About the Course

Schedule

- October 25 to December 19.
- 9:30h to 14:40h (5 hours/day).
- Holidays:
  - November 1 (Wed).
  - December 6, 7, and 8 (Wed, Thu, Fri).
- 170h for ML.
- 10h for “Formació Complementària”.

Daily Schedule

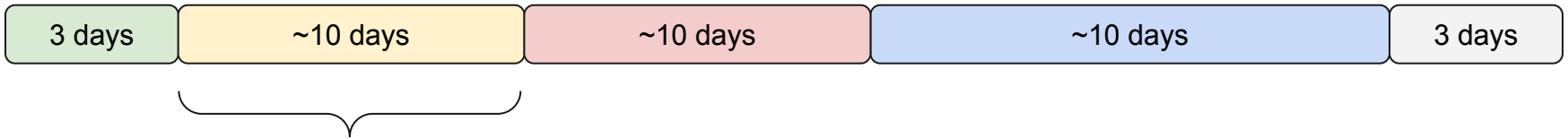
- 9:30h – 10:40h Finish previous day exercises.
  - Break
- 11h – 13h Theory + exercises.
  - Break
- 13:20h – 14:30h Exercises.

Syllabus

Intro:

- Python
- Numpy
- Pandas

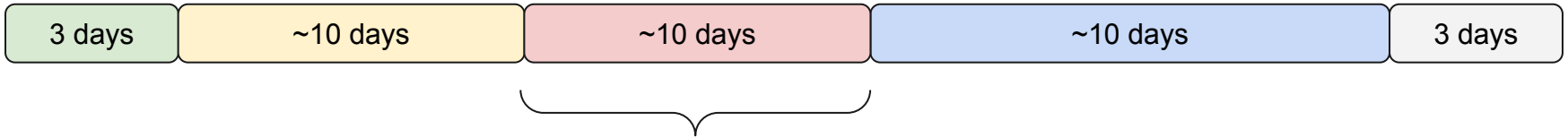
Syllabus



ML:

- Data cleaning.
- Algorithms: LR, kNN, DT, RF, GBDT, SVM, NB, NN
- Concepts: Train/Test, Cross-validation, Hyperparameter tuning, etc.
- Exercises: Price prediction, hotel cancelation

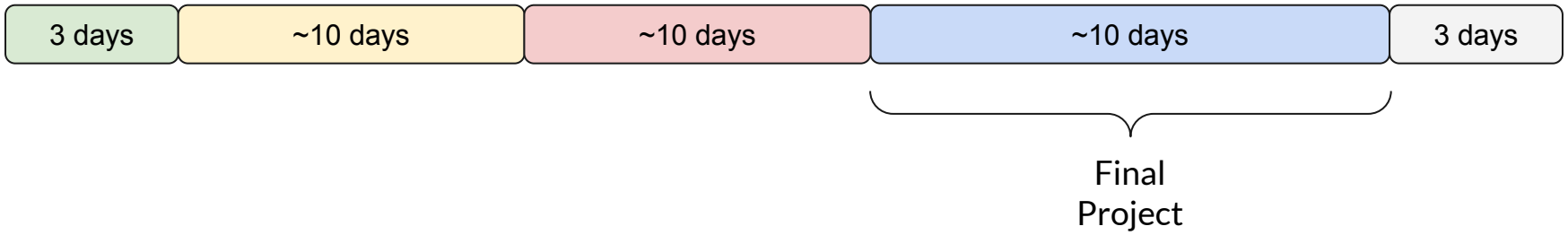
Syllabus



DL:

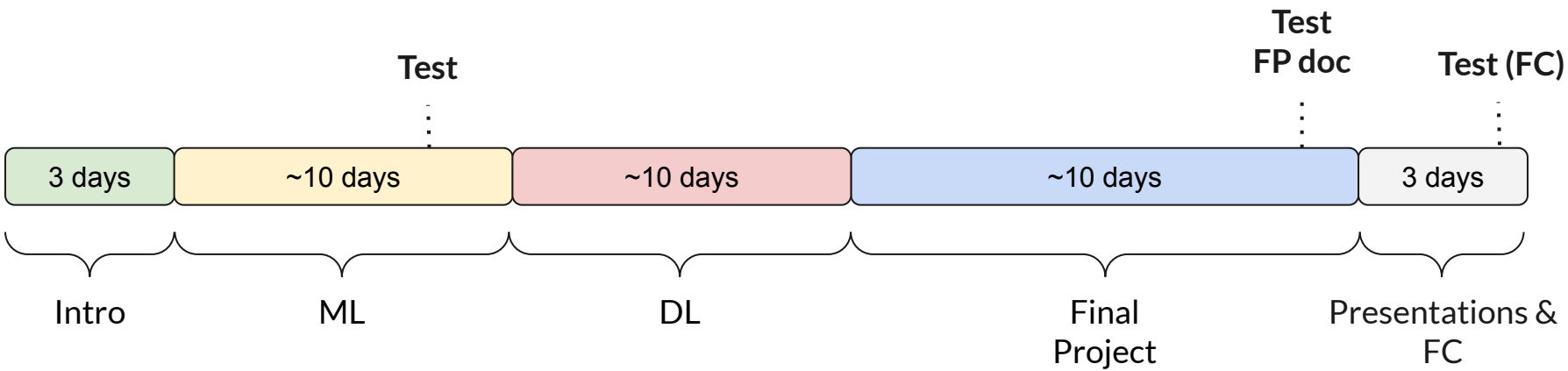
- DNN
- CNN
- RNN/LSTM
- GANs

Syllabus



Syllabus

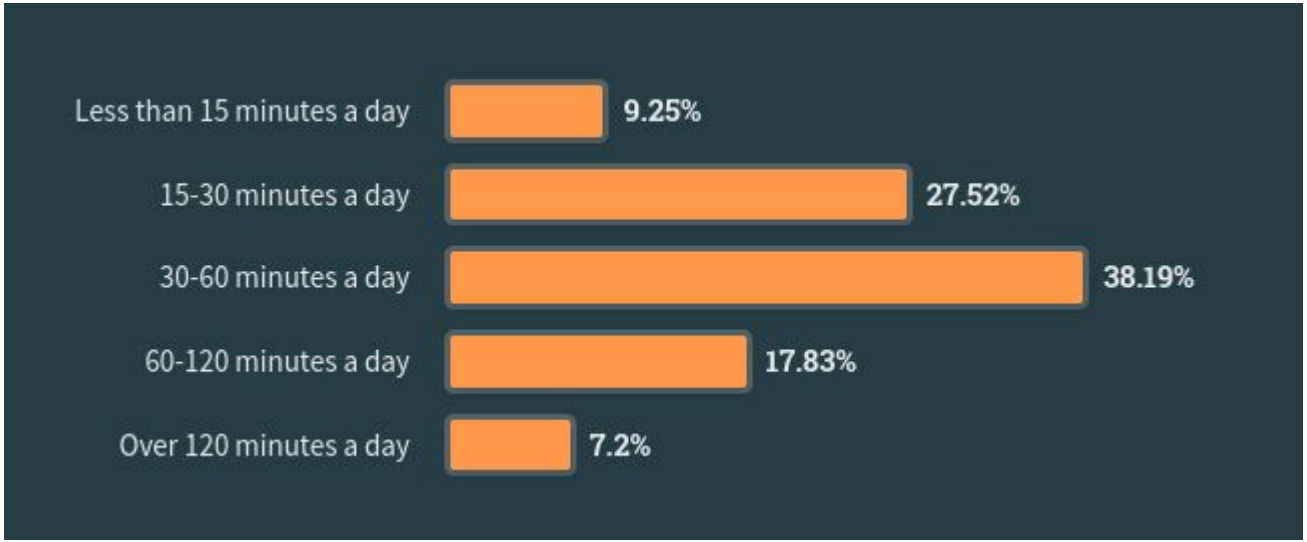
Syllabus


Test
Test
FP doc
Test (FC)
3 days
~10 days
~10 days
~10 days
3 days
Intro
ML
DL
Final
Project
Presentations &
FC

Cross-curricular

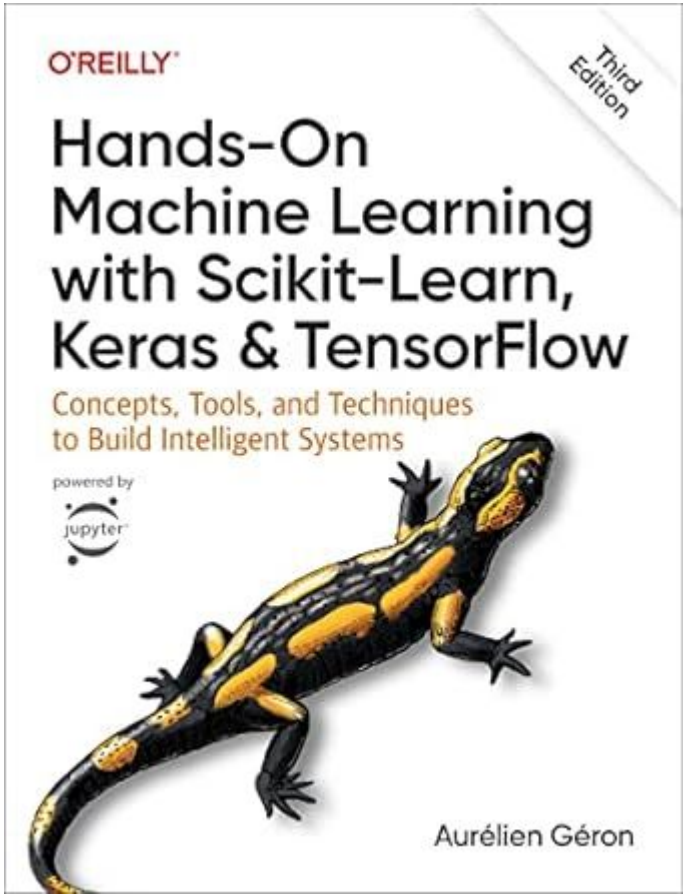
- Information search and comprehension.

Daily time spent searching for answers/solutions



Stackoverflow Survey 2023

Bibliography



Extra Bibliography

- AI
  - Artificial Intelligence: A Modern Approach (Stuart Russell et al.) (HARD)
- ML
  - The StatQuest illustrated guide to machine learning (Josh Starmer) (MEDIUM)
  - An Introduction to Statistical Learning (Gareth James et al.) (MEDIUM+)
  - The Elements of Statistical Learning (Trevor Hastie et al.) (HARD)
- DL
  - Deep Learning with Python (François Chollet) (MEDIUM)
  - Deep Learning (Ian Goodfellow et al.) (HARD)
- Maths
  - Mathematics for Machine Learning (Marc Peter Deisenroth et al.) (MEDIUM)

Python 3.X

- Numpy
- Pandas
- Matplotlib / Seaborn
- Scipy
- Sklearn
- Tensorflow / Pytorch

Why Python?

- Open source.
- Simple syntax (data science problems tend to be small scripts).
- Very useful libraries, specially for data science.
- Active community.
- Slow… but not so much with the appropriate libraries.

Jupyter Notebooks

In [1]:
```
%matplotlib inline
import pandas as pd
import numpy as np
import plotly
from IPython.display import display, Markdown as md
```

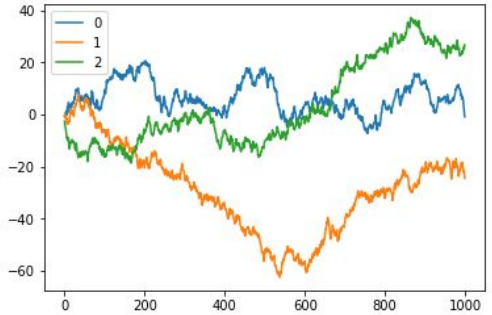
In [2]:
```
title = "My Shiny Report"
x = 1000
y = 3
```

In [3]:
```
display(md("# Just look at this graph from {}".format(title)))
```

Just look at this graph from My Shiny Report

In [4]:
```
df = pd.DataFrame(np.random.randn(x, y))
df.cumsum().plot()
```

Out[4]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f127adda278>
```



Jupyter Notebooks

- Visual Studio Code
- Browser
- JupyterLab desktop
- Google colab

Install

- Recommendation: Anaconda (www.anaconda.com).
  - Python, pip and jupyter notebooks should be included there.

- If not using Anaconda, follow the next slides for installing all the components separately.

Install Python

- Python (www.python.org).
- Pip (included with Python).
- Libraries (numpy, pandas, matplotlib, seaborn, scipy, sklearn, tensorflow)
  - `pip install numpy`
  - etc.

Jupyter Notebooks Installation

- `pip install jupyter`
- `jupyter notebook`

  or

  `jupyter notebook --no-browser`

- NB extensions (https://jupyter-contrib-nbextensions.readthedocs.io/en/latest/install.html):
  - `pip install jupyter_contrib_nbextensions`
  - `jupyter contrib nbextension install --user`

Course Drive

http://bit.ly/ml_oct_23