



**User Manual for
Analyzing Machine Data and
Sensor Data using
Hortonworks Sandbox**

Contents

1. Introduction	3
1.1 Getting Started	
2. Installing and Configuring the Hortonworks ODBC Driver on Windows	5
2.1 Download and Install the Hortonworks ODBC Driver	
2.2 Configure the ODBC connection in Windows	
3. Load Data into Hive	13
3.1 Sample Data File	
3.2 Loading Sample Data into Hive	
3.2.1 Connecting to the Sandbox welcome screen	
4. Reporting	21
4.1 Reporting with Excel	
4.2 Visualizing the sensor data using Excel Power View	
5. Conclusion	38

1. Introduction

In this test drive we are going to explore the Hortonworks Sandbox by using components like Apache Hive, Hortonworks ODBC driver and Apache Zeppelin.

Hortonworks Sandbox is a personal, portable Apache Hadoop® environment that comes with dozens of interactive Hadoop and its ecosystem tutorials and the most exciting developments from the latest HDP distribution

1.1 Getting Started

In test drive portal page click on Hortonworks and in the Hortonworks test drive page click on **Start Free Test Drive**, test drive will be started and the access information for the test drive will be displayed as shown below.

Details

Hortonworks Sandbox is a personal, portable Apache Hadoop® environment that comes with dozens of interactive Hadoop and its ecosystem tutorials and the most exciting developments from the latest HDP distribution.

[hortworks-user-guide](#)

TEST DRIVE DURATION
3 hours

ESTIMATED DEPLOYMENT DURATION
2 minutes to 20 minutes

[See details on Azure marketplace](#)

[Add this to your Azure account](#)

Details

Hortonworks Sandbox is a personal, portable Apache Hadoop® environment that comes with dozens of interactive Hadoop and its ecosystem tutorials and the most exciting developments from the latest HDP distribution.

[hortworks-user-guide](#)

TEST DRIVE DURATION
3 hours

ESTIMATED DEPLOYMENT DURATION
2 minutes to 20 minutes

[See details on Azure marketplace](#)

[Add this to your Azure account](#)

[Stop Test Drive](#)

The screenshot shows the Microsoft Azure Test Drive preview UI for a Hortonworks Sandbox. The title bar says "Microsoft Azure | Test Drive | Preview UI". The main content area displays the "Hortonworks" logo and "by Sysgain". It indicates the test drive is "Active (3 hours left)". Below this, there's a section titled "Access information" with three steps:

- Step 1: Access the test drive using the following details:
 - DNS name of Hortonworks Sandbox: hortdnsy3msh.centralus.cloudapp.azure.com
 - sandbox web console: <http://hortdnsy3msn.centralus.cloudapp.azure.com:8888>
- Step 2: Use the following information to login:
 - AdminUsername : hortonuser
 - AdminPassword : Sysga1n4205!
- Step 3: Explore [key use scenarios](#) or refer to the test drive documentation to try more complex scenarios.

Below this, a "Details" section provides information about the Hortonworks Sandbox, including its purpose ("a personal, portable Apache Hadoop® environment"), its duration ("TEST DRIVE DURATION: 3 hours"), and deployment duration ("ESTIMATED DEPLOYMENT DURATION: 2 minutes to 20 minutes").

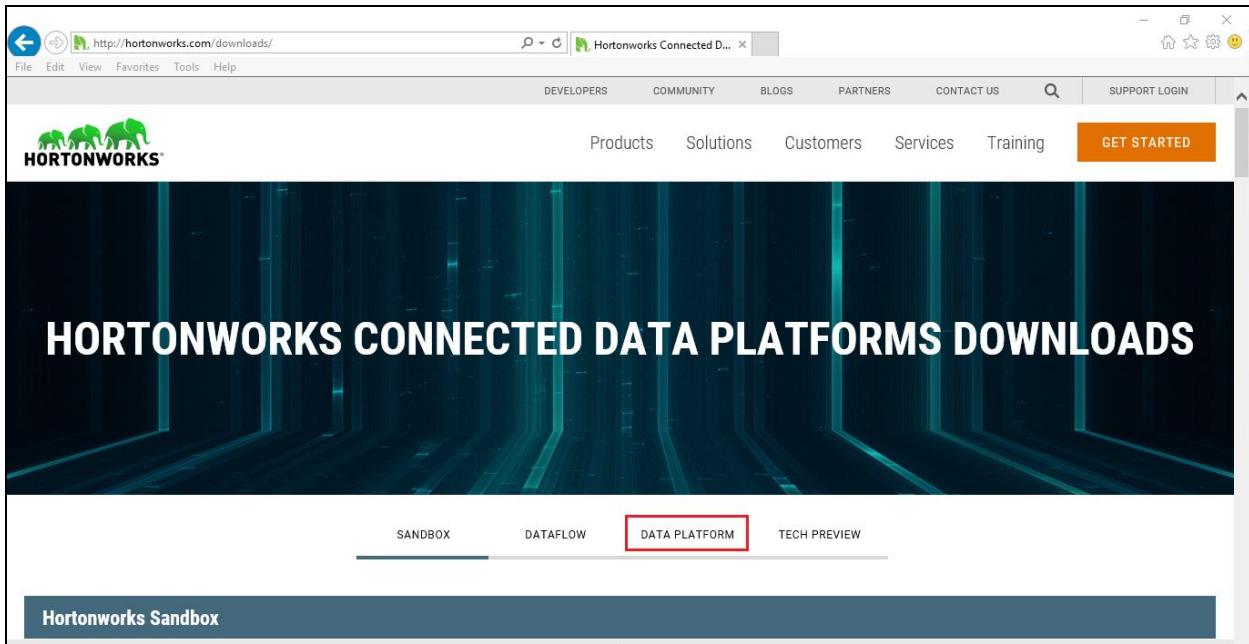
Once we done we can end the test drive or once time is up test drive will end automatically.

The screenshot shows the Microsoft Azure Test Drive preview UI after the test drive has ended. The title bar says "Microsoft Azure | Test Drive | Preview UI". The main content area displays the "Hortonworks" logo and "by Sysgain". It now says "Your test drive has ended." Below this, there are buttons to "Add this to your Azure account" or "Or sign up for a free Azure account". A note at the bottom says "See the Azure marketplace for [more product details](#), including pricing and deployment information." The "Details" section remains the same as in the previous screenshot.

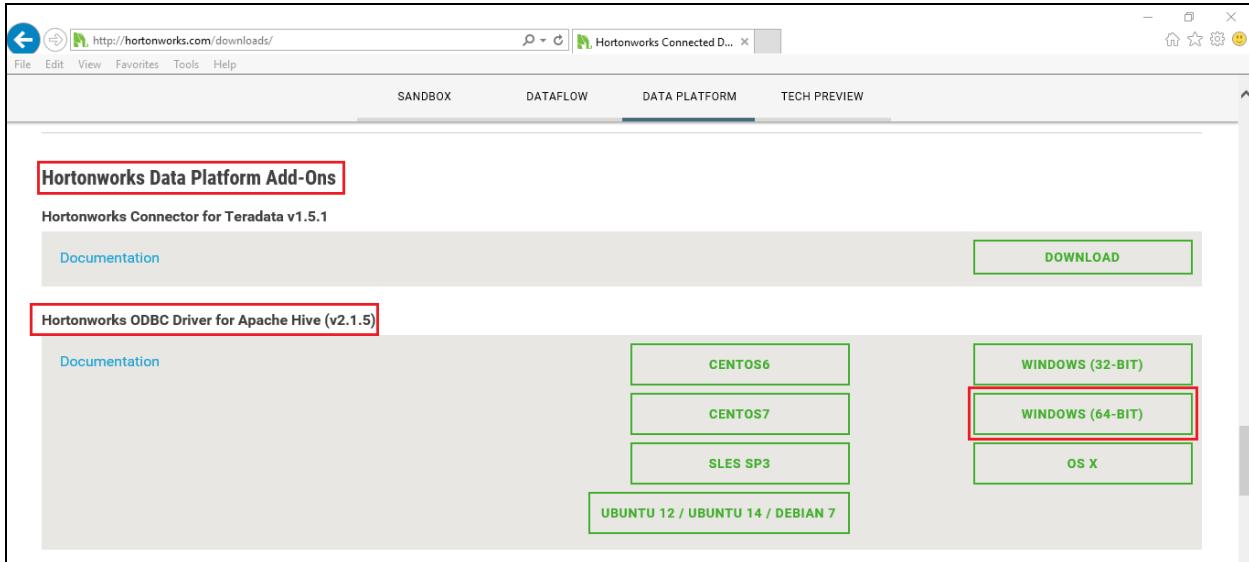
2. Installing and Configuring the Hortonworks ODBC Driver on Windows

2.1 Download and Install the Hortonworks ODBC Driver

In Windows 10, open a web browser and navigate to <http://hortonworks.com/download/>. Click the Data Platform then scroll down to Hortonworks Data Platform Add-ons and click on the Windows (64-BIT) box under Hortonworks ODBC Driver for Apache Hive (v2.1.5).



The screenshot shows the Hortonworks Connected Data Platforms Downloads page. At the top, there's a navigation bar with links for Developers, Community, Blogs, Partners, Contact Us, and a Support Login button. Below the navigation bar is a main menu with links for Products, Solutions, Customers, Services, Training, and a prominent orange GET STARTED button. The central part of the page features a large banner with the text "HORTONWORKS CONNECTED DATA PLATFORMS DOWNLOADS". Below the banner, there are four tabs: SANDBOX, DATAFLOW, DATA PLATFORM (which is highlighted with a red box), and TECH PREVIEW. A blue bar at the bottom contains the text "Hortonworks Sandbox".



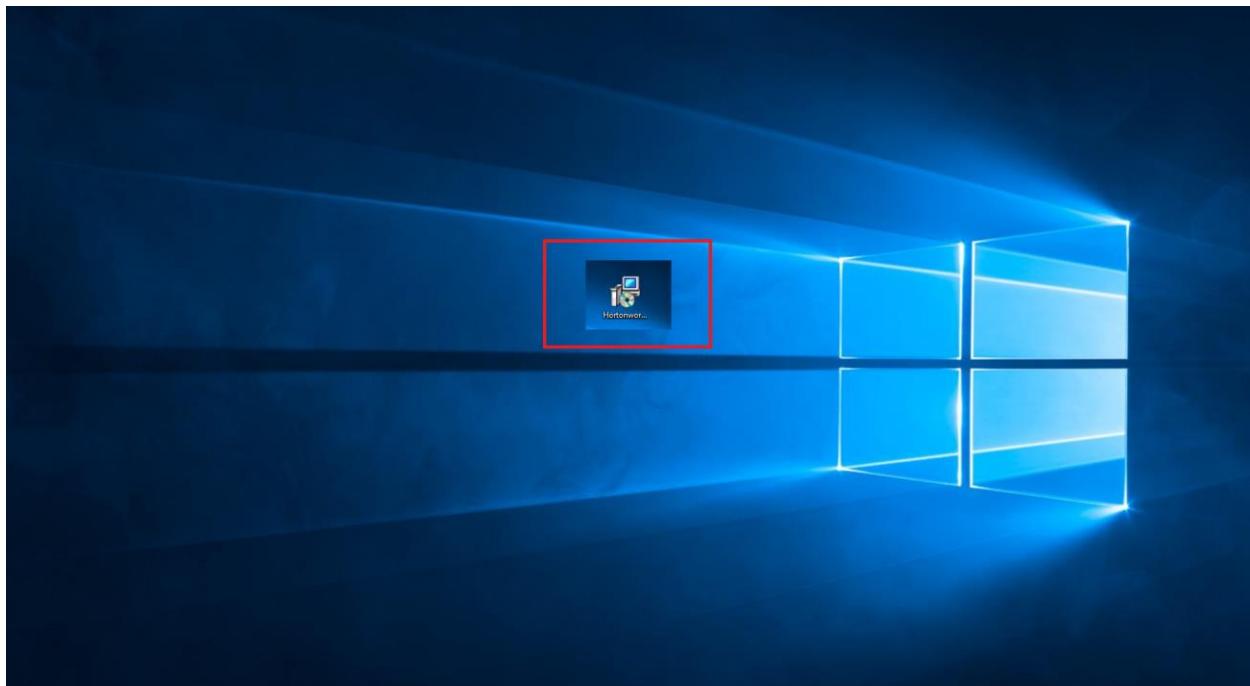
The screenshot shows the Hortonworks Data Platform Add-Ons page. At the top, there's a navigation bar with links for File, Edit, View, Favorites, Tools, and Help. Below the navigation bar, there are four tabs: SANDBOX, DATAFLOW, DATA PLATFORM (which is highlighted with a red box), and TECH PREVIEW. A section titled "Hortonworks Data Platform Add-Ons" contains a link for "Hortonworks Connector for Teradata v1.5.1". Below this, there's a "Documentation" link and a green "DOWNLOAD" button. Another section titled "Hortonworks ODBC Driver for Apache Hive (v2.1.5)" contains a "Documentation" link and a grid of download options. The options are arranged in two columns: CENTOS6, CENTOS7, SLES SP3 in the first column, and WINDOWS (32-BIT), WINDOWS (64-BIT), OS X in the second column. The "WINDOWS (64-BIT)" option is highlighted with a red box.

Once we click on the Windows (64 BIT) box, a file will be downloaded as shown below, save it on the desktop.

The screenshot shows the Hortonworks Data Platform Add-Ons page. It lists three add-ons:

- Hortonworks Connector for Teradata v1.5.1 (Documentation, DOWNLOAD button)
- Hortonworks ODBC Driver for Apache Hive (v2.1.5) (Documentation, CENTOS6, CENTOS7, SLES SP3, UBUNTU 12 / UBUNTU 14 / DEBIAN 7, WINDOWS (32-BIT), WINDOWS (64-BIT), OS X)
- Hortonworks JDBC Driver for Apache Hive (v1.0.36) (User Guide, MENU, Do you want to run or save HortonworksHiveODBC64.msi (14.5 MB) from public-repo-1.hortonworks.com? (Run, Save, Cancel, Contact Sales?))

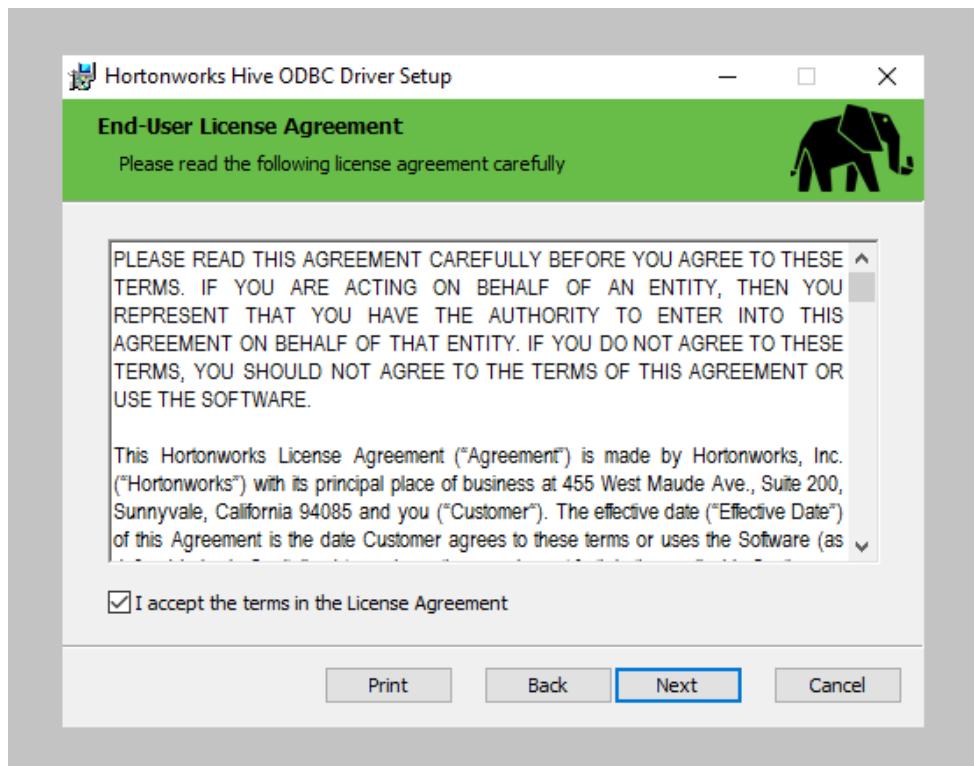
A red box highlights the download dialog for the HortonworksHiveODBC64.msi file.

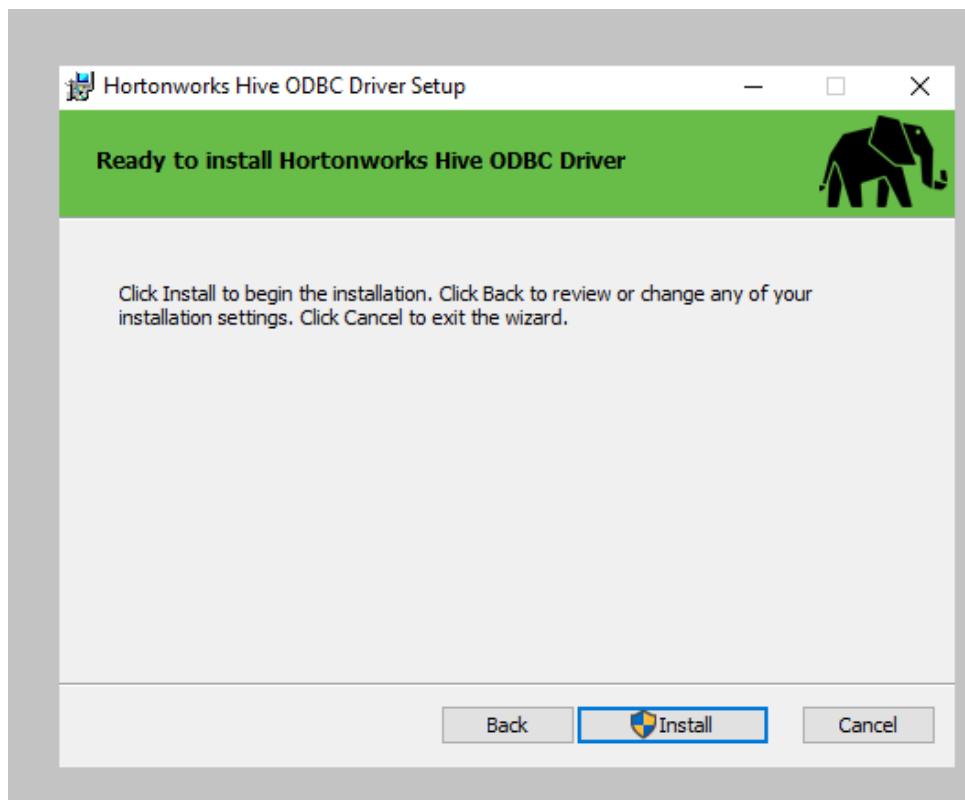
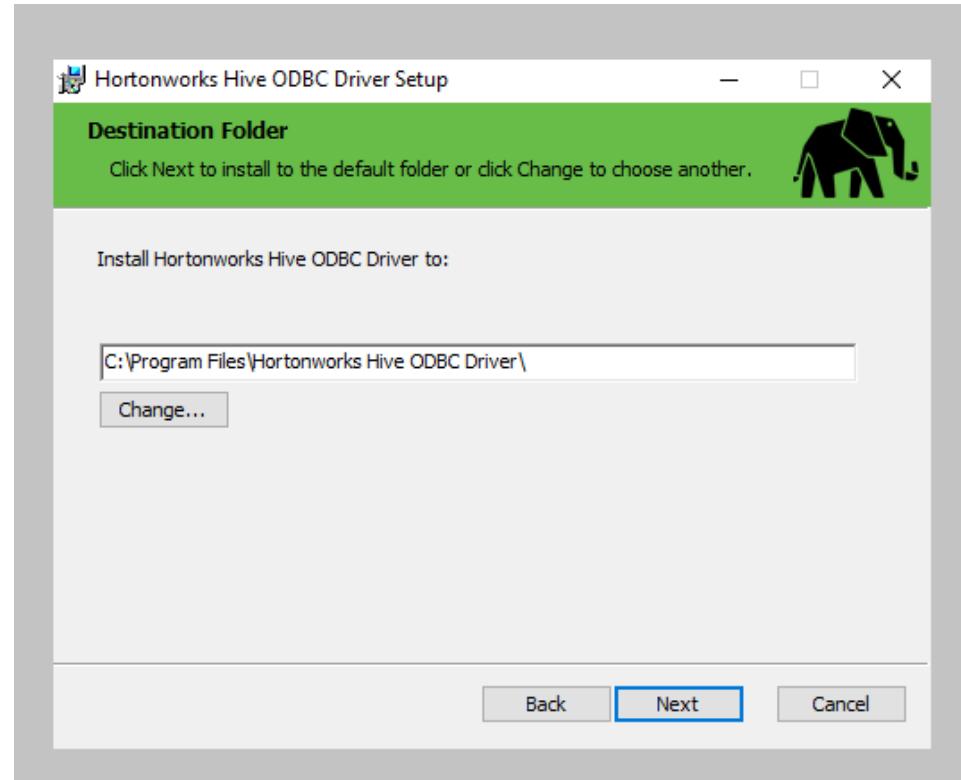


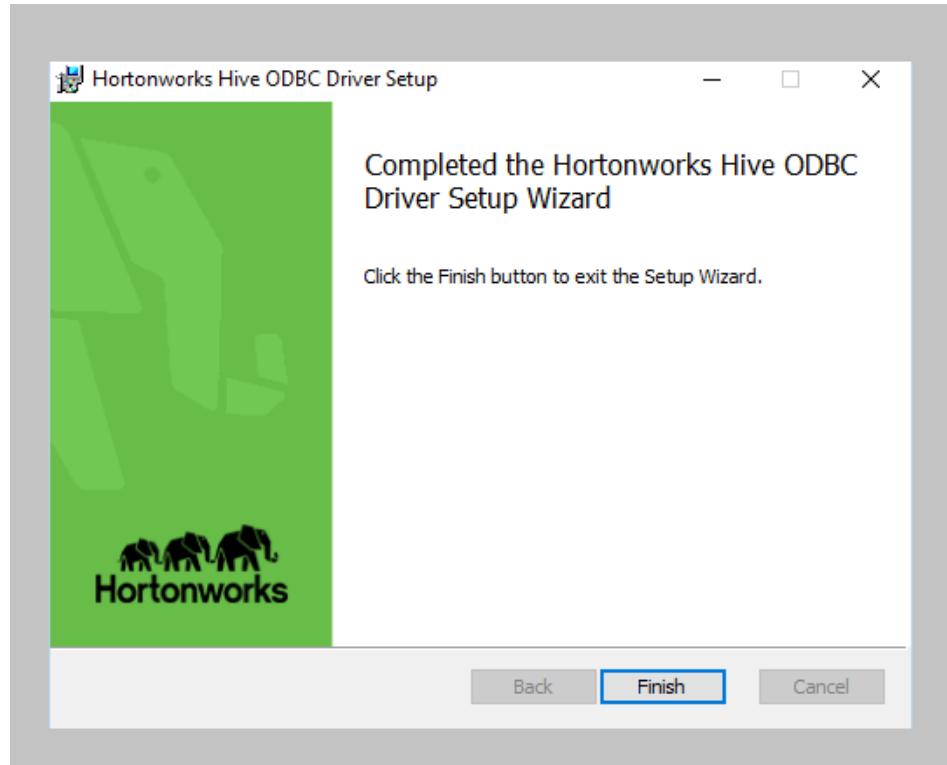
After the download is complete, double-click to open the downloaded HortonworksHiveODBC64.msi file, then click **Run** on the pop-up security message to open the setup wizard. Follow the screens below to finish the installation of the Hortonworks ODBC driver.



Review the license agreement, then select the checkbox to accept the license terms. Click **Next** to continue.



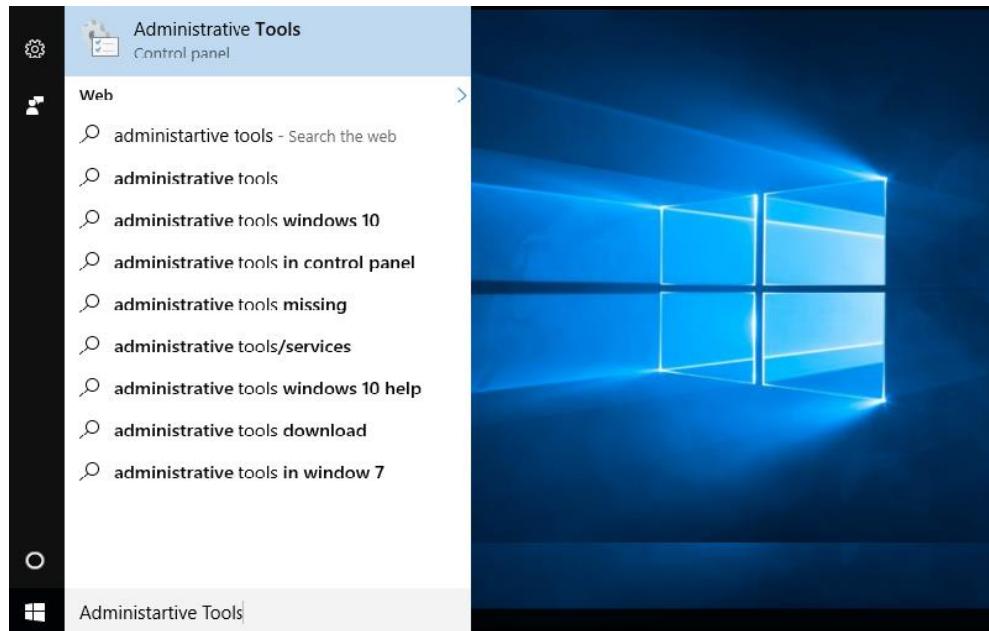


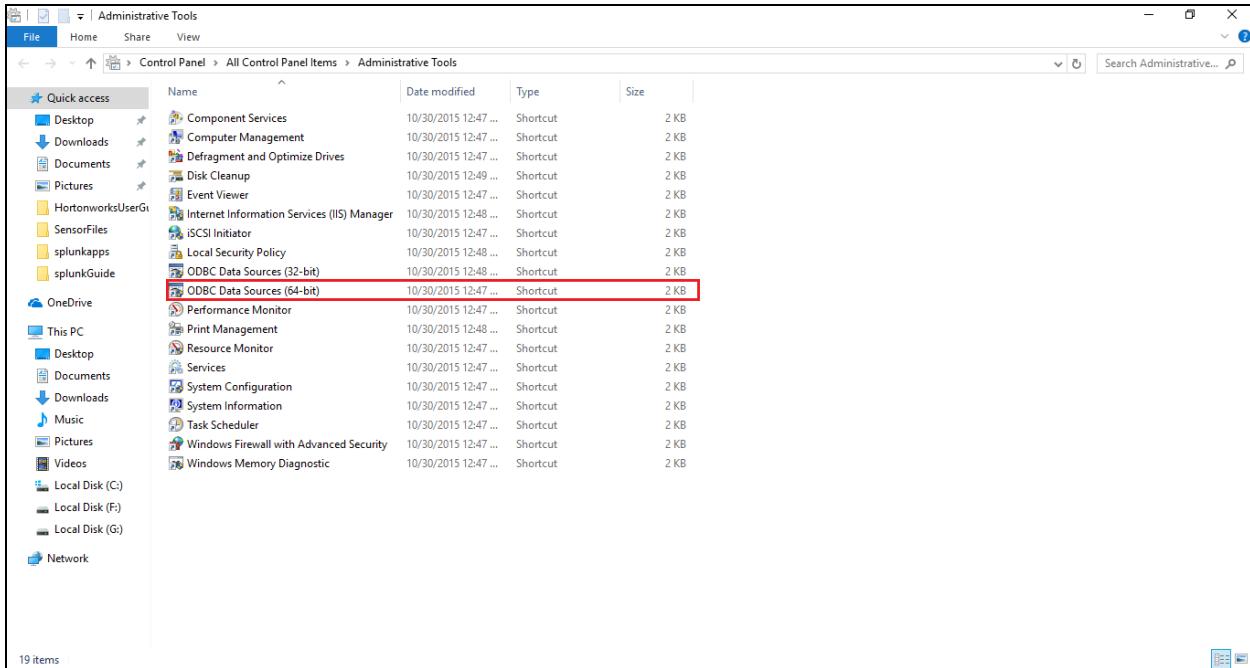


When the installation is complete, the setup wizard displays a confirmation message. Click **Finish** to close the setup wizard

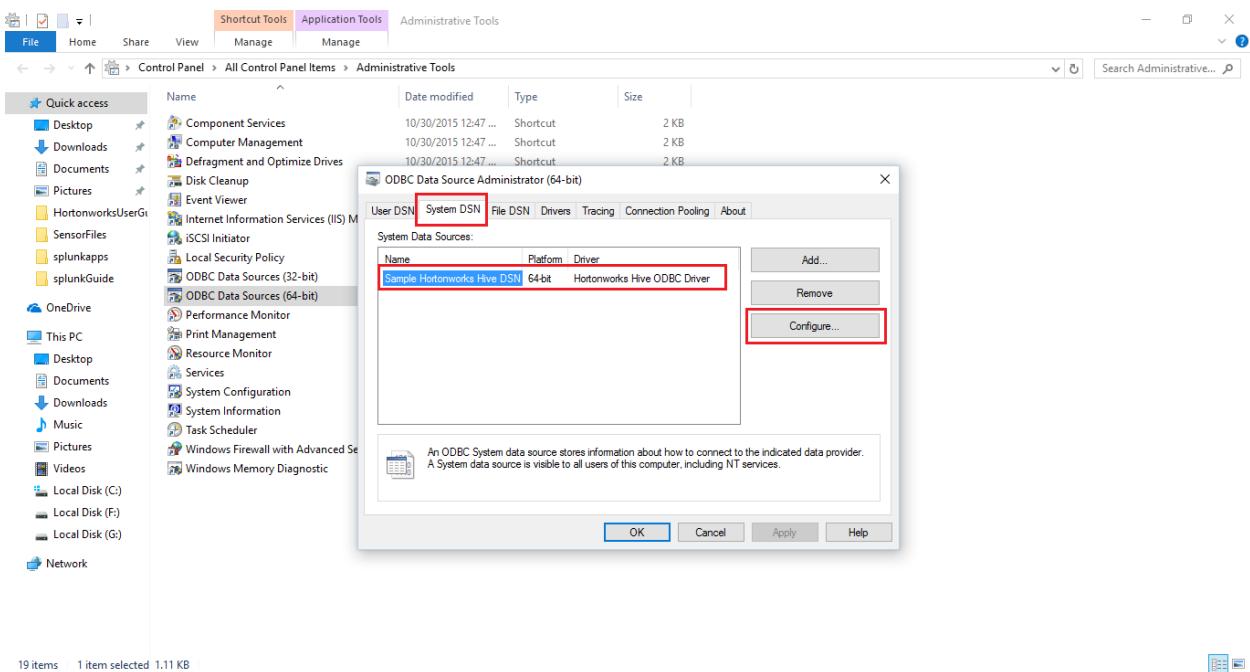
2.2 Configure the ODBC connection in Windows

In the Windows Control Panel, select **Administrative Tools**, then double-click **Data Sources (ODBC)** to open the ODBC Data Source Administrator



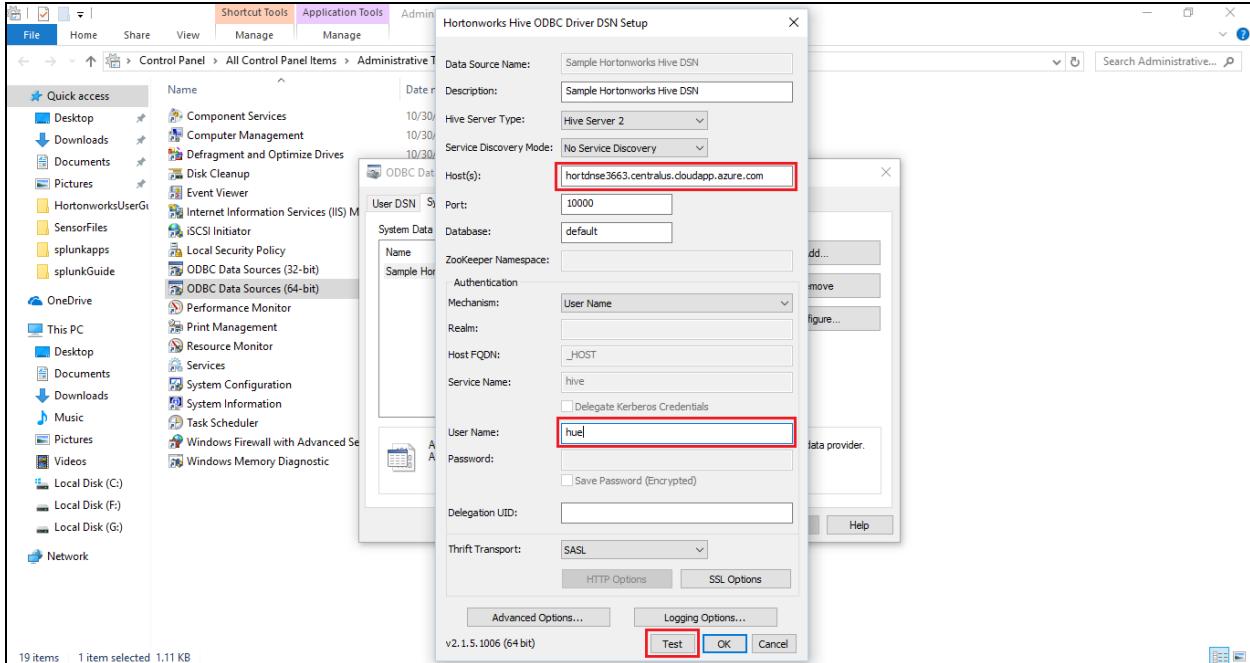


Select the **System DSN** tab. The Sample Hortonworks Hive DSN should be selected by default; if not, select it. Click **Configure** to continue.

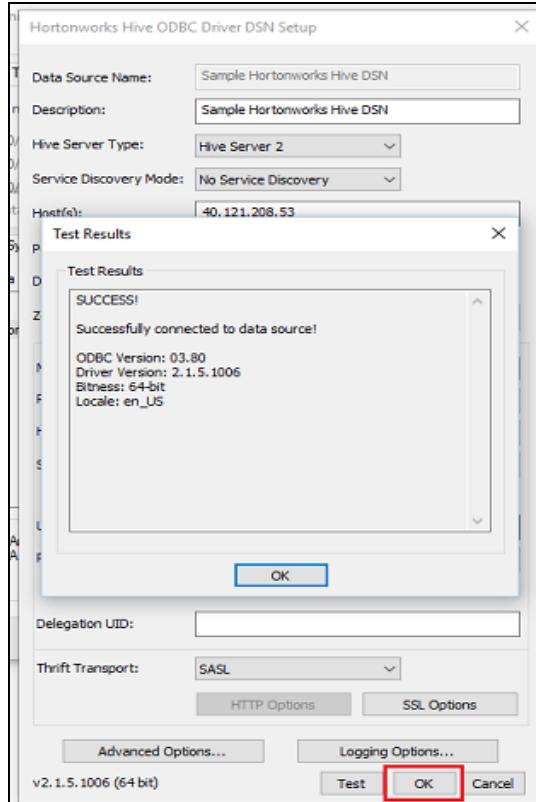


On the Hortonworks Hive ODBC Driver DSN Setup window, type in the settings as shown in the image below. Type the IP address/DNS name (the one which provided in the access information in the azure test drive portal ex: hortdnse3663.centralus.cloudapp.azure.com) of the Hortonworks sandbox in the Host box.

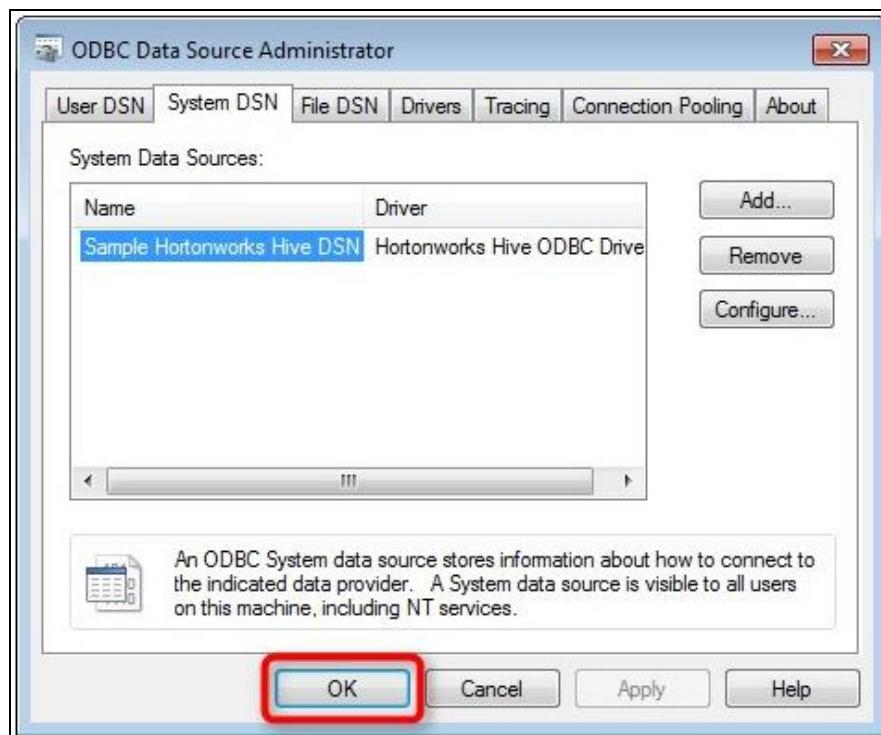
The Authentication mechanism should be set to User Name, and the sandbox user name should be entered in the User Name box (in this case the default user name, "hue").



Click **Test** to test the configuration settings. If the test is successful, a confirmation message appears.
Click **OK** to close the message box.



Click **OK** to close the ODBC Data Source Administrator window.



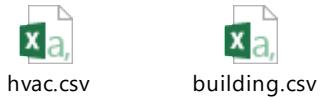
Now that we have configured the Hortonworks ODBC driver, we can enable ODBC connections in BI tools such as Microsoft Excel, then use those applications to access data in the Hortonworks Platform.

3. Load Data into Hive

3.1 Sample Data Files

In this test drive we are going to see how to refine data from heating, ventilation, and air conditioning (HVAC) systems using the Hortonworks Data Platform, and how to analyze the refined sensor data to maintain optimal building temperatures.

Once we refine the data we are going to see how to generate reports using Excel and Zeppelin. To use data for this test drive we attached two data files named “hvac.csv” and “building.csv”.



Copy the above files (double click on the above Excel file and Save As hvac.csv and building.csv respectively on the Desktop. Make sure that the file type is CSV (comma delimited)) in the Desktop.

3.2 Loading Sample Data into Hive

3.2.1 Connecting to the Sandbox welcome screen

After the deployment the access information is displayed in the test drive portal, from the access information get the host address of the Hortonworks sandbox append the port number: 8888 to host address, open your browser, and access Sandbox Welcome page at <http://<host/dns name>:8888/>

Welcome to Sandbox!

Zero to Big Data in 15 mins!

Register to get hands-on tutorials, demos, and videos. Harness the power of Open Enterprise Hadoop.

Test Drive User

testdriveuser@test.com

(800) 233-2742

Test Inc. Other

I agree to [Terms of Use](#)

Submit

Once we fill the data in the above form and click on submit, it will navigate to another page which contains the Apache Ambari login details.

Hortonworks Sandbox

With **HDP 2.4** in a few simple steps

1 get started 2 try 3 what's new

Develop queries for data & manage your HDP cluster

Apache Ambari is the best way to get started with your HDP journey. It provides a user driven wizard interface to interact with HDP.

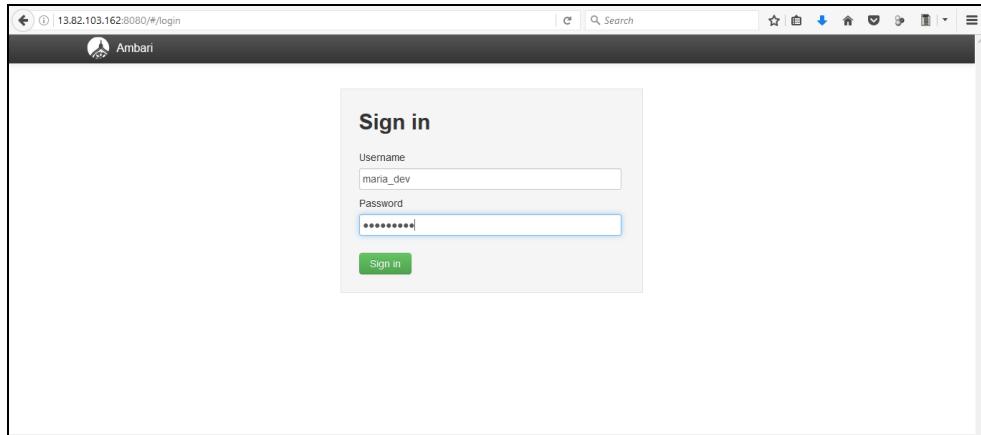
Try this simple tutorial to get started with HDP
[Hello HDP!](#)

url: <http://hortdnse3663.centralus.cloudapp.azure.com:8080>
username: maria_dev
password: maria_dev

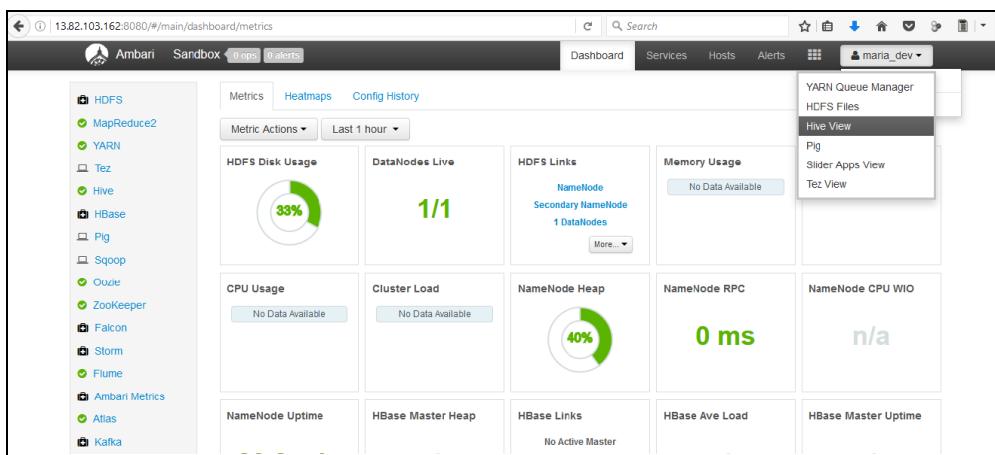
[View Advanced Options](#)

To load data into Hive we need to navigate to the Ambari login by using the Ambari login details provided in the above page.

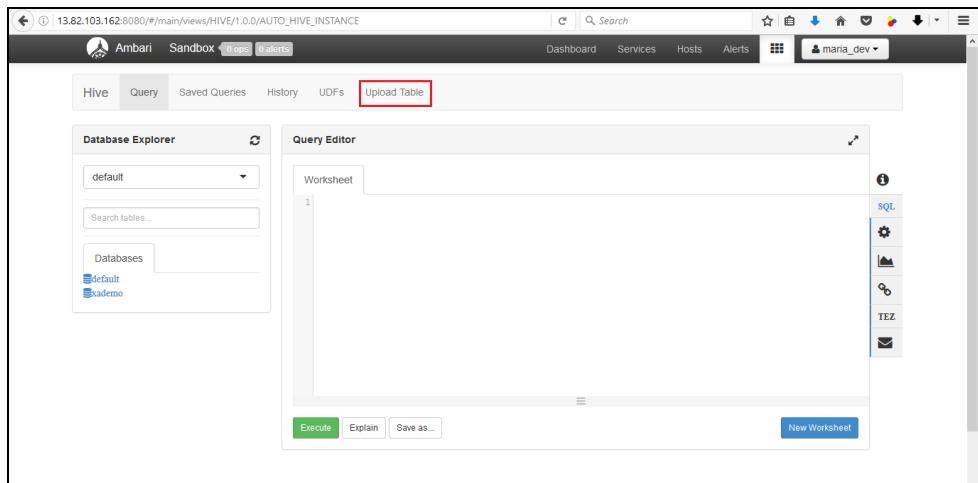
Login with the username ***maria_dev*** and password ***maria_dev***



Head on over to the Hive view using the dropdown menu in the top-right corner.



Then use the Upload Table tab to upload the two csv files.



Browse the hvac.csv file from the Desktop and click on the Upload Table button. Once the table is uploaded a success message will be displayed in green color as shown below.

Hive Database: default Table Name: hvac Is First Row Header? RecordDate Time TargetTemp ActualTemp STRING STRING INT INT 1-Jun-13 0:00:01 66 58 2-Jun-13 1:00:01 69 68 3-Jun-13 2:00:01 70 73 4-Jun-13 3:00:01 67 63 5-Jun-13 4:00:01 68 74 6-Jun-13 5:00:01 67 56 7-Jun-13 6:00:01 70 58 8-Jun-13 7:00:01 70 73 9-Jun-13 8:00:01 66 69

Hive Database: default Table Name: building Is First Row Header? BuildingID BuildingMgr BuildingAge HVACproduct INT STRING INT STRING 1 M1 25 AC1000 2 M2 27 FN39TG 3 M3 28 JDNS77 4 M4 17 GG1919 5 M5 3 ACMAX22 6 M6 9 AC1000 7 M7 13 FN39TG 8 M8 25 JDNS77 9 M9 11 GG1919

✓ Uploaded Successfully

✓ Uploaded Successfully

Now we can see the tables under default database. Click on the default under the Databases. We can see the following

```
1 create table if not exists hvac_building
2 as select h.* , b.country , b.hvacproduct , b.buildingage , b.buildingmgr
3 from buildings_orc b join hvac_temperatures h on b.buildingid = h.buildingid;
```

Now that we have both tables loaded in, we want to get better performance in Hive, so we're going to create new tables that utilize the highly efficient ORC file format. This will allow for faster queries when our datasets are much larger.

Execute the following query to create a new table `hvac Orc` that is stored as an ORC file.

CREATE TABLE hvac Orc STORED AS ORC AS SELECT * FROM hvac;

```
1 CREATE TABLE hvac Orc STORED AS ORC AS SELECT * FROM hvac;
```

Repeat the previous step, except this time we will make a table for buildings.

```
CREATE TABLE buildings_orc STORED AS ORC AS SELECT * FROM building;
```

Create a new worksheet in the Hive view and paste the following Hive query into your window. This query creates a new table hvac_temperatures and copies data from the hvac table

After you paste the query use Execute to create the new table.

```
CREATE TABLE hvac_temperatures as
```

```
select *, (targettemp - actualtemp) as temp_diff,
IF((targettemp - actualtemp) > 5, 'COLD',
IF((targettemp - actualtemp) < -5, 'HOT', 'NORMAL'))
AS temprange,
IF((targettemp - actualtemp) > 5, '1',
IF((targettemp - actualtemp) < -5, '1', 0))
AS extremetemp from hvac_orc;
```

The screenshot shows the Ambari Hive Query Editor. On the left, the Database Explorer lists databases: default, Building, Buildings orc, HVAC, HVAC orc, Sample 07, Sample 08, and Xademo. The main area is the Query Editor, which contains two worksheets. Worksheet (1) is active and displays the following Hive query:

```

1 CREATE TABLE hvac_temperatures AS
2 select *, (targetTemp - actualtemp) as temp_diff,
3 IF((targettemp - actualtemp) > 5, 'COLD',
4 IF((targettemp - actualtemp) < -5, 'HOT', 'NORMAL'))
5 AS temprange,
6 IF((targettemp - actualtemp) > 5, '1',
7 IF((targettemp - actualtemp) < -5, '1', 0))
8 AS extremetemp from hvac_orc;

```

Below the query are buttons for Execute, Explain, Save as..., and New Worksheet. To the right of the editor is a sidebar with icons for SQL, TEZ, and other tools.

On the Query Results page, use the slider to scroll to the right. You will notice that two new attributes appear in the hvac_temperatures table.

The data in the temprange column indicates whether the actual temperature was:

- NORMAL – within 5 degrees of the target temperature.
- COLD – more than five degrees colder than the target temperature.
- HOT – more than 5 degrees warmer than the target temperature.

If the temperature is outside of the normal range, extremetemp is assigned a value of 1; otherwise its value is 0.

Next we will combine the buildings_orc and hvac_temperatures data sets by using the below hive query.

```

create table if not exists hvac_building
as select h.* , b.country, b.hvacproduct, b.buildingage, b.buildingmgr
from buildings_orc b join hvac_temperatures h on b.buildingid = h.buildingid;

```

The screenshot shows the Ambari Hive Query Editor interface. In the Database Explorer on the left, the 'default' database is selected. The Query Editor panel contains a code editor with the following SQL query:

```

1 create table if not exists hvac_building
2 as select h.* , b.country , b.hvacproduct , b.buildingage , b.buildingmgr
3 from buildings_orc b join hvac_temperatures h on b.buildingid = h.buildingid;

```

Below the code editor are three buttons: 'Execute', 'Explain', and 'Save as...'. To the right of the code editor is a vertical toolbar with icons for 'SQL', 'TEZ', and other options. A red box highlights the 'TEZ' icon.

After you've successfully executed the query, use the database explorer to load a sample of the data from the new hvac_building table.

The screenshot shows the Ambari Hive Query Editor interface after the query has been executed. The Database Explorer on the left now lists the 'hvac_building' table, which is highlighted with a red box. The Query Process Results panel on the right displays the data from the 'hvac_building' table. The table has the following columns: recorddate, hvac_building.time, hvac_building.targettemp, hvac_building.actualtemp, and hvac_b. The data is as follows:

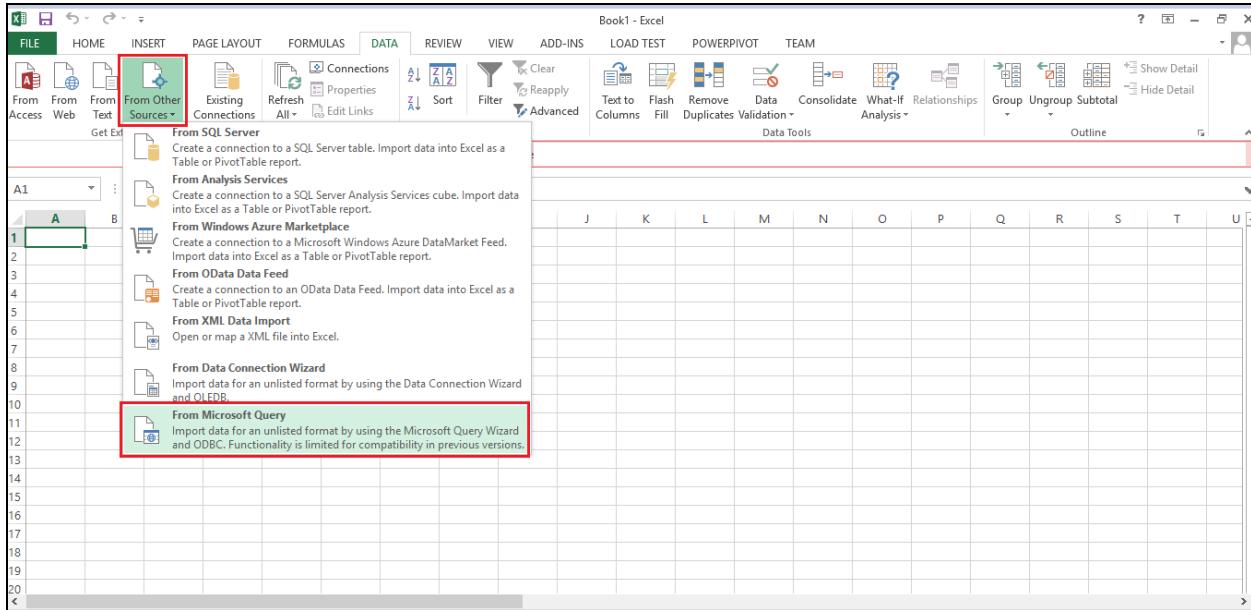
recorddate	hvac_building.time	hvac_building.targettemp	hvac_building.actualtemp	hvac_b
1-Jun-13	0:00:01	66	58	13
2-Jun-13	1:00:01	69	68	3
3-Jun-13	2:00:01	70	73	17
4-Jun-13	3:00:01	67	63	2
5-Jun-13	4:00:01	68	74	16

Now that we've constructed the data into a useful format, we can use different reporting tools to analyze the results.

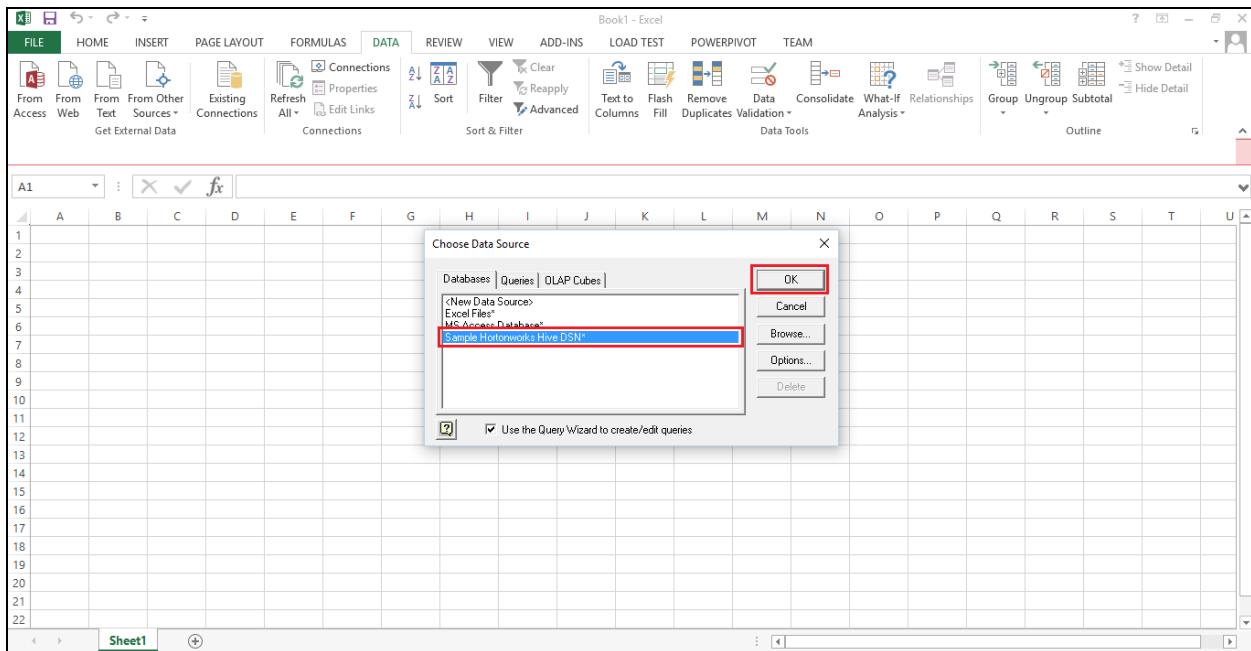
4. Reporting

4.1 Reporting with Excel

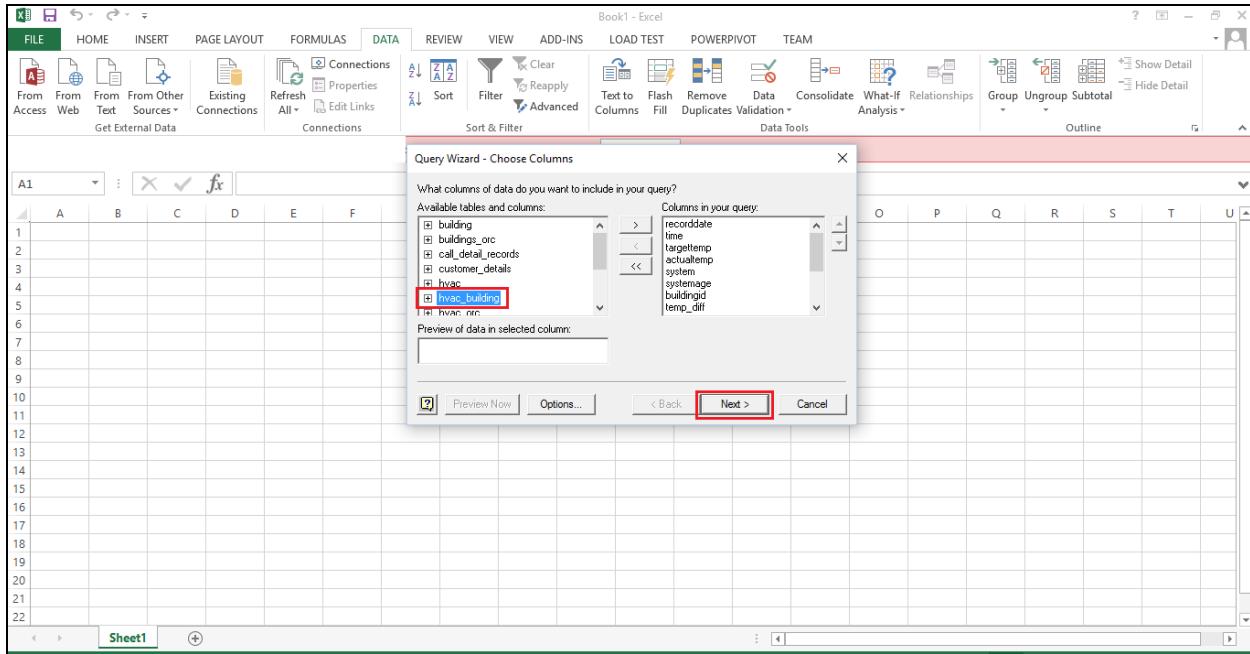
In Windows, open a new Excel workbook, then select Data > From Other Sources > From Microsoft Query



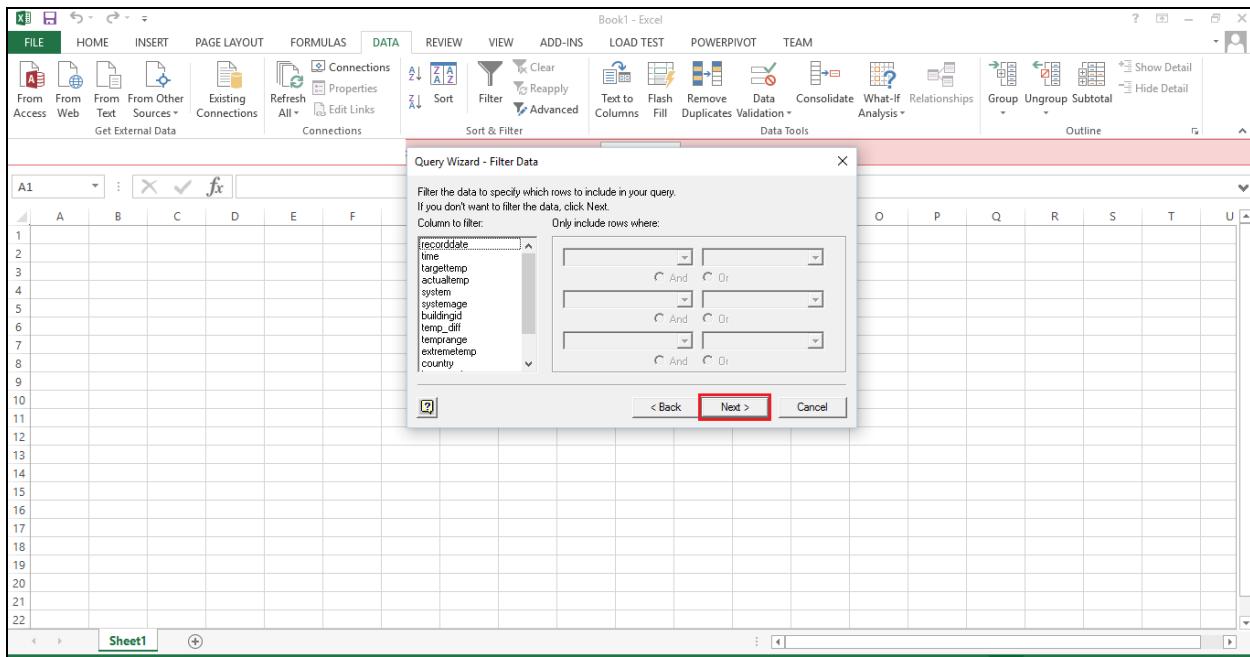
On the Choose Data Source pop-up, select the Hortonworks ODBC data source you installed previously, then click **OK**



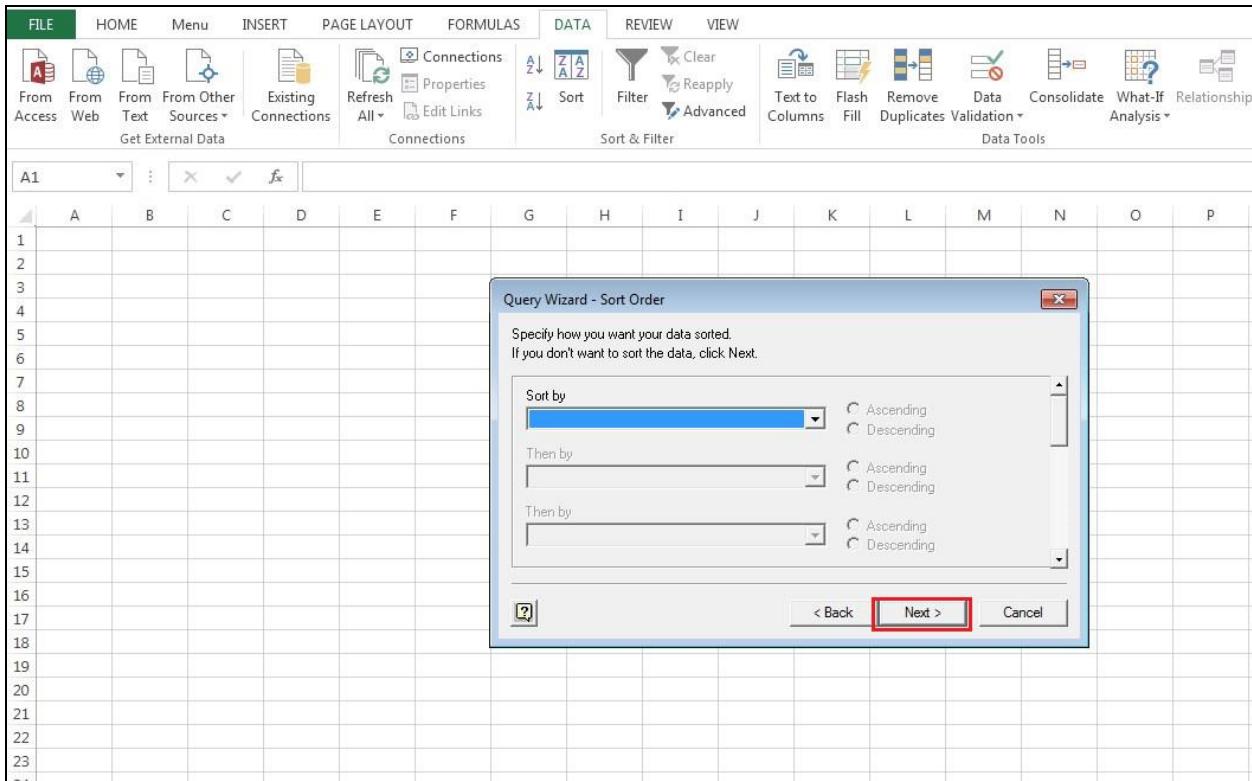
After the connection to the Sandbox is established, the Query Wizard appears. Select the “hvac_building” table in the Available tables and columns box, then click the right arrow button to add the entire “hvac_building” table to the query. Click **Next** to continue.



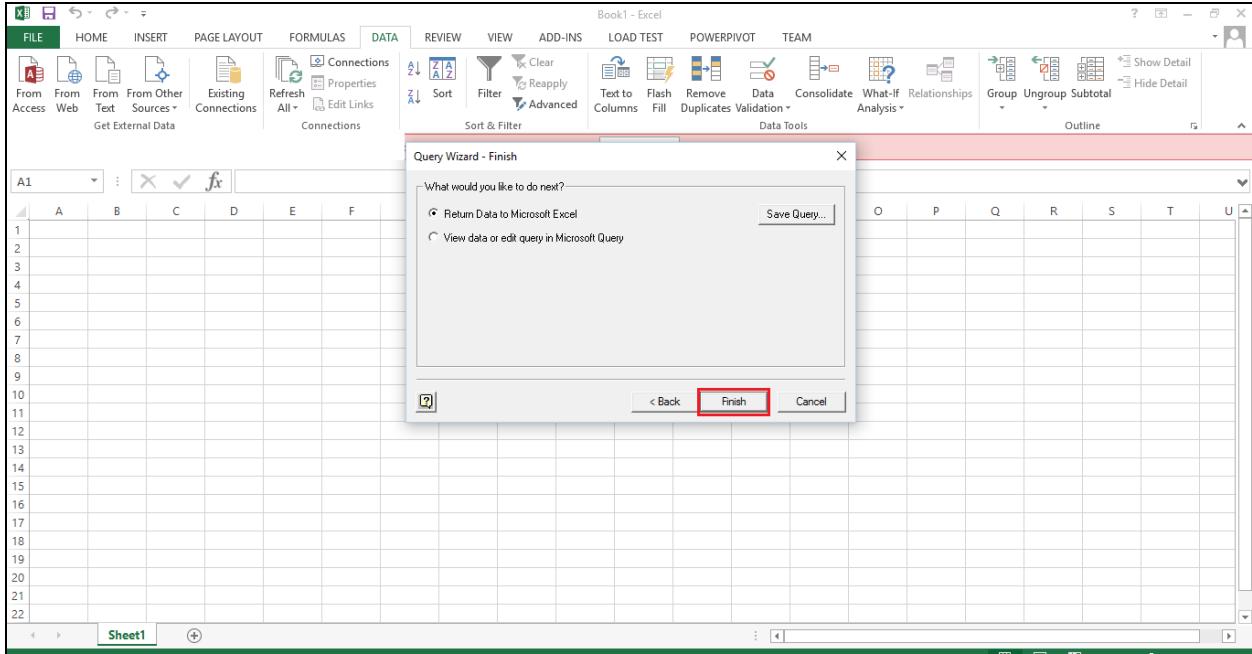
On the Filter Data screen, click **Next** to continue without filtering the data.



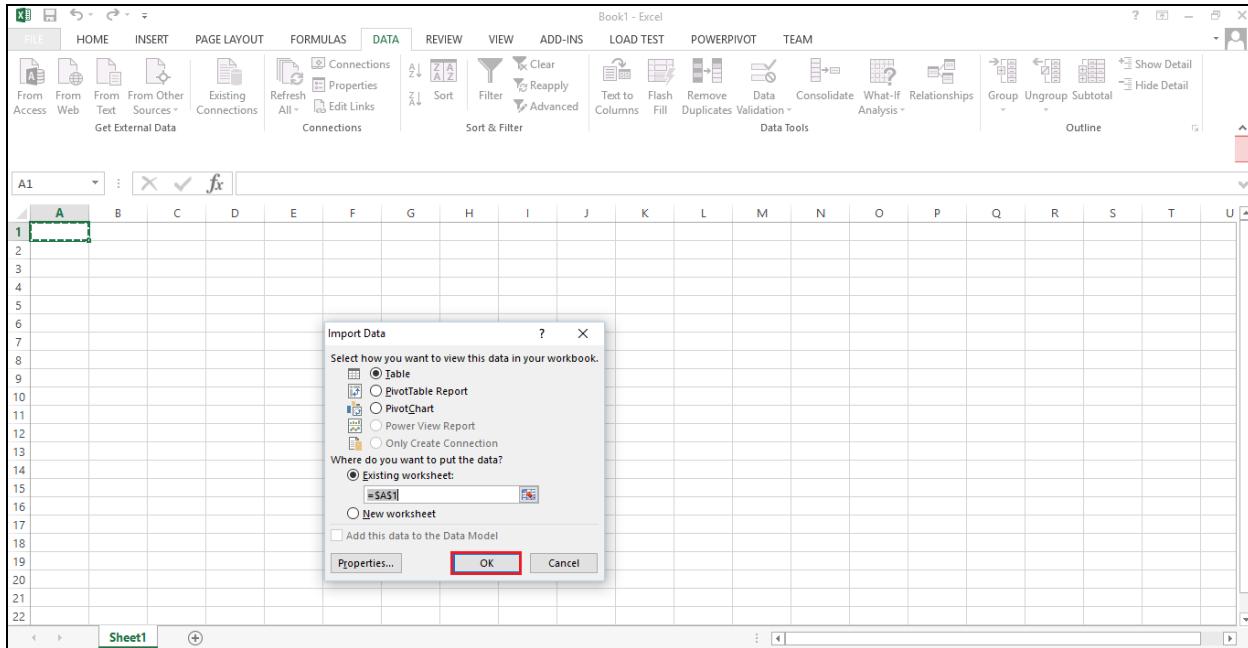
On the Sort Order screen, click **Next** to continue without setting a sort order



Click **Finish** on the Query Wizard Finish screen to retrieve the query data from the Sandbox and import it into Excel



On the Import Data dialog box, click **OK** to accept the default settings and import the data as a table



The imported query data appears in the Excel workbook.

	recorddate	time	targettemp	actualtemp	system	systemage	buildingid	temp	diff	temprange	extremetemp	country	hvacproduct	buildingage	buildingmgr
2	1-Jun-13	0:00:01	66	58	13	20	4	8	COLD	1	Finland	GG1919	17	M4	
3	2-Jun-13	1:00:01	69	68	3	20	17	1	NORMAL	0	Egypt	FN39TG	11	M17	
4	3-Jun-13	2:00:01	70	73	17	20	18	-3	NORMAL	0	Indonesia	JDNS77	25	M18	
5	4-Jun-13	3:00:01	67	63	2	23	15	4	NORMAL	0	Israel	ACMAX22	19	M15	
6	5-Jun-13	4:00:01	68	74	16	9	3	-6	HOT	1	Brazil	JDNS77	28	M3	
7	6-Jun-13	5:00:01	67	56	13	28	4	11	COLD	1	Finland	GG1919	17	M4	
8	7-Jun-13	6:00:01	70	58	12	24	2	12	COLD	1	France	FN39TG	27	M2	
9	8-Jun-13	7:00:01	70	73	20	26	16	-3	NORMAL	0	Turkey	AC1000	23	M16	
10	9-Jun-13	8:00:01	66	69	16	9	9	-3	NORMAL	0	Mexico	GG1919	11	M9	
11	10-Jun-13	9:00:01	65	57	6	5	12	8	COLD	1	Finland	FN39TG	26	M12	
12	11-Jun-13	10:00:01	67	70	10	17	15	-3	NORMAL	0	Israel	ACMAX22	19	M15	
13	12-Jun-13	11:00:01	69	62	2	11	7	7	COLD	1	South Africa	FN39TG	13	M7	
14	13-Jun-13	12:00:01	69	73	14	2	15	-4	NORMAL	0	Israel	ACMAX22	19	M15	
15	14-Jun-13	13:00:01	65	61	3	2	6	4	NORMAL	0	Singapore	AC1000	9	M6	
16	15-Jun-13	14:00:01	67	59	19	22	20	8	COLD	1	Argentina	ACMAX22	19	M20	
17	16-Jun-13	15:00:01	65	56	19	11	8	9	COLD	1	Australia	JDNS77	25	M8	
18	17-Jun-13	16:00:01	67	57	15	7	6	10	COLD	1	Singapore	AC1000	9	M6	
19	18-Jun-13	17:00:01	66	57	12	5	13	9	COLD	1	Saudi Arabia	JDNS77	25	M13	
20	19-Jun-13	18:00:01	69	58	8	22	4	11	COLD	1	Finland	GG1919	17	M4	
21	20-Jun-13	19:00:01	67	55	17	5	7	12	COLD	1	South Africa	FN39TG	13	M7	
22	21-Jun-13	20:00:01	69	72	7	5	17	-3	NORMAL	0	Egypt	FN39TG	11	M17	

4.1.1 Visualizing the sensor data using Excel Power View

In the Excel worksheet with the imported “hvac_building” table, select **Insert > Power View** to open a new Power View report

The screenshot shows the Microsoft Excel ribbon with the 'POWERVIEW' tab highlighted. A tooltip 'Insert a Power View Report' is displayed over the 'Power View' button. The main area of the screen displays a data table titled 'Book1 - Excel' with columns including recorddate, time, targettemp, actualtemp, system, systemage, buildingid, temp_diff, temprange, extremetemp, country, hvacproduct, buildingage, and buildingmgr.

Progress bar will appear on the screen, wait for some time the Power View will appear.

The screenshot shows the Microsoft Excel ribbon with the 'POWERVIEW' tab highlighted. A progress dialog box titled 'Opening Power View' is centered on the screen, displaying the message 'Working on opening Power View sheet...' and a green progress bar. The data table from the previous screenshot is visible in the background.

The Power View Fields area appears on the right side of the window, with the data table displayed on the left. Drag the handles or click the Pop Out icon to maximize the size of the data table

The screenshot shows the Microsoft Excel ribbon with the 'POWER VIEW' tab selected. On the left, there's a table of data with columns: recorddate, time, targettemp, actualtemp, system, and systemage. In the center, there's a 'Filters' pane. On the right, the 'Power View Fields' pane is open, listing various fields from a table named 'Table_Query_from_Sample_Hort'. The 'FIELDS' section contains 'recorddate', 'time', 'targettemp', 'actualtemp', 'system', and 'systemage'. The 'ACTIVE' section lists 'actualtemp', 'buildingage', 'buildingid', 'buildingmgr', 'country', 'extremetemp', 'hvacproduct', and 'recorddate'. The 'country' and 'extremetemp' checkboxes are checked.

In the Power View Fields area, select the checkboxes next to the **country** and **extremetemp** fields, and clear all of the other checkboxes. You may need to scroll down to see all of the check boxes

This screenshot is identical to the one above, but with a red box highlighting the 'country' and 'extremetemp' checkboxes in the 'ACTIVE' section of the 'Power View Fields' pane. All other checkboxes are cleared.

In the **FIELDS** box, click the down-arrow at the right of the **extremetemp** field, then select **Count (Not Blank)**

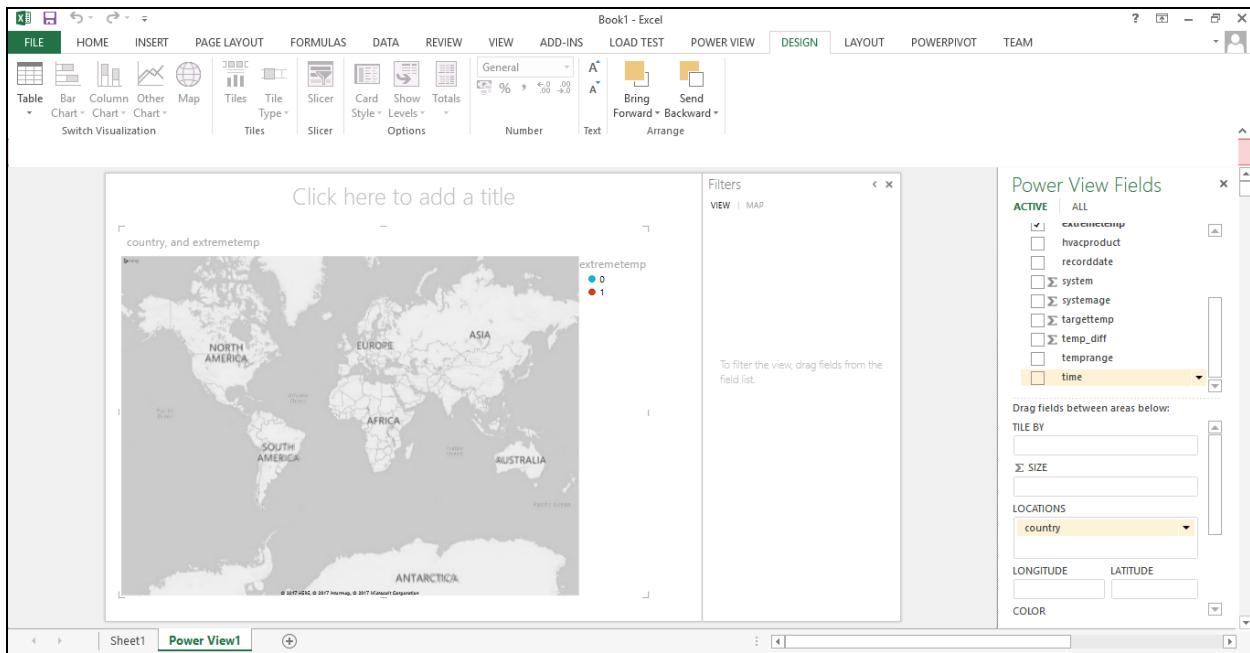
The screenshot shows the Microsoft Excel ribbon with the 'POWER VIEW' tab selected. In the main area, there is a table with columns 'country' and 'extremetemp'. The 'extremetemp' column contains binary values (0 or 1). To the right, the 'Power View Fields' ribbon tab is active, showing a list of fields from a table named 'Table_Query_from_Sample_Hortonworks'. The 'extremetemp' field is checked. A dropdown menu is open for this field, with the option 'Count (Not Blank)' highlighted and surrounded by a red box.

country	extremetemp
Argentina	0
Argentina	1
Australia	0
Australia	1
Belgium	0
Belgium	1
Brazil	0
Brazil	1
Canada	0
Canada	1
China	0
China	1
Egypt	0
Egypt	1
Finland	0
Finland	1
France	0
France	1
Germany	0
Germany	1
Hong Kong	0
Hong Kong	1
Indonesia	0
Indonesia	1
Israel	0
Israel	1
Mexico	0
Mexico	1
Saudi Arabia	0

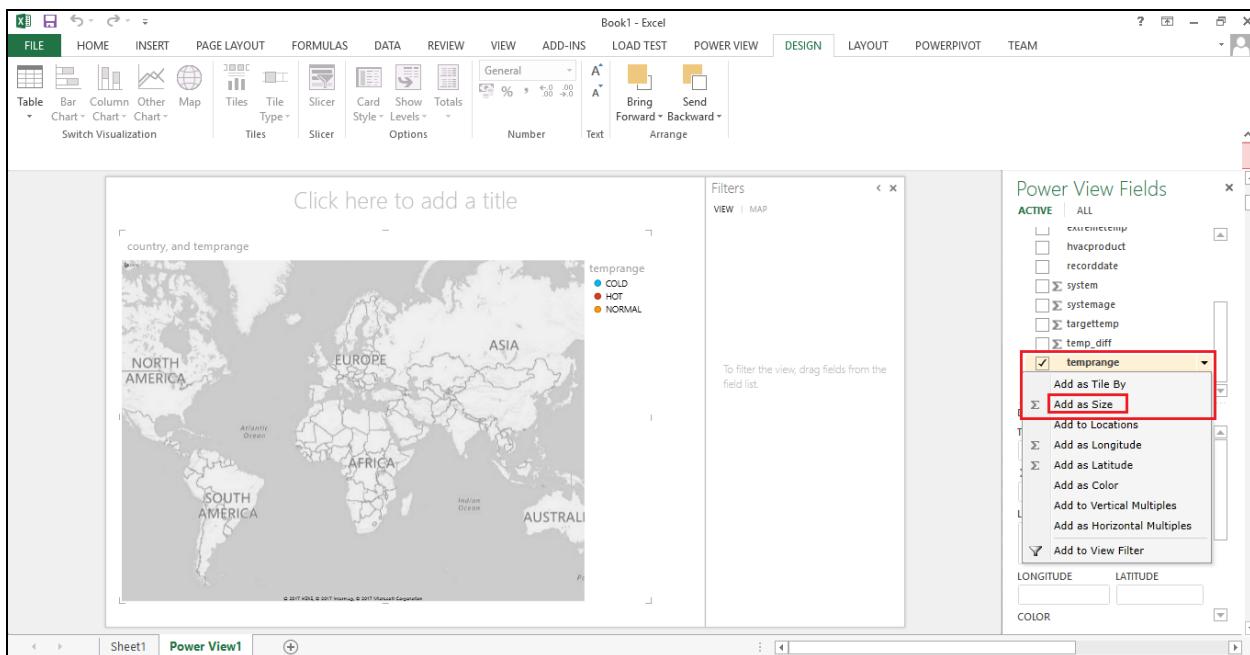
Click Map on the Design tab in the top menu

The screenshot shows the Microsoft Excel ribbon with the 'DESIGN' tab selected. In the main area, there is a map visualization titled 'Click here to add a title'. The map shows various locations with colored dots representing data points. To the right, the 'Power View Fields' ribbon tab is active, showing a list of fields from the same table. The 'extremetemp' and 'hvacproduct' fields are checked. A dropdown menu is open for the 'hvacproduct' field in the 'FIELDS' list, with the option '# Count of extremetemp' highlighted and surrounded by a red box.

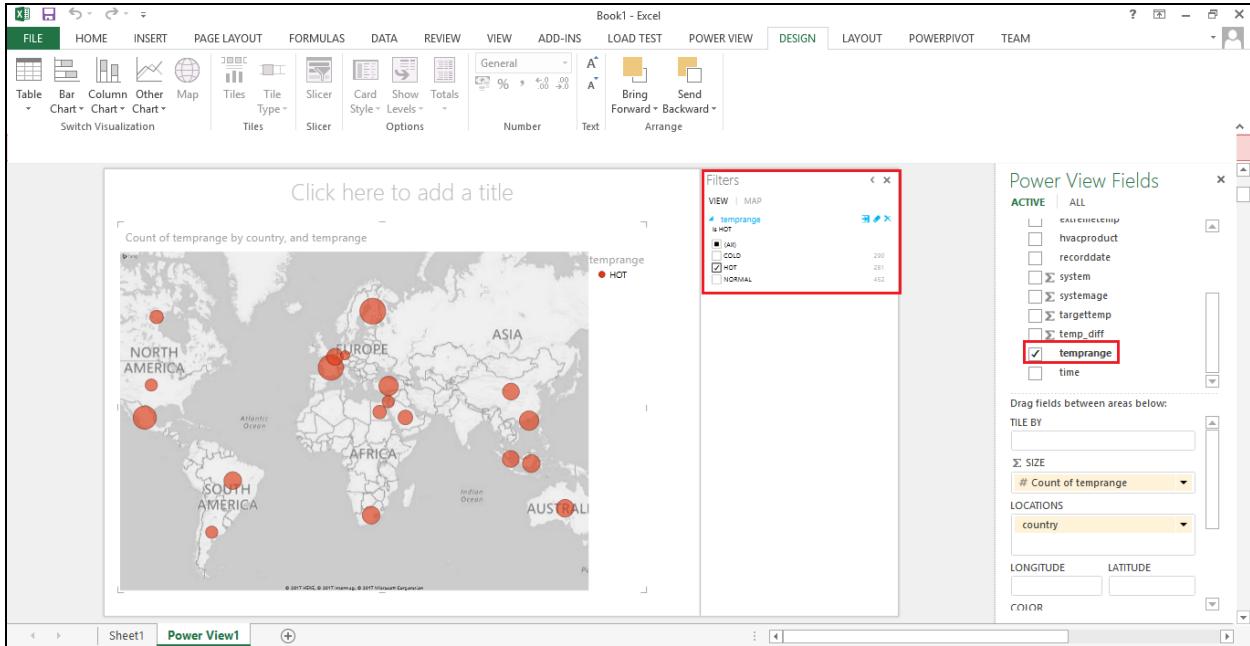
The map view displays a global view of the data



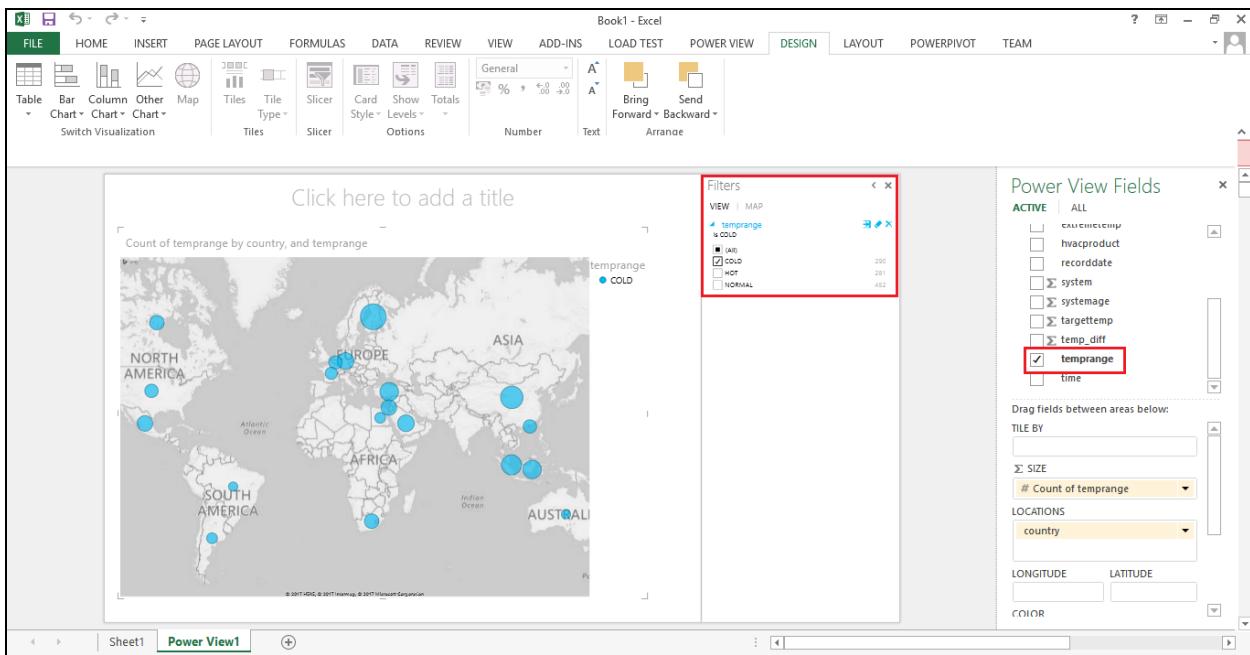
In the Power View Fields area, clear the **extremetemp** checkbox and select the **temprange** checkbox.
Click the down-arrow at the right of the **temprange** field, then select **Add as Size**



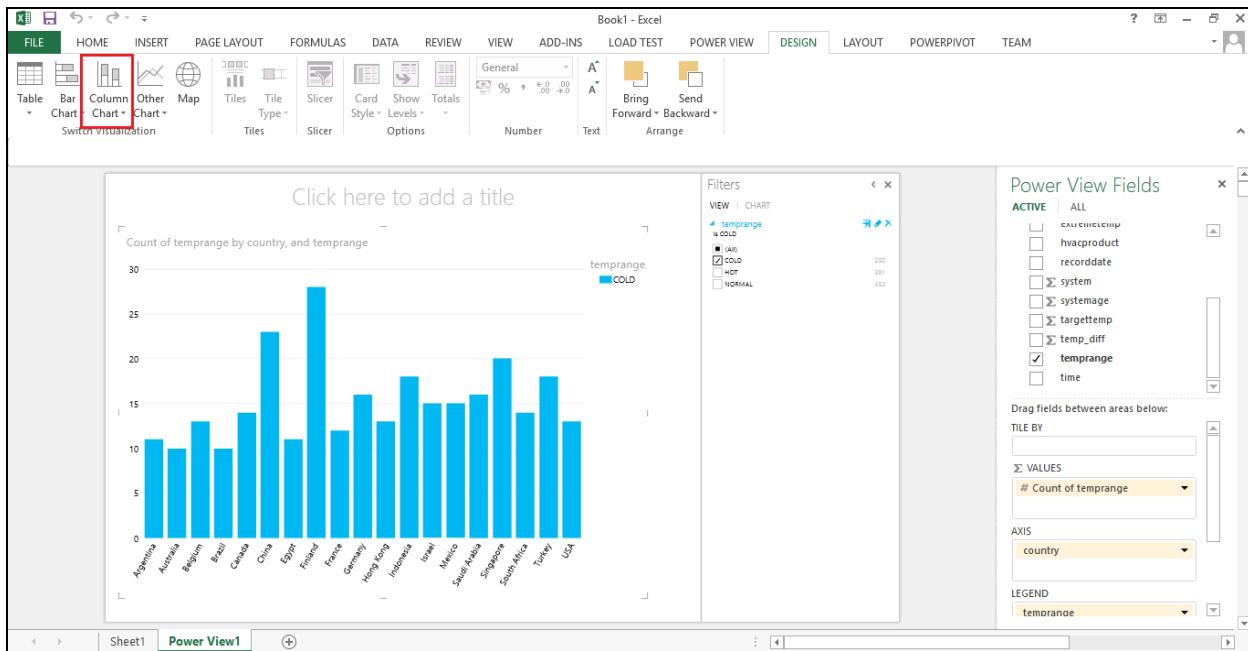
Drag **temprange** from the Power View Fields area to the Filters box, then select the **HOT** checkbox. We can see that the buildings in hot.



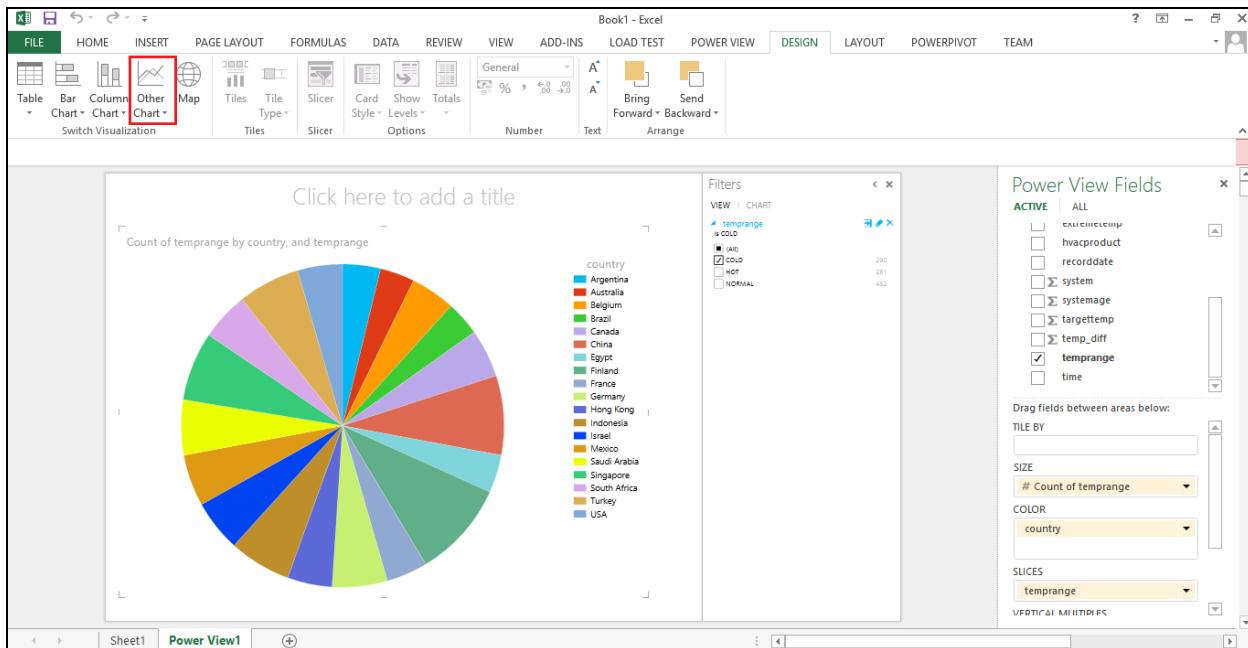
In the Filters box, clear the **HOT** checkbox and select the **COLD** checkbox. We can see that the buildings in cold.



Click on the Column Chart to visualize the data

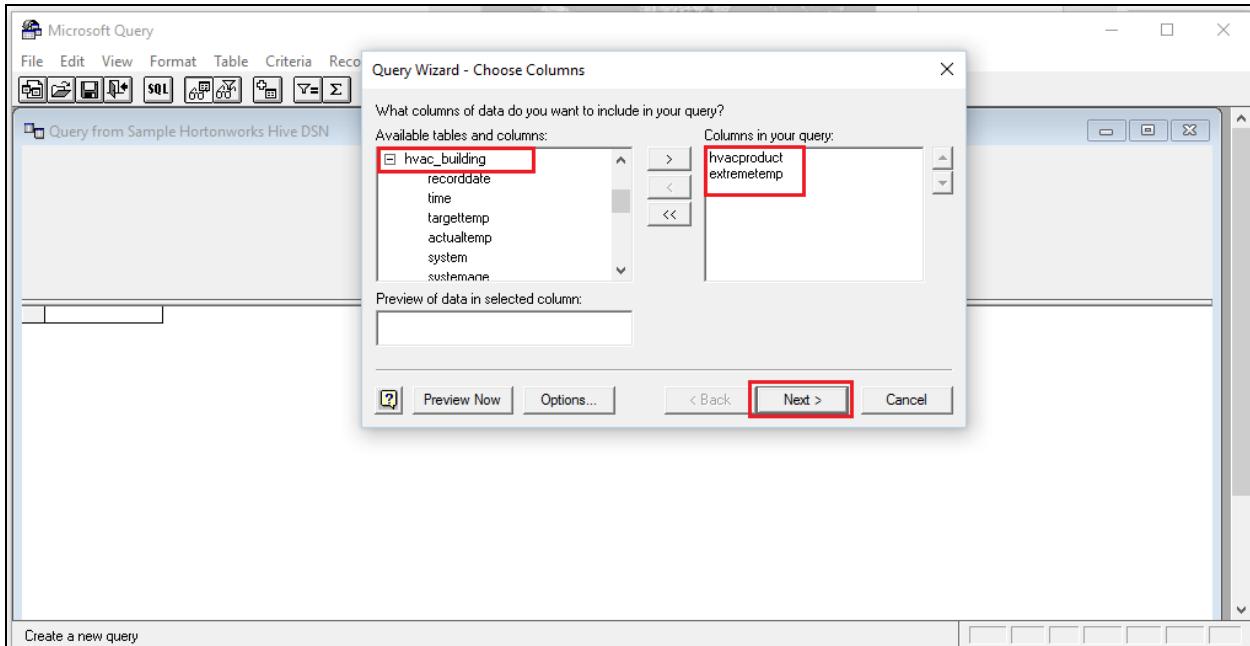


Click on the Other Charts and select pie Chart to visualize the data in Pie chart.



Open a new Excel worksheet, then select **Data > From Other Sources

From Microsoft Query to access the hvac_building table. Follow the same procedure as before to import the data, but this time only select the “hvaproduct” and “extremetemp” columns then click **Next**



In the Excel worksheet with the imported “hvaproduct” and “extremetemp” columns, select **Insert > Power View** to open a new Power View report.

	A	B
1	extremetemp	hvaproduct
2	1	GG1919
3	0	FN39TG
4	0	JDNS77
5	0	ACMAX22
6	1	JDNS77
7	1	GG1919
8	1	FN39TG
9	0	AC1000
10	0	GG1919
11	1	FN39TG
12	0	ACMAX22
13	1	FN39TG
14	0	ACMAX22
15	0	AC1000
16	1	ACMAX22
17	1	JDNS77
18	1	AC1000
19	1	JDNS77
20	1	GG1919
21	1	FN39TG
22	0	FN39TG
23	0	GG1919

Click the Pop Out icon to maximize the size of the data table. In the FIELDS box, click the down-arrow at the right of the **extremetemp** field, then select **Count (Not Blank)**, we will get the following screen.

The screenshot shows the Microsoft Excel ribbon with the 'POWER VIEW' tab selected. A Power View Fields pane is open on the right side. In the 'FIELDS' section, there is a dropdown menu next to the 'extremetemp' field. A red box highlights this dropdown menu. The data table in the main area shows the count of extreme temperatures for various HVAC products.

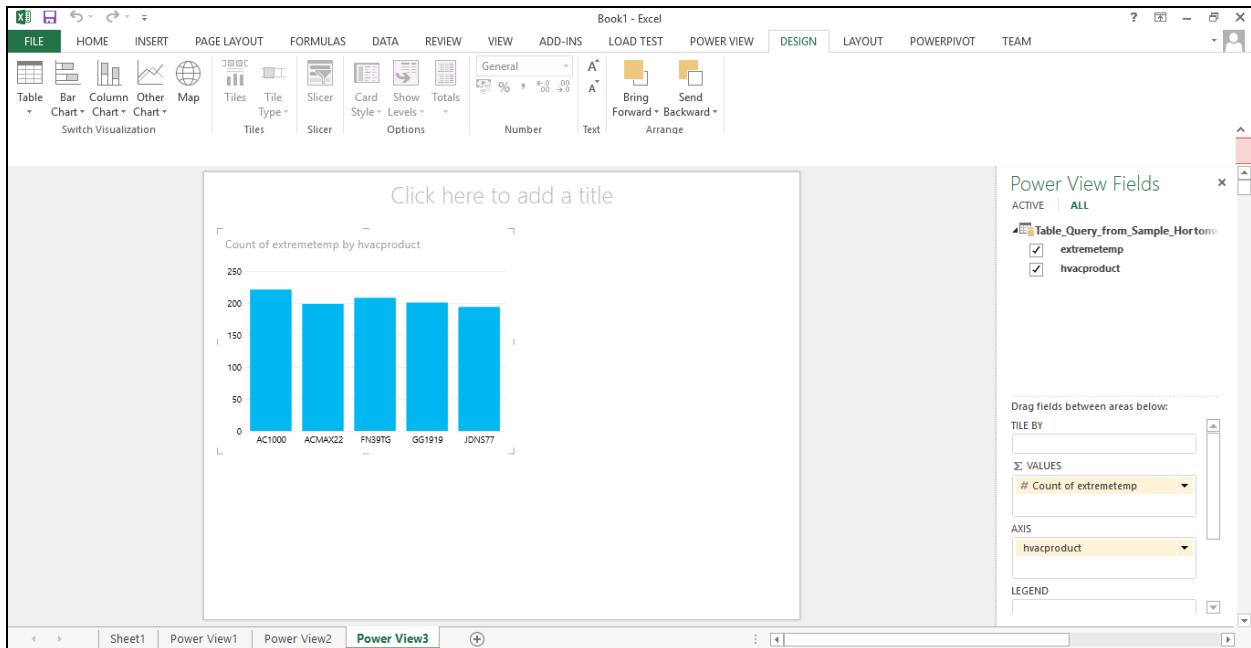
hvacproduct	Count of extremetemp
AC1000	221
ACMAX22	199
FN3RTG	208
GG1919	201
JDN577	194
Total	1,023

Select **Column Chart > Stacked Column** in the top menu

The screenshot shows the Microsoft Excel ribbon with the 'POWER VIEW' tab selected. A Power View Fields pane is open on the right side. In the 'FIELDS' section, there is a dropdown menu next to the '# Count of extremetemp' field. A red box highlights this dropdown menu. The data table in the main area shows the count of extreme temperatures for various HVAC products.

hvacproduct	Count of extremetemp
AC1000	221
ACMAX22	199
FN3RTG	208
GG1919	201
JDN577	194
Total	1,023

Click the down-arrow next to sort by hvacproduct in the upper left corner of the chart area, then select Count of extremetemp. We can see the following chart.



4.2 Visualizing the sensor data using Zeppelin

Apache Zeppelin makes data reporting easy on Hadoop. It has direct connections to Apache Spark and Hive in your cluster and allows you to create visualizations and analyze your data on the fly.

To start the Zeppelin go to <http://<host /DNS name of Hortonworks Sandbox>:9995>

Zeppelin generally runs on the 9995 port number, to know the exact port number navigate to the Zeppelin Notebook service click on the Config and go to Advanced Zeppelin Configurations.

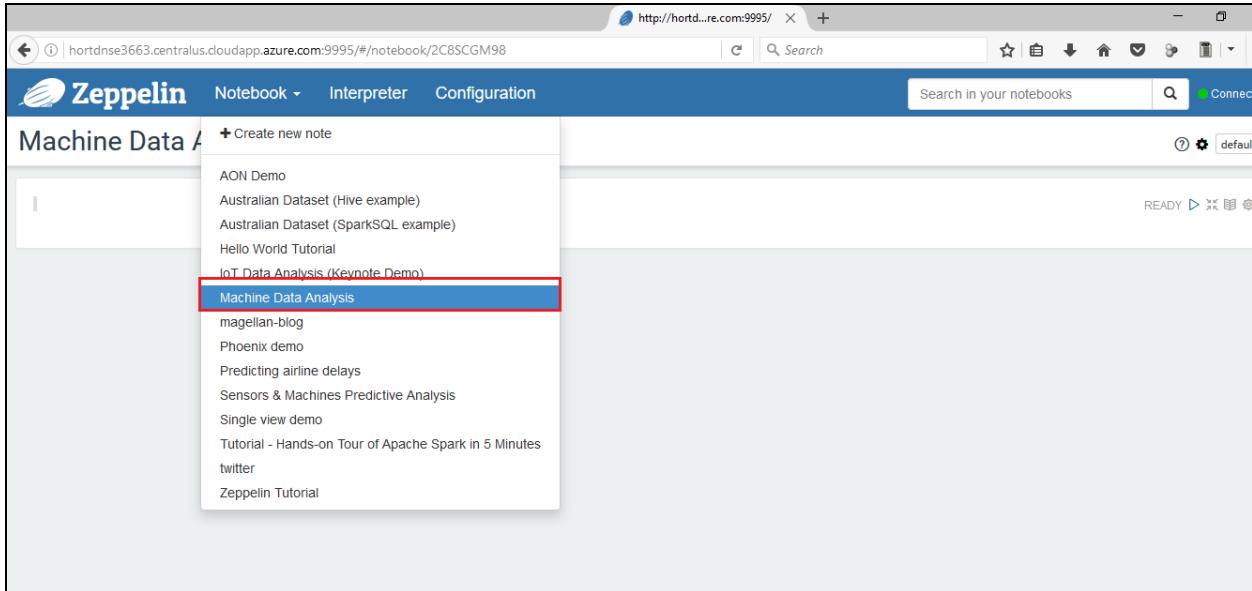
Once we open the Zeppelin we will get the following page.

The screenshot shows the Zeppelin web interface at the URL <http://hortd...re.com:9995/>. The page title is "Welcome to Zeppelin!". The left sidebar contains a "Notebook" section with a "Create new note" button, a "Help" section with a "Get started with Zeppelin documentation" link, and a "Community" section with links to "Mailing list", "Issues tracking", and "Github". A large blue hot air balloon illustration is on the right. A search bar at the top right says "Search in your notebooks".

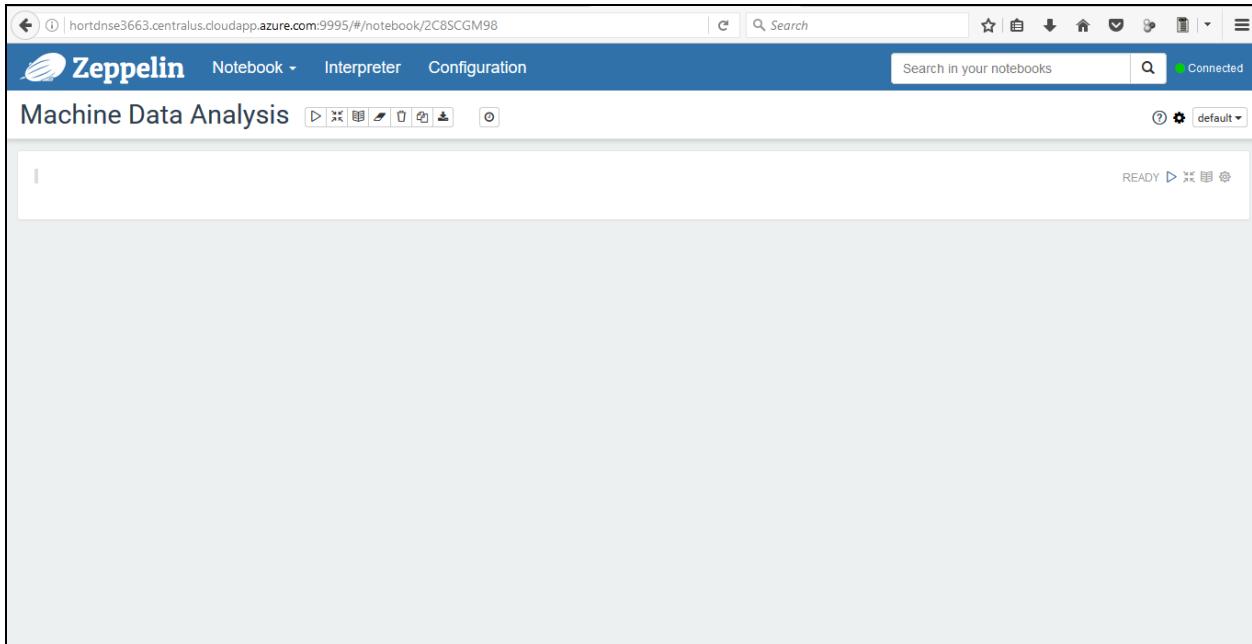
Hove over **Notebook**. Use the dropdown menu and **Create a new note**.

This screenshot is identical to the one above, but the "Create new note" button in the "Notebook" sidebar has been highlighted with a red box.

Head back to the **Zeppelin** homepage. Use the Notebook dropdown menu to open the new notebook **Machine Data Analysis**.



We will get the following screen



Type the following query into the note, then run it by clicking the **Run arrow** or by using the shortcut **Shift+Enter**.

```
%hive
```

```
select country, extremetemp, temprange from hvac_building
```

The screenshot shows the Zeppelin Notebook interface running on a browser. In the top navigation bar, the URL is `http://hortd...re.com:9995/#/notebook/2C8SCGM98`. The main area displays a code cell containing the following Hive query:

```
%hive
select country, extremetemp, temprange from hvac_building
```

To the right of the code cell, there is a toolbar with several icons. A red box highlights the "Run this paragraph (Shift+Enter)" button, which is the second icon from the left in the toolbar.

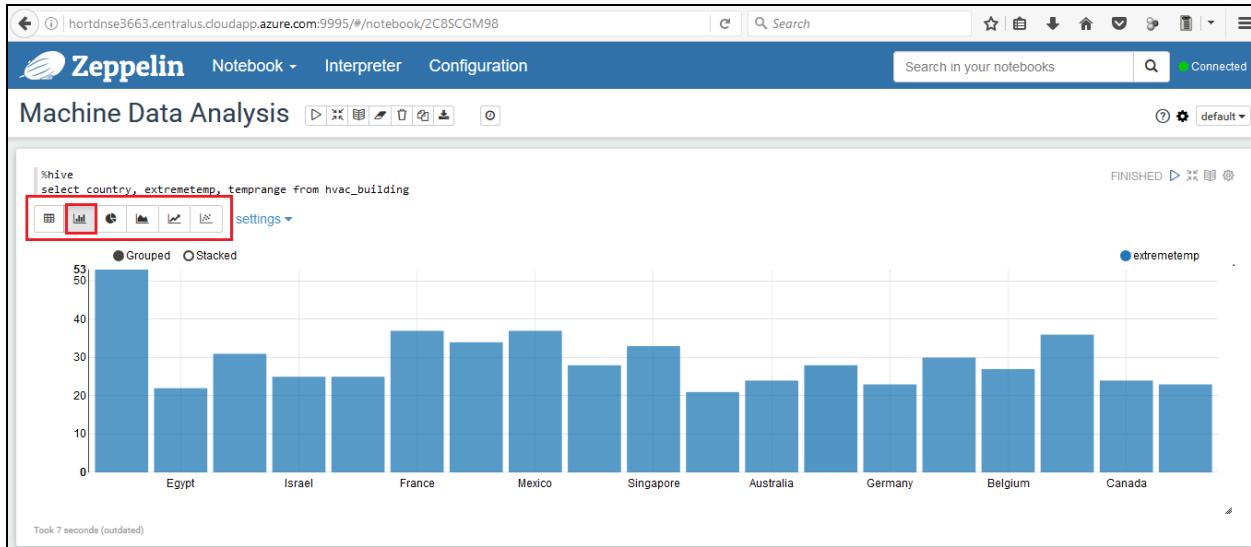
Once we run the query we will get the following details

The screenshot shows the Zeppelin Notebook interface after the query has been executed. The status bar at the top right indicates "FINISHED". The main area displays the results of the query as a table:

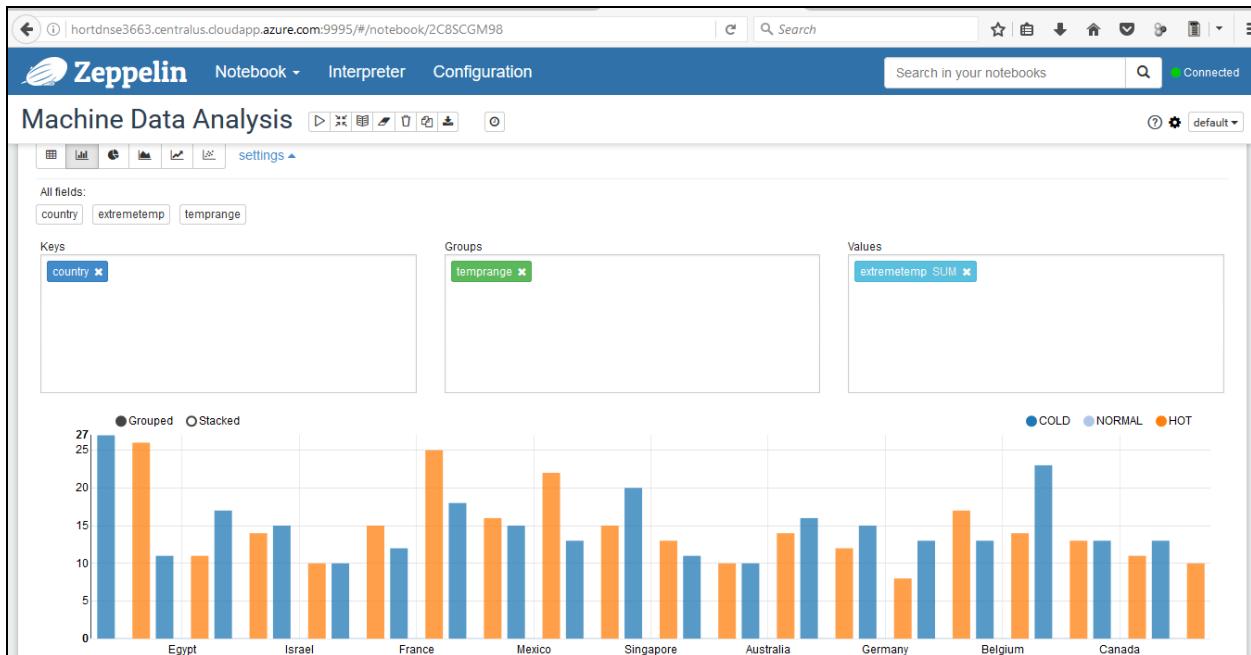
country	extremetemp	temprange
Finland	1	COLD
Egypt	0	NORMAL
Indonesia	0	NORMAL
Israel	0	NORMAL
Brazil	1	HOT
Finland	1	COLD
France	1	COLD
Turkey	0	NORMAL
Mexico	0	NORMAL

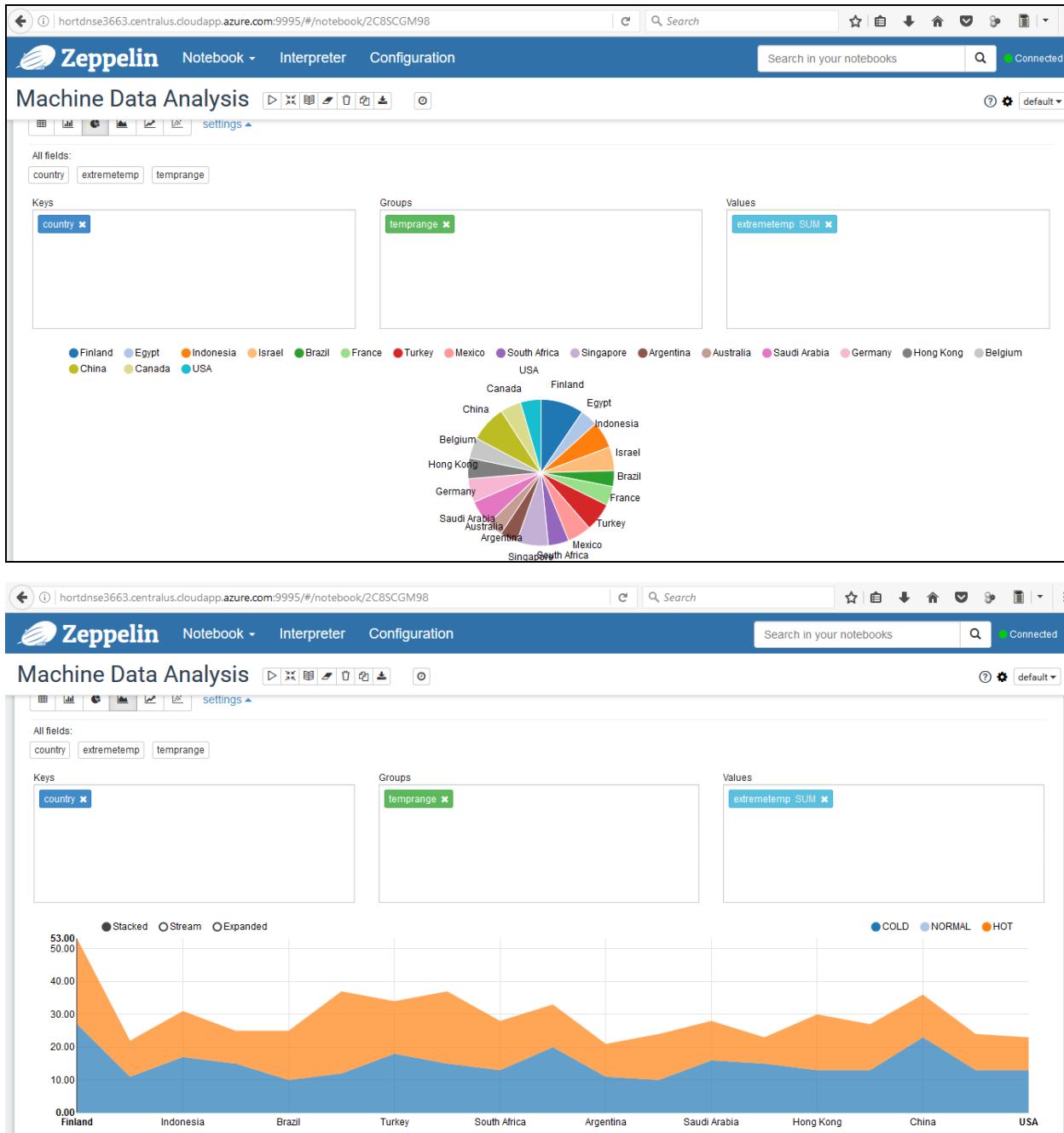
A message at the bottom of the table says "Took 7 seconds". To the right of the table, there is a toolbar with several icons. A red box highlights the "FINISHED" status indicator, which is the first icon in the toolbar.

We can view a chart of the data by clicking the chart button located just under the query, following screens will show us different reports for the above data.



Click **settings** to open up more advanced settings for creating the chart. Here you can experiment with different values and columns to create different types of charts





5. Conclusion

In this test drive we have seen how the Hortonworks Data Platform (HDP) can store and analyze sensor data. With real-time access to massive amounts of temperature and other types of data.