# IST 718 Final Project
## Cryptocurrency Price Prediction

# Database

**Asset**

| PK | Ticker |
|----|--------|
|    | Slug   |

**Youtube**

| PK | Ticker |
|----|--------|
| PK | Date |
|    | VideoCount |

**Twitter**

| PK | Ticker |
|----|--------|
| PK | Date |
|    | TweetCount |

**Google**

| PK | Ticker |
|----|--------|
| PK | Date |
|    | InterestLevel |

**Price**

| PK | Ticker |
|----|--------|
| PK | Date |
|    | High |
|    | Low |
|    | Close |
|    | Volume |

**CoinMarket**

| PK | Ticker |
|----|--------|
| PK | Date |
|    | CirculatingSupply |
|    | MaximumSupply |
|    | CMCRank |
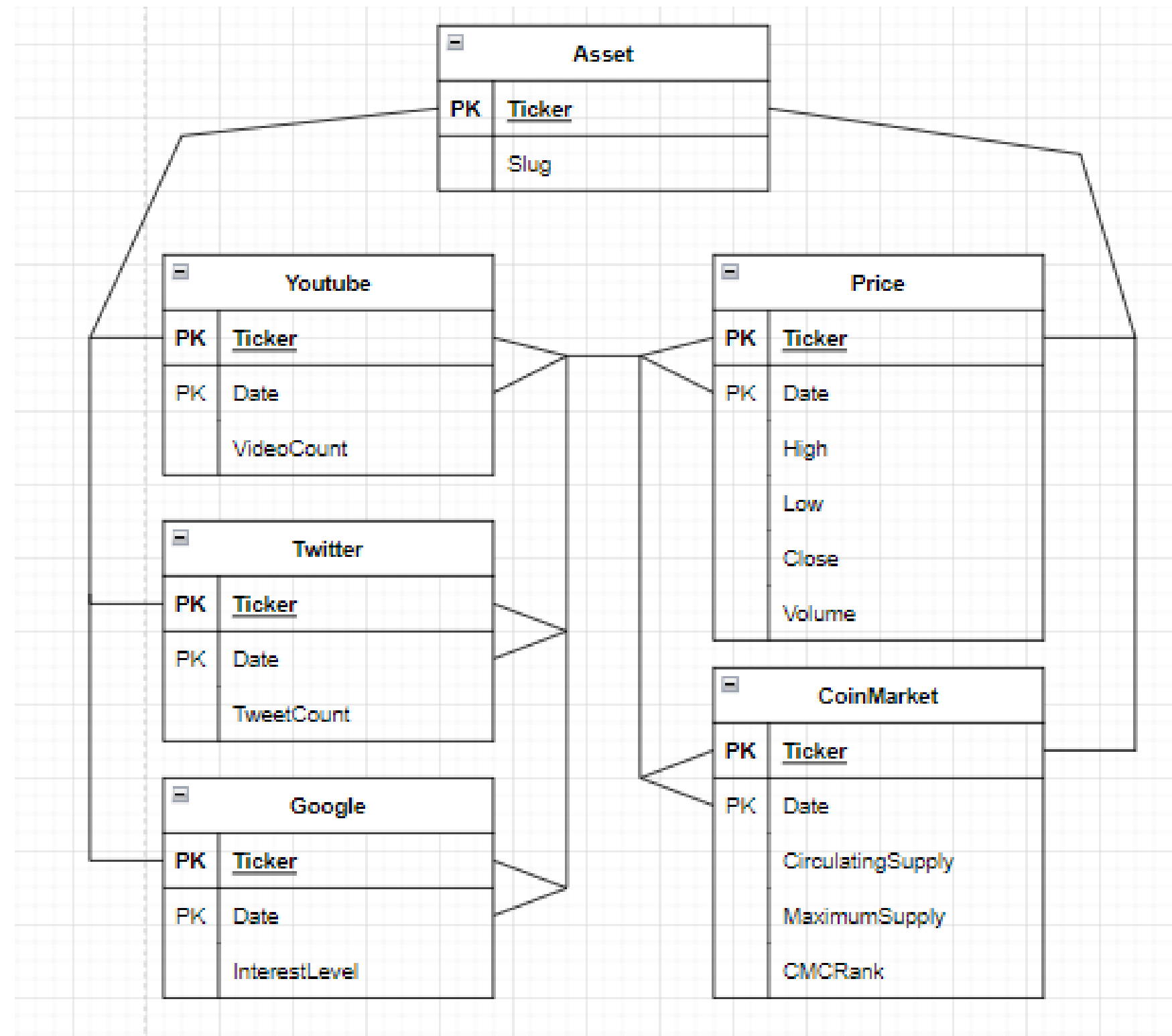
# Asset Table
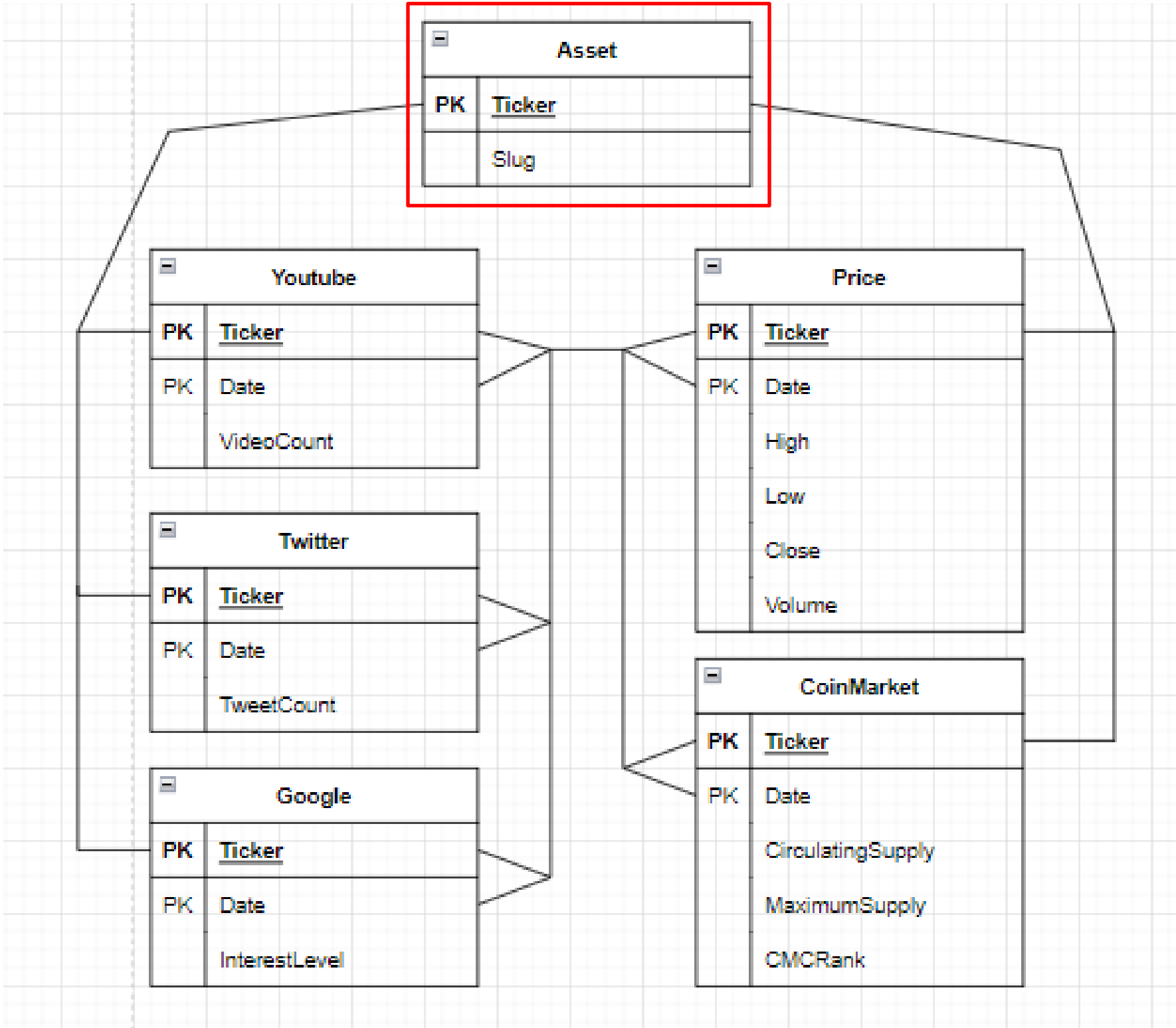


- Collected by web scraping Coinbase asset directory page
- This page has a list of all tradeable cryptocurrencies on Coinbase
- Currently about 158 cryptocurrencies that can be traded

### Top 10 Rows

| Ticker | Slug |
|--------|------|
| ZRX | 0x |
| 1INCH | 1inch |
| AAVE | Aave |
| ALCX | Alchemix |
| ACH | Alchemy P |
| AGLD | Adventure |
| ALGO | Algorand |
| AMP | Amp |
| FORTH | Amplefort |
| ANKR | Ankr |

# Price Table



- Collected through the yfinance Python package
- Historical price and volume data by day for all data available on Yahoo Finance

### Top 10 Rows

| Ticker | Date | High | Low | Close | Volume | Open |
|---|---|---|---|---|---|---|
| ZRX | 8/16/2017 | 0.280031 | 0.103962 | 0.224399 | 5232600 | 0.111725 |
| ZRX | 8/17/2017 | 0.238935 | 0.206735 | 0.206735 | 2752410 | 0.223022 |
| ZRX | 8/18/2017 | 0.35026 | 0.205558 | 0.293387 | 12793800 | 0.205558 |
| ZRX | 8/19/2017 | 0.543728 | 0.284357 | 0.478768 | 52677500 | 0.294578 |
| ZRX | 8/20/2017 | 0.475262 | 0.403409 | 0.424265 | 16016500 | 0.471296 |
| ZRX | 8/21/2017 | 0.44368 | 0.327742 | 0.351503 | 15013800 | 0.422273 |
| ZRX | 8/22/2017 | 0.375577 | 0.315831 | 0.364505 | 8045240 | 0.351547 |
| ZRX | 8/23/2017 | 0.410874 | 0.36401 | 0.386673 | 8663860 | 0.36401 |
| ZRX | 8/24/2017 | 0.392537 | 0.372134 | 0.38531 | 6227850 | 0.382705 |
| ZRX | 8/25/2017 | 0.394479 | 0.345851 | 0.361609 | 6146620 | 0.387431 |

# CoinMarket Table



- Collected through the CoinMarketCap API
- Daily snapshots of circulating supply, max supply, and coin market cap rank
- Cannot get historical data can only take the daily snapshots

Top 10 Rows

| Ticker | Date | CirculatingSupply | MaximumSupply | CMCRank |
|--------|------|-------------------|---------------|---------|
| 1INCH | 2/13/2022 | 415744664.5 | | 97 |
| AAVE | 2/13/2022 | 13514062.93 | 16000000 | 51 |
| ACH | 2/13/2022 | 4008728073 | 10000000000 | 275 |
| ADA | 2/13/2022 | 33592646022 | 45000000000 | 7 |
| AGLD | 2/13/2022 | 70170001 | | 455 |
| ALCX | 2/13/2022 | 1107093.861 | 1562945 | 283 |
| ALGO | 2/13/2022 | 6615850698 | 10000000000 | 28 |
| AMP | 2/13/2022 | 42227702186 | 92547638199 | 69 |
| ANKR | 2/13/2022 | 8162899378 | 10000000000 | 107 |
| API3 | 2/13/2022 | 36907810 | | 285 |
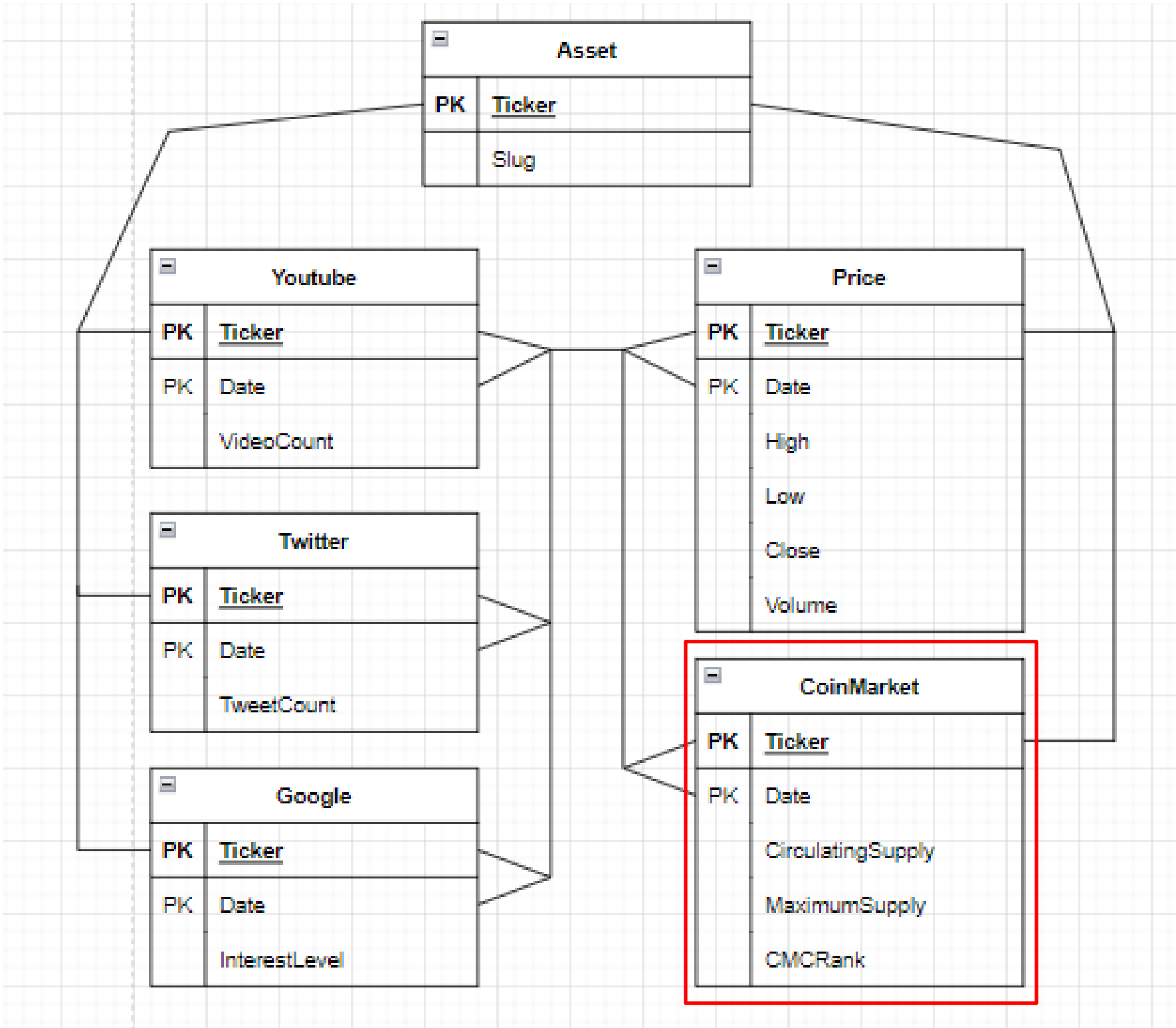
# Youtube Table
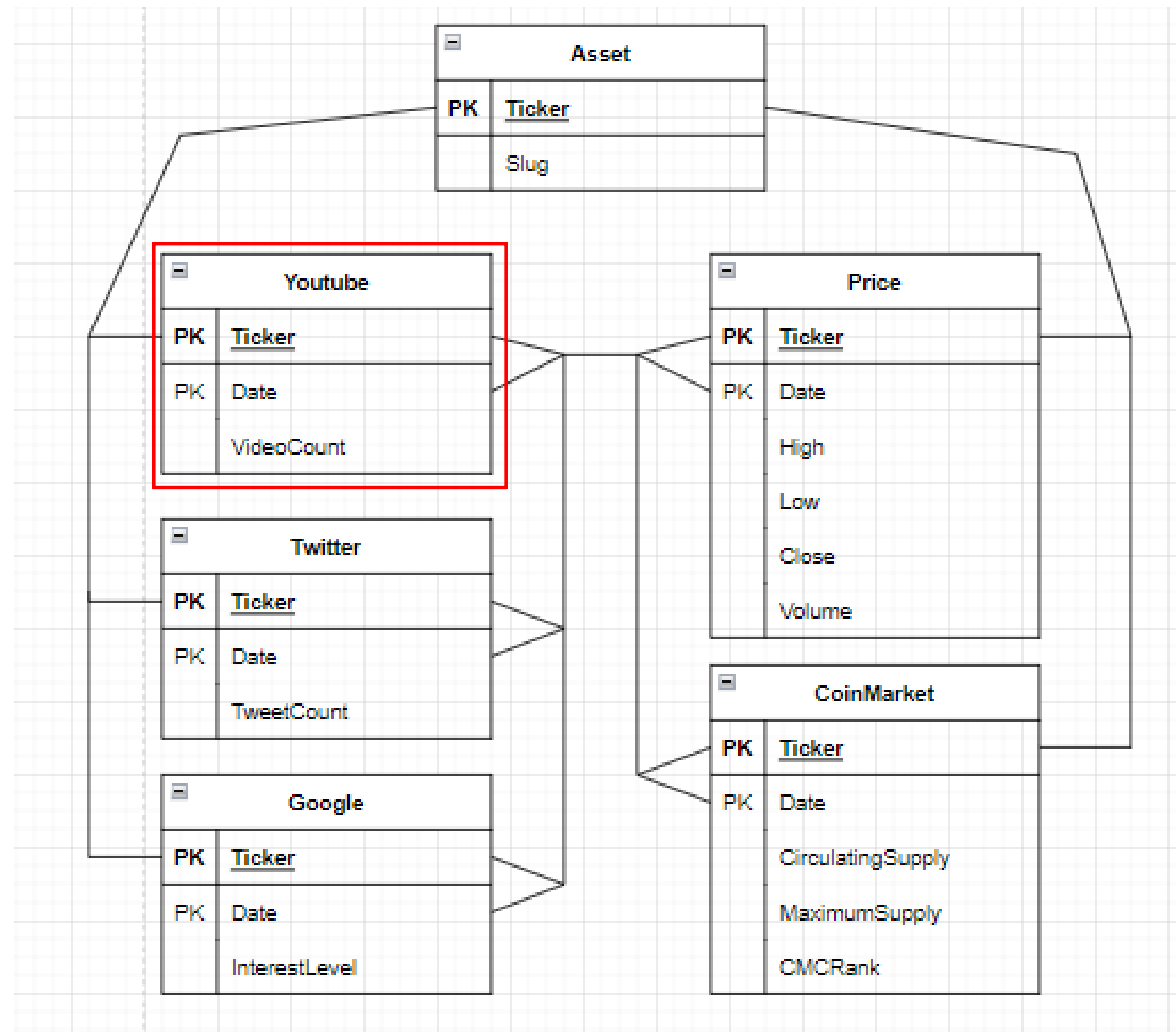


- Collected through Google Cloud / Youtube API
- Retrieve a count of published youtube videos for crypto slugs by day
- Capped at 1000 per day (100 for 10 API keys), but can get historical data
- Starting from the top down and building the history up overtime

Top 10 Rows

| Ticker | Date | VideoCount |
|--------|------|------------|
| ZRX | 2/13/2022 | 495 |
| ZEN | 2/13/2022 | 27 |
| ZEC | 2/13/2022 | 129 |
| YFII | 2/13/2022 | 9 |
| YFI | 2/13/2022 | 26 |
| XYO | 2/13/2022 | 64 |
| XTZ | 2/13/2022 | 1498 |
| XLM | 2/13/2022 | 71 |
| WLUNA | 2/13/2022 | 29 |
| WCFG | 2/13/2022 | 1 |

# Twitter Table
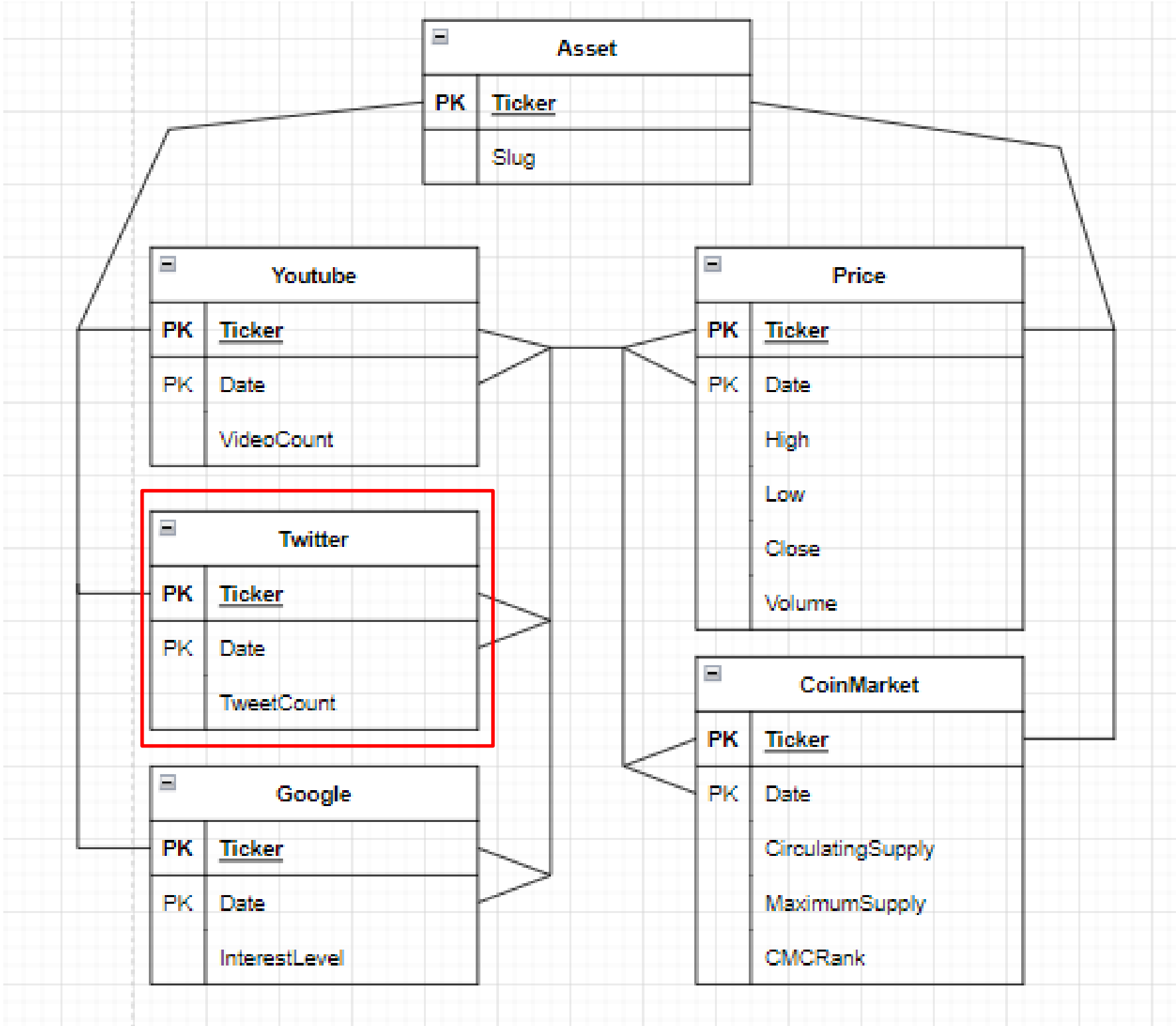


- Collected through Twitter API
- Retrieve a count of Tweets that contain hashtag of the ticker by day
- Cannot get historical data without elevated access. Snapshot taken every day

Top 10 Rows

| Ticker | Date | TweetCou |
|--------|------|----------|
| ZRX | 2/13/2022 | 68 |
| 1INCH | 2/13/2022 | 349 |
| AAVE | 2/13/2022 | 355 |
| ALCX | 2/13/2022 | 16 |
| ACH | 2/13/2022 | 148 |
| AGLD | 2/13/2022 | 17 |
| ALGO | 2/13/2022 | 4502 |
| AMP | 2/13/2022 | 675 |
| FORTH | 2/13/2022 | 66 |
| ANKR | 2/13/2022 | 143 |

# Google Table
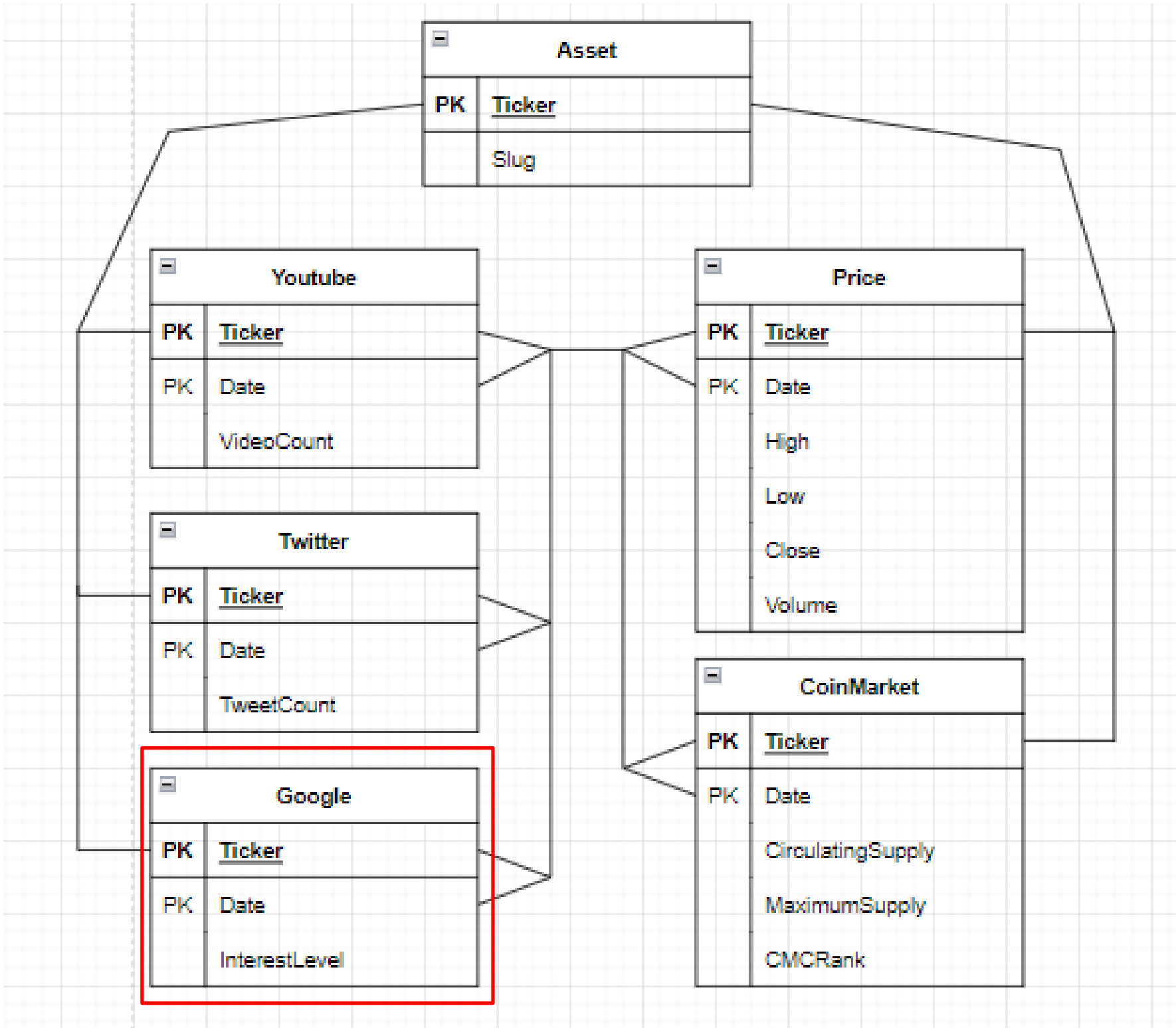


- Collected through pytrends python package
- Retrieve data from Google Trends about cryptocurrency searches by day
- Might need to use some extrapolation methods due to data availability

Top 10 Rows

In progress

# Additional Data?

- Population



- Dow jones / Nasdaq



- Sentiment



- Interest rate

# Project Objective

- Finish collecting data

- Exploratory data analysis / data cleaning

- Create a forecast for each cryptocurrency

- Invest 1000$ based on my predictions

- Measure the forecast and investment performance

- Make improvements