Cryptocurrency Volatility Analyzer - Detailed Project Report

# 1. Executive Summary

The Cryptocurrency Volatility Analyzer is a data analytics project built to understand, visualize, and classify volatility trends in the cryptocurrency market using historical price data. The project covers end-to-end data processing — from collection and storage to visualization and machine learning — for Bitcoin and Ethereum.

# 2. Data Collection & Storage

Data is collected using the api of Yahoo Finance with the help of Python libraries, which retrieves OHLCV data for Bitcoin (BTC-USD) and Ethereum (ETH-USD) from Yahoo Finance. The data is stored in a MySQL database using SQLAlchemy. The database is named `crypto\_db` and has separate tables for each cryptocurrency.

# 3. Feature Engineering

* Daily\_Return: Percentage change in closing price between two days.
* Log\_Return: Logarithmic return, preferred in finance for time-additive returns.
* Rolling\_Std\_7d: 7-day rolling standard deviation (short-term volatility).
* Rolling\_std\_30d: 30-day rolling standard deviation (long-term volatility).
* Rolling\_mean\_7d: 7-day rolling average of closing price.
* Bollinger upper/lower: Bollinger Bands indicating overbought/oversold conditions.
* Volume change: Daily percentage change in trading volume.
* Z - score: Z-score for identifying price outliers relative to 30-day mean.

# 4. Visualizations

* Bitcoin Daily Closing Price Chart

A graph showing the price of bitcoin

AI-generated content may be incorrect.

* Daily Return Line Chart

A graph showing a line of orange lines

AI-generated content may be incorrect.

* Rolling Volatility Comparison (7d vs 30d)

A graph showing a graph of a bitcoin rolling volatility

AI-generated content may be incorrect.

* Bollinger Bands

A graph showing a line

AI-generated content may be incorrect.

* Volume Change Over Time (Bar Chart)

A graph showing a graph of a bitcoin trading

AI-generated content may be incorrect.

* Z-score Over Time (Line Chart)

A graph showing a graph

AI-generated content may be incorrect.

# 5. Machine Learning Models

* Logistic Regression: Classifies high vs low volatility days using engineered features.

A screenshot of a graph

AI-generated content may be incorrect.

* Random Forest Classifier: Handles non-linearity and improves classification accuracy.

A screenshot of a graph

AI-generated content may be incorrect.

* KMeans Clustering: Unsupervised model to group similar volatility patterns.

A graph showing a number of small colored dots

AI-generated content may be incorrect.

* Linear Regression: Predicts future price movement based on current indicators.

A graph showing the results of a graph

AI-generated content may be incorrect.

* Isolation Forest: Detects price anomalies or outliers in time series data.

A graph showing a graph of data

AI-generated content may be incorrect.

# 6. Tools & Technologies

* Python 3 (Anaconda)
* Pandas, NumPy, Matplotlib, Scikit-learn, Yahoo Finance(API)
* MySQL with SQL Alchemy
* Jupyter Notebook for scripting