MSDS610 Data Engineering

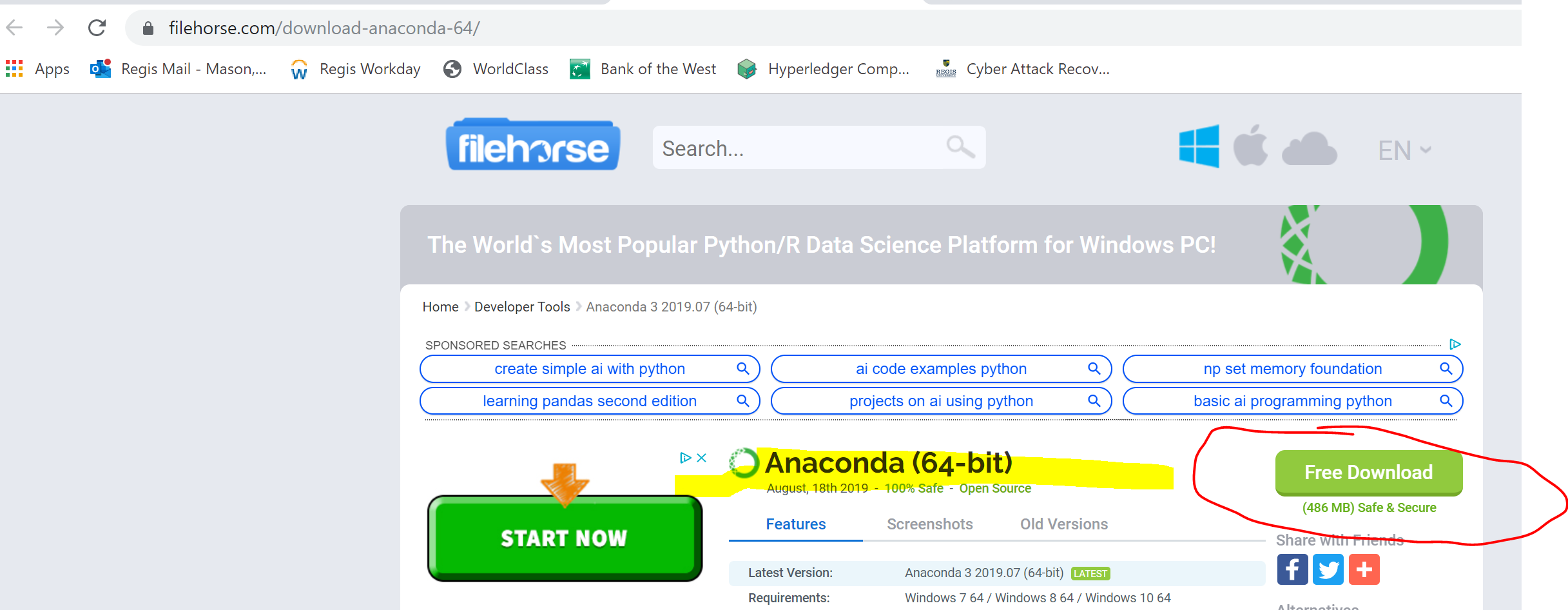
Week 4 Lab Supplemental – Part 2

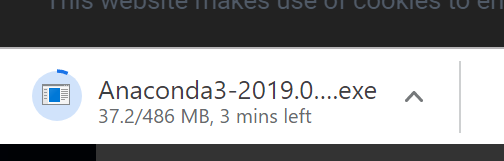
as of 9/20/2019

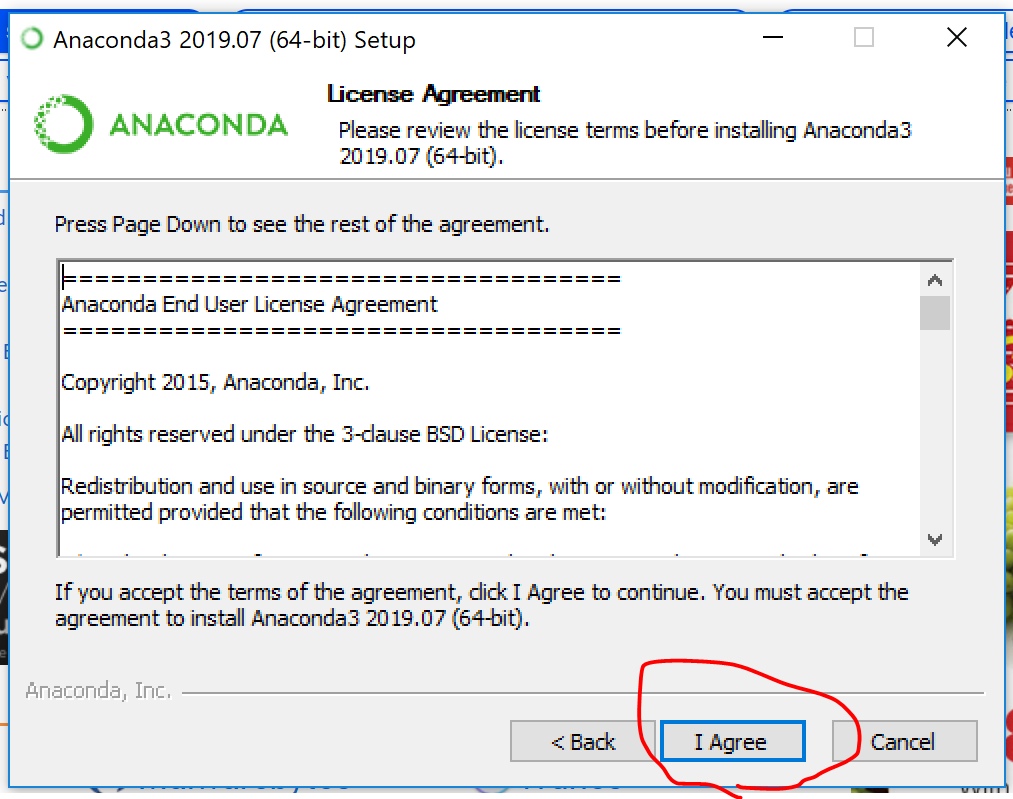
**In this document, you will:**

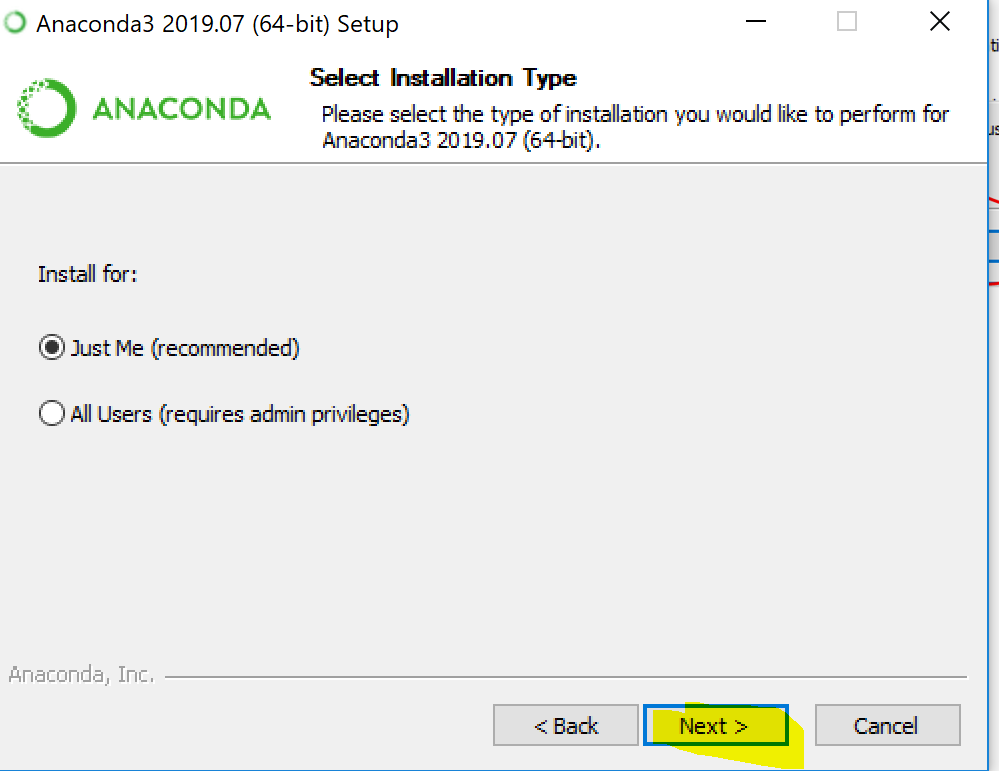
* Install Anaconda in Windows which automatically installs Jupyter Notebook
* Connect to the MongoDB with Python using your Jupyter Notebook
* Query data with Python from MongoDB

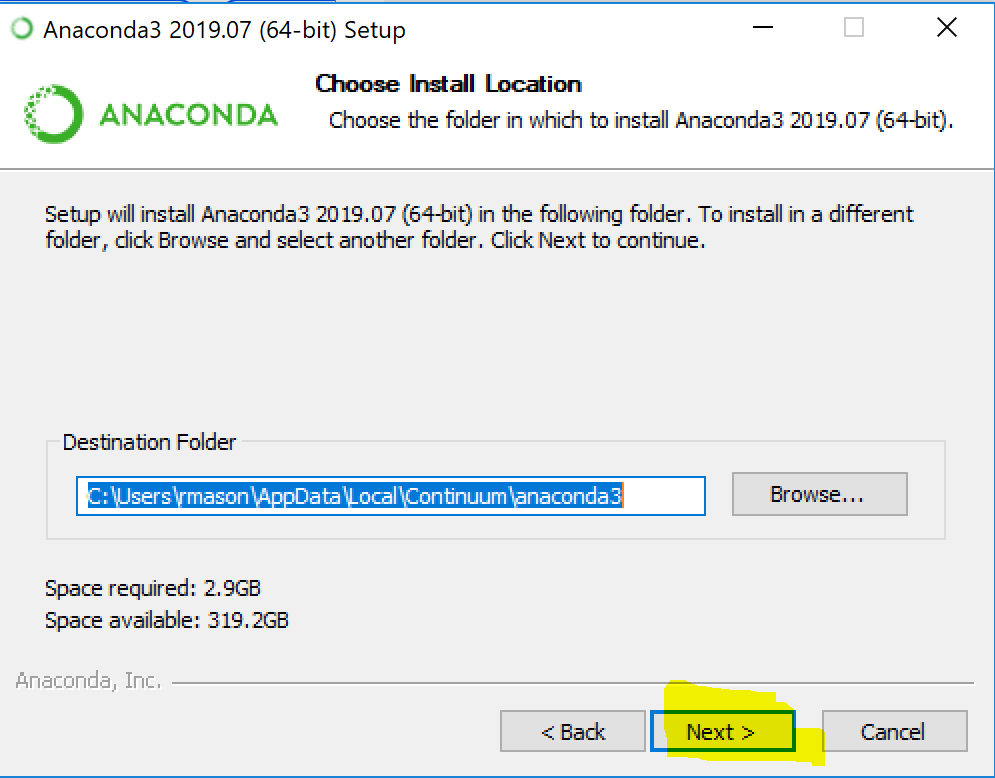
**Install Anaconda on your Windows PC, if you have not installed it already.**



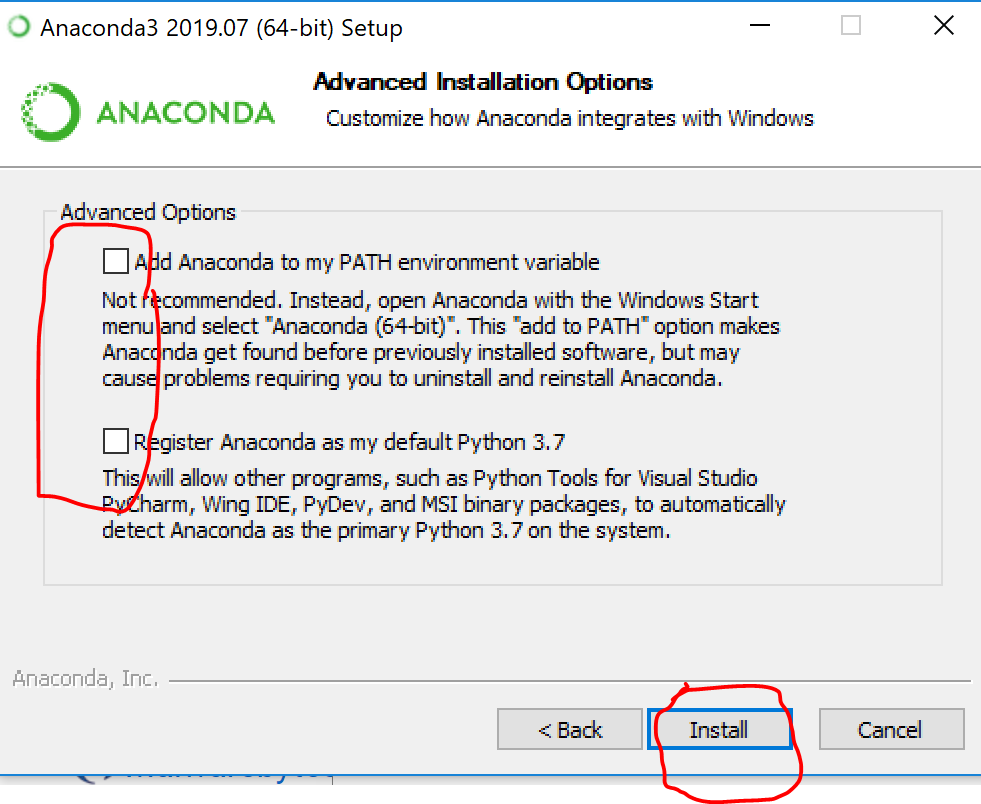


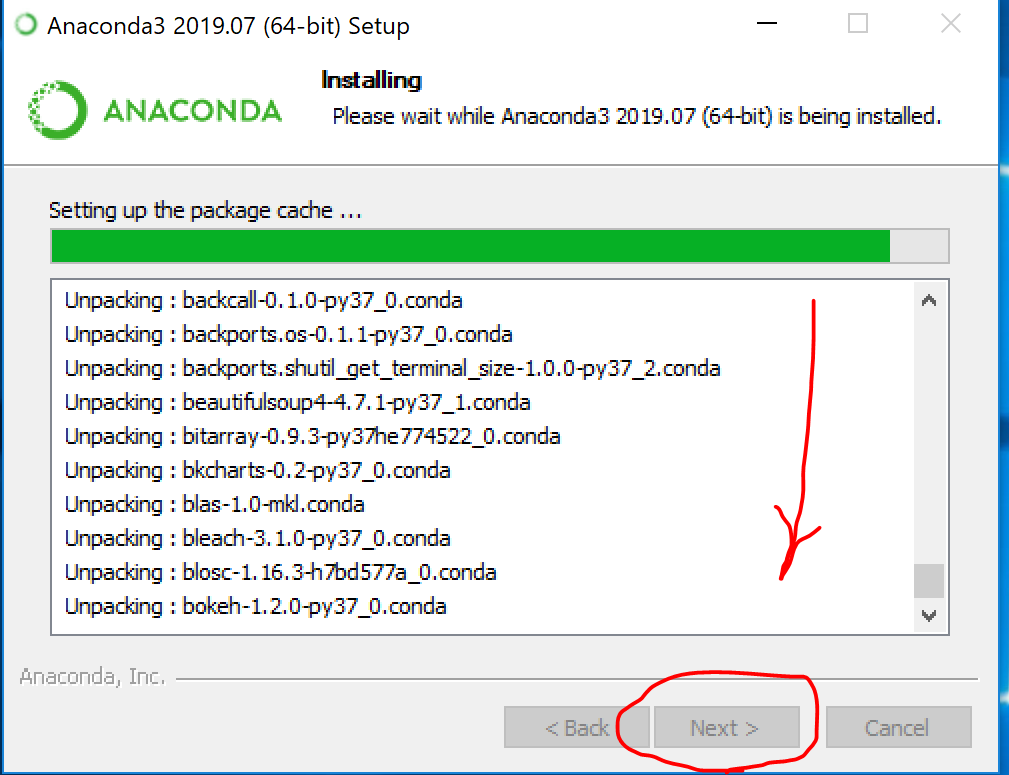






**Before checking boxes, read about these options. I left them unchecked.**

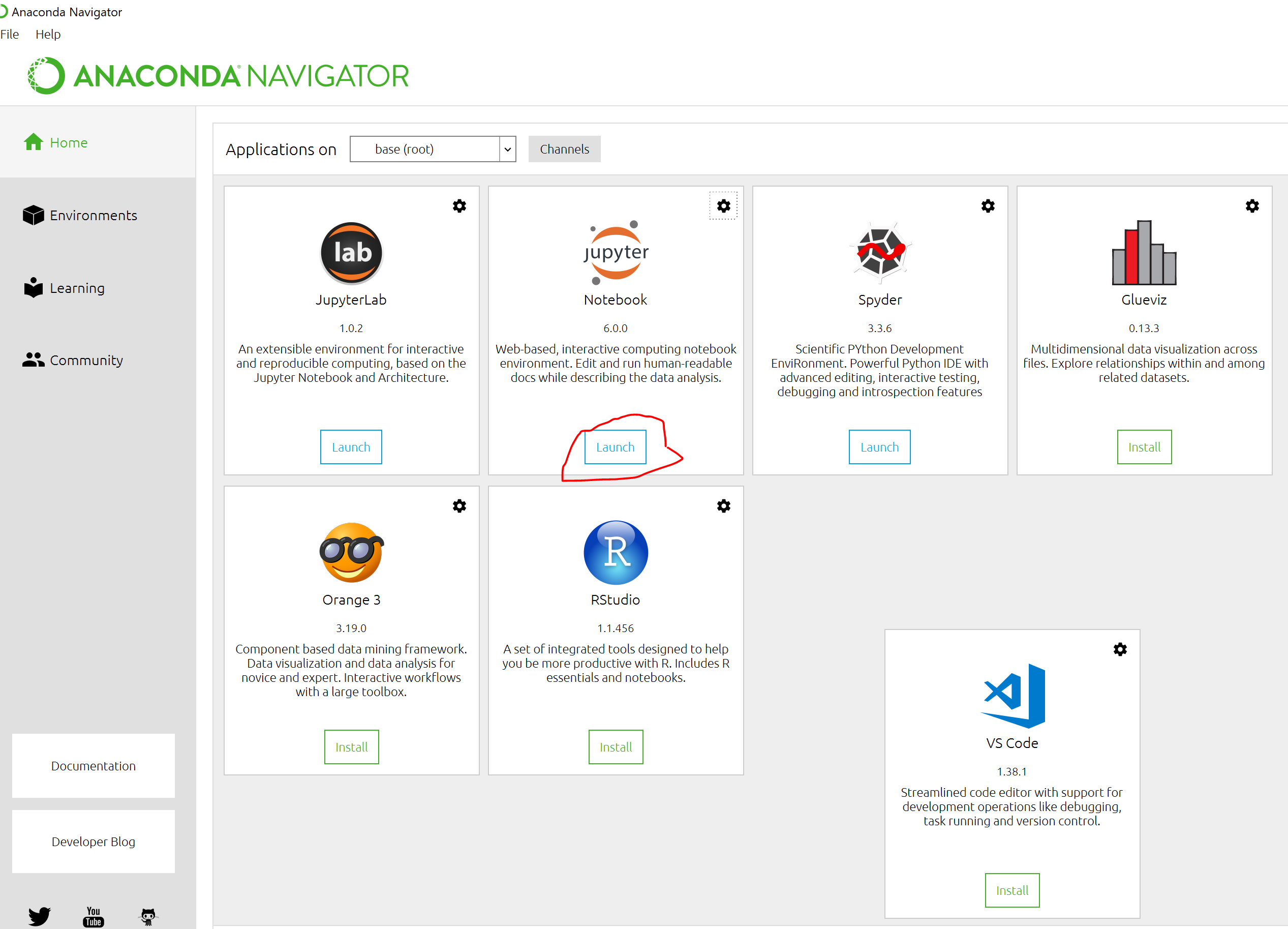




**After the installation is complete, click on the Next > button.**

**Note: The installation of Anaconda on your Windows PC can take at least 10 minutes.**

**Go to your PC menu and select the Anaconda Navigator. Then, Launch Jupyter Notebook.**



**After Jupyter notebooks starts, you will see a window as shown below. Select New and pick the Python 3.**



**I clicked on the File Menu option and then Save as… I called my new notebook “DataEng”.**

**You should not have to install pymongo, however it you get an error when you run about it missing, try installing it by putting this into a cell and run the cell.**

***conda install pymongo***

**I entered into the cells a few print commands and then a simple python program that accesses mongoDB. Notice that I changed the name of the database and collection name to match what I did in the Week 4 Supplement Part 1.**

*from pymongo import MongoClient*

*# create connection*

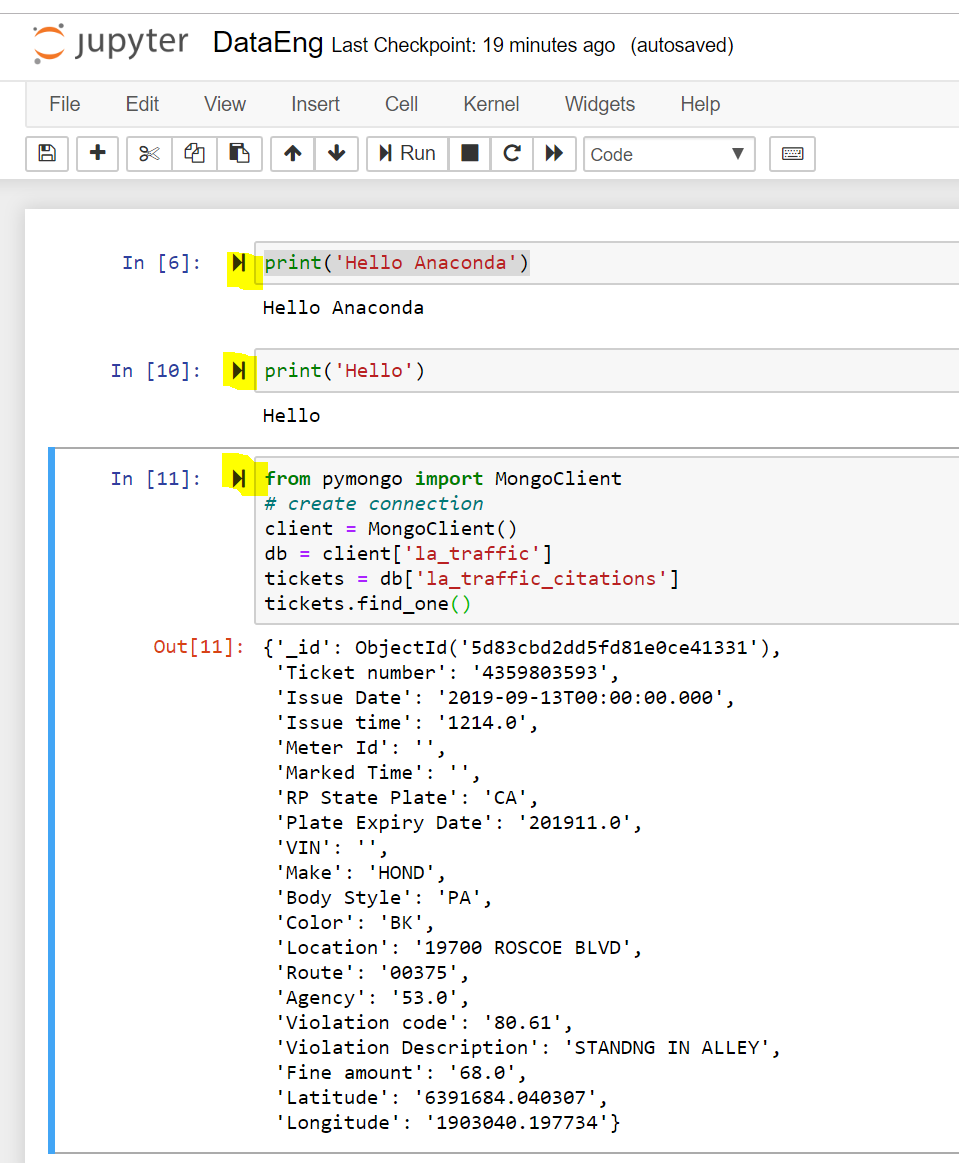
*client = MongoClient()*

*db = client[****'la\_traffic'****]*

*tickets = db[****'la\_traffic\_citations'****]*

*tickets.find\_one()*

**To add new cells, click on the Insert Menu option and select Insert Cell Below. I entered three test cells as shown below. You can run each cell by clicking on the run indicator in front of the cell. Move the mouse above the indicator to see what it does. Notice that the results will be displayed below the cell.**



**At this point, you have a working Python program that can access a MongoDB database and collection. When you are done testing, save your Notebook and logout.**

**Please continue with the lab exercise and figure out how to plot results. Please provide screen shots of your code and graph plot.**

**This is the end of the Week 4 Supplement – Part 2.**