SQL Data Analyst interview Question and Answer

1. Mention the difference between Data Mining and Data Profiling?

Data mining

Data mining is the process of discovering relevant information that has not yet been identified before. In data mining, raw

Data is converted into valuable information

Data profiling

Data profiling is done to evaluate a dataset for its uniqueness, logic, and consistency. It cannot identify inaccurate or incorrect data value

1. Define the term Data wrangling in data analytics.

It involves discovering, structure, cleaning, enriching, validating, and analyzing data. This process can turn, and map out

Larges amounts of data extracted from various sources into a more useful format

1. What are the common problems that data analyst encounter during analysis?

Handling duplicates

Collecting the meaningful right data and the right time

Handling data purging and storage problems.

Making data secure and dealing with compliance issues.

1. What are the best methods for data cleaning?

Create a data cleaning plan by understanding where the common errors take place and keep all the communications open

Before working with data , identify and remove the duplicates. This will leads to an easy and effective data analysis process

Focus on the accuracy of the data. Set cross-field validation, maintain the value types of data, and provide mandatory constraints.

Normalize the data at the entry point so that it is less chaotic. You will be able to ensure that all information is standardized, leading to fewer errors on entry.

1. What are the various steps involved in any analytics project?

Understanding the problems

An entire record is executed from analysis if any single value is missing

Collecting data

Gather the right data from various sources and other information based on your priorities.

Cleaning data

Clean the data to remove unwanted, redundant, and missing values, and make it ready for analysis.

Exploring and Analyzing data

Use data visualization and business intelligence tools, data mining techniques, and predictive modelling to analyze data.

Interpreting the results

Interpret the results to find out hidden patterns, future trends, and gain insights.

1. What is the significant of Exploratory Data Analysis (EDA)

Exploratory data analysis (EDA) helps to understand the data better

It helps you obtain confidence in your data to a point where you’re ready to engage a machine learning algorithm

It allows you to refine your selection of feature variables that will be used later for model building.

You can discover hidden trends and insights from the data.

1. Explain descriptive, predictive, and prescriptive analytics.

Descriptive:

It provides insight into the past to answer “what has happened”

Uses data aggregation and data mining techniques

Predictive

Understands the future to answer “what could happen”

Uses statistical models and forecasting techniques.

Prescriptive

Suggest various courses of action to answer “what should you do”

Uses simulation algorithms and optimization techniques to advise possible outcomes.

1. What are the different types of sampling technique used by data analyst.

There are majorly five types of sampling methods:

Simple random sampling

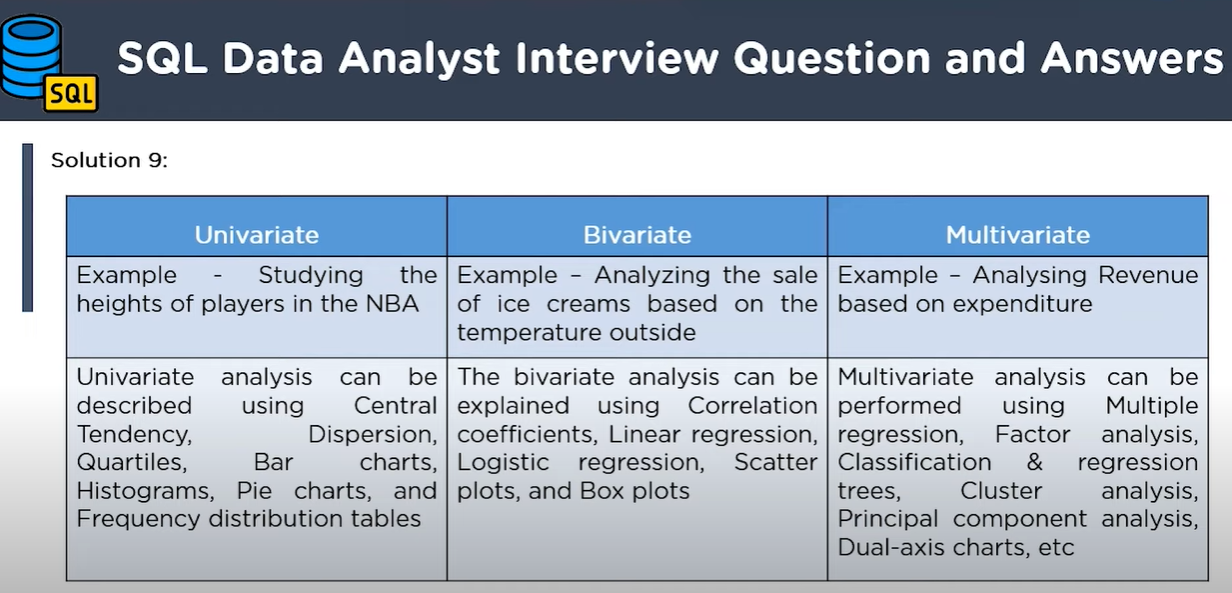
Systematic sampling

Cluster sampling

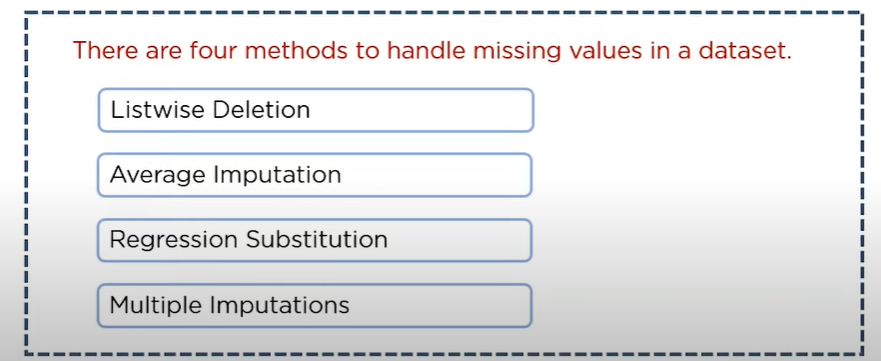
Stratified sampling

Judgmental or purposive sampling

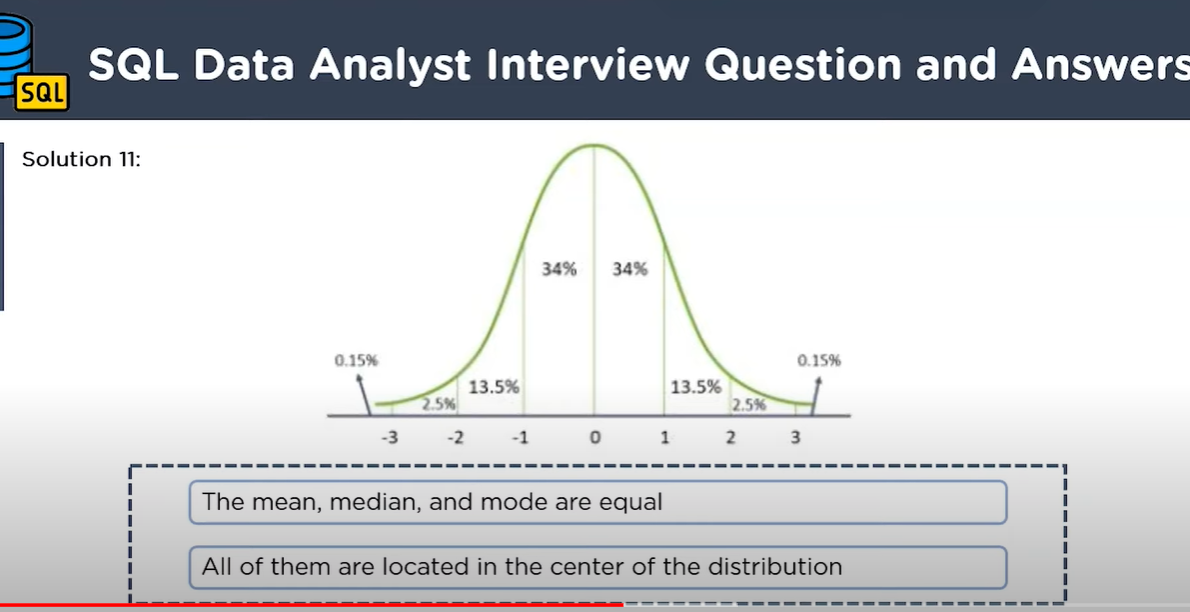
1. Describe univariate, bivariate, and multivariate analysis?

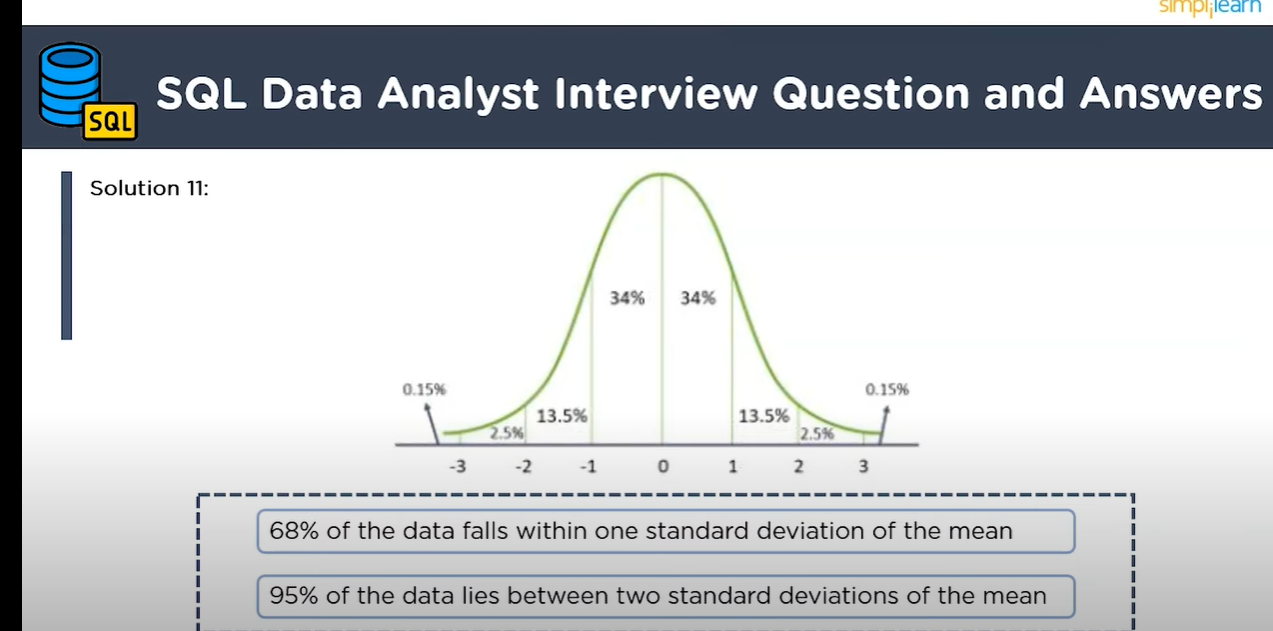


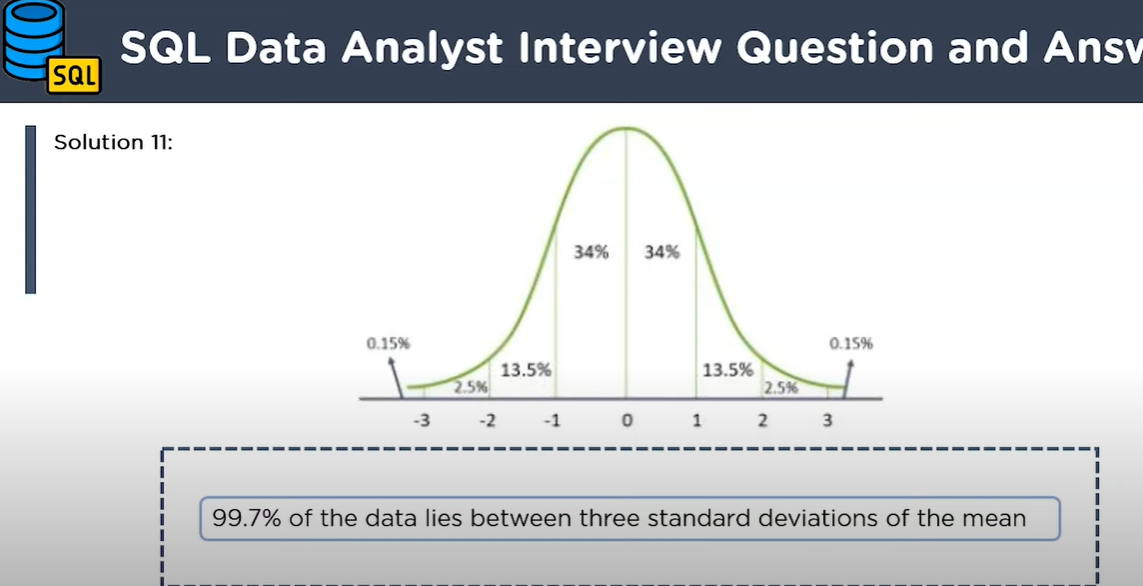
1. How can you handle missing value in a dataset?



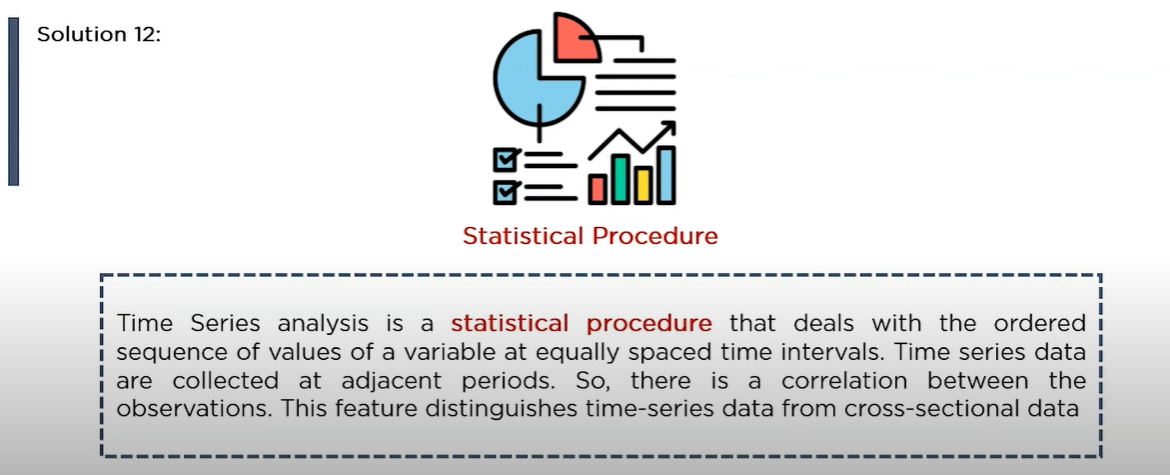
1. Explain the term normal distribution ??



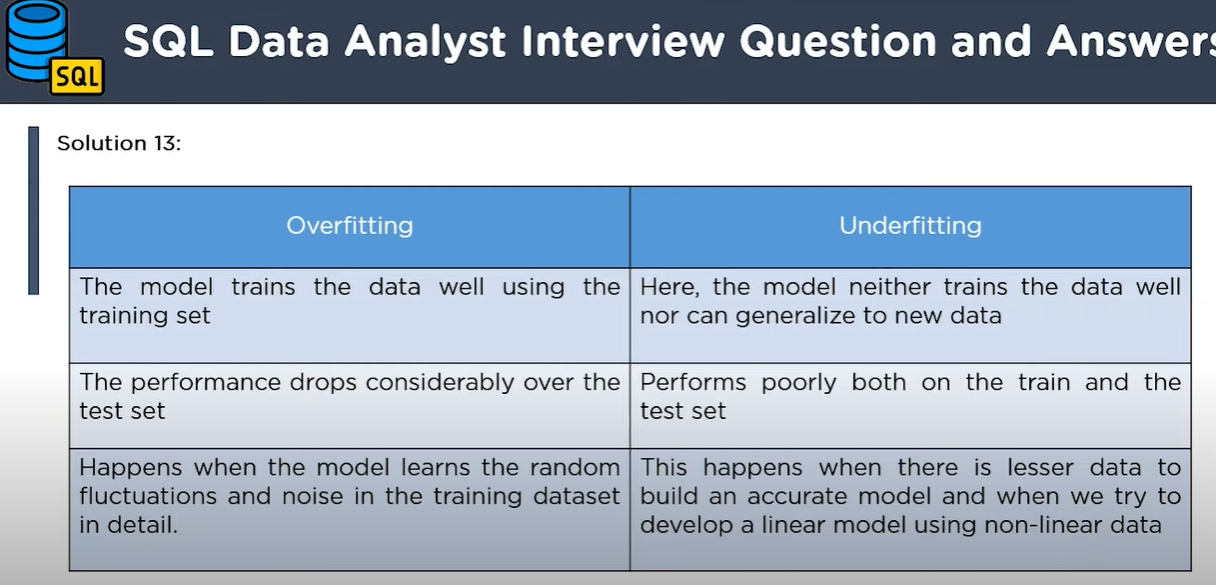




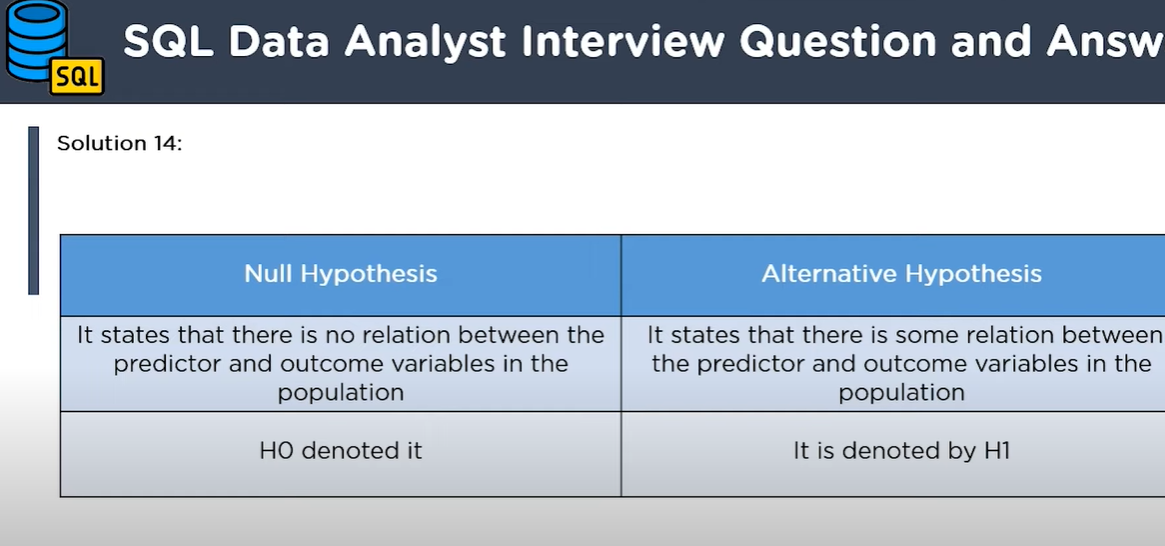
1. What is time series Analysis.



1. How is overfitting different from underfitting?



1. What are different types of Hypothesis Testing?



1. Explain the Type 1 and type 2 errors in statistics?

