

## HOMEWORK 10

Each question 1-2-3 is worth 20 points. Question 4 is worth 40 points.

When working on this homework, please make sure to include markdowns at appropriate places. Think of this as a report to your manager or advisor. You want the report to look professional.

**Problem 1.** Given the following numpy array

```
my_arr = np.array([2, 3, -4, 5, 7, -5, 10])
```

Use boolean indexing to answer the following questions.

- (1) Create a new array that contains only the positive elements from the original array.
- (2) Create a new array that contains only the elements that are negative and even from the original array.
- (3) Create a new array that contains elements that are multiple of 5 or multiple of 3.
- (4) How many elements are even, positive, and less than 20?

**Problem 2.** The following numpy array represents the daily electricity usage (in kWh) for a week

```
daily_use = np.array([10, 7, 14, 9, 11, 8, 10])
```

- (1) Find the mean, the standard deviation, and the range of this data set.
- (2) How many days is the usage above 10?
- (3) Assuming a rate of \$0.12 per kWh, find the cost for each day.
- (4) Find the total cost for this week, assuming that the customer has to pay a 10 % tax as well.

**Problem 3.** The following numpy array represents the daily step count for a week

```
steps = np.array([7000, 8500, 9000, 7500, 10000, 6500, 8000, 9000, 8000, 7000])
```

- (1) Calculate the total steps for the week.
- (2) Finds the average daily steps.
- (3) How many days does the user meet the goal of 8000 steps.
- (4) Create a line plot and a box plot for this data set.

**Problem 4.** A biologist is studying red maple trees in the Chicago area. He recorded the heights (measured in feet) of these trees in a numpy file named tree\_heights.

- (1) How many trees did the biologist record?
- (2) Calculate the mean, the standard deviation, and the range of this data set.

- (3) Create a histogram and a boxplot for this data set.
- (4) What percentage of trees are taller than 60 feet?
- (5) The biologist wants to use the tree's height to (approximately) calculate its diameter. The model that he uses is

$$\text{diameter} = 0.54 \log(\text{height}) + 0.37.$$

Here  $\log$  is the logarithm function. In Python, you can calculate by using `np.log`. Use this model to calculate the diameter of each tree in the dataset.

- (6) Create a histogram for the diameter data.
- (7) Estimate the volume of each tree in the dataset, assuming each tree has the shape of a cylinder. The volume can be calculated using the formula:

$$V = \frac{\pi}{4}(\text{diameter})^2 \times (\text{height}).$$

- (8) Create a boxplot for the volume.