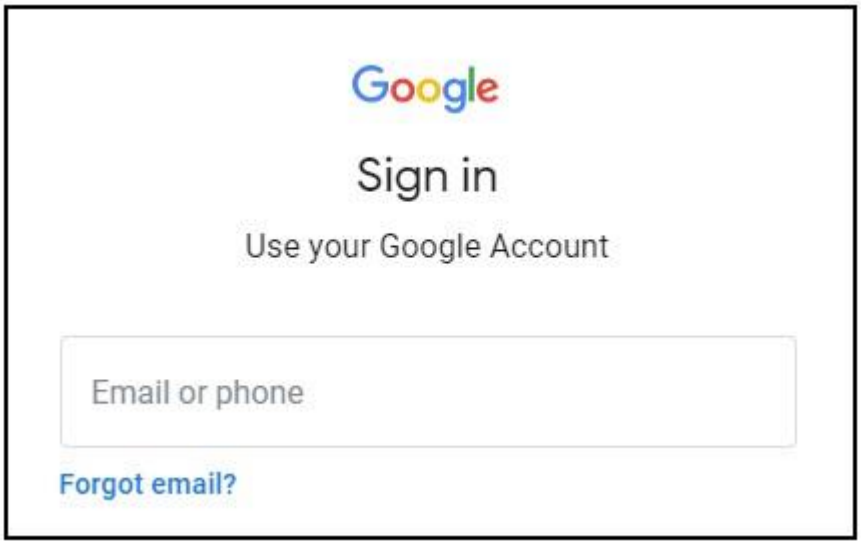
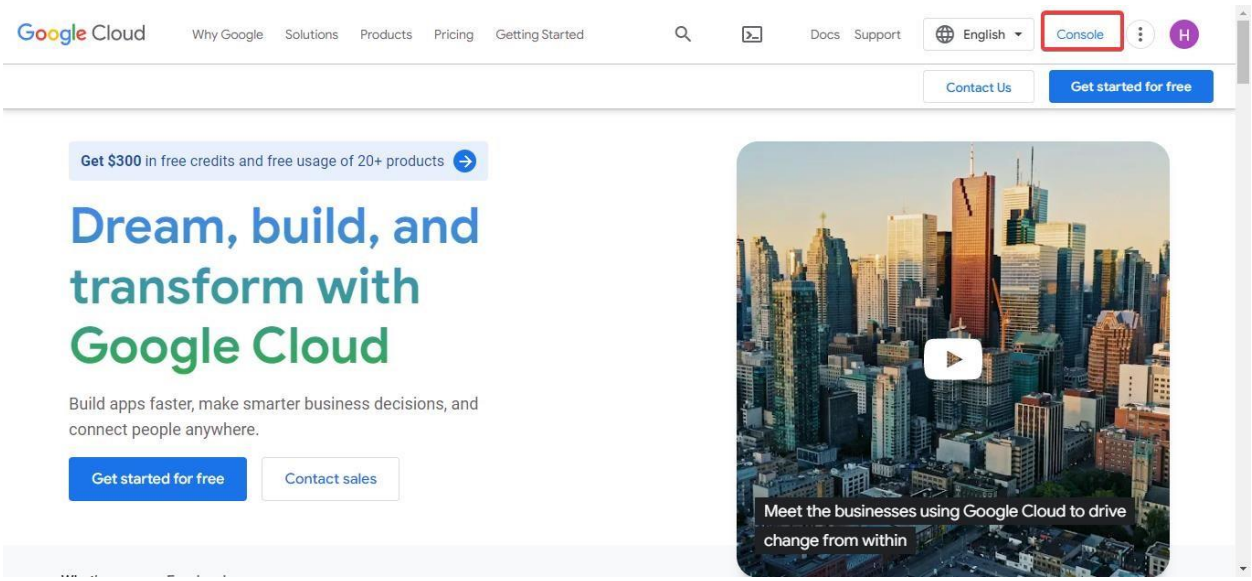
Exploring the Ethereum dataset Available in BigQuery

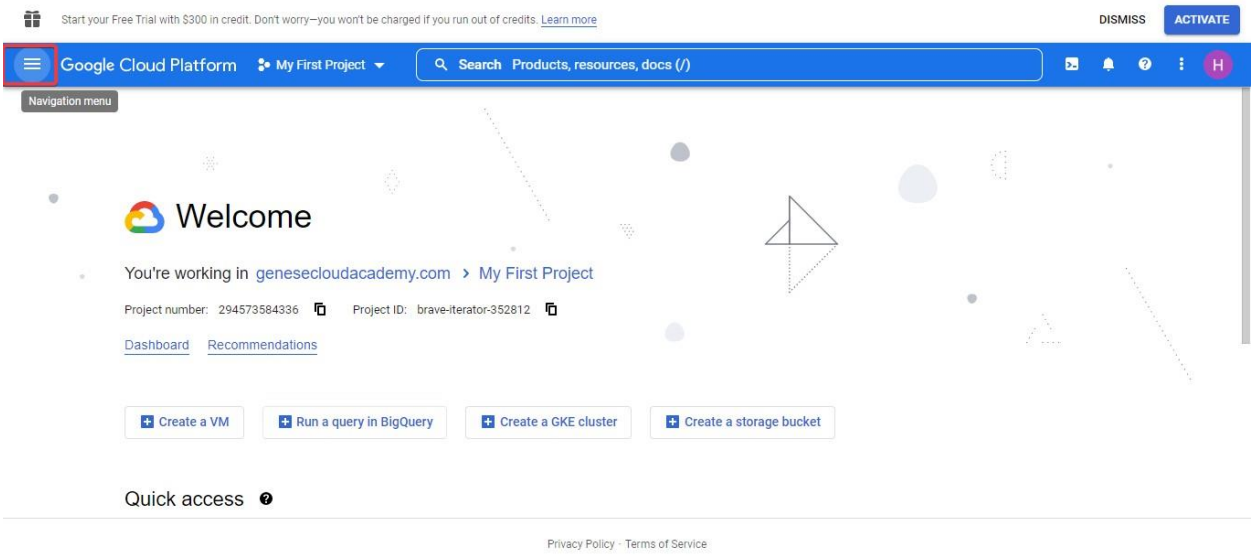
- **Log In** to google cloud with your **Google Account**

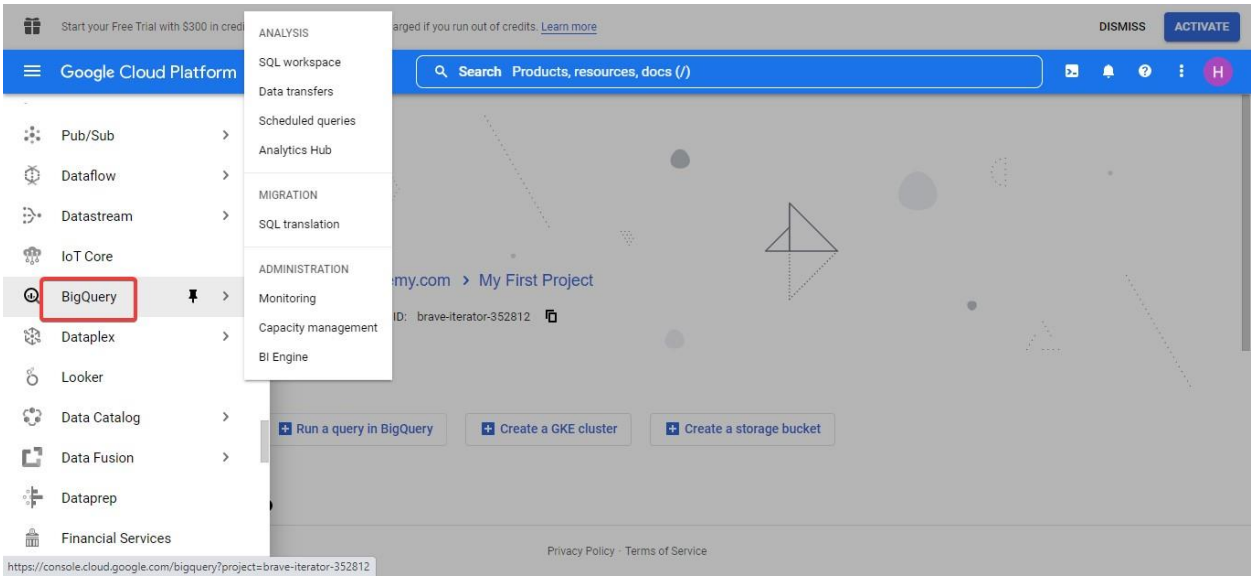


- Go to **Console.**

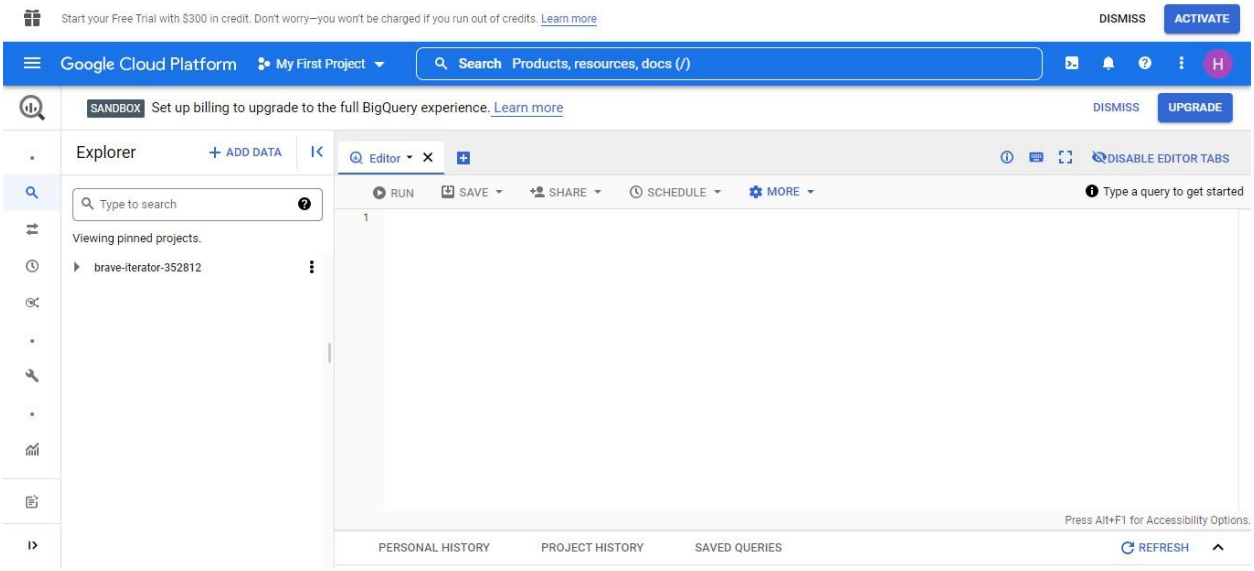


- Open **Navigation Menu > BigQuery**.

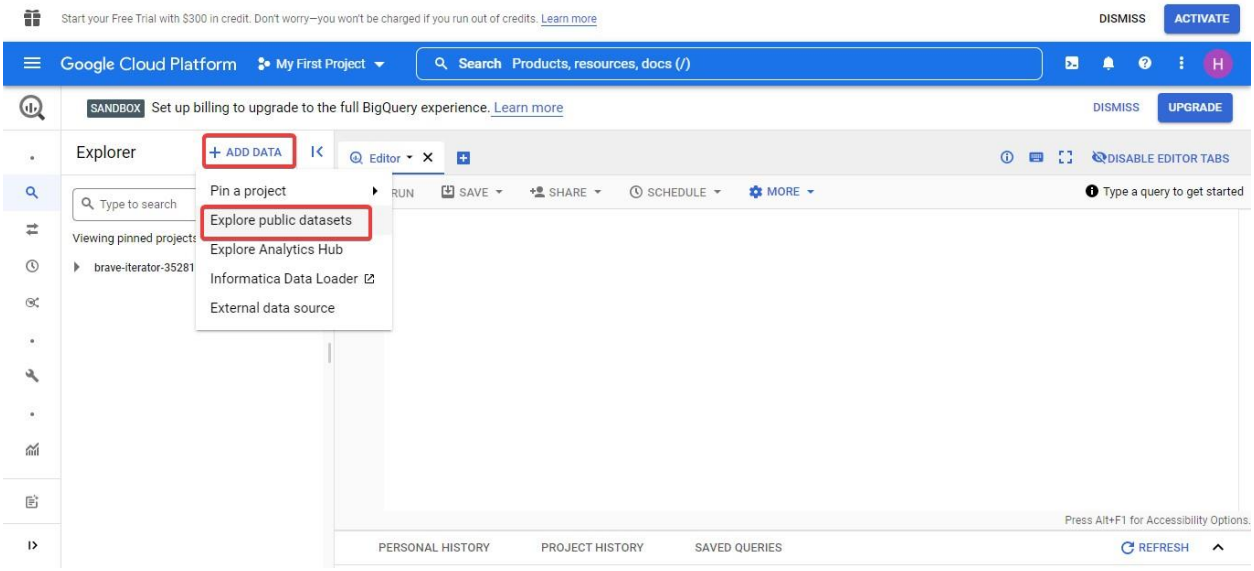




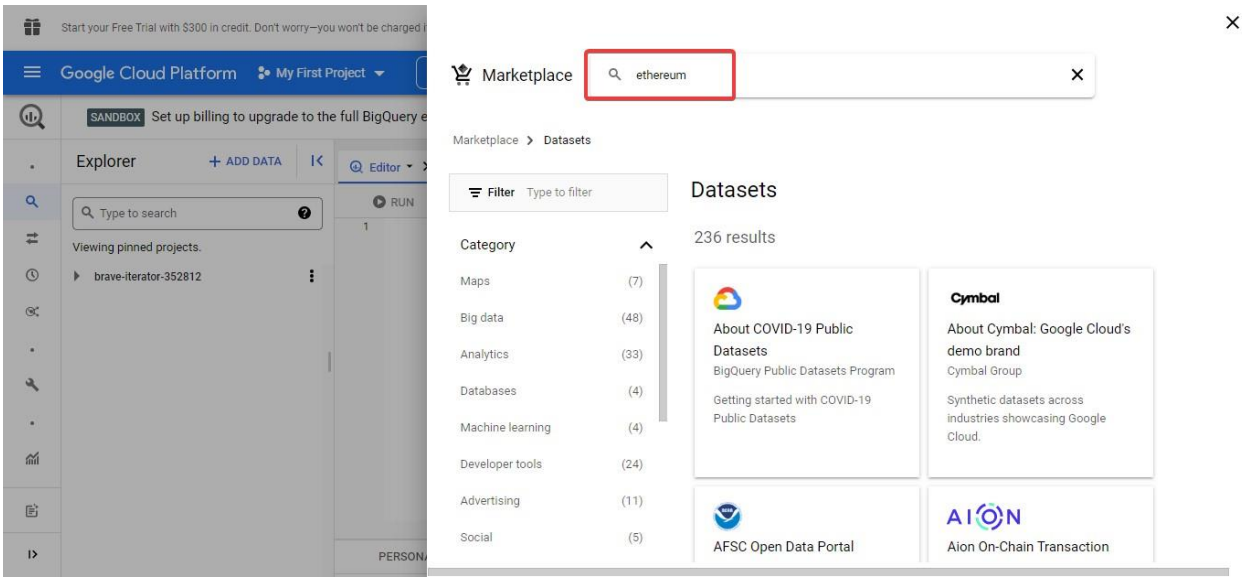
- The **Welcome to BigQuery in the Cloud Console** dailog box opens and click **DONE**.



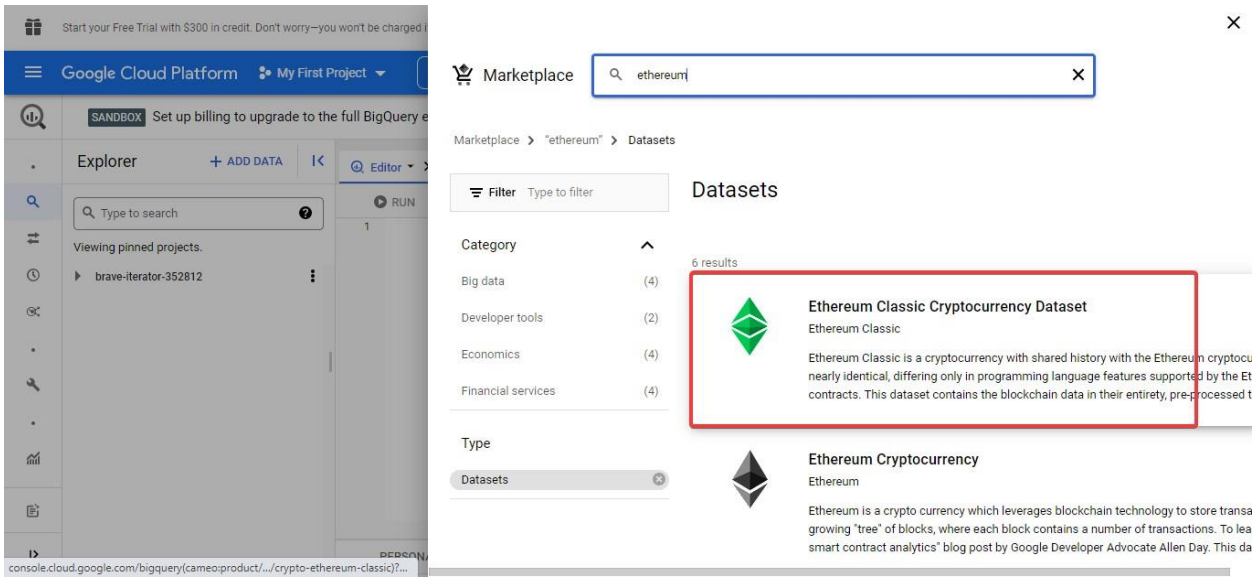
- Click + **ADD DATA > Explore public datasets** .



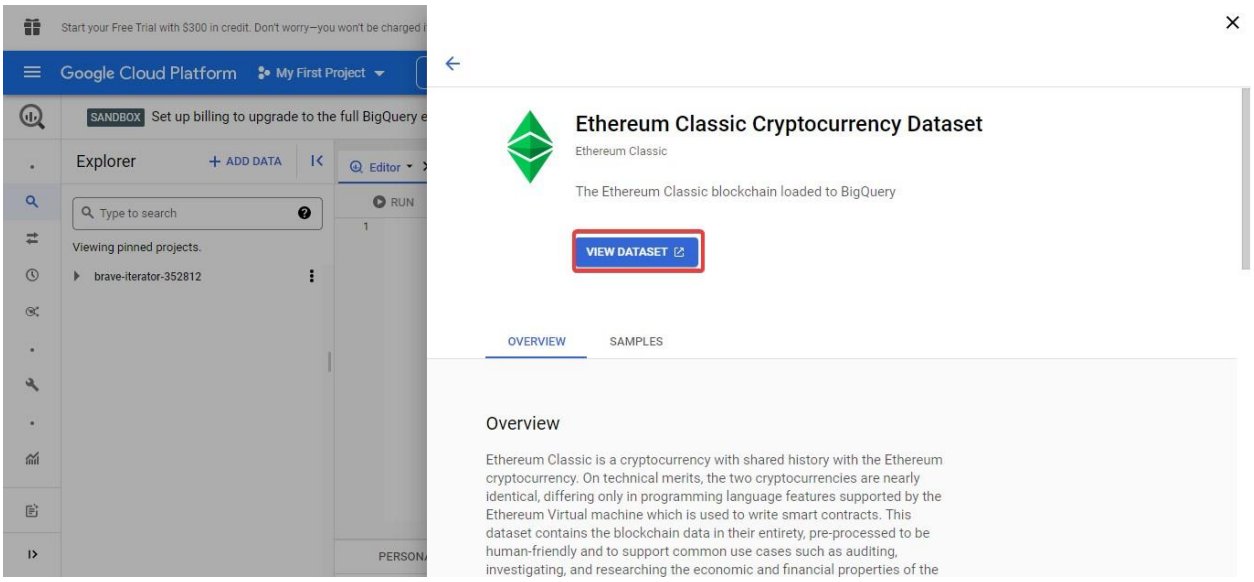
- In **Search for solutions**, type ethereum and press **enter**.



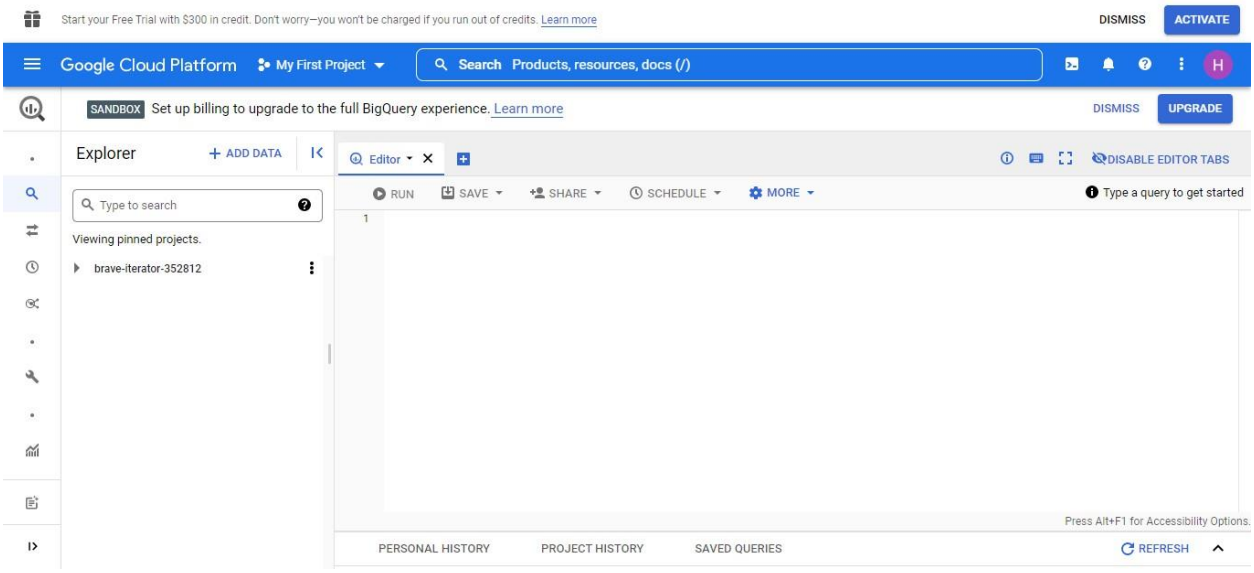
- Click Ethereum **Classic Cryptocurrency Dataset**



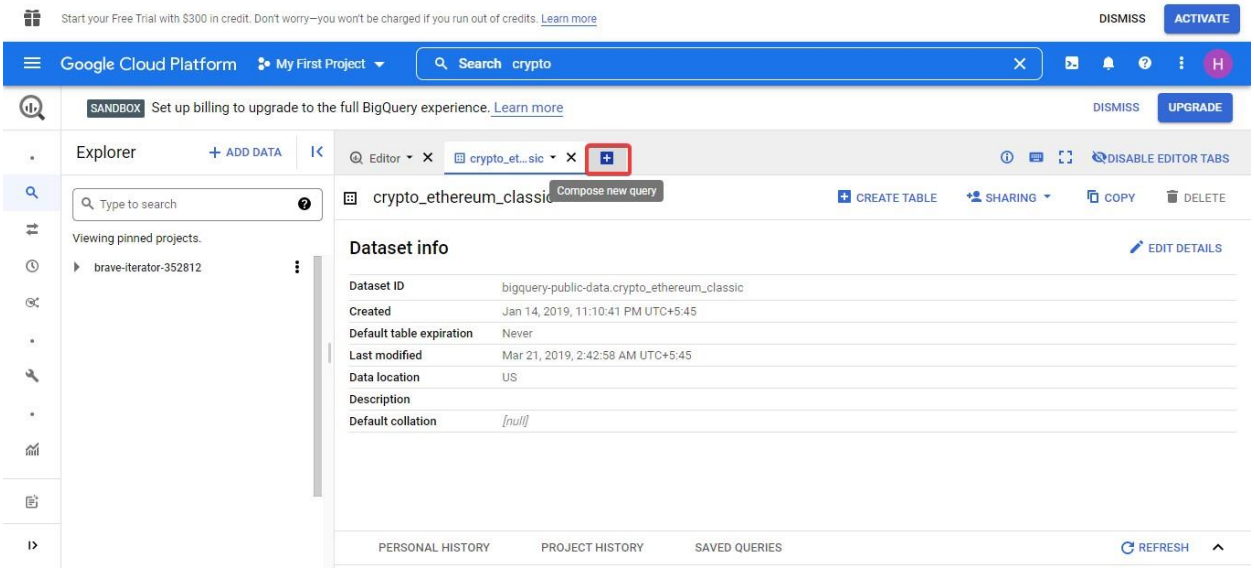
- Click **VIEW DATASET**



A new tab will open with BigQuery, and you should be on the bigquery-publicdata.crypto_ethereum_classic dataset.

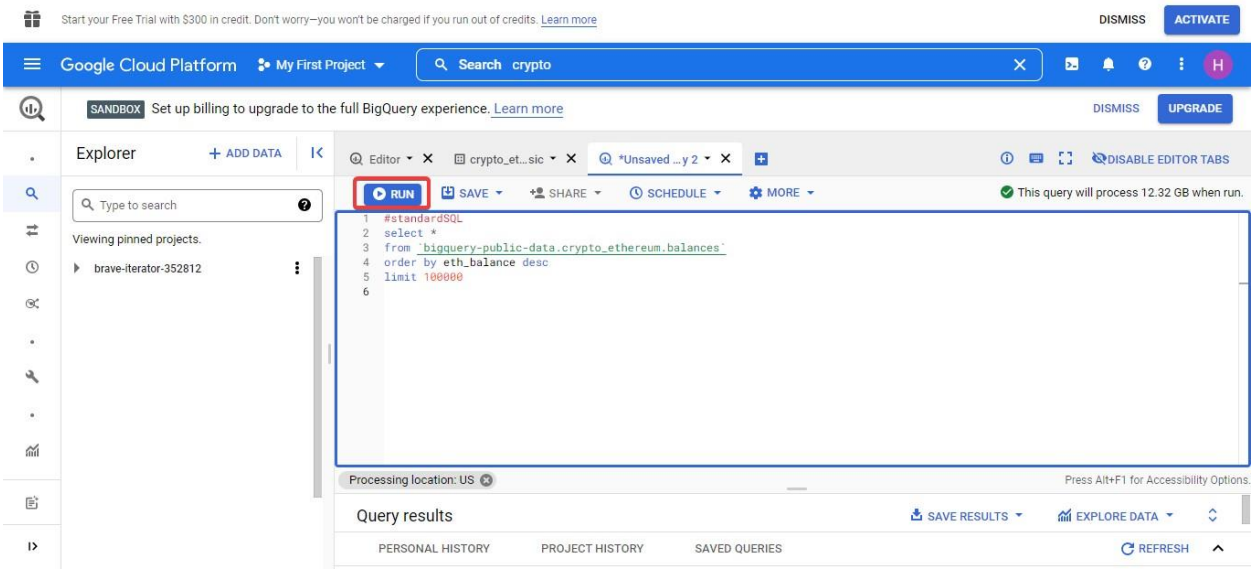


- Click + **Compose new query**.

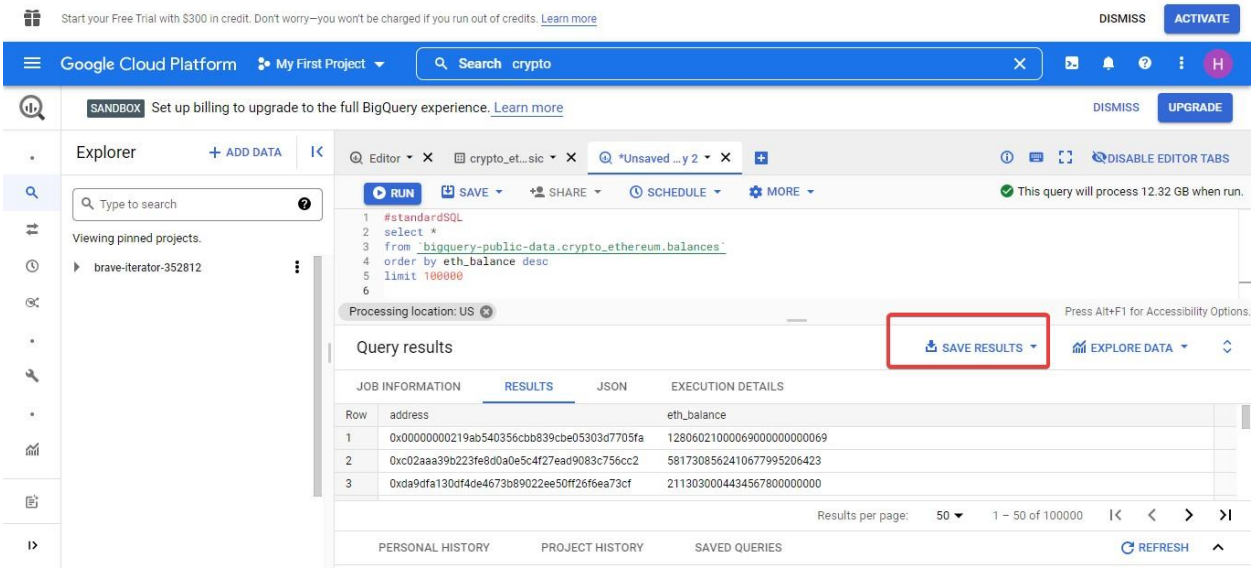


Copy and paste this query into the query window and then press Run.

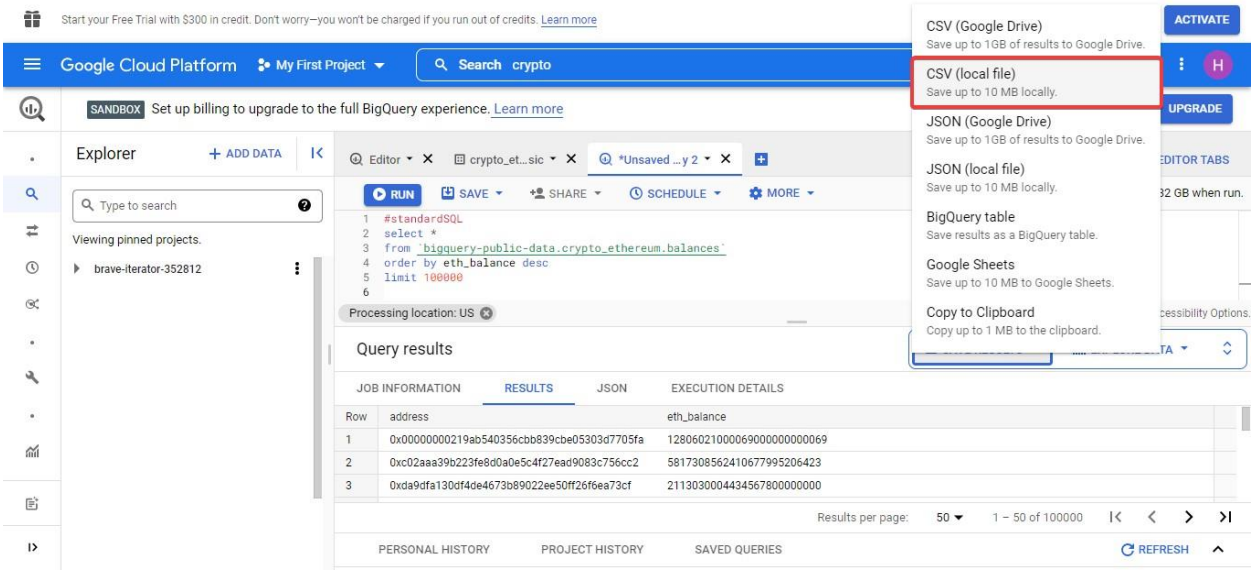
```
#standardSQL select * from `bigquery-public-data.crypto_ethereum.balances` order by eth_balance desc limit 100000
```



- Click **SAVE RESULTS**.



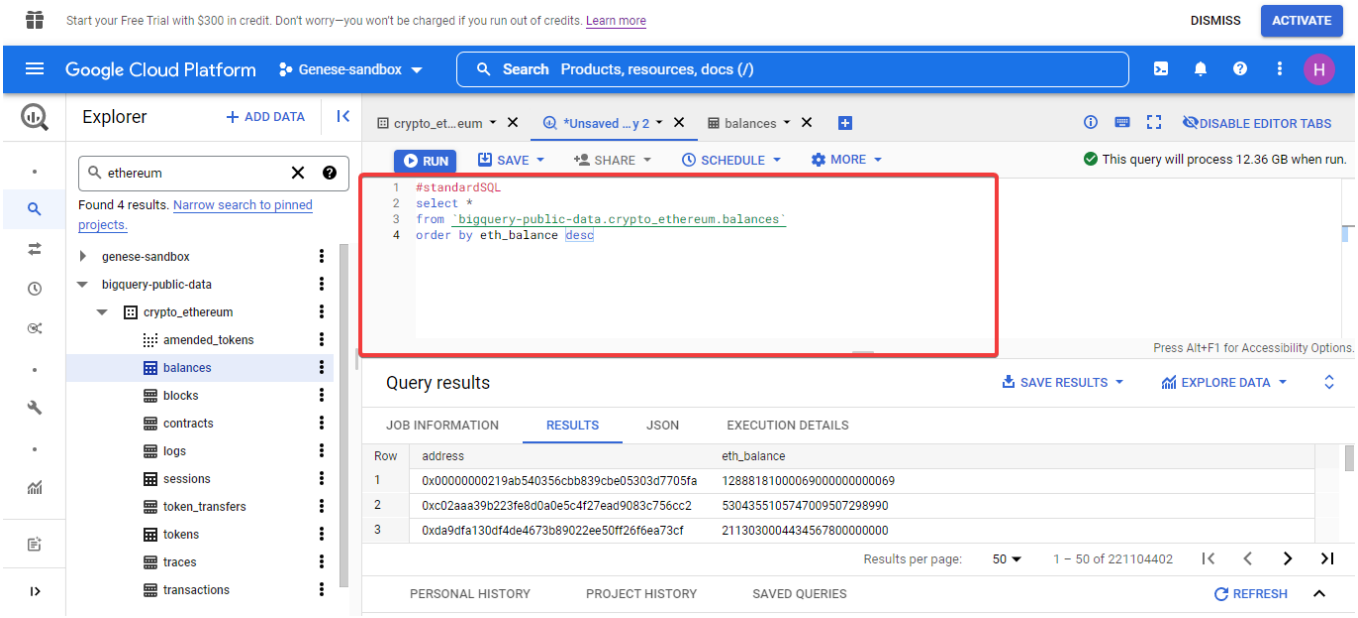
- Click CSV **(local file)** for download.



For unlimited query result,

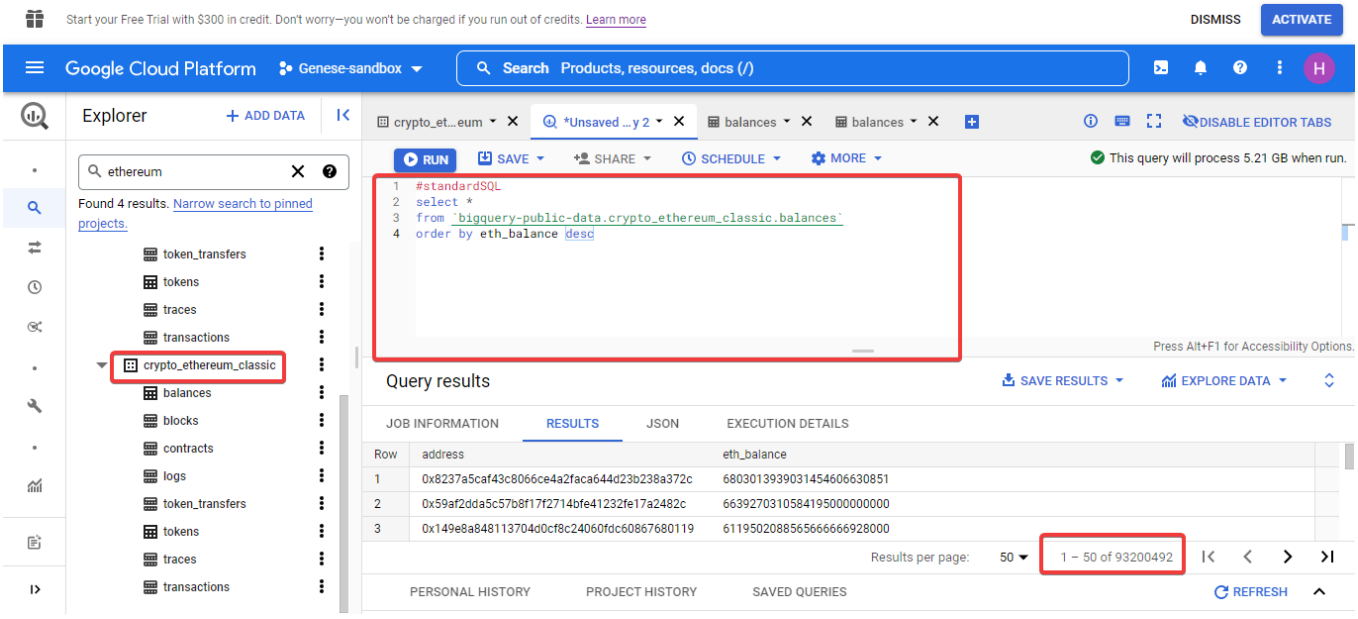
bigquery-public-data.crypto_ethereum.balances

```
#standardSQL
select *
from `bigquery-public-data.crypto_ethereum.balances`
order by eth_balance desc
```



bigquery-public-data.crypto_ethereum_classic.balances

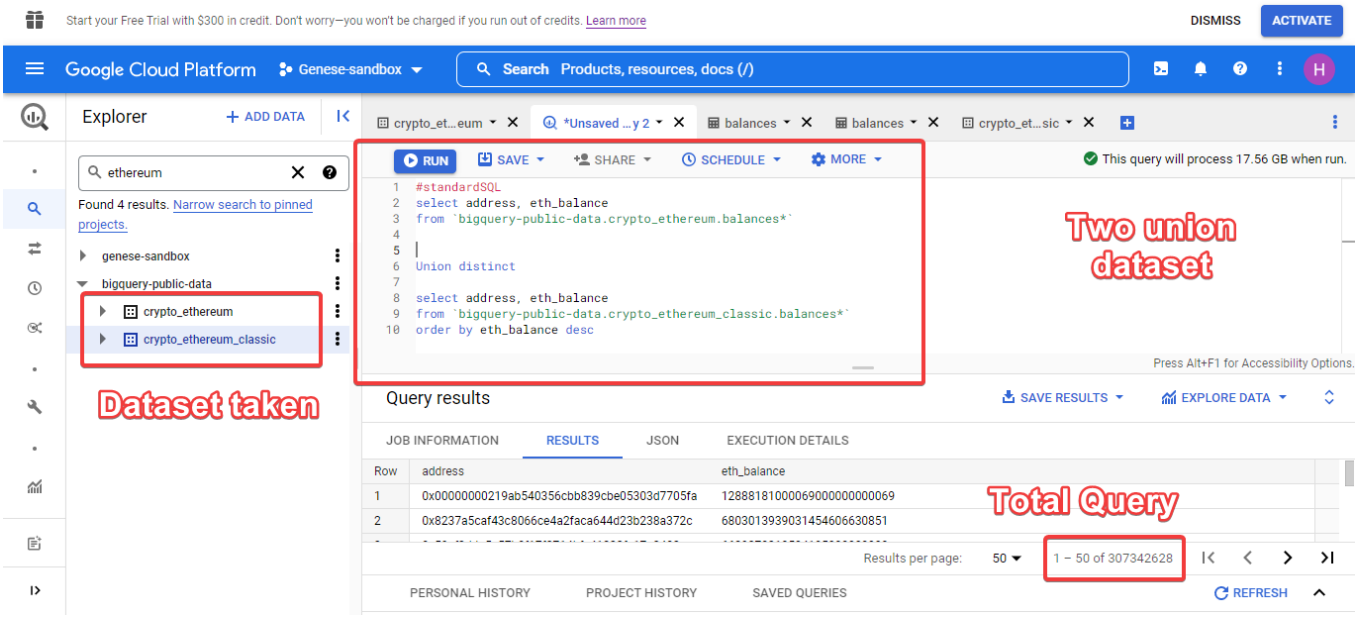
```
#standardSQL
select *
from `bigquery-public-data.crypto_ethereum_classic.balances`
order by eth_balance desc
```



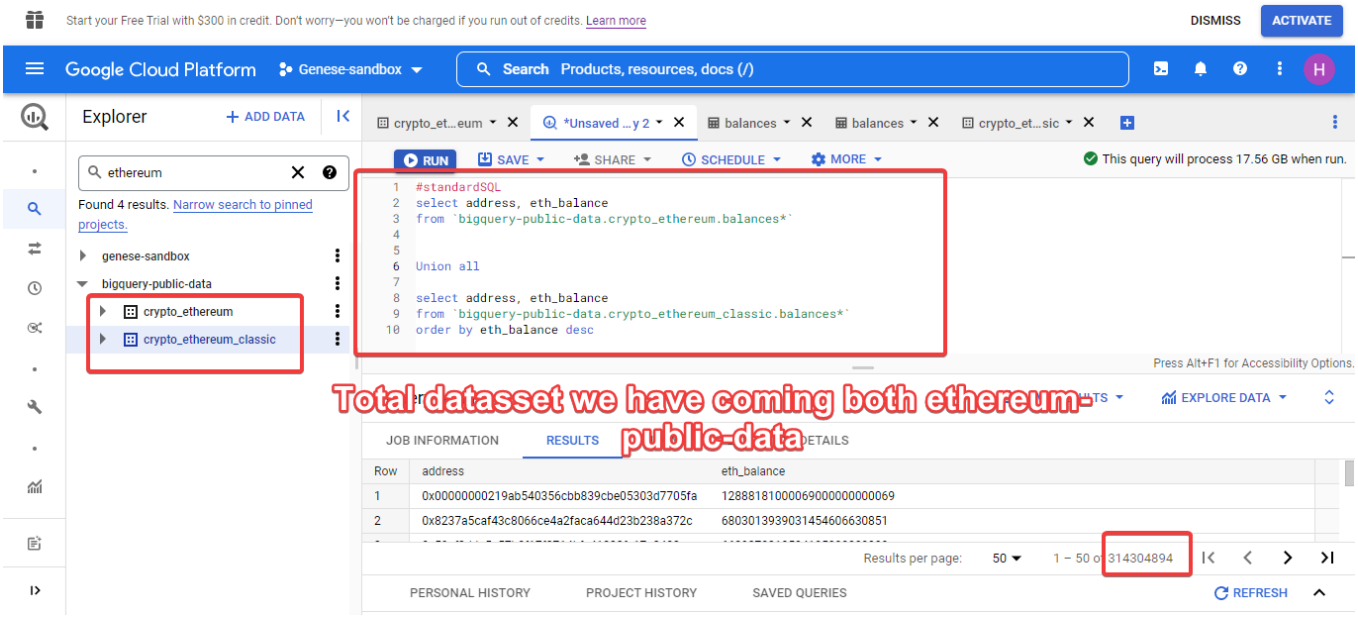
In order to prevent overlapping **bigquery-public-data.crypto_ethereum_classic** and **bigquery-public-data.crypto_ethereum.balances** in single output

Run the following code

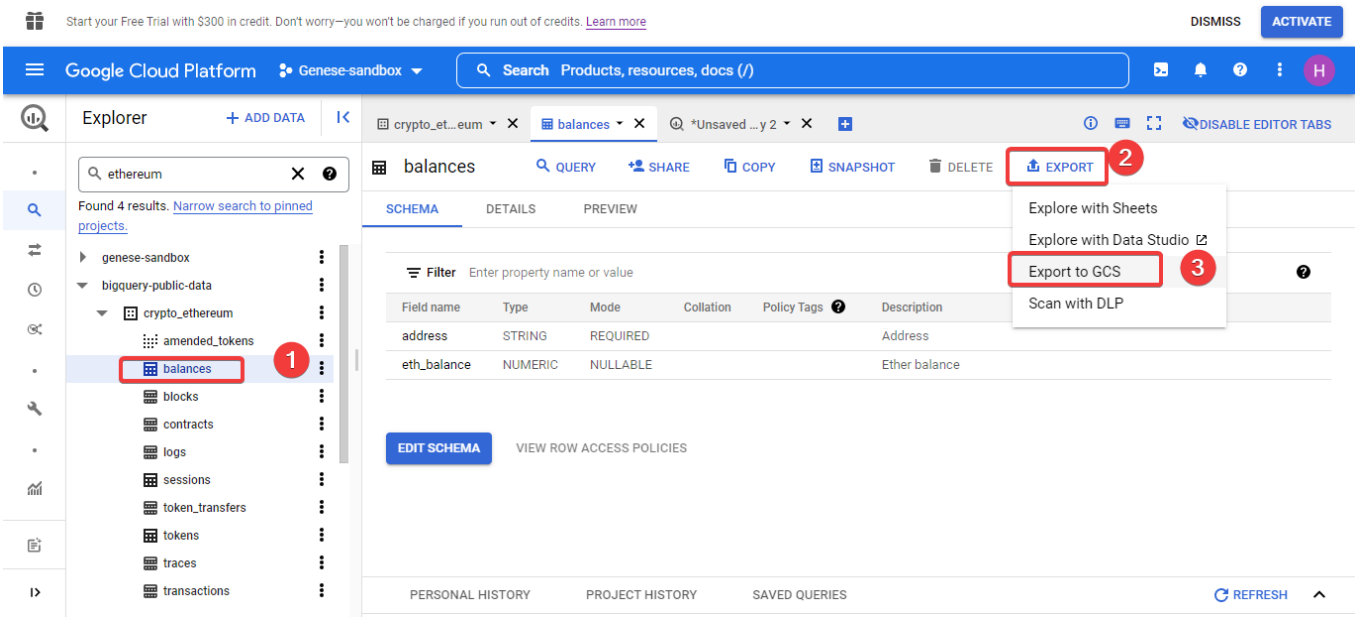
```
#standardSQL
select address, eth_balance
from `bigquery-public-data.crypto_ethereum.balances*`

Union distinct

select address, eth_balance
from `bigquery-public-data.crypto_ethereum_classic.balances*`
order by eth_balance desc
```



In order to overlapping, **bigquery-public-data.crypto_ethereum_classic** and **bigquery-public-data.crypto_ethereum.balances** in single output

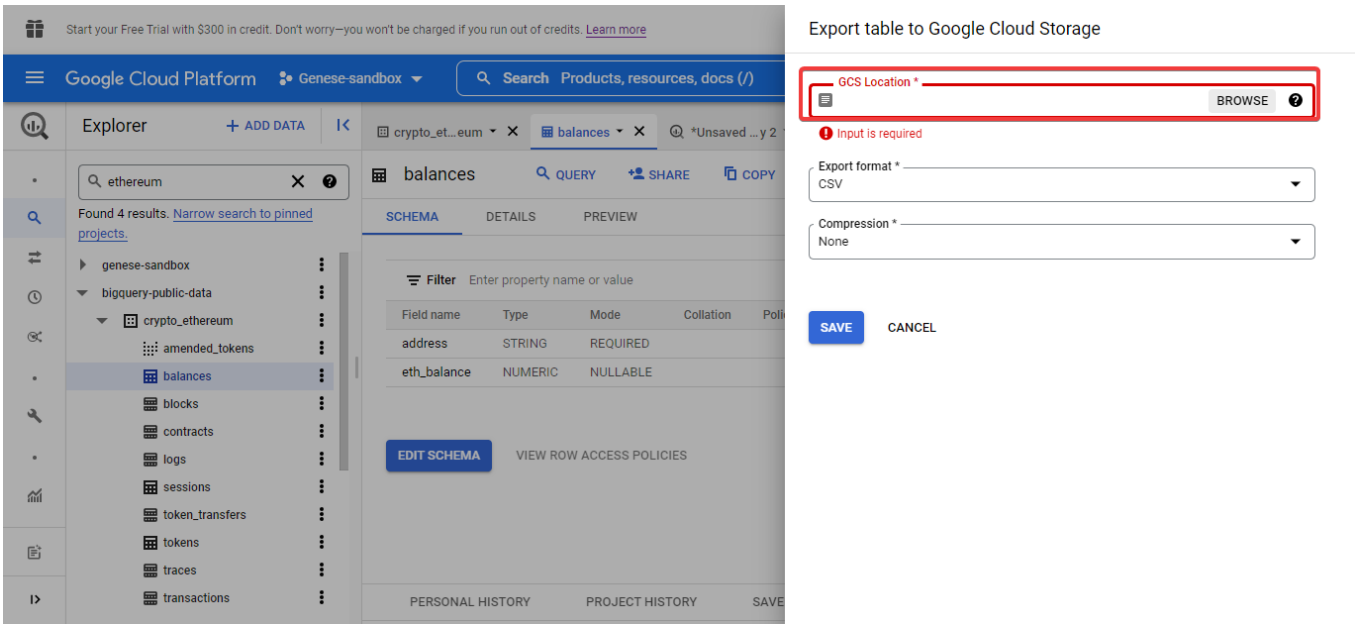


- Click **balances** on crypto_ethereum bigquery-public-data
- Click **Export**
- Click Export to GCS

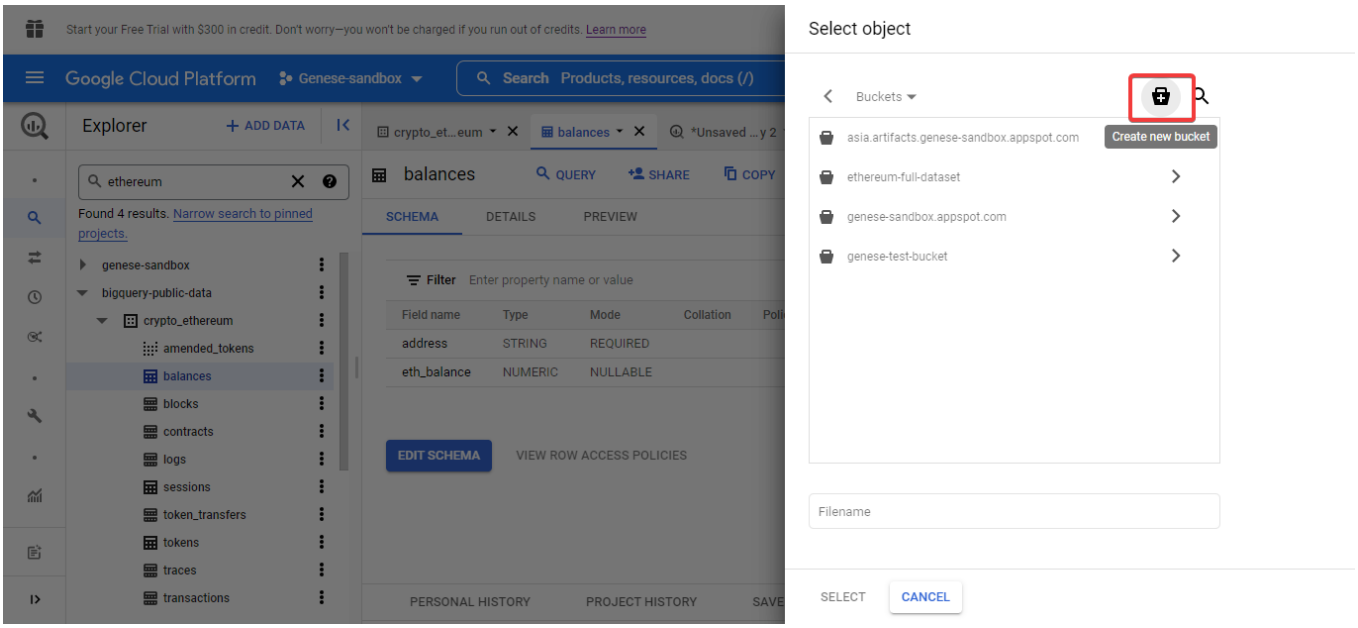


A dialogue box will appear

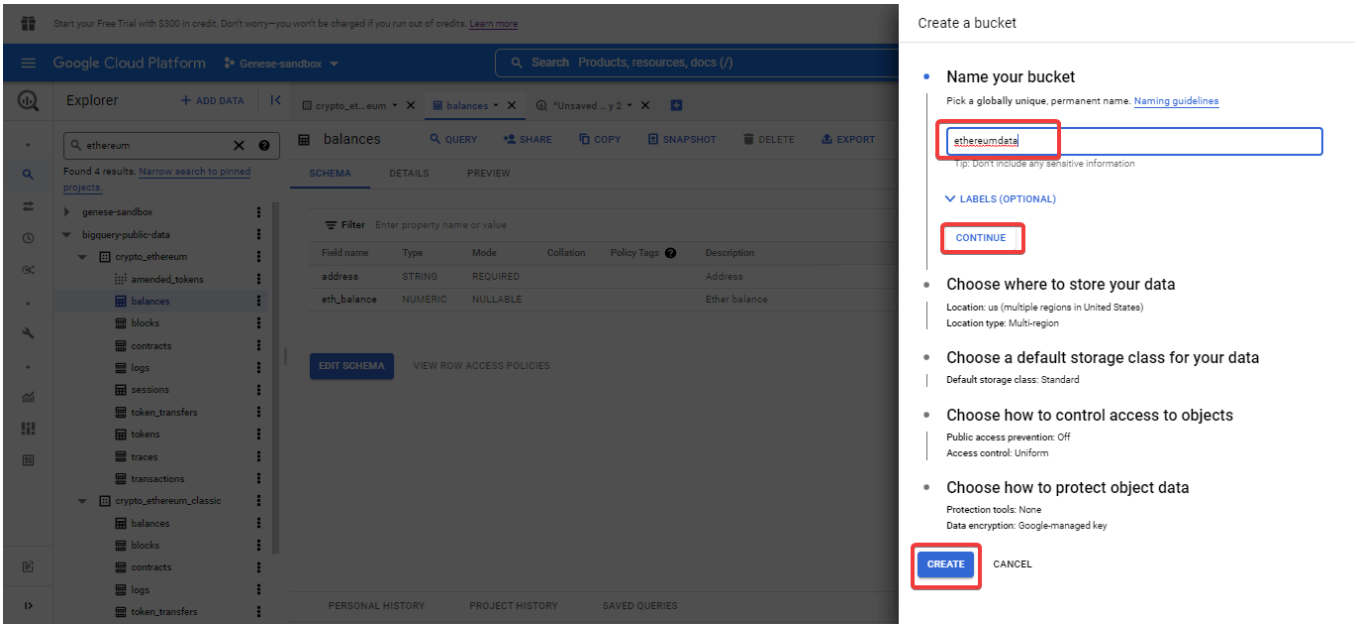
- Click **browse** to choose your bucket where you want to store your file.



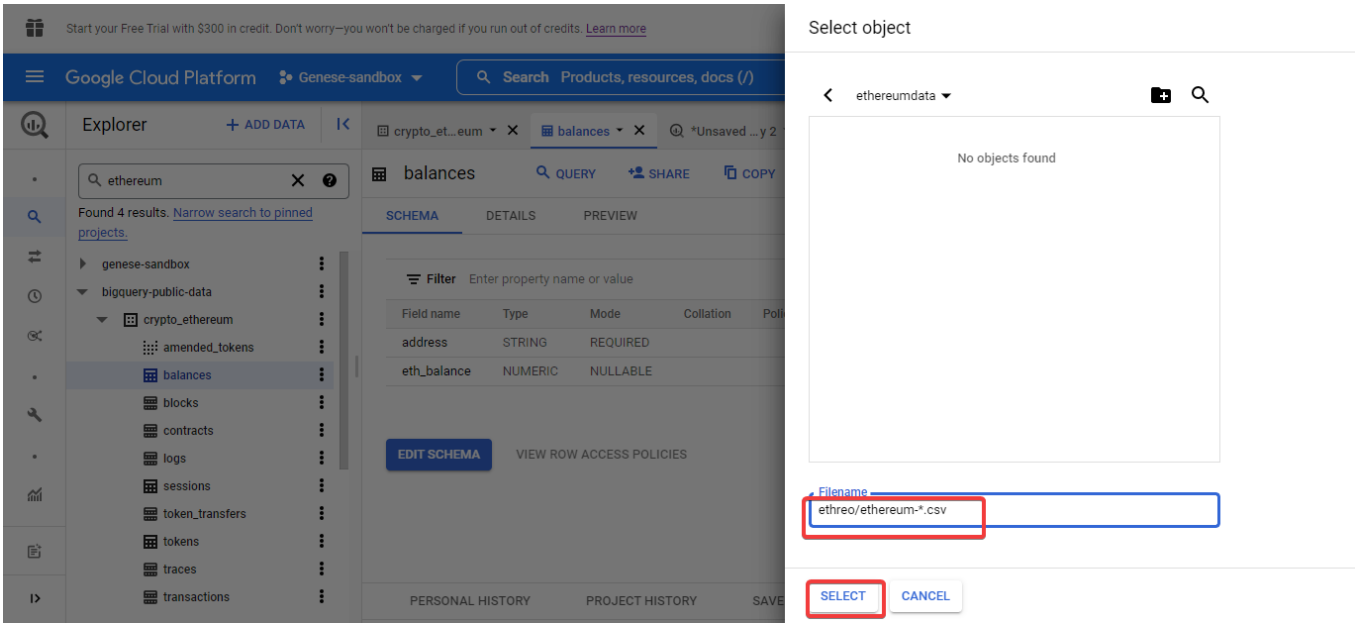
- Create new bucket.



- Name your bucket and continue. Then.click **CREATE.**

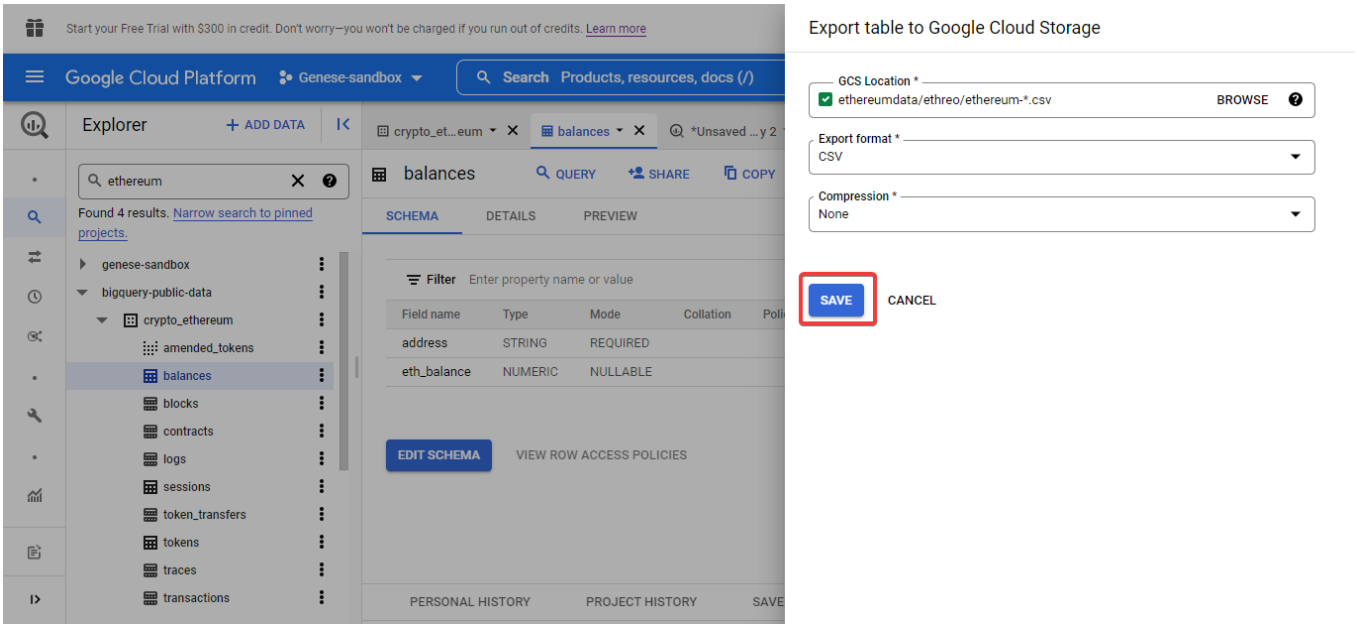


- Give a folder and file format you want to save on CSV or json file.
- Once completed Click Select.

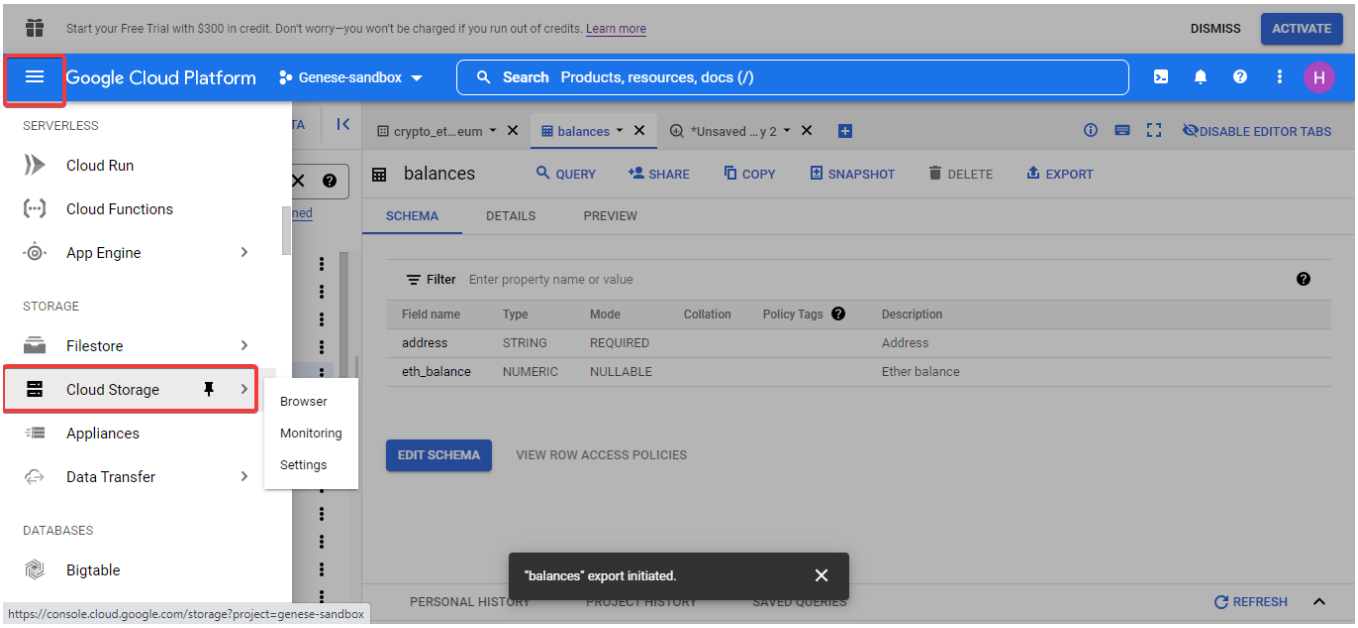


Lastly, we come into initial phase.

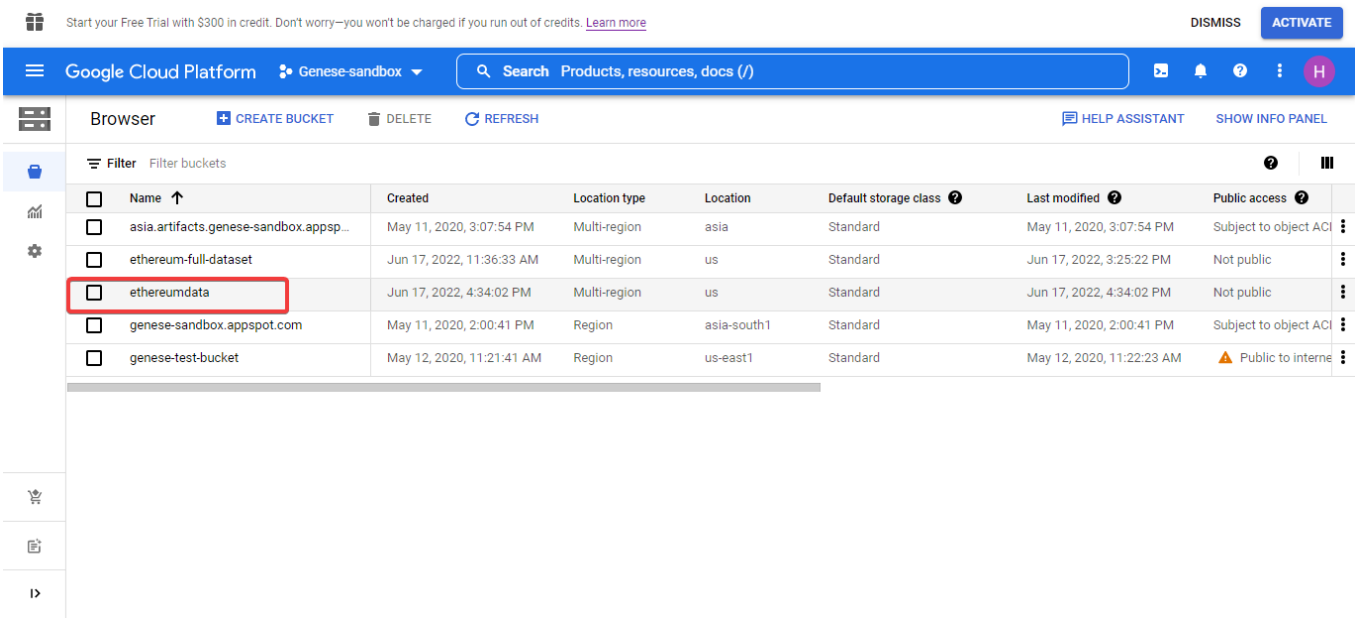
- Click **Save.**



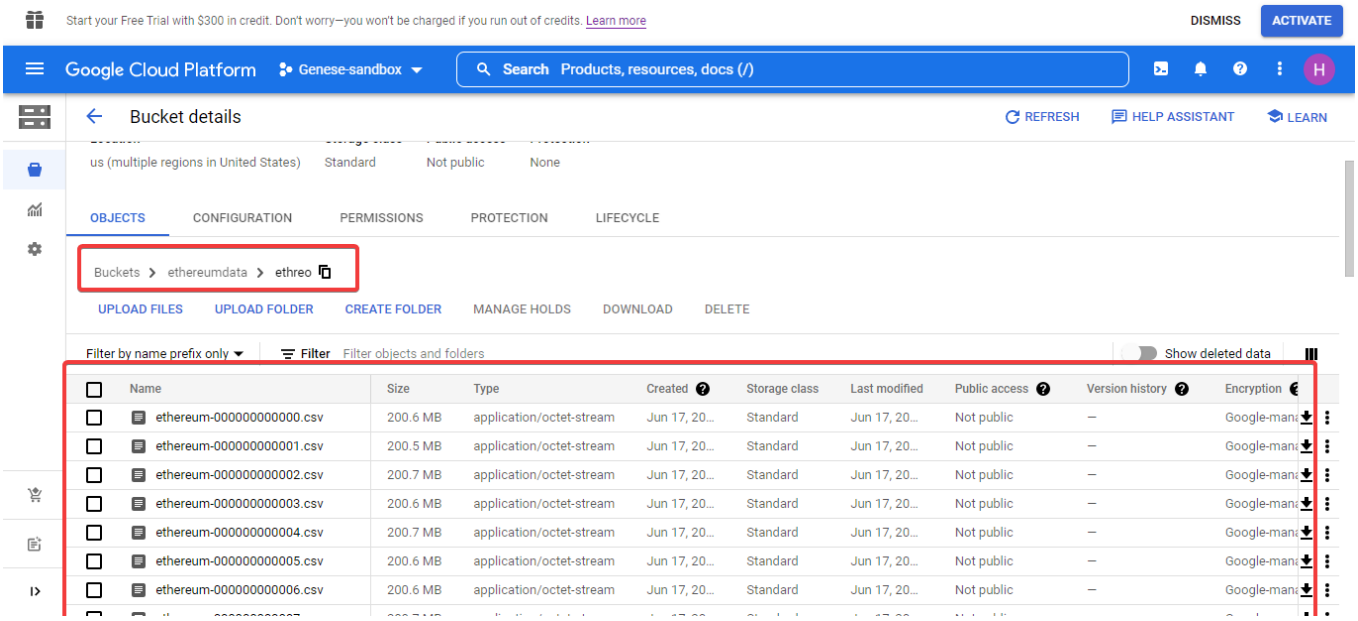
- Go to **Navigation** and Click **Cloud Storage.**



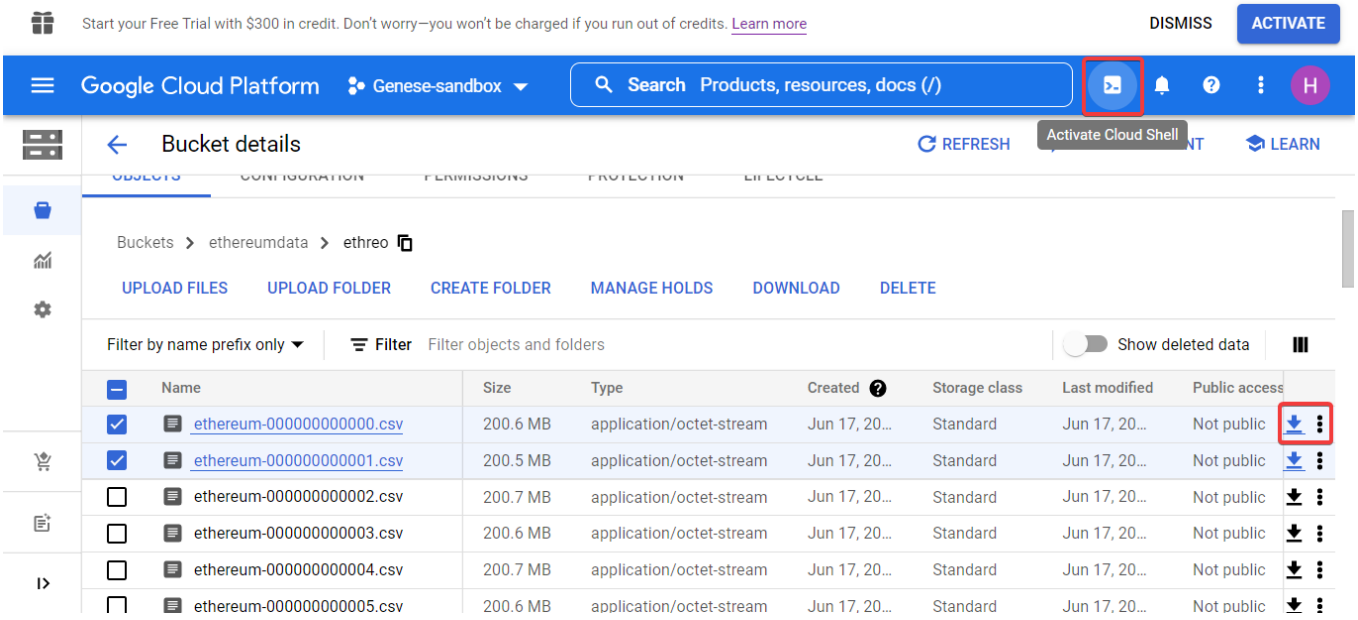
- Go to your bucket once you created and open the folder on it .



Finally you will see your files downloaded in segmented and numerical CSV files .



You can download file either using Activate Cloud Shell down-arrow button on corner.



Resources

EthereumETL | Grants (no date) *EthereumETL | Grants*. Available at: https://gitcoin.co/grants/233/ethereumetl (Accessed: 17 June 2022).

A Brief History of NFTs (no date). Available at: https://www.web3.university/tracks/build-your-first-nft/brief-history-of-nfts (Accessed: 17 June 2022).

Adrianus Yoga (2022) *Coursera - Exporting data from BigQuery to Cloud Storage*. Available at: https://www.youtube.com/watch?v=b6cV2W65s84 (Accessed: 17 June 2022).

Analyse Ethereum transactions with Google BigQuery (2020) *Anyblock Analytics*. Available at: https://www.anyblockanalytics.com/blog/analyse-transactions-with-ethereum-google-bigquery-data-set/ (Accessed: 17 June 2022).

Banik, R. (2021) 'Tutorial: Building a web3 frontend with React', *Scrappy Squirrels*, 9 December. Available at: https://medium.com/scrappy-squirrels/tutorial-building-a-web3-frontend-with-react-e0a87ea3bad (Accessed: 17 June 2022).

BharatiDWConsultancy (2020) *GCP - BigQuery - Union ALL, DISTINCT & Wildcard *_Table_Suffix - DIY#6*. Available at: https://www.youtube.com/watch?v=rAAeYt7Y7wQ (Accessed: 17 June 2022).

BoostUpStation (2021) *Google Cloud Platform Credentials | Using OAuth 2.0 to Access Google APIs | Client ID, access token*. Available at: https://www.youtube.com/watch?v=6MBLrH1j6Tg (Accessed: 17 June 2022).

Brossard, E. (2017) 'Answer to "Wildcard % operator not working on bigquery"', *Stack Overflow*. Available at: https://stackoverflow.com/a/44135182 (Accessed: 17 June 2022).

Buterin, V. (no date) 'Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform.', p. 36.

Cloud Storage client libraries (no date) *Google Cloud*. Available at: https://cloud.google.com/storage/docs/reference/libraries (Accessed: 17 June 2022).

CommandException: One or more URLs matched no objects. when I ran gsutil from Circle CI · Issue #1179 · GoogleCloudPlatform/gsutil (no date) *GitHub*. Available at: https://github.com/GoogleCloudPlatform/gsutil/issues/1179 (Accessed: 17 June 2022).

Community guides and resources (no date) *ethereum.org*. Available at: https://ethereum.org (Accessed: 17 June 2022).

Creating New BigQuery Datasets and Visualizing Insights (no date) *Coursera*. Available at: https://www.coursera.org/learn/gcp-creating-bigquery-datasets-visualizing-insights?action=enroll (Accessed: 17 June 2022).

CryptoKitties (no date) *CryptoKitties | Collect and breed digital cats!*, *CryptoKitties*. Available at: https://www.cryptokitties.co (Accessed: 17 June 2022).

DappReview - Discover your next favourite decentralized App (no date) *DappReview*. Available at: https://dapp.review (Accessed: 17 June 2022).

Data to CSV, execution failed due to file size - bytemeta (no date). Available at: https://bytemeta.vip/repo/google/crmint/issues/150 (Accessed: 17 June 2022).

DataExpert (2021) *Extracting Data from Google Cloud BigQuery to your Database | CSV | Excel Using Talend*. Available at: https://www.youtube.com/watch?v=Mv-DI8SmD84 (Accessed: 17 June 2022).

Ethereum ETL Airflow (2022). Blockchain ETL. Available at: https://github.com/blockchain-etl/ethereum-etl-airflow (Accessed: 17 June 2022).

Ethereum (execution layer) (no date). Available at: https://docs.infura.io/infura/networks/ethereum (Accessed: 17 June 2022).

Ethereum Historical Dataset (no date). Available at: https://www.kaggle.com/prasoonkottarathil/ethereum-historical-dataset (Accessed: 17 June 2022).

Ethereum price today, ETH to USD live, marketcap and chart (no date) *CoinMarketCap*. Available at: https://coinmarketcap.com/currencies/ethereum/ (Accessed: 17 June 2022).

ethereum/go-ethereum (2022). ethereum. Available at: https://github.com/ethereum/go-ethereum (Accessed: 17 June 2022).

‘Export BigQuery Table to CSV: 3 Easy Methods’ (no date) *Learn | Hevo*. Available at: https://hevodata.com/learn/export-bigquery-table-to-csv/ (Accessed: 17 June 2022).

Exporting Data From BigQuery - Google BigQuery — Google Developers (no date). Available at: https://download.huihoo.com/google/gdgdevkit/DVD1/developers.google.com/bigquery/exporting-data-from-bigquery.html (Accessed: 17 June 2022).

Exporting data from BigQuery into Athena (2018) *Big Data Demystified*. Available at: https://big-data-demystified.ninja/2018/05/27/how-to-export-data-from-google-big-query-into-aws-s3-emr-hive/ (Accessed: 17 June 2022).

Exporting table data | BigQuery (no date) *Google Cloud*. Available at: https://cloud.google.com/bigquery/docs/exporting-data (Accessed: 17 June 2022).

Exporting the Blockchain - Ethereum ETL (no date). Available at: https://ethereum-etl.readthedocs.io/en/latest/exporting-the-blockchain/ (Accessed: 17 June 2022).

gcloud auth login error Code Example (no date). Available at: https://www.codegrepper.com/code-examples/whatever/gcloud+auth+login+error (Accessed: 17 June 2022).

Getting started with authentication | Authentication (no date a) *Google Cloud*. Available at: https://cloud.google.com/docs/authentication/getting-started (Accessed: 17 June 2022).

Getting started with authentication | Authentication (no date b) *Google Cloud*. Available at: https://cloud.google.com/docs/authentication/getting-started (Accessed: 17 June 2022).

Google Cloud Tech (2015) *Getting started with Google Cloud Trace*. Available at: https://www.youtube.com/watch?v=NCFDqeo7AeY (Accessed: 17 June 2022).

Google Cloud Tech (2020) *BigQuery Web UI quickstart*. Available at: https://www.youtube.com/watch?v=PqGqmzLYb8A (Accessed: 17 June 2022).

‘google-auth Documentation’ (no date), p. 134.

google.auth._default — google-auth 1.30.0 documentation (no date). Available at: https://google-auth.readthedocs.io/en/master/_modules/google/auth/_default.html (Accessed: 17 June 2022).

google.oauth2.service_account_async module — google-auth 1.30.0 documentation (no date). Available at: https://google-auth.readthedocs.io/en/master/reference/google.oauth2._service_account_async.html (Accessed: 17 June 2022).

Hoffa, F. (2018) 'Answer to "Exporting large file from BigQuery to Google cloud using wildcard"', *Stack Overflow*. Available at: https://stackoverflow.com/a/51415646 (Accessed: 17 June 2022).

How to buy ETH (no date) *ethereum.org*. Available at: https://ethereum.org (Accessed: 17 June 2022).

How to Export BigQuery Data On a Schedule (2021) *Coupler.io Blog*. Available at: https://blog.coupler.io/bigquery-data-export/ (Accessed: 17 June 2022).

How to export data from the BigQuery API with Python | Census (2021) *Census Blog*. Available at: https://blog.getcensus.com/how-to-hack-it-extracting-data-from-google-bigquery-with-python-2/ (Accessed: 17 June 2022).

How to Export Data to a File in Google BigQuery (no date) *Chartio*. Available at: https://chartio.com/resources/tutorials/how-to-export-data-to-a-file-in-google-bigquery/ (Accessed: 17 June 2022).

How to import Google BigQuery tables to AWS Athena (2019) *freeCodeCamp.org*. Available at: https://www.freecodecamp.org/news/how-to-import-google-bigquery-tables-to-aws-athena-5da842a13539/ (Accessed: 17 June 2022).

Hwang, Y. (2021) 'Streaming Ethereum On-Chain Data to QuestDB', *Geek Culture*, 12 April. Available at: https://medium.com/geekculture/streaming-ethereum-on-chain-data-to-questdb-ea6b51d990ab (Accessed: 17 June 2022).

Isan-Rivkin (2018) 'How do you get all ERC20 tokens balances an address has? How do you sign transactions locally?', *r/ethdev*. Available at: www.reddit.com/r/ethdev/comments/84le1j/how_do_you_get_all_erc20_tokens_balances_an/ (Accessed: 17 June 2022).

jz22 (2021) 'Google Cloud Load Balancer with wildcard', *Server Fault*. Available at: https://serverfault.com/q/1088518 (Accessed: 17 June 2022).

Kasireddy, P. (2019) 'How does Ethereum work, anyway?', *Medium*, 29 October. Available at: https://preethikasireddy.medium.com/how-does-ethereum-work-anyway-22d1df506369 (Accessed: 17 June 2022).

Liang, J. (2018) 'Ethereum'. Harvard Dataverse. doi:10.7910/DVN/XIXSPR.

LLC, G. (no date) *google-cloud-bigquery: Google BigQuery API client library* [OS Independent]. Available at: https://github.com/googleapis/python-bigquery (Accessed: 17 June 2022).

Lopp, J. (2019) 'The Challenges of Building Ethereum Infrastructure', *Medium*, 14 July. Available at: https://medium.com/@lopp/the-challenges-of-building-ethereum-infrastructure-87e443e47a4b (Accessed: 17 June 2022).

Marius (2017) 'Answer to "Set GOOGLE_APPLICATION_CREDENTIALS in Python project to use Google API"', *Stack Overflow*. Available at: https://stackoverflow.com/a/46651026 (Accessed: 17 June 2022).

Medvedev, E. (2018) 'Exporting and Analyzing Ethereum Blockchain', *Medium*, 23 November. Available at: https://evgemedvedev.medium.com/exporting-and-analyzing-ethereum-blockchain-f5353414a94e (Accessed: 17 June 2022).

Medvedev, E. (2020) 'How to Query Balances for all Ethereum Addresses in BigQuery', *Google Cloud - Community*, 22 January. Available at: https://medium.com/google-cloud/how-to-query-balances-for-all-ethereum-addresses-in-bigquery-fb594e4034a7 (Accessed: 17 June 2022).

Medvedev, E. (2021) 'Query Ethereum Blockchain with SQL in Google BigQuery', *Medium*, 24 February. Available at: https://evgemedvedev.medium.com/ethereum-blockchain-on-google-bigquery-283fb300f579 (Accessed: 17 June 2022).

N, J. (2020) 'How to export more than 16000 rows in bigquery table to local machine as CSV file?', *Stack Overflow*. Available at: https://stackoverflow.com/q/65108688 (Accessed: 17 June 2022).

Numpy Ninja (2021) *GCPDE-BigQuery wildcard tables*. Available at: https://www.youtube.com/watch?v=u7Pd7fpZO4I (Accessed: 17 June 2022).

Peace (2022) *A Labeled Transactions-Based Dataset on the Ethereum Network*. Available at: https://github.com/salam-ammari/Labeled-Transactions-based-Dataset-of-Ethereum-Network (Accessed: 17 June 2022).

Puccetti, A. (2020) 'BigQuery Wildcards', *Medium*, 21 January. Available at: https://alepuccetti.medium.com/bigquery-wildcards-3aa155157484 (Accessed: 17 June 2022).

Query multiple tables using a wildcard table | BigQuery (no date) *Google Cloud*. Available at: https://cloud.google.com/bigquery/docs/querying-wildcard-tables (Accessed: 17 June 2022).

Quickstart: Using client libraries | BigQuery (no date) *Google Cloud*. Available at: https://cloud.google.com/bigquery/docs/quickstarts/quickstart-client-libraries (Accessed: 17 June 2022).

Qwiklabs (2019) 'Hey, BigQuery!', *Medium*, 19 February. Available at: https://qwiklabs.medium.com/hey-bigquery-e9b5bdf786b0 (Accessed: 17 June 2022).

Siddiqui, I. (2020) 'Solution To Extract Data From Ethereum', *Coinmonks*, 1 September. Available at: https://medium.com/coinmonks/solution-to-extract-data-from-ethereum-52d0b8007d1b (Accessed: 17 June 2022).

'[Solved] ImportError: No module named "google"' (2021) *Exception Error*, 22 October. Available at: https://exerror.com/importerror-no-module-named-google/ (Accessed: 17 June 2022).

Supermetrics (2019) *Supermetrics for BigQuery - How to extract the data you need with 3 query types*. Available at: https://www.youtube.com/watch?v=5ZaOqEkyp8E (Accessed: 17 June 2022).

Svanevik, A. (2020) 'How to get any Ethereum smart contract into BigQuery (in 8 mins)', *Medium*, 5 May. Available at: https://medium.com/@ASvanevik/how-to-get-any-ethereum-smart-contract-into-bigquery-in-8-mins-bab5db1fdeee (Accessed: 17 June 2022).

Tech Vine (2022) *ETL Processing on Google Cloud Using Dataflow and BigQuery || [GSP290] || Solution*. Available at: https://www.youtube.com/watch?v=s7LufrmQsGQ (Accessed: 17 June 2022).

Three methods for exporting CSV files from BigQuery | Census (2021) *Census Blog*. Available at: https://blog.getcensus.com/3-ways-to-export-csv-files-from-google-bigquery/ (Accessed: 17 June 2022).

Tutorial: How to query multiple 'events_' tables with '_table_suffix' (GA4) (2021) *GA4BigQuery*. Available at: https://www.ga4bigquery.com/tutorial-how-to-query-multiple-analytics-events-tables-with-table-suffix-ga4/ (Accessed: 17 June 2022).

Understanding Ethereum (no date) *districtOx Education Portal*. Available at: https://education.districtox.io/general-topics/understanding-ethereum/ (Accessed: 17 June 2022).

Use table wildcards for easy merges - Joining and Merging Datasets (no date) *Coursera*. Available at: https://www.coursera.org/lecture/gcp-creating-

bigquery-datasets-visualizing-insights/use-table-wildcards-for-easy-merges-mKWIf (Accessed: 17 June 2022).

Using OAuth 2.0 for Web Server Applications | YouTube Data API (no date) *Google Developers*. Available at: https://developers.google.com/youtube/v3/guides/auth/server-side-web-apps (Accessed: 17 June 2022).

What Is Ethereum? (no date) *district0x Education Portal*. Available at: https://education.district0x.io/general-topics/understanding-ethereum/what-is-ethereum/ (Accessed: 17 June 2022).

Zheng, P., Zheng, Z. and Dai, H. (2019) 'XBlock-ETH: Extracting and Exploring Blockchain Data From Ethereum'. arXiv. doi:10.48550/arXiv.1911.00169.