

# Applied AI with Python

## Course Introduction

FH-Ass.Prof. DI Martin Uray, BSc

Josef Ressel Center for Intelligent and Secure Industrial Automation  
Department for Information Technologies and Digitalisation  
Salzburg University of Applied Sciences

September 30, 2024



Salzburg University  
of Applied Sciences

# About Me

Lecturer



## **Martin Uray**

- ▶ Lecturer / Researcher at the Salzburg University of Applied Sciences
- ▶ Open Student Hours:
  - ▶ Monday, between 09.30 am - 4.30 pm,
  - ▶ and before / after class,
  - ▶ online via MS TEAMS
  - ▶ book slots using the booking tool
  - ▶ Exceptions may apply!

# Very Brief Overview of Artificial Intelligence

# Artificial Intelligence I

*Computer systems that have some of the qualities that the human brain has, such as the ability to understand and produce language, recognize or create pictures, solve problems, and learn; the study or creation of systems*

– Cambridge Dictionary

# Artificial Intelligence II

*Defining AI is not easy; in fact, there is no generally accepted definition of the concept ...*

*In its strictest definition, AI stands for the imitation by computers of the intelligence inherent in humans ....*

*AI is that it is a technology that enables machines to imitate various complex human skills. This, however, does not give us much to go on*

...

– Mission AI. Springer

# Artificial Intelligence III

*The field of artificial intelligence, or AI, is concerned with not just understanding but also building intelligent entities—machines that can compute how to act effectively and safely in a wide variety of novel situations. . . .*

*Some have defined intelligence in terms of fidelity to human performance, while others prefer an abstract, formal definition of intelligence called rationality—loosely speaking, doing the “right thing”.*

– Artificial Intelligence: A modern approach

# Artificial Intelligence IV



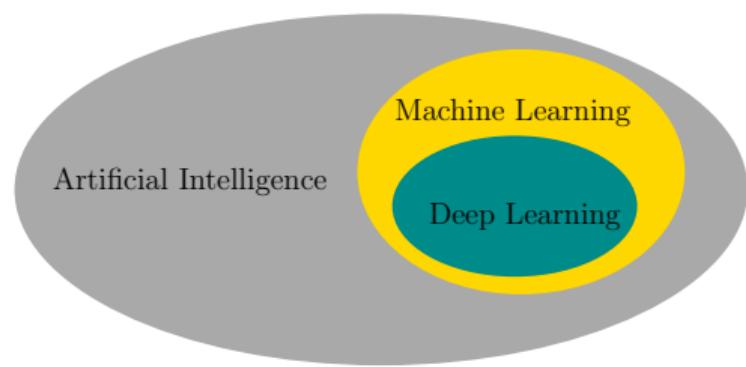
I make decisions based on  
**experience.**



I make decisions based on  
**data.**

# Artificial Intelligence V

- ▶ **Artificial Intelligence:** The set of all tasks in which a computer can make decisions
- ▶ **Machine Learning:** The set of all tasks in which a computer can make decisions based on data
- ▶ **Deep Learning:** The field of machine learning that uses certain objects called neural networks



## **Sample Applications:**

- ▶ Recommender Systems
- ▶ Dimensionality Reduction
- ▶ Autonomous Driving

# Introduction to Data Analysis

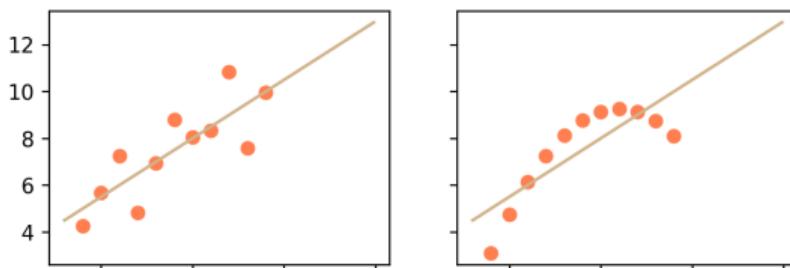
# Motivation I

## Anscombe's Quartet

Four datasets, 10 points each

	X1	Y1	X2	Y2	X3	Y3	X4	Y4
0	10.000000	8.040000	10.000000	9.140000	10.000000	7.460000	8.000000	6.580000
1	8.000000	6.950000	8.000000	8.140000	8.000000	6.770000	8.000000	5.760000
2	13.000000	7.580000	13.000000	8.740000	13.000000	12.740000	8.000000	7.710000
3	9.000000	8.810000	9.000000	8.770000	9.000000	7.110000	8.000000	8.840000
4	11.000000	8.330000	11.000000	9.260000	11.000000	7.810000	8.000000	8.470000
5	14.000000	9.960000	14.000000	8.100000	14.000000	8.840000	8.000000	7.040000
6	6.000000	7.240000	6.000000	6.130000	6.000000	6.080000	8.000000	5.250000
7	4.000000	4.260000	4.000000	3.100000	4.000000	5.390000	19.000000	12.500000
8	12.000000	10.840000	12.000000	9.130000	12.000000	8.150000	8.000000	5.560000
9	7.000000	4.820000	7.000000	7.260000	7.000000	6.420000	8.000000	7.910000
10	5.000000	5.680000	5.000000	4.740000	5.000000	5.730000	8.000000	6.890000

Anscombe's quartet



Properties of all four datasets<sup>a</sup>:

Property	Value
Mean of $x$	9
Sample variance of $x$ : $s_x^2$	11
Mean of $y$	7.50
Sample variance of $y$ : $s_y^2$	4.125
Correlation between $x$ and $y$	0.816
Linear regression line	$y = 3.00 + 0.500x$

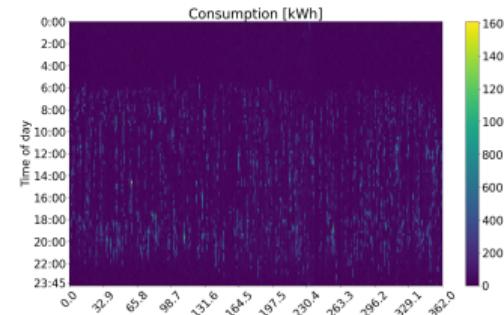
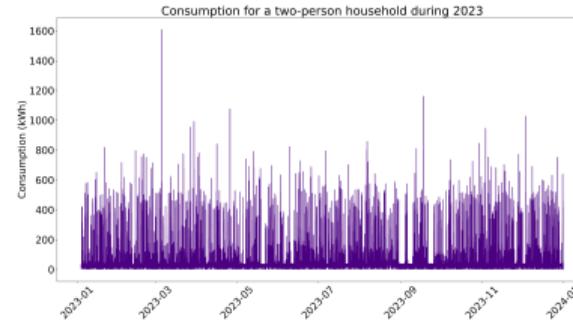
<sup>a</sup> F. J. Anscombe. "Graphs in Statistical Analysis". In: *The American Statistician* 27.1 (Feb. 1973), pp. 17–21. ISSN: 0003-1305, 1537-2731. DOI: 10.1080/00031305.1973.10478966. (Visited on

# Motivation II

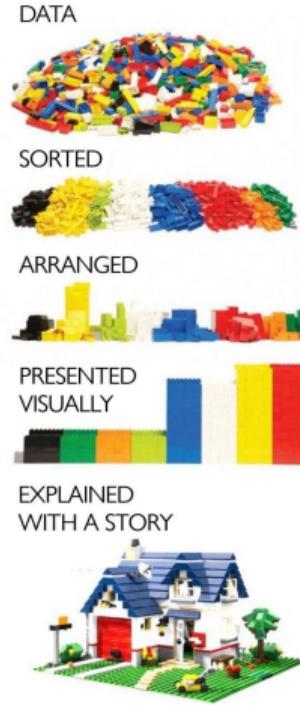
On the example of Smart Meter data

Datum	Restverbrauch (kWh)
2023-01-04 00:00:00	34.000000
2023-01-04 00:15:00	23.000000
2023-01-04 00:30:00	13.000000
2023-01-04 00:45:00	13.000000
2023-01-04 01:00:00	13.000000
2023-01-04 01:15:00	28.000000
2023-01-04 01:30:00	33.000000
2023-01-04 01:45:00	25.000000
2023-01-04 02:00:00	13.000000
2023-01-04 02:15:00	12.000000
2023-01-04 02:30:00	13.000000
2023-01-04 02:45:00	22.000000
2023-01-04 03:00:00	33.000000
2023-01-04 03:15:00	3.000000
2023-01-04 03:30:00	13.000000
2023-01-04 03:45:00	13.000000
2023-01-04 04:00:00	12.000000
2023-01-04 04:15:00	13.000000
2023-01-04 04:30:00	35.000000
2023-01-04 04:45:00	33.000000

VS.



# Motivation III



---

Source: <https://twitter.com/alvinfoo/status/1431458670684954629>

# Data Analysis

## Exploratory data analysis field of statistics that<sup>1</sup>

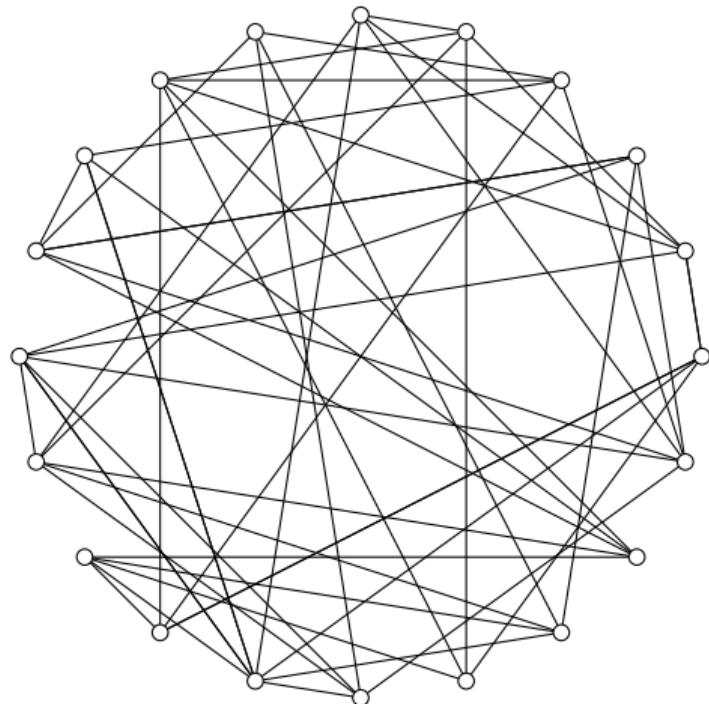
- ▶ “...has been an influential back-to-basics movement, eschewing probability models and focusing on graphical visualization of data.”
- ▶ “...along with a general view of data science as going beyond statistical theory ...”
- ▶ “...focused on discovery ...”

---

<sup>1</sup> Andrew Gelman and Aki Vehtari. “What are the most important statistical ideas of the past 50 years?” In: *arXiv:2012.00174 [stat]* (June 2021). (Visited on 07/06/2021).

# Data Sources

- ▶ Databases
- ▶ APIs
- ▶ Web Scraping
- ▶ Data Streams
- ▶ (Flat Files)



# Data Science with Python

# Data Science with Python?

Scientific Computing Languages: *R*, *SAS*, *Stat*, *MATLAB*, *SPSS*, ...

## Which Language to use?

- ▶ <https://www.tiobe.com/tiobe-index/>
- ▶ <https://bootcamp.berkeley.edu/blog/most-in-demand-programming-languages/>
- ▶ <https://statisticstimes.com/tech/top-computer-languages.php>
- ▶ ...

# Language Introduction

## Python

Python is an easy-to-use language that makes it simple to get your program working. This makes Python ideal for prototype development and other ad-hoc programming tasks. However, Python as well supports object-oriented programming with classes and multiple inheritance. Code can be grouped into modules and packages. – *python.org*<sup>2</sup>

---

<sup>2</sup> Guido van Rossum and Fred L. Drake. *The Python Language Reference*. <https://docs.python.org/3/reference/index.html>. 2011. (Visited on 09/11/2023).

# Language Introduction

Python is an **interpreted language**. Thus, it is similar to Matlab, but opposed to C, for instance, which is a compiled language.

## A SIDE-BY-SIDE COMPARISON OF COMPILED LANGUAGES AND INTERPRETED LANGUAGES



A look at how compilers and interpreters work, and how their differences affect memory, runtime speed, and computer workload.

	A COMPILER	AN INTERPRETER
Input	... takes an entire program as its input.	... takes a single line of code, or instruction, as its input.
Output	... generates intermediate object code.	... does not generate any intermediate object code.
Speed	... executes faster.	... executes slower.
Memory	... requires more memory in order to create object code.	... requires less memory (doesn't create object code).
Workload	... doesn't need to compile every single time, just once.	... has to convert high-level languages to low-level programs at execution.
Errors	... displays errors once the entire program is checked.	... displays errors when each instruction is run.

# Language Introduction

The two major Python versions, Python 2 and **Python 3**, are quite different from each other.

## Python3

In this course we only use Python 3! Python2 is EOL. Be aware of the two versions when searching for code snippets online.

For an in-depth overview about the differences between the two major versions, be referred to Sebastian Raschka's Blog.

## Alternate Language Implementations:

- ▶ CPython
- ▶ Jython
- ▶ Python for .NET
- ▶ IronPython
- ▶ PyPy

Each of these implementations varies in some way from the language as documented in this manual, or introduces specific information beyond what's covered in the standard Python documentation. Please refer to the implementation-specific documentation to determine what else you need to know about the specific implementation you're using.

# Python: Pro and Cons

## Advantages:

- ▶ Scripting Language
- ▶ General Purpose Language
- ▶ Very active community
- ▶ improved support for libraries
- ▶ FOSS

## Drawbacks:

- ▶ slow
- ▶ concurrent applications
- ▶ not very “green”<sup>3</sup>

---

<sup>3</sup> <https://www.green-coding.io/case-studies/energy-efficiency-python/>

# Environments

- ▶ Environments manage the packages for a certain project or application
- ▶ when activated only using packages from environment
- ▶ possible to have different versions for same package on different projects
- ▶ e.g. *conda* or *virtualenv*

# Package Manager

- ▶ Have you ever struggled with your packages installed?
- ▶ Package manager take care about installing and managing packages (or libraries).
- ▶ e.g. *conda* or *pip*

# IDE vs. Notebooks I

## Jupyter Notebooks

- ▶ All in One:
  - ▶ Code
  - ▶ Visualizations
  - ▶ Text (Markdown and  $\text{\LaTeX}$  support)
- ▶ Report look-and-feel

## Integrated Development Environment

- ▶ similar to what known already (Eclipse, Spyder, etc.)
- ▶ Development of applications
- ▶ integration of external dependencies
- ▶ can take care about environments
- ▶ Notebook Support

# IDE vs. Notebooks II

## Attention

Notebooks are cool for quick tests and data exploration, but shall not be applied for development, see this video.

# Python

## Community and Conferences/Meetups

### **Communities**

pydata (google group)  
pystatsmodels  
scikit-learn mailing-list  
etc.

### **Meetups**

PyCon and EuroPython  
regional PyCon conferences  
SciPy and EuroSciPy  
PyData

# To Dos for you

- Install Conda and create environment (see instructions on course website)
- Watch the video (Reading; see course website)

## Working Setup

A missing or non-working environment leads to a not account of attendance.