Assignment II

Programming with Python

Introduction

This is the second of the two assignments for this course. It will be graded as pass/fail and you and your group of up to 3 students will need to submit it before the start of the last lecture. Please hand in one assignment per group via the e-mail address vlcek@beyondsimulations.com.

In the assignment, you will practice the concepts you learned in the second part of the course. You will find a lot of information online and you are welcome to use generative AI to help you with the assignment. However, you are not allowed to copy the code from other groups and you have to indicate where and how you used AI to help you. Try to use comments to structure and to explain your code. Furthermore, use descriptive variable names and format your code nicely in order to make it more readable.

Redact secret information

In this excercise, you will create a program that is able to redact secret information in a text. The program should be able to redact the following: **zip codes, names, email addresses and phone numbers.** The program should ask the user **for a filename** and then read the file and **redact the secret information**. The program should then print the redacted text to the console and write it to a new file called `redacted.txt`. You can find a file with secret information in the git repository under `assignments/secret-text.txt`.

```
# Secret information redactor
# TODO: Create a program that is able to redact secret information in a text.
# Your code here
```

Dice roll simulator

In this excercise, you will create a program that is able to simulate dice rolls and visualizes their distribution. The program should ask the user for **the number of sides on the dice and the number of dices to roll**. Then, the program should simulate the dice rolls 10000 times and **visualize the distribution** of the dice rolls using a histogram.

```
# Dice roll simulator
# TODO: Create a program that is able to simulate a dice roll.
# Your code here
```

Regression analysis

In this excercise, you will create a program that performs linear regression on a dataset with housing prices. The dataset can be found in the git repository under `assignments/hamburg_housing_prices.csv`. The dataset contains the following columns:

- Price: The price of the house in Euros.
- Area: The area of the house in square meters.
- Rooms: The number of rooms in the house.
- Year_Built: The year the house was built.
- Distance_to_Center: The distance to the center of Hamburg in kilometers.

Your task is to perform a **linear regression analysis** on the dataset and to predict the price of a house based on its area, number of rooms, year it was built and its distance to the center of Hamburg. You can use the `pandas` library to read the dataset and the `numpy` library to perform the linear regression analysis. After performing the linear regression analysis, use the `matplotlib` library to visualize the relationship between the price and the other features as well as the predicted prices for a given set of features.

```
# Regression analysis
# TODO: Create a program that performs linear regression analysis on the dataset.
# Your code here
```