

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

PRINT

Article



Import Job Output From HDInsight Into Microsoft Excel

By **Abdul Rasheed Feroz Khan** on **Sep 17 2016**

Introduction

This article will help you import your job output from HDInsight into Microsoft Excel. Make sure that you have worked on my previous articles before you step into this one.

- [Kickstart Hadoop in HDInsight on Window](#)
- [Run a Hive Query from Azure Portal](#).

Pre-requisites

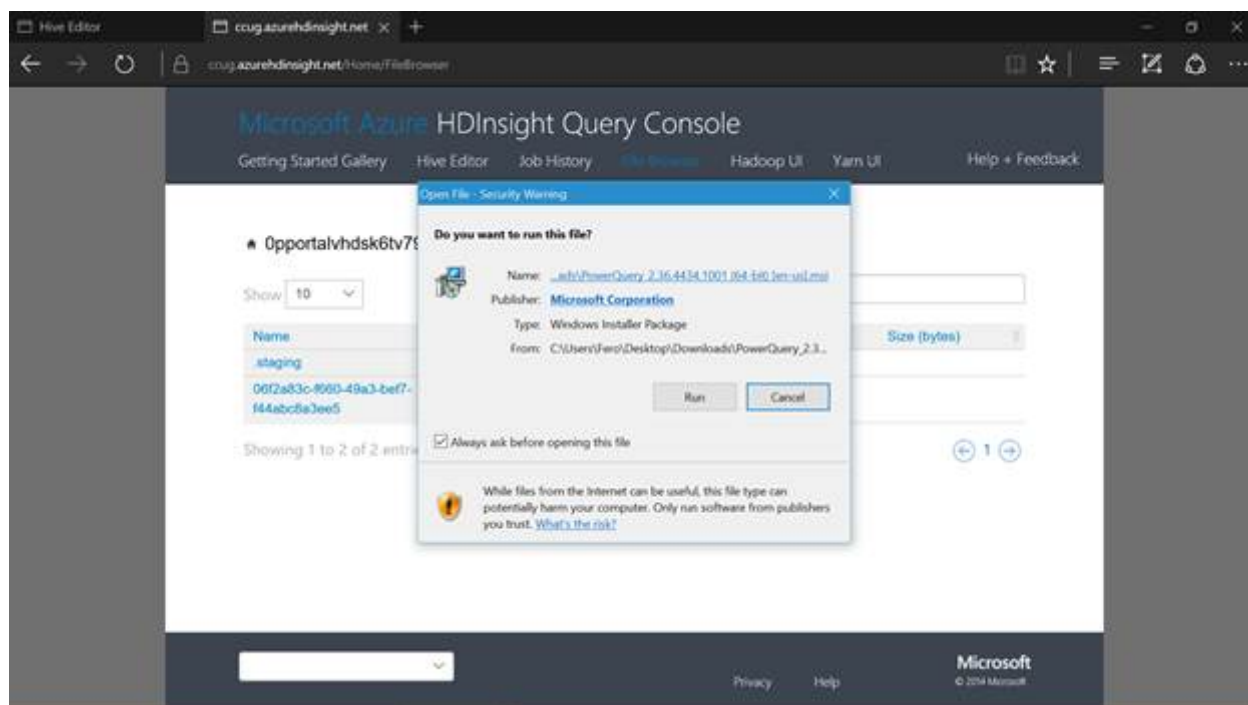
1. Microsoft Excel 2010 or 2013.
2. Microsoft Power Query installed.

Tools needed

Download Microsoft Power Query by clicking [here](#). (Select the appropriate 32 bit or 64 bit version, whichever you need for your Office package).

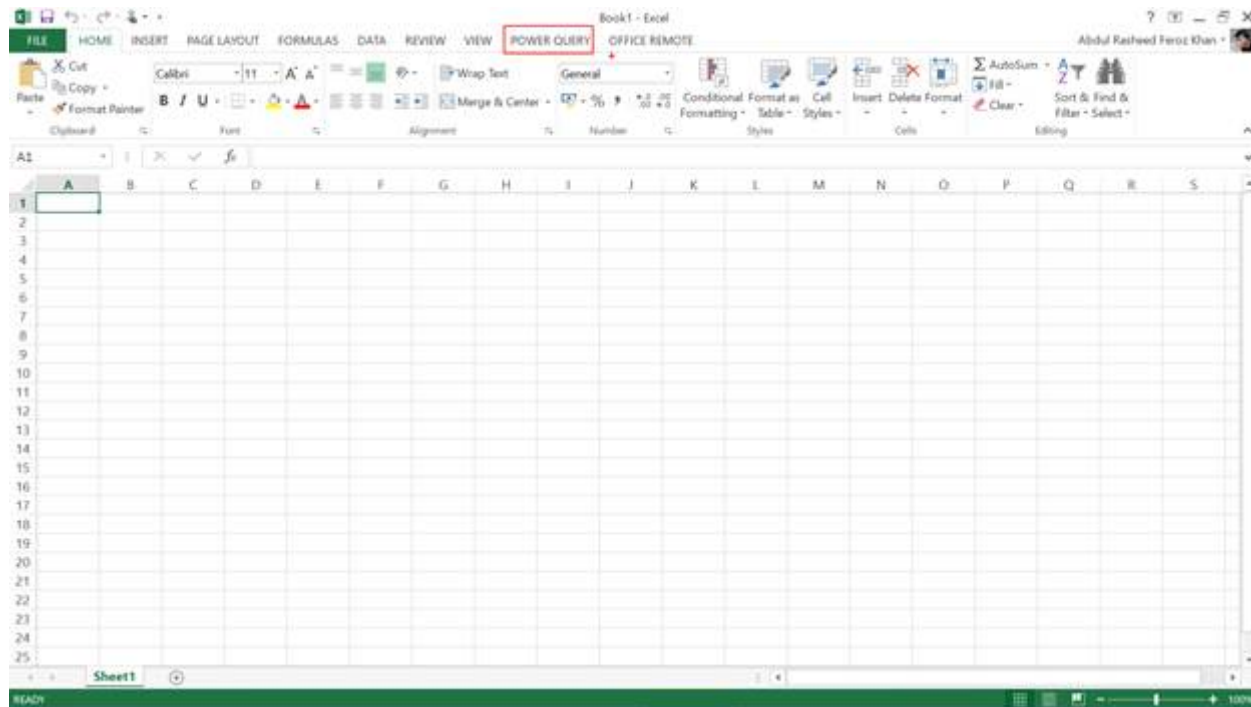
Follow the below steps now -

Step 1 - Install Microsoft Power Query if you don't have it already in your machine.

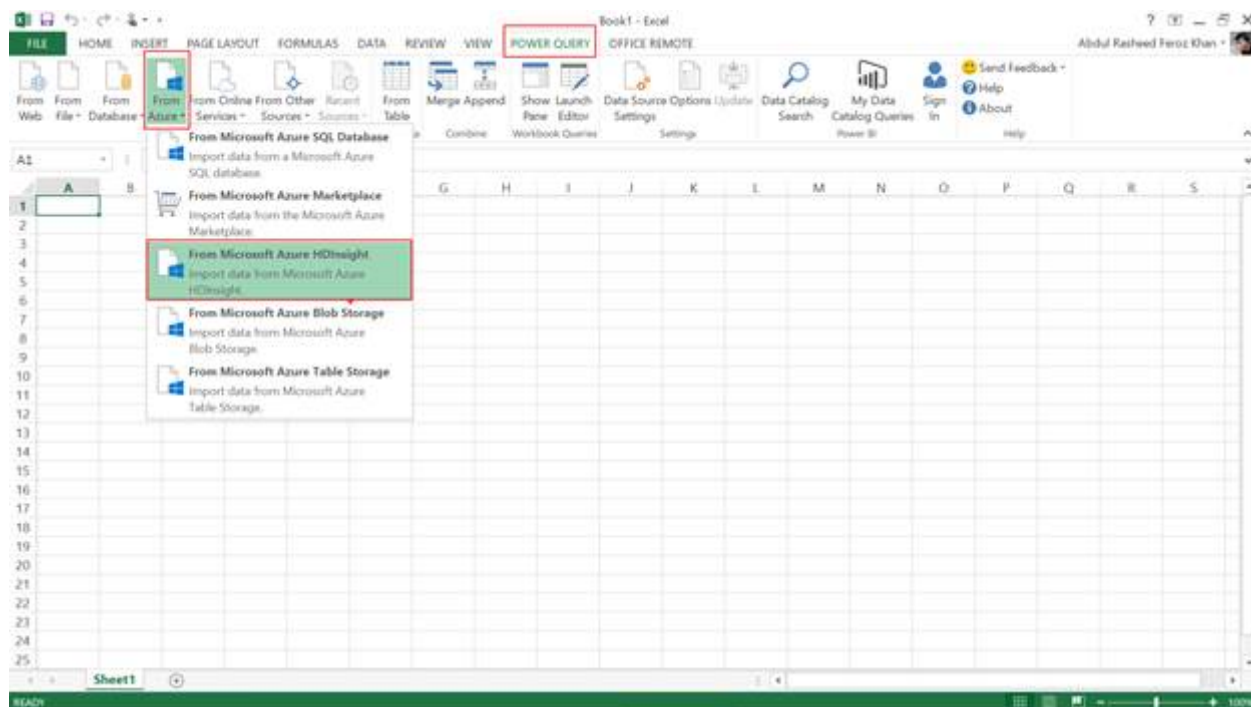


Meanwhile, open your cluster and run the job with your Hive Query. Click [here](#) to get knowledge about it.

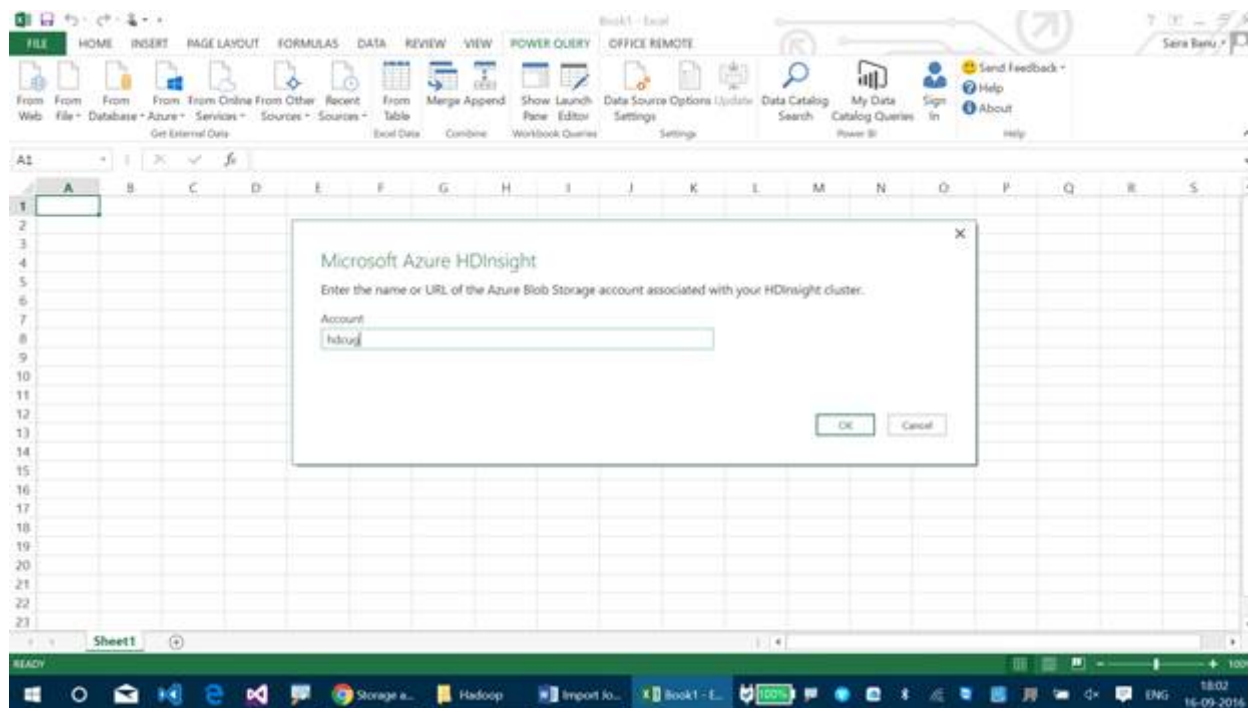
Step 2 - Open a new workbook in Microsoft Excel 2013. Here, you can find the Power Query added.



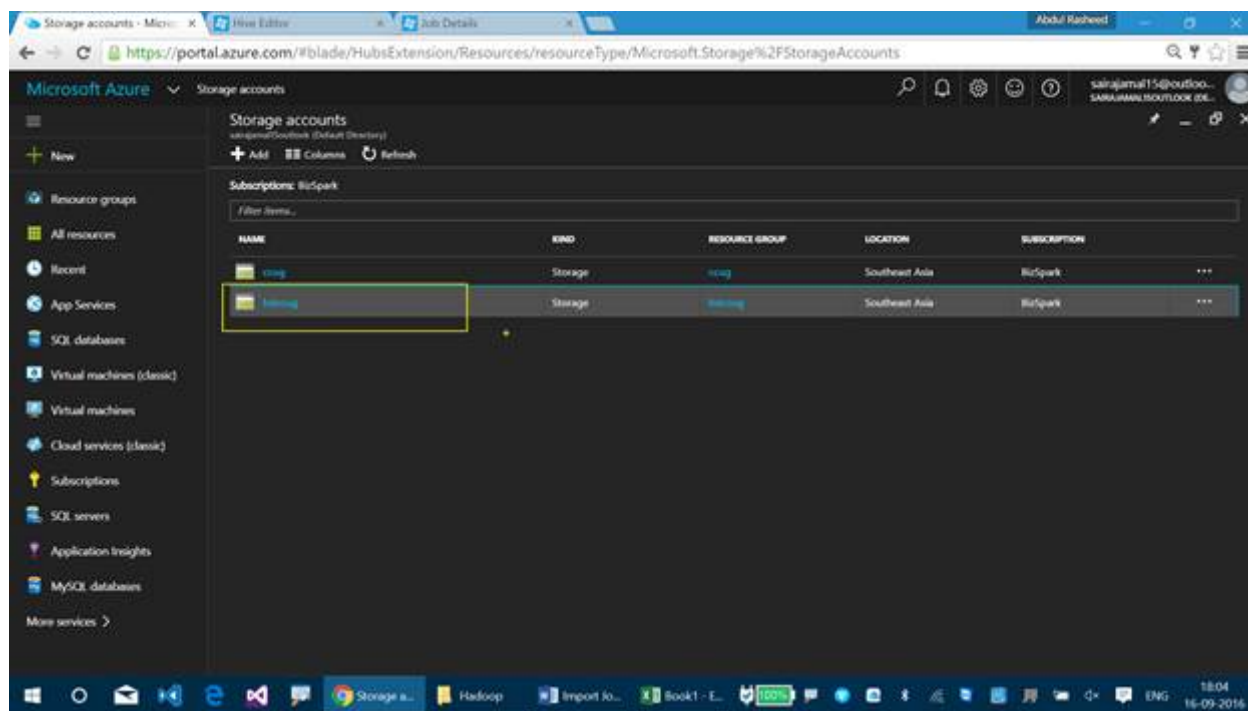
Step 3 - Click on Power Query -> From Azure -> From Microsoft Azure HDInsight,



Step 4 - Enter the Storage account name that we created in our previous article, and click OK.



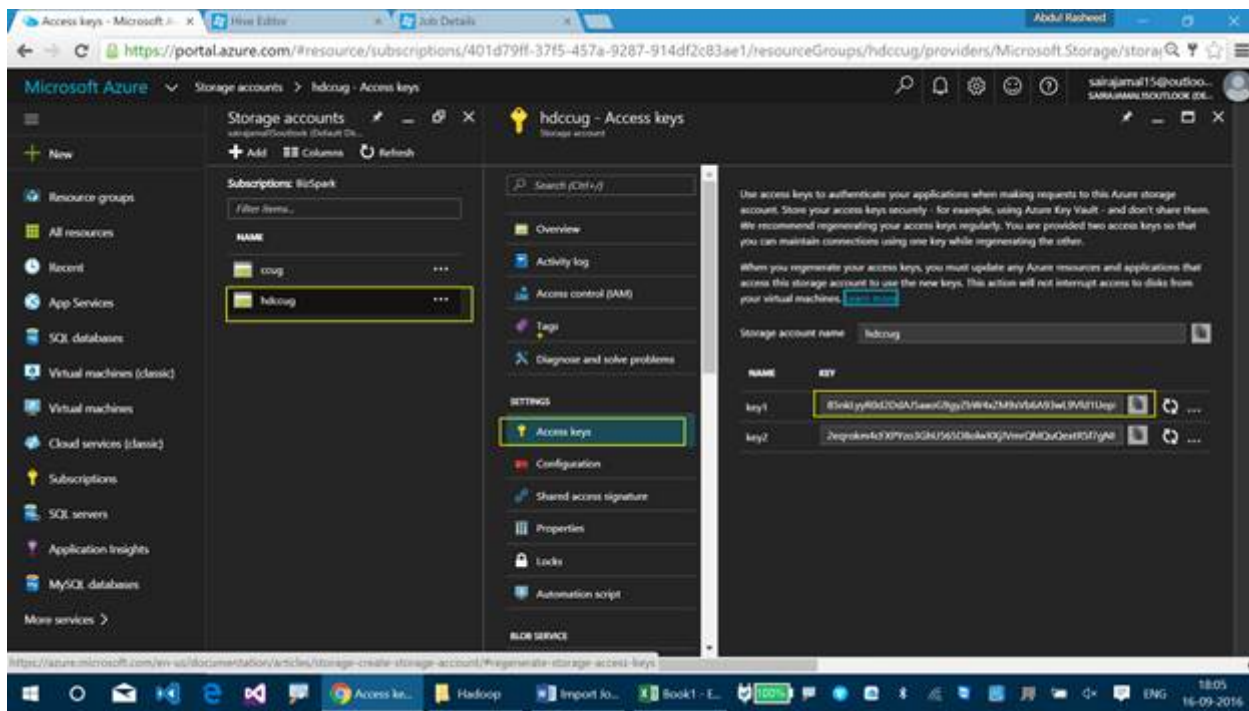
Step 5 - Enter the Account Key to access the Azure HDInsight and click on Connect. To know about the Account Key details, go to your storage account.



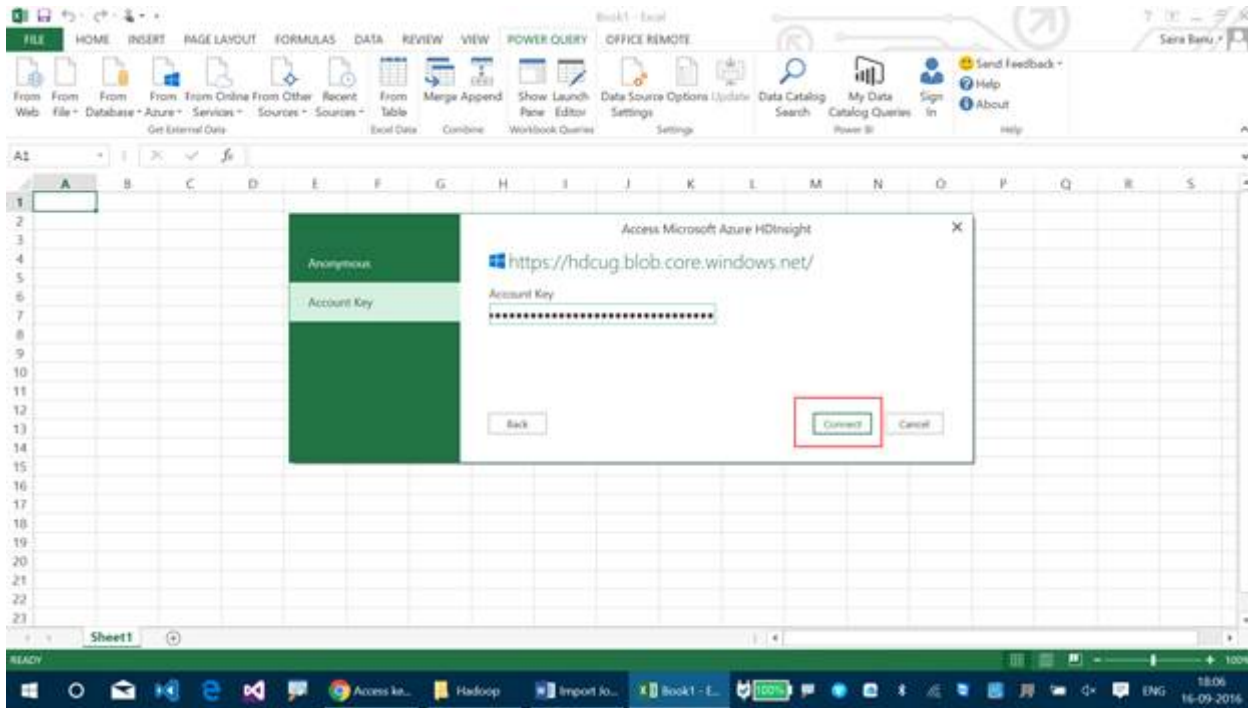
Click on Access Keys.



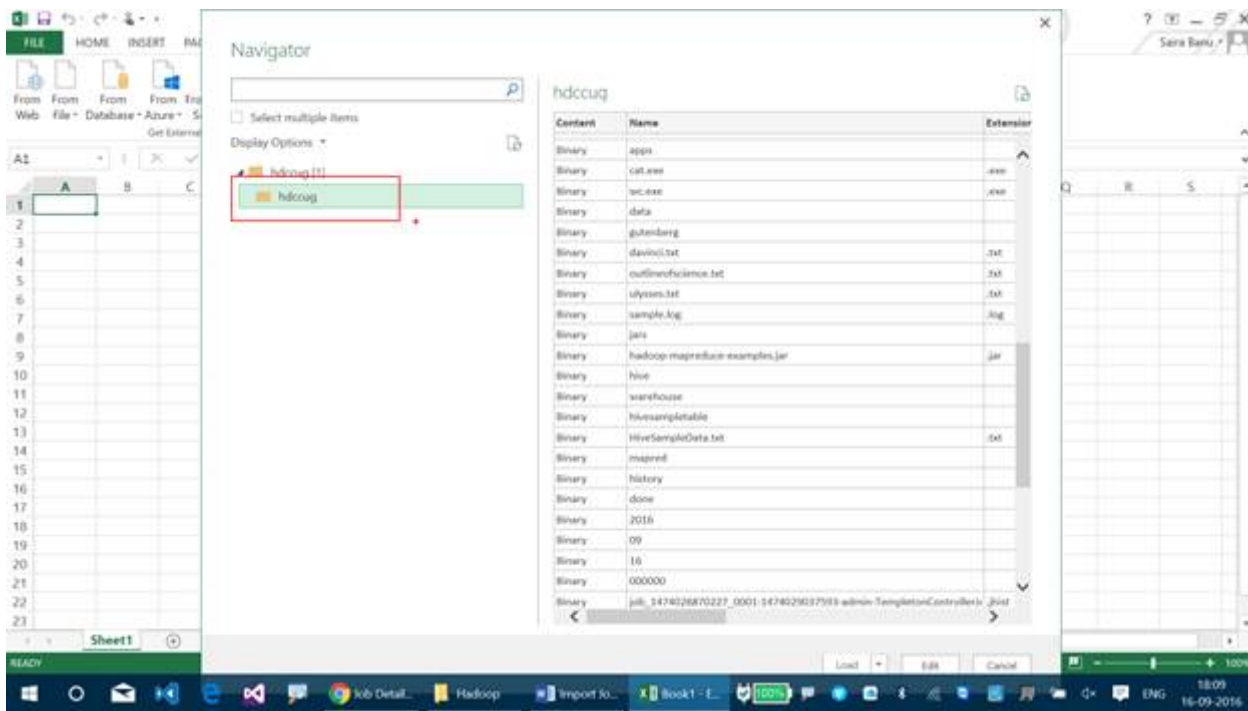
Copy the Key 1 from here and paste it in your Excel workbook.



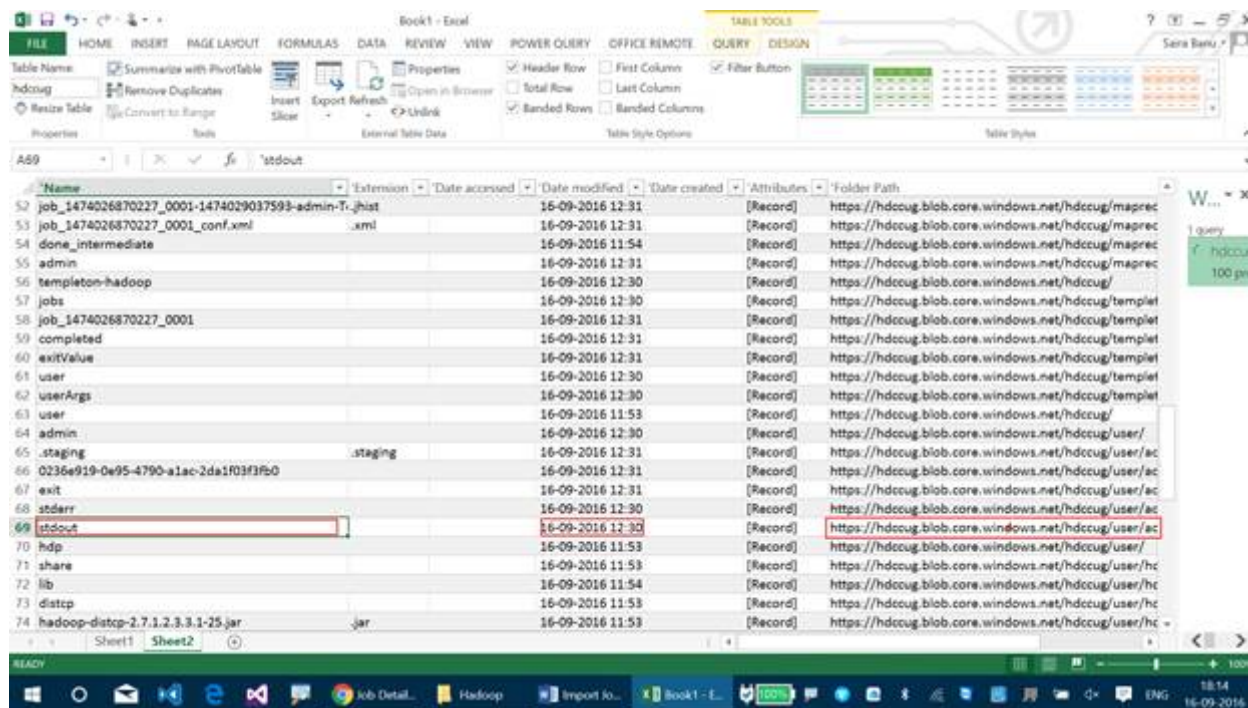
Now, click on Connect.



Step 6 - Now, select the Blob name from the right pane of the Navigator window and click on Load.



Step 7 - Locate stdout from the Name column and verify that the GUID in the corresponding folder path column matches the GUID you copied earlier in previous demos. This match suggests that the output data corresponds to the job you have submitted.



	Name	Extension	Date accessed	Date modified	Date created	Attributes	Folder Path
52	job_1474026870227_0001-1474029037593-admin-tr_hist			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/maprec
53	job_1474026870227_0001_conf.xml	.xml		16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/maprec
54	done_intermediate			16-09-2016 11:54		[Record]	https://hdccug.blob.core.windows.net/hdccug/maprec
55	admin			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/maprec
56	templeton-hadoop			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/
57	jobs			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
58	job_1474026870227_0001			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
59	completed			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
60	exitValue			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
61	user			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
62	userArgs			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/temple
63	user			16-09-2016 11:53		[Record]	https://hdccug.blob.core.windows.net/hdccug/
64	admin			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/
65	.staging	.staging		16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/ac
66	0236e919-0e95-4790-a1ac-2da1f03f3fb0			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/ac
67	exit			16-09-2016 12:31		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/ac
68	stderr			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/ac
69	stdout			16-09-2016 12:30		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/ac
70	hdp			16-09-2016 11:53		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/
71	share			16-09-2016 11:53		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/hc
72	lib			16-09-2016 11:54		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/hc
73	distcp			16-09-2016 11:53		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/hc
74	hadoop-distcp-2.7.1.2.3.3.1-25.jar	.jar		16-09-2016 11:53		[Record]	https://hdccug.blob.core.windows.net/hdccug/user/hc

Once done, click on Close & Load button in the upper-left corner, to import the Hive job output to your Excel sheet.

Thank you for using C# Corner