



or@DESKTOP-FBRFLFL: ~



```
sent 145000 records...
sent 145200 records...
sent 145400 records...
sent 145600 records...
sent 145800 records...
sent 146000 records...
sent 146200 records...
sent 146400 records...
sent 146600 records...
sent 146800 records...
sent 147000 records...
sent 147200 records...
sent 147400 records...
sent 147600 records...
sent 147800 records...
sent 148000 records...
sent 148200 records...
sent 148400 records...
sent 148600 records...
sent 148800 records...
sent 149000 records...
sent 149200 records...
sent 149400 records...
sent 149600 records...
sent 149800 records...
sent 150000 records...
sent 150200 records...
sent 150400 records...
done. total sent: 150482
(venv) or@DESKTOP-FBRFLFL:~$ |
```

```
or@DESKTOP-FBRFLFL:~$ docker exec -it ce72c078bf99 /opt/kafka/bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic wikimedia_topic_1 --from-beginning --max-messages 5
done. total sent: 150482
{"VIN (1-10)": "KM8K33AGXL", "County": "King", "City": "Seattle", "State": "WA", "Postal Code": "98103", "Model Year": "2020", "Make": "HYUNDAI", "Model": "KONA", "Electric Vehicle Type": "Battery Electric Vehicle (BEV)", "Clean Alternative Fuel Vehicle (CAEV) Eligibility": "Clean Alternative Fuel Vehicle Eligible", "Electric Range": "258", "Base MSRP": "0", "Legislative District": "43", "DOL Vehicle ID": "249675142", "Vehicle Location": "POINT (-122.34301 47.659185)", "Electric Utility": "CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)", "2020 Census Tract": "53033004800"}
{"VIN (1-10)": "1C4RJYB61N", "County": "King", "City": "Bothell", "State": "WA", "Postal Code": "98011", "Model Year": "2022", "Make": "JEEP", "Model": "GRAND CHEROKEE", "Electric Vehicle Type": "Plug-in Hybrid Electric Vehicle (PHEV)", "Clean Alternative Fuel Vehicle (CAEV) Eligibility": "Not eligible due to low battery range", "Electric Range": "25", "Base MSRP": "0", "Legislative District": "1", "DOL Vehicle ID": "233928502", "Vehicle Location": "POINT (-122.20578 47.762405)", "Electric Utility": "PUGET SOUND ENERGY INC|CITY OF TACOMA - (WA)", "2020 Census Tract": "53033021804"}
{"VIN (1-10)": "1C4RJYD61P", "County": "Yakima", "City": "Yakima", "State": "WA", "Postal Code": "98908", "Model Year": "2023", "Make": "JEEP", "Model": "GRAND CHEROKEE", "Electric Vehicle Type": "Plug-in Hybrid Electric Vehicle (PHEV)", "Clean Alternative Fuel Vehicle (CAEV) Eligibility": "Not eligible due to low battery range", "Electric Range": "25", "Base MSRP": "0", "Legislative District": "14", "DOL Vehicle ID": "229675939", "Vehicle Location": "POINT (-120.6027202 46.5965625)", "Electric Utility": "PACIFICORP", "2020 Census Tract": "53077002900"}
{"VIN (1-10)": "5YJ3E1EA7J", "County": "King", "City": "Kirkland", "State": "WA", "Postal Code": "98034", "Model Year": "2018", "Make": "TESLA", "Model": "MODEL 3", "Electric Vehicle Type": "Battery Electric Vehicle (BEV)", "Clean Alternative Fuel Vehicle (CAEV) Eligibility": "Clean Alternative Fuel Vehicle Eligible", "Electric Range": "215", "Base MSRP": "0", "Legislative District": "45", "DOL Vehicle ID": "104714466", "Vehicle Location": "POINT (-122.209285 47.71124)", "Electric Utility": "PUGET SOUND ENERGY INC|CITY OF TACOMA - (WA)", "2020 Census Tract": "53033021903"}
{"VIN (1-10)": "WBY7Z8C5XJ", "County": "Thurston", "City": "Olympia", "State": "WA", "Postal Code": "98501", "Model Year": "2018", "Make": "BMW", "Model": "I3", "Electric Vehicle Type": "Plug-in Hybrid Electric Vehicle (PHEV)", "Clean Alternative Fuel Vehicle (CAEV) Eligibility": "Clean Alternative Fuel Vehicle Eligible", "Electric Range": "97", "Base MSRP": "0", "Legislative District": "22", "DOL Vehicle ID": "185498386", "Vehicle Location": "POINT (-122.89692 47.043535)", "Electric Utility": "PUGET SOUND ENERGY INC", "2020 Census Tract": "53067010700"}
Processed a total of 5 messages
(venv) or@DESKTOP-FBRFLFL:~$
```

Mor-Greenberg/Spark-Stream: x Home x spark_streaming_wikimedia x +

localhost:8888/notebooks/spark_streaming_wikimedia.ipynb

mama Outlook ChatGPT Coliru google colab (4) Feed | LinkedIn AWS Academy Learn... NotebookLM GitHub

Jupyter spark_streaming_wikimedia Last Checkpoint: 3 minutes ago

File Edit View Run Kernel Settings Help Trusted

JupyterLab Python 3 (pykernel)

```
.start()

query

26/01/08 18:39:26 WARN ResolveWriteToStream: Temporary checkpoint location created which is deleted normally when the query didn't fail: /tmp/temporary-92c842de-67a0-4c11-81ee-e19996fca76b. If it's required to delete it under any circumstances, please set spark.sql.streaming.forceDeleteTempCheckpointLocation to true. Important to know deleting temp checkpoint folder is best effort.
26/01/08 18:39:26 WARN ResolveWriteToStream: spark.sql.adaptive.enabled is not supported in streaming DataFrames/Datasets and will be disabled.

[3]: <pyspark.sql.streaming.query.StreamingQuery at 0x796851fa1f40>

26/01/08 18:39:27 WARN MicroBatchExecution: Disabling AQE since AQE is not supported in stateful workloads.

Batch: 0

+-----+
|City|count|
+-----+
|Seattle|5610|
|Bellevue|2088|
|Tukwila|1773|
+-----+
```

6:40 PM 1/8/2026