Project Documentation: Automated Cryptocurrency Data Pipeline

Project Overview

This project implements an automated end-to-end data pipeline for cryptocurrency coins. It consists of two independent scripts:

1. ETL Script: This script extracts cryptocurrency data, cleans and transforms columns, and loads the processed data into a PostgreSQL mini data warehouse.

2. Visualisation Script: This script retrieves updated data from the central database and generates a scatter plot, which is saved to a folder that saves all previously generated graphs.

A 10-second delay between the two scripts ensures smooth execution. The process is designed to run automatically every hour for 12 hours.

Testing Procedure

The pipeline was tested to verify automation and reliability. The first test was executed manually at 17:12, and then tests were executed automatically every hour.

- The first test, conducted manually, was successful.

- The second test, automated, ran at 18:05 and was successful.

- The third test, automated, ran at 19:05 and was successful.

- The fourth test, automated, ran at 20:05 and was successful.

- The fifth test, automated, ran at 21:05 and was successful.

- The sixth test, automated, ran at 22:05 and was successful.

- The seventh test, automated, ran at 23:05 and was successful.

- The eighth test, automated, ran at 00:05 and was successful.

- The ninth test, automated, ran at 01:05 and was successful.

- The tenth test, automated, ran at 02:05 and was successful.

- The eleventh test, automated, ran at 03:05 and was successful.

- The twelfth and final test, automated, ran at 04:05 and was successful.

All 12 tests completed successfully.