Marketing Data Analyst (m/f/x) - Data Challenge

About Clark Germany

Clark is a digital insurance broker. Customers can upload their existing insurance contracts to our platform, obtain an overview of their current insurance situation and receive recommendations to improve their cover by getting cheaper and better policies or buying additional cover to cover insurance gaps. Insurance experts are always available for customers via phone, Email and in-app messenger.

Clark User Journey

Clark is a digital insurance broker that operates on both web and app platforms. After landing at our homepage or installing our app, users initiate a registration funnel where they need to tell Clark what contracts they already have, share some personal information, provide an email address and set up a password. As the last step of the funnel, users are expected to give a signature to authorise Clark to become their insurance broker.

After being verified, users can start using Clark. Below are some activities they can with Clark app/website:

- Upload existing contracts into my Clark account, so that I can have an overview of my policies.
- Request a new offer from Clark if I don't like one current insurance policy
- Do a demand check to provide more personal information and get an assessment of my current insurance situation
- Reach Clark customer service team via Email, phone or in-app messenger if I have any questions about my insurances
- and more …

Clark Business Model

Our business generates two types of revenue.

On one hand, by managing the existing insurance contracts for the customers, we obtain management commission fees every year as long as the customers do not revoke and do not terminate their customers with the insurers. Hence, it is the so-called recurring revenue.

On the other hand, by selling new contracts of certain categories (e.g. dental insurance), we receive sales commission fees from the insurers. We receive this revenue only once when the sale is closed. While for some other insurance categories (e.g. personal liability insurance), when we sell the new contracts, we don't get any sales fee.

Once we sell a new contract, regardless of the insurance category, we are able to also take it under management and thus receive recurring management fee.

Data Description

Dataset 1: facts_customer_extended

In this dataset, each entry is a customer and can be uniquely identified with column "Customer ID". Tabel 1 outlines the structure of table *facts_customer_extended*.

Dataset 2: facts_contract

In this dataset, each entry is a contract and can be uniquely identified with column "Contract ID". Tabel 2 outlines the structure of table *facts_contract*.

Questions

Please read through the following questions and get back to us in the form of slides. Please also share any scripts that you have used when you solve the challenge.

Question 1

From the two datasets, we would like to have an overview about the contract generation throughout customer lifetime.

The expected output is the following table:

Customer Created Month	**1**	**2**	**3**	**4**	**...**
2017/01	A	B	x%	x%	x%
2017/02	x%	x%	x%	x%	x%
2017/03	x%	x%	x%	x%	x%
2017/04	x%	x%	x%	x%	x%
2017/05	x%	x%	x%	x%	x%

For example, the highlighted A means that, out of 733 customers that were created in the month of 2017/01, K1 unique customers created at least one contracts within the **first 30 days** after their individual sign-up date and we have A = K1 / 733 *100%. B means that, out of 733 customers that were created in the month of 2017/01, K2 unique customers created at least one contracts within the **2nd 30 days** after their individual sign-up date and we have B = K2 / 733 *100%.

Please provide a SQL query that could achieve the desired output above. Note: please attach your code/script in addition to the final output.

<u>Other requirements</u>

- Please specify which SQL dialects you use.
- Make sure your code is clean, well-formatted and well-commented.
- It is sufficient to have an output with the following format - it is not necessary to present it in a matrix format.

Customer Created Month	Period	Percentage
2017/01	1	x%
2017/01	2	x%
...	...	...

Question 2

The management team at Clark is interested in one question: what kind of customers are the Most Valuable Customer (MVC) for Clark?

Please play around with the table *facts_customer_extended* and defend your hypotheses with data.

Other requirements

- First define what it means to be a high-value Clark customer (including what metrics you would look at to evaluate the value of a Clark customer) and describe your framework/methodology, then show us how different these MVC customers are compared against average Clark customers in terms of their demographic traits, portfolio composition and customer value.
- One example hypothesis could be: the older the customer, the higher the revenue he could generate for Clark. The rationale behind this hypothesis is that, the older the customer, the more existing contracts he has in his portfolio and the higher the management fee Clark can receive from insurers.
- Given that the datasets are sample data, it is very likely that you are unable to end up with any significant findings. The purpose of this challenge is not for you to come up with the perfect answers but to show us your process of tackling real-life, ambiguous projects.
- Please use a programming language of your choice (R, Python, etc) to answer this question.

Question 3

Apart from the dimensions included in table facts_customer_extended, can you suggest some additional attributes that could affect the value of Clark customers and provide a brief explanation for each proposal?

Question 4

Consider the following KPIs and elaborate on the importance of each:

- Customer Lifetime Value (CLV)
- Return on Ads Spend (ROAS)
- Cost per Acquisition (CPA)
- Value ratio (CLV/CPA)

Which amongst the above KPIs is the most fitting indicator according to you and why?

Trick Questions

For answering these questions, please try to think outside the box.

Question 5

You have to explain the difference between a table, a view, and a materialized view to your bright, curious, 12-year old niece who doesn't know anything about databases. How would you do it?

Question 6

Clark provides you a rope that takes 1 minute to burn out. The rope is non-uniform in nature i.e. it has different burning rates at different areas. You have been given the task of measuring 30 seconds just by burning this rope. How will you achieve this? Kindly, brief the process.

Question 7

You lost your Euro 2020 bet to your friend. The only way he agreed to give you a waiver is if you can solve this trick. He gives you eight coins and puts them on a surface in front of you. The coins are made as such that the surfaces on both sides are same, and it is not possible to tell the difference between Heads and Tails by touching it alone.
Now out of those eight coins, 4 are heads up, and 4 are tails up. He asks you to flip some coins and make two bundles and pile them up such that each of the bundles has an equal number of heads up.

Appendix

Table 1. Table Structure of *facts_customer_extended*

Column Name	Description	Comments
Customer ID	The unique identifier of the customer in our DB	This is the primary key of table *facts_customer_extended*.
Registration Date	The date on which the customer gave Clark his insurance mandate and became a Clark customer	N/A
Network	The source from which the customer came to Clark	For example, if a customer searches for branded or generic keywords on Google, click on the link to our website and register as a user, he or she will be labeled as "Search" in our DB.
Platform	A flag indicating whether the customer is an app user or a web user	"app" means that the customer completed the registration in our app and "web" means that the customer did that on our website.
Age	The age of the customer	N/A
Gender	The gender of the customer	N/A
Marital Status	The marital status of the customer	N/A
Kids	A flag indicating whether the customer has kids or not	"YES" means that the customer has at least one kids and "NO" means that the customer does not have kids yet.
Occupation	The occupation of the customer	N/A
Automobile	A flag indicating whether the customer has automobiles or not	"YES" means that the customer has at least one automobiles and "NO" means that the customer does not have any automobiles yet.
Property	A flag indicating whether the customer has properties or not	"YES" means that the customer has at least one automobiles and "NO" means that the customer does not have any automobiles yet.
Number of Logins	How many times the customer has logged in on our app or website since registration	N/A
Last Login Date	The last time when the customer logged in with his account in our app or on our website	N/A

Number of Interactions	The number of inbound and outbound interactions between the customer and Clark in various forms incl. phone calls, sms, email and in-app messenger.	N/A
Number of Exist. Contracts	The number of existing contracts the customer uploaded to Clark	N/A
Number of New Contracts	The number of new contracts the customer bought via Clark	N/A
Sum of Yearly Management Fee	The sum of yearly management fee from the contracts of the customer that are currently under management by Clark	Note: both existing contracts and new contracts can generate management revenue.
Sum of Sales Fee	The sum of sales fee from the sales of contracts to the customer	N/A

Table 2. Table Structure of *facts_contract*

Column Name	Description	Comments
Contract ID	The unique identifier of the contract in our DB	This is the primary key of table *facts_contract_extended*.
Created Date	The date on which the contract was entered into our DB	N/A
Type	A flag indicating whether the contract is an existing one or a new one	N/A
Insurance Category	The insurance category of this contract	N/A
Yearly Management Fee	The yearly management fee from the contract that Clark receives every year	N/A
Sales Fee	The sales fee from the sales of this contract	N/A
Customer ID	The identifier of the customer behind this contract	N/A