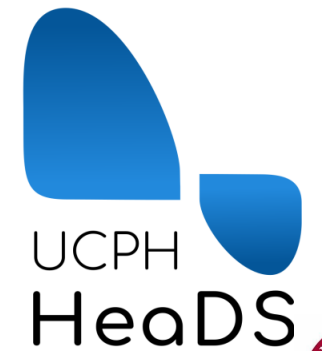


# Python Tsunami

– June 7<sup>th</sup>-9<sup>th</sup> –



# Center for Health Data Science (HeaDS)

<https://heads.ku.dk>

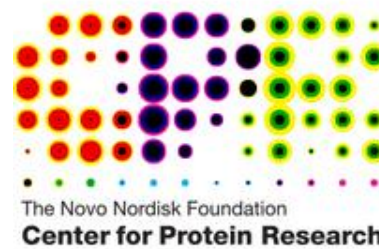
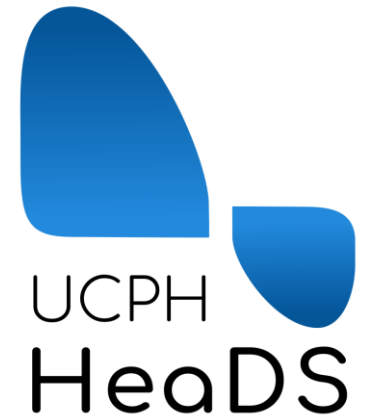
- **The Data Science lab:**
  - Provides data science support for all research groups at SUND
  - Organizes courses
- **Research units:**
  - work on different areas and topics within the field of health data science





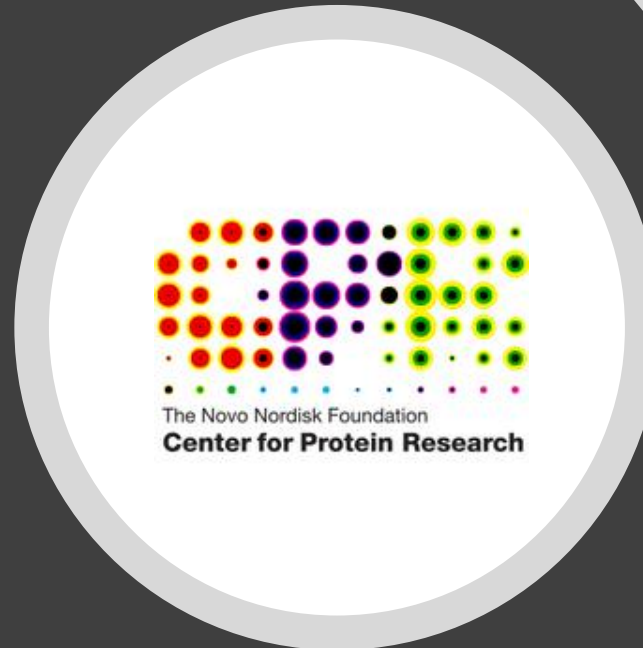
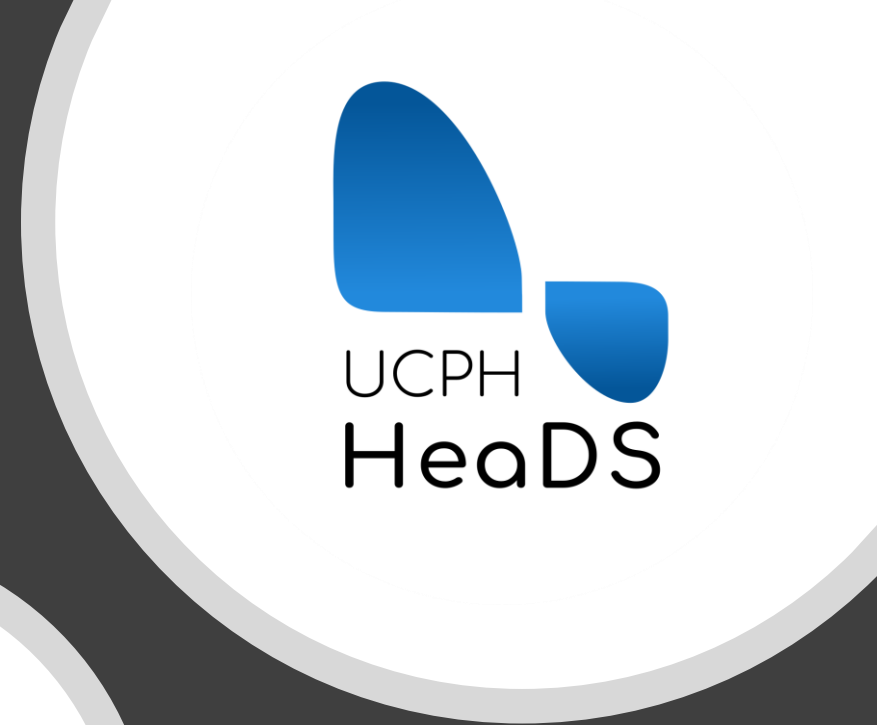
# The Team

1. Alberto Santos (HeaDS)
2. Jose Alejandro Romero Herrera (HeaDS)
3. Annelaura Bach Nielsen (NNF CPR)
4. Davide Placido (NNF CPR)
5. Henry Webel (NNF CPR)
6. Philip Charles (DBI (Oxford, UK))
7. Rita Colaço (PRI)
8. Roc Reguant (NNF CPR)
9. Thilde Terkelsen (HeaDS)



# Other Members of the Team

1. Dhouha Grissa (NNF CPR)
2. Grzegorz Jerzy Maciag (BRIC)
3. Katerina Nastou (NNF CPR)
4. Kübra Altinel (BRIC)
5. Marilena Hohmann (HeaDS)
6. Marta Matos (GENOME Center)
7. Nicholas Luke Cowie (DTU)



# Practical Things about the Course

---

Program

---

Coffee and Q&A

---

Teams

---

Breakout rooms

---

Datathon

# Practical Things about the Course

## -- Program --

	Monday 7th June	Tuesday 8th June	Wednesday 9th June
8:45-9:00		Coffee and the day before (optional)	
9:00-09:45	Introduction and motivation	Conditions	Visualization I
9:45-10:00	Coffee break		
10:00-11:00	Tools	Loops	Visualization II
11:00-12:00	Variables and data types	Functions	
	Numbers and operators		
12:00-13:00	Lunch break		
13:00-14:00	Importing data	Numpy	Datathon
14:00-14:45	Data structures		
14:45-15:15	Coffee break		
15:15-16:00	Data structures	Pandas	Datathon
16:00-17:00			Presentations

# Practical Things about the Course

## -- Teams --

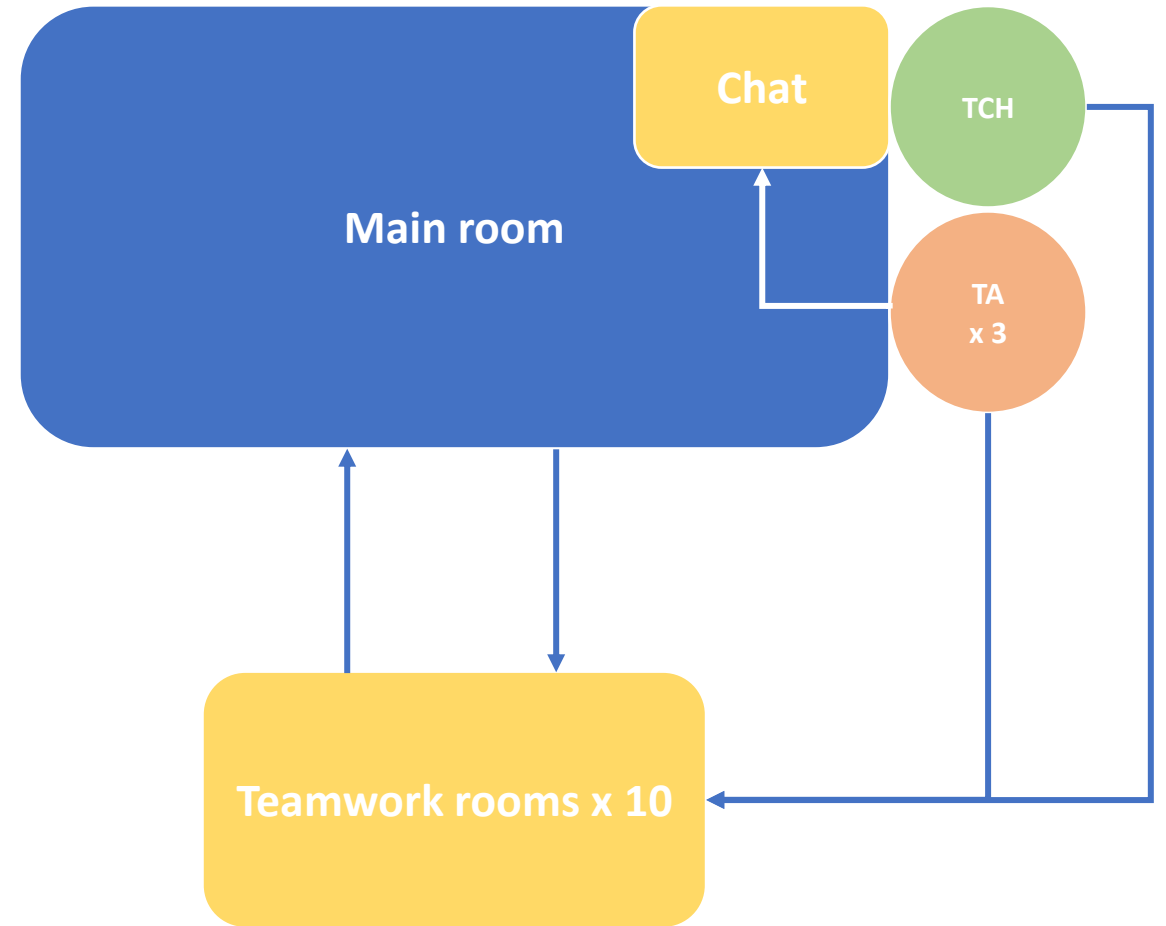
52 participants divided into 8 fixed teams

2 working modes:

- **Individual:** exercises
- **Teamwork:** discussions, practice and the Datathon

# Practical Things about the Course

## -- Breakout Rooms --

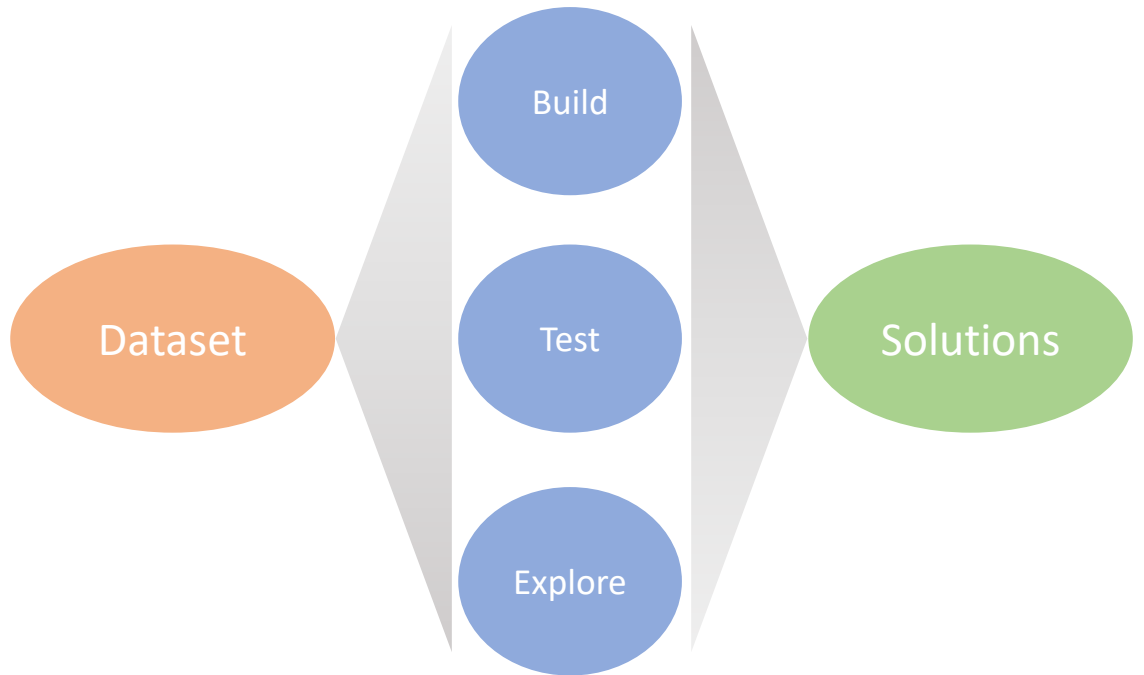




# Practical Things about the Course

## -- Datathon --

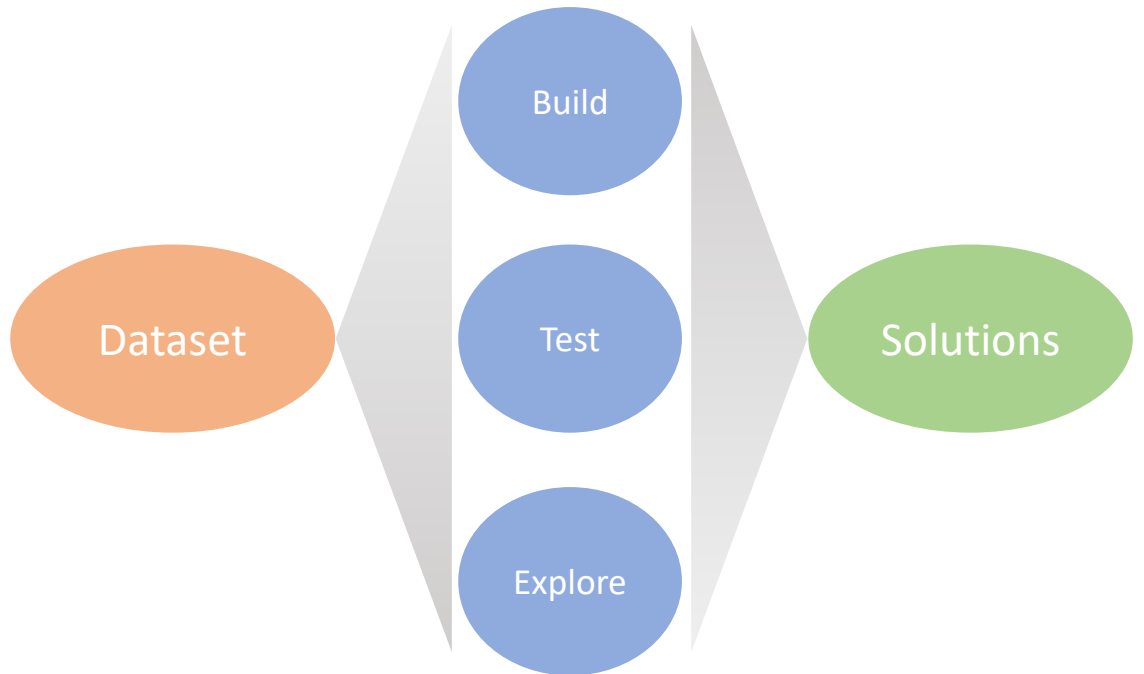
A **Datathon** is a **data-focused competition** — given a **dataset** and a limited amount of time, participants are challenged to use their **creativity** and **data science skills** to:



# Practical Things about the Course

## -- Datathon --

A **Datathon** is a **data-focused competition** — given a **dataset** and a limited amount of time, participants are challenged to use their **creativity** and **data science skills** to:



# What will you learn in this course?

---

Tools to work with Python

---

The basics of Python

---

Some of the most relevant scientific libraries

---

Visualization

---

Good practice

# Motivation

---

## Why will programming help you?

---

**Programming** is yet another **laboratory technique**

---

It helps you **automate processes** that you need to repeat again and again

---

It will **save you time**

---

It gives you **freedom** to process, analyze and plot your data as you want

---

It will help you **demystify** bioinformatics

---

It will facilitate **communication** with bioinformaticians

---

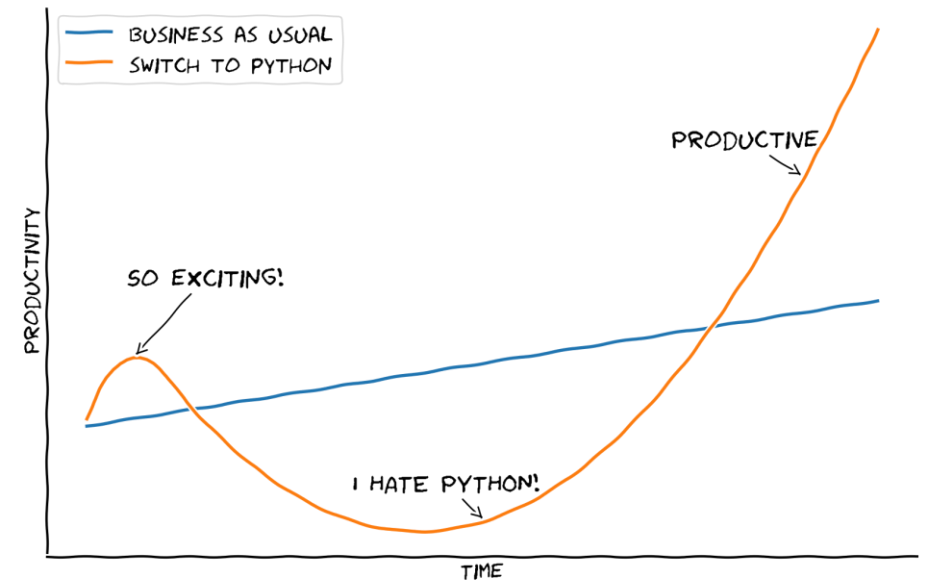
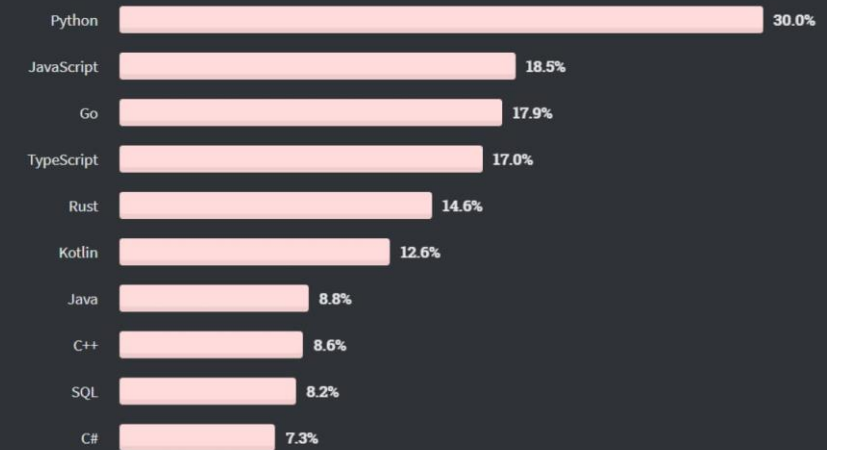
It will **improve** your **CV**

---

# Why Python?

- Python is **easy to use, powerful, and versatile**
- A great choice for **beginners and experts** alike
- Python's readability makes it a great **first programming language**
- It has a **huge community** behind developing useful libraries in many different fields (i.e biology, imaging, etc.)

The most wanted top programming languages





# Basic concepts

What's programming?

Variables, Functions

How do you approach a problem?

Car example

Class, Object, attribute, functions

# What's programming

- Programming is a way of **communicating** with a device: computer, cellphone, ..., machine of any kind
- This communication is possible if you speak a **language** that the machine understands
- Programming is **not difficult**, mastering it might be a bit more **challenging**



# Variables

*variable\_name = value*

**Variable** is a way of **storing values** that you want to use later

- To define variables, we use **name** of variable and '=' to assign values:

*my\_first\_variable = 3*

- Variables can have **different types**:

*my\_first\_variable = 3 # integer*

*my\_second\_variable = "This is my second variable" # string*

*my\_third\_variable = 3.0 # float*

# Functions

*def function\_name (parameters)*

**Function** is the way to **define actions**, i.e sum, print on the screen.

To define functions, we use the **reserved word def**:

```
def sum_two_numbers(a, b):
```

```
    return a + b
```

```
def say_hi():
```

```
    print("Hi")
```

Functions can be **called** by their **name** and specifying the **parameters**:

```
sum_two_numbers(a=7, b=5)
```

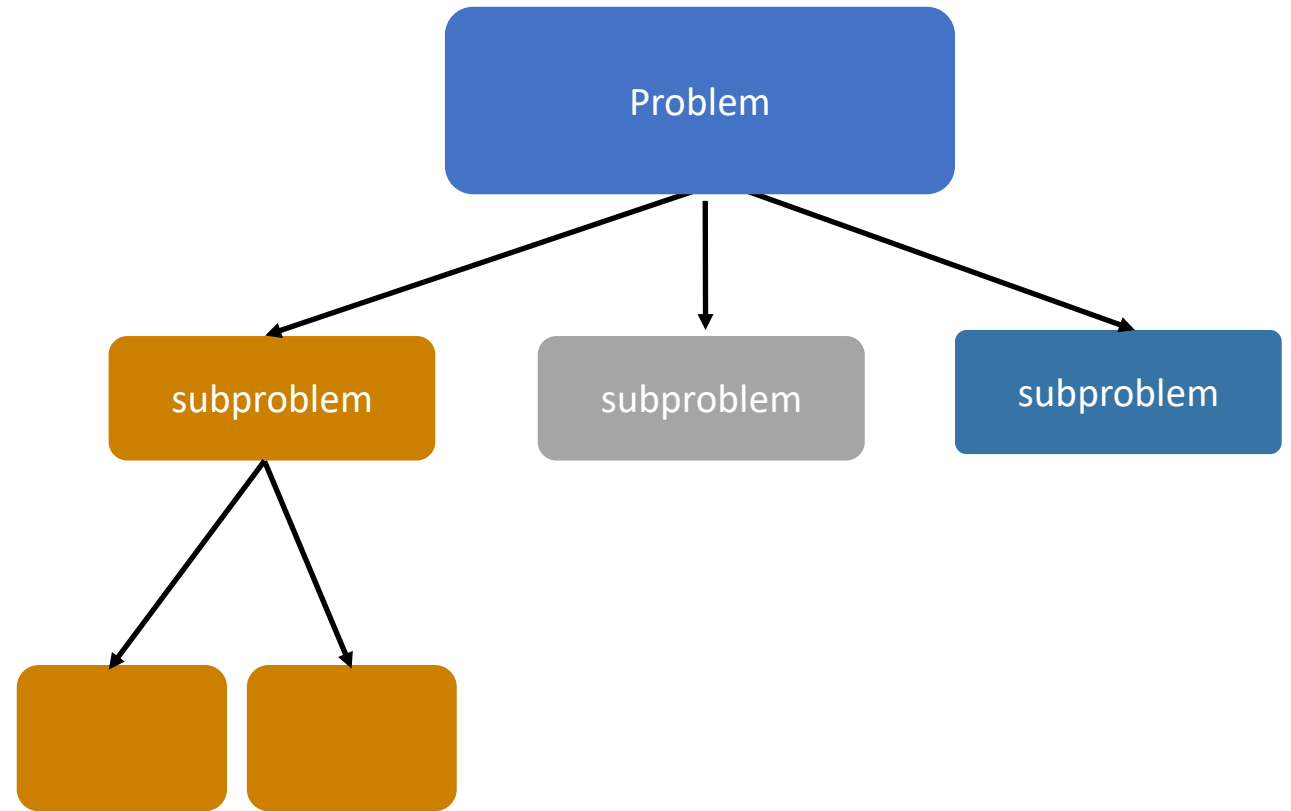
```
> 12
```

```
say_hi()
```

```
> Hi
```

# Strategy for Programming

*Divide and conquer*



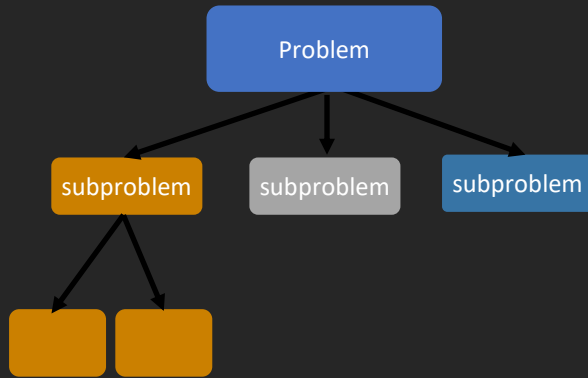




# The Car Problem

*Describe a car*

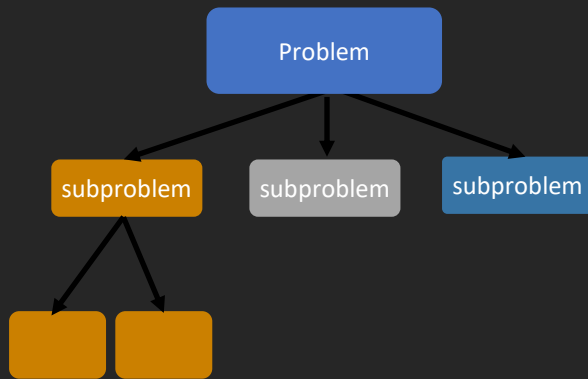
- Describe this **object**:
  - **Parts**: wheels, a stirring wheel, a frame, etc.
  - **Actions**: moves, breaks, etc.





# The Car Problem

*Describe a car*



- Describe this **object** → **Class**:
  - **Parts**: wheels, a stirring wheel, a frame, etc. → **variables or attributes**
  - **Actions**: start, change gear, etc. → **functions**

## Variables:

```
color = "blue"  
number_of_wheels = 4  
motor = True  
power = "gas"  
gear = None
```

## Functions:

```
def start_engine():  
...  
def change_gear(gear):  
...
```

# Coffee with your team

- **Introduce** yourself and **what you do**
- Explain your **motivation** to take the course
- Discuss what **data** could be relevant for you



