MVP

Cláudia Jacy Barenco Abbas

MOTIVAÇÃO

Possuo um Startup de Auditoria e Segurança em aplicações Web 3.0 então procurei desenvolver um MVP relacionado com esta área.

O objetivo deste MVP é o estudo das plataformas de classificação de reputação de investidores da moeda Bitcoin (plataforma Alpha e Plataforma OTC) para verificar se estas avaliações são consistentes e assim poderiam ser usada na detecção de nós fraudadores relacionando os resultados deste MVP com as informações publicadas nas mídias sobre ataques ao blockchain do Bitcoin.

PERGUNTAS

- QUANTOS E QUAIS NÓS FORAM MAL AVALIADOS EM AMBAS PLATAFORMAS (< 5)
- EM QUE ÉPOCA HOUVE MAIS NÓS MAL AVALIADOS ?
- COM O PASSAR DO TEMPO HOUVE MENOS NÓS MAL AVALIADOS ?
- ESTAS PLATAFORMAS PODEM SER UMA FERRAMENTA PARA APONTAR UMA TENDÊNCIA NO AUMENTO DOS NÓS ATACADOS ?
- PARA OS ANOS MAIS RECENTES. TEMOS OS MESMOS RESULTADOS ?

BUSCA PELOS DADOS

SNAP: Signed network datasets: Bitcoin Alpha web of trust network (stanford.edu)

BITCOIN ALPHA

This is who-trusts-whom network of people who trade using Bitcoin on a platform called Bitcoin Alpha. Since Bitcoin users are anonymous, there is a need to maintain a record of users' reputation to prevent transactions with fraudulent and risky users. Members of Bitcoin Alpha rate other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1. This is the first explicit weighted signed directed network available for research.

```
SOURCE, TARGET, RATING, TIME
```

- S. Kumar, B. Hooi, D. Makhija, M. Kumar, V.S. Subrahmanian, C. Faloutsos. REV2: Fraudulent User Prediction in Rating Platforms. 11th ACM International Conference on Web Searchand Data Mining (WSDM), 2018.

where

- `SOURCE`: node id of source, i.e., rater
- `TARGET`: node id of target, i.e., ratee
- `RATING`: the source's rating for the target, ranging from -10 to +10 in steps of 1
- `TIME`: the time of the rating, measured as seconds since Epoch. (This can be converted to human readable data easily as described here)

Dataset statistics	
Nodes	3,783
Edges	24,186
Range of edge weight	-10 to +10
Percentage of positive edges	93%

Similar network from another Bitcoin platform, Bitcoin OTC, is available here.

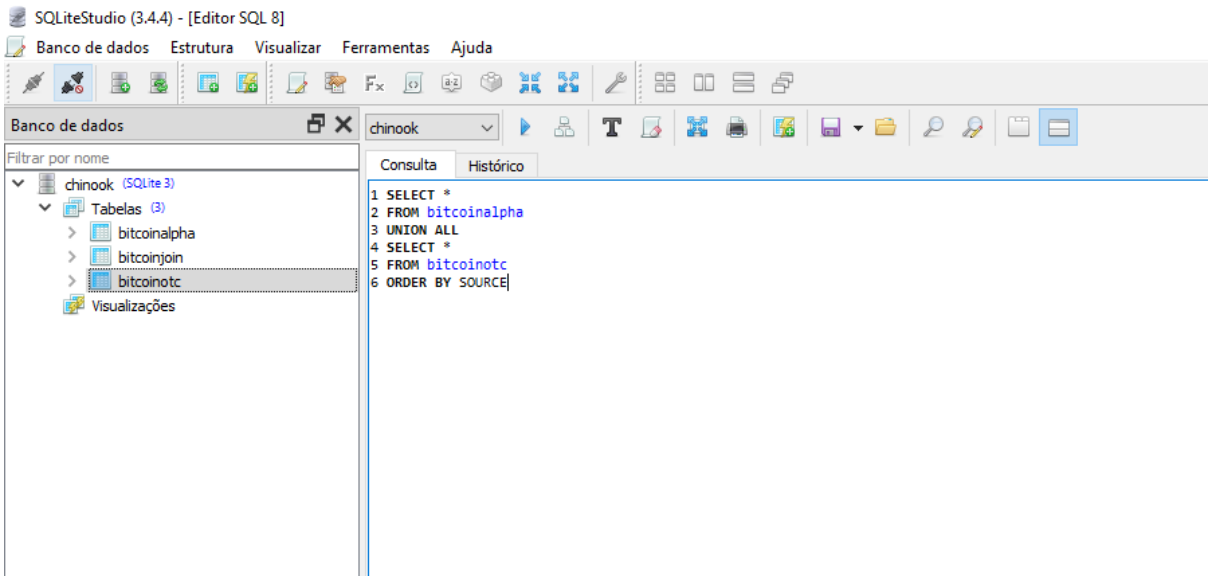
BITCOIN OTC

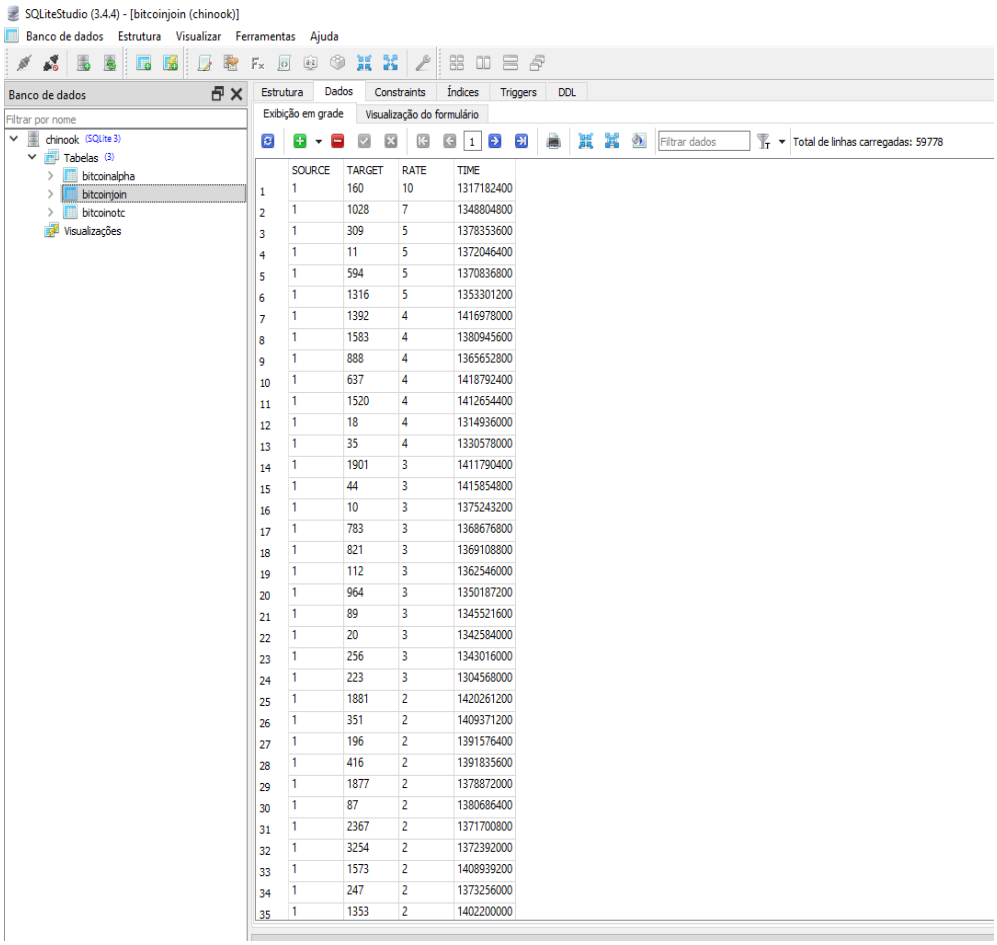
#bitcoin-otc

This is who-trusts-whom network of people who trade using Bitcoin on a platform called Bitcoin OTC. Since Bitcoin users are anonymous, there is a need to maintain a record of users' reputation to prevent transactions with fraudulent and risky users. Members of Bitcoin OTC rate other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1. This is the first explicit weighted signed directed network available for research

Dataset statistics	
Nodes	5,881
Edges	35,592
Range of edge weight	-10 to +10
Percentage of positive edges	89%

CRIAÇÃO DO BANCO DE DADOS COM O JOIN DAS DUAS TABELAS NO SQLLITESTUDIO

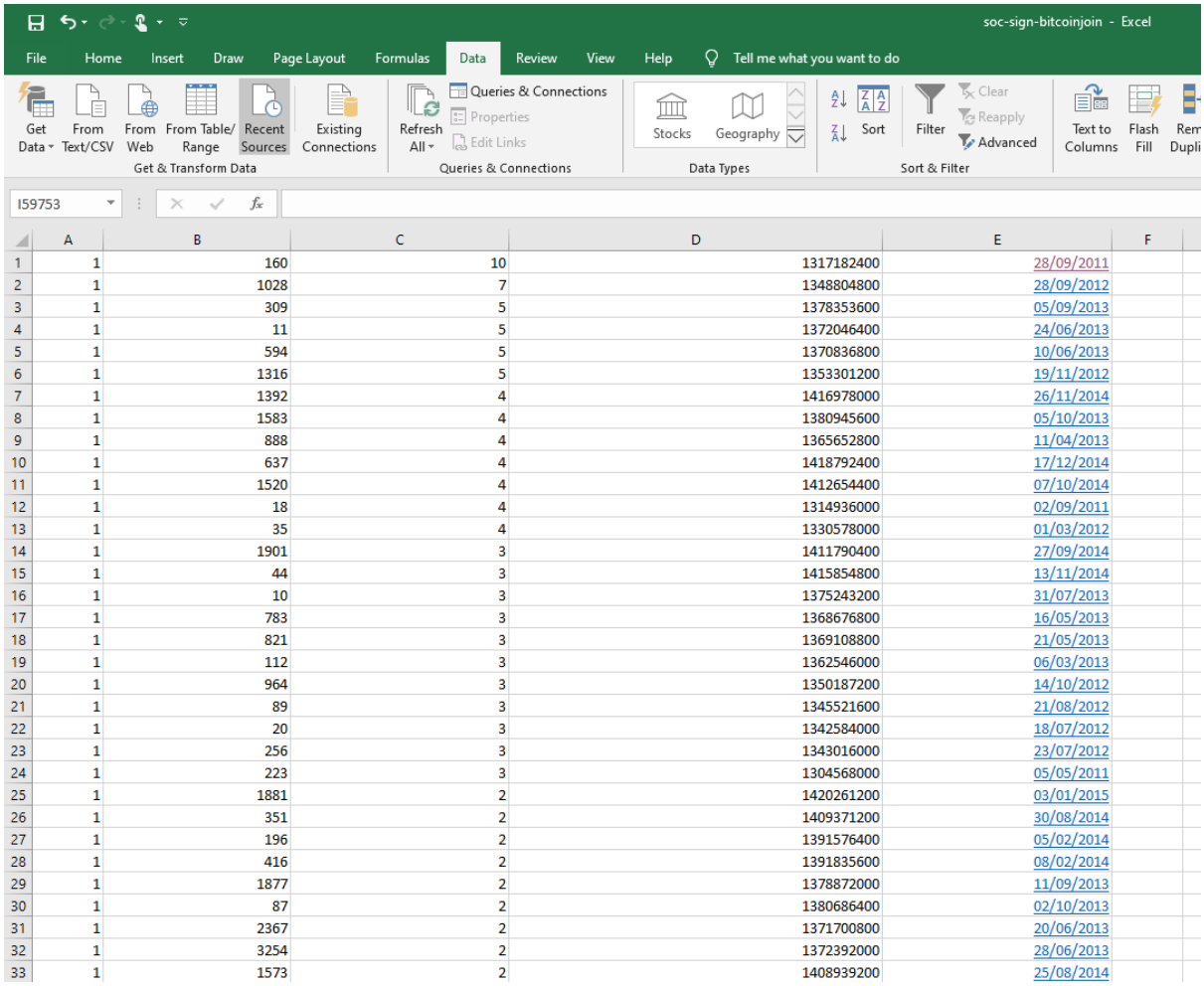




	SOURCE	TARGET	RATE	TIME
1	1	160	10	1317182400
2	1	1028	7	1348804800
3	1	309	5	1378353600
4	1	11	5	1372046400
5	1	594	5	1370836800
6	1	1316	5	1353301200
7	1	1392	4	1416978000
8	1	1583	4	1380945600
9	1	888	4	1365652800
10	1	637	4	1418792400
11	1	1520	4	1412654400
12	1	18	4	1314936000
13	1	35	4	1330578000
14	1	1901	3	1411790400
15	1	44	3	1415854800
16	1	10	3	1375243200
17	1	783	3	1368676800
18	1	821	3	1369108800
19	1	112	3	1362546000
20	1	964	3	1350187200
21	1	89	3	1345521600
22	1	20	3	1342584000
23	1	256	3	1343016000
24	1	223	3	1304568000
25	1	1881	2	1420261200
26	1	351	2	1409371200
27	1	196	2	1391576400
28	1	416	2	1391835600
29	1	1877	2	1378872000
30	1	87	2	1380686400
31	1	2367	2	1371700800
32	1	3254	2	1372392000
33	1	1573	2	1408939200
34	1	247	2	1373256000
35	1	1353	2	1402200000

TRATAMENTO DOS DADOS

Foi necessária a conversão do dado time para um dado que seja de fácil leitura.



	A	B	C	D	E	F
1	1	160	10	1317182400	28/09/2011	
2	1	1028	7	1348804800	28/09/2012	
3	1	309	5	1378353600	05/09/2013	
4	1	11	5	1372046400	24/06/2013	
5	1	594	5	1370836800	10/06/2013	
6	1	1316	5	1353301200	19/11/2012	
7	1	1392	4	1416978000	26/11/2014	
8	1	1583	4	1380945600	05/10/2013	
9	1	888	4	1365652800	11/04/2013	
10	1	637	4	1418792400	17/12/2014	
11	1	1520	4	1412654400	07/10/2014	
12	1	18	4	1314936000	02/09/2011	
13	1	35	4	1330578000	01/03/2012	
14	1	1901	3	1411790400	27/09/2014	
15	1	44	3	1415854800	13/11/2014	
16	1	10	3	1375243200	31/07/2013	
17	1	783	3	1368676800	16/05/2013	
18	1	821	3	1369108800	21/05/2013	
19	1	112	3	1362546000	06/03/2013	
20	1	964	3	1350187200	14/10/2012	
21	1	89	3	1345521600	21/08/2012	
22	1	20	3	1342584000	18/07/2012	
23	1	256	3	1343016000	23/07/2012	
24	1	223	3	1304568000	05/05/2011	
25	1	1881	2	1420261200	03/01/2015	
26	1	351	2	1409371200	30/08/2014	
27	1	196	2	1391576400	05/02/2014	
28	1	416	2	1391835600	08/02/2014	
29	1	1877	2	1378872000	11/09/2013	
30	1	87	2	1380686400	02/10/2013	
31	1	2367	2	1371700800	20/06/2013	
32	1	3254	2	1372392000	28/06/2013	
33	1	1573	2	1408939200	25/08/2014	

VISUALIZAÇÃO GRÁFICA DOS DADOS

UTILIZAÇÃO NEO4J AURA DB para para importar e manusear os dados a partir de arquivos no formato JSON.

Import results

Total time: 00:00:14

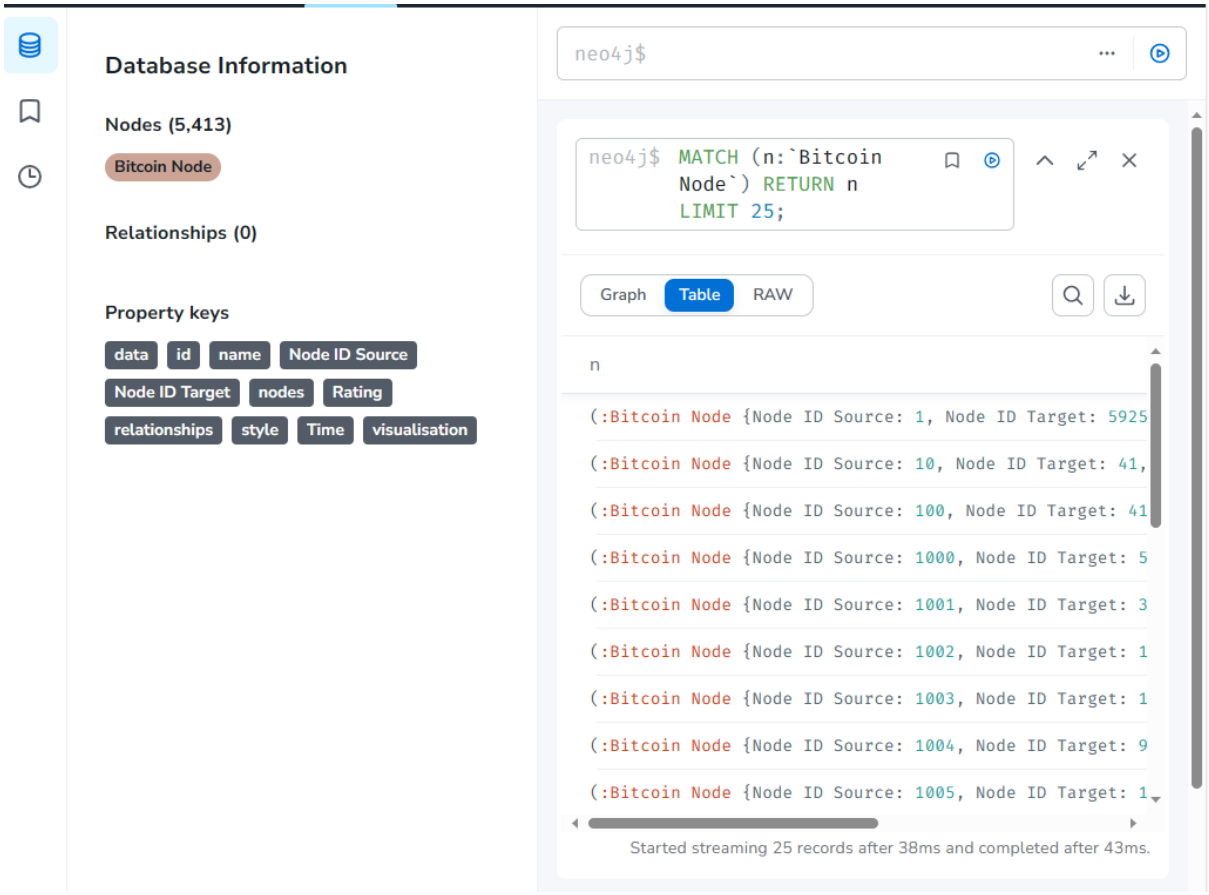
Run import completed

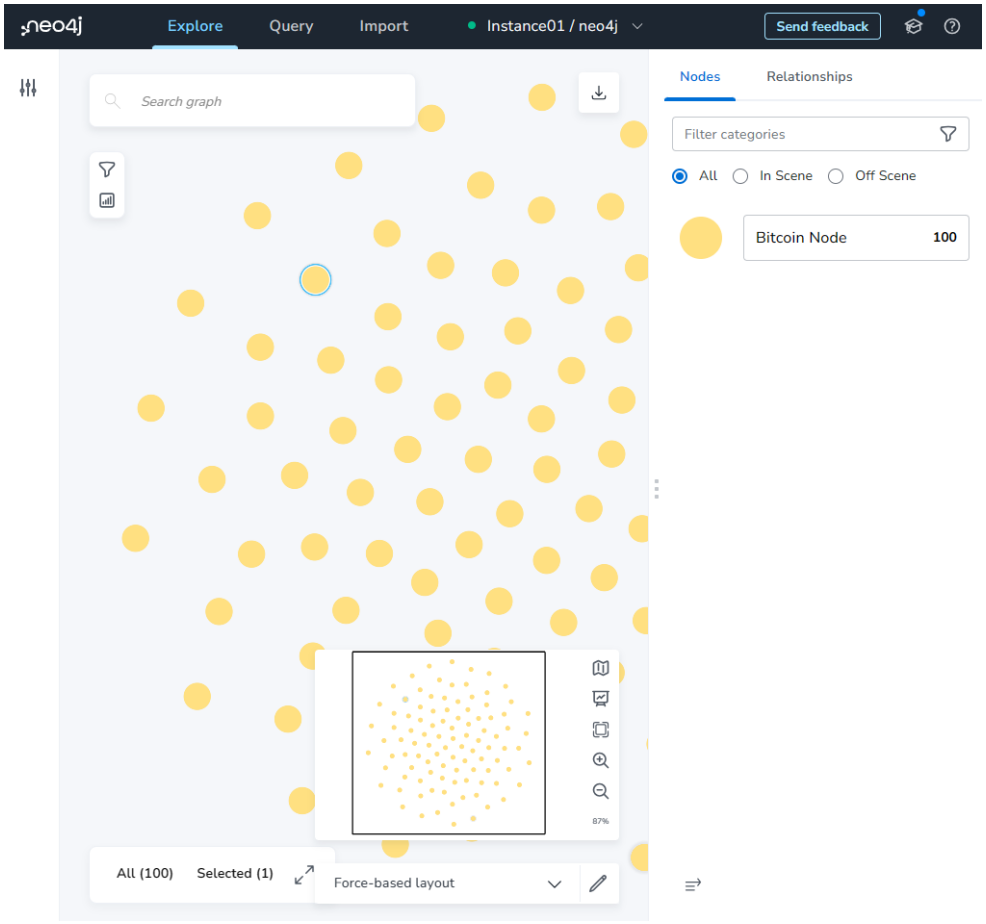
Import completed successfully.

Bitcoin Node soc-sign-bitcoinjoin.csv Show Cypher

Time taken	File size	File rows	Nodes created	Properties set	Labels added	Query count	Query time
00:00:14	1.9 MiB	59,777	5,413	157,217	5,413	3	00:00:13

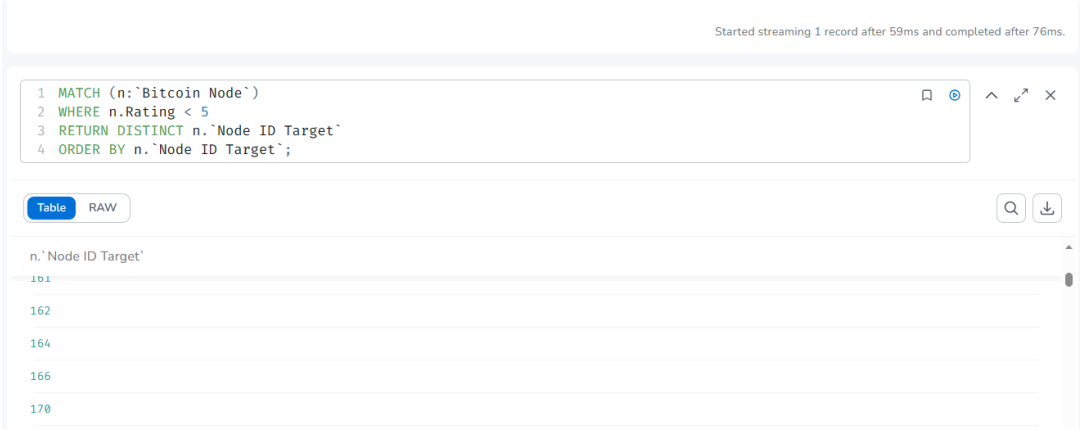
Close Explore results



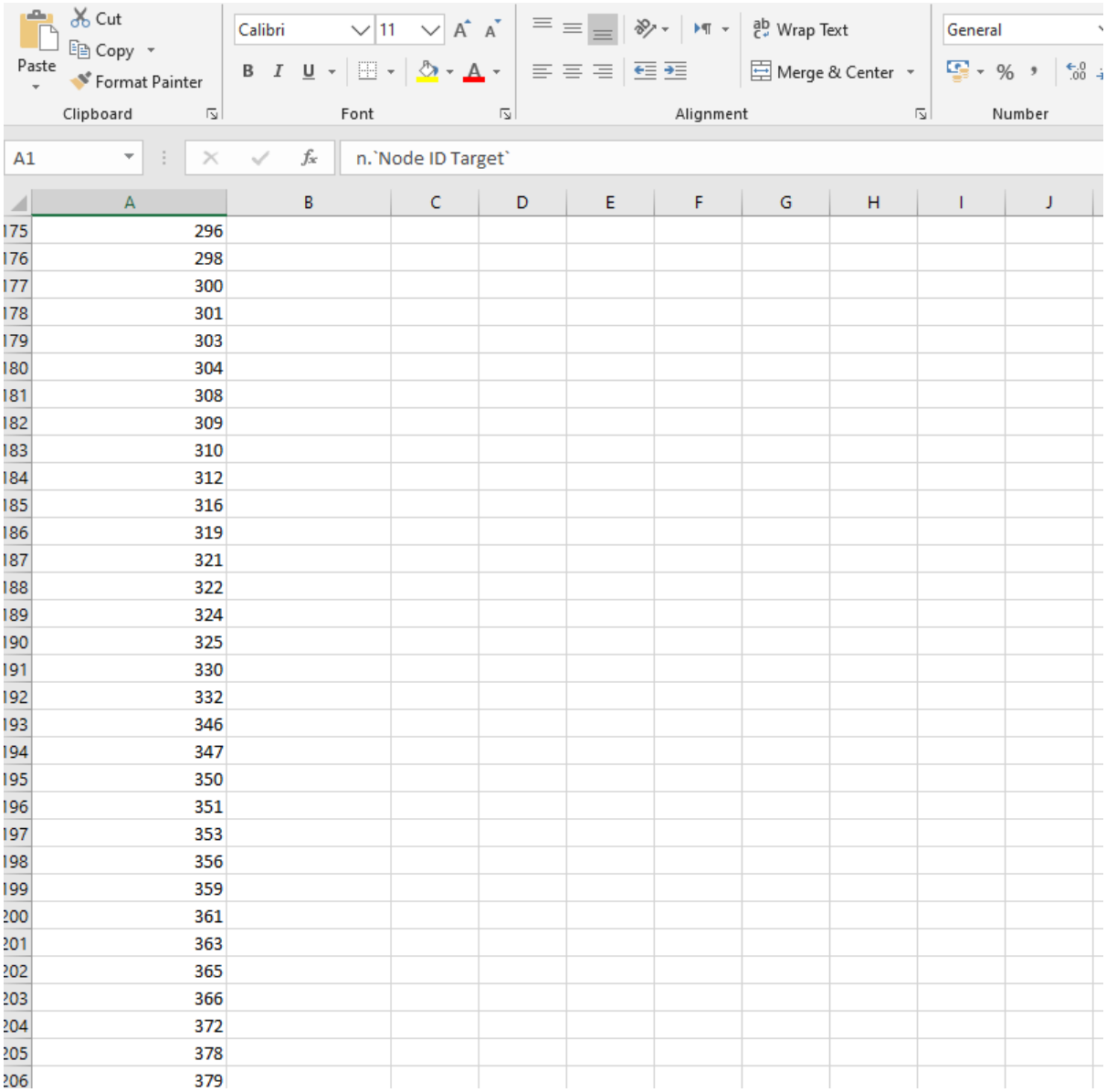


Realização dos Cypher Query

- **QUAIS NÓS FORAM MAL AVALIADOS EM AMBAS PLATAFORMAS (RATING < 5)**

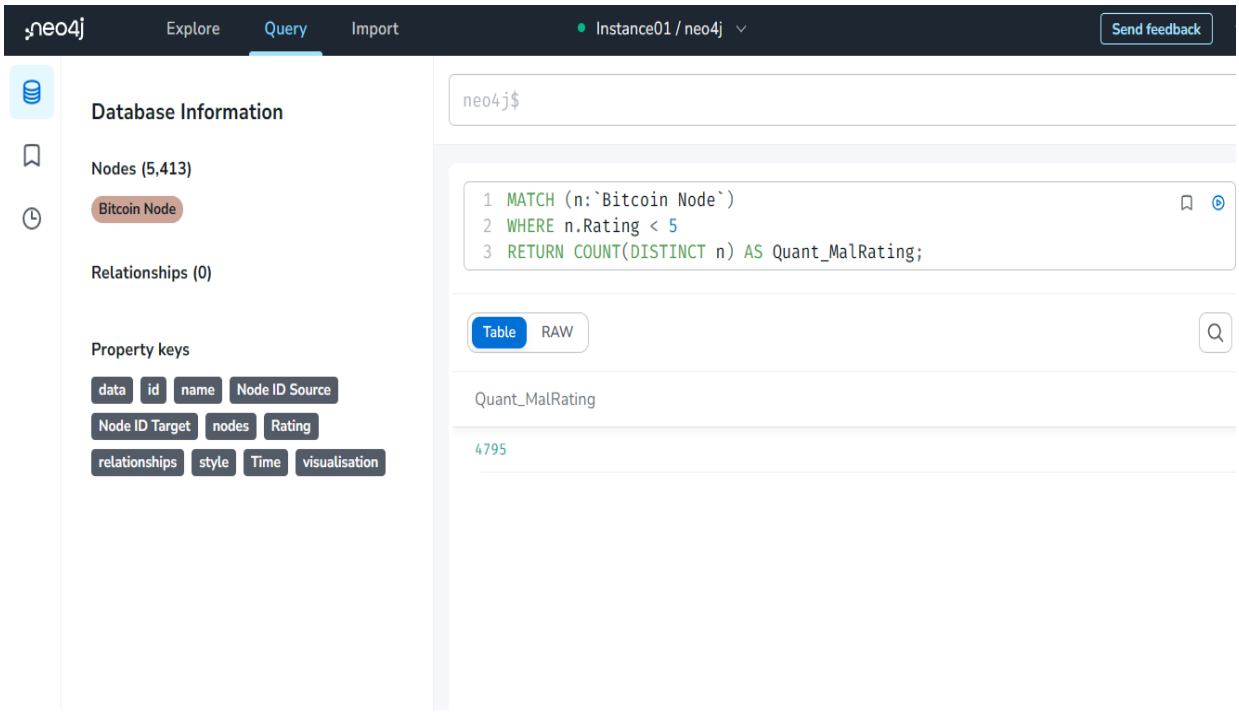


RESULTADOS



	A	B	C	D	E	F	G	H	I	J
175	296									
176	298									
177	300									
178	301									
179	303									
180	304									
181	308									
182	309									
183	310									
184	312									
185	316									
186	319									
187	321									
188	322									
189	324									
190	325									
191	330									
192	332									
193	346									
194	347									
195	350									
196	351									
197	353									
198	356									
199	359									
200	361									
201	363									
202	365									
203	366									
204	372									
205	378									
206	379									

- **QUANTOS NÓS FORAM MAL AVALIADOS ?**



Foram encontrados 4795 nós Bitcoin mal avaliados.

- **EM QUE ÉPOCA HOUVE MAIS NÓS MAL AVALIADOS ?**

Para responder esta pergunta foi criado um relacionamento DATE (TEM_RATING_EM)

Import results

Total time: 00:00:09

Run import completed

Import completed successfully.

00:00:03	1.9 MiB	59,778	0	119,418	0	3	00:00:03

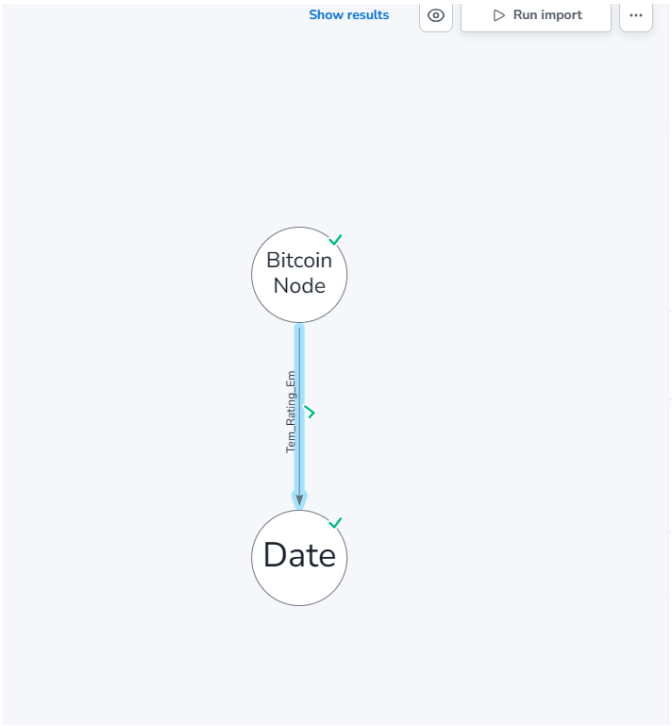
Date soc-sign-bitcoinjoin.csv Show Cypher

Time taken	File size	File rows	Nodes created	Properties set	Labels added	Query count	Query time
00:00:01	1.9 MiB	59,778	709	709	709	3	00:00:01

Tem_Rating_Em soc-sign-bitcoinjoin.csv Show Cypher

Time taken	File size	File rows	Relationships created	Properties set	Query count	Query time
00:00:04	1.9 MiB	59,778	17,687	0	3	00:00:04

Close Explore results



Relationship type
Reverse direction

Name

Tem_Rating_Em

File

Name

soc-sign-bitcoinjoin.csv

Filter file

Node ID mapping

	Node	ID	ID column
From	Bitcoin Node	Node ID Source	SOURCE
To	Date	Date_Rating	DATE

Need help?

Properties
Map from file

Database Information

Nodes (6,122)

Bitcoin Node
Date

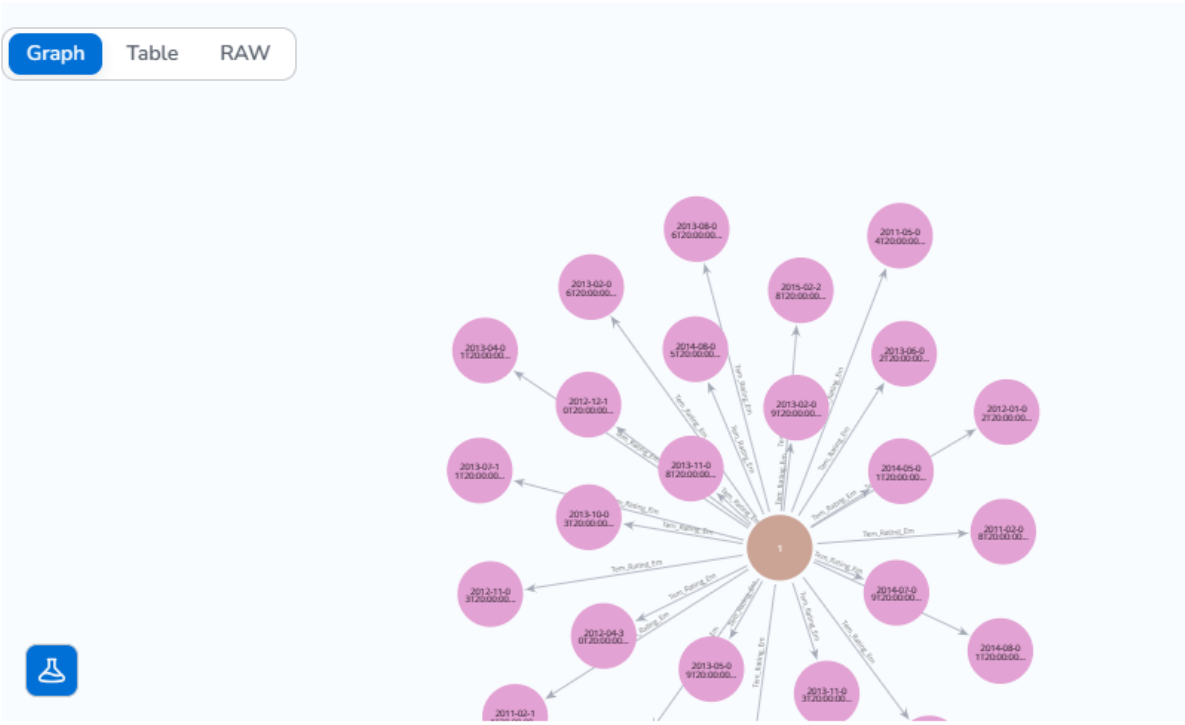
Relationships (17,687)

Tem_Rating_Em

Property keys

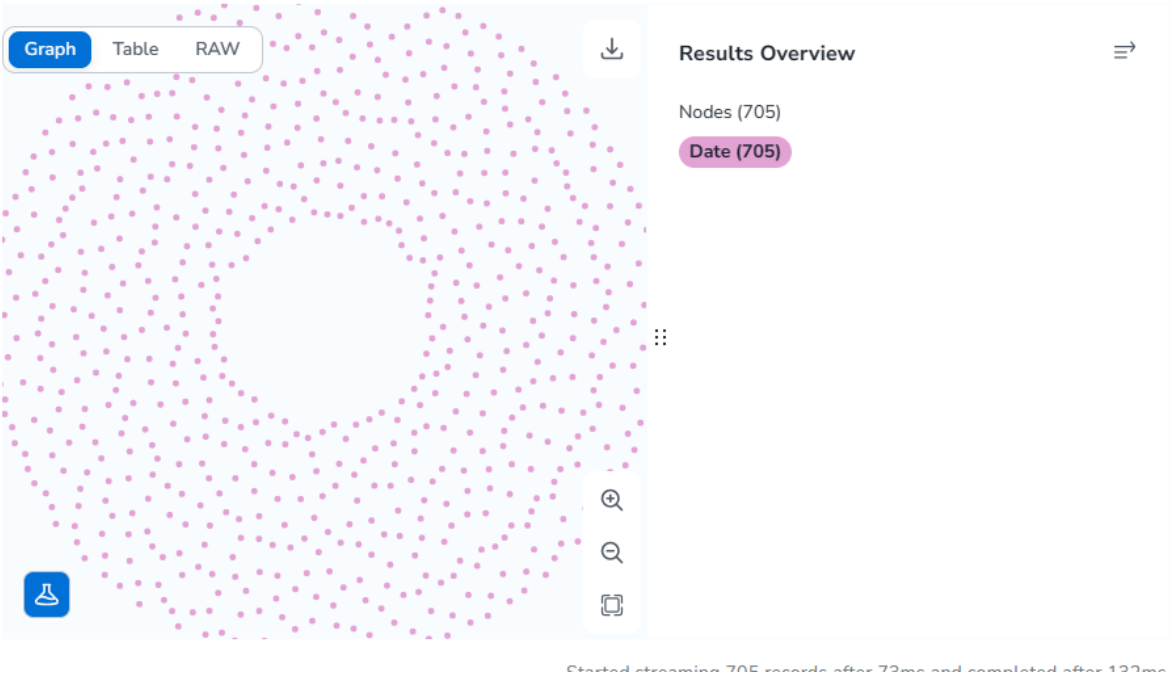
data
Date_Rating
id
name
Node ID Source
Node ID Target
nodes
Rating
relationships
style
Time
visualisation

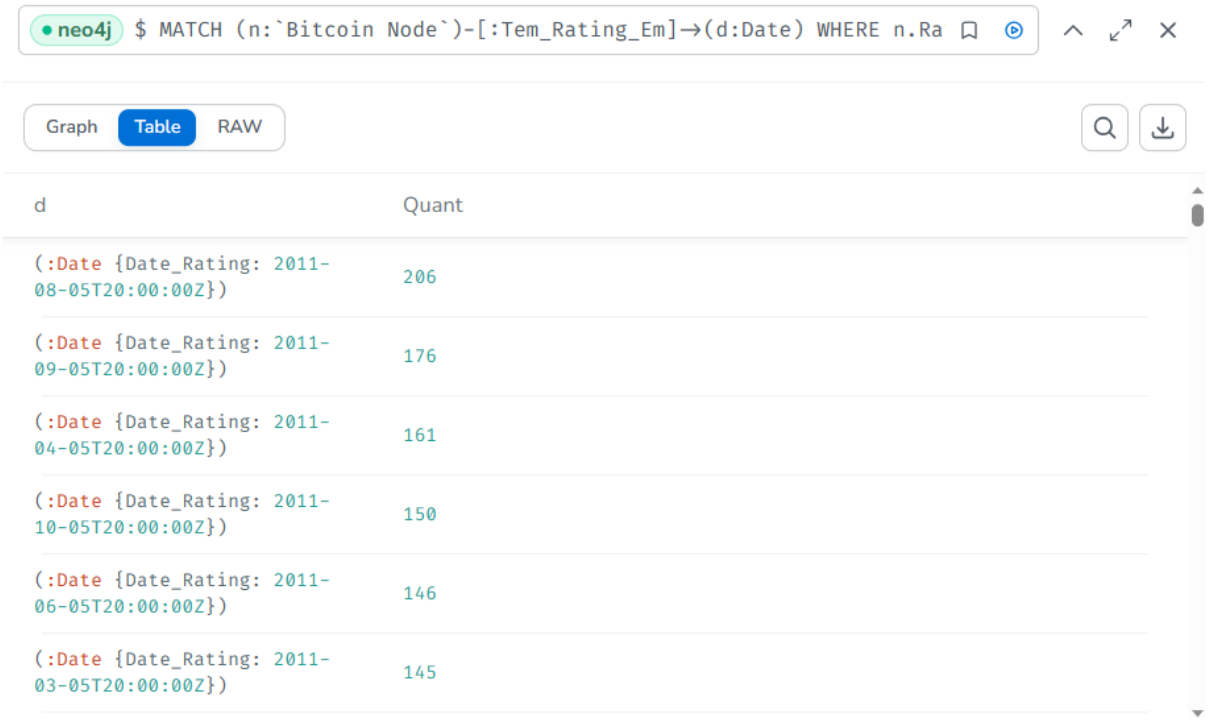
```
neo4j $ MATCH p=()-[:Tem_Rating_Em]→() RETURN p LIMIT 25;
```



QUERY PARA OBTER O RESULTADO

```
1 MATCH (n:`Bitcoin Node`)-[:Tem_Rating_Em]→(d:Date)
2 WHERE n.Rating < 5
3 WITH d, COUNT(n) AS Quant
4 ORDER BY Quant DESC
5 RETURN d, Quant;
```





O ano de 2011 foi o que teve mais nós mal avaliados com 206 avaliações em 05 de agosto de 2011.

- **COM O PASSAR DO TEMPO HOUVE MENOS NÓS MAL AVALIADOS ?**

QUERY PARA OBTER O RESULTADO

```
MATCH (n:`Bitcoin Node`)-[:Tem_Rating_Em]→(d:Date)
WHERE n.Rating < 5
WITH d.Date_Rating AS Data, COUNT(n) AS Quant
WITH Data.year AS Year, Quant
RETURN Year, SUM(Quant) AS TotalQuant
ORDER BY Year DESC;
```

Table RAW

Year	TotalQuant
2016	19
2015	363
2014	1817
2013	4775
2012	4183
2011	4311
2010	47

Como pode-se ver em 2016 temos somente 19 más avaliações de nós mas em 2015 363 **e o pico foi no ano de 2013. Então conclui-se que não houve um padrão de diminuissão no número de ataques.**

- **ESTAS PLATAFORMAS PODEM SER UMA FERRAMENTA PARA APONTAR UMA TENDÊNCIA NO AUMENTO DOS NÓS ATACADOS ?**

BUSCA NA INTERNET POR NOTÍCIAS SOBRE ATAQUES A REDE DO BITCOIN NO ANO DE 2013

Basenado-se nos dados mostrados neste site pode-se notar que de 2011 a 2016 o ano de 2013 tivemos mais advertências de ataques ao blockchain do Bitcoin e o segundo ano foi 2012 tendo uma relação positiva com os dados da quantidade de nós mal avaliados no estudo anterior. (Referência: Common Vulnerabilities and Exposures - Bitcoin Wiki)

2011 -> 1

2012 -> 6

2013 -> 11

2014 -> 2

2015 -> 2

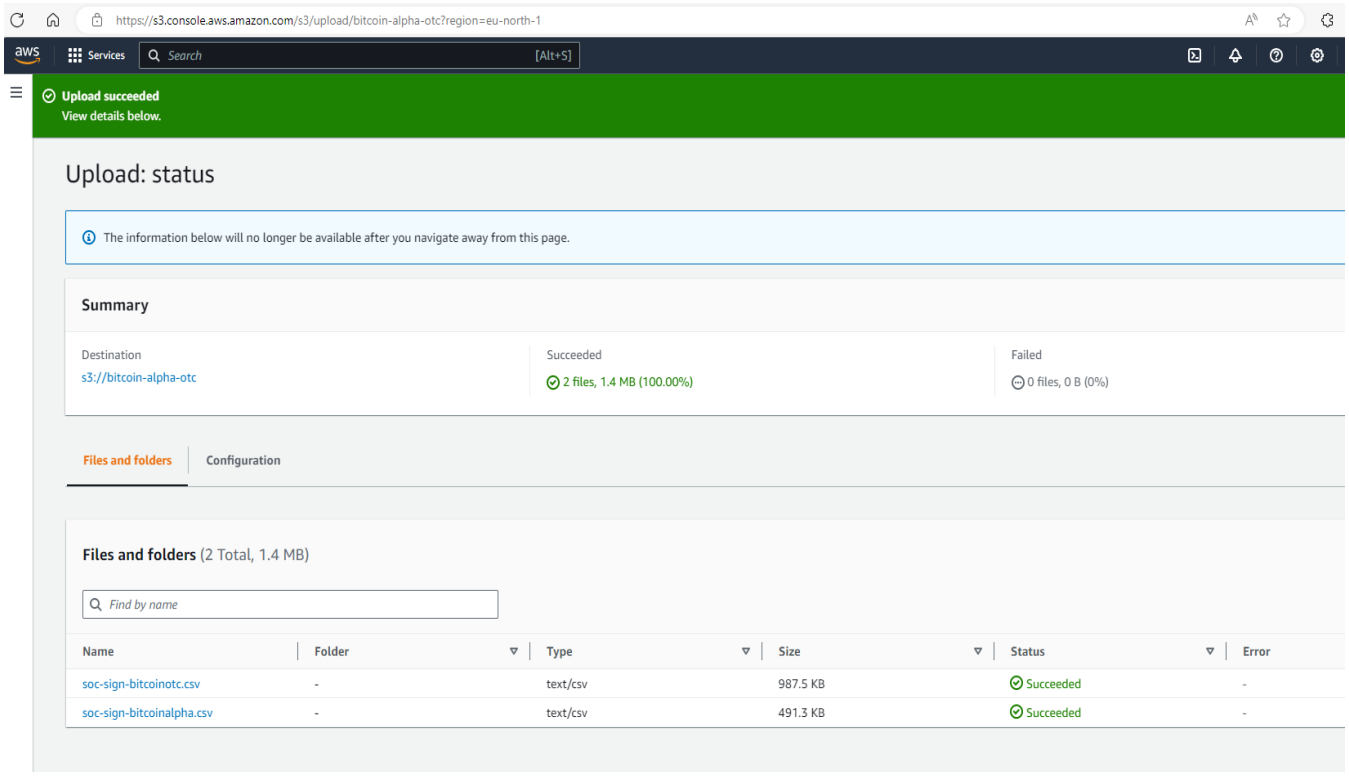
2016 -> 1

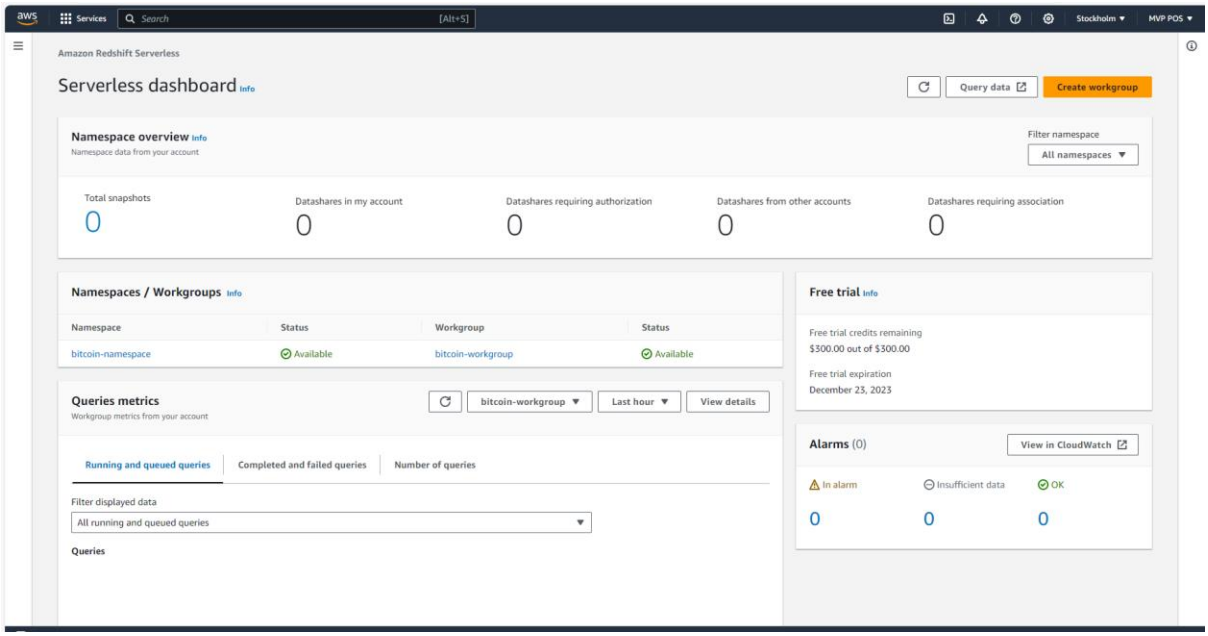
CVE-2011-4447	2011-11-11	wxBitcoin and bitcoind	Exposure[6]	Hard	Wallet non-encryption
CVE-2012-1909	2012-03-07	Bitcoin protocol and all clients	Netsplit[1]	Very hard	Transaction overwriting
CVE-2012-1910	2012-03-17	bitcoind & Bitcoin-Qt for Windows	Unknown[7]	Hard	Non-thread safe MingW exceptions
BIP 0016	2012-04-01	All Bitcoin clients	Fake Conf[8]	Miners[9]	Softfork: P2SH
CVE-2012-2459	2012-05-14	bitcoind and Bitcoin-Qt	Netsplit[1]	Easy	Block hash collision (via merkle root)
CVE-2012-3789	2012-06-20	bitcoind and Bitcoin-Qt	DoS[3]	Easy	(Lack of) orphan txn resource limits
CVE-2012-4682		bitcoind and Bitcoin-Qt	DoS[3]		
CVE-2012-4683	2012-08-23	bitcoind and Bitcoin-Qt	DoS[3]	Easy	Targeted DoS by CPU exhaustion using alerts
CVE-2012-4684	2012-08-24	bitcoind and Bitcoin-Qt	DoS[3]	Easy	Network-wide DoS using malleable signatures in alerts
CVE-2013-2272	2013-01-11	bitcoind and Bitcoin-Qt	Exposure[6]	Easy	Remote discovery of node's wallet addresses
CVE-2013-2273	2013-01-30	bitcoind and Bitcoin-Qt	Exposure[6]	Easy	Predictable change output
CVE-2013-2292	2013-01-30	bitcoind and Bitcoin-Qt	DoS[3]	Hard	A transaction that takes at least 3 minutes to verify
CVE-2013-2293	2013-02-14	bitcoind and Bitcoin-Qt	DoS[3]	Easy	Continuous hard disk seek
CVE-2013-3219	2013-03-11	bitcoind and Bitcoin-Qt 0.8.0	Fake Conf[8]	Miners[9]	Unenforced block protocol rule
CVE-2013-3220	2013-03-11	bitcoind and Bitcoin-Qt	Netsplit[1]	Hard	Inconsistent BDB lock limit interactions
BIP 0034	2013-03-25	All Bitcoin clients	Fake Conf[8]	Miners[9]	Softfork: Height in coinbase
BIP 0050	2013-05-15	All Bitcoin clients	Netsplit[1]	Implicit[2]	Hard fork to remove txid limit protocol rule
CVE-2013-4627	2013-06-??	bitcoind and Bitcoin-Qt	DoS[3]	Easy	Memory exhaustion with excess tx message data
CVE-2013-4165	2013-07-20	bitcoind and Bitcoin-Qt	Theft[10]	Local	Timing leak in RPC authentication
CVE-2013-5700	2013-09-04	bitcoind and Bitcoin-Qt 0.8.x	DoS[3]	Easy	Remote p2p crash via bloom filters
CVE-2014-0160	2014-04-07	Anything using OpenSSL for TLS	Unknown[7]	Easy	Remote memory leak via payment protocol
CVE-2015-3641	2014-07-07	bitcoind and Bitcoin-Qt prior to 0.10.2	DoS[3]	Easy	(Yet) Unspecified DoS
BIP 66	2015-02-13	All Bitcoin clients	Fake Conf[8]	Miners[9]	Softfork: Strict DER signatures
BIP 65	2015-11-12	All Bitcoin clients	Fake Conf[8]	Miners[9]	Softfork: OP_CHECKLOCKTIMEVERIFY
BIPs 68, 112 & 113	2016-04-11	All Bitcoin clients	Fake Conf[8]	Miners[9]	Softforks: Rel locktime, CSV & MTP locktime

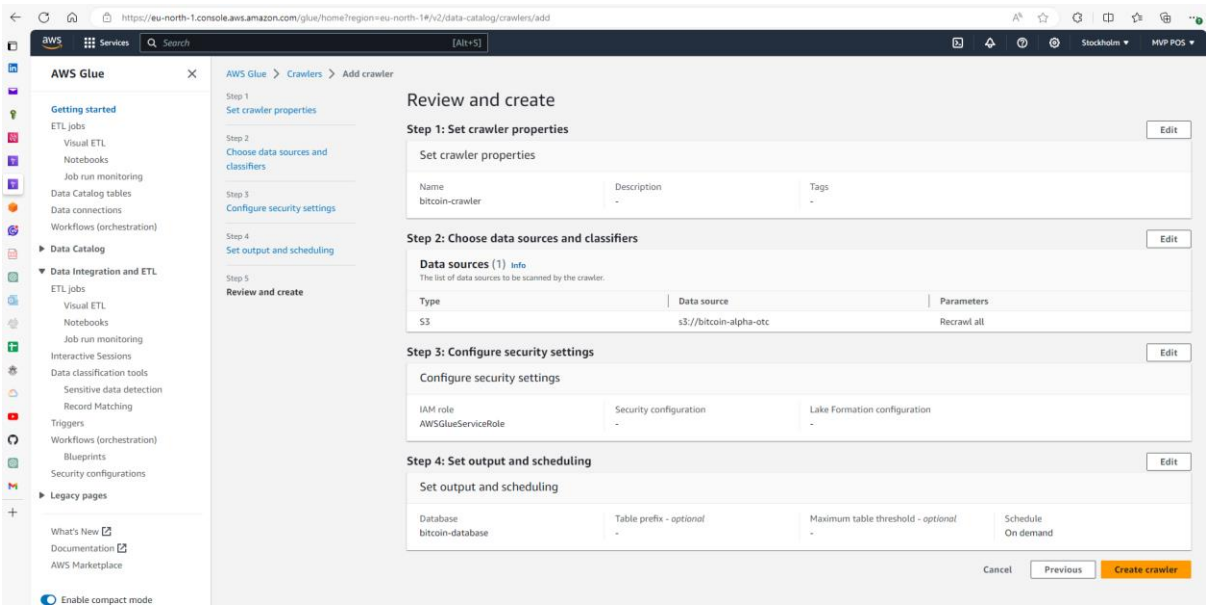
- **PARA OS ANOS MAIS RECENTES. TEMOS OS MESMOS RESULTADOS ?**

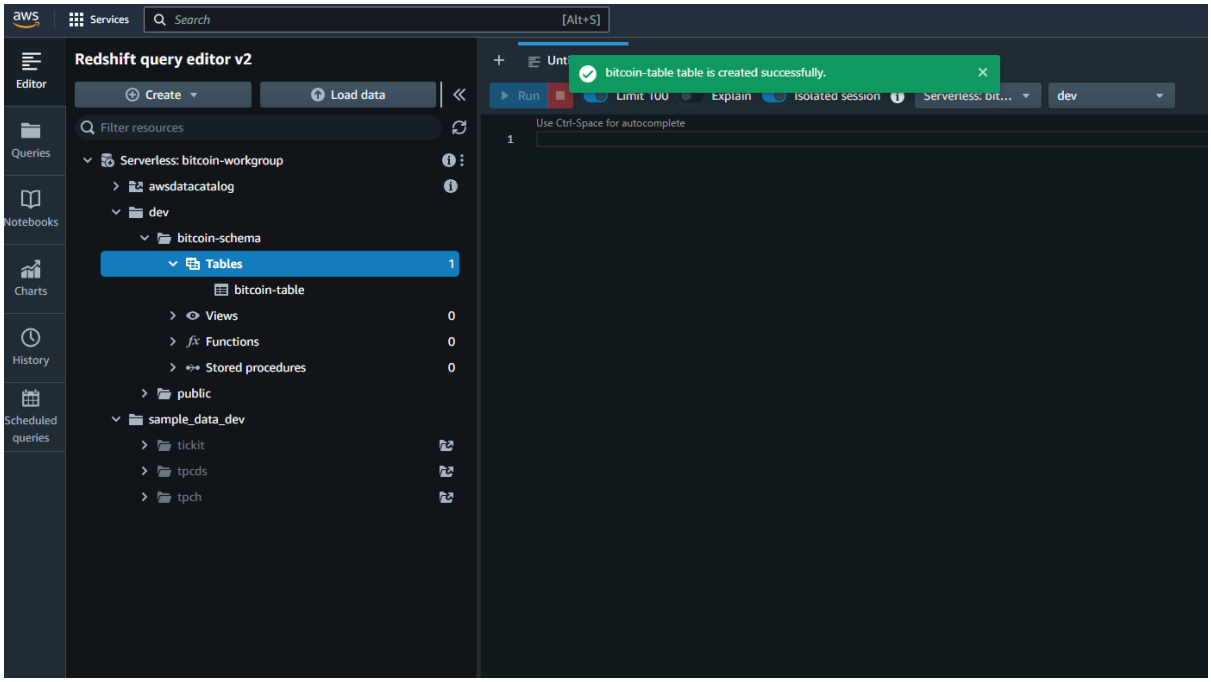
Para este estudo utilizou-se o serviço AWS Glue para realizar o ETL (Extract, Transform and Load). Foram criadas a seguintes etapas :

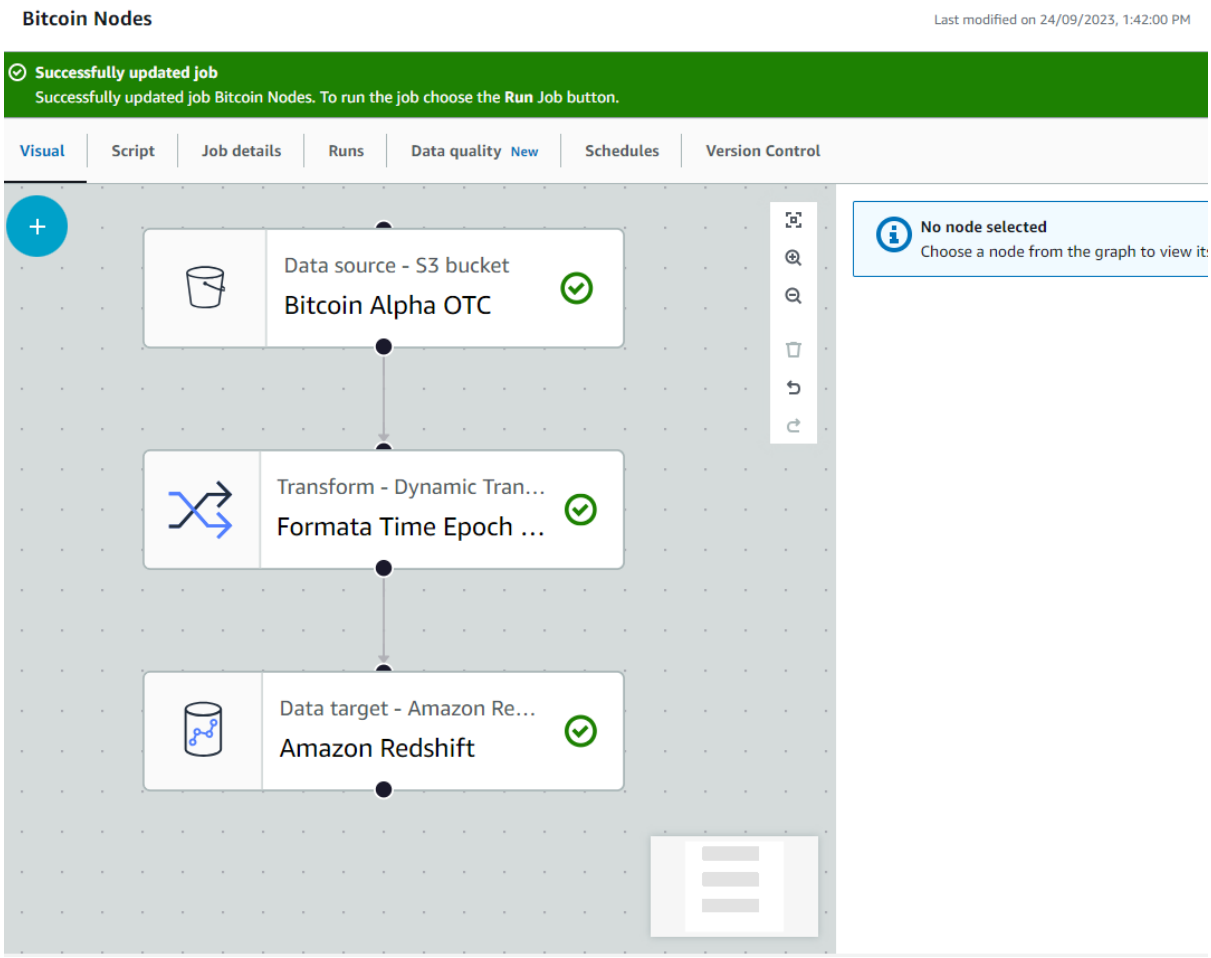
1. Foi criado um bucket S3 chamado bitcoin-alpha-otc (s3://bitcoin-alpha-otc) e os dois arquivos .csv foram carregados ao bucket.

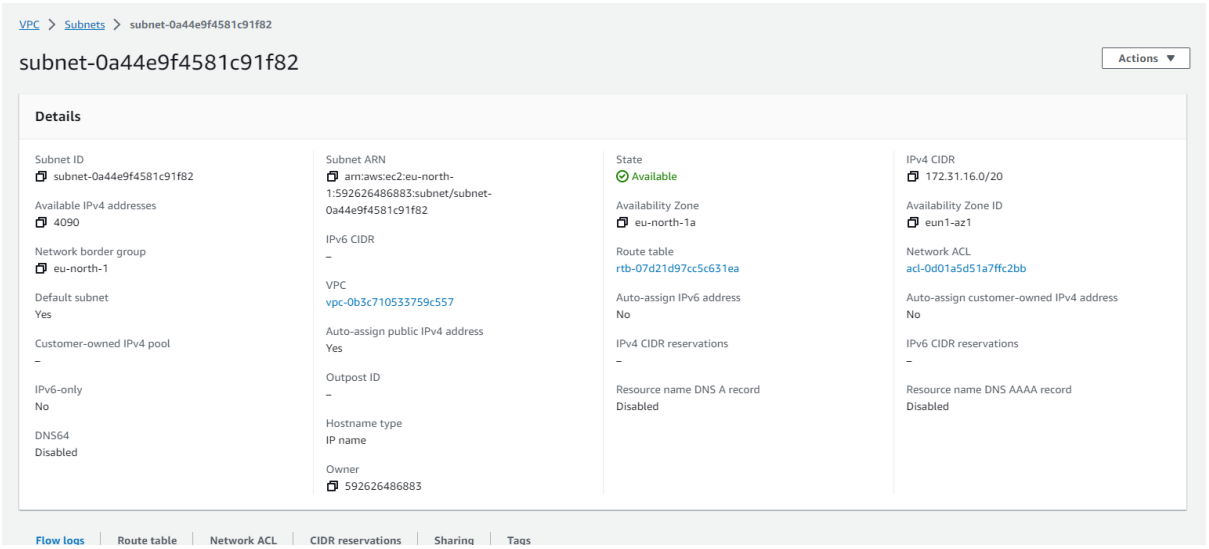



Services
Search
[Alt+S]
Stockholm
MVP POS
Amazon Redshift Serverless
Serverless dashboard Info
Query data
Create workgroup
Namespace overview Info
Namespace data from your account
Filter namespace
All namespaces
Total snapshots
0
Datashares in my account
0
Datashares requiring authorization
0
Datashares from other accounts
0
Datashares requiring association
0
Namespaces / Workgroups Info
Namespace
Status
Workgroup
Status
bitcoin-namespace
Available
bitcoin-workgroup
Available
Free trial Info
Free trial credits remaining
$300.00 out of $300.00
Free trial expiration
December 23, 2023
Queries metrics
Workgroup metrics from your account
bitcoin-workgroup
Last hour
View details
Running and queued queries
Completed and failed queries
Number of queries
Filter displayed data
All running and queued queries
Queries
Alarms (0)
View in CloudWatch
In alarm
Insufficient data
OK
0
0
0


https://eu-north-1.console.aws.amazon.com/glue/home?region=eu-north-1#/v2/data-catalog/crawlers/add
Services
Search
[Alt+S]
Stockholm
MVP POS
AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Data Catalog
Data Integration and ETL
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Interactive Sessions
Data classification tools
Sensitive data detection
Record Matching
Triggers
Workflows (orchestration)
Blueprints
Security configurations
Legacy pages
What's New
Documentation
AWS Marketplace
Enable compact mode
AWS Glue > Crawlers > Add crawler
Step 1
Set crawler properties
Step 2
Choose data sources and classifiers
Step 3
Configure security settings
Step 4
Set output and scheduling
Step 5
Review and create
Review and create
Step 1: Set crawler properties
Edit
Set crawler properties
Name
bitcoin-crawler
Description
-
Tags
-
Step 2: Choose data sources and classifiers
Edit
Data sources (1) Info
The list of data sources to be scanned by the crawler.
Type
Data source
Parameters
S3
s3://bitcoin-alpha-otc
Recrawl all
Step 3: Configure security settings
Edit
Configure security settings
IAM role
AWSGlueServiceRole
Security configuration
-
Lake Formation configuration
-
Step 4: Set output and scheduling
Edit
Set output and scheduling
Database
bitcoin-database
Table prefix - optional
-
Maximum table threshold - optional
-
Schedule
On demand
Cancel
Previous
Create crawler

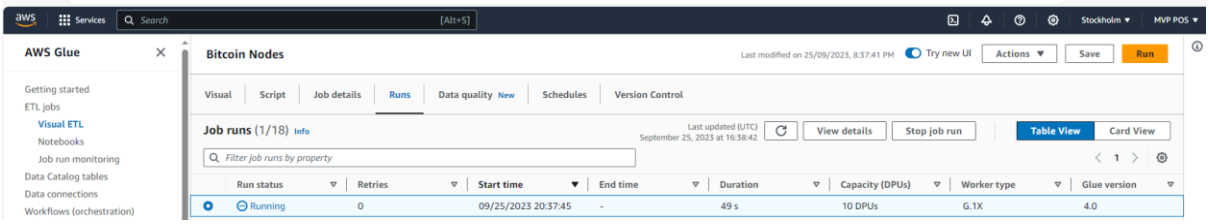

Services
Search
[Alt+S]
Redshift query editor v2
Editor
Queries
Notebooks
Charts
History
Scheduled queries
Create
Load data
Filter resources
Serverless: bitcoin-workgroup
awsdatacatalog
dev
bitcoin-schema
Tables 1
bitcoin-table
Views 0
Functions 0
Stored procedures 0
public
sample_data_dev
tickit
tpcds
tpch
bitcoin-table table is created successfully.
Run
Limit 100
Explain
Isolated session
dev
Use Ctrl-Space for autocomplete
1


Bitcoin Nodes
Last modified on 24/09/2023, 1:42:00 PM
Successfully updated job
Successfully updated job Bitcoin Nodes. To run the job choose the Run Job button.
Visual
Script
Job details
Runs
Data quality New
Schedules
Version Control
Data source - S3 bucket
Bitcoin Alpha OTC
Transform - Dynamic Tran...
Formata Time Epoch ...
Data target - Amazon Re...
Amazon Redshift
No node selected
Choose a node from the graph to view it



AWS Regions Info

Region	Status
Africa (Cape Town)	Disabled
Asia Pacific (Hong Kong)	Disabled
Asia Pacific (Hyderabad)	Disabled
Asia Pacific (Jakarta)	Disabled
Asia Pacific (Melbourne)	Disabled
Europe (Zurich)	Disabled
Europe (Milan)	Disabled
Europe (Spain)	Disabled
Israel (Tel Aviv)	Disabled
Middle East (UAE)	Disabled
Middle East (Bahrain)	Disabled
Asia Pacific (Tokyo)	Enabled by default



CloudWatch continuous logs

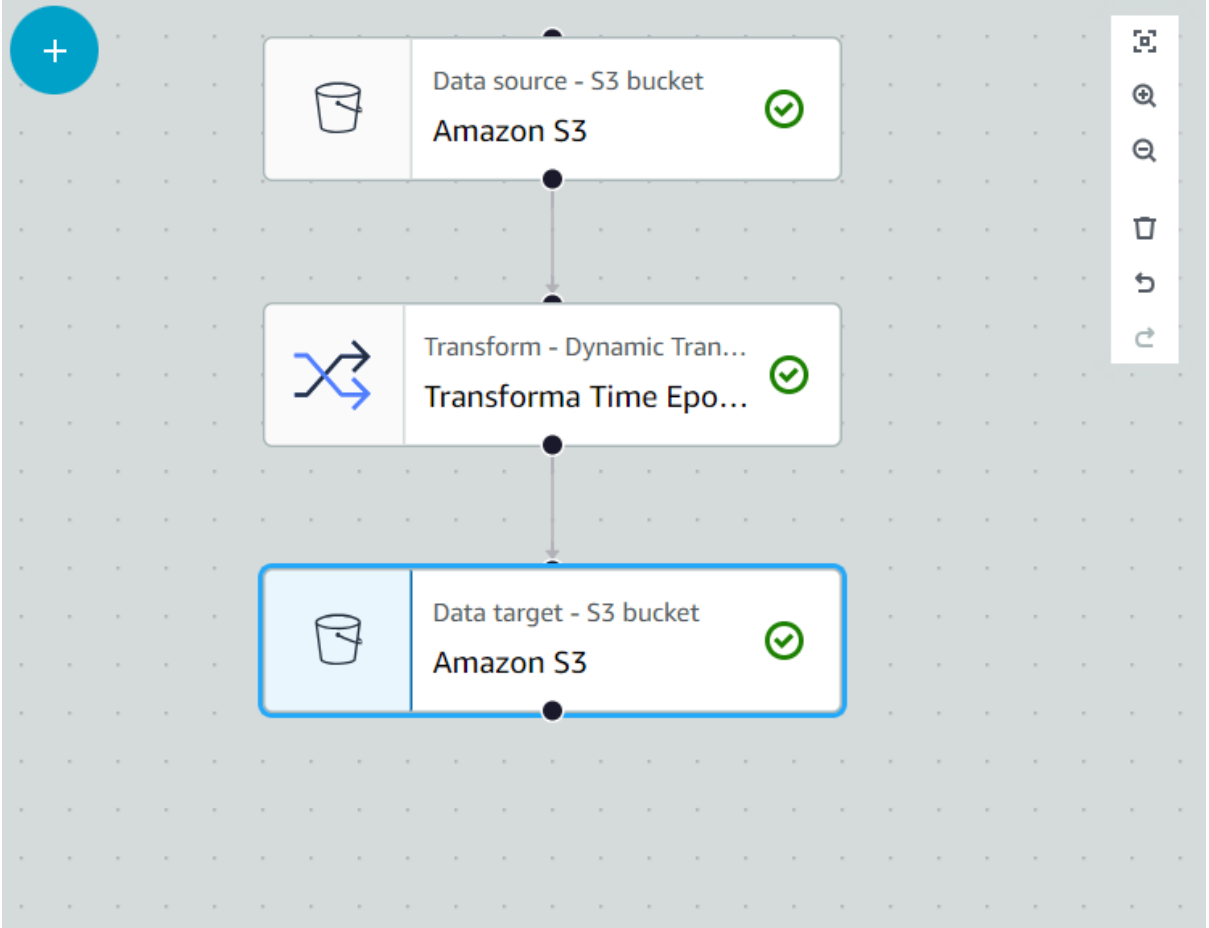
Driver logs

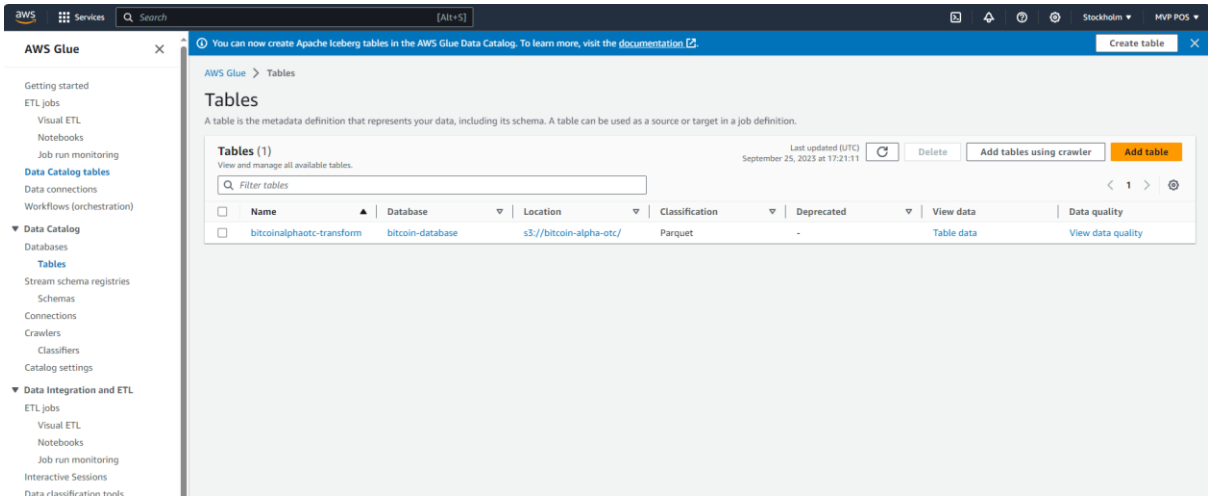
Driver and executor log streams

```
23/09/25 17:20:17 INFO LogPusher: stopping
23/09/25 17:20:17 INFO ProcessLauncher: postprocessing
23/09/25 17:20:17 INFO HadoopDataSink: Succeeded to retrieve created table bitcoinalphaotc-transform in database bitcoin-database after job run with catalogId
23/09/25 17:20:17 INFO LakeformationRetryWrapper$: Lakeformation: API call succeeded
23/09/25 17:20:17 INFO HadoopDataSink: Table bitcoinalphaotc-transform created successfully in database bitcoin-database with catalogId of
23/09/25 17:20:17 INFO HadoopDataSink: UpdateCatalog: creating table with input : {Name: bitcoinalphaotc-transform,StorageDescriptor: {Columns: [{Name: col0,Type: string,}, {Name:
23/09/25 17:20:17 INFO HadoopDataSink: Table bitcoinalphaotc-transform not found in database bitcoin-database with catalogId of , creating a new one.
23/09/25 17:20:16 INFO AWSConnectionUtils$: AWSConnectionUtils: use proxy in glue client configuration. Host: null, Port: -1
23/09/25 17:20:16 INFO BlockManagerInfo: Removed broadcast_6_piece0 on 172.35.142.45:36859 in memory (size: 56.8 KiB, free: 5.8 GiB)
23/09/25 17:20:16 INFO BlockManagerInfo: Removed broadcast_6_piece0 on 172.34.158.214:41875 in memory (size: 56.8 KiB, free: 5.8 GiB)
```


Bitcoin Nodes

Visual | Script | Job details | Runs | Data quality New | Schedules | Version Control





=

09/25/2023 21:19:19

Job name	Id	Run status	Glue version
Bitcoin Nodes	jr_7e76faa7287e6e538091879cacfb0bfba7dfc22c51f1d1709f3a4070c8ba235e	Succeeded	4.0
Retry attempt number	**Start time**	**End time**	**Start-up time**
Initial run	September 25, 2023 9:19:19 PM	September 25, 2023 9:20:28 PM	7 seconds
Execution time	**Last modified on**	**Trigger name**	**Security configuration**
1 minute 2 seconds	September 25, 2023 9:20:28 PM	-	-
Timeout	**Max capacity**	**Number of workers**	**Worker type**
2880 minutes	10 DPUs	10	G.1X