



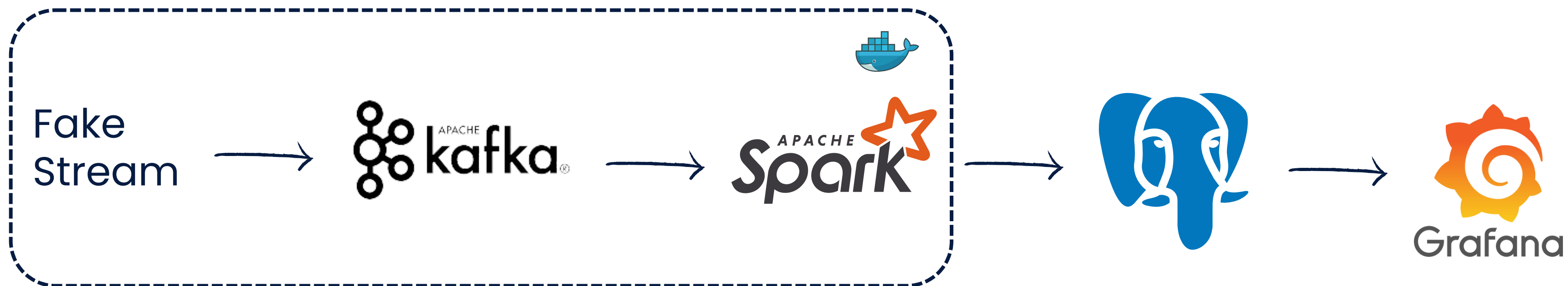
Social Media Analysis

Streaming processing using fake stream
data, Kafka, Spark, Postgresql, and Grafana

Farah Shadra



Stack



Ingestion

Processing

Storage

Dataviz

Ingestion

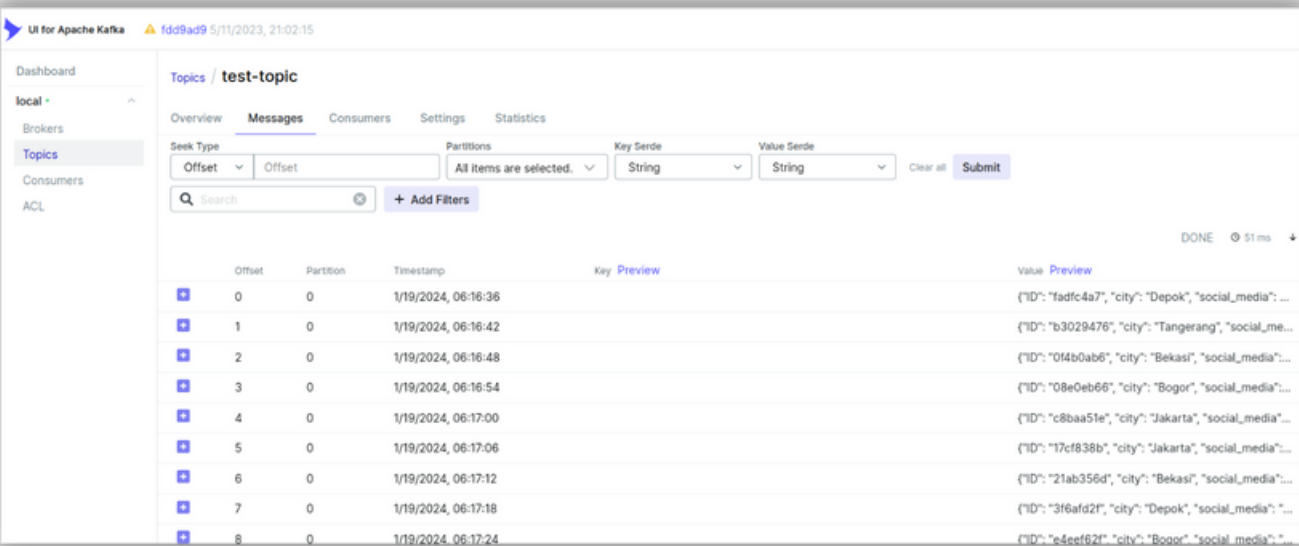
Fake Stream

| ID | Kota | Sosial_Media | Usage_c ount | Sales | Ts |
|-----------|-----------|--------------|-----------------|-------|----------------|
| 1dd1 2 | Jakarta | X | 800 | 45 | 1705323 357 |
| 1dd1 3 | Tangerang | Instagram | 200 | 123 | 1705323 357 |
| 1dd1 3 | Bogor | Facebook | 300 | 5 | 1705323 357 |

Event Produce

```
reh471@reh471: ~/Documents/dibimbing_spark_airflow
reh471@reh471: ~/Documents/d... x reh471@reh471: ~/Documents/d... x reh471@reh471: ~/Documents/d... x
e=-1, serialized_value_size=144, serialized_header_size=-1)
-----
{"ID": "889d2b23-b544-4ea9-9477-567b2cb90314", "Kota": "Depok", "Sosial_Media": "Telegram", "Usage_C
unt": 77, "Sales": 98, "Ts": 1705417287}'
-----
RecordMetadata(topic='test-topic', partition=0, topic_partition=TopicPartition(topic='test-topic', par
ition=0), offset=15070, timestamp=1705697835586, log_start_offset=0, checksum=None, serialized_key_si
e=-1, serialized_value_size=141, serialized_header_size=-1)
-----
{"ID": "7e81e426-0081-448a-b0c8-87b5a228f254", "Kota": "Bekasi", "Sosial_Media": "X", "Usage_Count":
250, "Sales": 118, "Ts": 1705161177}'
-----
RecordMetadata(topic='test-topic', partition=0, topic_partition=TopicPartition(topic='test-topic', par
ition=0), offset=15071, timestamp=1705697840594, log_start_offset=0, checksum=None, serialized_key_si
e=-1, serialized_value_size=137, serialized_header_size=-1)
-----
{"ID": "f49294df-44fd-4daf-bada-b1d2b5e0d5d2", "Kota": "Depok", "Sosial_Media": "X", "Usage_Count":
11, "Sales": 94, "Ts": 1705623516}'
-----
RecordMetadata(topic='test-topic', partition=0, topic_partition=TopicPartition(topic='test-topic', par
ition=0), offset=15072, timestamp=1705697845605, log_start_offset=0, checksum=None, serialized_key_si
e=-1, serialized_value_size=135, serialized_header_size=-1)
-----
```

To Kafka Topic



Processing

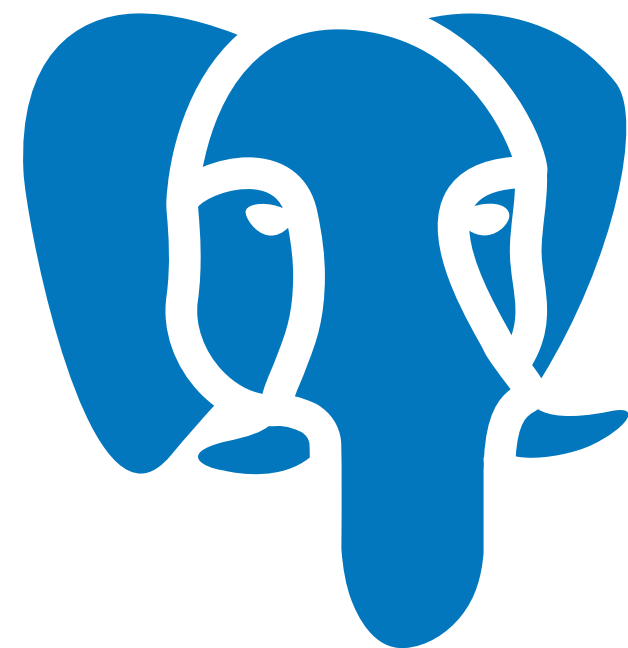
Listen from Kafka Topic, cleaning, transformation, load to DB, etc

```
reh471@reh471: ~/Documents/dibimbing_spark_airflow
reh471@reh471: ~/Documents/di... x reh471@reh471: ~/Documents/di... x reh471@reh471: ~/Documents/di... x
de05aabd-99dc-433d-a18e-d894d5846e27|Jakarta|X|327|63|2024-01-16 01:14:35|
-----+-----+-----+-----+-----+-----+
rocessing batch 1441
-----+-----+-----+-----+-----+-----+
ID|Kota|Sosial_Media|Usage_Count|Sales|Ts|
-----+-----+-----+-----+-----+-----+
cbefd251-d21c-4f93-a931-f0360786990a|Jakarta|Telegram|122|81|2024-01-15 11:59:24|
-----+-----+-----+-----+-----+-----+
rocessing batch 1442
-----+-----+-----+-----+-----+-----+
ID|Kota|Sosial_Media|Usage_Count|Sales|Ts|
-----+-----+-----+-----+-----+-----+
ae24abc4-1f07-4446-add1-7425aade435e|Bekasi|TikTok|612|65|2024-01-17 06:11:57|
-----+-----+-----+-----+-----+-----+
rocessing batch 1443
-----+-----+-----+-----+-----+-----+
ID|Kota|Sosial_Media|Usage_Count|Sales|Ts|
-----+-----+-----+-----+-----+-----+
13891473-a896-4a35-b2cf-b3be22010723|Bogor|X|324|50|2024-01-15 17:08:03|
-----+-----+-----+-----+-----+-----+
```

| ID | Kota | Sosial_Media | Usage_count | Sales | Ts |
|-------|-----------|--------------|-------------|-------|------------|
| 1dd12 | Jakarta | X | 800 | 45 | 12-01-2024 |
| 1dd13 | Tangerang | Instagram | 200 | 123 | 12-01-2024 |
| 1dd13 | Bogor | Facebook | 300 | 5 | 12-01-2024 |

Storage

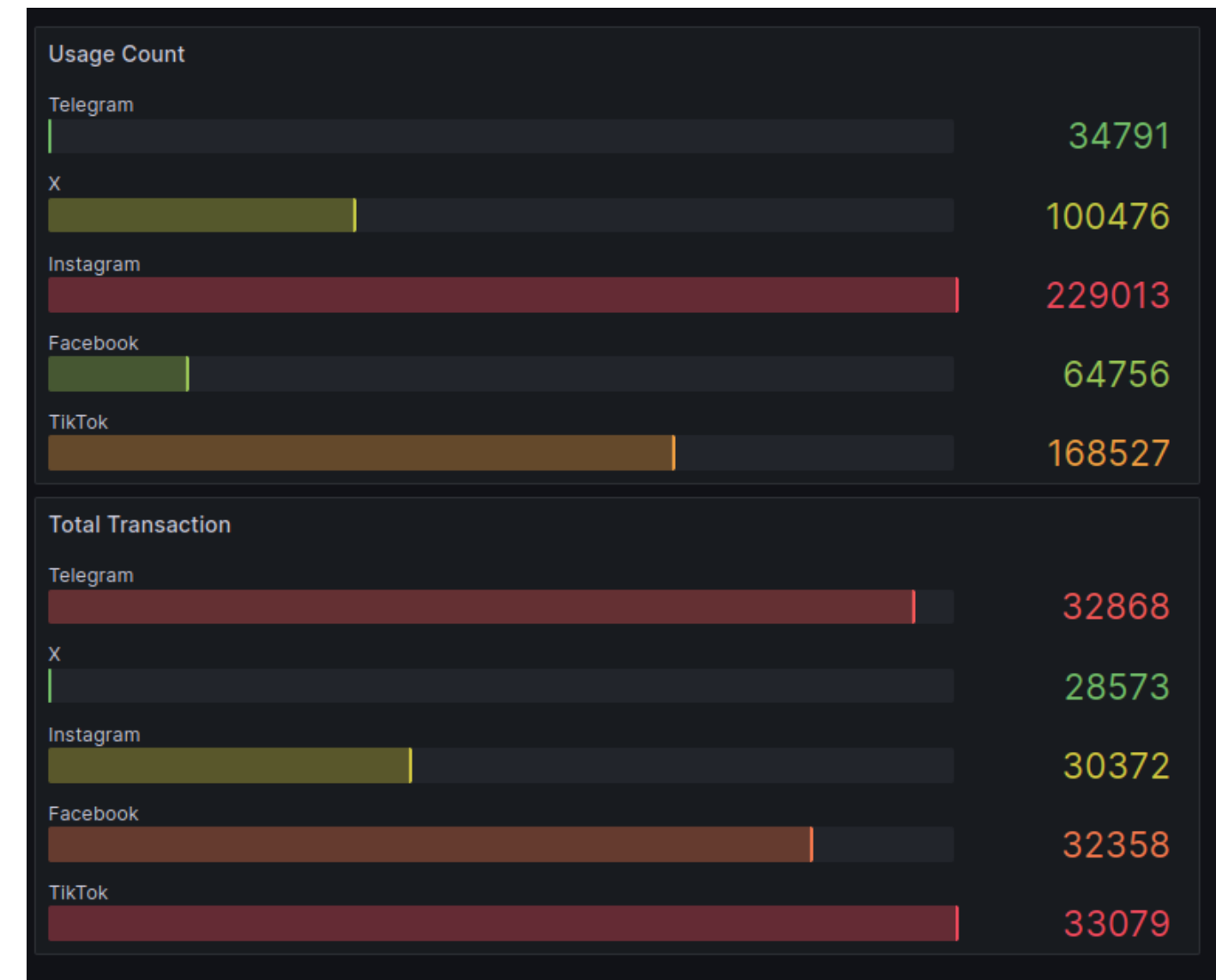
Neon Serverless Postgresql Postgres instance



| socmed_user_sales | | | | | | |
|-------------------|--------------------------------------|-----------|--------------|-------------|-------|---------------------|
| # | ID | Kota | Sosial_Media | Usage_Count | Sales | Ts |
| 501 | 17b42ad2-46a6-4e59-9128-be00d2e689b9 | Jakarta | Facebook | 141 | 118 | 2024-01-19 07:16:38 |
| 502 | 8bb813c5-3f24-4807-932f-4f48f2ff2d4a | Bogor | TikTok | 337 | 63 | 2024-01-13 16:34:05 |
| 503 | c887fa5a-f7a2-4444-bf21-5db1f070be89 | Bekasi | Facebook | 246 | 147 | 2024-01-15 00:35:07 |
| 504 | d447d7a1-85c7-4e7c-b9b0-cd587777077e | Depok | TikTok | 532 | 122 | 2024-01-19 04:33:50 |
| 505 | 888911c9-2cb9-46ac-a79b-020c15b980b1 | Bogor | Instagram | 812 | 67 | 2024-01-16 12:33:32 |
| 506 | fa74a295-5311-4998-8253-fb451920fd29 | Jakarta | Telegram | 82 | 73 | 2024-01-12 21:24:20 |
| 507 | b0574607-0371-48e0-9387-2d0beb39560c | Tangerang | Telegram | 74 | 96 | 2024-01-19 07:47:32 |
| 508 | bf07674c-e187-44de-a636-12590975679e | Jakarta | Telegram | 126 | 138 | 2024-01-17 23:12:07 |
| 509 | 09dc7325-f900-4019-9a77-f0cc4b92cbce | Tangerang | Facebook | 287 | 80 | 2024-01-13 00:21:13 |
| 510 | adb9b8c4-86d7-4d43-9156-4d04a5594a35 | Tangerang | Telegram | 58 | 141 | 2024-01-13 03:19:10 |
| 511 | cdad381f-9bc3-4fe1-ad2e-bc8c6dbe44ee | Jakarta | Telegram | 63 | 82 | 2024-01-18 14:37:53 |
| 512 | 9371782f-ac6d-437f-9f21-7dd873357299 | Depok | Telegram | 148 | 85 | 2024-01-15 00:51:17 |
| 513 | f78626fe-fcf6-418f-b18e-6e6f2cf7b49d | Tangerang | Facebook | 297 | 133 | 2024-01-18 05:28:52 |
| 514 | 9b897622-c6a2-4027-a8b9-a0258c795cc4 | Jakarta | Instagram | 671 | 85 | 2024-01-19 07:02:46 |
| 515 | c48641a6-17f7-4130-b073-a0bd213cc627 | Bekasi | Facebook | 125 | 146 | 2024-01-18 05:11:18 |
| 516 | eccee5c6-a0a8-4162-b905-489fbf9eee38 | Tangerang | X | 359 | 137 | 2024-01-16 05:40:33 |
| 517 | 518a117b-3300-49cc-9772-47abb4a93004 | Depok | Instagram | 781 | 145 | 2024-01-13 19:36:09 |

Dataviz

Lightweight, real-time, open source...



Refresh every 5 seconds to display dynamic data.



Thanks!

Any feedback?

Farah Shadra

farahduta7@gmail.com

