This is a more basic and detailed instruction of what needs to be installed to run Vicky’s smart data analytics script. I have used the laptop provided by the ChemE department (windows 10 operating system), so some of the instructions are focused on windows OS, but there should be equivalent steps for a MAC OS. The steps should be executed in the same order.

Download Vicky’s software from: <http://web.mit.edu/braatzgroup/SPA.zip>

**Python**

* Download and install Python3 and a Python IDE (integrated development environment). There is a variety of different IDEs (overview here: <https://www.guru99.com/python-ide-code-editor.html>)

**R and RStudio**

* Install R on your computer (<https://cran.r-project.org/bin/windows/base/>)
* Add the folder of R to your path (e.g. C:\Program Files\R\R-4.0.0\bin\x64). How to add a variable to your system path is described here: <https://www.architectryan.com/2018/03/17/add-to-the-path-on-windows-10/>
* RStudio is not necessarily required, but it is a good way to check whether the r packages have been installed properly, so I would recommend installing this as well (<https://rstudio.com/products/rstudio/download/#download>)

**Rpy2**

* Install rpy2 package with the command “python –m pip install rpy2” in the command prompt. The command prompt can be opened by searching for command prompt in the start menu.
* Restart required for package activation

**Acepack**

* Download package from https://cran.r-project.org/web/packages/acepack/index.html
* Add package to library folder of your R installation (e.g. C:\Program Files\R\R-4.0.0\library)
* Adjust package path in ace\_R.py (lib\_loc = x) to “C:\Program Files\R\R-4.0.0\library” (I assume this is all you need to do for MAC OS)
* Adjusting the package path did actually not work properly for me, so instead I replaced the lines 19-25 in Vicky’s code by the simple command ace = importr('acepack') as used before for the utils package. Windows utilizes “\” instead of “/” for paths and “\U” as used in “\Users” is a python command (in case you stored the package somewhere under user), so this leads to some trouble. If the packages are properly installed, ace = importr('acepack') should work as well.

**SPLS**

* Download package from https://cran.r-project.org/web/packages/spls/index.html
* Add package to library folder of your R installation (e.g. C:\Users\chemegrad2018\anaconda3\Lib\R\library when using the department laptop)
* Adjust package path in SPLS.py
* Similar to acepack, the second option can be used for Windows OS
* Typing library() in Rstudio shows the different libraries installed. This can be used to check whether acepack and spls are installed

**Additional packages**

* Install sklearn package with the command “python –m pip install sklearn” in the command prompt
* Install seaborn package with the command “python –m pip install seaborn” in the command prompt
* Install statsmodels package with the command “python –m pip install statsmodels” in the command prompt
* Install xlrd package with the command “python –m pip install xlrd” in the command prompt
* The following two packages were already preinstalled with my python version (python 3.6), but since Vicky mentioned them specifically in her e-mail, the same goes for:
  + Install matlab package with the command “python –m pip install matlab” in the command prompt
  + Install tensorflow package with the command “python –m pip install tensorflow” in the command prompt