Database Basics

Installation and setup guide

AUTHOR

Table of Contents

[**Introduction** 2](#_Toc443629899)

[**Setup Introduction** 2](#_Toc443629900)

[**Anaconda Installation** 3](#_Toc443629901)

[**IPython and PyMySQL Installation** 4](#_Toc443629902)

[**MySQL Installation** 5](#_Toc443629903)

[**Opening and Working in Jupyter Notebook** 12](#_Toc443629904)

[**Loading Workbooks from GitHub to Jupyter Notebook** 14](#_Toc443629905)

[**Downloading and Importing Data into MySQL Workbench** 17](#_Toc443629906)

# **Introduction**

TBA

# **Setup Introduction**

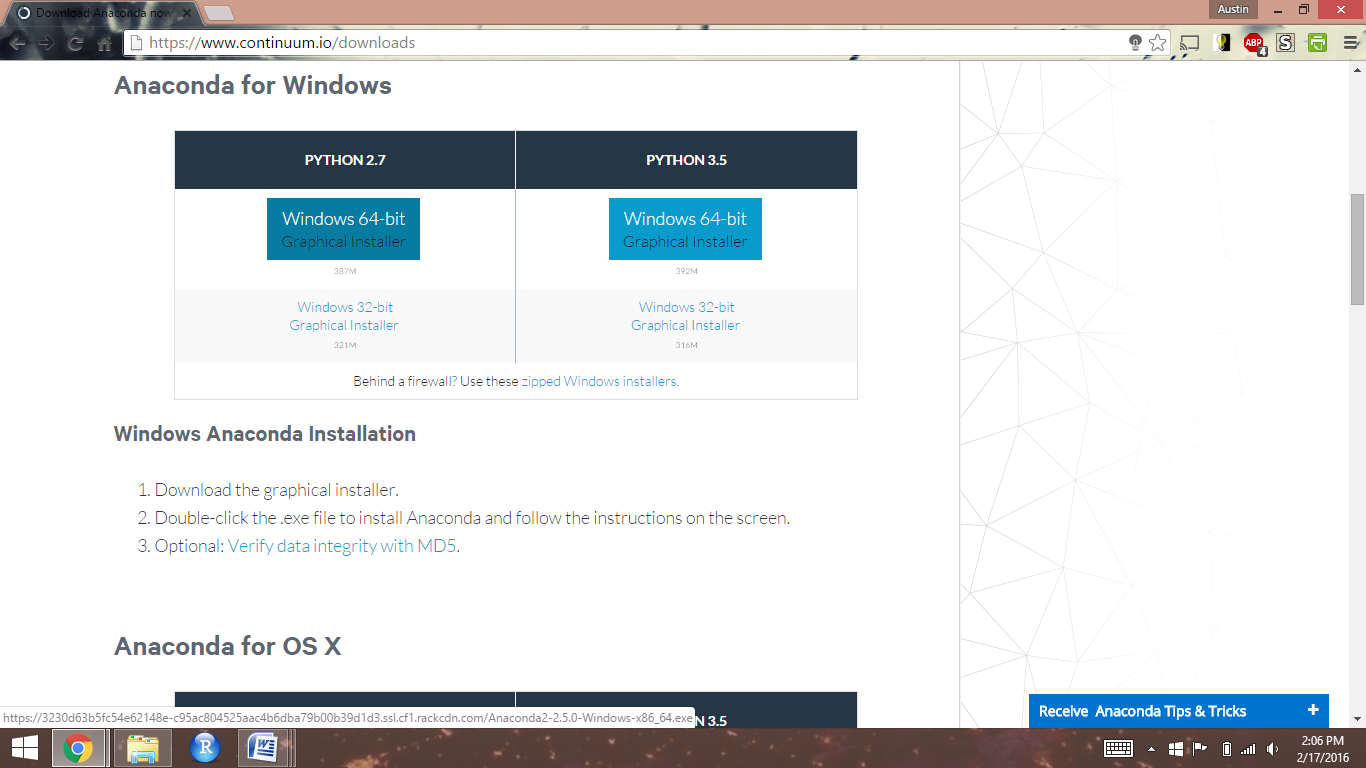
In the Database Basics exercises, as well as throughout this course, you will download and install several different software packages to interact with and analyze data. For the Database Basics exercises you will be interacting with a "homework" database that you will be downloading and installing in your local MySQL database. These exercises will help you to learn the SQL query language and practice writing your own SQL queries within a Python programming language environment.

When downloaded directly from the Python website, Python includes a command line Python interpreter that you can use to type in and run Python programs. This shell is useful, but limits your ability to interact with data – nested loops are difficult to write and it is can be difficult to re-use code from session to session. In order to avoid these constraints, you will be installing **Anaconda** and **IPython.** These software packages provide the basic Python programming language, a more full-featured and robust command line Python tool, as well as the easy-to-use **Jupyter Notebook** software environment in which you will be writing your code and viewing the output.

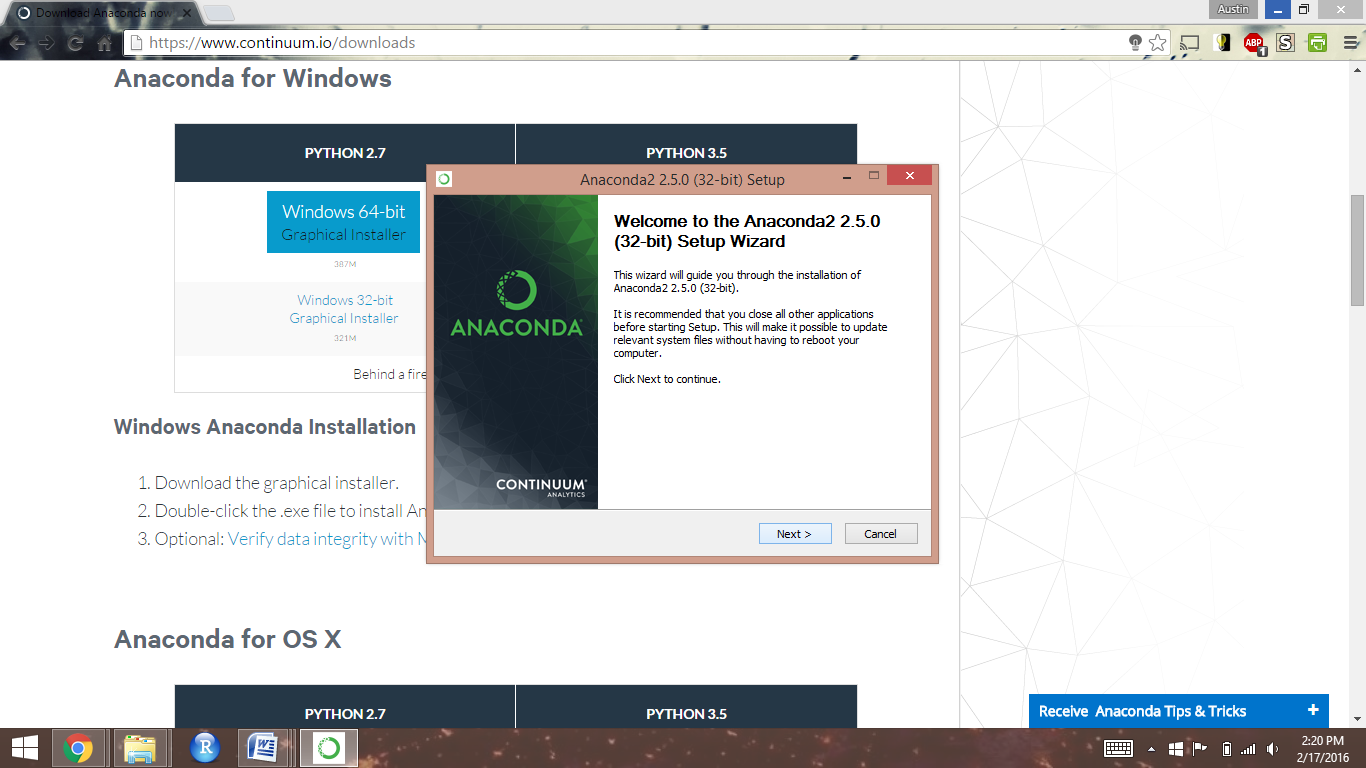
While Anaconda provides a comprehensive programming environment, Python alone cannot interpret queries written in SQL syntax. SQL must be run against a relational database, and you have to tell Python how to connect to and interact with your database of choice. For this class, you will install a local copy of the **MySQL** suite of software that will serve as your database, then the **PyMySQL** package that tells Python how to run SQL against a MySQL database.

Below are a list of steps to install all of the necessary components to get you up and running with Anaconda and MySQL. Once you have successfully completed these steps, you will be ready to begin working through the Database Basics Workbook!

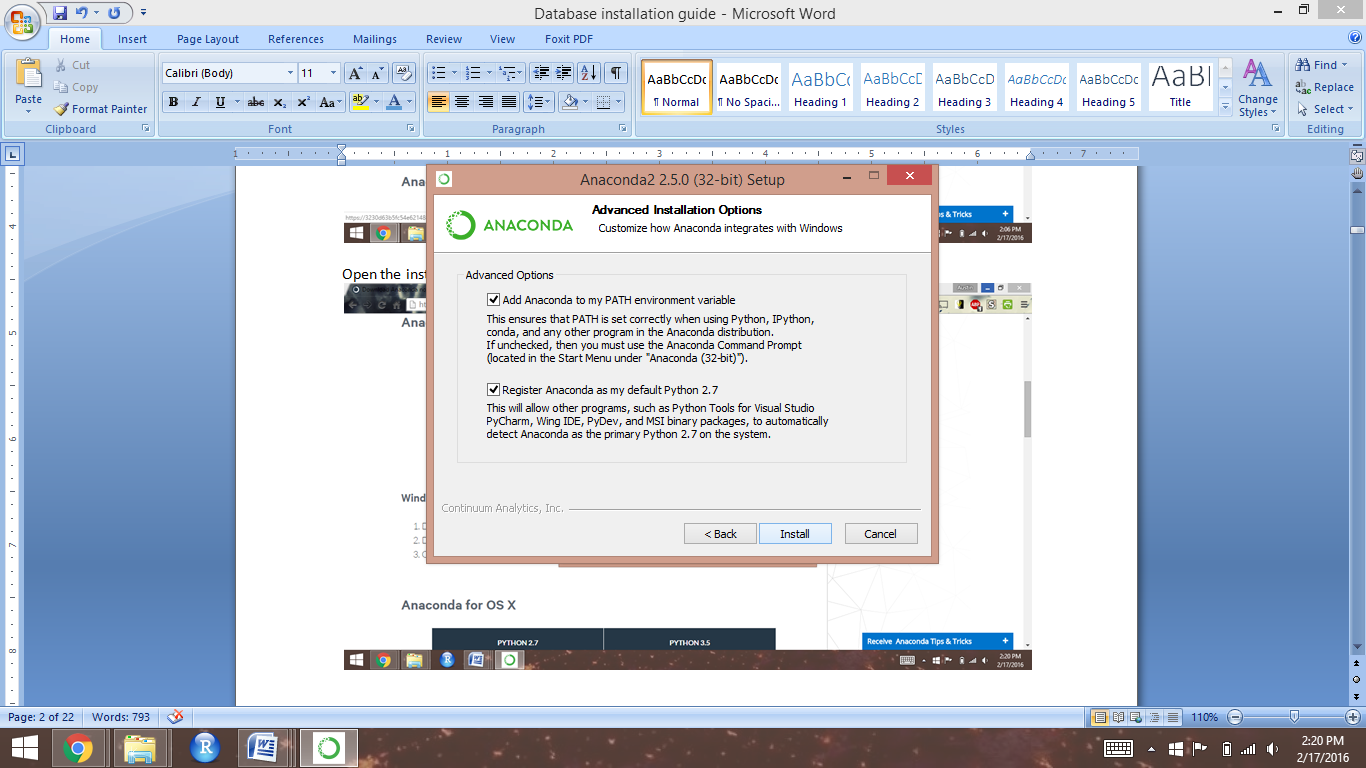
# **Anaconda Installation**

Go to [Continuum Analytics](https://www.continuum.io/downloads) and select Anaconda Python 2.7 for the appropriate OS

Open the downloaded executable file and move forward with installation

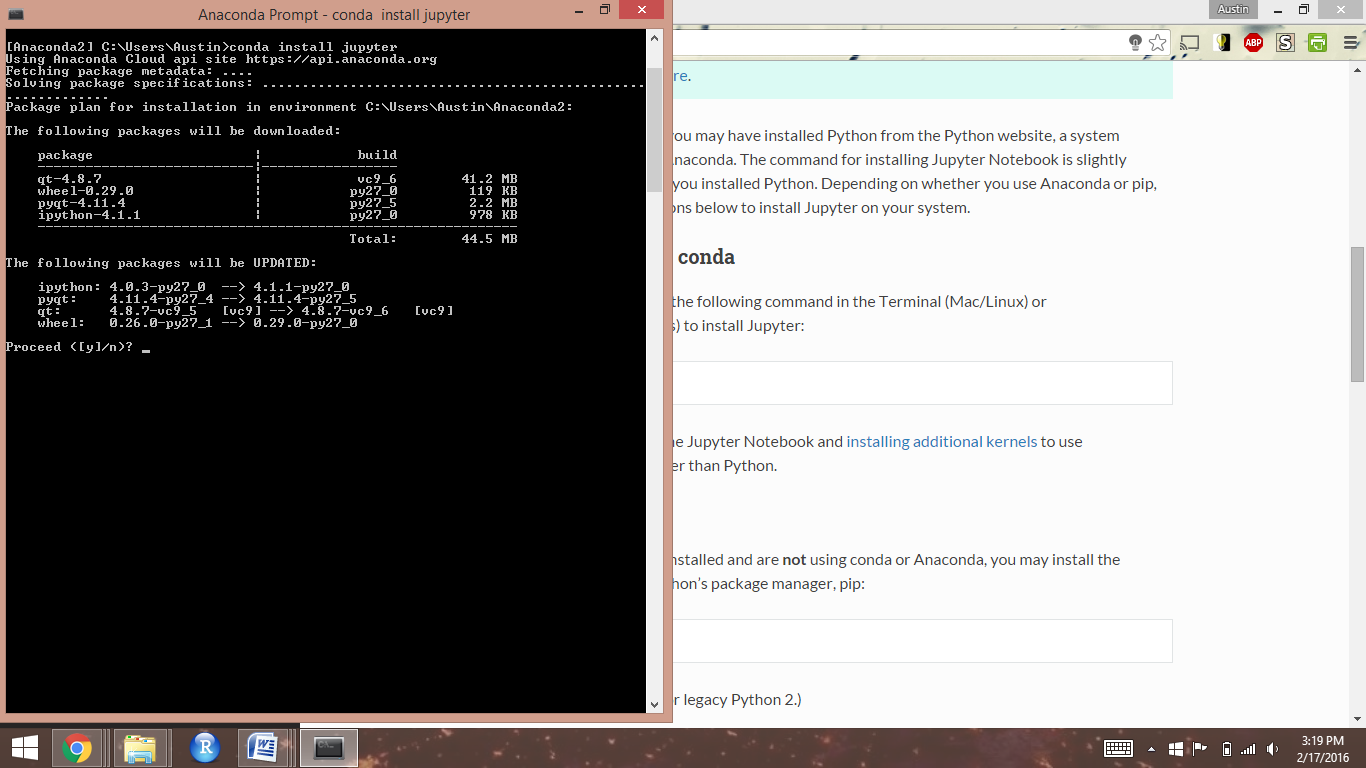


Keep the default advanced options checked, and click next until installation has finished

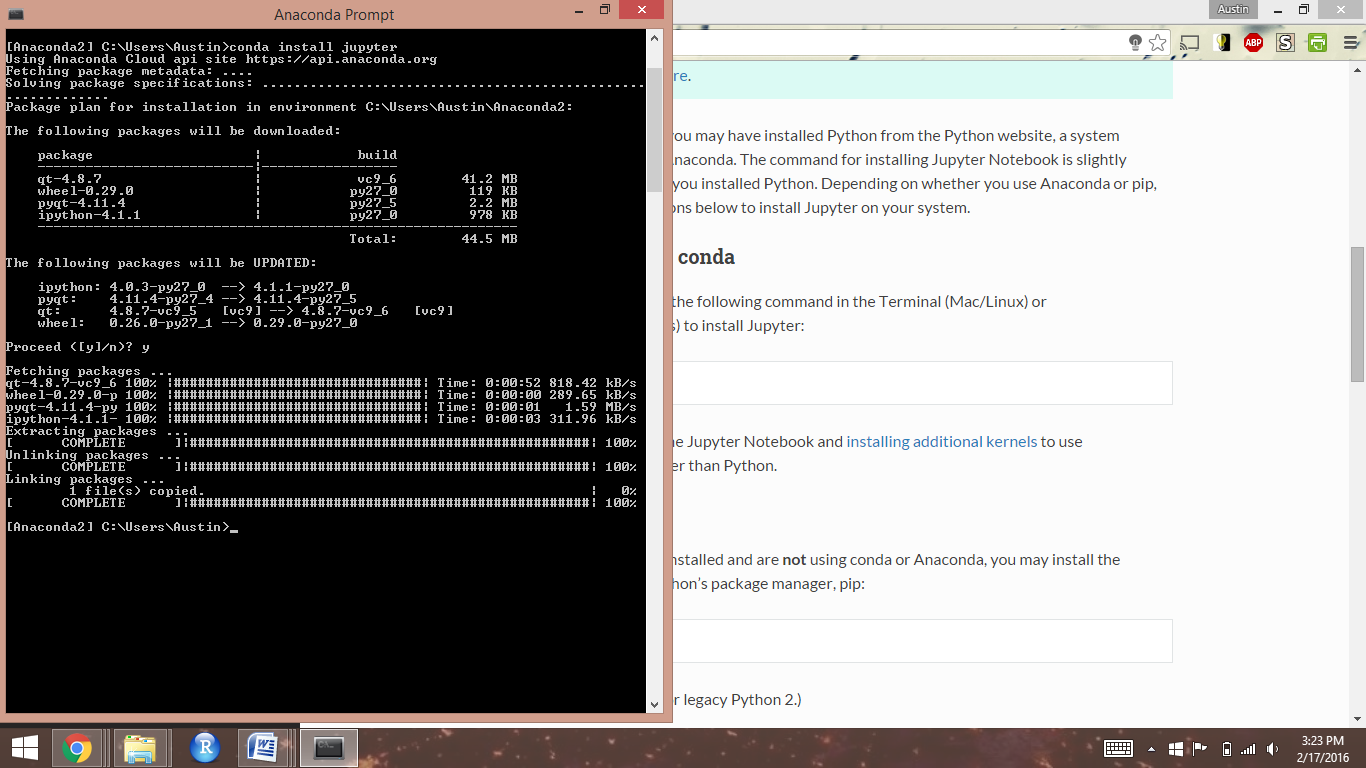


# **IPython and PyMySQL Installation**

Open Anaconda Prompt, 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.



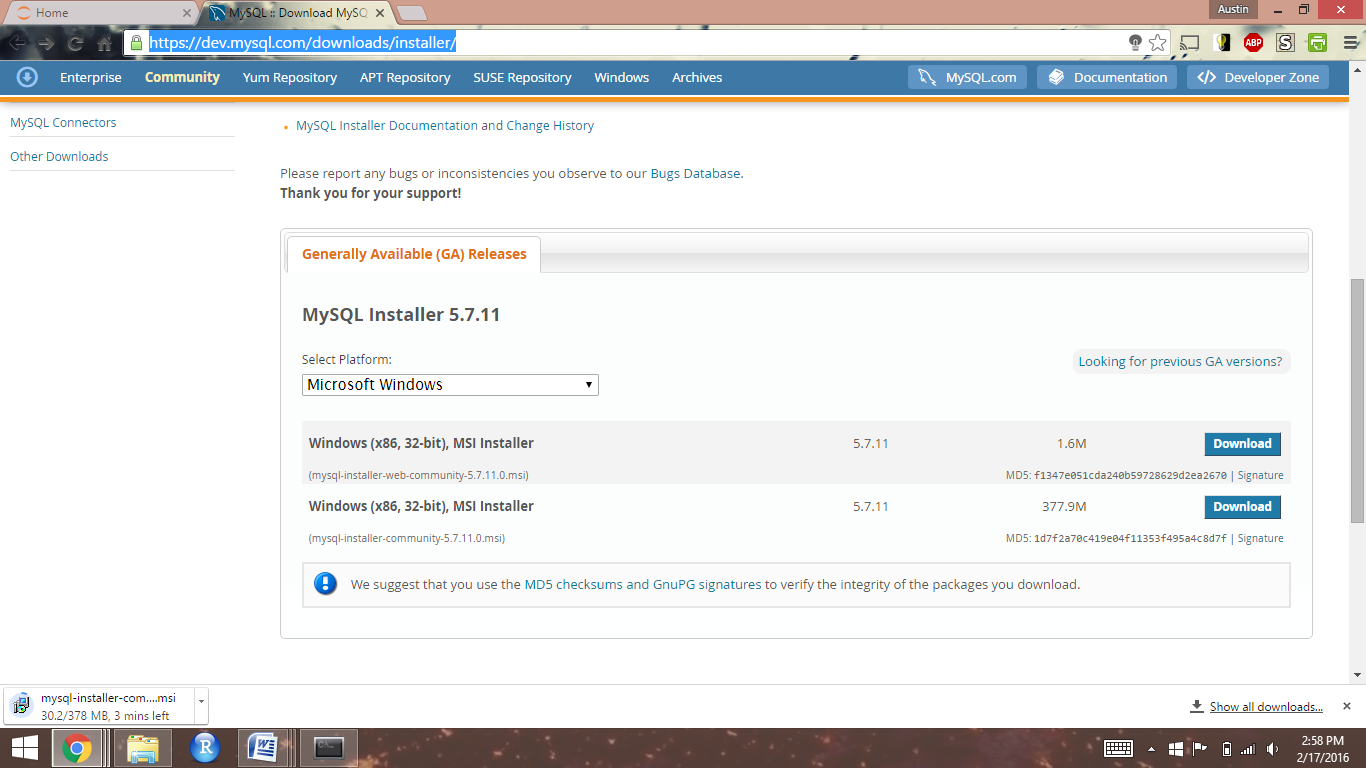
Once the prompt returns to your original directory the installation is complete:



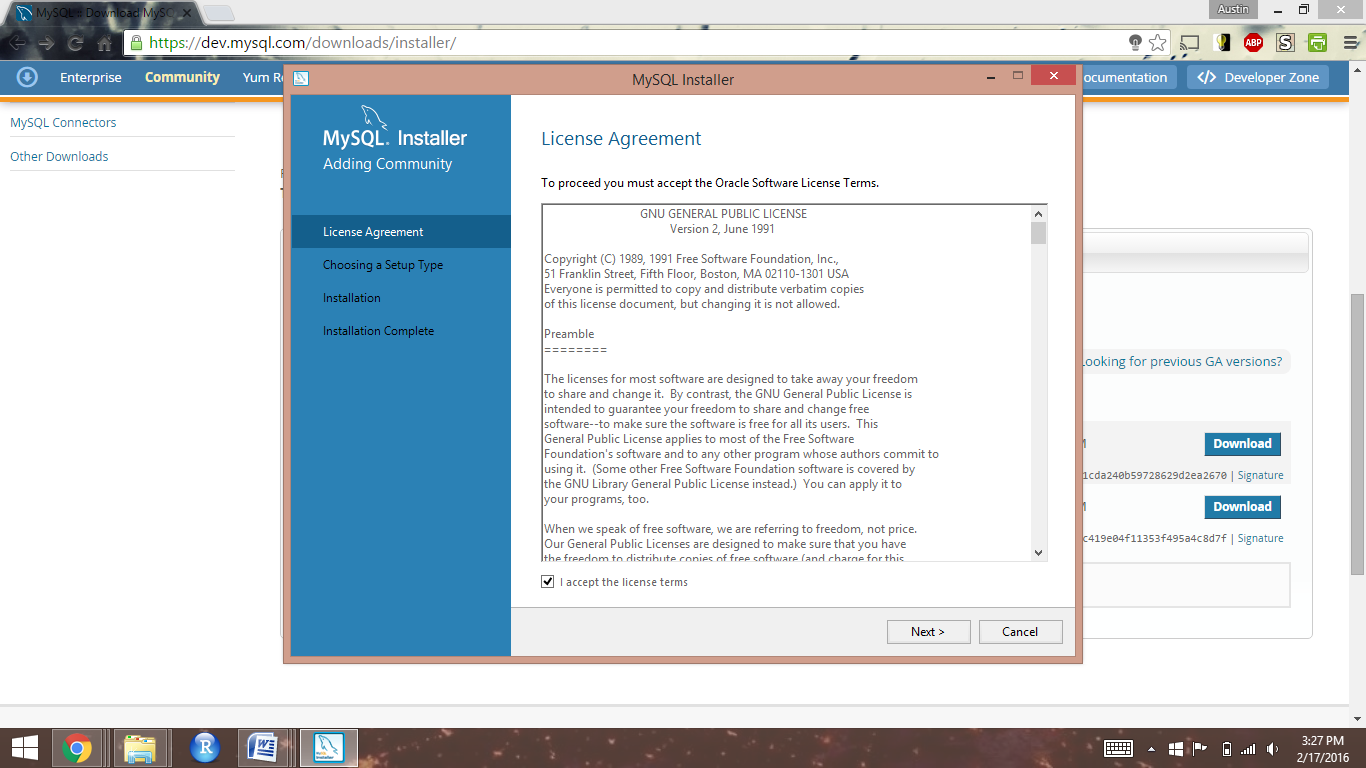
Do the same for PyMySQL: type “conda install pymysql”, then type “Y” to continue. Once the prompt returns to your original directory this time, you can close the Anaconda Prompt box.

# **MySQL Installation**

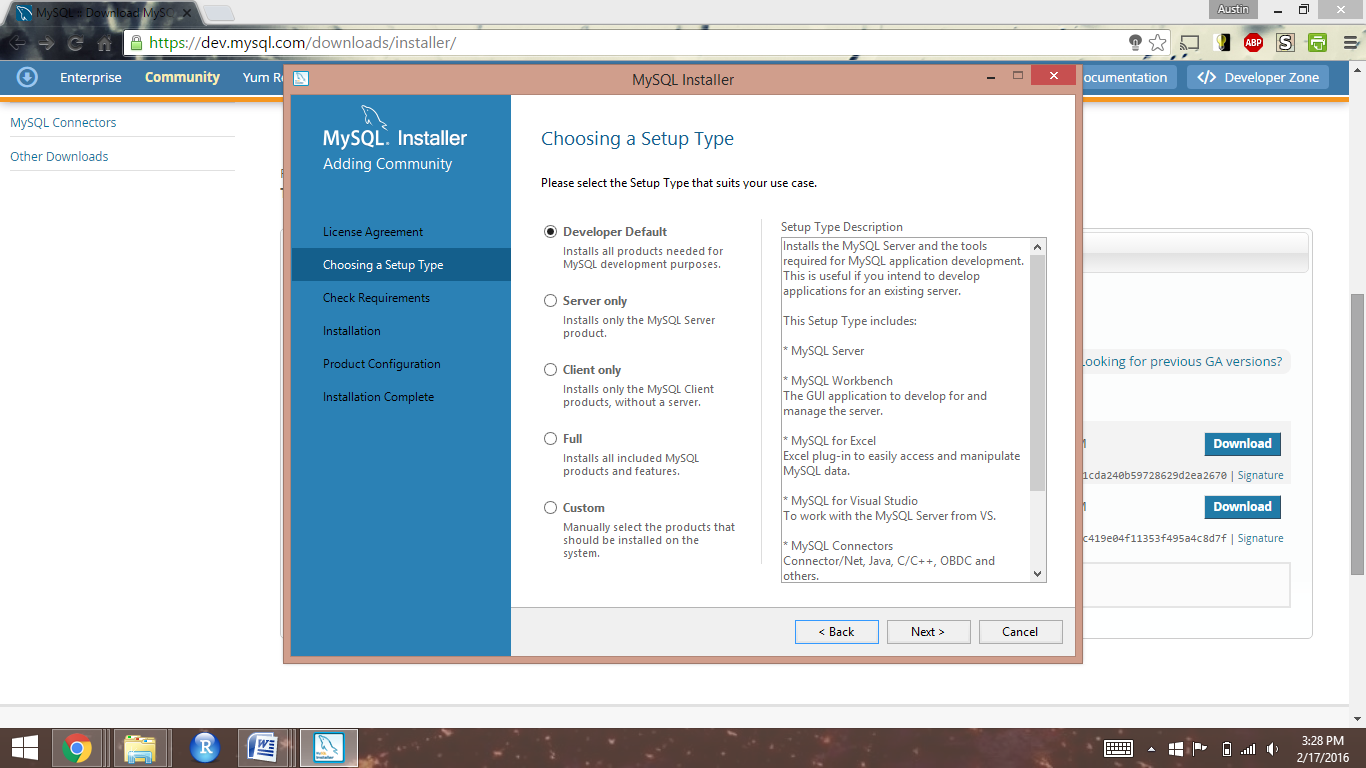
Go to the [Oracle website](https://dev.mysql.com/downloads/installer/) and download the MySQL Installer package (if you don’t want to create an Oracle ID, just click the “No thanks, just start my download” link at the bottom of the page).



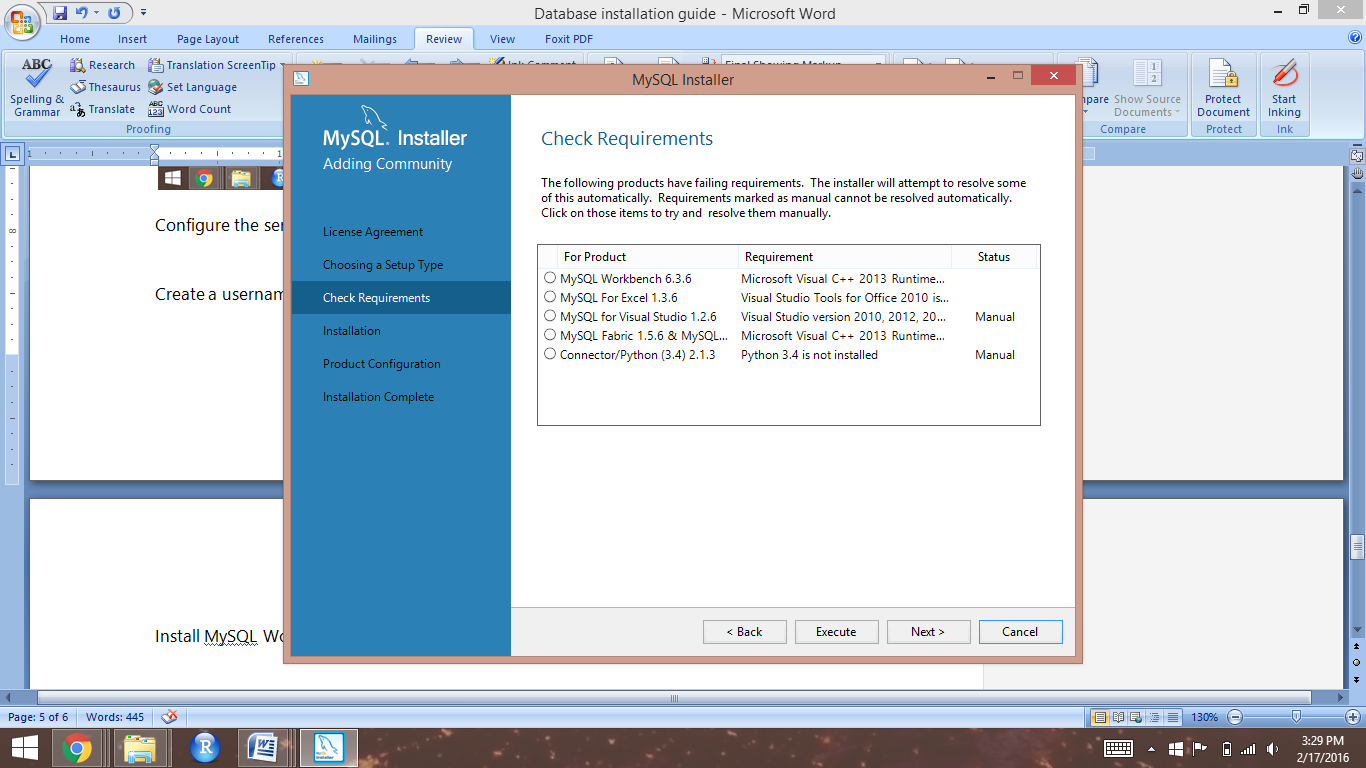
Open the downloaded executable file and move forward with installation



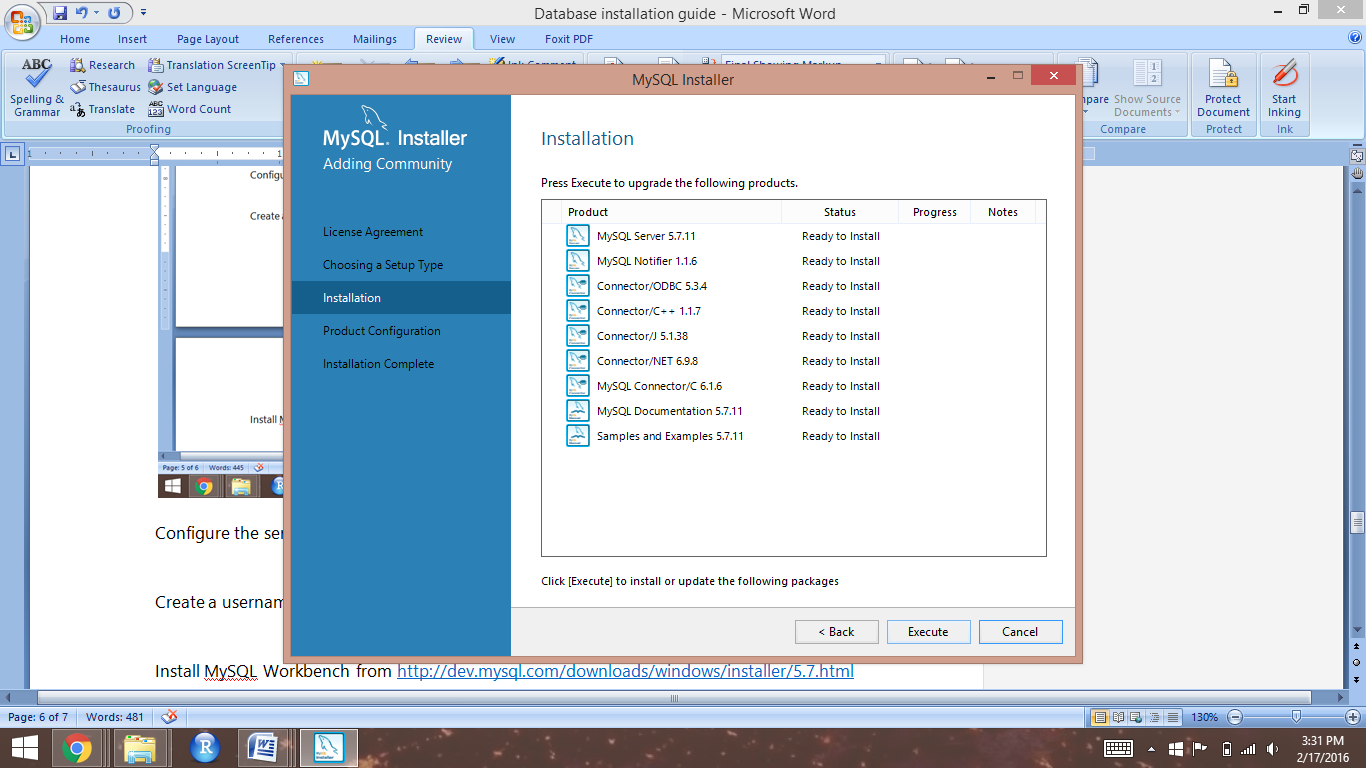
Select "Developer Default" and continue



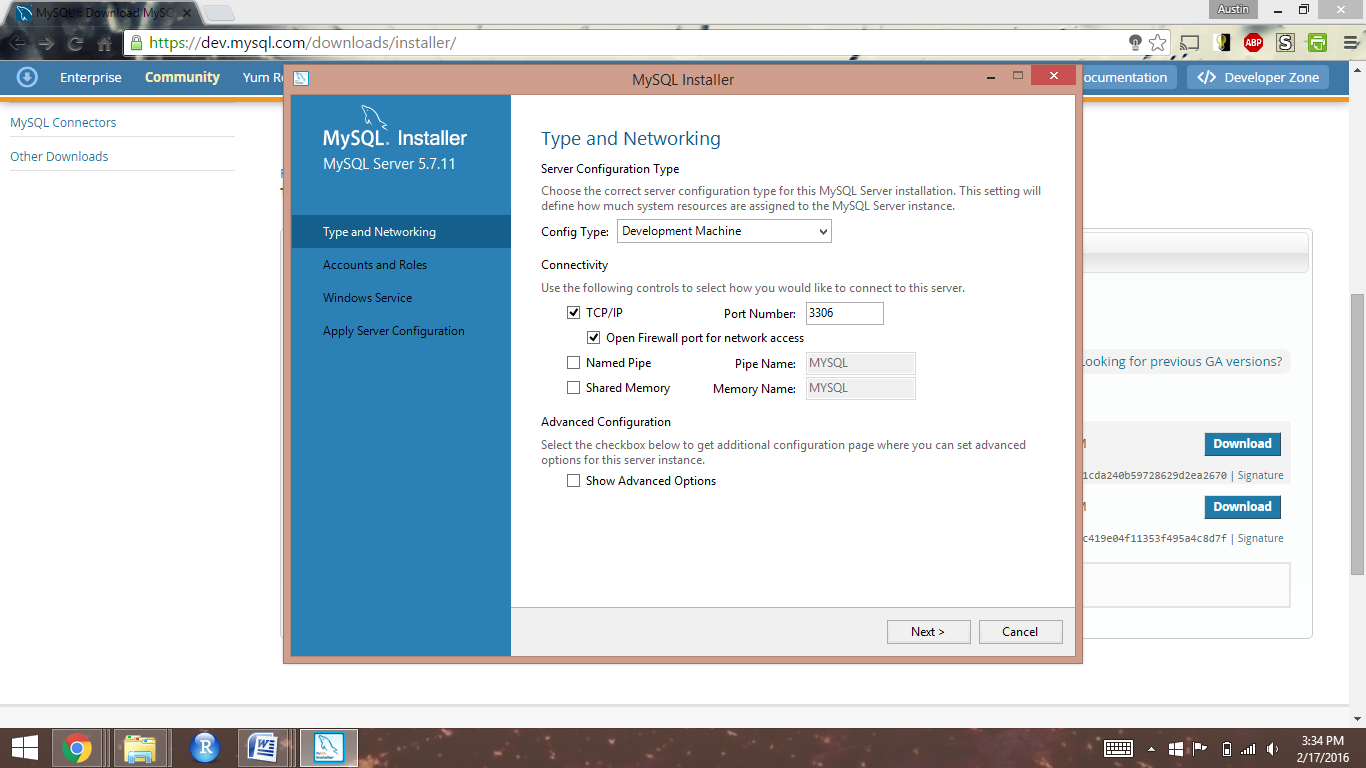
When performing the requirement check, you may encounter that your system lacks requirements for the most up-to-date versions of the MySQL software. This should not be a problem, and you should move forward with the installation.



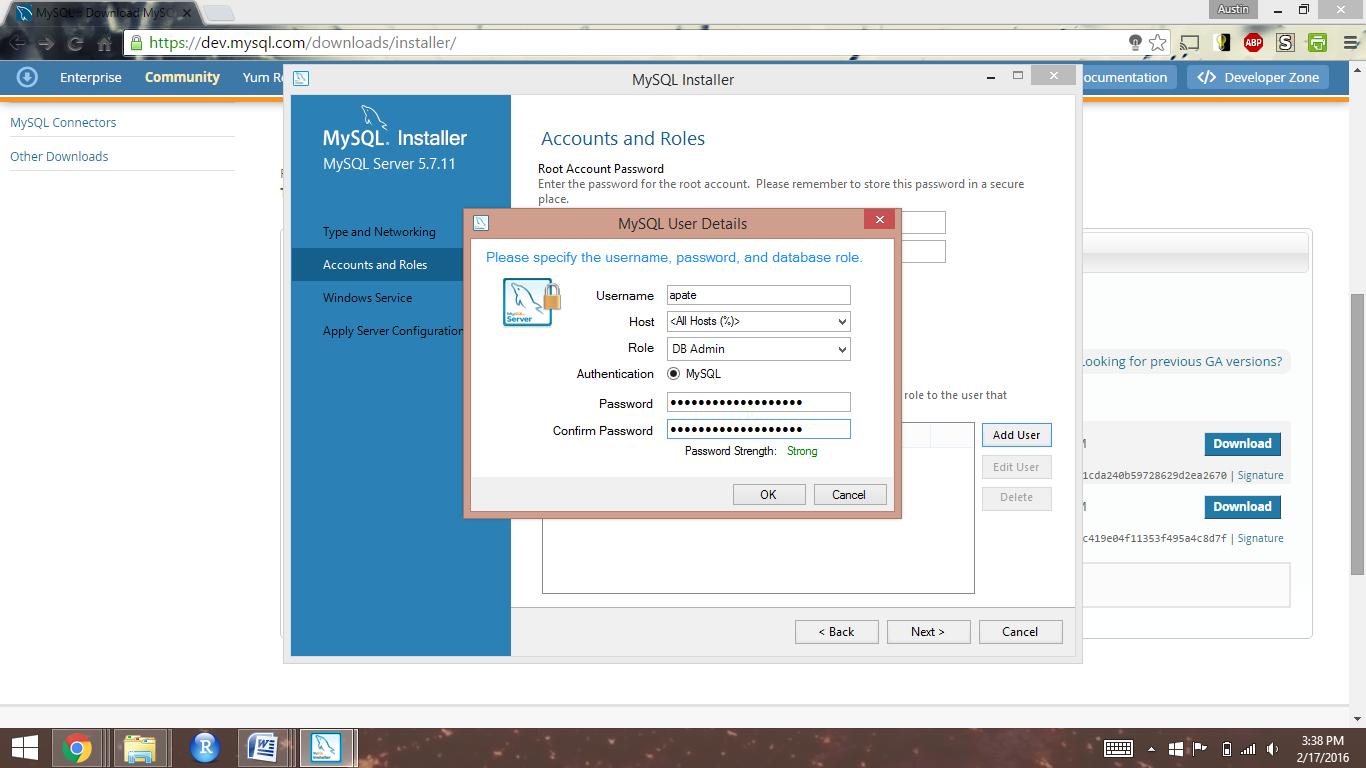
Select "Execute" to install the components of the MySQL developer package



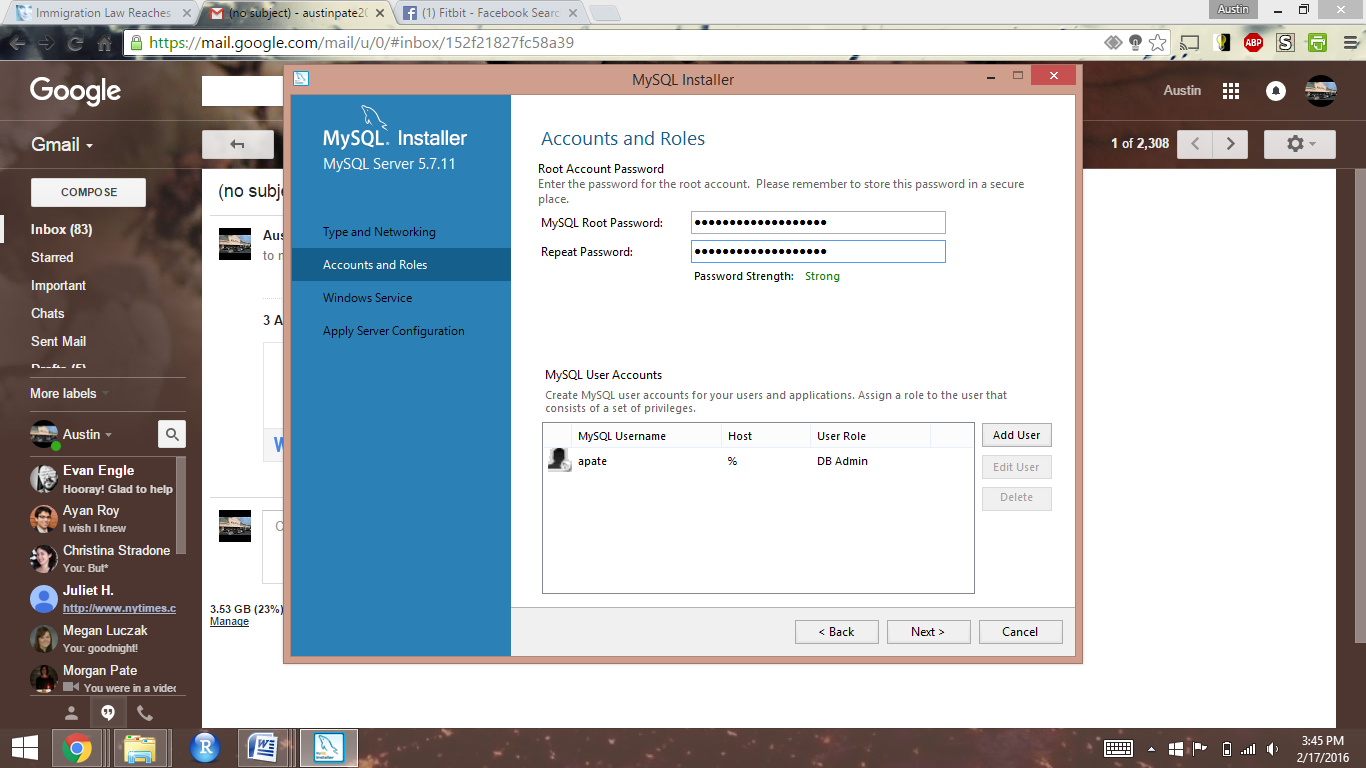
Keep the server configuration default settings and click "Next"



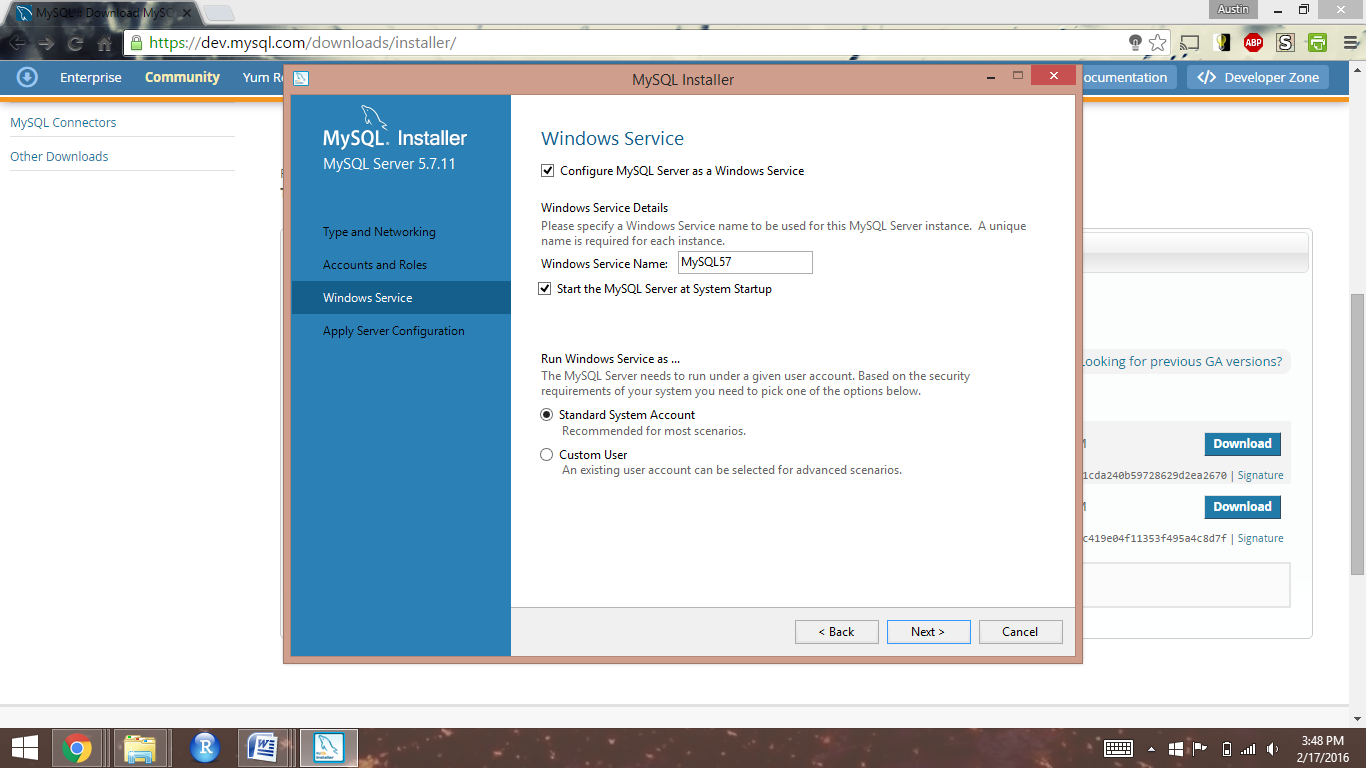
Click "Add User" then create a username and password for your server admin account. Click "OK"



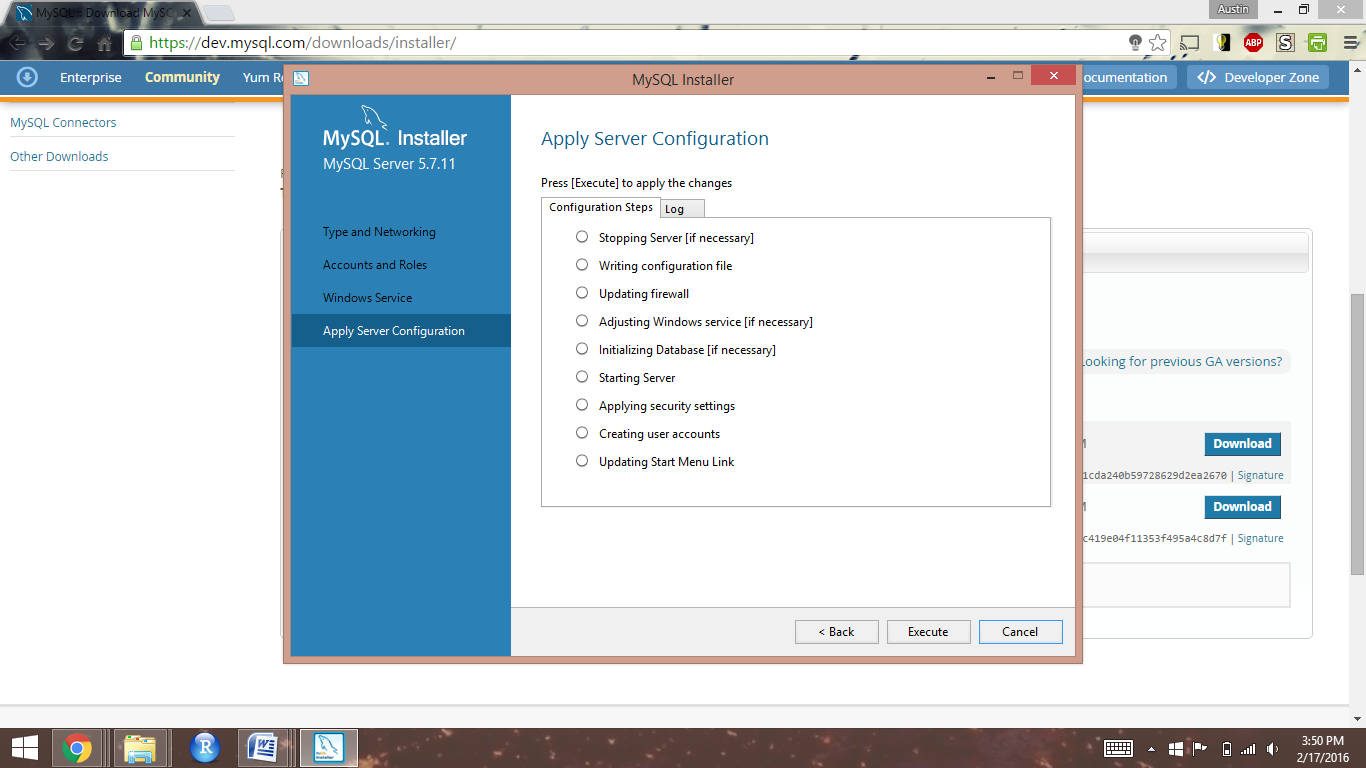
Create a password for your MySQL root account, then click "Next"



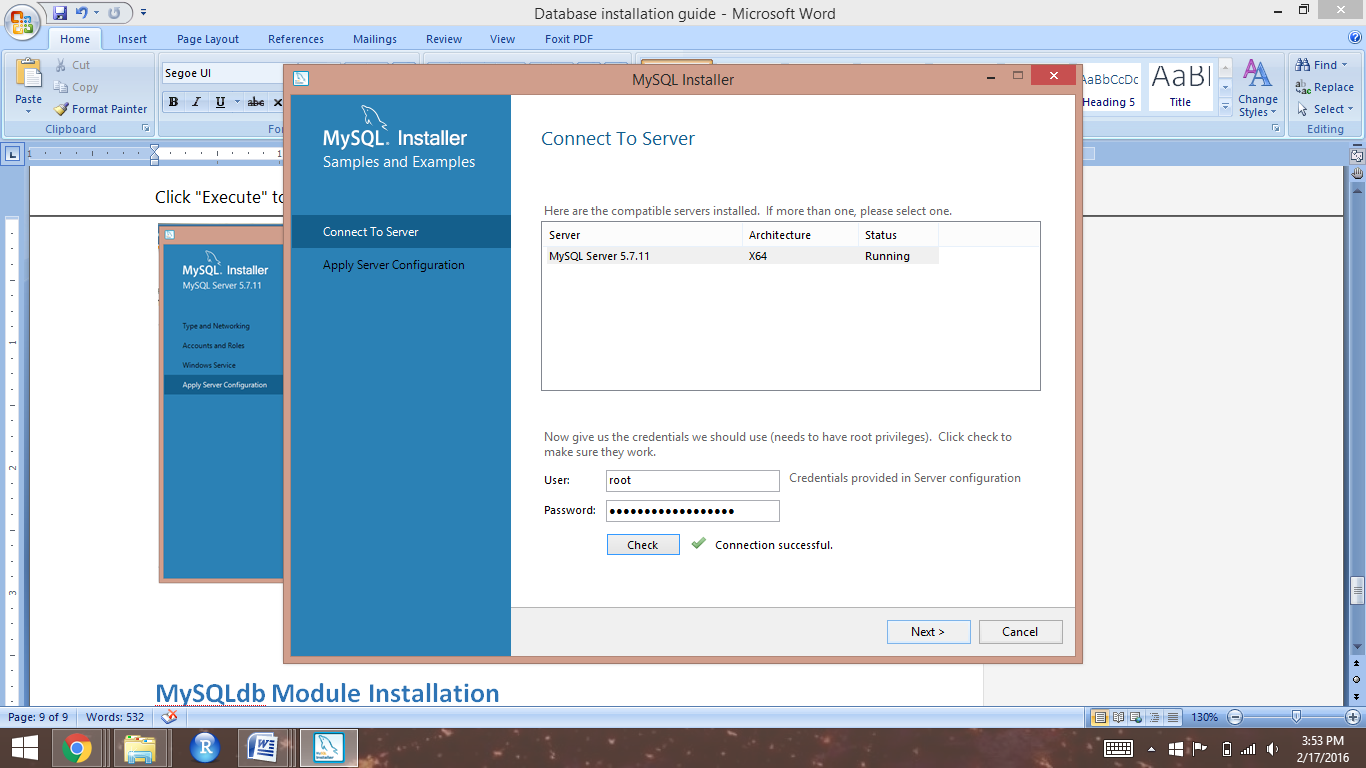
Keep the default settings for Windows Service and the "Standard System Account", then click "Next"



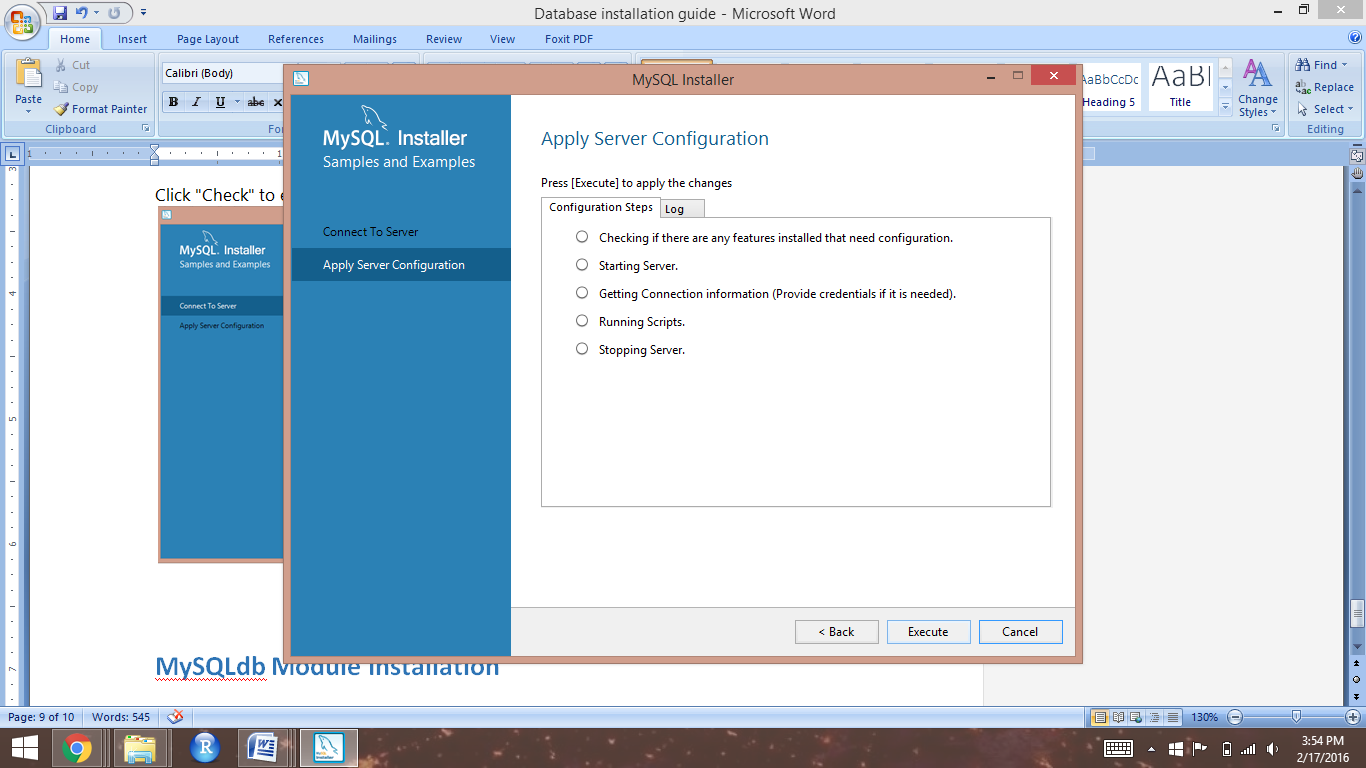
Click "Execute" to apply your server configuration, the "Finish"



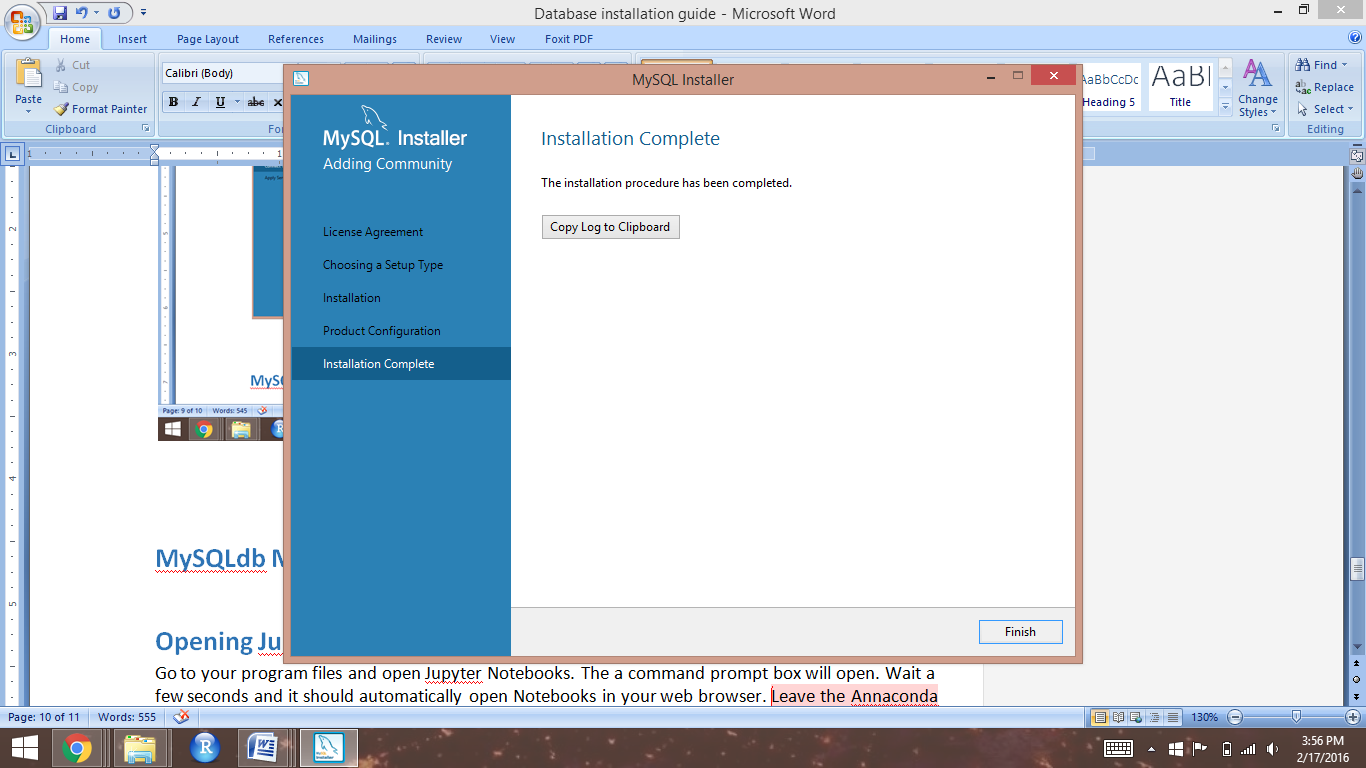
Click "Check" to ensure you can connect to your server, then click "Next"



Click "Execute" to finalize your server configuration, then click "Finish"

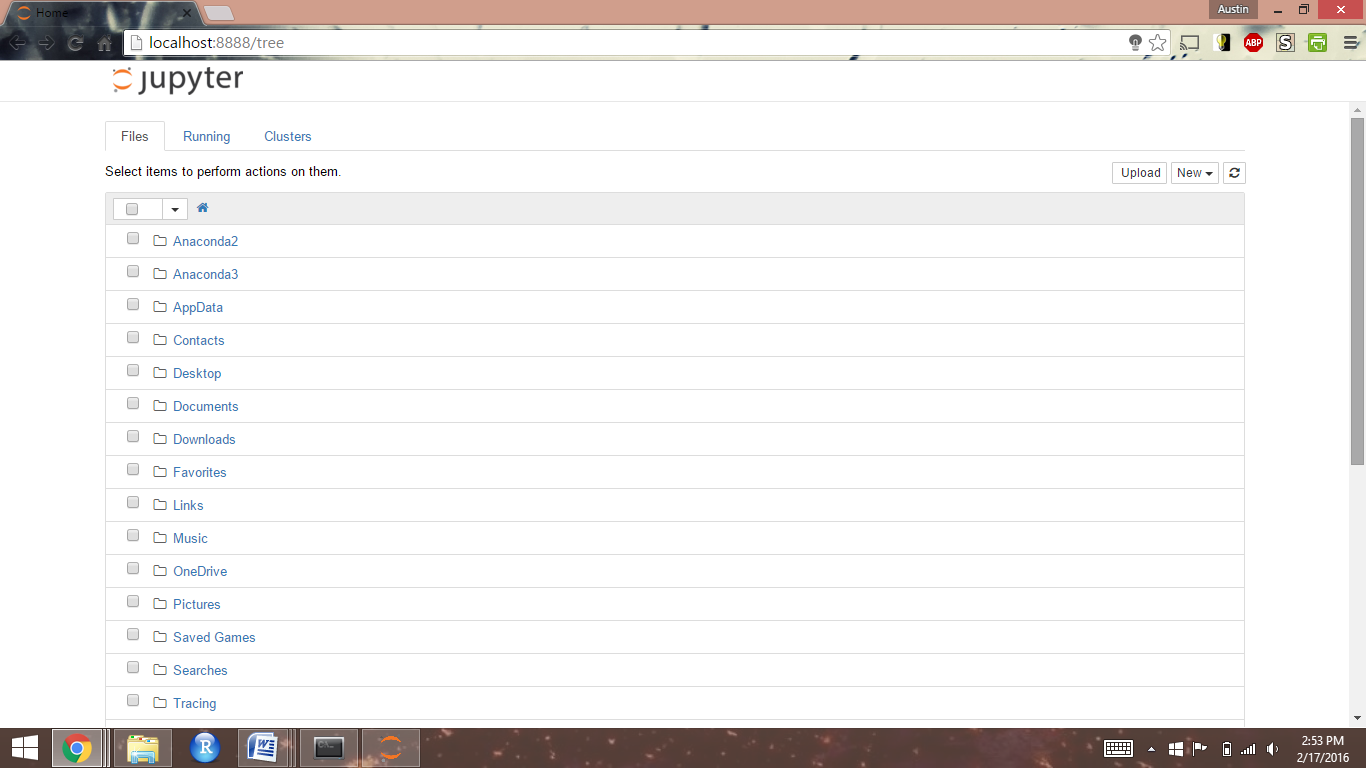


Click "Finish" to complete the MySQL installation

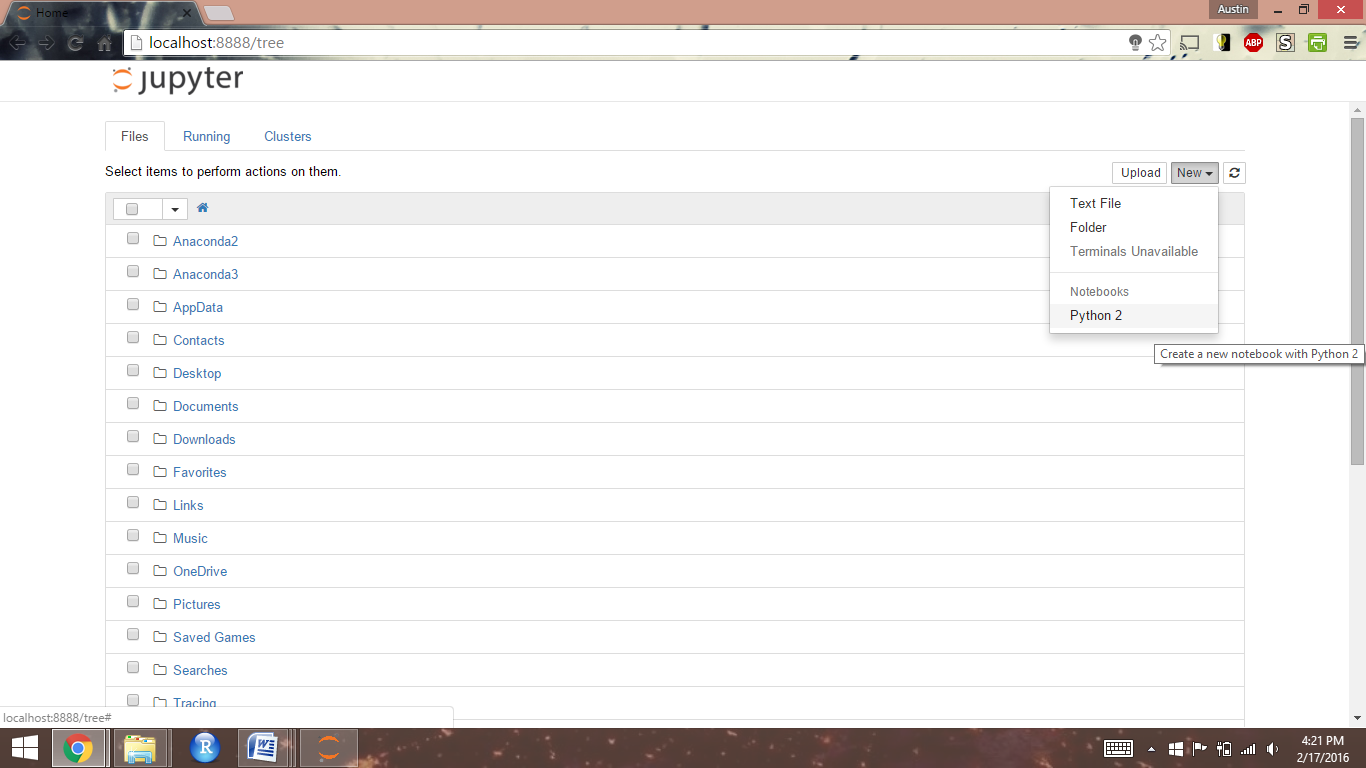


# **Opening and Working in Jupyter Notebook**

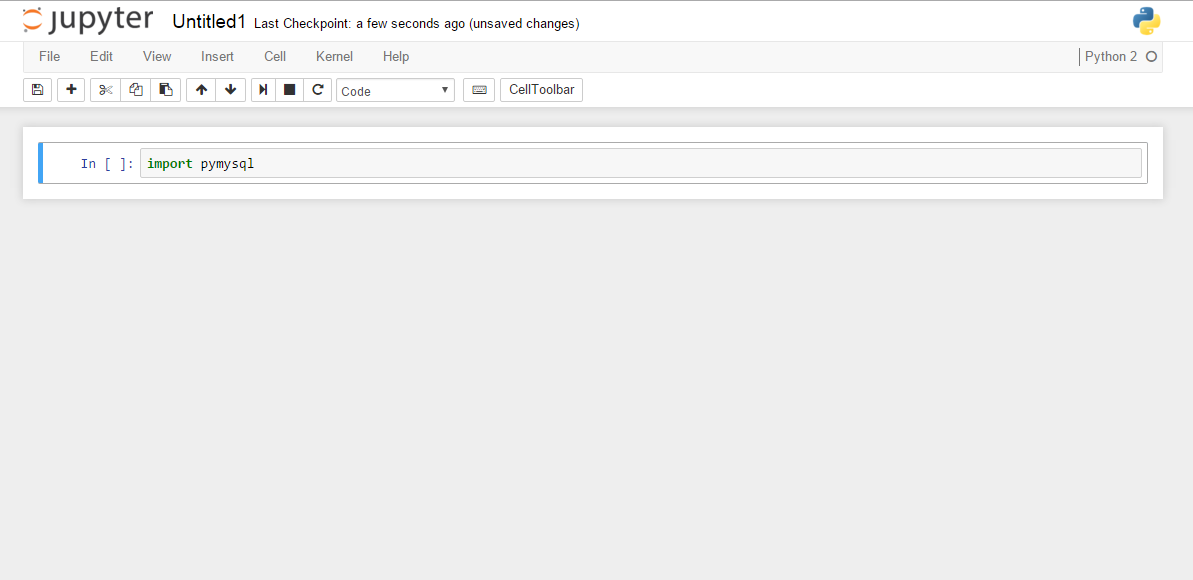
Go to your program files and open Jupyter Notebook. A command prompt box will open. Wait a few seconds and it should automatically open the Notebook in your web browser. Leave the Annaconda Prompt box open while you use your Notebooks.



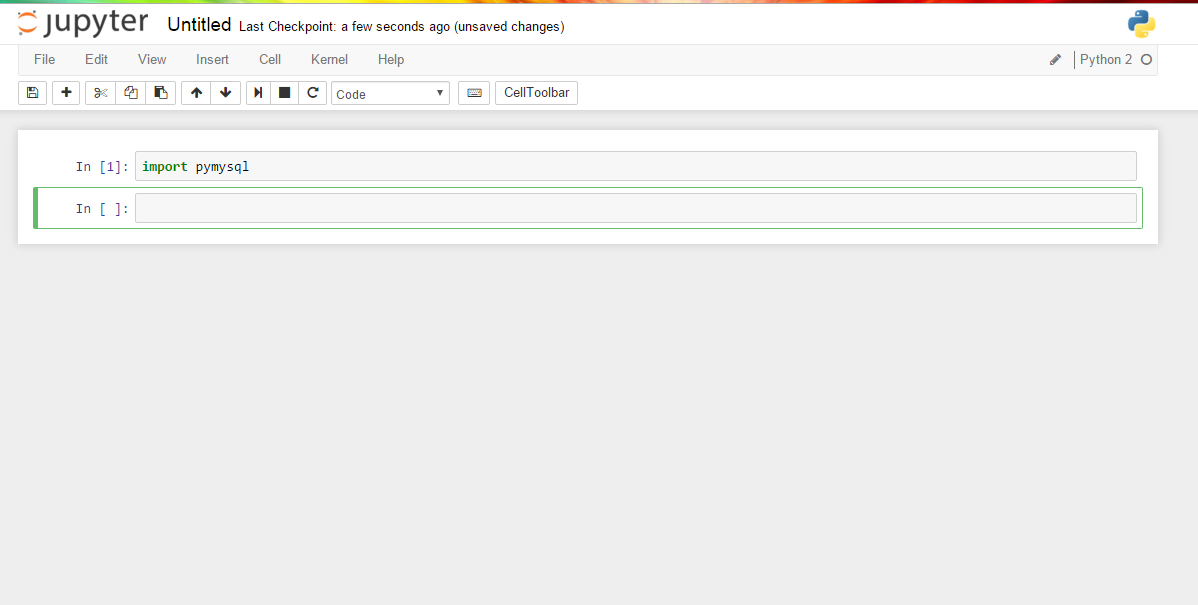
To create a new notebook, click the "New" tab, then click "Python 2" under "Notebooks"



To confirm the successful installation of the PyMySQL module, type "import pymysql" then click the Run icon



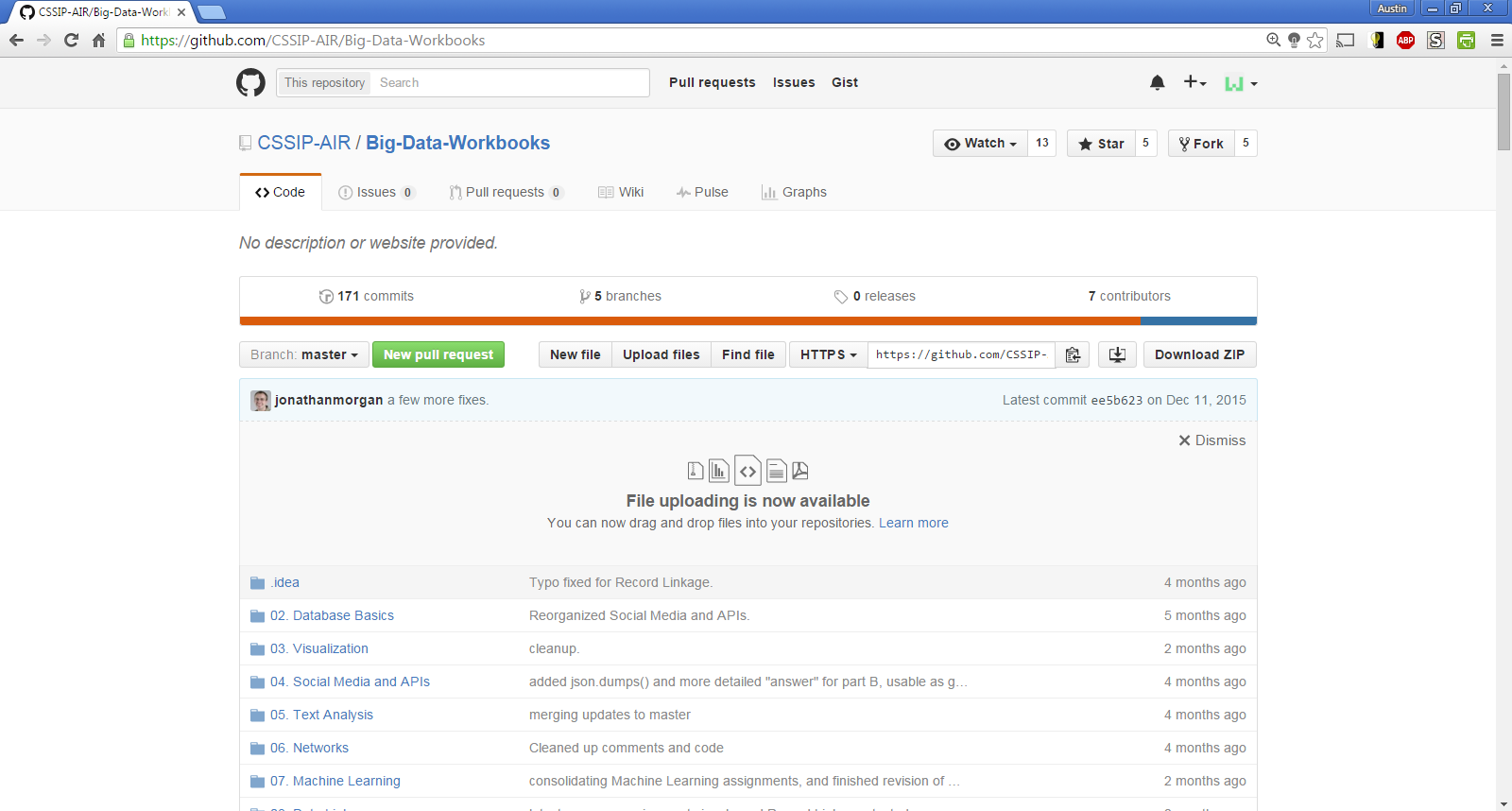
If the module was successfully installed, a new empty cell will appear below



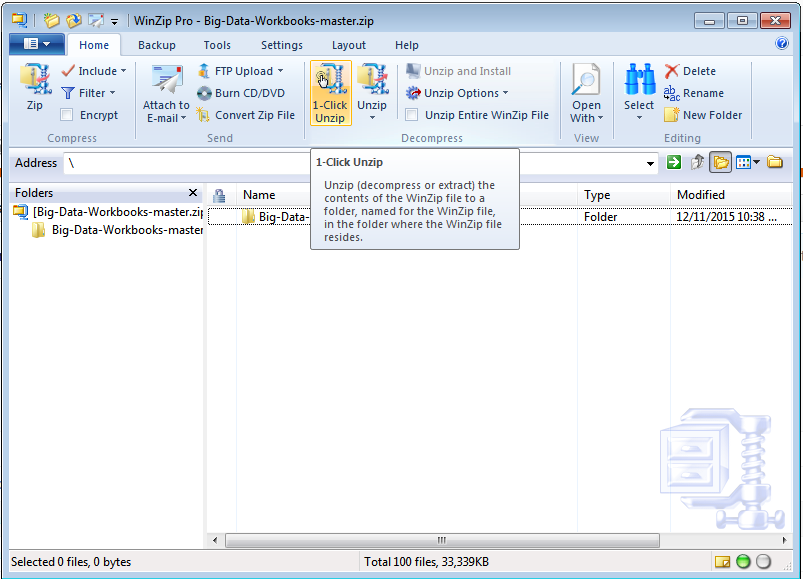
# 

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

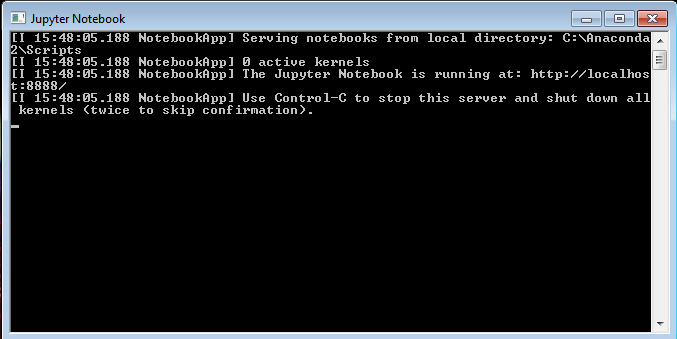
Go to Workbooks repository on GitHub and download the ZIP file



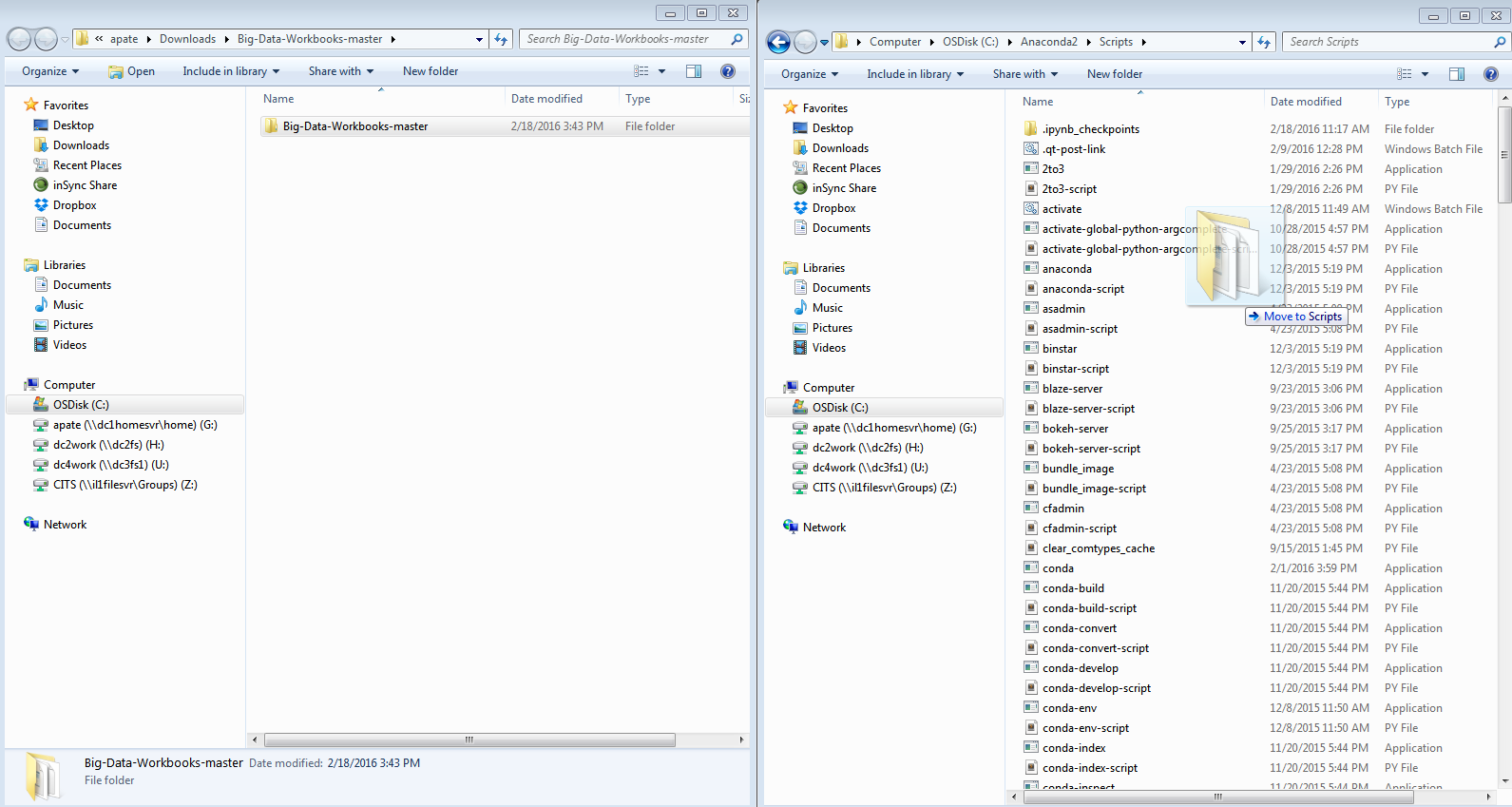
Open the ZIP file and unzip the folder. There should now be an unzipped workbook file in your downloads folder.



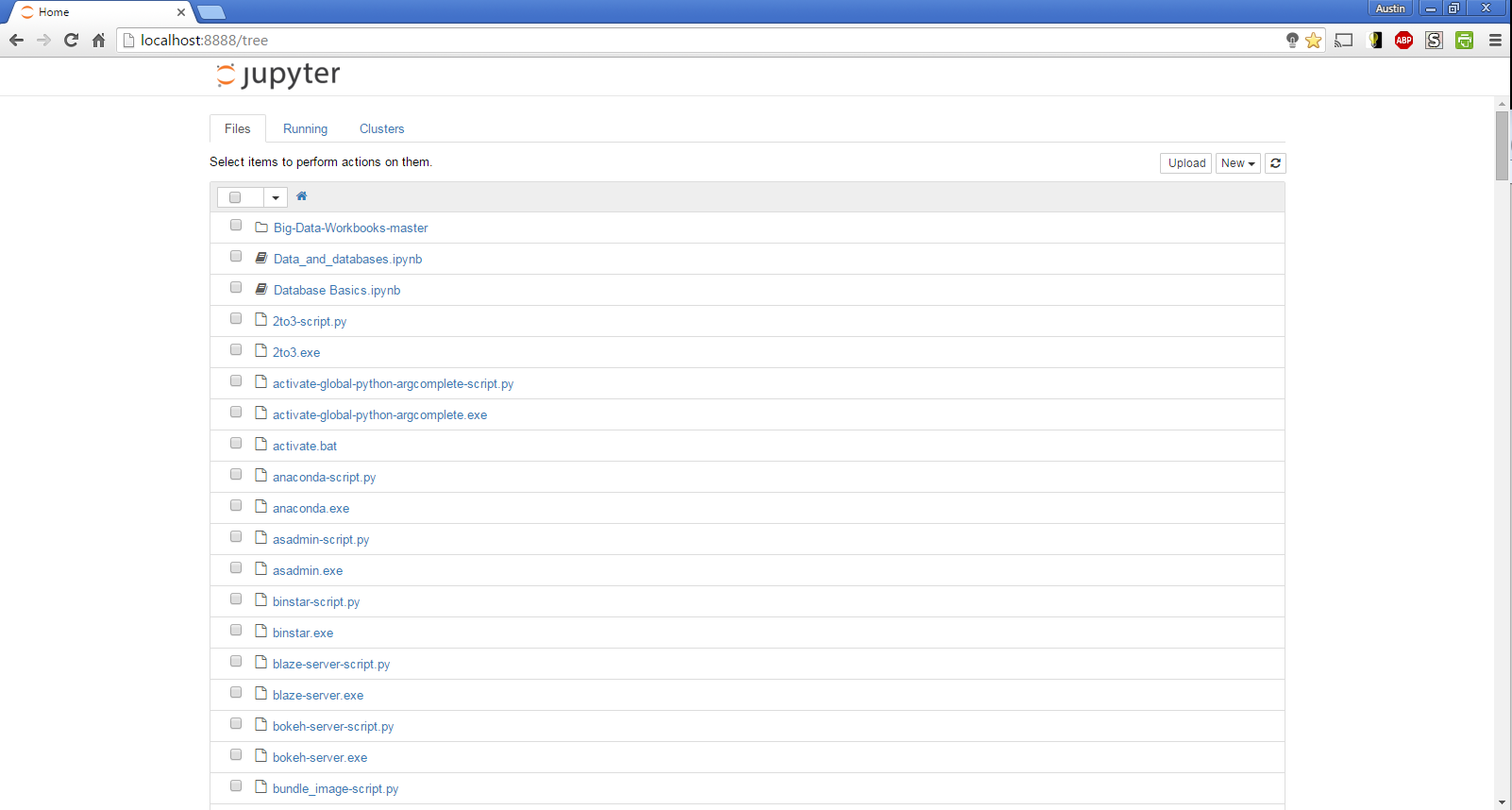
If you don’t have Notebooks open already, do so. Once you have them open, toggle back to the Jupyter Notebook prompt box. Use the first line of text to determine where your Notebook directory is located. In this example, the local directory is **C:\Anaconda2\Scripts**.



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



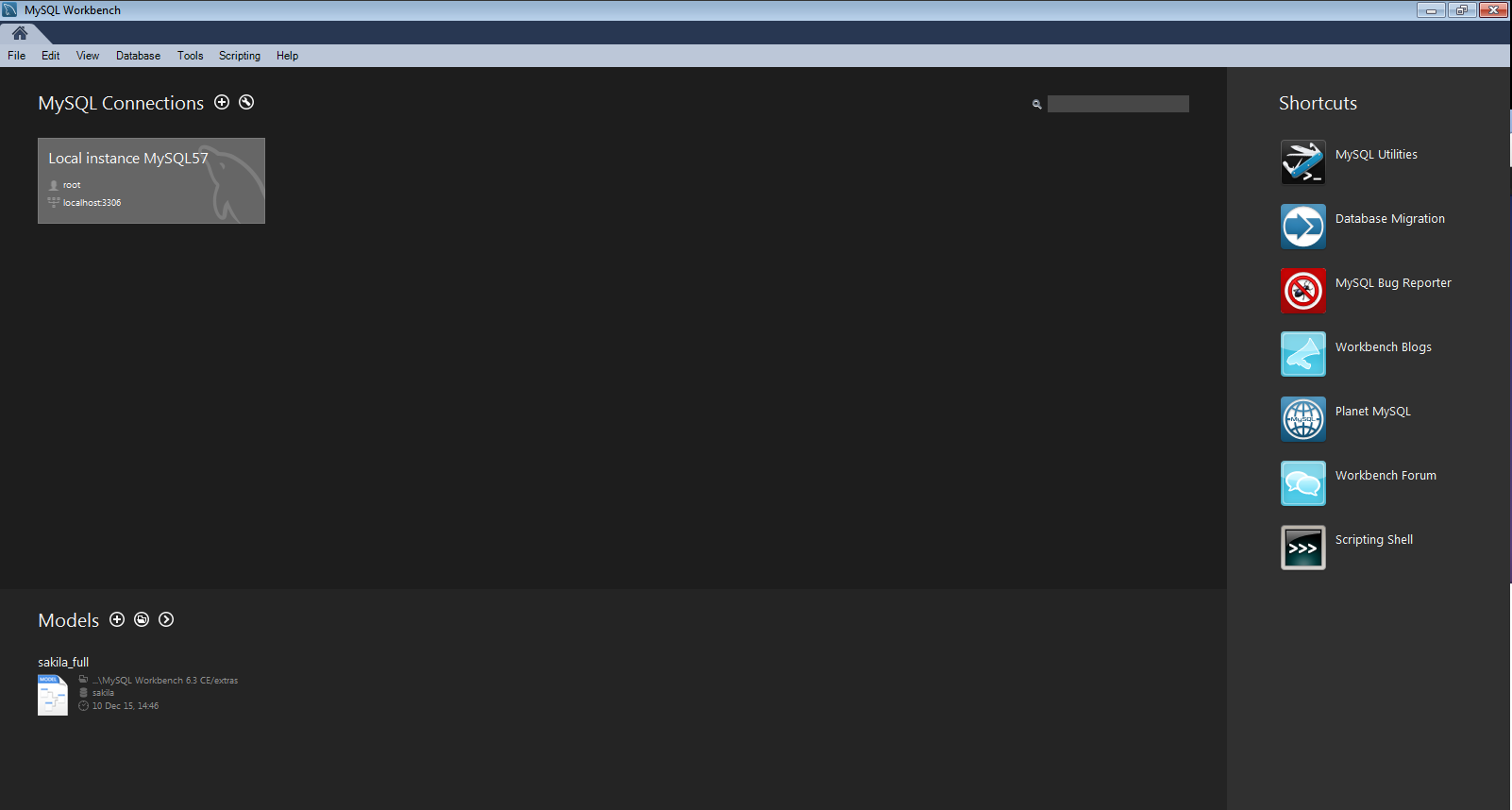
If your transfer was successful, you should be able to see the workbooks folder in your Notebook



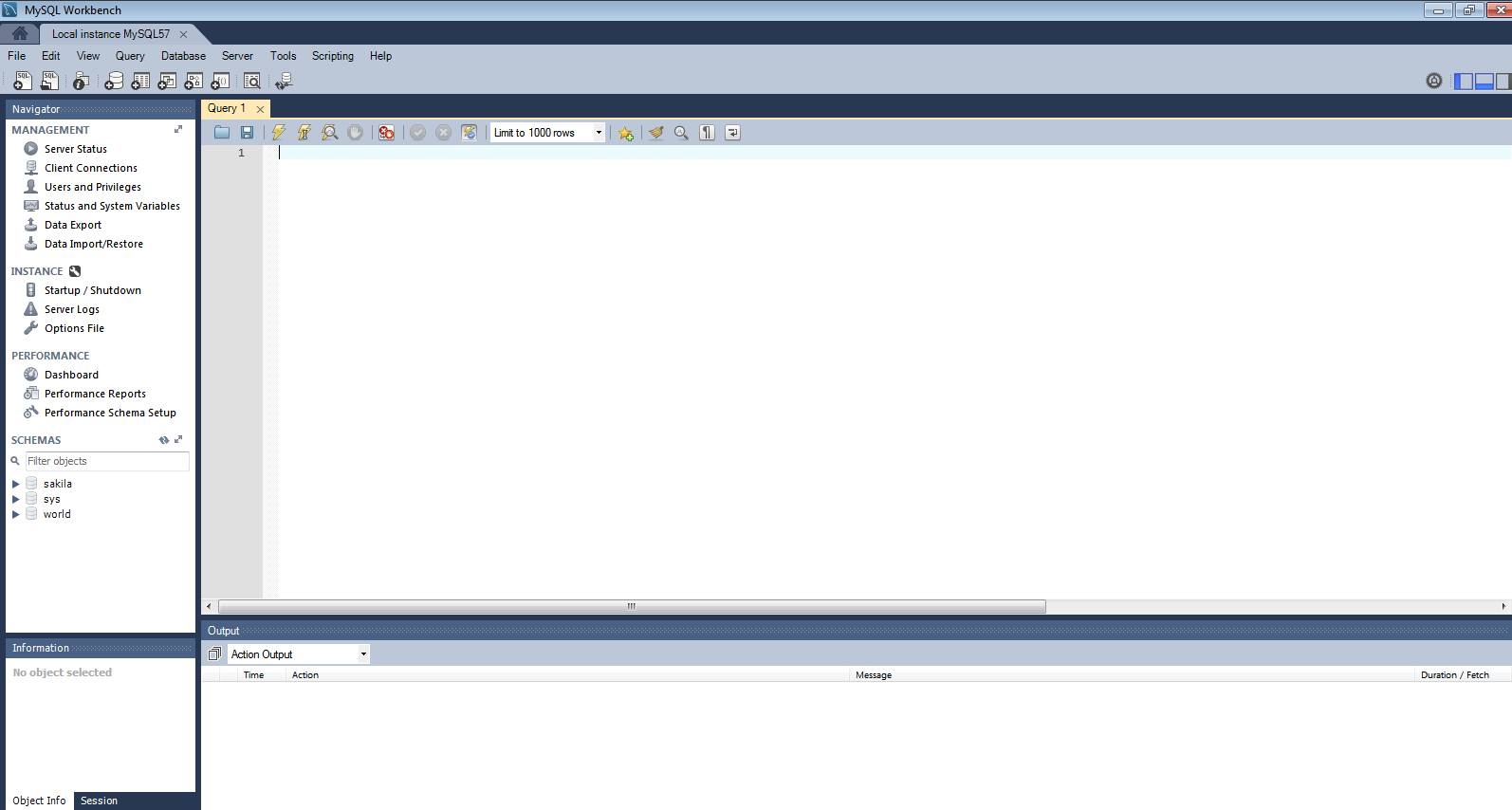
# **Downloading and Importing Data into MySQL Workbench**

Go to <http://cssip.org/docs/publications/bigdataworkbook.zip> and download the dump file containing the databases you will be working with

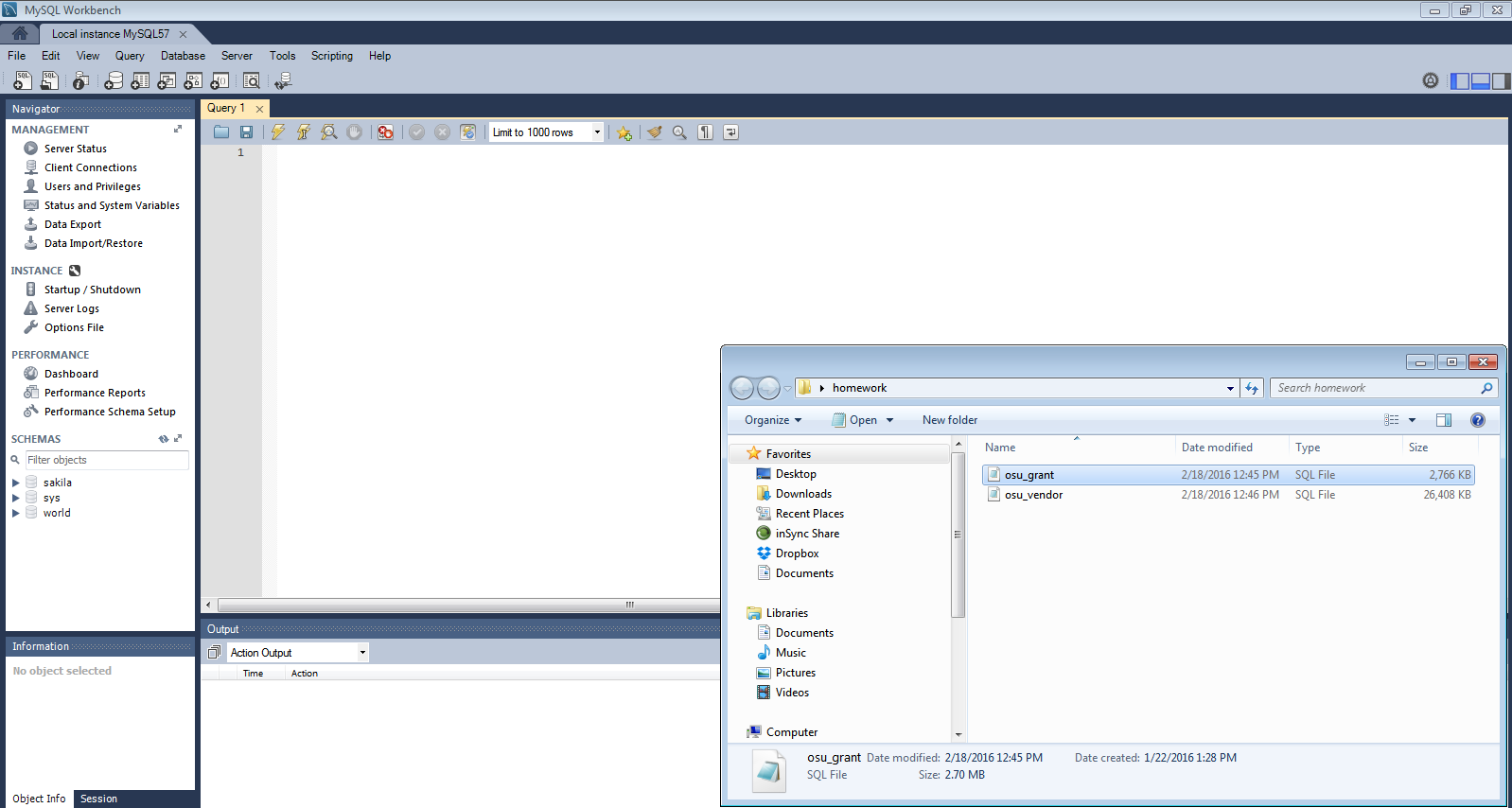
Open MySQL Workbench. There should be an available connection based upon your initial MySQL setup. Double click the “Local instance” icon to open your local server.



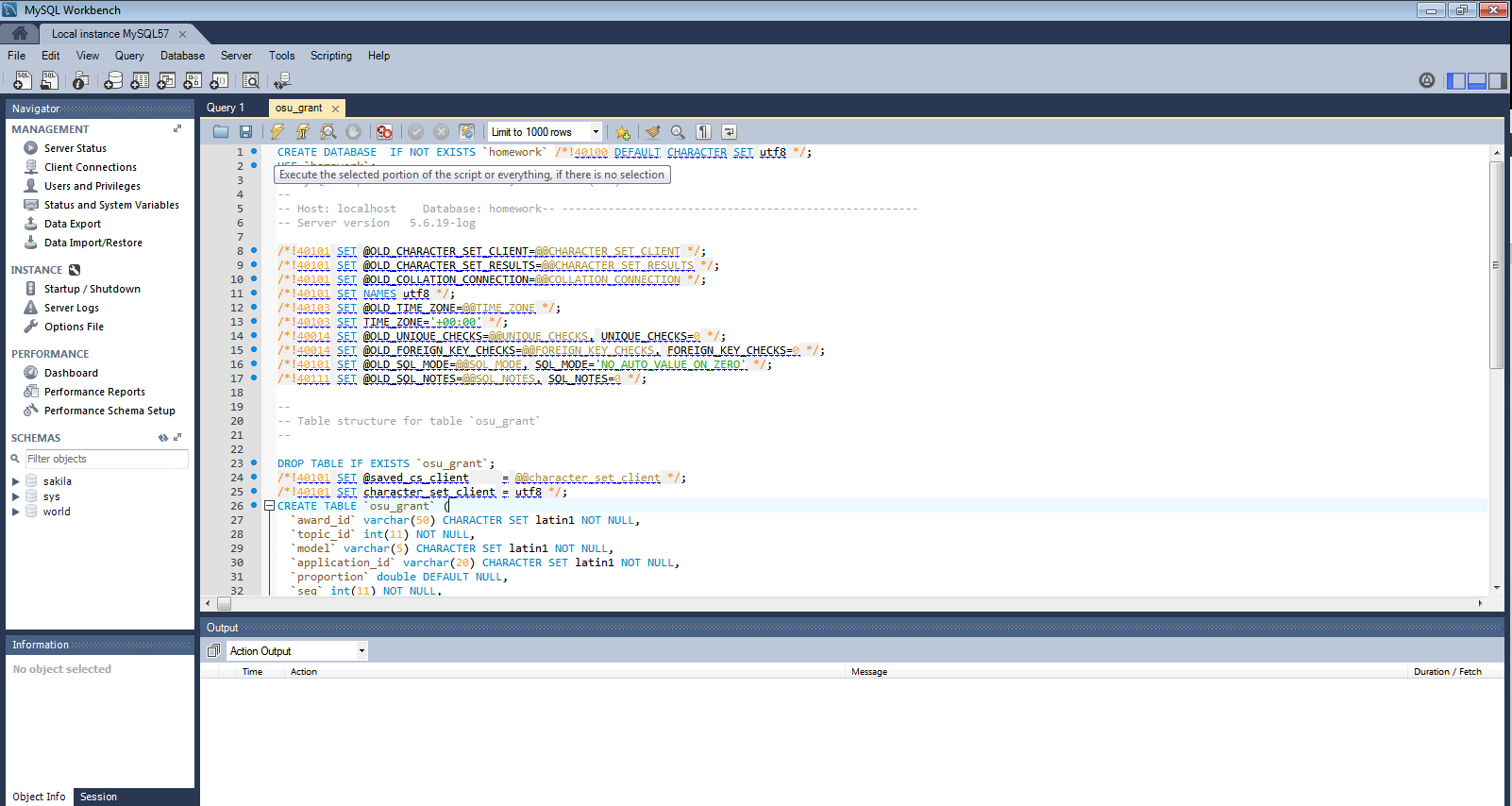
Once you open the connection to your local server, you should be directed to the screen below



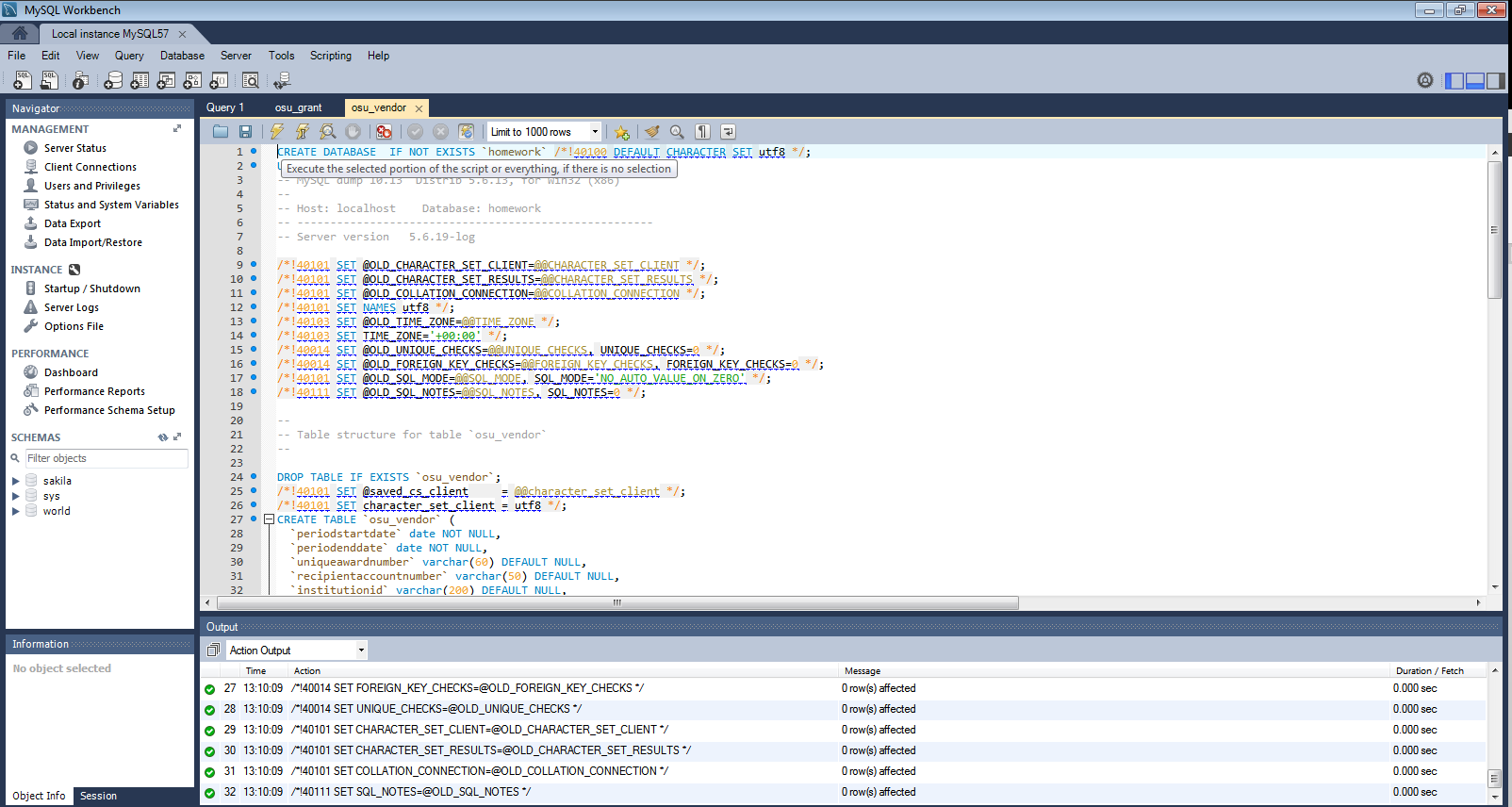
Toggle back to the downloaded database files, then drag-and-drop the osu\_grant file into the “Query 1” window



Once the window has populated with the file script, click the first lightning bolt button on at the top to execute the script, creating the homework database and osu\_grant table



Perform the same process as before by dragging, dropping, and then running the script for the osu\_vendor file



Confirm you have successfully created and imported the homework database by clicking the refresh button next to “Schemas” on the left-hand toolbar. You should be able to see the osu\_grant and osu\_vendor tables listed under the homework database.

