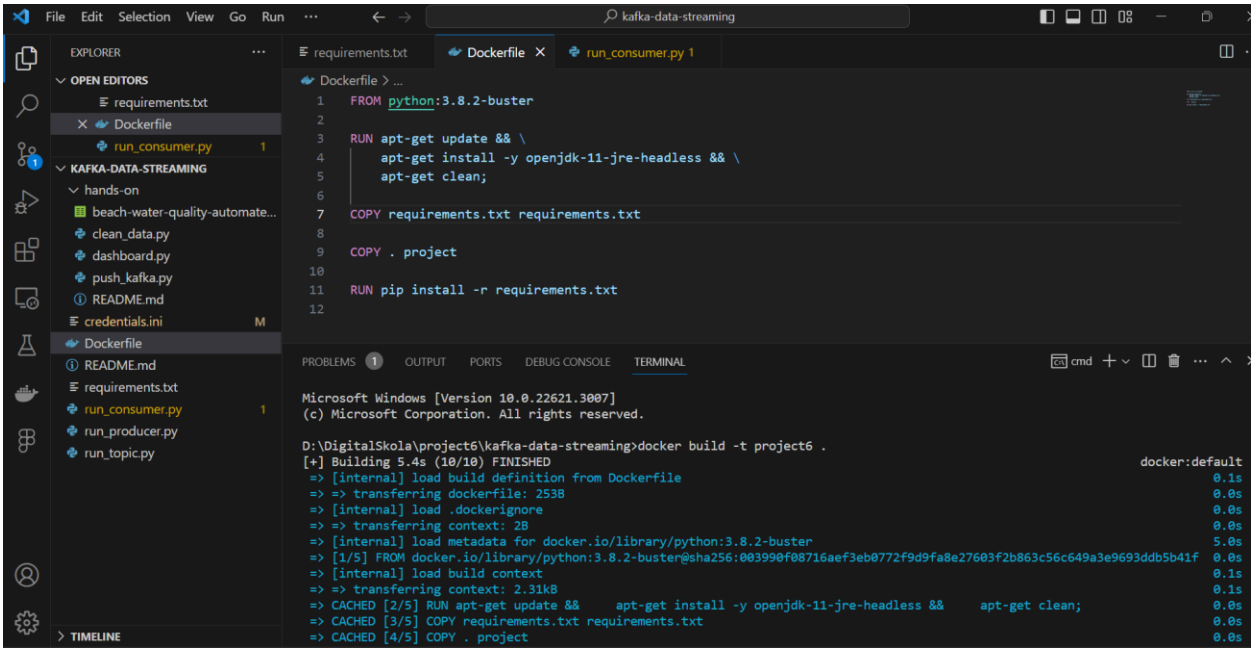
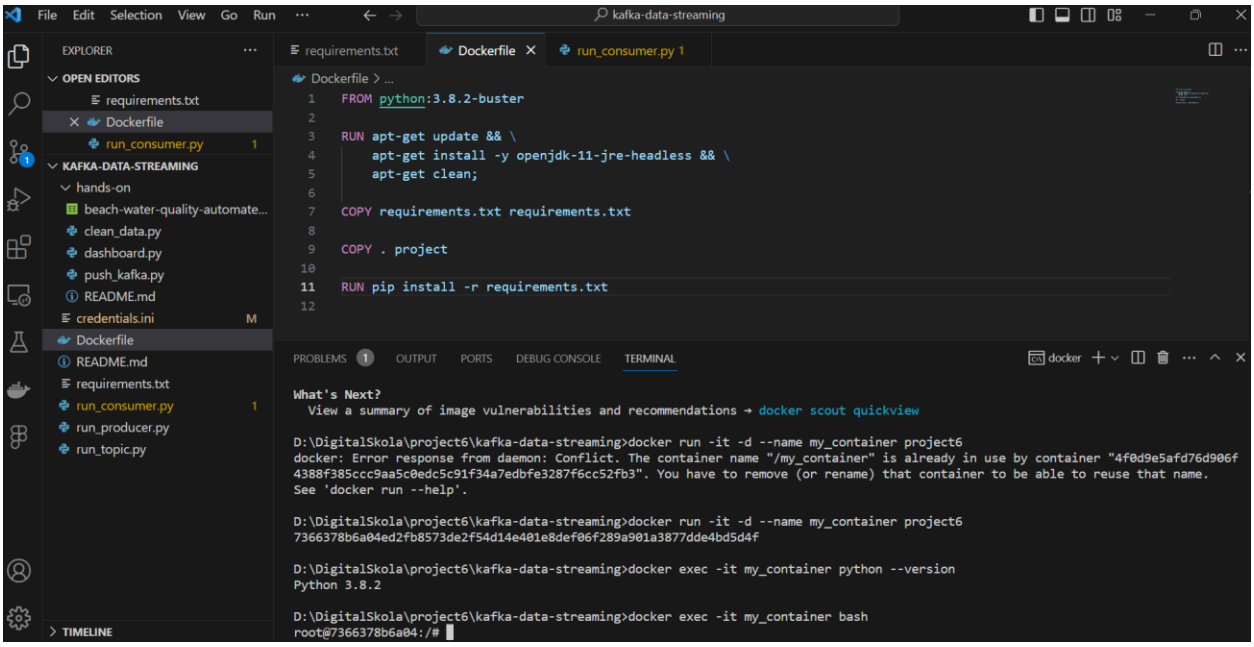
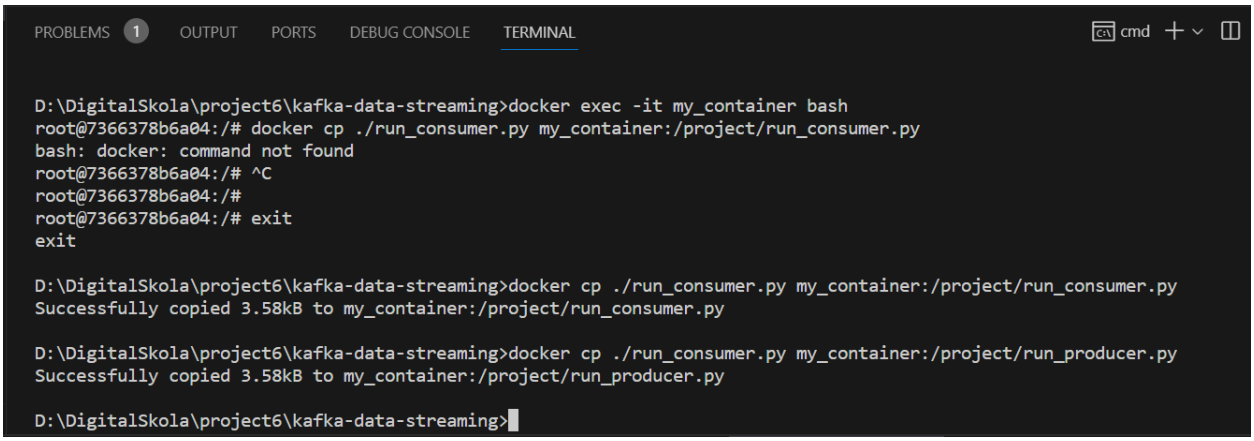
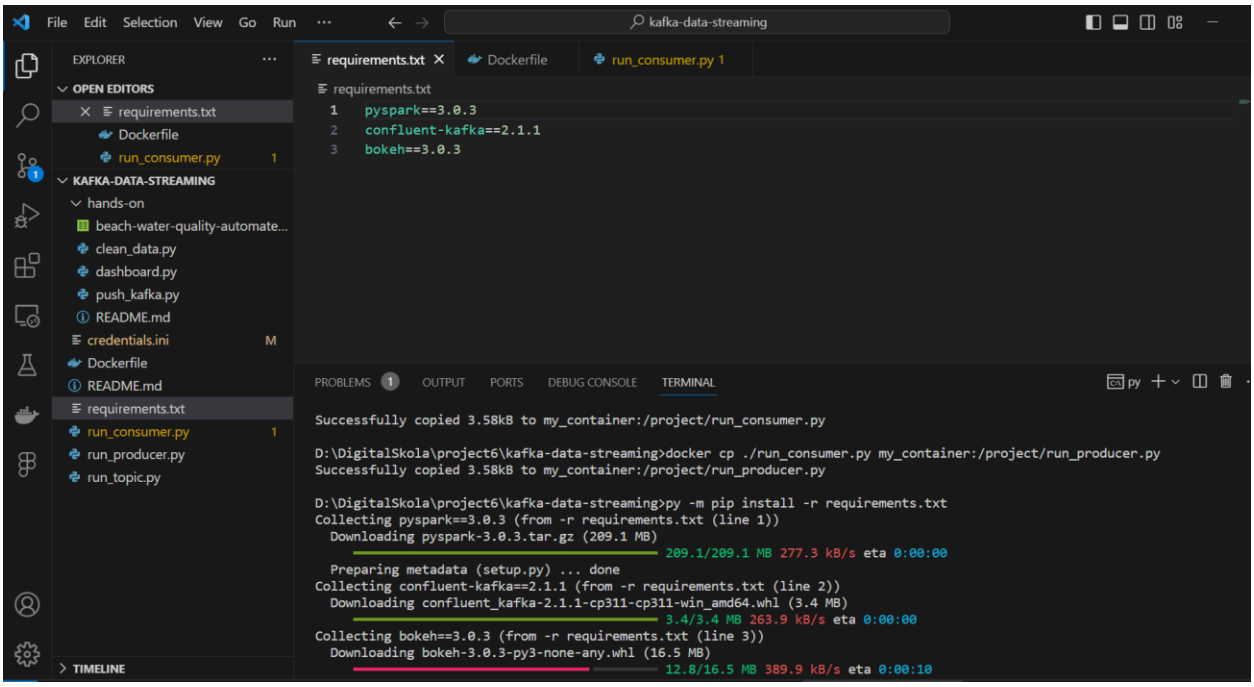
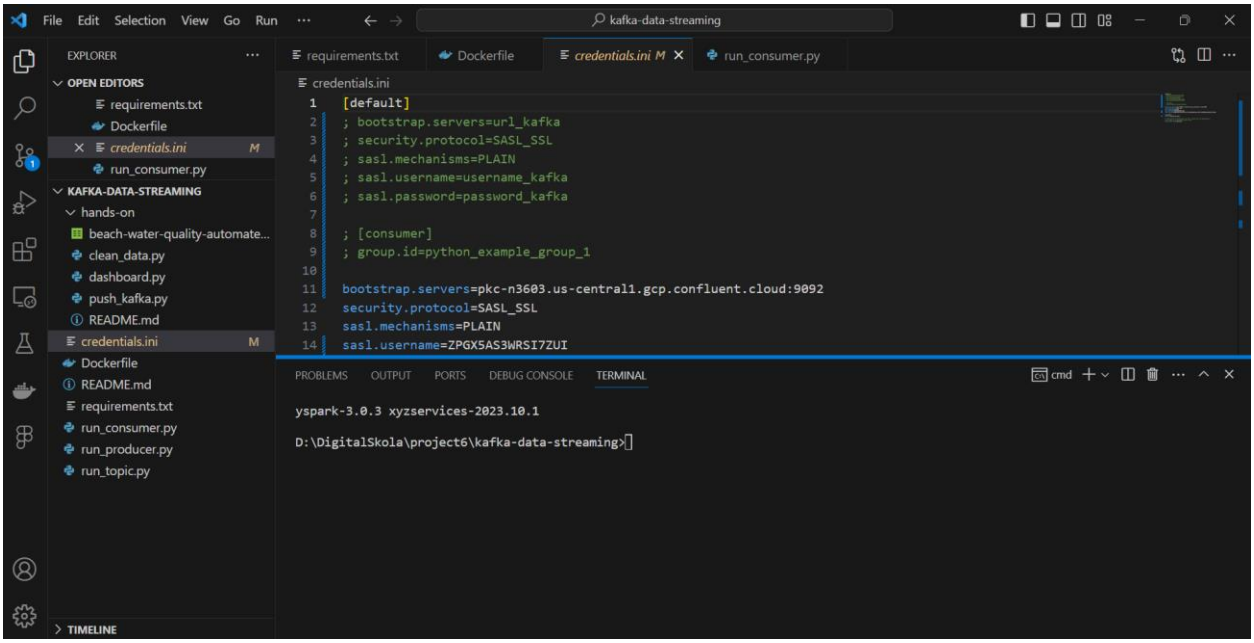
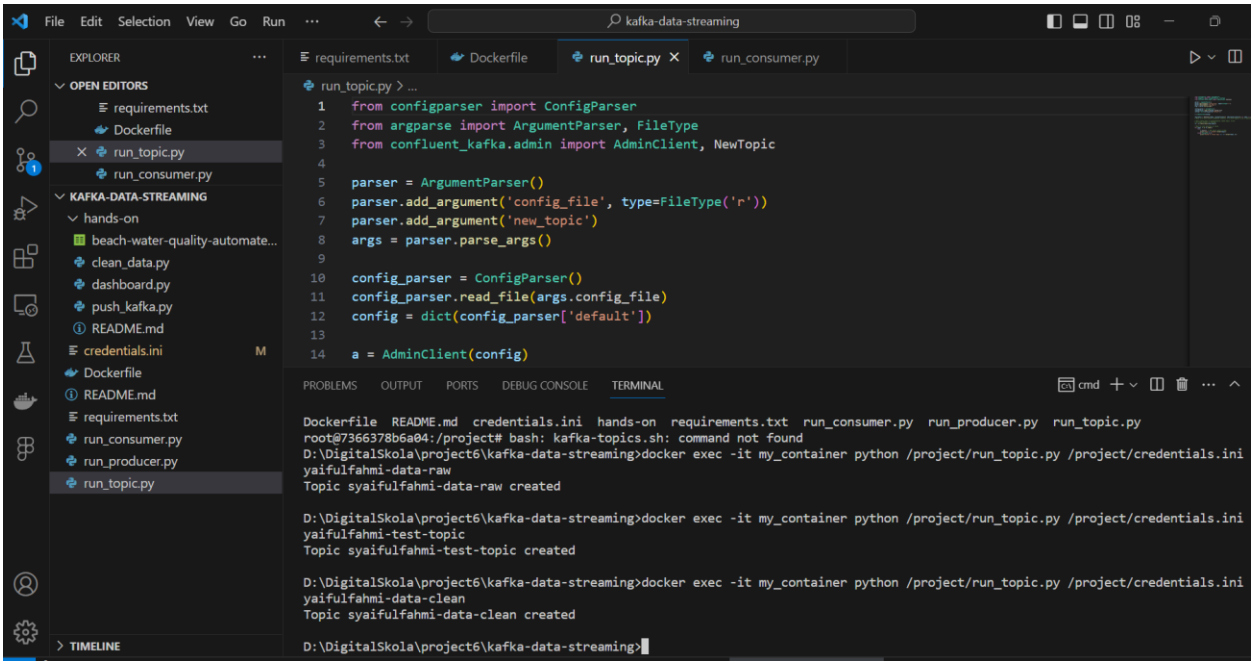

File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_consumer.py 1
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-automate...
clean_data.py
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py 1
run_producer.py
run_topic.py
TIMELINE
requirements.txt
Dockerfile
run_consumer.py 1
Dockerfile > ...
1 FROM python:3.8.2-buster
2
3 RUN apt-get update && \
4 apt-get install -y openjdk-11-jre-headless && \
5 apt-get clean;
6
7 COPY requirements.txt requirements.txt
8
9 COPY . project
10
11 RUN pip install -r requirements.txt
12
PROBLEMS 1 OUTPUT PORTS DEBUG CONSOLE TERMINAL
cmd
Microsoft Windows [Version 10.0.22621.3007]
(c) Microsoft Corporation. All rights reserved.
D:\DigitalSkola\project6\kafka-data-streaming>docker build -t project6 .
[+] Building 5.4s (10/10) FINISHED docker:default
=> [internal] load build definition from Dockerfile 0.1s
=> => transferring dockerfile: 253B 0.0s
=> [internal] load .dockerignore 0.0s
=> => transferring context: 2B 0.0s
=> [internal] load metadata for docker.io/library/python:3.8.2-buster 5.0s
=> [1/5] FROM docker.io/library/python:3.8.2-buster@sha256:003990f08716aef3eb0772f9d9fa8e27603f2b863c56c649a3e9693ddb5b41f 0.0s
=> [internal] load build context 0.1s
=> => transferring context: 2.31kB 0.1s
=> CACHED [2/5] RUN apt-get update && apt-get install -y openjdk-11-jre-headless && apt-get clean; 0.0s
=> CACHED [3/5] COPY requirements.txt requirements.txt 0.0s
=> CACHED [4/5] COPY . project 0.0s

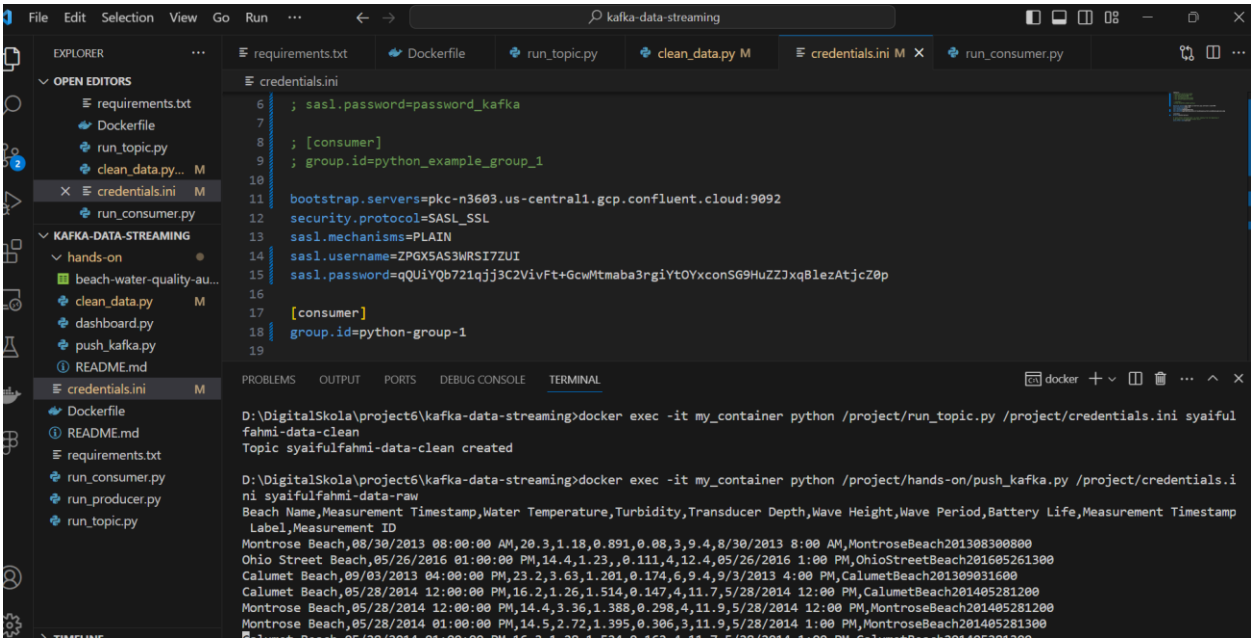

File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_consumer.py 1
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-automate...
clean_data.py
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py 1
run_producer.py
run_topic.py
TIMELINE
requirements.txt
Dockerfile
run_consumer.py 1
Dockerfile > ...
1 FROM python:3.8.2-buster
2
3 RUN apt-get update && \
4 apt-get install -y openjdk-11-jre-headless && \
5 apt-get clean;
6
7 COPY requirements.txt requirements.txt
8
9 COPY . project
10
11 RUN pip install -r requirements.txt
12
PROBLEMS 1 OUTPUT PORTS DEBUG CONSOLE TERMINAL
docker
What's Next?
View a summary of image vulnerabilities and recommendations → docker scout quickview
D:\DigitalSkola\project6\kafka-data-streaming>docker run -it -d --name my_container project6
docker: Error response from daemon: Conflict. The container name "/my_container" is already in use by container "4f0d9e5afd76d906f4388f385ccc9aa5c0edc5c91f34a7edbfe3287f6cc52fb3". You have to remove (or rename) that container to be able to reuse that name.
See 'docker run --help'.
D:\DigitalSkola\project6\kafka-data-streaming>docker run -it -d --name my_container project6
7366378b6a04ed2fb8573de2f54d14e401e8def06f289a901a3877dde4bd5d4f
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python --version
Python 3.8.2
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container bash
root@7366378b6a04:/#

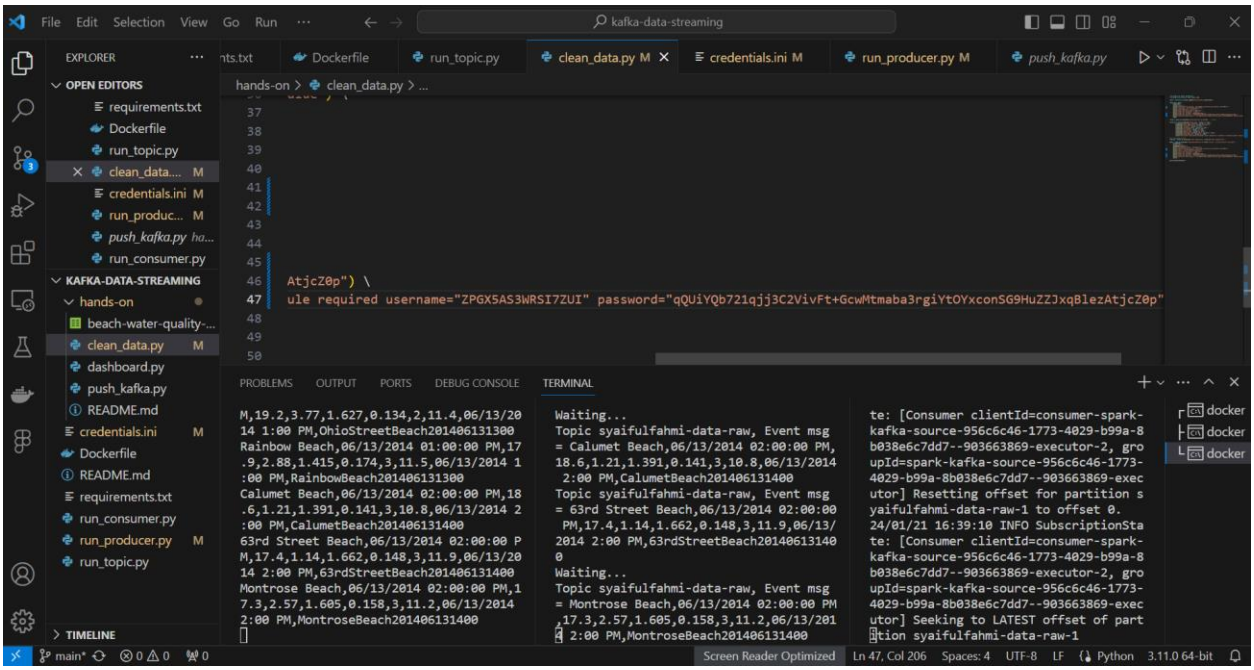

PROBLEMS 1
OUTPUT
PORTS
DEBUG CONSOLE
TERMINAL
cmd
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container bash
root@7366378b6a04:/# docker cp ./run_consumer.py my_container:/project/run_consumer.py
bash: docker: command not found
root@7366378b6a04:/# ^C
root@7366378b6a04:/#
root@7366378b6a04:/# exit
exit
D:\DigitalSkola\project6\kafka-data-streaming>docker cp ./run_consumer.py my_container:/project/run_consumer.py
Successfully copied 3.58kB to my_container:/project/run_consumer.py
D:\DigitalSkola\project6\kafka-data-streaming>docker cp ./run_consumer.py my_container:/project/run_producer.py
Successfully copied 3.58kB to my_container:/project/run_producer.py
D:\DigitalSkola\project6\kafka-data-streaming>

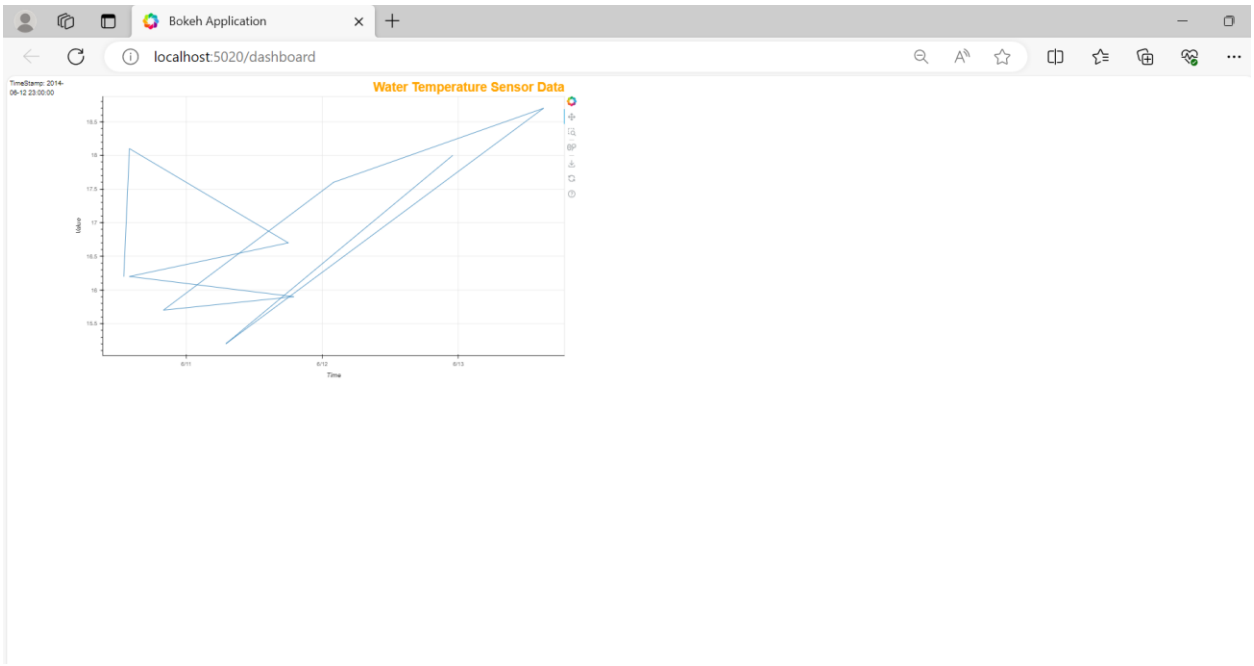

File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_consumer.py 1
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-automate...
clean_data.py
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py 1
run_producer.py
run_topic.py
TIMELINE
requirements.txt
Dockerfile
run_consumer.py 1
requirements.txt
1 pyspark==3.0.3
2 confluent-kafka==2.1.1
3 bokeh==3.0.3
PROBLEMS 1
OUTPUT
PORTS
DEBUG CONSOLE
TERMINAL
py
Successfully copied 3.58kB to my_container:/project/run_consumer.py
D:\DigitalSkola\project6\kafka-data-streaming>docker cp ./run_consumer.py my_container:/project/run_producer.py
Successfully copied 3.58kB to my_container:/project/run_producer.py
D:\DigitalSkola\project6\kafka-data-streaming>py -m pip install -r requirements.txt
Collecting pyspark==3.0.3 (from -r requirements.txt (line 1))
Downloading pyspark-3.0.3.tar.gz (209.1 MB)
209.1/209.1 MB 277.3 kB/s eta 0:00:00
Preparing metadata (setup.py) ... done
Collecting confluent-kafka==2.1.1 (from -r requirements.txt (line 2))
Downloading confluent_kafka-2.1.1-cp311-cp311-win_amd64.whl (3.4 MB)
3.4/3.4 MB 263.9 kB/s eta 0:00:00
Collecting bokeh==3.0.3 (from -r requirements.txt (line 3))
Downloading bokeh-3.0.3-py3-none-any.whl (16.5 MB)
12.8/16.5 MB 389.9 kB/s eta 0:00:10


File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
credentials.ini M
run_consumer.py
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-automate...
clean_data.py
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py
run_producer.py
run_topic.py
TIMELINE
requirements.txt
Dockerfile
credentials.ini M
run_consumer.py
credentials.ini
1 [default]
2 ; bootstrap.servers=url_kafka
3 ; security.protocol=SASL_SSL
4 ; sasl.mechanisms=PLAIN
5 ; sasl.username=username_kafka
6 ; sasl.password=password_kafka
7
8 ; [consumer]
9 ; group.id=python_example_group_1
10
11 bootstrap.servers=pkc-n3603.us-central1.gcp.confluent.cloud:9092
12 security.protocol=SASL_SSL
13 sasl.mechanisms=PLAIN
14 sasl.username=ZPGX5AS3WRSI7ZUI
PROBLEMS OUTPUT PORTS DEBUG CONSOLE TERMINAL
cmd
yspark-3.0.3 xyzservices-2023.10.1
D:\DigitalSkola\project6\kafka-data-streaming>


File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_topic.py
run_consumer.py
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-automate...
clean_data.py
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py
run_producer.py
run_topic.py
TIMELINE
requirements.txt
Dockerfile
run_topic.py
run_consumer.py
run_topic.py > ...
1 from configparser import ConfigParser
2 from argparse import ArgumentParser, FileType
3 from confluent_kafka.admin import AdminClient, NewTopic
4
5 parser = ArgumentParser()
6 parser.add_argument('config_file', type=FileType('r'))
7 parser.add_argument('new_topic')
8 args = parser.parse_args()
9
10 config_parser = ConfigParser()
11 config_parser.read_file(args.config_file)
12 config = dict(config_parser['default'])
13
14 a = AdminClient(config)
PROBLEMS OUTPUT PORTS DEBUG CONSOLE TERMINAL
cmd
Dockerfile README.md credentials.ini hands-on requirements.txt run_consumer.py run_producer.py run_topic.py
root@7366378b6a04:/project# bash: kafka-topics.sh: command not found
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python /project/run_topic.py /project/credentials.ini
yaifulfahmi-data-raw
Topic syaifulfahmi-data-raw created
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python /project/run_topic.py /project/credentials.ini
yaifulfahmi-test-topic
Topic syaifulfahmi-test-topic created
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python /project/run_topic.py /project/credentials.ini
yaifulfahmi-data-clean
Topic syaifulfahmi-data-clean created
D:\DigitalSkola\project6\kafka-data-streaming>


File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_topic.py
clean_data.py... M
credentials.ini M
run_consumer.py
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-au...
clean_data.py M
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py
run_producer.py
run_topic.py
requirements.txt
Dockerfile
run_topic.py
clean_data.py M
credentials.ini M
run_consumer.py
credentials.ini
; sasl.password=password_kafka
; [consumer]
; group.id=python_example_group_1
bootstrap.servers=pkc-n3603.us-central1.gcp.confluent.cloud:9092
security.protocol=SASL_SSL
sasl.mechanisms=PLAIN
sasl.username=ZPGX5AS3WRSI7ZUI
sasl.password=qQUiYQb721qjj3C2VivFt+GcwMtmaba3rgiYtOYxconSG9HuZZJxqBlezAtjcZ0p
[consumer]
group.id=python-group-1
PROBLEMS OUTPUT PORTS DEBUG CONSOLE TERMINAL
docker
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python /project/run_topic.py /project/credentials.ini syaifulfahmi-data-clean
Topic syaifulfahmi-data-clean created
D:\DigitalSkola\project6\kafka-data-streaming>docker exec -it my_container python /project/hands-on/push_kafka.py /project/credentials.ini syaifulfahmi-data-raw
Beach Name,Measurement Timestamp,Water Temperature,Turbidity,Transducer Depth,Wave Height,Wave Period,Battery Life,Measurement Timestamp Label,Measurement ID
Montrose Beach,08/30/2013 08:00:00 AM,20.3,1.18,0.891,0.08,3,9.4,8/30/2013 8:00 AM,MontroseBeach201308300800
Ohio Street Beach,05/26/2016 01:00:00 PM,14.4,1.23,,0.111,4,12.4,05/26/2016 1:00 PM,OhioStreetBeach201605261300
Calumet Beach,09/03/2013 04:00:00 PM,23.2,3.63,1.201,0.174,6,9.4,9/3/2013 4:00 PM,CalumetBeach201309031600
Calumet Beach,05/28/2014 12:00:00 PM,16.2,1.26,1.514,0.147,4,11.7,5/28/2014 12:00 PM,CalumetBeach201405281200
Montrose Beach,05/28/2014 12:00:00 PM,14.4,3.36,1.388,0.298,4,11.9,5/28/2014 12:00 PM,MontroseBeach201405281200
Montrose Beach,05/28/2014 01:00:00 PM,14.5,2.72,1.395,0.306,3,11.9,5/28/2014 1:00 PM,MontroseBeach201405281300


File Edit Selection View Go Run
kafka-data-streaming
EXPLORER
OPEN EDITORS
requirements.txt
Dockerfile
run_topic.py
clean_data.... M
credentials.ini M
run_produc... M
push_kafka.py ha...
run_consumer.py
KAFKA-DATA-STREAMING
hands-on
beach-water-quality-...
clean_data.py M
dashboard.py
push_kafka.py
README.md
credentials.ini M
Dockerfile
README.md
requirements.txt
run_consumer.py
run_producer.py M
run_topic.py
TIMELINE
Dockerfile
run_topic.py
clean_data.py M
credentials.ini M
run_producer.py M
push_kafka.py
hands-on > clean_data.py > ...
AtjcZ0p") \
ule required username="ZPGX5AS3WRSI7ZUI" password="qQUiYQb721qjj3C2VivFt+GcwMtmaba3rgiYtOYxconSG9HuZZJxqBlezAtjcZ0p"
PROBLEMS OUTPUT PORTS DEBUG CONSOLE TERMINAL
M,19.2,3.77,1.627,0.134,2,11.4,06/13/2014 1:00 PM,OhioStreetBeach201406131300
Rainbow Beach,06/13/2014 01:00:00 PM,17.9,2.88,1.415,0.174,3,11.5,06/13/2014 1:00 PM,RainbowBeach201406131300
Calumet Beach,06/13/2014 02:00:00 PM,18.6,1.21,1.391,0.141,3,10.8,06/13/2014 2:00 PM,CalumetBeach201406131400
63rd Street Beach,06/13/2014 02:00:00 PM,17.4,1.14,1.662,0.148,3,11.9,06/13/2014 2:00 PM,63rdStreetBeach201406131400
Montrose Beach,06/13/2014 02:00:00 PM,17.3,2.57,1.605,0.158,3,11.2,06/13/2014 2:00 PM,MontroseBeach201406131400
Waiting...
Topic syaifulfahmi-data-raw, Event msg = Calumet Beach,06/13/2014 02:00:00 PM,18.6,1.21,1.391,0.141,3,10.8,06/13/2014 2:00 PM,CalumetBeach201406131400
Topic syaifulfahmi-data-raw, Event msg = 63rd Street Beach,06/13/2014 02:00:00 PM,17.4,1.14,1.662,0.148,3,11.9,06/13/2014 2:00 PM,63rdStreetBeach201406131400
Waiting...
Topic syaifulfahmi-data-raw, Event msg = Montrose Beach,06/13/2014 02:00:00 PM,17.3,2.57,1.605,0.158,3,11.2,06/13/2014 2:00 PM,MontroseBeach201406131400
te: [Consumer clientId=consumer-spark-kafka-source-956c6c46-1773-4029-b99a-8b038e6c7dd7--903663869-executor-2, groupId=spark-kafka-source-956c6c46-1773-4029-b99a-8b038e6c7dd7--903663869-executor] Resetting offset for partition syaifulfahmi-data-raw-1 to offset 0.
24/01/21 16:39:10 INFO SubscriptionState: [Consumer clientId=consumer-spark-kafka-source-956c6c46-1773-4029-b99a-8b038e6c7dd7--903663869-executor-2, groupId=spark-kafka-source-956c6c46-1773-4029-b99a-8b038e6c7dd7--903663869-executor] Seeking to LATEST offset of partition syaifulfahmi-data-raw-1
docker
docker
docker
main*
Screen Reader Optimized
Ln 47, Col 206
Spaces: 4
UTF-8
LF
Python
3.11.0 64-bit


Bokeh Application
localhost:5020/dashboard
TimeStamp: 2014-06-12 23:00:00
Water Temperature Sensor Data
Value
Time
6/11
6/12
6/13