Lab 1 – Install & Setup your Environment

Brendan Tierney

TU 257 – Data Analytics

Follow along – Don’t rush ahead

Follow along with the steps I get for each Lab

Don’t rush ahead

Don’t ask Question about something we are covering later

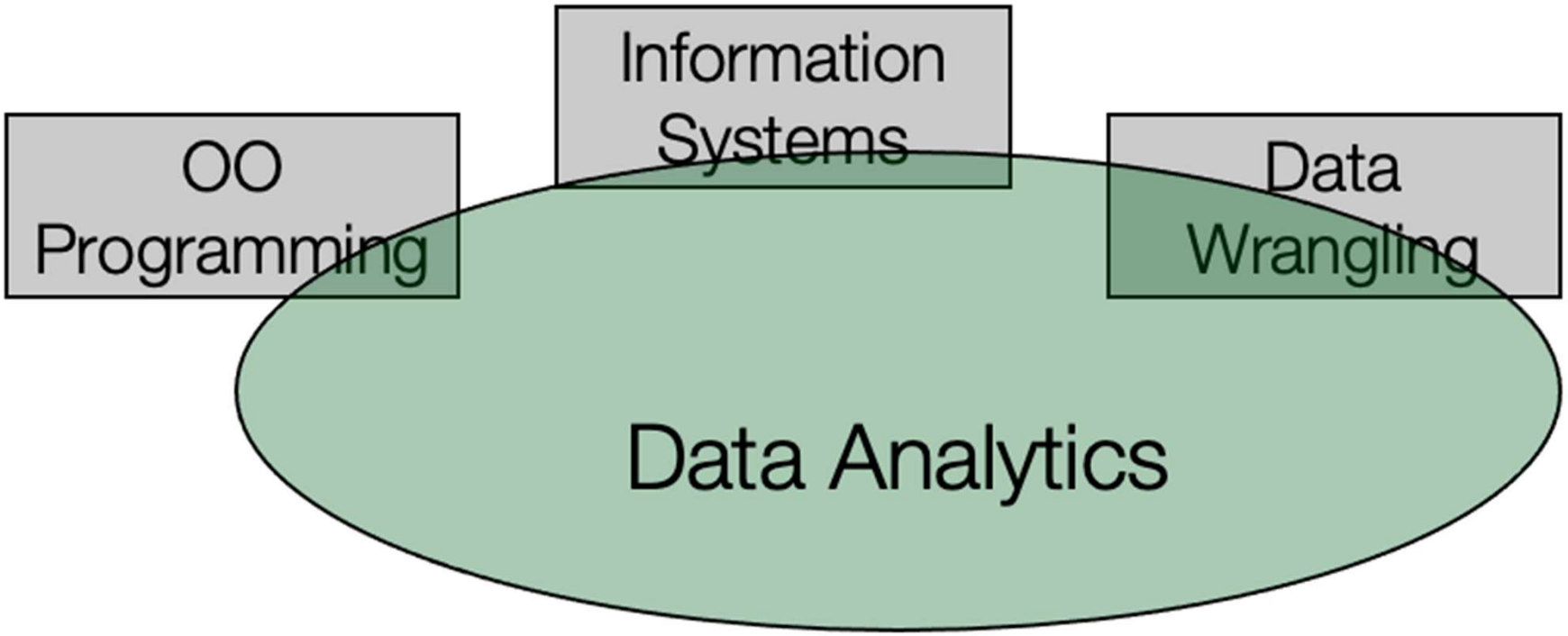
Wait

Labs are designed to be followed in a certain order

Follow that order

Don’t skip steps - if you do, things will go wrong

If something doesn’t work, look back on what you did and see if you missed a step(s)


Information Systems
OO Programming
Data Wrangling
Data Analytics

Agenda

Install Python

Install IDE

- Jupyter Notebook
- Jupyter Lab
- Google Colabs

- Explore these and Pick One
- If you are unsure, install Jupyter Notebook

Run some simple code


python™

+


jupyter

ou might have Python
eady installed – Great
ou have nothing to do!

nstalled it last Semester.
Great, but can you
emember how to use it?

I need to install it, How
do that?

Install Python yourself,
and all the libraries
yourself

Or use Anaconda, wh
it installs lots of relat
libraries

Or use a website with
everything done for yo

You will need to do the following for

Data Wangling module


STOP
LOOK
LISTEN

Anaconda

Anaconda provides a prebuilt environment

Download

- https://www.anaconda.com/

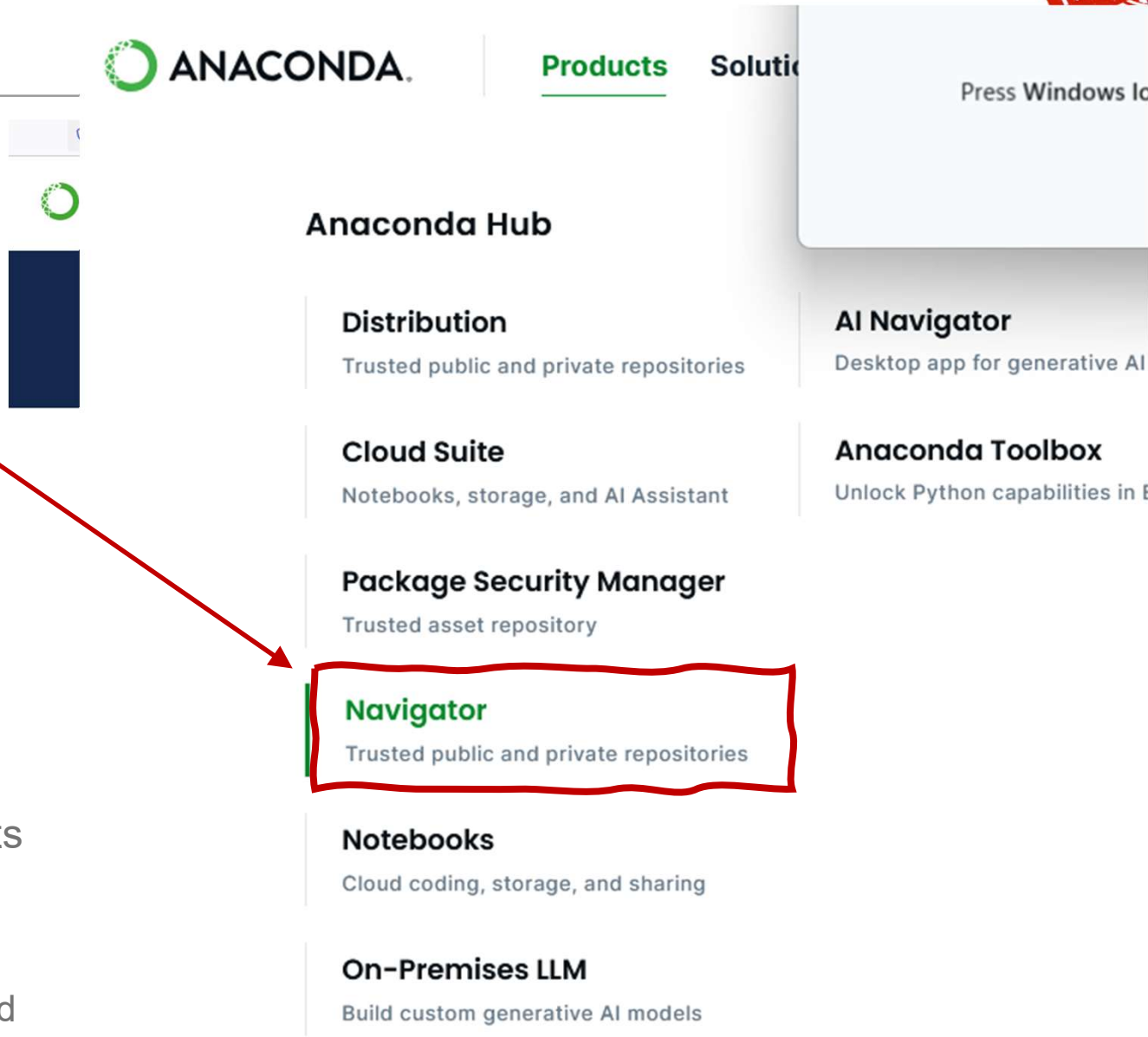
Free to use

Will include Python and lots of libraries for doing Data Science, Machine Learning, etc work

Will have correct version of libraries to run correctly with each other and version of Python

Can create multiple Anaconda environments on your Computer, with each one for a different project

- Allows separation of work which might need different version of libraries etc



Anaconda

Anaconda provides a prebuilt environment

Download

- https://www.anaconda.com/

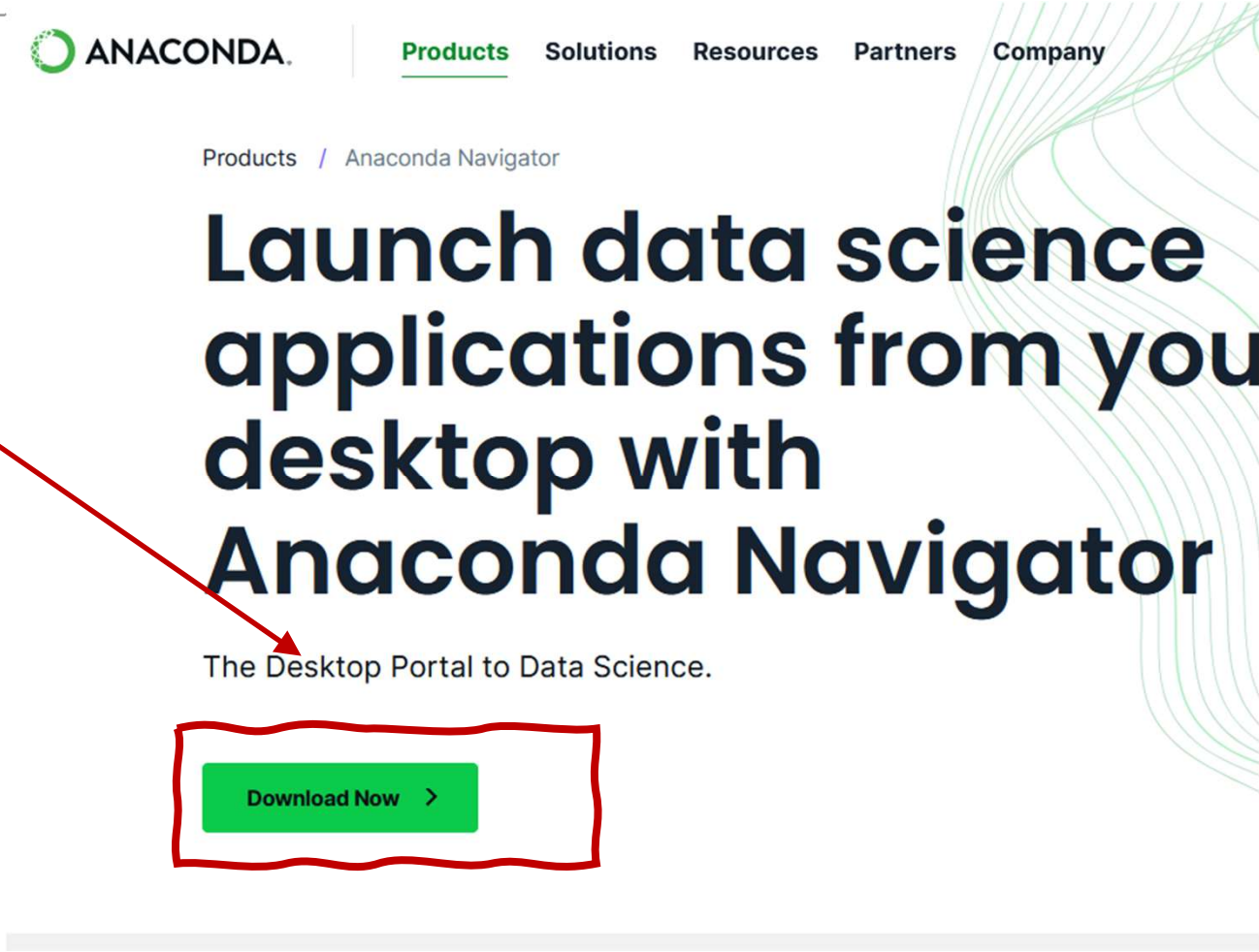
Free to use

Will include Python and lots of libraries for doing Data Science, Machine Learning, etc work

Will have correct version of libraries to run correctly with each other and version of Python

Can create multiple Anaconda environments on your Computer, with each one for a different project

- Allows separation of work which might need different version of libraries etc



Anaconda

Anaconda provides a prebuilt environment

Download

- https://www.anaconda.com/

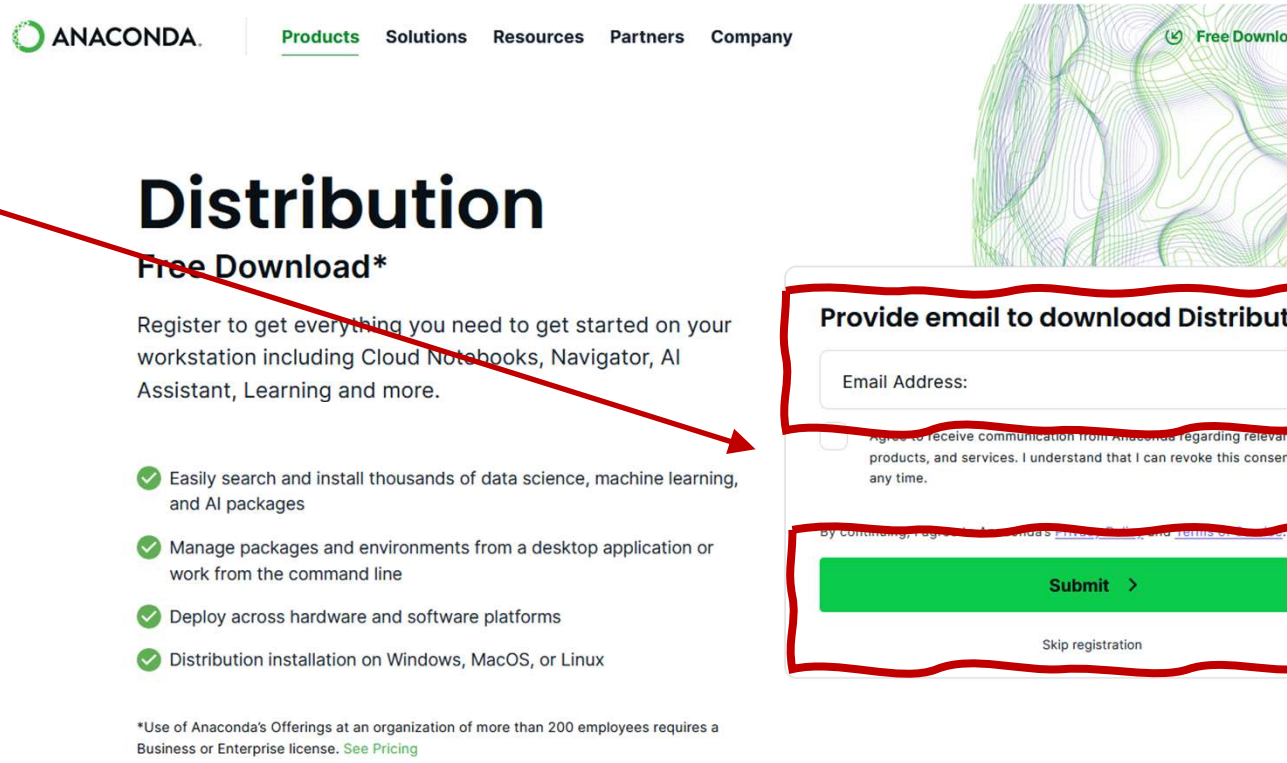
Free to use

Will include Python and lots of libraries for doing Data Science, Machine Learning, etc work

Will have correct version of libraries to run correctly with each other and version of Python

Can create multiple Anaconda environments on your Computer, with each one for a different project

- Allows separation of work which might need different version of libraries etc



Anaconda

Anaconda provides a prebuilt environment

Download

- https://www.anaconda.com/

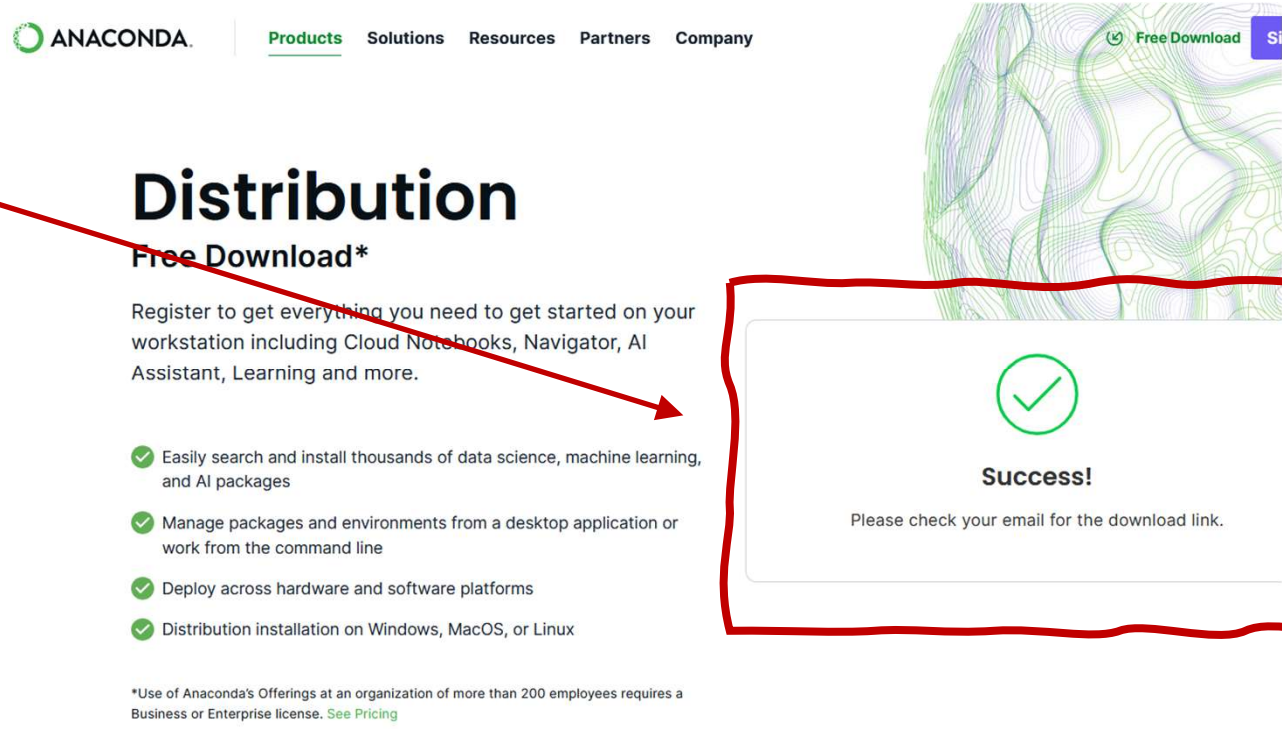
Free to use

Will include Python and lots of libraries for doing Data Science, Machine Learning, etc work

Will have correct version of libraries to run correctly with each other and version of Python

Can create multiple Anaconda environments on your Computer, with each one for a different project

- Allows separation of work which might need different version of libraries etc



Download Now

For installation assistance, refer to Troubleshooting.

Download Anaconda Distribution or Miniconda by choosing the proper installer for your machine. Learn the difference from our Documentation.

Anaconda Installers

Download

Windows

Python 3.12

Mac

Python 3.12

Linux

Python 3.12

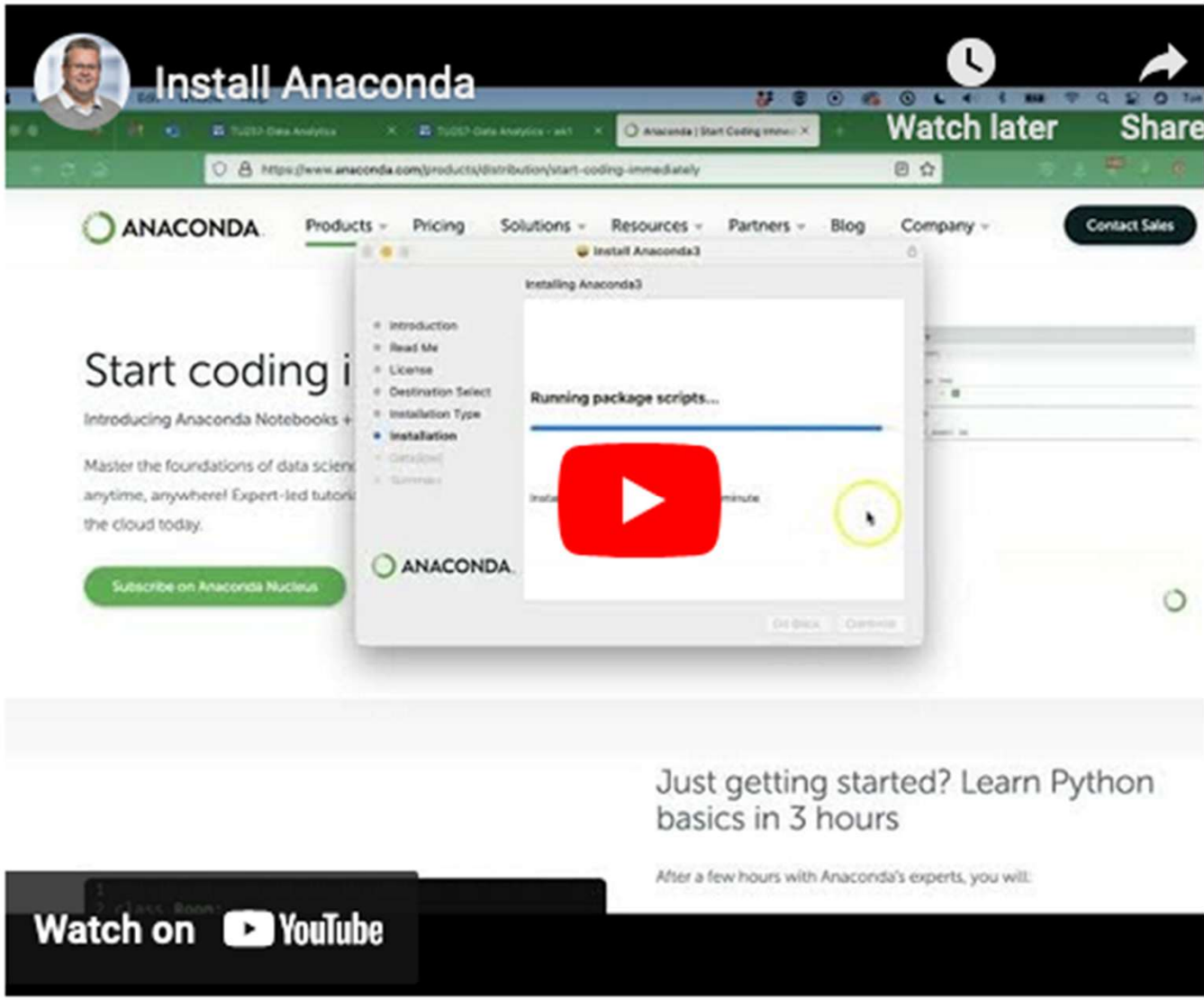
Anaconda – Watch the video

This video shows you how to install Anaconda

Video on Module webpage week 1

Can take from 4-10 minutes to install

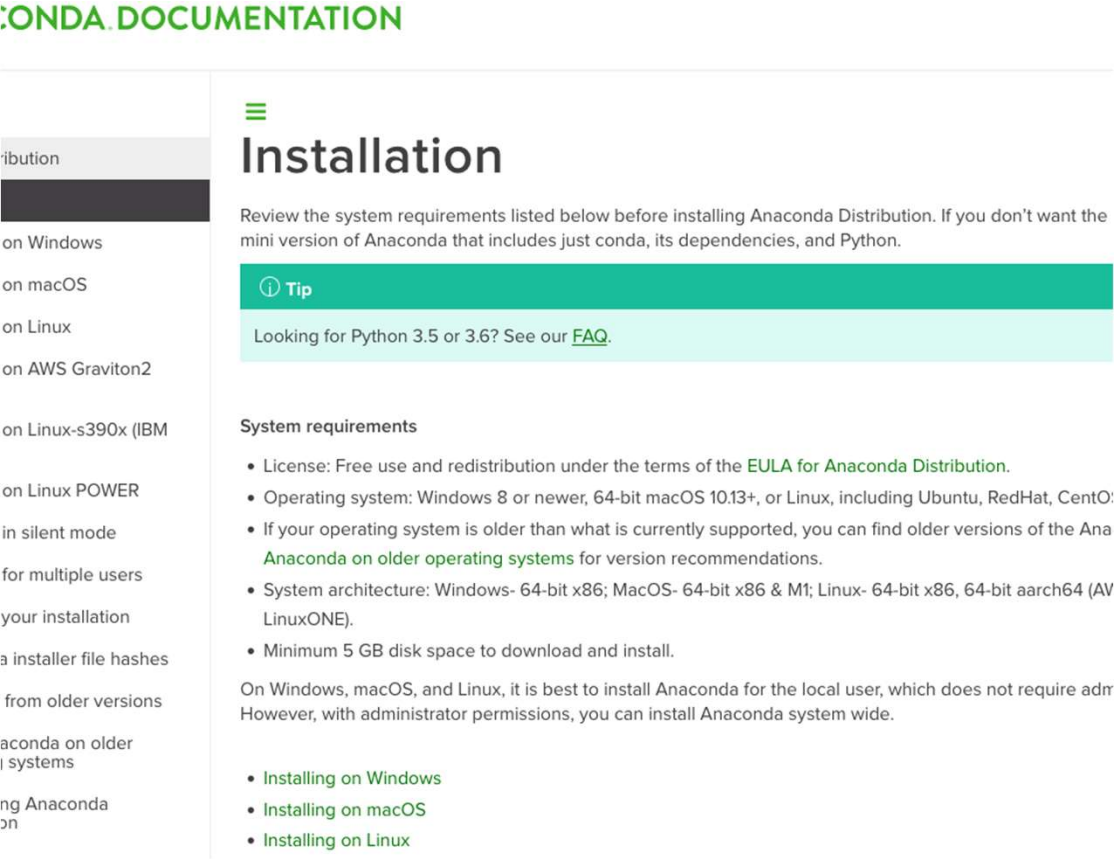
- Sometimes it can take 40+ minutes to install !!!



Anaconda

Check out the Installation Instructions for Step-by-Step guide

https://docs.anaconda.com/anaconda/install/index.html



- Windows Installation Instructions

https://docs.anaconda.com/anaconda/install/windo

- Mac Installation Instructions

https://docs.anaconda.com/anaconda/install/mac-c

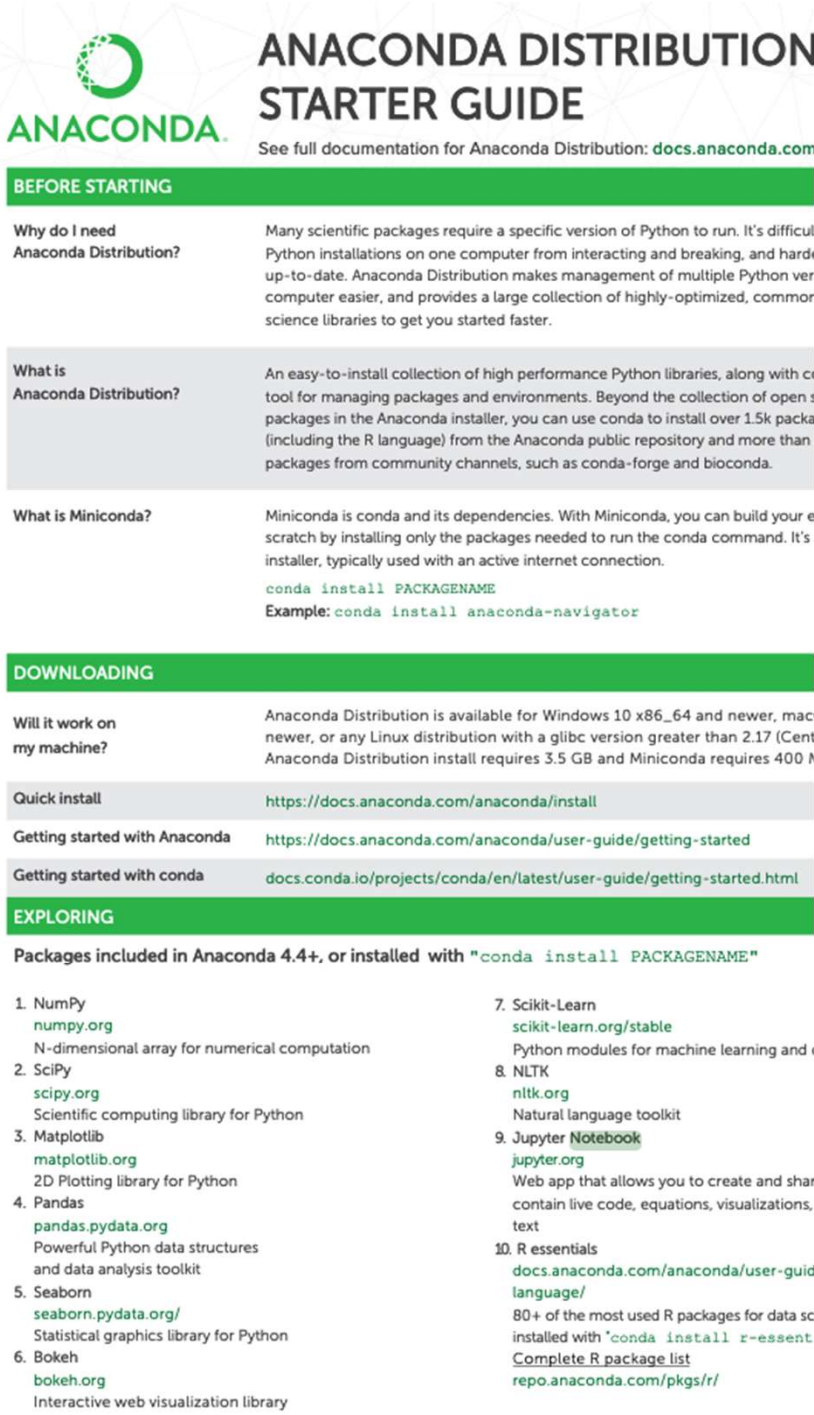
- Linux Installation Instructions

https://docs.anaconda.com/anaconda/install/linux/

Anaconda

Anaconda Starter Guide

https://docs.anaconda.com/_downloads/3613d324acc0a4b3c203fd79c71a2b45/Anaconda-Starter-Guide.pdf



ANACONDA

ANACONDA DISTRIBUTION STARTER GUIDE

See full documentation for Anaconda Distribution: docs.anaconda.com

BEFORE STARTING

Why do I need Anaconda Distribution?

Many scientific packages require a specific version of Python to run. It's difficul Python installations on one computer from interacting and breaking, and hard up-to-date. Anaconda Distribution makes management of multiple Python ver computer easier, and provides a large collection of highly-optimized, commor science libraries to get you started faster.

What is Anaconda Distribution?

An easy-to-install collection of high performance Python libraries, along with c tool for managing packages and environments. Beyond the collection of open packages in the Anaconda installer, you can use conda to install over 1.5k packa (including the R language) from the Anaconda public repository and more than packages from community channels, such as conda-forge and bioconda.

What is Miniconda?

Miniconda is conda and its dependencies. With Miniconda, you can build your e scratch by installing only the packages needed to run the conda command. It's installer, typically used with an active internet connection.

`conda install PACKAGENAME`

Example: `conda install anaconda-navigator`

DOWNLOADING

Will it work on my machine?

Anaconda Distribution is available for Windows 10 x86_64 and newer, mac newer, or any Linux distribution with a glibc version greater than 2.17 (Cent Anaconda Distribution install requires 3.5 GB and Miniconda requires 400 M

Quick install	https://docs.anaconda.com/anaconda/install
Getting started with Anaconda	https://docs.anaconda.com/anaconda/user-guide/getting-started
Getting started with conda	docs.conda.io/projects/conda/en/latest/user-guide/getting-started.html

EXPLORING

Packages included in Anaconda 4.4+, or installed with "`conda install PACKAGENAME`"

1. NumPy
   numpy.org
   N-dimensional array for numerical computation
2. SciPy
   scipy.org
   Scientific computing library for Python
3. Matplotlib
   matplotlib.org
   2D Plotting library for Python
4. Pandas
   pandas.pydata.org
   Powerful Python data structures and data analysis toolkit
5. Seaborn
   seaborn.pydata.org/
   Statistical graphics library for Python
6. Bokeh
   bokeh.org
   Interactive web visualization library
7. Scikit-Learn
   scikit-learn.org/stable
   Python modules for machine learning and
8. NLTK
   nltk.org
   Natural language toolkit
9. Jupyter Notebook
   jupyter.org
   Web app that allows you to create and shar contain live code, equations, visualizations, text
10. R essentials
    docs.anaconda.com/anaconda/user-guid language/
    80+ of the most used R packages for data sc installed with "`conda install r-essent`
    Complete R package list
    repo.anaconda.com/pkgs/r/


ALTERNATIVE

nstall Python

This Module assumes you have a Basic Knowledge

- From your own experience
- From another module
- From your own learning
- Or you can lean some bits yourself

Lots of resources available

- See Note, website, Python website, Google!

RTFM

- Online Python Documentation https://docs.python.org/3/

nstall Python

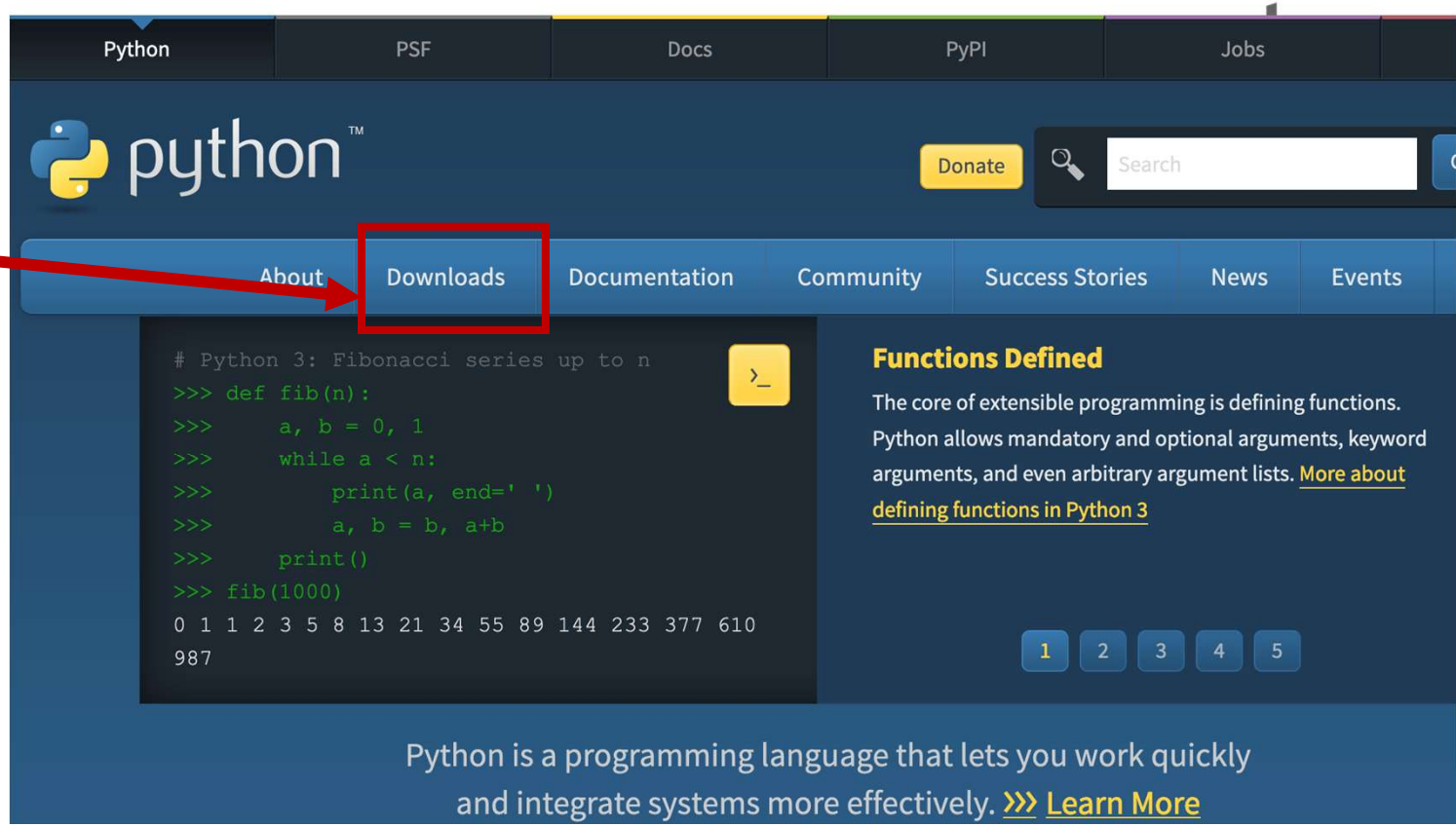
Go install Python

- You might have this done already
- Or already installed from another module

Go to https://www.python.org/

Install latest version

Follow Installation Instructions



Python Useful Links

Python website. https://www.python.org/

Python Online Documentation. https://docs.python.org/3/

PyPi – Python Package Index. https://pypi.org/

Google's Python Tutorials. https://developers.google.com/edu/python/

Lots and lots of resources available.

Data Science Notebooks

Lots of other similar products/notebooks

https://datasciencenotebook.org

Jupyter Notebook
– on Anaconda

Starter Guide

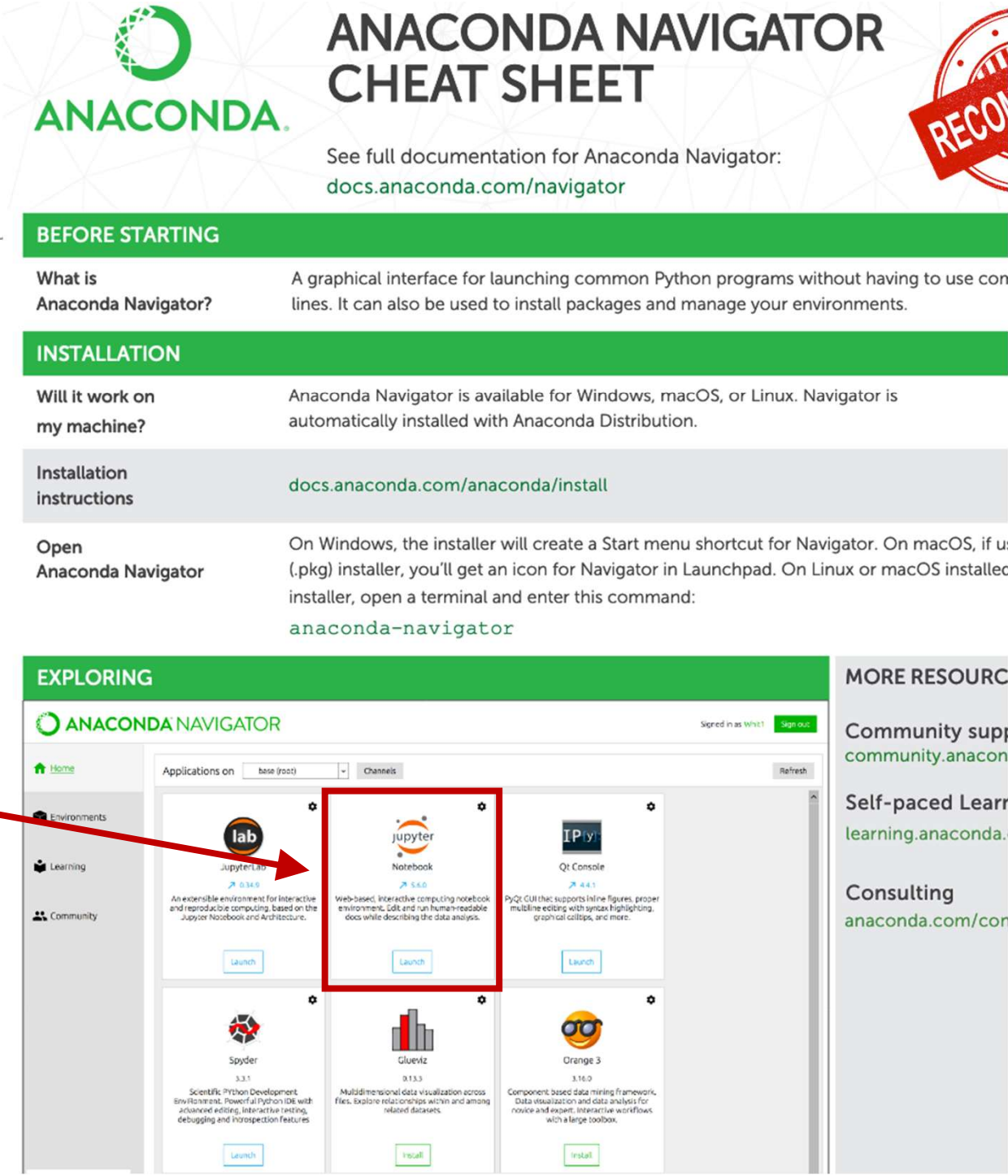
https://docs.anaconda.com/_downloads/3613d324acc0a4b3c203fd79c71a2b45/Anaconda-Starter-Guide.pdf

Jupyter Notebook

r just type the following on command line

```
jupyter notebook
```

ee next few slides



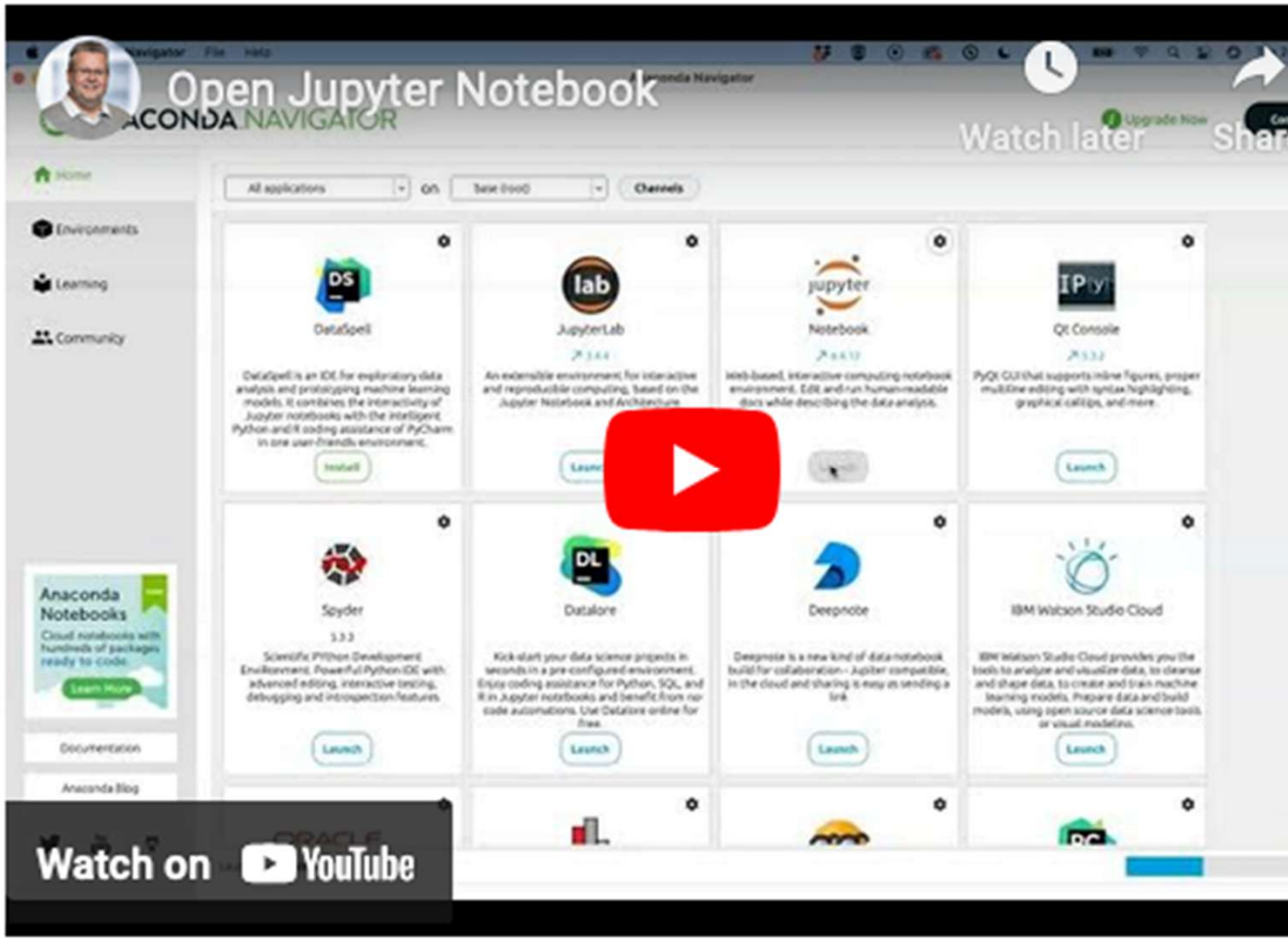
Jupyter Notebook – on Anaconda – Watch the Video

Video available on module webpage

Steps through how to open Jupyter Notebook in Anaconda Navigator

Also shows how to open Lab Exercise Notebook

- Notebook can be downloaded from webpage




ALTERNATIVE

nstall Jupyter Notebook

Very Popula

I'll be using this
Demo Code

Home page : https://jupyter.org/

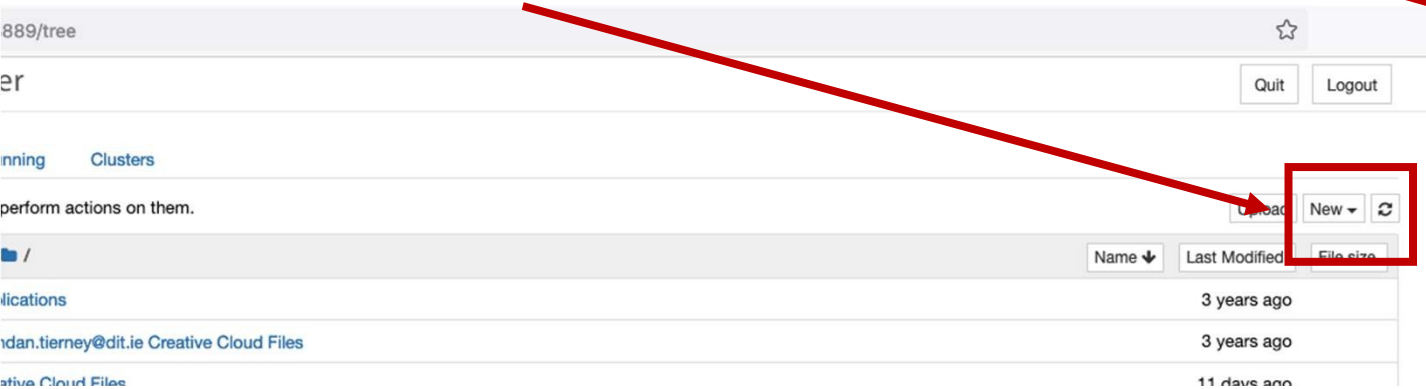
Run the install command in a terminal window

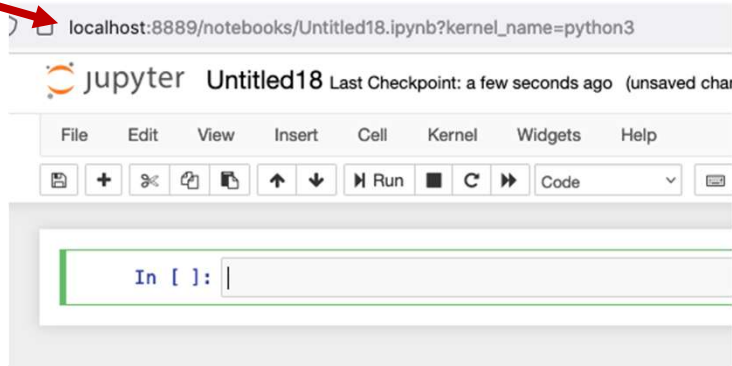
- `pip3 install notebook`

To run Jupyter Notebook, run the following in a terminal window

- `jupyter notebook`

- A webpage will open in your default browser
- Under new, select Python. A Jupyter Notebook will Open



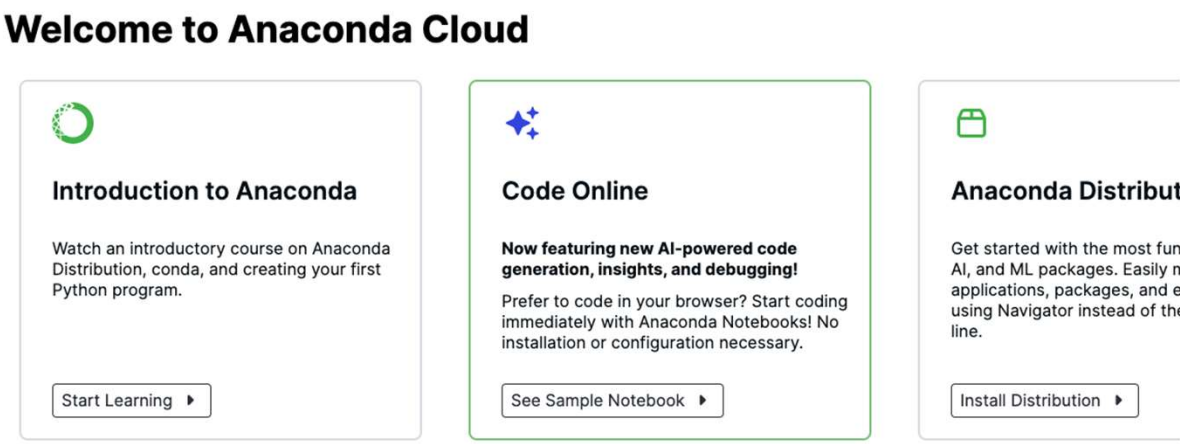


ATIVE

Anaconda Cloud - Notebook

Jupyter Notebo
on the Cloud

Home page : https://anaconda.cloud/

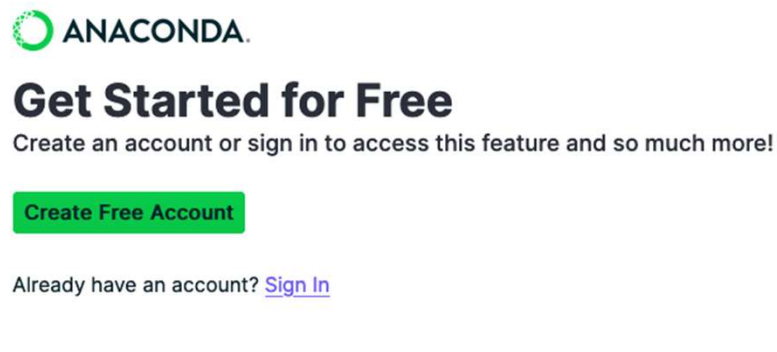


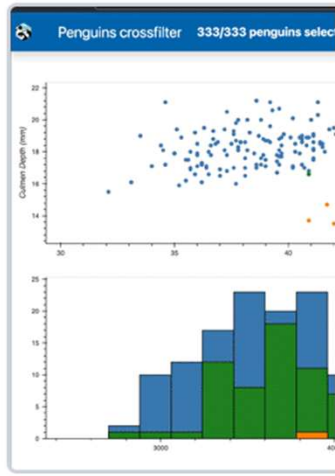
Create an Account

It's FREE to use

- With some limits – It's FREE
- Storage, Compute resources, etc restrictions

Lots of online resources, examples, etc.





It's hosted on the Cloud – Need to be online to use it!

- No internet -> No access to it.

ATIVE

nstall Jupyter Lab

Newer Version of Jupyter Notebook

I might use this from time to

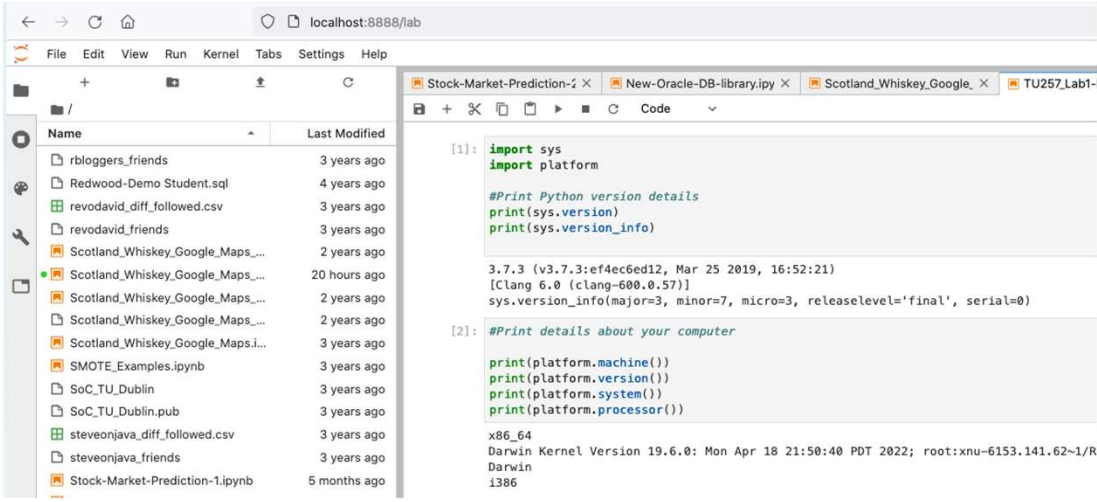
Home page : https://jupyter.org/

Run the install command in a terminal window

- `pip3 install jupyterlab`

To run Jupyter Notebook, run the following in a terminal window

- `jupyter-lab`
  - `or  jupyter lab`



- Open Notebook in similar way to Jupyter Notebook
- For now, Jupyter Notebook as slightly more features. Lab will catch up at some point/

Google Colab

Cloud hosted Notebook environment by Google

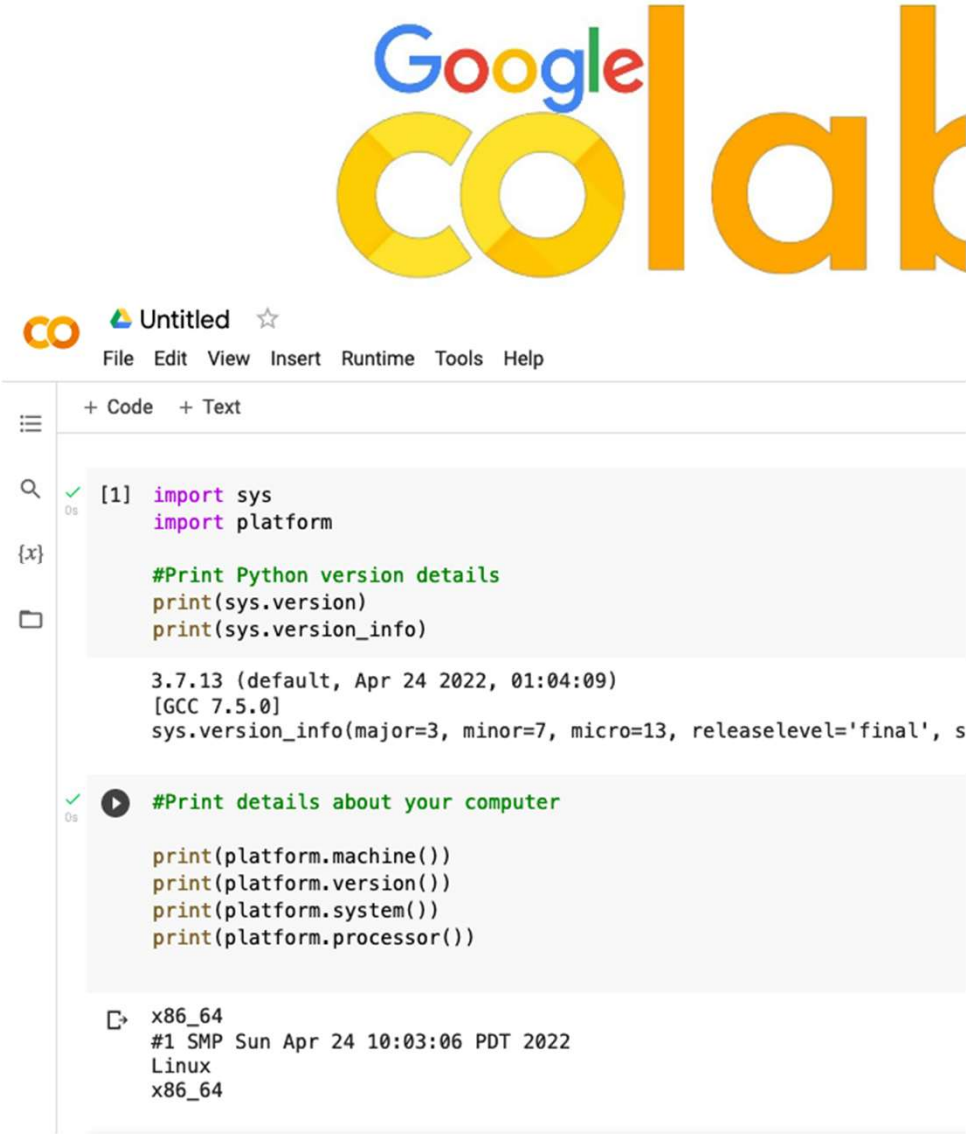
No local install of Python, Jupyter, etc

Need a Google account.

- Notebooks are saved to your Google Drive or Github.

Free to use, but has limited computing resources

- Although you can use limited GPUs on this for free
- Pay for more computing power. (This isn't needed for this course



If unsure, install Jupyter Notebook

mportant

s module will not look at or go into detail
out

- Python basics
- Syntax
- Structure of Python
- Different types of objects
- Writing functions

hese things are/were covered elsewhere

r you can pick these up quickly

r can find these things elsewhere

We will concentrate on the different Libraries
Functions available in Python

Will focus on using these Functions to analys
data

Live programming, in this module, is not real
possible

- Error prone
- Time consuming
- Boring to watch
- Wasted time

To avoid these issues Prepared Notebooks a
made available

- Focuses on the important Tasks and Step
- Is Reproducible => VERY Important for Ar
- Templates for you to use


How does it work?

nload the Notebook (see webpage). Then Run it. Write code for questions.

```
t sys
t platform


t Python version details
(sys.version)
(sys.version_info)


t details about your computer
(platform.machine())
(platform.version())
(platform.system())
(platform.processor())
```

[1]:
```
import sys
import platform

#Print Python version details
print(sys.version)
print(sys.version_info)
```

```
3.7.3 (v3.7.3:ef4ec6ed12, Mar 25 2019, 16:52:21)
[Clang 6.0 (clang-600.0.57)]
sys.version_info(major=3, minor=7, micro=3, releaselevel='final', serial=0)
```

[2]:
```
#Print details about your computer

print(platform.machine())
print(platform.version())
print(platform.system())
print(platform.processor())
```

```
x86_64
Darwin Kernel Version 19.6.0: Mon Apr 18 21:50:40 PDT 2022; root:xnu-6153.141.62~1/RELEA
Darwin
i386
```

webpage for notebook. You can download and Open this notebook on your machine – See web

r Coding Exercises : The notebook will ask you to write some simple Python code

Lab Exercise

Download the Notebook from webpage

Open it in Jupyter Notebook

Run the sample code

Can you write some simple code?

Can you print your name

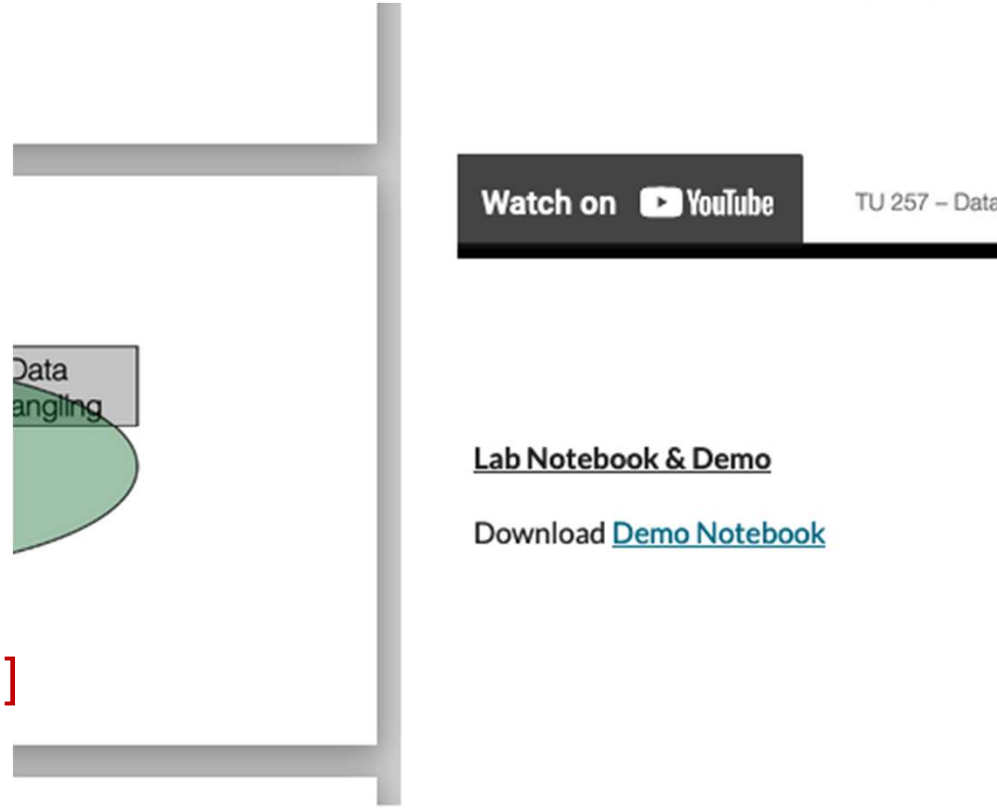
Can you create a variable and print it

… [what else can you do? From what you learned in Semester 1]

Save the notebook

That's it

- You do not need to upload or submit it anywhere



Complete all Exercises
before
Next Week

- Install Anaconda/Python
- Decide what tool/notebook you will use
- Install tool/notebook
- Do a quick check to make sure it is working (see previous slides)