

NC STATE UNIVERSITY

CROP AND SOIL SCIENCES

CS 590 Special Topics

Programming and Data Science for Applied Research

Course Concepts and Workflow

WHAT DO YOU HOPE TO LEARN IN THIS CLASS?

Underlying Course Concepts in R and Python

- Basic Programming Skills
- Advanced Programming Topics and Packages
- Data Visualizations and Interactive Packages
- Machine Learning (Supervised and Unsupervised) Algorithms
- Section Exercises and Capstone Project

Environment Setup – Anaconda (Prompt)

Objectives

- Install the Anaconda Working Environment
- Install R and Python3 Kernels for Jupyter Notebooks
- Open and Explore Jupyter Notebooks
- GitHub Exercise

Let's Browse the Syllabus!

Anaconda and Jupyter Notebooks

- Anaconda is a distributions of R and Python with working environments for many programming languages – via Jupyter Notebooks
- Jupyter Notebooks is an **I**ntegrated **W**orking **E**nvironment (IDE) that manages multiple programming languages (kernels) under one environment
- Jupyter Notebooks gives programmers the ability to write code, display images, and write markdown notes for future references all in one window
- Jupyter Notebooks is the most popular IDE in data science and it is a great learning tool

Disclaimer...

Jupyter Notebooks is not the only IDE available for R and Python programming

Please feel free to use any IDE for R or Python (e.g. Sublime, TextWrangler, R Studio, Vi/Vim, Visual Studio, Atom, etc.)

It won't hurt my feelings... much

Anaconda Setup and Installation

Determine the Computer for Setup and Installation

- Local Machine (Lab Computer)
- Remote Desktop Connection (IP Address Required)
- Online Only (Binder/Docker Setup)

Let's Download Anaconda!

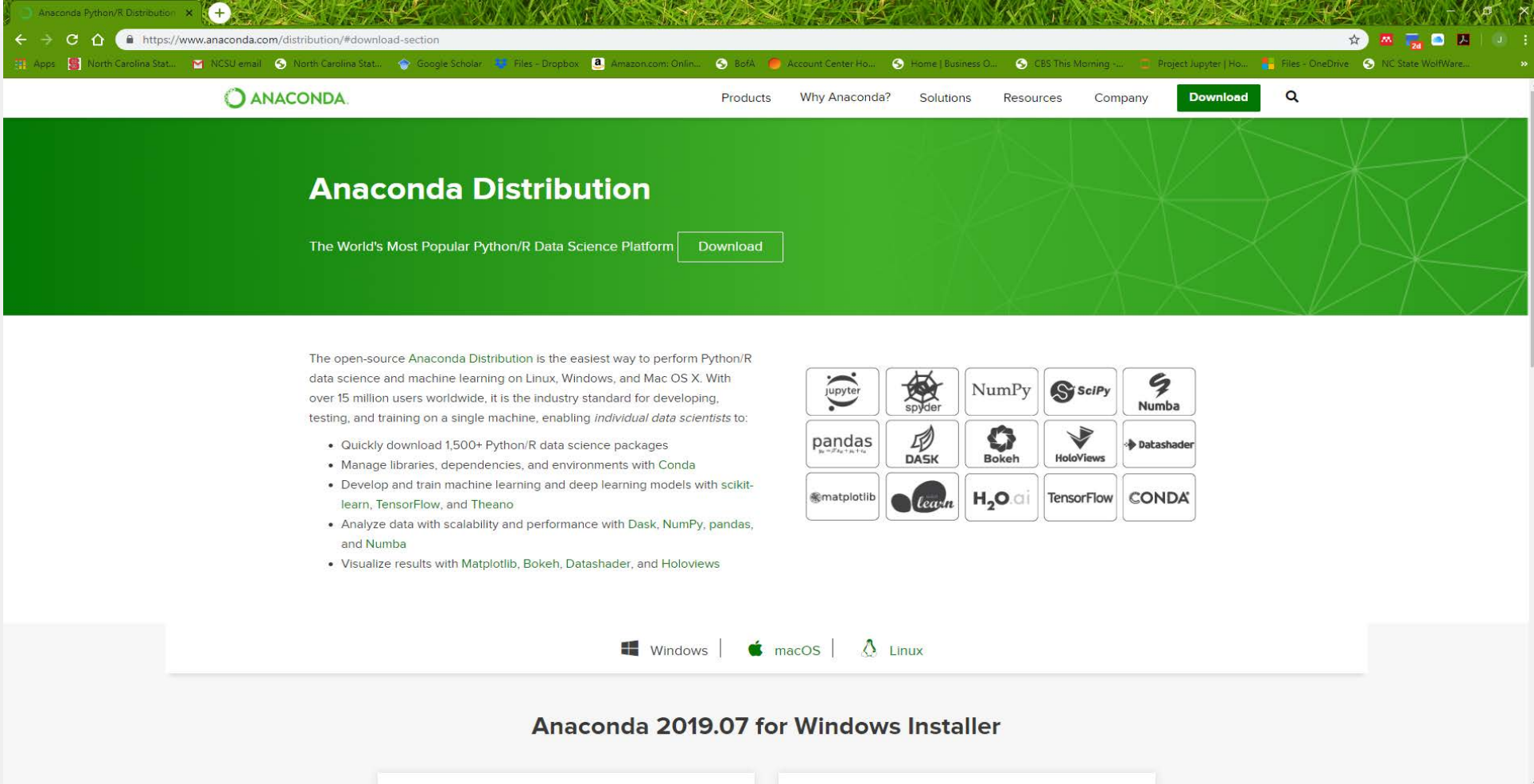
- Go To:

<https://www.anaconda.com/distribution/>

- Or Google Search

Anaconda 3 – Click the first link (Anaconda Python/R Distribution – Free Download)

Download Anaconda 3 – Home Page



The screenshot shows the Anaconda Distribution home page in a web browser. The browser's address bar displays the URL <https://www.anaconda.com/distribution/#download-section>. The page features a green header with the Anaconda logo and navigation links: Products, Why Anaconda?, Solutions, Resources, Company, and a prominent green Download button. Below the header, a large green banner contains the text "Anaconda Distribution" and "The World's Most Popular Python/R Data Science Platform" with a Download button. The main content area describes the open-source distribution as the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. It lists benefits such as quickly downloading 1,500+ Python/R data science packages, managing libraries and dependencies with Conda, and developing machine learning models with scikit-learn, TensorFlow, and Theano. A grid of logos for various data science libraries is displayed, including Jupyter, Spyder, NumPy, SciPy, Numba, pandas, Dask, Bokeh, HoloViews, Datashader, matplotlib, and TensorFlow. At the bottom, there are links for Windows, macOS, and Linux, and a section for the "Anaconda 2019.07 for Windows Installer".

ANACONDA

Products Why Anaconda? Solutions Resources Company **Download**

Anaconda Distribution

The World's Most Popular Python/R Data Science Platform **Download**

The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 15 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with Conda
- Develop and train machine learning and deep learning models with scikit-learn, TensorFlow, and Theano
- Analyze data with scalability and performance with Dask, NumPy, pandas, and Numba
- Visualize results with Matplotlib, Bokeh, Datashader, and HoloViews

jupyter spyder NumPy SciPy Numba

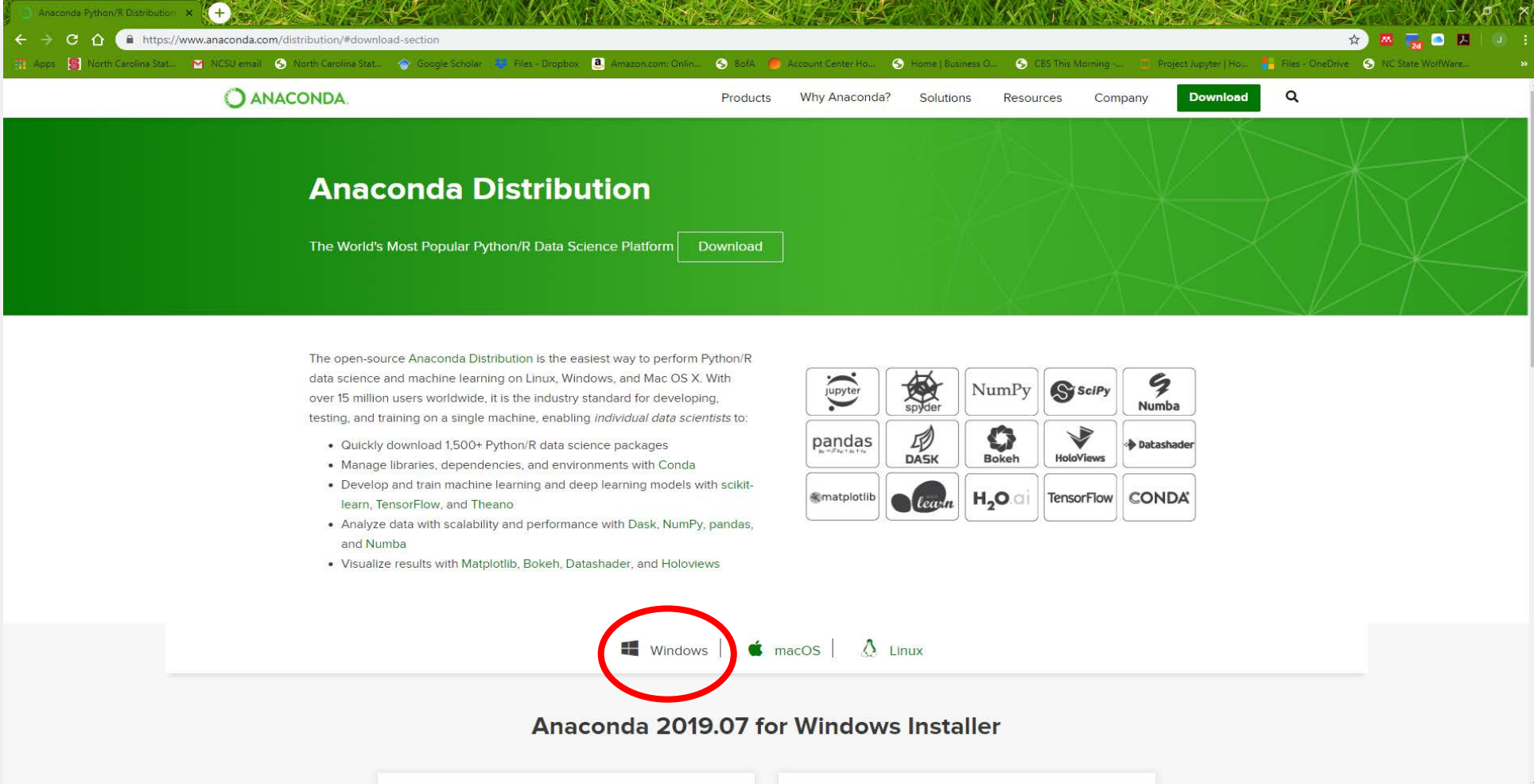
pandas DASK Bokeh HoloViews Datashader

matplotlib tensorflow H2O.ai TensorFlow CONDA

Windows | macOS | Linux

Anaconda 2019.07 for Windows Installer

Download Anaconda 3 – Home Page



The screenshot shows the Anaconda Distribution website. The browser's address bar displays <https://www.anaconda.com/distribution/#download-section>. The page features a green header with the Anaconda logo and navigation links: Products, Why Anaconda?, Solutions, Resources, Company, and a prominent green Download button. Below the header, a large green banner contains the text "Anaconda Distribution" and "The World's Most Popular Python/R Data Science Platform" with a Download button. The main content area describes the open-source distribution as the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. It lists benefits such as quickly downloading 1,500+ Python/R data science packages, managing libraries and dependencies with Conda, and developing machine learning models with scikit-learn, TensorFlow, and Theano. A grid of logos for various data science libraries is displayed, including Jupyter, Spyder, NumPy, SciPy, Numba, pandas, Dask, Bokeh, HoloViews, Datashader, matplotlib, and H2O.ai. At the bottom, a navigation bar shows icons for Windows, macOS, and Linux. The Windows icon is circled in red, and the text "Anaconda 2019.07 for Windows Installer" is visible below it.

ANACONDA

Products Why Anaconda? Solutions Resources Company **Download**

Anaconda Distribution

The World's Most Popular Python/R Data Science Platform **Download**

The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 15 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with Conda
- Develop and train machine learning and deep learning models with scikit-learn, TensorFlow, and Theano
- Analyze data with scalability and performance with Dask, NumPy, pandas, and Numba
- Visualize results with Matplotlib, Bokeh, Datashader, and HoloViews

jupyter spyder NumPy SciPy Numba

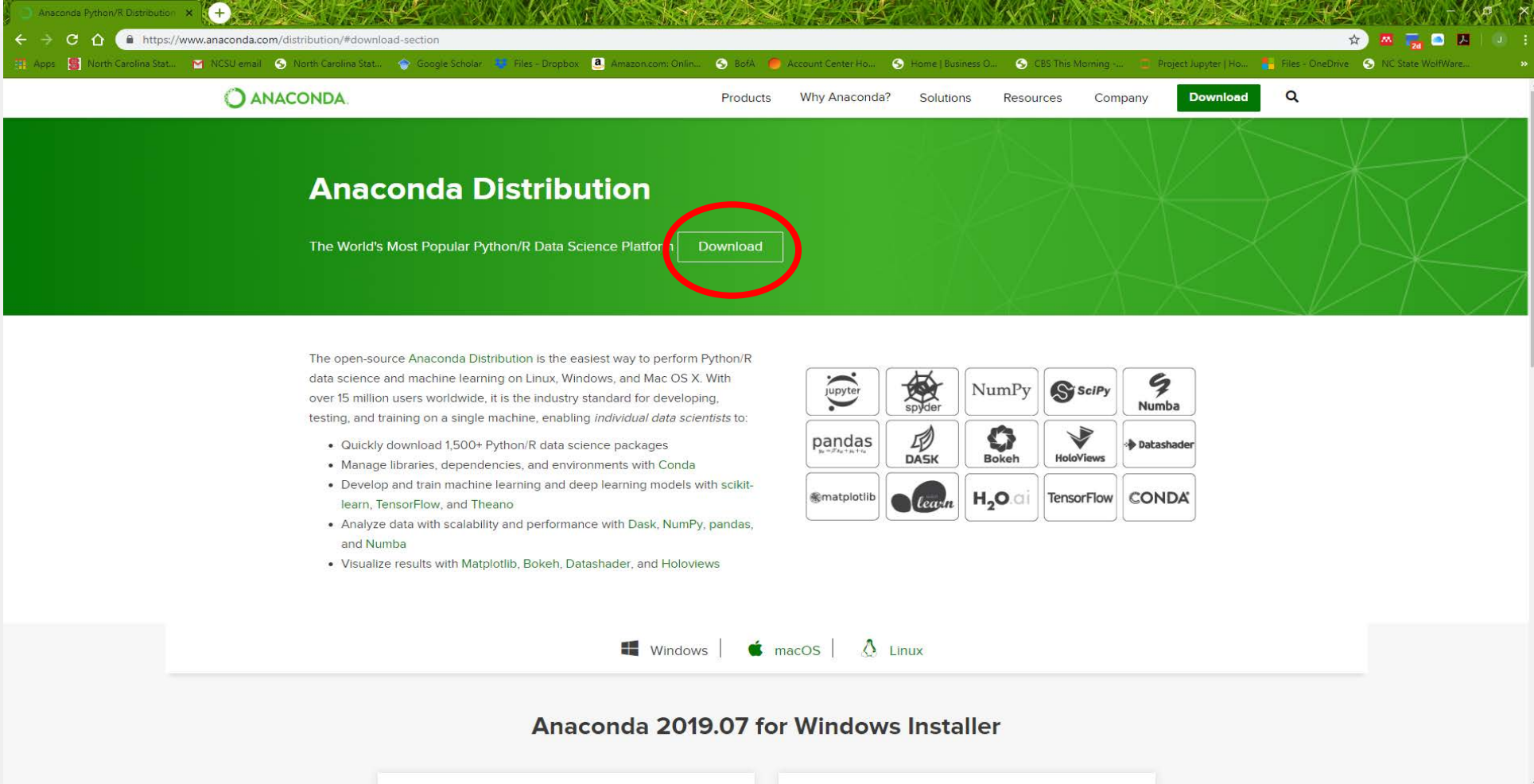
pandas DASK Bokeh HoloViews Datashader

matplotlib H2O.ai TensorFlow CONDA

Windows | macOS | Linux

Anaconda 2019.07 for Windows Installer

Download Anaconda 3 – Home Page



The screenshot shows the Anaconda Distribution website. The browser's address bar displays `https://www.anaconda.com/distribution/#download-section`. The page features a green header with the Anaconda logo and navigation links: Products, Why Anaconda?, Solutions, Resources, Company, and a prominent green **Download** button. Below the header, a large green banner contains the text "Anaconda Distribution" and "The World's Most Popular Python/R Data Science Platform", with a white **Download** button circled in red. The main content area describes the open-source distribution and lists its benefits. To the right, a grid of logos for various data science libraries is shown. At the bottom, there are links for Windows, macOS, and Linux, followed by the text "Anaconda 2019.07 for Windows Installer".

ANACONDA

Products Why Anaconda? Solutions Resources Company **Download**

Anaconda Distribution

The World's Most Popular Python/R Data Science Platform **Download**

The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 15 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with Conda
- Develop and train machine learning and deep learning models with scikit-learn, TensorFlow, and Theano
- Analyze data with scalability and performance with Dask, NumPy, pandas, and Numba
- Visualize results with Matplotlib, Bokeh, Datashader, and Holoviews

jupyter spyder NumPy SciPy Numba
pandas DASK Bokeh HoloViews Datashader
matplotlib scikit-learn H2O.ai TensorFlow CONDA

Windows | macOS | Linux

Anaconda 2019.07 for Windows Installer

Download Anaconda 3 – Windows Installer

The screenshot shows the Anaconda website's download section for Windows. The browser's address bar shows the URL <https://www.anaconda.com/distribution/#download-section>. The page has a green header with navigation links for Windows, macOS, and Linux. The main heading is "Anaconda 2019.07 for Windows Installer". Below this, there are two white boxes for different Python versions. The "Python 3.7 version" box contains a green "Download" button and lists two installers: "64-Bit Graphical Installer (486 MB)" and "32-Bit Graphical Installer (418 MB)". The "Python 2.7 version" box contains a green "Download" button and lists two installers: "64-Bit Graphical Installer (427 MB)" and "32-Bit Graphical Installer (361 MB)". Below these boxes is a section titled "Get Started with Anaconda Distribution" with five white boxes containing links to documentation, blog, community support, webinars, and training. The bottom of the page features a green taskbar with a taskbar icon for "Anaconda3-2019.0...exe" and a "Show all" button.

Windows | macOS | Linux

Anaconda 2019.07 for Windows Installer

Python 3.7 version

[Download](#)

64-Bit Graphical Installer (486 MB)
32-Bit Graphical Installer (418 MB)

Python 2.7 version

[Download](#)

64-Bit Graphical Installer (427 MB)
32-Bit Graphical Installer (361 MB)

Get Started with Anaconda Distribution

Documentation

Installation and user guide for Anaconda Distribution 5

[Read More](#)

Anaconda Blog

News, software releases, and developer best practices

[Read More](#)

Community Support

Solutions and knowledge from the community

[Read More](#)

Anaconda Webinars

Industry trends and tutorials from Anaconda

[Read More](#)

Anaconda Training

Learn Python for Data Science with DataCamp

[Start Learning](#)

Anaconda3-2019.0...exe [Show all](#)

Download Anaconda 3 – Python 3.7 Version

The screenshot shows the Anaconda website's download section for Windows. The browser's address bar indicates the URL is <https://www.anaconda.com/distribution/#download-section>. The page title is "Anaconda 2019.07 for Windows Installer". There are two main download options:

- Python 3.7 version**: This option is highlighted with a red circle around its "Download" button. Below the button, it lists "64-Bit Graphical Installer (486 MB)" and "32-Bit Graphical Installer (418 MB)".
- Python 2.7 version**: This option has a "Download" button and lists "64-Bit Graphical Installer (427 MB)" and "32-Bit Graphical Installer (361 MB)".

Below the download options is a section titled "Get Started with Anaconda Distribution" with five links:

- Documentation**: Installation and user guide for Anaconda Distribution 5. [Read More](#)
- Anaconda Blog**: News, software releases, and developer best practices. [Read More](#)
- Community Support**: Solutions and knowledge from the community. [Read More](#)
- Anaconda Webinars**: Industry trends and tutorials from Anaconda. [Read More](#)
- Anaconda Training**: Learn Python for Data Science with DataCamp. [Start Learning](#)

The bottom of the browser window shows a taskbar with the file "Anaconda3-2019.0...exe" open.

Download Anaconda 3 – Python 3.7 Version

Windows | macOS | Linux

Anaconda 2019.07 for Windows Installer

Python 3.7 version

[Download](#)

64-Bit Graphical Installer (486 MB)
32-Bit Graphical Installer (418 MB)

Python 2.7 version

[Download](#)

64-Bit Graphical Installer (427 MB)
32-Bit Graphical Installer (361 MB)

Get Started with Anaconda Distribution

Documentation

Installation and user guide for Anaconda Distribution 5

[Read More](#)

Anaconda Blog

News, software releases, and developer best practices

[Read More](#)

Community Support

Solutions and knowledge from the community

[Read More](#)

Anaconda Webinars

Industry trends and tutorials from Anaconda

[Read More](#)

Anaconda Training

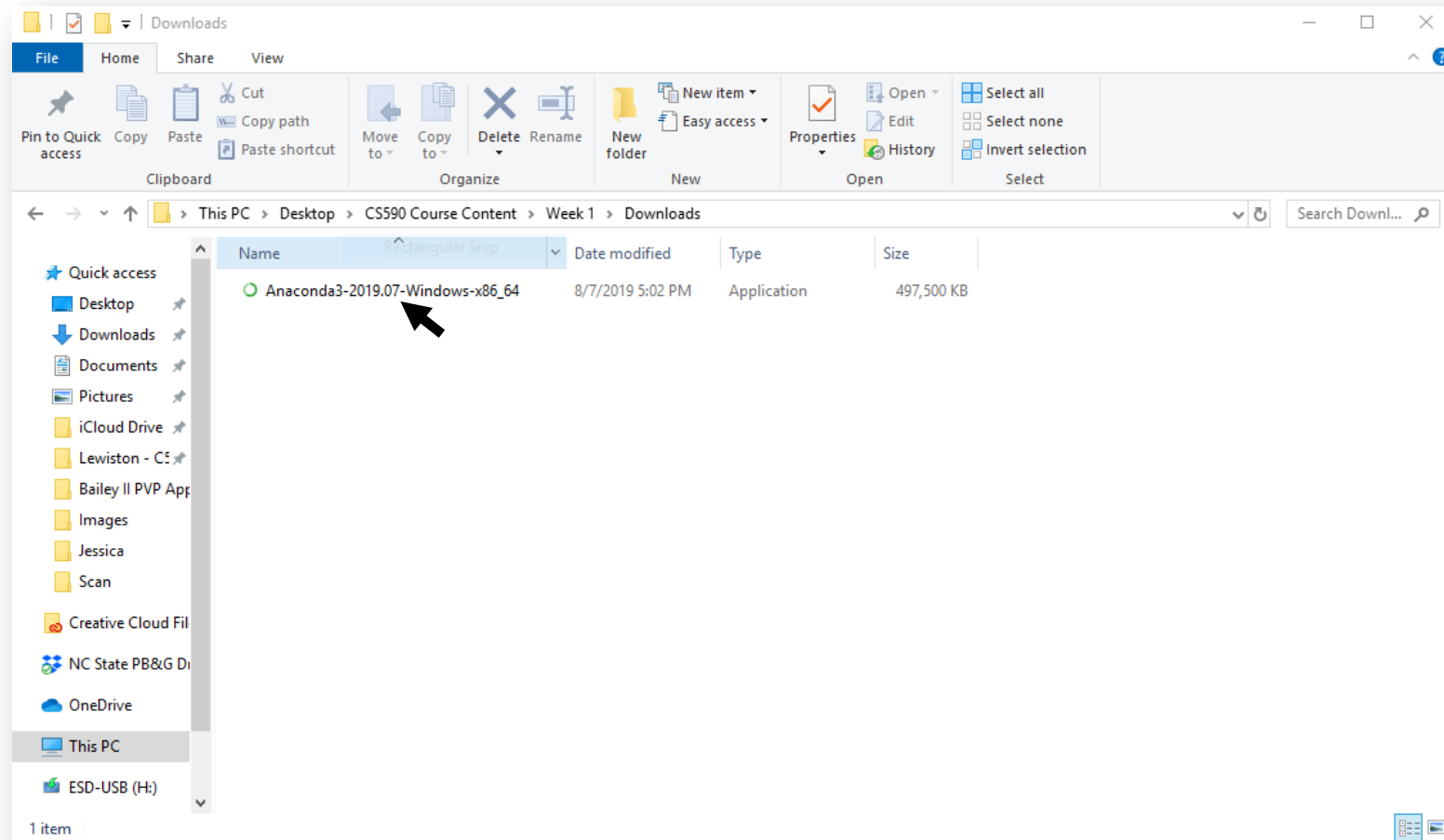
Learn Python for Data Science with DataCamp

[Start Learning](#)

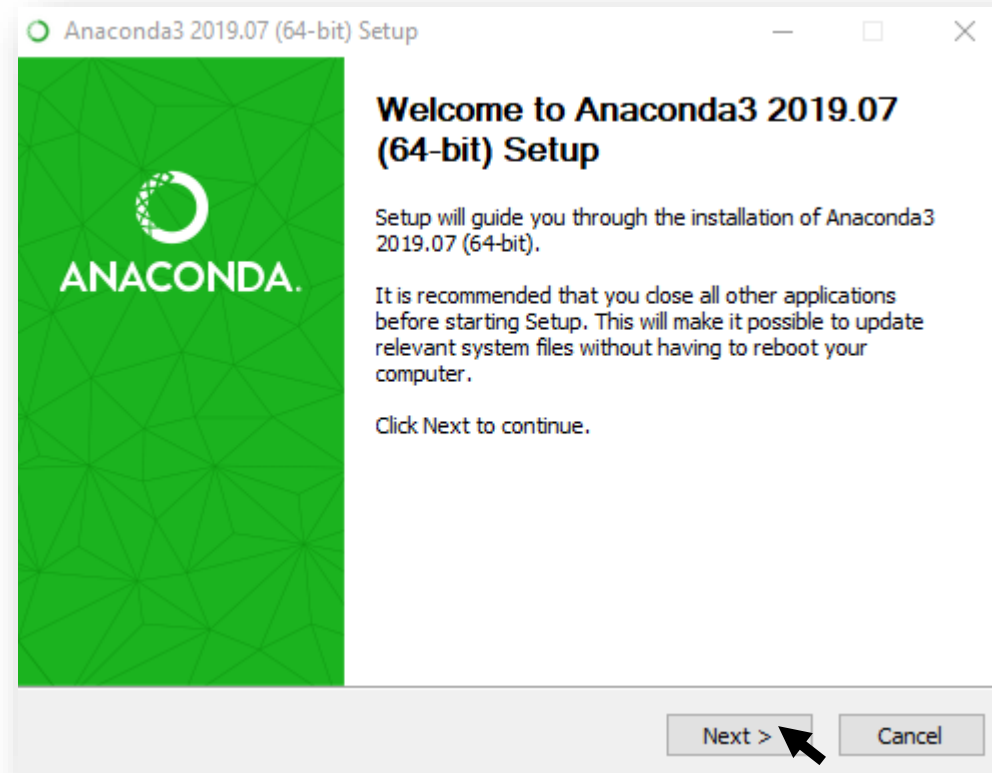
Anaconda3-2019.0...exe

Show all

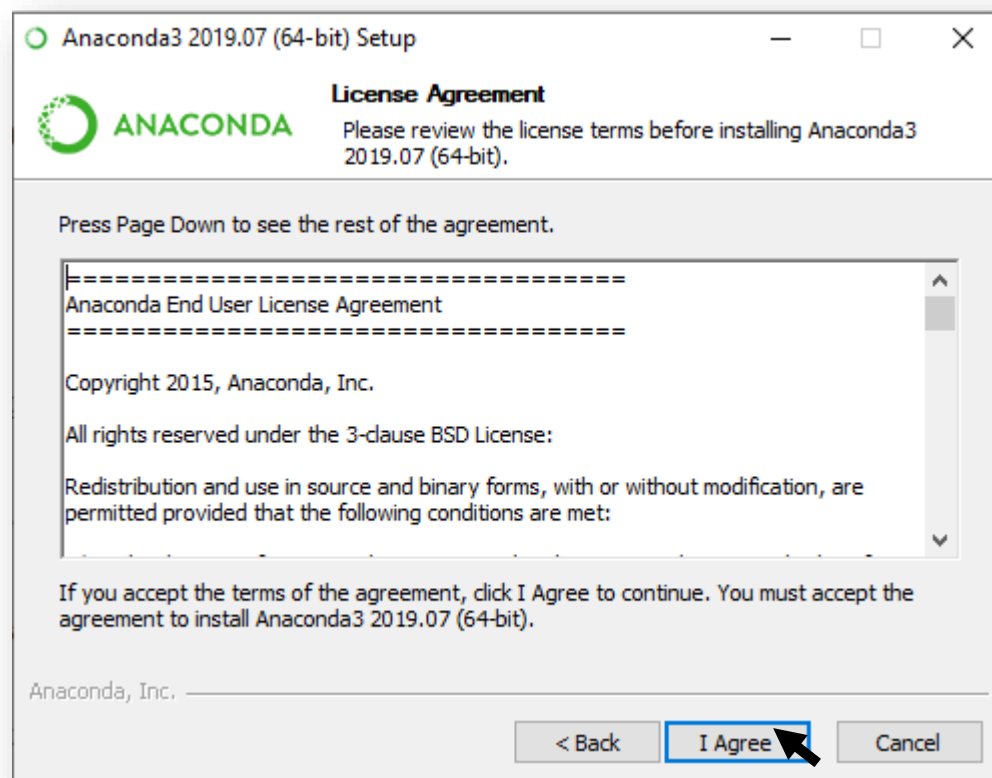
Anaconda 3 Executable File



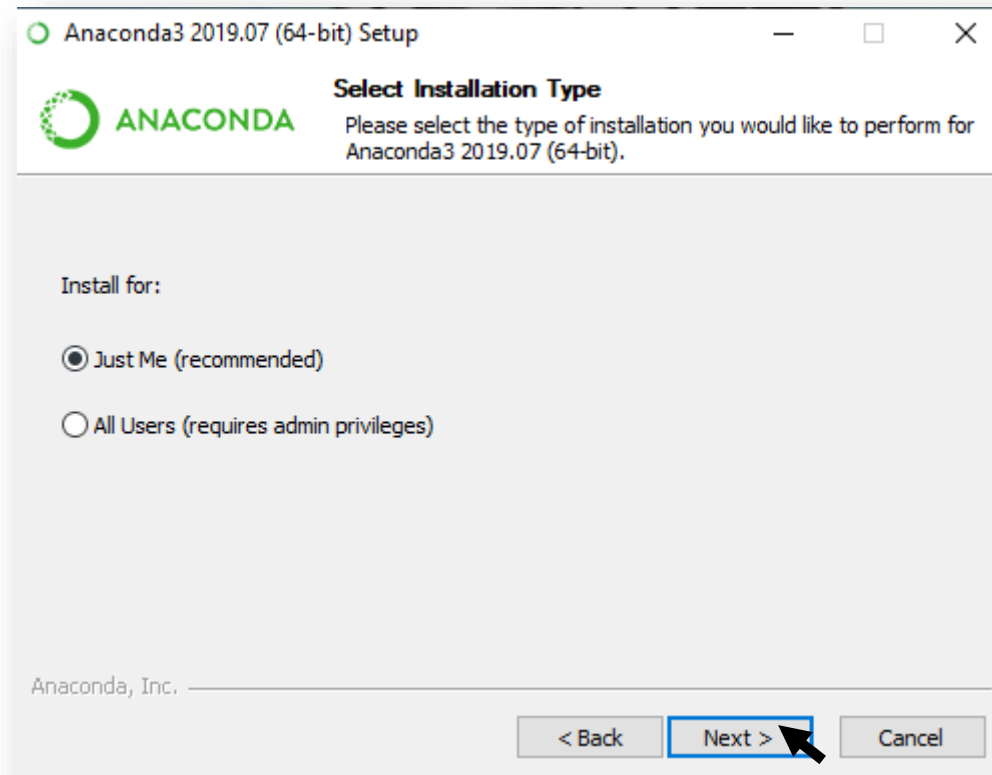
Anaconda 3 Installation and Setup



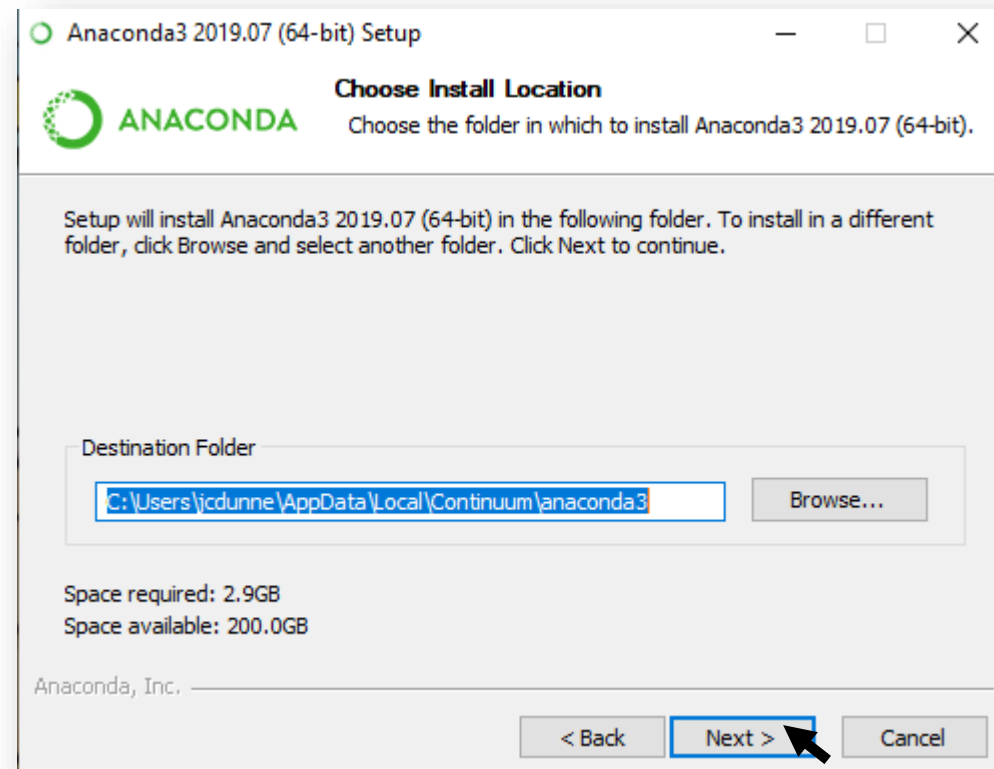
Anaconda 3 Installation and Setup



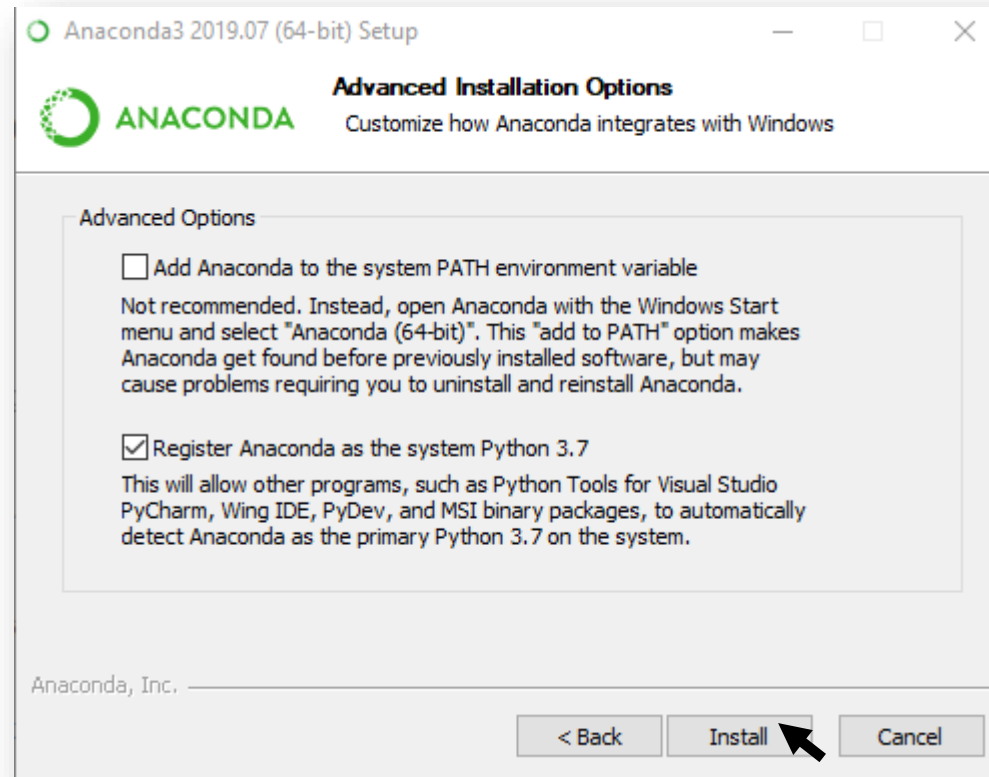
Anaconda 3 Installation and Setup



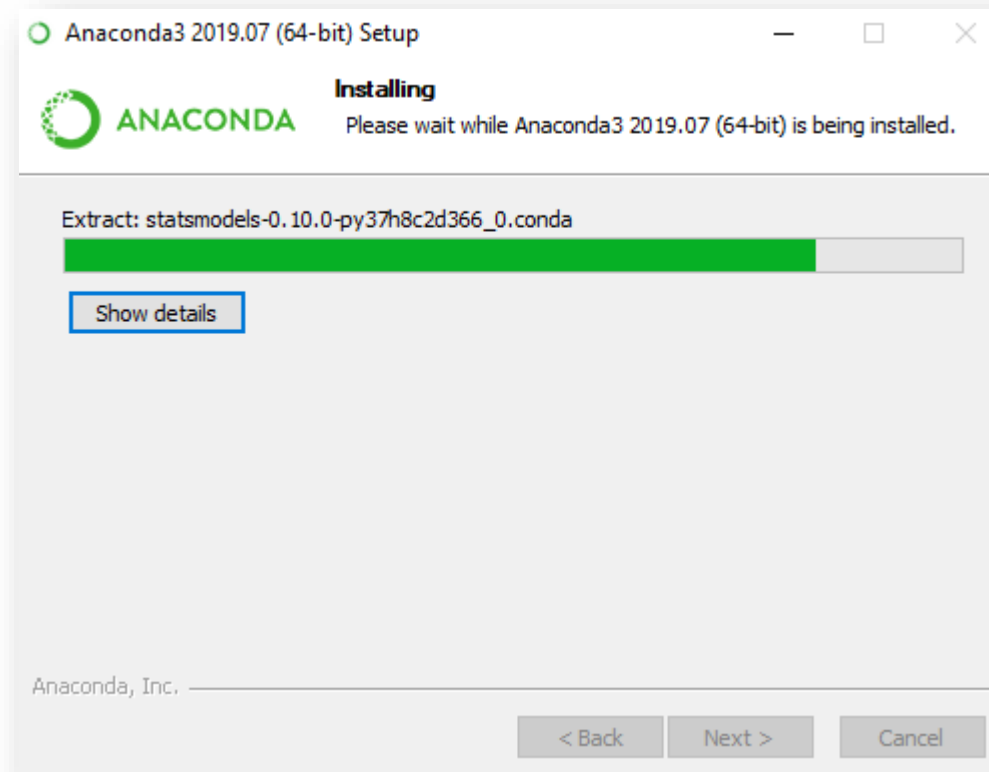
Anaconda 3 Installation and Setup



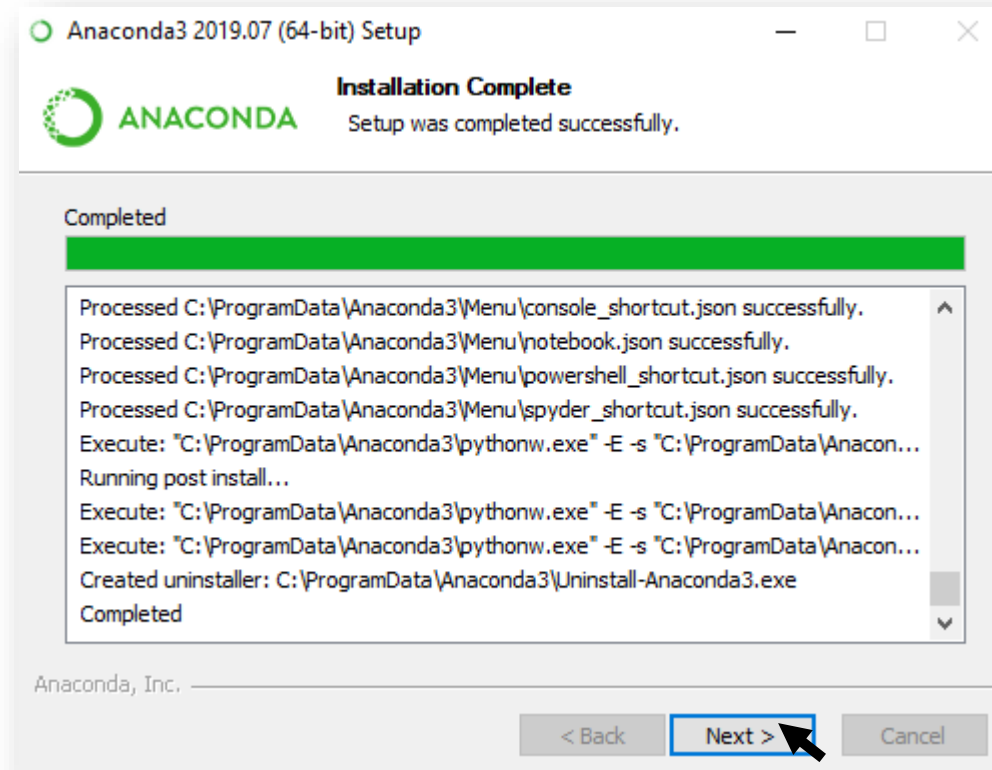
Anaconda 3 Installation and Setup



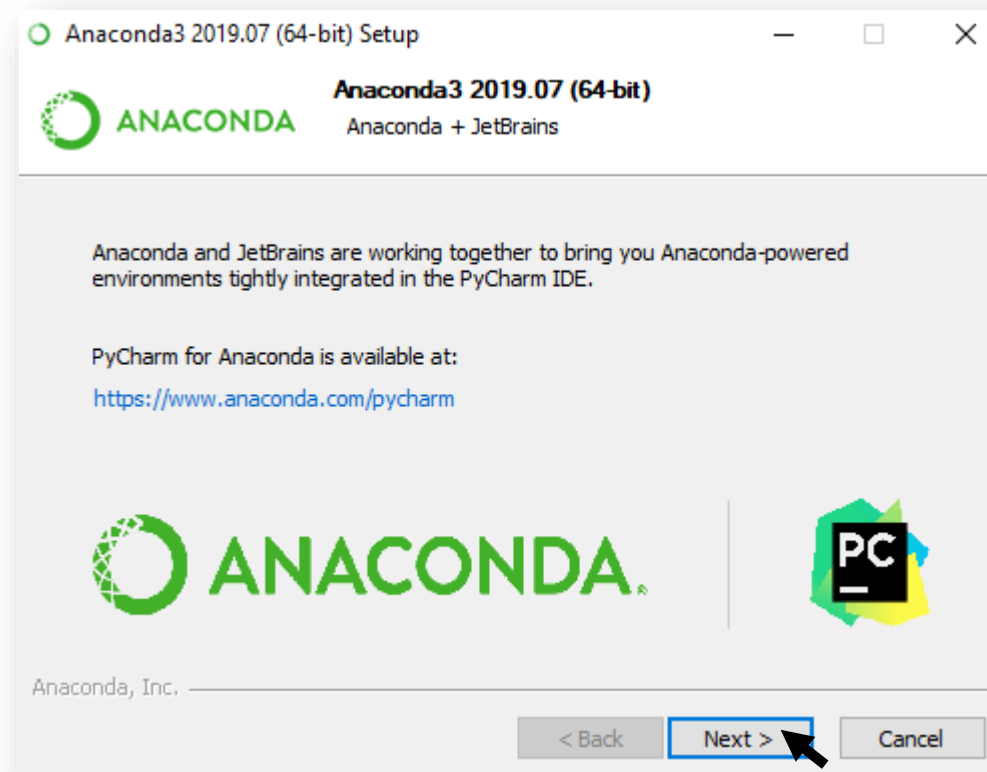
Anaconda 3 Installation and Setup



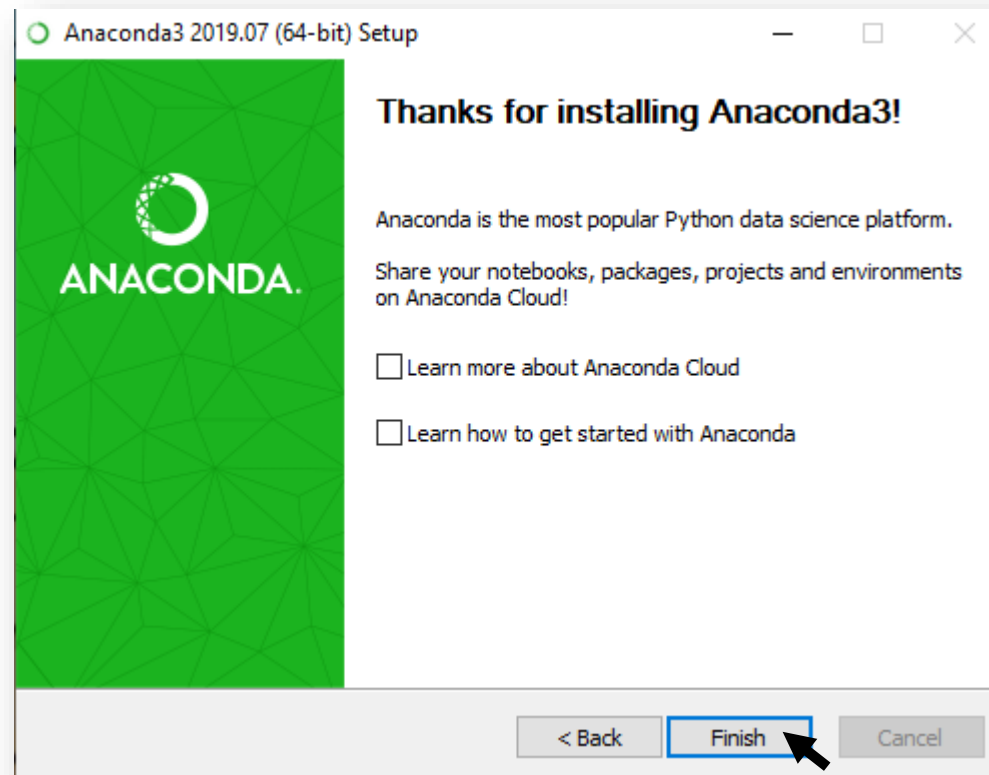
Anaconda 3 Installation and Setup



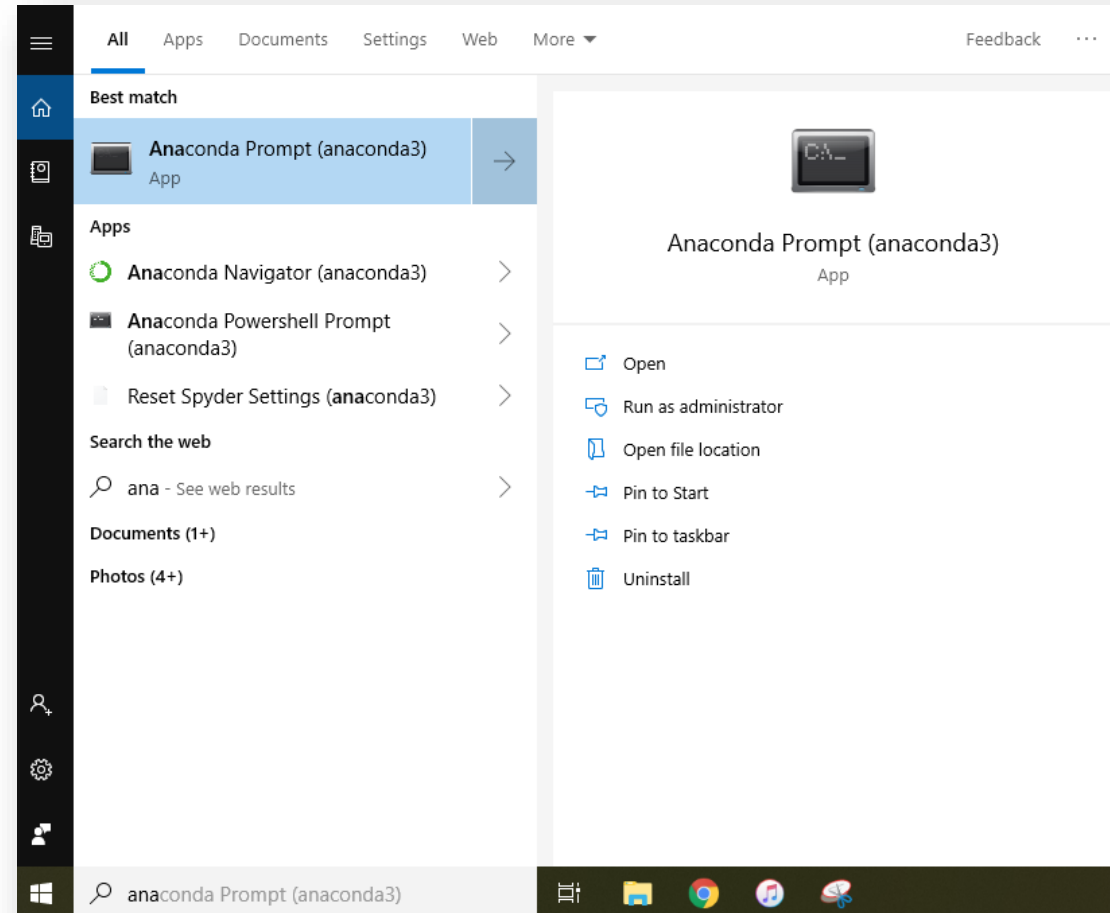
Anaconda 3 Installation and Setup



Anaconda 3 Installation and Setup



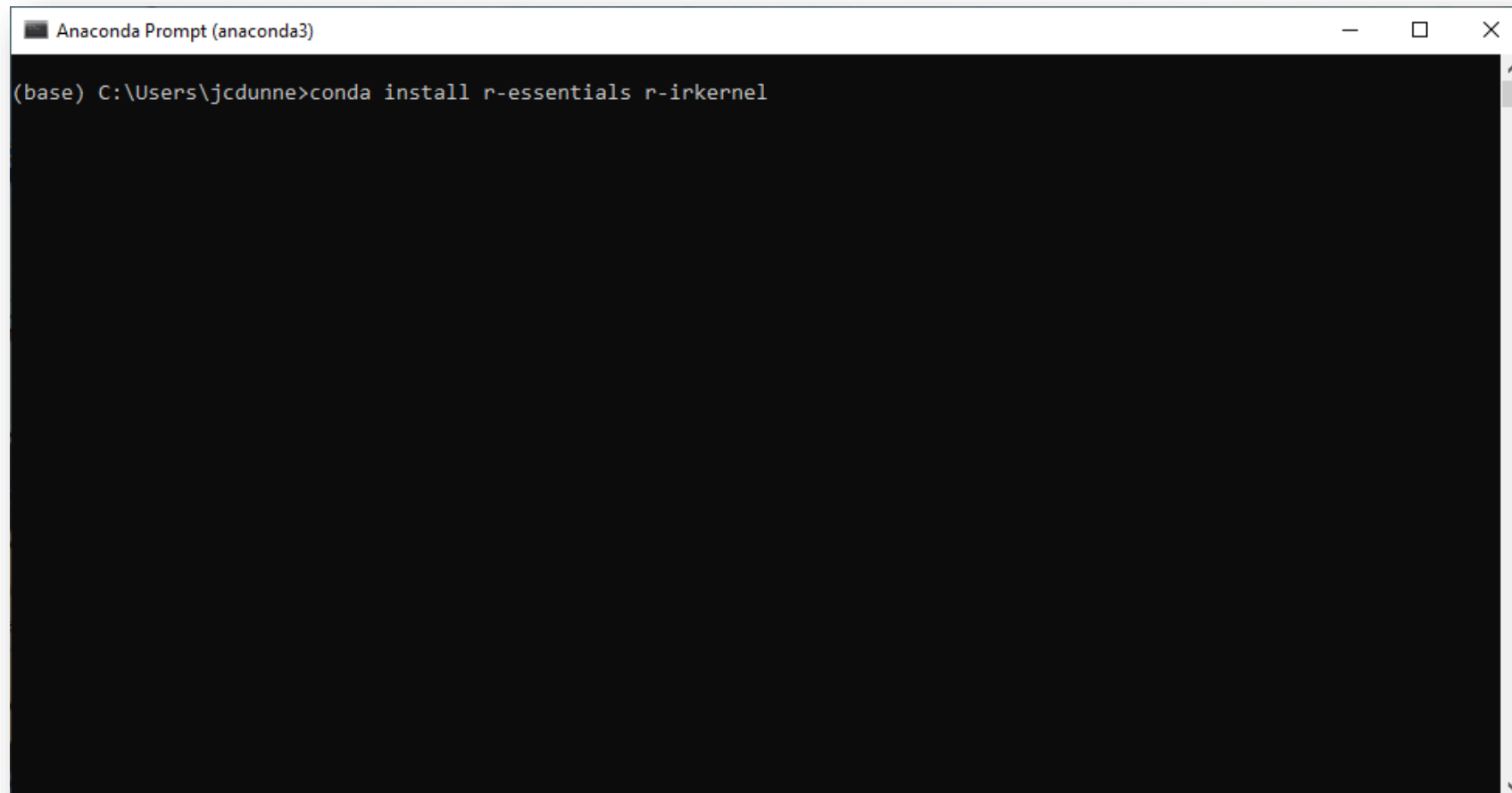
Installing R / R-Kernel for Jupyter Notebooks



*** MAC Users > Finder > Search “Terminal” > Open Terminal



Installing R / R-Kernel for Jupyter Notebooks

A screenshot of an Anaconda Prompt window. The title bar reads "Anaconda Prompt (anaconda3)". The command prompt shows the command `conda install r-essentials r-irkernel` being entered. The prompt is in a dark-themed environment with a black background and white text. The window has standard Windows window controls (minimize, maximize, close) in the top right corner.

```
Anaconda Prompt (anaconda3)
(base) C:\Users\jcdunne>conda install r-essentials r-irkernel
```

> conda install r-essentials r-irkernel

Installing R / R-Kernel for Jupyter Notebooks

```
Anaconda Prompt (anaconda3) - conda install r-essentials

r-tidyr                conda-forge/win-64::r-tidyr-0.8.3-r36h796a38f_1
r-tidysselect          conda-forge/win-64::r-tidysselect-0.2.5-r36h796a38f_1001
r-tidyverse            conda-forge/noarch::r-tidyverse-1.2.1-r36h6115d3f_1002
r-timedate             conda-forge/noarch::r-timedate-3043.102-r36h6115d3f_1001
r-tinytex              conda-forge/noarch::r-tinytex-0.14-r36h6115d3f_1
r-ttr                  conda-forge/win-64::r-ttr-0.23_4-r36h17dded8_1001
r-utf8                 conda-forge/win-64::r-utf8-1.1.4-r36hda5aaf8_1001
r-uuid                 conda-forge/win-64::r-uuid-0.1_2-r36hda5aaf8_1002
r-vctrs                conda-forge/win-64::r-vctrs-0.2.0-r36hda5aaf8_1
r-viridislite          conda-forge/noarch::r-viridislite-0.3.0-r36h6115d3f_1002
r-whisker              conda-forge/noarch::r-whisker-0.3_2-r36h6115d3f_1002
r-withr               conda-forge/noarch::r-withr-2.1.2-r36h6115d3f_1001
r-xfun                 conda-forge/noarch::r-xfun-0.8-r36h6115d3f_1
r-xml2                 conda-forge/win-64::r-xml2-1.2.1-r36h796a38f_0
r-xtable              conda-forge/noarch::r-xtable-1.8_4-r36h6115d3f_2
r-xts                  conda-forge/win-64::r-xts-0.11_2-r36hda5aaf8_1
r-yaml                 conda-forge/win-64::r-yaml-2.2.0-r36hda5aaf8_1002
r-zeallot              conda-forge/noarch::r-zeallot-0.1.0-r36h6115d3f_1001
r-zoo                  conda-forge/win-64::r-zoo-1.8_6-r36hda5aaf8_1

The following packages will be UPDATED:

conda                pkgs/main::conda-4.7.10-py37_0 --> conda-forge::conda-4.7.11-py37_0

The following packages will be SUPERSEDED by a higher-priority channel:

certifi              pkgs/main --> conda-forge

Proceed ([y]/n)?
```

> Proceed ([y]/n)? y

Installing R / R-Kernel for Jupyter Notebooks

```
Anaconda Prompt (anaconda3)

r-utf8      conda-forge/win-64::r-utf8-1.1.4-r36hda5aaf8_1001
r-uuid      conda-forge/win-64::r-uuid-0.1_2-r36hda5aaf8_1002
r-vctrs      conda-forge/win-64::r-vctrs-0.2.0-r36hda5aaf8_1
r-iridislite conda-forge/noarch::r-iridislite-0.3.0-r36h6115d3f_1002
r-whisker    conda-forge/noarch::r-whisker-0.3_2-r36h6115d3f_1002
r-withr      conda-forge/noarch::r-withr-2.1.2-r36h6115d3f_1001
r-xfun       conda-forge/noarch::r-xfun-0.8-r36h6115d3f_1
r-xml2       conda-forge/win-64::r-xml2-1.2.1-r36h796a38f_0
r-xtable     conda-forge/noarch::r-xtable-1.8_4-r36h6115d3f_2
r-xts        conda-forge/win-64::r-xts-0.11_2-r36hda5aaf8_1
r-yaml       conda-forge/win-64::r-yaml-2.2.0-r36hda5aaf8_1002
r-zeallot    conda-forge/noarch::r-zeallot-0.1.0-r36h6115d3f_1001
r-zoo        conda-forge/win-64::r-zoo-1.8_6-r36hda5aaf8_1

The following packages will be UPDATED:

conda                pkgs/main::conda-4.7.10-py37_0 --> conda-forge::conda-4.7.11-py37_0

The following packages will be SUPERSEDED by a higher-priority channel:

certifi              pkgs/main --> conda-forge

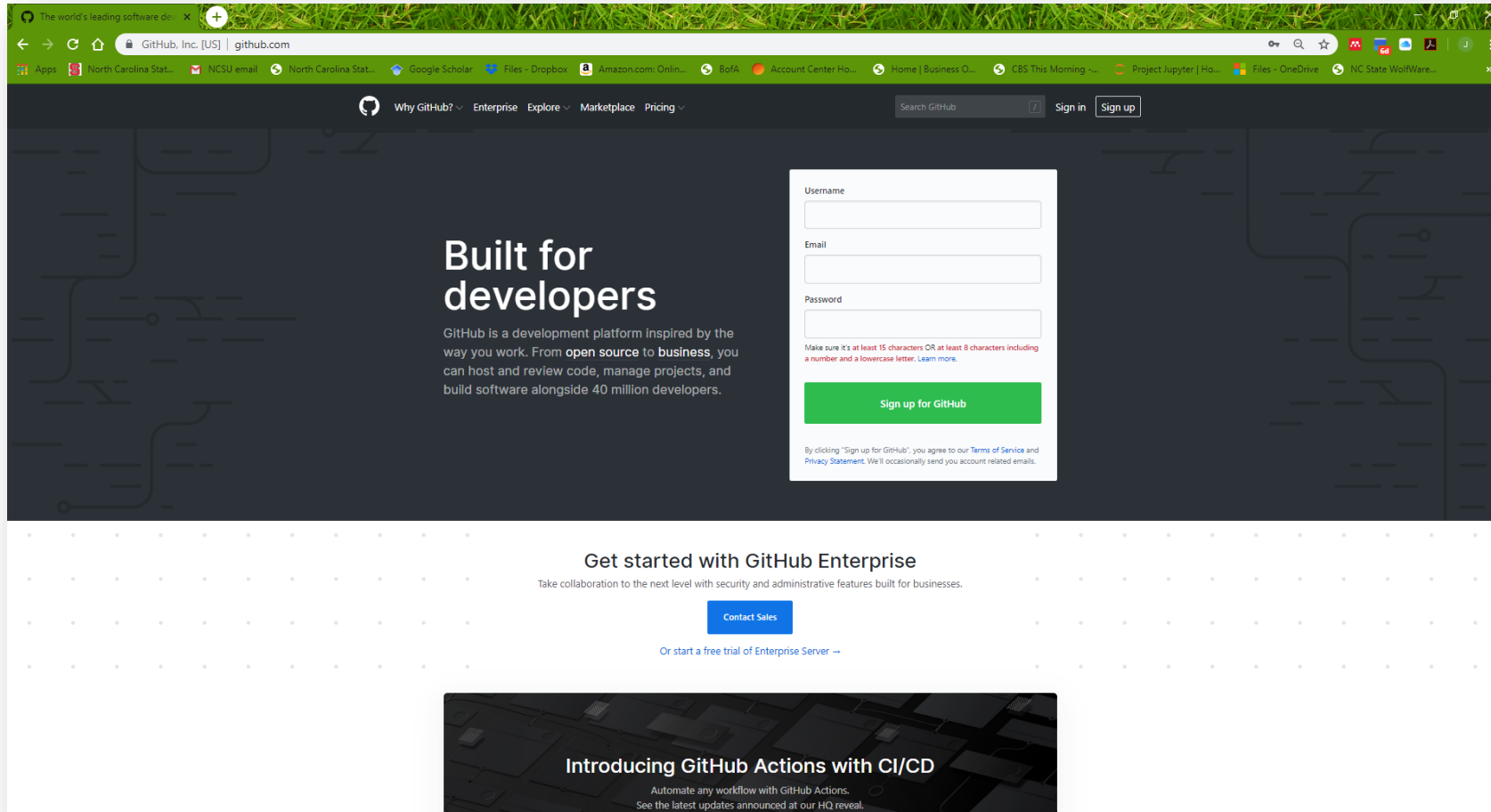
Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: done
Executing transaction: done

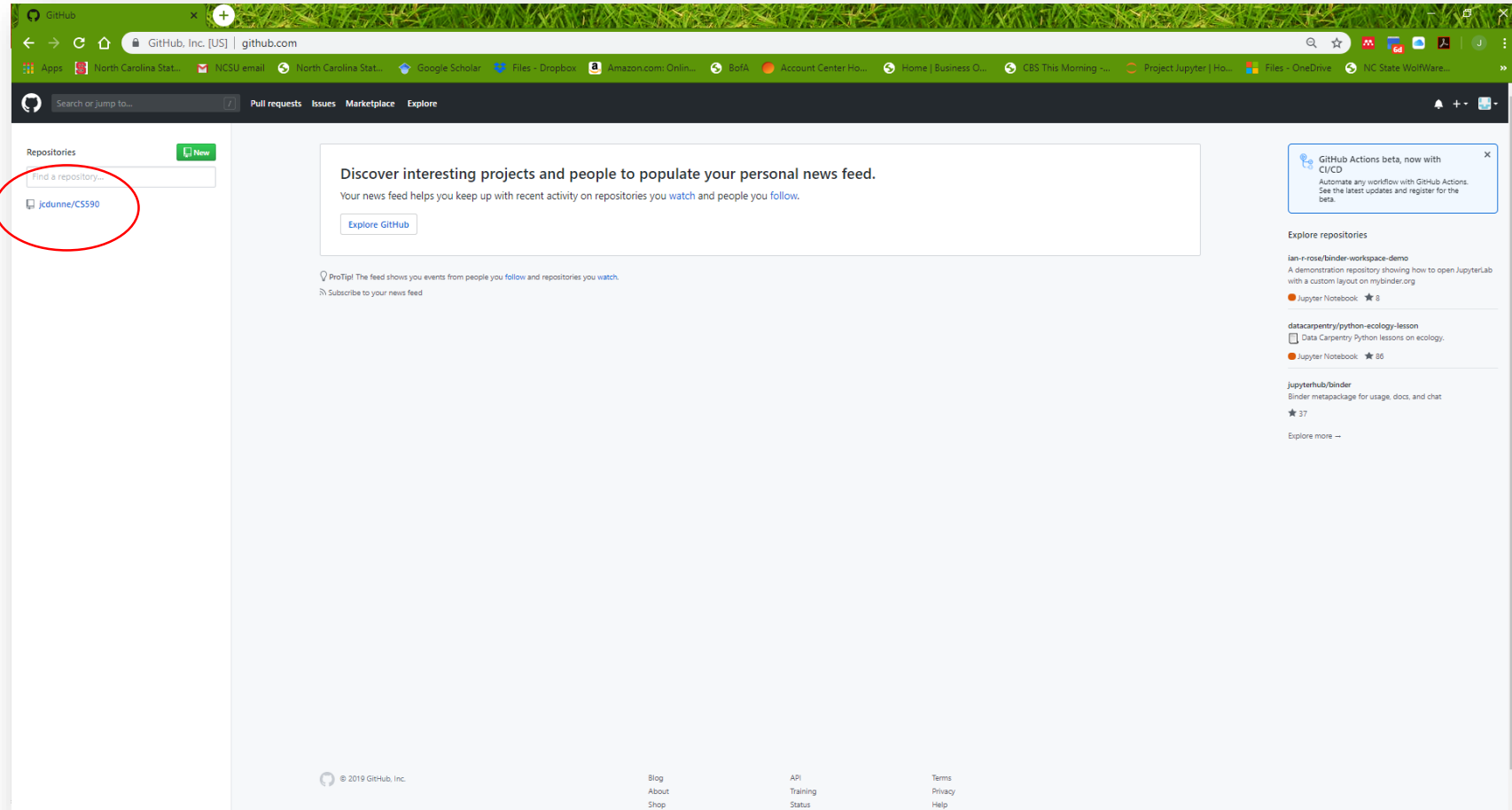
(base) C:\Users\jcdunne>
```

> exit

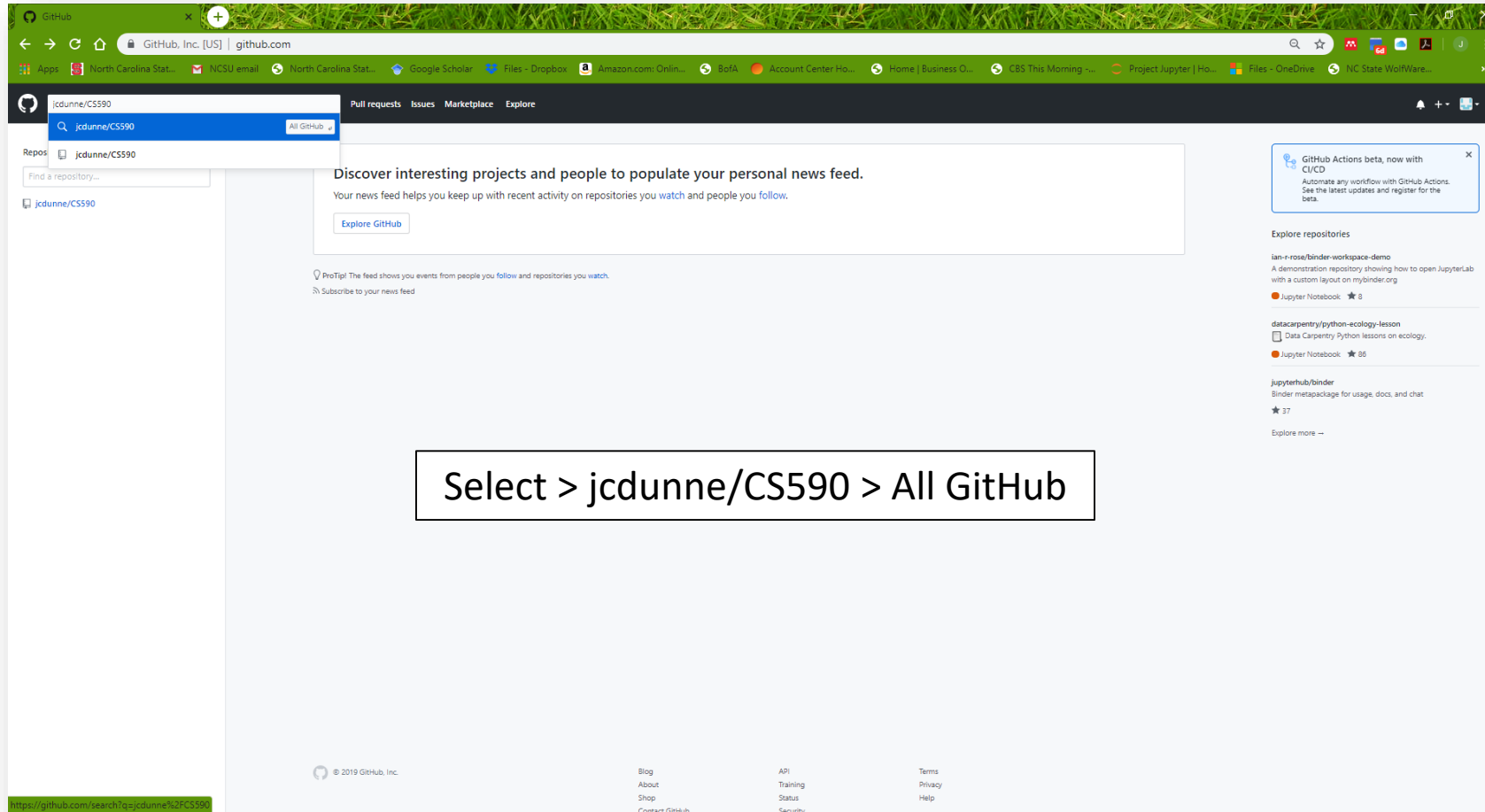
GitHub Exploration



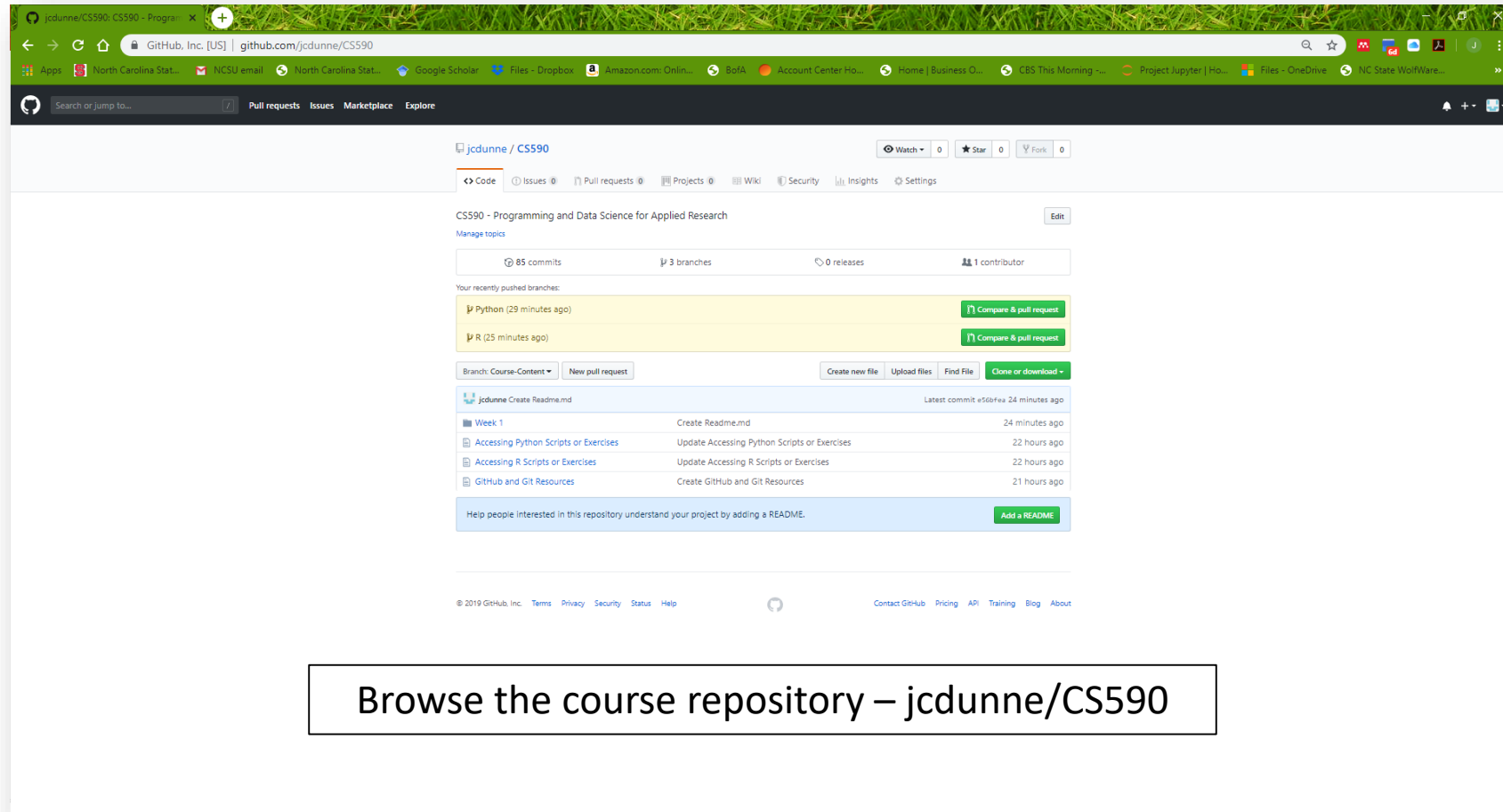
GitHub Repository – CS590



GitHub Repository Search – jcdunne/CS590



GitHub Course Repository – jcdunne/CS590



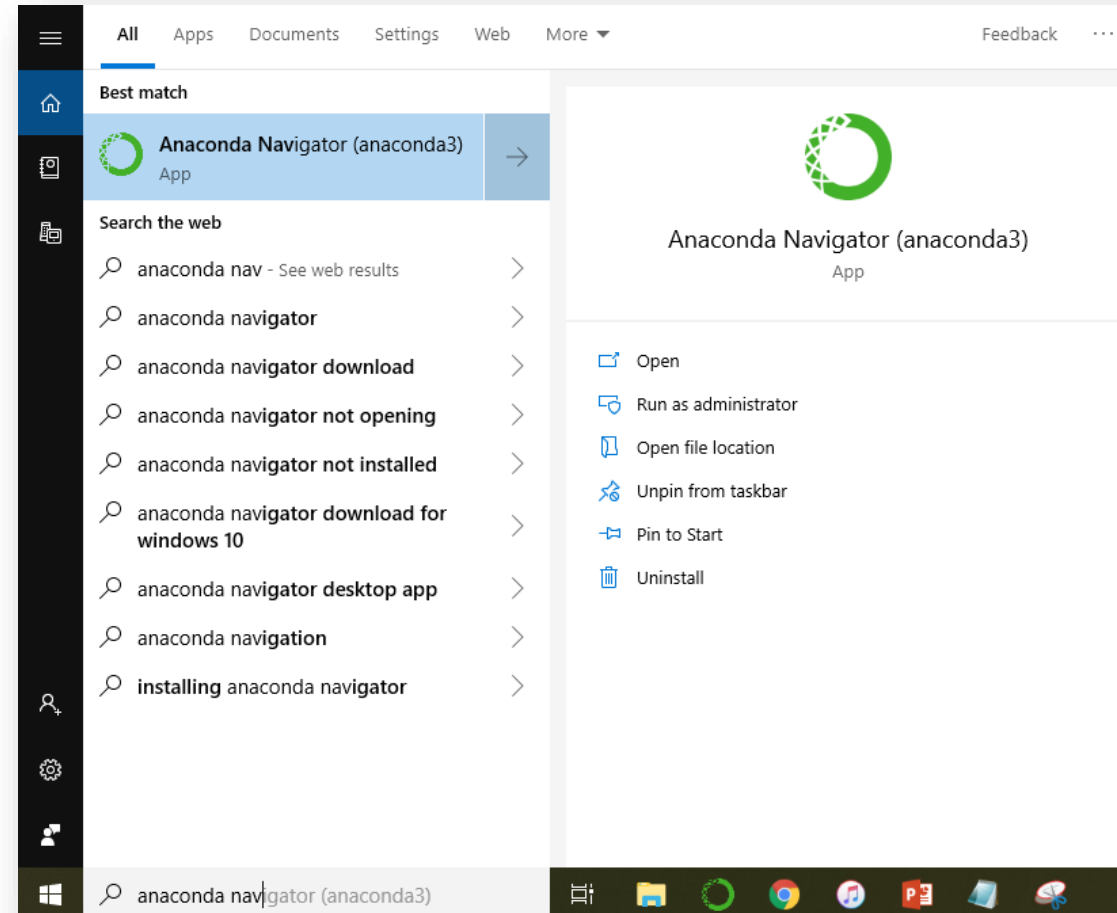
Jupyter NBViewer Exploration



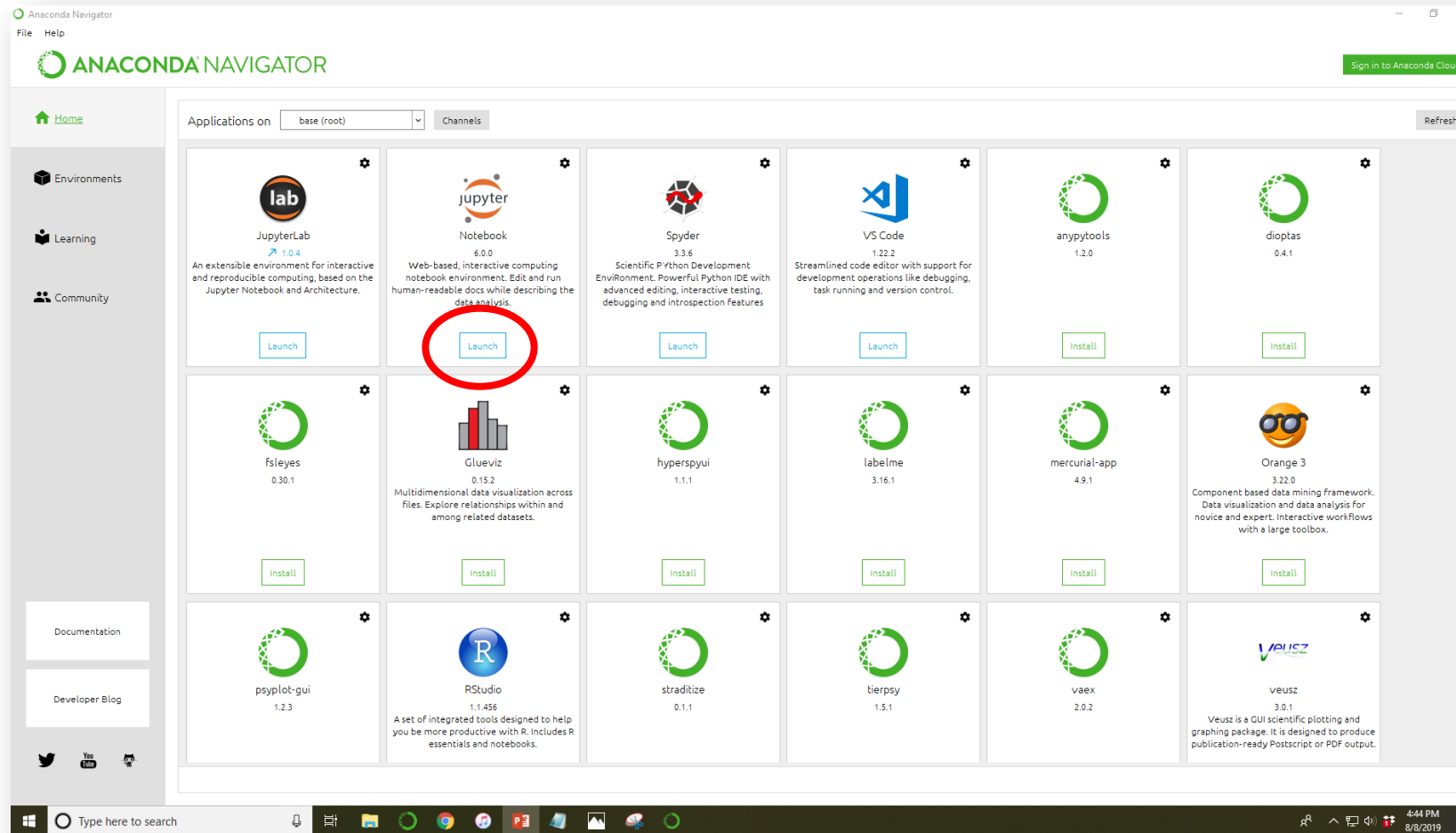
Jupyter NBViewer Search



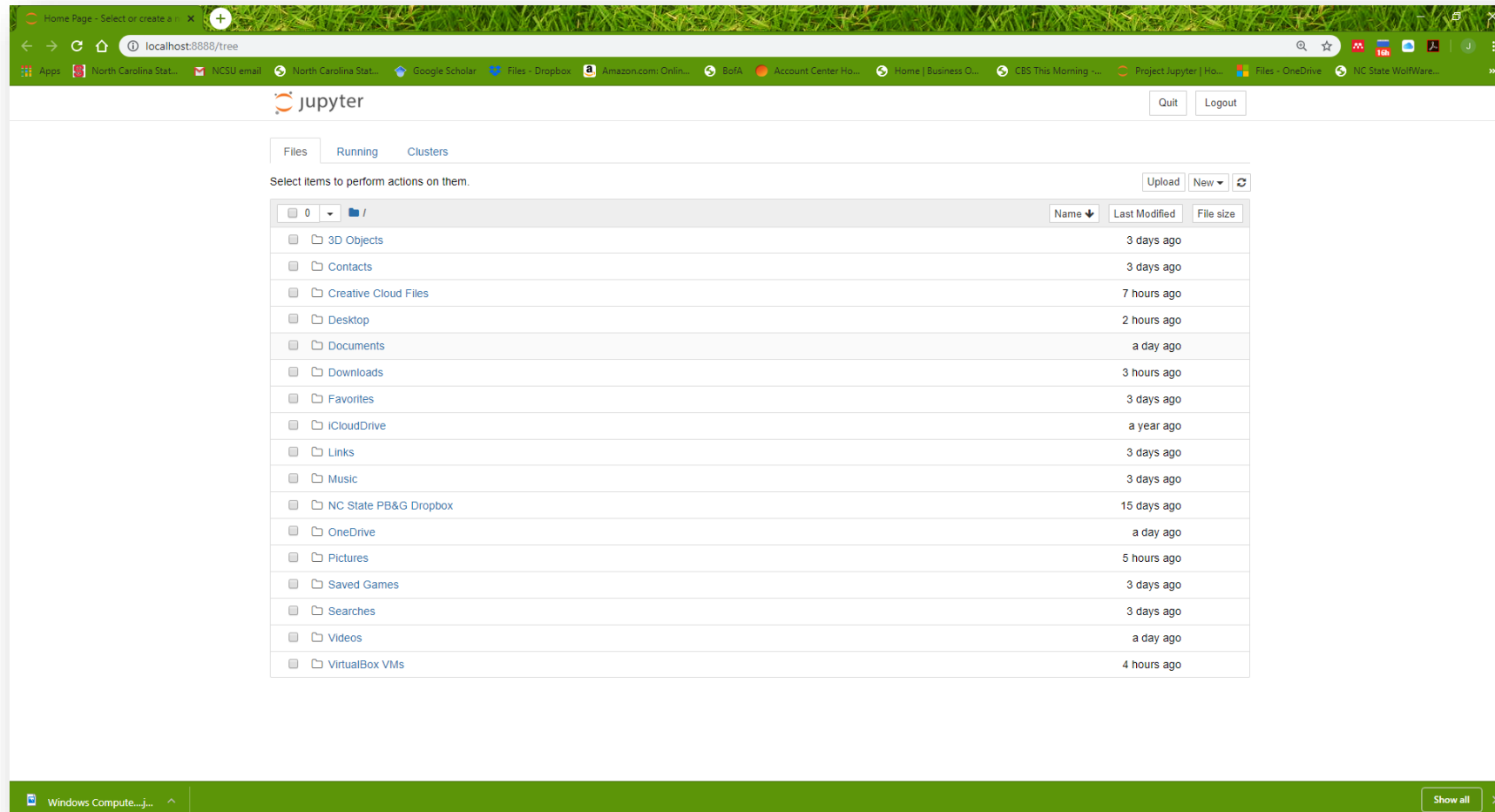
Launch Anaconda Navigator



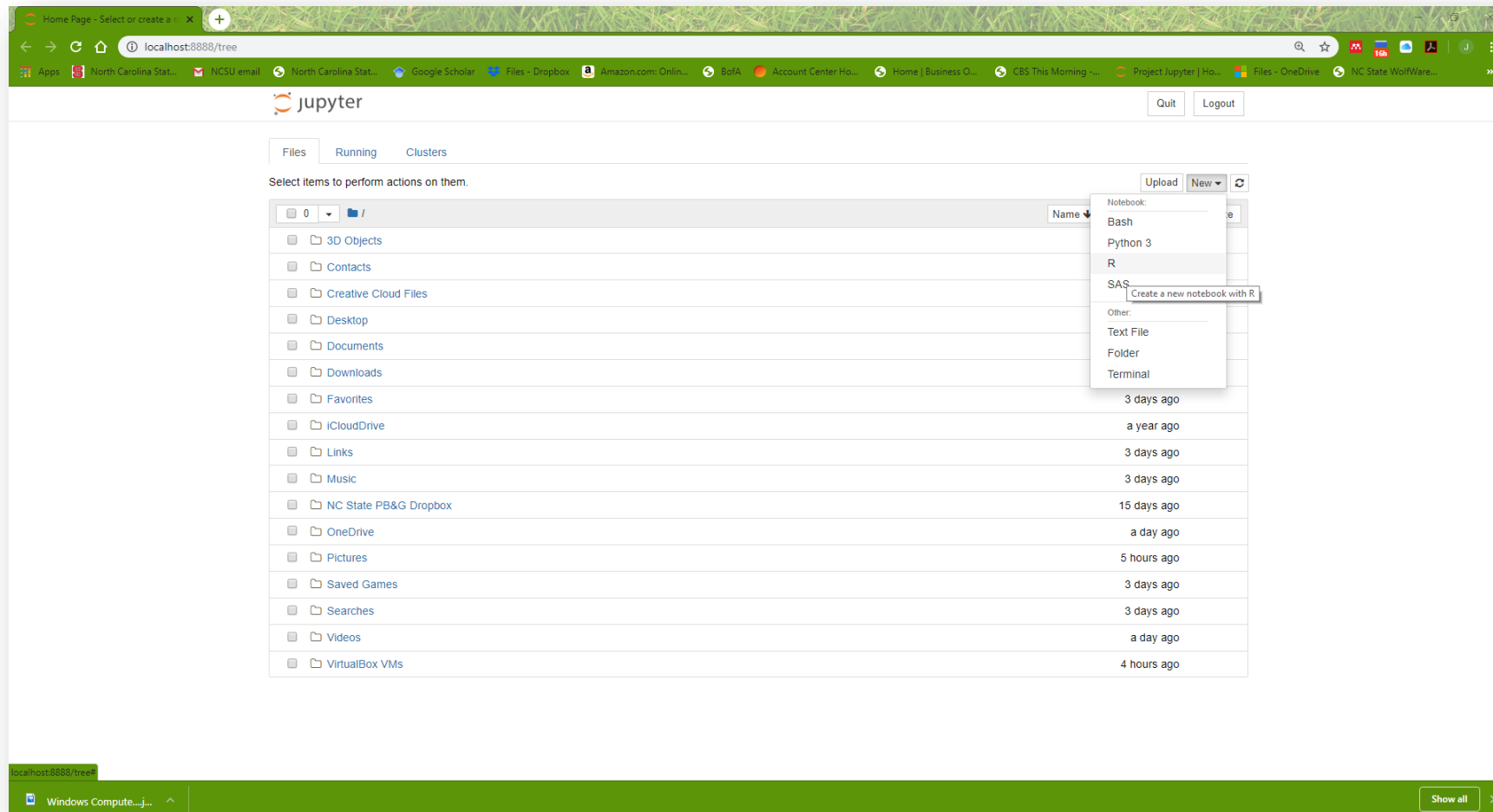
Launch Anaconda Navigator



Launch Anaconda Navigator



Launch Anaconda Navigator



Launch Anaconda Navigator

