We are confident that this new payment dashboard will provide valuable insights for your business and facilitate a better understanding of your payment data. Additionally, we will explore the possibility of adding a filter to view data by payment type, providing greater flexibility in data visualization.

Regarding potential improvement suggestions, we will collaborate closely with you during the development process. As we progress in the project, we will present concrete ideas tailored to your business needs.

In terms of the possibility of using other dashboard visualization tools, we are open to exploring alternative options that may be more suitable for your requirements. If you have any specific preferences or would like to discuss other tools, please feel free to let us know for evaluation by our technical team.

We remain at your disposal for any questions or concerns you may have throughout the process.

Sincerely,
Ana

ENDNOTE

It’s possible that the client may not be satisfied with the timeframe we have provided for the project implementation. Therefore, I believe it is crucial to maintain open and effective communication. In such a scenario, I would suggest considering, in collaboration with the technical team, the possibility of shortening the delivery time while taking into account the client's needs and expectations. This could involve actions such as considering the reallocation of resources from other teams or exploring alternative solutions to meet the requirements within a shorter timeframe. Our priority would be to ensure that the client is content with the final outcome, and we are willing to do whatever it takes to achieve that.