## 3. How to use this book

This book is a <u>Jupyter Book</u>, which allows you to interactively execute, reproduce, or modify the sections of this book that contain code. This chapter provides an overview of what Jupyter books are, and how they allow their content to be reproducible.

# 3.1. Jupyter book project

The Jupyter Book project is an open-source project. It is developed by the <u>Executable Book Project</u>, which is an <u>international collaboration between several universities and other open-source projects</u>.

Under the hood, a Jupyter Book is essentially a collection of markdown files and Jupyter notebooks. These files are compiled together according to the configuration and table of content specifications. You can follow the Github icon at the top of this page to see the source code for this book.

## 3.2. Book layout

The book layout provides five main zones:

- 1. The search tool (top left): It allows finding all sections of the book that contains a particular term.
- 2. The table of content (left panel): It allows navigating between the different chapters and main sections of the book.
- 3. The external links (top): It contains icons for launching notebooks, switching to full-screen mode, accessing the Github repository, or downloading the content of the current section.
- 4. The main panel (center): The content of the current section.
- 5. The section content (Right panel): It allows navigating in the subsections of the current section.

### 3.3. Reproducibility

All sections that contain code (Jupyter notebooks), such as this section, will have a shuttle icon in the external links, allowing to run the notebook on the cloud. Two services are provided: Google Colab, and Binder. We briefly describe these two services below, with their pros and cons.

#### 3.3.1. Google Colab

<u>Google Colab</u> is a service provided by Google Research and requires to have a Google account. It allows to run Python notebooks for free on Google cloud.

The computational resources provided by Google Colab can fluctuate. They depend on available Google resources, as well as your usage of the service (intense use may lead to lower resources). You may however expect to have one core, and a few gigabytes of RAM. Google Colab also provides a free GPU, whose model may vary depending on available resources. Provided GPU models often include Nvidia K80s, T4s, P4s, or P100s. See this <u>link</u> for more details on the resources that Google can provide.

#### 3.3.2. Binder

<u>Binder [BTK+21]</u> allows to create, use and share custom computing environments. It is powered by <u>BinderHub</u>, which is an open-source tool that deploys the Binder service in the cloud.

Notebooks may be launched and run for free, thanks to the <u>BinderHub Federation</u>, currently supported by Google Cloud Platform, OVH, GESIS notebooks, and the Alan Turing Institute. BinderHub relies on <u>Repo2Docker</u> to generate docker images from a Git repository, and <u>JupyterHub</u> to run and share

- **E** Contents
- 3.1. Jupyter book project
- 3.2. Book layout
- 3.3. Reproducibility
  - 3.3.1. Google Colab
  - 3.3.2. Binder
  - 3.3.3. Pros and cons
  - 3.3.4. Try it!
- 3.4. Local execution and book compilation

notebooks.

It takes usually around one or two minutes for Binder to create a Docker container and start serving a notebook. The environment.yml file, present at the root of the Github book repository, is used to specify which Python libraries must be installed in the docker container.

As with Google Colab, the resources provided for free by the BinderHub Federation fluctuate. See the <u>Binder user guide</u> and <u>Binder user guidelines</u> for more details on the resources that the BinderHub Federation can provide.

#### 3.3.3. Pros and cons

A summary of the features of Google Colab and Binder is provided below.

	Google Colab	Binder
Registration	Requires a Google login	None
CPU resources	Fluctuate	Fluctuate
RAM resources	Fluctuate	Fluctuate
GPU resources	Fluctuate. Often includes Nvidia K80s, T4s, P4s, or P100s	None
Session duration	Up to 12 hours  Print to PDF	Up to 12 hours
Environment	Python 3.6	Customizable with environment.yml file

In our experience, Google Colab execution times are about twice as fast as Binder.

This summary reflects the status of the two services at the time of the writing of this section (January 2021). The features of these two services are however expected to evolve in the short future.

#### 3.3.4. Try it!

Open this notebook using either Google Colab or Binder to execute the following cell:

```
print('Hello world!')

Hello world!
```

# 3.4. Local execution and book compilation

This is an open-source book, and the code is made <u>available on GitHub</u> (or by following the GitHub icon in the external links at the top of this page). You may run the book notebooks on your computer by cloning the GitHub repository.

You may also recompile the book. Compiling the book will first require you to install Jupyter Book.

Once done, this is a two-step process:

1. Clone the book repository:

git clone https://github.com/Fraud-Detection-Handbook/fraud-detection-handbook

1. Compile the book

jupyter-book build fraud-detection-handbook

The book will be available locally at fraud-detection-handbook/\_build/html/index.html.