**Project Goal / Guiding Question**

**Can we observe meaningful patterns in Bitcoin’s on-chain OP\_RETURN messages around the time of major global events?**

This project investigates whether transaction output messages or volume patterns show correlation with events such as elections, economic crises, or regulatory announcements.

**Tools Used (plan – may change if unforeseen circumstances occur)**

* Python (Google Colab / VS Code)
* Google BigQuery
* pandas, matplotlib, seaborn (for analysis and visualization)
* CSV export and text parsing for deeper message analysis

**Milestones**

| **Milestone** |  |
| --- | --- |
| Set up BigQuery access & test query |  |
| Export outputs with OP\_RETURN messages |  |
| Join timestamps via transactions → blocks |  |
| Generate summary statistics |  |
| Compare against real-world events |  |
| Visualizations & findings |  |
| Final presentation / README polish |  |

**Dataset Selection Overview**

**Primary Dataset:**  
bigquery-public-data.crypto\_bitcoin.outputs  
**Supplementary Datasets (via join):**  
bigquery-public-data.crypto\_bitcoin.transactions  
bigquery-public-data.crypto\_bitcoin.blocks

This dataset provides a record of **Bitcoin transaction outputs**, including:

* Value transferred
* Output scripts (script\_asm)
* Potential embedded messages (OP\_RETURN)
* Timestamps (joined through transactions → blocks)

**Data Dictionary**

**Here’s a starter set for the most relevant fields:**

| **Variable Name** | **Type** | **Description** |
| --- | --- | --- |
| script\_asm | STRING | Output script in assembly format, may contain embedded text via OP\_RETURN |
| value | INTEGER | Value of the output in satoshis (1 BTC = 100,000,000 satoshis) |
| transaction\_hash | STRING | Unique identifier of the transaction |
| block\_hash | STRING | Unique identifier of the block the transaction is included in (joined from transactions) |
| block\_timestamp | TIMESTAMP | Date/time when the block was mined (joined from blocks) |