

Test your skills - python, numpy and pandas

Task 1

Let us define a simple list:

```
x = list(range(5))
```

Now, perform the following operations without a computer (on a paper) and then check if the results are consistent with python. Should some operation raise an Error, explain why does this happen.

a

```
x[3:4]
x[:-1]
type(x[3:4])
len(x[3:4])
```

b

```
x2 = x[0:5]
x2[1] = 10
sum(x2)
sorted(x[::-1])
x3 = x
x3[1] = 10
sum(x3)
tmp = x.sort()
tmp
x
```

c

```
def fun(x):
    x[0] = 10
    return x[:2], x[2:]

tmp = [1, 2, 3]
result = fun(tmp)
sum(tmp)
result
result[0] = result[1]
result[1][0] += 3
result
```

d

```
[str(e) + '0' if e < 5 else e for e in range(10) if e > 3]
```

e

```
for i, e in enumerate(x):
    print(f'{i:02d}:{e:3d}')
```

f

```
z = {'a' : 1, 'b' : 2}
z[1]
z['a']
z['c']
z.get(1)
z.get('c')
z['c'] = 3
z.keys()
```

Task 2

Create a vector **x** (1 dimensional `numpy.array`) with 20 random values.

- Divide **x** into two vectors **x1** – consisting of all values greater than the mean value of **x** and **x2** – consisting of values smaller or equal to the mean.
- Standardize vector **x**, i.e. transform its values so that the mean of a new vector is equal 0 and standard deviation is equal to 1.

Task 3

Create a random matrix **x** with 4 rows and 5 columns.

- Standardize the columns of **x**.
- Standardize the rows of **x**.
- Find the maximal value in each row of **x**.

Task 4

Load `grades.csv` file into `pandas.DataFrame`.

- Select only records for which a column `school` has value `SP1`.
- Find the average grade for `math` and `eng` in the whole school.
- Find the average grades for each `class` for each school subject.
- Sort the frame obtained in (c) by the average `math` grade.
- Add a new column to the initial data frame indicating if a student's math grade is above the average grade in school or not.