Take-home exercise for Data Analyst position

1) Backstory

In **Mock Company**, we use two systems to manage our customer and sales data.

Our CRM (Customer Relationship Management) system is the single source of truth for our customer master data. The data can be found in the file: customers.csv. Our E-commerce system contains information on all the purchases and can be found in the file purchases.csv.

Our Stakeholders in Marketing want to launch a new targeted campaign and they need more insights into different segments of our customers. Using the two datasets provided we want to create a dashboard based on their requirements. The requirements are listed below.

2) Loading the data

At ClimatePartner we work with Power BI as our BI tool, and we would therefore prefer it if the challenge were done in Power BI.

* Load the data into a workbook.
* Create a data model from the two datasets by joining them on a unique id.

3) Requirements & Analysis

Here are the requirements from our colleagues in Marketing:

1. We would like a tabular overview of the customers with basic information about them along with the number of purchases they have placed, the revenue they have generated, and the average revenue per purchase.
2. We would like to know which departments our customers work in, and how they are distributed geographically within the US.
   1. Create a stacked bar chart showing which departments customers are working in, and color coded by their location.
   2. For more details on the above visualization, please build a matrix that shows the monthly income we make from different departments with the possibility to expand on the locations, and the individual customers.
3. We would like to see a pie chart of the income per customer city showing the percentage each city is responsible for.
4. We would like to see the cumulative income for the last 3 years to determine if there is a month where customers purchase more/less than other months. We would like to be able to slice this by department and location.
5. Are there any other customer characteristics in the data that could help us segment the customers?
6. BONUS question (If you are not able to solve this one, don’t worry.): We are also considering if we should rather be segmenting by what the customers buy. Could you provide the top 3 product categories based on income?

4) Submitting your results

We appreciate your efforts to do this take-home exercise in your free time, and we do not want it to be too heavy a burden on you. We tried to keep it simple, so our estimate is that it should be doable in a couple of hours. However, if you can only afford to spend less time on this exercise, please do not be discouraged. Submit your partial results, and just mention the amount of time you spent on it.

* If you have any questions or thoughts during the exercise, feel free to mention them in your submission.
* Please submit your results to: jobs@climatepartner.de
* The results should be sent as a .pbix file if using Power BI or as a pdf of the visualizations otherwise.

We will review your submission as soon as possible and get back to you. During the review we are going to focus on the following aspects:

* You built a usable data model with the raw data provided.
* You show that you have understood the information contained in the data.
* You show that you can visualize the data in a meaningful manner.
* You perform the necessary transformation steps to provide the requested analysis.