Business Analytics with Power BI

Module 3 – Predictive Analytics with Power BI and R

Student Lab Manual – Lab 3 – Using R with 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.

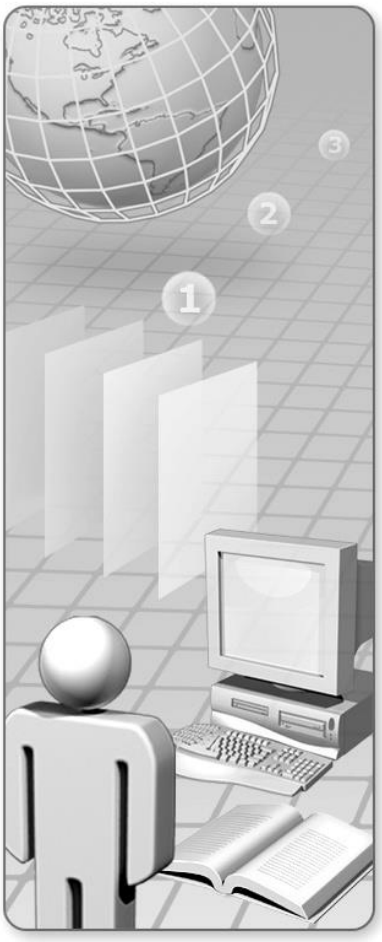
The contents of this package are for informational and training purposes only and are provided "as is" without warranty of any kind, whether express or implied, including but not limited to the implied warranties of merchantability, fitness for a particular purpose, and non-infringement.

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.

Module 3

Predictive Analytics with Power BI and R

Lab 3 – Using R with Power BI



Lab 3: Using R with Power BI

Introduction

As seen in the lab 2, R language is very powerful. In this lab, you will use R as a data source for your analysis and also to create rich visualizations by using R libraries (**ggplot2** and **corrplot**).

Objectives

After completing this lab, you will be able to:

- Create an analysis in Microsoft Power BI by using R data.
- Show R plots directly into a Power BI Desktop file.

Estimated time to complete this lab

45 minutes (depends on experience)

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**

Exercise 1: Using R as a Data Source

Introduction

In this exercise, you will create an analysis where the source is a script in R.

Objectives

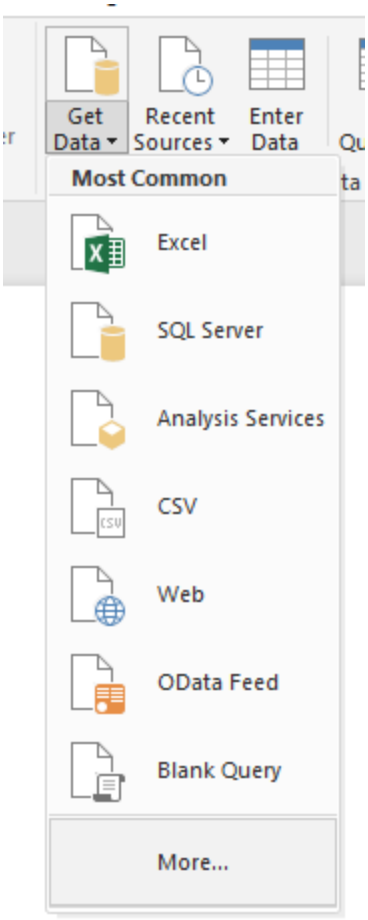
After completing this exercise, you will be able to:

- Create an analysis in Power BI using R data.

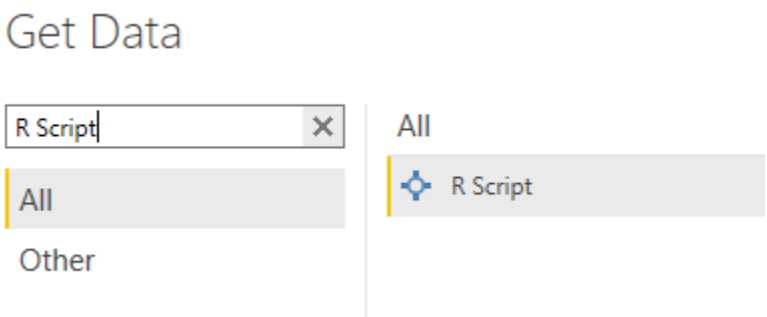
Using R as a Data Source

In this task, you will create a Power BI Desktop report which data will come from a R script.

1. To open **Power BI Desktop**, n the taskbar, click the **Power BI Desktop** shortcut.
2. Click **Get Data** (ribbon menu), and then select **More**.



3. In the Search box, type **R Script**, and then click **Connect**.

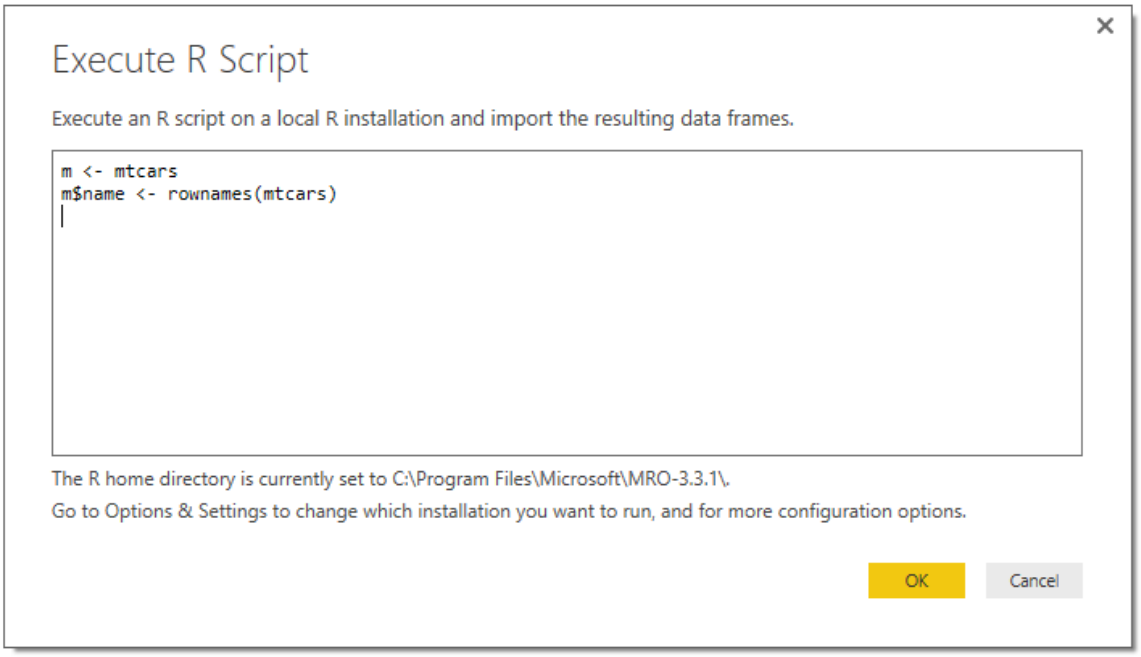


4. You should now see a window called **Execute R Script**. Input the following script to it.

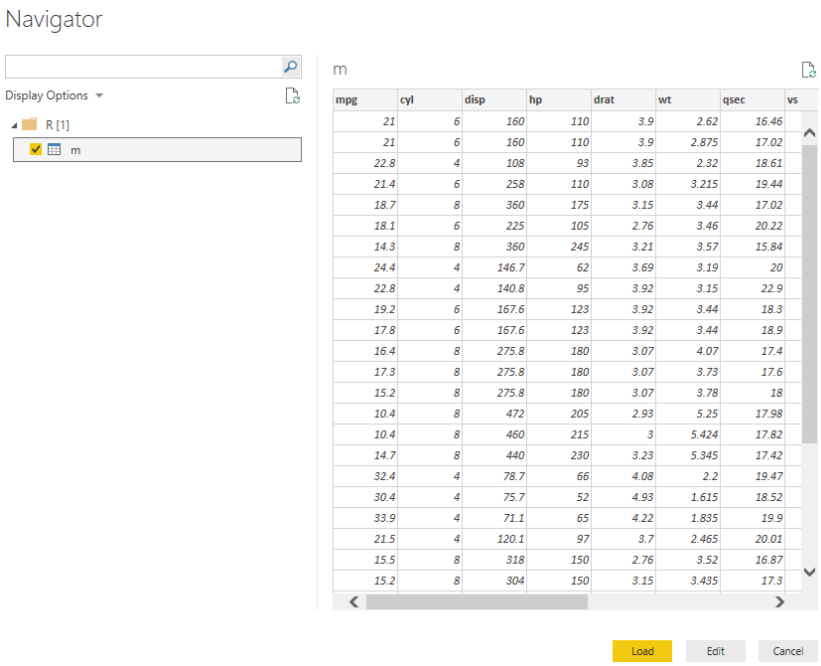
```
m <- mtcars
m$name <- rownames(mtcars)
```

Note: You can input any R script here, as long you have R installed on your machine and the referenced libraries are installed.

5. Your screen should be similar to the following screenshot. R is already installed on your virtual machine. Click **OK**.



6. In the navigator window, you should see **m** as an available dataset to use. Click the check box (left side of "m") to select it. After that, you will be able to see a preview of your data.



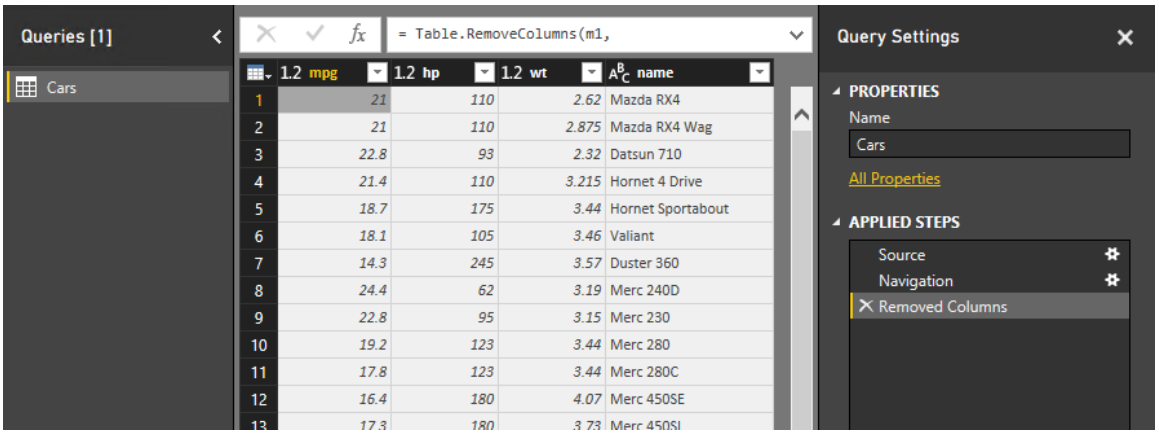
7. Click **Edit**.

8. In the **Query Settings** pane, change the name from **m** to **Cars**.

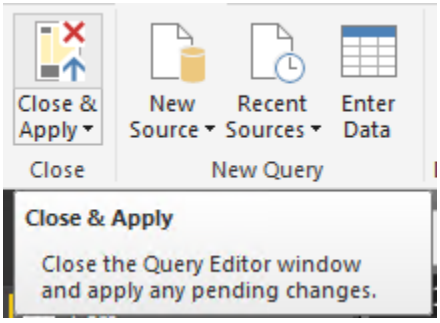
9. Remove the following columns from your query:

   cyl, **disp**, **drat**, **qsec**, **vs**, **am**, **gear**, and **carb**

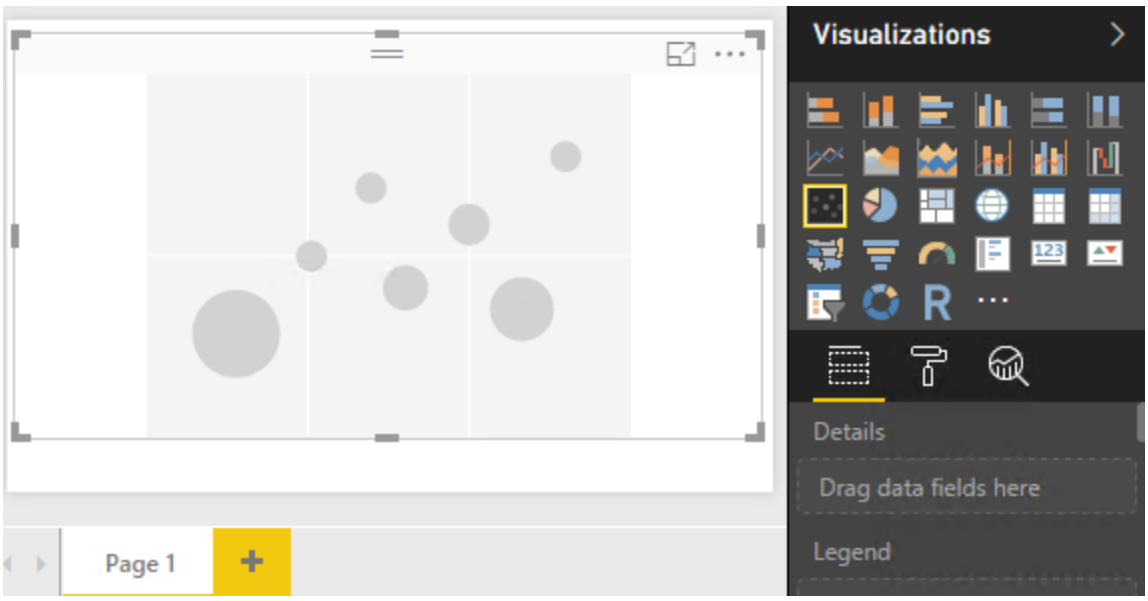
10. Your screen should be similar to the following screenshot:



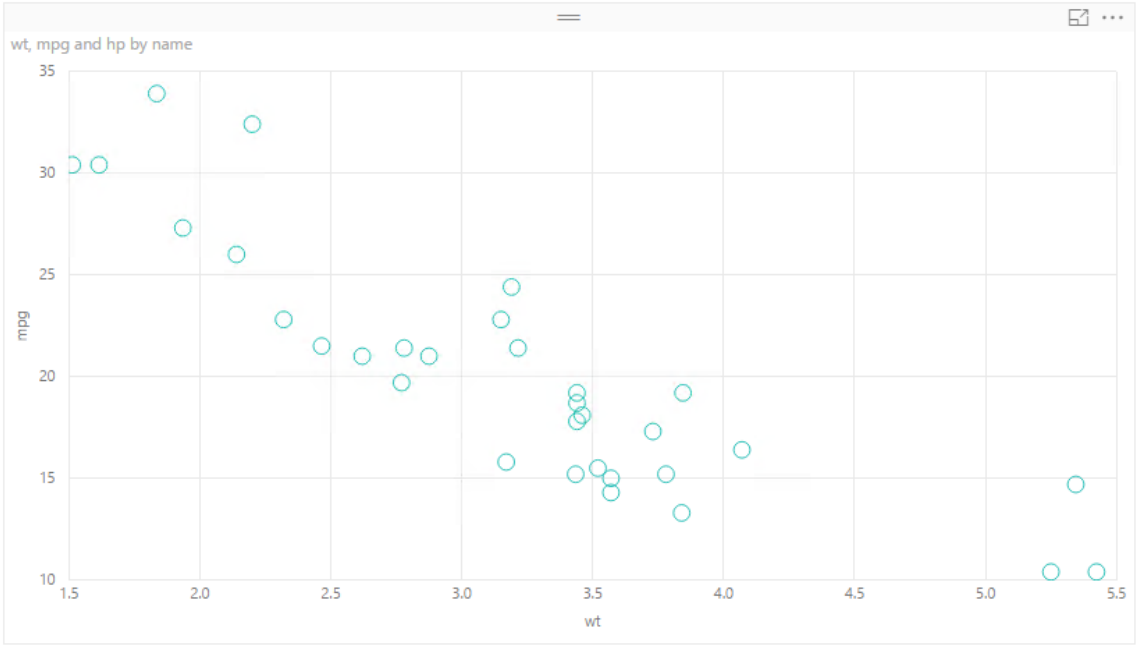
11. Click **Close & Apply**.



12. Add a **Scatter Chart** visual, and resize it to a bigger size (to occupy all the area of report).



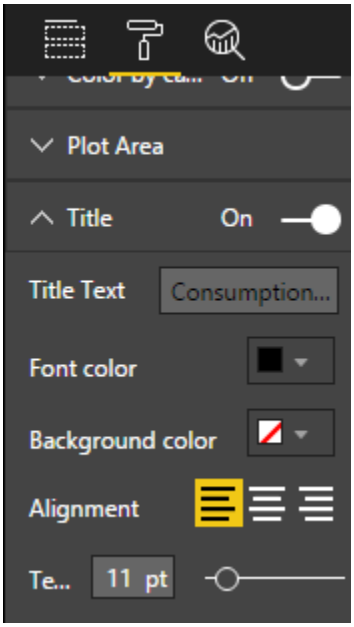
13. Add the field **name** to **Details**.
14. Add the field **wt** to **X Axis**.
15. Add the field **mpg** to **Y Axis**.
16. Add the field **hp** to **Tooltips**.
17. You should now have the following chart:



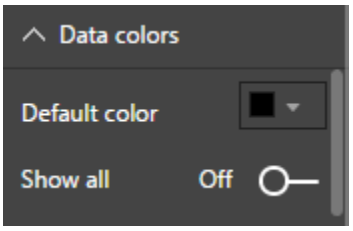
18. With the scatter chart selected, go to **Format** tab, expand **Title**, and change the **Title Text** property to **Consumption by weight**.

19. Change the **Font color** to **black**.

20. Change the **Text size** property to **11 pt**.

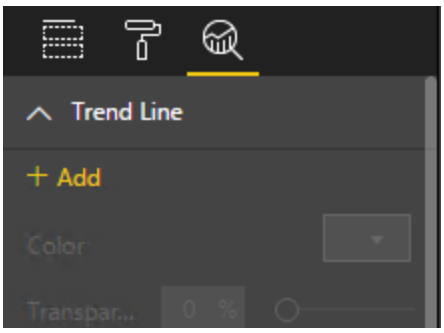


21. Still on **Format**, go to **Data colors**, and change the **Default color** to **black**,



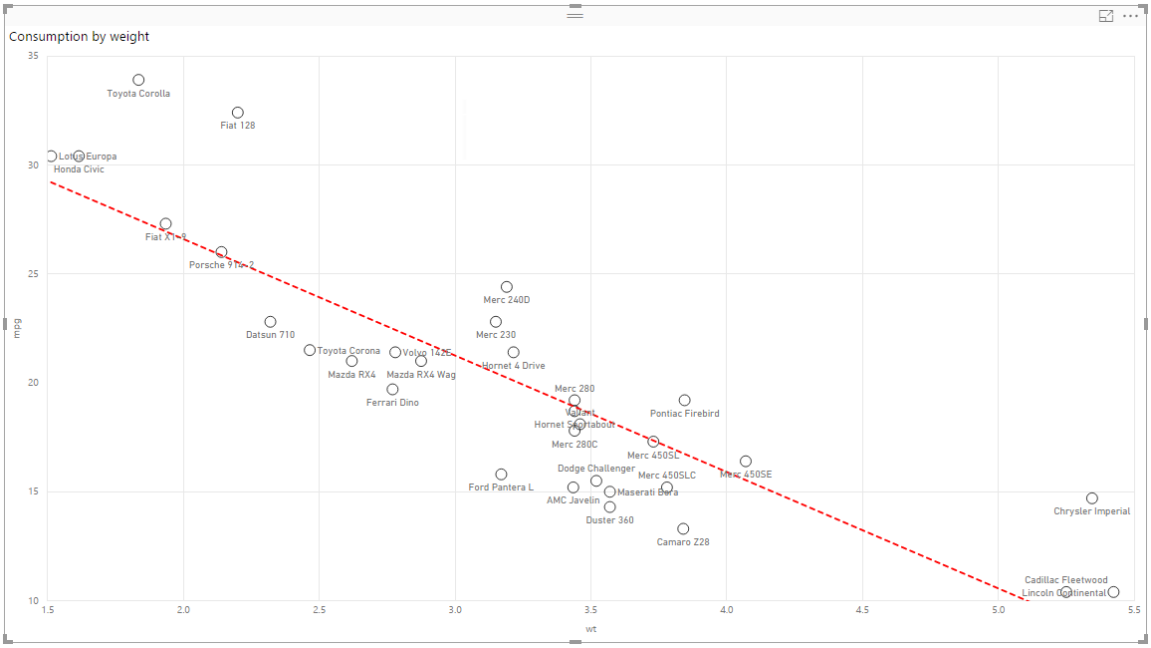
22. Locate **Category labels** and change it to **On**.

23. Select the **Analytics** pane (you must have the scatter chart selected). Search for **Trend Line** and click **+ Add**.



24. Change the color to **Red**.

25. Now you have a scatter chart that uses R to generate data, and is plotting a trend line (similar to a linear regression).



26. You can save the report. Create and use the **E:\Labs\M3Lab3** folder. Name the report as **RDataSource.pbix** and close **Power BI Desktop**.

Exercise 2: Using R Visuals

Introduction

In this exercise, you will use R Visual to show two charts generated by R code.

Objectives

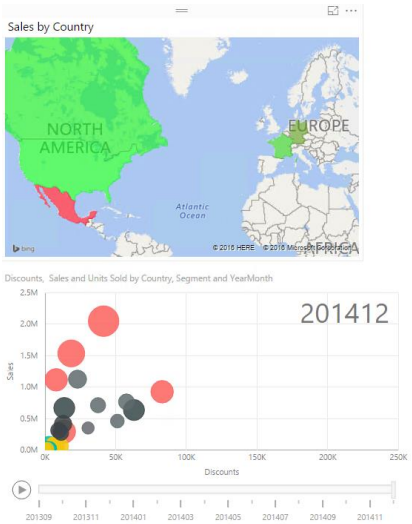
After completing this exercise, you will be able to:

- Show R plots directly into a Power BI Desktop file.

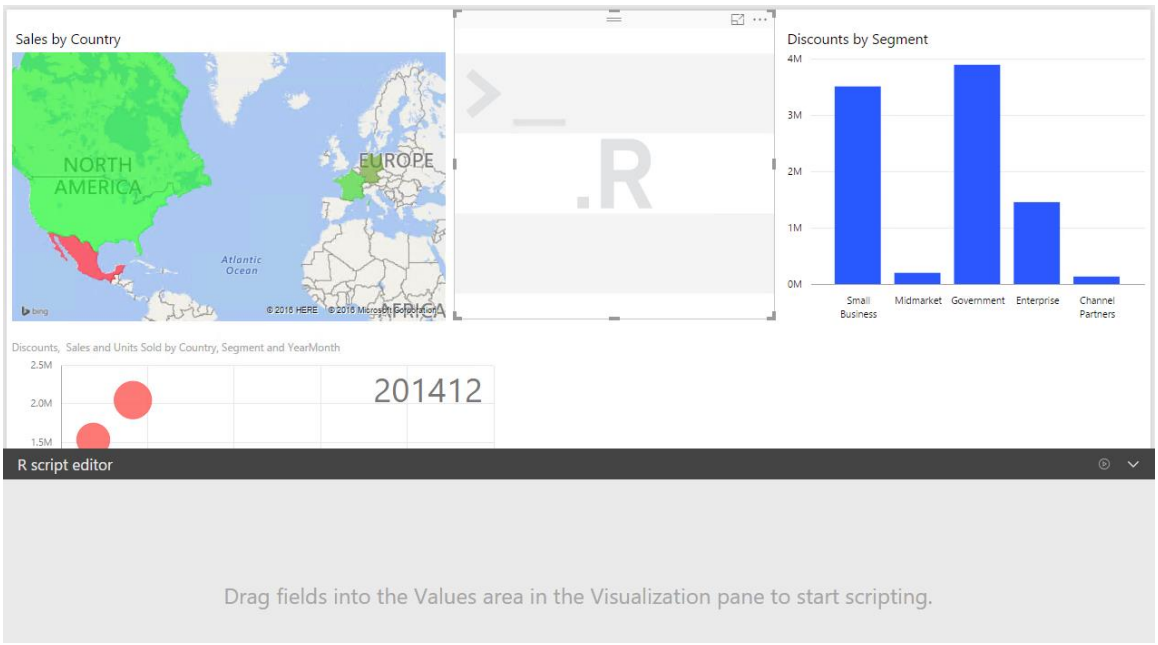
Using R Visuals

In this task, you will edit an existing Power BI Desktop file to add two new visualizations created by using R code.

1. Go to **E:\Assets\M3 – Lab 3.** Create a copy from **Lab3_RVisuals_Initial.pbix** in **E:\Labs\M3Lab3**.
2. Open the file you just copied. You should see the following report:



3. Locate the R visual in the **Visualizations** pane. Add it to your report. Position it between the map and bar chart.

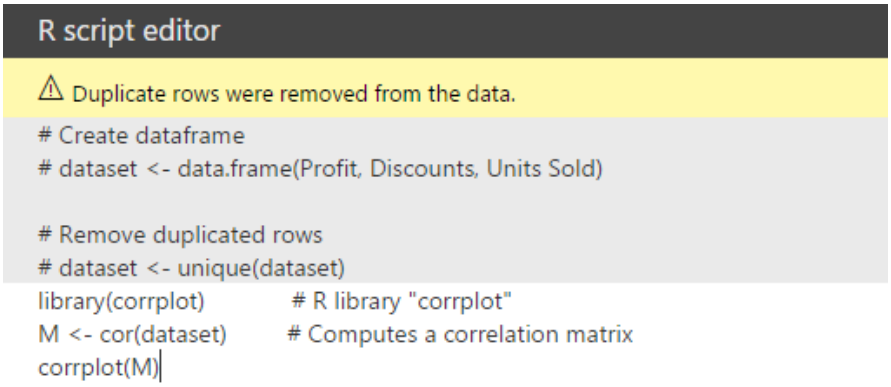


4. Drag the following fields to the **Values** property of the R visual. These values will compose the data frame available (called dataset) to you in the R script window:

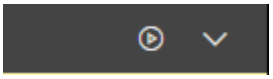
   Profit, **Discounts** and **Units Sold**.

5. Add the following R script to the **R script editor** window.

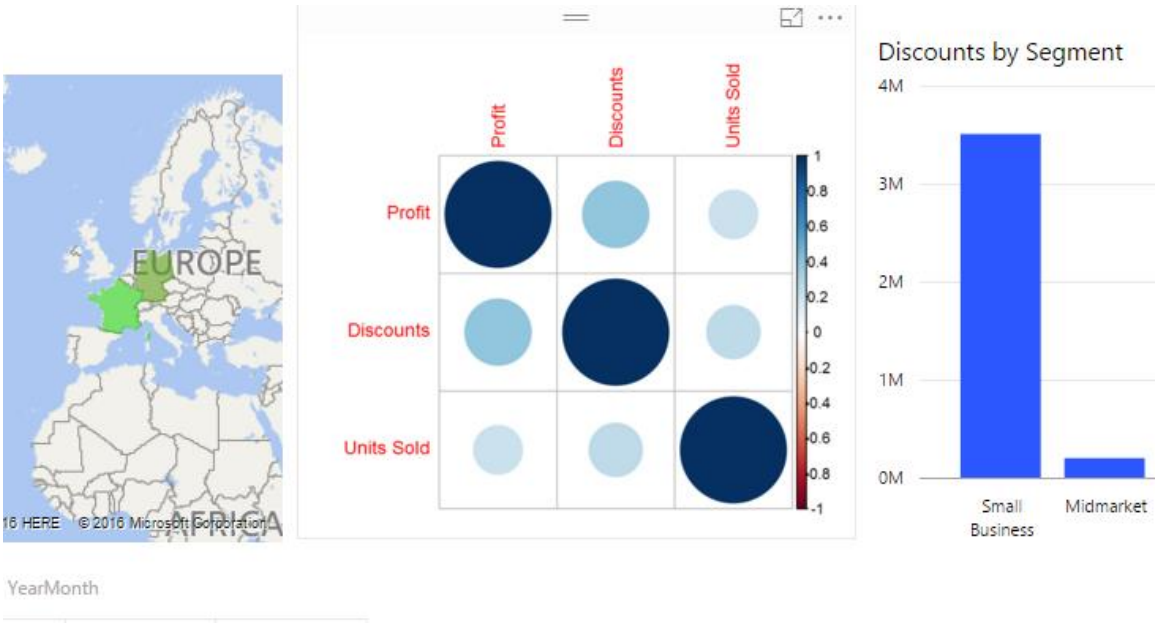
```
library(corrplot)          # R library "corrplot"
M <- cor(dataset)          # Computes a correlation matrix
corrplot(M)
```



6. Click the play icon.



7. You should see the following result:



8. Add another R visual to the region below the existing R visual and bar chart.

9. Drag the following fields to the **Values** property from the R visual you just added:

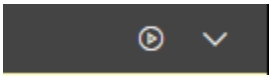
   Sales and **Segment**

10. Add the following code to the **R script editor** window.

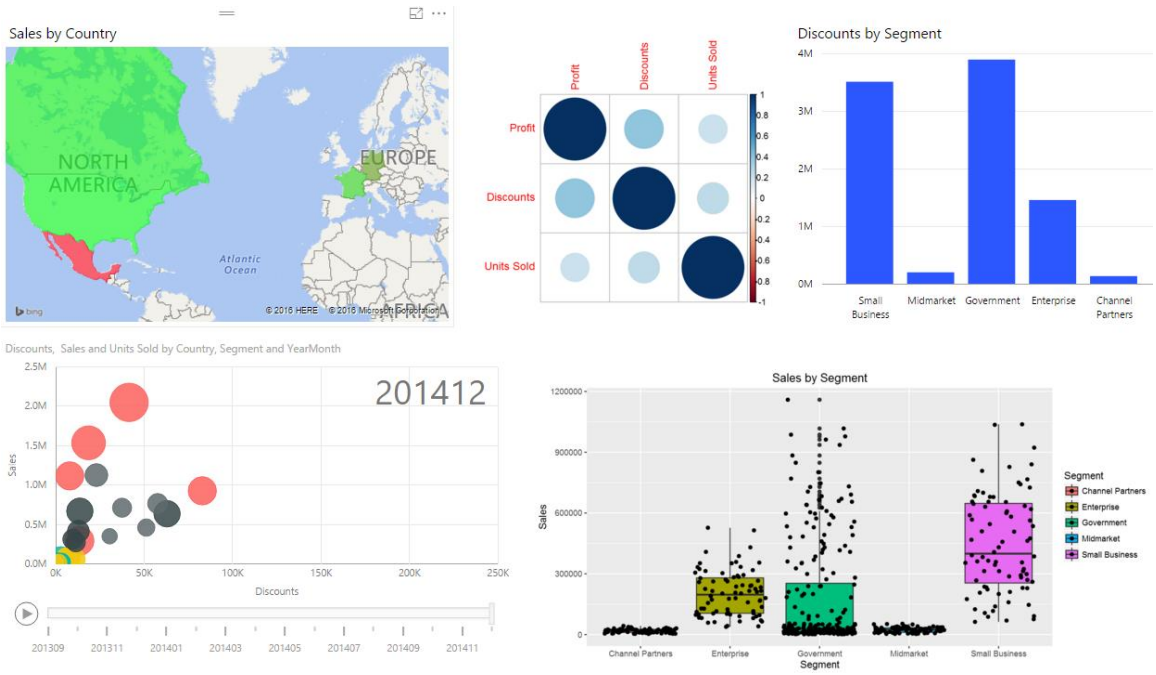
```
library(ggplot2)

qplot(Segment, Sales, data=dataset, geom=c("boxplot", "jitter"),
fill=Segment, main="Sales by Segment", xlab="Segment",
ylab="Sales")
```

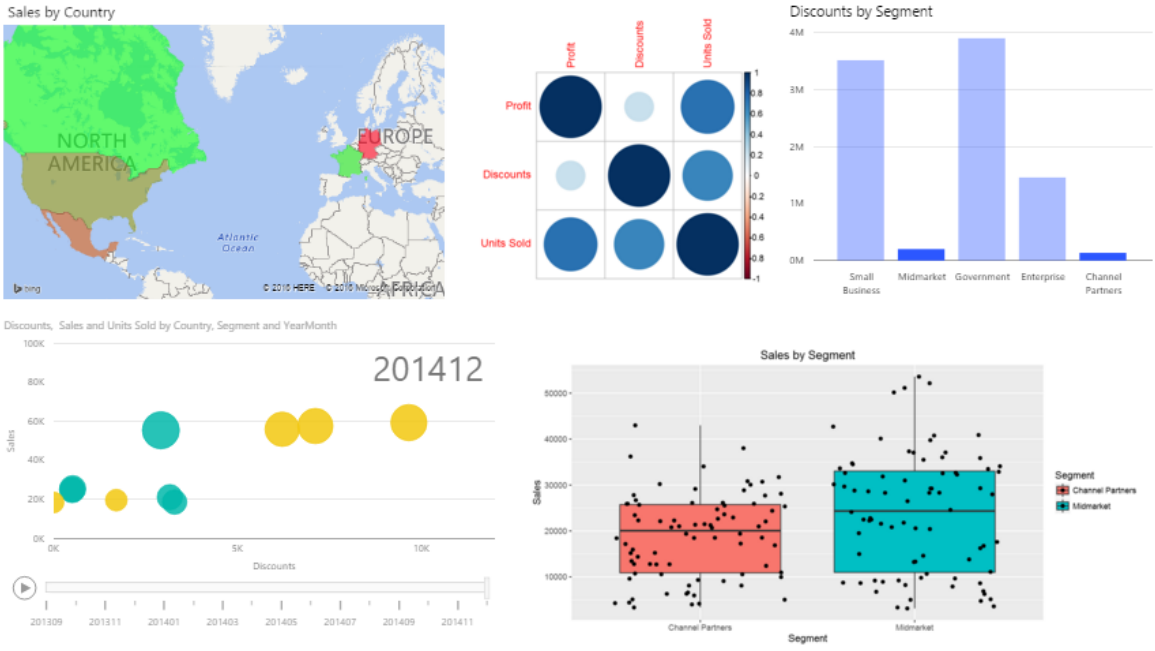
11. Click the play icon.



12. You should now have the following report, which uses regular Power BI charts with charts created by using R code.



13. Click the map and/or the others Power BI visuals to see that the R source is filtered and the visuals are updated to reflect the applied filter.



14. Save your report as **Lab3_R_Visials_Final.pbix** in the **E:\Labs\M3Lab3** directory.

15. Close **Power BI Desktop**.