Project Description — One-Page Summary

The Crypto Exchange Analytics Platform is a web-based application developed by three students — **Faik, Erjon and Jasin** — for their Software Design and Architecture course. Its goal is to collect, process and analyse historical cryptocurrency market data using a pipe-and-filter architecture. The platform retrieves daily price and volume data for the top 1,000 active cryptocurrencies, stores the processed data in a database and provides analytical tools such as charts, statistics and comparison options.

Cryptocurrency markets are highly fragmented, with data spread across many exchanges and trading pairs. Researchers and investors often lack a reliable, long-term and unified source of historical OHLCV data. Existing tools may only cover recent years or a limited set of coins. This project addresses that gap by automatically gathering, cleaning and harmonising data from multiple sources into a consistent, accessible format.

The main data providers used are the CoinGecko API, the CoinMarketCap API and the CCXT library. CoinGecko offers extensive coverage of thousands of coins with a generous free tier, making it suitable as a primary source. CoinMarketCap provides additional historical coverage and is used for verification or missing data. CCXT enables direct access to exchanges like Binance, Kraken and Coinbase, which is useful when aggregator APIs have data-length limitations. These sources were chosen for their coverage, historical depth and free access.

The project uses Python 3.11 due to its strong data processing ecosystem (pandas, numpy, ccxt, async libraries). Flask will serve as the initial web framework, with the possibility of switching to FastAPI for improved async performance. SQLite will be used during prototyping, while PostgreSQL will support scalability. All development and documentation will be maintained in a shared GitLab repository.

The data ingestion system is built using the pipe-and-filter architectural style. The first filter retrieves and validates the top 1,000 cryptocurrencies. The second filter determines the last available date in the database. The third filter downloads missing OHLCV and market metrics, normalises the data and stores it efficiently. Multiple pipelines can run in parallel for improved scalability and reuse.

By the end of the semester, the project will produce a clean and complete 10-year dataset for 1 000 top cryptocurrencies, along with a web interface that allows users to visualise trends, compare coins and export data. The solution offers transparency, wide coverage and useful analytical tools while demonstrating real-world application of

software architecture principles. Its modular design allows for easy future extensions, such as sentiment analysis, on-chain metrics or machine-learning-based predictions.