Aditya Pitchuka  
INFO 5502.002 (14676)  
Spring 2020

**Assignment – 1**

Write a Python program that can take different numbers of strings (each string may include digits, characters, and special symbols) and then sorted them. You need first define the rule for sorting and then implement the sorting function using Python (DO NOT use any existing sort function). Visualize a list of input strings (the list must include at least 500 strings) using a scatter plot (one dimension will be the length of each string, and the other dimension could be the order of the string in the list).

**Input** 🡪 Reading a file which contains strings and storing into a list.

For sorting the list of strings, I am using **Selection sort** technique.

**Behaviour of technique**  
It virtually divides the list into unsorted and sorted and then sorts the rest of the list. In this technique first it starts by comparing the first element with the rest of the list for the least value’s index. Once it is found, it swaps the value with the starting element in the list. This means, the least element will be sorted first. And then it checks the second element with the rest of the unsorted list and finds for the least value’s index. On finding it, swapping will be done and the process will follow-on.

**Sample Illustration of technique:**

Consider a list of numbers which are unsorted as  
list\_numbers = [4,7,1,3,0,5,2,6]

1. Least\_index = 0 and its value is 4.  
   Now the comparison’s starts from 7 to run till the end. It breaks at ‘0’ as it is lesser value than our key and its will be swapped with it. So, the list will be:  
   [0,7,1,3,4,5,2,6]
2. Now the index is at 1 and value is 7. Comparison starts from the next position of index i.e., 2 and the its corresponding value is 1. As it is lesser, swapping will be done.

[0,1,7,3,4,5,2,6]

1. Index is at 2 and the comparison will be done till the end of the list. Once the least element is found, swapping is done.

[0,1,2,3,4,5,7,6]

1. Rest of the patterns follows on, As the list already looked sorted (in this selected case) the index will moved until the condition breaks. This happens at the index 6 and the corresponding value is 7 and comparison will be with rest of the list, only one element in this case i.e., 6. Swap will be done.

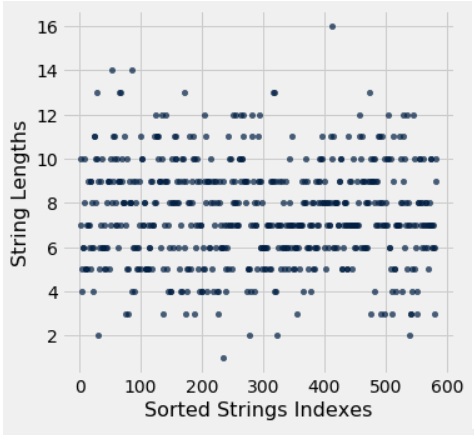
[0,1,2,3,4,5,6,7]

Once the sorting is done, will take the lengths of each string in sorted list and save in another list which is used for plotting a graph used as Y-axis.   
On the X axis we will use the order of the strings for which we will define another list.

Table is defined with these two lists as columns and then it is used to plot a graph.

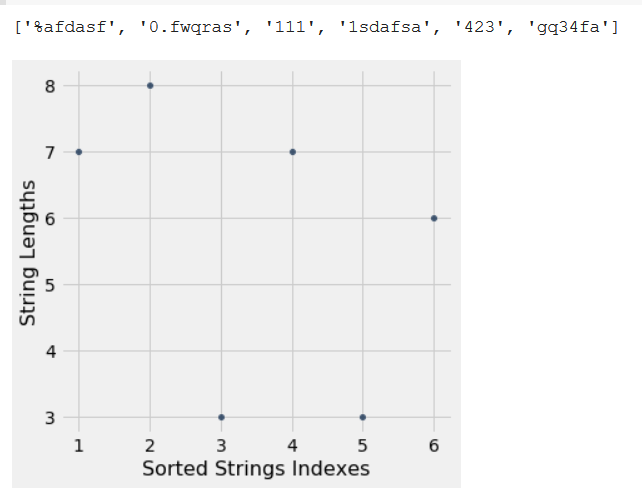
**Visualization Results:**

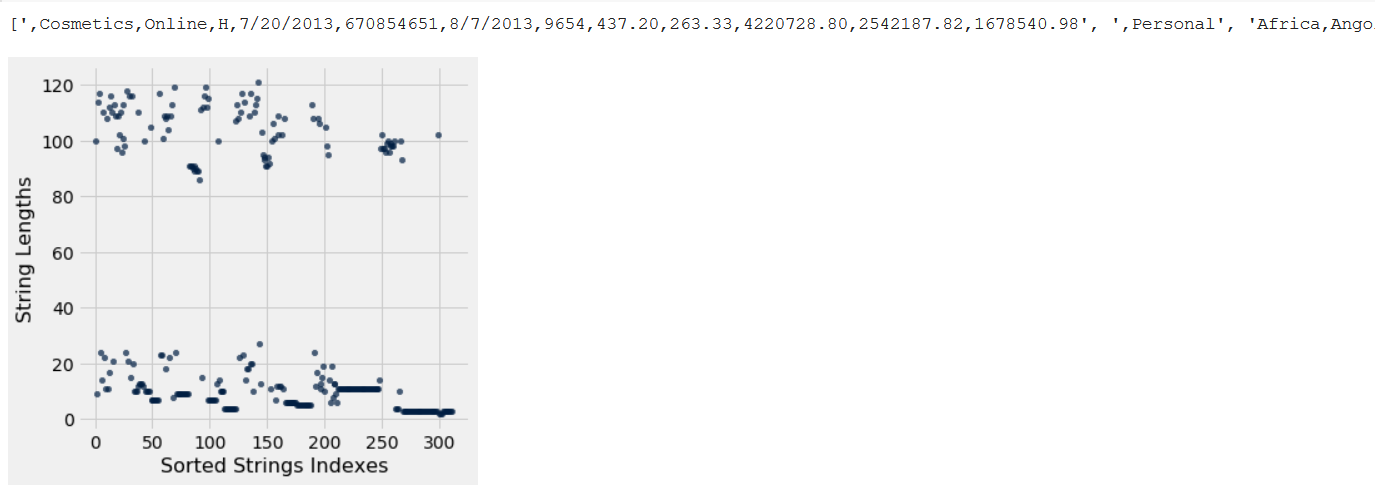
For the input file,  the visualization scatter graph is like this.



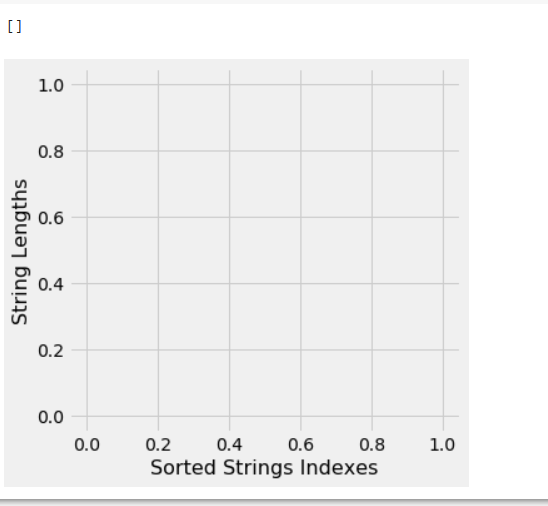
**Test Cases**

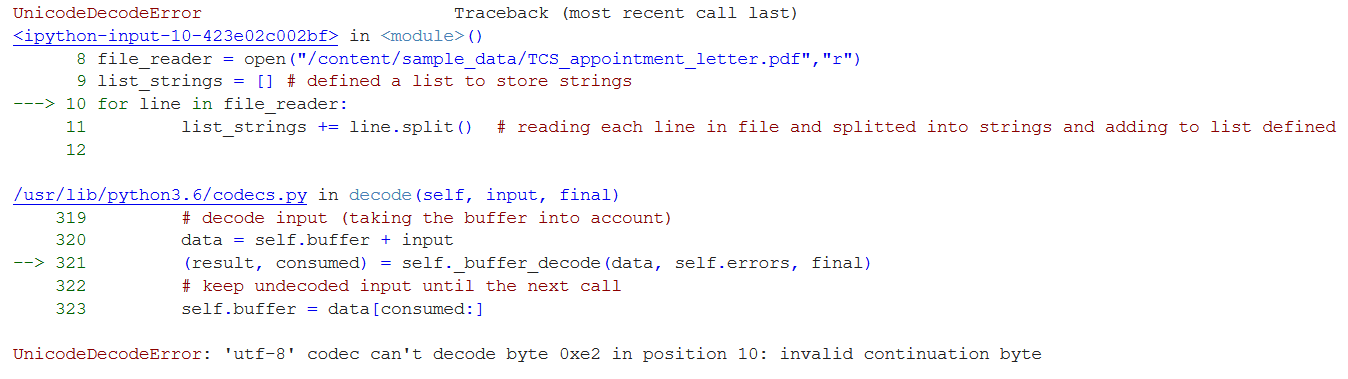
For the below input list of strings, output will be like:  
['1sdafsa','%afdasf','423','111','0.fwqras','gq34fa']

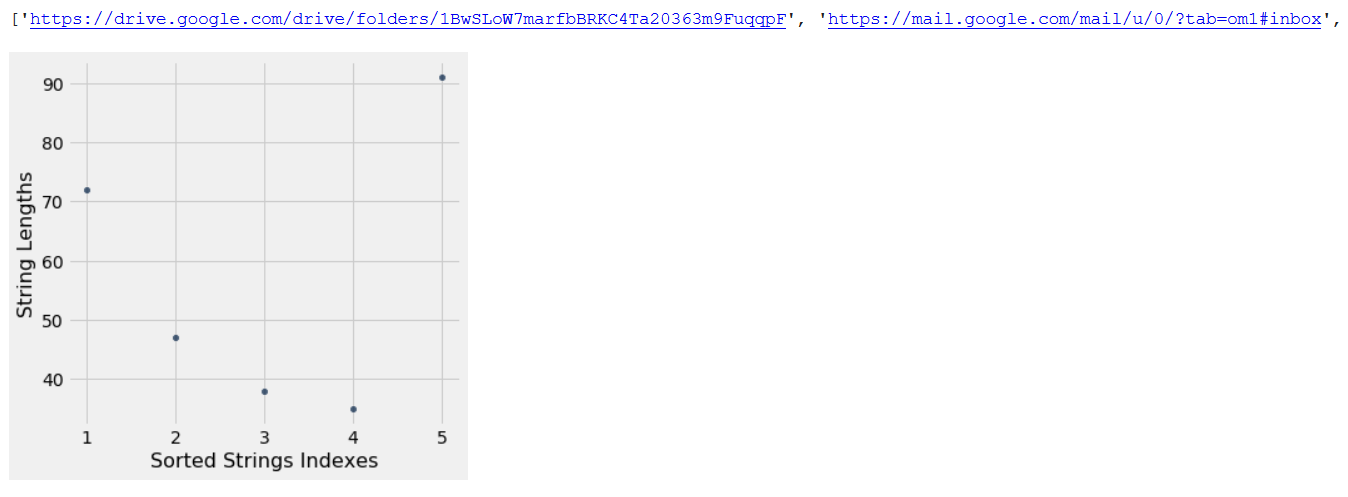


For a sample csv file type input, we see below output.

For an empty file/list.



When uploading a pdf file, it throws an error.

When URLs are entered.