

Q1. Write codes for following executions

- append a list to the second list
 - access the index of a list
 - difference between the two lists
 - removing the 0th, 4th and 5th elements. Sample List : ['Red', 'Green', 'White', 'Black', 'Pink', 'Yellow']
- Counts the number of occurrences of item 50 from a tuple

Q2. Write Numpy code to generate a random list of 100 integers (range of 55 to 150) and identify the index of the largest element and smallest element. Change this list into a 10 x 10 matrix and replace all diagonal elements with 1

Q3. Create a DataFrame of 7 rows and 7 columns containing random integers in the range of 1 to 100. Compute the correlation of each row with the preceding row.

Consider the following DataFrame **EXERCISE** to answer the given questions where 'Kind' attribute indicates the type of exercise regime followed.

ID	Name	Diet	Pulse	Time (min)	Kind
0	A	low fat	85	40	walking
1	A	low fat	85	45	walking
2	A	no fat	88	30	running
3	B	no fat	90	10	walking
4	B	no fat	92	15	rest
5	B	low fat	93	30	rest
6	C	low fat	97	15	rest
7	C	low fat	97	15	rest
8	C	low fat	94	30	walking
9	D	low fat	80	10	walking
10	D	low fat	82	15	rest
11	D	low fat	83	30	rest
12	E	no fat	91	10	rest
13	E	low fat	92	15	running
14	E	low fat	91	30	running

- Assuming the data is stored in a csv file "Exercise.csv", give appropriate commands to read this file, indexed on 'Name' and 'Diet' into a dataframe named EXERCISE. Modify this command to read only the first 5 rows of the file. If the file contains millions of records, then give the command to read the file in small pieces of uniform size.
- Differentiate between qcut and cut methods. Use the appropriate method to create 4 bins on the 'Pulse' attribute. Store the corresponding bin value of 'Pulse' attribute as a new attribute 'Pbin' in the original DataFrame. Display the count of values of each bin.

Q4. For the following DataFrame write commands for following instructions

	ord_no	purch_amt	ord_date	customer_id	salesman_id
0	70001	150.5	?	3002	5002
1	NaN	270.65	2012-09-10	3001	5003
2	70002	65.26	NaN	3001	?
3	70004	110.5	2012-08-17	3003	5001
4	NaN	948.5	2012-09-10	3002	NaN
5	70005	2400.6	2012-07-27	3001	5002
6	--	5760	2012-09-10	3001	5001
7	70010	?	2012-10-10	3004	?
8	70003	12.43	2012-10-10	--	5003
9	70012	2480.4	2012-06-27	3002	5002
10	NaN	250.45	2012-08-17	3001	5003
11	70013	3045.6	2012-04-25	3001	--

- i. to detect missing values of a given DataFrame. Display True or False.
- ii. to count the number of missing values in each column of a given DataFrame
- iii. to drop the rows where at least one element is missing in a given DataFrame
- iv. to keep the rows with at least 2 NaN values in a given DataFrame
- v. to replace NaNs with median or mean of the specified columns in a given DataFrame

Q5. For the given DataFrame write commands for loc() & iloc()

	Weather	Temperature	Wind	Humidity
Day				
Mon	Sunny	12.79	13	30
Tue	Sunny	19.67	28	96
Wed	Sunny	17.51	16	20
Thu	Cloudy	14.44	11	22
Fri	Shower	10.51	26	79
Sat	Shower	11.07	27	62
Sun	Sunny	17.50	20	10

- i. Selecting via a single value
- ii. To get all rows
- iii. Multiple rows for a columns
- iv. Slicing column labels
- v. Selecting rows if humidity>50