

PYTHON – WORKSHEET 1

Q1 to Q8 have only one correct answer. Choose the correct option to answer your question.

1. Which of the following operators is used to calculate remainder in a division? A) # B) & C) % D) \$
2. In python $2//3$ is equal to? A) 0.666 B) 0 C) 1 D) 0.67
3. In python, $6<<2$ is equal to? A) 36 B) 10 C) 24 D) 45
4. In python, $6\&2$ will give which of the following as output? A) 2 B) True C) False D) 0
5. In python, $6|2$ will give which of the following as output? A) 2 B) 4 C) 0 D) 6
6. What does the finally keyword denote in python? A) It is used to mark the end of the code B) It encloses the lines of code which will be executed if any error occurs while executing the lines of code in the try block. C) The finally block will be executed no matter if the try block raises an error or not. D) None of the above
7. What does raise keyword is used for in python? A) It is used to raise an exception. B) It is used to define lambda function C) It's not a keyword in python. D) None of the above
8. Which of the following is a common use case of yield keyword in python? A) in defining an iterator B) while defining a lambda function C) in defining a generator D) in for loop. Q9 and Q10 have multiple correct answers. Choose all the correct options to answer your question.
9. Which of the following are the valid variable names? A) _abc B) labc C) abc2 D) None of the above
10. Which of the following are the keywords in python? A) yield B) raise C) look-in D) all of the above Q11 to Q15 are programming questions. Answer them in Jupyter Notebook.
11. Write a python program to find the factorial of a number.
12. Write a python program to find whether a number is prime or composite.
13. Write a python program to check whether a given string is palindrome or not.
14. Write a Python program to get the third side of right-angled triangle from two given sides.
15. Write a python program to print the frequency of each of the characters present in a given string

In [1]:

```
#1) Which of the following operators is used to calculate remainder in a division?  
#A) # B) & C) % D) $
```

In [2]:

```
# Answer = C) %
```

In [3]:

```
2%2
```

Out[3]:

```
0
```

In [4]:

```
3%2
```

Out[4]:

```
1
```

In [5]:

```
#2) In python 2//3 is equal to? A) 0.666 B) 0 C) 1 D) 0.67
```

In [6]:

```
# Answer = B) 0
```

In [7]:

```
2//3
```

```
Out[7]:  
0
```

```
In [8]:
```

```
# 3) In python, 6<<2 is equal to? A) 36 B) 10 C) 24 D) 45
```

```
In [9]:
```

```
# Answer = C) 24
```

```
In [10]:
```

```
6<<2
```

```
Out[10]:
```

```
24
```

```
In [11]:
```

```
# 4) In python, 6&2 will give which of the following as output? A) 2 B) True C) False D) 0
```

```
In [12]:
```

```
# Answer = A) 2
```

```
In [13]:
```

```
6&2
```

```
Out[13]:
```

```
2
```

```
In [14]:
```

```
# 5) In python, 6/2 will give which of the following as output? A) 2 B) 4 C) 0 D) 6
```

```
In [15]:
```

```
# Answer = D) 6
```

```
In [16]:
```

```
6|2
```

```
Out[16]:
```

```
6
```

```
In [17]:
```

```
# 6 ) what does the finally keyword denotes in python? A) It is used to mark the end of the
```

```
In [18]:
```

```
# Answer = C) the finally block will be executed no matter if the try block raises an error
```

```
In [19]:
```

```
# 7) what does raise keyword is used for in python? A) It is used to raise an exception. B)
```

```
In [20]:
```

```
# Answer = A) It is used to raise an exception
```

```
In [21]:
```

```
raise?
```

```
Object `raise` not found.
```

```
In [22]:
```

```
# 8) which of the following is a common use case of yield keyword in python? A) in defining
```

```
In [23]:
```

```
# Answer = C) in defining a generator
```

```
In [24]:
```

```
def fun_generator():  
    yield "Hello world!!"  
    yield "WELCOME HOME "
```

```
obj = fun_generator()
```

```
print(type(obj))
print(next(obj))
print(next(obj))
```

```
<class 'generator'>
Hello world!!
WELCOME HOME
```

In [25]:

```
# 9) Choose all the correct options to answer your question.
# Which of the following are the valid variable names? A) _abc B) 1abc C) abc2 D) None of th
```

In [26]:

```
# Answers = A) AND C)
```

In [27]:

```
_abc=[1,2]
```

```
abc2=[5,6]
```

In [28]:

```
_abc
```

Out[28]:

```
[1, 2]
```

In [29]:

```
abc2
```

Out[29]:

```
[5, 6]
```

In [30]:

```
1abc=[3,4]
```

```
1abc
```

```
Input In [30]
```

```
1abc=[3,4]
```

```
^
```

```
SyntaxError: invalid syntax
```

In [31]:

```
# 10) Which of the following are the keywords in python? A) yield B) raise C) look-in D) all
```

In [32]:

```
# Answers = A) yield and B) raise
```

In [33]:

```
#11. Write a python program to find the factorial of a number.
```

In [34]:

```
def factorial(n):
    return 1 if(n==0 or n==1) else n * factorial(n-1)
```

In [35]:

```
num=5
print('Factorial of num ',num , 'is',factorial(num))
```

```
Factorial of num 5 is 120
```

In [36]:

```
def fact(n):
    if n<0:
        return 0
    elif n==0 or n==1:
        return 1
    else:
```

```

f=1
while(n>1):
    f *= n
    n -=1
return f

```

In [37]:

```

n=6
print('Factorial of number ',n, 'is ',fact(n))

```

Factorial of number 6 is 720

In [38]:

#12. Write a python program to find whether a number is prime or composite

In [39]:

```

from math import sqrt

```

In [40]:

```

def prime(num):
    if num>1:
        for i in range (2,num):
            if (num%i)==0:
                print(num,'is not a prime number but a composite number')
                break
            else:
                print(num,'is a prime number')
                break

```

In [41]:

```

prime(6)

```

6 is not a prime number but a composite number

In [42]:

```

prime(4)

```

4 is not a prime number but a composite number

In [43]:

```

prime(3)

```

3 is a prime number

In [44]:

```

prime(5)

```

5 is a prime number

In [45]:

```

prime(11)

```

11 is a prime number

In [46]:

```

prime(8)

```

8 is not a prime number but a composite number

In [47]:

#13. Write a python program to check whether a given string is palindrome or not

In [48]:

```
def ispalindrome(w):  
    return w==w[::-1]
```

In [49]:

```
w='malayalam'  
ispalindrome(w)
```

Out[49]:

True

In [50]:

```
w='radar'  
ispalindrome(w)
```

Out[50]:

True

In [51]:

```
w='civic'  
ispalindrome(w)
```

Out[51]:

True

In [52]:

```
w='cake'  
ispalindrome(w)
```

Out[52]:

False

In [53]:

#14. Write a Python program to get the third side of right-angled triangle from two given si

In [54]:

```
import math  
a=float(input('Enter Base = '))  
b=float(input('Enter Height = '))  
c=math.sqrt(a**2 + b**2)  
print('Hypotenuse = ',c)
```

Enter Base = 10

Enter Height = 20

Hypotenuse = 22.360679774997898

In [55]:

15. Write a python program to print the frequency of each of the characters present in a g

In [56]:

```
test_string='RAMAA DEVI'  
ans={i:test_string.count(i) for i in test_string}  
print('The count of all characters in RAMAA DEVI is ',  
      ans)
```

The count of all characters in RAMAA DEVI is {'R': 1, 'A': 3, 'M': 1, ' ': 1, 'D': 1, 'E': 1,

In [58]:

```
from collections import Counter  
test_string='RAMADEVI'  
ans=Counter(test_string)  
print('The count of all characters in RAMADEVI is ',ans)
```

The count of all characters in RAMADEVI is Counter({'A': 2, 'R': 1, 'M': 1, 'D': 1, 'E': 1, 'I': 1,

STATISTICS WORKSHEET-8

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

1. In hypothesis testing, type II error is represented by β and the power of the test is $1-\beta$ then β is: a. The probability of rejecting H_0 when H_1 is true b. The probability of failing to reject H_0 when H_1 is true c. The probability of failing to reject H_1 when H_0 is true d. The probability of rejecting H_0 when H_1 is true
2. In hypothesis testing, the hypothesis which is tentatively assumed to be true is called the a. correct hypothesis b. null hypothesis c. alternative hypothesis d. level of significance
3. When the null hypothesis has been true, but the sample information has resulted in the rejection of the null, a _ has been made a. level of significance b. Type II error c. critical value d. Type I error
4. For finding the p-value when the population standard deviation is unknown, if it is reasonable to assume that the population is normal, we use a. the z distribution b. the t distribution with $n - 1$ degrees of freedom c. the t distribution with $n + 1$ degrees of freedom d. none of the above
5. A Type II error is the error of a. accepting H_0 when it is false b. accepting H_0 when it is true c. rejecting H_0 when it is false d. rejecting H_0 when it is true
6. A hypothesis test in which rejection of the null hypothesis occurs for values of the point estimator in either tail of the sampling distribution is called a. the null hypothesis b. the alternative hypothesis c. a one-tailed test d. a two-tailed test
7. In hypothesis testing, the level of significance is a. the probability of committing a Type II error b. the probability of committing a Type I error c. the probability of either a Type I or Type II, depending on the hypothesis to be tested d. none of the above
8. In hypothesis testing, β is a. the probability of committing a Type II error b. the probability of committing a Type I error c. the probability of either a Type I or Type II, depending on the hypothesis to be test d. none of the above
9. When testing the following hypotheses at an α level of significance $H_0: p = 0.7$ $H_1: p > 0.7$ The null hypothesis will be rejected if the test statistic Z is a. $z > z_\alpha$ b. $z < z_\alpha$ c. $z < -z$ d. none of the above
10. Which of the following does not need to be known in order to compute the P-value? a. knowledge of whether the test is one-tailed or two-tail b. the value of the test statistic c. the level of significance d. All of the above are needed
11. The maximum probability of a Type I error that the decision maker will tolerate is called the a. level of significance b. critical value c. decision value d. probability value
12. For t distribution, increasing the sample size, the effect will be on a. Degrees of Freedom b. The t-ratio c. Standard Error of the Means d. All of the Above
13. Q13 to Q15 are subjective answers type questions. Answers them in their own words briefly.
12. What is Anova in SPSS?
13. What are the assumptions of Anova?
14. What is the difference between one way Anova and two way Anova?

Q1) In hypothesis testing, type II error is represented by β and the power of the test is $1-\beta$ then β is:

- a. The probability of rejecting H_0 when H_1 is true
- b. The probability of failing to reject H_0 when H_1 is true
- c. The probability of failing to reject H_1 when H_0 is true
- d. The probability of rejecting H_0 when H_1 is true

ANSWER: a) The probability of rejecting H_0 when H_1 is true

Q2) In hypothesis testing, the hypothesis which is tentatively assumed to be true is called the

- a. correct hypothesis
- b. null hypothesis
- c. alternative hypothesis
- d. level of significance

ANSWER: b) null hypothesis

Q3) When the null hypothesis has been true, but the sample information has resulted in the rejection of the null, a

_ has been made a. level of significance b. Type II error c. critical value d. Type I error

ANSWER: d) Type I error

Q4) For finding the p-value when the population standard deviation is unknown, if it is reasonable to assume that the population is normal, we use

- a. the z distribution
- b. the t distribution with $n - 1$ degrees of freedom
- c. the t distribution with $n + 1$ degrees of freedom
- d. none of the above

ANSWER: a) the z distribution

Q5) A Type II error is the error of

a. accepting H_0 when it is false b. accepting H_0 when it is true c. rejecting H_0 when it is false d. rejecting H_0 when it is true

ANSWER: a) accepting H_0 when it is false

Q6) A hypothesis test in which rejection of the null hypothesis occurs for values of the point estimator in either tail of

the sampling distribution is called a. the null hypothesis b. the alternative hypothesis c. a one-tailed test d. a two-tailed test

ANSWER: d) a two-tailed test

Q7) In hypothesis testing, the level of significance is

a. the probability of committing a Type II error b. the probability of committing a Type I error c. the probability of either a Type I or Type II, depending on the hypothesis to be tested d. none of the above

ANSWER: c) the probability of either a Type I or Type II, depending on the hypothesis to be tested.

Q8) In hypothesis testing, β is

a. the probability of committing a Type II error b. the probability of committing a Type I error c. the probability of either a Type I or Type II, depending on the hypothesis to be test d. none of the above

ANSWER: a) the probability of committing a Type II error.

Q9) When testing the following hypotheses at an α level of significance

$H_0: p = 0.7$ $H_1: p > 0.7$ The null hypothesis will be rejected if the test statistic Z is a. $z > z_\alpha$ b. $z < z_\alpha$ c. $z < -z$ d. none of the above

ANSWER: a) $z > z_\alpha$

Q10) which of the following does not need to be known in order to compute

the P-value?

a. knowledge of whether the test is one-tailed or two-tail b. the value of the test statistic c. the level of significance d. All of the above are needed

ANSWER: c) the level of significance.

Q11) The maximum probability of a Type I error that the decision maker will tolerate is called the

a. level of significance b. critical value c. decision value d. probability value

ANSWER: a) level of significance

Q12) For t distribution, increasing the sample size, the effect will be on

a. Degrees of Freedom b. The t-ratio c. Standard Error of the Means d. All of the Above

ANSWER: d) ALL of the Above

Q13) what is Anova in SPSS?

ANSWER: Analysis of Variance, (ANOVA) in SPSS (Statistical Software Suite), is used for examining the differences in the mean values of the dependent variable associated with the effect of the controlled independent variables, after taking into account the influence of the uncontrolled independent variables. Essentially, ANOVA in SPSS is used as the test of means for two or more populations

Q14) what are the assumptions of Anova?

ANSWER: Anova is a parametric test. There are some criteria your data must meet if you want to use it. 1) The data should be normally distributed. 2) Data should be independantly sampled. 3) If you have an independant group, there must be homogeneity of variance. 4) If you have a repeated measures, it must have sphericity.

Q15) what is the difference between one way Anova and two way Anova?

ANSWER: The only difference between one-way and two-way ANOVA is the number of independent variables. A one-way ANOVA has one independent variable, while a two-way ANOVA has two.

MACHINE LEARNING

ASSIGNMENT - 10 In Q1 to Q8, only one option is correct, Choose the correct option:

1. In the linear regression equation $y = \theta_0 + \theta_1 x$, θ_0 is the: A) Slope of the line B) Independent variable C) y intercept D) Coefficient of determination
 2. True or False: Linear Regression is a supervised learning algorithm. A) True B) False
 3. In regression analysis, the variable that is being predicted is: A) the independent variable B) the dependent variable C) usually denoted by x D) usually denoted by r
 4. Generally, which of the following method(s) is used for predicting continuous dependent variables? A) Logistic Regression B) Linear Regression C) Both D) None of the above
 5. The coefficient of determination is: A) the square root of the correlation coefficient B) usually less than zero C) the correlation coefficient squared D) equal to zero
 6. If the slope of the regression equation is positive, then: A) y decreases as x increases B) y increases as x increases C) y decreases as x decreases D) None of these
 7. Linear Regression works best for: A) linear data B) non-linear data C) both linear and non-linear data D) None of the above
 8. The coefficient of determination can be in the range of: A) 0 to 1 B) -1 to 1 C) -1 to 0 D) 0 to infinity In Q9 to Q13, more than one options are correct, Choose all the correct options:
 9. Which of the following evaluation metrics can be used for linear regression?
A) Classification Report B) RMSE C) ROC curve D) MAE
 1. Which of the following is true for linear regression? A) Linear regression is a supervised learning algorithm. B) Linear regression supports multi-collinearity. C) Shape of linear regression's cost function is convex. D) Linear regression is used to predict discrete dependent variable.
 2. Which of the following regularizations can be applied to linear regression? A) Ridge B) Lasso C) Pruning D) Elastic Net
 3. Linear regression performs better for: A) Large amount of training samples with small number of features. B) Same number of features and training samples C) Large number of features D) The variables which are drawn independently, identically distributed
 10. Which of the following assumptions are true for linear regression? A) Linearity B) Homoscedasticity C) Non-Independent D) Normality
- MACHINE LEARNING ASSIGNMENT - 10 Q14 and Q15 are subjective answer type questions, Answer them briefly.
11. Explain Linear Regression?
 12. What is difference between simple linear and multiple linear regression?

Q1) In the linear regression equation $y = \theta_0 + \theta_1 x$, θ_0 is the:

A) Slope of the line B) Independent variable C) y intercept D) Coefficient of determination

ANSWER: C) y intercept

Q2) True or False: Linear Regression is a supervised learning algorithm.

A) True B) False

ANSWER: A) True

Q3) In regression analysis, the variable that is being predicted is:

A) the independent variable B) the dependent variable C) usually denoted by x D) usually denoted by r

ANSWER: B) the dependant variable

Q4) Generally, which of the following method(s) is used for predicting continuous

dependent variables? A) Logistic Regression B) Linear Regression C) Both D) None of the above

ANSWER: B) Linear Regression

Q5) The coefficient of determination is:

A) the square root of the correlation coefficient B) usually less than zero C) the correlation coefficient squared D) equal to zero

ANSWER: C) The correlation coefficient squared

Q6) 6. If the slope of the regression equation is positive, then:

A) y decreases as x increases B) y increases as x increases C) y decreases as x decreases D) None of these

ANSWER: B) y increases as x increases.

Q7) Linear Regression works best for:

A) linear data B) non-linear data C) both linear and non-linear data D) None of the above

ANSWER: C) both linear and non-linear data

Q8) The coefficient of determination can be in the range of:

A) 0 to 1 B) -1 to 1 C) -1 to 0 D) 0 to infinity

ANSWER: 0 to 1

Q9) which of the following evaluation metrics can be used for linear regression?

A) Classification Report B) RMSE C) ROC curve D) MAE

ANSWER: B) RMSE D) MAE

Q10) which of the following is true for linear regression?

A) Linear regression is a supervised learning algorithm. B) Linear regression supports multi-collinearity. C) Shape of linear regression's cost function is convex. D) Linear regression is used to predict discrete dependent variable

ANSWER: A) Linear regression is a supervised learning algorithm. B) Linear regression supports multi-collinearity. C) Shape of linear regression's cost function is convex.

Q11) which of the following regularizations can be applied to linear regression?

A) Ridge B) Lasso C) Pruning D) Elastic Net

ANSWER: A) Ridge B) Lasso D) Elastic Net

Q12) Linear regression performs better for:

A) Large amount of training samples with small number of features. B) Same number of features and training samples C) Large number of features D) The variables which are drawn independently, identically distributed

ANSWER: A) Large amount of training samples with small number of features. C) Large number of features D) The variables which are drawn independently, identically distributed

Q13) which of the following assumptions are true for linear regression?

A) Linearity B) Homoscedasticity C) Non-Independent D) Normality

ANSWER: A) Linearity B) Homoscedasticity D) Normality

Q14) Explain Linear Regression.

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables they are considering, and the number of independent variables getting used. There are many names for a regression's dependent variable. It may be called an outcome variable, criterion variable, endogenous variable, or regressand. The independent variables can be called exogenous variables, predictor variables, or regressors.

Q15) what is difference between simple linear and multiple linear regression?

Simple linear regression has only one x and one y variable. Multiple linear regression has one y and two or more x variables. For instance, when we predict rent based on square feet alone that is simple linear regression. When we predict rent based on square feet and age of the building that is an example of multiple linear regression.

In []: