

# Connect PowerBI with Spark



Waqas Rafiq

Jan 28, 2019 · 2 min read

## 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>
<name> hive.server2.transport.mode </name>
<value> http </value>
</property>

<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:

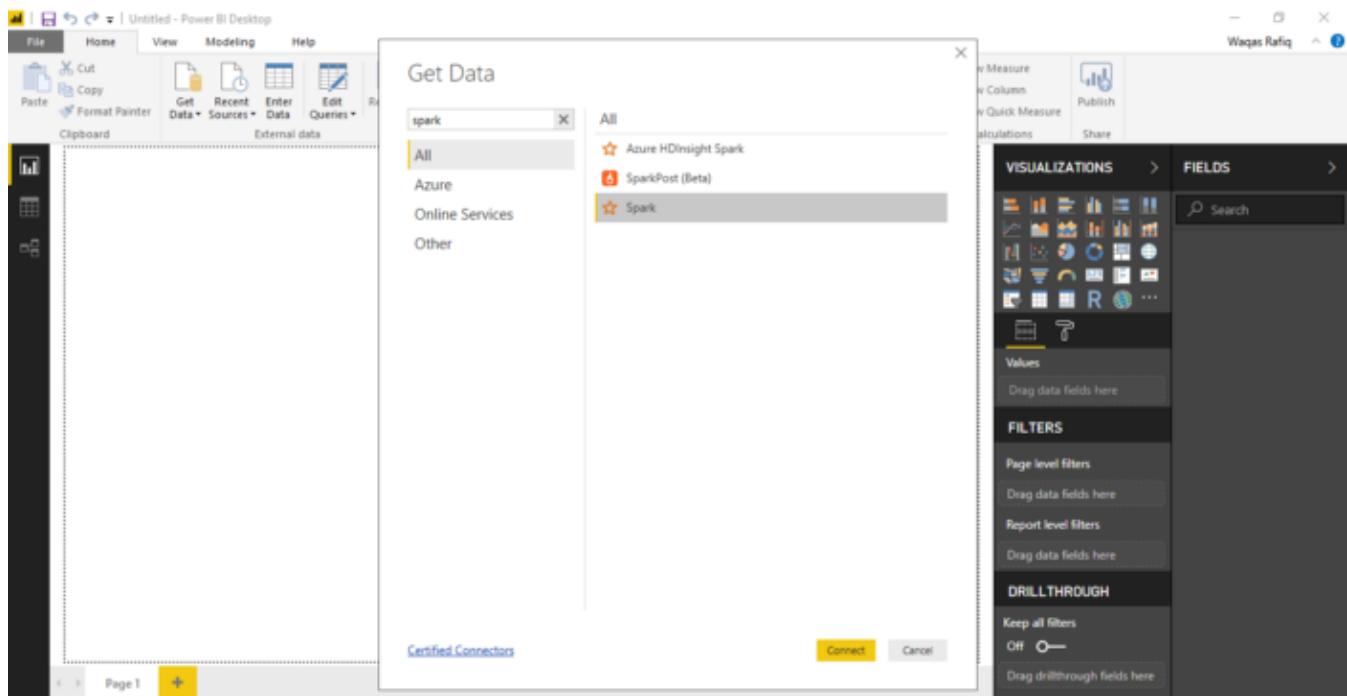
Start hive metastore => start hive => start spark thriftserver

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.

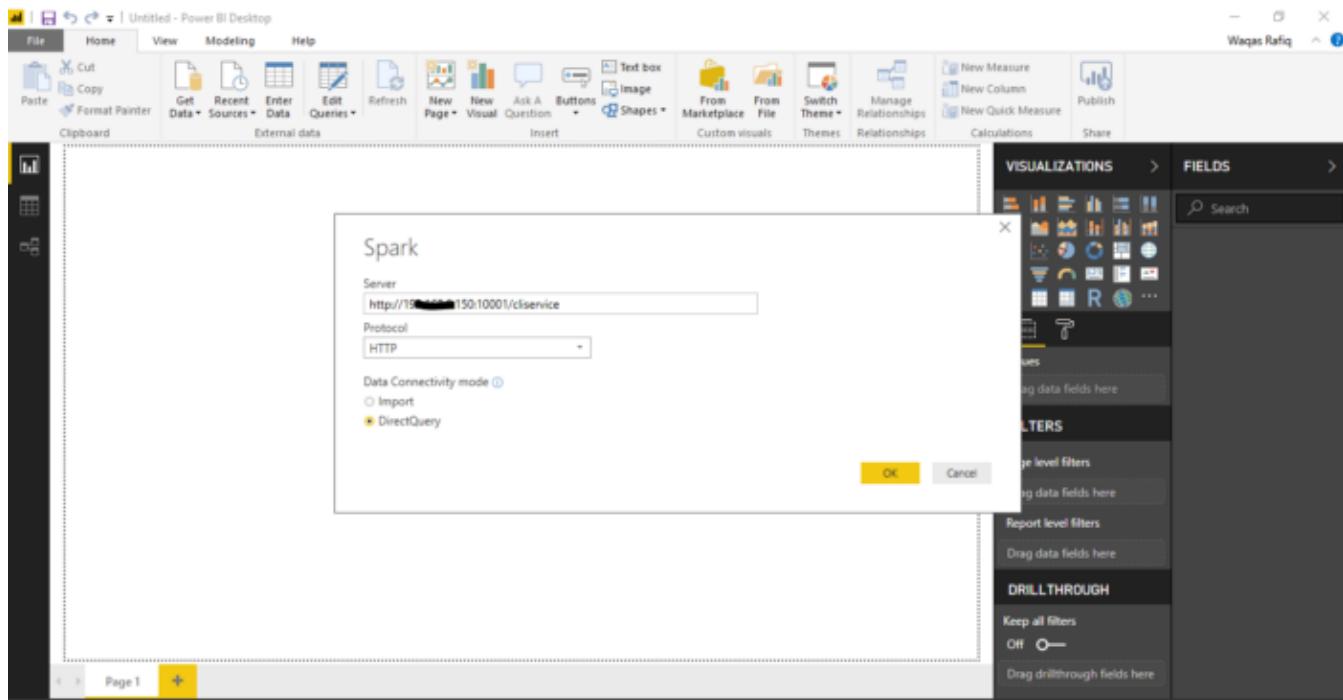


## Step 4:

Now in the server put <http://hostipaddress: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 )



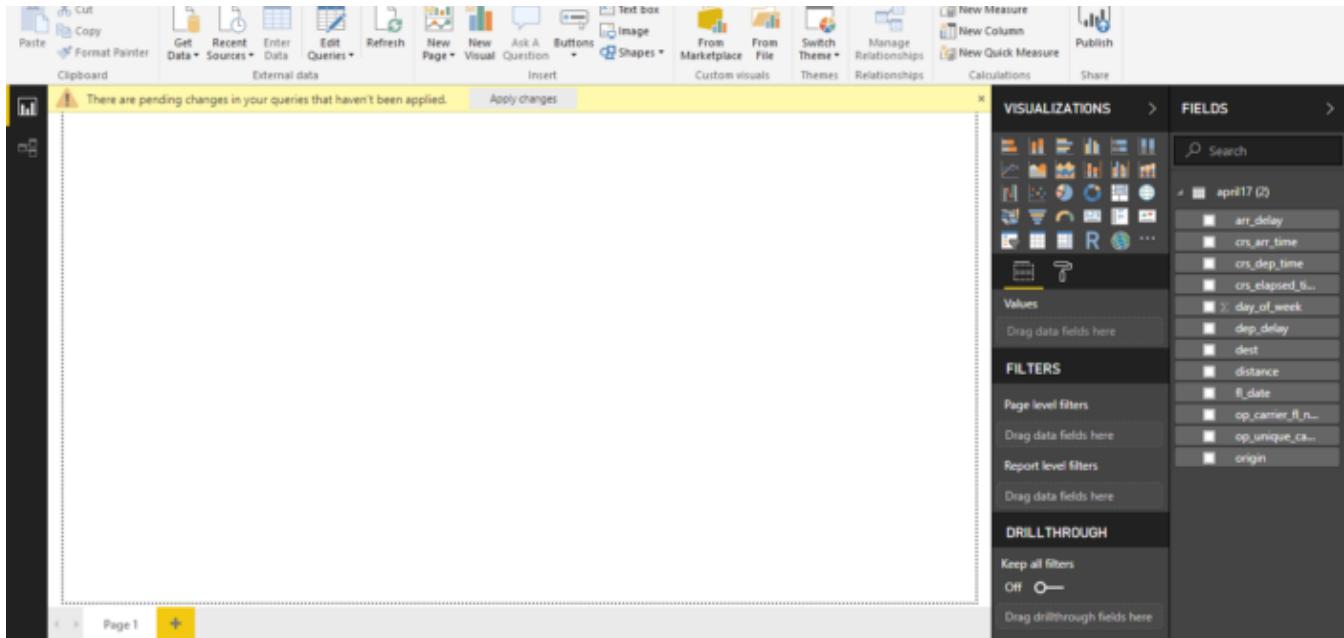
## Step 5:

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

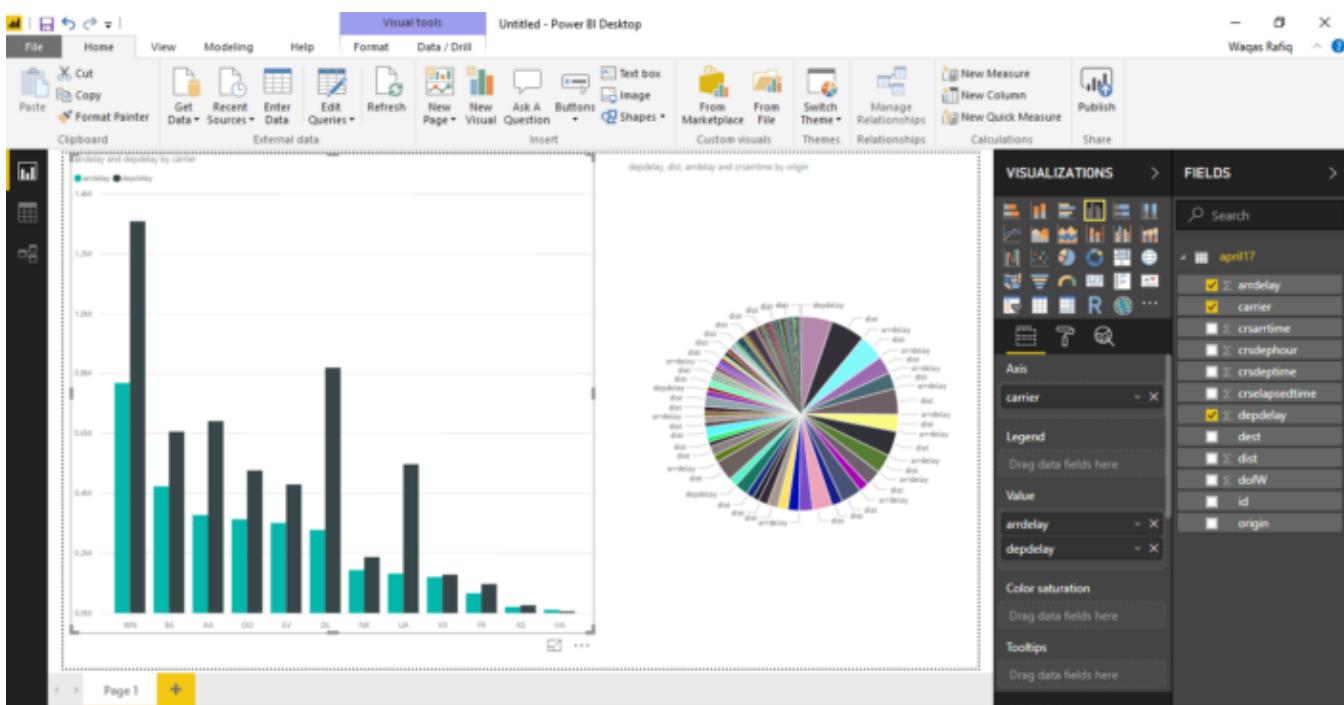
day_of_week	fl_date	op_unique_carrier	op_carrier_fl_num	origin
null	"FL_DATE"	"OP_UNIQUE_CARRIER"	"OP_CARRIER_FL_NUM"	"O"
J	"2017-05-22"	"DL"	"2"	"JF"
J	"2017-05-22"	"DL"	"4"	"U"
J	"2017-05-22"	"DL"	"6"	"U"
J	"2017-05-22"	"DL"	"7"	"SL"
J	"2017-05-22"	"DL"	"10"	"D"
J	"2017-05-22"	"DL"	"13"	"M"
J	"2017-05-22"	"DL"	"14"	"SL"
J	"2017-05-22"	"DL"	"15"	"A"
J	"2017-05-22"	"DL"	"16"	"U"
J	"2017-05-22"	"DL"	"17"	"D"
J	"2017-05-22"	"DL"	"18"	"B"
J	"2017-05-22"	"DL"	"19"	"D"
J	"2017-05-22"	"DL"	"22"	"M"
J	"2017-05-22"	"DL"	"23"	"D"
J	"2017-05-22"	"DL"	"24"	"R"
J	"2017-05-22"	"DL"	"25"	"A"
J	"2017-05-22"	"DL"	"30"	"T"
J	"2017-05-22"	"DL"	"31"	"A"
J	"2017-05-22"	"DL"	"32"	"D"
J	"2017-05-22"	"DL"	"33"	"A"
J	"2017-05-22"	"DL"	"38"	"SE"

Successfully loaded the data in powerbi via spark.





Here you can see some visualizations using the same data which I have loaded from spark.



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

Get the Medium app

