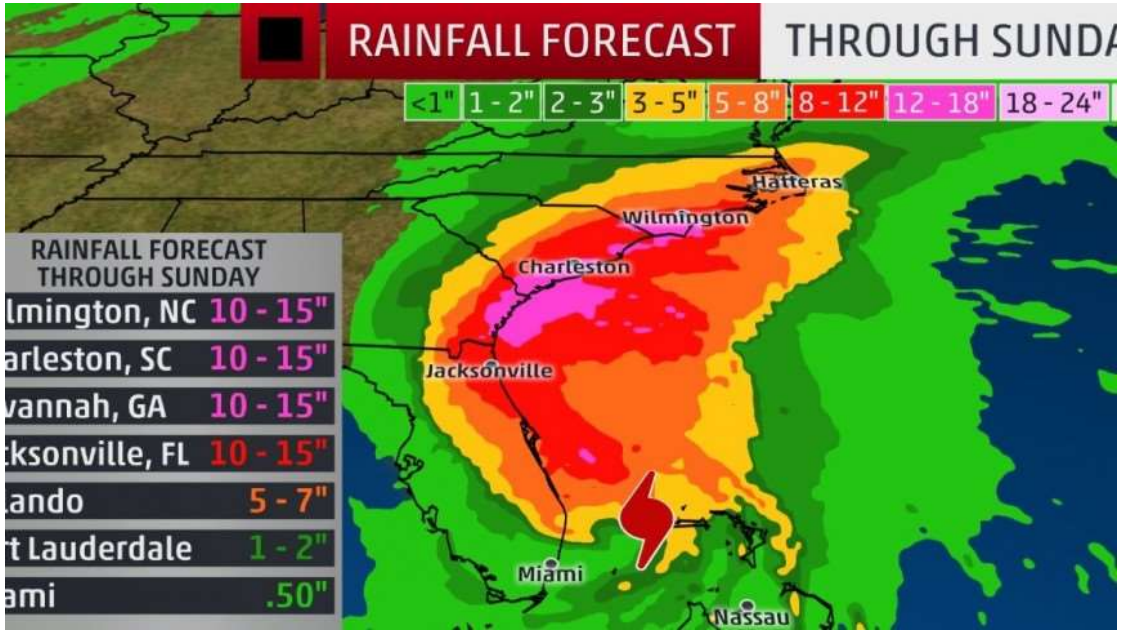
# MACHINE LEARNING IN PYTHON

NUS National University of Singapore | School of Computing




**Amirhassan Monajemi**

# MACHINE LEARNING IN PYTHON

# 2022

AMIRHASSAN MONAJEMI

# ABOUT THE LECTURER

Dr. Amirhassan Monajemi (aka Monadjemi) is a Senior Lecturer in AI and Data Science with the School of Computing at the National University of Singapore. Before joining the NUS, he was with the Faculty of Computer Engineering, University of Isfahan, Iran, where he was serving as a professor of AI and Machine Learning.

Dr. Monajemi has taught diverse computer courses for years, registered a few patents in the fields of AI, Machine Vision, and Signal Processing applications, published more than a hundred research papers in peer-reviewed, indexed journals, and supervised several Data Science, IoT, and AI industrial projects in various scales.

# AGENDA, DAY 1

| Day | Time | Topics |
|---|---|---|
| 1 | 8.45am | Registration |
| | 9.00am – 10.45am | Python Data Structures: List, Tuple, and Dictionary |
| | 10.45am – 11am | Break |
| | 11am – 1pm | Numpy Data Structures: ndarray<br>pandas Data Structures: Series and DataFrame |
| | 1pm – 2pm | Lunch |
| | 2pm – 3.45pm | Introduction to Machine Learning |
| | 3.45am – 4pm | Break |
| | 4pm – 5.30pm | Supervised Learning:<br>- Linear Regression<br>- Logistic Regression<br>- Practical |

# AGENDA, DAY 2

NUS National University of Singapore | School of Computing

| Day | Time | Topics |
|---|---|---|
| 2 | 8.45am | Registration |
| | 9.00am – 10.45am | Supervised Learning:<br>- Support Vector Machines<br>- K-Nearest Neighbors<br>- Practical |
| | 10.45am – 11am | Break |
| | 11am – 1pm | Unsupervised Learning:<br>- K-Means<br>- Hierarchical Clustering<br>- Practical |
| | 1pm – 2pm | Lunch |
| | 2pm – 3.45pm | Artificial Neural Networks |
| | 3.45am – 4pm | Break |
| | 4pm – 5.30pm | Practical |

# AGENDA, DAY 3

| Day | Time | Topics |
|---|---|---|
| 3 | 8.45am | Registration |
| | 9.00am – 10.45am | Artificial Neural Networks<br>Practice |
| | 10.45am – 11am | Break |
| | 11am – 1pm | Deep Learning<br>- Practical |
| | 1pm – 2pm | Lunch |
| | 2pm – 3.15pm | Case Study Presentations |
| | 3.15am – 3.30pm | Break |
| | 3.30pm – 5.30pm | Assessment |

# PYTHON PROGRAMMING

Python is an interpreted, high-level, general-purpose programming language. Created in 1991.

It supports many OSs, functional, structured, and object oriented programing.

Python 3 was introduced in 2009.

It is the most popular PL for AI, ML, and DA applications.

It is a FOSS

Python's core philosophy

- Explicit is better than implicit
- Beautiful is better than ugly
- Readability counts.
- Simple is better than complex.
- Complex is better than complicated.

**WWW.PYTHON.ORG**

# PYTHON PROGRAMMING

Python features are a combination of high level PLs, along with functional PLs (e.g. LISP), and matrix-oriented PLs (e.g. MATLAB).

Personal and organizational investment in learning Python are rational.

A big open source community is supporting that, you wont walk alone…

Many examples, toolboxes, and packages available for free

# PACKAGES AND TOOLS WE'RE GOING TO USE

- Anaconda
- Tensorflow
- Keras (in TF)
- Pandas
- Numpy
- Matplotlib
- Seaborn
- Sklearn
- Jupyter Notebook

… and a few more

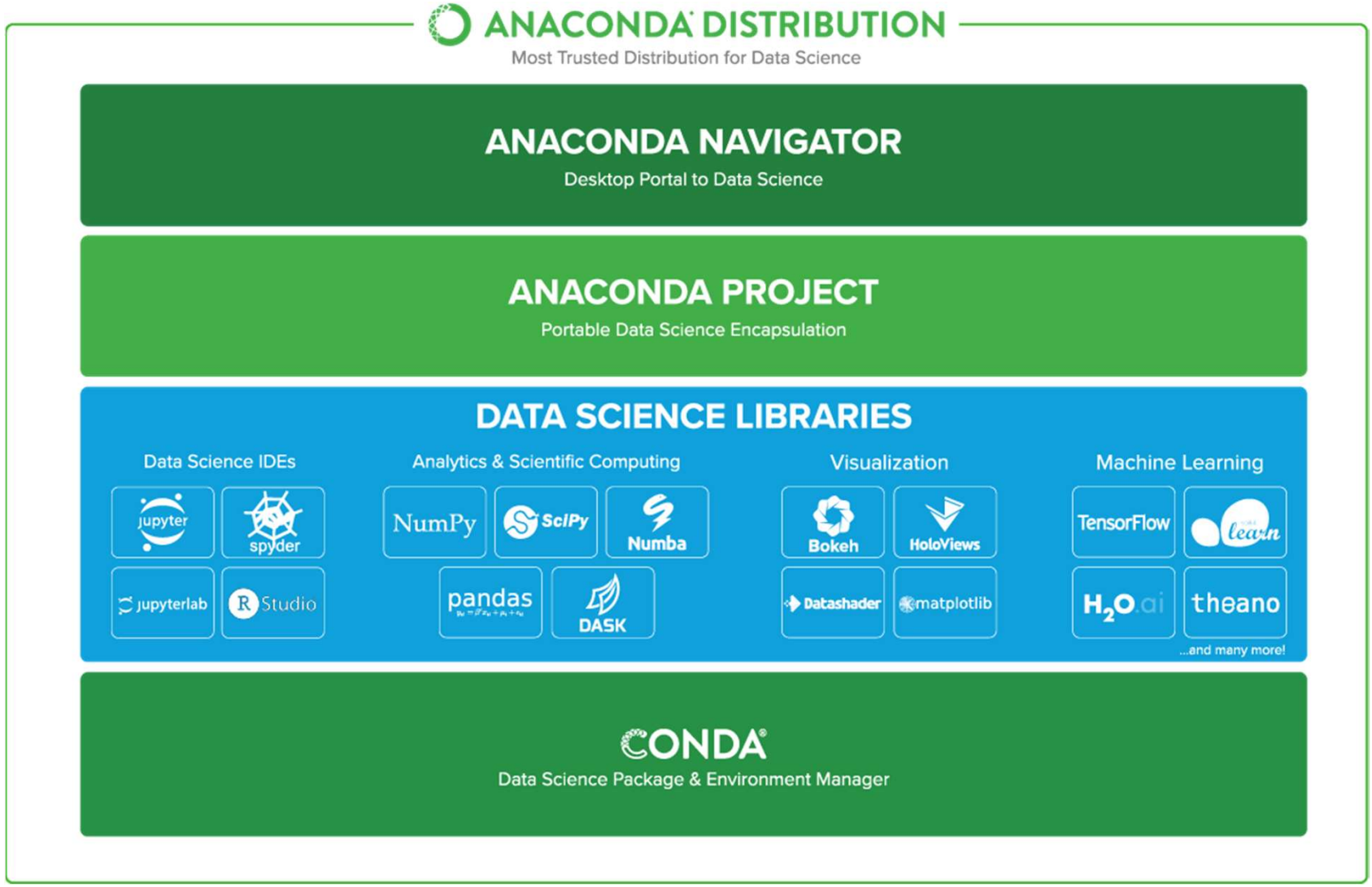
# WHAT ARE THESE?

- Anaconda: A free and open-source distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment.
- Developers describe Anaconda as "The Enterprise Data Science Platform for Data Scientists, IT Professionals and Business Leaders".
- Anaconda functions as a wrapper or a Swiss army knife, but a good one.
- It is an extra shell or a package for ML and DA.



HTTPS://STACKOVERFLOW.COM/

# WHAT ARE THESE?




**HTTPS://STACKOVERFLOW.COM/**

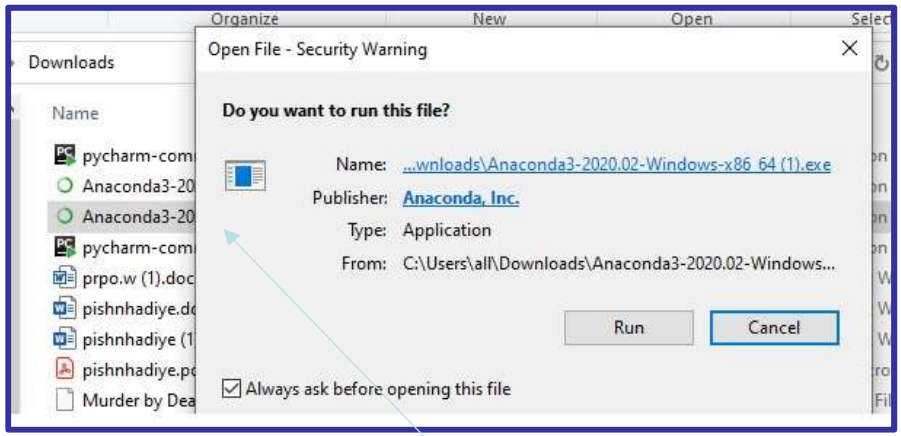
# WHAT ARE THESE?

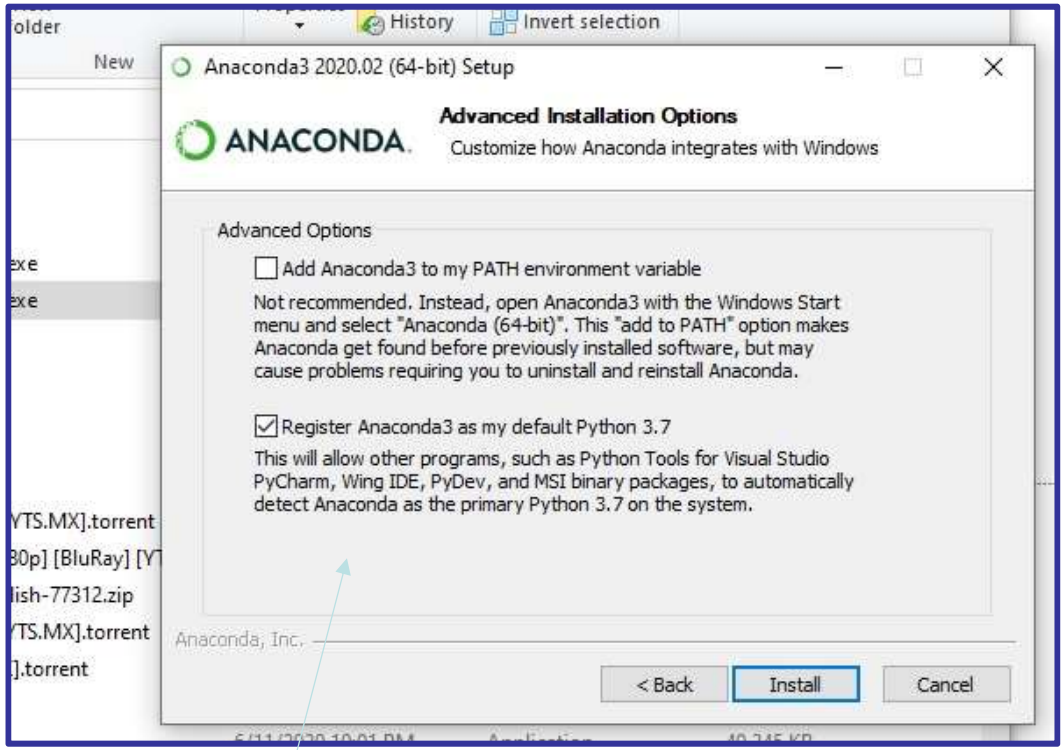
- Tensorflow: An end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community. It's very useful for ANN and Deep Learning.
- Keras: An open-source neural-network library written in Python. It is capable of running on top of TensorFlow, Microsoft Cognitive Toolkit, R, Theano, or PlaidML.
- Jupyter Notebook: an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

# PROCEDURE

1. Install Anaconda
    1. Go to: Anaconda → Products → individual, https://www.anaconda.com/products/individual
    2. Find a proper latest version for your OS and mode (64bits, in some rare cases 32bits)
    3. What is it? Well, to me Anaconda is something between an OS and a toolbox, necessary however for many Python programs and apps. Should we get back to older ICT terminology, it would be defined as a shell or a distribution. It will support many AI and Data Science programs written in Python or R. it is a FOSS.
    4. In July 2020, the advisable version is for Pyhton3.7

# ANACONDA INSTALLATION



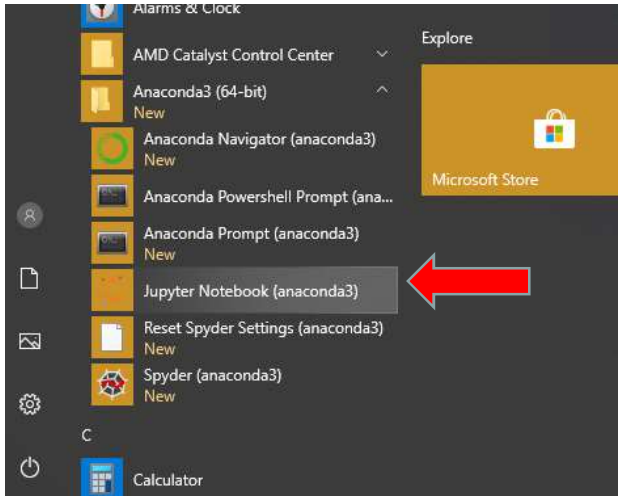



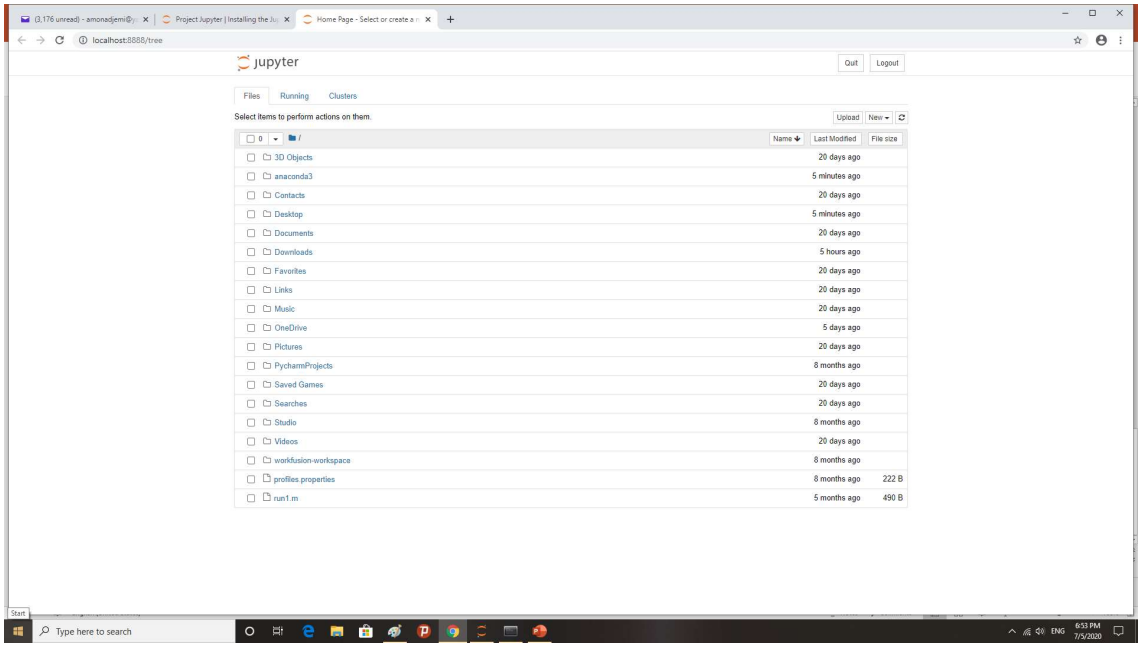

1- Run the installer file, 2- Keep it like this

# JUPYTER INSTALLATION

Typically, when you install Anaconda, Jupyter notebook would be installed automatically.

Run it, and you can have the IDE interface in you web browser.

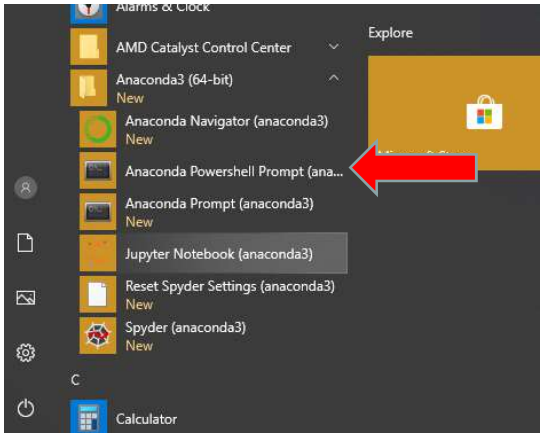


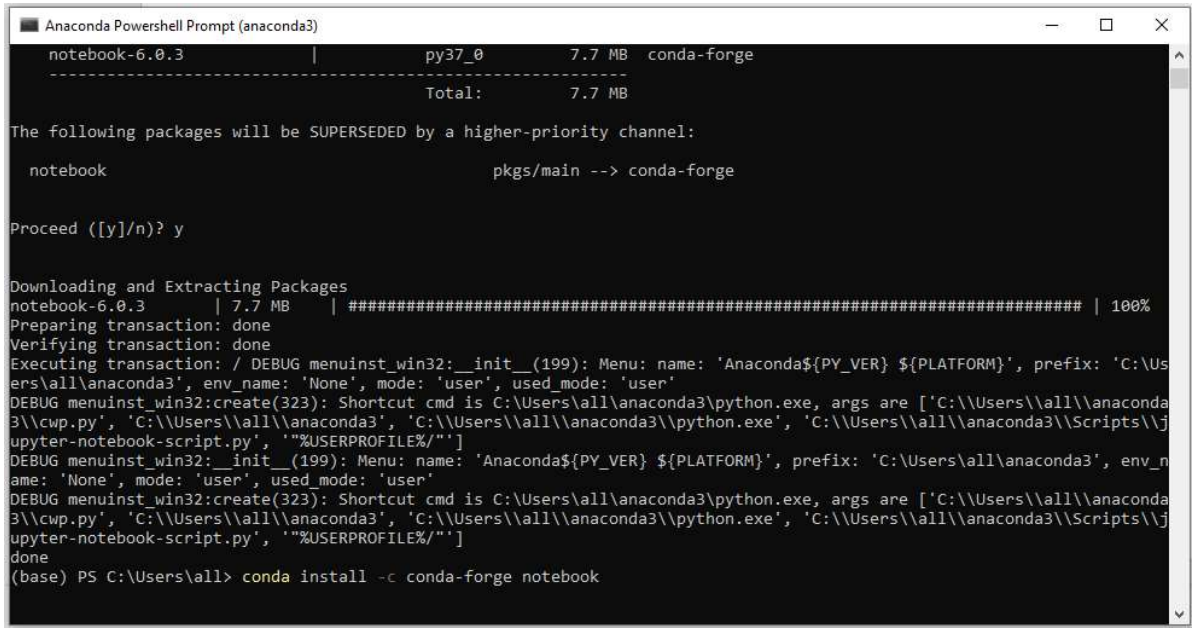



# JUPYTER INSTALLATION

Or, run the Anaconda Powershell, and type this two commands:

- >>conda install -c conda-forge jupyterlab
- >>conda install -c conda-forge notebook



# PYTHON PACKAGES INSTALLATION

REMEMBER: All through all setup steps, you must be connected to the Internet.

You may need several Python packages. Here you are some examples:

```
>>conda install -c conda-forge  tensorflow
>>conda install -c conda-forge   keras
```

To install Tensorflow and Keras packages/toolboxes/libraries.

Alternatively, you may use pip command:

```
pip  install  tensorflow
```

To check which packages have been already installed in your Python/Anaconda system, use:

```
>>conda list
```

# PYTHON PACKAGES INSTALLATION

NUS National University of Singapore | School of Computing

## To install packages from your local hard drive:

conda install --offline package path and name

For example:

conda install --offline "C:\Users\sleam\Downloads\SimpSOM-1.3.4.tar.gz"

## To install packages from a web source:

conda install -c URL

For example:

conda install -c https://conda.binstar.org/pymc pymc

**Instead, you may try:**

```
pip  install  package_name      #(from anaconda powershell prompt)
!pip  install  package_name     #(from jupyter notebook)
```

# NOTEBOOKS TO COVER

| Notebook | Comments |
|---|---|
| p1_review1.ipynb | Exercise 1 & 2 included |
| p2_numpy1.ipynb | |
| p3_Pandas1.ipynb | |
| p4_pandas2.ipynb | |
| p5_visual1_plot.ipynb | |
| Exercise 3.ipynb | Data Preparation |
| p6_housing1.ipynb | |
| p7_read1.ipynb | |
| p8_MV_regression1.ipynb | |
| p9_decision_tree1.ipynb | |
| p10_log_reg_class1.ipynb | |
| p11_hierrarchical_cluster1.ipynb | |
| p12_c_means_cluster1.ipynb | |
| Exercise 4.ipynb | Prediction |
| p13_ann1p.ipynb | |
| p14_ann_func_est_best.ipynb | |
| p15_deep_func_est_best2.ipynb | |
| Mini Project.pynb | Neural Networks |

Extra:
deep5.ipynb
deep5_gpu.ipynb

# USEFUL REFERENCES

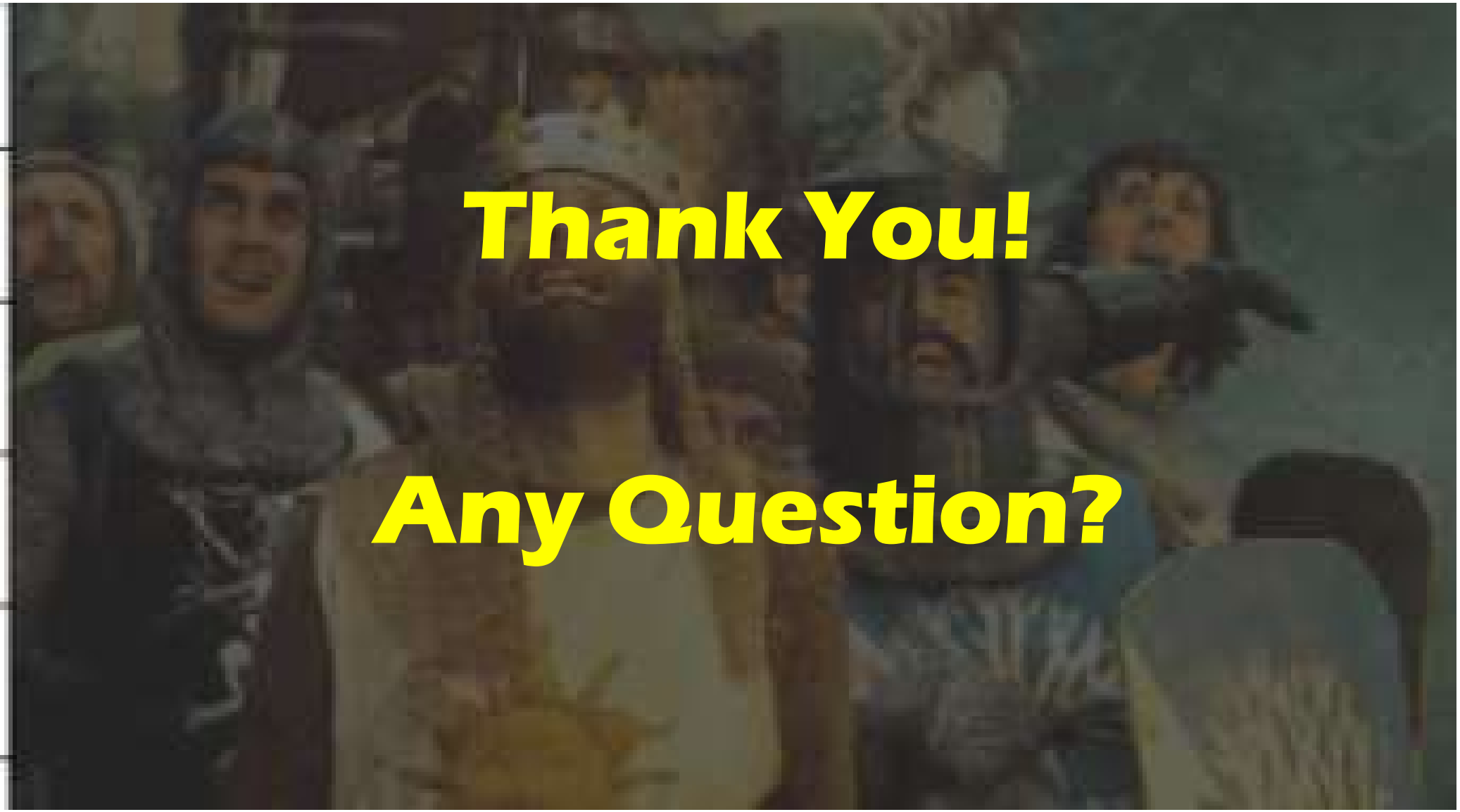

1. Artificial Intelligence, A Modern Approach, Stewart Rusell and Peter Norvig, Pearson Pub, 3rd ed. 2009.
2. Artificial Intelligence For Dummies, John Paul Mueller, Luca Massaron, Wiley, 1st ed., 2018.
3. Deep Learning For Dummies, John Paul Mueller, Luca Massaron, Wiley, 1st ed., 2019.
4. Geeksforgeeks
5. Towardsdatascience
6. Stackoverflow

# Programme Evaluation

**Please complete the SSG Post Course evaluation survey:**

https://forms.gle/2az5wni2U3nmAv7bA

# THAT'S ALL …



# LIST OF DATASETS USED

- 2.survey.csv
- Automobile_data.csv
- celebs1-bad.csv
- celebs1.csv
- celebs2.csv
- celebs3_country.csv
- collegedataold.csv
- College_Data.csv
- company_sales_data.csv
- creditcard2_smaller.csv
- housepricedata.csv
- loan_data.csv
- Mall_Customers.csv
- resale-sample.csv
- survey.csv
- titanic_train.csv
- ToyotaCorolla.csv
- USA_housing.csv
- wine-w1.csv
- winequality-red.csv
- winequality-white.csv
- wine_r1.csv
- wine_r2_t.csv

- boston.xlsx
- Sample-Superstore.xls
- Technologies(1-26) (002).xlsx
- wine_data.xlsx