

Source: C# Corner (www.c-sharpcorner.com)

PRINT

Article



Run A Hive Query From Azure Portal

By [Abdul Rasheed Feroz Khan](#) on **Sep 10 2016**

Introduction

This article will help you learn Hadoop using HDInsight on Windows. Here, we will be running a Hive Query from the Azure Portal. Look at my previous article to understand what is Hadoop, what is HDInsight, and how to create a cluster in Hadoop.

Links

Click [here](#) for Kickstart Hadoop in HDInsight on Windows.

What is Hive?

Hive provides a Database Query Interface to Apache Hadoop. It is a component of HDP (Hortonworks Data Platform). Hive provides an interface like SQL, for storing the data in HDP.

In the previous article, we have created an HDInsight Cluster. Now, we will be running a Hive job to query a sample Hive table. We will be using hivesampletable which is available with HDInsight clusters. The table contains data about mobile device manufacturers, models, and platforms. A Hive query, on this table, retrieves the data for mobile devices by a specific manufacturer.

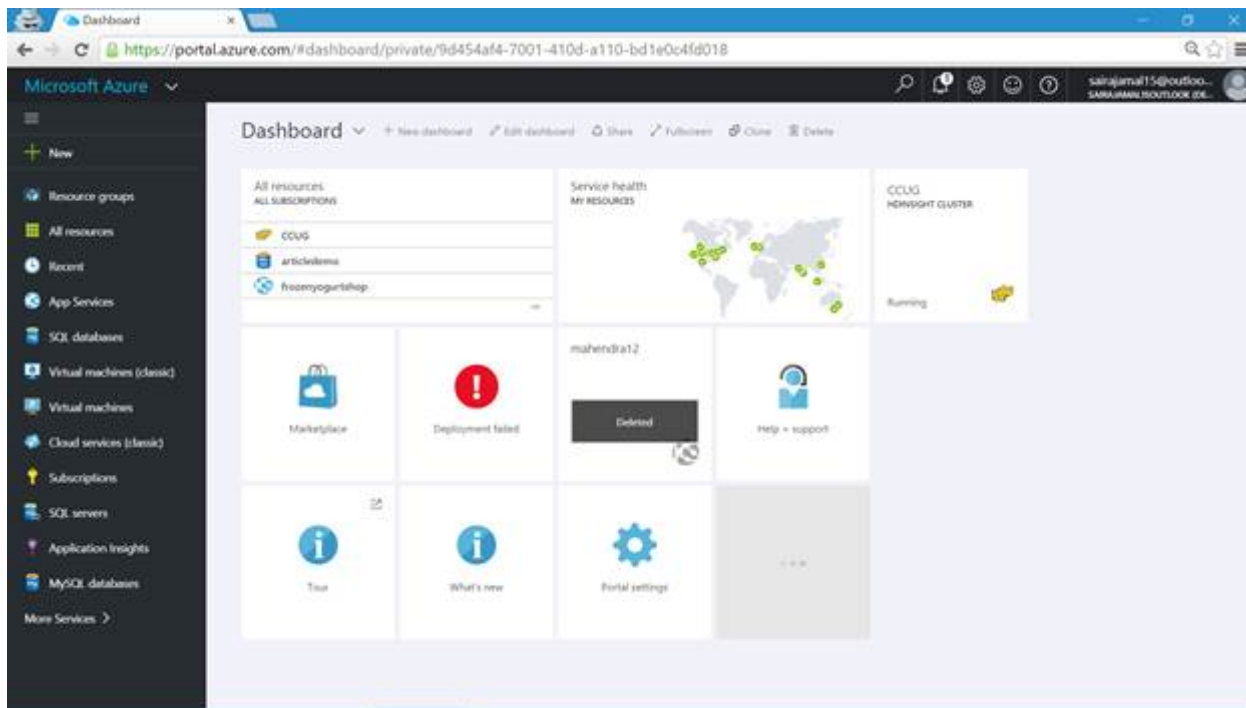
Pre-requisites

1. An Azure Subscription.

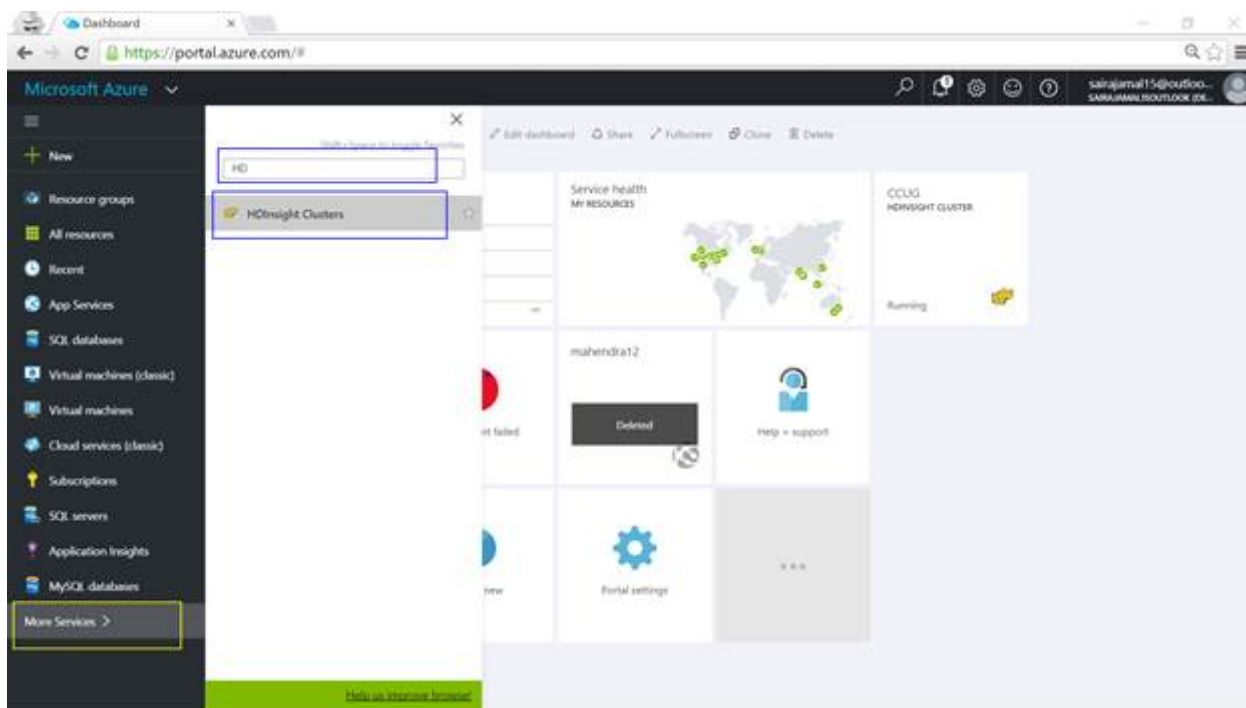
Note: Click [here](#) to get a free trial account of Azure.

Follow the below-mentioned steps now.

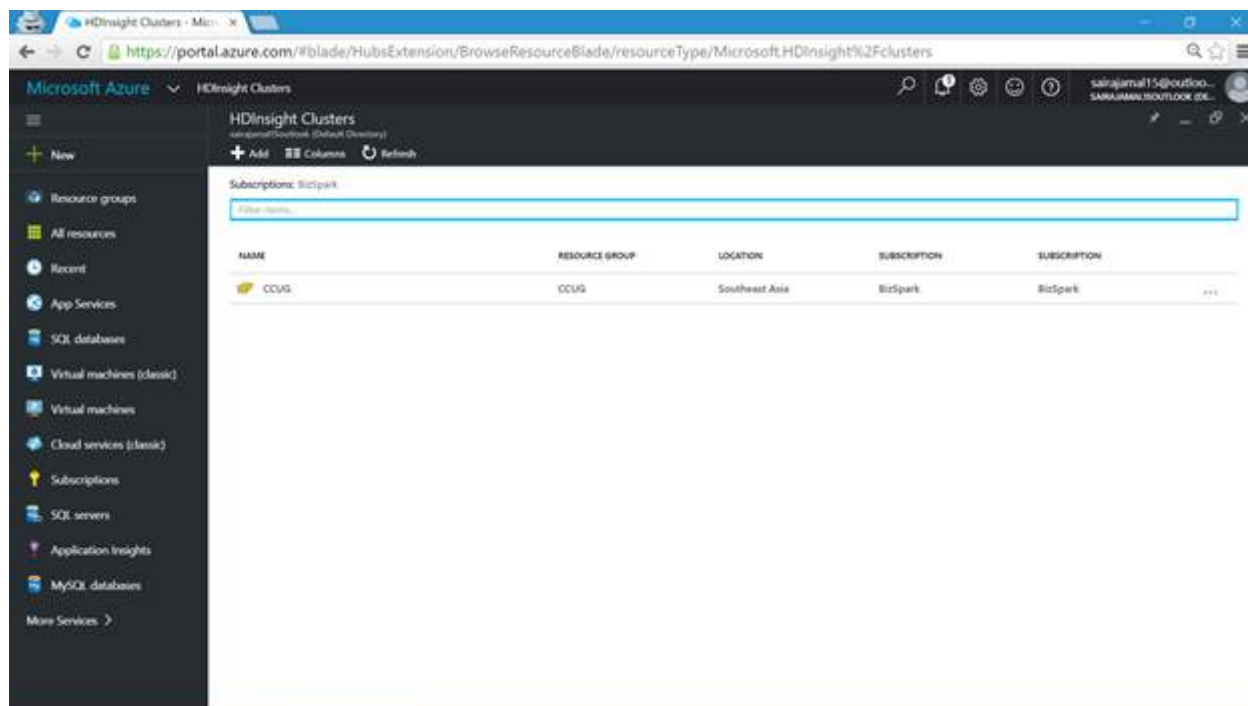
Step 1 - Go to the [Azure Portal](#) and sign in to your account.



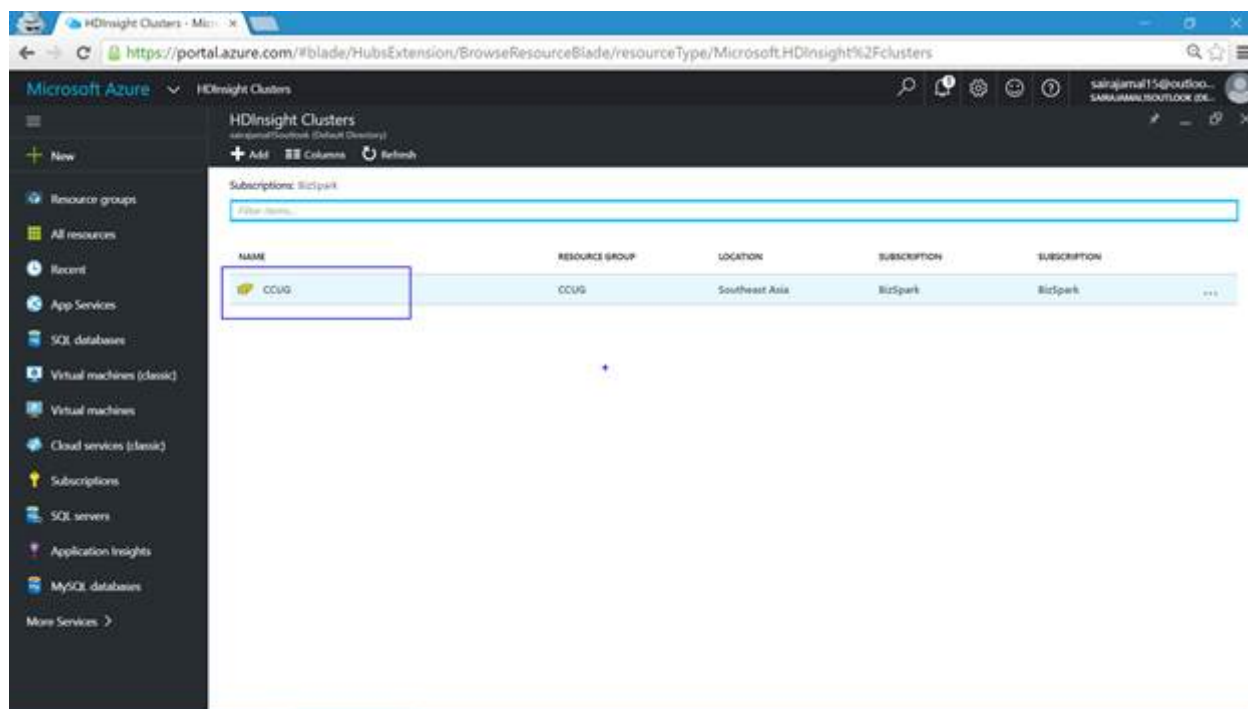
Step 2 - Click on More Services - HDInsight Clusters and click on HDInsight Clusters.



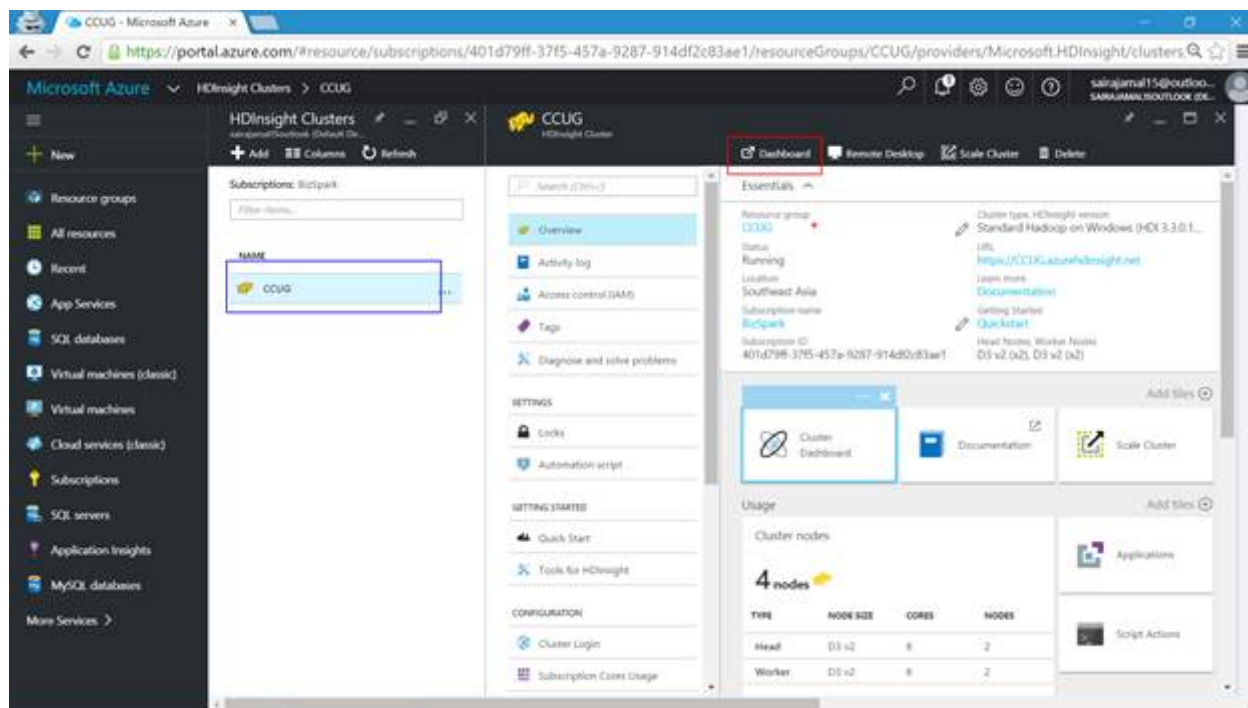
You can find the list of the Clusters that are available.



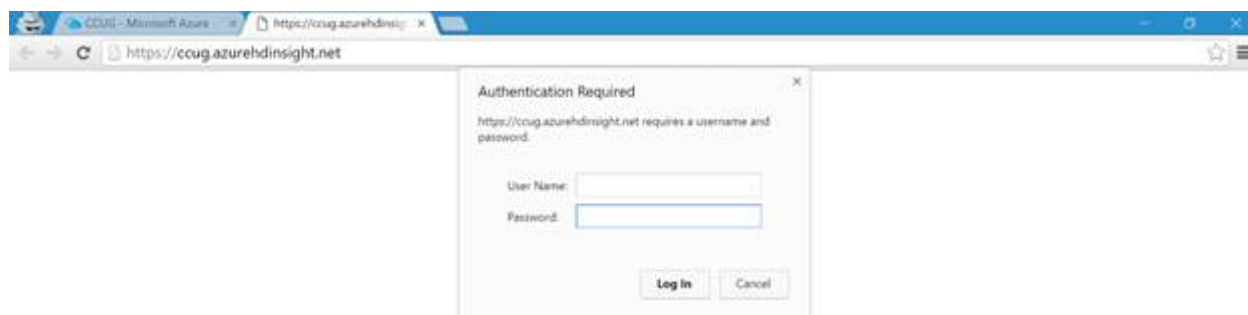
Step 3 - Click the name of the cluster that you want to use. Here, I will be using CCUG Cluster that I have created.



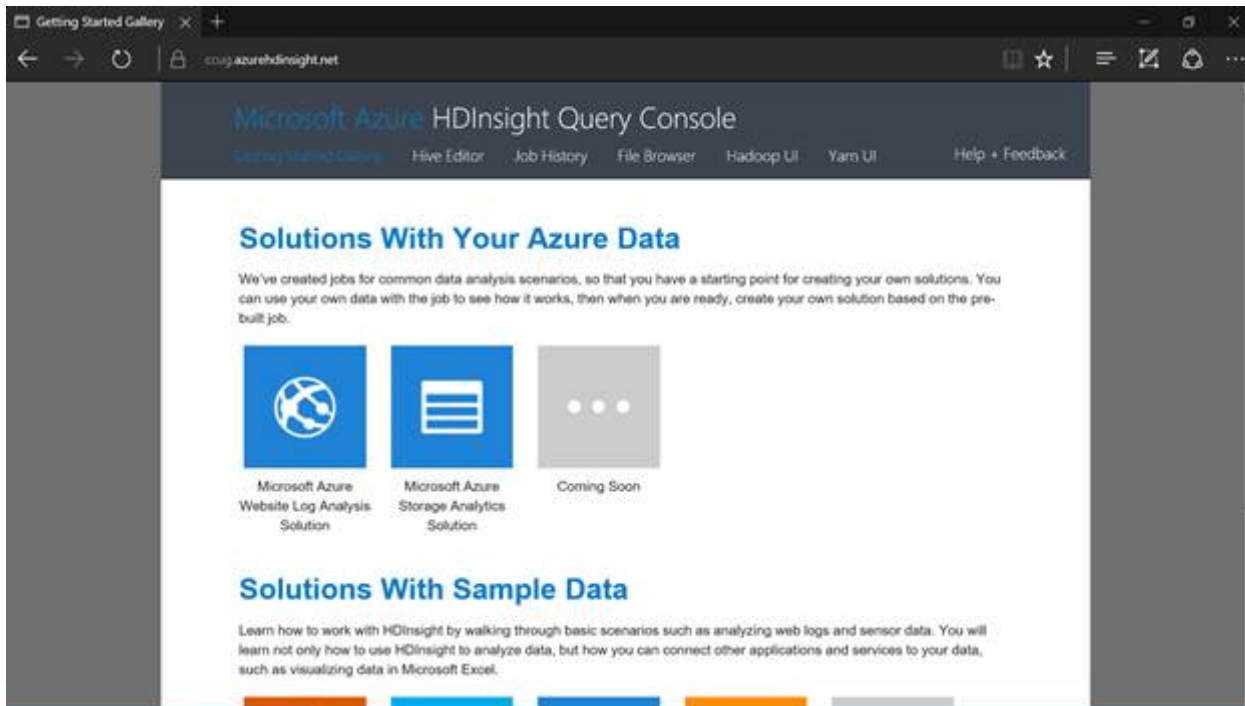
Step 4 - Go on the Dashboard of the cluster.



Clicking on the Dashboard will open a new webpage for you. Enter your login credentials for the cluster and login here.

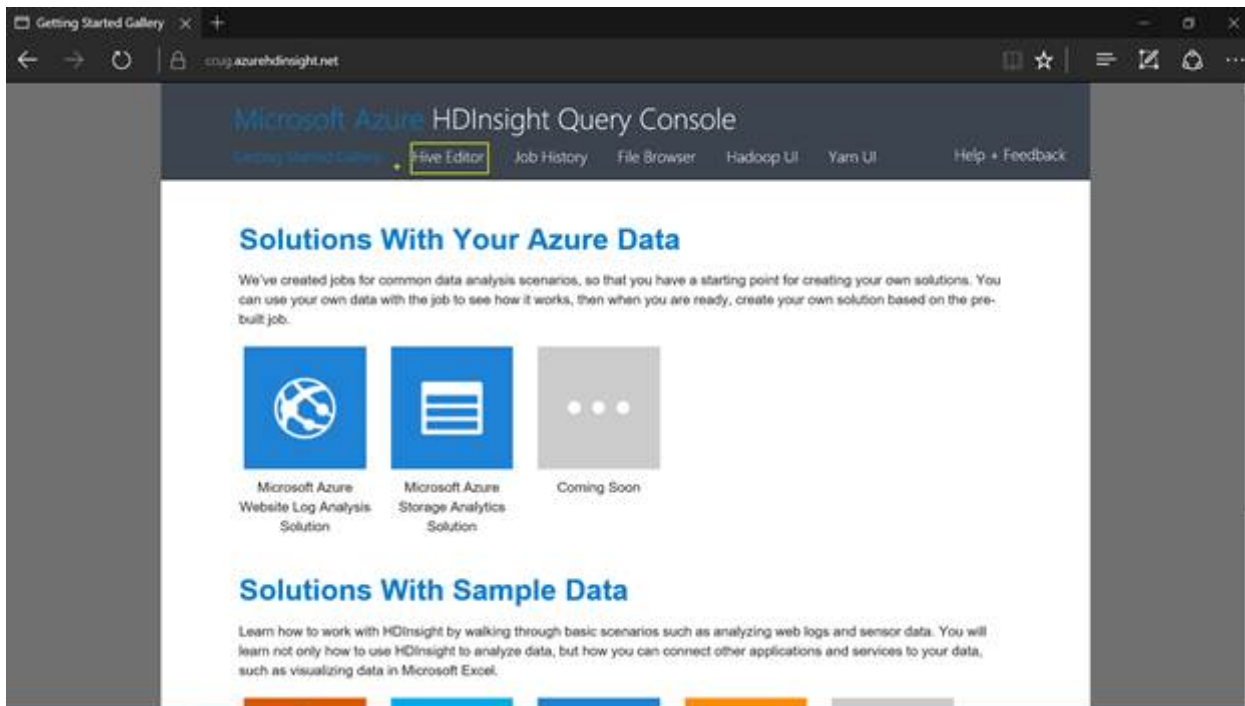


Note: Use the default name of “admin” for login and password that you gave while creating the particular cluster.



From the above home screen of your cluster, click on Hive Editor under the Dashboard. Here, you can find the URL – www.ccug.azurehdinsight.net.

This means, you can also login to the webpage directly without using Azure portal.



The Dashboard at the top of the page can help us submit the Hive queries. Check the Hadoop job logs and browse the files in the storage.

Step 5 - For Query name, enter "HTC20". This is the Job Title. Followed by this, enter the following Hive Query.

1. Select * from hivesampletable
2. WHERE devicemake LIKE "HTC%"
3. LIMIT 20;

Click on Submit button after entering the Hive Query.

Microsoft Azure HDInsight Query Console

Getting Started Gallery | Job History | File Browser | Hadoop UI | Yarn UI | Help + Feedback

Use this editor and submit a hive query to your cluster.

Query Name: HTC20

```
1 Select * from hivesampletable
2 WHERE devicename LIKE 'HTC*'
3 LIMIT 20;
```

Job Session: Updating Jobs **Submit**

Query Name	Date	Id	Action	Status
No data available in table				

Under the Job Session, you can find that the job has been completed.

Microsoft Azure HDInsight Query Console

Getting Started Gallery | Job History | File Browser | Hadoop UI | Yarn UI | Help + Feedback

Use this editor and submit a hive query to your cluster.

Query Name: HTC20

```
1 Select * from hivesampletable
2 WHERE devicename LIKE 'HTC*'
3 LIMIT 20;
```

Job Session: Updating Jobs **Submit**

Query Name	Date	Id	Action	Status
HTC20	08-09-2016 19:24:52	job_1473341675318_0001		Completed

Step 6 - Click on the query name, now, to see the output. Make a note of Job Start Time (UTC). We will be working on it later.

The screenshot shows the Microsoft Azure HDInsight Query Console interface. At the top, there's a navigation bar with links like 'Getting Started Gallery', 'Hive Editor', 'File Browser', 'Hadoop UI', and 'Yarn UI'. Below this, a 'Job Details' section displays a table with job information:

Hadoop Job Id	job_1473341675318_0001
Query Name	HTC20
Status	Completed
Job Start Time (UTC)	8/8/2018 1:54:52 PM

Below the job details, there's a 'Job Query' section with a text area containing the following SQL query:

```
Select * from hivesampletable  
WHERE devicename LIKE "HTC"  
LIMIT 20;
```

To the right of the query is a 'Download File' button. Below the query, there's a 'Job Output' section displaying a table of results:

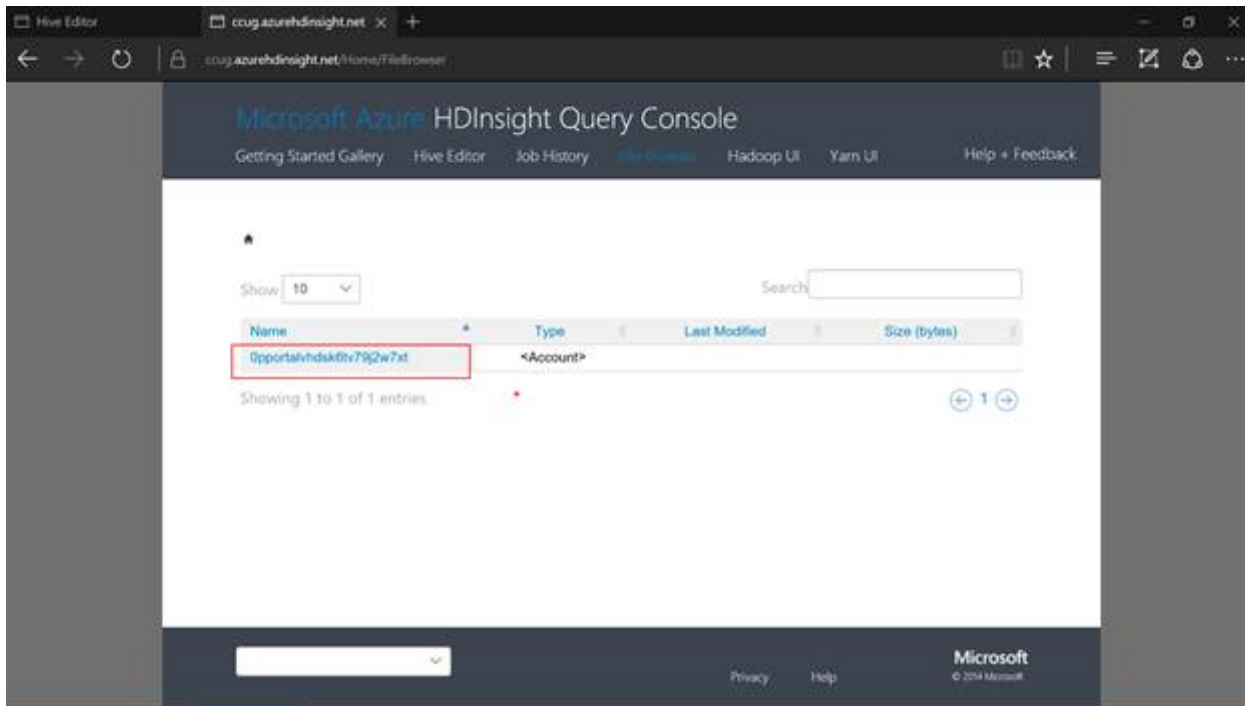
23	19:19:44	en-US	Android HTC	Incredible	Pennsylvania	United States	NULL	0
23	19:19:46	en-US	Android HTC	Incredible	Pennsylvania	United States	1.4757422	0

This page shows you the Job Output and Job Log. You have an option here where you can download the output file and log file.

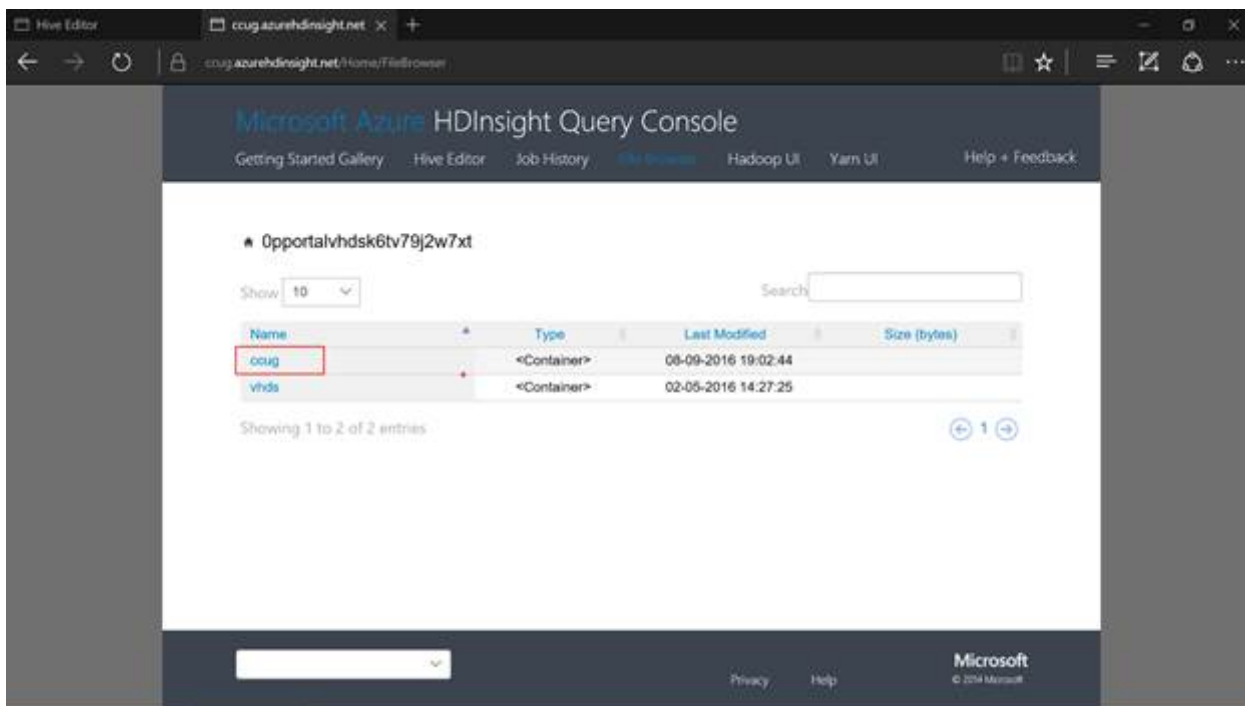
Step 7 - Now, let's browse the output file. Click on File Browser.

This screenshot is similar to the previous one, but the 'File Browser' tab in the navigation bar is highlighted with a yellow box. The rest of the interface, including the job details, query, and output, remains the same.

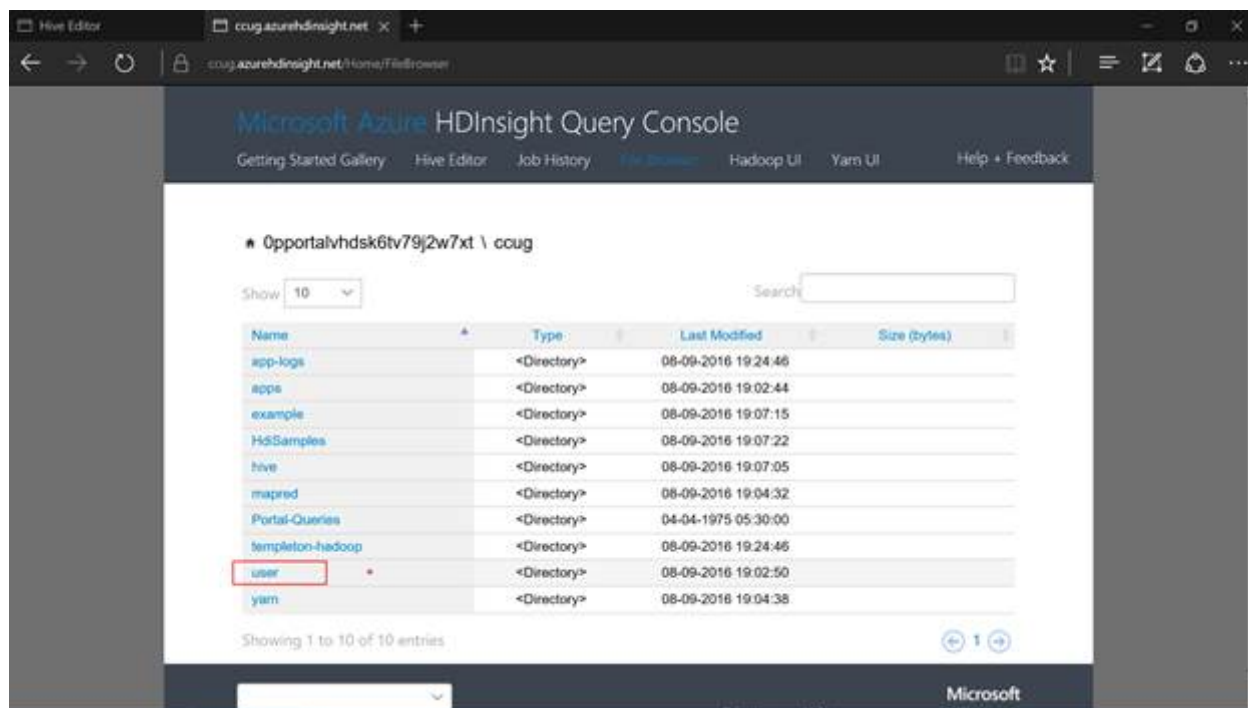
Click on the Storage account name over here, followed by the container name.



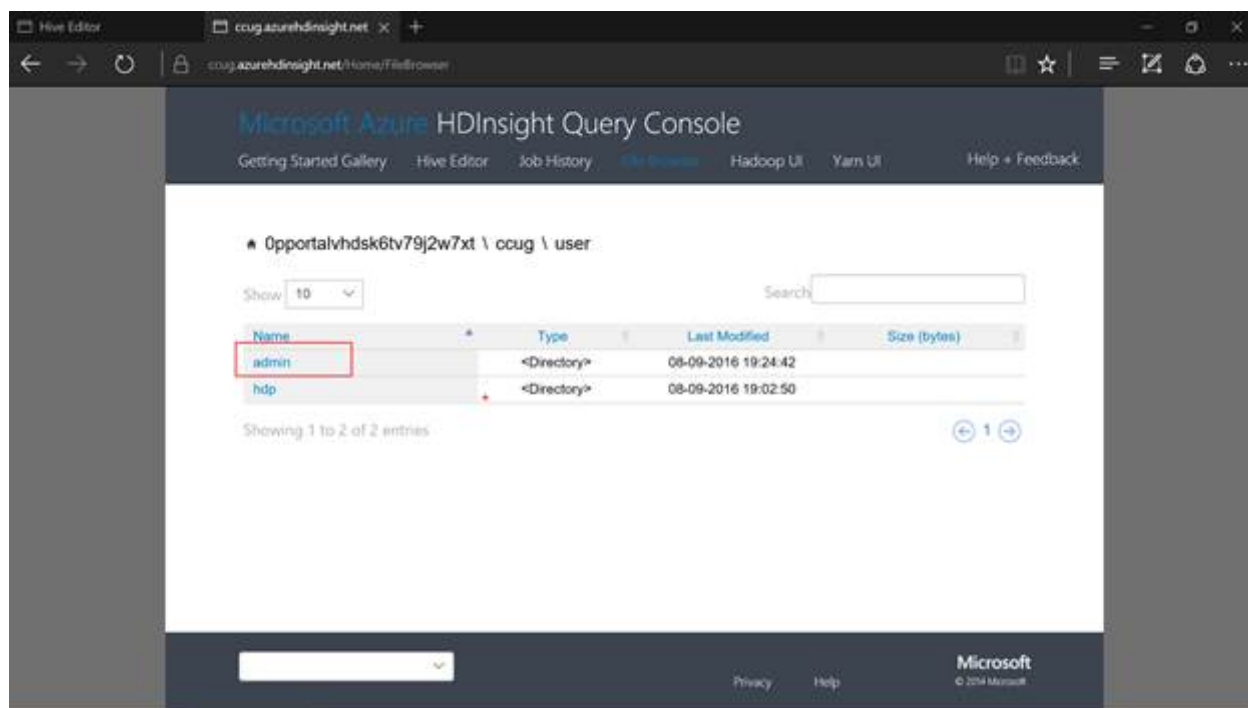
Container name



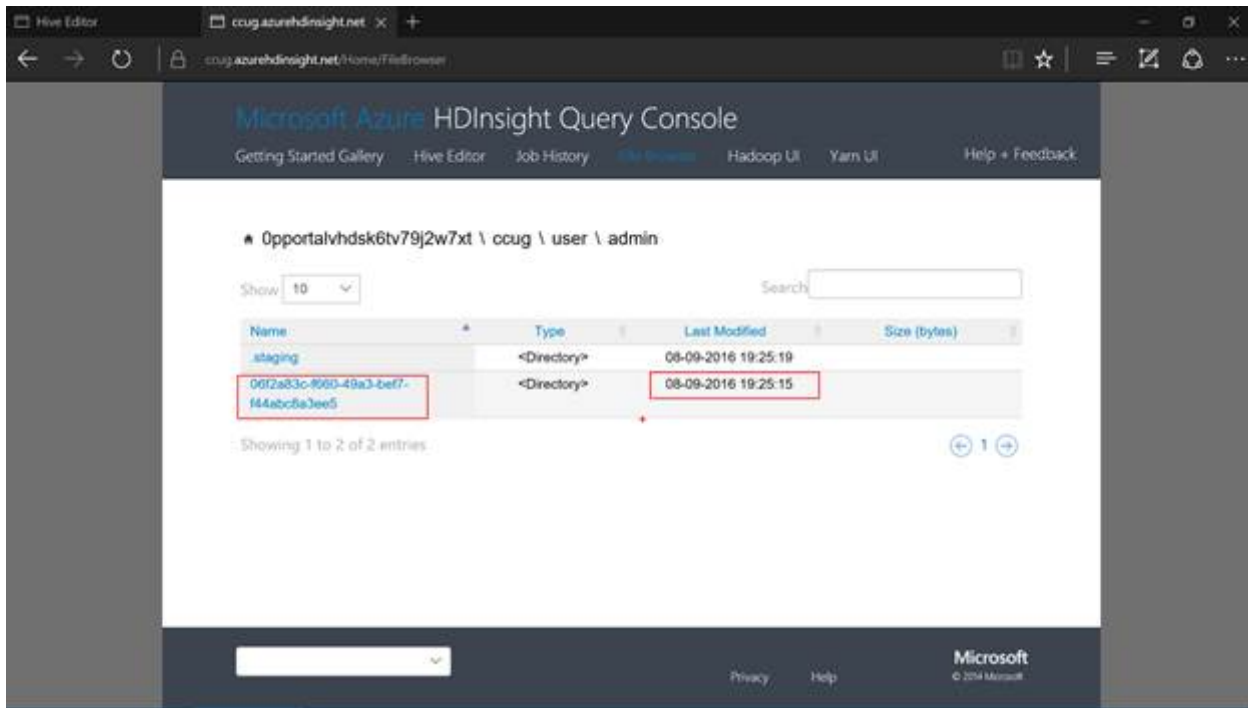
Step 8 - Click on User.



Step 9 - Click on Admin now.



Step 10 - Click on GUID that has the last modified time (the job which started after the job time that you noted earlier). Again, make a note of this GUID. We will be using this in our next module of work.



Follow my next article to learn connecting to Microsoft Business Intelligence tools for Excel.

Thank you for using C# Corner