**SQL:**

**Question 1**:

**Answer**:From runners table id 1,4 & 5 will be picked with their corresponding name

**Question 2**:Write a query to fetch values in table test\_a that are and not in test\_b without using the NOT keyword.

**Answer**: SELECT \* FROM test\_a a

LEFT JOIN test\_b b ON a.value = b.value

WHERE b.value IS NULL;

**Question: 3:**Write a query to to get the list of users who took the a training lesson more than once in the same day, grouped by user and training lesson, each ordered from the most recent lesson date to oldest date.

**Answer:**

select username from users where id in (SELECT user\_id, user\_training\_id,training\_date, COUNT(\*) AS num\_lessons

FROM training\_details

GROUP BY user\_id, user\_training\_id, training\_date

HAVING COUNT(\*) > 1

ORDER BY training\_date DESC;)

**Question4:**

**Answer:**

Select Emp\_name,Avg(salary) group by Manager\_id;

**Statistics:**

**Question: 1:What is the meaning of six sigma in statistics? Give proper example**

**Answer:**Six Sigma is a methodology used in quality management and process improvement. It aims to improve the quality of process outputs by identifying and eliminating the causes of defects and minimizing variability in manufacturing and business processes. The term "Six Sigma" refers to a level of quality that is statistically defined as 3.4 defects per million opportunities (DPMO).

example:

1.Here's how Six Sigma could be applied to improve call center efficiency:

2.Define the process: Identify the steps involved in handling customer calls, including call routing, agent availability, and customer service interactions.

3.Measure: Collect data on the average wait time for customers before their calls are answered. Measure other relevant metrics such as call volume, agent utilization, and customer satisfaction scores.

4.Analyze: Analyze the data to identify factors contributing to longer wait times. This could include peak call times, staffing levels, call routing inefficiencies, and agent performance.

5.Improve: Implement changes to the call center operations to reduce wait times and improve customer service. This could involve optimizing call routing algorithms, adjusting staffing levels to match call volume fluctuations, providing additional training for agents to handle calls more efficiently, and implementing self-service options for common inquiries.

**Question: 2:What type of data does not have a log-normal distribution or a Gaussian distribution? Give proper example**

**Answer:**Data that doesn't follow a log-normal distribution or a Gaussian distribution is often referred to as non-normal data. There are various types of non-normal distributions, and data can exhibit different distributions depending on its nature.

Here are a few examples**:Uniform Distribution,Exponential Distribution,Poisson Distribution,Binomial Distribution**

**Question: 3:What is the meaning of the five-number summary in Statistics? Give proper example**

**Answer:**The five-number summary is a descriptive statistics technique used to summarize the distribution of a dataset. It consists of five values: the minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum. These values provide insights into the spread and central tendency of the data.

Here's what each component of the five-number summary represents:

Minimum: The smallest value in the dataset.

First Quartile (Q1): The value below which 25% of the data falls. It is the median of the lower half of the dataset.

Median (Q2): The middle value of the dataset when it is sorted in ascending order. Also known as the second quartile, it divides the dataset into two equal halves.

Third Quartile (Q3): The value below which 75% of the data falls. It is the median of the upper half of the dataset.

Maximum: The largest value in the dataset.

**Question: 4:What is correlation? Give an example with a dataset & graphical representation on jupyter Notebook**

**Answer:**

**Deep learning**

**Question: 1**

**(a) Explain how you can implement DL in a real-world application.**

**Answer:**Implementing deep learning (DL) in a real-world application involves several key steps, including problem definition, data collection and preprocessing, model selection and architecture design, training and optimization, deployment, and monitoring. Here's a high-level overview of the process

**(b) What is the use of Activation function in Artificial Neural Networks? What would be the problem if we don't use it in ANN networks.**

**Answer:**

Activation functions play a crucial role in artificial neural networks (ANNs) by introducing non-linearity to the model. The primary purpose of activation functions is to introduce non-linear transformations to the input data, allowing neural networks to learn complex patterns and relationships in the data.

If we don't use activation functions in ANNs, several problems can arise:

1)Loss of Expressive Power: Without non-linear activation functions, ANNs would only be able to represent linear transformations of the input data. This limitation would severely restrict their ability to learn complex patterns and relationships in the data.

2)Inability to Learn Complex Functions: Many real-world problems require ANNs to learn non-linear relationships between input and output variables. Without activation functions, ANNs would be unable to capture the non-linearities present in the data, leading to poor performance on complex tasks.

3)Gradient Descent Issues: Activation functions are essential for gradient-based optimization algorithms like backpropagation to work effectively. Without activation functions, gradients would become constant or vanish, making it challenging to update the network parameters and optimize the loss function effectively.