Business Analytics with Power BI

Module 3 – Predictive Analytics with Power BI and R

Student Lab Manual – Lab 4 – Using Azure Machine Learning and Power BI

Version 1.0

Conditions and Terms of Use

Microsoft Confidential

This training package is proprietary and confidential, and is intended only for uses described in the training materials. Content and software is provided to you under a Non-Disclosure Agreement and cannot be distributed. Copying or disclosing all or any portion of the content and/or software included in such packages is strictly prohibited.

The contents of this package are for informational and training purposes only and are provided "as is" without warranty of any kind, whether express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, and non-infringement.

Training package content, including URLs and other Internet Web site references, is subject to change without notice. Because Microsoft must respond to changing market conditions, the content should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication. Unless otherwise noted, the companies, organizations, products, domain names, e-mail addresses, logos, people, places, and events depicted herein are fictitious, and no association with any real company, organization, product, domain name, e-mail address, logo, person, place, or event is intended or should be inferred.

Copyright and Trademarks

© 2016 Microsoft Corporation. All rights reserved.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

For more information, see Use of Microsoft Copyrighted Content at

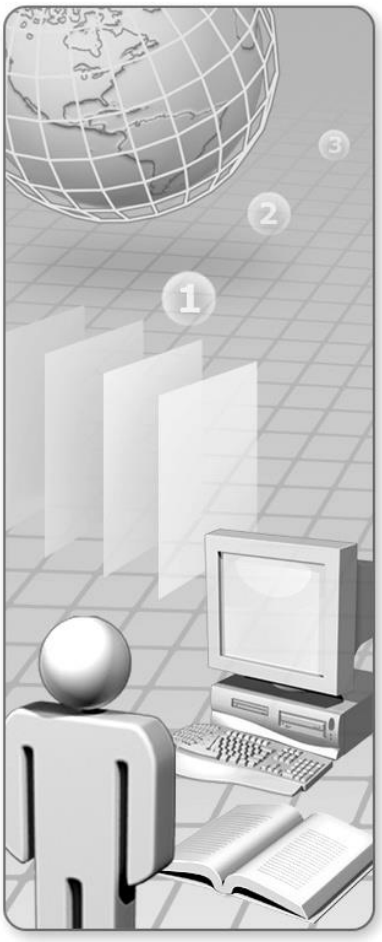
http://www.microsoft.com/en-us/legal/intellectualproperty/Permissions/default.aspx

DirectX, Hyper-V, Internet Explorer, Microsoft, Outlook, OneDrive, SQL Server, Windows, Microsoft Azure, Windows PowerShell, Windows Server, Windows Vista, and Zune are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. Other Microsoft products mentioned herein may be either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. All other trademarks are property of their respective owners.

Module 3

Predictive Analytics with Power BI and R

Lab 4 – Using Azure Machine Learning and Power BI



© 2016 Microsoft Corporation

Microsoft Confidential

Lab 4: Using Azure Machine Learning and Power BI

Introduction

In this lab, you will use Microsoft Power BI to show results based on Microsoft Azure Machine Learning predictions. Our context is the Titanic scenario, where we need to determinate if a specific person would have survived the disaster (classification problem) or not.

Objectives

After completing this lab, you will be able to:

- Use Azure Machine Learning to create predictive models.
- Use Azure SQL Database as a repository for predicted data.
- Consume predicted data with Microsoft Power BI.

Estimated time to complete this lab

1.5 hours

Prerequisites

You will need the following to complete this lab successfully:

- A Microsoft account. It will be used for Azure Machine Learning and Azure.
- An Azure Machine Learning account. If you do not have one, you can create at: http://studio.azureml.net (select **Sign Up** and **Free Workspace**).
- An Azure account. If you do not have one, you can create one for free at: https://azure.microsoft.com/en-us/free/. Note: a valid credit card will be required. You get an Azure account, plus a $200 credit to spend on Azure services for 30 days. You can use this $200 credit to try out any combination of Azure services. You can cancel the account after the labs for this course. You can find more information at: https://azure.microsoft.com/en-us/free/free-account-faq/.

© 2016 Microsoft Corporation

Resources

Virtual machine (VM) Name	**Business Analytics with Power BI - Module 1**
Domain	**POWERBI-WIN10**
User	**POWERBI-WIN10\LabUser**
Password	**P@ssw0rd1!**
Lab Files	**E:\Labs**
Asset Files	**E:\Assets**

© 2016 Microsoft Corporation

Exercise 1: Using Azure Machine Learning and Power BI

Introduction

In this exercise, you will create an analysis based on predicted data from Azure Machine Learning model.

The scenario for prediction will be the Titanic disaster. You will plot information about the survival rate. The database will contain information about people that were not in the disaster, and we will try to determine if they would have survived or not.

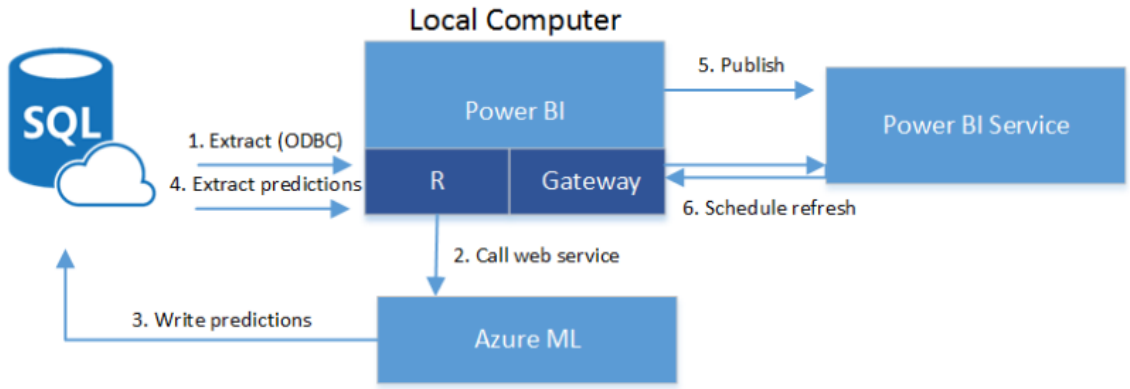
What you will do is:

1. Use a R script to extract data out of an Azure SQL Database that has not yet been scored by machine learning model. This script will use Open Database Connectivity (ODBC) to retrieve the data.
2. Use R script to call the Azure Machine Learning web service and send it the unscored data.
3. Write the output of the Azure ML model back into Azure SQL Database.
4. Use R to read scored data into Power BI Desktop.

(Challenge – if time permits)

5. Publish the Power BI file to the Power BI service.
6. Schedule a refresh of the data using the Personal Gateway (which triggers a scheduled re-run of our R script and brings in new predictions).

In the end, you will have the following architecture:



© 2016 Microsoft Corporation

Microsoft Confidential

Objectives

After completing this exercise, you will be able to:

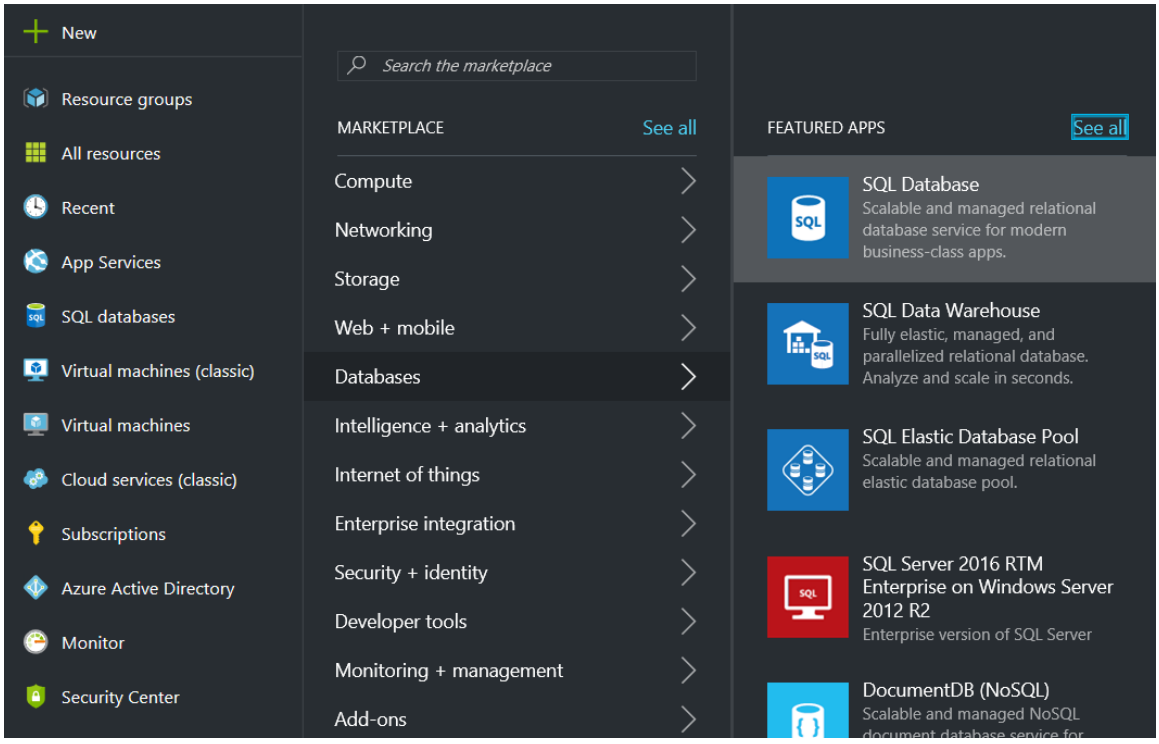
- Create an Azure SQL Database
- Create an Azure Machine Learning Model
- Create a Power BI report and interoperate with Azure Machine Learning models.

Creating a Repository (Azure Database)

In this task, you will create an Azure SQL database. If you already have one, you can use it. Just remember to create a firewall rule to enable access using your local IP address (step 8 of this task), and you can then skip to the next task.

1. Go to the Azure portal: http://portal.azure.com.

2. Select **+ New**, select **Databases**, and then select **SQL Database**.



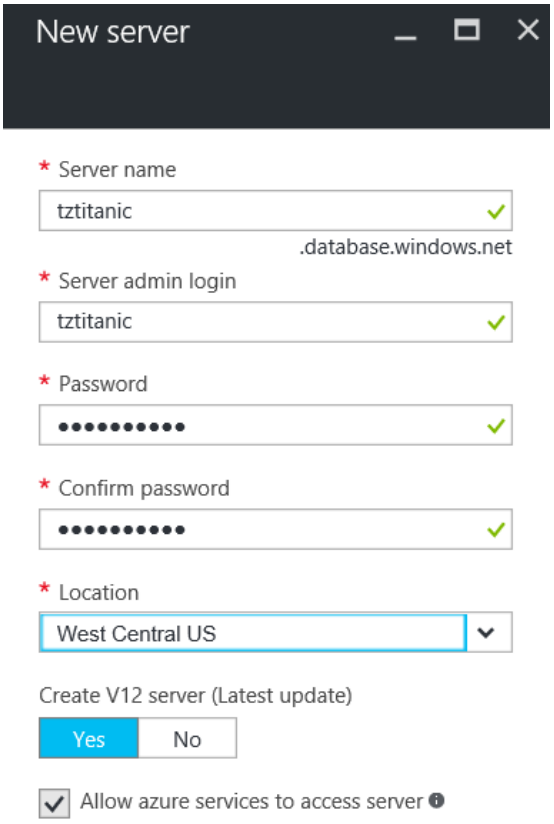
3. Enter the data to create your database:

   - Database Name: **Titanic**
   - Subscription: **Select your own subscription**.
   - Resource group: **Create New**, and name it as **TitanicRG**

© 2016 Microsoft Corporation

Microsoft Confidential

- Select source: **Blank database**.
- Server: **Create a new server**.
- Server name: It needs to be a unique name. Try to use your own **name initials + titanic**.
- Server admin login: Use the same name that you gave the server (for simplicity).
- Password: Use one that meets the requirements. A suggestion would be **P@ssw0rd1!**. Note that you can use your own password, but just make sure that you can remember it later.
- Location: **Select one near you**.
- Create V12 server (Latest update): **Yes**.
- Allow azure services to access server: **Selected**.

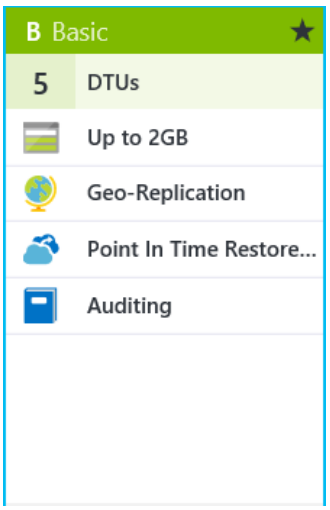


4. Click **Select**.

5. In the Pricing Tier, select **B Basic**. You can select another tier if you want. For our lab, this will be enough.

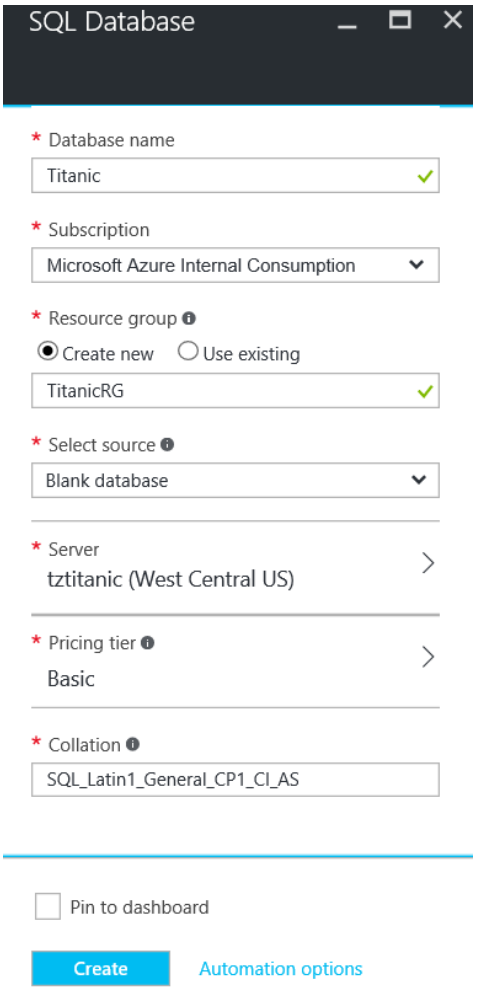
© 2016 Microsoft Corporation

Microsoft Confidential



Collation: **SQL_Latin1_General_CP1_CI_AS**.

6. You should now see a screen that is similar to the following screenshot:

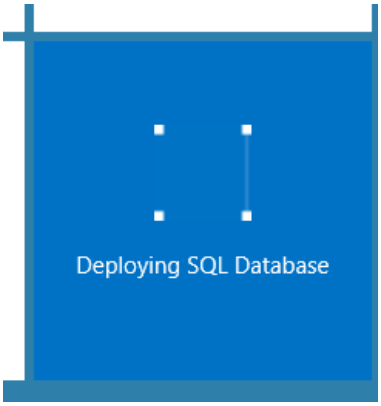


© 2016 Microsoft Corporation

Microsoft Confidential

7. Select the **Pin to dashboard** check box, and then click **Create**.

8. This will take a few minutes. You can view the progress in your Azure dashboard.

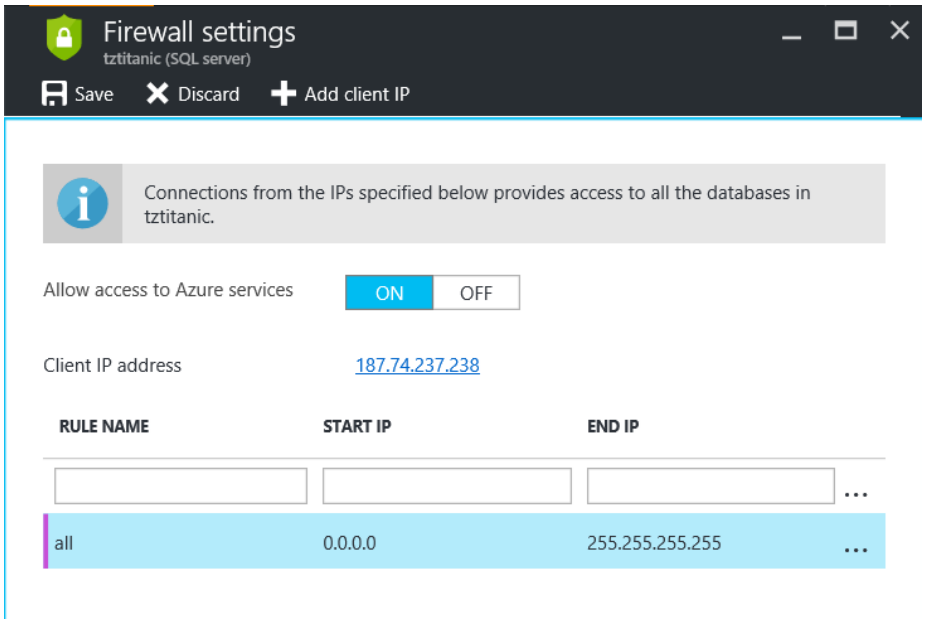


9. When the process finishes, you will be redirected to your database screen.

10. Click the **Set server firewall** tab. We need to allow our local R (using Power BI Desktop) to access the database we just created. For simplicity, we will enable access from any IP. In your production system, you should always enable firewall rules for the actual IPs.

11. Enter the following data, and then click **Save**. Click **Ok**.

    Rule name: **all**
    Start IP: **0.0.0.0**
    End IP: **255.255.255.255**



© 2016 Microsoft Corporation

Microsoft Confidential

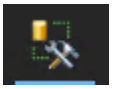
12. Click + **Add client IP**, click **Save**, and then click **OK**.

13. Leave the portal open. You will need to get information about your database later.

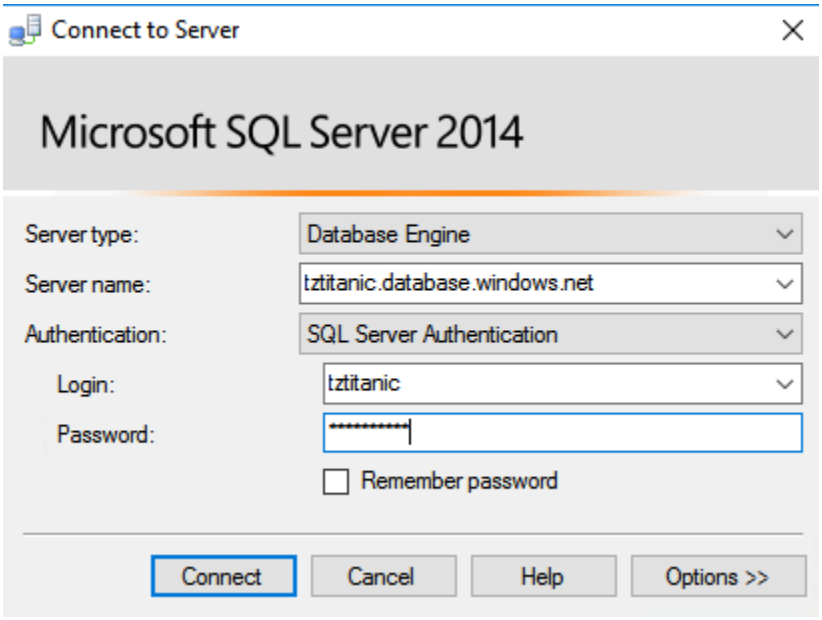
Populating Data

In this task, you will create two tables in the Titanic database and add data to it.

1. Open **Microsoft SQL Server Management Studio**. You can find it on the taskbar.



2. In the **Connect to Server** window, fill in the information about the database server you just created. After this, your screen should look like the following screenshot. Click **Connect**. If you see connectivity issues, verify whether your client IP was added to the firewall rules (last task).



3. Click **New Query**.

4. Open the file located at: **E:\Assets\M3 - Lab 4\TitanicTablesData.txt**. Copy all its contents.

5. Paste the script to the new query window you created in step 3. **Run** the entire script (you can press F5).

6. Click **New Query** again. Add the following script and **run** it. The query needs to return all values you just inserted.

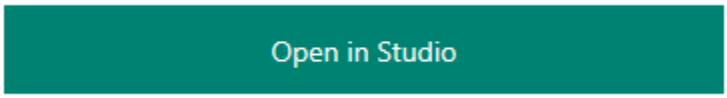
```
SELECT * FROM Titanic
```

7. Close **SQL Server Management Studio**.

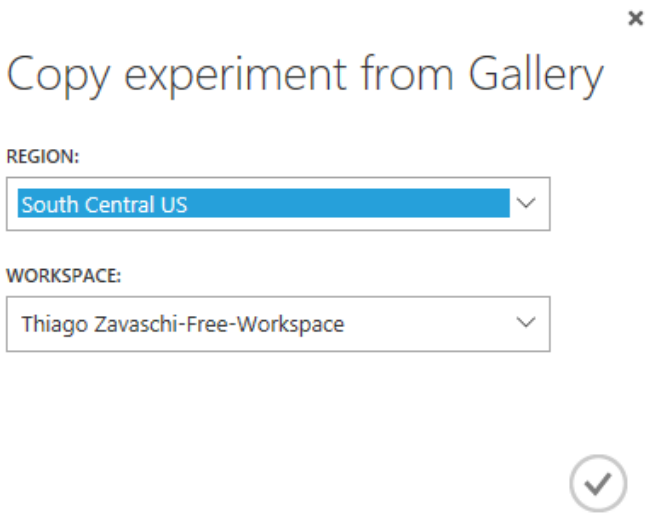
Creating an AzureML Model

In this task, you will create an Azure Machine Learning model and publish a web service to access it.

1. Open Azure ML Studio at http://studio.azureml.net, and then sign in. If you do not have an account, you can create one (select **Sign Up Here**, select **Free Workspace**, and then follow the on-screen instructions).
2. Open another tab in your browser and go to https://aka.ms/azuremltitanic.
3. Click **Open in Studio**.



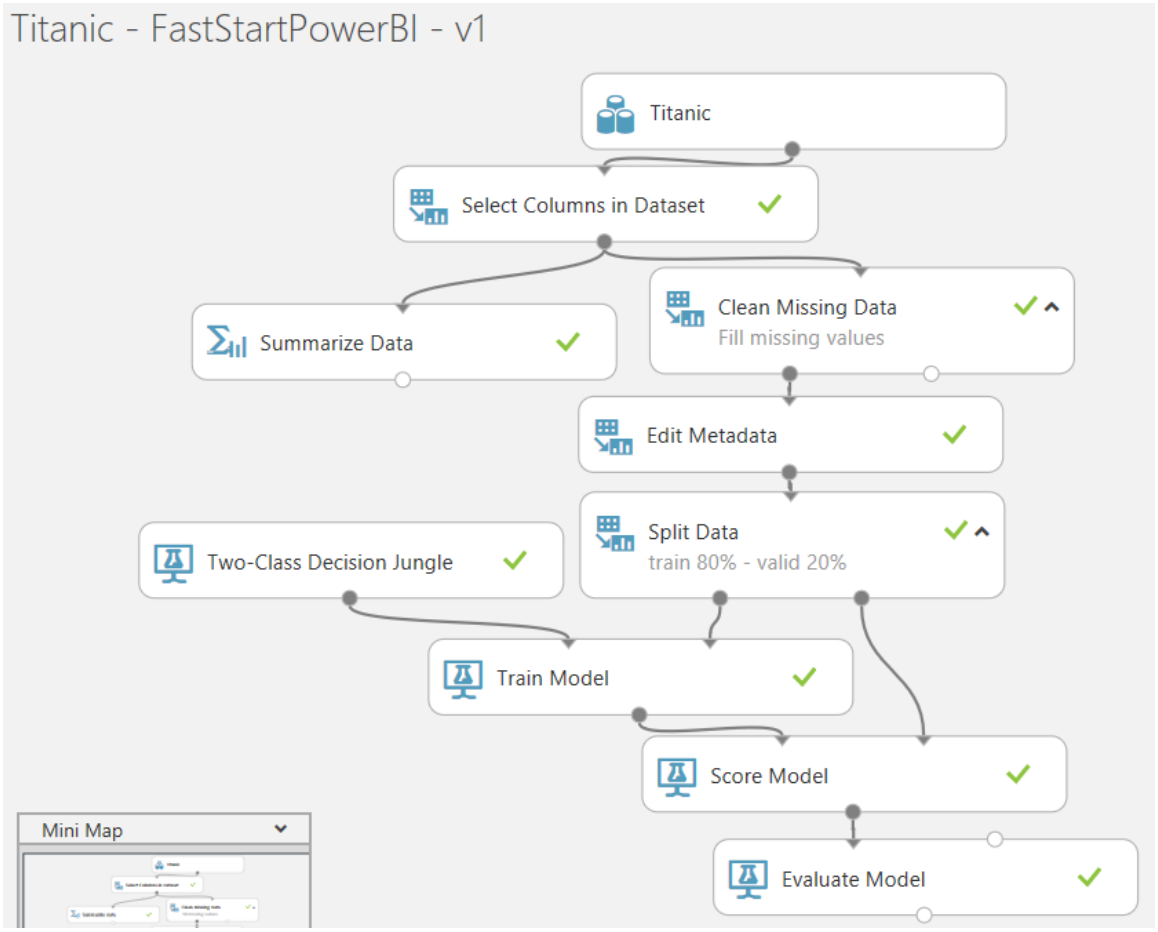
4. You will be redirected to you Azure Machine Learning studio to create a copy from this experiment. Select a region near to you.



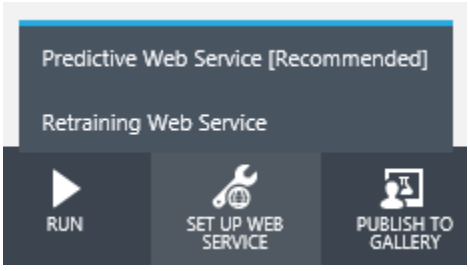
5. After the experiment is created, you can click **RUN**, and wait for process to finish. This operation can take a few moments to finish.

© 2016 Microsoft Corporation

6. Your experiment should look like the following:



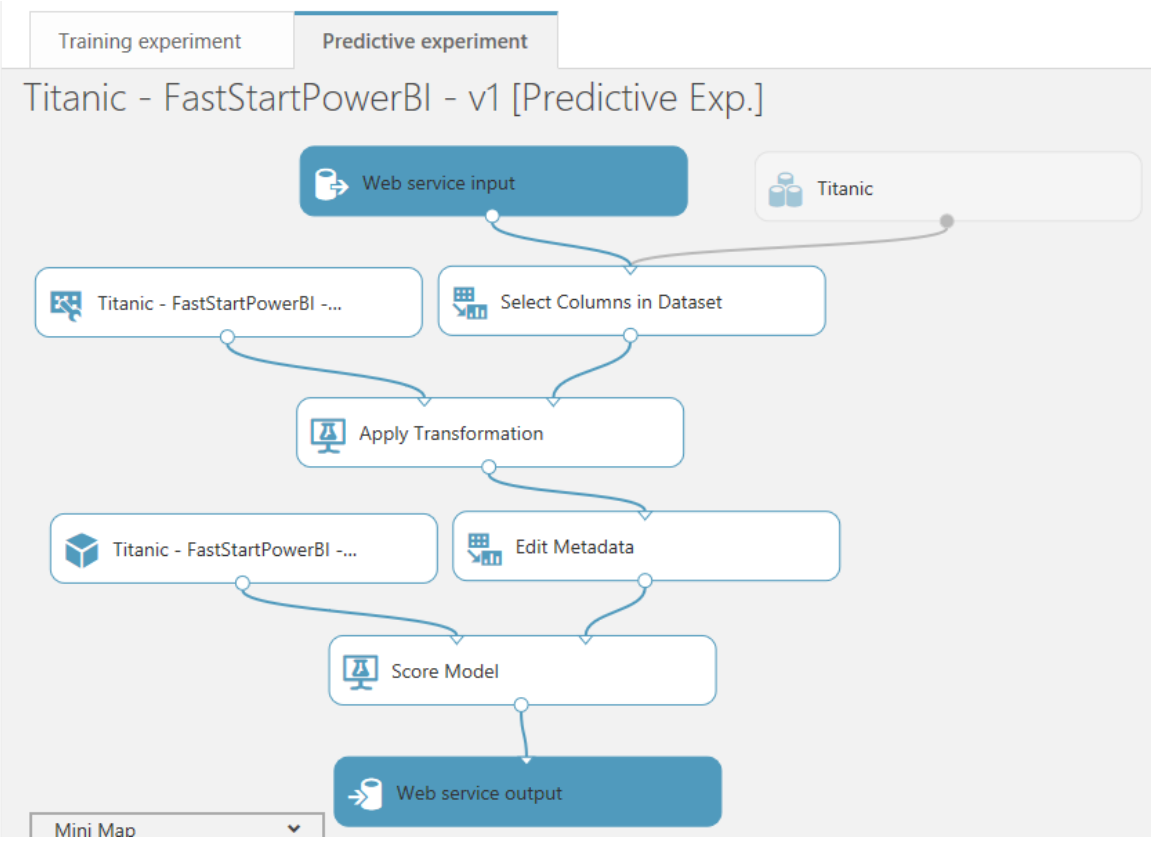
7. Move the cursor over **SET UP WEB SERVICE**, and then select **Predictive Web Service [Recommended]**. The process will take a few moments to finish.



8. Reorganize the modules so you can see your experiment better. You should see something similar to the following screenshot. Notice that now we have two tabs: **Training experiment** and **Predictive experiment** (scoring experiment). You also have the modules for web service input and output.

© 2016 Microsoft Corporation

Microsoft Confidential



9. Now you need to change your predictive experiment so that it can write your predictions to your Azure SQL Database table called **TitanicScored**.

10. Remove the **connection** between **Web service input** and **Select Columns in Dataset** (select the arrow and press the Delete key).

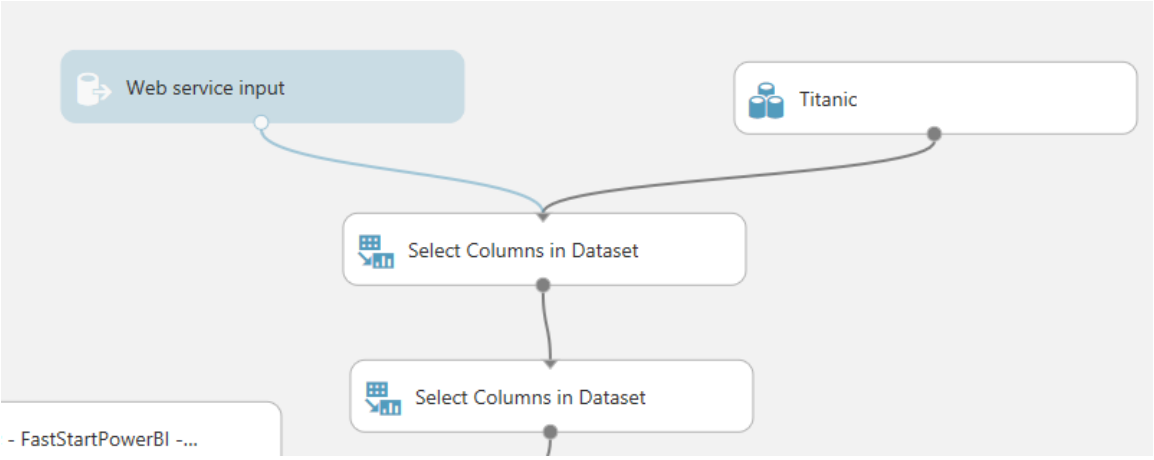
11. Remove the **connection** between **Titanic** and **Select Columns in Dataset** (select the arrow and press the Delete key).

12. Add a new **Select Columns in Dataset** module and connect the output of the **Web service input** module to this new module. Do the same for the **Titanic** dataset.

13. Connect the output from the new **Select Columns in Dataset** module to the input of the old **Select Columns in Dataset** module. This part of your experiment should be similar to the following screenshot:

© 2016 Microsoft Corporation

Microsoft Confidential



14. Select the new **Select Columns in Dataset** module. In the properties pane, click **Launch column selector**.

15. Select **WITH RULES**. Begin with: **ALL COLUMNS**, and then click – button. Your screen should look like the following screenshot:

16. Click the check mark to confirm.

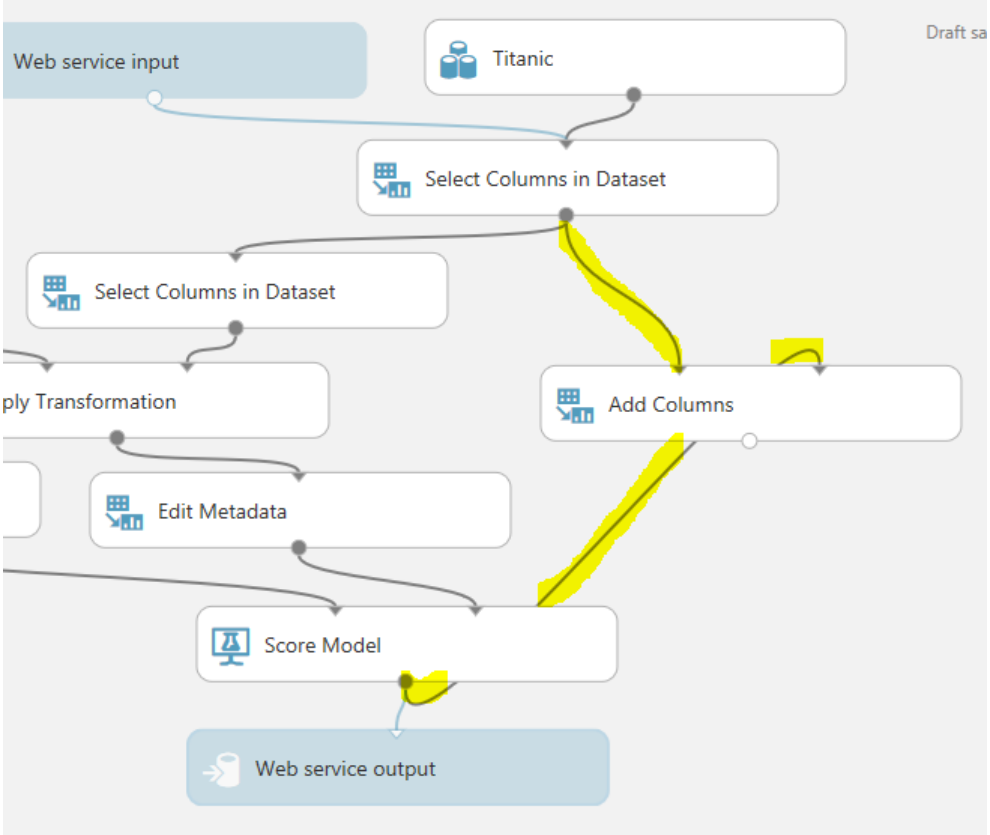
17. Add an **Add Columns** module.

18. Connect the output of the recently added **Select Columns in Dataset** module to the first input (leftmost) of the **Add Columns** module.

19. Connect the output of **the Score Model** module to the second input (rightmost) of the **Add Columns** module. Make sure this part of your experiment looks like the following screenshot:

© 2016 Microsoft Corporation

Microsoft Confidential



20. Click **Run**. Everything should run smoothly.

21. Add an **Export Data** module.

22. Connect the output of the **Add Columns** module to the input of the **Export Data** module.

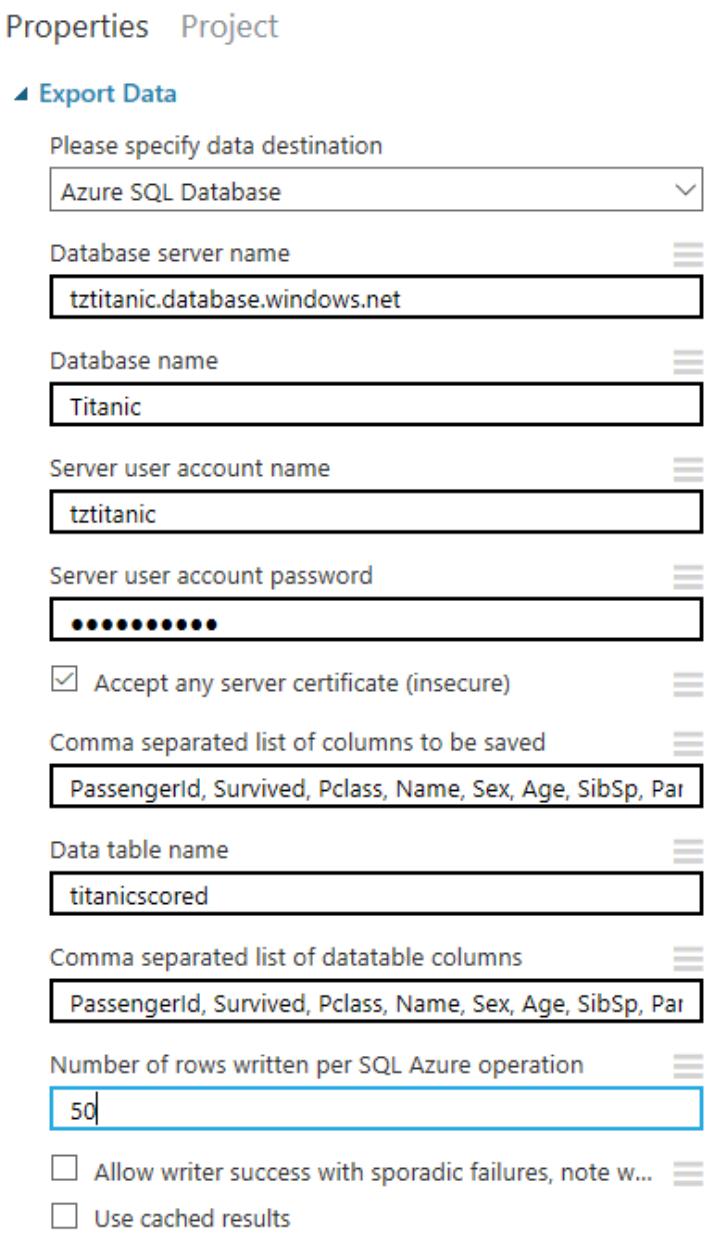
23. Select the **Export Data** module. In the **Properties** pane, specify the following values.

    - Please specify data destination: **Azure SQL Database**.
    - Database server name: Use your Azure Database server name. It should be similar to: **tztitanic.database.windows.net**.
    - Database name: **Titanic**.
    - Server user account name: Use the user you specified when you created the Azure database. Should be similar to: **tztitanic**.
    - Server user account name: Use your password.
    - Accept any server certificate (insecure): **Selected**.
    - Comma separated list of columns to be saved: **PassengerId**, **Survived**, **Pclass,** **Name**, **Sex**, **Age**, **SibSp**, **Parch**, **Ticket**, **Fare**, **Cabin**, **Embarked**, **Scored Labels**, and **Scored Probabilities**
    - Data table name: **titanicscored**

© 2016 Microsoft Corporation

Microsoft Confidential

- Comma separated list of datatable columns: **PassengerId**, **Survived**, **Pclass**, **Name**, **Sex**, **Age**, **SibSp**, **Parch**, **Ticket**, **Fare**, **Cabin**, **Embarked**, **[Scored Labels]**, and **[Scored Probabilities]**
- Number of rows written per SQL Azure operation: **50**
- Allow writer success with sporadic failures: **Cleared**
- Use cached results: **Cleared**



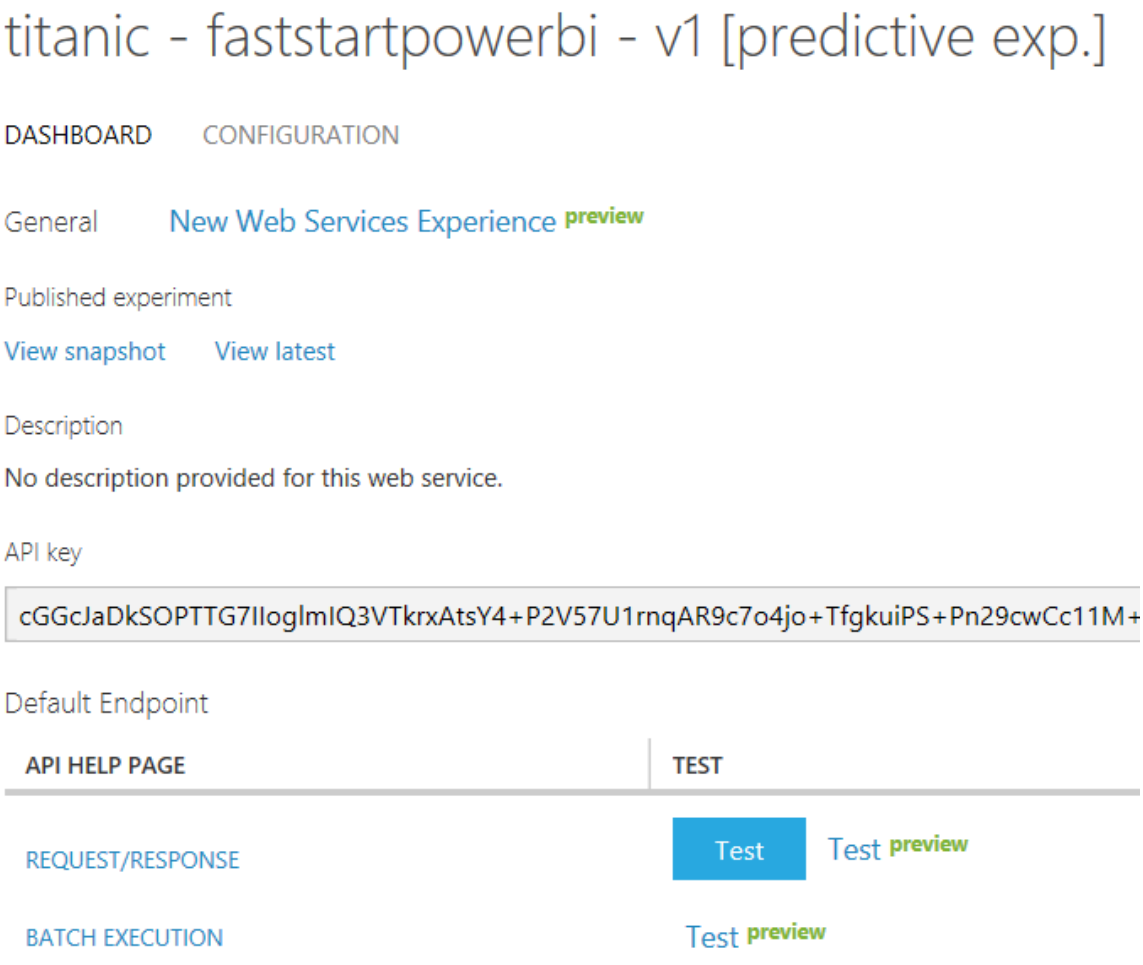
24. Click on **Run** button and wait the process to finish.
25. Save your experiment.

© 2016 Microsoft Corporation

Microsoft Confidential

26. Click on **DEPLOY WEB SERVICE** button.

27. You will be redirected to the web service page. Leave this website open, because you will need to get information about API keys from it later.



28. Open **SQL Server Management Studio**. Connect to your Azure database server.

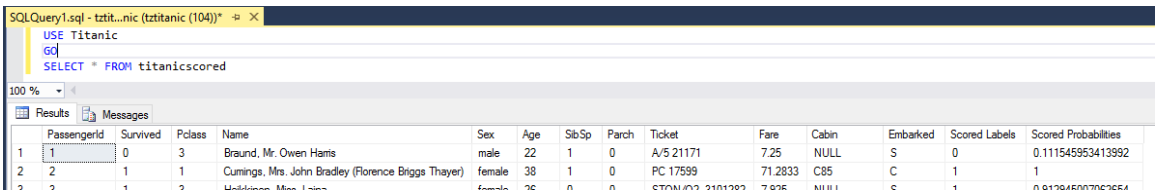
29. Run the following query:

```
USE Titanic
GO
SLECT * FROM titanicscored
```

30. You should see the result of your prediction.

© 2016 Microsoft Corporation

Microsoft Confidential



	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Scored Labels	Scored Probabilities
1	1	0	3	Braund, Mr. Owen Harris	male	22	1	0	A/5 21171	7.25	NULL	S	0	0.111545953413992
2	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Thayer)	female	38	1	0	PC 17599	71.2833	C85	C	1	1

31. Close **SQL Server Management Studio**.

Creating Power BI Desktop Report

In this task, you will create a Power BI report that uses an R script as a data source. This script will connect them both: the Azure Machine Learning and Microsoft Power BI report.

1. Open **Microsoft Visual Studio 2015**. You can click the shortcut on the taskbar.

2. Open the R Interactive pane and execute the following commands. This can take a few moments to finish.

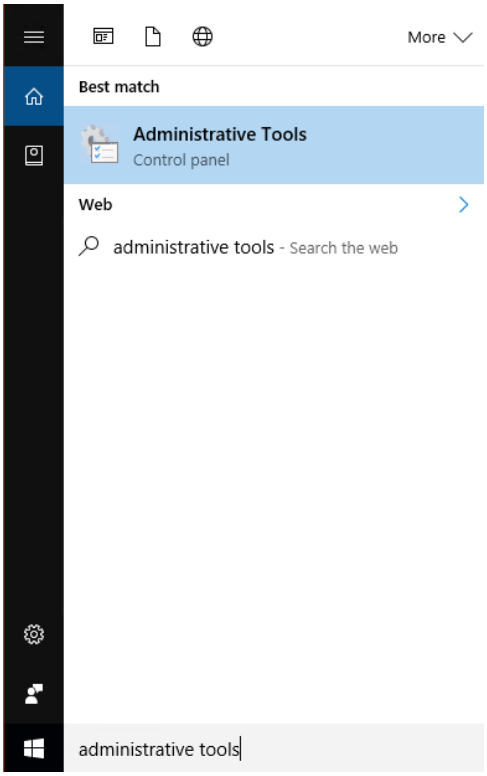
```
install.packages(c("RODBC","RCurl", "rjson"))
```

3. Close Visual Studio.

4. From the taskbar, in Cortana search, type **Administrative Tools**.

5. You should see Administrative Tools in the result. Click **Administrative Tools**.

© 2016 Microsoft Corporation



6. Open **ODBC Data Sources (64-bit)**.

7. Click **Add** in the **ODBC Data Source Administrator (64-bit)** window.

8. Select **SQL Server Native Client 11.0**, and then click **Finish**.

9. Enter the following information in the **Create a New Data Source to SQL Server** window:

    - Name: **titanicserver**
    - Description: **titanicserver**
    - Server: Your server name. It should be similar to: **tztitanic.database.windows.net**.

10. Click **Next**.

11. Select the **With SQL Server authentication using a login ID and password entered by the user** check box. Type in your credentials that you created with Azure database. Your Login ID should be similar to: **tztitanic**.

12. Click **Next**.

13. Select the **Change the default database to** option, and then select **Titanic**.

© 2016 Microsoft Corporation

14. Click **Next**, and then click **Finish**.

15. Click **Test Data Source**. You should receive the following message: **TESTS COMPLETED SUCCESSFULLY!**

16. Click **OK**, and then click **OK** again.

17. Open the file located at **E:\Assets\M3 - Lab 4\RDataSource.txt**. Select and copy all its contents .

18. Open **Power BI Desktop**. You can click the shortcut on the taskbar.

19. Add a new **R Script** data source.

20. Paste the script you just copied in step 17 to the R script window.

21. Change the following script line by replacing the **uid** and **pwd** parameters with your user and password credentials that you defined when you created the Azure SQL database.

```
conn <- odbcConnect("titanicserver", uid = "userid", pwd =
"password!" )
```

22. Replace the API key with your own API key from the Azure Machine Learning web service page that you left open in the last task.

    This code line should look like this:

    ```
    api_key =
    "cGGcJaDkSOPTTG7IIoglmIQ3VTkrxAtsY4+P2V57U1rnqAR9c7o4jo+TfgkuiPS
    +Pn29cwCc11M+mjh6TSnu5A=="
    ```

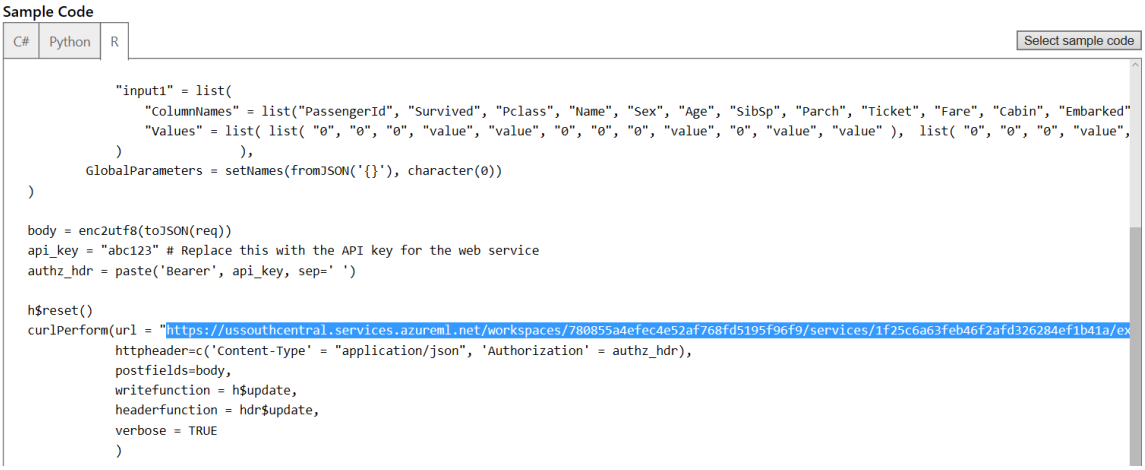
23. To get the final piece of information that you need, go to the same Azure Machine Learning webpage on which you got the API key, and then click **REQUEST/RESPONSE**.

24. Go to the end of the opened page. Click the **R** tab to see a sample code in R. You should see a similar code than the one you are using in your Power BI Desktop report. Look for

© 2016 Microsoft Corporation

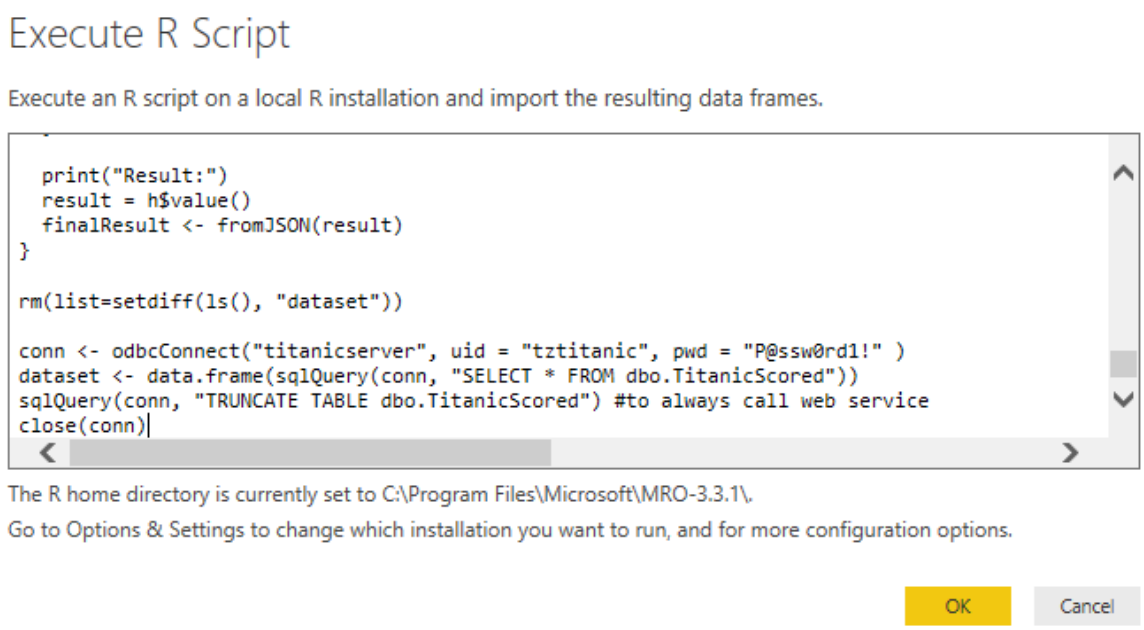
Microsoft Confidential

the **curlPerform()** function and copy the **url** parameter to your experiment (search for "urlgoeshere").



25. After you copy the **url** to the **curlPerform()** function, go to the end of your script and change the parameters for the **odbcConnect()** function. You should type in the same information from step 20.

26. Your screen should be similar to the following screenshot. Click **OK**.



This script does the following (every refresh, either from Power BI Desktop or Power BI service through Personal gateway):

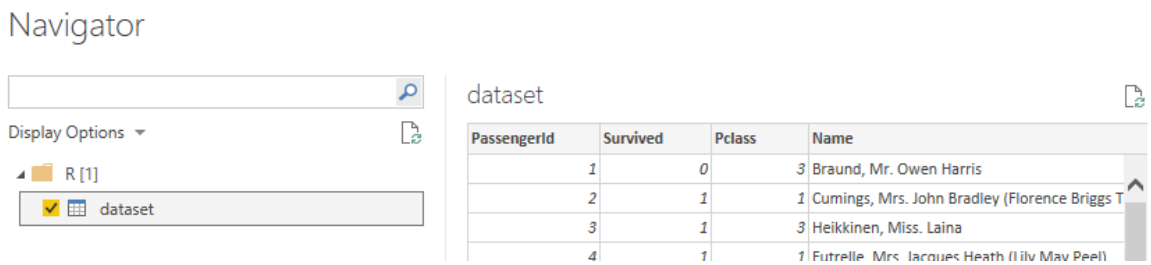
1. Get all the data that is not scored yet (from the **Titanic** table).

© 2016 Microsoft Corporation

Microsoft Confidential

2. If there is data in this condition, then call Azure Machine Learning and classify it.
3. AzureML will write the output of the prediction back to the Azure database (**TitanicScored** table).
4. Retrieve data from the **TitanicScored** table.
5. Truncate the **TitanicScore** table so it will always call the model on the cloud (for lab purposes only).

27. Select the dataset in the **Navigator** window, and click **Load**.

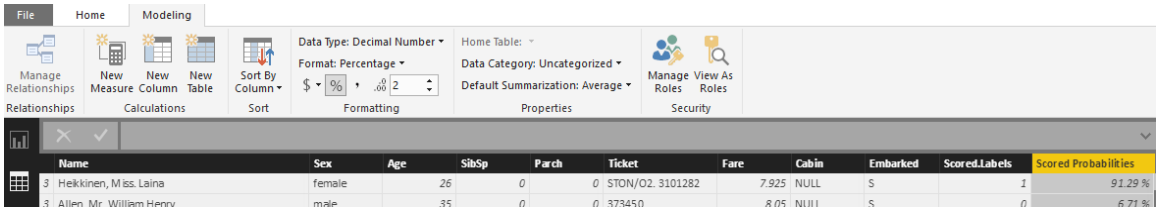


Note: Now you have the scored data and can use all the Power BI visual you already know.

28. Click to see data.

29. Select the **Scored.Probabilities** column and change its name to **Scored Probabilities**.

30. Click the **Modeling** tab (with the **Scored Probabilities** column selected). Change the **Format** parameter to **Percentage**, and **Default Summarization** to **Average**.

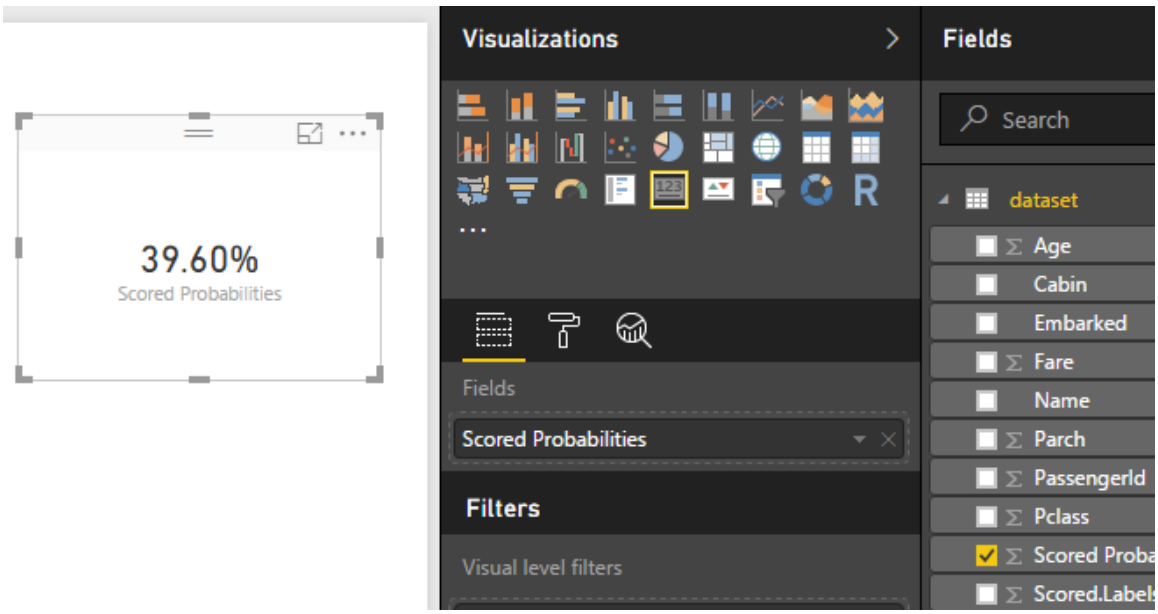


31. Go back to **Report**. Add a **card** visual.

32. Add **Scored Probabilities** from the **Fields** pane to the **Fields** property of the card visual.

© 2016 Microsoft Corporation

Microsoft Confidential



33. Add the **Sex** field as a **Slicer** visual.

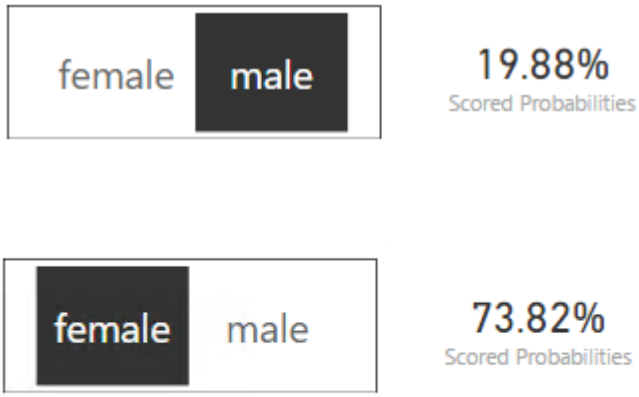
34. Select the slicer and go to **Format**. Change the following properties:

    - General > Orientation: **Horizontal**
    - Header > **Off**
    - Items > Text Size: **22 pt**
    - Border > **On**

    Note: You should have now a very simple report that can show Titanic survival rate by Sex based on a prediction made by Azure Machine Learning.

© 2016 Microsoft Corporation

Microsoft Confidential



35. Save your report at **E:\Labs\M3Lab4\AzureML.pbix**. You can close Power BI Desktop.

Challenge – Auto Refreshing Data

To complete our scenario (if time permits), you still need to perform a couple more tasks, which are now up to you. You can ask your instructor for help).

Perform the following tasks:

1. Publish to the Power BI Service the report you created in the last task.
2. Install and configure a Personal Gateway. Hint: https://powerbi.microsoft.com/en-us/guided-learning/powerbi-learning-4-6-install-configure-personal-gateway/.
3. Configure auto refresh options to your report. You can also force a refresh from the Power BI Service.

Note: Help your colleagues if they are having difficulties and you already finished.

Finishing the Lab

To avoid being charged, remember to release all used resources from Azure (Azure SQL Database, Azure Machine Learning) after you finish the exercises.

© 2016 Microsoft Corporation