DataScientest

CryptoBot with Binance

Relevant training course : Analytics Engineer

Level of difficulty : 9/10

Description of the project :
Nowadays, the world of cryptocurrencies is starting to take an important place and is growing. These are simply quite volatile and unstable financial markets based on Blockchain technology.

Step	Description	Objectives	Modules / Masterclasses / Templates	Conditions for project validation
1	Unstructured Data Exploration	Collect two types of data through the Binance API. Thanks to this API, we can retrieve information on the prices of different markets (BTC-USDT, BTC-ETH, etc.). The goal will be to create a generic data retrieval function in order to have data from any market.	103 - Data API Fundamentals	An explanatory file of the processing and the different data accessible (doc / pdf) An example of data collected.
2	Data Modeling	Design a database schema that follows third normal form (3NF) or **applies the principles of denormalization**.	145 - Fundamentals of Data Integration 142 - SQL	Data stored in a relational database (PostgreSQL/MySQL) **UML Diagram** A SQL query file to show that the database is functional
3	Construction of the ETL/ELT pipeline	Create ETL/ELT pipeline to populate the database : - **Extraction** : Load data from original sources - **Chargement** : Denormalize (star schema) and load data into target system (SQL Database) - **Transformation** : Fusion, Encodage	Snowflake BigQuery Snaplogic 144 - ETL with pyspark	DataWarehouse Demo Star Schema

		Run the pipeline to integrate as much data as possible from the original dataset		
4	Final Dashboard	Creation of the Power BI Dashboard (or other) and preparation of the defense.	181 - Power BI 182 - Tableau 185 - Looker Studio	Defense Report

Bibliographie :

https://www.binance.com/fr/binance-api
https://binance-docs.github.io/apidocs/spot/en/#change-log
https://docs.binance.us/?python#introduction