ML-Worksheet2

1. Movie Recommendation systems are an example of:

Ans: a

2. Sentiment Analysis is an example of:

Ans: d

3.Can decision trees be used for performing clustering?

Ans: a

4.Which of the following is the most appropriate strategy for data cleaning before performing clustering analysis, given less than desirable number of data points:

Ans: a

5.What is the minimum no. of variables/ features required to perform clustering?

Ans: b

6.For two runs of K-Mean clustering is it expected to get same clustering results?

Ans: b

7.Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?

Ans: a

8.Which of the following can act as possible termination conditions in K-Means?

Ans: d

9. Which of the following algorithms is most sensitive to outliers?

Ans: a

10. How can Clustering (Unsupervised Learning) be used to improve the accuracy of Linear Regression model (Supervised Learning):

Ans: d

11. What could be the possible reason(s) for producing two different dendrograms using agglomerative clustering algorithms for the same dataset?

Ans: d

12. 12. Is K sensitive to outliers?

Ans: Yes K is sensitive to outliers because it is a mean and is easily influenced by extremes.

13. Why is K means better?

Ans:K means is better as it can gerealize to clusters of various shapes & sizes, can easily adapt to new examples, guarantees convergence, can warm start the positions of centroids, and is relatively simeple to use.

14.Is K means a deterministic algorithm?

Ans:The basic K means is a non deterministic alogrithm, i.e. running the alogirithm many times on the same data can give different results.

WORKSHEET 2 SQL

Ql to Q13 have only one correct answer. Choose the correct option to answer your question.

1. Which of the following constraint requires that there should not be duplicate entries?

D) Unique

2. Which of the following constraint allows null values in a column?

D) None of them

3. Which of the following statements are true regarding Primary Key?

A) Each entry in the primary key uniquely identifies each entry or row in the table

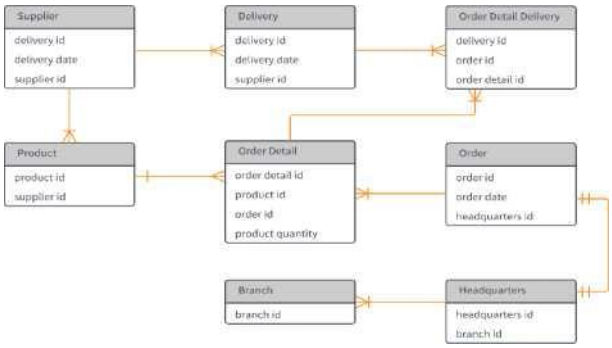
4. Which of the following statements are true regarding Unique Key?

D)All of the above

5. Which of the following is/are example of referential constraint?

B) Foreign Key

For Questions 6-13 refer to the below diagram and answer the questions:



6. How many foreign keys are there in the Supplier table?

D)1

7. The type of relationship between Supplier table and Product table is:

D) many to many

8. The type of relationship between Order table and Headquarter table is:

B) many to one

9. Which of the following is a foreign key in Delivery table?

D) None of them

10. The number of foreign keys in order details is:

D)2

11. The type of relationship between Order Detail table and Product table is:

B) many to one

12. DDL statements perform operation on which of the following database objects?

D) None of them

13. Which of the following statement is used to enter rows in a table?

A) Insert in to

Q14 and Q15 have one or more correct answer. Choose all the correct option to answer your question.

14. Which of the following is/are entity constraints in SQL?

B) Unique C) Primary Key D) Null

15. Which of the following statements is an example of semantic Constraint?

A) A blood group can contain one of the following values - A, B, AB and O.

B) A blood group can only contain characters

STATISTICS WORKSHEET-2

1. What represent a population parameter?

C) both

2. What will be median of following set of scores (18,6,12,10,15)?

C) 12

3. What is standard deviation?

D)All of the above

4.The intervals should be _____ in a grouped frequency distribution

C)Both of these

5.What is the goal of descriptive statistics?

D) All of these

6.A set of data organized in a participant by variables format is called

B) Data set

7. In multiple regression,_____ independent variables are used

A) 2 or more

8. Which of the following is used when you want to visually examine the relationship between 2 quantitative variables?

B) Scatterplot

9.Two or more groups means are compared by using

D) Analysis of variance

10. _____is a raw score which has been transformed into standard deviation units?

A) Z-score

11. _____is the value calculated when you want the arithmetic average?

A) mean

12. Find the mean of these set of number (4,6,7,9,2000000)?

D) 400005.2

13. _____ is a measure of central tendency that takes into account the magnitude of scores?

D) Mean

14. _____ focuses on describing or explaining data whereas _____involves going beyond immediate data and making inferences

A) Descriptive and inferences

15. What is the formula for range?

D)H-L