Streaming Finance Data with AWS Lambda

This project simulates the process of working with streaming data by leveraging AWS Lambda functions and yfinance python library. This process consists of four parts:

1. `DataTransformer - a lambda function that gathers the data`
2. `DataCollector - a Kinesis stream that holds the data`
3. `DataAnalyzer - a serverless process to query the S3 data`
4. `DataVisualization - data visualization of the query results`

Historical data for 10 stocks - Amazon (AMZN), Alibaba Group (BABA), Walmart (WMT), eBay (EBAY), Shopify (SHOP), Target (TGT), Best Buy (BBY), The Home Depot (HD), Costco (COST), Kroger (KR) - between Nov Oct 24th, 2022, and Nov 04th, 2022 was collected via a lambda function and streamed to S3 buckets. Then a database was created in Athena using Glue to allow for serverless query of the collected data. The data was then exported to csv to be analysed in Juypter notebook.

Answers to Analysis Questions

1) Graph the average volatility trend per company (A single Line Chart: Each line refers to a company) Which company is the most volatile?

Costco is the most volatile company based on average volatility trend.

2) Graph the daily highest volatility per company (A Grouped Bar Chart: Each group refers to a company and the bars refer to the daily highest volatility) Do the findings from this graph support your conclusion from the first graph?

This graph does support the findings from the first graph. Costco is the most volatile company based on using daily highest volatility.

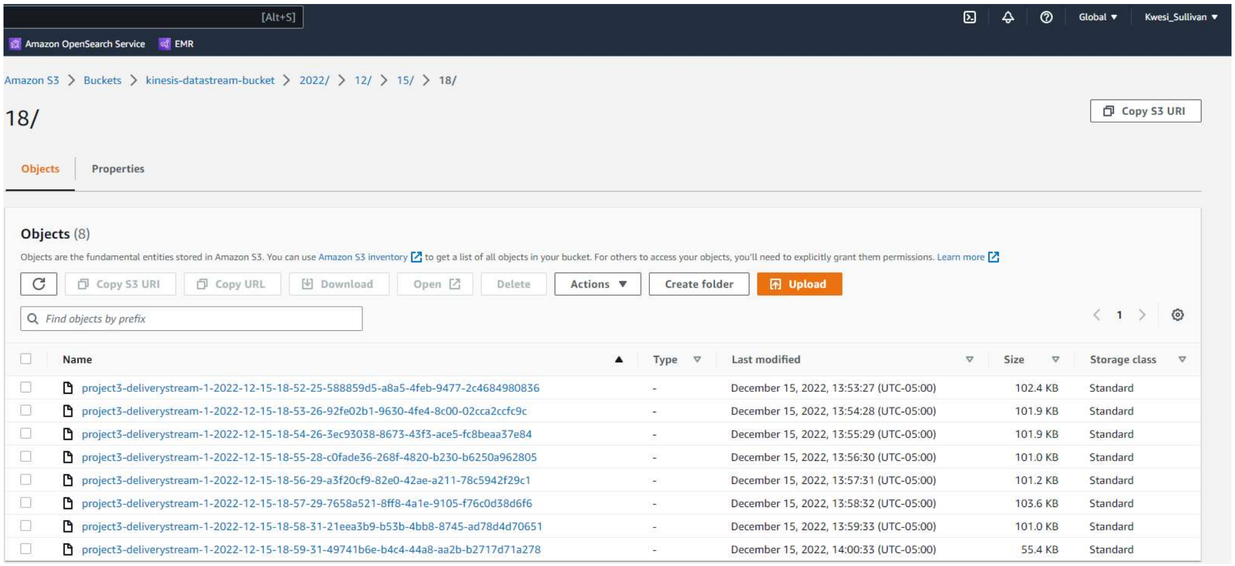
3) Graph the normalized average volatility per company (A Bar Chart: Each bar refers to a company). Normalization allows the comparison of quantities or objects on an appropriate scale. In this case, which company is actually the most volatile?

Amazon is the most volatile company when using normalized volatility.

4) Graph the normalized average volatility on October 24th, 2022 (or Any Date) (A Bar Chart: Each bar refers to a company) Which company is the most volatile on a given day?

On October 24th, 2022, Kroger is the most volatile company.

S3 Bucket



Amazon S3 > Buckets > kinesis-datastream-bucket > 2022/ > 12/ > 15/ > 18/

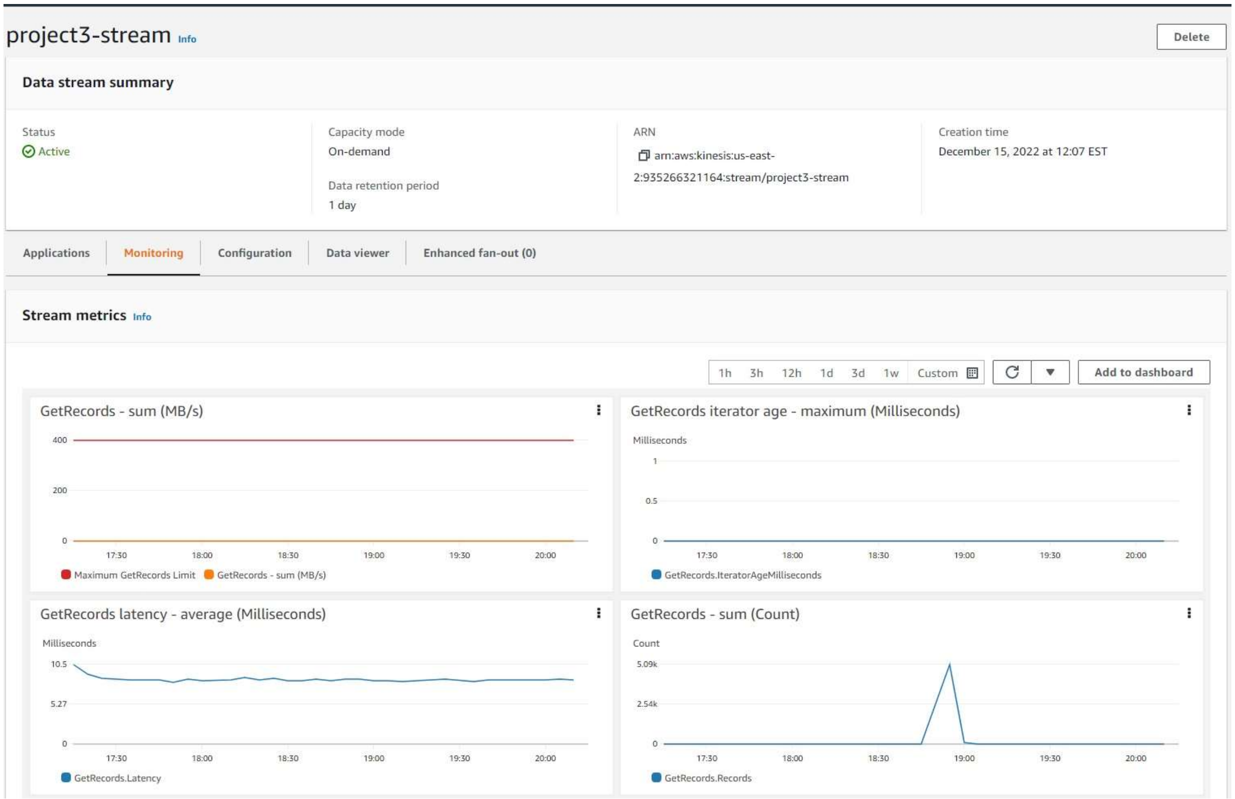
18/

Objects (8)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

□	Name ▲	Type	Last modified	Size	Storage class
□	project3-deliverystream-1-2022-12-15-18-52-25-588859d5-a8a5-4feb-9477-2c4684980836	-	December 15, 2022, 13:53:27 (UTC-05:00)	102.4 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-53-26-92fe02b1-9630-4fe4-8c00-02cca2ccfc9c	-	December 15, 2022, 13:54:28 (UTC-05:00)	101.9 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-54-26-3ec93038-8673-43f3-ace5-fc8beaa37e84	-	December 15, 2022, 13:55:29 (UTC-05:00)	101.9 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-55-28-c0fade36-268f-4820-b230-b6250a962805	-	December 15, 2022, 13:56:30 (UTC-05:00)	101.0 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-56-29-a3f20cf9-82e0-42ae-a211-78c5942f29c1	-	December 15, 2022, 13:57:31 (UTC-05:00)	101.2 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-57-29-7658a521-8ff8-4a1e-9105-f76c0d38d6f6	-	December 15, 2022, 13:58:32 (UTC-05:00)	103.6 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-58-31-21eea3b9-b53b-4bb8-8745-ad78d4d70651	-	December 15, 2022, 13:59:33 (UTC-05:00)	101.0 KB	Standard
□	project3-deliverystream-1-2022-12-15-18-59-31-49741b6e-b4c4-44a8-aa2b-b2717d71a278	-	December 15, 2022, 14:00:33 (UTC-05:00)	55.4 KB	Standard

AWS Kinesis Monitoring



project3-stream Info

Data stream summary

Status: Active

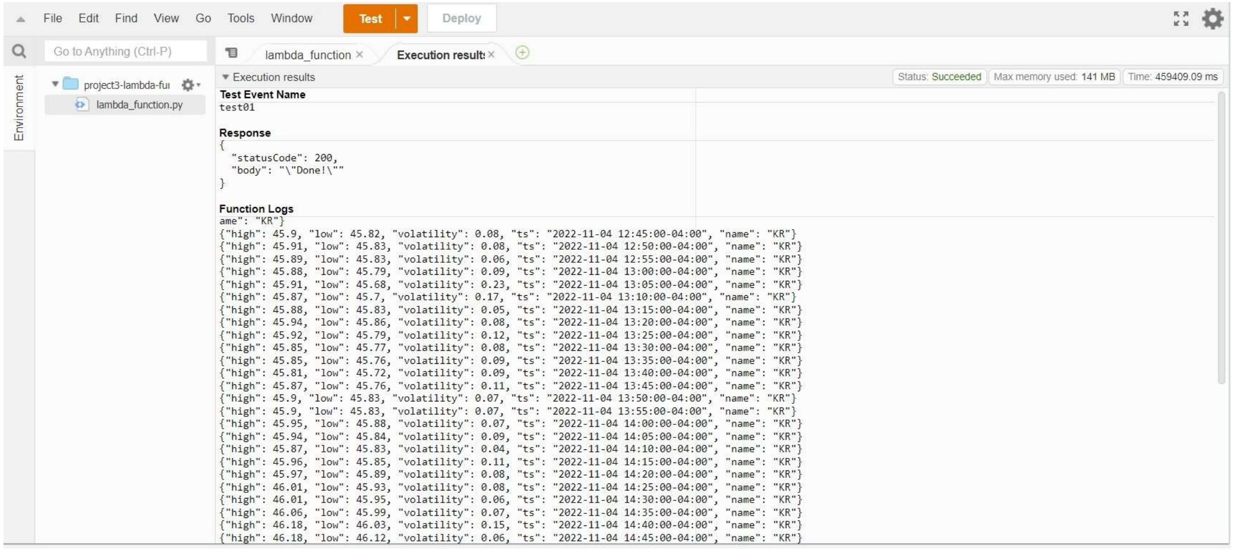
Capacity mode: On-demand

Data retention period: 1 day

ARN: arn:aws:kinesis:us-east-2:935266321164:stream/project3-stream

Creation time: December 15, 2022 at 12:07 EST

Execution Results in AWS Lambda Management Console



Test Event Name

```
test01
```

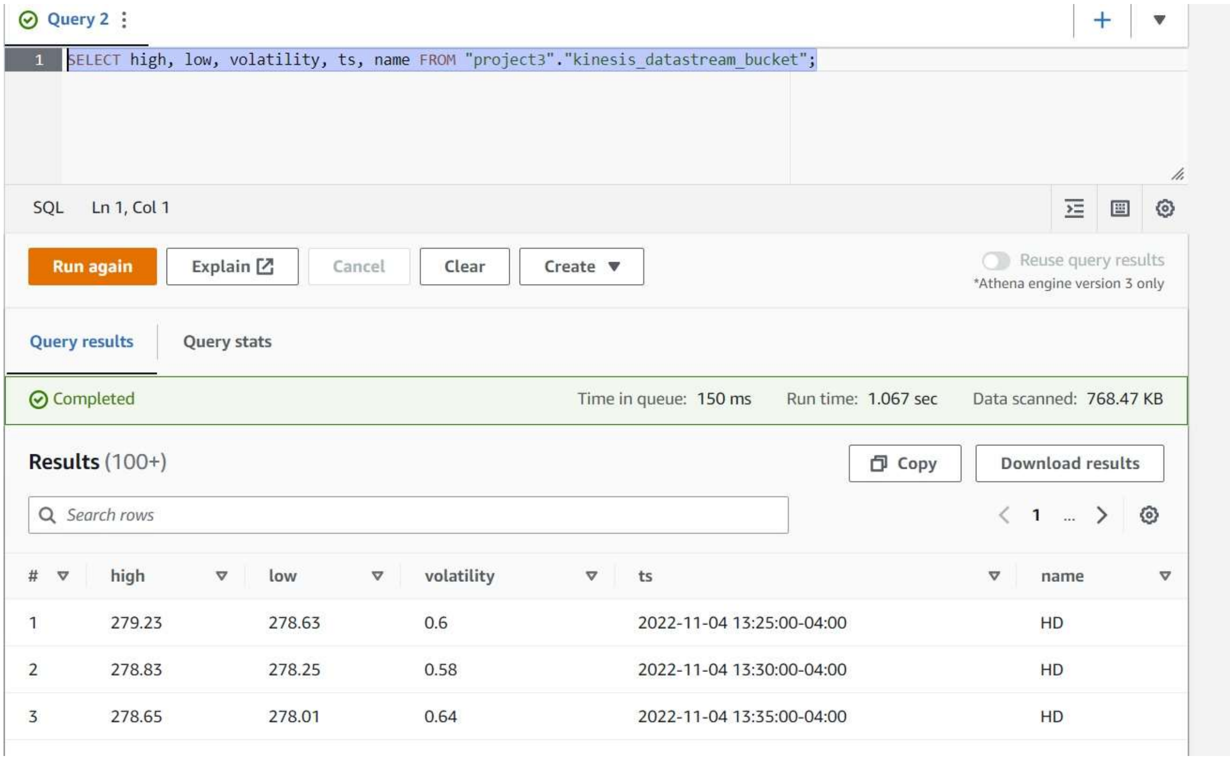
Response

```
{
  "statusCode": 200,
  "body": "\"Done!\""
}
```

Function Logs

```
ame": "KR"}
{"high": 45.9, "low": 45.82, "volatility": 0.08, "ts": "2022-11-04 12:45:00-04:00", "name": "KR"}
{"high": 45.91, "low": 45.83, "volatility": 0.08, "ts": "2022-11-04 12:50:00-04:00", "name": "KR"}
{"high": 45.89, "low": 45.83, "volatility": 0.06, "ts": "2022-11-04 12:55:00-04:00", "name": "KR"}
{"high": 45.88, "low": 45.79, "volatility": 0.09, "ts": "2022-11-04 13:00:00-04:00", "name": "KR"}
{"high": 45.91, "low": 45.68, "volatility": 0.23, "ts": "2022-11-04 13:05:00-04:00", "name": "KR"}
{"high": 45.87, "low": 45.7, "volatility": 0.17, "ts": "2022-11-04 13:10:00-04:00", "name": "KR"}
{"high": 45.88, "low": 45.83, "volatility": 0.05, "ts": "2022-11-04 13:15:00-04:00", "name": "KR"}
{"high": 45.94, "low": 45.86, "volatility": 0.08, "ts": "2022-11-04 13:20:00-04:00", "name": "KR"}
{"high": 45.92, "low": 45.79, "volatility": 0.12, "ts": "2022-11-04 13:25:00-04:00", "name": "KR"}
{"high": 45.85, "low": 45.77, "volatility": 0.08, "ts": "2022-11-04 13:30:00-04:00", "name": "KR"}
{"high": 45.85, "low": 45.76, "volatility": 0.09, "ts": "2022-11-04 13:35:00-04:00", "name": "KR"}
{"high": 45.81, "low": 45.72, "volatility": 0.09, "ts": "2022-11-04 13:40:00-04:00", "name": "KR"}
{"high": 45.87, "low": 45.76, "volatility": 0.11, "ts": "2022-11-04 13:45:00-04:00", "name": "KR"}
{"high": 45.9, "low": 45.83, "volatility": 0.07, "ts": "2022-11-04 13:50:00-04:00", "name": "KR"}
{"high": 45.9, "low": 45.83, "volatility": 0.07, "ts": "2022-11-04 13:55:00-04:00", "name": "KR"}
{"high": 45.95, "low": 45.88, "volatility": 0.07, "ts": "2022-11-04 14:00:00-04:00", "name": "KR"}
{"high": 45.94, "low": 45.84, "volatility": 0.09, "ts": "2022-11-04 14:05:00-04:00", "name": "KR"}
{"high": 45.87, "low": 45.83, "volatility": 0.04, "ts": "2022-11-04 14:10:00-04:00", "name": "KR"}
{"high": 45.96, "low": 45.85, "volatility": 0.11, "ts": "2022-11-04 14:15:00-04:00", "name": "KR"}
{"high": 45.97, "low": 45.89, "volatility": 0.08, "ts": "2022-11-04 14:20:00-04:00", "name": "KR"}
{"high": 46.01, "low": 45.93, "volatility": 0.08, "ts": "2022-11-04 14:25:00-04:00", "name": "KR"}
{"high": 46.01, "low": 45.95, "volatility": 0.06, "ts": "2022-11-04 14:30:00-04:00", "name": "KR"}
{"high": 46.06, "low": 45.99, "volatility": 0.07, "ts": "2022-11-04 14:35:00-04:00", "name": "KR"}
{"high": 46.18, "low": 46.03, "volatility": 0.15, "ts": "2022-11-04 14:40:00-04:00", "name": "KR"}
{"high": 46.18, "low": 46.12, "volatility": 0.06, "ts": "2022-11-04 14:45:00-04:00", "name": "KR"}
```

Athena results



```
SELECT high, low, volatility, ts, name FROM "project3"."kinesis_datastream_bucket";
```

Completed   Time in queue: 150 ms   Run time: 1.067 sec   Data scanned: 768.47 KB

Results (100+)

#	high	low	volatility	ts	name
1	279.23	278.63	0.6	2022-11-04 13:25:00-04:00	HD
2	278.83	278.25	0.58	2022-11-04 13:30:00-04:00	HD
3	278.65	278.01	0.64	2022-11-04 13:35:00-04:00	HD