

# Leeswijzer Data Mining & Exploration

## Table of contents

<b>Afkortingen</b>	<b>1</b>
<b>Les 1 - Introductie</b>	<b>2</b>
Libraries en tools: . . . . .	2
pyenv . . . . .	2
poetry . . . . .	2
git . . . . .	2
pandas . . . . .	2
polars . . . . .	2
loguru . . . . .	3
click . . . . .	3
pydantic . . . . .	3
<b>Les 2: Visualisatie</b>	<b>3</b>
<b>Les 3 - Statistiek</b>	<b>3</b>
<b>Les 4: Evaluatie</b>	<b>3</b>
<b>Les 5: Integratie</b>	<b>3</b>
<b>Les 6: Testen, presenteren</b>	<b>3</b>
<b>Les 7: Presentaties</b>	<b>3</b>

## Afkortingen

PDA: Python for Data Analysis MT: the model thinker

## Les 1 - Introductie

Les 1 heeft vooral een heleboel setup van randvoorwaarden. Kijk op zn minst even op alle paginas die ik hieronder heb neergezet; in een uurtje zou je alles moeten kunnen bekijken en de stukjes die ik aan heb gegeven doorgelezen.

### Libraries en tools:

#### pyenv

- management van python versies.
- [docs](#)
- Lees “what pyenv does”, “pyenv does not” en “usage”

#### poetry

- virtual environment management. Alternatief voor pip en conda.
- [docs](#)
- Lees [Basic usage](#) hele pagina en [version constraints](#)

#### git

- versiemangement van je code
- Lees de handleiding (git.pdf) in de repo

#### pandas

- Data frames. Zou bekend moeten zijn. Als naslag, hoofdstuk 5 en 7 van PDA.

#### polars

- Data frames. [docs]<sup>9</sup><https://github.com/pola-rs/polars>). Blazingly fast dataframes, met een backend in Rust.
- Optioneel, maar aan te raden, vooral voor zwaardere workloads.
- De [user guide](#) geeft een goede introductie.

## **loguru**

- logging.
- Gebruik nooit meer print statements, behalve in je jupyter notebooks. Verder gebruik je altijd loguru in je programmas, tenzij je een hele goede reden hebt.
- [docs](#)

## **click**

- command line argument parsing
- [docs](#)

## **pydantic**

- setting management en dataclasses.
- Hier is een vergelijking met andere opties: [click](#). Lees even door, het is een 5 minute read, want pydantic is veel rijker dan settings alleen.

## **Les 2: Visualisatie**

PDA hoofdstuk 9

## **Les 3 - Statistiek**

MT: hoofdstuk 3, hoofdstuk 5.

## **Les 4: Evaluatie**

Dashboarding [flet.dev](#)

## **Les 5: Integratie**

## **Les 6: Testen, presenteren**

## **Les 7: Presentaties**