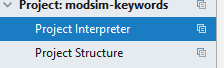
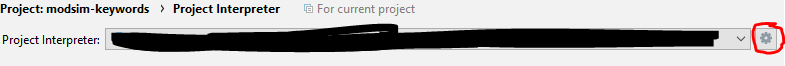
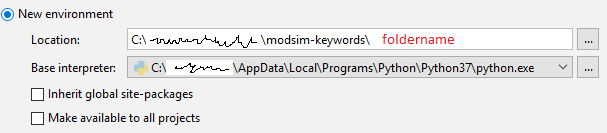
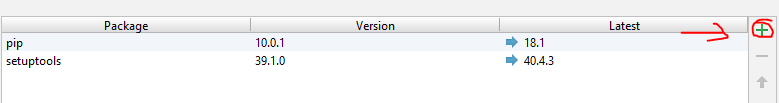
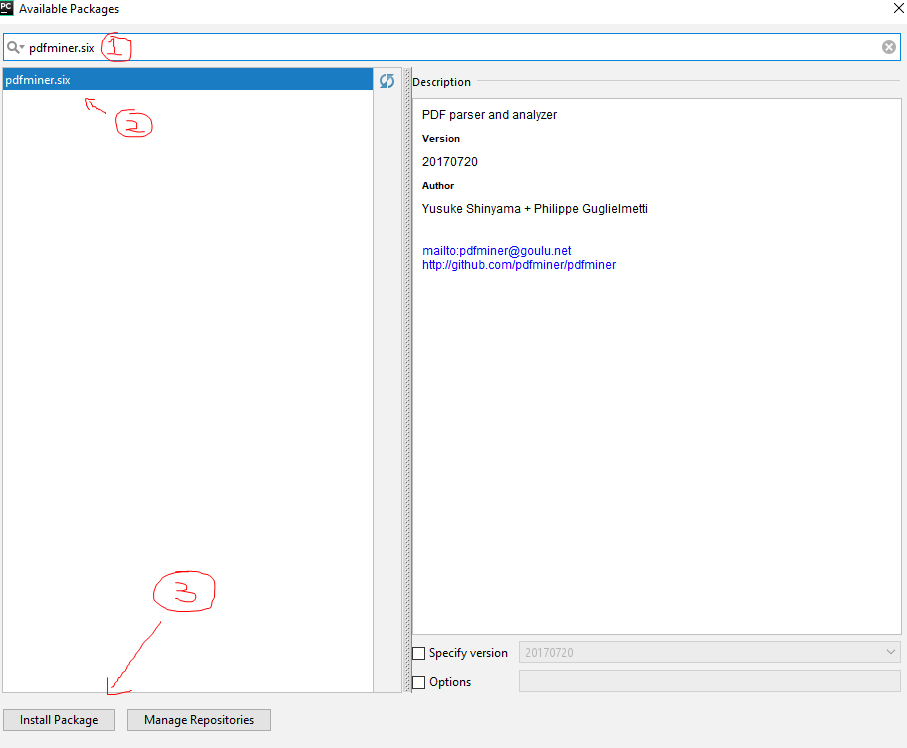
**Setting up interpreter (pyCharm)**

1. Enter **settings** (ctrl + alt + s)
2. Inside Project: modsim-Keywords, select **Project Interpreter**  
     
   
3. Click on **gears** and click on **show all**  
     
     
     
   
4. Click on the green +. To add a python interpreter  
   
5. Attempt to use the existing environment, if that doesn’t work you’ll have to start a new environment. (the existing environment is located in the **intenviroment** folder.)

**New Environment**

1. Select a location to setup the environment, inside the project folder is an ideal place. In the example below we have placed it in the project folder under foldername. However anywhere would be fine (as long as folder is empty).****
2. Press “OK”
3. Now we need to install the required packages, to install a package click on the green + 
4. To install a package, search for the package, select it, and then install it. Below is an example for pdfminer.six  
   

**Install the following packages using this method**

*[pdf keywords extractor]*

pdfminer.six

PyPDF2

bs4

magickwand

Wand

Pillow  
pytesseract

Html5Lib

Cv2

**NOTE:** You will also be prompt to install the popular MagicK wand software. Which can be found here <http://www.imagemagick.org/script/download.php>. MagicK wand is used to convert the pdf’s to image files for pytesseract (OCR) to read.

*[web file extractor]*

requests

wget

*[SkosReaderWriter]*

openpyxl

xmltodict