Database Basics

Installation and setup guide

AUTHOR

Table of Contents

[**Introduction** 2](#_Toc461725448)

[**Setup Introduction** 2](#_Toc461725449)

[**Anaconda Installation** 3](#_Toc461725450)

[Windows 3](#_Toc461725451)

[MacOS 4](#_Toc461725452)

[**IPython/Jupyter Installation** 7](#_Toc461725453)

[**Opening and Working in Jupyter Notebook** 8](#_Toc461725454)

[**Loading Workbooks from GitHub to Jupyter Notebook** 10](#_Toc461725455)

# **Introduction**

The goal of the “BIG DATA AND SOCIAL SCIENCE RESEARCH: THEORY AND PRACTICAL APPROACHES” book is to bring computer scientists and social scientists together to provide a practically oriented overview of the analytical and statistical tools associated with big data for social science students. We use real data used for real world policy problems to analyze some of the vast new sets of data on human beings. We show how this data can be collected, integrated, and analyzed in a scientific fashion. In the accompanying exercise notebooks, we walk students through key programming and analysis techniques so they have working examples they can apply to their own work.

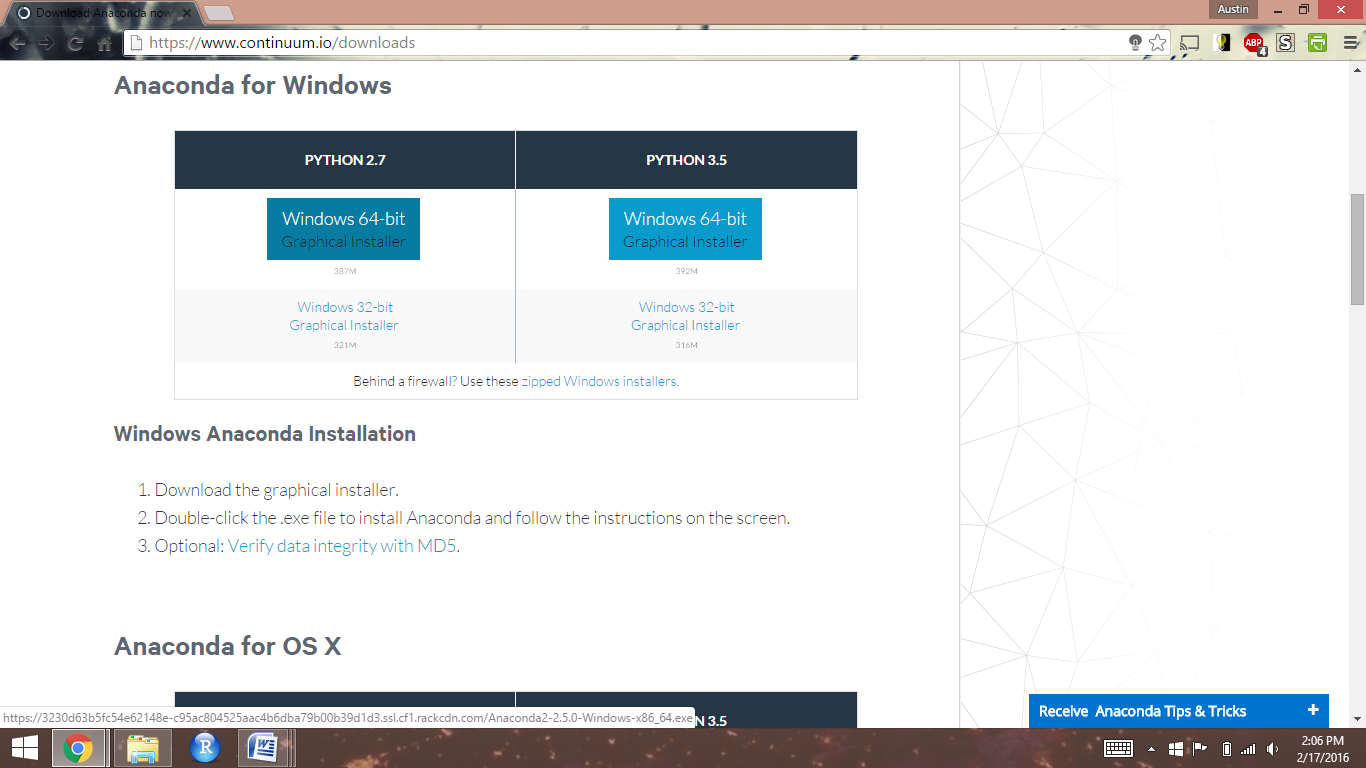
# **Setup Introduction**

Before you start working on exercises, you will need to download and install several different software packages to interact with and analyze data.

The exercises are written in IPython notebooks, interactive web pages that contain a combination of explanation and code that walk you through key programming and analysis concepts, providing explanation and context along with code samples you can edit and run as you progress through each exercise. To run these notebooks, you will also install Python and Jupyter. You can install these packages directly, but it is much easier to install them together in the Anaconda Python distribution. Anaconda includes the basic Python programming language, a more full-featured and robust command line Python tool named IPython, as well as the easy-to-use **Jupyter Notebook** software environment in which you will be writing your code and viewing the output.

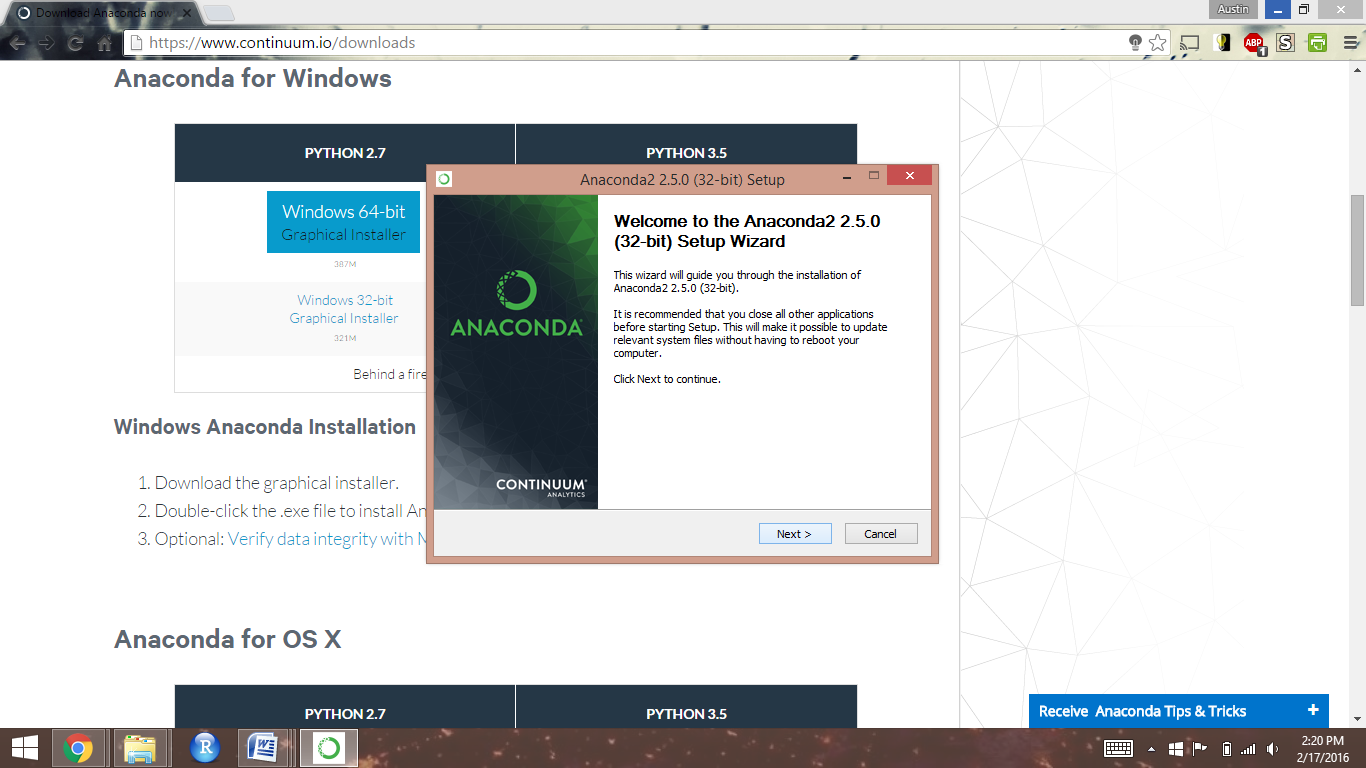
*Note: The installation instructions below use Windows screenshots. Installation on a Mac follows the same steps.*

# **Anaconda Installation**

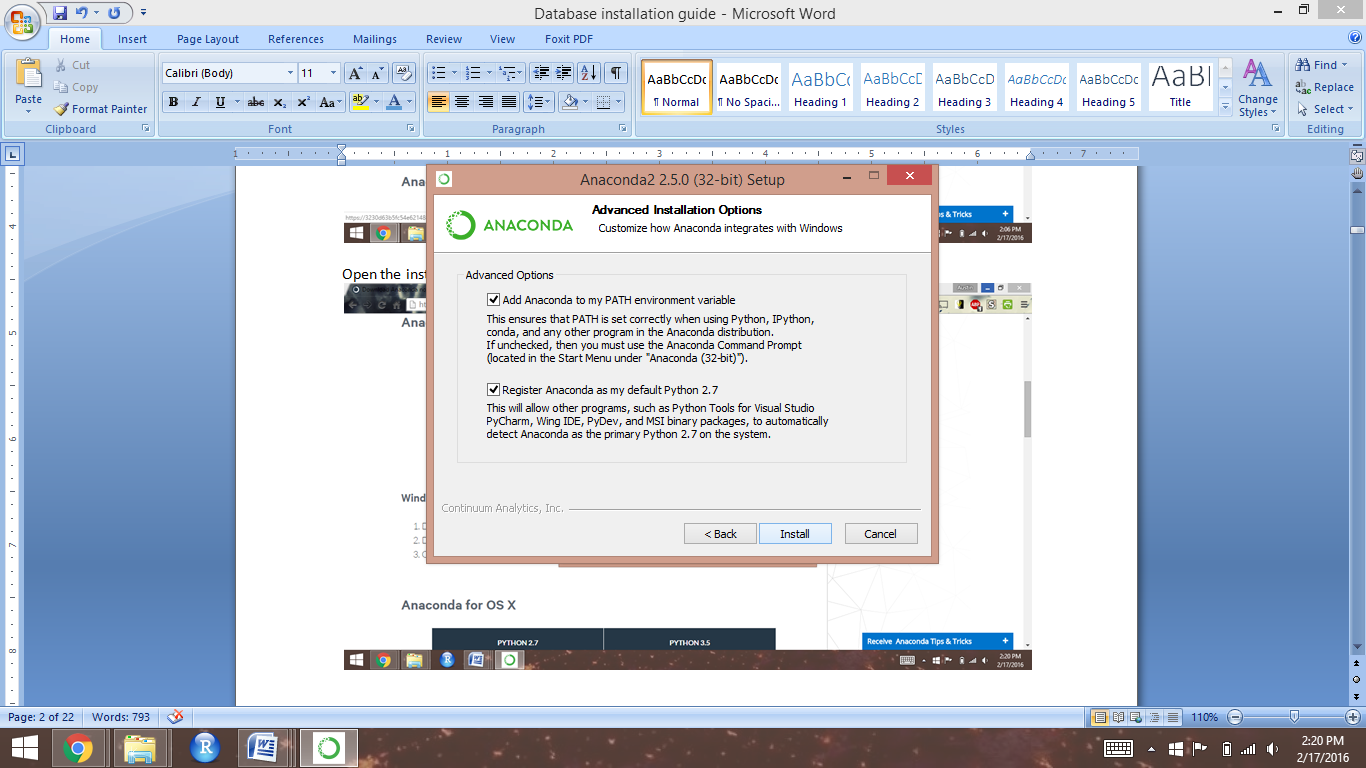
Go to [Continuum Analytics](https://www.continuum.io/downloads) and select Anaconda Python 2.7 for the appropriate OS (we are updating the exercise notebooks to support Python 3 as well, but not all have been updated yet).

## Windows

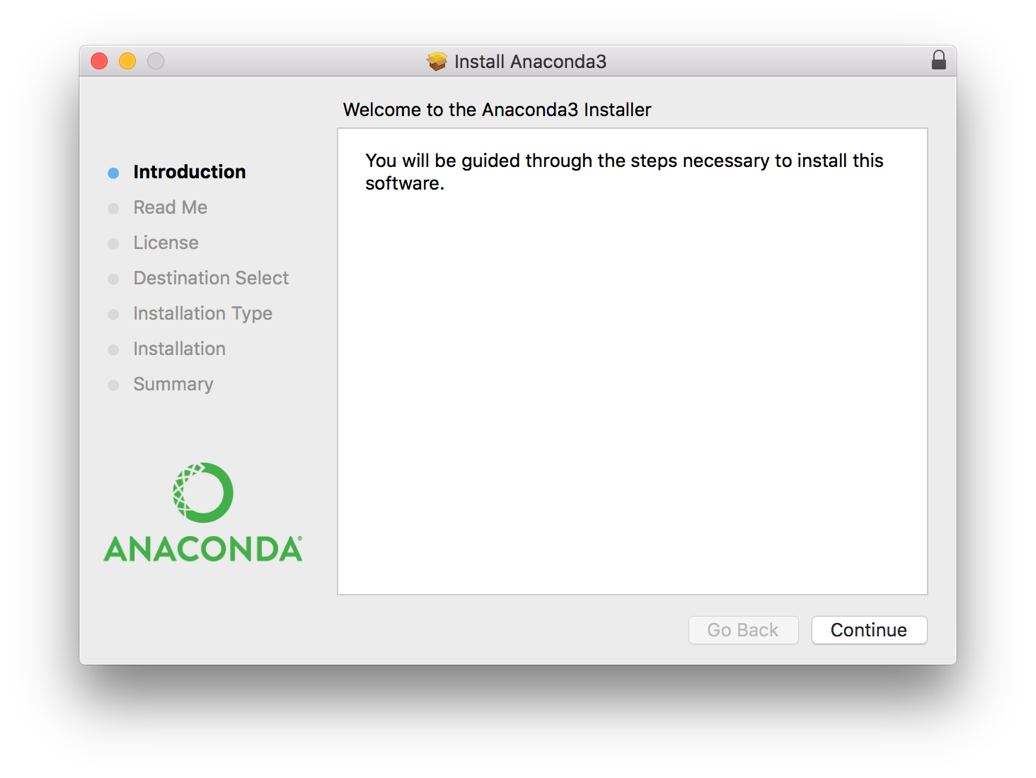
Open the downloaded executable file and move forward with installation



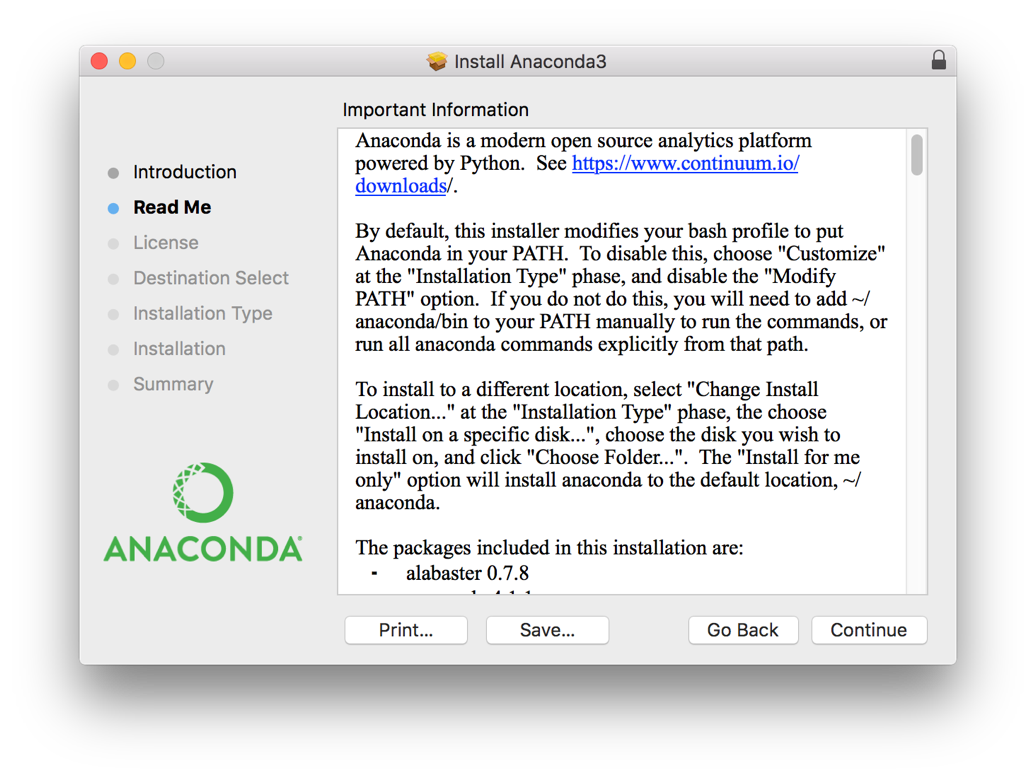
Keep the default advanced options checked, and click next until installation has finished



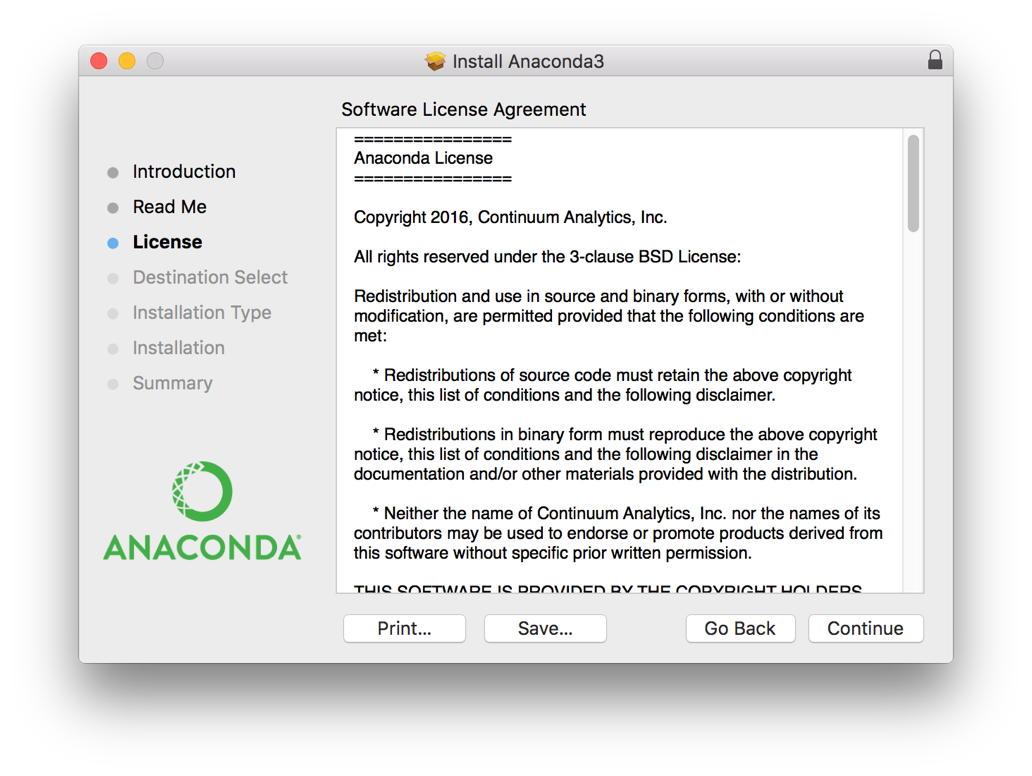
## MacOS

Open the downloaded “.pkg” file and click “Continue”.

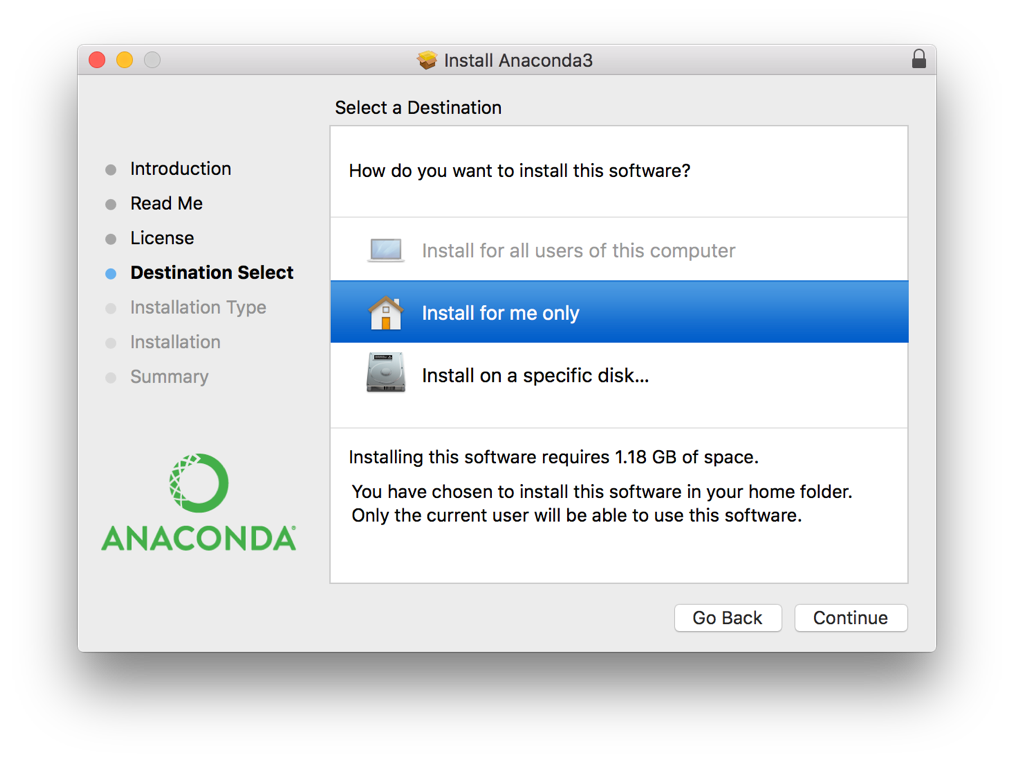
View the Read Me then click “Continue”:



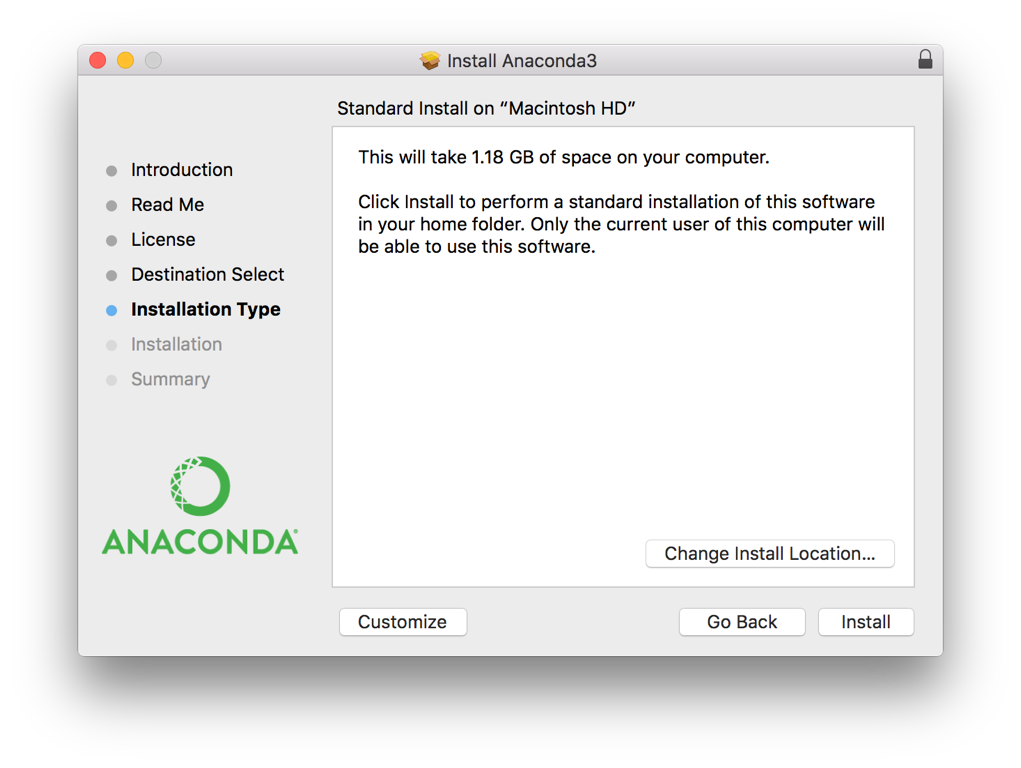
At the License screen, click continue, then accept the license:



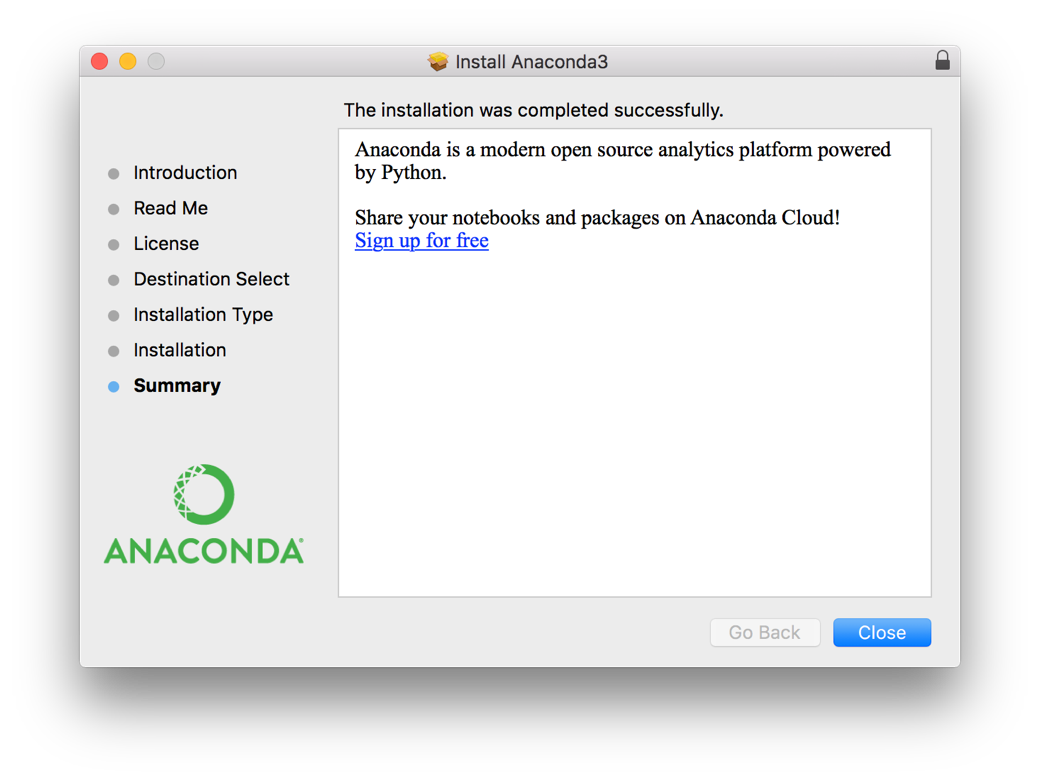
Select a destination for the installation (we generally recommend “Install for me only”, which makes a local copy of all of the Python files that is only for the current user, separate from the computer’s Python installation).



Click “Install”:



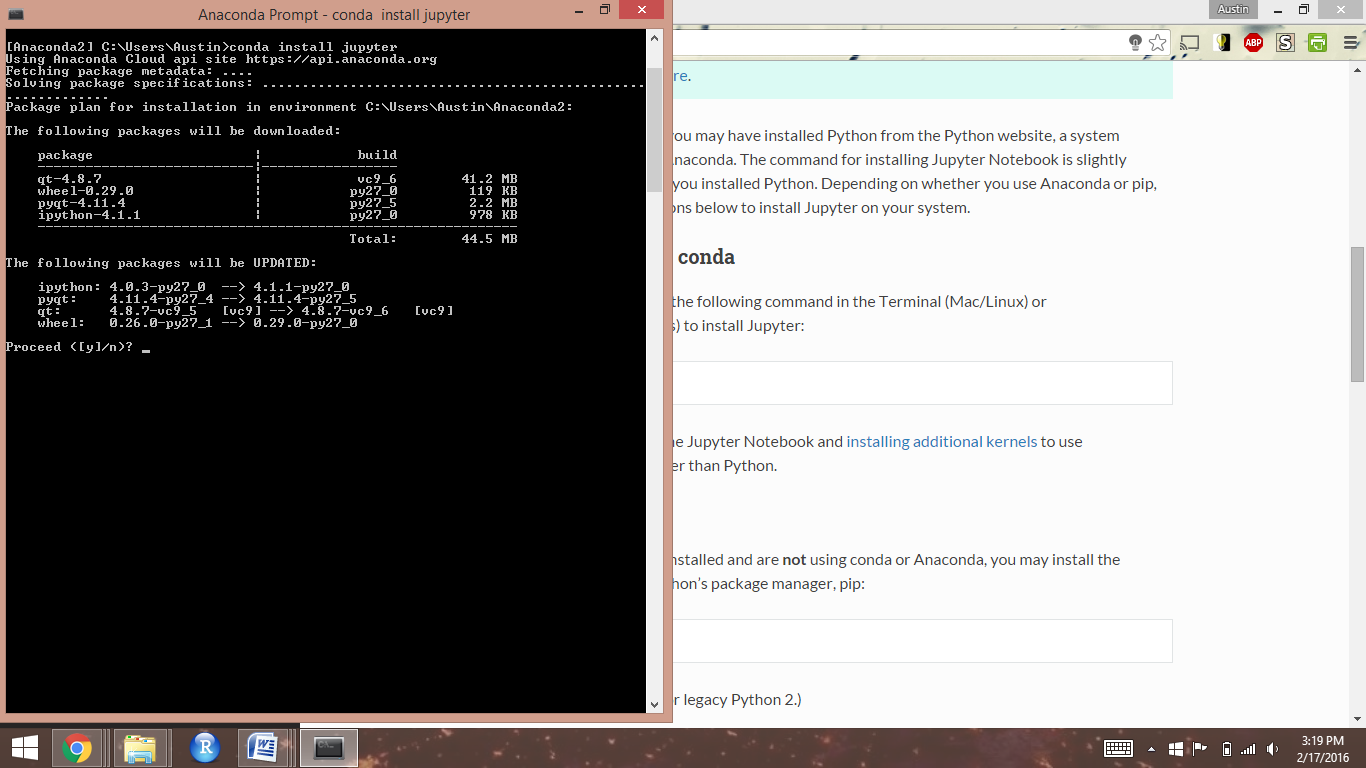
And you’re done!



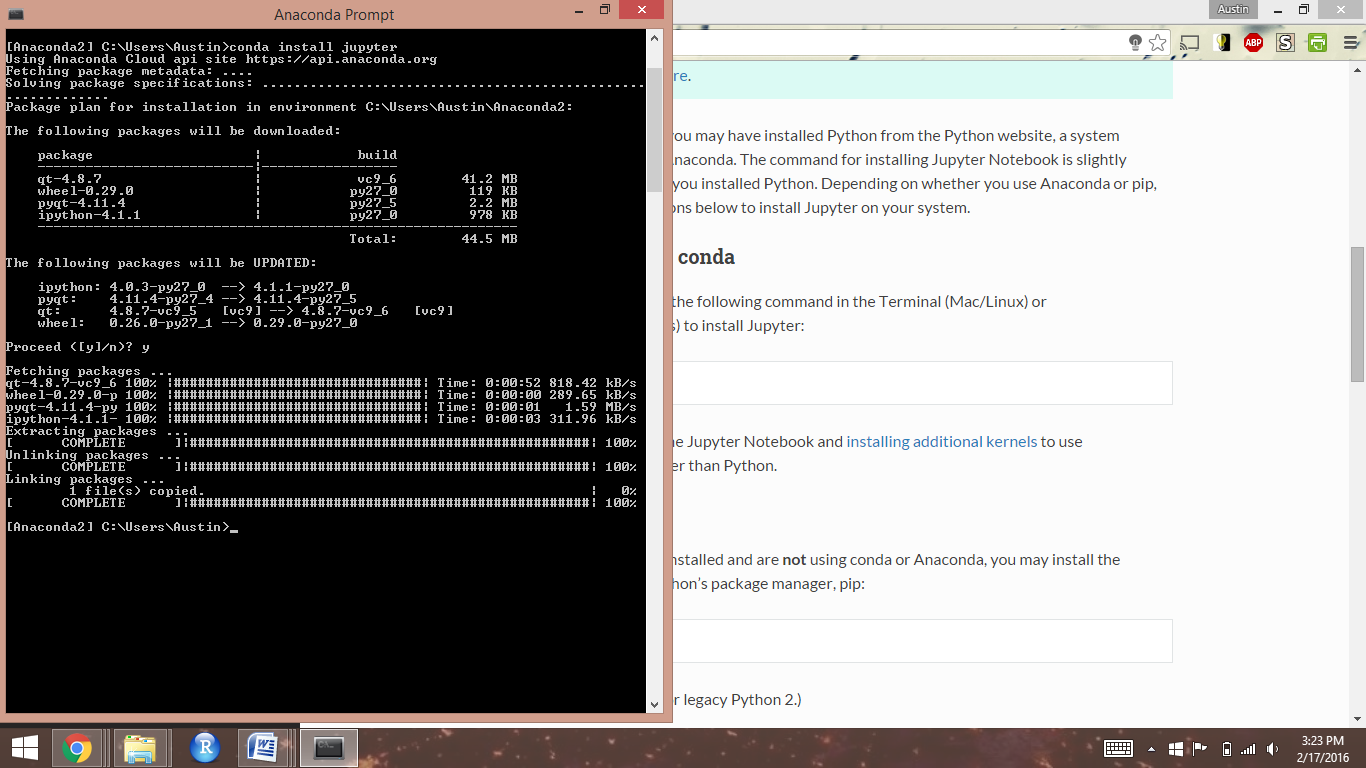
# **IPython/Jupyter Installation**

To install python packages, you’ll open a shell on your computer and use the conda command, provided by Anaconda. First, open a command shell on your computer (the windows command shell is pictured, but the commands are the same whether on Windows or Mac).

To make sure that jupyter is installed, type "conda install jupyter" and then hit enter. Once it has run the prompt will ask you to proceed. Type "y" and hit enter to finish the installation.



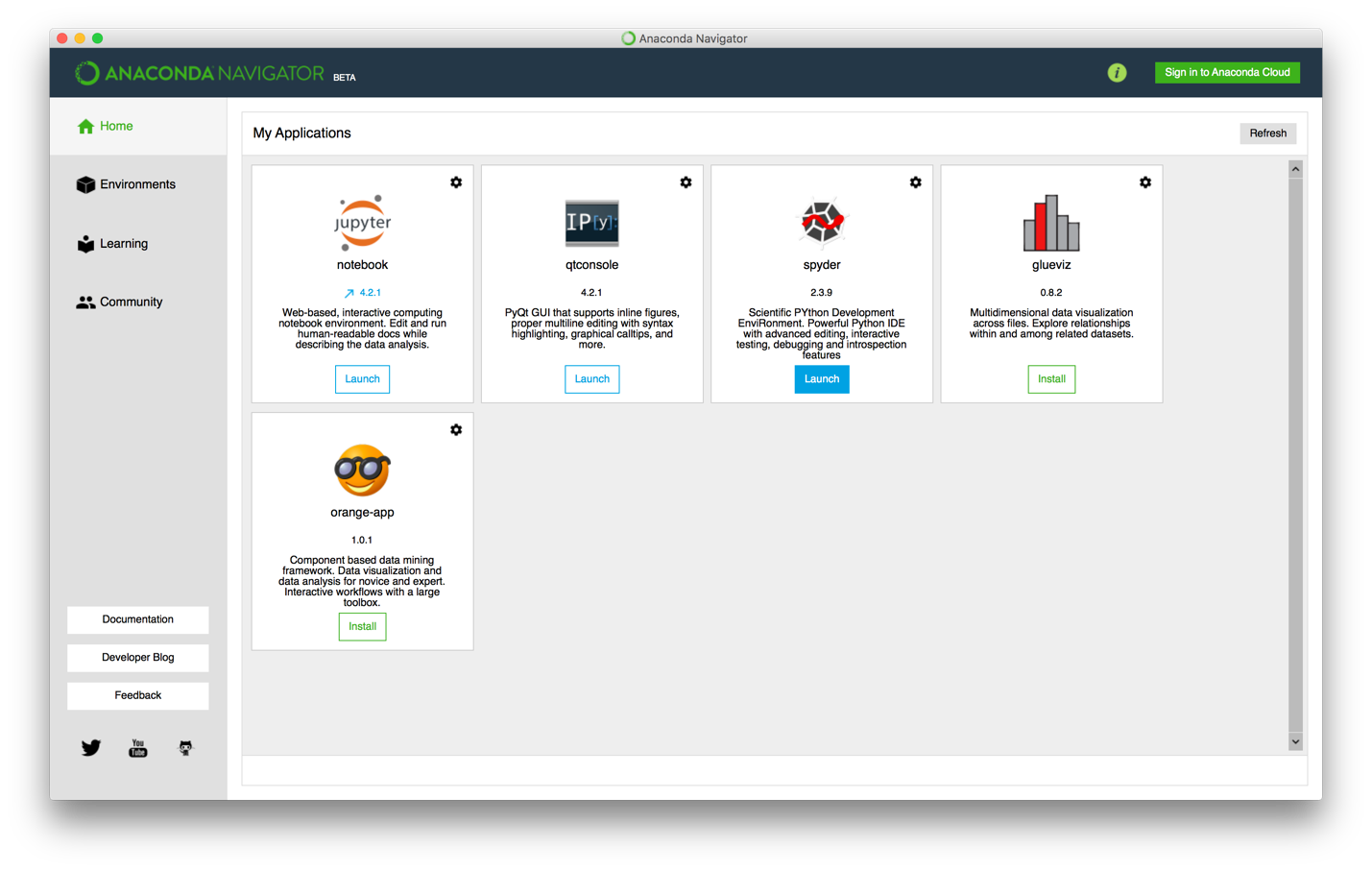
Once the prompt returns to your original directory the installation is complete and you can close your command prompt window:



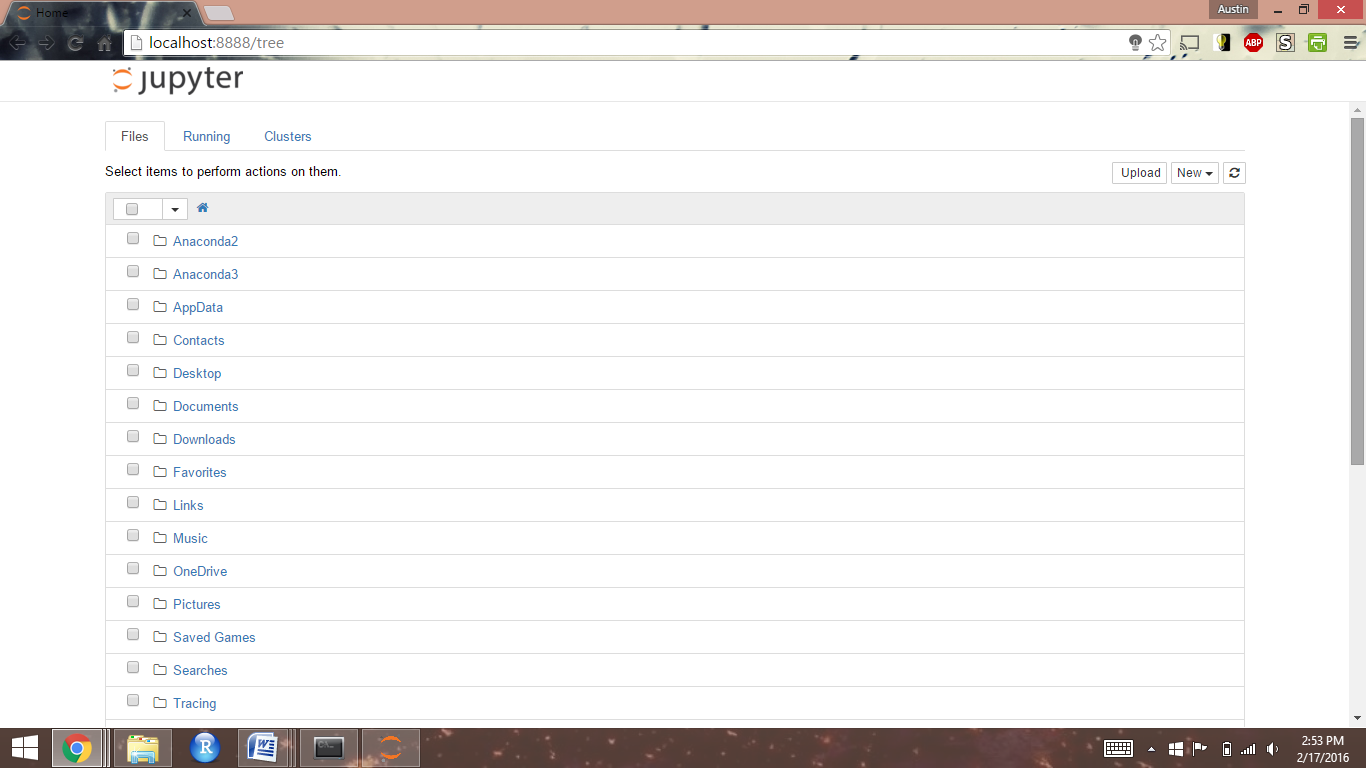
# **Opening and Working in Jupyter Notebook**

To open Jupyter Notebook, go to your Program Files (Windows) or Applications (Mac OS) and open the Navigator program installed by Anaconda.

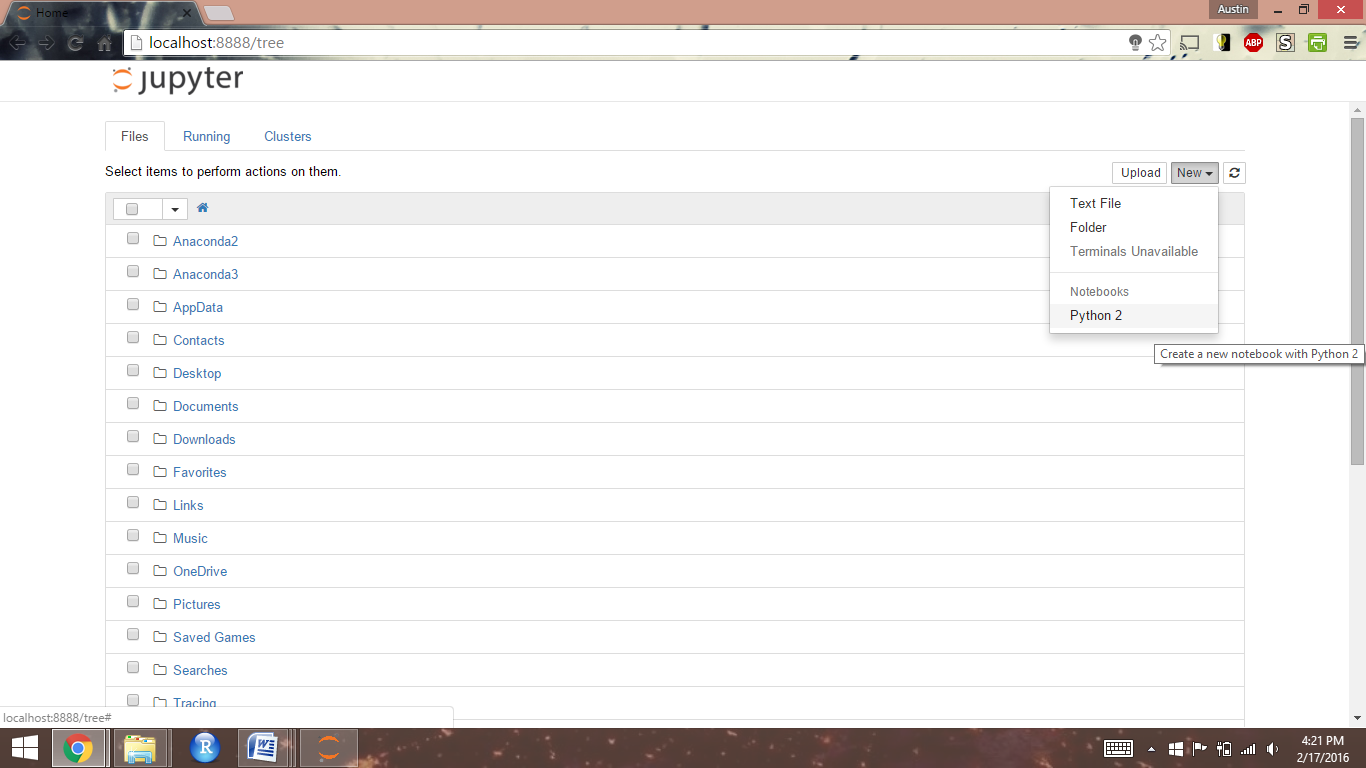
In the Navigator window that opens (same on Windows and Mac), click the “launch” button for “Jupyter notebook”:



A command prompt box will open. Wait a few seconds and it should automatically open the Jupyter Notebook main page in your web browser. Leave the command prompt box open while you use your Notebooks.

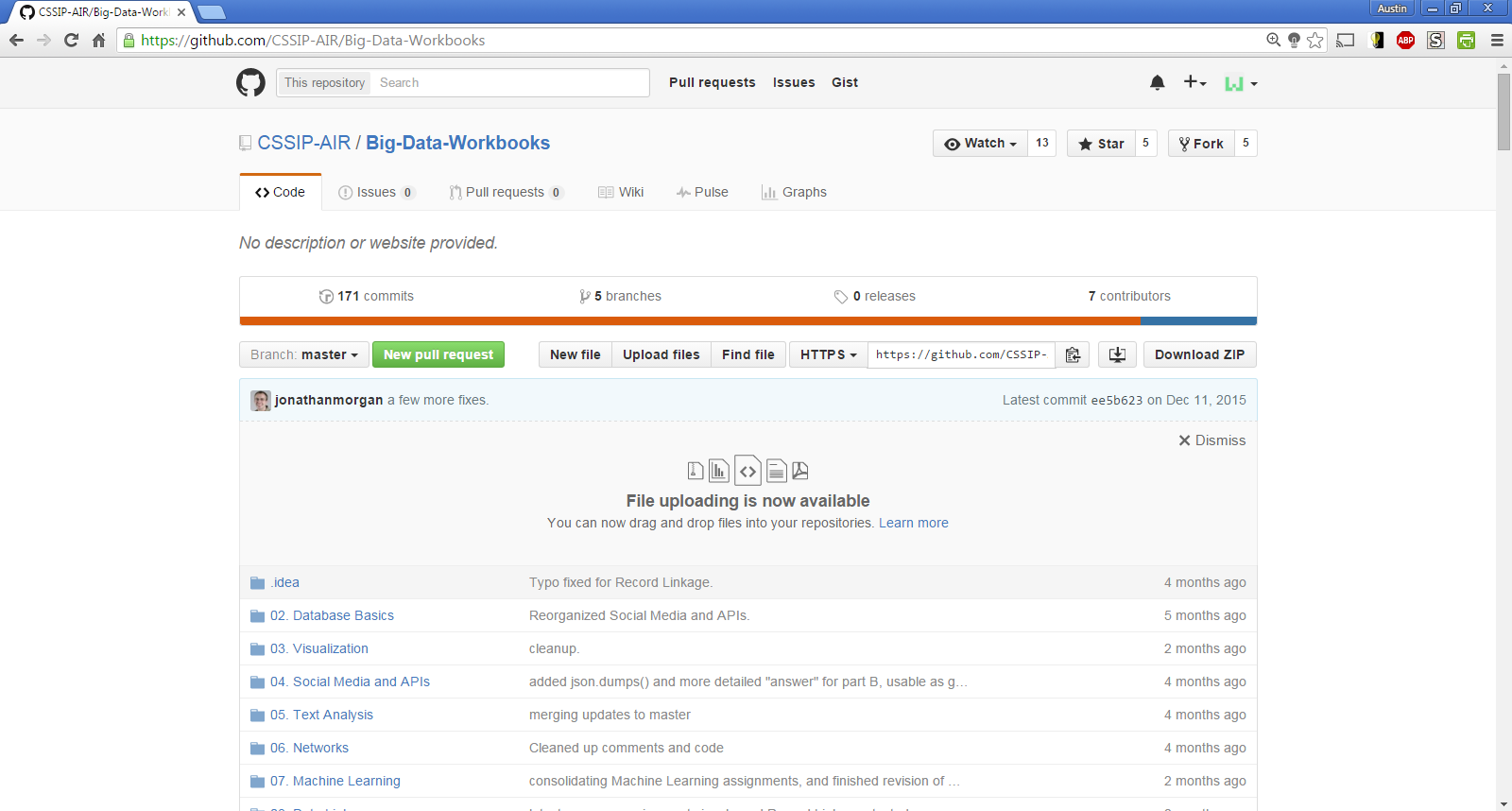


To create a new notebook, click the "New" tab, then click "Python 2" under "Notebooks"

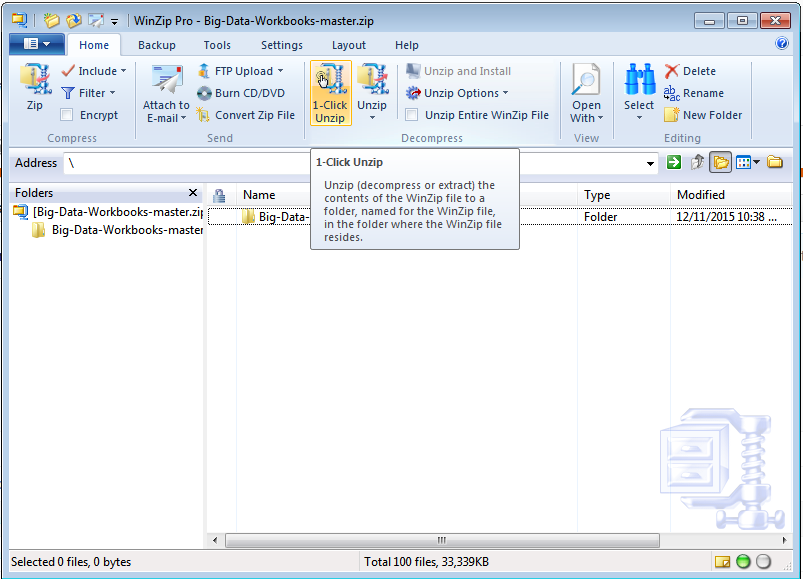


# **Loading Workbooks from GitHub to Jupyter Notebook**

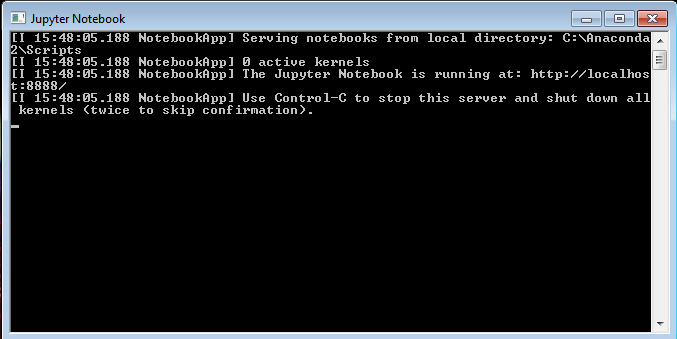
Go to Workbooks repository on GitHub and download the ZIP file



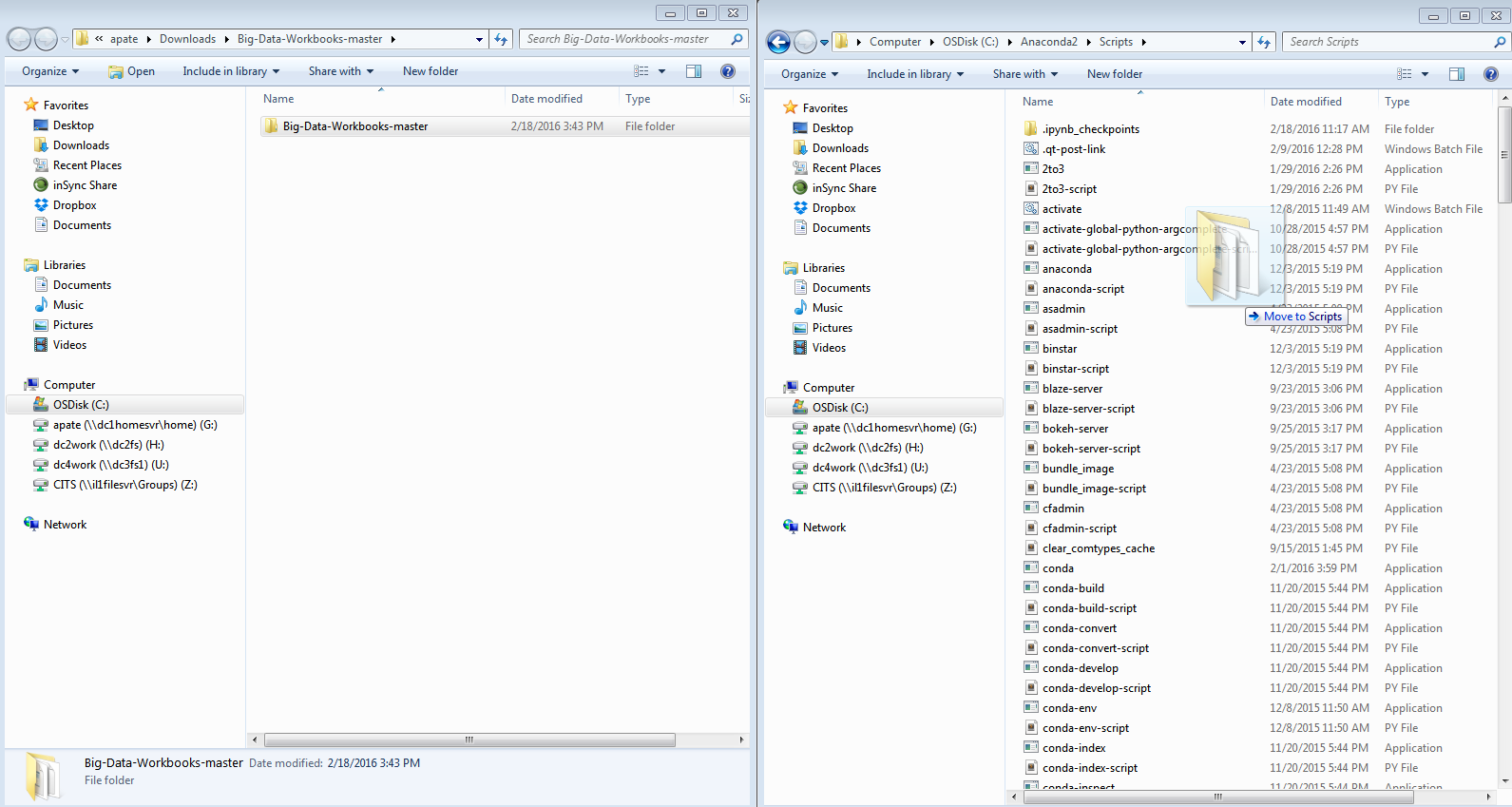
Open the ZIP file and unzip the folder. There should now be an unzipped workbook file in your downloads folder.



If you don’t have Notebooks open already, do so. Once you have them open, toggle back to the Jupyter Notebook prompt box. Use the first line of text to determine where your Notebook directory is located. In this example, the local directory is **C:\Anaconda2\Scripts**.



Open the file path of your Notebook directory, then drag-and-drop your unzipped workbooks folder into the directory.



If your transfer was successful, you should be able to see the workbooks folder in your Notebook

