

Core Module 5

Points: 61/90

Time: 55:04

1

Write you name: *

Seema

2

Write you Email: *

dksainiaggarwal@gmail.com

3

Your NSTI Location: *

☐ *Patna*☐ *Kanpur*☒ *Dehradun*☐ *Allahabad*

- ☐ *Ramanthapur*
- ☐ *Vidyanagar*
- ☐ *Mumbai*
- ☐ *Jodhpur*
- ☐ *Noida*
- ☐ *Bengaluru*
- ☐ *Bhubaneswar*
- ☐ *SaltLake(Kolkata)*
- ☐ *Howrah*
- ☐ *Calicut*
- ☐ *Trivandrum*

✓ **Correct** 1/1 Points

4

With the increase in sample size, the error also____

- ☒ a) decreases
- ☐ b) increases
- ☐ c) remains same
- ☐ d) all of the above

✓ **Correct** 1/1 Points

5

Which is the correct command to install Matplotlib?

- ☒ pip install matplotlib
- ☐ pip install matplotlib.pz
- ☐ pip install matplotlib.*
- ☐ All of the above

✗ **Incorrect** 0/1 Points

6

What is the purpose of a data warehouse in the context of business intelligence?

- ☒ a) Storing unstructured data
- ☐ b) Storing all data generated by a business
- ☐ c) Aggregating and organizing data for analysis and reporting
- ☐ d) Running machine learning algorithms

✗ **Incorrect** 0/1 Points

7

What is a correct syntax to check the number of dimensions in an

array?

- ☒ a) arr.ndim()
- ☐ b) arr.dim()
- ☐ c) arr.dim
- ☐ d) arr.ndim

✓ **Correct** 1/1 Points

8

Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of

- ☒ Reinforcement Learning
- ☐ Supervised Learning: Classification
- ☐ Unsupervised Learning: Regression
- ☐ None of the above

✓ **Correct** 1/1 Points

9

Which function from the options given below can read the dataset from a large text file?

- ☐ read_json
- ☐ read_pickle

☐ read_hdf

☒ read_csv

✓ **Correct** 1/1 Points

10

Which of the following does not visualize data?

☐ Charts

☐ Maps

☐ Graphs

☒ Shapes

✓ **Correct** 1/1 Points

11

When the possible outcomes of an experiment are equally likely to occur, this we apply:

☐ Relative Probability

☐ Subjective probability

☐ Conditional probability

☒ Classical probability

✓ **Correct** 1/1 Points

12

In OLS linear regression, if the RSS is zero, what does this indicate?

- ☐ A) The model has no predictive power.
- ☒ B) The model is perfect and predicts all data points without error.
- ☐ C) The model has a very high bias.
- ☐ D) The model is overfitting the data.

✓ **Correct** 1/1 Points

13

Which of these are the steps in the sampling process_____?

- ☐ a) choosing the sampling frame
- ☐ b) defining the target population
- ☐ c) identifying and selecting the method of sample
- ☒ d) all of these

✓ **Correct** 1/1 Points

14

The independent variable is used to explain the dependent variable in _____.

- ☒ Linear regression analysis
- ☐ Multiple regression analysis
- ☐ Non-linear regression analysis
- ☐ None of the above

✗ **Incorrect** 0/1 Points

15

Which of the following is wrong statement about the maximum likelihood approach?

- ☐ This method doesn't always involve probability calculations
- ☒ It finds a tree that best accounts for the variation in a set of sequences
- ☐ The method is similar to the maximum parsimony method
- ☐ The analysis is performed on each column of a multiple sequence alignment

✗ **Incorrect** 0/1 Points

16

Suppose that R-square value of a bivariate regression of Y on X is 0.64. Which of the following is necessarily true?

- ☐ The correlation coefficient between X and Y is 0.8
- ☒ The correlation coefficient is -0.8
- ☐ The correlation coefficient is 0

☐ The correlation coefficient is either 0.8 or -0.8

✓ **Correct** 1/1 Points

17

Which of the following is a commonly used NoSQL database for handling unstructured or semi-structured data in Big Data Analytics?

- ☐ a) MySQL
- ☒ c) MongoDB
- ☐ d) Oracle Database
- ☐ d) SQL Server

✓ **Correct** 1/1 Points

18

What will be output for the following code?

```
import pandas as pd
import numpy as np
s = pd.Series(np.random.randn(2))
print s.size
```

- ☐ 0
- ☐ 1
- ☒ 2
- ☐ 3

✓ **Correct** 1/1 Points

19

What do we call a complete and proper survey of a population?

- ☐ a) report
- ☒ b) census
- ☐ c) sample
- ☐ d) none of these

✗ **Incorrect** 0/1 Points

20

Which of the following distributions is characterized by a constant probability density over its entire range?

- ☐ a) Binomial
- ☒ b) Normal
- ☐ c) Uniform
- ☐ d) Exponential

✗ **Incorrect** 0/1 Points

21

The most important object defined in NumPy is an N-dimensional

array type called?

- ☐ a) ndarray
- ☐ b) narray
- ☒ c) nd_array
- ☐ d) darray

✗ **Incorrect** 0/1 Points

22

Find the standard deviation of the given data sets 7,47,8,42,47,95,42,96,3

- ☒ 29.09
- ☐ 30.09
- ☐ 31.09
- ☐ 32.09

✓ **Correct** 1/1 Points

23

Plot which is used to given statistical summary is

- ☒ Box plot
- ☐ Bar
- ☐ Line

☐ Histogram

✗ **Incorrect** 0/1 Points

24

To import pyplot module we can write

- ☐ a. import pyplot as plt
- ☐ b. import matplotlib.pyplot
- ☒ c. import matplotlib.pyplot as plt
- ☐ Both a and c
- ☐ Both b and c

✗ **Incorrect** 0/1 Points

25

What do we call the population value?

- ☐ a) statistic
- ☐ b) parameter
- ☒ c) data
- ☐ d) variable

✓ **Correct** 1/1 Points

26

What is the difference between supervised and unsupervised learning?

- ☒ Supervised learning requires labeled data while unsupervised learning does not.
- ☐ Unsupervised learning requires labeled data while supervised learning does not.
- ☐ Supervised learning does not require data while unsupervised learning does.
- ☐ There is no difference between supervised and unsupervised learning.

✗ **Incorrect** 0/1 Points

27

What is the role of a KPI (Key Performance Indicator) in business intelligence?

- ☒ a) Identifying new business opportunities
- ☐ b) Measuring and tracking important business metrics
- ☐ c) Conducting market research
- ☐ d) Building data warehouses

✓ **Correct** 1/1 Points

28

Which metric indicates the proportion of the variance in the dependent variable that is explained by the independent variables in a regression model?

- ☐ a) Mean Absolute Error (MAE)

- ☐ b) Root Mean Square Error (RMSE)
- ☒ c) R-squared (R^2)
- ☐ d) Mean Squared Error (MSE)

✗ **Incorrect** 0/1 Points

29

In simple linear regression, what is the formula for the slope of the regression line?

- ☐ a) β_0
- ☐ b) β_1
- ☐ c) σ
- ☒ d) R-squared

✓ **Correct** 1/1 Points

30

What does ROI stand for in business?

- ☒ a) Return on Investment
- ☐ b) Risk of Inflation
- ☐ c) Rate of Interest
- ☐ d) Revenue of Income

✓ **Correct** 1/1 Points

31

What is machine learning?

- ☐ The selective acquisition of knowledge through the use of manual programs
- ☐ The selective acquisition of knowledge through the use of computer programs
- ☐ The autonomous acquisition of knowledge through the use of manual programs
- ☒ The autonomous acquisition of knowledge through the use of computer programs

✓ **Correct** 1/1 Points

32

What is the probability of getting 1 and 5 if a dice is thrown once?

- ☐ 1/6
- ☒ 1/3
- ☐ 2/3
- ☐ 8/9

✓ **Correct** 1/1 Points

33

When the available population is _____, we use a stratified sample.

- ☐ a) too small
- ☐ b) very large
- ☐ c) homogeneous
- ☒ d) heterogeneous

✓ **Correct** 1/1 Points

34

What is a correct syntax to print the number 8 from the array below:

arr = np.array([[1,2,3,4,5], [6,7,8,9,10]])

- ☐ a) print(arr[3,0])
- ☒ b) print(arr[1,2])
- ☐ c) print(arr[7,2])
- ☐ d) print(arr[2,3])

✗ **Incorrect** 0/1 Points

35

If the R-squared (R²) score of a regression model is 0, what does it imply about the model's performance?

- ☒ a) The model has perfect predictions.
- ☐ b) The model has no predictive power.
- ☐ c) The model's predictions are 100% accurate.

- ☐ d) The model has overfit the data.

✓ **Correct** 1/1 Points

36

The difference between a statistic and the parameter is called:

- ☐ Non-random
- ☒ Sampling error
- ☐ Random
- ☐ None of these

✗ **Incorrect** 0/1 Points

37

Which of the following is an example of a non-parametric machine learning algorithm?

- ☐ Linear regression
- ☐ Logistic regression
- ☐ Decision tree
- ☒ Support vector machine

✓ **Correct** 1/1 Points

38

Which of the following is an example of a clustering algorithm?

- ☐ Decision tree
- ☐ Random forest
- ☒ K-means
- ☐ Gradient descent

✗ **Incorrect** 0/1 Points

39

Which of the following of a random variable is a measure of spread?

- ☒ standard deviation
- ☐ variance
- ☐ empirical mean
- ☐ All of the above

✓ **Correct** 1/1 Points

40

The p-value obtained from a classical hypothesis test is:

- ☐ the probability that the null hypothesis is true given the data
- ☐ the probability that the null hypothesis is false given the data

- ☒ the probability of observing the data or more extreme values if the null hypothesis is true
- ☐ the probability of observing the data or more extreme values if the alternative hypothesis is true

✓ **Correct** 1/1 Points

41

In inferential statistics, a p-value is used to determine:

- ☐ a) The mean of a dataset
- ☐ b) The spread of a dataset
- ☒ c) The significance of a result
- ☐ d) The mode of a dataset

✗ **Incorrect** 0/1 Points

42

What do we call sample value?

- ☐ a) variable
- ☐ b) parameter
- ☒ c) data
- ☐ d) statistic

✓ **Correct** 1/1 Points

43

If X and Y in a regression model are totally unrelated,

- ☐ the correlation coefficient would be -1
- ☒ the coefficient of determination would be 0
- ☐ the coefficient of determination would be 1
- ☐ the SSE would be 0

✓ **Correct** 1/1 Points

44

Which of the following is a measure of central tendency?

- ☐ a) Range
- ☐ b) Variance
- ☒ c) Median
- ☐ d) Standard Deviation

✓ **Correct** 1/1 Points

45

What we pass in DataFrame in pandas ?

- ☐ Integer
- ☐ String
- ☒ Pandas series
- ☐ All of the above

✓ **Correct** 1/1 Points

46

What is the range of values for the correlation coefficient (r)?

- ☒ a) -1 to 1
- ☐ b) 0 to 100
- ☐ c) $-\infty$ to ∞
- ☐ d) 0 to 1

✓ **Correct** 1/1 Points

47

NumPY stands for?

- ☐ a) Numbering Python
- ☐ b) Number In Python
- ☒ c) Numerical Python

☐ d) None Of the above

✓ **Correct** 1/1 Points

48

What is the purpose of cross-validation in machine learning?

- ☐ To evaluate the performance of a model on a held-out test set
- ☒ To evaluate the performance of a model on different subsets of the data
- ☐ To compare the performance of different models
- ☐ To tune the hyperparameters of a model

✗ **Incorrect** 0/1 Points

49

In linear regression, what is the role of the dependent variable (Y)?

- ☐ a) It is the variable we are trying to predict or explain.
- ☐ b) It is a constant value used for model calibration.
- ☐ c) It is not used in linear regression.
- ☒ d) It is the variable used for controlling the independent variables.

✓ **Correct** 1/1 Points

50

A Machine Learning technique that helps in detecting the outliers in data.

- ☐ Clustering
- ☐ Classification
- ☒ Anomaly Detection
- ☐ All of the above

✓ **Correct** 1/1 Points

51

Which technology framework is commonly used for distributed storage and processing in Big Data Analytics?

- ☐ a) SQL
- ☒ b) Hadoop
- ☐ c) NoSQL
- ☐ d) Spark

✓ **Correct** 1/1 Points

52

Which type of analytics focuses on predicting future outcomes based on historical data and statistical models?

- ☐ a) Descriptive Analytics

- ☒ b) Predictive Analytics
- ☐ c) Prescriptive Analytics
- ☐ d) Diagnostic Analytics

✓ **Correct** 1/1 Points

53

In the context of Big Data Analytics, what does the term "Data Mining" refer to?

- ☒ a) Extracting valuable insights from large datasets
- ☐ b) Storing data in a distributed database
- ☐ c) Transforming data into a structured format
- ☐ d) Running complex SQL queries

✗ **Incorrect** 0/1 Points

54

Which of the following distributions is commonly used to model the number of events occurring in a fixed interval of time or space?

- ☒ a) Uniform
- ☐ b) Normal
- ☐ c) Poisson

☐ D) Exponential

✗ **Incorrect** 0/1 Points

55

When is the correlation coefficient equal to 0?

- ☐ a) When there is a strong positive linear relationship.
- ☒ b) When there is a strong negative linear relationship.
- ☐ c) When there is no linear relationship.
- ☐ d) When there is a curvilinear relationship.

✓ **Correct** 1/1 Points

56

Which of the following is NOT a common data visualization tool used in business intelligence?

- ☐ a) Microsoft Excel
- ☐ b) Tableau
- ☒ c) Python
- ☐ d) Power BI

✓ **Correct** 1/1 Points

57

Consider a set of 20 samples from a standard normal distribution. We square each sample and sum all the squares. The number of degrees of freedom for a Chi Square distribution will be?

- ☐ 40
- ☒ 20
- ☐ 10
- ☐ None of these

✗ **Incorrect** 0/1 Points

58

Mean square error of estimator obtained by the method of minimum Chi-square is

- ☒ More than or equal to ML Estimators
- ☐ Less than ML Estimator
- ☐ More than ML Estimators
- ☐ Equal to ML Estimators

✗ **Incorrect** 0/1 Points

59

The method in which the previously calculated probabilities are revised with new probabilities is classified as

- ☒ Updating theorem
- ☐ Dependency theorem
- ☐ Central Limit Theorem
- ☐ Bayes theorem

✓ **Correct** 1/1 Points

60

Which of the following function is associated with a continuous random variable?

- ☐ pmf
- ☐ pmv
- ☒ pdf
- ☐ All of the above

✓ **Correct** 1/1 Points

61

What are the three primary characteristics of Big Data, often referred to as the "Three Vs"?

- ☐ a) Volume, Value, Velocity
- ☒ b) Volume, Variety, Velocity
- ☐ c) Variety, Velocity, Validation

☐ d) Value, Velocity, Validation

✓ **Correct** 1/1 Points

62

**Find the mode of the call received on 7 consecutive day
11,13,13,17,19,23,25**

☐ 11

☒ 13

☐ 17

☐ 23

✓ **Correct** 1/1 Points

63

Which function is used to add grid lines to the plot?

☐ gridlines()

☐ grids()

☒ grid()

☐ gridmarks()

✓ **Correct** 1/1 Points

64

VLOOKUP function is used to?

- ☒ a) find related records
- ☐ b) Looks up text can contain 'v'
- ☐ c) Check if two cells are identical
- ☐ d) None of the above

✓ **Correct** 1/1 Points

65

Which of the following is an example of unstructured data in the context of Big Data?

- ☐ a) A relational database
- ☐ b) Sales transaction records
- ☒ c) Textual social media posts
- ☐ d) A structured Excel spreadsheet

✓ **Correct** 1/1 Points

66

Which is the correct code statement to draw a plot without a line?

- ☐ plt.plot(xpoints, ypoints)

- ☐ plt.plot(xpoints, ypoints, 0)
- ☐ plt.plot(xpoints, ypoints, False)
- ☒ plt.plot(xpoints, ypoints, 'o')

✓ **Correct** 1/1 Points

67

Which of the following random variable that take on only a countable number of possibilities?

- ☒ Discrete
- ☐ Non discrete
- ☐ Continuous
- ☐ All of the above

✗ **Incorrect** 0/1 Points

68

The range is a measure of:

- ☒ a) Central tendency
- ☐ b) Spread or variability
- ☐ c) Position
- ☐ d) Probability

✗ **Incorrect** 0/1 Points

69

What is the primary objective of linear regression using OLS?

- ☐ a) Minimize the mean absolute error
- ☐ b) Minimize the mean squared error
- ☐ c) Maximize the R-squared value
- ☒ d) Maximize the correlation coefficient

✓ **Correct** 1/1 Points

70

In which of the following types of sampling the information is carried out under the opinion of an expert?

- ☐ quota sampling
- ☐ convenience sampling
- ☐ purposive sampling
- ☒ judgement sampling

✓ **Correct** 1/1 Points

71

A Null Hypothesis has Level of Significance 9%. For what values of Level of

Significances it will be rejected?

- ☐ 0.99
- ☒ 0.09
- ☐ 0.099
- ☐ 0.9

✓ **Correct** 1/1 Points

72

NumPy arrays can be ____.

- ☐ a) Indexed
- ☐ b) Sliced
- ☐ c) Iterated
- ☒ d) All of the mentioned above

✗ **Incorrect** 0/1 Points

73

Out of these, which is not a probability sampling?

- ☒ a) cluster sampling
- ☐ b) stratified sampling
- ☐ c) quota sampling

☐ d) simple random sampling

✓ **Correct** 1/1 Points

74

What is true about Data Visualization?

- ☐ Data Visualization is used to communicate information clearly and efficiently to users by the usage of information graphics such as tables and charts.
- ☐ Data Visualization helps users in analyzing a large amount of data in a simpler way.
- ☐ Data Visualization makes complex data more accessible, understandable, and usable.
- ☒ All of the above

✓ **Correct** 1/1 Points

75

In the regression equation $Y = 21 - 3X$, the slope is

- ☐ 21
- ☐ -21
- ☐ 3
- ☒ -3

✓ **Correct** 1/1 Points

76

Which of the following is not a type of non-probability sampling?

- ☐ Quota sampling
- ☐ Convenience sampling
- ☐ Snowball sampling
- ☒ Stratified random sampling

✓ **Correct** 1/1 Points

77

Any hypothesis which is tested for the purpose of rejection under the assumption that it is true is called:

- ☐ Alternative hypothesis
- ☐ Statistical hypothesis
- ☒ Null hypothesis
- ☐ None of these

✗ **Incorrect** 0/1 Points

78

What is the purpose of residual analysis in linear regression?

- ☒ a) To calculate the R-squared value.

- ☐ b) To determine the best-fit line.
- ☐ c) To assess the model's assumptions and identify potential problems.
- ☐ d) To find the p-values for the coefficients.

✓ **Correct** 1/1 Points

79

Which of the following is not a measure of spread or variability?

- ☐ a. Standard Deviation
- ☐ b. Variance
- ☒ c. Mean
- ☐ d. Range

✓ **Correct** 1/1 Points

80

Which of the following statement for creating dataframe is valid?

- ☐ df = pd.dataframe(dict1)
- ☐ df = pd.DataFrame(dict1)
- ☐ df = pd.dataFrame(dict1)
- ☒ df = pd.DataFrame(dict1)

✗ **Incorrect** 0/1 Points

81

One of the members of the population is known as the.....

- ☐ a) data
- ☒ b) family
- ☐ c) element
- ☐ d) group

✓ **Correct** 1/1 Points

82

$E(X) = \lambda$ is used for which distribution?

- ☐ Binomial distribution
- ☐ Bernoulli's distribution
- ☒ Poisson's distribution
- ☐ Laplace distribution

✗ **Incorrect** 0/1 Points

83

The distribution that is formed by all possible values of a statistics is known as:

- ☒ a) normal distribution
- ☐ b) hypergeometric distribution
- ☐ c) binomial distribution
- ☐ d) sampling distribution

✓ **Correct** 1/1 Points

84

The process of arranging data into rows and columns is called?

- ☐ a) Frequency distribution
- ☐ b) Classification
- ☒ c) tabulation
- ☐ d) Array

✗ **Incorrect** 0/1 Points

85

What does the central limit theorem state?

- ☒ if the sample size decreases then the sample distribution must approach normal distribution
- ☐ if the sample size decreases then the sampling distribution much approach an exponential distribution
- ☐ All the above

☐ None of these

✓ **Correct** 1/1 Points

86

Identify the following chart's type

- ☐ Bubble Chart
- ☐ Frequency Polygon
- ☐ Pie Chart
- ☒ Scatter Chart

✓ **Correct** 1/1 Points

87

Which of the following is not the parameter of pyplot's plot() method?

- ☐ Marker
- ☒ Lineheight
- ☐ Linestyle
- ☐ Color

✓ **Correct** 1/1 Points

88

This image represents a

- ☐ Frequency polygon
- ☐ Histogram
- ☒ Box plot
- ☐ Scatter plot

✓ **Correct** 1/1 Points

89

In business intelligence, what is ETL an acronym for?

- ☒ a) Extract, Transform, Load
- ☐ b) Efficient Technology Link
- ☐ c) Evaluate, Test, Launch
- ☐ d) Enterprise Tracking Log

✓ **Correct** 1/1 Points

90

Choose the correct option regarding the sampling method?

- ☐ a) the sample is the population's part
- ☐ b) it helps in determining sampling error

- ☐ c) sampling saves money, time, and energy
- ☒ d) all these options are correct

✗ **Incorrect** 0/1 Points

91

Which metric is used to evaluate the accuracy of regression models and penalize larger errors more heavily?

- ☐ a) Mean Absolute Error (MAE)
- ☒ b) Root Mean Square Error (RMSE)
- ☐ c) Mean Squared Error (MSE)
- ☐ d) R-squared (R^2)

✓ **Correct** 1/1 Points

92

The binomial distribution is applicable when:

- ☒ a) Events are independent, and the probability of success remains constant for each trial.
- ☐ b) Events are dependent on each other.
- ☐ c) There are infinitely many possible outcomes.
- ☐ d) The sample size is not known.

✖ **Incorrect** 0/1 Points

93

Which of the following type of chart is not supported by pyplot?

- ☐ Histogram
- ☐ Boxplot
- ☒ Pie
- ☐ All of the above

Keep the information with you by saving your response.

[Save my response](#)

This content is created by the owner of the form. The data you submit will be sent to the form owner. Microsoft is not responsible for the privacy or security practices of its customers, including those of this form owner. Never give out your password.

Powered by Microsoft Forms |

The owner of this form has not provided a privacy statement as to how they will use your response data. Do not provide personal or sensitive information.

| [Terms of use](#)