

Producer

Dalam project ini producer menggunakan data penjualan obat di “Apotek Dibimbing Ajinusa”. Dengan melakukan simulasi membuat topic dengan nama **apotek-dibimbing-ajinusa** pada cluster Kafka, membaginya menjadi **5 Partition**, membuat **1 Replikasi**, **1 jam Retensi** data disimpan dalam Kafka, dan max penyimpanan 1 GB.

The screenshot shows the Apache Kafka UI for the topic 'apotek-dibimbing-ajinusa'. The 'Overview' tab is selected, displaying various metrics and a table of partitions. The metrics include 5 Partitions, Replication Factor 1, URP 0, 5 of 5 In Sync Replicas, Type External, and Segment Size 130 KB. The Segment Count is 5, and the Message Count is 389. The table below lists the partitions with their IDs, Replicas, First Offset, Next Offset, and Message Count.

Partition ID	Replicas	First Offset	Next Offset	Message Count
0	1	0	76	76
1	1	0	78	78
2	1	0	78	78
3	1	0	81	81
4	1	0	76	76

The screenshot shows the 'Topics / Create' form in the Apache Kafka UI. The form is for creating a new topic named 'apotek-dibimbing-ajinusa'. The configuration includes 5 partitions, a Cleanup policy of Delete, 1 Min In Sync Replicas, and a Replication Factor of 1. The Time to retain data (in ms) is set to 3600000 (1 hour). The Max size on disk in GB is set to 1 GB, and the Maximum message size in bytes is set to Maximum message size. The Custom parameters section is empty, and the 'Create topic' button is visible.

Topic Name *

apotek-dibimbing-ajinusa

Number of partitions *

5

Cleanup policy

Delete

Min In Sync Replicas

1

Replication Factor

1

Time to retain data (in ms) 1h

3600000

12 hours 1 day 2 days 7 days 4 weeks

Max size on disk in GB

1 GB

Maximum message size in bytes

Maximum message size

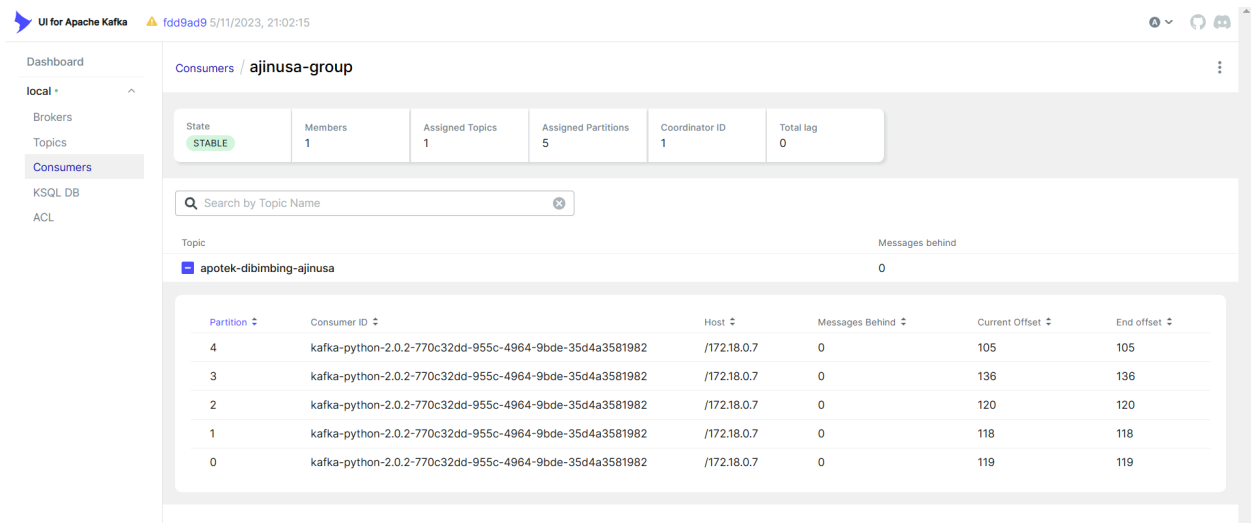
Custom parameters

+ Add Custom Parameter

Cancel Create topic

Consumer

Dalam project ini dilakukan simulasi dengan 3 consumer di dalam 1 consumer group yang sama yaitu “ajinusa-group”. Consumer membaca topic **apotek-dibimbing-ajinusa**

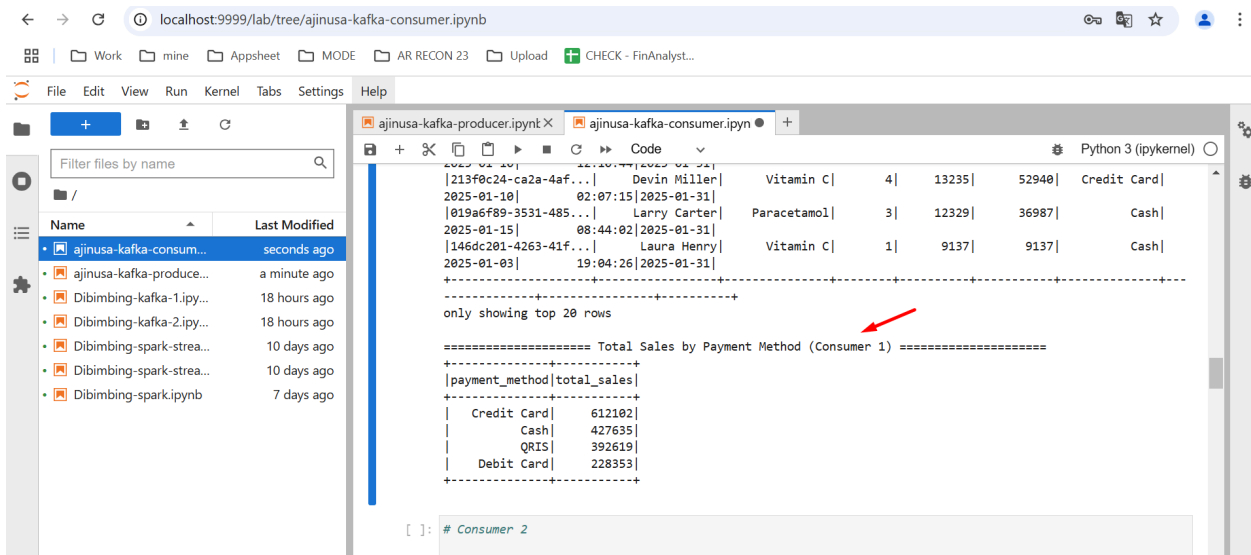


State	Members	Assigned Topics	Assigned Partitions	Coordinator ID	Total lag
STABLE	1	1	5	1	0

Topic	Messages behind
apotek-dibimbing-ajinusa	0

Partition	Consumer ID	Host	Messages Behind	Current Offset	End offset
4	kafka-python-2.0.2-770c32dd-955c-4964-9bde-35d4a3581982	/172.18.0.7	0	105	105
3	kafka-python-2.0.2-770c32dd-955c-4964-9bde-35d4a3581982	/172.18.0.7	0	136	136
2	kafka-python-2.0.2-770c32dd-955c-4964-9bde-35d4a3581982	/172.18.0.7	0	120	120
1	kafka-python-2.0.2-770c32dd-955c-4964-9bde-35d4a3581982	/172.18.0.7	0	118	118
0	kafka-python-2.0.2-770c32dd-955c-4964-9bde-35d4a3581982	/172.18.0.7	0	119	119

Berikut ini adalah Consumer 1, melakukan analisis data dengan menampilkan data Sales berdasarkan *Payment Method* . Consumer 1 mengambil data dari topic **apotek-dibimbing-ajinusa** dari semua Partition yang ada



```
ajinusa-kafka-consumer.ipynb
```

Timestamp	Name	Product	Partition	Offset	Payment Method
2025-01-10 02:07:15	Devin Miller	Vitamin C	4	13235	Credit Card
2025-01-10 08:44:02	Larry Carter	Paracetamol	3	12329	Cash
2025-01-15 19:04:26	Laura Henry	Vitamin C	1	9137	Cash

only showing top 20 rows

===== Total Sales by Payment Method (Consumer 1) =====

payment_method	total_sales
Credit Card	612102
Cash	427635
QRIS	392619
Debit Card	228353

```
[ ]: # Consumer 2
```

Berikut ini adalah Consumer 2, melakukan analisis data dengan menampilkan data Sales per Bulan transaksi. Consumer 2 mengambil data dari topic **apotek-dibimbing-ajinusa** dari semua Partition yang ada

localhost:9999/lab/tree/ajinusa-kafka-consumer.ipynb

File Edit View Run Kernel Tabs Settings Help

ajinusa-kafka-producer.ipynb X ajinusa-kafka-consumer.ipynb +

Python 3 (ipykernel)

ID	Name	Product	Quantity	Price	Total Price	Payment Method
202b806b-aa27-44b...	Connie West	Paracetamol	5	8052	40260	Cash
2025-01-10	05:47:31 2025-01-31					
7c5fcf8b-845e-4cc...	Craig Dalton	Vitamin C	2	6210	12420	Cash
2025-01-09	15:27:08 2025-01-31					
eaf59d98-1f41-474...	John Guerrero	Amoxicillin	1	15788	15788	QRIS
2025-01-18	02:47:32 2025-01-31					
c185edd2-fed0-4af...	Christopher Murphy	Amoxicillin	5	18879	94395	Credit Card
2025-01-18	16:31:10 2025-01-31					
c8783659-4b92-442...	Brittany Francis	Vitamin C	1	16215	16215	Cash
2025-01-08	19:17:46 2025-01-31					
e111c2c3-b786-442...	Richard Davis DVM	Antibiotik	3	2844	8532	Cash
2025-01-11	13:09:37 2025-01-31					

only showing top 20 rows

===== Total Sales by Month (Consumer 2) =====

Month	Total Sales
2025-01-31	1622863

Berikut ini adalah Consumer 3, melakukan analisis data dengan menampilkan data *Total Sales by Payment Method* dan *Total Sales by Month*. Consumer 2 mengambil data dari topic **apotek-dibimbing-ajinusa** hanya dari Partition nomor 2 saja

ajinusa-kafka-producer.ipynb X ajinusa-kafka-consumer.ipynb +

Python 3 (ipykernel)

```
[5]: # Consumer 3: coba ambil data dari partition tertentu

from kafka import TopicPartition
from pyspark.sql import SparkSession
from pyspark.sql.types import *
from pyspark.sql.functions import *
from pyspark.sql import functions as F

partition = 2

consumer3 = KafkaConsumer(group_id='ajinusa-group',
                           bootstrap_servers=[f'{kafka_host}:9092'],
                           value_deserializer=lambda m: json.loads(m.decode('ascii')),
                           # auto_offset_reset='earliest',
                           enable_auto_commit=False)

consumer3.assign([TopicPartition(kafka_topic_partition, partition)])

# Membaca pesan dan menyimpannya ke dalam List
```

localhost:9999/lab/tree/ajinusa-kafka-consumer.ipynb

WorkmineAppsheetMODEAR RECON 23UploadCHECK - FinAnalyst...

FileEditViewRunKernelTabsSettingsHelp

Filter files by name

/

Name	Last Modified
ajinusa-kafka-consum...	seconds ago
ajinusa-kafka-produce...	a minute ago
Dibimbing-kafka-1.ipy...	20 hours ago
Dibimbing-kafka-2.ipy...	20 hours ago
Dibimbing-spark-strea...	10 days ago
Dibimbing-spark-strea...	10 days ago
Dibimbing-spark.ipynb	7 days ago

ajinusa-kafka-producer.ipyn

ajinusa-kafka-consumer.ipyn

Python 3 (ipykernel)

Code

```
[6ca07074-3565-4e7...| John Flores| Antibiotik| 1| 16915| 16915| Credit Card|
2025-01-17| 19:23:26|2025-01-31|
+-----+-----+-----+-----+-----+
===== Consumer 3 : Partition 2 =====
===== Total Sales by Payment Method =====
+-----+
|payment_method|total_sales|
+-----+
|      Cash|      61260|
|      QRIS|      34569|
| Credit Card|      16915|
|   Debit Card|      6620|
+-----+

===== Total Sales by Month =====
+-----+
|      eom|total_sales|
+-----+
|2025-01-31|      119364|
+-----+
```

KSQL DB

Di sini disimulasikan membaca data di cluster Kafka melalui fitur KSQL DB. Dengan cara membuat Stream **apotek_dibimbing_ajinusa_stream** dan Tabel **apotek_dibimbing_ajinusa_ksql** dan **payment_method_sales_ksql** yang bersumber dari topic 'apotek-dibimbing-ajinusa'

The screenshot shows the KSQL DB interface with a query executed. The query is: `SELECT * FROM apotek_dibimbing_ajinusa_stream;`. The results are displayed in a table with the following schema:

TRANSACTION_ID	BUYER_NAME	MEDICATION_NAME	QUANTITY	UNIT PRICE	TOTAL PRICE	PAYMENT METHOD	TRANSACTION DATE	TRANSACTION TIME
4718e0cd-a3b9-495d-9c46-27e0fcd3fb2	Jared Conrad	Amoxicillin	4	5828	23312	Debit Card	2025-01-16	20:28:11
94e0d472-5e8f-463f-90df-6b7f5f165407	Andre Smith	Cough Syrup	2	3168	6336	Debit Card	2025-01-09	09:43:45
0d5ba2ce-a239-416c-a3e1-c7af15165407	Tracy Dorsey	Cough Syrup	5	2753	13765	Debit Card	2025-01-07	14:20:49
4fc8ac41-4777-4533-9e16-666a6c9300f8	Shannon Hunt	Paracetamol	2	5534	11068	Credit Card	2025-01-13	12:38:39
b0ff8d10-53b0-467f-89a6-751ac8bee90a	Matthew Bender	Amoxicillin	2	11004	22008	Cash	2025-01-19	12:25:03
709f9c5f-0a36-4fc3-be65-216a1b63b22	John Adams	Ibuprofen	2	13241	26482	Credit Card	2025-01-17	06:23:54
c8773241-c203-4780-aed9-ba2e13e49f41	Joseph Back	Paracetamol	5	13079	65395	Debit Card	2025-01-18	10:45:17

The screenshot shows the KSQL DB interface with a query executed. The query is: `SELECT * FROM apotek_dibimbing_ajinusa_stream;`. The results are displayed in a table with the following schema:

TRANSACTION_ID	BUYER_NAME	MEDICATION_NAME	QUANTITY	UNIT PRICE	TOTAL PRICE	PAYMENT METHOD	TRANSACTION DATE	TRANSACTION TIME
6aa70dec-6bee-483b-a87f-94d9e3429cdf	Danny Armstrong	Amoxicillin	5	18601	93005	Debit Card	2025-01-01	23:19:01
b0ff8d10-53b0-467f-89a6-751ac8bee90a	Matthew Bender	Amoxicillin	2	11004	22008	Cash	2025-01-19	12:25:03
300990d6-fb50-41af-9721-013048b3e30e	Jessica Ward	Paracetamol	1	7957	7957	QRIS	2025-01-02	10:36:06
4718e0cd-a3b9-495d-9c46-27e0fcd3fb2	Jared Conrad	Amoxicillin	4	5828	23312	Debit Card	2025-01-16	20:28:11
94e0d472-5e8f-463f-90df-6b7f5f165407	Andre Smith	Cough Syrup	2	3168	6336	Debit Card	2025-01-09	09:43:45
0d5ba2ce-a239-416c-a3e1-c7af15165407	Tracy Dorsey	Cough Syrup	5	2753	13765	Debit Card	2025-01-07	14:20:49
29513e31-272e-4861-91f1-8239f9e6b6b6	Robert Sales	Cough Syrup	1	19532	19532	Credit Card	2025-01-07	14:01:42

Dashboard

local ▾

Brokers

Topics

Consumers

KSQL DB

ACL

TablesStreams

KSQL

Clear

1 select * from payment_method_sales_ksql;

Clear resultsExecute

Schema

'PAYMENT_METHOD' STRING KEY	'TOTAL_PRICE' INTEGER
QRIS	324145
Debit Card	248204
Credit Card	177256
Cash	74504

Stream properties:

Key

+ Add