***python\_data\_matcher.py***



Luzie U. Wingen

**python\_data\_matcher.py** is a python script to match or join the content of two spread sheets by their first column items. The initial purpose was to join marker information from different lists, but it has been proven useful beyond that.

# Basic instructions:

Install Python 3.5 or higher (<https://www.python.org>).

**Unzip python\_data\_matcher.zip**

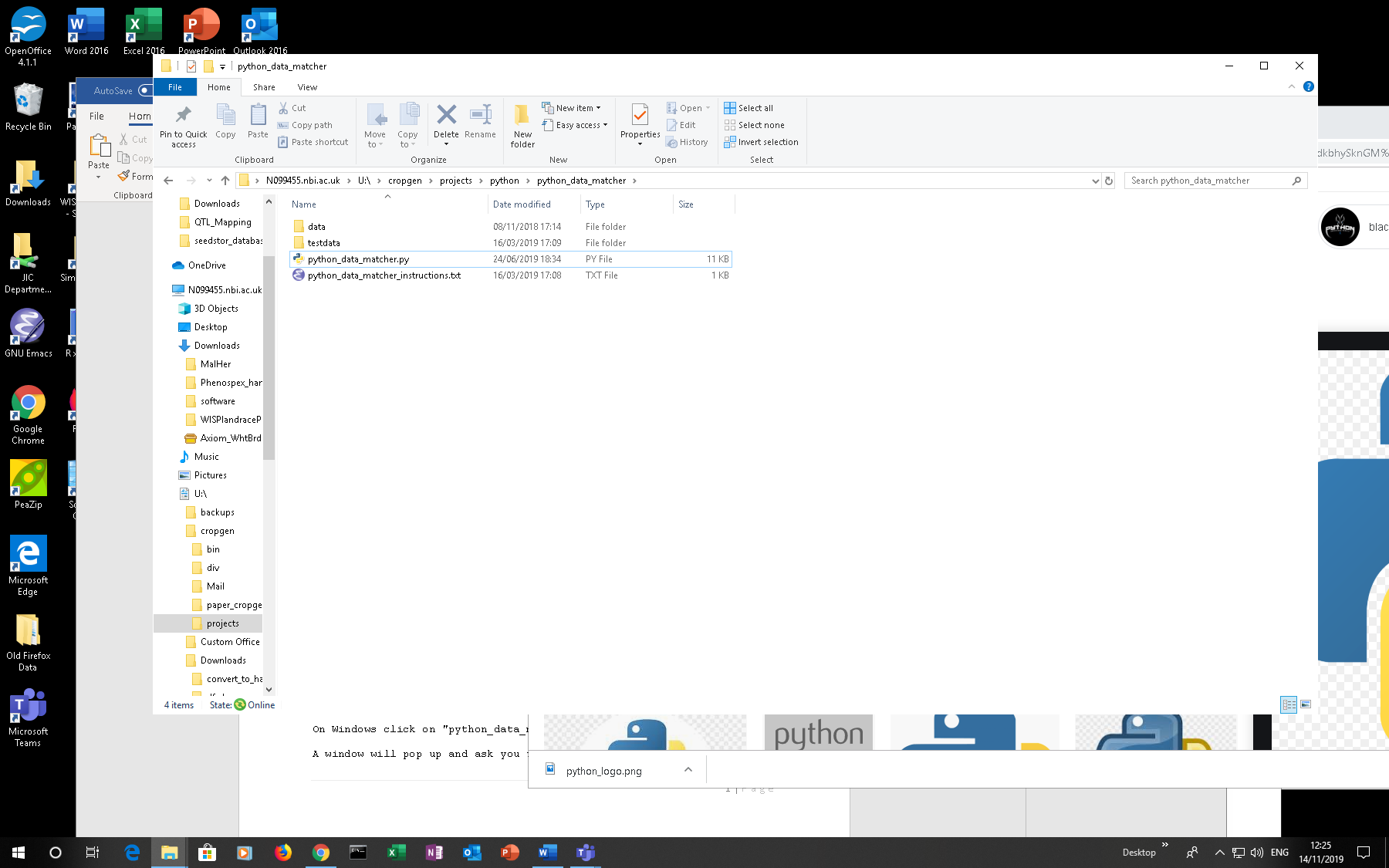
Unzip the zip file on your computer and navigate into the directory **python\_data\_matcher**, which contains the script **python\_data\_matcher.py**.

**Save spread sheets from Excel as "csv".**

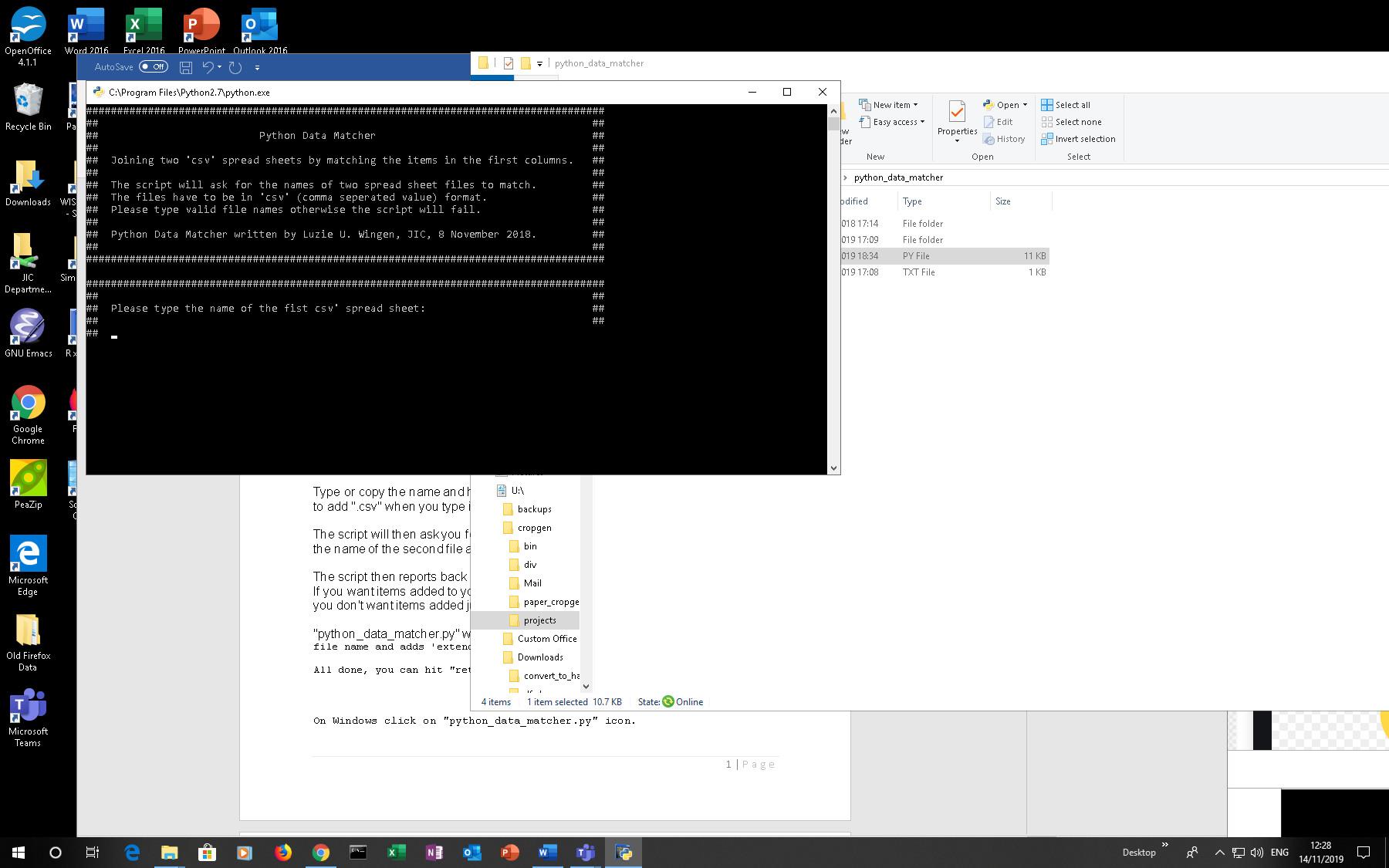
The script works on comma-separated spread sheets or "csv" files. Any Excel spread sheet needs to be saved as “csv” first. “csv”-files are best put into the directory with the python script - less typing of directory names.

**Start the python\_data\_matcher.py script**

On Windows click on **python\_data\_matcher.py** icon to execute the script.



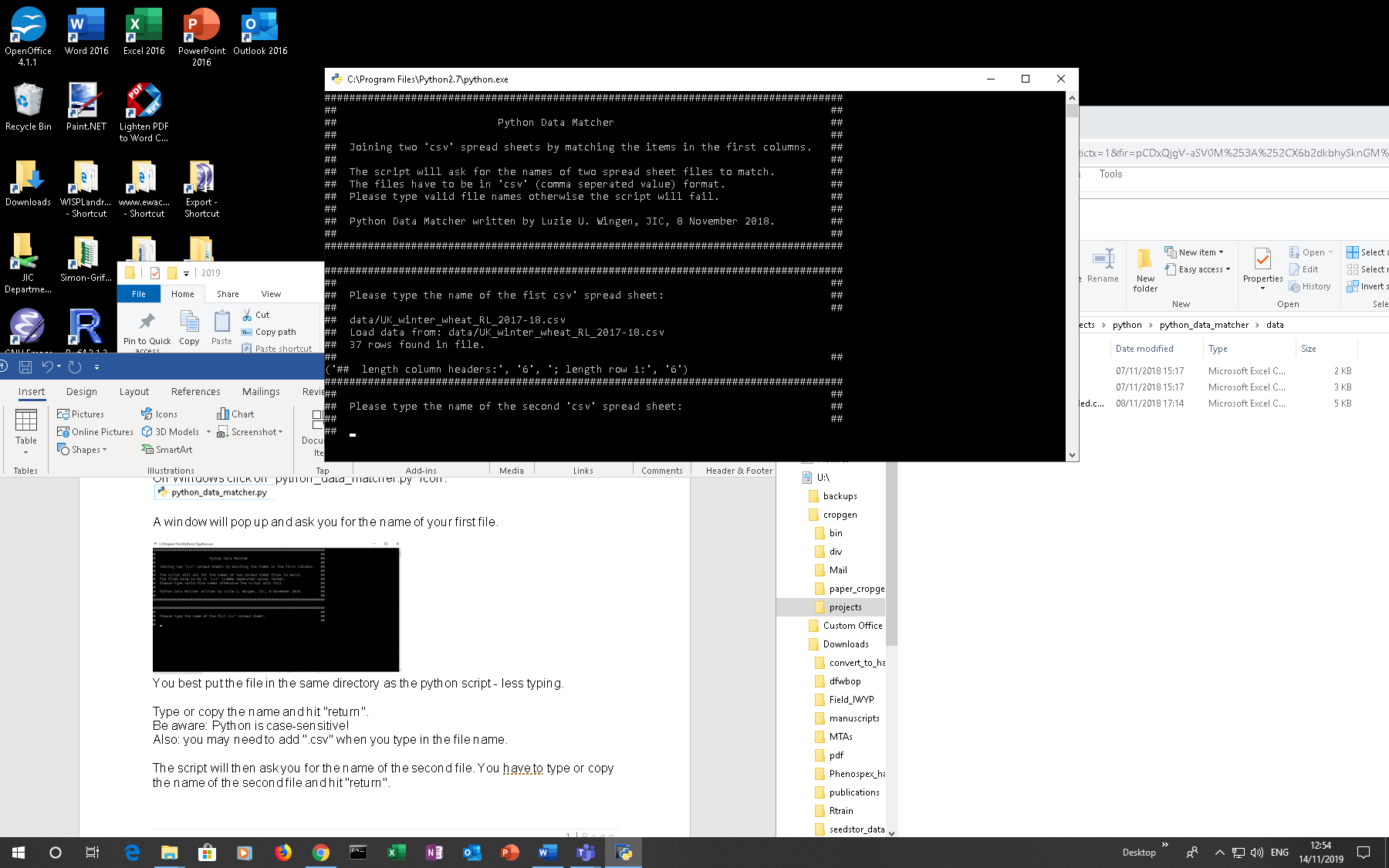
A window will pop up and ask for the name of the first file.



**Type or copy the name of the first file and hit "return".**

Be aware: python is case-sensitive!

Also: The file extension ".csv" is part of the file name.



The script will then ask for the name of the second file.

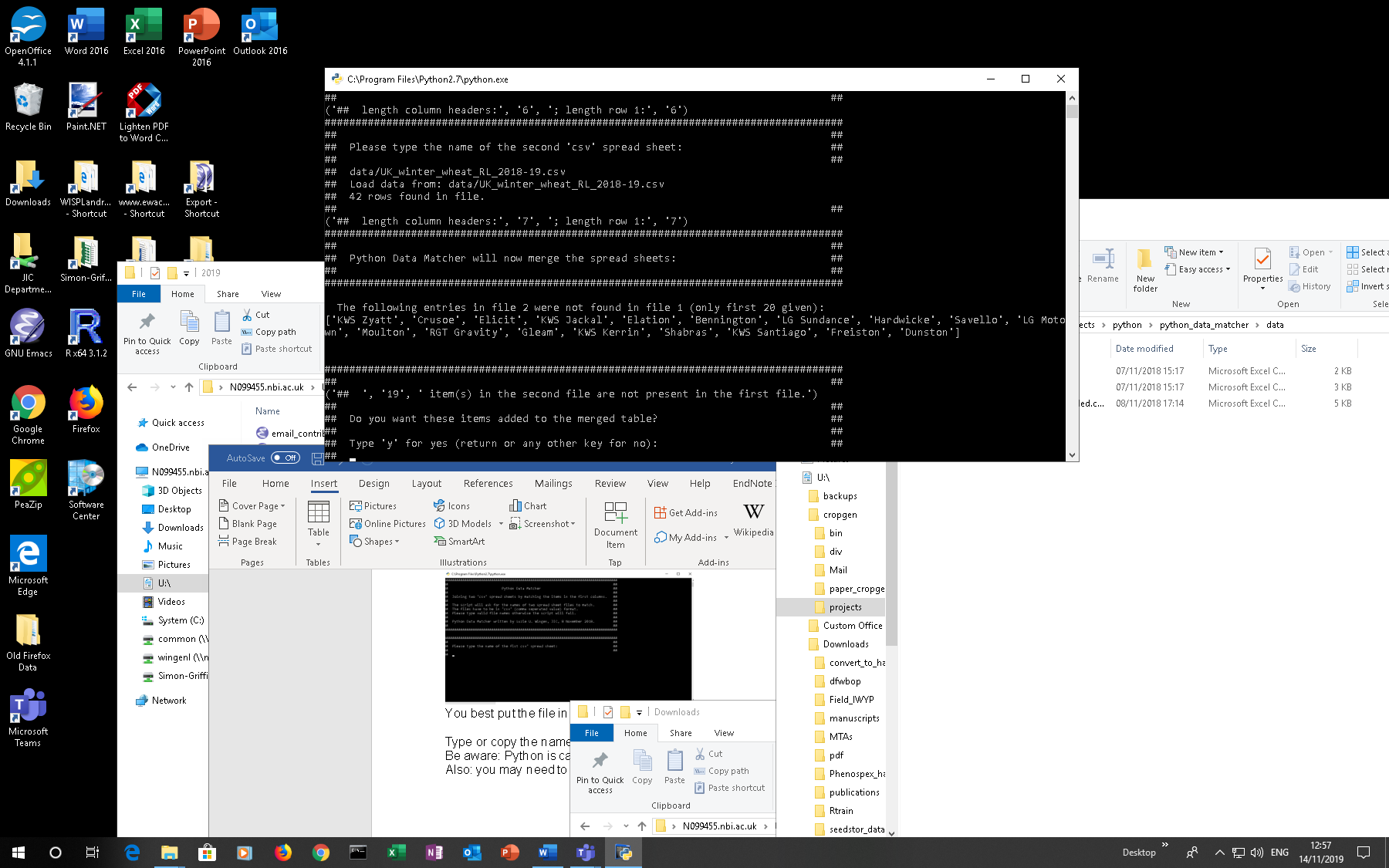
**Type or copy the name of the second file and hit "return".**

The script reports back if it finds items in the first column of the second file that are not present in the first column of the first file.

**Adding items from the second file to the first**

If you want items added answer with: "y" + "return".

If you don't want items added just hit "return".



**python\_data\_matcher.py** produces a new file with the joint items.

As a filename it will use the first file name with an added '\_extended'.

**All done**

Hit "return" to leave the script and inspect the new file which will be in the **python\_data\_matcher** directory.

14/11/2019