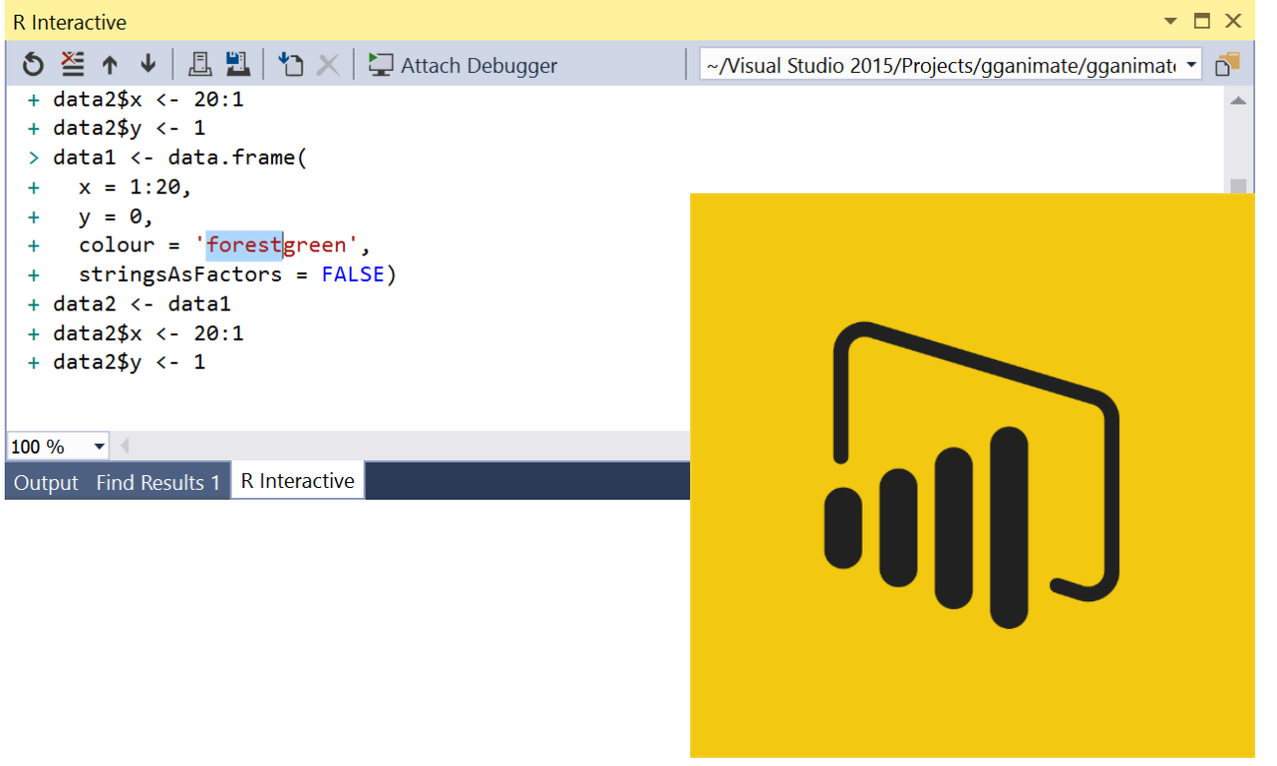
Microsoft

Business Analytics with Power BI

Microsoft Services

Module 3: Predictive Analytics with Power BI and R

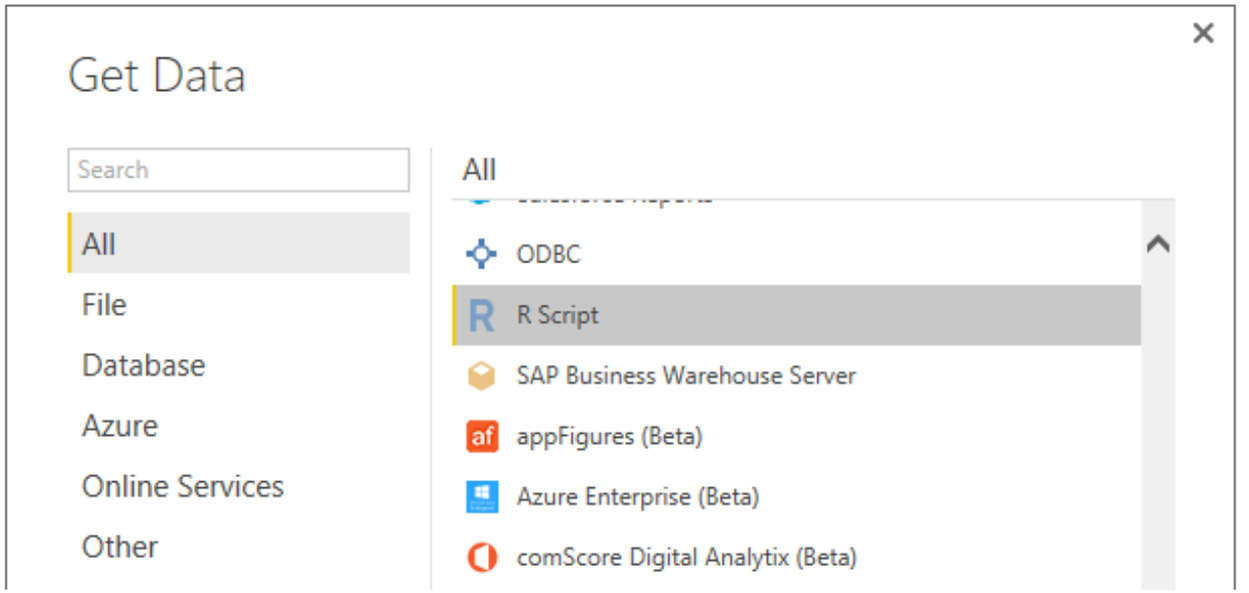
Lesson 4: Using R with Power BI



R as a Data Source

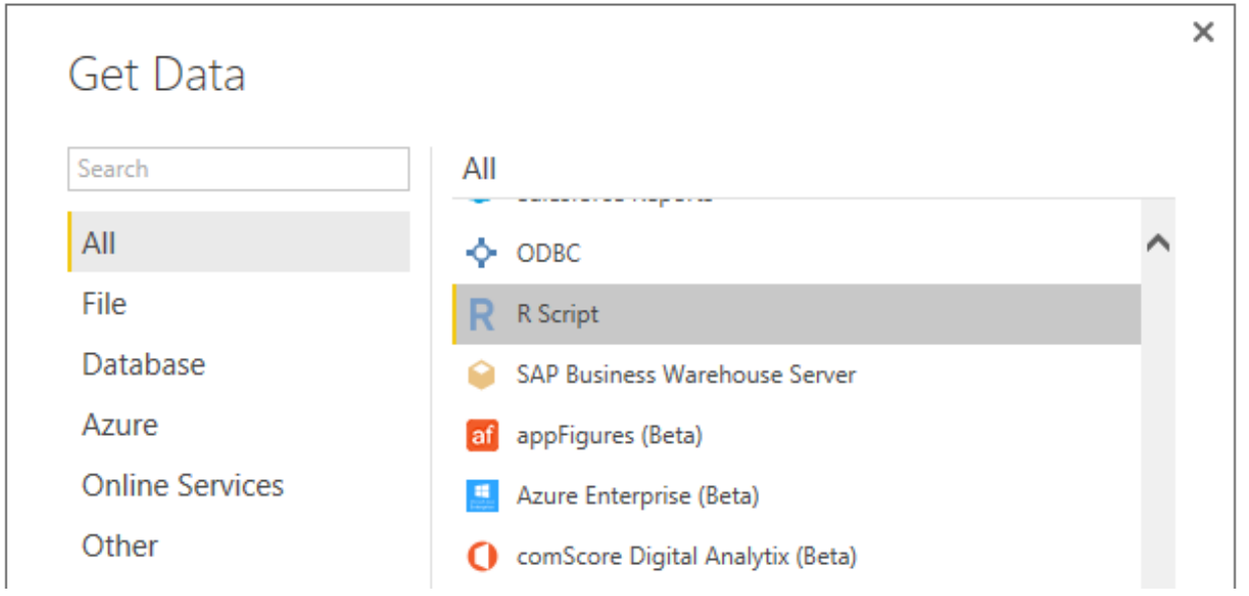
- Power BI Desktop does not include, deploy or install the R engine
- To run R scripts in Power BI Desktop, you must separately install R on your local computer

https://msdn.microsoft.com/en-us/microsoft-r/
https://cran.r-project.org/

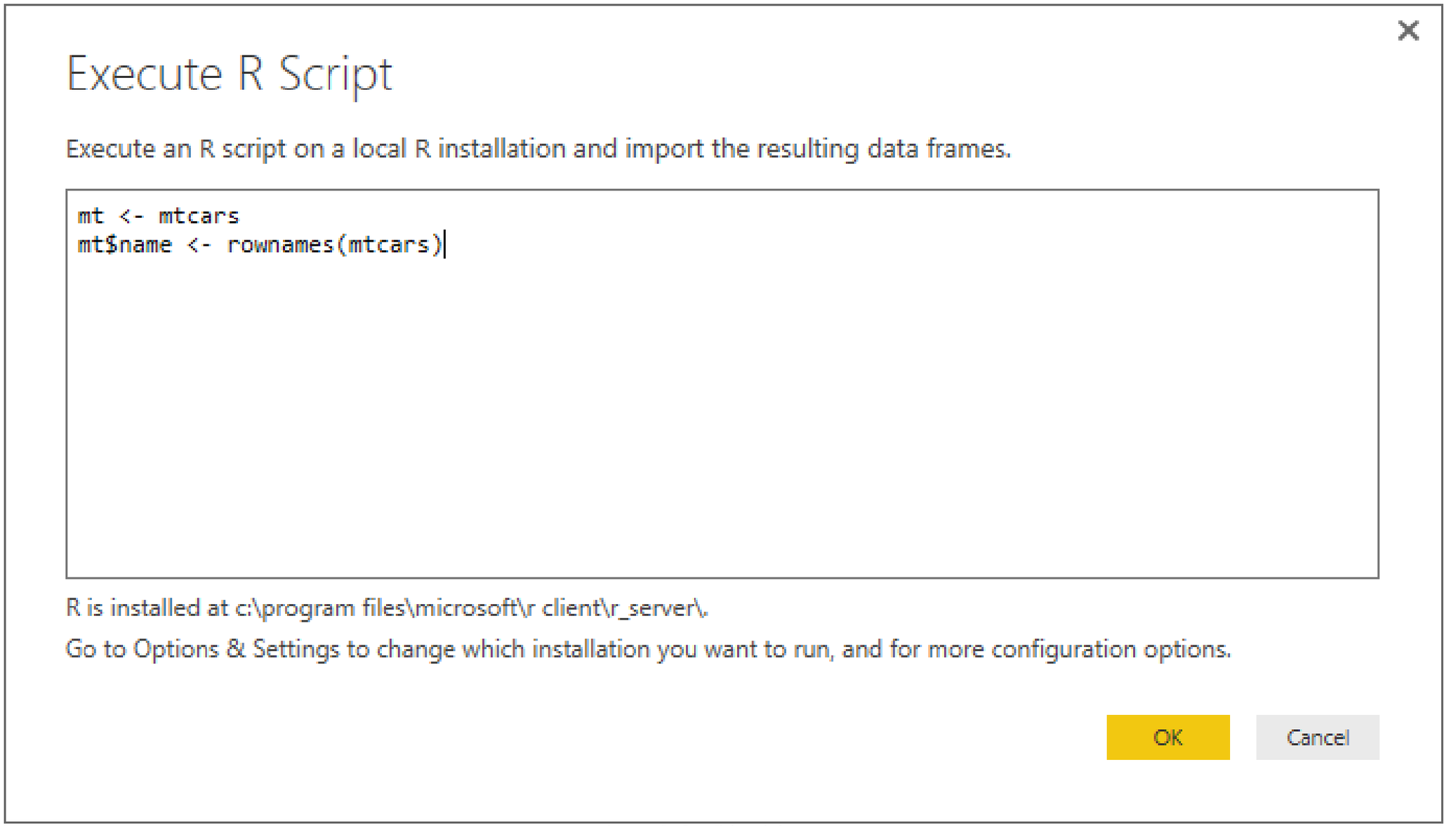


R as a Data Source

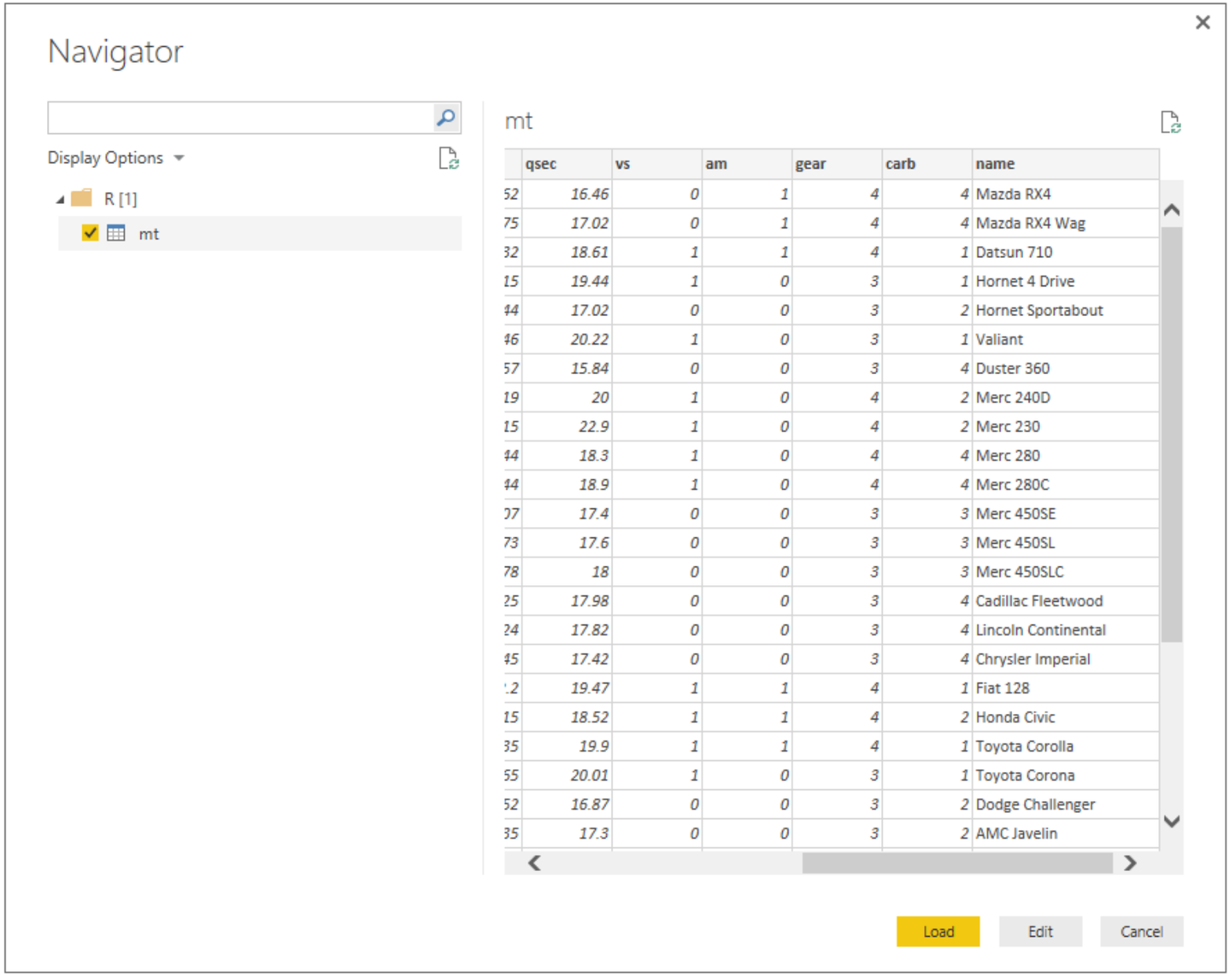
- Power BI has R as a data source. It receives a R script as a “query”
- The R script must return a data frame
- This data frame will be seen by Power BI as a dataset (like any other)
- The R engine that process the script must be installed on the same machine that has Power BI Desktop

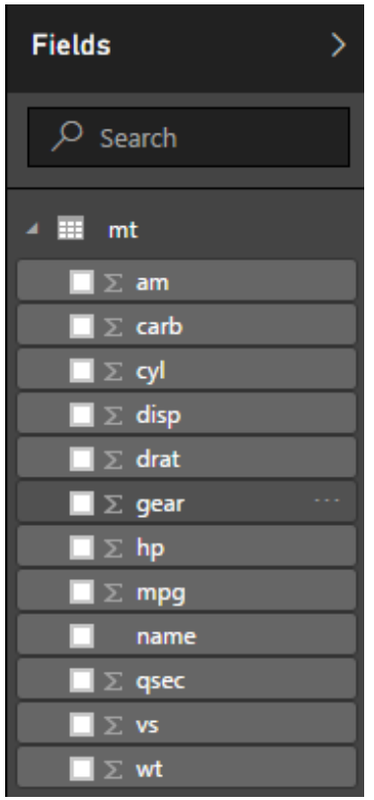


R as a Data Source

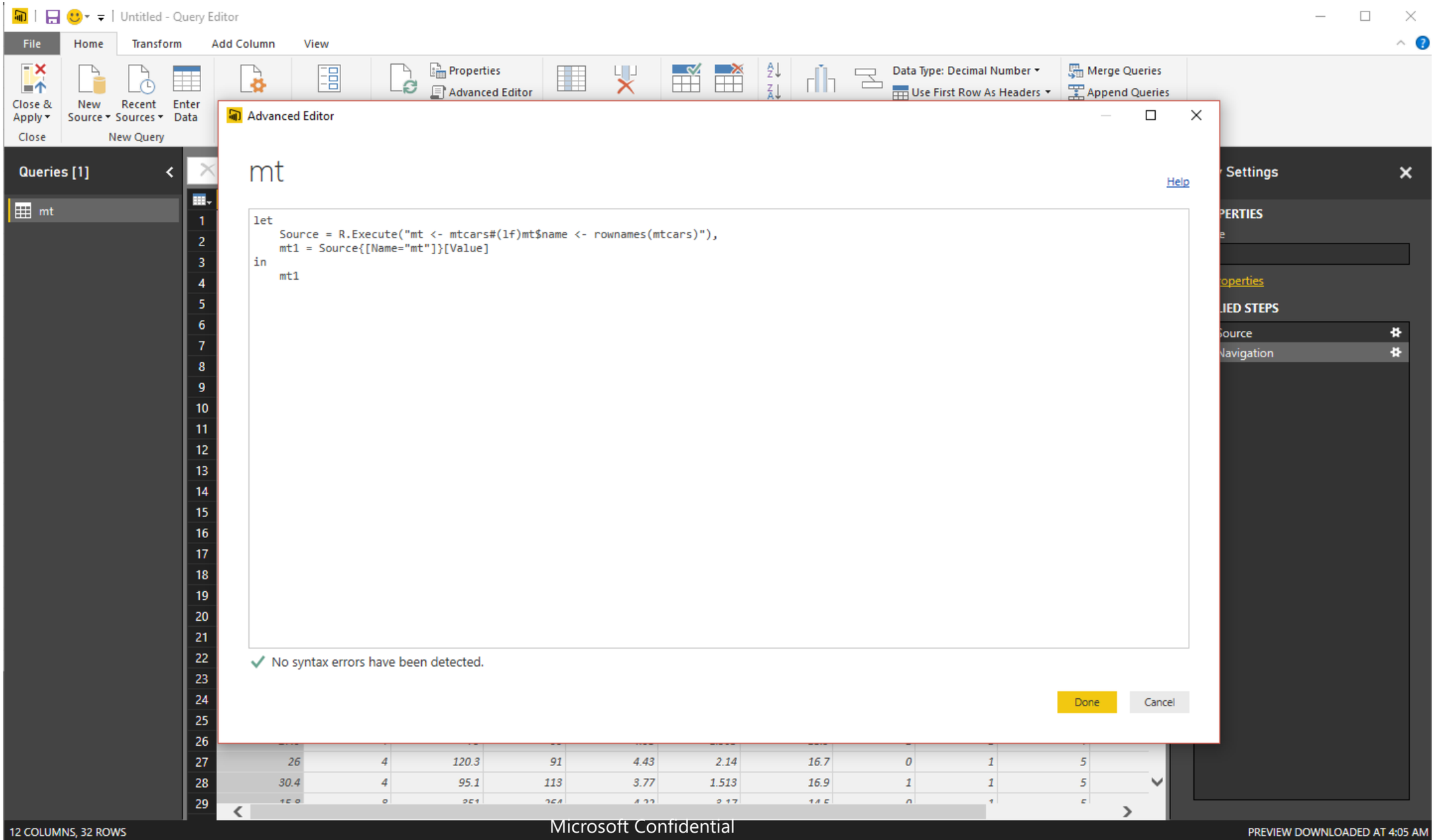


R as a Data Source

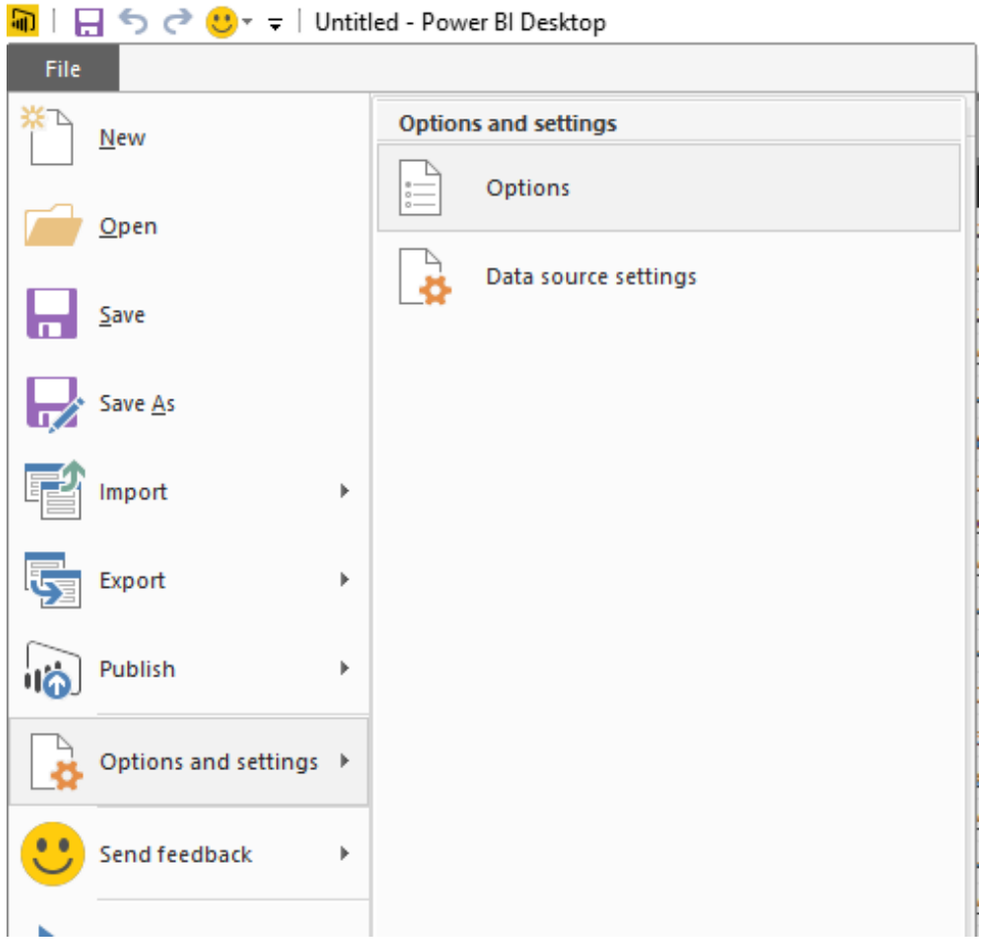


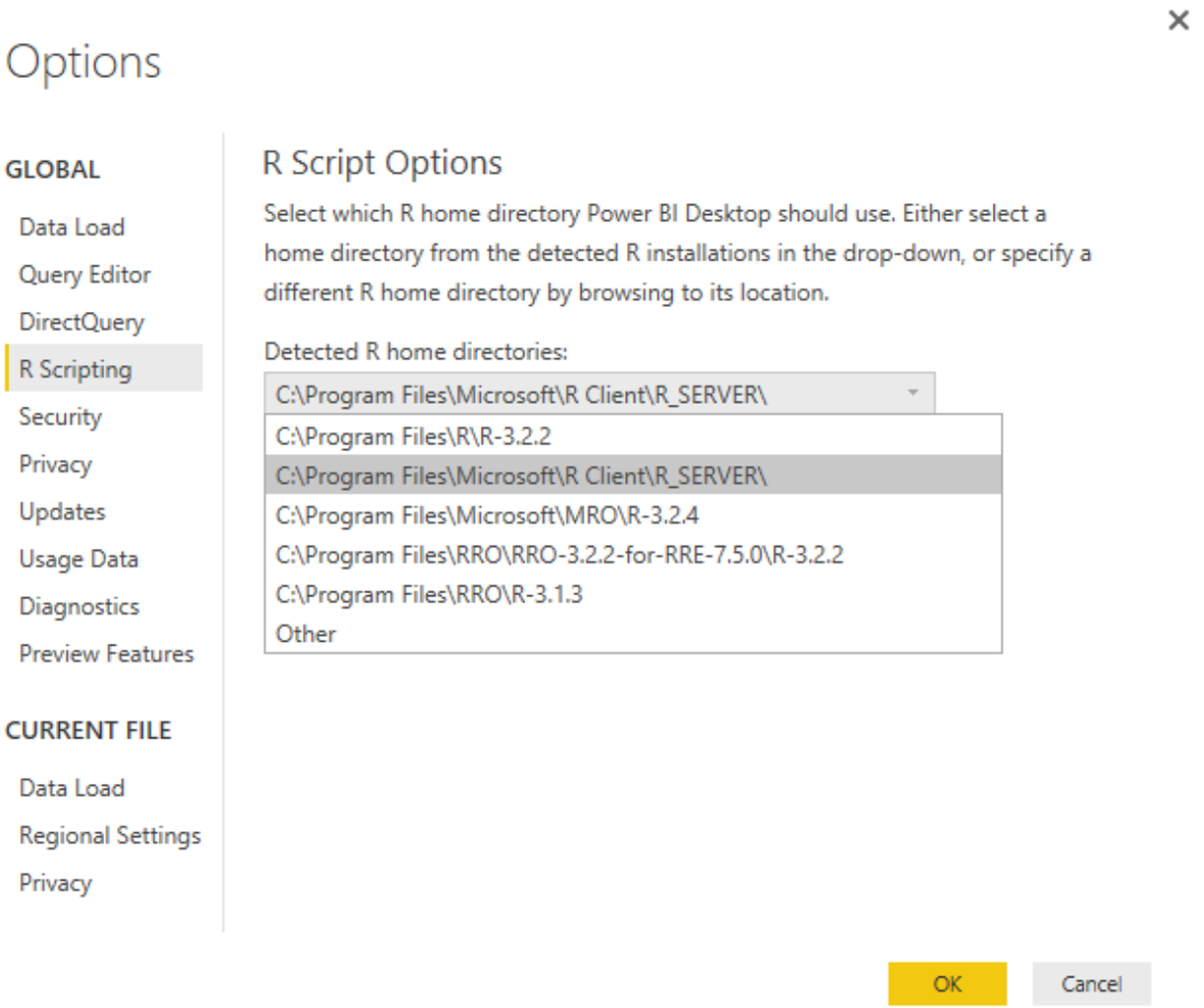


R as a Data Source



Changing R Location / Installation





Demonstration: R as a Data Source

R as a Data Source

Lab: Using R with Power BI

Exercise 01 – Using R as a Data Source

R Visuals

We can use a script in R to generate a chart in a Power BI report

This chart will be connected (by default) with other graphical elements

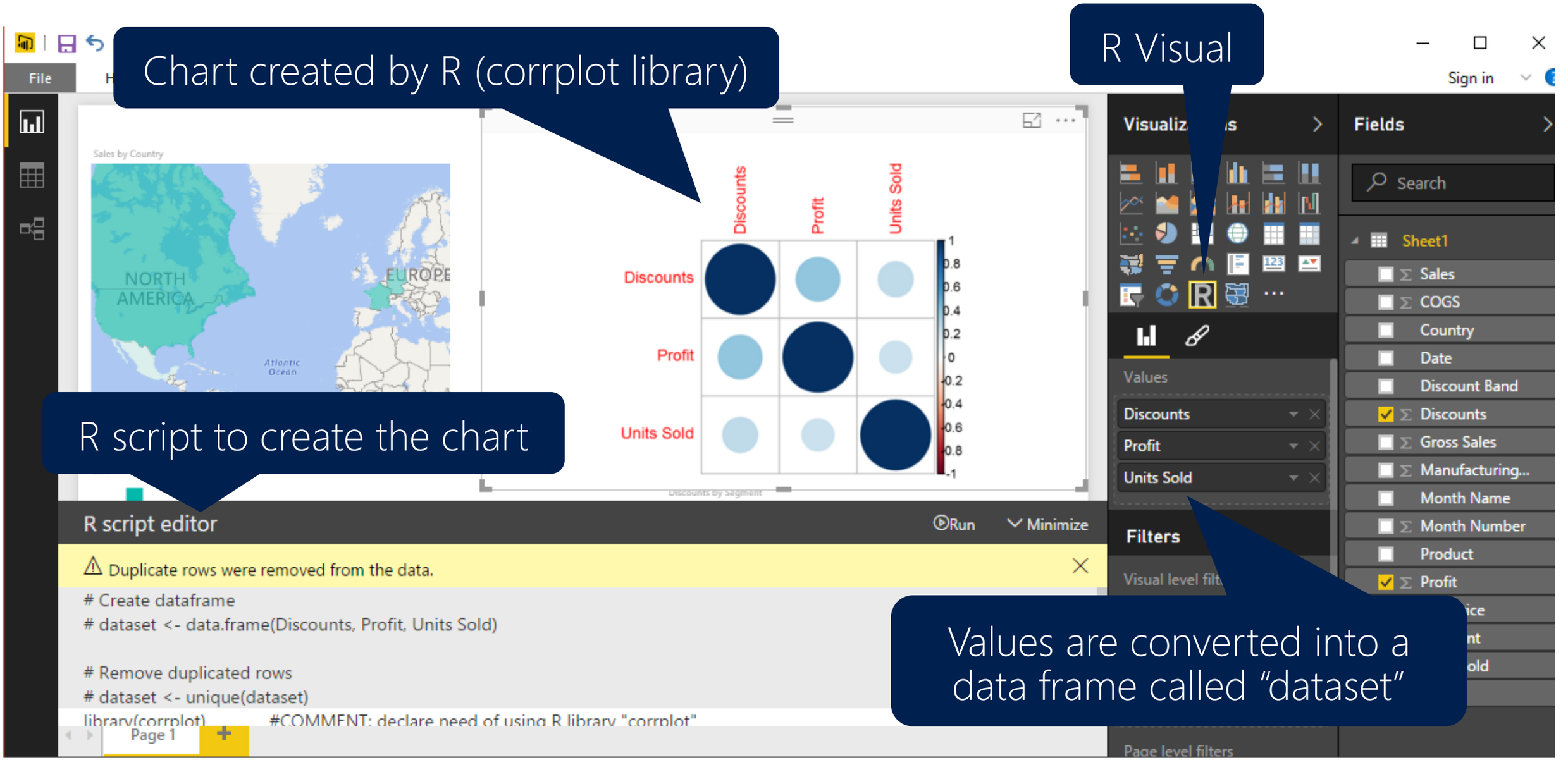
You can also see it in browser (Power BI Service) and mobile apps

After the report is published, the R visual is processed by a R service inside the Power BI Service

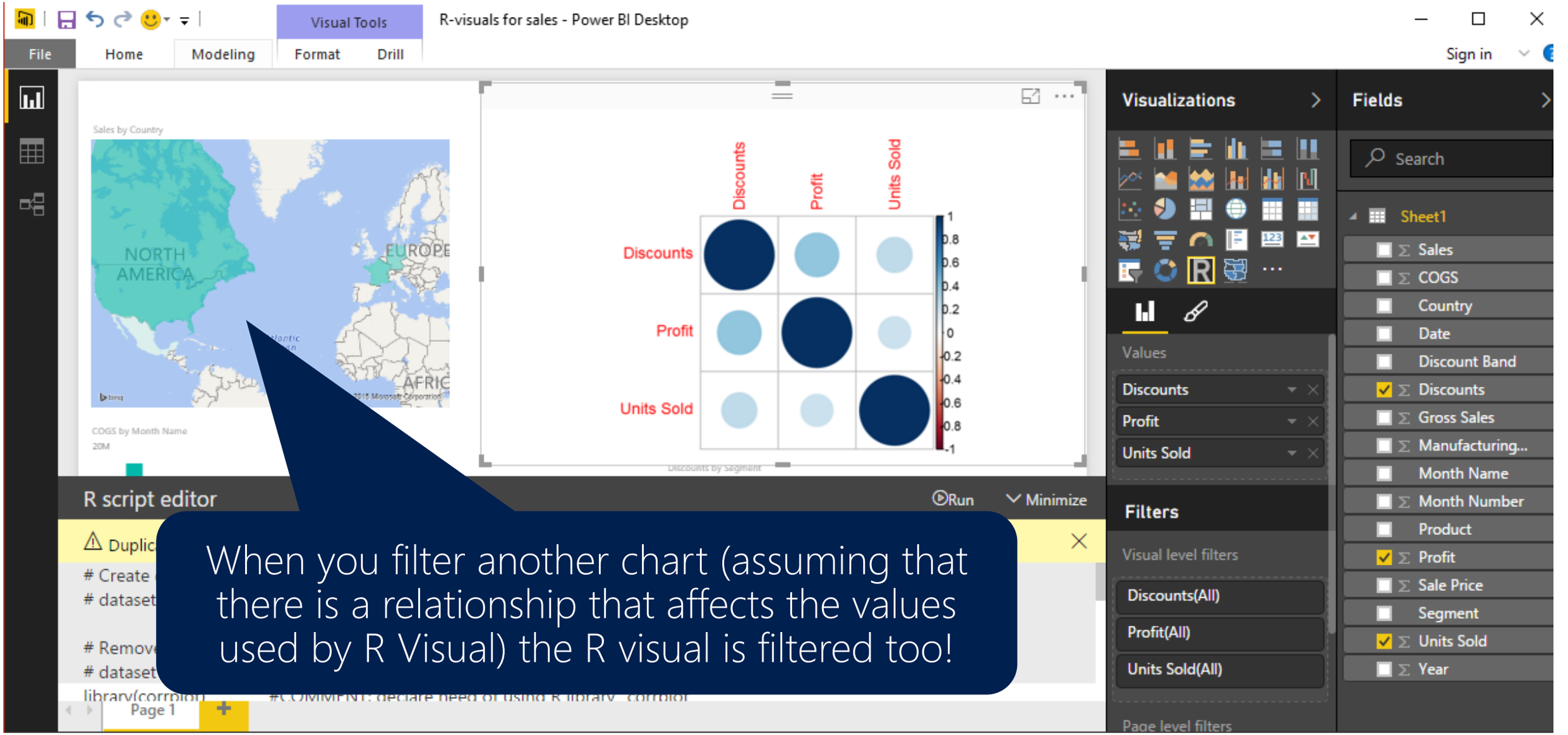
You must verify if the libraries being used are available in the service

Libraries that need internet connection will not work

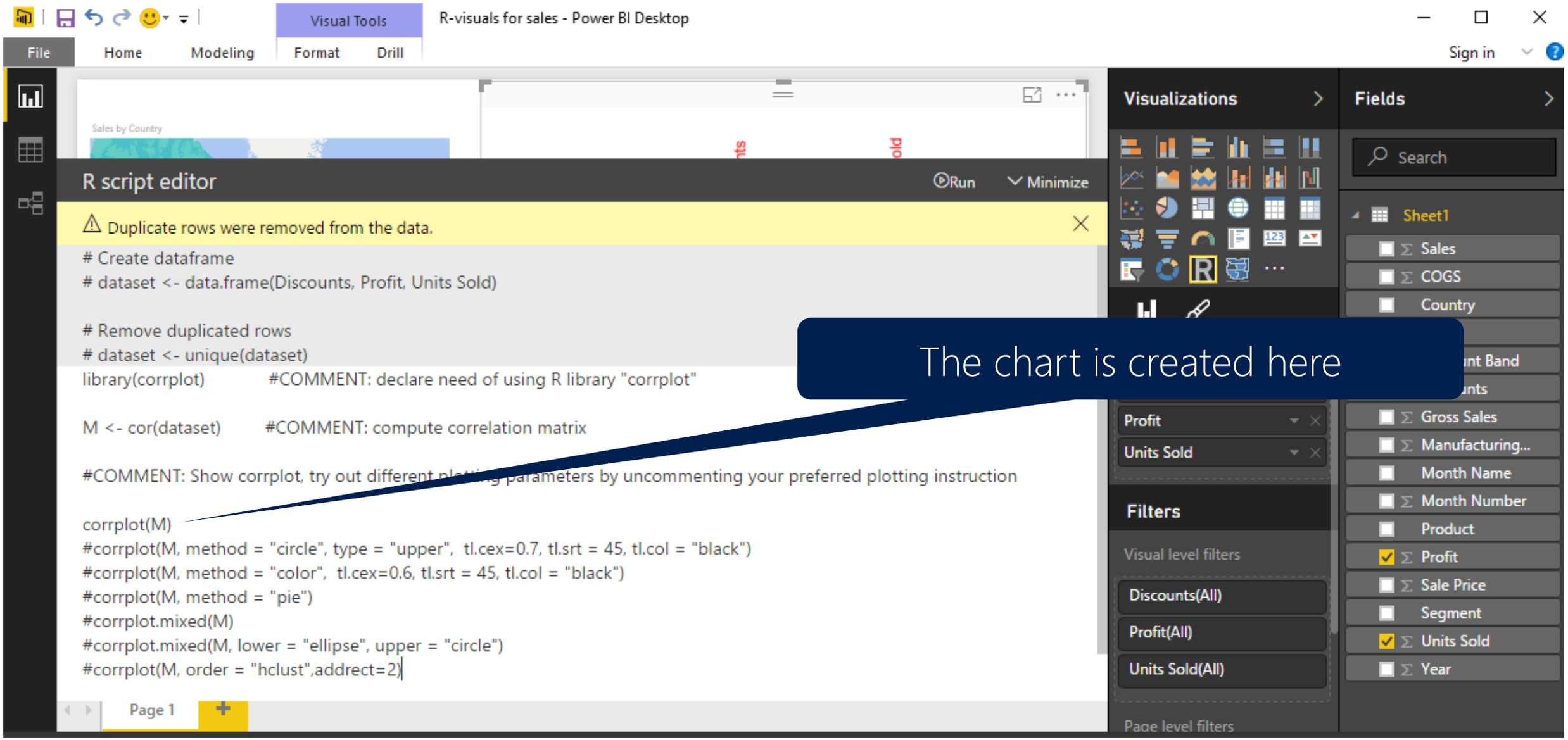
R Visuals



R Visuals



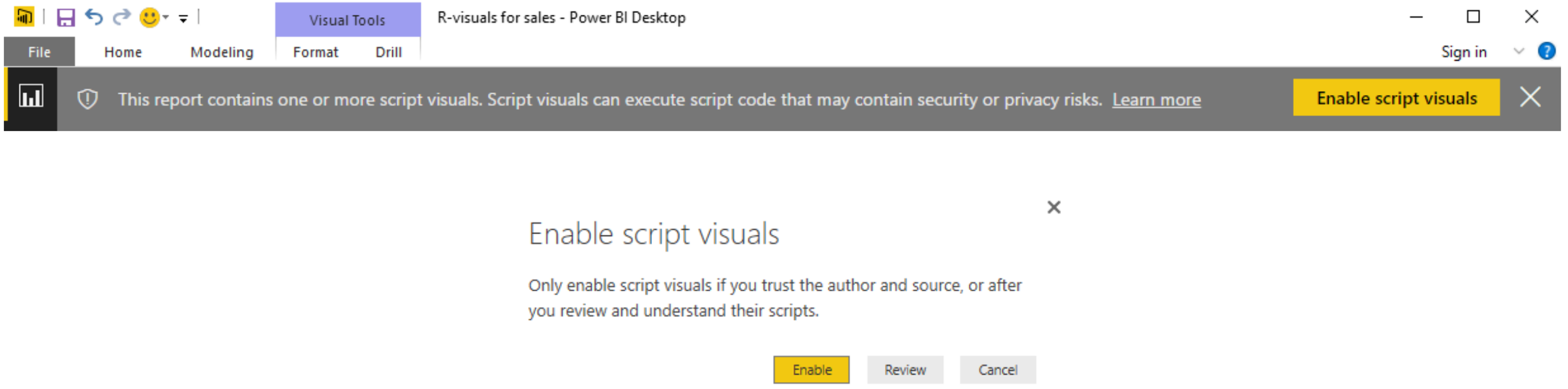
R Visuals



R Visuals

When you open a Power BI Desktop file that contains at least one R Visual, you are prompted to confirm the execution of script visuals

Since scripts can have malicious code, you must confirm the execution



R Visuals

Limitations

- Data size limitations – data used by the R visual for plotting is limited to 150,000 rows. If more than 150,000 rows are selected, only the top 150,000 rows are used and a message is displayed on the image

- Calculation time limitation – if an R visual calculation exceeds 5 minutes the execution times out, resulting in an error

- Relationships – as with other Power BI Desktop visuals, if data fields from different tables with no defined relationship between them are selected, an error occurs

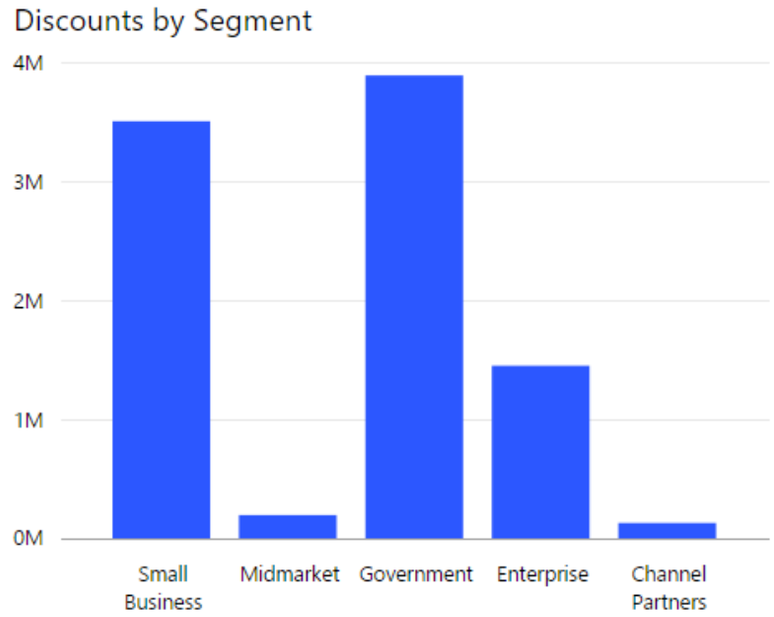
- R visuals are refreshed upon data updates, filtering, and highlighting. However, the image itself is not interactive and cannot be the source of cross-filtering

R Visuals

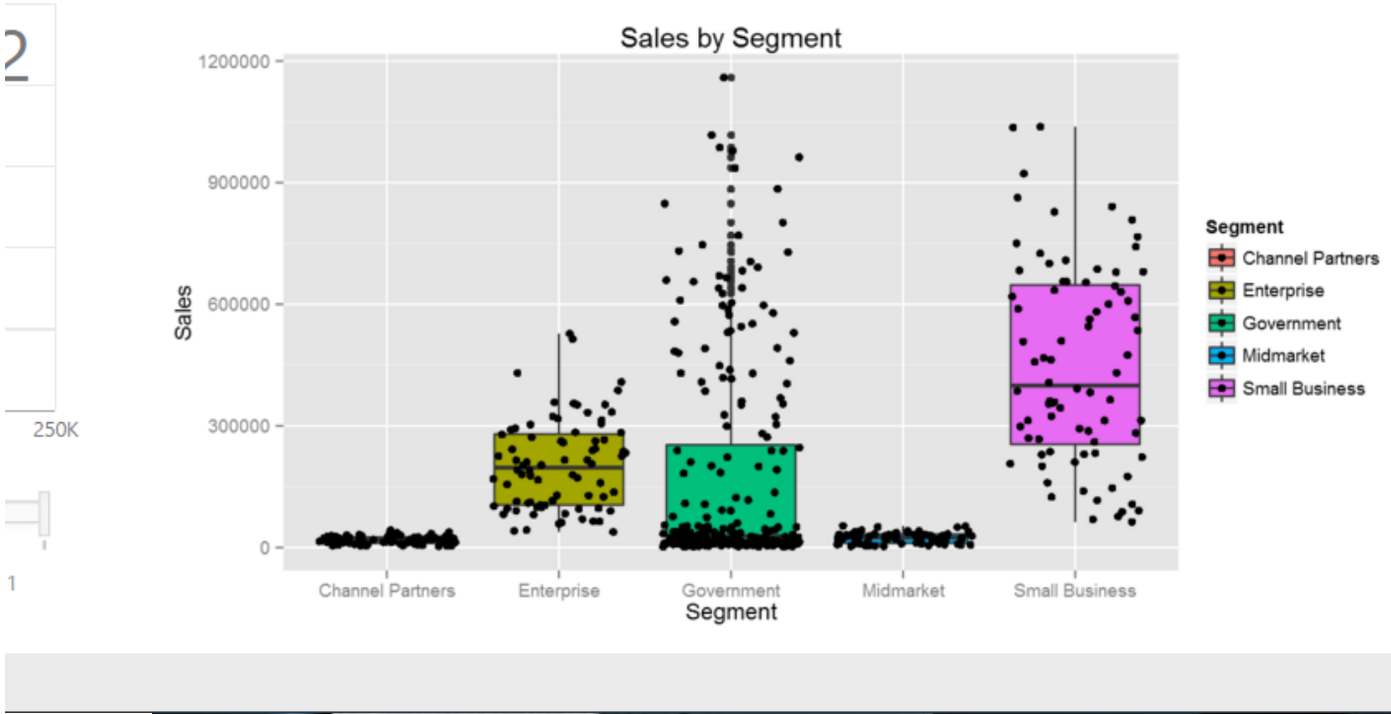
Limitations (cont..)

- R visuals respond to highlighting other visuals, but you cannot click on elements in the R visual in order to cross filter other elements
- Only plots that are plotted to the R default display device are displayed correctly on the canvas. Avoid explicitly using a different R display device

Demonstration: R Visuals



R Visuals



Refreshing R Data

For now, the only supported method to refresh R data (manually or automatically) is using Personal Gateway

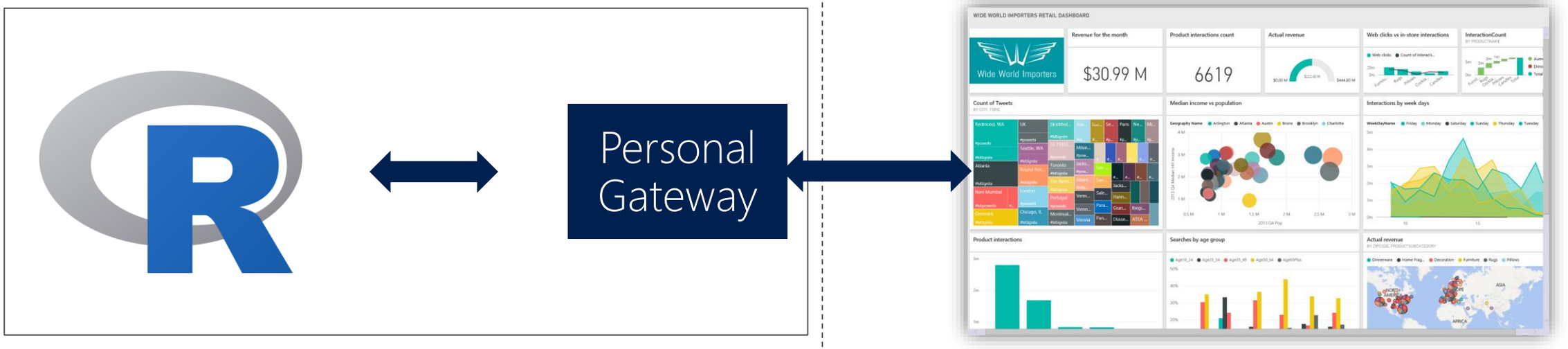
If you don't have admin permissions, the machine must be on and user **logged in**

	With Administrator permissions	Without Administrator permissions
Power BI Gateway - Personal runs as a	Service	Application
Scheduled Refresh	As long as your computer and the gateway service is running, you do not have to be logged in at the scheduled refresh time.	You must be logged in to your computer at the scheduled refresh time.

Refreshing R Data

For now, the only supported method to refresh R data (manually or automatically) is through Power BI Personal Gateway.

Notice that this is required *only when using R as a data source*. R Visuals are rendered entirely on server.



Customer Network

Internet (Power BI)

Lab: Using R with Power BI

Exercise 02 – Using R Visuals