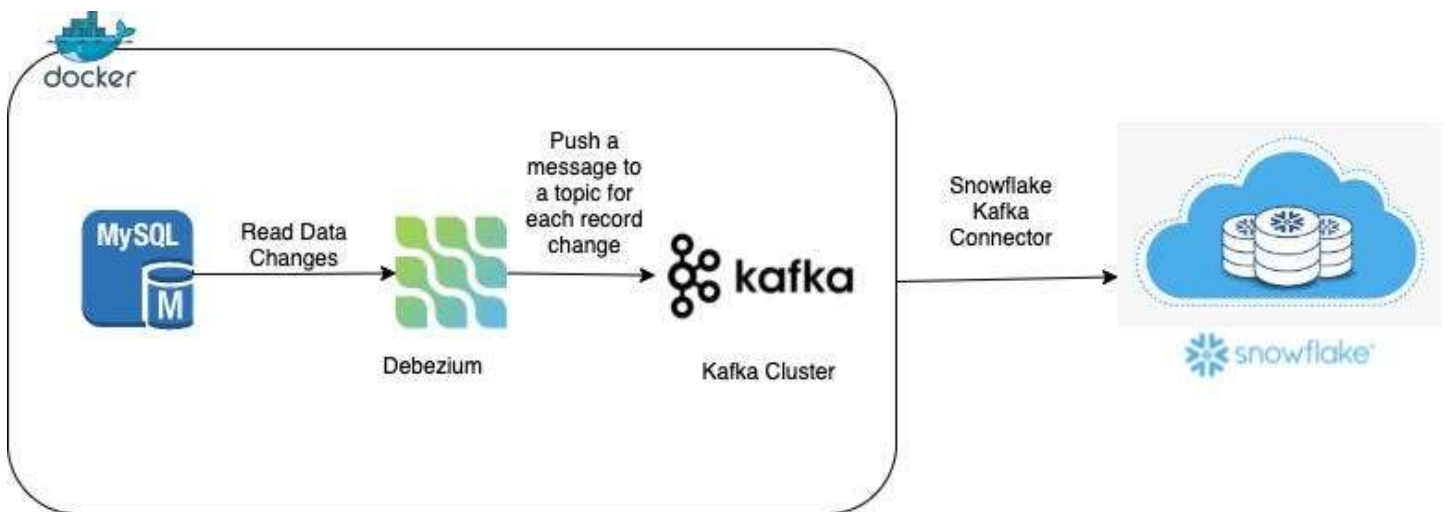


Kafka Workshop

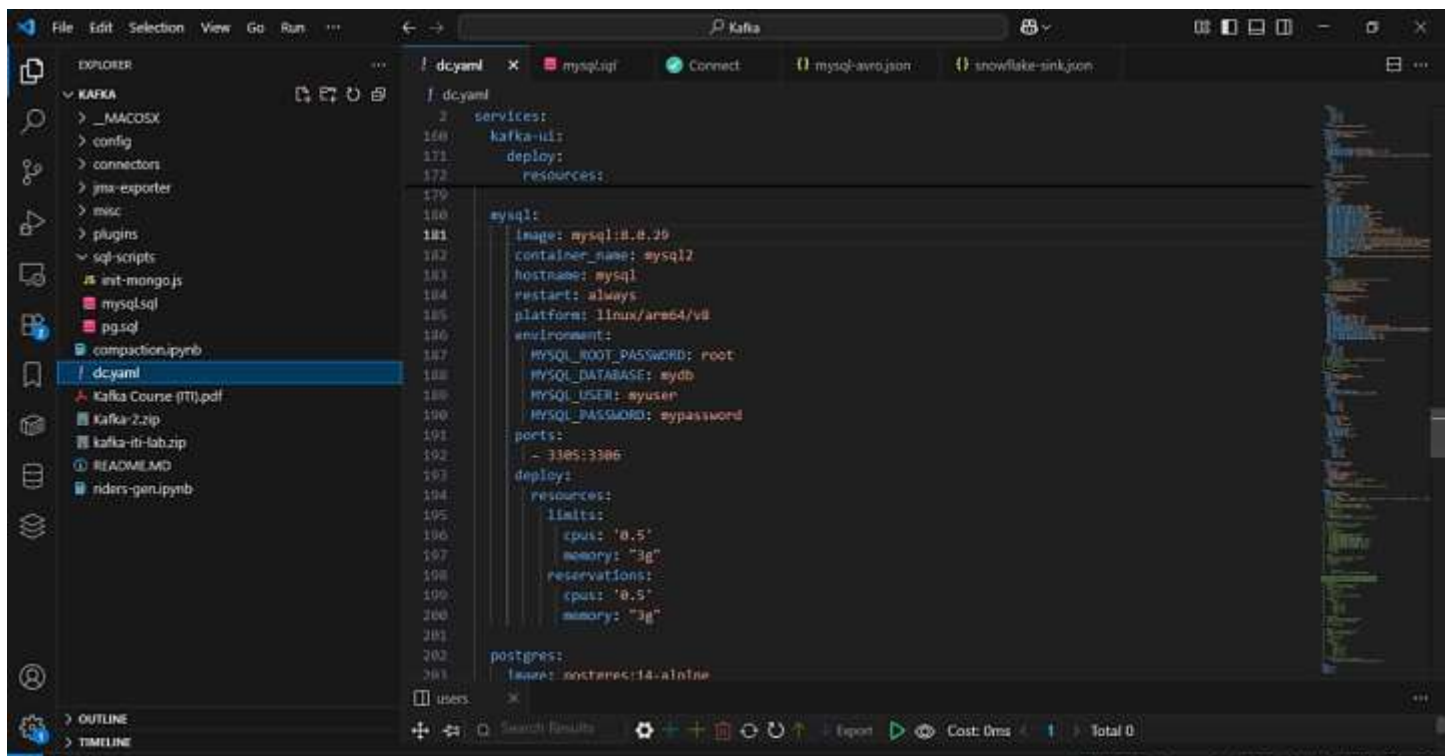
By

Moataz Gamal Ali

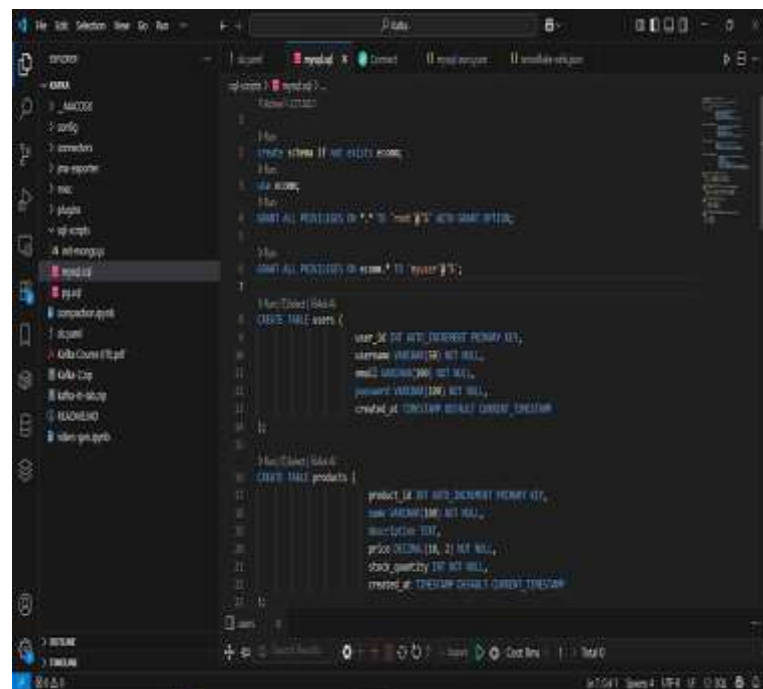
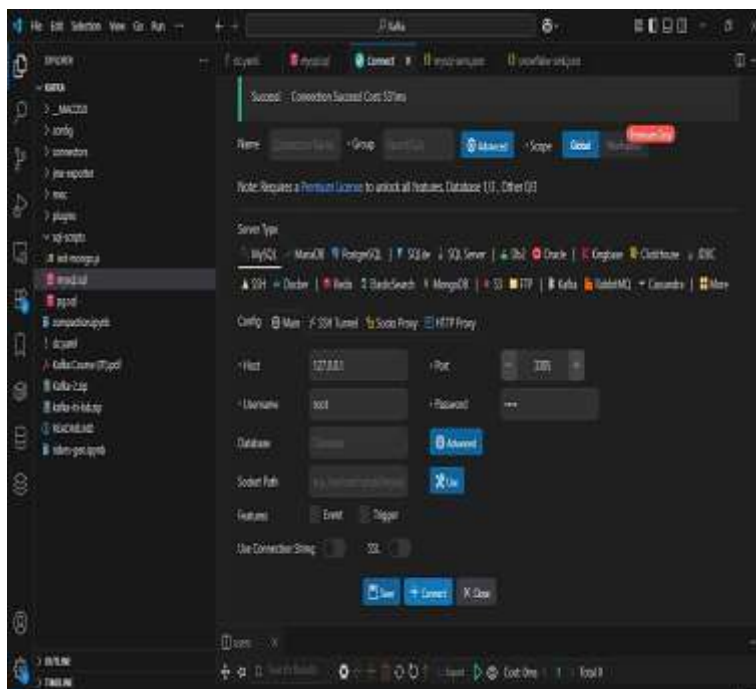
Data Engineering Zagazig Track



1. Un-commented MySQL image from yaml file and build the file again



2. Connect to MySQL container from localhost and create the database and insert values in each table using script in /sql-scripts/mysql.sql file



3. Used the json configuration file inside the connectors folder to configure the connection from mysql database.

```
connectors > mysql > {} mysql-avro.json > {} config > key.converter
1  {
2    "name": "mysql-avro-connector",
3    "config": {
4      "connector.class": "io.debezium.connector.mysql.MySqlConnector",
5      "database.hostname": "mysql2",
6      "database.port": "3306",
7      "database.user": "root",
8      "database.password": "root",
9      "database.server.id": "18405",
10     "database.include.list": "ecomm",
11     "topic.prefix": "kafka-workshop",
12     "key.converter": "org.apache.kafka.connect.json.JsonConverter",
13     "key.converter.schema.registry.url": "http://schema-registry:8081",
14     "value.converter": "io.confluent.connect.avro.AvroConverter",
15     "value.converter.schema.registry.url": "http://schema-registry:8081",
16     "schema.history.internal.kafka.bootstrap.servers": "broker:29092",
17     "schema.history.internal.kafka.topic": "ecomm.schemahistory",
18     "include.schema.changes": "true",
19     "event.processing.failure.handling.mode": "ignore",
20     "transforms": "unwrap",
21     "transforms.unwrap.type": "io.debezium.transforms.ExtractNewRecordState",
22     "transforms.unwrap.drop.tombstones": false,
23     "time.precision.mode": "connect"
```

4. Downloaded the snowflake sink connector and put it inside the plugins, made my sink connector "snowflake" configurations and put it in connectors.

```
connectors > snowflake > {} snowflake-sink.json > {} config
1  {
2    "name": "snowflake-sink-ecomm",
3    "config": {
4      "connector.class": "com.snowflake.kafka.connector.SnowflakeSinkConnector",
5      "tasks.max": "1",
6
7      "topics.regex": "kafka-workshop.*",
8
9      "snowflake.url.name": "YREDIGW-DB28773.snowflakecomputing.com",
10     "snowflake.user.name": "MOATAZGAMAL710",
11     "snowflake.private.key": "MIIEvAIBADANBgkqhkiG9w0BAQEFAASCByggSiAgEAAoIBAQDKhzsGk5lLgWx3F19HjLJVU",
12     "snowflake.database.name": "NOSQL_ECOMM",
13     "snowflake.schema.name": "PUBLIC",
14     "snowflake.stage.name": "KAFKA_STAGE",
15     "snowflake.role.name": "ACCOUNTADMIN",
16
17     "buffer.count.records": "500",
18     "buffer.size.bytes": "5000000",
19     "buffer.flush.time": "60",
20
21     "key.converter": "org.apache.kafka.connect.json.JsonConverter",
22     "key.converter.schemas.enable": "false",
```

5. Configured the key pair that used for authentication for snowflake using the snowflake document with rsa.

```
NOSQL_ECOMM_PUBLIC Settings Open in Workspaces Code Versions
1 ALTER USER MOATAZGAMAL710 SET rsa_public_key='MIIBIjANBgkqhkiG9w0BAQEFAAOCAQ8AMIIBCgKCAQEAyoc7BpOZSxsF9xZfR4yy
2 VVEGL2fIhs5ZJB0oy2H1lsNfKFwdbgAUmx5kZpZPhXZu5FbFcbD171ZLyEX7a1zx
3 RaU5G7IBjQ2hAnjXev9MJJ16C5V3vnSKa6Cek6MN1HBXxfidNqfYRP0NfTH3S1MY
4 8a06QsrXVFzjft5+ghGFvNFUnZsx5ouGTITMjWf4kr3EXUUXMBpm6pXAQBEE6YL
5 ra1YK+Lo6y+WKD1BJa/Y9IomtG/BnG7LtnrEPFJwhcBD/Jp4EkrgQceNtRZUdG/7
6 tAtMoAV2IXJLnuo6g1m64lejUZLOx/ELcrI0f+K7prPL8Wk5Crd7uHz/OHLrWQI
7 owIDAQAB'
8
```

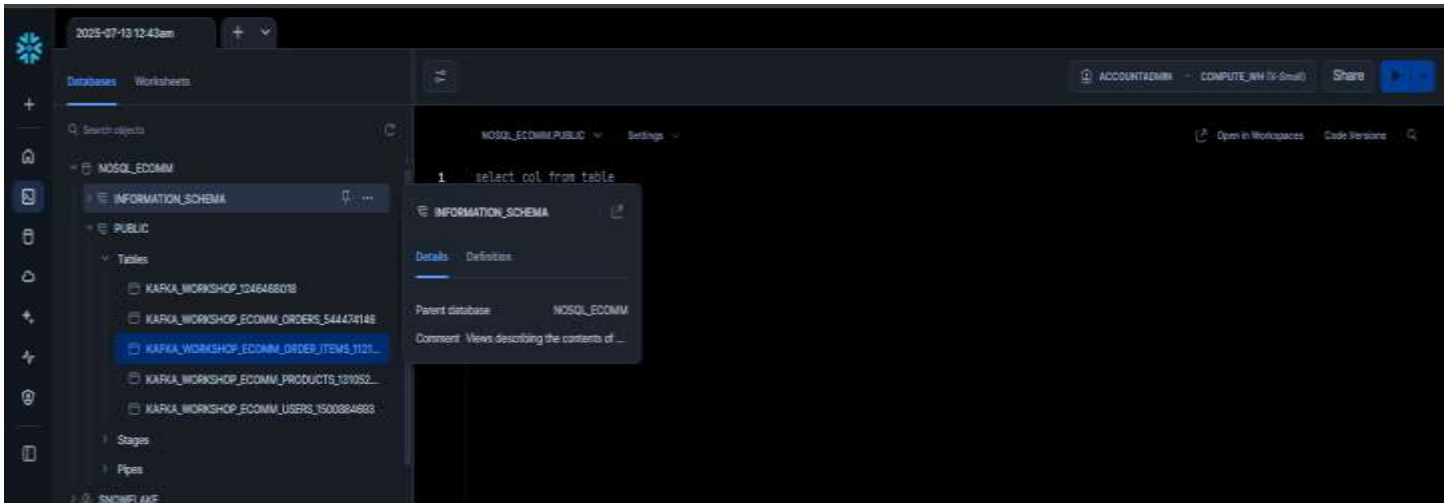
6. Run the commands for the connectors and checked that they are running

```
motaz@DESKTOP-V9KGL86 MINGW64 /d/ITI/Kafka
$ curl http://localhost:8083/connectors/?expand=status
{"mysql-avro-connector":{"status":{"name":"mysql-avro-connector","connector":{"state":"RUNNING","worker_id":"connect:8083"},"tasks":[{"id":0,"state":"RUNNING","worker_id":"connect:8083"},"type":"source"},"snowflake-sink-ecomm":{"status":{"name":"snowflake-sink-ecomm","connector":{"state":"RUNNING","worker_id":"connect:8083"},"tasks":[{"id":0,"state":"RUNNING","worker_id":"connect:8083"},"type":"sink"]}}}
motaz@DESKTOP-V9KGL86 MINGW64 /d/ITI/Kafka
$
```

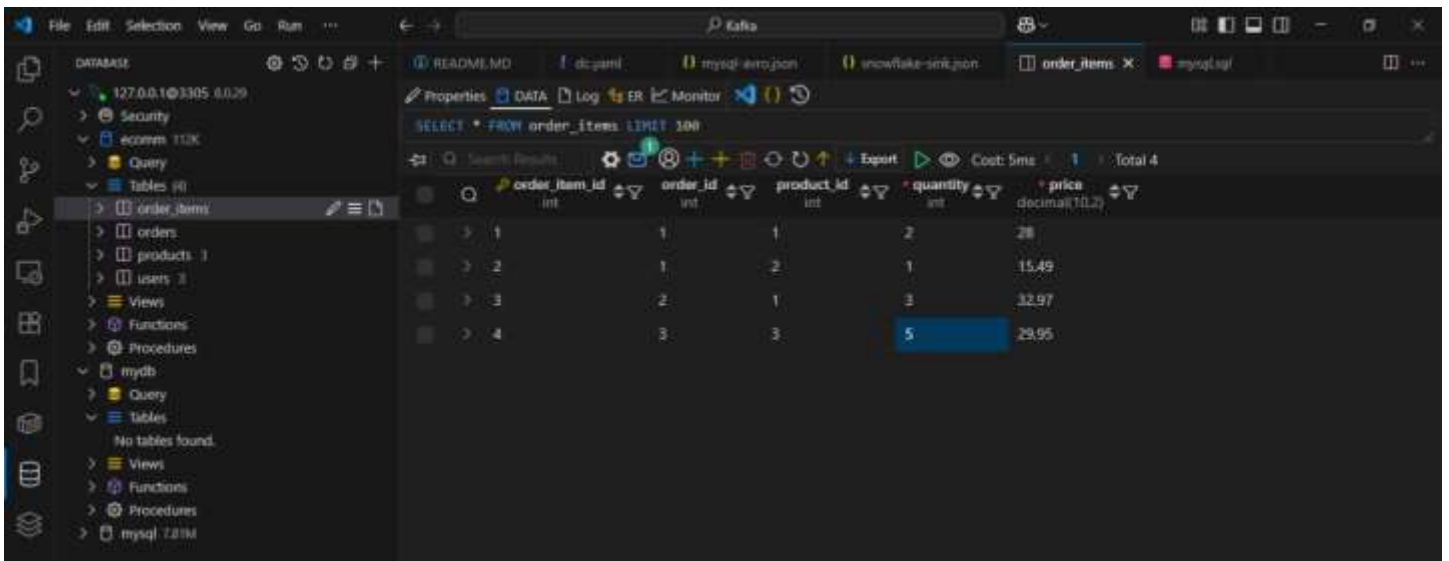
7. Checking Kafka topics

Topic Name	Partitions	Out of sync replicas	Replication Factor	Number of messages	Size
__consumer_offsets	50	0	1	22	6 KB
__transaction_state	50	0	1	4	568 Bytes
_confluent-ksq-default_command_topic	1	0	1	2	7 KB
_confluent-monitoring	1	0	1	2561	191 KB
_schemas	1	0	1	7	7 KB
default_ksql_processing_log	1	0	1	0	0 Bytes
docker-connect-configs	1	0	1	9	9 KB
docker-connect-offsets	25	0	1	8	2 KB
docker-connect-status	5	0	1	32	8 KB
ecomm.schemahistory	1	0	1	9	13 KB
kafka-workshop	1	0	1	9	6 KB
kafka-workshop.ecomm.order_items	1	0	1	5	1 KB
kafka-workshop.ecomm.orders	1	0	1	3	764 Bytes
kafka-workshop.ecomm.products	1	0	1	3	904 Bytes
kafka-workshop.ecomm.users	1	0	1	3	851 Bytes

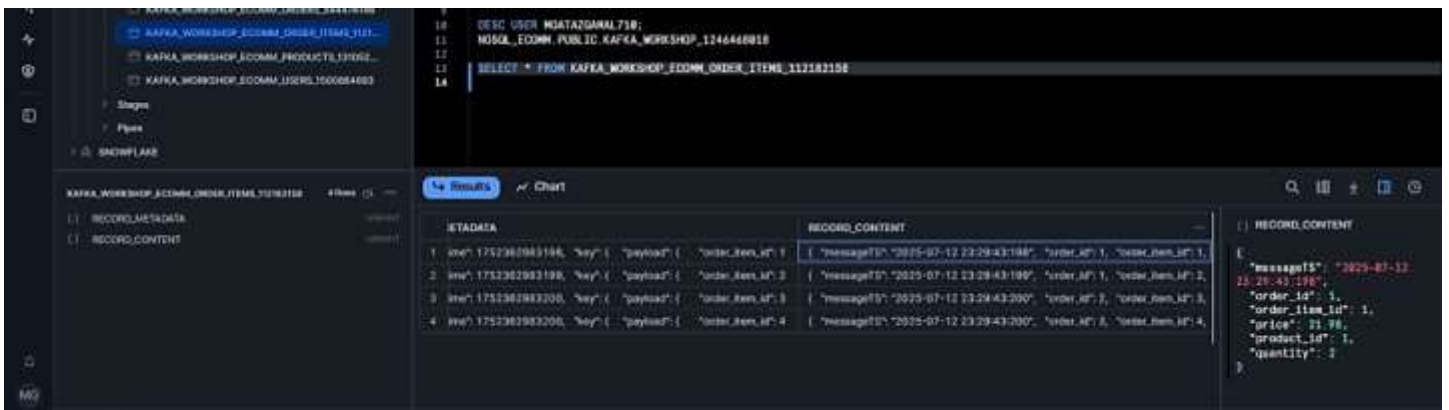
8. Checking created tables on snowflake



9. updating a record in order_item table – Changing the price to 28



10. Price before updating is 21.90 on snowflake



11. Price after updating is 28

The screenshot shows a Snowflake SQL query results interface. The query executed is:

```
SELECT * FROM KAFKA_WORKSHOP_EDDMM_ORDER_ITEMS_112182158
```

The results are displayed in a table with two columns: **METADATA** and **RECORD_CONTENT**. The table contains 5 rows of data. The first four rows have a **price** of 20, while the fifth row has a **price** of 28, indicating an update.

LINE	KEY	PAYLOAD	MESSAGE	ORDER_ID	ORDER_ITEM_ID	PRICE
1	1752362983198	{ "payload": { "order_item_id": 1	{ "messageTS": "2025-07-12 23:29:43:198", "order_id": 1, "order_item_id": 1,	1	1	20
2	1752362983199	{ "payload": { "order_item_id": 2	{ "messageTS": "2025-07-12 23:29:43:199", "order_id": 1, "order_item_id": 2,	1	2	20
3	1752362983200	{ "payload": { "order_item_id": 3	{ "messageTS": "2025-07-12 23:29:43:200", "order_id": 2, "order_item_id": 3,	2	3	20
4	1752362983200	{ "payload": { "order_item_id": 4	{ "messageTS": "2025-07-12 23:29:43:200", "order_id": 3, "order_item_id": 4,	3	4	20
5	1752363759792	{ "payload": { "order_item_id": 1	{ "messageTS": "2025-07-12 23:42:39:792", "order_id": 1, "order_item_id": 1,	1	1	28

The **RECORD_CONTENT** column shows the full JSON message for the selected row, including the updated **price** of 28.