Створення топіків

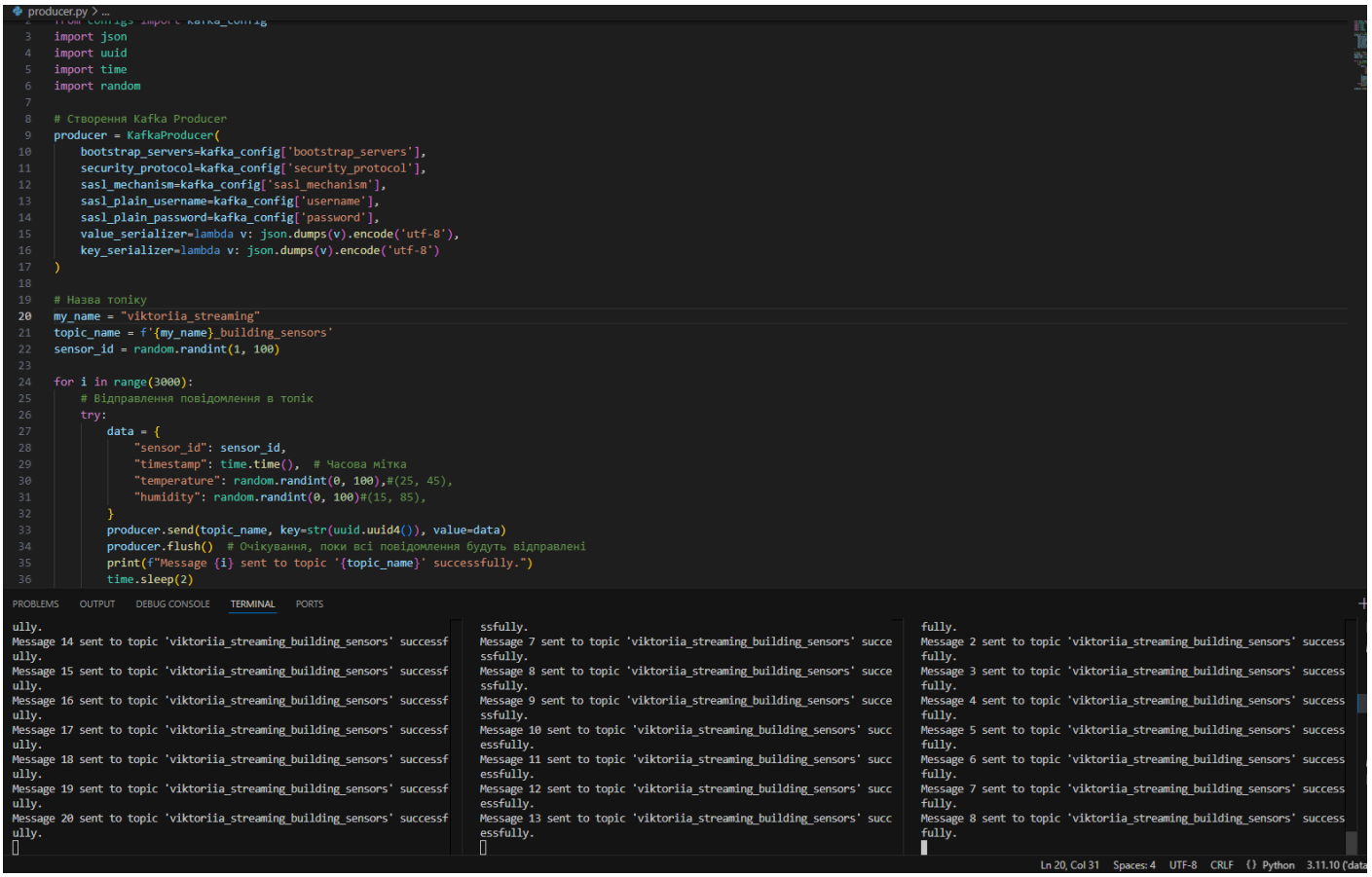
create_topics.py > ...

```python
from kafka.admin import KafkaAdminClient, NewTopic
from configs import kafka_config

# Створення клієнта Kafka
admin_client = KafkaAdminClient(
    bootstrap_servers=kafka_config['bootstrap_servers'],
    security_protocol=kafka_config['security_protocol'],
    sasl_mechanism=kafka_config['sasl_mechanism'],
    sasl_plain_username=kafka_config['username'],
    sasl_plain_password=kafka_config['password']
)


# Визначення нового топіку
my_name = "viktoriia_streaming"
topic_names= [f'{my_name}_building_sensors', f'{my_name}_alerts']#, f'{my_name}_humidity_alerts'

building_sensors = NewTopic(name=topic_names[0], num_partitions=2, replication_factor=1)
temperature_alerts = NewTopic(name=topic_names[1], num_partitions=1, replication_factor=1)
#humidity_alerts = NewTopic(name=topic_names[2], num_partitions=1, replication_factor=1)

# Створення нового топіку
try:
    admin_client.create_topics(new_topics=[building_sensors, temperature_alerts], validate_only=False)
    print(f"Topics {topic_names} created successfully.")
except Exception as e:
    print(f"An error occurred: {e}")

# Перевіряємо список існуючих топіків
[print(topic) for topic in admin_client.list_topics() if my_name in topic]

# Закриття зв'язку з клієнтом
admin_client.close()
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
Message 524 sent to topic 'viktori_building_sensors' successfully.
Traceback (most recent call last):
  File "c:\Projects\data_engineering\producer.py", line 36, in <module>
    time.sleep(2)
KeyboardInterrupt
PS C:\Projects\data_engineering> & C:/Users/kusiy/anaconda3/envs/data_engineering/python.exe c:/Projects/data_engineering/create_topics.py
Traceback (most recent call last):
  File "c:\Projects\data_engineering\create_topics.py", line 20, in <module>
    humidity_alerts = NewTopic(name=topic_names[2], num_partitions=1, replication_factor=1)
                                    ~~~~~~~~~~~^^^
IndexError: list index out of range
PS C:\Projects\data_engineering> & C:/Users/kusiy/anaconda3/envs/data_engineering/python.exe c:/Projects/data_engineering/create_topics.py
Topics ['viktoriia_streaming_building_sensors', 'viktoriia_streaming_alerts'] created successfully.
viktoriia_streaming_building_sensors
viktoriia_streaming_alerts
PS C:\Projects\data_engineering>
```

Відправка повідомлень

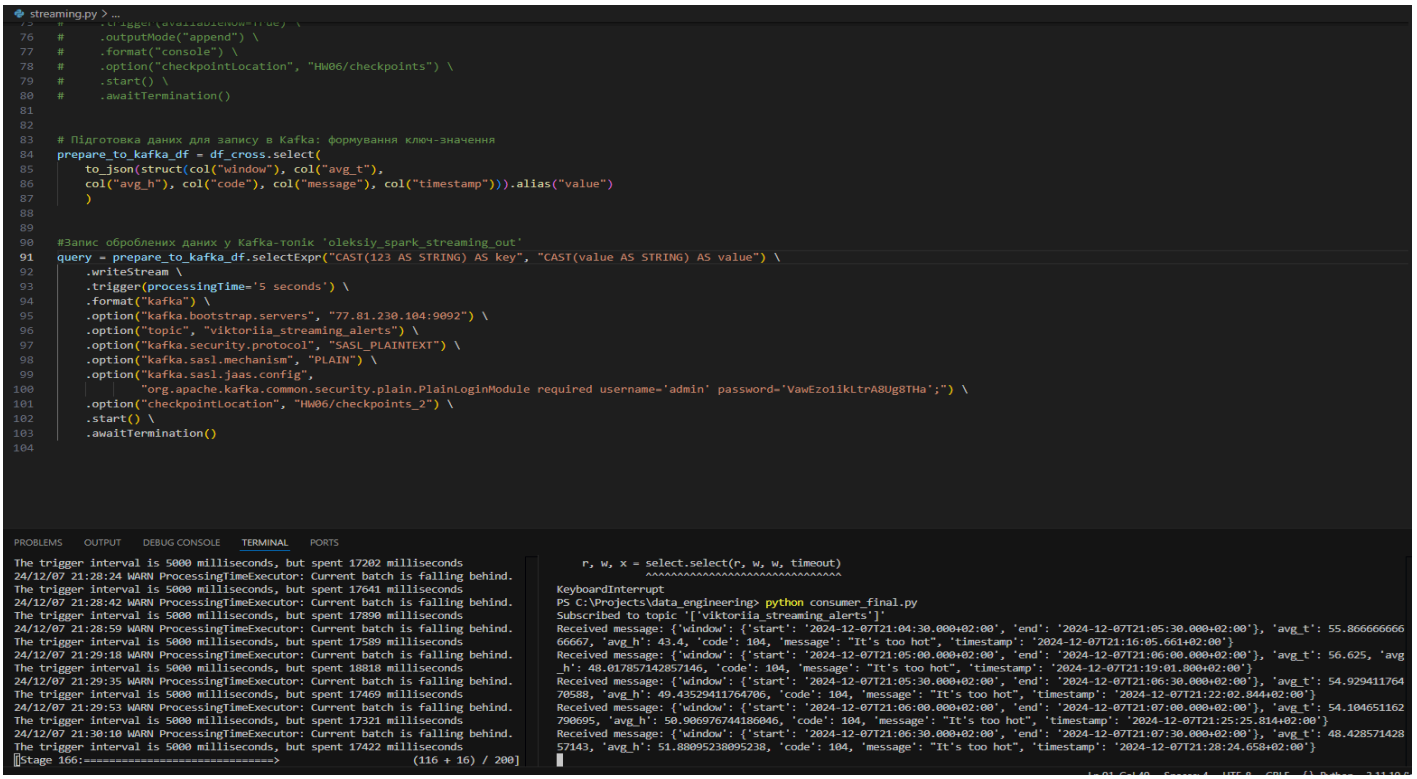


```python
import json
import uuid
import time
import random

# Створення Kafka Producer
producer = KafkaProducer(
    bootstrap_servers=kafka_config['bootstrap_servers'],
    security_protocol=kafka_config['security_protocol'],
    sasl_mechanism=kafka_config['sasl_mechanism'],
    sasl_plain_username=kafka_config['username'],
    sasl_plain_password=kafka_config['password'],
    value_serializer=lambda v: json.dumps(v).encode('utf-8'),
    key_serializer=lambda v: json.dumps(v).encode('utf-8')
)

# Назва топіку
my_name = "viktoriia_streaming"
topic_name = f'{my_name}_building_sensors'
sensor_id = random.randint(1, 100)

for i in range(3000):
    # Відправлення повідомлення в топік
    try:
        data = {
            "sensor_id": sensor_id,
            "timestamp": time.time(),  # Часова мітка
            "temperature": random.randint(0, 100),#(25, 45),
            "humidity": random.randint(0, 100)#(15, 85),
        }
        producer.send(topic_name, key=str(uuid.uuid4()), value=data)
        producer.flush()  # Очікування, поки всі повідомлення будуть відправлені
        print(f"Message {i} sent to topic '{topic_name}' successfully.")
        time.sleep(2)
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS

ully.
Message 14 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 15 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 16 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 17 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 18 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 19 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.
Message 20 sent to topic 'viktoriia_streaming_building_sensors' successf
ully.

ssfully.
Message 7 sent to topic 'viktoriia_streaming_building_sensors' succe
ssfully.
Message 8 sent to topic 'viktoriia_streaming_building_sensors' succe
ssfully.
Message 9 sent to topic 'viktoriia_streaming_building_sensors' succe
ssfully.
Message 10 sent to topic 'viktoriia_streaming_building_sensors' succ
essfully.
Message 11 sent to topic 'viktoriia_streaming_building_sensors' succ
essfully.
Message 12 sent to topic 'viktoriia_streaming_building_sensors' succ
essfully.
Message 13 sent to topic 'viktoriia_streaming_building_sensors' succ
essfully.

fully.
Message 2 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 3 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 4 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 5 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 6 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 7 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
Message 8 sent to topic 'viktoriia_streaming_building_sensors' success
fully.
```

Ln 20, Col 31 Spaces: 4 UTF-8 CRLF {} Python 3.11.10 ('data

Зчитування повідомлень з топіку, фільтрація, відправка повідомлень з тривогою та їх вивід на екран



```python
#     .outputMode("append") \
#     .format("console") \
#     .option("checkpointLocation", "HW06/checkpoints") \
#     .start() \
#     .awaitTermination()


# Підготовка даних для запису в Kafka: формування ключ-значення
prepare_to_kafka_df = df_cross.select(
    to_json(struct(col("window"), col("avg_t"),
    col("avg_h"), col("code"), col("message"), col("timestamp"))).alias("value")
    )


#Запис оброблених даних у Kafka-топік 'oleksiy_spark_streaming_out'
query = prepare_to_kafka_df.selectExpr("CAST(123 AS STRING) AS key", "CAST(value AS STRING) AS value") \
    .writeStream \
    .trigger(processingTime='5 seconds') \
    .format("kafka") \
    .option("kafka.bootstrap.servers", "77.81.230.104:9092") \
    .option("topic", "viktoriia_streaming_alerts") \
    .option("kafka.security.protocol", "SASL_PLAINTEXT") \
    .option("kafka.sasl.mechanism", "PLAIN") \
    .option("kafka.sasl.jaas.config",
            "org.apache.kafka.common.security.plain.PlainLoginModule required username='admin' password='VawEzo1ikLtrA8Ug8THa';") \
    .option("checkpointLocation", "HW06/checkpoints_2") \
    .start() \
    .awaitTermination()
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS

The trigger interval is 5000 milliseconds, but spent 17202 milliseconds
24/12/07 21:28:24 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17641 milliseconds
24/12/07 21:28:42 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17890 milliseconds
24/12/07 21:28:59 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17589 milliseconds
24/12/07 21:29:18 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 18818 milliseconds
24/12/07 21:29:35 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17469 milliseconds
24/12/07 21:29:53 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17321 milliseconds
24/12/07 21:30:10 WARN ProcessingTimeExecutor: Current batch is falling behind.
The trigger interval is 5000 milliseconds, but spent 17422 milliseconds
[Stage 166:============================>                    (116 + 16) / 200]

    r, w, x = select.select(r, w, w, timeout)
              ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
KeyboardInterrupt
PS C:\Projects\data_engineering> python consumer_final.py
Subscribed to topic '['viktoriia_streaming_alerts']'
Received message: {'window': {'start': '2024-12-07T21:04:30.000+02:00', 'end': '2024-12-07T21:05:30.000+02:00'}, 'avg_t': 55.866666666
66667, 'avg_h': 43.4, 'code': 104, 'message': "It's too hot", 'timestamp': '2024-12-07T21:16:05.661+02:00'}
Received message: {'window': {'start': '2024-12-07T21:05:00.000+02:00', 'end': '2024-12-07T21:06:00.000+02:00'}, 'avg_t': 56.625, 'avg
_h': 48.017857142857146, 'code': 104, 'message': "It's too hot", 'timestamp': '2024-12-07T21:19:01.800+02:00'}
Received message: {'window': {'start': '2024-12-07T21:05:30.000+02:00', 'end': '2024-12-07T21:06:30.000+02:00'}, 'avg_t': 54.929411764
70588, 'avg_h': 49.435294117647056, 'code': 104, 'message': "It's too hot", 'timestamp': '2024-12-07T21:22:02.844+02:00'}
Received message: {'window': {'start': '2024-12-07T21:06:00.000+02:00', 'end': '2024-12-07T21:07:00.000+02:00'}, 'avg_t': 54.104651162
790695, 'avg_h': 50.906976744186046, 'code': 104, 'message': "It's too hot", 'timestamp': '2024-12-07T21:25:25.814+02:00'}
Received message: {'window': {'start': '2024-12-07T21:06:30.000+02:00', 'end': '2024-12-07T21:07:30.000+02:00'}, 'avg_t': 48.428571428
57143, 'avg_h': 51.88095238095238, 'code': 104, 'message': "It's too hot", 'timestamp': '2024-12-07T21:28:24.658+02:00'}
```