WORKSHEET SET 2: SOLUTIONS

STATISTICS WORKSHEET-2

Q1. (C) Both

Q2. (C) 12

Q3. (D) All of the above

Q4. (C) Both of these

Q5. (D) All of the above

Q6. (B) Dataset

Q7. (A) 2 or more

Q8. (B) Scatterplot

Q9. (D) Analysis of variance

Q10. (A) Z-Score

Q11. (C) Mean

Q12. (D) 400005.2

Q13. (B) Mode

Q14. (A) Descriptive & Inferences

Q15. (D) H – L

SQL WORKSHEET-2

Q1. (D) Unique

Q2. (A) Primary Key

Q3. (C) There can be null values in Primary key.

Q4. (A) There should not be any duplicate entries.

Q5. (A) Foreign Key

Q6. (B) 3

Q7. (D) Many to Many.

Q8. (C) One to One

Q9. (B) Supplier ID.

Q10. (C) 3

Q11. (B) Many to One.

Q12. (C) Table.

Q13. (A) Insert into.

Q14. (C) Primary Key, (B) Unique

Q15. (A) A blood group can contain one of the following values - A, B, AB and O.

(B) A blood group can only contain characters.

(C) A blood group cannot have null values.

MACHINE LEARNING WORKSHEET-2

Q1. (B) 1 & 2

Q2. (D) 1, 2 & 4

Q3. (A) True

Q4. (A) 1 only.

Q5. (B) 1

Q6. (B) No

Q7. (A) Yes

Q8. (D) All of the above.

Q9. (A) K-Means

Q10. (D) All of the above.

Q11. (D) All of the above.

Q12. Yes, K- means is sensitive to outliers. An outlier is a point of data which is different from other data points. K-means is a clustering algorithm, outliers stay far from these clusters of data. Due to this K-Means is highly influenced by these outliers.

Q13. K-Means is an unsupervised learning algorithm. It is better because it is an easy and relatively simple algorithm to implement. It can be used on large datasets. It also adapts to changes in data easily.

Q14. No, K-Means is not a deterministic algorithm. Since, K-Means begins with a random set of data points, this influences the quality of the clusters which makes determination difficult.