BUX

Analytics Engineer Assignment

Intro

As an analytics engineer at BUX we expect from you that you are able to translate data requirements from other teams into an easy-to-use data model. Also, we expect that you know how to create insightful reports using best practises in data visualisation and dashboard design. The goal of this assignment is to demonstrate that you possess these capabilities.

Part 1: Data modelling

In the first part of the assignment you will focus on creating a data model based on raw event data that we've provided to you. The data is generated in the form of JSON events, and is ingested in raw form into our data warehouse by our data engineers. At BUX we use Snowflake for all our data warehousing activities.

The data that you will be working with is generated by the Trading Team that supports our BUX Zero app. BUX Zero is a mobile app that allows users to buy and sell stocks. The Trading Team consists of backend developers who build and maintain the "broker" - systems that are responsible for handling all trading activities (i.e. the buy/sell orders of our clients). Orders can come in different forms (e.g. market/limit, and can have different states (e.g. created, filled, cancelled, rejected). When an order is executed by the broker, the trading team will create order partials, which are the actual transactions that happen in the outside world. One order can result in one or multiple order partials. The true price of the order will be known at the moment of execution.

To get you started we prepared an environment for you in Snowflake, which you can use for your assignment. **The login credentials can be found in the invite email for this assignment**.

Your environment will have two databases. The EXT_ANALYTICS_ENGINEER_RAW database is read-only, and is populated with the raw JSON event data from our Trading Team (see Table 1). The second database (EXT_ANALYTICS_ENGINEER_{SHORTNAME}) is your workspace. Here you can store all the modelled data, including temporary objects if needed.

Table 1: Overview of data in the EXT_ANALYTICS_ENGINEER_RAW database

Schema	**Table**	**Description**
INGESTED	BROKER_ACCOUNTS_DOMAINEVENTS	Contains BrokerAccountCreated events*, which are triggered when a new account has been created.
INGESTED	BROKER_MASTERDATA_DOMAINEVENTS	Contains event data about the instruments (=securities) that we offer in BUX Zero, and their states.
INGESTED	BROKER_ORDERS_DOMAINEVENTS	Contains event data about orders and order partial executions

** some of the properties in the BrokerAccountCreated events contain sensitive personal data like email, name, address, passport number, etc. These columns are encrypted by default, and should be ignored for this assignment.*

Part 2: Data visualization

In the second part of the assignment you will create a dashboard based on the data model that you created before. At BUX we use Tableau for all our reporting, so we expect that you will create your dashboard using Tableau Desktop, which can be downloaded here if needed. The dashboard should provide insight in the trading behaviour of our clients over time.

Business users would like to use your dashboard to answer questions like:

- What securities / market categories are traded the most?
- What kind of order types (market/limit/basic) are clients placing?
- What is the sentiment in terms of buying/selling?
- What is the total, average/median order volume per day?
- Who are the most engaged clients, and where do they come from?

Our business users would also like to be able to filter the data based on the following dimensions:

- Period
- Country (based on country of residence)
- EU vs US securities
- Latest order state

Deliverables

As a deliverable we expect a Git repository that will contain all the code that was used in Part 1 for the data modelling, and the Tableau report that was created in Part 2.

We believe a good analytics engineer should be able to do this assignment within a reasonable amount of time, so we expect results within two weeks after receiving the assignment invite.

Once the Git repo has been delivered, we will plan an evaluation session where you can present your solution. During this session you will have the ability to explain the design decisions that you made.

If you have questions about the assignment, or you run into problems, please reach out to me by sending an email to derek.willemsen@getbux.com.