



Quién soy.

- **Consultor en Data** en Paradigma Digital, con más de 8 años como docente para importantes escuelas de negocios y profesor colaborador en la UOC.
- **Especializado** en data mining, optimización de modelos y machine learning en área del Marketing, Retail y Banca-Finanzas entre otras. Además de especialista en analítica digital, SEO y PPC en digital marketing y visualización de datos - BI.
- **Apasionado** de IoT, datos y robótica, dedico el tiempo con mi familia y a mi deporte favorito, bici de carretera.



Marco Russo (aka marcusRB)



[@rb_marcus](https://twitter.com/rb_marcus)



github.com/marcusRB



[marcusRB](https://www.linkedin.com/in/marcusRB)

Qué vamos a ver.

1. Configuración de los entornos (Windows)
2. Get Data (cloud, datasets, opendata)
3. Transformation (limpieza, imputaciones, manipulación)
4. Visualization (elementos nativos, librerías)
5. Limitation (Power BI PRO - Premium)
6. Conclusiones

01.01

...

Introducción

Configuración entornos.

Power BI Desktop: entorno Windows

Power BI

Important! Selecting a language below will dynamically change the complete page content to that language.

Select Language:

English

Download



Microsoft Power BI Desktop is built for the analyst. It combines state-of-the-art interactive visualizations, with industry-leading data query and modeling built-in. Create and publish your reports to Power BI. Power BI Desktop helps you empower others with timely critical insights, anytime, anywhere.

+ Details

– System Requirements

Supported Operating System

Windows 10, Windows Server 2012 R2, Windows Server 2008 R2, Windows Server 2012, Windows 7, Windows 8, Windows 8.1

Microsoft Power BI Desktop requires Internet Explorer 10 or greater.

Microsoft Power BI Desktop is available for 32-bit (x86) and 64-bit (x64) platforms.

– Install Instructions

Download the version of Power BI Desktop that matches the architecture (x86 or x64) of your Windows OS. Run the MSI installer and follow the setup steps.

Power BI Desktop: entorno virtual

Power BI

Una máquina virtual replica un entorno Windows dentro de otro sistema operativo (sea Mac OS o Linux, e incluso otro Windows) y permite realizar test o pruebas como en este caso, de tener nuestra aplicación de Power BI .

Actualmente podemos instalar esta versión de Windows10 para las aplicaciones:

- Virtualbox
- VMware
- Parallels
- Vagrants
- HyperV

Y disponible a este enlace:

<https://developer.microsoft.com/en-us/microsoft-edge/tools/vms/>



Microsoft | Microsoft Edge Developer

Reference

Extensions

Origin Trials

Support

Careers

Home \ Tools \ VMs

Virtual Machines

Test IE11 and Microsoft Edge Legacy using free Windows 10 virtual machines you download and manage locally

Windows 10 with Legacy Microsoft Edge and Internet Explorer 11

Choose a VM platform:

VirtualBox >

Vagrant >

HyperV (Windows) >

VMware (Windows, Mac) >

Parallels (Mac) >

①Before installing, please note:

These virtual machines expire after 90 days. We recommend setting a snapshot when you first install the virtual machine which you can roll back to later. Mac users will need to use a tool that supports zip64, like [The Unarchiver](#), to unzip the files. The password to your VM is "Passw0rd!"

[View installation instructions](#)

The Microsoft Software License Terms for the Microsoft Edge and IE VMs are included in the [release notes](#) and supersede any conflicting Windows license terms included in the VMs. By downloading and using this software, you agree to these [license terms](#).



Power BI Desktop: entorno virtual Virtualbox

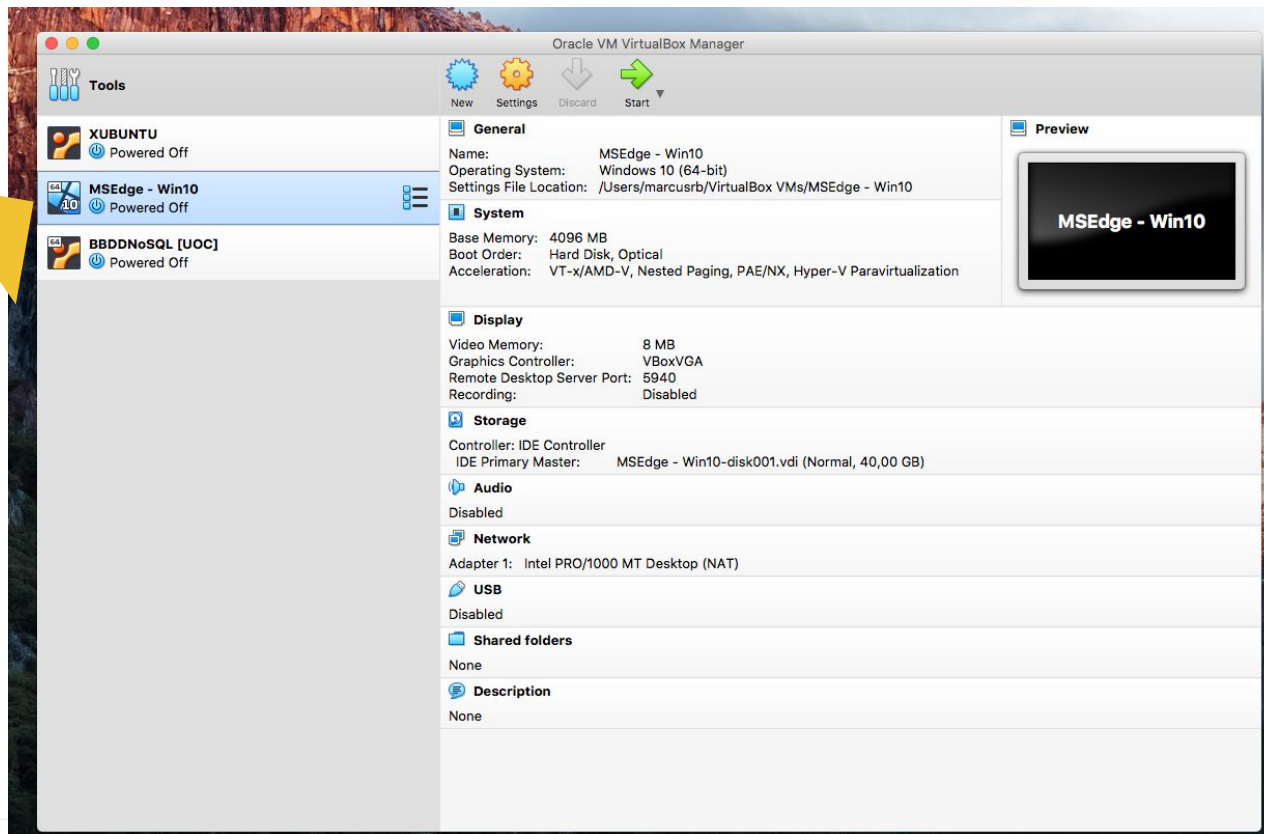
Power BI



En caso de Virtualbox asignamos un mínimo de:

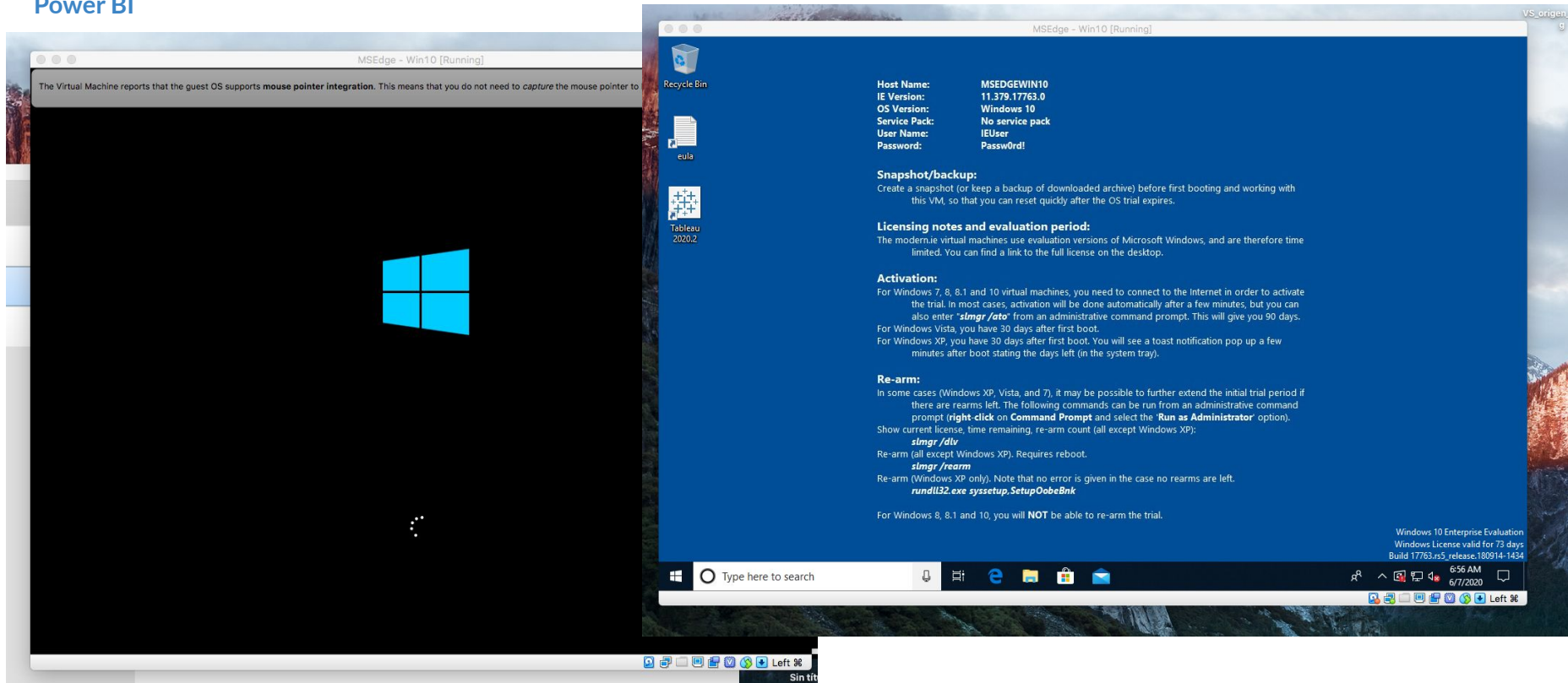
- 4gb de memoria RAM
- 2 CPU

El resto de opciones vienen por defecto siendo el espacio disponible en disco duro de un mínimo de 50 gb

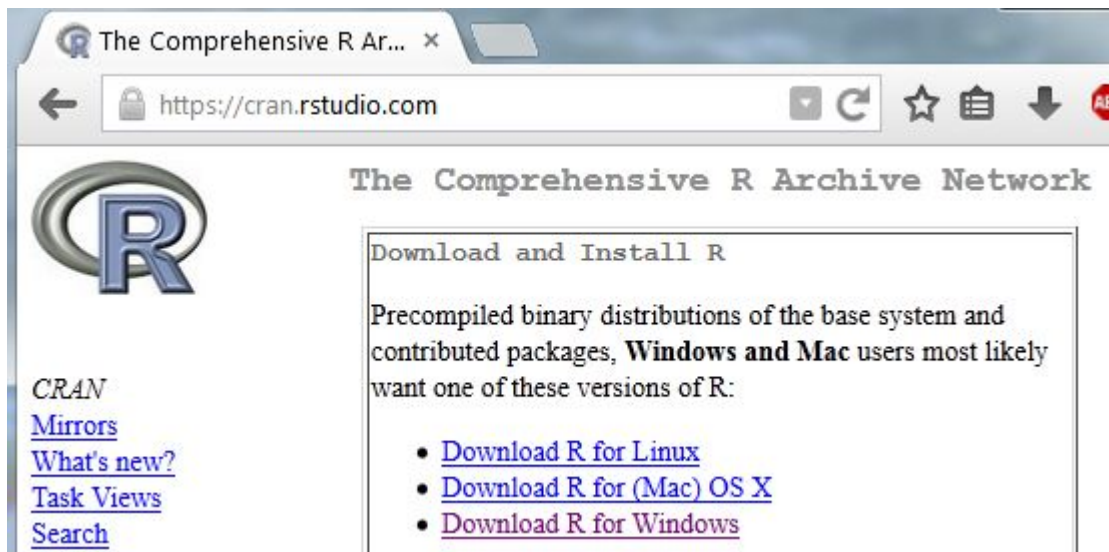


Power BI Desktop: entorno virtual Virtualbox

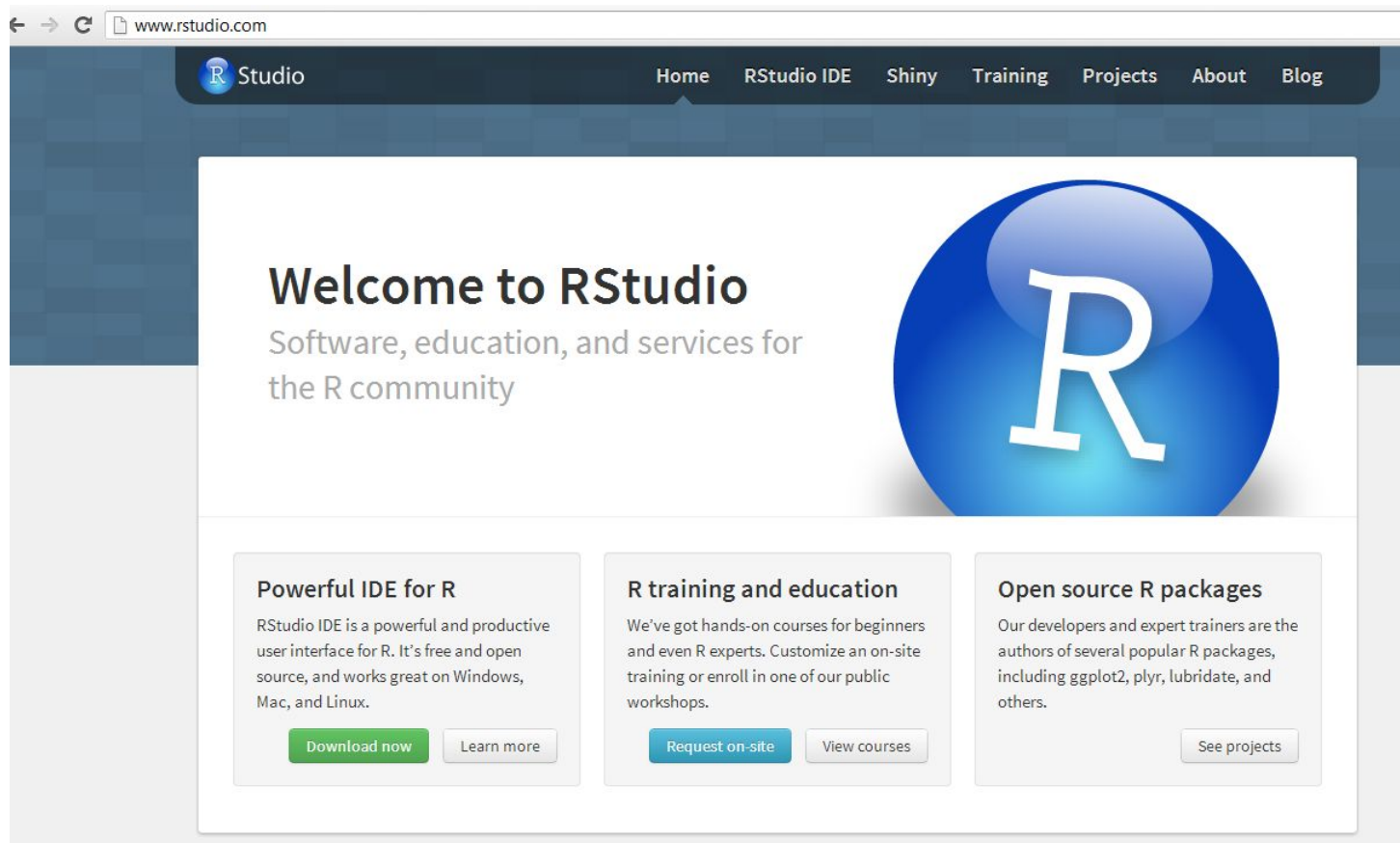
Power BI



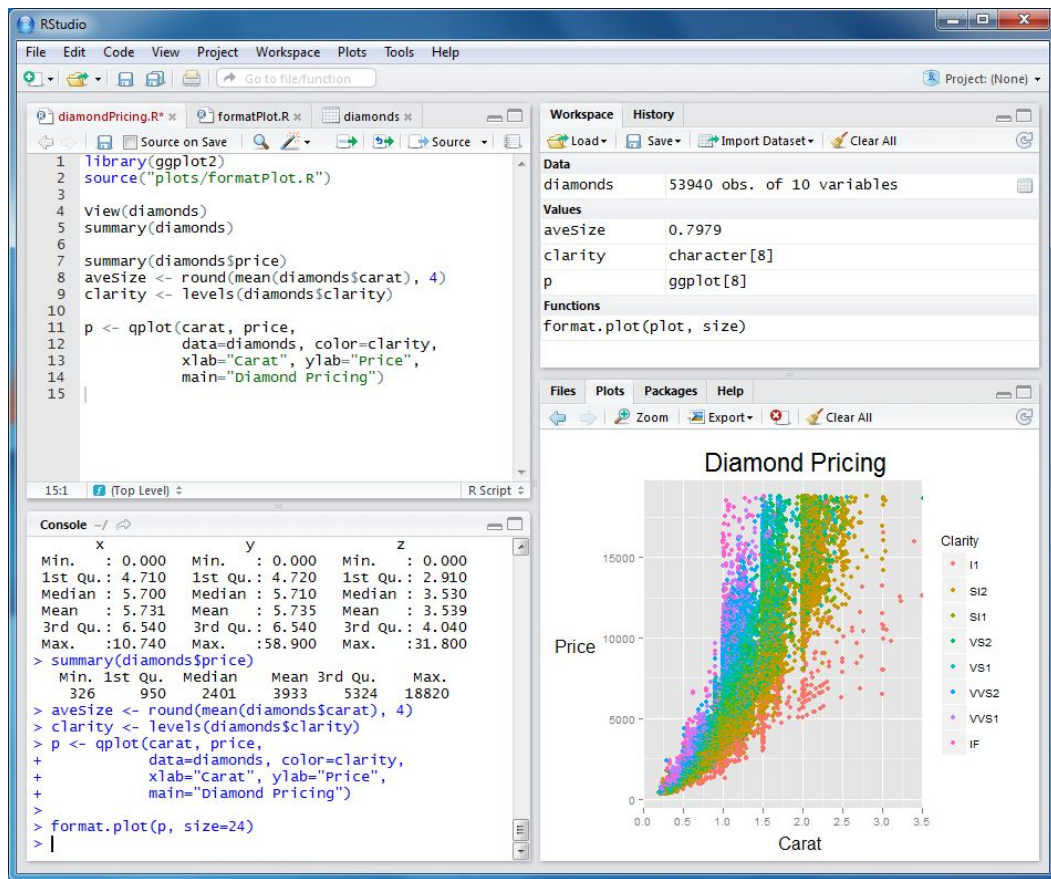
Configuración R commander.



Configuración Rstudio.



Comprobación de R Studio.

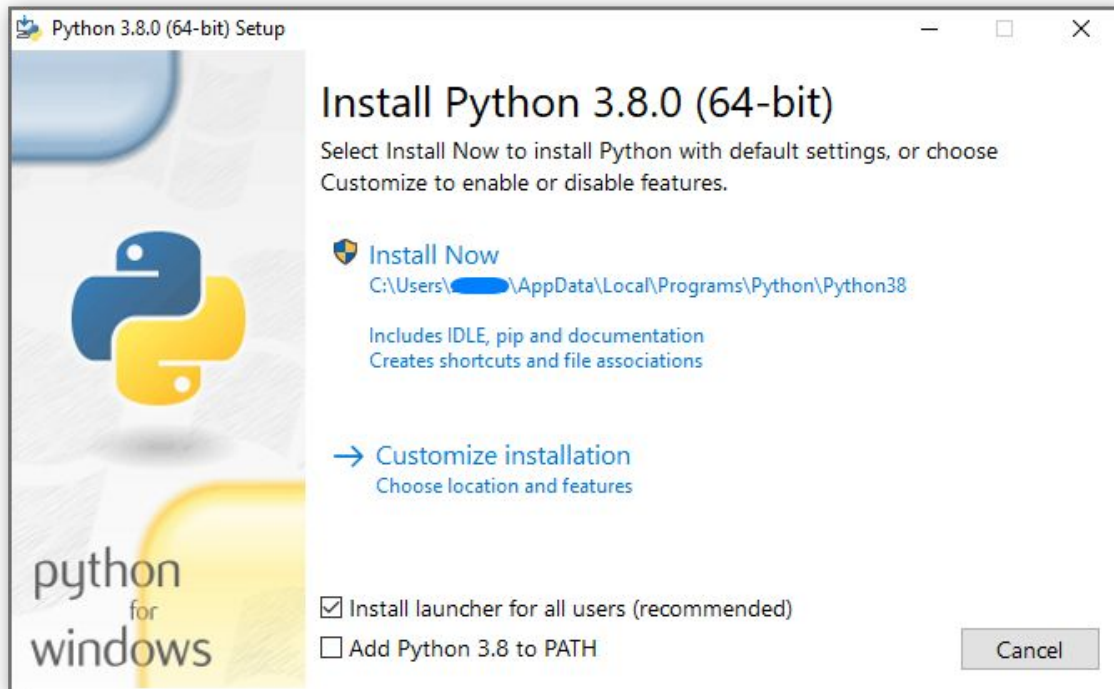




Power BI

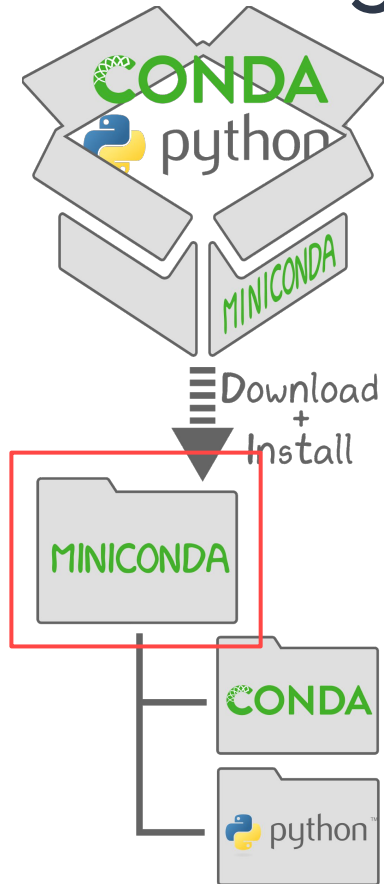
Demo 1.

Configuración entorno Python.



<https://www.microsoft.com/en-us/p/python-38/>

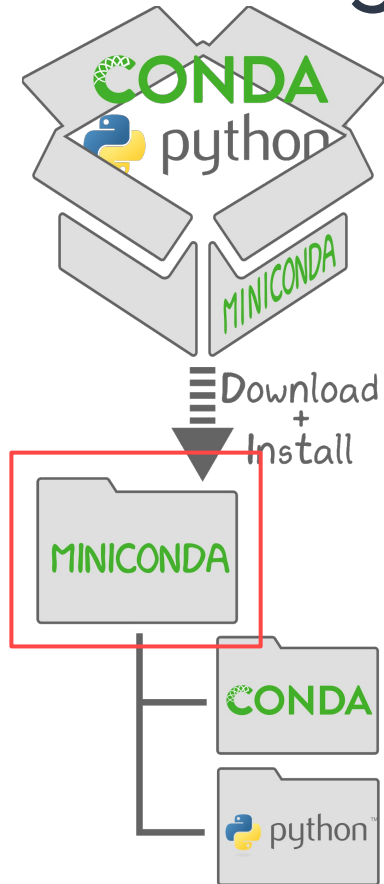
Configuración Miniconda for Win10.



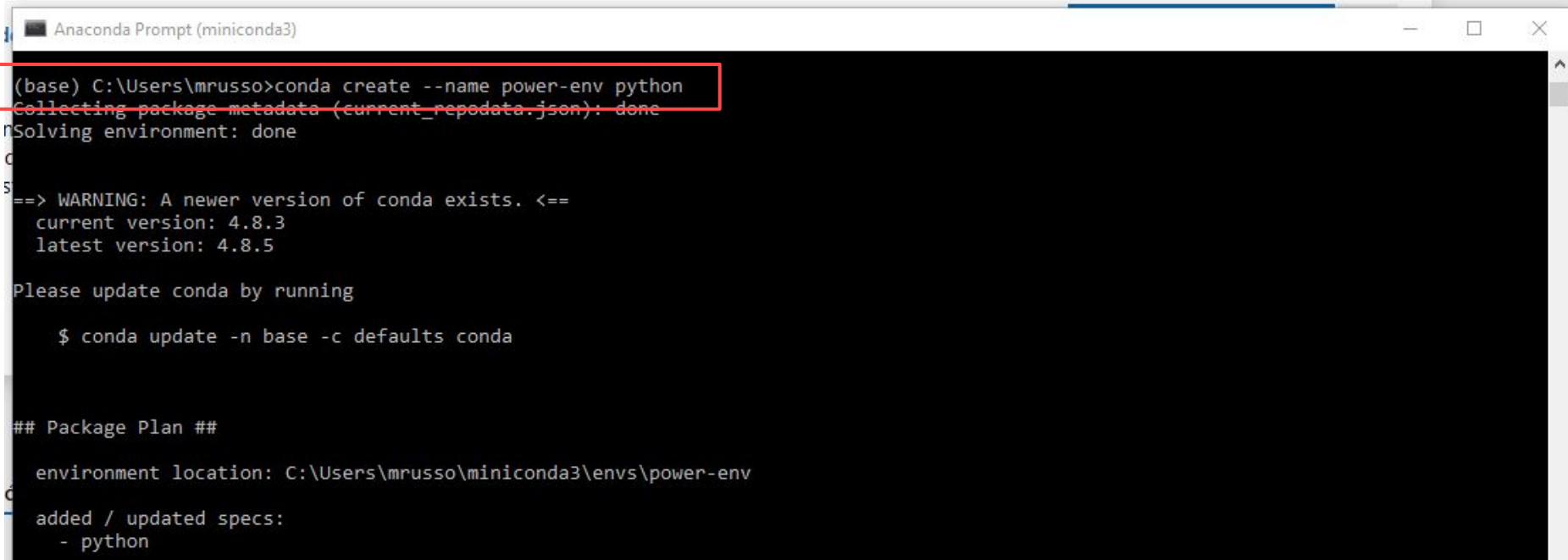
System requirements

- 32- or 64-bit computer.
- For Miniconda---400 MB disk space.
- For Anaconda---Minimum 3 GB disk space to download and install.
- Windows, macOS, or Linux.

Configuración Miniconda for Win10.



Configuración Miniconda for Win10.



```
Anaconda Prompt (miniconda3)

(base) C:\Users\mrusso>conda create --name power-env python
Collecting package metadata (current_repodata.json): done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
  current version: 4.8.3
  latest version: 4.8.5

Please update conda by running

  $ conda update -n base -c defaults conda

## Package Plan ##

  environment location: C:\Users\mrusso\miniconda3\envs\power-env

  added / updated specs:
    - python
```


Configuración Miniconda for Win10.

```
Proceed ([y]/n)? y

Downloading and Extracting Packages
python-3.8.5          | 15.7 MB | ##### | 100%
sqlite-3.33.0        | 809 KB  | ##### | 100%
certifi-2020.6.20    | 157 KB  | ##### | 100%
openssl-1.1.1h       | 4.8 MB  | ##### | 100%
vs2015_runtime-14.16 | 1.2 MB  | ##### | 100%
wheel-0.35.1         | 37 KB   | ##### | 100%
setuptools-50.3.0    | 741 KB  | ##### | 100%
ca-certificates-2020 | 122 KB  | ##### | 100%
pip-20.2.3           | 1.8 MB  | ##### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate power-env
#
# To deactivate an active environment, use
#
#     $ conda deactivate

(base) C:\Users\mrusso>conda activate power-env
```

<https://conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html>



Power BI

Demo 2.

Configuración en Power BI.

The screenshot shows the 'Options' dialog box with the 'R scripting' tab selected. A red rectangle highlights the 'R script options' section. The 'Detected R home directories' dropdown is set to 'C:\Program Files\Power BI Desktop\Power BI Desktop\'. The 'Detected R IDEs' dropdown is set to 'R Studio'. The 'How to install R' link is visible. The 'Change temporary storage location' link is also visible. The 'OK' and 'Cancel' buttons are at the bottom right.

Options

GLOBAL

- Data Load
- Power Query Editor
- DirectQuery
- R scripting**
- Python scripting
- Security
- Privacy
- Regional Settings
- Updates
- Usage Data
- Diagnostics
- Preview features
- Auto recovery
- Report settings

R script options

To choose a home directory for R, select a detected R installation from the drop-down list, or select Other and browse to the location you want.

Detected R home directories:

C:\Program Files\Power BI Desktop\Power BI Desktop\

[How to install R](#)

To choose which R integrated development environment (IDE) you want Power BI Desktop to launch, select a detected IDE from the drop-down list, or select Other to browse to another IDE on your machine.

Detected R IDEs:

R Studio

[Learn more about R IDEs](#)

[Change temporary storage location](#)

Note: Sometimes, R custom visuals automatically install additional packages. For those to work, the temporary storage folder name must be written in Latin characters (letters in the English alphabet).

OK Cancel

The screenshot shows the 'Options' dialog box with the 'Python scripting' tab selected. A red rectangle highlights the 'Python script options' section. The 'Detected Python home directories' dropdown is set to 'Other'. The 'Set a Python home directory' field is set to 'C:\Users\mrusso\miniconda3\envs\power-env'. The 'Browse' button is visible. The 'How to install Python' link is visible. The 'Detected Python IDEs' dropdown is set to 'Other'. The 'Browse to the Python IDE you want' field is set to 'C:\Users\mrusso\AppData\Local\Programs\Microsoft\Power BI Desktop\'. The 'Browse' button is visible. The 'Learn more about Python IDEs' link is visible. The 'Change temporary storage location' link is also visible. The 'OK' and 'Cancel' buttons are at the bottom right.

Options

GLOBAL

- Data Load
- Power Query Editor
- DirectQuery
- R scripting
- Python scripting**
- Security
- Privacy
- Regional Settings
- Updates
- Usage Data
- Diagnostics
- Preview features
- Auto recovery
- Report settings

Python script options

To choose a home directory for Python, select a detected Python installation from the drop-down list, or select Other and browse to the location you want.

Detected Python home directories:

Other

Set a Python home directory:

C:\Users\mrusso\miniconda3\envs\power-env

[How to install Python](#)

To choose which Python integrated development environment (IDE) you want Power BI Desktop to launch, select a detected IDE from the drop-down list, or select Other to browse to another IDE on your machine.

Detected Python IDEs:

Other

Browse to the Python IDE you want:

C:\Users\mrusso\AppData\Local\Programs\Microsoft\Power BI Desktop\

[Learn more about Python IDEs](#)

[Change temporary storage location](#)

Note: Sometimes, Python custom visuals automatically install additional packages. For those to work, the temporary storage folder name must be written in Latin characters (letters in the English alphabet).

OK Cancel



Power BI

Demo 3.

01.02



Power BI

Get Data con R.



Power BI

Demo 4.

01.03



Power BI

Get Data con Python.



Power BI

Demo 5.

02.01



Power BI

Transformation.

Transformación e imputaciones.





Power BI

Demo 6.

03.01



Power BI

Visualization.

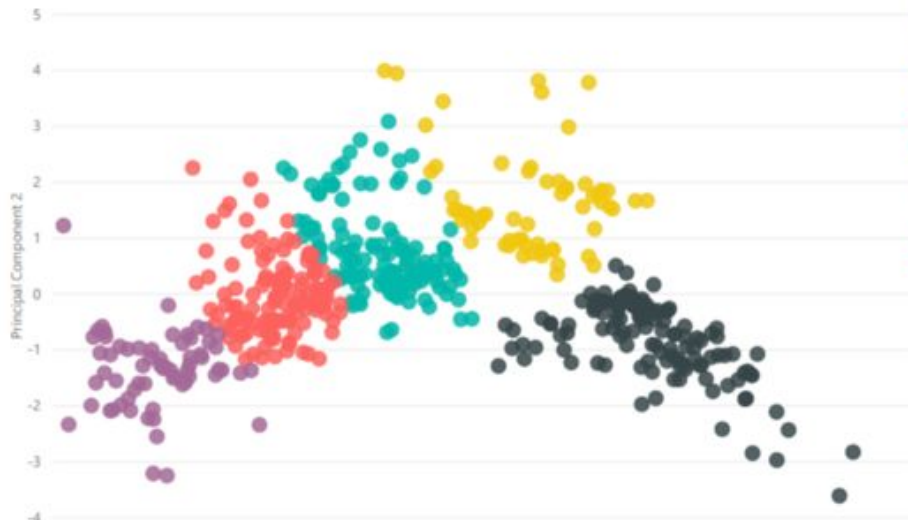
Visualización con R y Python.

Clustering Boston Housing Data

<https://archive.ics.uci.edu/ml/machine-learning-databases/housing/>

Variance explained

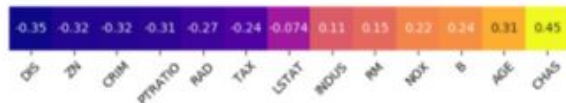
0,11



Principal Component 1

Variance explained

0,47



House value in \$1000's



CRIM per capita crime rate by town
ZN proportion of residential land zoned for lots over 25,000 sq.ft.
INDUS proportion of non-retail business acres per town
CHAS Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
NOX nitric oxides concentration (parts per 10 million)
RM average number of rooms per dwelling
AGE proportion of owner-occupied units built prior to 1940
DIS weighted distances to five Boston employment centres
RAD index of accessibility to radial highways
TAX full-value property-tax rate per \$10,000
PTRATIO pupil-teacher ratio by town
B $1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town
LSTAT % lower status of the population
MEDV Median value of owner-occupied homes in \$1000's

05.01



Power BI

Limitation.

Limitaciones con Python.

The Power BI service also imposes other limits on Python script execution.

CONSTRAINT	LIMIT
Dataframe size	150,000 rows 250MB when serialized in CSV format
Execution time	60 seconds
Memory consumption	1 GB
Disk I/O	1 GB/sec
Image size for a visual in PNG format	2 MB

These limits ensure that user-provided scripts are performant, and moderate in the consumption of resources.

Limitaciones con R.

Known Limitations

R visuals in the Power BI service have a few limitations:

- R visuals support is limited to the packages identified in [Learn which R packages are supported](#). There currently is no support for custom packages.
- 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. Additionally, the input data has a limit of 250 MB.
- Resolution - all R visuals are displayed at 72 DPI.
- Plotting device - only plotting to the default device is supported.
- Calculation time limitation – if an R visual calculation exceeds 60 seconds the script times out, resulting in an error.
- R visuals are refreshed upon data updates, filtering, and highlighting. However, the image itself is not interactive and does not support tool tips.
- R visuals respond to highlighting other visuals, but you cannot click on elements in the R visual in order to cross filter other elements.
- R visuals are currently not supported for the *Time* data type. Please use Date/Time instead.
- R visuals do not display when using **Publish to web**.
- R visuals do not support renaming input columns. Columns will be referred to by their original name during script execution.
- R visuals currently do not print with dashboard and reports printing
- R visuals are currently not supported in the DirectQuery mode of Analysis Services

<https://docs.microsoft.com/en-us/power-bi/visuals/service-r-visuals#known-limitations>



**¡Muchas
Gracias!**

Recursos del webinar.

github.com/marcusRB/WEBINAR_PoweR_PY

<https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-r-visuals>

<https://powerbi.microsoft.com/es-es/blog/python-visualizations-in-power-bi-service/>

<https://powerbi.microsoft.com/en-us/blog/pythonblogepisode1/>

<https://powerbi.microsoft.com/fr-fr/blog/data-cleansing-with-r-in-power-bi/>



Paradigma

- https://youtu.be/AFTanu_iBD4

- <https://www.paradigmadigital.com/dev/5-tips-power-bi-sencillos-utiles/>

- <https://www.paradigmadigital.com/dev/bigquery-y-power-bi-emplatado-datos/>

- <https://www.paradigmadigital.com/techbiz/cocinando-datos-bigquery-powerbi-primeros-pasos/>