PYTHON WORKSHEET

Q1 to Q7 have only one correct answer. Choose the correct option to answer your question.

1. which of the following is not a core datatype in python? A) list B) struct C) tuple C) set
2. which of the following is an invalid variable name in python? A) *init* B) no_1 C) 1_no D) _1
3. which one of the following is a keyword in python? A) in B) *init* C) on D) foo
4. In which of the following manner are the operators of the same precedence executed in python? A) Left to Right B) BODMAS C) Right to Left D) None of these
5. Arrange the following in decreasing order of the precedence when they appear in an expression in python? i) Multiplication ii) Division iii) Exponential iv) Parentheses A) iii – iv – ii – i B) iii – iv – i – ii C) iv – iii – ii – i D) iii – ii – i – iv
6. (28//6)**3/3%3 = ? A) 7.1111… B) 0 C) 0.3333… D) 1
7. a = input("Enter an integer"). What will be the data type of a? A) int B) str C) float D) double Q8 and Q10 have multiple correct answers. Choose all the correct options to answer your question.
8. which of the following statements are correct? A) Division and multiplication have same precedence in python B) Python's operators' precedence is based on PEDMAS C) Python's operators' precedence is based on VBODMAS D) In case of operators' having the same precedence, the one on the left side is executed first.
9. which of the following is(are) valid statement(s) in python? A) abc = 1,000,000 B) a b c = 1000 2000 3000 C) a,b,c = 1000, 2000, 3000 D) a_b_c = 1,000,000 ASSIGNMENT
10. which of the following is not equal to x16 in python? A) x44 B) x16 C) x^16 D) (x4)**4 Q11 to Q13 are subjective questions, answer them briefly
11. Differentiate between a list, tuple, set and dictionary.
12. Are strings mutable in python? Suppose you have a string "I+Love+Python", write a small code to replace '+' with space in python.
13. what does the function ord() do in python? Explain with an example. Also, write down the function for getting the data type of a variable in python. Q14 and Q15 are programming questions. Answer them in Jupyter Notebook.
14. Write a python program to solve a quadratic equation of the form ax^2+bx+c=0. Where a, b and c are to be taken by user input. Handle the erroneous input, such as 'a' should not be equal to 0.
15. Write a python program to find the sum of first 'n' natural numbers without using any loop. Ask users to input the value of 'n'.

Q1. which of the following is not a core datatype in python?

A) list B) struct C) tuple C) set

ANSWER: B)struct

Q2. which of the following is an invalid variable name in python?

A) *init* B) no_1 C) 1_no D) _1

ANSWER: C) 1_no

Q3) Which one of the following is a keyword in python?

A) in B) *init* C) on D) foo

ANSWER: A) in

Q4) In which of the following manner are the operators of the same precedence executed in python?

A) Left to Right B) BODMAS C) Right to Left D) None of these

ANSWER: A) Left to Right

Q5) Arrange the following in decreasing order of the precedence when they appear in an expression in python?

i) Multiplication ii) Division iii) Exponential iv) Parentheses A) iii – iv – ii – i B) iii – iv – i – ii C) iv – iii – ii – i D) iii – ii – i – iv

ANSWER: C) iv – iii – ii – i

Q6) (28//6)**3/3%3 = ?

A) 7.1111… B) 0 C) 0.3333… D) 1

ANSWER: C) 0.3333...

Q7) a = input("Enter an integer"). What will be the data type of a?

A) int B) str C) float D) double

ANSWER: B) str

Q8) Which of the following statements are correct?

A) Division and multiplication have same precedence in python B) Python's operators' precedence is based on PEDMAS C) Python's operators' precedence is based on VBODMAS D) In case of operators' having the same precedence, the one on the left side is executed first.

ANSWER: A) Division and multiplication have same precedence in python B) Python's operators' precedence is based on PEDMAS D) In case of operators' having the same precedence, the one on the left side is executed first.

Q9. Which of the following is(are) valid statement(s) in python?

A) abc = 1,000,000 B) a b c = 1000 2000 3000 C) a,b,c = 1000, 2000, 3000 D) a_b_c = 1,000,000

ANSWER: A) abc = 1,000,000 C) a,b,c = 1000, 2000, 3000 D) a_b_c = 1,000,000

Q10) which of the following is not equal to x**16 in python?

A) x44 B) x16 C) x^16 D) (x4)**4

ANSWER: A) x44 D) (x4)4

Q11) Differentiate between a list, tuple, set and dictionary.

List, tuple, dictionary, and set are different types of collections in Python. List and tuple are ordered collections of items, while dictionary and set are unordered collections of items. List and dictionary are mutable, meaning they can be changed or replaced, while tuple and set are immutable, meaning they cannot be changed or replaced. Dictionary holds key:value pairs, while the other collections hold single values. Set has no duplicate elements, while the other collections can have duplicates. Set and dictionary are faster for checking if an item is present than list and tuple.

Q12) Are strings mutable in python? Suppose you have a string "I+Love +Python", write a small code to replace '+'

with space in python.

YES, strings are mutable in python.

The following program completes the answer

In [1]:
```
str="I+Love+Python"
s="I Love Python"
str=s
```
In [2]:
```
str
```

Q13) What does the function ord() do in python? Explain with an example. Also, write down the function for getting the data type of a variable in python.

Python ord() function returns the Unicode code from a given character. This function accepts a string of unit length as an argument and returns the Unicode equivalence of the passed argument. In other words, given a string of length 1, the ord() function returns an integer representing the Unicode code point of the character when an argument is a Unicode object, or the value of the byte when the argument is an 8-bit string.

In [3]:
```
v=ord('a')
v1=ord('z')
print(v)
print(v1)
print(type(v))
print(type(v1))
```

```
97
122
<class 'int'>
<class 'int'>
```

Q14) Write a python program to solve a quadratic equation of the form ax^2+bx+c=0. Where a, b and c are to be taken by user input. Handle the erroneous input, such as 'a' should not be equal to 0.

In [4]:
```
import cmath
a=int(input('Enter a = '))
b=int(input('Enter b = '))
c=int(input('Enter c = '))
if a != 0:
    d = (b**2) - (4*a*c)
    sol1 = (-b-cmath.sqrt(d))/(2*a)
    sol2 = (-b+cmath.sqrt(d))/(2*a)
    print('The solution are {0} and {1}'.format(sol1,sol2))

else:
    print('This is not a quadratic equation')
```

```
Enter a = 2
Enter b = 3
Enter c = 4
The solution are (-0.75-1.1989578808281798j) and (-0.75+1.1989578808281798j)
```

Q15) Write a python program to find the sum of first 'n' natural numbers without using any loop. Ask users to input the value of 'n'.

In [5]:
```python
# LOOP METHOD
n=int(input('Enter any natural number = '))
sum=0
for i in range(0,n+1):
    sum+=i
print(sum)
```

```
Enter any natural number = 8
36
```

In [6]:
```python
# FORMULA METHOD
n=int(input('Enter any natural number ='))
sum=(n*(n+1)/2)
print(sum)
```

```
Enter any natural number =7
28.0
```

STATISTICS WORKSHEET-9

Q1 to Q12 have only one correct answer. Choose the correct option to answer your question.

1. The owner of a travel agency would like to determine whether or not the mean age of the agency's customers is over
2. If so, he plans to alter the destination of their special cruises and tours. If he concludes the mean age is over 24 when it is not, he makes a ___ error. If he concludes the mean age is not over 24 when it is, he makes a __error. a. Type II; Type II b. Type I; Type I c. Type I; Type II d. Type II; Type I
3. Suppose we wish to test H0: μ =53 vs H1: μ > 53. What will result if we conclude that the mean is greater than 53 when its true value is really 55? a. We have made a Type I error b. We have made a correct decision c. We have made a Type II error d. None of the above are correct
4. The value that separates a rejection region from an acceptance region is called a ___. a. parameter b. critical value c. confidence coefficient d. significance level
5. A hypothesis test is used to prevent a machine from under filling or overfilling quart bottles of beer. On the basis of sample, the machine is shut down for inspection. A thorough examination reveals there is nothing wrong with the filling machine. From a statistical point of view: a. Both Type I and Type II errors were made. b. A Type I error was made. c. A Type II error was made. d. A correct decision was made.
6. Suppose we wish to test H0 : μ =21 vs H1 : μ > 21. which of the following possible sample results gives the most evidence to support H1 (i.e., reject H0)? Hint: Compute Z-score. a. x = 23 s , = 3 b. x = 19 s , = 4 c. x = 17

s , = 7 d. x = 18 s , = 6

7. Given H0: μ = 25, H1: μ ≠ 25, and P-value = 0.041. Do you reject or fail to reject H0 at the 0.01 level of significance? a. fail to reject H0 b. not sufficient information to decide c. reject H0 WORKSHEET

8. A bottling company needs to produce bottles that will hold 12 ounces of liquid. Periodically, the company gets complaints that their bottles are not holding enough liquid. To test this claim, the bottling company randomly samples 36 bottles. Suppose the p-value of this test turned out to be 0.0455. State the proper conclusion. a. At α = 0.085, fail to reject the null hypothesis. b. At α = 0.035, accept the null hypothesis. c. At α = 0.05, reject the null hypothesis. d. At α = 0.025, reject the null hypothesis.

9. If a hypothesis test were conducted using α = 0.05, for which of the following p-values would the null hypothesis be rejected? a. 0.100 b. 0.041 c. 0.055 d. 0.060 9 . For H1: μ > μ0 p-value is 0.042. What will be the p-value for Ha: μ < μo? a. 0.084 b. 0.021 c. 0.958 d. 0.042

10. The test statistic is t = 2.63 and the p-value is 0.9849. What type of test is this? a. Right tail b. Two tail c. Left tail d. Can't tell

11. The test statistic is z =2.75, the critical value is z = 2.326. The p- value is … a. Less than the significance level b. Equal to the significance level c. Large than the significance level

12. The area to the left of the test statistic is 0.375. What is the probability value if this is a left tail test? a. 0.750 b. 0.375 c. 0.1885 d. 0.625 Q13 to Q15 are subjective answers type questions, Answers them in their own words briefly. 13.What is T distribution and Z distribution? 14.Is the T distribution normal? 15.What does the T distribution tell us?

Q1)The owner of a travel agency would like to determine whether or not the mean age of the agency's customers is over 24. If so, he plans to alter the destination of their special cruises and tours. If he concludes the mean age is over 24 when it is not, he makes a ___ error. If he concludes the mean age is not over 24 when it is, he makes a __error.

a. Type II; Type II b. Type I; Type I c. Type I; Type II d. Type II; Type I

ANSWER: C) Type I ; Type II

Q2) Suppose we wish to test H0: μ =53 vs H1: μ > 53. What will result if we conclude that the mean is greater than 53 when its true value is really 55?

a. We have made a Type I error b. We have made a correct decision c. We have made a Type II error d. None of the above are correct

ANSWER: b) We have made a correct decision.

Q3) The value that separates a rejection region from an acceptance region is called a ___.

a. parameter b. critical value c. confidence coefficient d. significance level

ANSWER: b) critical value

Q4) A hypothesis test is used to prevent a machine from under filling or overfilling quart bottles of beer. On the basis of sample, the machine is shut down for inspection. A thorough examination reveals there is nothing wrong with the filling machine. From a statistical point of view:

a. Both Type I and Type II errors were made. b. A Type I error was made. c. A Type II error was made. d. A correct decision was made.

ANSWER: b) A Type I error was made.

Q5) Suppose we wish to test H0 : μ =21 vs H1 : μ > 21. Which of the following possible sample results gives the most evidence to support H1 (i.e., reject H0)? Hint: Compute Z-score.

a. x = 23 s , = 3 b. x = 19 s , = 4 c. x = 17 s , = 7 d. x = 18 s , = 6

ANSWER: c) x = 17 s , = 6

Q6) Given H0: μ = 25, H1: μ ≠ 25, and P-value = 0.041. Do you reject or fail to reject H0 at the 0.01 level of significance?

a. fail to reject H0 b. not sufficient information to decide c. reject H0

ANSWER: a) fail to reject H0

Q7) A bottling company needs to produce bottles that will hold 12 ounces of liquid. Periodically, the company gets complaints that their bottles are not holding enough liquid. To test this claim, the bottling company randomly samples 36 bottles. Suppose the p-value of this test turned out to be 0.0455. State the proper conclusion.

a. At α = 0.085, fail to reject the null hypothesis. b. At α = 0.035, accept the null hypothesis. c. At α = 0.05, reject the null hypothesis. d. At α = 0.025, reject the null hypothesis.

ANSWER: c) At α = 0.05, reject the null hypothesis.

Q8) If a hypothesis test were conducted using α = 0.05, for which of the following p-values would the null hypothesis be rejected?

a. 0.100 b. 0.041 c. 0.055 d. 0.060

ANSWER: b) 0.041

Q9) For H1: $\mu > \mu_0$ p-value is 0.042. What will be the p-value for Ha: $\mu < \mu_0$?

a. 0.084 b. 0.021 c. 0.958 d. 0.042

ANSWER: c) 0.958

Q10) The test statistic is t = 2.63 and the p-value is 0.9849. What type of test is this?

a. Right tail b. Two tail c. Left tail d. Can't tell

ANSWER: c)Left tail.

Q11) The test statistic is z =2.75, the critical value is z = 2.326. The p-value is …

a. Less than the significance level b. Equal to the significance level c. Large than the significance level

ANSWER: a) Less than the significance level.

Q12) The area to the left of the test statistic is 0.375. what is the probability value if this is a left tail test?

a. 0.750 b. 0.375 c. 0.1885 d. 0.625

ANSWER: b) 0.375

Q13) What is T distribution and Z distribution?

The z-distribution and the t-distribution are both used for inference, but they have different assumptions and applications. The z-distribution, also called the standard normal distribution, has a mean of 0 and a standard deviation of 1. It is used when the population standard deviation is known or the sample size is large (at least 30). The t-distribution, also called the Student's t-distribution, is similar to the z-distribution, but has fatter tails and is sensitive to sample size. It is used when the population standard deviation is unknown or the sample size is small (less than 30). As the sample size increases, the t-distribution approaches the z-distribution.

Q14) Is the T distribution normal?

It is a type of normal distribution used for smaller sample sizes, where the variance in the data is unknown

Q15) What does the T distribution tell us?

In statistics ,the t-distribution is most often used to : 1)Find the critical values for a confidence interval when the data is approximately normally distributed. 2)Find the corresponding p-value from a statistical test that uses the t-distribution (t-tests, regression analysis).

MACHINE LEARNING

1. Which of the following are disadvantages of using Hard Margin SVM classifier? A) They allow misclassifications, that's why they are not optimal. B) They cannot be used when the data is not completely linearly separable while allowing no errors. C) They are not optimal to use in case of outliers. D) None of the above.
2. Which of the following statements are true regarding maximal margin classifier? A) It is the most optimal classifier in a completely linearly separable data. B) It's the classifier for which the margin length or the distance between the closest data-point on either side of the classifier and the classifier is maximized. C) Any possible classifier which can linearly separate the data of two classes is called maximal margin classifier. D) All of the above.
3. Which of the following statements are true regarding soft margin SVM classifier? A) They are less sensitive to outliers and can be used even in their presence. B) They make sure that there is no data point present in the margin area. C) They allow some degree of errors or misclassification. D) They can be used in case data is not completely linearly separable.
4. Which of the following statements are true regarding SVMs? A) They take the data from lower dimensional space to some higher dimensional space in case the data is not likely to be linearly separable. B) They use the kernel tricks to escape the complex computations required to transform the data. C) If the data is not linearly separable SVM technique cannot be used. D) All of the above.
5. Which of the following Statements are true regarding the Kernel functions used in SVM? A) These functions gives value of the dot product of pairs of data-points in the desired higher. dimensional space without even explicitly converting the whole data in to higher dimensional space. B) We have to first convert the whole data in to the higher dimensional space before applying the kernel function. C) The data product values given by the kernel functions are used to find the classifier in the higher dimensional space. D) None of the above
6. How can SVM be classified? A) It is a model trained using unsupervised learning. It can be used for classification and regression. B) It is a model trained using unsupervised learning. It can be used for classification but not for regression C) It is a model trained using supervised learning. It can be used for classification and regression. D) It is a model trained using supervised learning. It can be used for classification not for regression.
7. The quality of an SVM model depends upon: A) Selection of Kernel B) Kernel Parameters C) Soft Margin Parameter C D) All of the above
8. The SVM's are less effective when: A) The data is linearly separable. B) The data is clean and ready to use. C) The data is noisy and contains overlapping points. D) None of these.
9. What would happen when you use very small C (C~0)? A) Misclassification would happen. B) Data will be correctly classified. C) Can't say D) None of these.
10. What do you mean by generalization error in terms of the SVM? A) How far the hyperplane is from the support vectors. B) How accurately the SVM can predict outcomes for unseen data. C) The threshold amount of error in an SVM. D) None of these.

Q1) Which of the following are disadvantages of using Hard Margin SVM classifier?

A) They allow misclassifications, that's why they are not optimal. B) They cannot be used when the data is not completely linearly separable while allowing no errors. C) They are not optimal to use in case of outliers. D) None of the above.

ANSWER: B) They cannot be used when the data is not completely linearly separable while allowing no errors.

Q2) Which of the following statements are true regarding maximal margin classifier?

A) It is the most optimal classifier in a completely linearly separable data. B) It's the classifier for which the margin length or the distance between the closest data-point on either side of the classifier and the classifier is maximized. C) Any possible classifier which can linearly separate the data of two classes is called maximal margin classifier. D) All of the above.

ANSWER: D) All of the above.

Q3) Which of the following statements are true regarding soft margin SVM classifier?

A) They are less sensitive to outliers and can be used even in their presence. B) They make sure that there is no data point present in the margin area. C) They allow some degree of errors or misclassification. D) They can be used in case data is not completely linearly separable.

ANSWER: A) They are less sensitive to outliers and can be used even in their presence. C) They allow some degree of errors or misclassification. D) They can be used in case data is not completely linearly separable.

Q4) Which of the following statements are true regarding SVMs?

A) They take the data from lower dimensional space to some higher dimensional space in case the data is not likely to be linearly separable. B) They use the kernel tricks to escape the complex computations required to transform the data. C) If the data is not linearly separable SVM technique cannot be used. D) All of the above.

ANSWER: D) All of the above.

Q5) Which of the following Statements are true regarding the Kernel functions used in SVM?

A) These functions gives value of the dot product of pairs of data-points in the desired higher. dimensional space without even explicitly converting the whole data in to higher dimensional space. B) We have to first convert the whole data in to the higher dimensional space before applying the kernel function. C) The data product values given by the kernel functions are used to find the classifier in the higher dimensional space. D) None of the above

ANSWER: A) These functions gives value of the dot product of pairs of data-points in the desired higher. dimensional space without even explicitly converting the whole data in to higher dimensional space. C) The data product values given by the kernel functions are used to find the classifier in the higher dimensional space.

Q6) How can SVM be classified?

A) It is a model trained using unsupervised learning. It can be used for classification and regression. B) It is a model trained using unsupervised learning. It can be used for classification but not for regression C) It is a model trained using supervised learning. It can be used for classification and regression. D) It is a model trained using supervised learning. It can be used for classification not for regression.

ANSWER: C) It is a model trained using supervised learning. It can be used for classification and regression.

Q7) The quality of an SVM model depends upon:

A) Selection of Kernel B) Kernel Parameters C) Soft Margin Parameter C D) All of the above

ANSWER: D) All of the above.

Q8) The SVM's are less effective when:

A) The data is linearly separable. B) The data is clean and ready to use. C) The data is noisy and contains overlapping points. D) None of these.

ANSWER: None of these

Q9) What would happen when you use very small C (C~0)?

A) Misclassification would happen. B) Data will be correctly classified. C) Can't say D) None of these.

ANSWER: A)Misclassification would happen.

Q10) What do you mean by generalization error in terms of the SVM?

A) How far the hyperplane is from the support vectors. B) How accurately the SVM can predict outcomes for unseen data. C) The threshold amount of error in an SVM. D) None of these.

ANSWER: B) How accurately the SVM can predict outcomes for unseen data.

In [ ]: