



PES UNIVERSITY
100 feet Ring Road, BSK 3rd Stage
Bengaluru 560085
Department of Computer Science and Engineering

Department of Computer Science and Engineering
B. Tech. CSE - 6th Semester
Jan – May 2024

UE21CS343BB3
DATABASE TECHNOLOGIES (DBT)

PROJECT REPORT
on

"CryptoStream: Real-time Cryptocurrency Data Analysis".

Submitted by : Team #: 01

Chethan R	PES1UG21CS156	6 C
Chiranth T S	PES1UG21CS159	6 C
Praneeth S	PES1UG21CS437	6 H
Vighnesh V R	PES1UG21CS710	6 L
Vishnu Jalan	PES1UG21CS717	6L

Class of Prof. Vinay Joshi

Table of Contents

Sl. No	Topic	Page No.
1.	Introduction	3
2.	Installation of Software	3-4
3.	Input Data a. Source b. Description	4
4.	Streaming Mode Experiment a. Description b. Windows c. Results	5-7
5.	Batch Mode Experiment a. Description b. Data Size c. Results	7-9
6.	Comparison of Streaming & Batch Modes a. Results and Discussion	11
7.	Conclusion	12
8.	References	12

1.Introduction:

CryptoStream a cutting-edge platform for real-time cryptocurrency data analysis. Leveraging Apache Spark, Kafka, and SQLite,

CryptoStream enables users to access live cryptocurrency data from a leading exchange via WebSocket connections. Through seamless integration with Spark's distributed computing capabilities, users can perform advanced data processing, transformations, and aggregations, gaining valuable insights into cryptocurrency price movements, trading volumes, and market trends.

With CryptoStream, users can stay ahead of the curve in the dynamic world of cryptocurrencies.

2.Installation of Software

Spark installation:

```
vighnesh@vighnesh-VirtualBox:~/Downloads$ spark-shell
24/04/20 18:30:13 WARN Utils: Your hostname, vighnesh-VirtualBox resolves to a loopback address: 127.0.1.1; using 10.0.2.15 instead
(on interface enp0s3)
24/04/20 18:30:13 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
24/04/20 18:30:35 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
applicable
Spark context Web UI available at http://10.0.2.15:4040
Spark context available as 'sc' (master = local[*], app id = local-1713618039059).
Spark session available as 'spark'.
Welcome to

  ____  __
 / ___/ /_
/_  /_ / __ \
 \___/ \___/

version 3.5.1

Using Scala version 2.12.18 (OpenJDK 64-Bit Server VM, Java 1.8.0_402)
Type in expressions to have them evaluated.
Type :help for more information.

scala> █
```

spark running:

```
vighnesh@vighnesh-VirtualBox: ~/Downloads
vighnesh@vighnesh-VirtualBox: ~/Downloads
vighnesh@vighnesh-VirtualBox:~/Downloads$ /home/vighnesh/Downloads/spark/sbin/start-all.sh
starting org.apache.spark.deploy.master.Master, logging to /home/vighnesh/Downloads/spark/logs/spark-vighnesh-org.apache.spark.deplo
y.master.Master-1-vighnesh-VirtualBox.out
localhost: starting org.apache.spark.deploy.worker.Worker, logging to /home/vighnesh/Downloads/spark/logs/spark-vighnesh-org.apache.
spark.deploy.worker.Worker-1-vighnesh-VirtualBox.out
vighnesh@vighnesh-VirtualBox:~/Downloads$ █
```

Kafka installed and running:

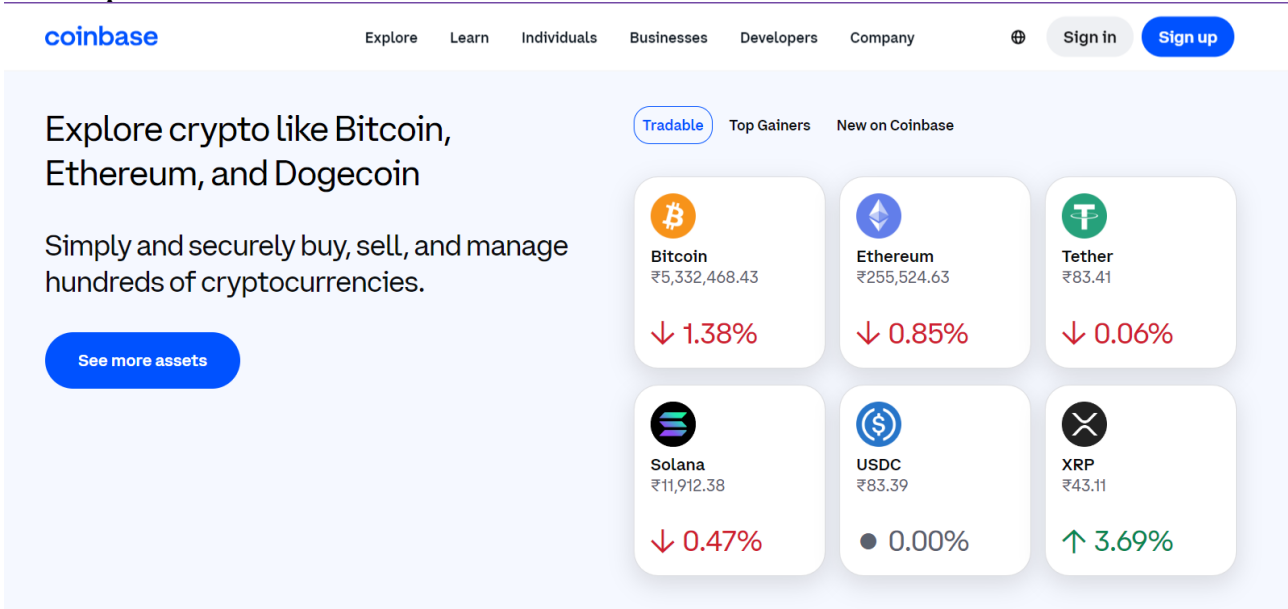
```
vighnesh@vighnesh-VirtualBox: ~/Downloads
vighnesh@vighnesh-VirtualBox: ~/Downloads$ sudo systemctl start kafka
vighnesh@vighnesh-VirtualBox: ~/Downloads$ sudo systemctl status kafka
● kafka.service - Apache Kafka Server
   Loaded: loaded (/etc/systemd/system/kafka.service; disabled; vendor preset: enabled)
   Active: active (running) since Sat 2024-04-20 17:36:43 IST; 1h 2min ago
     Docs: http://kafka.apache.org/documentation.html
   Main PID: 2443 (java)
    Tasks: 71 (limit: 5611)
   Memory: 415.4M
      CPU: 1min 21.764s
   CGroup: /system.slice/kafka.service
           └─2443 /usr/lib/jvm/java-8-openjdk-amd64/bin/java -Xmx1G -Xms1G -server -XX:+UseG1GC -XX:MaxGCPauseMillis=20 -XX:Initia

Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,361] INFO [LogLoader partition=ETH-USD-0, dir>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,376] INFO Created log for partition ETH-USD-0>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,393] INFO [Partition ETH-USD-0 broker=0] No c>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,397] INFO [Partition ETH-USD-0 broker=0] Log >
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,647] INFO Creating topic LTC-USD with configu>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,792] INFO [ReplicaFetcherManager on broker 0]>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,846] INFO [LogLoader partition=LTC-USD-0, dir>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,856] INFO Created log for partition LTC-USD-0>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,859] INFO [Partition LTC-USD-0 broker=0] No c>
Apr 20 17:38:55 vighnesh-VirtualBox kafka-server-start.sh[2443]: [2024-04-20 17:38:55,860] INFO [Partition LTC-USD-0 broker=0] Log >
lines 1-21/21 (END)
```

3.Input Data:

Source:

We are using the Coinbase Pro WebSocket API as the source for our input data.



The input data for our project consists of real-time ticker data for various cryptocurrencies, including Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC). This data is sourced from the Coinbase Pro WebSocket API, which provides a

continuous feed of market updates, including price changes, trading volume, best bid/ask prices, and other relevant metrics.

The Coinbase Pro WebSocket API offers a reliable and efficient way to access cryptocurrency market data in real time, making it an ideal data source for our project. By subscribing to specific channels and product IDs, we can tailor the data stream to our requirements and capture relevant information for analysis and visualization.

4.Streaming Mode Experiment

In the streaming mode experiment, we aimed to analyze real-time cryptocurrency market data using Apache Spark Streaming, Kafka, and SQLite. We set up a WebSocket connection to the Coinbase Pro API to receive live ticker data for Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC). This data was then ingested into Kafka topics using a Kafka producer. Meanwhile, Spark Streaming was utilized to consume the data from Kafka topics, perform various transformations and aggregations, and finally, store the processed data in a SQLite database.

Streaming data using spark and kafka:

```
vighnesh@vighnesh-VirtualBox:~/Downloads$ streamlit run test22.py

You can now view your Streamlit app in your browser.

Local URL: http://localhost:8501
Network URL: http://10.0.2.15:8501

24/04/20 18:52:27 WARN Utils: Your hostname, vighnesh-VirtualBox resolves to a loopback address: 127.0.1.1; using 10.0.2.15 instead
(on interface enp0s3)
24/04/20 18:52:27 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
24/04/20 18:52:30 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
applicable
24/04/20 18:52:34 WARN Utils: Service 'SparkUI' could not bind on port 4040. Attempting port 4041.
The socket is open
```

1. **Data Ingestion Window:** This window represents the duration in which the WebSocket connection is open, and data is continuously received from the Coinbase Pro API.

test22 - Streamlit

localhost:8501

90%

RUNNING... Stop Deploy

WebSocket Data Display

Connecting to WebSocket...

Price: 82.82 Product ID: LTC-USD Time: 2024-04-20T13:22:36.943196Z 24h Volume: 126487.10669629 Best Bid: 82.79 Best Ask: 82.82

Price: 3061.59 Product ID: ETH-USD Time: 2024-04-20T13:22:38.641979Z 24h Volume: 54280.01104255 Best Bid: 3061.38 Best Ask: 3061.59

Price: 63828.84 Product ID: BTC-USD Time: 2024-04-20T13:22:38.981783Z 24h Volume: 11196.83318878 Best Bid: 63828.65 Best Ask: 63828.84

Price: 63828.88 Product ID: BTC-USD Time: 2024-04-20T13:22:39.536801Z 24h Volume: 11196.83326555 Best Bid: 63828.87 Best Ask: 63828.88

Price: 3061.66 Product ID: ETH-USD Time: 2024-04-20T13:22:39.789826Z 24h Volume: 54280.02590359 Best Bid: 3061.37 Best Ask: 3061.67

Price: 3061.67 Product ID: ETH-USD Time: 2024-04-20T13:22:39.789826Z 24h Volume: 54280.04241235 Best Bid: 3061.37 Best Ask: 3061.67

Price: 63828.88 Product ID: BTC-USD Time: 2024-04-20T13:22:39.842278Z 24h Volume: 11196.83339980 Best Bid: 63828.87 Best Ask: 63828.88

Price: 82.8 Product ID: LTC-USD Time: 2024-04-20T13:22:40.302862Z 24h Volume: 126487.33926507 Best Bid: 82.79 Best Ask: 82.81

Processing Window: Spark Streaming processes the incoming data in micro-batches, typically with a window size of a few seconds, to perform transformations and aggregations.

Kafka topic1 output Bitcoin (BTC)

```
vighnesh@vighnesh-VirtualBox:~/Downloads$ python3 test23.py
{"type": "ticker", "sequence": 78858178254, "product_id": "BTC-USD", "price": "63834.64", "open_24h": "64964.88", "volume_24h": "11207.37307376", "low_24h": "62964.25", "high_24h": "65034.41", "volume_30d": "448271.98385499", "best_bid": "63834.63", "best_bid_size": "0.00018003", "best_ask": "63834.64", "best_ask_size": "0.26660894", "side": "buy", "time": "2024-04-20T13:24:20.653777Z", "trade_id": "634407067", "last_size": "0.00251561"}
{"type": "ticker", "sequence": 78858178320, "product_id": "BTC-USD", "price": "63834.64", "open_24h": "64964.88", "volume_24h": "11207.37358008", "low_24h": "62964.25", "high_24h": "65034.41", "volume_30d": "448271.98436131", "best_bid": "63834.63", "best_bid_size": "0.00018003", "best_ask": "63834.64", "best_ask_size": "0.26610262", "side": "buy", "time": "2024-04-20T13:24:21.269945Z", "trade_id": "634407068", "last_size": "0.00050632"}
{"type": "ticker", "sequence": 78858178387, "product_id": "BTC-USD", "price": "63834.63", "open_24h": "64964.88", "volume_24h": "11207.37362950", "low_24h": "62964.25", "high_24h": "65034.41", "volume_30d": "448271.98441073", "best_bid": "63834.63", "best_bid_size": "0.00013061", "best_ask": "63834.64", "best_ask_size": "0.26610262", "side": "sell", "time": "2024-04-20T13:24:21.415422Z", "trade_id": "634407069", "last_size": "0.00004942"}
{"type": "ticker", "sequence": 78858178496, "product_id": "BTC-USD", "price": "63834.64", "open_24h": "64964.88", "volume_24h": "11207.37440482", "low_24h": "62964.25", "high_24h": "65034.41", "volume_30d": "448271.98518605", "best_bid": "63834.63", "best_bid_size": "0.00013061", "best_ask": "63834.64", "best_ask_size": "0.12380899", "side": "buy", "time": "2024-04-20T13:24:21.917376Z", "trade_id": "634407070", "last_size": "0.00077532"}
{"type": "ticker", "sequence": 78858178504, "product_id": "BTC-USD", "price": "63834.63", "open_24h": "64964.88", "volume_24h": "11207.37453543", "low_24h": "62964.25", "high_24h": "65034.41", "volume_30d": "448271.98531666", "best_bid": "63832.00", "best_bid_size": "0.00017458", "best_ask": "63834.64", "best_ask_size": "0.12380899", "side": "sell", "time": "2024-04-20T13:24:22.050027Z", "trade_id": "634407071", "last_size": "0.00013061"}
```


Kafka topic2 output Ethereum (ETH),

```
vighnesh@vighnesh-VirtualBox:~/Downloads$ python3 test24.py
{"type": "ticker", "sequence": 59333158729, "product_id": "ETH-USD", "price": "3061.72", "open_24h": "3100.12", "volume_24h": "54306.83245172", "low_24h": "3019.11", "high_24h": "3108", "volume_30d": "3054285.60656624", "best_bid": "3061.52", "best_bid_size": "0.19597500", "best_ask": "3061.72", "best_ask_size": "1.15466673", "side": "buy", "time": "2024-04-20T13:24:34.935620Z", "trade_id": "516353350", "last_size": "0.01544612"}
{"type": "ticker", "sequence": 59333159370, "product_id": "ETH-USD", "price": "3061.53", "open_24h": "3100.12", "volume_24h": "54307.10124887", "low_24h": "3019.11", "high_24h": "3108", "volume_30d": "3054285.87536339", "best_bid": "3061.40", "best_bid_size": "0.04355051", "best_ask": "3061.53", "best_ask_size": "0.58264492", "side": "buy", "time": "2024-04-20T13:24:37.529839Z", "trade_id": "516353351", "last_size": "0.26879715"}
{"type": "ticker", "sequence": 59333159682, "product_id": "ETH-USD", "price": "3061.62", "open_24h": "3100.12", "volume_24h": "54307.11717030", "low_24h": "3019.11", "high_24h": "3108", "volume_30d": "3054285.89128482", "best_bid": "3061.44", "best_bid_size": "0.37658475", "best_ask": "3061.62", "best_ask_size": "0.84006970", "side": "buy", "time": "2024-04-20T13:24:39.141945Z", "trade_id": "516353352", "last_size": "0.01592143"}
{"type": "ticker", "sequence": 59333160327, "product_id": "ETH-USD", "price": "3061.41", "open_24h": "3100.12", "volume_24h": "54307.14918075", "low_24h": "3019.11", "high_24h": "3108", "volume_30d": "3054285.92329527", "best_bid": "3061.40", "best_bid_size": "0.04355051", "best_ask": "3061.41", "best_ask_size": "0.79496145", "side": "buy", "time": "2024-04-20T13:24:40.176654Z", "trade_id": "516353353", "last_size": "0.03201045"}
{"type": "ticker", "sequence": 59333160460, "product_id": "ETH-USD", "price": "3061.47", "open_24h": "3100.12", "volume_24h": "54307.14929249", "low_24h": "3019.11", "high_24h": "3108", "volume_30d": "3054285.92340701", "best_bid": "3061.40", "best_bid_size": "0.04355051", "best_ask": "3061.47", "best_ask_size": "0.81649021", "side": "buy", "time": "2024-04-20T13:24:40.834554Z", "trade_id": "516353354", "last_size": "0.00011174"}
```

Kafka topic3 output Litecoin (LTC)

```
vighnesh@vighnesh-VirtualBox:~/Downloads$ python3 test25.py
{"type": "ticker", "sequence": 16406155247, "product_id": "LTC-USD", "price": "82.81", "open_24h": "81.3", "volume_24h": "126684.10048751", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072087.61353963", "best_bid": "82.79", "best_bid_size": "0.80280000", "best_ask": "82.81", "best_ask_size": "40.83934664", "side": "buy", "time": "2024-04-20T13:24:45.190094Z", "trade_id": "128959497", "last_size": "0.17669605"}
{"type": "ticker", "sequence": 16406155952, "product_id": "LTC-USD", "price": "82.84", "open_24h": "81.3", "volume_24h": "126685.01705727", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072088.53010939", "best_bid": "82.82", "best_bid_size": "17.63111624", "best_ask": "82.84", "best_ask_size": "23.36563018", "side": "buy", "time": "2024-04-20T13:24:57.206266Z", "trade_id": "128959498", "last_size": "0.91656976"}
{"type": "ticker", "sequence": 16406155965, "product_id": "LTC-USD", "price": "82.82", "open_24h": "81.3", "volume_24h": "126689.44752867", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072092.96058079", "best_bid": "82.82", "best_bid_size": "13.20064484", "best_ask": "82.84", "best_ask_size": "23.36563018", "side": "sell", "time": "2024-04-20T13:24:57.898082Z", "trade_id": "128959499", "last_size": "4.4304714"}
{"type": "ticker", "sequence": 16406155977, "product_id": "LTC-USD", "price": "82.84", "open_24h": "81.3", "volume_24h": "126690.65463288", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072094.16768500", "best_bid": "82.82", "best_bid_size": "13.20064484", "best_ask": "82.84", "best_ask_size": "22.15852597", "side": "buy", "time": "2024-04-20T13:24:58.169543Z", "trade_id": "128959500", "last_size": "1.20710421"}
{"type": "ticker", "sequence": 16406156110, "product_id": "LTC-USD", "price": "82.84", "open_24h": "81.3", "volume_24h": "126692.15764892", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072095.67070104", "best_bid": "82.82", "best_bid_size": "13.20064484", "best_ask": "82.84", "best_ask_size": "6.58032960", "side": "buy", "time": "2024-04-20T13:25:01.638548Z", "trade_id": "128959501", "last_size": "1.50301604"}
{"type": "ticker", "sequence": 16406156112, "product_id": "LTC-USD", "price": "82.84", "open_24h": "81.3", "volume_24h": "126693.41307910", "low_24h": "79.67", "high_24h": "83.13", "volume_30d": "11072096.92613122", "best_bid": "82.82", "best_bid_size": "13.20064484", "best_ask": "82.84", "best_ask_size": "5.32489942", "side": "buy", "time": "2024-04-20T13:25:01.638548Z", "trade_id": "128959502", "last_size": "1.25543018"}
```

Storage Window: Once processed, the data is stored in the SQLite database, allowing for retrieval and analysis at a later time.

 **SQLite Viewer**
view sqlite file online


Drop file here to load content or click on this box to open file dialog.

sqlite_sequence (1 rows) Export

Execute

name	seq
ticker_data	2605

1 / 1

 **SQLite Viewer**
view sqlite file online

Drop file here to load content or click on this box to open file dialog.

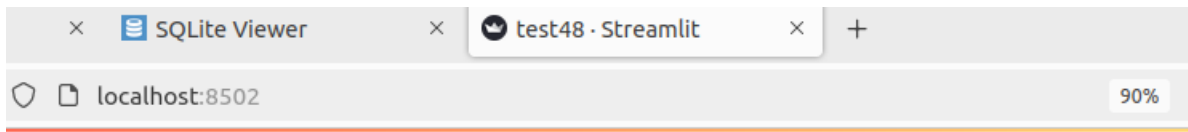
ticker_data (2605 rows) Export ▾

SELECT * FROM 'ticker_data' LIMIT 0,30 Execute

id	price	product_id	time	volume_24h	best_bid	best_ask
1	82.75	LTC-USD	2024-04-20T06:48:39.784001Z	139513.08377774	82.74	82.76
2	3072	ETH-USD	2024-04-20T06:48:40.933252Z	88809.07295661	3071.78	3072
3	64261.23	BTC-USD	2024-04-20T06:48:42.876443Z	15771.65364779	64259.18	64261.23
4	82.76	LTC-USD	2024-04-20T06:48:44.044406Z	139514.52164628	82.75	82.76
5	64261.26	BTC-USD	2024-04-20T06:48:44.222478Z	15771.65591261	64258.87	64261.26
6	64258.26	BTC-USD	2024-04-20T06:48:45.373394Z	15771.6614079	64258.26	64261.25
7	64258.77	BTC-USD	2024-04-20T06:48:47.126288Z	15771.66596197	64255.92	64258.77
8	3071.58	ETH-USD	2024-04-20T06:48:48.094401Z	88809.09652026	3071.58	3071.77
9	64258.8	BTC-USD	2024-04-20T06:48:48.407653Z	15771.66723332	64256.42	64258.8
10	3071.63	ETH-USD	2024-04-20T06:48:48.607626Z	88809.11172026	3071.63	3071.76
11	64256.93	BTC-USD	2024-04-20T06:48:49.607626Z	15771.74223332	64256.92	64258.79

5. Batch Mode Experiment

In the batch mode experiment, we conducted offline analysis of historical cryptocurrency market data using Apache Spark Batch Processing. We retrieved a large dataset containing historical price, volume, and other relevant metrics for Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC) from a reliable cryptocurrency exchange. The dataset covered a specific time period, allowing us to perform batch processing on a static snapshot of the data.



batch mode analysis

Serial No	Min Price	Max Price	Avg Price
1 to 5	82.75	64261.26	26352.0
2 to 6	82.76	64261.26	39187.102
3 to 7	82.76	64261.26	51424.456
4 to 8	82.76	64261.26	39186.526
5 to 9	3071.58	64261.26	52021.734
6 to 10	3071.58	64258.8	39783.808
7 to 11	3071.58	64258.8	39783.542
8 to 12	3071.58	64258.8	39783.172
9 to 13	3071.63	64258.8	52020.200000000004
10 to 14	3071.63	64256.93	52019.623999999996
11 to 15	64255.0	64256.93	64256.297999999995
12 to 16	64255.0	64256.92	64255.912
13 to 17	82.76	64256.72	51421.08

here we are using 5 window size and sliding window mechanism

test48 - Streamlit			
localhost:8502			
2589 to 2593	63852.19	63852.19	63852.19
2590 to 2594	63852.19	63852.19	63852.19
2591 to 2595	63852.01	63852.19	63852.154
2592 to 2596	63852.01	63857.29	63853.174
2593 to 2597	63852.01	63857.29	63854.192
2594 to 2598	63852.01	63857.29	63855.21000000001
2595 to 2599	63852.01	63857.29	63856.23
2596 to 2600	63857.28	63857.29	63857.286
2597 to 2601	3062.4	63857.29	51698.308000000005
2598 to 2602	3062.4	63857.29	39539.332
2599 to 2603	3062.38	63857.29	27380.352000000003
2600 to 2604	3062.38	63857.29	15221.37
2601 to 2605	3062.38	3062.47	3062.406

6.Comparison of Streaming & Batch Modes

SQLite Viewer

https://inloop.github.io/sqlite-viewer/

SELECT * FROM 'ticker_data' LIMIT 0,30

id	price	product_id	time	volume
1	82.75	LTC-USD	2024-04-20T06:48:39.78...	139
2	3072	ETH-USD	2024-04-20T06:48:40.93...	888
3	64261.23	BTC-USD	2024-04-20T06:48:42.87...	157
4	82.76	LTC-USD	2024-04-20T06:48:44.04...	139
5	64261.26	BTC-USD	2024-04-20T06:48:44.22...	157
6	64258.26	BTC-USD	2024-04-20T06:48:45.37...	157
7	64258.77	BTC-USD	2024-04-20T06:48:47.12...	157
8	3071.58	ETH-USD	2024-04-20T06:48:48.09...	888
9	64258.8	BTC-USD	2024-04-20T06:48:48.48...	157
10	3071.63	ETH-USD	2024-04-20T06:48:48.66...	888
11	64256.93	BTC-USD	2024-04-20T06:48:49.60...	157
12	64256.92	BTC-USD	2024-04-20T06:48:49.60...	157

test48 - Streamlit

localhost:8502

batch mode analysis

Serial No	Min Price	Max Price	Avg Price
1 to 5	82.75	64261.26	26352.0
2 to 6	82.76	64261.26	39187.102
3 to 7	82.76	64261.26	51424.456
4 to 8	82.76	64261.26	39186.526
5 to 9	3071.58	64261.26	52021.734
6 to 10	3071.58	64258.8	39783.808
7 to 11	3071.58	64258.8	39783.542
8 to 12	3071.58	64258.8	39783.172

SQLite Viewer

https://inloop.github.io/sqlite-viewer/

2592	63852.19	BTC-USD	2024-04-20T13:25:12.69...	1
2593	63852.19	BTC-USD	2024-04-20T13:25:12.69...	1
2594	63852.19	BTC-USD	2024-04-20T13:25:12.69...	1
2595	63852.01	BTC-USD	2024-04-20T13:25:12.69...	1
2596	63857.29	BTC-USD	2024-04-20T13:25:12.93...	1
2597	63857.28	BTC-USD	2024-04-20T13:25:13.73...	1
2598	63857.28	BTC-USD	2024-04-20T13:25:13.77...	1
2599	63857.29	BTC-USD	2024-04-20T13:25:13.79...	1
2600	63857.29	BTC-USD	2024-04-20T13:25:13.96...	1
2601	3062.4	ETH-USD	2024-04-20T13:25:14.11...	5
2602	3062.4	ETH-USD	2024-04-20T13:25:14.39...	5
2603	3062.38	ETH-USD	2024-04-20T13:25:14.85...	5
2604	3062.38	ETH-USD	2024-04-20T13:25:14.85...	5
2605	3062.47	ETH-USD	2024-04-20T13:25:14.85...	5

test48 - Streamlit

localhost:8502

2591 to 2595	63852.01	63852.19	63852.154
2592 to 2596	63852.01	63857.29	63853.174
2593 to 2597	63852.01	63857.29	63854.192
2594 to 2598	63852.01	63857.29	63855.21000000001
2595 to 2599	63852.01	63857.29	63856.23
2596 to 2600	63857.28	63857.29	63857.286
2597 to 2601	3062.4	63857.29	51698.308000000001
2598 to 2602	3062.4	63857.29	39539.332
2599 to 2603	3062.38	63857.29	27380.352000000001
2600 to 2604	3062.38	63857.29	15221.37
2601 to 2605	3062.38	3062.47	3062.406

7. Conclusion:

1. Real-time vs. Offline Analysis:

- **Streaming Mode:** Provides real-time insights into cryptocurrency market dynamics, allowing for immediate response to changing market conditions and opportunities.
- **Batch Mode:** Conducts offline analysis on historical data snapshots, enabling in-depth exploration of long-term trends, patterns, and anomalies in the cryptocurrency market.

2. Data Freshness:

- **Streaming Mode:** Offers up-to-date information by processing data as it arrives, ensuring the freshest insights but may be subject to processing delays or latency.
- **Batch Mode:** Analyzes historical data in bulk, providing a comprehensive view of past market behavior but lacks immediacy in responding to current market events.

3. Resource Utilization:

- **Streaming Mode:** Requires continuous processing resources to handle incoming data streams in real-time, which may result in higher resource utilization and operational costs.
- **Batch Mode:** Utilizes resources intermittently for scheduled batch processing tasks, potentially resulting in lower overall resource consumption and cost-effectiveness.

8. References:

- Spark twitter streaming online course