(Senior) Business Intelligence Analyst case study

Table of Contents

Context

As a BI Analyst, you have been tasked with performing an exploratory analysis on a dataset of customer orders in a time period. The goal is to identify variance in the performance across different dimensions, which can be used to understand potential improvement in our product and business:

- What major insights can you identify for the management team using the data?
- Which patterns do you feel would be worth investigating further?
- How do you recommend to grow our revenue?

You can use tools of your own choice to analyse, visualise and present your findings. In addition to presenting your findings, you should be able to show how you performed your analysis.

Note: We are interested in learning about your data exploration process and analytical reasoning, and not in any specific correct answer

Data Exploration

We are looking to evaluate two important aspects of your work as a BI Analyst

- **Technical/Analytical:** Your technical skills in working with tools and your analytical skills - Presented to Data & Insight team members
- **Business**: Your ability to communicate findings and translate them into business implications - Presented to key stakeholders.

Prepare to present your take on the questions above and discuss trade-offs and your analytical process. Choose how you would like to present this, but please limit your time to

15 minutes. The rest of the time will be spent on questions and discussion. We don't expect you to spend more than a couple of hours at most preparing for this case. You can choose to focus on specific parts instead of covering all possible paths. Feel free to make any assumptions you feel make sense

Attached datasets

The attached file contains 3 datasets as CSV files:

orders.csv

It contains a dataset of orders from a random subset of our customers made in a period. The file contains the following fields:

- order_id: An anonymous identifier for the order an order line belongs to
- account_id: An anonymous identifier for the customer who placed the order
- booking_date: The date the order was booked
- service_date: The date the order order was serviced/dispatched
- transport_mode: The mode of transport used for the services
- origin: The starting point of the shipment
- destination: The ending point of the shipment
- metric: The unit of cargo in cbm, teu, kg
- quantity: The amount of cargo
- revenue: The order value of the shipment
- margin_percent: The percentage of margin made on the order

customers.csv

It contains additional information on the customers in the orders dataset. The file contains the following fields:

- account_id: Anonymous identifier for the customer
- account_manager: The responsible account manager for this account from forto
- company_type: The identifier to check the persona of a company
- sub_team_id: The identifier of the sub team responsible for this account
- geography: The region of the shipment
- Industry: Industry of the account/customer
- sub_industry: The sub industry of the account/customer

eurostat_volumes.csv

We are also providing external data from Eurostat that mirrors the industry trends across various transport modes, regions, and industries. Extra points will be awarded to the candidates answering the following questions:

- What trends you can identify from the data?
- How will those trends benefit forto?
- What steps should be taken by forto based on those trends?

Good Luck