



# Python Programming Setup Guide

Deep Learning for Text Analytics

Chair of Information Systems, Humboldt-Universität zu Berlin

Updated: April 2024



# Introduction

The tutorial materials typically consist of Jupyter notebooks. We will make these materials available via our [GitHub repository](#).

You are encouraged to familiarize yourself with the concepts of GitHub and Jupyter notebooks. Introductory materials are available on the [course Moodle page](#) as well as on the web and on YouTube.

The easiest way to open Jupyter notebooks is to use cloud-based environments like Google Colab. Colab is the right choice if you are hesitant to install software on your private computer and/or feel less comfortable with installing and configuring programs and packages.

The second option is to use your own computer for the tutorial. All the software we need is free to use and you do not need an account with Google or any other cloud provider. However, you should be ready to do some configuration work.



# Option 1: Use GitHub and Colab

Go to our GitHub repository

<https://github.com/Humboldt-WI/delta>

Open the first notebook in Colab

1. Click on “exercises”
2. Click on “tut1\_recap\_bads\_student.ipynb”
3. Click on the button in the top left “Open in Colab”

Save changes in your Google Drive

After making changes you can save a copy of the notebook in your personal Google Drive.



# Option 1: Use GitHub and Colab

The screenshot shows a Google Colab notebook interface with the following content:

**File menu:** Datei, Bearbeiten, Anzeige, Einfügen, Laufzeit, Tools, Hilfe.

**Toolbar:** + Code, + Text, In Google Drive kopieren, Teilen, Bearbeiten.

**Search bar:** 🔎

**Table of contents:**

- BADS Exercise 1 on Python programming
  - We covered a lot of concepts in the first tutorial on [Python programming](#). Solving the exercises allows you to test your familiarity with these concepts.
- Variables, assignments, and comparisons
  - 1. Create two variables *a* and *b* and assign values of 3 and 4.5.
  - 2. Query the type of variable *a*.
  - 3. Check whether variable *b* is a text variable.
  - 4. Calculate  $a^2 + \frac{1}{b}$ ,  $\sqrt{a * b}$ , and  $\log_2(a)$ .
- Matrix algebra
  - Create three additional variables as follows:
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$



## Option 2: Use your own machine

### Install Python3 and git

Python: <https://www.python.org/downloads/> (It is recommended to install the second-latest version, since new releases tend to have compatibility issues with some modules)

Git: <https://git-scm.com/>

### Open terminal

Search for “cmd.exe” on Windows.

On Mac, iterm2 is a good alternative to the default “terminal.app”.

### Clone repository

Download the [GitHub Desktop App](#) and clone the repository from our URL ...

... or type into the terminal:

```
git clone https://github.com/Humboldt-WI/delta
```



## Option 2: Use your own machine

### Install requirements

Type into the terminal:

```
cd delta  
pip install --upgrade pip  
pip install -r requirements.txt
```

### Open Jupyter

Type into the terminal:

```
jupyter lab
```

A tab with Jupyter Lab should open in your default browser.

You can now browse the repository and select the first exercise notebook.

Any changes to the exercise notebooks should be saved outside of the cloned git repo to avoid merge conflicts when the online repository is updated.



## Option 2: Use your own machine

```
(base) Vincents-Air-2:bads-repo VincentGurgul$ jupyter lab
Fail to get yarn configuration. dyld: Library not loaded: /usr/local/opt/icu4c/lib/libicui18n.68.dylib
Referenced from: /usr/local/bin/node
Reason: image not found

[I 2022-10-05 16:07:08.145 ServerApp] jupyterlab | extension was successfully linked.
[I 2022-10-05 16:07:08.160 ServerApp] nbclassic | extension was successfully linked.
[I 2022-10-05 16:07:08.822 ServerApp] notebook_shim | extension was successfully linked.
[I 2022-10-05 16:07:08.823 ServerApp] panel.io.jupyter_server_extension | extension was successfully linked.
[I 2022-10-05 16:07:08.892 ServerApp] notebook_shim | extension was successfully loaded.
[I 2022-10-05 16:07:08.895 LabApp] JupyterLab extension loaded from /Users/VincentGurgul/opt/anaconda3/lib/python3.8/site-packages/jupyterlab
[I 2022-10-05 16:07:08.895 LabApp] JupyterLab application directory is /Users/VincentGurgul/opt/anaconda3/share/jupyter/lab
[I 2022-10-05 16:07:08.904 ServerApp] jupyterlab | extension was successfully loaded.
[I 2022-10-05 16:07:08.914 ServerApp] nbclassic | extension was successfully loaded.
/panel-preview/ /panel-preview
[I 2022-10-05 16:07:08.918 ServerApp] panel.io.jupyter_server_extension | extension was successfully loaded.
[I 2022-10-05 16:07:08.920 ServerApp] Serving notebooks from local directory: /Users/VincentGurgul/Documents/Arbeit/Lessmann/work/bads/bads-repo
[I 2022-10-05 16:07:08.920 ServerApp] Jupyter Server 1.19.1 is running at:
[I 2022-10-05 16:07:08.920 ServerApp] http://localhost:8888/lab?token=efaedd8ca7414941b393e02769aab9cd07c289e6e7059444
[I 2022-10-05 16:07:08.920 ServerApp] or http://127.0.0.1:8888/lab?token=efaedd8ca7414941b393e02769aab9cd07c289e6e7059444
[I 2022-10-05 16:07:08.920 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 2022-10-05 16:07:08.929 ServerApp]

To access the server, open this file in a browser:
file:///Users/VincentGurgul/Library/Jupyter/runtime/jpserver-18983-open.html
Or copy and paste one of these URLs:
http://localhost:8888/lab?token=efaedd8ca7414941b393e02769aab9cd07c289e6e7059444
or http://127.0.0.1:8888/lab?token=efaedd8ca7414941b393e02769aab9cd07c289e6e7059444
```



# Option 2: Use your own machine

Screenshot of a Jupyter Notebook interface titled "BADS Exercise 1 on Python programming".

The sidebar shows a list of exercises:

- 1\_ex\_python.ipynb (4 days ago)
- 2\_ex\_desc... (4 days ago)
- 3\_ex\_pred... (4 days ago)
- 4\_ex\_data... (4 days ago)
- 5\_ex\_supe... (4 days ago)
- 6\_ex\_mod... (4 days ago)
- 8\_ex\_ense... (4 days ago)

The main content area displays the following text and exercises:

We covered a lot of concepts in the first tutorial on Python programming. Solving the exercises allows you to test your familiarity with these concepts.

### Variables, assignments, and comparisons

1. Create two variables  $a$  and  $b$  and assign values of 3 and 4.5.
2. Query the type of variable  $a$ .
3. Check whether variable  $b$  is a text variable.
4. Calculate  $a^2 + \frac{1}{3} \cdot \sqrt{a * b}$ , and  $\log_2(a)$ .

### Matrix algebra

Create three additional variables as follows:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

```
[1]: # Code to create variables A, B, and y
```

At the bottom, status bar: Simple 0 1 Python 3 | idle Mode: Edit Ln 1, Col 1 1\_ex\_python.ipynb

# Virtual Environments



- It is recommended (but not required) to set up a virtual environment when running Python on your local machine. There are two options, the built-in Python environment manager and the free third-party software Anaconda.
- For a guide on Python virtual environments see:  
<https://www.dataquest.io/blog/a-complete-guide-to-python-virtual-environments/>
- For a guide on Anaconda see:  
<https://docs.anaconda.com/free/navigator/tutorials/manage-environments/>

# Useful resources and further reading



- Colab:  
<https://www.kdnuggets.com/2020/06/google-colab-deep-learning.html>
- Jupyter:  
<https://towardsdatascience.com/getting-the-most-out-of-jupyter-lab-9b3198f88f2d>
- More on Jupyter:  
<https://pabloinsente.github.io/intro-jupyter-ide>

# Thank you!

Vincent Gurgul

Chair of Information Systems  
School of Business and Economics  
Humboldt-University of Berlin, Germany

Telephone: +49 30 2093-99543  
E-Mail: [vincent.gurgul@hu-berlin.de](mailto:vincent.gurgul@hu-berlin.de)

<https://bit.ly/hu-wi>  
<https://www.hu-berlin.de>

