HOMEWORK

NumPy

1. **How to get the common items between two arrays?**

   a = np.array([1,2,3,2,3,4,3,4,5,6])

   b = np.array([4,7,15,3,5,6,9,1,12,14])

2. **How to remove from array 'a' the items that exist in 'b'?**

   a = np.array([1,3,5,7,9])

   b = np.array([2,4,6,8,10])

3. **Replace all the even numbers in array 'a' with the new value '-2':**

   a = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9])

4. **Swap the last two rows of the following array:**

   a = np.arange(25).reshape(5,5)

matplotlib

5. **Draw a scatter plot with the following data: s_x and s_y**

   s_x = range(1, 25)

   s_y = [1 / x for x in s_x]

   s_z = [1 / (25 - x) for x in s_x]

   1. Overlap in this plot the additional dataset s_z, and add a proper legend to the plot, axes labels and a title
   2. Generate 2 subplots in a figure using datasets s_y and s_z, and add a proper legend to the plots and a title

pandas

6. **Prepare the plot "Root Mean Square Fluctuation vs Residue Number" (Potasium Ion Channel Kv1.2) by using the following input file 'rmsf.xvg'.**

   Here we propose some steps for the exercise:

   1. read the input file and:
   - skip the lines that start with # or @
   - rename columns as Residue_Number and RMSF passing the name of columns from an external list
   2. inspect the created dataframe
   3. show a snapshot of dafaframe showing a set of rows randomly selected
   4. get basic statistics about the loaded data
   5. get the datetype for each elements in colums
   6. add a new colum with the $(RMDF)^2$ and save a new dataframe
   7. make the plot of residue_name vs RMSF and save a jpeg pic