Lab8 Exercises

Q1) This is an employee's (doctors) data that works on hospital, Convert the following to data frame then sort the data according to the names in ascending order and if they are equals sort them according to id then clean the data based on your analysis and understand for the data.

input:

```
data = {'names': ['ahmed', 'mohammed', 'anas', 'foad', 'loay', 'gamal', 'ahmed'],
 'ids': ['987643776', '846993756', '778354653', '892347346', '312278543', '773536473', '783526234'],
 'ages': [62, 27, 20, 50, 43, 22, 34],
     'number of children' : [4, 2, 1, 4, 5, 1, 3],
     'years of service':[30, 2, 1, 22 , 15, 1, 12]}
```

output of the first requirement (sorting) :

	names	ids	ages	number of children	years of service
0	ahmed	987643776	62	4	30
1	mohammed	846993756	27	2	2
2	anas	778354653	20	1	1
3	foad	892347346	50	4	22
4	loay	312278543	43	5	15
5	gamal	773536473	22	1	1
6	ahmed	783526234	34	3	12

Q2) This is a products data, convert the following to data frame then clean the data based on your understanding of the data, after that sort the data by the "sold quantity per month" column descending order.
Show me the products names and the remains quantity of the products in ascending order.
Show me what are a 3 products that I should call the storekeeper to give me more quantity, and is there a product that I should stop to ask from the storekeepers. To do this you should show a percentage of the purchase for the product and any product have a purchase under the 30% should stop request it.

Input:

```
data = {'names': ['sugar', 'milk', 'soda', 'chocolate', 'meat', 'yogurt', 'cake'],
     'price': [15, 32, 3, 6, 35, 4, 12],
     'quantity': [150, 243, 300, 60, 30, 456, 470],
     'sold quantity per month' : [120, 150, 313, 58, 8, 359, 60]}
```

Q3) Check if the indices is unique or not then get the index of the max value then show this information about the series

input:

```
pd.Series(range(5), index=['a', 'a', 'b', 'b', 'c'])
```

output:

```
count    5.000000
mean     2.000000
std      1.581139
min      0.000000
25%      1.000000
50%      2.000000
75%      3.000000
max      4.000000
dtype: float64
```

Q4) For the given excel file get the maximum value and the minimum value of the SepalLengthCm then sort the data according to the SepalLengthCm in descending order after that print the correlation between SepalLengthCm and PetalLengthCm then get the mean of SepalLengthCm values at the end get the count of Virginica species