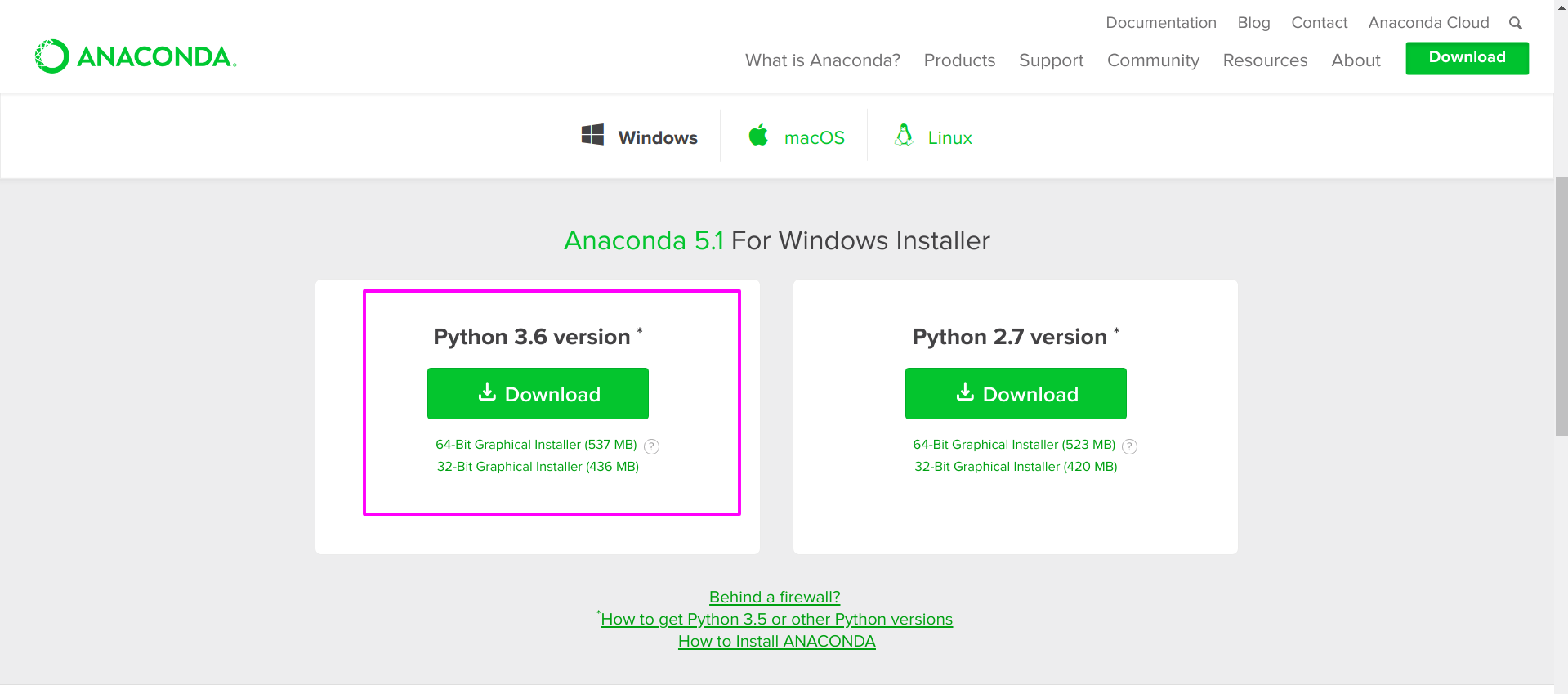
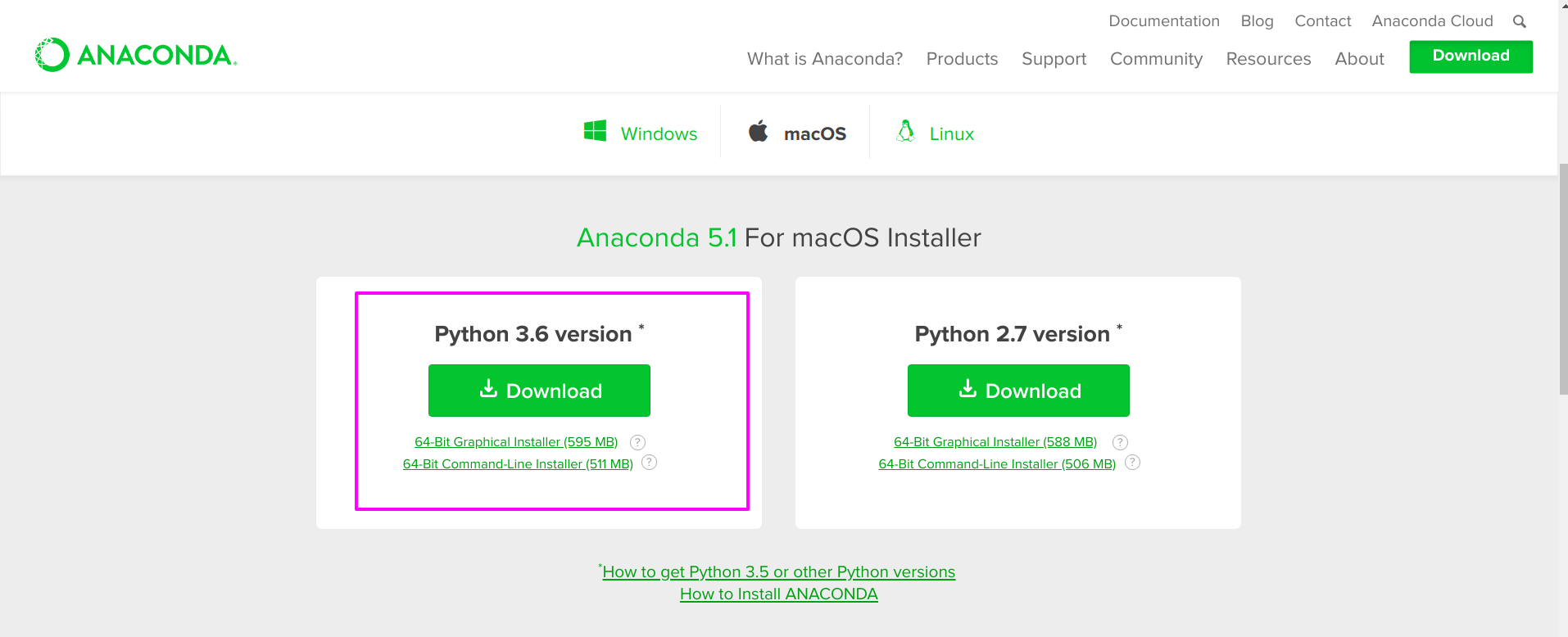
**User tutorial:**

1. install python and third-party libs(**bs4, xlwt**)

if you are not familiar with python, recommend download Anaconda at <https://www.anaconda.com/download/> and install Python 3.6 version. So the third-party libs will be included.

For windows:

For mac:

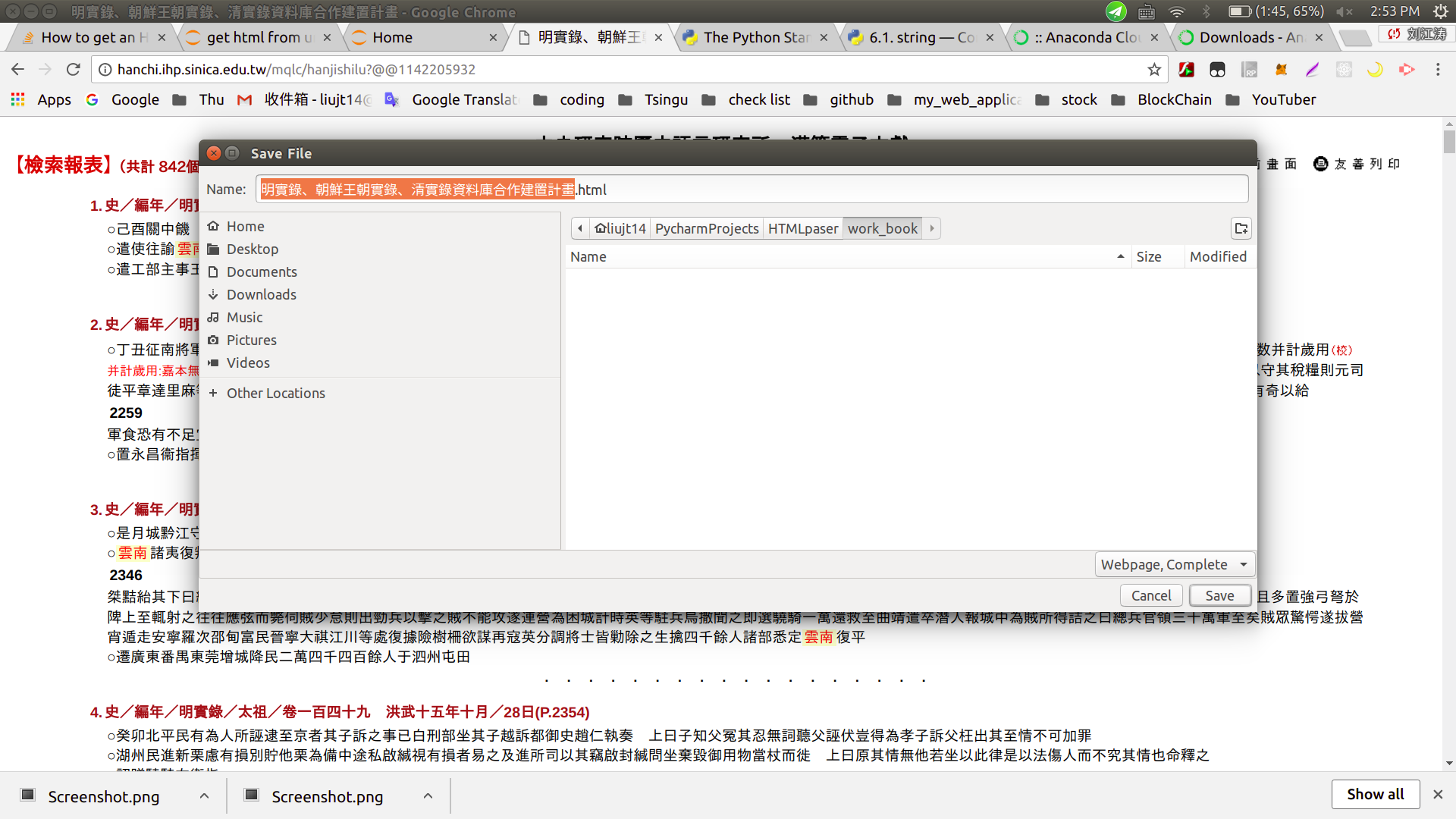
linux the same step.

2. Save the website as ***html***file.

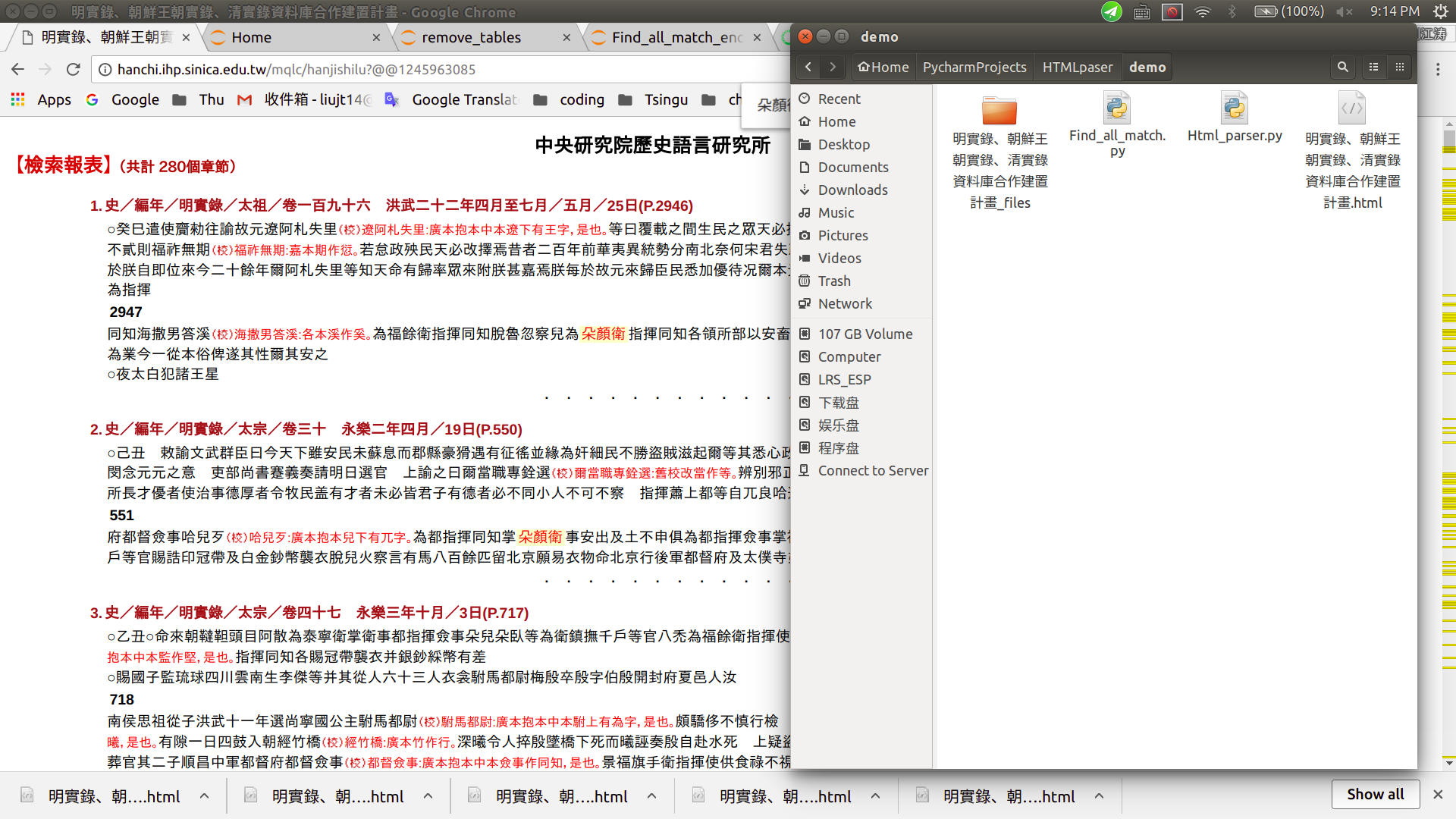
Right click the website and choose ***Save as...***



Save to *one emopty folder* and **DO NOT change the name**.

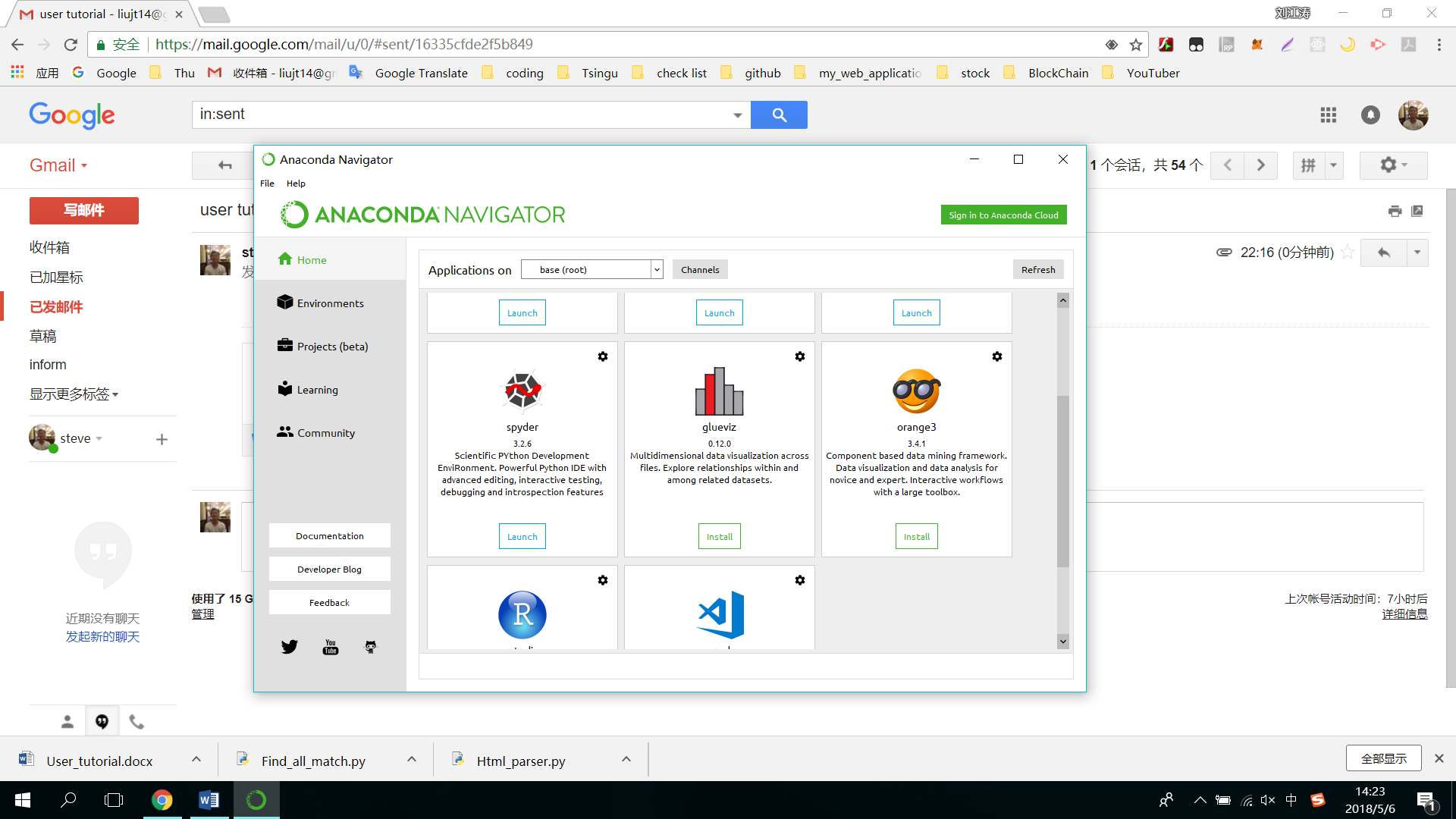


Copy the python files –Html\_parser.py and Find\_all\_match.py to this folder

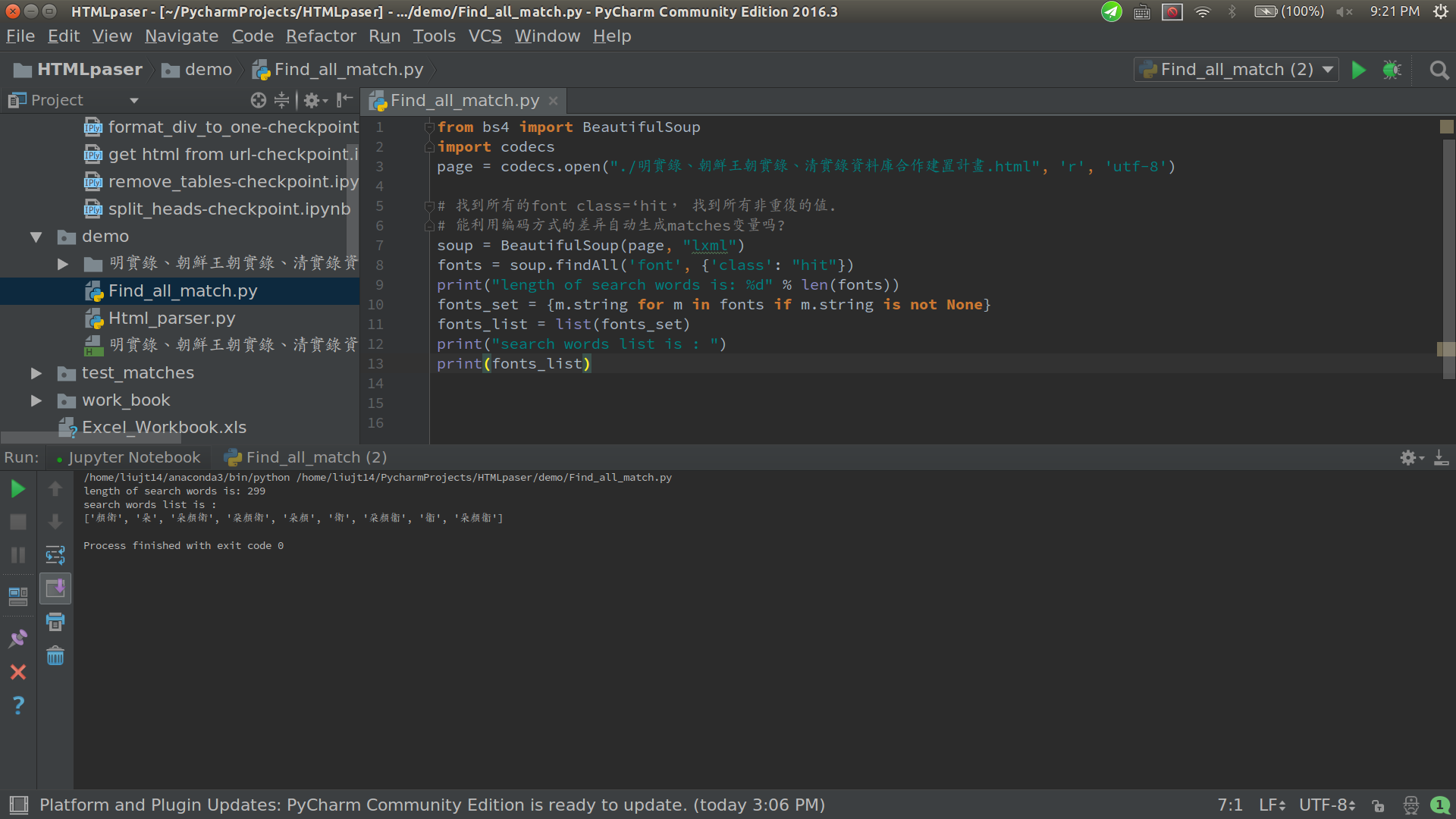


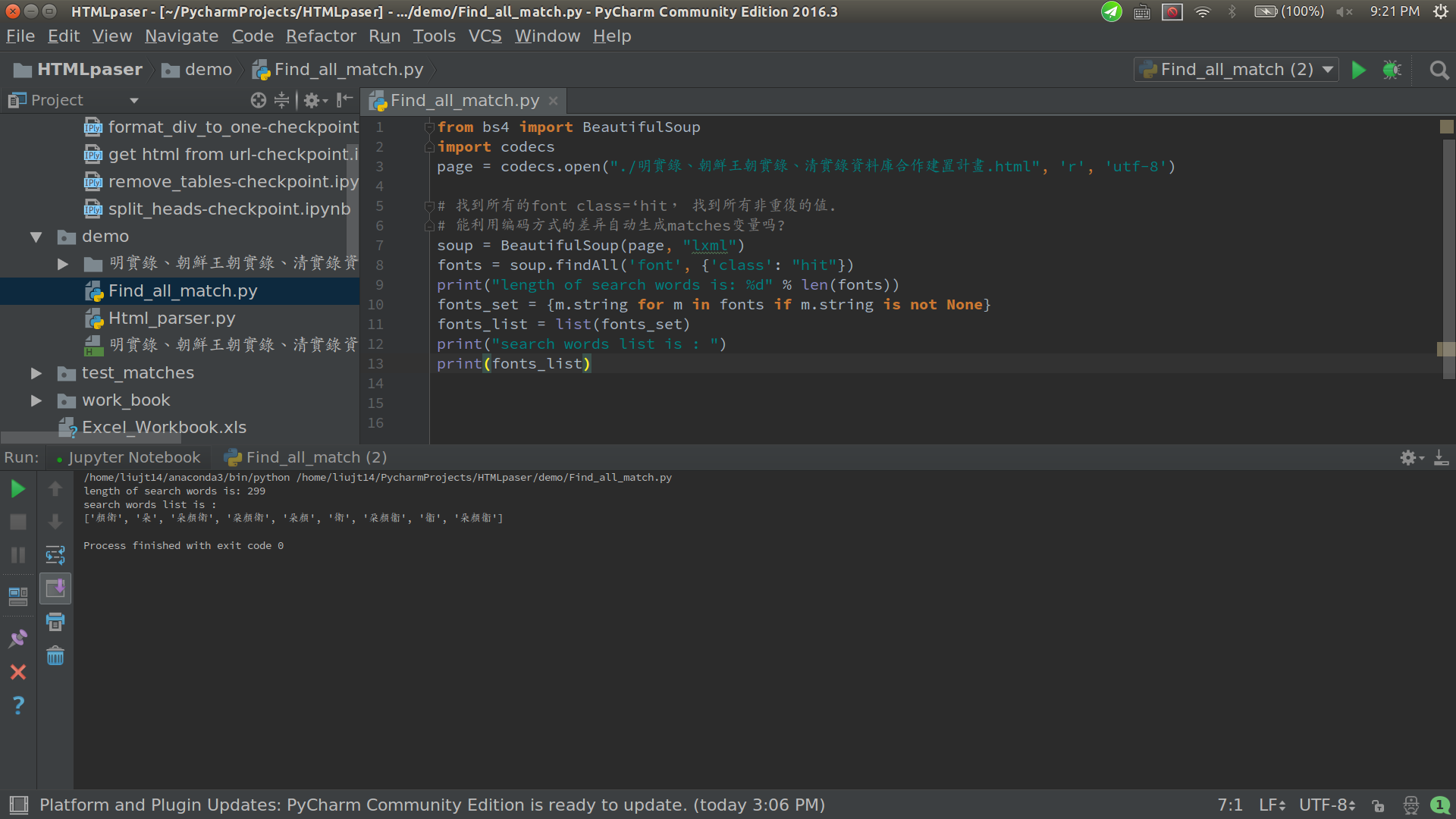
3. Run python file

After installed Anaconda, you can launch ***Spyder,*** and use it to open python files.



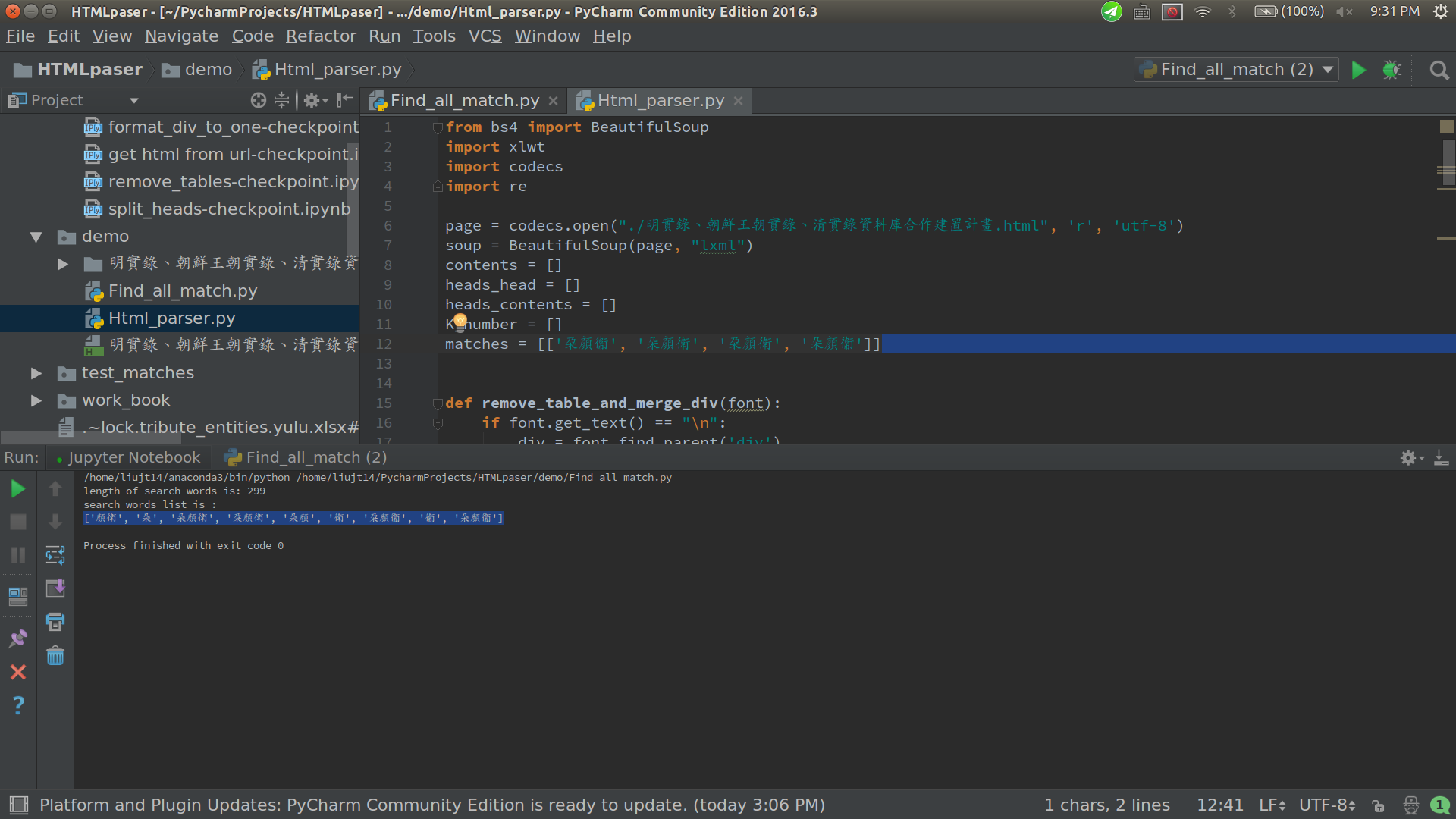
There are two .py files, First run Find\_all\_match.py, it will print all found words that we will use later.

I am using PyCharm.



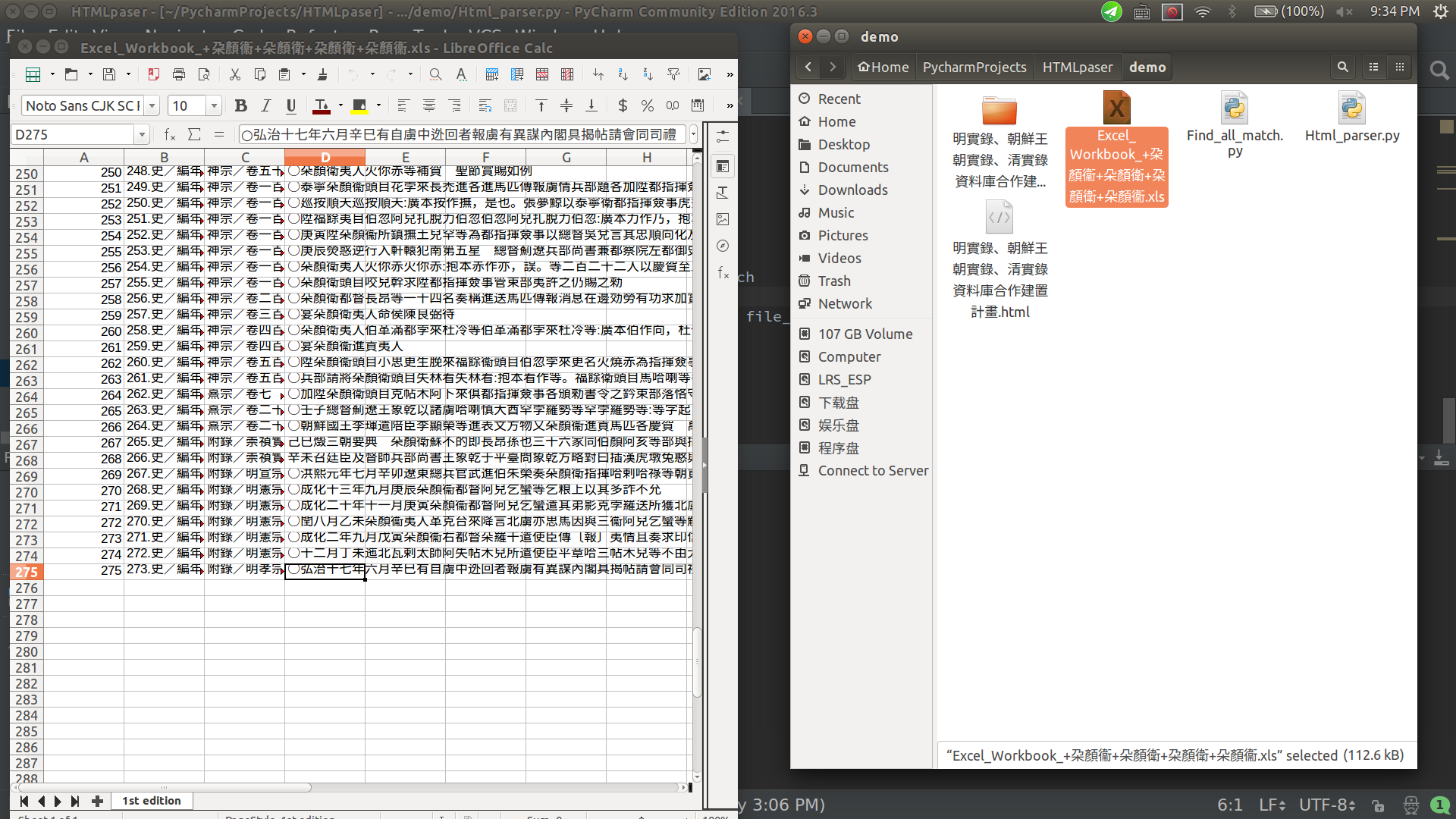
The list printed is the different types of same word.

['顏衛', '朵', '朵顏衛', '朶顏衛', '朵顏', '衛', '朶顏衞', '衞', '朵顏衞'], we need to search '朵顏衛', '朶顏衛','朶顏衞' or '朵顏衞' .

So we need to open Html\_parser.py, find the **matches** varible in *line 12* and type the words in it.

**Run** this .py file.

All done, enjoy!



**Notice:**

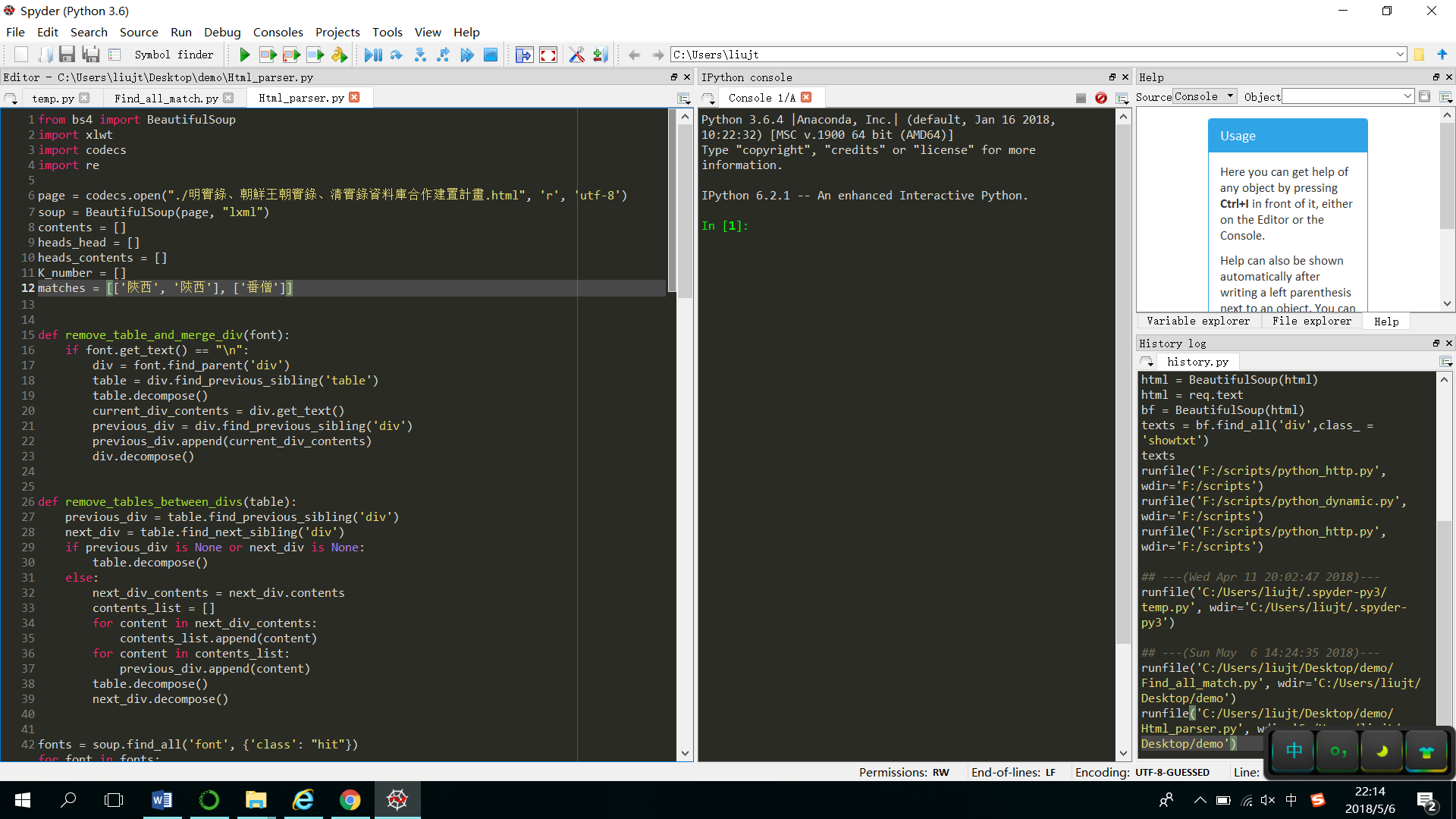
When there is more than one list of words. for example, search ‘陝西’run Find\_all\_match.py and got :

['番', '僧', '番僧', '陜', '西', '陜西', '陝西']

The matches varible in Html\_parser.py will looks like this:

matches = **[[**'陜西', '陝西'**],[**'番僧'**]]**

**Use Anaconda->Spyder to run python file.**



**Results:**

