Take-Home challenge

## Data Analyst / BI Engineer

## Context

First of all, thank you for taking the time to apply at Astrafy and being willing to complete this technical take-home challenge. We are well aware that your time is precious and are grateful that you are willing to take the time to demonstrate to us your skills via this challenge.

The challenge is made up of a design challenge and a coding challenge. Once you submit your solution to those challenges, we will review it upfront and then set a meeting for you to go through your solution and explain the process that led you to this design and code architecture.

## Generic Guidelines

There is not a perfect solution for each of those challenges and so is the case for the projects you would be working on at Astrafy. Problems can be tackled in many different ways and we are as much interested in the thought process that led you to those designs and code than in the final results (“the journey matters more than the destination”).

Some guidelines we highly recommend you to follow:

* **Documentation is as important as the code**: we should be able to understand everything you did by reading the README, Docstrings in your code, eventual short tutorials (using Loom for instance), etc.
* **Do not reinvent the wheel**: some say that savvy developers excel at “copy-pasting”. Chances are that a library or something similar to what you are trying to do has already been done. A developer's job is sometimes putting together different Lego blocks and making them work all together in a seamless way. We therefore encourage you to look on the web and take inspiration from similar problems.

## Submission

Once you are done with the challenge, please follow those guidelines:

1. For the design challenge, save your deliverables as pdf.
2. For the coding challenge, put your code on a public Github repository and put the link of your Looker Studio in the README.md. You can share the dashboard with the google group [founders@astrafy.io](mailto:founders@astrafy.io) and [bi@astrafy.io](mailto:bi@astrafy.io).
3. Reply to the email you received with this “Take-Home challenge” with the pdf from step 1 attached and a link to the Github repository (step 2).

## Data:

* Orders: sales lines between 01-01-2022 and 31-12-2023 (1 line per order)
* Sales: sales lines per product between 01-01-2022 and 31-12-2023 (1 line per order and per product)

## Part 1: Coding Challenge

### Introduction

* Implement using dbt and Bigquery
  + <https://docs.getdbt.com/docs/introduction>
  + <https://cloud.google.com/bigquery/docs/introduction>
* Organize the dbt project in a way to present it during the technical interview
* Document
* Save in a git repository

### Exercice 1:

What is the number of orders in the year 2023?

### Exercice 2:

What is the number of orders per month in the year 2023?

### Exercice 3:

What is the average number of products per order for each month of the year 2023?

### Exercice 4:

Create a table (1 line per order) for all orders in the year 2022 and 2023; this table is similar to orders with an additional column: the qty\_product column that gives the quantity of products in the order, for all orders in 2022 and 2023

### Exercice 5: Order segmentation

Orders are segmented into 3 groups:

* New: it's the 1st order of the customer (client\_id) in the past 12 months. In the 12 months prior to this order, the customer did not place any orders
* Returning: it's between the 2nd and the 4th order of the customer in the past 12 months. In the 12 months prior to this order, the customer had already placed between 1 and 3 orders
* VIP: it's the 5th or more order of the customer in the past 12 months. In the 12 months prior to this order, the customer had already placed at least 4 orders or more

Calculate for each order placed in 2023, the segment of this order.

### Exercice 6:

Create a table (1 line per order) for all orders of the year 2023 only; with an additional column: the order\_segmentation column which gives the segment of this order

## Part 2: LookML: Semantic Layer

Implemented the LookML semantic layer based on your previous work (the code must be stored in a git repository). **The final result must be usable within Looker.**

The semantic layer must enable the creation of a business-relevant dashboard by including all necessary dimensions and measures. It should support customer segmentation (Part 1 Ex. 5) directly within the semantic model to allow for meaningful visualizations.

The semantic layer should be easy to understand and maintainable.

## Part 3: Design Challenge

The marketing team of an e-commerce site has requested your help in creating a dashboard for daily analysis. The goal is to track key performance indicators (KPIs) that will allow them to monitor business activity, identify trends in revenue fluctuations, and better understand their customer base.

Propose a dashboard that includes relevant KPIs to achieve these objectives.

(Bonus): The marketing team would also like to have predictions for the upcoming week or month regarding sales performance. Please include any forecasts or predictive insights that could help them anticipate future trends.