Group Project Proposal: Team 1

Ariel Hsieh, Jiaqi Yang, Neha Mrutyunjaya Bhadragoudar, Niranjan Rao Saraf Srinivas Rao

End to End Data Analytics Implementation

1. Abstract:

In this project, we aim to build a comprehensive data analytics solution using both historical and real-time Bitcoin price data to generate actionable insights for market analysis and trend prediction. We will integrate archived Bitcoin price data with real-time updates using CoinAPI to maintain a continuous flow of data. The system will employ Snowflake as the data warehouse for storage and transformation, Apache Airflow for scheduling and managing data pipelines, and dbt for ELT processes. Finally, we will visualize the insights, including real-time price trends, historical patterns, and predictive analysis on a dynamic dashboard using Preset.io. This project demonstrates a real-world application of modern data engineering techniques and provides a data-driven approach for decision-making in cryptocurrency trading.

---

2. Dataset Links and Short Description:

Historical Dataset:

- **Source**: CoinAPI.io
- **Description**: Historical Bitcoin price data including open, close, high, low, and volume metrics. This data will be pulled as a CSV or JSON file and stored in a cloud-based storage service like AWS S3 or Google Cloud Storage for integration with Snowflake.

Real-Time Dataset:

- **Source**: CoinAPI.io REST API

    Description: Real-time Bitcoin price data updated on a daily or hourly basis. The data will include current price, volume, and other related metrics. We will leverage the REST API provided by CoinAPI to stream or pull real-time data directly into Snowflake for timely analysis.

---

3. Reference Links:

- CoinAPI Documentation - https://docs.coinapi.io/
- Snowflake Documentation - https://docs.snowflake.com/
- Apache Airflow Documentation - https://airflow.apache.org/docs/
- dbt Documentation - https://docs.getdbt.com/
- Preset.io Documentation - https://docs.preset.io/