Visualization with Excel Tools and Windows Azure

# Introduction

Power Query and Power Map are two add-ins that are available as free downloads from Microsoft to enhance the data access and data visualization capabilities of Microsoft Excel 2013. You can use Power Query to retrieve data from a variety of data sources and integrate that data as part of an Excel data model. In this demonstration, you learn how to work with these tools to analyze data in your Windows Azure Storage account.

# Prerequisites

Your computer must be running one of the following operating systems:

* Windows Server 2008 R2 (with Microsoft .NET Framework 4.0)
* Windows 7
* Windows 8

In addition, you must install one of the following versions of Microsoft Office on your computer:

* Office Professional Plus 2013
* Office 365 ProPlus
* Office 365 Midsize
* Office 365 E3, E4, A3, A4, G3, or G4

Although a 32-bit version is available, you should use a 64-bit computer if you will be analyzing large volumes of data. With a 32-bit computer, you need a minimum of 1 GB of RAM, but a 64-bit computer should have 2 GB of RAM.

Your computer must have Internet connectivity to use Power Map.

# Windows Azure Storage

You can use a Windows Azure Storage account to store data that you want to analyze with Excel. To create a new storage account, open the Windows Azure Management Portal, click the New Button, click Data Services, click Storage, and then click Quick Create. Type in a URL for your storage account, select a location or affinity group, and assign the storage to a Windows Azure subscription. Last, click Create Storage Account.

An easy way to work with your storage account is to download and install [Azure Storage Explorer](http://azurestorageexplorer.codeplex.com/). When you launch the application, click the Add Account button, type in your storage account name and storage account key, as shown in Figure 1, and click Add Storage Account. You can optionally select the Use HTTPS check box to use a secure connection for sensitive data. You can locate your storage account key in the Windows Azure Management Portal by accessing the Storage page, selecting the storage account, and then click the Manage Access Keys button at the bottom of the page. Click the icon to the right of the Primary Access Key to copy it to your clipboard, and then paste it into the Add Storage Account dialog box in Azure Storage Explorer.

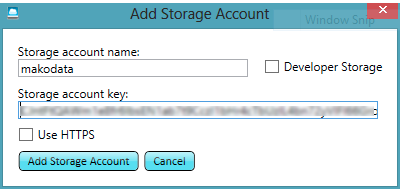


Figure 1: Add Storage Account in Azure Storage Explorer

Next, you need to create a container in your storage account to hold one or more blobs. Each blob is a file of any type. You use block blobs for files up to 200 GB and page blobs for files up to 1TB. For this example, the data is small enough to fit into a block blob. To create the blob container, click the New button in the Container section of the ribbon, type in a name for the container, and click Create Container.

Next you need to access data to put into storage. For this demonstration, download the [Mako\_Real\_Actual\_Sharks](http://www.layerscape.org/File/Download/27eb8120-d53b-4f14-a9e3-d5bd04f67dd9/Mako_Real_Actual_Sharks.csv) file to your computer. Then, in Azure Storage Explorer, select the container you created, click the Upload button, and select the CSV file that you just downloaded. When the upload completes, you can view the newly created blob in the selected container, as shown in Figure 2. Now you’re ready to use this file in Excel.

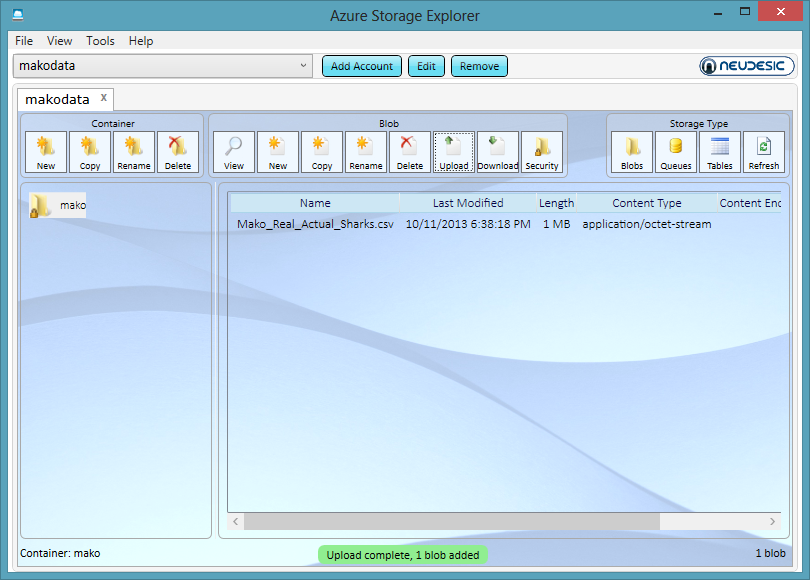


Figure 2: Blob in Windows Azure Storage Container

# Installation of Power Map and Power Query

The first step is to download the add-ins from the Microsoft Download Center. Use the following links:

* Power Map: <http://www.microsoft.com/en-us/download/details.aspx?id=38395>
* Power Query: <http://www.microsoft.com/en-us/download/confirmation.aspx?id=39933>

Locate the downloaded files on your computer’s file system and double-click each file to start the respective installation wizard. Follow the instructions to complete each wizard. Your computer must be connected to the Internet during installation as additional files might need to be downloaded by the installer.

# Power Query

Power Query is a tool that allows you to find, merge, and manipulate data from many different types of data sources, including both structured and semi-structured data. You can even search for data by entering a keyword. Power Query will suggest web sites that contain data related to your query and can scrape the data from a web page. If you already know where data is located, you can import it from a URL, your network, or your computer. Power Query can import many different file types or retrieve data from a database, a SharePoint list, an OData feed, Hadoop, Windows Azure HDInsight, Windows Azure Data Marketplace, Windows Azure Blob Storage, Active Directory, or Facebook. After importing the data, you can manipulate the data in a variety of ways, applying transformations, filtering, and splitting columns, to name just a few operations. If you import multiple data sets that share data in common, you can combine these datasets into a common table for easy analysis. Think of Power Query as fast and easy data integration and cleansing tool that you can use within Excel.

## Import Data

In this demonstration, you use Power Query to retrieve data from your Windows Azure Storage account and prepare the data for visualization in Power Map. Start by creating a new workbook in Excel 2013. On the Power Query tab of the ribbon, click From Other Sources and select From Windows Azure Blob Storage. Type the name of your storage account and paste in the access key (which you obtain as described earlier in the Windows Azure Storage section of this document).

When you click Save, the Query Editor window displays the containers in your storage account. When you click the container, you can see the blobs in the selected container, as shown in Figure 3.

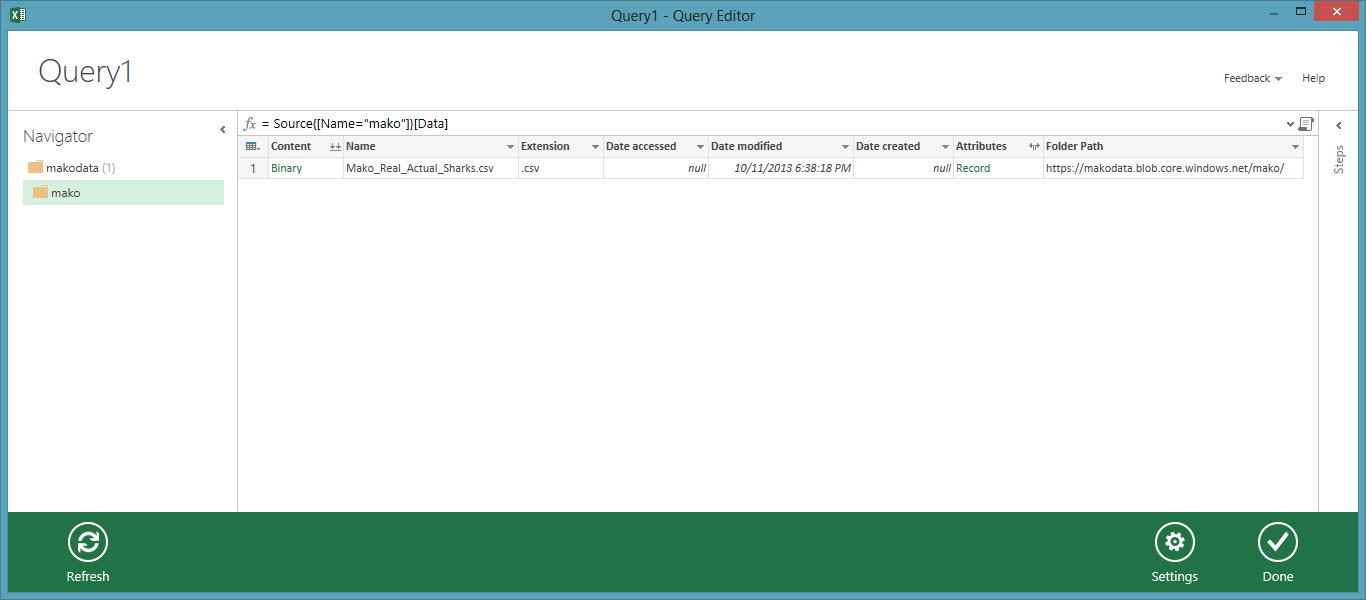


Figure 3: Blob in Query Editor Window

When you click the Binary link in the Content column, Power Query imports the data into the Query Editor window where you can view the data. In this case, the data imports as multiple columns, as shown in Figure 4. If necessary, you can transform the data, reshaping it as needed to support your analysis. In this example, no transformations are required before using the data, so click Done.

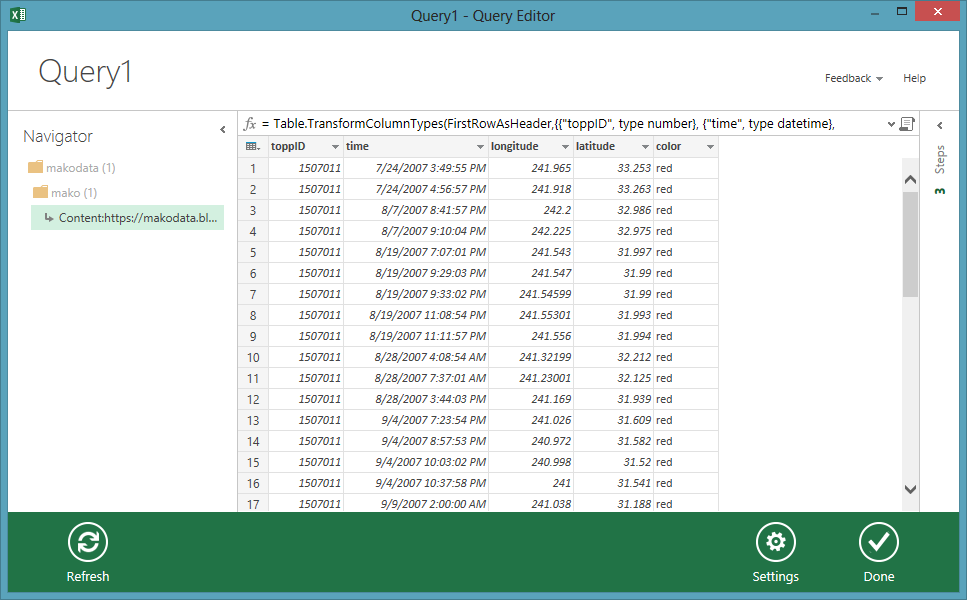


Figure 4: Query Editor Window with Imported Data

To learn more about Power Query, refer to the documentation at <http://office.microsoft.com/en-us/excel-help/microsoft-power-query-for-excel-help-HA104003813.aspx>.

# Power Map

Power Map provides a new perspective for your data by plotting geocoded data onto a three-dimensional view of the earth and optionally showing changes to that data over time. To use Power Map, you import raw data into a Microsoft Excel 2013 workbook, add the data to an Excel data model, and enhance the data in the data model if necessary. Once the data model is prepared, you can then insert a Power Map into the workbook and add items from the data model for viewing on the map. You start by identifying the geographical elements in your data model, select the numeric data to display as well as an aggregation function like sum or count, and then you specify a visualization method such as a column chart, bubble map, or heat map. If your data model includes a date data type, you can identify the date as a time element and then use the map’s play axis to view location changes to the numeric data over time.

To use the data imported by using Power Query, open the Insert tab of the ribbon, click the Map button, and then click Launch Power Map. A new window opens after a few seconds. Power Map attempts to detect geocoded data in your data model, as shown in Figure 5, and performs a simple mapping. If necessary, you can change the check box selections for geographical fields in the Task Panel on the right side of the screen. In addition, if Power Map fails to correctly identify the type of data that the selected fields represent, you can assign the field to an applicable type from the drop-down list to the right of the field in the Geography And Map Level section of the Task Panel.

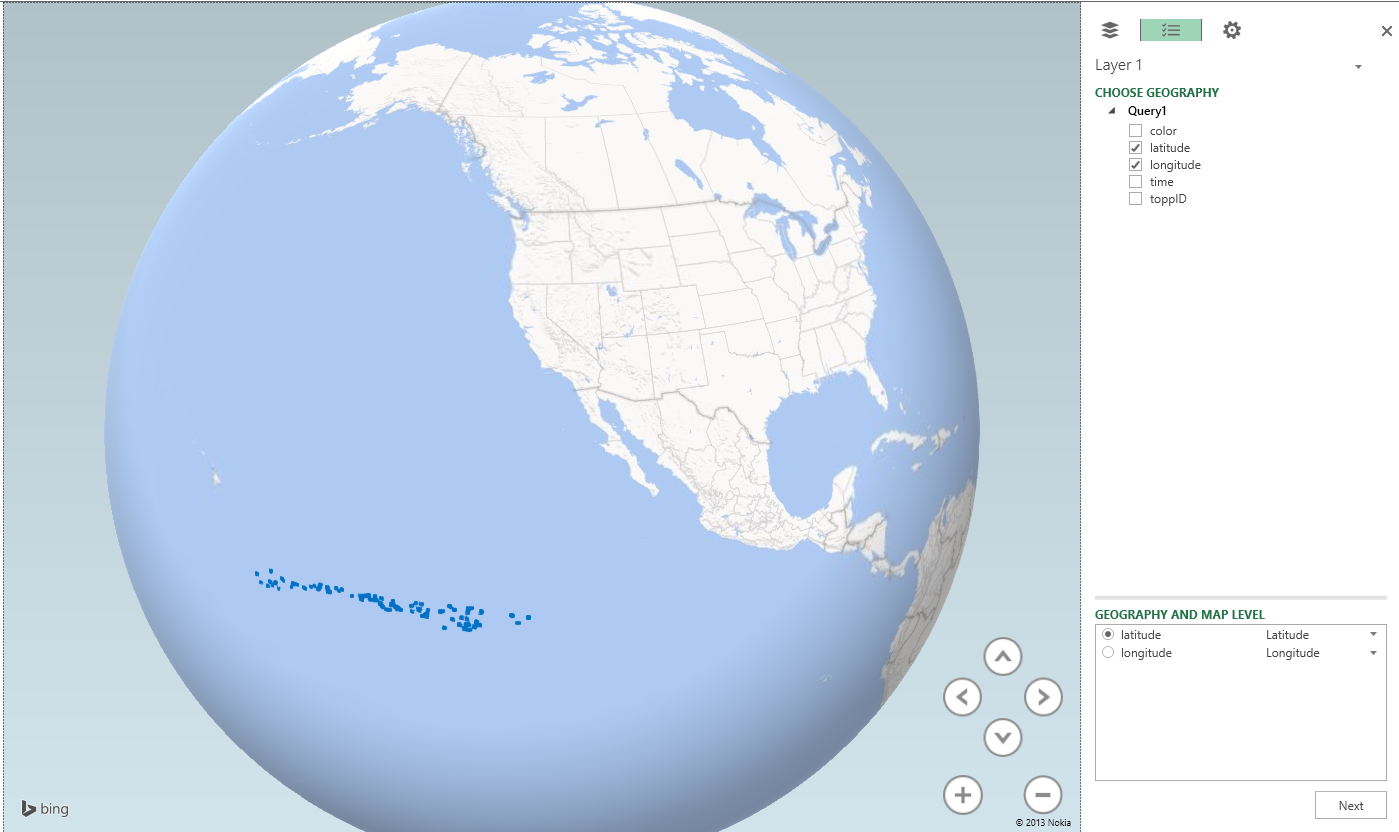


Figure 4: Geography Selection

Figure 5: Geography and Map Level Selection

Click the Next button in the bottom right corner to continue defining the map properties. You now specify the type of chart to display by choosing one of the following options from the Type drop-down list: Column, Bubble, HeatMap, or Region. For this demonstration, keep the default selection of Column.

Your next step is to specify the field containing the numeric value to aggregate and visualize on the map as a column. In the case of the shark data, there is no numeric value available, but you can select the toppID field, which represents an individual shark, to add it to the Height section of the Task panel. The default aggregation is Sum, but you can click the arrow icon to the right of toppID and change the aggregation to Count (Distinct).

The map adjusts to show columns representing the count of sharks in each location, but the imported data contains another variable to help us better visualize the shark population. Select the color check box and notice that it now displays as a category. The map now displays columns with multiple colors and a legend lists each color next to the color to which it corresponds on the map, as shown in Figure 6. You can close the legend by hovering the mouse over the legend, and then clicking the X that displays in the top right corner.

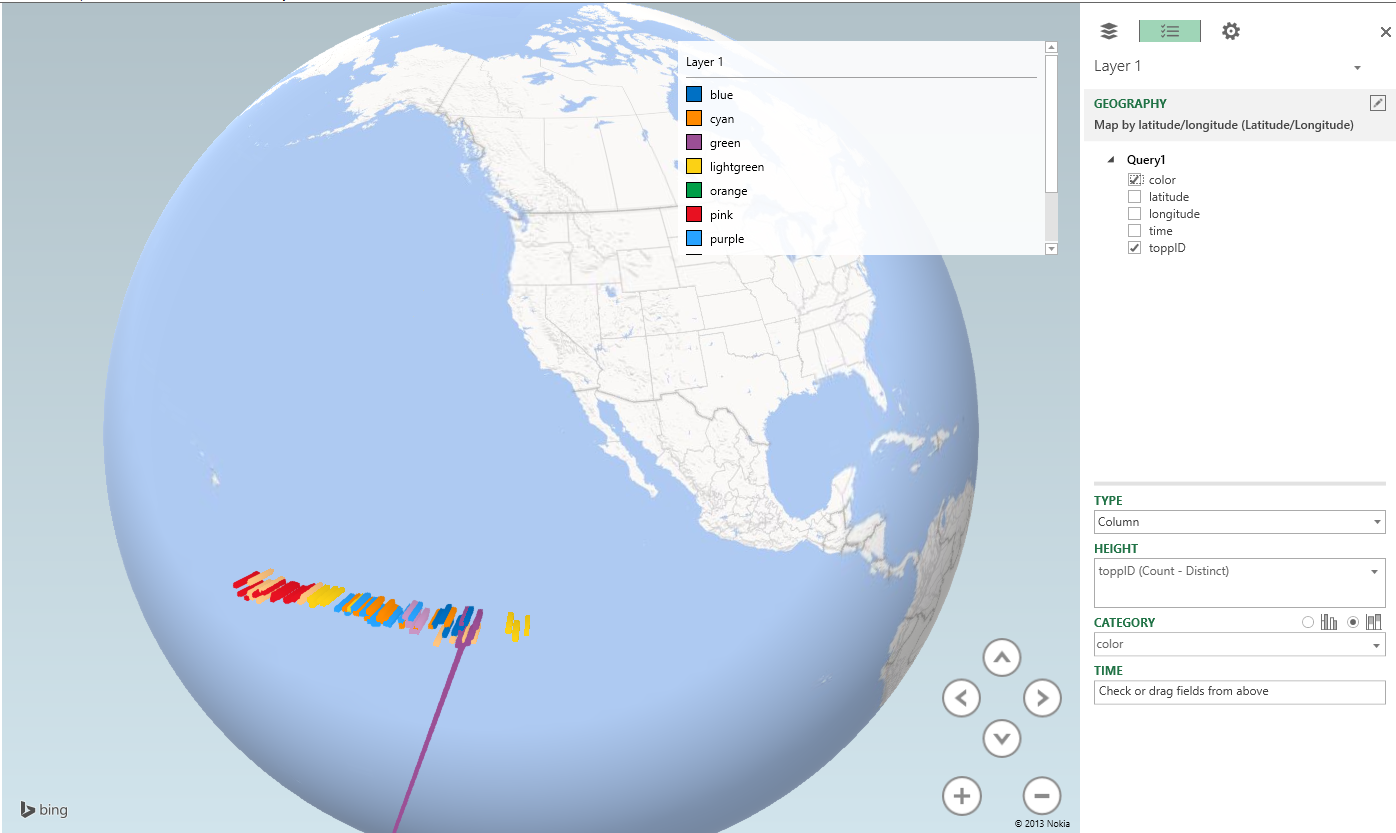


Figure 6: Layer Definition

Figure 6: Layer Definition

**Note:** In this example, the colors in the legend are not consistent with the labels because Power Map understands the color values from the data set as text labels with no meaning rather than specific colors.

The settings you configure here are associated with Layer 1. Click the Add Layer button in the toolbar to add a new layer and begin the process again by selecting geographical fields first. You can click the icon in the top left corner of the Task Panel to control which layers are visible. For example, you could obtain data related to sea temperatures by date and include that in the model, allocating the shark movement data to Layer 1 and the temperature data to Layer 2.

You can also visualize changes to the data over time by adding a play axis to the map. To do this, drag Date from the model’s field list to the Time box in the bottom right corner of the window. A play axis displays below the map. When you click the play button on the left side of the play axis, the map renders the data points applicable to the current date playing. In addition, you can use the arrow buttons in the bottom right corner of the map to rotate the angle of the view or use the plus or minus buttons to change the zoom factor.

Try changing the Type to a different selection to observe the results.