**Statistic Worksheet -2**

1. What represents the population parameter?
2. SD
3. What will be the meadian of following set of scores (18, 6,12,10,15) ?
4. 12
5. What is standard deviation?
6. All of the above
7. The intervals should be \_\_\_\_\_\_\_\_\_ in grouped frequency distribution?

c. both of these.

5. What is the goal of descriptive statistics?

b) summarizing and explaining specific set of data.

6. A set of data organized in a participants 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 two 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?

d) SDU score

11. \_\_\_\_\_\_\_ is the value calculated when you want the airthmatic average?

a) meadian

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 score?

b) mode

14. \_\_\_\_\_\_\_ focuses on describing or explaining data whereas \_\_\_\_\_\_ involves going beyond immediate data and making inferences?

A) descriptive and interfernces

15. what is the formula for range?

D) H-L

**SQL Worksheet -2**

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

D) Unique

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

C) Null

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?

c) Multuple columns can have single unique key together.

5. which of the following is / are referential comstraint?

b) Foreign Key

6. How many foreign keys are there in supplier table

b)3

7.The type of relationship between supplier table and product table?

b) many to one

8. The type of relationship between oreder table and headquater table?

d) many to many

9. which of the followimg is foreign key in delvery table?

a) delivery ID

10. The number of foreign keys in order details ?

b)1

11. Tge type of relationship between order detail table and product detail table?

a) One to many

12. DDL statement operation on which of followimg database object?

c) tables

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

a) insert into

14. hich of the following is entity constraint in sql?

c)primary key

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

a) a blood group can contain one of the following values A, B, AB, O

**Machine Learning Worksheet 2**

1. Movie Reccomdation are an example of ?

b) Classification and clustering

2. Sentiment analysis is an example?

d) 1,2 and 4

3. can decision tree used for clustering?

a)True

4. Which of the following is most appropriate data cleaning statergies before plannind clustering?

a) capping and flooring of variables

5. what is the minimum numbers of variables required to perform clustering?

b)1

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

a)yes

7. Is it possible that assignments of observations to clusters does not change iterations in k means?

a) Yes

8.Which of the following can act as possible termination conditiond in k means?

d)All of the above

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

a) K means clustering algorithm

10. How can clustering(unsupervised ML) can used to improve the accuracy of Linear Regression Model(Supervised ML)

d) All of the above

11. What ould be the possible reason for producing two different dendographs clustering algorithm for the same dataset

d) all of the above

12. Is K sensitive to outliers?

-- The K- Means Clustering algorithms is sensitive to outliers because mean is easily influenced by extrame values. K medoids clustering is a variant of k meansthat is more robust to noises and outliers.instaed of using the mean points as the center of clusters , K-medoids uses an actual point in the cluster to represent it. Medoid is the most centrally located object in the cluster with minimum sum of distance to other points.

13. Why is K means better?

* Relatively simple to implement.
* Scales to large datasets.
* Gurantee convergence
* Can warm start the positins of centroids.
* Easily adapts to new examples.
* Generalises to cluster to different shapes and sizes,

14. Is K means is a deterministic algorithm?

K means has many drawbacks too. One of the significant drawback of K means is its non-deterministic nature. K means start with the random set of data pointsas initial centroids. This random selection influences the quality of the resulting clusters.