Dr. Sandro Andreotti

Programming for Data Science

Winter term 2023/2024

Assignment 8
Issued: 2023-10-26

1   NumPy masking

A leap year is each year that is a multiple of 4 (except for years that are a multiple of 100 but not of 400). Given a NumPy array of integers (years) compute:

- The number of leap years.

- The largest and smallest leap year.

- (only) The leap years sorted in reverse (descending) order.

2   Pandas basics

From a list of dictionaries create a Pandas DataFrame containing information on persons. These include (to be column names):

- first_name

- last_name

- phone

- zip_code

- age (in years)

- height (in cm)

- weight (in kg)

With the resulting DataFrame perform the following tasks:

1. Get the first_name, last_name and age (in this order) for all persons younger than 28.

2. Get the last_name, first_name, height, age (in this order) for all persons younger than 28 AND taller than 175 cm.

3. One year has passed, time to update your table! Increment the age of all persons by one year.

4. Add a new column (bmi) to your DataFrame which contains the body mass index calculated from the height and weight for each person.

5. Berlin's zip codes are in the range 10115 - 14199. Return first_name, last_name and zip_code for all persons that appear to live in Berlin.