

Using R in Power BI

Sudarshana A

August 11, 2020

Contents

Prerequisite	1
R Installation	1
R IDE	1
Power BI Options	2
Import Data	2
Example 1 - Connect to available dataset in R	3
Example 2 - import csv, xlsx, sql etc. to R / transform	7
Tranform Data	8
Visualisations	10

Prerequisite

R Installation

- R must be installed on the same Windows computer as Power BI Desktop

<https://www.r-project.org/>

Australia - mirrors

https://cran.csiro.au/	CSIRO
https://mirror.aarnet.edu.au/pub/CRAN/	AARNET
https://cran.ms.unimelb.edu.au/	School of Mathematics and Statistics, University of Melbourne
https://cran.curtin.edu.au/	Curtin University

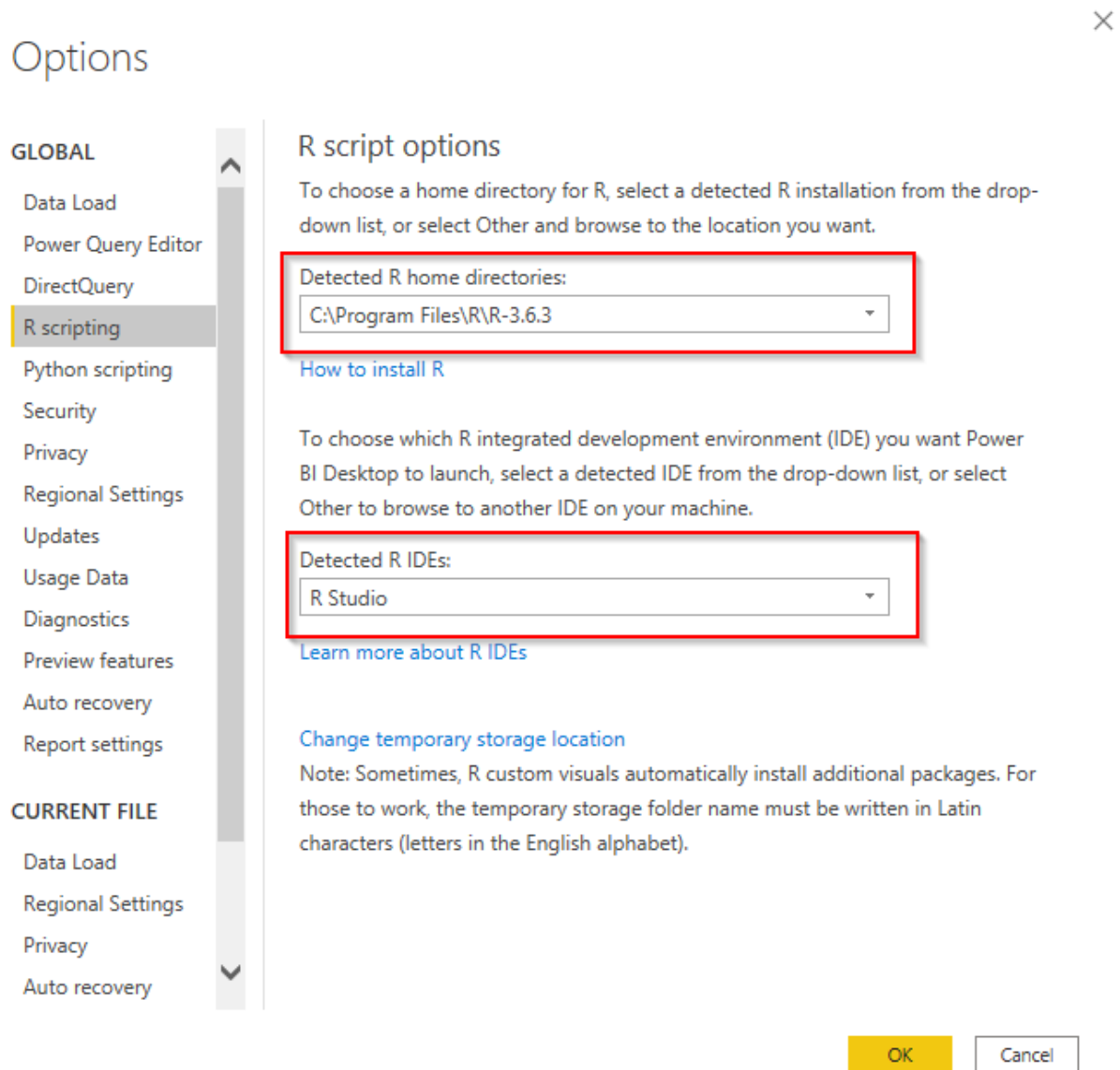
R IDE

- Good to have

IDE to use with R, Notepad++, R Studio, VS etc.

Power BI Options

We can verify correct R installation and detected IDEs by opening *Options* dialog box. Also using this we can provide a specific R installation and R IDE to use with Power BI.



Import Data

To import data you need to create a dataframe. Power BI can access the dataframes and you can select required dataframes to import.

- Connect to available dataset in R
- import csv, xlsx, sql etc. data to R then connect
- import data to R, transform, clean up data and then connect

Example 1 - Connect to available dataset in R

We can use available dataset in R and import this to Power BI
cars dataset in R consist of breaking distances vs. speed data.
we can preview the data using `head()` function.

```
head(cars)
```

```
##    speed dist
## 1      4     2
## 2      4    10
## 3      7     4
## 4      7    22
## 5      8    16
## 6      9    10
```

Lets import this dataset to Power BI by assigning to a dataframe.

In **Power BI**, click *Get Data -> Other -> R script*

Get Data



- All
- File
- Database
- Power Platform
- Azure
- Online Services
- Other

Other

- Web
- SharePoint list
- OData Feed
- Active Directory
- Microsoft Exchange
- Hadoop File (HDFS)
- Spark
- Hive LLAP (Beta)
- R script
- Python script
- ODBC
- OLE DB
- Acterys : Model Automation & Planning (Beta)
- Automation Anywhere (Beta)
- Solver
- Cherwell (Beta)

[Certified Connectors](#)

Connect

Cancel

Click Connect

Use the following R script to get the data from *cars* dataset.

```
cars_data <- cars
```

R script



Script

```
cars_data <- cars
```

The script will run with the following R installation C:\Program Files\R\R-3.6.3.

To configure your settings and change which R installation you want to run, go to Options and settings.

OK

Cancel


Click OK

We can select required dataframe to Load/Tranform in Power BI

Navigator

Display Options ▾

▲  R [1]

☒  cars_data

cars_data

speed	dist
4	2
4	10
7	4
7	22
8	16
9	10
10	18
10	26
10	34
11	17
11	28
12	14
12	20
12	24
12	28
13	26
13	34
13	34
13	46
14	26
14	36
14	60
14	80
15	20

Load

Transform Data

Cancel

Queries [1] <

cars_data

fx = Source([Name="cars_data"])[Value]

	1.2 speed	1.2 dist
1	4	2
2	4	10
3	7	4
4	7	22
5	8	16
6	9	10
7	10	18
8	10	26
9	10	34
10	11	17
11	11	28
12	12	14
13	12	20
14	12	24
15	12	28
16	13	26
17	13	34
18	12	24

2 COLUMNS, 50 ROWS Column profiling based on top 1000 rows

Query Settings

PROPERTIES

Name

cars_data

All Properties

APPLIED STEPS

Source

X Navigation

PREVIEW DOWNLOADED AT 3:48 PM

Example 2 - import csv, xlsx, sql etc. to R / transform

```
library(readxl)

CompletedWORwA <- read_excel("C:/<directory path>/CompletedWORwA.xlsx")

HFC_NSW <- subset(CompletedWORwA, TECHNOLOGY == 'HFC' & STATE == 'NSW')
```

Power BI Navigator

Navigator

Display Options ▾

▲

R [2]

☒ ☐ CompletedWORwA

☒ ☐ HFC_NSW

HFC_NSW

WORKORDERNO	ACTIVITYID	TECHNOLOGY	WORKORI
BSA100002604551-1	WOR700078998916	HFC	REMEI
BSA100002414188	WOR700071899856	HFC	HFCCN
BSA100002415859	WOR700071964242	HFC	HFCCN
BSA100002432671	WOR700072594067	HFC	HFCCN
BSA100002439329-1	WOR700072839451	HFC	HFCCN
BSA100002458866	WOR700073558516	HFC	HFCCN
BSA100002489478	WOR700074601158	HFC	HFCCN
BSA100002521493	WOR700075737337	HFC	HFCCN
BSA100002495090-2	WOR700075983700	HFC	HFCCN
BSA100002492936-1	WOR700076060192	HFC	HFCCN
BSA100002491779-2	WOR700076203633	HFC	HFCCN
BSA100002528070-2	WOR700076435180	HFC	HFCCN
BSA100002543872	WOR700076505639	HFC	HFCCN
BSA100002546758	WOR700076601643	HFC	HFCCN
BSA100002557842	WOR700076957974	HFC	HFCCN
BSA100002560601	WOR700077049867	HFC	HFCCN
BSA100002540295-2	WOR700077207188	HFC	HFCCN
BSA100002517060-2	WOR700077246585	HFC	HFCCN
BSA100002567886	WOR700077293682	HFC	HFCCN
BSA100002568138	WOR700077302392	HFC	HFCCN
BSA100002568144	WOR700077302578	HFC	HFCCN
BSA100002550578-1	WOR700077325522	HFC	HFCCN
BSA100002569882	WOR700077360777	HFC	HFCCN

Load

Transform Data

Cancel

Tranform Data

We can transform Power BI data using R scripts. Power BI Query Editor can apply R Scripts to a dataset in order to transform the data.

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Data Type: Text ▾
Detect Data Type
Rename

Any Column

ABC 123
Text Column ▾

Statistics ▾
Standard
Scientific

Trigonometry ▾
Rounding ▾
Information ▾

Number Column

Date ▾
Time ▾
Duration ▾

Date & Time Column

Structured Column ▾

Run R script

Run Python script

Scripts

Queries [2]
CompletedWORwA
HFC_NSW

fx = Source{[Name="CompletedWORwA"]}[Value]

	WORKORDERNO	ACTIVITYID	TECHNOLOGY
1	BSA100002262507	WOR100155979271	FTTP
2	BSA100002569828-1	WOR700077803974	HFC
3	BSA100002567705-1	WOR100172879250	FTTC
4	BSA100002567145	WOR700077187614	HFC
5	BSA100002612673	WOR100173032363	FTTC
6	BSA100002598922-1	WOR100173228097	FTTC
7	BSA100002613552-1	WOR100173627929	FTTC
8	BSA100002565559-1	WOR700078495439	HFC
9	BSA100002615817	WOR700078802810	HFC
10	BSA100002615813	WOR700078802950	HFC
11	BSA100002606529-1	WOR700078934989	HFC
12	BSA100002607313-1	WOR100173398425	FTTC
13	BSA100002602526-1	WOR100173424451	FTTC
14	BSA100002615289-1	WOR100173438600	FTTC
15	BSA100002610527-1	WOR100173537123	FTTC
16	BSA100002604551-1	WOR700078998916	HFC
17	BSA100002600383-1	WOR700079021544	HFC
18	BSA100002564160-2	WOR100173523313	FTTN
19	BSA100002583491-1	WOR100173572034	FTTC
20	BSA100002553323-1	WOR70007772508	HFC
21			

Query Settings

PROPERTIES
Name
CompletedWORwA
All Properties

APPLIED STEPS
Source
Navigation

14 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 1:56 PM

```
library(dplyr)
```

```
output <- summarise(group_by(dataset,TECHCODE), freq = n())
```

HFC_NSW

1	BSA100002604551-1	WOR700078998916	HFC	REMED	OPERATE \ DEMAND INSTALL \ HF
2	BSA100002414188	WOR700071899856	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
3	BSA100002415859	WOR700071964242	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
4	BSA100002432671	WOR700072594067	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
5	BSA100002439329-1	WOR700072839451	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
6	BSA10000				ISTALL \ HF
7	BSA10000				ISTALL \ HF
8	BSA10000				ISTALL \ HF
9	BSA10000				ISTALL \ HF
10	BSA10000				ISTALL \ HF
11	BSA10000				ISTALL \ HF
12	BSA10000				ISTALL \ HF
13	BSA10000				ISTALL \ HF
14	BSA10000				ISTALL \ HF
15	BSA10000				ISTALL \ HF
16	BSA10000				ISTALL \ HF
17	BSA10000				ISTALL \ HF
18	BSA10000				ISTALL \ HF
19	BSA10000				ISTALL \ HF
20	BSA10000				ISTALL \ HF
21	BSA10000				ISTALL \ HF
22	BSA10000				ISTALL \ HF
23	BSA10000				ISTALL \ HF
24	BSA10000				ISTALL \ HF
25	BSA10000				ISTALL \ HF
26	BSA100002573572	WOR700077492035	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
27	BSA100002576289	WOR700077577539	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
28	BSA100002576717	WOR700077593453	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
29	BSA100002561767-1	WOR700077640334	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF
30	BSA100002560506-1	WOR700077654565	HFC	HFCN	OPERATE \ DEMAND INSTALL \ HF

Run R script

Enter R scripts into the editor to transform and shape your data.

Script

```
# 'dataset' holds the input data for this script

library(dplyr)

output <- summarise(group_by(dataset,TECHCODE), freq = n())
```

The script will run with the following R installation C:\Program Files\R\R-3.6.3.
To configure your settings and change which R installation you want to run, go to Options and settings.

OK Cancel

Query Settings

PROPERTIES
Name
HFC_NSW
All Properties

APPLIED STEPS
Source
Navigation

Queries [2]	fx	= #"Run R script"{{Name="output"}}	Query Settings
CompletedWORwA	A ^B C TECHCODE	1 ² 3 freq	PROPERTIES
HFC_NSW			Name
			HFC_NSW
			All Properties
			APPLIED STEPS
			Source
			Navigation
			Run R script
			X "output"
2 COLUMNS, 145 ROWS	Column profiling based on top 1000 rows		

Visualisations

Use R button on the Visualizations pane, this adds a graphic placeholder to the report and opens the R script editor pane.

We need to select required columns from the Fields pane, or drag columns to Values section of the Visualisations pane in Power BI.

The screenshot shows the Power BI Desktop interface. The top ribbon includes File, Home, Insert, Modeling, View, Help, External Tools, Format, Data / Drill, and Table tools. The main workspace displays a scatter plot titled "dist and speed". The right-hand pane is split into "Visualizations" and "Fields". The "Fields" pane shows the "cars" dataset with "dist" and "speed" selected. The "Visualizations" pane shows the "Values" section with "dist" and "speed" added. The "R script editor" pane is open at the bottom, displaying the following R code:

```

1 # Duplicate rows will be removed from the data.
2 # The following code to create a dataframe and remove duplicated rows is always executed and acts as a
3 # preamble for your script:
4 # dataset <- data.frame(dist, speed)
5 # dataset <- unique(dataset)
6 # Paste or type your script code here:

```

Select **dist** and **speed** from the cars dataset and use below script to show scatter plot of dist vs speed.

The following code to create a dataframe and remove duplicated rows is always executed and acts as a preamble for your script:

```
# dataset <- data.frame(dist, speed)
# dataset <- unique(dataset)
```

Paste or type your script code here:

```
plot(dataset$speed, dataset$dist)
```

