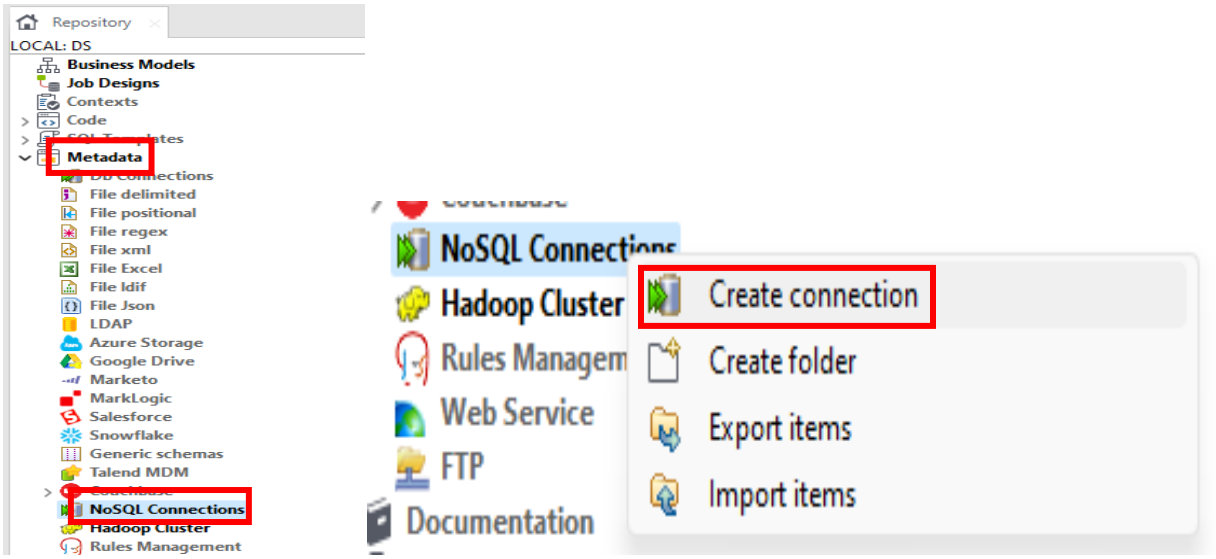
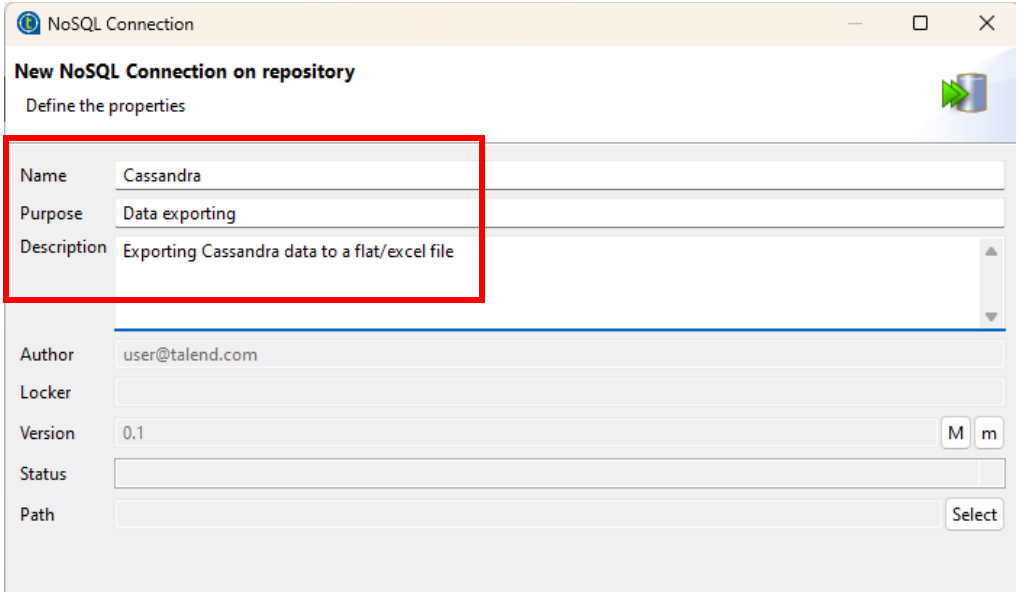
Miguel Soniel & Jan Edgar Tupas CS172 BM1

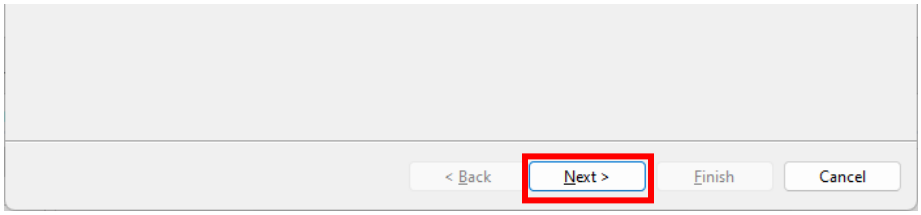
SA1 DOCUMENTATION

We will export data from a Cassandra keyspace to a flat/excel file using Talend. First, we create a NoSQL connection. Open the Metadata tab. Then we will right-click on the NoSQL Connections option and click Create connection.

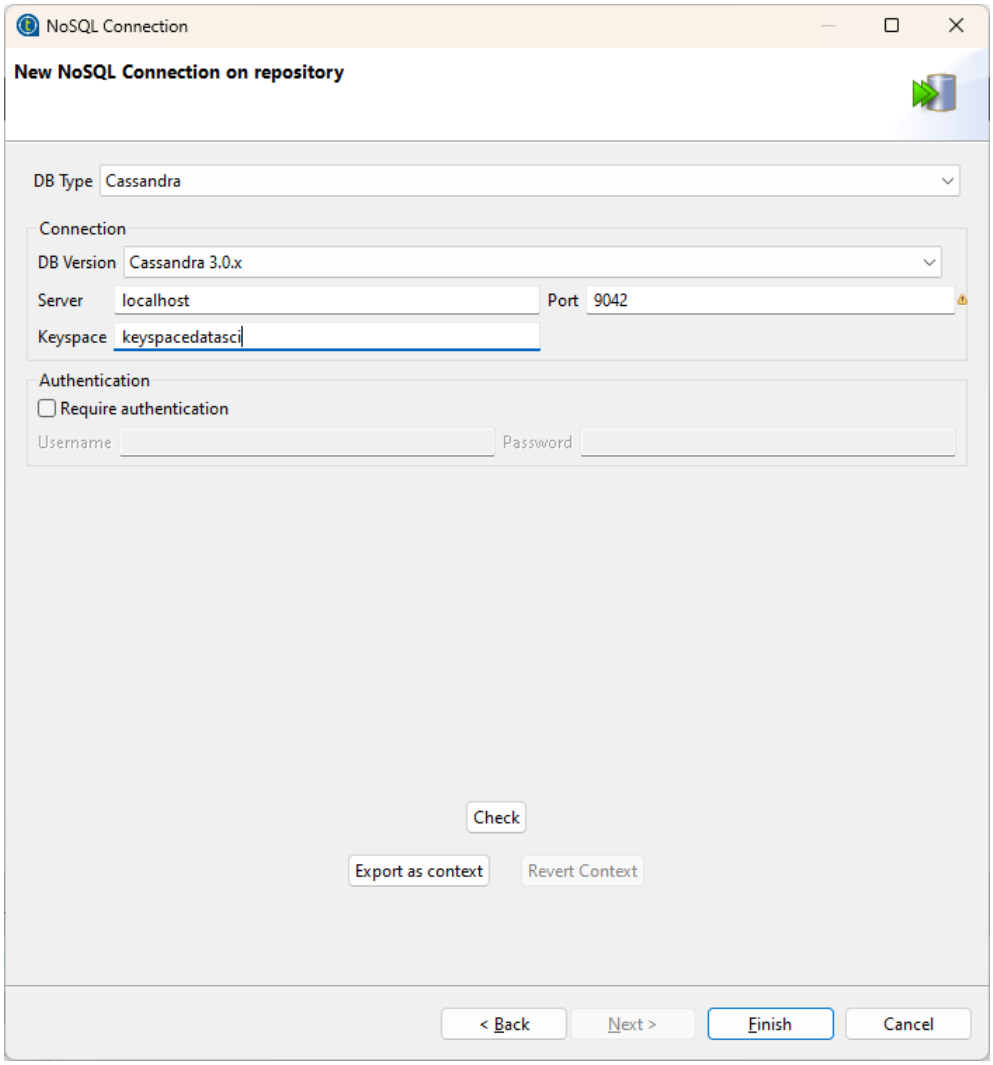


We will enter a name, purpose, and description for the connection. Here we used Cassandra for the name, then Data exporting for the purpose, and Exporting Cassandra data to a flat/excel file for the description. Then we click next.

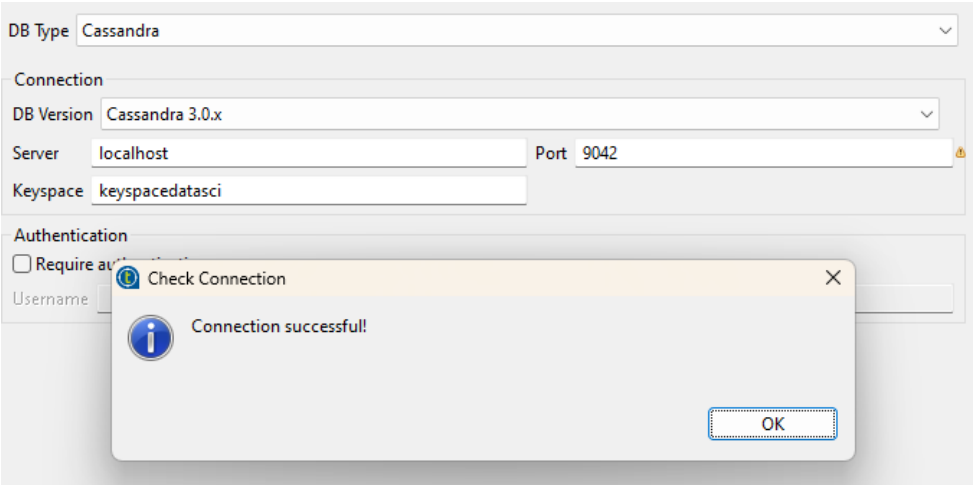




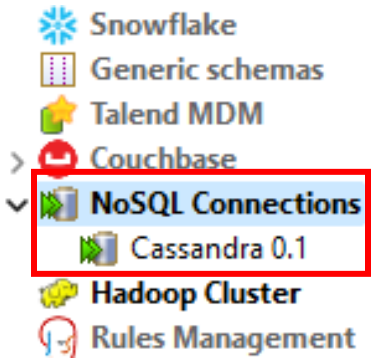
Then in the next tab, we set the DB Type to Cassandra, the DB Version to 3.0.x because our Cassandra version is 3.11.16, the server to localhost, and the port to 9042. Then we set the keyspace to the target keyspace where we want to get data. We set it to keyspacedatasci because that is the keyspace where the table of the data is located.



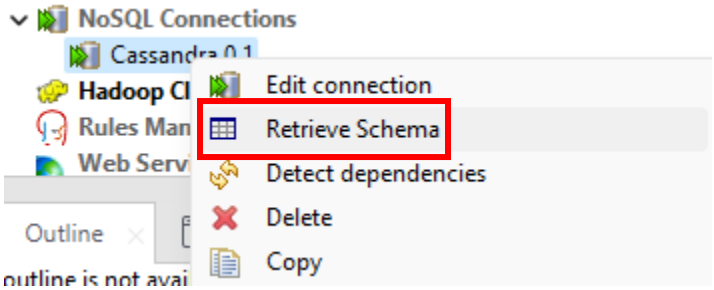
We click on check to test if the connection works. It shows that the connection is successful.



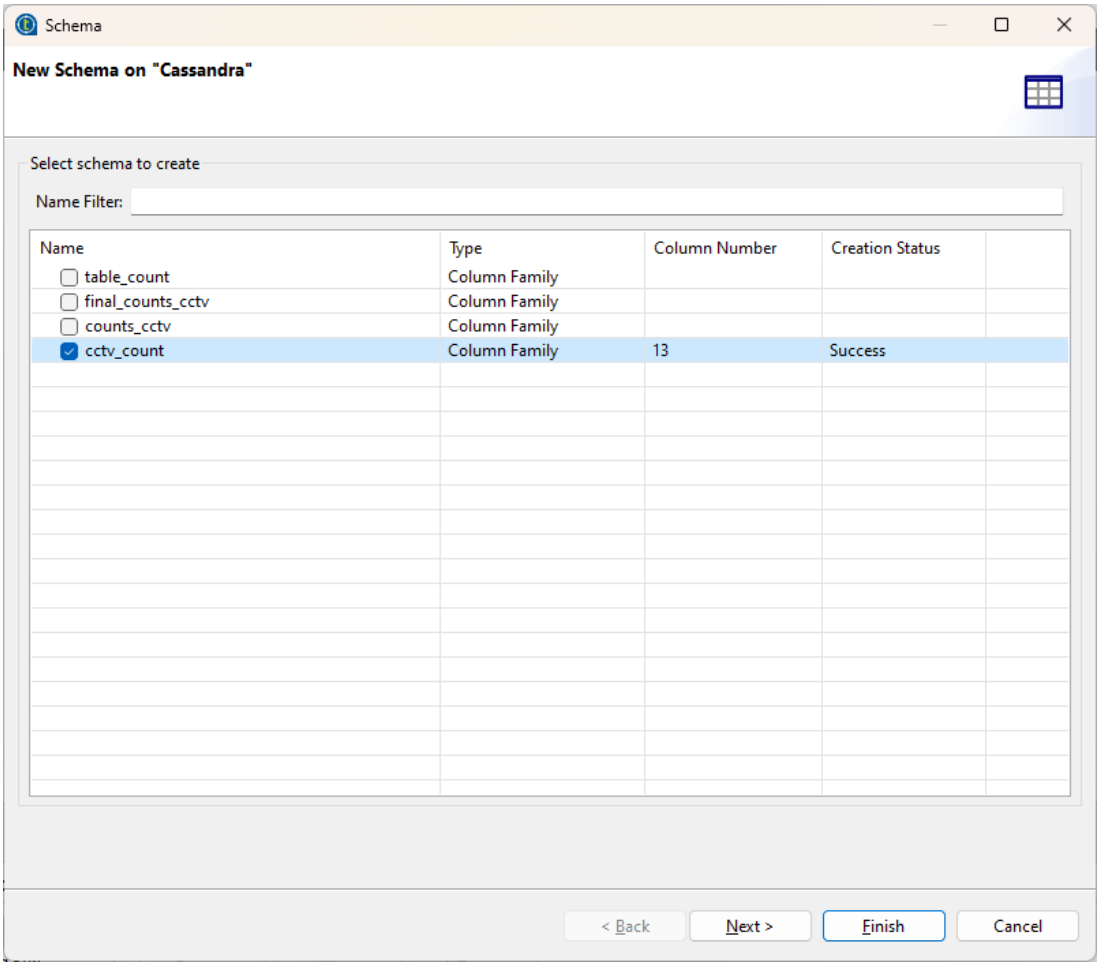
Now we have our Cassandra connection. It can be seen below the NoSQL Connections tab.



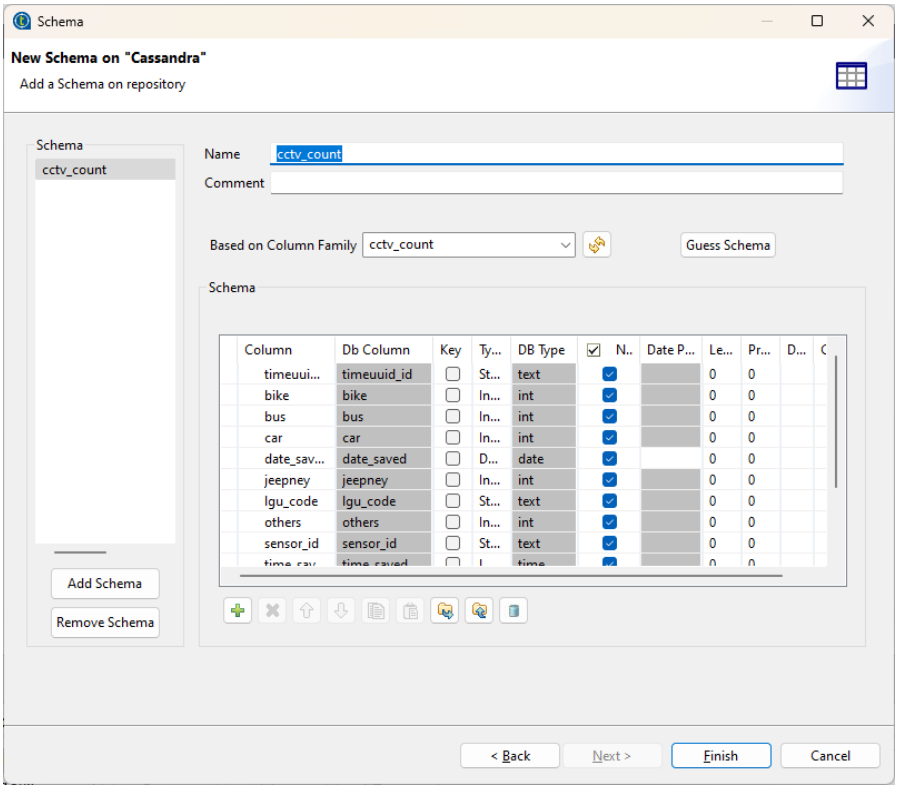
We now need to get the schema. We will right click on the Cassandra connection then click retrieve schema.



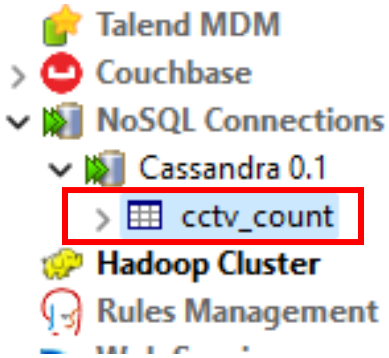
We can see the tables inside our Cassandra keyspace. We will get data from cctv_count table. So we select it and click next.



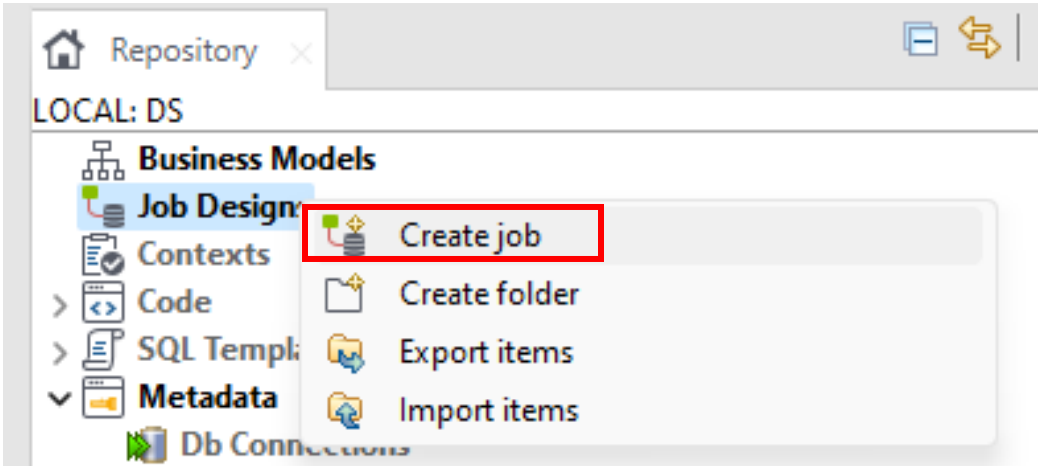
The columns inside the table are shown. We will complete the retrieval by clicking finish.



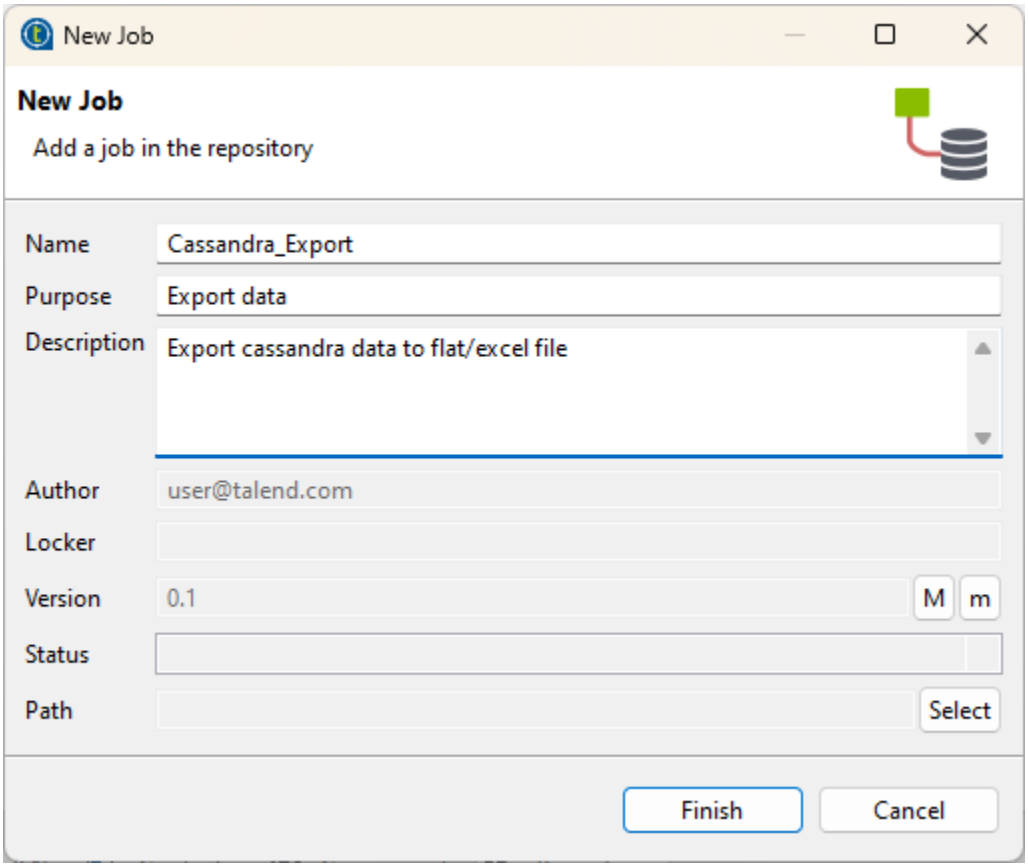
Once it is done, we can see the cctv_count schema under the Cassandra connection. This means it was successful.



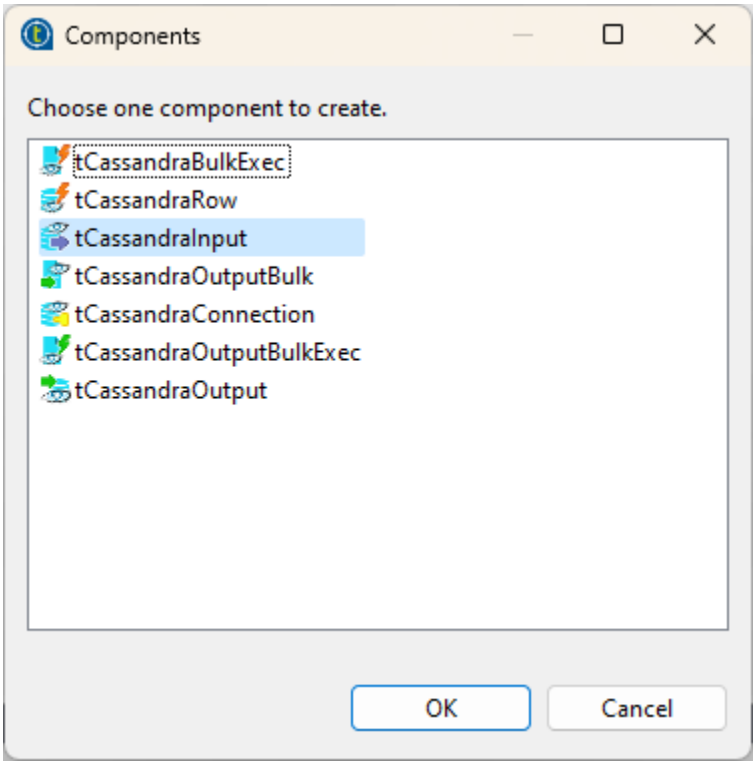
Now we need to create a job for exporting the data. We will right click on Job Designs and click create a job.



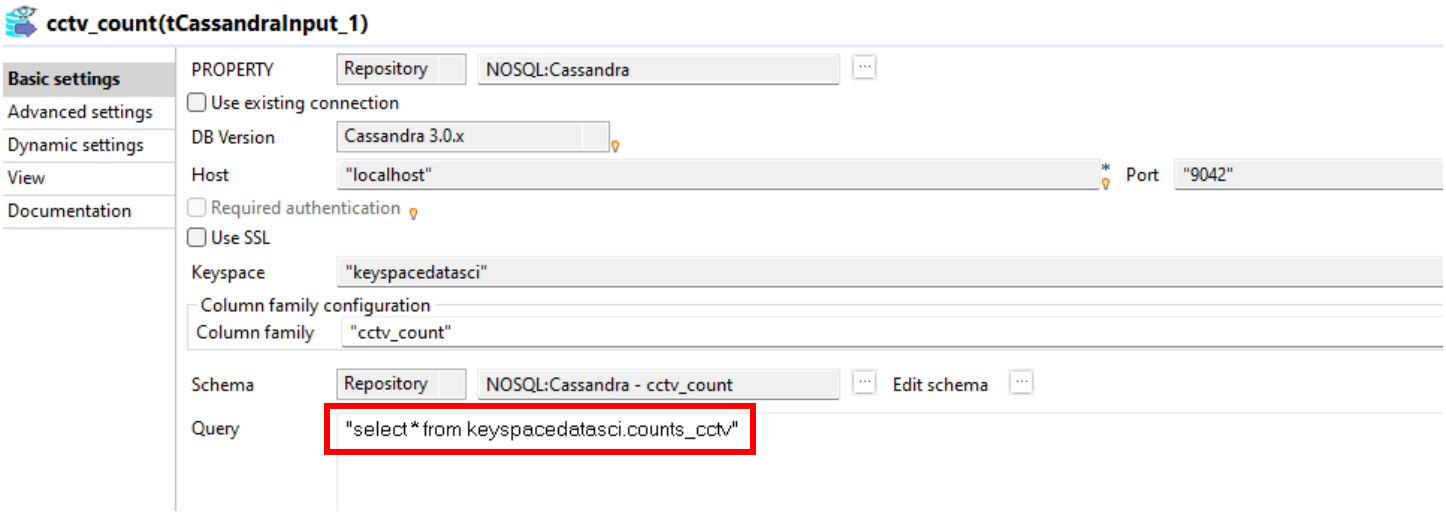
We have to name the job. Here we used Cassandra_Export for the name, Export data for the purpose, and Export Cassandra data to flat/excel file for description. Click finish to create the job.



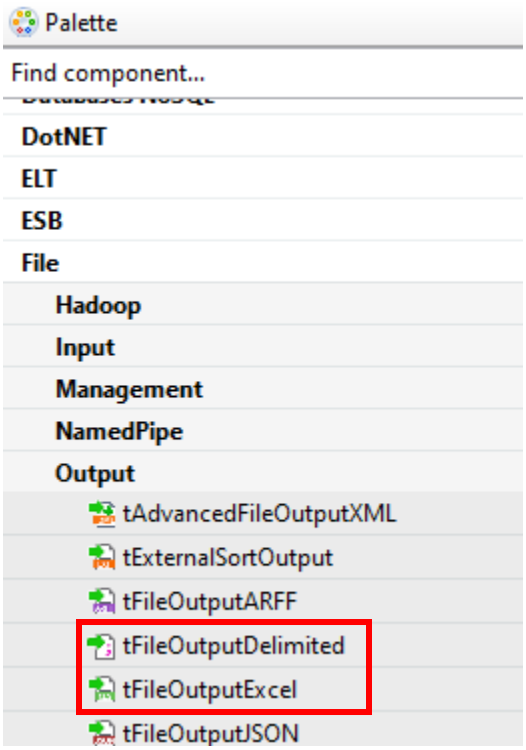
Now we drag the cctv_count schema inside the designer of the job we created.Then we click on tCassandraInput for the component because we need to get data.



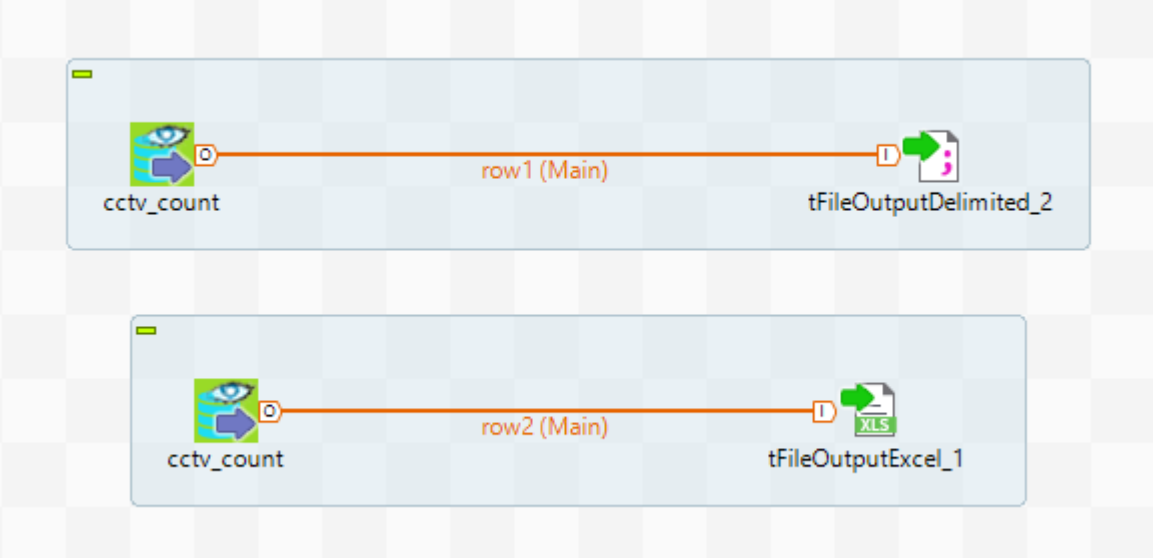
In the tCassandraInput component settings, we will change the query to “select * from keyspacedatasci.counts_cctv”. Then we save the settings.

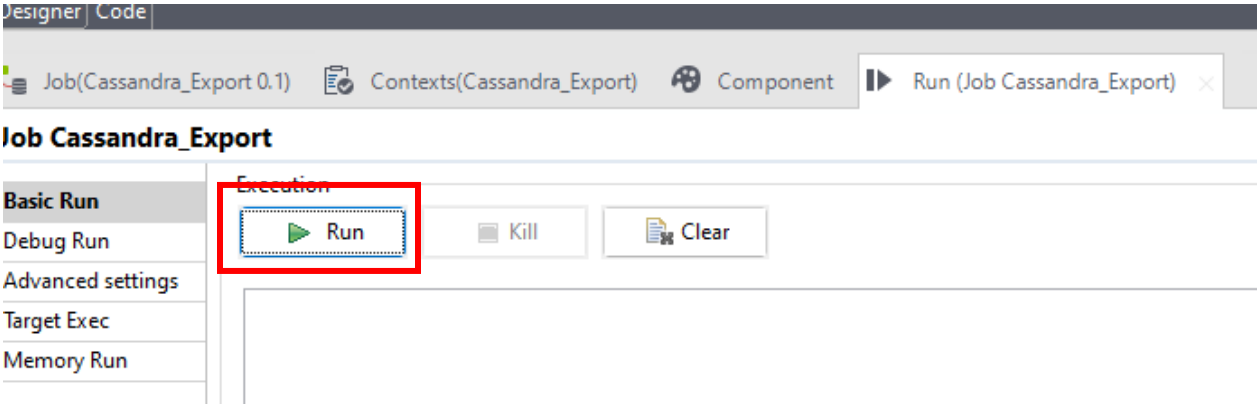


Then in the palette we will click on File, then Output, then tFileOutputDelimited for the flat file. We will also do it for tFileOutputExcel for the excel file. We will drag both components to the designer.

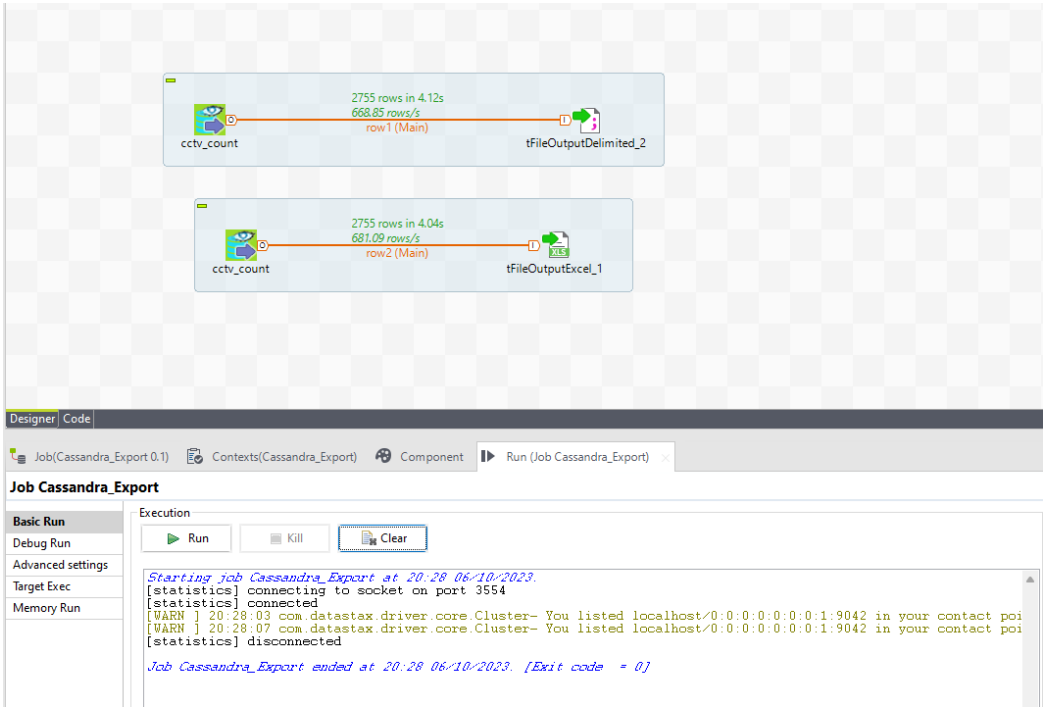


Then we will connect the tCassandraInput components to the tFileOutputDelimited and tFileOutputExcel components. In the Run tab, we click run to begin exporting.

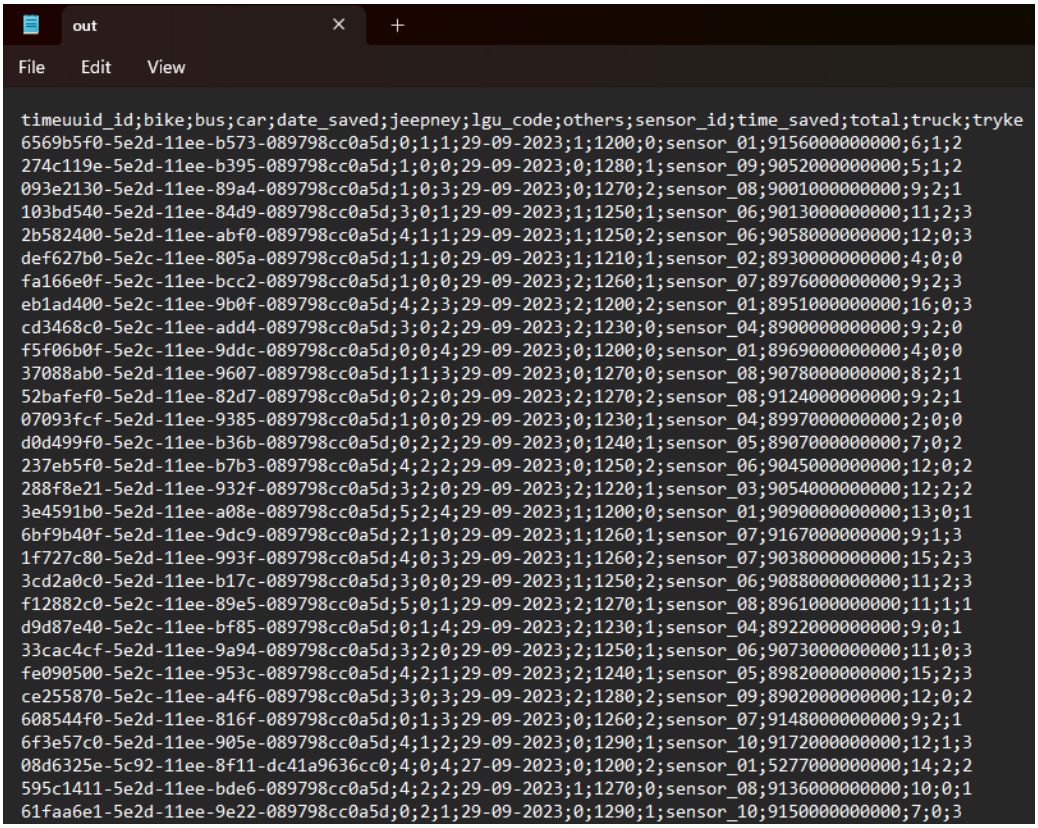




Here we can see that the data is successfully exported to a flat file and an excel file. All rows inside the table are exported.



The flat file with the exported data can be seen below. We successfully exported data from Cassandra to flat file.

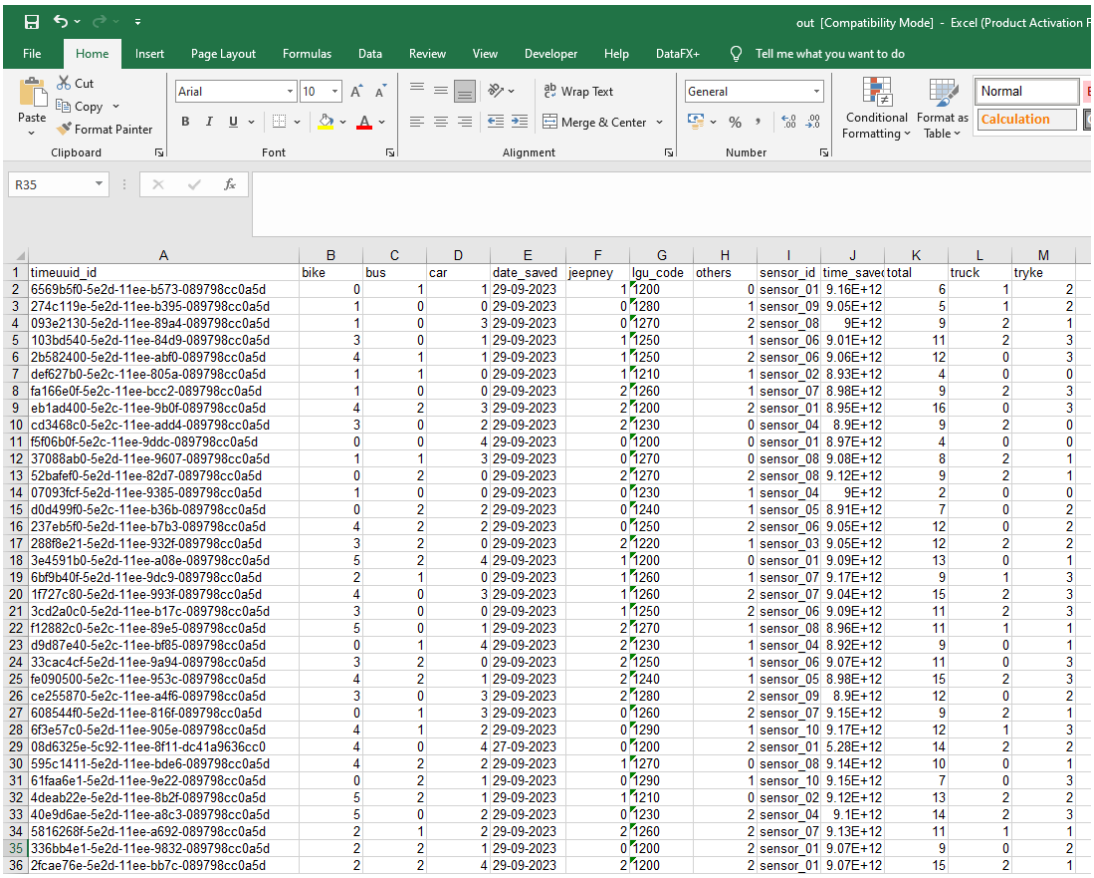


out

File Edit View

```
timeuuid_id;bike;bus;car;date_saved;jeepney;lgu_code;others;sensor_id;time_saved;total;truck;tryke
6569b5f0-5e2d-11ee-b573-089798cc0a5d;0;1;1;29-09-2023;1;1200;0;sensor_01;9156000000000;6;1;2
274c119e-5e2d-11ee-b395-089798cc0a5d;1;0;0;29-09-2023;0;1280;1;sensor_09;9052000000000;5;1;2
093e2130-5e2d-11ee-89a4-089798cc0a5d;1;0;3;29-09-2023;0;1270;2;sensor_08;9001000000000;9;2;1
103bd540-5e2d-11ee-84d9-089798cc0a5d;3;0;1;29-09-2023;1;1250;1;sensor_06;9013000000000;11;2;3
2b582400-5e2d-11ee-abf0-089798cc0a5d;4;1;1;29-09-2023;1;1250;2;sensor_06;9058000000000;12;0;3
def627b0-5e2c-11ee-805a-089798cc0a5d;1;1;0;29-09-2023;1;1210;1;sensor_02;8930000000000;4;0;0
fa166e0f-5e2c-11ee-bcc2-089798cc0a5d;1;0;0;29-09-2023;2;1260;1;sensor_07;8976000000000;9;2;3
eb1ad400-5e2c-11ee-9b0f-089798cc0a5d;4;2;3;29-09-2023;2;1200;2;sensor_01;8951000000000;16;0;3
cd3468c0-5e2c-11ee-add4-089798cc0a5d;3;0;2;29-09-2023;2;1230;0;sensor_04;8900000000000;9;2;0
f5f06b0f-5e2c-11ee-9ddc-089798cc0a5d;0;0;4;29-09-2023;0;1200;0;sensor_01;8969000000000;4;0;0
37088ab0-5e2d-11ee-9607-089798cc0a5d;1;1;3;29-09-2023;0;1270;0;sensor_08;9078000000000;8;2;1
52bafef0-5e2d-11ee-82d7-089798cc0a5d;0;2;0;29-09-2023;2;1270;2;sensor_08;9124000000000;9;2;1
07093fcf-5e2d-11ee-9385-089798cc0a5d;1;0;0;29-09-2023;0;1230;1;sensor_04;8997000000000;2;0;0
d0d499f0-5e2c-11ee-b36b-089798cc0a5d;0;2;2;29-09-2023;0;1240;1;sensor_05;8907000000000;7;0;2
237eb5f0-5e2d-11ee-b7b3-089798cc0a5d;4;2;2;29-09-2023;0;1250;2;sensor_06;9045000000000;12;0;2
288f8e21-5e2d-11ee-932f-089798cc0a5d;3;2;0;29-09-2023;2;1220;1;sensor_03;9054000000000;12;2;2
3e4591b0-5e2d-11ee-a08e-089798cc0a5d;5;2;4;29-09-2023;1;1200;0;sensor_01;9090000000000;13;0;1
6bf9b40f-5e2d-11ee-9dc9-089798cc0a5d;2;1;0;29-09-2023;1;1260;1;sensor_07;9167000000000;9;1;3
1f727c80-5e2d-11ee-993f-089798cc0a5d;4;0;3;29-09-2023;1;1260;2;sensor_07;9038000000000;15;2;3
3cd2a0c0-5e2d-11ee-b17c-089798cc0a5d;3;0;0;29-09-2023;1;1250;2;sensor_06;9088000000000;11;2;3
f12882c0-5e2c-11ee-89e5-089798cc0a5d;5;0;1;29-09-2023;2;1270;1;sensor_08;8961000000000;11;1;1
d9d87e40-5e2c-11ee-bf85-089798cc0a5d;0;1;4;29-09-2023;2;1230;1;sensor_04;8922000000000;9;0;1
33cac4cf-5e2d-11ee-9a94-089798cc0a5d;3;2;0;29-09-2023;2;1250;1;sensor_06;9073000000000;11;0;3
fe090500-5e2c-11ee-953c-089798cc0a5d;4;2;1;29-09-2023;2;1240;1;sensor_05;8982000000000;15;2;3
ce255870-5e2c-11ee-a4f6-089798cc0a5d;3;0;3;29-09-2023;2;1280;2;sensor_09;8902000000000;12;0;2
608544f0-5e2d-11ee-816f-089798cc0a5d;0;1;3;29-09-2023;0;1260;2;sensor_07;9148000000000;9;2;1
6f3e57c0-5e2d-11ee-905e-089798cc0a5d;4;1;2;29-09-2023;0;1290;1;sensor_10;9172000000000;12;1;3
08d6325e-5c92-11ee-8f11-dc41a9636cc0;4;0;4;27-09-2023;0;1200;2;sensor_01;5277000000000;14;2;2
595c1411-5e2d-11ee-bde6-089798cc0a5d;4;2;2;29-09-2023;1;1270;0;sensor_08;9136000000000;10;0;1
61faa6e1-5e2d-11ee-9e22-089798cc0a5d;0;2;1;29-09-2023;0;1290;1;sensor_10;9150000000000;7;0;3
```

The excel file that contains all exported data is also shown. We successfully transferred data from Cassandra cctv_count table to excel file.



out [Compatibility Mode] - Excel (Product Activation F

File Home Insert Page Layout Formulas Data Review View Developer Help DataFX+ Tell me what you want to do

Cut Copy Format Painter — Paste — Clipboard

Arial 10 — Font

Wrap Text Merge & Center — Alignment

General — Number

Conditional Formatting Format as Table Normal Calculation

R35

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	timeuuid_id	bike	bus	car	date_saved	jeepney	lgu_code	others	sensor_id	time_saved	total	truck	tryke
2	6569b5f0-5e2d-11ee-b573-089798cc0a5d	0	1	1	29-09-2023	1	1200	0	sensor_01	9.16E+12	6	1	2
3	274c119e-5e2d-11ee-b395-089798cc0a5d	1	0	0	29-09-2023	0	1280	1	sensor_09	9.05E+12	5	1	2
4	093e2130-5e2d-11ee-89a4-089798cc0a5d	1	0	3	29-09-2023	0	1270	2	sensor_08	9E+12	9	2	1
5	103bd540-5e2d-11ee-84d9-089798cc0a5d	3	0	1	29-09-2023	1	1250	1	sensor_06	9.01E+12	11	2	3
6	2b582400-5e2d-11ee-abf0-089798cc0a5d	4	1	1	29-09-2023	1	1250	2	sensor_06	9.06E+12	12	0	3
7	def627b0-5e2c-11ee-805a-089798cc0a5d	1	1	0	29-09-2023	1	1210	1	sensor_02	8.93E+12	4	0	0
8	fa166e0f-5e2c-11ee-bcc2-089798cc0a5d	1	0	0	29-09-2023	2	1260	1	sensor_07	8.98E+12	9	2	3
9	eb1ad400-5e2c-11ee-9b0f-089798cc0a5d	4	2	3	29-09-2023	2	1200	2	sensor_01	8.95E+12	16	0	3
10	cd3468c0-5e2c-11ee-add4-089798cc0a5d	3	0	2	29-09-2023	2	1230	0	sensor_04	8.9E+12	9	2	0
11	f5f06b0f-5e2c-11ee-9ddc-089798cc0a5d	0	0	4	29-09-2023	0	1200	0	sensor_01	8.97E+12	4	0	0
12	37088ab0-5e2d-11ee-9607-089798cc0a5d	1	1	3	29-09-2023	0	1270	0	sensor_08	9.08E+12	8	2	1
13	52bafef0-5e2d-11ee-82d7-089798cc0a5d	0	2	0	29-09-2023	2	1270	2	sensor_08	9.12E+12	9	2	1
14	07093fcf-5e2d-11ee-9385-089798cc0a5d	1	0	0	29-09-2023	0	1230	1	sensor_04	9E+12	2	0	0
15	d0d499f0-5e2c-11ee-b36b-089798cc0a5d	0	2	2	29-09-2023	0	1240	1	sensor_05	8.91E+12	7	0	2
16	237eb5f0-5e2d-11ee-b7b3-089798cc0a5d	4	2	2	29-09-2023	0	1250	2	sensor_06	9.05E+12	12	0	2
17	288f8e21-5e2d-11ee-932f-089798cc0a5d	3	2	0	29-09-2023	2	1220	1	sensor_03	9.05E+12	12	2	2
18	3e4591b0-5e2d-11ee-a08e-089798cc0a5d	5	2	4	29-09-2023	1	1200	0	sensor_01	9.09E+12	13	0	1
19	6bf9b40f-5e2d-11ee-9dc9-089798cc0a5d	2	1	0	29-09-2023	1	1260	1	sensor_07	9.17E+12	9	1	3
20	1f727c80-5e2d-11ee-993f-089798cc0a5d	4	0	3	29-09-2023	1	1260	2	sensor_07	9.04E+12	15	2	3
21	3cd2a0c0-5e2d-11ee-b17c-089798cc0a5d	3	0	0	29-09-2023	1	1250	2	sensor_06	9.09E+12	11	2	3
22	f12882c0-5e2c-11ee-89e5-089798cc0a5d	5	0	1	29-09-2023	2	1270	1	sensor_08	8.96E+12	11	1	1
23	d9d87e40-5e2c-11ee-bf85-089798cc0a5d	0	1	4	29-09-2023	2	1230	1	sensor_04	8.92E+12	9	0	1
24	33cac4cf-5e2d-11ee-9a94-089798cc0a5d	3	2	0	29-09-2023	2	1250	1	sensor_06	9.07E+12	11	0	3
25	fe090500-5e2c-11ee-953c-089798cc0a5d	4	2	1	29-09-2023	2	1240	1	sensor_05	8.98E+12	15	2	3
26	ce255870-5e2c-11ee-a4f6-089798cc0a5d	3	0	3	29-09-2023	2	1280	2	sensor_09	8.9E+12	12	0	2
27	608544f0-5e2d-11ee-816f-089798cc0a5d	0	1	3	29-09-2023	0	1260	2	sensor_07	9.15E+12	9	2	1
28	6f3e57c0-5e2d-11ee-905e-089798cc0a5d	4	1	2	29-09-2023	0	1290	1	sensor_10	9.17E+12	12	1	3
29	08d6325e-5c92-11ee-8f11-dc41a9636cc0	4	0	4	27-09-2023	0	1200	2	sensor_01	5.28E+12	14	2	2
30	595c1411-5e2d-11ee-bde6-089798cc0a5d	4	2	2	29-09-2023	1	1270	0	sensor_08	9.14E+12	10	0	1
31	61faa6e1-5e2d-11ee-9e22-089798cc0a5d	0	2	1	29-09-2023	0	1290	1	sensor_10	9.15E+12	7	0	3
32	4deab22e-5e2d-11ee-8b2f-089798cc0a5d	5	2	1	29-09-2023	1	1210	0	sensor_02	9.12E+12	13	2	2
33	40e9d6ae-5e2d-11ee-a8c3-089798cc0a5d	5	0	2	29-09-2023	0	1230	2	sensor_04	9.1E+12	14	2	3
34	5816268f-5e2d-11ee-a692-089798cc0a5d	2	1	2	29-09-2023	2	1260	2	sensor_07	9.13E+12	11	1	1
35	336bb4e1-5e2d-11ee-9832-089798cc0a5d	2	2	1	29-09-2023	0	1200	2	sensor_01	9.07E+12	9	0	2
36	2fcae76e-5e2d-11ee-bb7c-089798cc0a5d	2	2	4	29-09-2023	2	1200	2	sensor_01	9.07E+12	15	2	1

References

CData. (n.d.). Connect to cassandra data and transfer data in talend. https://www.cdata.com/kb/tech/cassandra-jdbc-talend.rst

Kattimani, S. (2017, January 8). Store data in cassandra. https://www.youtube.com/watch?v=KMCzJkgHv5s

Talend. (n.d.) Guide to talend. https://www.tutorialspoint.com/talend/talend_quick_guide.htm

Talend. (2023, June 9). Talend | A complete, scalable data management solution. https://www.talend.com/

Talend. (n.d.). Cassandra cql. https://github.com/Talend/apache-camel/blob/master/components/camel-cassandraql/src/main/docs/cql-component.adoc