

[Open in app](#)[Get started](#)

Waqas Rafiq

[Follow](#)Jan 28, 2019 · 2 min read · [Listen](#)[Save](#)

Connect PowerBI with Spark

Microsoft Power BI e Apache Spark



Hi this tutorial is for those people who wants to connect their powerbi with spark but they couldn't find any helpful tutorial on internet.

Step 1:

Add below properties to /home/user/spark/conf/hive-site.xml and /home/user/apachehive/conf/hive-site.xml

```
<property>
```



[Open in app](#)[Get started](#)

```
<property>
<name> hive.server2.thrift.http.port </name>
<value> 10001 </value>
</property>

<property>
<name>  hive.server2.http.endpoint </name>
<value> cliservice </value>
</property>
```

Step 2:

Start hive metastore => start hive => start spark thriftserver

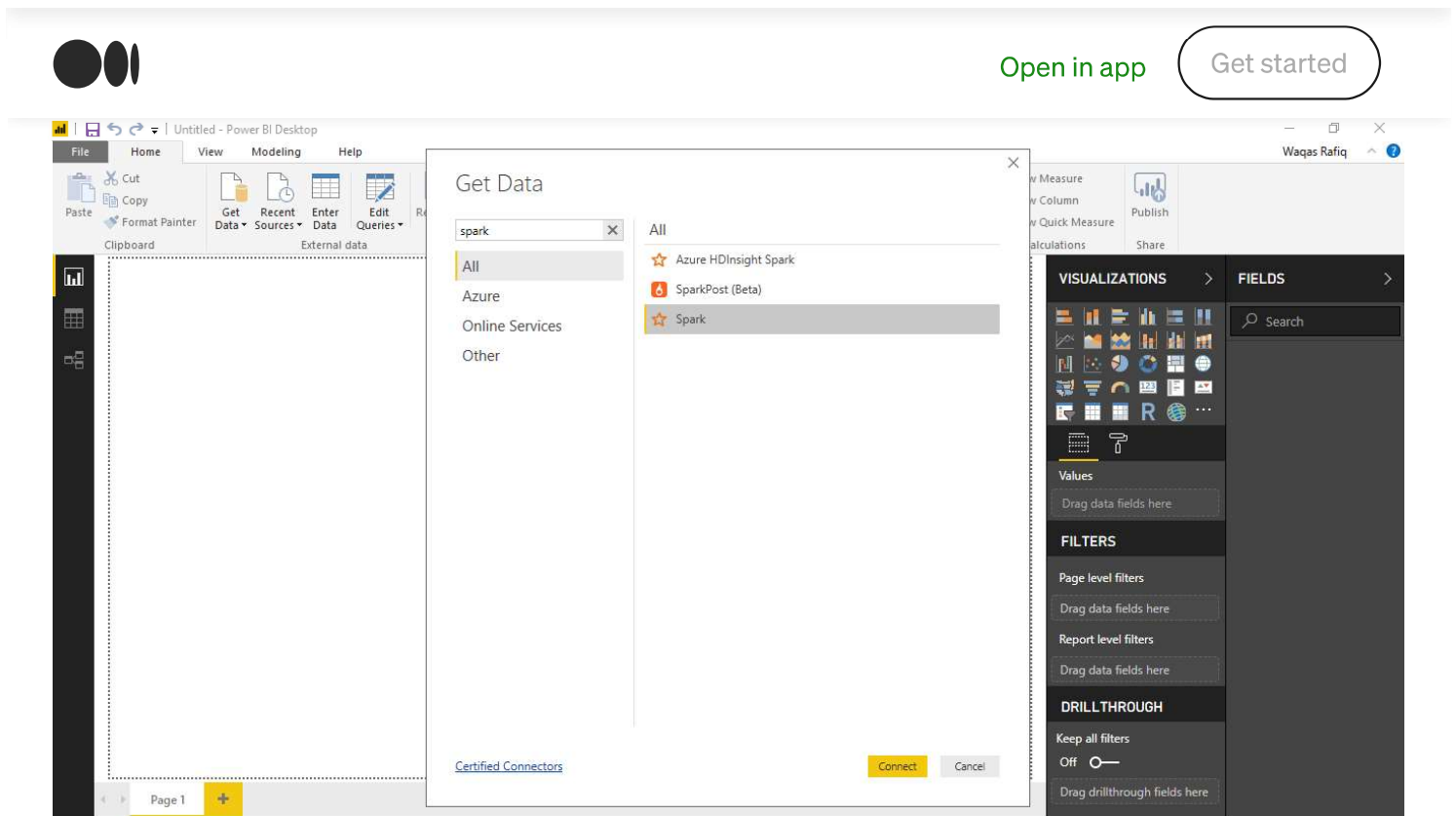
The commands are available below

```
cd /home/user/apachehive/bin/ //get into the directory of hive
./hive --service metastore&    //start metastore
./hive                        //to start hive
cd /home/user/spark/sbin/      //get into the directory of spark
./start-thriftserver           //to start spark thrift server
```

Step 3:

Open the powerbi and click on get data after that go to the spark and click on connect button.



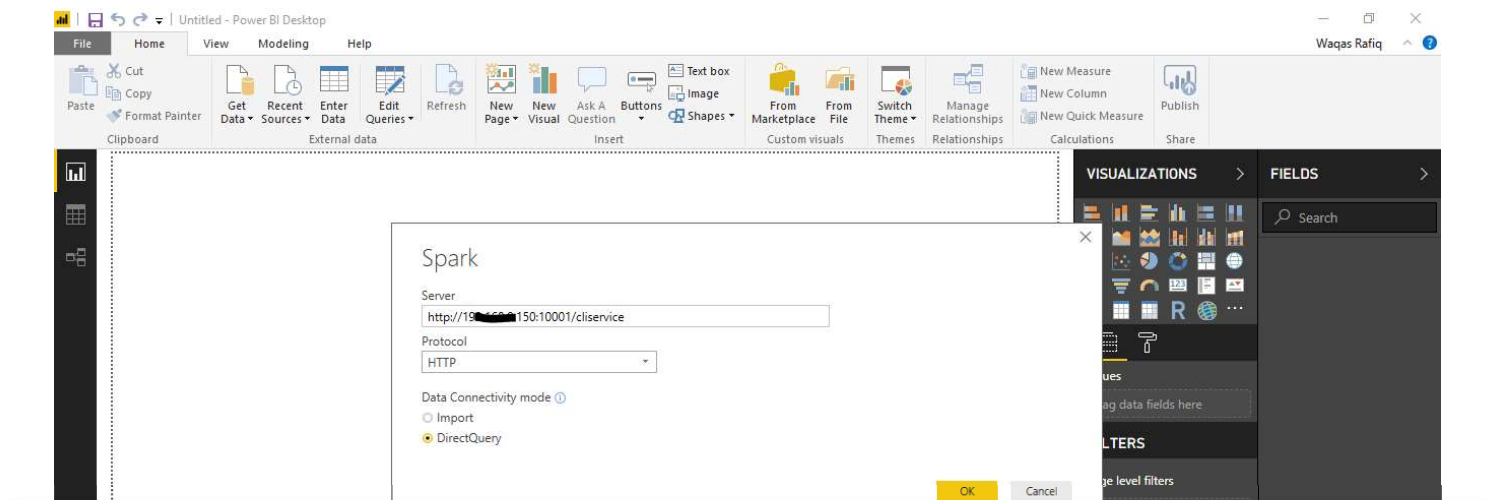


Step 4:

Now in the server put <http://hostipaddres:10001/cliservice>

Select HTTP in Protocol

In Data Connectivity mode check on DirectQuery (means dont want to import data just directly work on it) or Import (means import the data and then work on it)




[Open in app](#)
[Get started](#)

Step 5:

At this time check that table which you want and load it.

The screenshot shows the Power BI Desktop interface. The 'Navigator' window is open, displaying a list of tables: 'http://192.168.0.150:10001/cliservice: 2 [3]', 'april17', 'flight', and 'students'. The 'april17' table is selected. The data preview for 'april17' is shown, indicating it was 'Preview downloaded on Friday'. The table has the following columns: 'day_of_week', 'fl_date', 'op_unique_carrier', 'op_carrier_fl_num', and 'origin'. The data is as follows:

day_of_week	fl_date	op_unique_carrier	op_carrier_fl_num	origin
1	2017-05-22	"DL"	"2"	"JF"
1	2017-05-22	"DL"	"4"	"LA"
1	2017-05-22	"DL"	"6"	"LJ"
1	2017-05-22	"DL"	"7"	"SL"
1	2017-05-22	"DL"	"10"	"DI"
1	2017-05-22	"DL"	"11"	"M"
1	2017-05-22	"DL"	"14"	"S"
1	2017-05-22	"DL"	"15"	"A"
1	2017-05-22	"DL"	"16"	"LJ"
1	2017-05-22	"DL"	"17"	"D"
1	2017-05-22	"DL"	"18"	"BI"
1	2017-05-22	"DL"	"19"	"D"
1	2017-05-22	"DL"	"22"	"M"
1	2017-05-22	"DL"	"23"	"D"
1	2017-05-22	"DL"	"24"	"R"
1	2017-05-22	"DL"	"25"	"A"
1	2017-05-22	"DL"	"30"	"T"
1	2017-05-22	"DL"	"31"	"A"
1	2017-05-22	"DL"	"32"	"DI"
1	2017-05-22	"DL"	"33"	"A"
1	2017-05-22	"DL"	"38"	"SE"

Buttons at the bottom of the preview window include 'Load', 'Edit', and 'Cancel'. The background shows the Power BI Desktop ribbon and the 'FIELDS' pane.

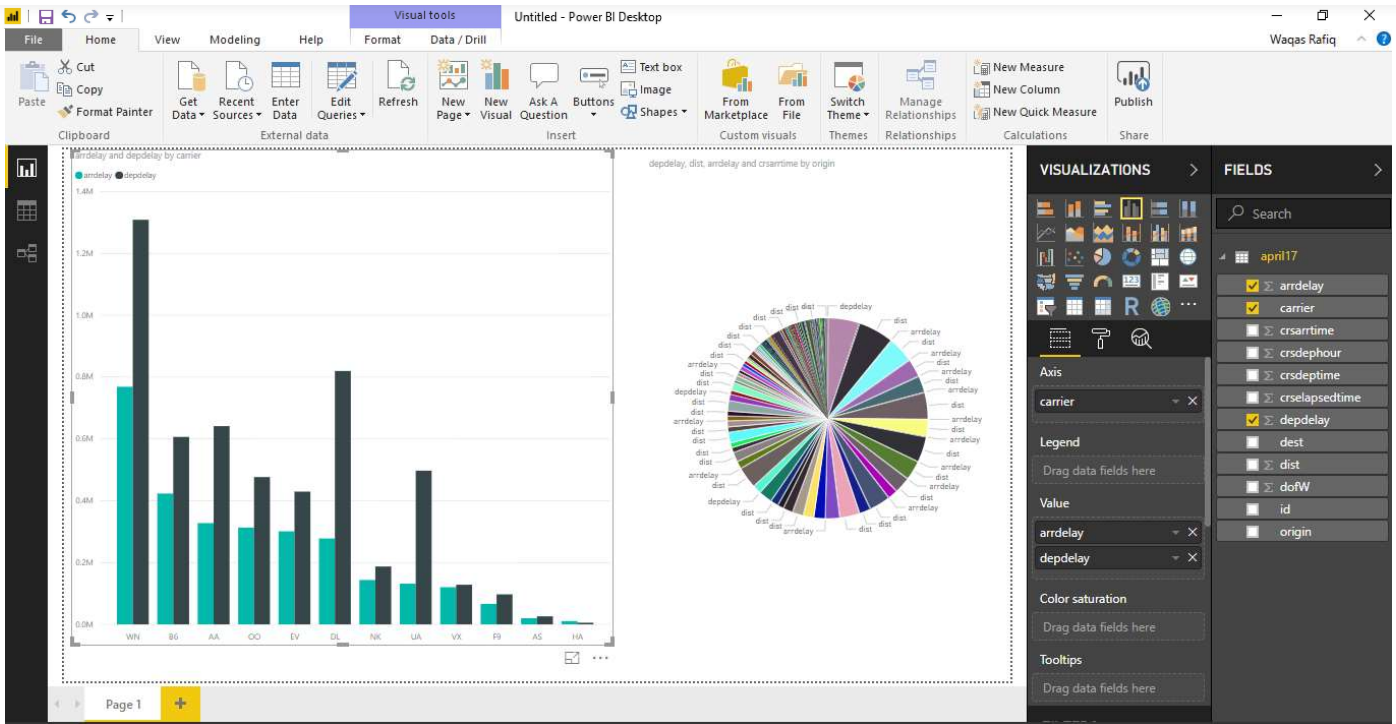
Successfully loaded the data in powerbi via spark.

The screenshot shows the main Power BI Desktop interface. A yellow banner at the top states: 'There are pending changes in your queries that haven't been applied. Apply changes'. The ribbon includes 'File', 'Home', 'View', 'Modeling', and 'Help'. The 'FIELDS' pane on the right shows the 'april17 (2)' table with the following fields: 'arr_delay', 'crs_arr_time', 'crs_dep_time', 'crs_elapsed_ti...', 'day_of_week', 'dep_delay', 'dest', 'distance', 'fl_date', 'op_carrier_fl_n...', 'op_unique_ca...', and 'origin'. The 'VISUALIZATIONS' pane is also visible on the left.



Open in app

Get started



Hopefully this tutorial was helpful for you guys, if you have any query or any issue feel free to ask me in comment section.

[About](#) [Help](#) [Terms](#) [Privacy](#)

Get the Medium app

