Documentation for BankRisk GitHub repository:

5/31/16

Special note on installing Anaconda:

* This is the software that I have been using. Recently, it has undergone major modifications in what was formerly called ipython notebook. It is now called jupyter notebook. (I think this switch must exist with Canopy Express, too.) The changes are informally referred to as The Big Switch because the developers are reconfiguring the logic of the software. As such, there are some changes that have affected how the software runs. In particular, there have been changes to the default working directories. This note shows how to return to the prior notebook working directory. (Sorry, I only ran into this issue a couple of weeks ago.)
* Choosing between Python 2 and Python 3 versions. This is actually a fairly big deal to many, many python users. I am agnostic. Choose either one. Believe it or not, it appears to me that all of the python veterans seem to stick with Python 2. There are two main differences. (1) The print command requires parentheses in Python 3. Before, parentheses would cause errors in Python 2, but I think that is fixed now. (2) Division: In Python 3, division defaults to floating point. In Python 2, division defaults to integer. (In Python 2, 5 divided by 2 is 2, not 2.5.) To fix this, simply add “from \_\_future\_\_ import division” (without quotes) at the top of the code. You will see this in several of my code modules.
* After installing Anaconda 2 or 3, go to a Windows command prompt window (click on “Command Prompt” after right-clicking the Windows icon in the lower left corner of the desktop) and type: “conda update conda”, “conda update --all”, and “conda install pandas-datareader”. (Don’t type the quotes, do these one at a time and follow the prompts.) This will update all of the software packages and add the pandas-datareader package used for downloading data from the FRB St. Louis FRED database.
* Also in the command prompt window, type “jupyter notebook --generate-config”. This will create a configuration file called “jupyter\_notebook\_config.py” in your “C://Users/yourname/.jupyter directory. Open that file in Windows notebook (or any other text editor). Do an “edit, find” for “notebook\_dir”. You will see:

# c.NotebookApp.notebook\_dir = '’

This line must be modified to look like:

c.NotebookApp.notebook\_dir = 'C://Users/denni'

This is my Users directory in Windows explorer. You should use the name for your Users directory. Notice that I have removed the # symbol at the start of the line. The # symbol is a comment in Python. Removing the # symbol activates the line. (Also remove the space after the # symbol.) Now, when you start Jupyter notebook, it should open a browser window showing your “C://Users/yourname” directory. If not, then the reconfiguration described here failed. Please try again.

Files:

* BankRisk\_Demos\_pt1.zip: Contains demo and exhibit excel files for chapters 1-6 (credit risk). These are the files I used to develop all of the exhibits for the book. Many use excel data tables (which I find to be indispensable.) Only one file is incomplete: chap05\_exhib\_2011allbanks\_pds\_template.xlsx. Two worksheets within that file were too big to upload to GitHub. The first worksheet needs to be populated with all of the data in 2011allbanks\_pds\_20111231\_roe0mm2.csv. (I left the first 20 rows or so to show how the data must be arranged.) There are roughly 7000 rows here. Next, there is a huge worksheet named 200903\_201412 that must be populated with all of the data in allbanks\_f2468\_roe0mm2\_2009\_2014.csv (74MB and 165,000+ rows). Again, I left 20 or so rows to show how the data must be arranged. (Note: Those files will be generated by running the ipynb code for the PD/LGD models.) Place these files into the “C://BankRisk” directory.
* BankRisk\_Demos\_pt2.zip: Contains demo and exhibit excel files for chapters 7-11 (interest rate risk). Place these files into the “C://BankRisk” directory.
* BankRisk\_ipynb.zip: All of the ipython (Jupyter) notebook code files. These files should be placed into a subdirectory called: C://Users/yourname/bankrisk. The python software will look for the code files in the C://Users/yourname directory. When it opens you should see bankrisk listed as a subdirectory. Just click on it to work with the code files. (Do not attempt to run these files if you are not proficient in python and ipython notebook. This is not the material to use if you are just beginning to learn python.)
* BankRisk\_processingtables.csv: These are all of the ancillary data tables that are referenced by the ipynb files. They should all be placed into a root directory called “C://BankRisk”.
* df2\_2001\_2009.7z and df2\_2010\_2015.7z: These are the Call Report raw data extracts for the indicated years for all banks in the country. (I also provided the python code that created these files from the bulk data files in case you are curious.) You must use 7-Zip open source software (which is superior in many ways to WinZip) to un-compress these files. Put them in the “C://BankRisk” directory.

How to get started:

1. Run br001b\_CallDataExtr\_concat.ipynb. This concatenates the two large df2… files into one even larger one which is used for everything.
2. Run br001c\_CallDataCalcs\_AllBanks\_20160511.ipynb. This processes the raw Call Report data stored in the df2… files into calculated fields. For example, the Call Report shows year-to-date income and expense numbers. This processing will convert them into quarterly numbers to allow rolling 4q averages to be calculated. (This takes a long time.)
3. Definitely study: br003ma, br003mb, and br003mc code files. They are the same code, but they are configured to illustrate how to set up specialized BHCs, large aggregates, or small aggregates. There is a large block of code near the top of the files where all of the selections are made. The output from these three files is discussed in detail in the book.
4. Generally, if you scan through any of the ipynb notebook files, you should recognize the outputs from the book. This will give you a good idea about what each module is used for. Please note that the “univstress” modules provide output for the full models, the 3-variable models, and the 4-variable models. You can select any of those you like.
5. Good luck! Let me know if you run into problems.