**"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?

Answer1) The mean is the average of data set, it is found by the sum of all values in data set and divide the sum by the number of values in that set. The median is the middle value when we order values ascending or descending. But if the number of values is even, we add the middle two numbers and divide by 2. The mode is the most frequent value in data set.

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

Answer2) Left skewness means that the top of distribution is on the left side than its right. In another meaning, the curve shape of distribution is on the left side. Right skewness means that the top of distribution is on the right side than its left. In another meaning, the curve shape of distribution is on the right.

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

Answer3) List and tuple are the same thing , a collection of various data types. But the list is mutable and tuple is immutable because tuple object doesn’t support item assignment.

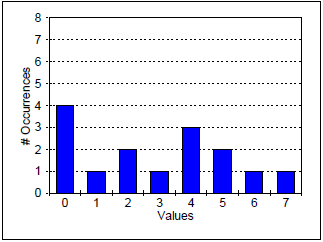
Q4) How can you generate random numbers in Python?

Answer4) By using Random library , by keyword (import random) and use its functions.

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

Answer5) Break stops and terminates the loop if a certain condition is met. Continue skips the code inside the loop if a specific condition is met and then completes the rest of loop. Pass doesn’t do anything.

Figure 1 :



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

Answer6) n= 4+1+2+1+3+2+1+1=15 .

After calculating the frequency of the data , we find that the value of 3 is the 8th repetition.

The median of the data=3

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

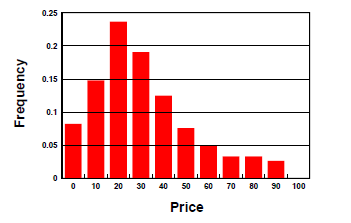
Answer7)

The mean==43/15=2.86



The standard deviation==1.42

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 that’s a right skewed distribution and the values in the tail affect the mean more than the median.

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.

Answer9)

n=[3,7,2,8,1]

for l in n:

    print(l)

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

Answer10) Outlier is a value that differs from the rest of the data. To explain, say that if the number of sales per day is in range(10-30) and on day it became 80. That’s an outlier.

PS Problems :

<https://codeforces.com/problemset/problem/791/A>

a = int(input("Enter the weight of Limak: "))

b = int(input("Enter the weight of Bob: "))

for i in range(1, b+1):

    a = a \* 3

    b = b \* 2

    if b < a:

        print("Limak became larger than Bob after %d years"%(i))

        break

<https://codeforces.com/problemset/problem/58/A>

Bonus :

<https://codeforces.com/problemset/problem/1624/A>

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

Answer) Machine learning depends on a dataset completely, it learns patterns and relations from the data and enable to make decisions without programming for each possibility (input + output = program). Traditional programming relies on what programmer writes instructions to the computer to follow(input + program = output).

Q2) State some examples of machine learning models ?