Statistics Worksheet – 2

Q1) C → Both

Q2) C → 12

Q3) D → All of the above

Q4) C → Both

Q5) B → Summarizing and explaining a specific set of data

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) D → Mean

Q14) A → Descriptive and Inferences

Q15) D → H – L

SQL Worksheet – 2

Q1) D → Unique

Q2) A → Primary Key

Q3) A → Each entry in primary key uniquely identifies each entry or row in the table

Q4) D → All

Q5)

Q6) B → 3

Q7) B → many to one

Q8) B → many to one

Q9) A → Delivery ID

Q10) D → 2

Q11) D → Many to Many

Q12) C → Table

Q13) A → Insert into

Q14) B & C → Primary key and Unique

Q15) A,C & D

Machine Learning Worksheet – 2

Q1) A → 2 only

Q2) D → 1,2 and 4

Q3) A → True

Q4) A → Capping and Flouring of variable

Q5) B → 1

Q6) B → No

Q7) A → Yes

Q8) D → All

Q9) A → K-means clustering algorithm

Q10) D → All

Q11) D → All

Q12) K Means Clustering is most sensitive to outliers as it uses the mean of Cluster Data points, and means is most sensitive to outliers.

Q13) a) It is simple to implement k-means.

b) Can easily adjust to the changes

c) It is suitable for large no of datasets

d) Works well in hyper-spherical clusters

e) It doesn't take more time in classifying similar characteristics in data like hierarchical algorithms

Q14) k means is non-deterministic. As the algo is worked on several time on the data every time it gives different results. This is due to random selection of data points.