**"Note: It's okay if you don't know about some topics. Just search and try to give your best answer."**

For statistics questions this may help : [link](https://www.youtube.com/watch?v=k3aKKasOmIw&list=PLU5aQXLWR3_yYS0ZYRA-5g5YSSYLNZ6Mc&index=2&t=26s), [link](https://www.youtube.com/watch?v=2oJldeE4JcU)

Main Questions :

Q1) What is the difference between mean, median and mode?

mean:It is the average of the values ​​(sum over their number)

median:It is the number in the middle of the numbers after arranging them in ascending or descending order.

If there are two numbers in the middle, the median = the mean of these two numbers

mode:It is the number that appears the most among the numbers

Q2) What is the difference between left and right skewness of the data?

When the distribution is asymmetric, the distribution curve is skewed.

It is right skewed when the curve is longer on the right side of its peak, and it is left skewed when the curve is longer on the left side of its peak.

Q3) What is the difference between list and tuples in Python?

list and tuple in Python are the classes of Python Data Structures. The list is dynamic, whereas the tuple has static characteristics.

tuple is faster and tuple does not support item assigment but list support it.

Q4) How can you generate random numbers in Python?

using randint function from random module and it takes the lower and maximam value in range

**like:**

import random

random\_variable=random.randint(1,100)

Q5) What is break, continue and pass in Python?

**break:** used to exit from loop when a condition is met

**continue:** used to end the current iteration in a loop and continues to the next iteration

**pass:**used as a placeholder for future code to avoid errors from empty loop

Figure 1 :



Q6) Find the median of the data in Figure 1.

0,0,0,0,1,2,2,**3**,4,4,4,5,5,6,7

median=3

Q7) Find the standard deviation of the data in Figure 1.

mean=(1+2\*2+3+4\*3+5\*2+6+7)**/**15=**2.86667**



**standard Deviation=** 2**.**356349073

Figure 2 :



Q8) Which of the following best describes the data in Figure 2? (Base your answer on the appearance of the histogram.)

You do not need to do any calculations. Select just one statement below and complete the one you select.

a) The mean is greater than the median because

The figure represents right skewed, in this case most of the data points are on the left, and in this case the median is located at the left of the mean, i.e. the median is smaller than the mean

b) The median is greater than the mean because

c) The mean and median are roughly equal because

Q9) Given a list of integers numbers = [3, 7, 2, 8, 1], write a Python code snippet to find and print the maximum value in the list.

<https://ideone.com/GCVtMz>

Q10) What is an outlier in a dataset, and how would you explain it to someone who is new to statistics?

It is a value that is much greater or much less than the rest of the values, which has a negative impact and leads to incorrect conclusions. For example, if there was a company of six people and I wanted to calculate the average salaries in the company and the salaries of the six people were as follows:

1000,1500,2000,3000,2500,100000

Here 100000 is very much larger than the rest of the values, which will make the value of the mean equal to 168333.3333 is a number that does not express the true reality at all.

PS Problems :

[https://codeforces.com/problemset/problem/ HYPERLINK "https://codeforces.com/problemset/problem/791/A" HYPERLINK "https://codeforces.com/problemset/problem/791/A" HYPERLINK "https://codeforces.com/problemset/problem/791/A" HYPERLINK](https://codeforces.com/problemset/problem/791/A)

solution =><https://ideone.com/ZiU6W8>

[https://codeforces.com/problemset/problem/ HYPERLINK "https://codeforces.com/problemset/problem/58/A" HYPERLINK "https://codeforces.com/problemset/problem/58/A" HYPERLINK "https://codeforces.com/problemset/problem/58/A" HYPERLINK](https://codeforces.com/problemset/problem/58/A)

solution =><https://ideone.com/2gsS0F>

Bonus :

[https://codeforces.com/problemset/problem/ HYPERLINK "https://codeforces.com/problemset/problem/1624/A" HYPERLINK "https://codeforces.com/problemset/problem/1624/A" HYPERLINK "https://codeforces.com/problemset/problem/1624/A" HYPERLINK](https://codeforces.com/problemset/problem/1624/A)

solution =><https://ideone.com/pFXEMT>

Q1) What is the difference between machine learning and traditional programming?

**traditional programming:** programmer write the rules for the programm these rules dictate how the computer will process the input to produce output.

**input + rules => output**

**machine learning:** programmer trains a module The programmer gives the program the inputs and the resulting outputs, and then the program sets the rules according to which it will set the rules it will use to process the input to produce output.

**input + output => rules**

Q2) State some examples of machine learning models ?

**Supervised learning**

**Unsupervised learning**

**Random forest**