

Project Brief

Market ETL Data Project

A solo-built data engineering pipeline that ingests daily market price data, transforms it into analytics-ready datasets, and demonstrates its usability through a lightweight analytical dashboard.

Project Objective

To design and implement a reliable daily ETL pipeline that transforms raw market price data into high-quality analytical tables, and to validate these outputs via a simple read-only dashboard.

High-Level Functionality

- Daily ingestion of market price data from public APIs
- Clear separation of raw, cleaned, and analytics datasets
- Schema enforcement and basic data quality checks
- Calculation of derived financial metrics (e.g. returns, rolling volatility)
- Streamlit dashboard that visualises the data

Data Sources

- Yahoo Finance API
- CoinGecko

Tech Stack

- Python – requests, pandas
- SQL
- Local Database (DuckDB or PSQL)
- Streamlit

Delivery Constraints

- Designed to be completed within a short, fixed timeframe
- Uses only free and publicly available data sources
- Optimised for clarity and reliability rather than production scale