
Energy Informatics


<https://proglang.informatik.uni-freiburg.de/teaching/energy-informatics/2018ws/>

Exercise Sheet 5 – Sequences and data

2018-11-14

Exercise 1 (Movie database – Plotting)

In exercise sheet 4, we analyzed a movie database. We now use `matplotlib` to draw some associated plots.

 **Mathplotlib**

Mathplotlib is a generic and powerful plotting library for python. Many tutorials and documentation are available on <https://matplotlib.org/tutorials/index.html>.

```
import matplotlib.pyplot as plt

xdata = [ ... ]
ydata = [ ... ]

plt.plot(xdata,ydata,"r-") # plot the data in red dashes
plt.axis([0, 6, 0, 20]) # Delimit axis on 0-6, 0-20

plt.xlabel('data x') # Add some labels
plt.ylabel('data y')

plt.show() # Show the plot!
```

Draw the plots corresponding to the movie database questions. Propose additional question, and draw the appropriate plots.

Exercise 2 (Caesar cipher)

In caesar cipher, the text is encrypted by shifted the letters by a specified amount (which is the key). For example, if the key is 14, 'a' becomes 'o'.

Write two functions that takes a key and encrypt and decrypt text typed by the user. The code corresponding to a character is given by the `ord` function.

Exercise 3 (Breaking the cipher)

Caesar cipher can be broken by analysis the frequency of the letter in the text, and comparing them with known frequencies for the chosen language. The frequency of letters in the english language is given in the file `freq.csv`.

1. Write a function that counts the frequency of letters in a text. Try this function on various text found online. You can also try to plot the distribution obtained.
2. Write a function that computes the distance between two distribution of frequencies:

$$d(\bar{x}, \bar{y}) = \sqrt{\sum_i (x_i - y_i)^2}$$

3. Write a function to decrypt an arbitrary text encoded with Caesar cipher by comparing with the english distribution, shifted any number from 0 to 25.