

Load data from Kafka to Hadoop

<Steps to run the python file to load data from Kafka>

- ➔ Copy the python file to /user/livy folder using the following code
 - **Hadoop fs -ls -put spark_kafka_to_local /user/livy**
- ➔ Once the python file is copied we need to run the below code :
 - **export SPARK_KAFKA_VERSION=0.10**
spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.12:3.1.2
spark_kafka_to_local.py
- ➔ The above code is also providing the dependencies for spark to connect to the kafka {packages and export command}

<Steps to load the data into Hadoop>

- ➔ Copy the python file from local to hadoop to the /user/livy folder using the following code
 - **Hadoop fs -ls -put spark_local_flatten /user/livy**
- ➔ Once the python file is copied we need to run the below code :
 - **export SPARK_KAFKA_VERSION=0.10**
spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.12:3.1.2
spark_local_flatten.py
- ➔ The above code is also providing the dependencies for spark to connect to the kafka {packages and export command}

<Screenshot of the data>

➔ Screenshot of the copied data from kafka to hadoop

```

[{"value_str": "\"(\"customer_id\": \"f196728510\", \"app_version\": \"4.3.21\", \"OS_version\": \"iOS\", \"lat\": \"73.982053\", \"lon\": \"109.376113\", \"page_id\": \"e7bc5fb2-1231-11eb-adc1-0242ac120002\", \"button_id\": \"f1e99492-17ae-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"Yes\", \"is_scroll_up\": \"Yes\", \"is_scroll_down\": \"Yes\", \"timestamp\": \"2020-03-24 06:09:12\")\"", "value_str": "\"(\"customer_id\": \"f17156145\", \"app_version\": \"4.1.4\", \"OS_version\": \"Android\", \"lat\": \"-24.5899605\", \"lon\": \"67.624771\", \"page_id\": \"de545711-3914-4450-8c11-b17b8daab5e1\", \"button_id\": \"f1cfa68a-1231-11eb-adc1-0242ac120002\", \"is_button_click\": \"No\", \"is_page_view\": \"No\", \"is_scroll_up\": \"No\", \"is_scroll_down\": \"Yes\", \"timestamp\": \"2020-03-24 06:09:12\")\"", "value_str": "\"(\"customer_id\": \"f228999996\", \"app_version\": \"4.2.39\", \"OS_version\": \"iOS\", \"lat\": \"-69.1155955\", \"lon\": \"-143.163289\", \"page_id\": \"de545711-3914-4450-8c11-b17b8daab5e1\", \"button_id\": \"f1e99492-17ae-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"Yes\", \"is_scroll_up\": \"Yes\", \"is_scroll_down\": \"Yes\", \"timestamp\": \"2020-06-26 09:08:00\")\"", "value_str": "\"(\"customer_id\": \"f322444444\", \"app_version\": \"3.2.24\", \"OS_version\": \"Android\", \"lat\": \"-47.734696\", \"lon\": \"65.891429\", \"page_id\": \"f328829e-17ae-11eb-adc1-0242ac120002\", \"button_id\": \"f1cfa68a-1231-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"Yes\", \"is_scroll_up\": \"Yes\", \"is_scroll_down\": \"No\", \"timestamp\": \"2020-04-11 10:41:05\")\"", "value_str": "\"(\"customer_id\": \"f42165322\", \"app_version\": \"4.1.13\", \"OS_version\": \"iOS\", \"lat\": \"80.8403425\", \"lon\": \"-86.627180\", \"page_id\": \"de545711-3914-4450-8c11-b17b8daab5e1\", \"button_id\": \"f1e99492-17ae-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"No\", \"is_scroll_up\": \"No\", \"is_scroll_down\": \"No\", \"timestamp\": \"2020-01-27 14:03:00\")\"", "value_str": "\"(\"customer_id\": \"f21222347\", \"app_version\": \"1.2.27\", \"OS_version\": \"iOS\", \"lat\": \"-3.6452735\", \"lon\": \"173.446829\", \"page_id\": \"f328829e-17ae-11eb-adc1-0242ac120002\", \"button_id\": \"f1e99492-17ae-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"No\", \"is_scroll_up\": \"No\", \"is_scroll_down\": \"No\", \"timestamp\": \"2020-07-13 15:48:28\")\"", "value_str": "\"(\"customer_id\": \"f173363636\", \"app_version\": \"1.1.2\", \"OS_version\": \"iOS\", \"lat\": \"52.6143785\", \"lon\": \"-4.783003\", \"page_id\": \"e7bc5fb2-1231-11eb-adc1-0242ac120002\", \"button_id\": \"f1cfa68a-1231-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"No\", \"is_scroll_up\": \"No\", \"is_scroll_down\": \"Yes\", \"timestamp\": \"2020-08-16 13:27:47\")\"", "value_str": "\"(\"customer_id\": \"f25631522\", \"app_version\": \"2.1.6\", \"OS_version\": \"Android\", \"lat\": \"73.076750\", \"lon\": \"13.785813\", \"page_id\": \"f328829e-17ae-11eb-adc1-0242ac120002\", \"button_id\": \"f1e99492-17ae-11eb-adc1-0242ac120002\", \"is_button_click\": \"No\", \"is_page_view\": \"Yes\", \"is_scroll_up\": \"Yes\", \"is_scroll_down\": \"Yes\", \"timestamp\": \"2020-01-23 08:41:36\")\"", "value_str": "\"(\"customer_id\": \"f92937542\", \"app_version\": \"3.2.16\", \"OS_version\": \"iOS\", \"lat\": \"-7.0993865\", \"lon\": \"-40.763662\", \"page_id\": \"de545711-3914-4450-8c11-b17b8daab5e1\", \"button_id\": \"f1cfa68a-1231-11eb-adc1-0242ac120002\", \"is_button_click\": \"Yes\", \"is_page_view\": \"No\", \"is_scroll_up\": \"No\", \"is_scroll_down\": \"No\", \"timestamp\": \"2020-01-23 08:41:36\")"}]
```

➔ Screenshot of the flattened data [loading to Hadoop]

```
hadoop@ip-172-31-40-21:~$
47743967,3,3.20,iOS,-13.990733,-178.175295,b328829e-17ae-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,Yes,No,No,No,
46894536,4,2.31,iOS,-59.551650,115.343099,e7bc5fb2-1231-11eb-adc1-0242ac120002,fcba68aa-1231-11eb-adc1-0242ac120002,Yes,No,Yes,Yes,
87861168,4,4.27,iOS,-9.0299885,87.753166,de545711-3914-4450-8c11-b17b8dabb5e1,fcba68aa-1231-11eb-adc1-0242ac120002,Yes,No,No,Yes,
98750961,1,4.24,Android,-67.056611,-42.210252,b328829e-17ae-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,Yes,No,
34175709,2,4.1,iOS,-65.870590,112.347735,de545711-3914-4450-8c11-b17b8dabb5e1,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,No,Yes,Yes,
35577566,2,2.19,Android,29.5281265,28.317591,e7bc5fb2-1231-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,Yes,No,
39790044,3,3.32,iOS,60.0575415,-36.785274,b328829e-17ae-11eb-adc1-0242ac120002,fcba68aa-1231-11eb-adc1-0242ac120002,No,Yes,No,
49451715,2,3.8,iOS,77.0398825,174.942366,e7bc5fb2-1231-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,Yes,No,Yes,
87559592,2,2.17,Android,-72.2719865,-66.732794,e7bc5fb2-1231-11eb-adc1-0242ac120002,fcba68aa-1231-11eb-adc1-0242ac120002,No,Yes,No,
20011383,2,1.30,Android,38.3581915,27.925051,b328829e-17ae-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,No,No,Yes,
83519078,4,2.9,iOS,78.4456805,-83.204209,b328829e-17ae-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,No,No,Yes,No,
35953253,1,3.11,iOS,29.422319,106.217372,de545711-3914-4450-8c11-b17b8dabb5e1,fcba68aa-1231-11eb-adc1-0242ac120002,No,No,Yes,Yes,
79134831,1,3.25,iOS,-43.676692,-13.693995,de545711-3914-4450-8c11-b17b8dabb5e1,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,Yes,No,
74932660,2,4.11,Android,-16.258038,-67.903992,b328829e-17ae-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,No,No,No,Yes,
48948016,1,2.37,Android,28.116246,80.131224,b328829e-17ae-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,Yes,No,No,Yes,
70449656,2,3.23,Android,-38.571419,102.352612,de545711-3914-4450-8c11-b17b8dabb5e1,e1e99492-17ae-11eb-adc1-0242ac120002,No,Yes,No,Yes,
41398860,4,1.10,Android,-58.4709445,-56.966952,e7bc5fb2-1231-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,No,No,Yes,Yes,
35844721,2,3.22,Android,0.997151,-56.869833,de545711-3914-4450-8c11-b17b8dabb5e1,fcba68aa-1231-11eb-adc1-0242ac120002,Yes,Yes,No,Yes,
49195530,4,1.38,Android,-65.173138,-45.282215,de545711-3914-4450-8c11-b17b8dabb5e1,fcba68aa-1231-11eb-adc1-0242ac120002,No,No,No,Yes,
92467259,1,1.28,iOS,-13.8285225,166.976040,e7bc5fb2-1231-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,No,No,No,
39856207,4,4.21,iOS,-45.9151095,-93.708618,de545711-3914-4450-8c11-b17b8dabb5e1,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,No,No,Yes,
47664043,3,3.14,Android,-19.842446,117.930952,e7bc5fb2-1231-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,Yes,No,No,Yes,
38725955,1,1.32,iOS,-86.7711605,-133.680955,e7bc5fb2-1231-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,Yes,Yes,No,
52083661,1,3.6,Android,-20.051056,-113.408925,b328829e-17ae-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,No,No,Yes,No,
21248817,2,2.14,iOS,25.7465345,-115.096705,e7bc5fb2-1231-11eb-adc1-0242ac120002,e1e99492-17ae-11eb-adc1-0242ac120002,Yes,No,Yes,No,
32887975,2,3.23,Android,-14.898615,-168.408479,b328829e-17ae-11eb-adc1-0242ac120002,a95dd57b-779f-49db-819d-b6960483e554,Yes,No,Yes,No,
```