

Laptop Prep for “Hands-on: Cluster Analysis with Python”

Overview

Laptop preparation for the class consists of four steps, with detailed instructions below:

1. Download course files from GitHub
2. Installation of Anaconda Python
3. Package downloads
4. Verify installation

NOTE – When using a work laptop, please keep the following in mind:

- Administrator permission may be required to complete laptop prep – contact your IT department if you’re unsure you have sufficient permissions.
- It is often necessary to disable anti-virus software to allow for the installation. As such, disabling any anti-virus is recommended before laptop prep.
- Corporate proxy servers and firewalls can block the installation. Be sure to consult your IT department as needed.
- Lastly, installing the latest version of Anaconda Python is recommended – even if you have Python already installed.

The GitHub repository with all required course files is located here:

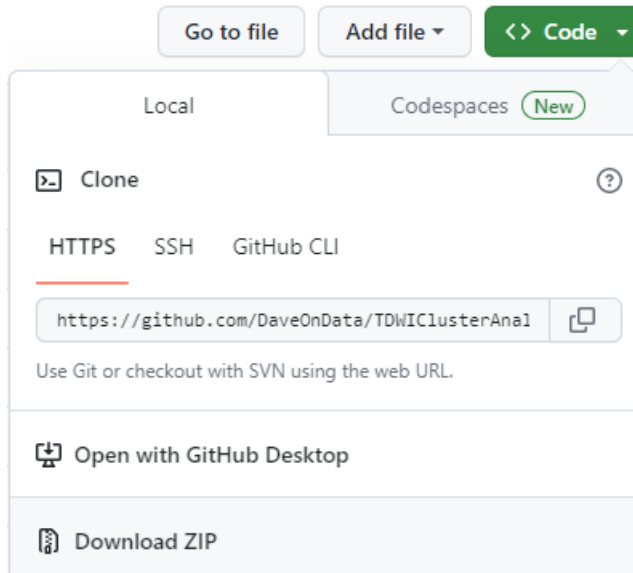
- <https://github.com/DaveOnData/TDWClusterAnalysisWithPython>

Hardware Requirements

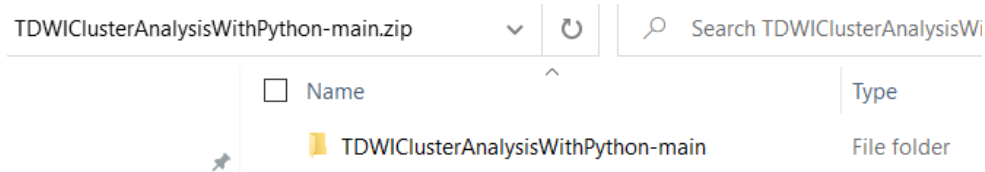
1. Windows or Mac OS X preferred (instructors have no experience with Linux)
2. 64-bit operating system
3. 8GB of RAM, 16GB preferred
4. 4GB of free drive space

Step 1 - Download the files from GitHub

1. Within the GitHub repository page, click on the “Code” button and select “Download ZIP”:



2. Copy the file folder within the downloaded ZIP to a well-known location on your laptop (e.g., the Desktop):



3. Open the file folder. You should see the following files:

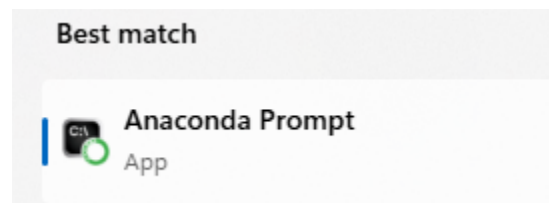
- Hands-On Lab 1 - K-Means.ipynb
- Hands-On Lab 2 - Optimizing K-Means.ipynb
- Hands-On Lab 3 - DBSCAN.ipynb
- Hands-On Lab 4 - PCA.ipynb
- Hands-On Lab 5 - Categorical Data.ipynb
- Heart.csv
- LaptopPrepClusterAnalysisWithPython.pdf
- README.md
- Verify Installation.ipynb

Step 2 – Anaconda Python Installation

1. Open your browser and navigate to <https://anaconda.com/download>.
2. You do not need to provide an email address. Click “Skip registration.”
3. Click the download the appropriate graphical installer for your laptop (e.g., the Windows installer).
4. When the installer has downloaded, start the installer and follow the instructions (accepting defaults) to complete the installation.

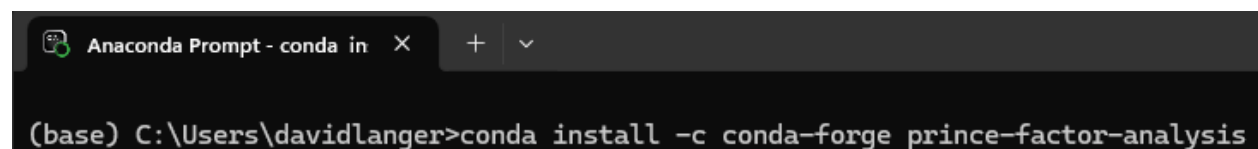
Step 3 – Package Downloads

1. NOTE – Some packages are used across multiple TDWI classes. If you are taking multiple classes as part of the same training (e.g., conference or bootcamp), you only need to install the packages once.
2. With Anaconda Python installed, start the Anaconda Prompt:



NOTE – If you’re using Mac, just use the Terminal app. All commands are the same.

3. At the command prompt type the following without quotes and hit <enter>:
 - a. “conda install -c conda-forge prince-factor-analysis”



NOTE – Installation can take quite a while to run. Please be patient.

4. If prompted, hit the “y” key and <enter> to proceed:

```
Anaconda Prompt - conda in X + v

added / updated specs:
- prince-factor-analysis

The following packages will be downloaded:

package | build
-----|-----
ca-certificates-2025.1.31 | h56e8100_0 155 KB conda-forge
certifi-2025.1.31 | pyhd8ed1ab_0 159 KB conda-forge
conda-25.1.1 | py312h2e8e312_1 1.1 MB conda-forge
openssl-3.4.1 | ha4e3fda_0 8.1 MB conda-forge
prince-factor-analysis-0.7.1 | pyhd8ed1ab_1 21 KB conda-forge
-----|-----
Total: 9.6 MB

The following NEW packages will be INSTALLED:

prince-factor-ana~ conda-forge/noarch::prince-factor-analysis-0.7.1-pyhd8ed1ab_1

The following packages will be UPDATED:

ca-certificates pkgs/main::ca-certificates-2024.12.31~ --> conda-forge::ca-certificates-2025.1.31-h56e8100_0
certifi 2024.12.14-pyhd8ed1ab_0 --> 2025.1.31-pyhd8ed1ab_0
conda 24.11.3-py312h2e8e312_0 --> 25.1.1-py312h2e8e312_1
openssl 3.4.0-ha4e3fda_1 --> 3.4.1-ha4e3fda_0

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

5. You should see something like the following:

```
Anaconda Prompt X + v

Downloading and Extracting Packages:

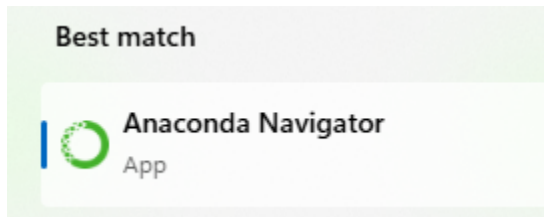
Preparing transaction: done
Verifying transaction: done
Executing transaction: done

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

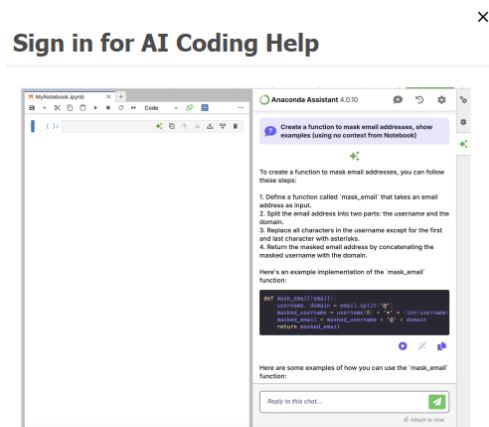
6. Repeat the above process by typing the following on the command line (without quotes) and hitting <enter>:
- “conda install -c conda-forge plotnine”

Step 4 – Verify Installation

1. With Anaconda Python installed, start the Anaconda Navigator application:



2. If you see the following dialog, close it (no need to sign in):



Access Anaconda Assistant in Jupyter Notebooks and JupyterLab.

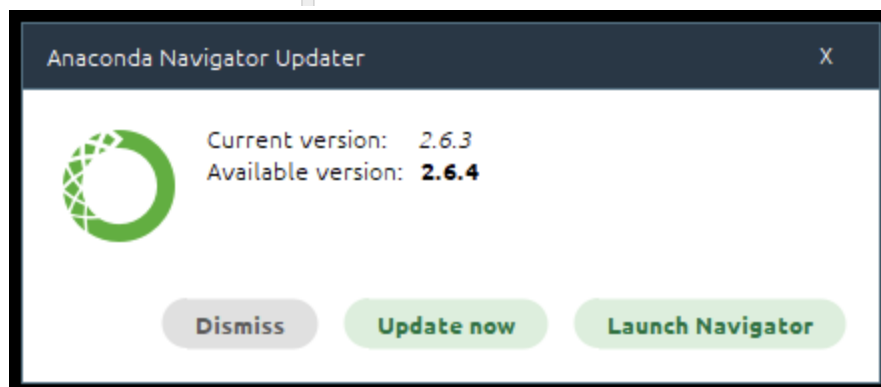
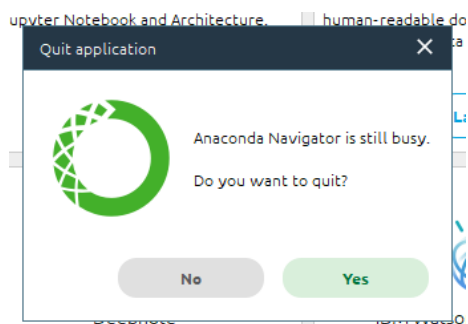
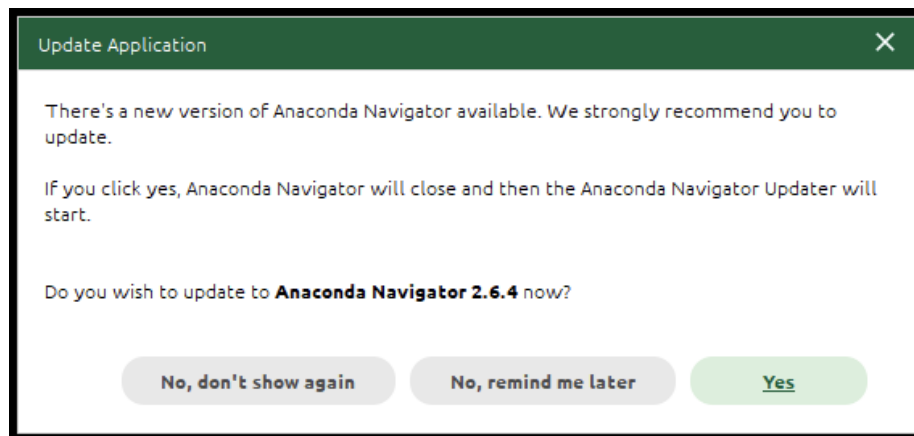
☐ Do not show again

[Connect to a Repository](#)

[Sign In](#)

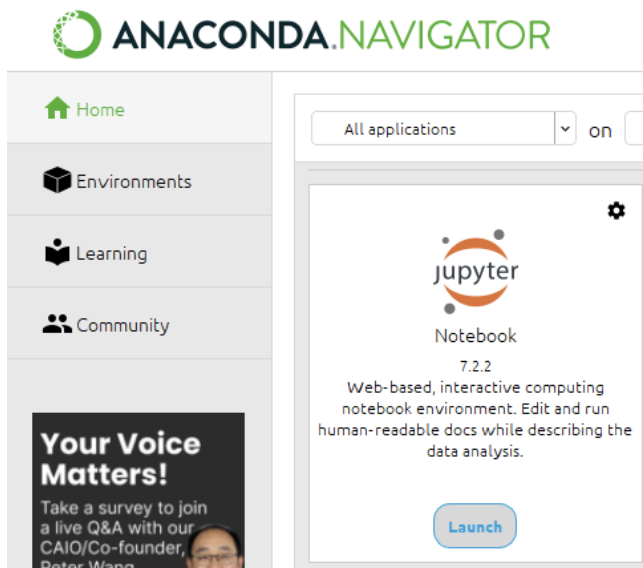
Don't have an account? [Sign Up](#)

3. You may be prompted to upgrade Anaconda Navigator. Follow the dialogs to do so:

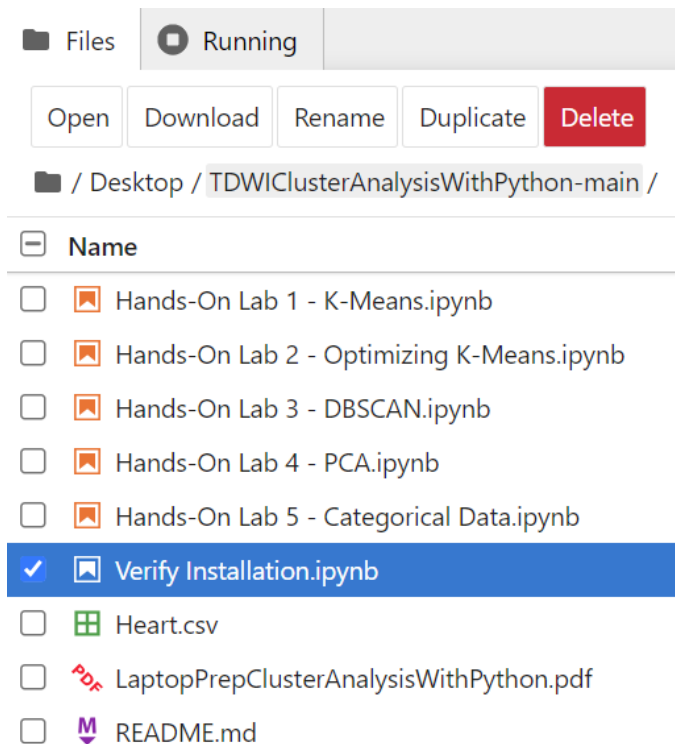


4. If needed, relaunch Anaconda Navigator.

5. NOTE – Your Anaconda Navigator window might not look exactly like the following. Within Anaconda Navigator, launch Jupyter Notebook:



6. Within the Jupyter browser, navigate to where you copied the course file folder:



7. Click on the “Verify Installation.ipynb” entry.

8. Run all the cells in the notebook:

Verify Installation Last Checkpoint: 2 minutes ago

Run	Kernel	Settings	Help
Run Selected Cell			Shift+Enter
Run Selected Cell and Insert Below			Alt+Enter
Run Selected Cell and Do not Advance			Ctrl+Enter
Run Selected Text or Current Line in Console			
Run All Above Selected Cell			
Run Selected Cell and All Below			
Render All Markdown Cells			
Run All Cells			

9. Your output should look like the following, with no errors:

Verify Installation

Run the following code cell you should see no errors as a result of the running the code.

```
[1]: from prince import FAMD
     from plotnine import ggplot
```

Congratulations! You are now ready for the class!