## **PYTHON**

All Question are done in .ipynb file under the name of Python Questions

## 

## 

## 

## 

## **SQL QUESTIONS**

QNO1.

ANS). The above query will return nothing because the races table contain null winner\_id.

This query selects all rows from the table "runners" where the "id" column does not exist in the result of the subquery, which selects the "winner\_id" from the table "races". In other words, it returns all the runners who have not appeared in the "winner\_id" column of the "races" table, implying that they have not won any race.

Alternative Query

SELECT \* FROM runners

WHERE NOT EXISTS (SELECT 1 FROM races WHERE races.winner\_id = runners.id);

QNO2.

Ans).

SELECT a.id

FROM test\_a a

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

WHERE b.id IS NULL;

QNO3.

Ans). SELECT u.username, td.training\_id, td.training\_date

FROM users u

INNER JOIN training\_details td ON u.user\_id = td.user\_id

GROUP BY u.username, td.training\_id

HAVING COUNT(\*) > 1

ORDER BY u.username, td.training\_date DESC;

QNO4.

SELECT manager\_id, AVG(salary) AS Average\_Salary\_Under\_Manager

FROM Employee

GROUP BY manager\_id;

## **STATISTICS**

QNO1. What is the meaning of six sigma in statistics? Give proper example

ANS). Six Sigma is a set of techniques and tools for process improvement. It seeks to improve the quality of the output of a process by identifying and removing the causes of defects and minimizing variability in manufacturing and business processes.

Let's consider a manufacturing company that produces bottles. They want to ensure consistent bottle volume to avoid under or overfilling.

They can measure the volume of each bottle produced and create a statistical distribution of the volumes.

If the distribution has a small standard deviation (sigma), it indicates minimal variation in bottle volume.

A Six Sigma process, in this context, would aim to achieve a very low standard deviation, resulting in almost all bottles having the desired volume, minimizing waste and ensuring product consistency.

QNO2. What type of data does not have a log-normal distribution or a Gaussian distribution? Give proper example

ANS). Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well.

EXAMPLE:- The number of apples on a tree. It can't be negative or a fraction, and there must be at least zero apples (not applicable). It would likely follow a Poisson distribution, which is suited for counting data with a non-negative integer range.

QNO3. What is the meaning of the five-number summary in Statistics? Give proper example

ANS). The five-number summary is a descriptive statistic in statistics that provides a concise summary of the distribution of a dataset. It consists of five key values:

* Minimum: The smallest value in the dataset.
* First Quartile (Q1): The value below which 25% of the data falls. It is also known as the lower quartile.
* Median (Q2): The middle value of the dataset when it is ordered from smallest to largest. It represents the value below which 50% of the data falls.
* Third Quartile (Q3): The value below which 75% of the data falls. It is also known as the upper quartile.
* Maximum: The largest value in the dataset.

Consider a dataset representing exam scores for 50 students: {45, 67, 88, 72, 58, 95, 82, 63, 71, 48, 90, 85, 78, 69, 100, 55, 87, 75, 61}

* Minimum Value: 45
* First Quartile (Q1): 58 (25% of students scored 58 or lower)
* Median (Q2): 72 (50% of students scored 72 or lower)
* Third Quartile (Q3): 87 (75% of students scored 87 or lower)
* Maximum Value: 100

QNO4. What is correlation? Give an example with a dataset & graphical representation on jupyter Notebook?

ANS). Answer is on .ipynb file under the name of statistics

## 

## 

## **DEEP LEARNING**

QNO1. a) Explain how you can implement DL in a real-world application.

(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.

ANS)

A). Deep Learning (DL) has found extensive applications in various real-world scenarios due to its ability to learn from large amounts of data and extract meaningful patterns. Here's a general overview of how you can implement DL in a real-world application:

Problem Definition: Start by clearly defining the problem you want to solve. Whether it's image recognition, natural language processing, recommendation systems, or any other task, having a well-defined problem is crucial.

Data Collection and Preprocessing: Gather relevant data for your problem domain. This could be images, text, numerical data, etc. Clean the data, handle missing values, and preprocess it to make it suitable for training.

Model Selection: Choose an appropriate deep learning architecture for your problem. This could be convolutional neural networks (CNNs) for image-related tasks, recurrent neural networks (RNNs) for sequence data, or transformer models for natural language processing (NLP).

Model Training: Split your data into training, validation, and testing sets. Train your model using the training data and validate its performance using the validation set. Fine-tune hyperparameters such as learning rate, batch size, and network architecture to improve performance. This step often requires significant computational resources, so training may be done on powerful GPUs or TPUs.

Model Evaluation: Evaluate your trained model on the test set to assess its performance. Metrics such as accuracy, precision, recall, F1-score, or mean squared error are commonly used depending on the nature of the problem.

Deployment: Once you have a satisfactory model, deploy it into production.

Monitoring and Maintenance: Continuously monitor the performance of your deployed model in real-world scenarios. Retrain the model periodically with new data to keep it up-to-date and maintain its performance. Monitor for concept drift and update the model accordingly.

B)

Activation functions play a crucial role in artificial neural networks (ANNs) by introducing non-linearity into the network, allowing it to learn complex patterns and relationships in the data. Without activation functions, ANNs would essentially be linear models, unable to capture the intricate features present in many real-world datasets.

If activation functions were not used in ANN networks:

* Linear Behavior: The network would behave as a linear combination of its inputs, limiting its ability to learn complex patterns and relationships in the data.
* Limited Representational Power: Without the ability to introduce non-linearity, ANNs would be limited in their representational power, severely restricting their ability to model real-world phenomena accurately.
* Gradient Vanishing or Explosion: Backpropagation, the process by which the network learns from data, would be hindered due to the absence of non-linear activation functions. This could lead to vanishing or exploding gradients, making training unstable or impractical.