Challenge 7: Visualizing in Power BI Desktop

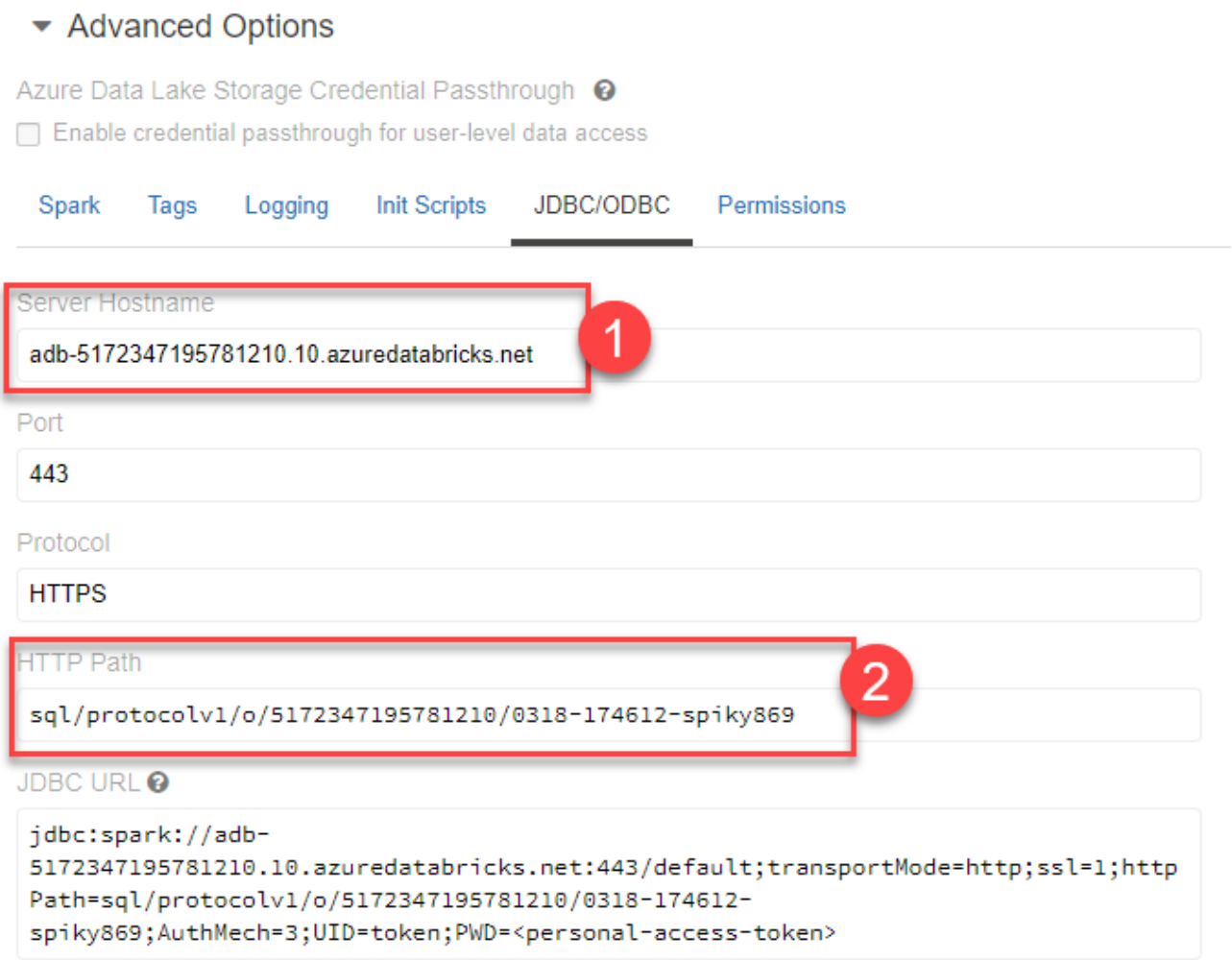
Duration: 20 minutes

In this exercise, you will create visualizations in Power BI Desktop.

Task 1: Obtain the JDBC connection string to your Azure Databricks cluster

Before you begin, you must first obtain the JDBC connection string to your Azure Databricks cluster.

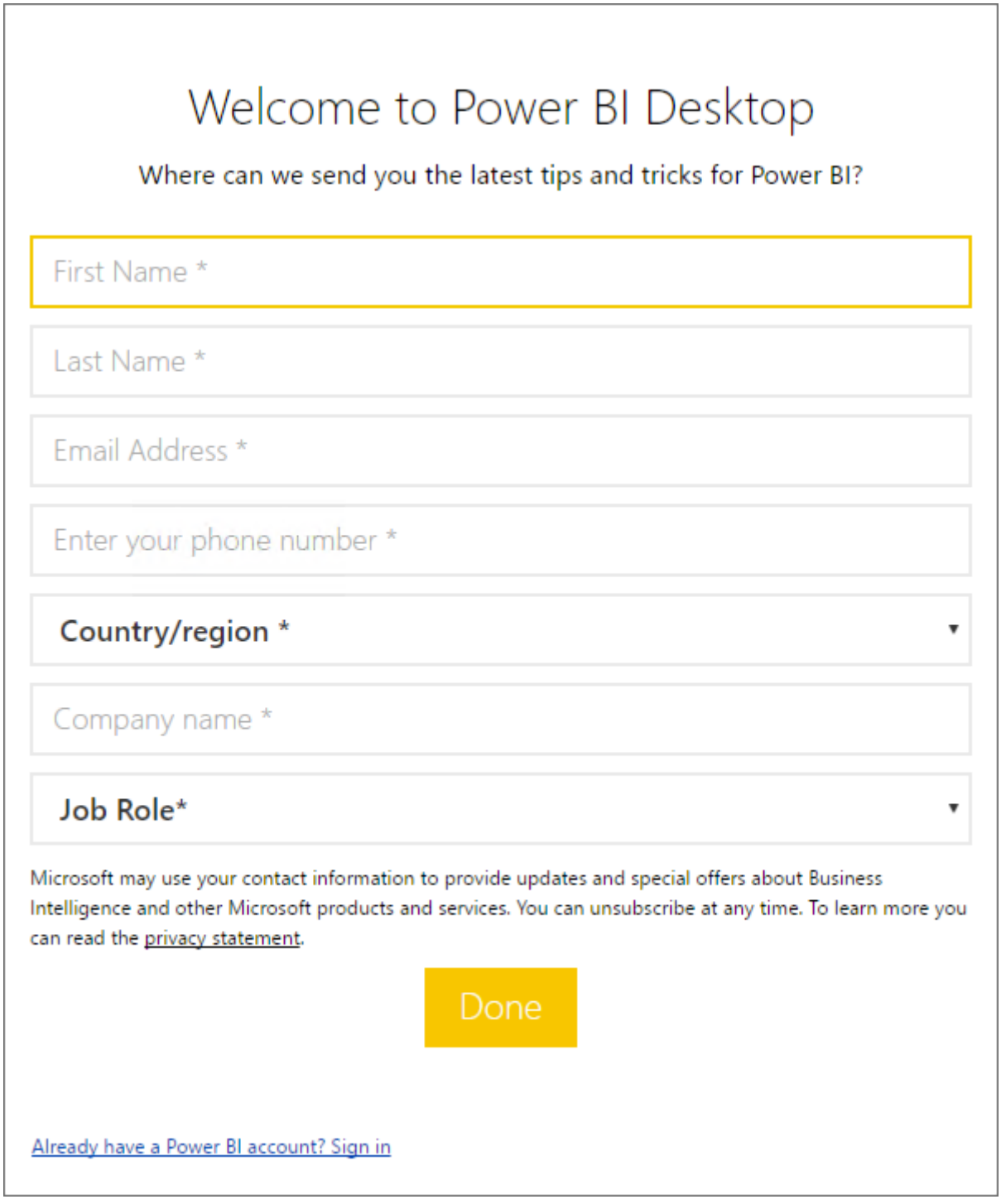
1. In Azure Databricks, go to **Compute** and select your cluster.
2. On the cluster edit page, in the **Configuration** tab, scroll down to the bottom of the page, expand **Advanced Options**, then select the **JDBC/ODBC** tab.
3. On the **JDBC/ODBC** tab, copy and save the **Server Hostname (1)** and **HTTP Path (2)** to be used during the next task. You can use a text editor such as Notepad to keep the values for later use.



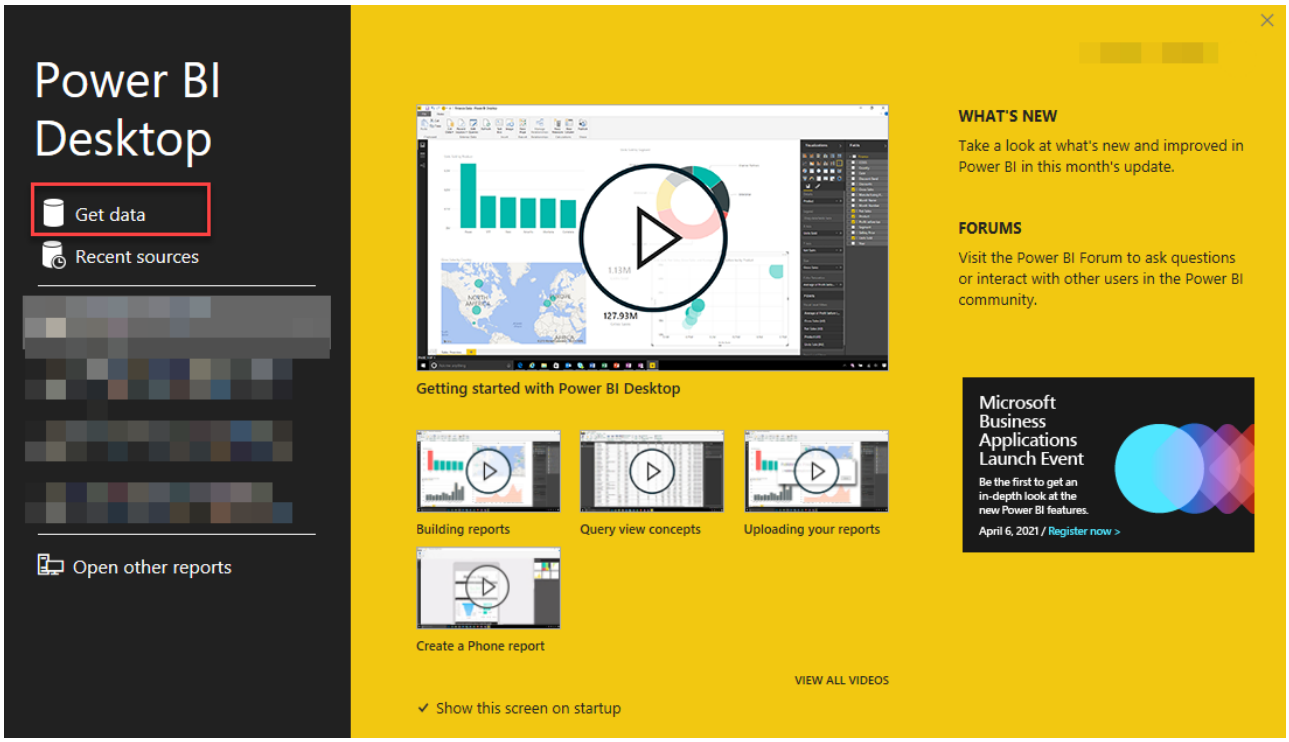
Task 2: Connect to Azure Databricks using Power BI Desktop

1. If you did not already do so during the before the hands-on lab setup, download Power BI Desktop from https://powerbi.microsoft.com/en-us/desktop/.

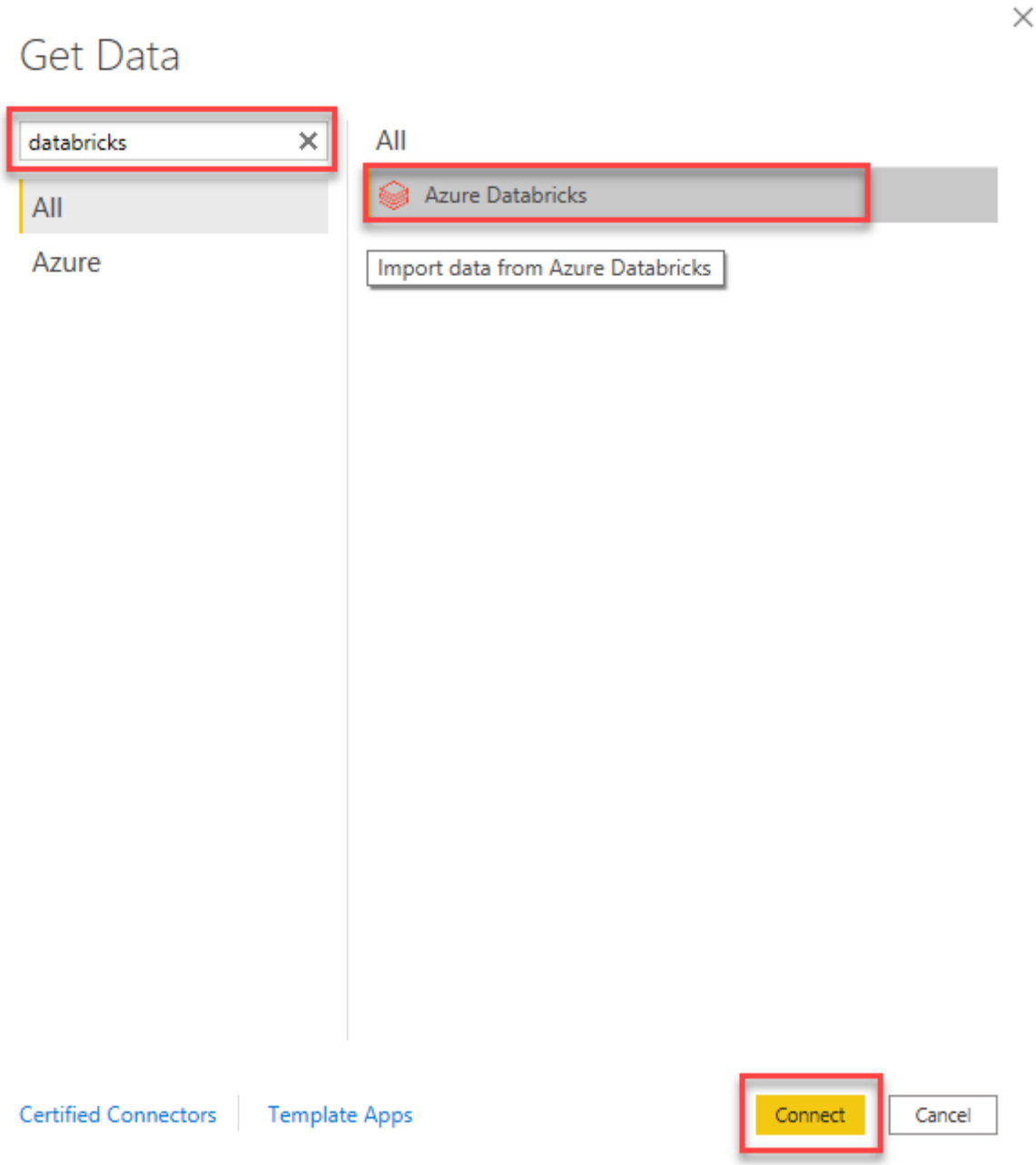
2. When Power BI Desktop starts, you will need to enter your personal information or Sign in if you already have an account.



3. Select Get data on the screen that is displayed next.



4. Select **Azure Databricks** from the list of available data sources. You may enter `databricks` into the search field to find it faster.

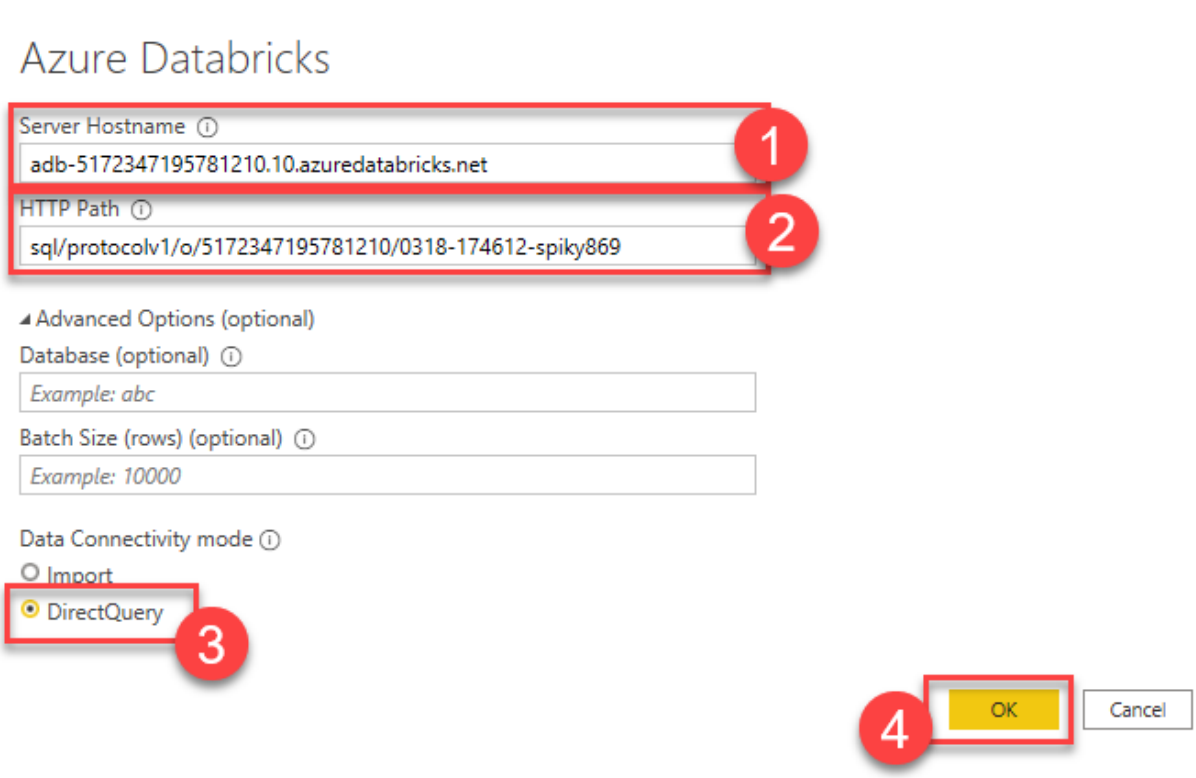


5. Select **Connect**.

6. On the next screen, you will be prompted for your Azure Databricks cluster information.

7. On the Azure Databricks connection information dialog, enter the following:

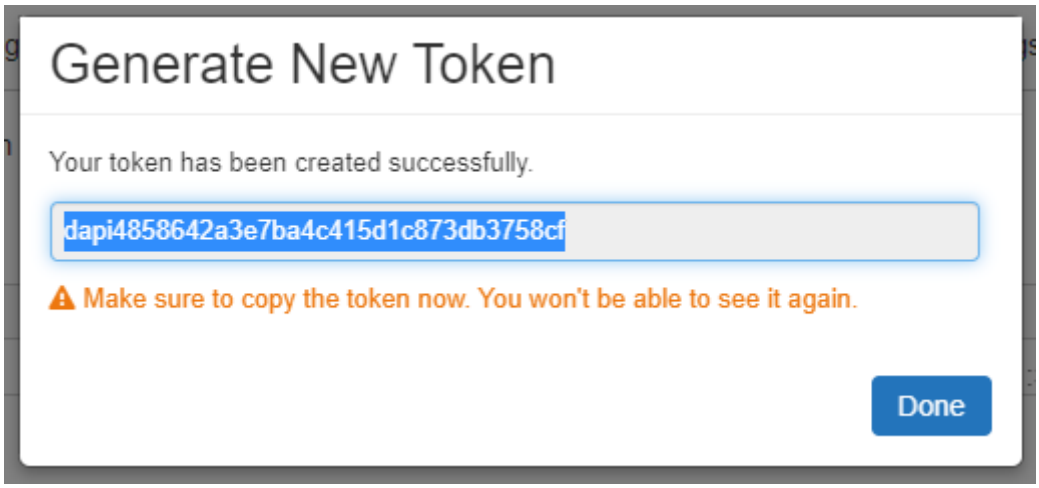
    - **Server Hostname (1)**: Paste the JDBC **Server Hostname (1)** value you copied in the previous task.
    - **HTTP Path (2)**: Paste the JDBC **HTTP Path** value you copied in the previous task.
    - **Data Connectivity mode**: Select **DirectQuery (3)** for the Data Connectivity mode. This option will offload query tasks to the Azure Databricks Spark cluster, providing near-real-time querying.

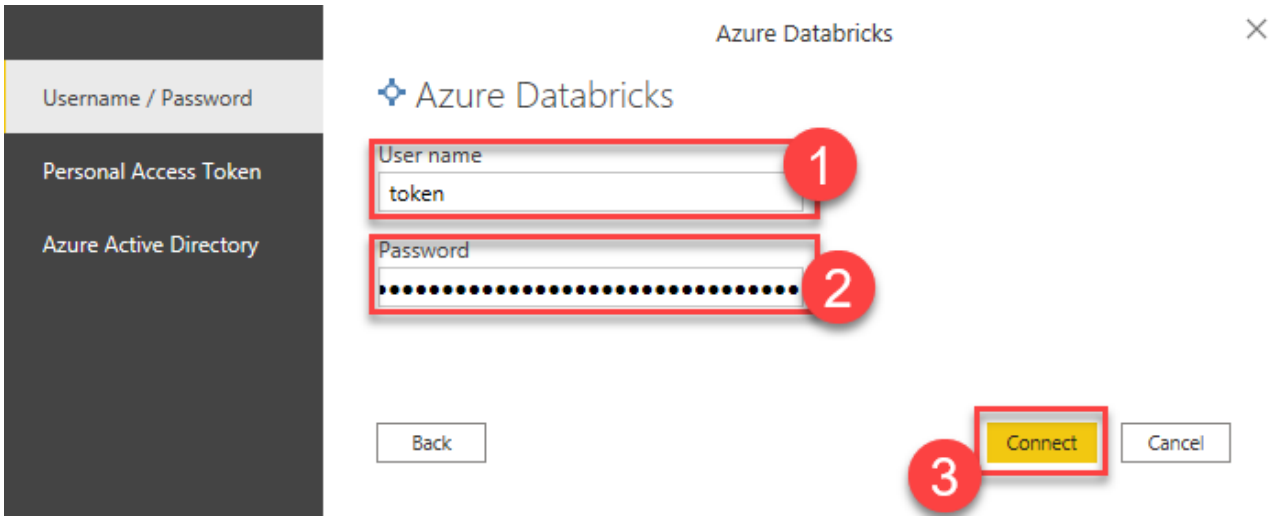


8. Select **OK (4)**.

9. Select the **Username/Password** option from the credentials menu and then enter your credentials on the next screen as follows:

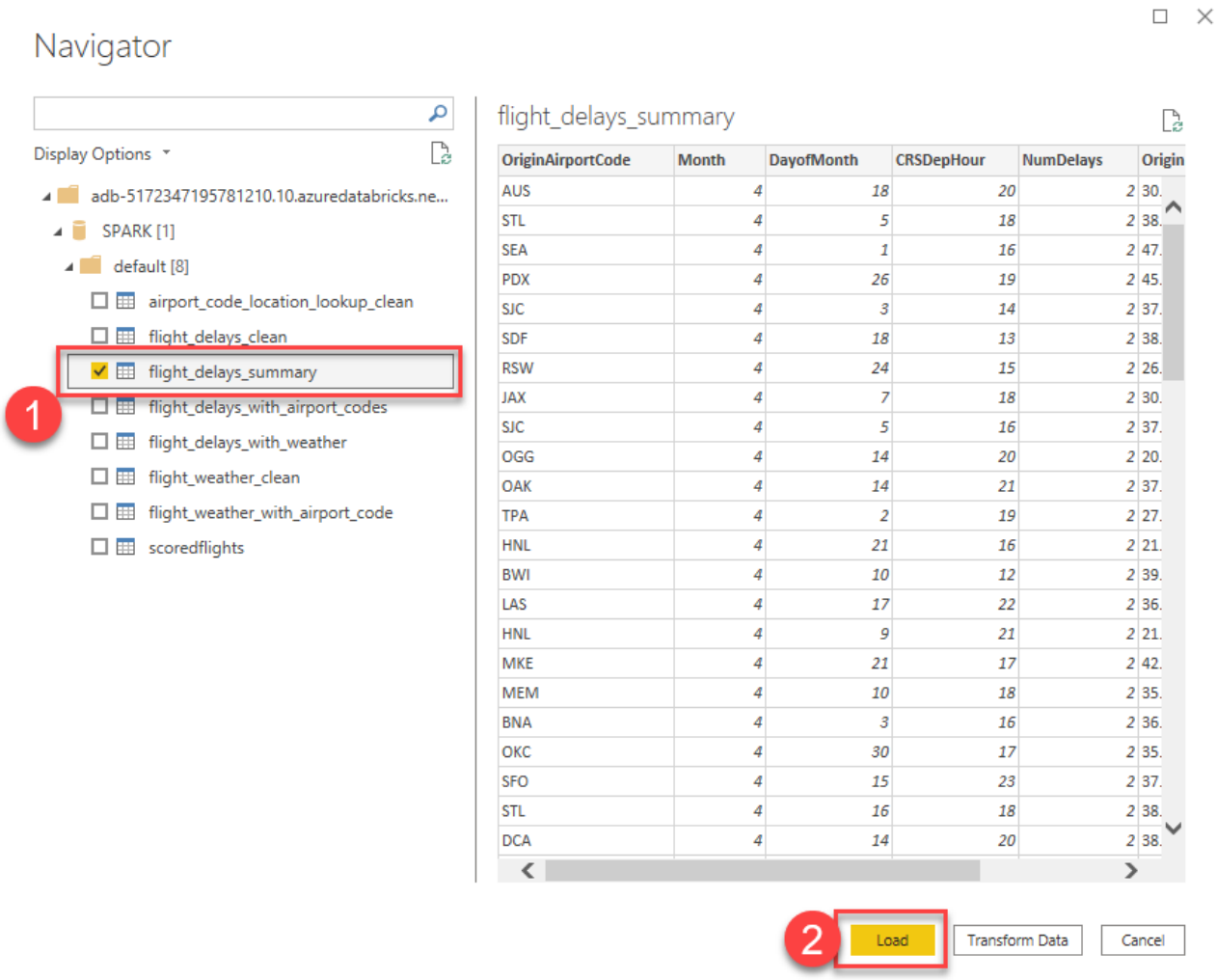
    - **User name (1)**: `token`
    - **Password (2)**: Remember that ADF Access token we generated for the Azure Data Factory notebook activity? Paste the same value here for the password.





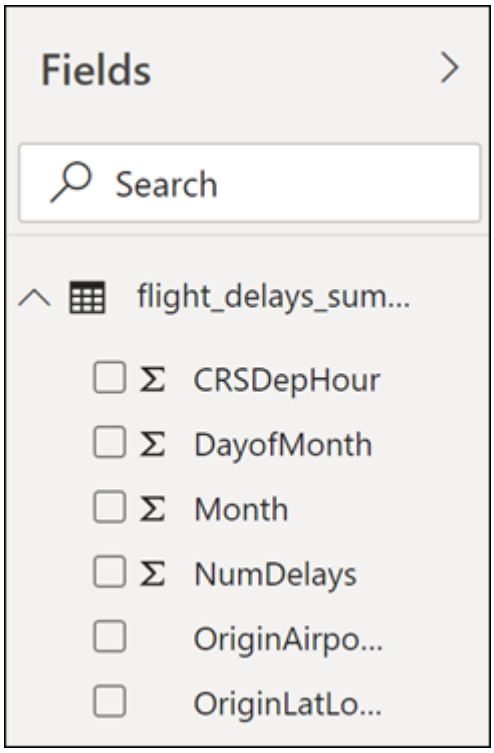
10. Select **Connect (3)**.

11. In the Navigator dialog, check the box next to **flight_delays_summary (1)**, and select **Load (2)**.

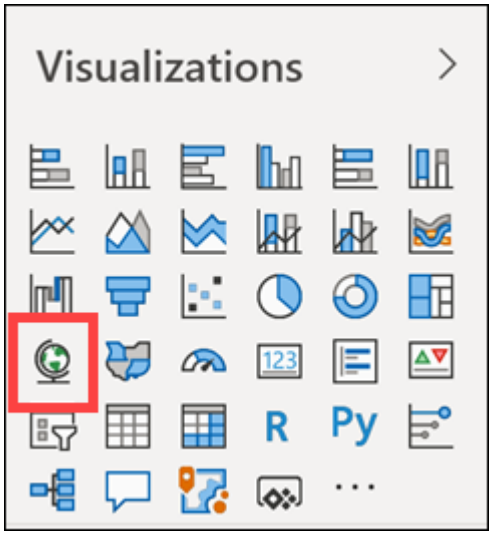


Task 3: Create Power BI report

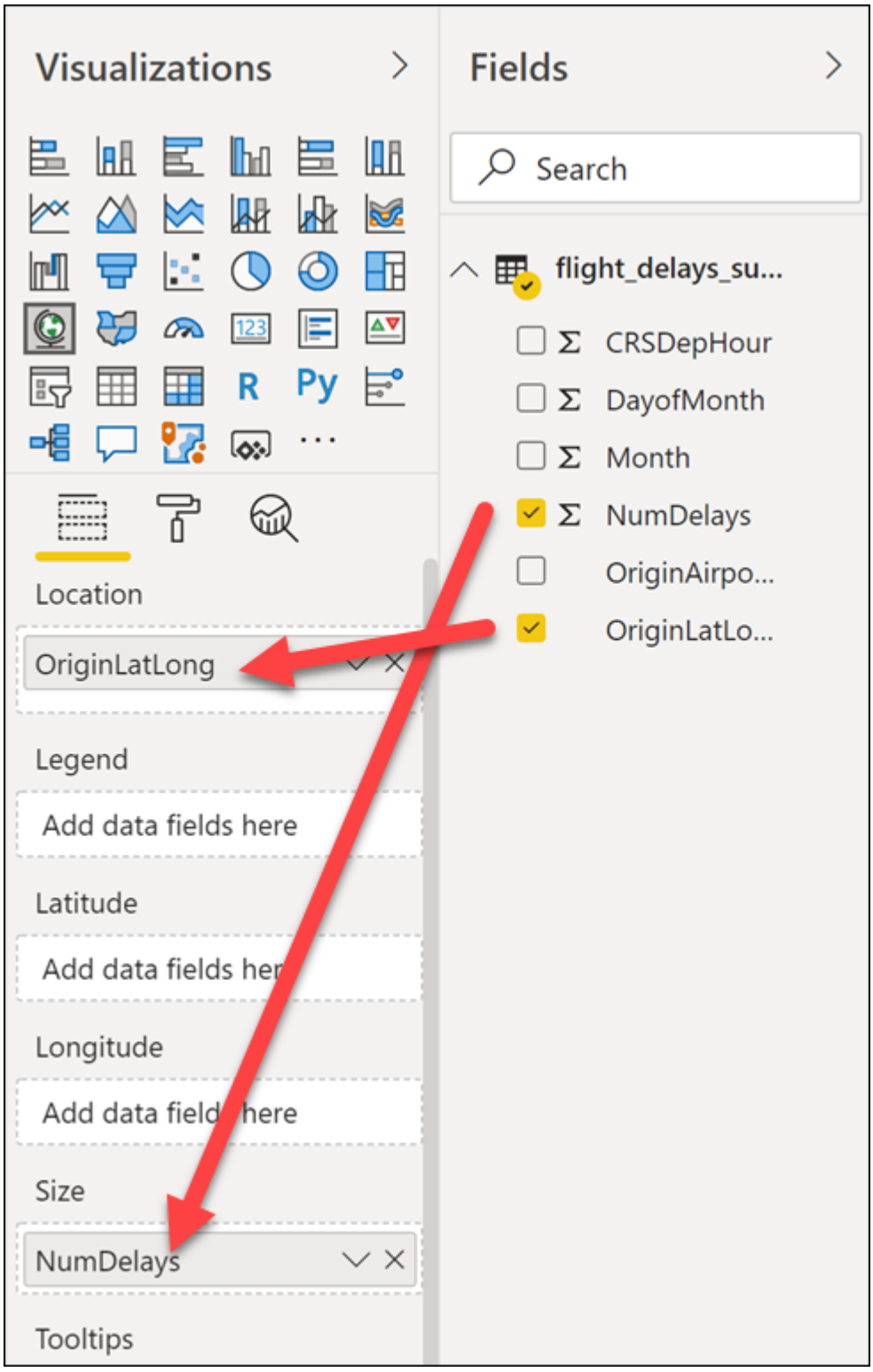
1. Once the data finishes loading, you will see the fields appear on the far side of the Power BI Desktop client window.



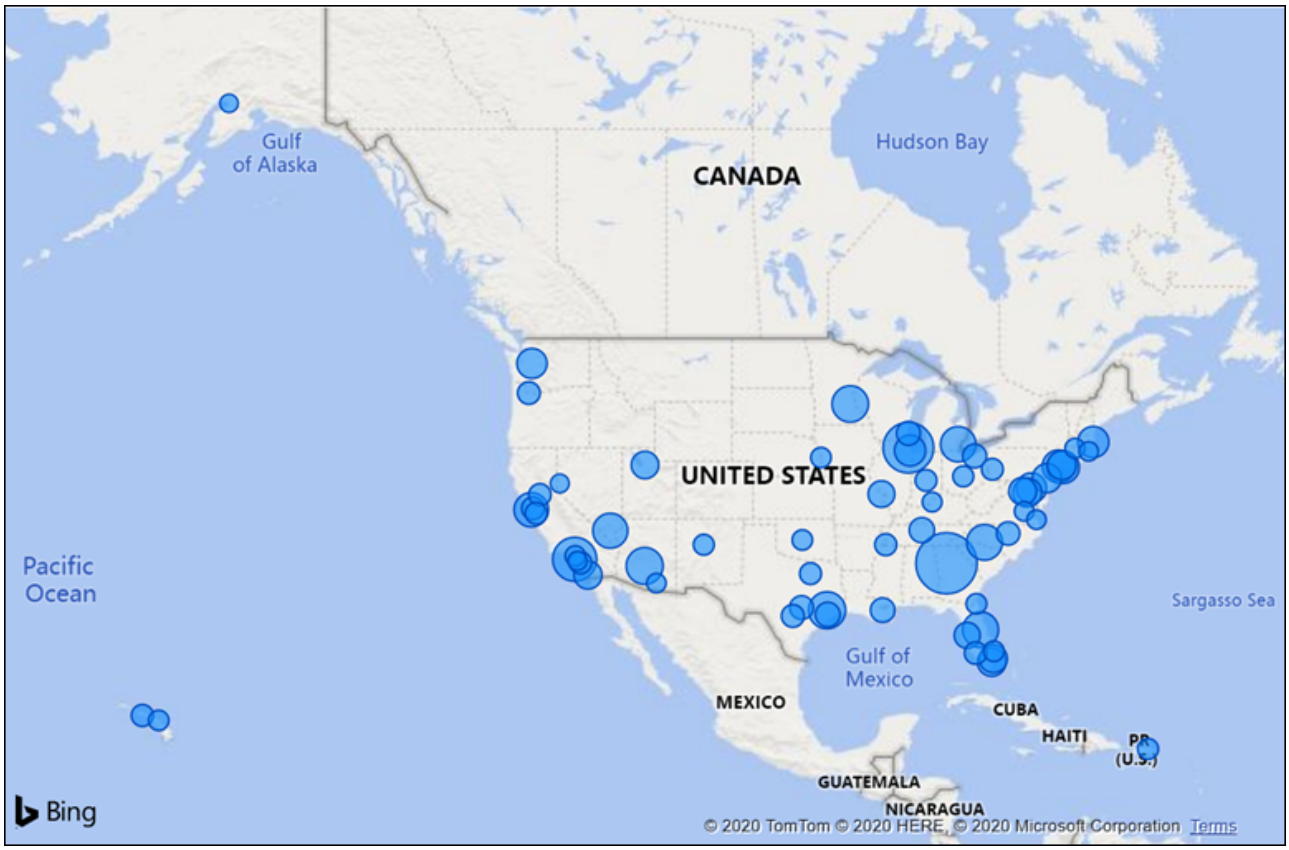
2. From the Visualizations area, next to Fields, select the Globe icon to add a Map visualization to the report design surface.



3. With the Map visualization still selected, drag the **OriginLatLong** field to the **Location** field under Visualizations. Then Next, drag the **NumDelays** field to the **Size** field under Visualizations.

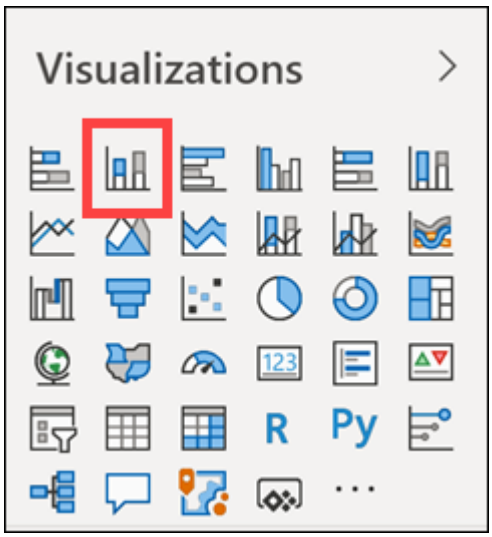


4. You should now see a map that looks similar to the following (resize and zoom on your map if necessary):



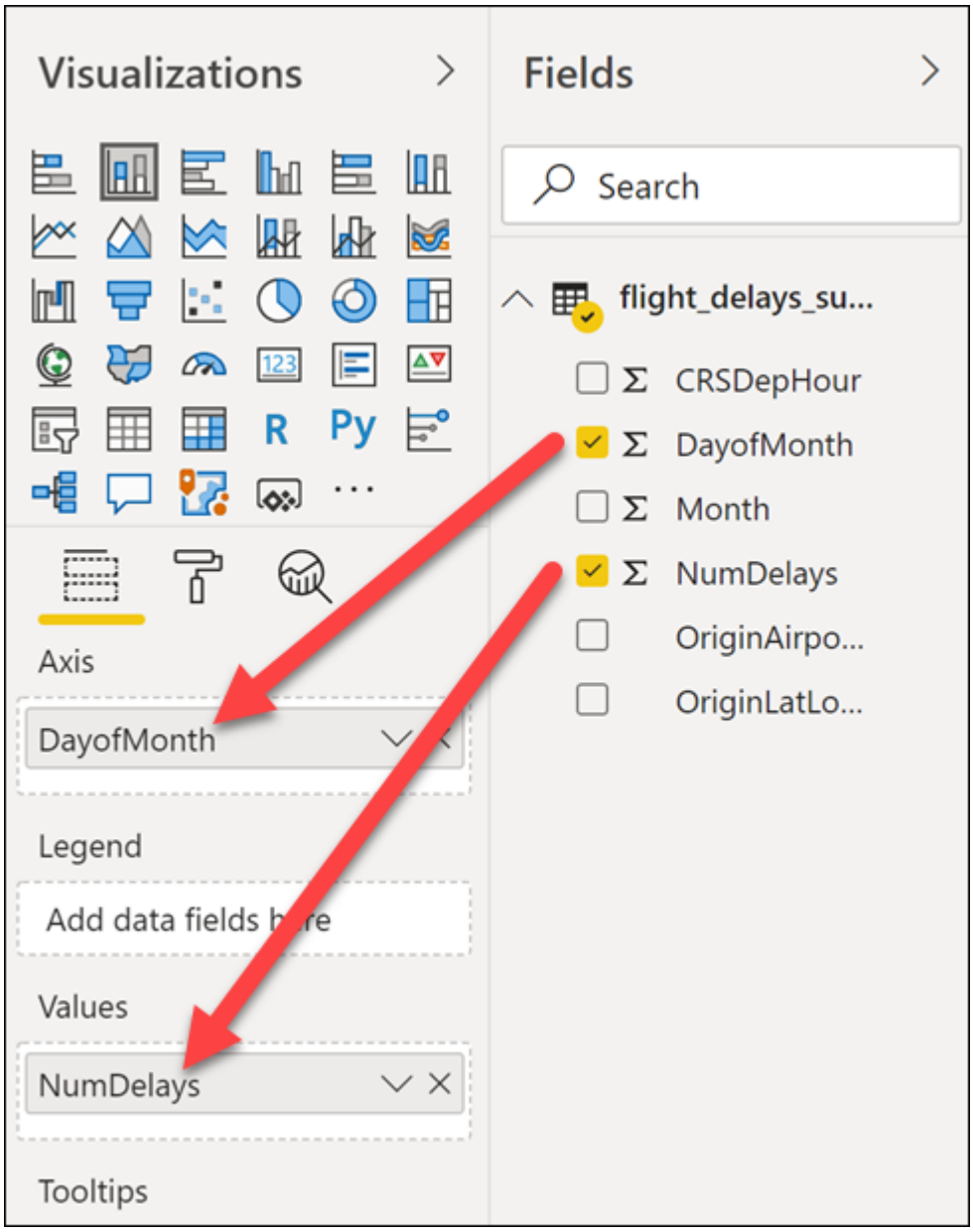
5. Unselect the Map visualization by selecting the white space next to the map in the report area.

6. From the Visualizations area, select the **Stacked Column Chart** icon to add a bar chart visual to the report's design surface.

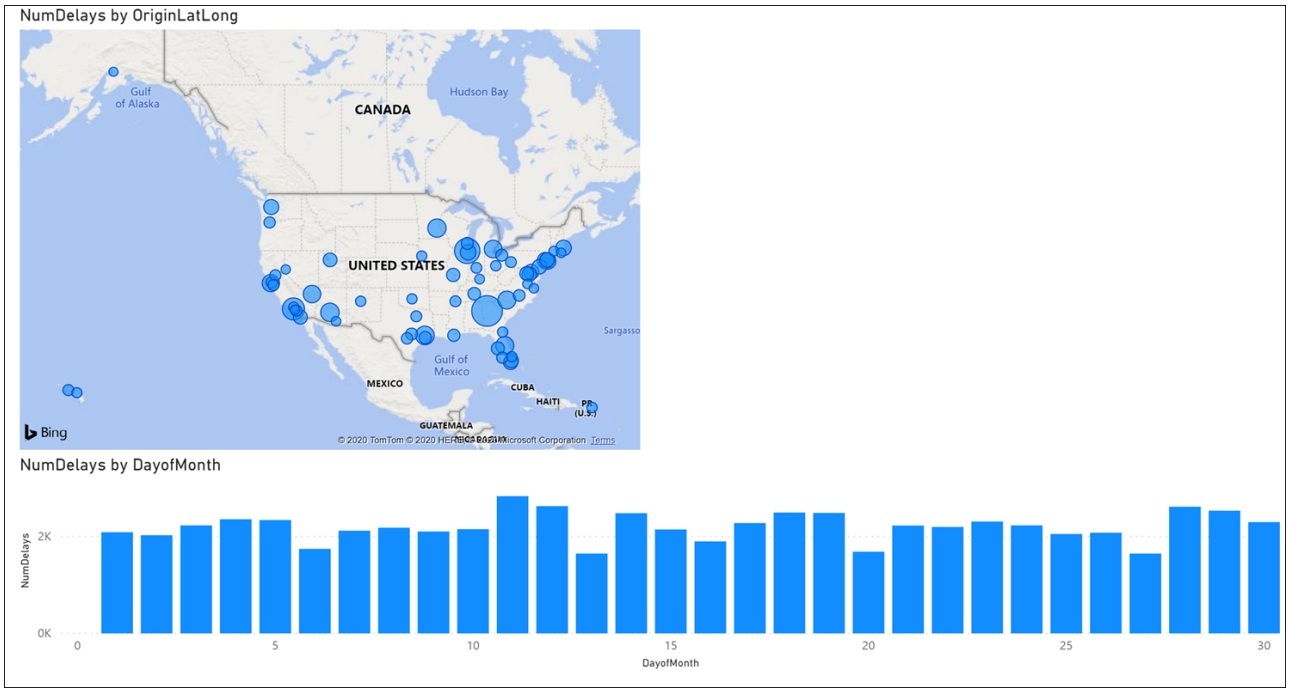


7. With the Stacked Column Chart still selected, drag the **DayofMonth** field and drop it into the **Axis** field located under Visualizations.

8. Next, drag the **NumDelays** field over, and drop it into the **Value** field.

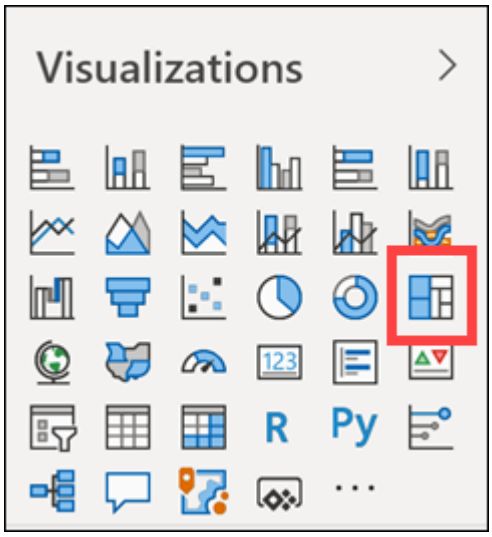


9. Grab the corner of the new Stacked Column Chart visual on the report design surface, and drag it out to make it as wide as the bottom of your report design surface. It should look something like the following.



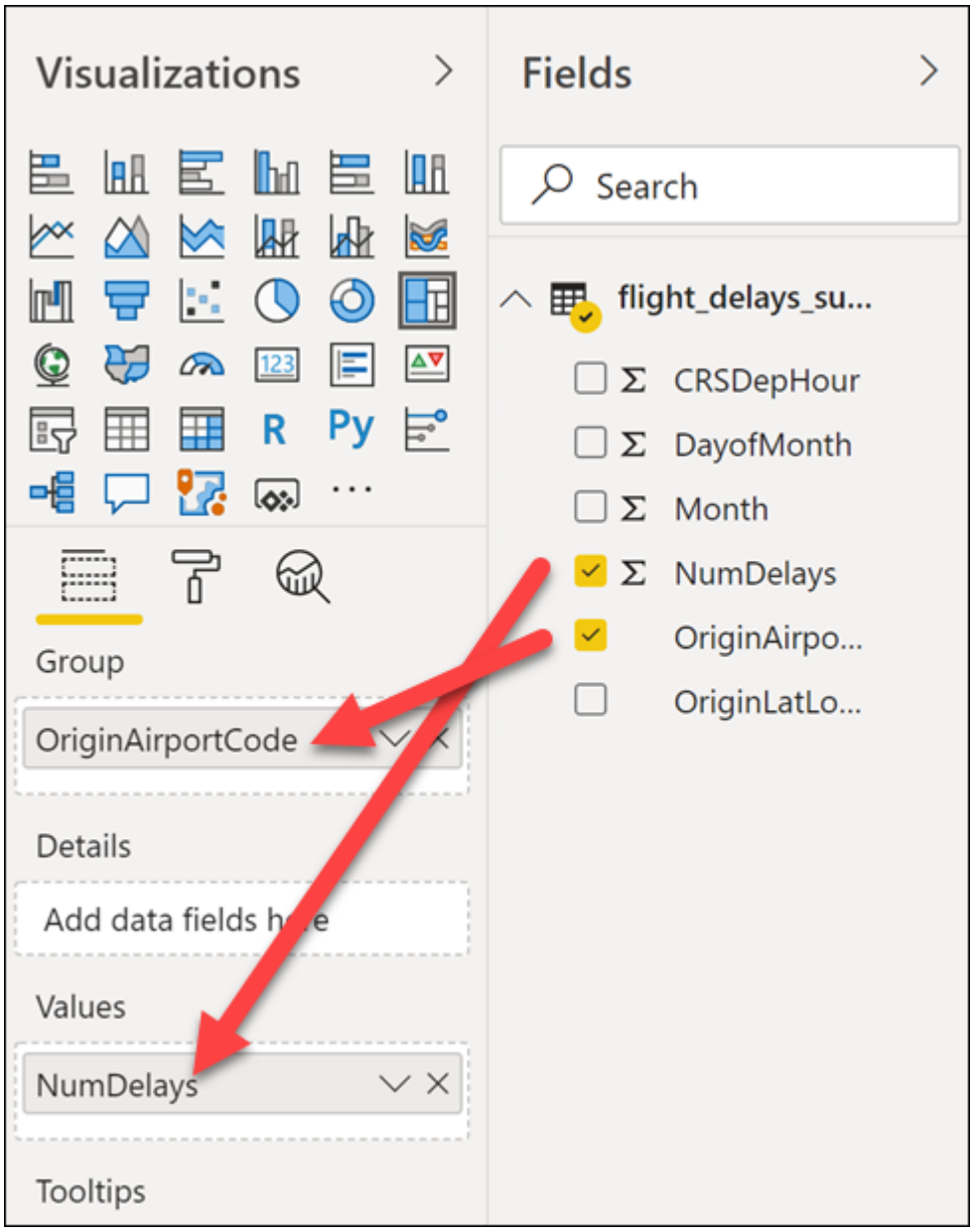
10. Unselect the Stacked Column Chart visual by selecting the white space next to the map on the design surface.

11. From the Visualizations area, select the Treemap icon to add this visualization to the report.

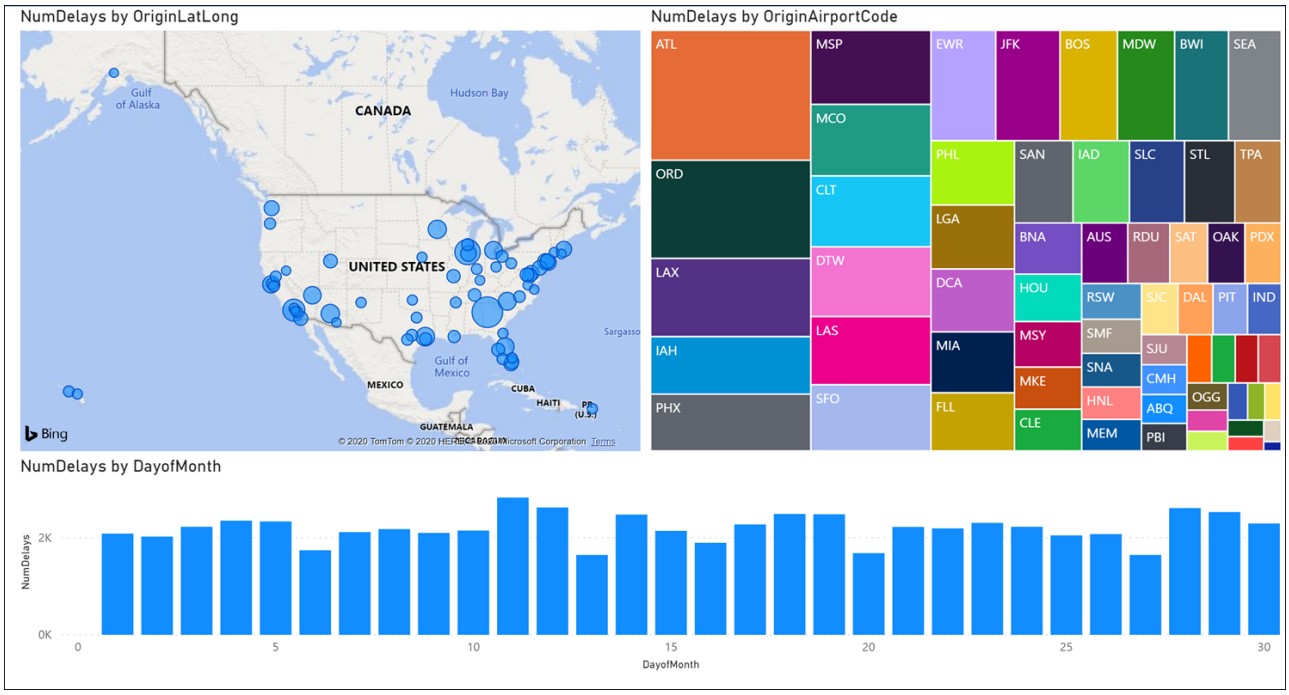


12. With the Treemap visualization selected, drag the **OriginAirportCode** field into the **Group** field under Visualizations.

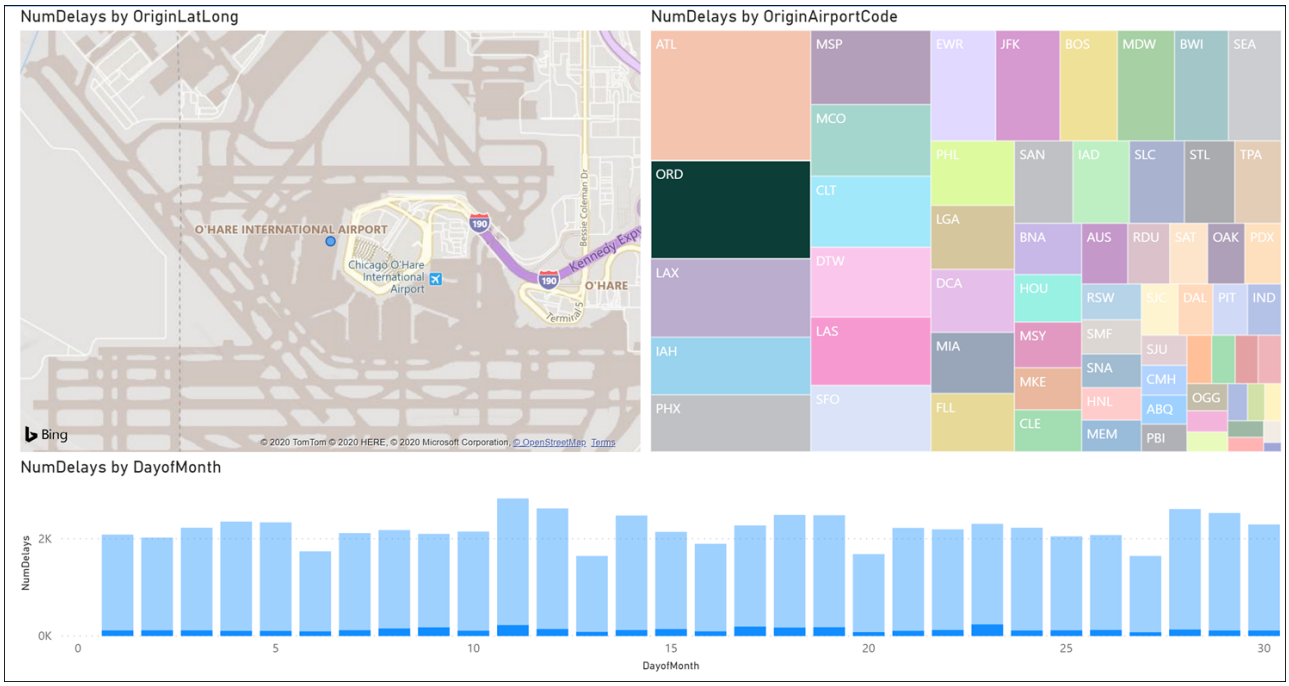
13. Next, drag the **NumDelays** field over, and drop it into the **Values** field.



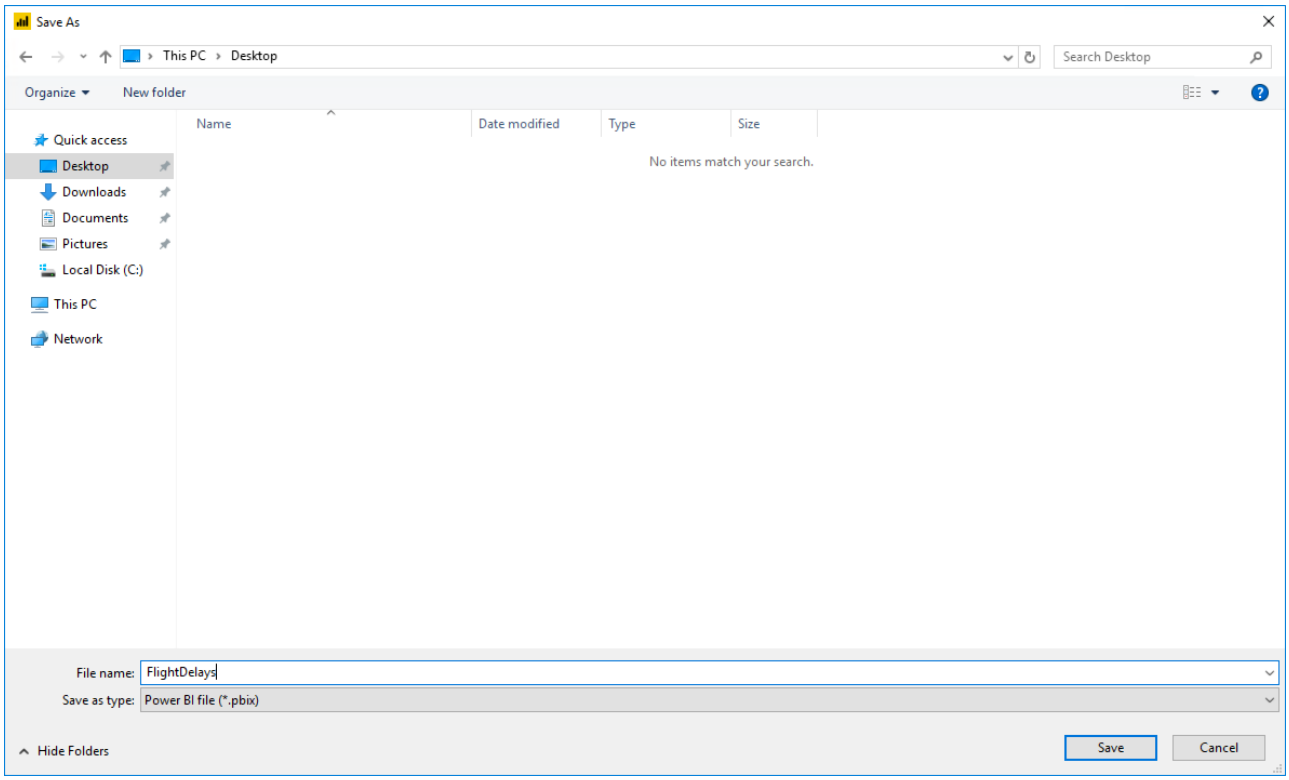
14. Grab the corner of the Treemap visual on the report design surface, and expand it to fill the area between the map and the side edge of the design surface. The report should now look similar to the following.



15. You can cross filter any of the visualizations on the report by selecting one of the other visuals within the report, as shown below (This may take a few seconds to change as the data is loaded).



16. You can save the report by choosing Save from the File menu and entering a name and location for the file.


Save As
This PC › Desktop
Search Desktop
Organize
New folder
Quick access
Desktop
Downloads
Documents
Pictures
Local Disk (C:)
This PC
Network
Name
Date modified
Type
Size
No items match your search.
File name: FlightDelays
Save as type: Power BI file (*.pbix)
Hide Folders
Save
Cancel