

BIG DATA

Q.No.1

Hadoop is a framework that works with a variety of related tools. Common cohorts include

- a) MapReduce, Hive and HBase
- b) MapReduce, MySQL and Google Apps
- c) MapReduce, Hummer and Iguana
- d) MapReduce, Heron and Trumpet

Q.No.2

A _____ node acts as the Slave and is responsible for executing a Task assigned to it by the JobTracker.

- a) MapReduce
- b) Mapper
- c) TaskTracker
- d) JobTracker

Q. No. 3

_____ can best be described as a programming model used to develop Hadoop-based applications that can process massive amounts of data.

- a) MapReduce
- b) Mahout
- c) Oozie
- d) All of the mentioned

Q. No. 4

_____ part of the MapReduce is responsible for processing one or more chunks of data and producing the output results.

- a) Maptask
- b) Mapper
- c) Task execution
- d) All of the mentioned

Q. No. 5

_____ function is responsible for consolidating the results produced by each of the Map() functions/tasks.

- a) Reduce
- b) Map
- c) Reducer
- d) All of the mentioned

Q. No. 6

_____ is a utility which allows users to create and run jobs with any executables as the mapper and/or the reducer.

- a) Hadoop Strdata
- b) Hadoop Streaming
- c) Hadoop Stream
- d) None of the mentioned

Q. No. 7

HDFS works in a _____ fashion.

- a) master-worker
- b) master-slave
- c) worker/slave
- d) all of the mentioned

Q. No. 8

_____ NameNode is used when the Primary NameNode goes down.

- a) Rack
- b) Data
- c) Secondary
- d) None of the mentioned

Q.No.9

_____ is the slave/worker node and holds the user data in the form of Data Blocks.

- a) DataNode
- b) NameNode
- c) Data block
- d) Replication

Q.No.10

HDFS provides a command line interface called _____ used to interact with HDFS.

- a) "HDFS Shell"
- b) "FS Shell"
- c) "DFS Shell"
- d) None of the mentioned

Q.No.10

HDFS is implemented in _____ programming language.

- a) C++
- b) Java
- c) Scala
- d) None of the mentioned

Q. No. 11

For YARN, the _____ Manager UI provides host and port information.

- a) Data Node
- b) NameNode
- c) Resource
- d) Replication

Q. No. 11

For _____ the HBase Master UI provides information about the HBase Master uptime.

- a) HBase
- b) Oozie
- c) Kafka
- d) All of the mentioned

Q. No. 12

During start up, the _____ loads the file system state from the fsimage and the edits log file.

- a) DataNode
- b) NameNode
- c) ActionNode
- d) None of the mentioned

Q. No. 13

Hive uses _____ for logging.

- a) logj4
- b) log4l
- c) log4i
- d) log4j

Q. No. 14

HiveServer2 introduced in Hive 0.11 has a new CLI called _____

- a) BeeLine
- b) SqlLine
- c) HiveLine
- d) CLilLine

Q. No. 14

In _____ mode HiveServer2 only accepts valid Thrift calls.

- a) Remote
- b) HTTP
- c) Embedded
- d) Interactive

Q. No. 15

In order to read any file in HDFS, instance of _____ is required.

- a) filesystem
- b) datastream
- c) outstream
- d) inputstream

Q. No. 16

Which of the following interface is implemented by Sqoop for recording?

- a) SqoopWrite
- b) SqoopRecord
- c) SqoopRead
- d) None of the mentioned

Q.No.17

Sqoop uses _____ to fetch data from RDBMS and stores that on HDFS.

- a) Hive
- b) Map reduce
- c) Imphala
- d) BigTOP

Q.No.18

Microsoft uses a Sqoop-based connector to help transfer data from _____ databases to Hadoop.

- a) PostgreSQL
- b) SQL Server
- c) Oracle
- d) MySQL

Q. No. 19

Apache Flume 1.3.0 is the fourth release under the auspices of Apache of the so-called _____ codeline.

- a) NG
- b) ND
- c) NF
- d) NR

Q. No. 20

_____ was created to allow you to flow data from a source into your Hadoop environment.

- a) Imphala
- b) Oozie
- c) Flume
- d) All of the mentioned

JAVA

Q. No. 1

Question:

Which of the following option leads to the portability and security of Java?

Answer Choices

A: Bytecode is executed by JVM

B: The applet makes the Java code secure and portable

C: Use of exception handling

D: Dynamic binding between objects

Q. No. 2

Question:

Which of these cannot be used for a variable name in Java?

Answer Choices

A: identifier

B: keyword

C: identifier & keyword

D: none of the mentioned

Q. No. 3

Question:

Which of the below is not a memory leak solution?

Answer Choices

A: Code changes

B: JVM parameter tuning

C: Process restart

D: GC parameter tuning

Q. No. 4

Question:

What is the output of the following program?

```
public class Test{  
    private String function(){  
        return ("Test");  
    }  
    public final static String function(int data){ return  
        ("Test123");  
    }  
    public static void main(String[] args){ Test  
        obj = new Test();  
        System.out.println(obj.function());  
    }  
}
```

Answer Choices

A: Test

A: Compilation error

B: Runtime error

D: None of these

Q. No. 5

Question:

Suppose a class has public visibility. In this class we define a protected method.

Which of the

following statements is correct?

Answer Choices

A: This method is only accessible from inside the class itself and from inside all subclasses.

B: In a class, you cannot declare methods with a lower visibility than the visibility of the class in which it is defined.

C: From within protected methods you do not have access to public methods.
D: This method is accessible from within the class itself and from within all classes defined in the same package as the class itself.

Q. No. 6

Question:

Which of the following is true about inheritance in Java?

- 1) Final methods can not be overridden
- 2) Protected members are accessible within a package and inherited classes outside the package.
- 3) Protected methods are final.
- 4) We cannot override private methods.

Answer Choices

A: 1, 2 and 4

B: Only 1 and 2

C: 1, 2 and 3

D: 2, 3 and 4

Q. No. 7

Question:

What is the output of the following program?

```
import java.io.IOException;
class Derived {
    public void getDetails() throws IOException //line 23 {
        System.out.println("Derived class");
    }
}
public class Test extends Derived {
    public void getDetails() throws Exception //line 24 {
        System.out.println("Test class");
    }
    public static void main(String[] args) throws IOException //line 25 {
        Derived obj = new Test();
        obj.getDetails();
    }
}
```

Answer Choices

A: Compilation error due to line 23

B: Compilation error due to line 24

C: Compilation error due to line 25

D: None of the above

Q. No. 8

Question:

Which of the following is used to make an Abstract class?

Answer Choices

A: Making atleast one-member function as pure virtual(abstract) function B:

Making atleast one-member function as virtual function

C: Declaring as Abstract class using virtual keyword

D: Declaring as Abstract class using static keyword

Q. No. 9

Question:

What is the output of the following program?

```
class Base {  
    public final void show() {  
        System.out.println("Base::show() called");  
    }  
}  
  
class Derived extends Base { public  
    void show() {  
        System.out.println("Derived::show() called");  
    }  
}  
  
public class Main {  
    public static void main(String[] args) { Base  
        b = new Derived(); b.show();  
    }  
}
```

Answer Choices

A: Derived::show() called

B: Base::show() called

C: Compiler Error

D: Exception

Q. No. 10

Question:

The concept of multiple inheritance is implemented in Java by

I. Extending two or more classes.

II. Extending one class and implementing one or more interfaces.

III. Implementing two or more interfaces.

Answer Choices

A: Only (II)

B: (I) and (II)

C: (II) and (III)

D: Only (III)

Q. No. 11

Question:

What is the output of the following program?

```
public class Outer {  
    private static int data = 10; private  
    static int LocalClass() {  
        class Inner {  
            public int data = 20;  
            private int getData() {  
                return data;  
            }  
        }  
    }  
};
```

```

Inner inner = new Inner();
return inner.getData();
}
public static void main(String[] args) {
System.out.println(data * LocalClass());
}
}

```

Answer Choices

A: Compilation error

B: Runtime Error

C: 200

D: None of the above

Q. No. 12

Question:

In which of the following package Exception class exist?

Answer Choices

A: java.util

B: java.io

C: java.lang

D: java.net

Q. No. 13

Question:

Which of the below statement is/are true about Error?

I. An Error is a subclass of Throwable.

II. An Error is a subclass of Exception.

III. Error indicates serious problems that a reasonable application should not try to catch. iv. An

Error is a subclass of IOException.

Answer Choices

A: (I) and (IV)

B: (I) and (II)

C: (I) and (III)

D: (II) and (III)

Q. No. 14

Question:

Which of these value are returns under the case of normal termination of a program?

Answer Choices

A: 0

B: 1

C: 2

D: 3

Q. No. 15

Question:

Which of these methods can be used to obtain set of all keys in a map?

Answer Choices

A: getAll()

B: getKeys()

C: keyall()

D: keySet()

Q. No. 16

Question:

Which of the following constructor of class Thread is valid one?

Answer Choices

A: Thread(Runnable threadOb, int priority)

B: Thread(int priority)

C: Thread(Runnable threadOb, String threadName)

D: Thread(String threadName, int priority)

Q. No.17

Question:

What is the default value of priority variable MIN_PRIORITY AND MAX_PRIORITY?

A: 0 & 256

B: 0 & 1

C: 1 & 10

D: 1 & 256

Q. No. 18

Question:

What notifyAll() method do?

Answer Choices

A: Wakes up one threads that are waiting on this object's monitor B:

Wakes up all threads that are not waiting on this object's monitor C:

Wakes up all threads that are waiting on this object's monitor

D: None of the above

Q. No. 19

Question:

What is synchronization in reference to a thread?

Answer Choices

A: It's a process of handling situations when two or more threads need access to a shared resource

B: It's a process by which many thread are able to access same shared resource simultaneously

C: It's a process by which a method is able to access many different threads simultaneously

D: It's a method that allow too many threads to access any information the require

Q. No. 20

Question:

Which of the following is method of JDBC batch process?

Answer Choices

A: setBatch()

B: deleteBatch()

C: removeBatch()

D: addBatch()

Statistics:

Q.No.1

What is a hypothesis?

- A. A statement that the researcher wants to test through the data collected in a study
- B. A research question the results will answer
- C. A theory that underpins the study
- D. A statistical method for calculating the extent to which the results could have happened by chance

Q.No.2

Least Square Method uses ____.

- A. Linear polynomial
- B. Linear regression
- C. Linear sequence
- D. None of the mentioned above

Q. No. 3

Alternative Hypothesis is also called as?

- A. Null Hypothesis
- B. Research Hypothesis
- C. Simple Hypothesis
- D. None of the mentioned above

Q. No. 4

Data Analysis is defined by the statistician?

- A. John Tukey
- B. Hans Peter Luhn
- C. Gregory Lon
- D. None of the mentioned above

Q. No. 5

All pandas data structures are ____ mutable but not always _____ mutable.

- a) size, value
- b) semantic, size
- c) value, size
- d) none of the mentioned

Q. No. 6

Which of the following object you get after reading CSV file?

- a) DataFrame
- b) Character Vector
- c) Panel
- d) All of the mentioned

Q. No. 7

Which of the following testing is concerned with making decisions using data?

- a) Probability
- b) Hypothesis
- c) Causal

d) None of the mentioned

Q. No. 8

The expected value or _____ of a random variable is the center of its distribution.

a) mode

b) median

c) mean

d) bayesian inference

Q.No.9

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

a) variance

b) standard deviation

c) empirical mean

d) all of the mentioned

Q.No.10

The square root of the variance is called the _____ deviation.

a) empirical

b) mean

c) continuous

d) standard

Q. No. 11

Which of the following inequality is useful for interpreting variances?

a) Chebyshev

b) Stautatory

c) Testory

d) All of the mentioned

Q. No. 12

Chebyshev's inequality states that the probability of a "Six Sigma" event is less than

a) 10%

b) 20%

c) 30%

d) 3%

Q. No. 13

The _____ function returns its argument with a modified shape, whereas the

_____ method modifies the array itself.

a) reshape, resize

b) resize, reshape

c) reshape2, resize

d) all of the mentioned

Q. No. 14

ndarray is also known as the alias array.

a) True

b) False

Q. No. 15

Which of the following things can be accomplished with linear model?

- a) Flexibly fit complicated functions
- b) Uncover complex multivariate relationships
- c) Build accurate prediction models
- d) All of the mentioned

Q. No. 16

Which of the following statement is incorrect with respect to outliers?

- a) Outliers can have varying degrees of influence
- b) Outliers can be the result of spurious or real processes
- c) Outliers cannot conform to the regression relationship
- d) None of the mentioned

Q.No.17

The _____ of a collection of data is the joint density evaluated as a function of the parameters with the data fixed.

- a) probability
- b) likelihood
- c) poisson distribution
- d) all of the mentioned

Q.No.18

Which of the following object has a method cov to compute covariance between series?

- a) Series
- b) DataFrame
- c) Panel
- d) None of the mentioned

Q. No. 19

The _____ of the Chi-squared distribution is twice the degrees of freedom.

- a) variance
- b) standard deviation
- c) mode
- d) none of the mentioned

Q. No. 20

Gosset's distribution is invented by which of the following scientist?

- a) William Gosset
- b) William Gosling
- c) Gosling Gosset
- d) All of the mentioned

DBT:

Q.No.1

Which one is a main feature of DBMS?

- A. Efficient storage and retrieval of data
- B. Efficient storage and safety of data

- C. Efficient storage and security of data
- D. Efficient storage and scalability of data

Q.No.2

Which is the most adapted DBMS model?

- A. Network Model
- B. Hierarchical Model
- C. Hybrid Model
- D. Relation Model

Q. No. 3

What is the main aspect of Relational data model?

- A. Usage of tables to store data
- B. Availability of foreign key to maintain relation
- C. Availability of data dictionary for metadata
- D. All of the above

Q. No. 4

What is common in all RDBMS software?

- A. All permit storage of data
- B. All permit usage of constraints for performance
- C. All permit SQL commands to interact with data
- D. All of the above

Q. No. 5

In Relational Database Model a table represents:

- A. An Entity
- B. Group of Records
- C. Information about an object
- D. None of the above

Q. No. 6

Which of the database model is followed by Oracle DB?

- A. Network
- B. Hierarchical
- C. Relational
- D. NoSQL

Q. No. 7

What is the full form of SQL?

- A. Structured Query Language
- B. Standard Query Language
- C. Simple Query Language
- D. Standard Question Language

Q. No. 8

Who proposed the Relation Model?

- A. Raymond Boyce
- B. E. F. Codd
- C. C J Date
- D. None of the above

Q.No.9

Normalization rules can be applied on which database model?

- A. Network

- B. Hierarchical
- C. NoSQL
- D. Relational

Q.No.10

Who proposed the initial Normalization forms?

- A. C J Date
- B. E. F. Codd
- C. Raymond Boyce
- D. None of the above

Q. No. 11

What is the main component of First Normalization Form?

- A. Elimination of repeating groups
- B. Composite primary key
- C. In-Direct dependency
- D. All of the above

Q. No. 12

Which normalization form talks about transitive dependency?

- A. 1NF
- B. 2NF
- C. 3NF
- D. BCNF

Q. No. 13

Which one is the most recent Normalization form?

- A. 4NF
- B. 5NF
- C. 6NF
- D. PJNF

Q. No. 14

What type of databases preferably witness denormalization?

- A. OLTP
- B. Data Warehouse
- C. Business Intelligence
- D. None of the above

Q. No. 15

Which one of the following statements is true?

- A. Normalization may result in splitting of table to two or more tables
- B. Normalization rules can be applied in any order
- C. Database must adhere to all Normalization rules
- D. All of the above

Q. No. 16

A table is assumed to be in which Normalization form if it does not contain composite candidate keys?

- A. 1NF
- B. 2NF
- C. 3NF

D. BCNF

Q.No.17

Evaluate the set of SQL statements:

DROP TABLE dept;

ROLLBACK;

DESCRIBE dept

Which one is true about the set?

A. DESCRIBE command will display the structure of dept table.

B. ROLLBACK statement successfully recovered dropped table.

C. DESCRIBE statement returns an error of dept table does not exist.

D. DESCRIBE command require COMMIT to be issued to display structure of dept table

Q.No.18

Which query is used to display customers who have not provided their phone number?

A. SELECT * FROM customer WHERE cust_contact=NULL;

B. SELECT * FROM customer WHERE cust_contact=NVL(cust_contact,0);

C. SELECT * FROM customer WHERE cust_contact IS = NULL;

D. SELECT * FROM customer WHERE cust_contact IS NULL;

Q. No. 19

COMMIT statement is a part of which SQL category?

A. DML

B. DDL

C. DCL

D. TCL

Q. No. 20

MongoDB provides high _____ with replica sets.

A. Performance

B. Availability

C. Scalability

D. none of the mentioned

DV:

Q.No.1

Business intelligence (BI) is a broad category of application programs which includes

a) Decision support

b) Data miningss

c) OLAP

d) All of the mentioned

Q.No.2

In MS-Excel which functions will calculate the number of workdays between

6/9/2013 and

8/12/2013

A.Workday

- B.Date
- C.Networkday
- D.Weekday

Q. No. 3

In MS-Excel data marker on a chart are linked to data points in a worksheet, therefore

- A.You can automatically apply formatting to a data series
- B.You can change the position of a data marker and that automatically changes the data point value in the worksheet
- C.You can change a data point value and automatically that is drawn in the chart
- D.Both (b) & (c)

Q. No. 4

MS-Excel uses the _____ Function when creating a data table

- A.Average
- B.Count
- C.Both (a) & (b)
- D.Table

Q. No. 5

In MS-Excel 3-D reference in a formula

- A.Cannot be modified
- B.Only appears on summary worksheets
- C.Both (a) & (b)
- D.Spans worksheets

Q. No. 6

Is it possible to deploy a URL action on a dashboard object to open a Web Page within a

dashboard rather than opening the system's web browser?

- a) True, with the use of Tableau Server
- b) True, with the use of a Web Page object
- c) False, not possible
- d) True, requires a plug-in

Q. No. 7

How do you identify a continuous field in Tableau?

- a) It is identified by a blue pill in the visualization.
- b) It is identified by a green pill in a visualization.
- c) It is preceded by a # symbol in the data window.
- d) When added to the visualization, it produces distinct values.

Q. No. 8

For creating variable size bins we use _____.

- a) Sets
- b) Groups
- c) Calculated fields
- d) Table Calculations

Q.No.9

A good reason to use a bullet graph.

- a) Analyzing the trend for a time period
- b) Comparing the actual against the target sales
- c) Adding data to bins and calculating count measure
- d) Displaying the sales growth for a particular year

Q.No.10

By definition, Tableau displays measures over time as a _____.

- a) Bar
- b) Line
- c) Histogram
- d) Scatter Plots

Q. No. 11

The icon associated with the field that has been grouped is a _____.

- a) Paper Clip
- b) Set
- c) Hash
- d) Equal To

Q. No. 12

A sheet cannot be used within a story directly. Either sheets should be used within a dashboard, or a dashboard should be used within a story.

- a) True
- b) False

Q. No. 13

How do you identify a continuous field in Tableau?

- a) It is identified by a blue pill in the visualization.
- b) It is identified by a green pill in a visualization.
- c) It is preceded by a # symbol in the data window.
- d) When added to the visualization, it produces distinct values.

Q. No. 14

In regards to separated value files such as .csv and .tsv, what is the delimiter?

- a) Any character such as the comma (,) or tab (\t) that is used to separate the column data.
- b) Any character such as the comma (,) or tab (\t) that is used to separate the row data
- c) Anywhere the comma (,) character is used in the file
- d) Delimiters are not used in separated value files

Q. No. 15

This is an arrangement in which a company outsources some or all of its customer relationship management functions to an application service provider (ASP)

- a) Customer Information Control System
- b) spend management
- c) hosted CRM
- d) online transaction processing

Q. No. 16

Data warehouse architecture is based on _____

- a) RDBMS
- b) Sybase

c) SQL Server

d) DBMS

Q.No.17

Which of the following does not form part of BI Stack in SQL Server ?

a) SSIS

b) OBIEE

c) SSAS

d) SSRS

Q.No.18

_____ is a system where operations like data extraction, transformation and loading

operations are executed

a) ETL

b) Data staging

c) Data integration

d) None

Q. No. 19

_____ is a performance management tool that recapitulates an organization's performance

from several standpoints on a single page

a) Data Cube

b) Balanced Scorecard

c) Dashboard

d) All of above

Q. No. 20

This is the processing of data about customers and their relationship with the enterprise in

order to improve the enterprise's future sales and service and lower cost_____

a) customer relationship management

b) CRM analytics

c) database marketing

d) customer relationship management

Linux and CC

Q.No.1

In the shell, by default, all variables are considered and stored as

a) string

b) integer

c) character

d) float

Q.No.2

Which command reads user input from the terminal and assign this value to a variable name?

- a) read
- b) get
- c) declare
- d) set

Q.No.3

Which one of the following statement is true about variables in shell?

- a) variables do not require declaration before assigning value to them
- b) variables are case sensitive
- c) to extract the contents of a variable, we have to provide the variable a preceding \$
- d) all of the mentioned

Q.No.4

srwxr-xrw- is a

- a) internet socket file
- b) unix domain socket file
- c) symbolic link
- d) shared file

Q.No.5

Binary or executable files are:

- a) Regular files
- b) Device files
- c) Special files
- d) Directory files

Q.No.6

Which directory contain device special files?

- a) /etc
- b) /etc/dev
- c) /root/bin
- d) /dev

Q.No.7

All device files are stored in which directory?

- a) /etc
- b) /bin
- c) /dev
- d) /usr

Q.No.8

Command used to count number of character in a file is

- a) grep
- b) wc
- c) count
- d) cut

Q.No.9

Command used to determine the path of an executable file is

- a) which
- b) where
- c) wexec

d) what

Q.No.10

Which of the following command output contains userid?

a) ls

b) help

c) date

d) ls -l

Q.No. 11

Which command is used to display all the files including hidden files in your current and its subdirectories ?

a) ls -aR

b) ls -a

c) ls -R

d) ls -l

Q.No. 12

Which of the following commands can be used to change default permissions for files and directories at the time of creation

a) Chmod

b) Chown

c) Umask

d) Chgrp

Q.No. 13

Which tar command option is used to list the files in a tape archive format?

a) cvf

b) tvf

c) xvf

d) ovf

Q.No. 14

Which of the following is a type of cloud computing service?

a) Service-as-a-Software (SaaS)

b) Software-and-a-Server (SaaS)

c) Software-as-a-Service (SaaS)

d) Software-as-a-Server (SaaS)

Q.No. 15

Applications and services that run on a distributed network using virtualized resources is known as _____

a) Parallel computing

b) Soft computing

c) Distributed computing

d) Cloud computing

Q.No. 16

Which of the following is an example of a PaaS cloud service?

a) Heroku

b) AWS Elastic Beanstalk

- c) Windows Azure
- d) All of the above

Q.No.17

Point out the wrong statement.

- a) Azure enables .NET Framework applications to run over the Internet
- b) Cloud Computing has two distinct sets of models
- c) Amazon has built a worldwide network of datacenters to service its search engine
- d) None of the mentioned

Q.No.18

Which of the following architectural standards is working with cloud computing industry?

- a) Web-application frameworks
- b) Service-oriented architecture
- c) Standardized Web services
- d) All of the mentioned

Q.No. 19.

Which of the following is the most essential element in cloud computing by CSA?

- a) Virtualization
- b) Multi-tenancy
- c) Identity and access management
- d) All of the mentioned

Q.No. 20.

Which of the following model consists of the service that you can access on a cloud computing platform?

- a) Deployment
- b) Service
- c) Application
- d) None of the mentioned

PML:

Q.No.1

The k-means algorithm is a

- (A) Supervised learning algorithm
- (B) Unsupervised learning algorithm
- (C) Semi-supervised learning algorithm
- (D) Weakly supervised learning algorithm

Q.No.2

For unsupervised learning we have ____ model.

- (A) interactive
- (B) predictive
- (C) descriptive

(D) prescriptive

Q. No. 3

Which of the following is true about SVM?

- (A) It is useful only in high-dimensional spaces
- (B) It requires less memory
- (C) SVM does not perform well when we have a large data set
- (D) SVM performs well when we have a large data set

Q. No. 4

When you find many noises in data, which of the following options would you consider in

KNN?

- (A) Increase the value of k
- (B) Decrease the value of k
- (C) Noise does not depend on k
- (D) $k = 0$

Q. No. 5

A multiple regression model has:

- (A) Only one independent variable
- (B) More than one independent variable
- (C) More than one dependent variable
- (D) None of the above

Q. No. 6

The average squared difference between classifier predicted output and actual output

- (A) mean squared error
- (B) root mean squared error
- (C) mean absolute error
- (D) mean relative error

Q. No. 7

Which of the following is not a benefit of using Grid search?

- (A) It can be applied to non-differentiable functions.
- (B) It can be applied to non-continuous functions
- (C) It is easy to implement
- (D) It runs reasonably fast for multiple linear regression

Q. No. 8

Data used to build a data mining model

- (A) validation data
- (B) training data
- (C) test data
- (D) hidden data

Q.No.9

Introducing a non-essential variable into a linear regression model may result in:

- (1).Increase in R-square, (2).Decrease in R-square

- (A) Only 1 is correct
- (B) Only 2 is correct

- (C) Either 1 or 2
- (D) None of these

Q.No.10

Which data is used to tune the parameters of supervised learning model

- (A) training
- (B) test
- (C) verification
- (D) validation

Q. No. 11

Which of the following is not TRUE for regression?

- (A) It relates inputs to outputs
- (B) It is used for prediction
- (C) It may be used for interpretation
- (D) It discovers causal relationships

Q. No. 12

Which of the following statements regarding outliers is correct?

- (A) Outliers should be identified and removed from a dataset
- (B) Outliers should be part of the training dataset but should not be present in the test data
- (C) Outliers should be part of the test dataset but should not be present in the training data
- (D) The nature of the problem determines how outliers are used

Q. No. 13

Each data instance is assigned a conditional probability value using this method.

- (A) linear regression
- (B) logistic regression
- (C) simple regression
- (D) multiple linear regression

Q. No. 14

Neural networks

- (A) optimize a convex cost function
- (B) always output values between 0 and 1
- (C) can be used for regression as well as classification
- (D) can be used in an ensemble

Q. No. 15

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

- (A) PCA
- (B) K-Means
- (C) KNN
- (D) None of the above

Q. No. 16

Which of the following is a performance measure for regression?

- (A) Accuracy
- (B) Recall
- (C) Error rate
- (D) RMSE

Q.No.17

Conversion of a text corpus to a numerical representation is done using ____ process.

- (A) Tokenization
- (B) Normalization
- (C) Vectorization
- (D) None of the above

Q.No.18

PCA is a technique for

- (A) Feature extraction
- (B) Feature construction
- (C) Feature selection
- (D) None of the above

Q. No. 19

Predicting whether a tumour is malignant or benign is an example of?

- (A) Unsupervised Learning
- (B) Supervised Regression Problem
- (C) Supervised Classification Problem
- (D) Categorical Attribute

Q. No. 20

For categorical data, ____ cannot be used as a measure of central tendency.

- (A) Median
- (B) Mean
- (C) Quartile
- (D) None of the above

PYTHON and R:

Q.No.1

Which type of Programming does Python support?

- a) object-oriented programming
- b) structured programming
- c) functional programming
- d) all of the mentioned

Q.No.2

All keywords in Python are in _____

- a) Capitalized
- b) lower case
- c) UPPER CASE
- d) None of the mentioned

Q. No. 3

Which keyword is used for function in Python language?

- a) Function
- b) Def
- c) Fun

d) Define

Q. No. 4

What is the order of precedence in python?

- a) Exponential, Parentheses, Multiplication, Division, Addition, Subtraction
- b) Exponential, Parentheses, Division, Multiplication, Addition, Subtraction
- c) Parentheses, Exponential, Multiplication, Division, Subtraction, Addition
- d) Parentheses, Exponential, Multiplication, Division, Addition, Subtraction

Q. No. 5

What does pip stand for python?

- a) unlimited length
- b) all private members must have leading and trailing underscores
- c) Preferred Installer Program
- d) none of the mentioned

Q. No. 6

Which of the following functions is a built-in function in python?

- a) factorial()
- b) print()
- c) seed()
- d) sqrt()

Q. No. 7

Which of the following is not a core data type in Python programming?

- a) Tuples
- b) Lists
- c) Class
- d) Dictionary

Q. No. 8

What is the order of namespaces in which Python looks for an identifier?

- a) Python first searches the built-in namespace, then the global namespace and finally the local namespace
- b) Python first searches the built-in namespace, then the local namespace and finally the global namespace
- c) Python first searches the local namespace, then the global namespace and finally the builtin namespace
- d) Python first searches the global namespace, then the local namespace and finally the builtin namespace

Q.No.9

Which one of the following is not a keyword in Python language?

- a) pass
- b) eval
- c) assert
- d) nonlocal

Q.No.10

Which of the following statements is used to create an empty set in Python?

- a) ()

- b) []
- c) { }
- d) set()

Q. No. 11

To add a new element to a list we use which Python command?

- a) list1.addEnd(5)
- b) list1.addLast(5)
- c) list1.append(5)
- d) list1.add(5)

Q. No. 12

Which one of the following is the use of function in python?

- a) Functions don't provide better modularity for your application
- b) you can't also create your own functions
- c) Functions are reusable pieces of programs
- d) All of the mentioned

Q. No. 13

Which of the following is a feature of Python DocString?

- a) In Python all functions should have a docstring
- b) Docstrings can be accessed by the `__doc__` attribute on objects
- c) It provides a convenient way of associating documentation with Python modules, functions, classes, and methods
- d) All of the mentioned

Q. No. 14

What will be the output of the following Python code?

```
print("Hello {0[0]} and {0[1]}".format(('foo', 'bin')))
```

- a) Hello ('foo', 'bin') and ('foo', 'bin')
- b) Error
- c) Hello foo and bin
- d) None of the mentioned

Q. No. 15

What is output of `print(math.pow(3, 2))`?

- a) 9.0
- b) None
- c) 9
- d) None of the mentioned

Q. No. 16

Which of the following is the use of `id()` function in python?

- a) Every object in Python doesn't have a unique id
- b) In Python `id` function returns the identity of the object
- c) None of the mentioned
- d) All of the mentioned

Q.No.17

The process of pickling in Python includes _____

- a) conversion of a Python object hierarchy into byte stream
- b) conversion of a datatable into a list
- c) conversion of a byte stream into Python object hierarchy

d) conversion of a list into a datatable

Q.No.18

The character Dot (that is, '.') in the default mode, matches any character other than

-
- a) caret
 - b) ampersand
 - c) percentage symbol
 - d) newline

Q. No. 19

Which of the following functions clears the regular expression cache?

- a) re.sub()
- b) re.pos()
- c) re.purge()
- d) re.subn()

Q. No. 20

Which of the following functions results in case insensitive matching?

- a) re.A
- b) re.U
- c) re.I
- d) re.X