

CS 115 - Introduction to Programming in Python

Lab Guide 8

Lab Objectives: Data visualization – matplotlib.

IMPORTANT:

For the questions below:

- Do not use any loops. All data extraction can be done using Boolean/relational indexing.
- You may assume that the indicators/countries are sorted by GDP in ascending order.
- You may not assume the set of countries or the position of the country in the files.
- You may only use functionality discussed in CS115.

TIPS:

- `np.set_printoptions(suppress=True)`: displays the float values in the console with n decimal places, suppresses the scientific/exponential notation.
- The subplot function returns an axis object. The hist function returns a tuple containing the counts and bin edge values of a histogram. You can set the `xtick` marks of the histogram to the edge values using the following statements:

```
ax = plt.subplot(...)
data = plt.hist(...)
ax.set_xticks( data[1] )
```

1. Download the files: `indicators.txt` and `countries.txt`, the data in the files is parallel, meaning the name and region of a country in the `countries.txt` file is at the same position/index as the indicator data in `indicators.txt`.
2. Create a Python script that does the following:
 - a. Using `loadtxt`, load the country/region data into a numpy array, `countries`, and the indicator data into a numpy array, `indicators`.
 - b. Select all rows from `indicators` where the GDP is below 1000 and store as `indicators_under_1000`.
 - c. Store the country/region data (`europe_countries`) and indicator data (`europe_data`) for all countries in Europe.
 - d. Select the male and female employment indicators for Turkey and store as `turkey_employment_data`.
 - e. Using the data from the appropriate numpy arrays, create the plots shown below.

