

Learning the basics of Python in Power BI

INTRODUCTION TO PYTHON IN POWER BI



Jacob H. Marquez

Data Scientist

Prerequisite knowledge

Power BI

[Introduction to DAX in Power BI](#)

Python

[Introduction to Data Visualization with Seaborn](#)

What we will be learning

Power BI

- Importing and joining data with Power Query
- Finding missing data with Power Query
- Calculating Correlation Coefficient with Quick Measures

Python

- Importing and joining data with pandas
- Finding missing data and imputation
- Creating visualizations with Seaborn

Getting started - downloading Python

<https://powerbi.microsoft.com/downloads/>

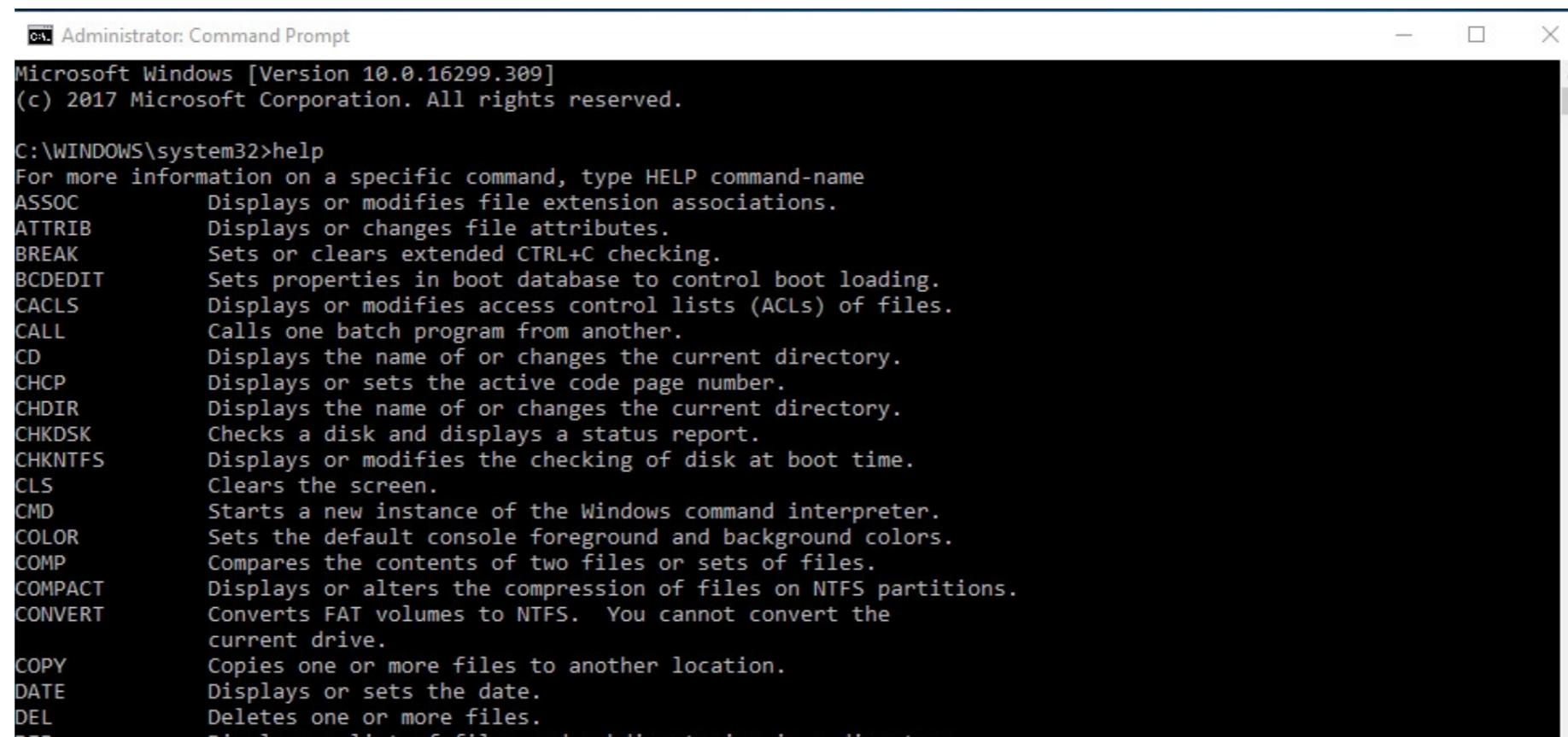
<https://www.python.org/downloads/>



Getting started - installing packages

The two required Python packages are:

- pandas
- matplotlib



A screenshot of a Windows Command Prompt window titled "Administrator: Command Prompt". The window shows a list of command names and their descriptions. The commands listed are: ASSOC, ATTRIB, BREAK, BCDEDIT, CACLS, CALL, CD, CHCP, CHDIR, CHKDSK, CHKNTFS, CLS, CMD, COLOR, COMP, COMPACT, CONVERT, COPY, DATE, DEL, and ECHO. Each command is followed by a brief description of its function.

```
Administrator: Command Prompt
Microsoft Windows [Version 10.0.16299.309]
(c) 2017 Microsoft Corporation. All rights reserved.

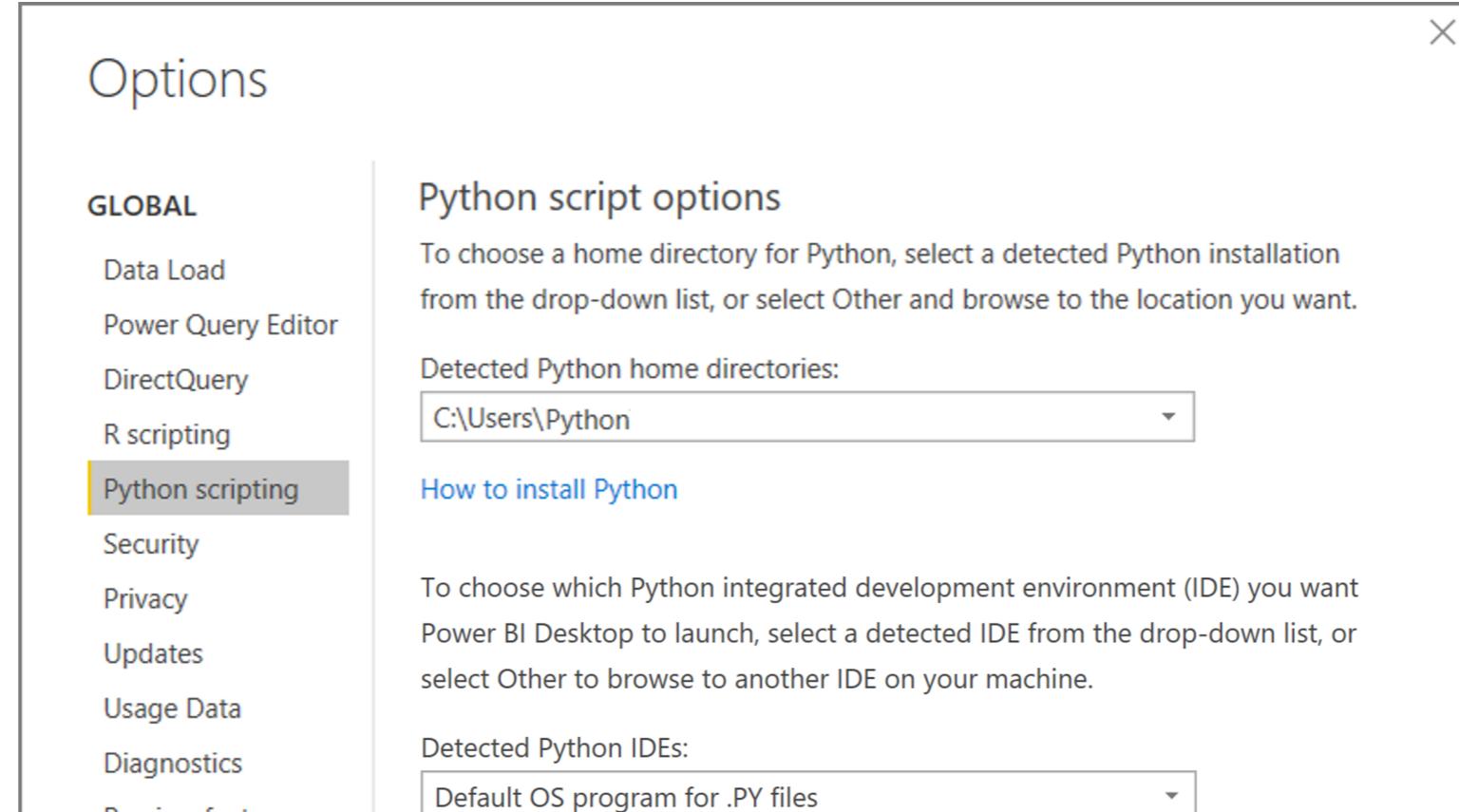
C:\WINDOWS\system32>help
For more information on a specific command, type HELP command-name
ASSOC      Displays or modifies file extension associations.
ATTRIB     Displays or changes file attributes.
BREAK      Sets or clears extended CTRL+C checking.
BCDEDIT    Sets properties in boot database to control boot loading.
CACLS      Displays or modifies access control lists (ACLs) of files.
CALL       Calls one batch program from another.
CD         Displays the name of or changes the current directory.
CHCP       Displays or sets the active code page number.
CHDIR     Displays the name of or changes the current directory.
CHKDSK    Checks a disk and displays a status report.
CHKNTFS   Displays or modifies the checking of disk at boot time.
CLS        Clears the screen.
CMD        Starts a new instance of the Windows command interpreter.
COLOR      Sets the default console foreground and background colors.
COMP       Compares the contents of two files or sets of files.
COMPACT    Displays or alters the compression of files on NTFS partitions.
CONVERT   Converts FAT volumes to NTFS. You cannot convert the
          current drive.
COPY      Copies one or more files to another location.
DATE      Displays or sets the date.
DEL       Deletes one or more files.
ECHO      Displays a list of files and their attributes.
```

¹ <https://packaging.python.org/en/latest/tutorials/installing-packages/>

Getting started - Power BI

"File" > "Options and settings" > "Options" > "Python scripting"

- Home Directories
- Integrated Development Environments (IDE)



¹ <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-scripts>

Getting started - Power BI

The screenshot shows the 'Options' dialog box in Power BI. On the left, a sidebar lists various global settings: Data Load, Power Query Editor, DirectQuery, R scripting, Python scripting (which is selected and highlighted in yellow), Security, Privacy, Updates, Usage Data, Diagnostics, Preview features, Auto recovery, and Report settings. The main pane, titled 'Python script options', contains instructions for choosing a Python home directory and a detected Python IDE. A green box highlights the 'Detected Python home directories:' dropdown, which shows 'C:\Users\Python'. Below it is a link to 'How to Install Python'. Another green box highlights the 'Detected Python IDEs:' dropdown, which shows 'Default OS program for .PY files'. Below it is a link to 'Learn more about Python IDEs'.

GLOBAL

Data Load

Power Query Editor

DirectQuery

R scripting

Python scripting

Security

Privacy

Updates

Usage Data

Diagnostics

Preview features

Auto recovery

Report settings

Options

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:

C:\Users\Python

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:

Default OS program for .PY files

Learn more about Python IDEs

¹ <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-scripts>

Getting started - Power BI

The screenshot shows the 'Options' dialog box in Power BI. On the left, a sidebar lists various global settings: Data Load, Power Query Editor, DirectQuery, R scripting, Python scripting (which is selected and highlighted in yellow), Security, Privacy, Updates, Usage Data, Diagnostics, Preview features, Auto recovery, and Report settings. The main pane is titled 'Python script options' and contains instructions for choosing a Python home directory. It shows a dropdown menu set to 'C:\Users\Python'. Below this is a section for choosing a Python IDE, with a dropdown menu set to 'Default OS program for .PY files'. A green rectangular box highlights this section. At the bottom of the main pane is a link 'Learn more about Python IDEs'.

GLOBAL

Data Load
Power Query Editor
DirectQuery
R scripting
Python scripting
Security
Privacy
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:

C:\Users\Python

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:

Default OS program for .PY files

[Learn more about Python IDEs](#)

¹ <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-scripts>

Limitations

- DataFrames temporarily saved
- 150,000 rows used in visualizations
- Datasets less than 250 MB
- Datasets set to "Public"

Benefits and reasons to learn...

Python as a Power BI user

- Flexibility in data transformations
- Easier method for difficult calculations
- Build custom visualizations

Power BI as a Python user

- Faster data visualization
- Build quick reporting dashboards
- Easier sharing with others

Datasets

fished.csv

entity	year	fished
Australia	1990	153148
United States	1990	502923
Japan	1990	567895

farmed.csv

entity	year	farmed
Australia	1990	23567
United States	1990	436758
Japan	1990	534978

Let's practice!

INTRODUCTION TO PYTHON IN POWER BI

Set up, import, and transform data

INTRODUCTION TO PYTHON IN POWER BI



Jacob H. Marquez

Data Scientist

Let's practice!

INTRODUCTION TO PYTHON IN POWER BI