**Introduction:**

Hello sir/ma’m,

Greeting of the day,

First of all, thank you for allowing me to introduce myself. My name is Shailesh Suthar. I am from Pali, Rajasthan. I have completed my graduation from MNIT, Jaipur. I have one and half years of full-time work experience, along with 8 months of paid internship experience.

In my current role at Saarthee, I’m working for COMCAST Smart solution team for building their products. This project is divided into three components:

* The first one includes complete end-to-end development of the Sales-Hub CRM application on Dynamic 365 helps to track their lead and conversions.
* The second component includes the development of power BI dashboard from whatever the data comes from sales-hub application on dynamic 365.
* The last one includes the implementation of campaign automation workflow using Knime and Google analytics platform.

Apart from these:

During my academic years, I completed a 6-month internship at the intelligent automation team of JP Morgan, where I have developed a daily risk report that was previously created manually, taking around five to six hours on a daily basis. For that, I designed a Python interface along with Python library which i named it as a PyautoPDF, through which anyone can automate their report in a shorter amount of time.

I also did a 2-month summer internship at Jaipur Clubfoot Pvt Ltd, where I built a machine-learning model to optimize values for minimum dimension variation in SLA printing. The same model has been published in one of the research papers for further enhancements.

I do have some of the achievement which includes:

Recently, in the review meeting I was the top performer in the whole 2023 batch, and received an employee rating of 3.5 out of 5

And during my academic year, me and my team secured first place in a national-level competition organized by SAE India, where we designed and developed an Autonomous 4 Wheeled 4-Seater Electric vehicle for Urban Mobility based on the Indian Traffic Scenario. Which took an about ten months-four stages to complete, and secure 1.5 lac of cash winning price.

In my free time, I enjoy playing chess and volleyball.

------------------------------------------------------------------------------------------------------------------------------------------

**Family:**

In my family, there are five members: my father, who working as a labor at ongoing construction in our village; my mother, who is a housewife; my elder sister, who got married; myself; and my younger sister, who is currently pursuing her final year of a BA degree.

**Reason for leaving the current firm:**

There are a couple of reasons why I want to switch to the firm:

* Once my supervisor left the company, the work environment just wasn’t the same as before. The person who replaced them started micromanaging the team not related to the work but also politics.
* While I was grateful for the opportunities at saarthee, I felt like because I wasn’t using my skills and experience to their full potential. I'm looking for an opportunity that will let me grow, so I can take my career to the next level.
* And I do want to add-up more skills/innovation to my profession which helps me to grow exponentially.

**Strength and Weaknesses:**

* Strength includes:
  + I do have a confident and collaborative nature to work as a team and completing the business objective before deadline.
  + I usually pre-plan everything on the business objective so that I can optimize the task properly to complete it within the given span of time.
  + Innovative nature to think over any problem
  + I have one law which I named it as “Star Law” which is “everything is possible in life.”
* Along with the strengths I do have some of the weaknesses which include:
  + I usually try to think deeply over the given objective which sometimes takes more time than normally. I’m trying to work upon it so that I can reduce my overthinking timing to a certain extent.

**Can you describe yourself in three to five words?**

* Punctual, passionate, innovative, loyal to work

**Trends and Developments in Compensation: (Suggestion)**

* A compensation program refers to the set of guidelines and policies that determine how employees are paid and rewarded within an organization, which also includes how effective compensation structures play a key role in employee satisfaction, retention, and organizational success.
* And the biggest thing I have ever seen is that compensation is inversely proportional the employee churn rate.
* For to analyze the same, we can build a dashboard which contains information about the trends of how many employees get churn during the specific span of time.

**Can you describe a challenging data analysis project you worked on, the steps you took to tackle it, and the outcome?**

* One of the challenging data analysis projects I have ever been worked on is to develop the 'Sales-Hub' app on Dynamics 365 for COMCAST to track their lead and conversions, which required a deep understanding of the sales process and dynamics 365 within the organization. To tackle this, I started by studying Micro-Soft articles and gathering the requirements from stakeholders and understand pain points. I then designed a data model that could accurately capture and analyze the necessary information. Next, I leveraged my skills in Power Automate to automate processes for email notifications, AI Builder and agreement reports, BPF stage time calculator as so on. Additionally, I enhanced Dynamics 365 with responsive forms and JavaScript to improve data quality and usability. The outcome of the project was highly successful, as the 'Sales-Hub' app streamlined the sales process for COMCAST, leading to increased efficiency and better decision-making based on real-time data insights.

**What is Compensation Program:**

* A compensation program refers to the set of guidelines and policies that determine how employees are paid and rewarded within an organization. It includes base pay, bonuses, equity, and other benefits.

**Why Compensation Program:**

* I am excited about the opportunity to become a Compensation Program Analyst, where I can leverage my analytical, innovative, and data visualization skills to support the organization in making strategic, forward-looking decisions. I strongly believe that effective compensation strategies are inversely related to employee churn rates. To analyze and address this, I envision creating a dashboard that tracks trends in employee turnover over specific time periods, providing valuable insights to help reduce churn and enhance retention.

**Scenario-based Questions:**

**If you identified a significant pay disparity in your analysis, how would you address it?**

* Access the data
* Analyze potential causes
  + Historical pay
  + Market comparison
  + Performance matrix
* Evaluate policy and compliance
* Recommend Adjustment
* Develop communication strategy
* Monitor and follow-up

**Imagine the company is restructuring its bonus program. How would you approach evaluating its impact on employee motivation and retention?**

* Understand the Objectives of the Bonus Program
  + Start by clarifying the primary goals of the bonus program. Is it designed to:
  + Increase employee motivation and productivity?
  + Improve retention rates, especially for high-performing employees?
  + Align employee goals with company objectives?
* Define Key Performance Indicators (KPIs)
  + Employee Retention Rates
  + Participation Rates
  + Performance Metrics
  + Employee Engagement or Satisfaction Scores
* Conduct Pre- and Post-Program Analysis
  + Historical Benchmarking
  + Trend Analysis
* Gather Employee Feedback
* Provide Recommendations for Improvement
* Present Findings to the leadership team

**Describe a situation where you had to explain complex compensation data to a non-technical audience.**

**How do you collaborate with HR, finance, and business leaders to develop compensation recommendations?**

* Identify Key Takeaways and Audience Needs
  + For HR teams, focus on how the data affects recruitment, retention, and fairness.
  + For finance or leadership, highlight financial impact, ROI, and alignment with business objectives.
* Use Visual Aids and Clear Formatting
* Break Down Technical Terminology (use agile methodology)
* Use Scenarios or Examples
* Summarize Key Insights and Actions
* Be Prepared for Questions

**How do you approach analyzing compensation data to ensure competitiveness and equity?**

**Can you describe a time when you recommended changes to a compensation structure? What was the outcome?**

* Situation and Background:
* Recommendation Process
* Proposed Solution
* Outcome
* Example:
* "I noticed during quarterly reviews that our turnover rate was high in a specific department, particularly among mid-level employees. After conducting a market analysis and benchmarking our pay levels, I found that our base salaries were below the industry median. I recommended a shift toward a more competitive base pay model with incremental bonuses for tenure. After presenting this to leadership and gathering input from department managers, we implemented the new structure, resulting in a 15% decrease in turnover over the next six months and a notable increase in employee engagement scores."

**How would you go about conducting a market salary analysis?**

* Conducting a market salary analysis requires a methodical approach to gather relevant data, analyze it, and make informed recommendations based on your findings. Here’s a step-by-step guide on how to approach it:
* 1. Define the Scope and Objectives
* 2. Collect Reliable Data Sources
* 3. Benchmark Comparable Roles
* 4. Adjust for Geographic and Industry Differences
* 5. Analyze the Data
* 6. Develop Recommendations
* 7. Communicate Findings and Implement Changes
* Example Answer:
* "I would begin by defining the scope of the analysis, identifying which roles and departments are the focus. I’d collect market salary data from reliable sources and benchmark similar roles. After adjusting for regional and industry factors, I’d analyze the data to identify any compensation gaps and then develop actionable recommendations. Finally, I would work with leadership to communicate the findings and discuss implementation strategies."

**What tools do you use to analyze compensation data, and how proficient are you in using them?**

* Use the power bi dashboard development examples.

**Describe your experience with data visualization and reporting for compensation insights.**

* Use the power bi dashboard development examples.

**How do you ensure data accuracy and consistency when preparing compensation reports?**

* Establish Clear Data Sources and Standards
* Implement Data Validation Procedures
* Use Version Control and Documentation
* Cross-Check with Relevant Teams
* Perform Consistency Checks Before Finalizing Reports

**How do you manage data within systems to support compensation reviews?**

**Data Analysis:**

* **Differences between Data Mining and Data Profiling**
  + Data Mining: A process that involves analyzing large datasets to discover patterns, trends, and relationships among data points, often used for predictive modeling and decision-making.
  + Data Profiling: The examination of a dataset to understand its structure, content, quality, and integrity, primarily focused on summarizing data attributes and assessing data quality.
  + Methods
    - Data Mining: Uses advanced techniques like clustering, classification, regression, and association rules to find patterns and correlations.
    - Data Profiling: Employs basic statistical summaries (e.g., mean, median, frequency distribution) and data quality checks (e.g., checking for null values, consistency).
  + Applications
    - Data Mining: Applied in fields like marketing, fraud detection, and customer relationship management to identify trends and improve decision-making.
    - Data Profiling: Commonly used in data warehousing, ETL (Extract, Transform, Load) processes, and data integration projects to ensure data is accurate and consistent.
* **Define the term 'Data Wrangling' in Data Analytics**
  + Data wrangling, also known as data munging, is the process of cleaning, transforming, and organizing raw data into a structured, usable format for analysis. It involves various tasks such as handling missing values, correcting errors, standardizing formats, and combining datasets. The goal is to prepare the data for accurate and efficient analysis, ensuring it’s suitable for generating insights or feeding into machine learning models. Data wrangling is a critical step in data analytics as it addresses data quality issues, making the data more accessible and reliable for decision-making processes or predictive modeling.
* **Steps involved in an analytics project**
  + Understanding the problem > collecting the data > cleaning data > Exploring and analyzing the data > interpreting results.
* **Common problems encountered during data analysis**
  + Handling duplicate
  + Collecting the meaningful right data and the right time
  + Handling data purging and storage problems
  + Making data secure and dealing with compliance issues
* **Best methods for data cleaning**
  + Remove Duplicate Data
  + Handle Missing Values
  + Standardized Data Formats
  + Correct Structural Errors
  + Filter Outliers
  + Validate Data Accuracy
  + Normalize Data
  + Remove Irrelevant Information
  + Use Data Profiling
  + Implement Consistent Naming Conventions
* **Significance of Exploratory Data Analysis (EDA)**
  + Understand Data Structure and Patterns
  + Identify Data Anomalies and Outliers
  + Uncovered Relationships and Correlations
  + Assess Data Quality and Cleanliness
  + Provide Initial Hypotheses for Modeling
  + Inform Data Transformation Decisions
  + Visualize Key Insights
* **Descriptive, predictive, and prescriptive analytics**
  + Descriptive Analytics
    - Focuses on understanding past data and identifying patterns to summarize what has happened.
    - Uses historical data to generate insights and reports, providing a clear view of past trends and outcomes.
  + Predictive Analytics
    - Uses statistical models, machine learning, and data mining to make forecasts about future events based on historical data.
    - Helps in understanding the likelihood of future outcomes and trends, supporting proactive decision-making.
  + Prescriptive Analytics
    - Provides recommendations on actions to achieve desired outcomes, using optimization, simulation, and advanced algorithms.
    - Aims to prescribe actions by predicting outcomes and suggesting the best course of action based on possible scenarios.
* **Types of sampling techniques used by data analysts**
  + Simple random sampling
  + Systematic sampling
  + Cluster sampling
  + Stratified sampling
  + Judgmental or purposive sampling
* **Ethical considerations of data analysis**
  + Data Privacy - Protecting personal and sensitive information.
  + Informed Consent - Obtaining permission to use data.
  + Bias and Fairness - Ensuring unbiased and equitable analysis.
  + Data Accuracy - Ensuring reliable and truthful data.
  + Transparency - Being open about methods and results.
  + Data Security - Safeguarding against unauthorized access.
  + Ethical Use of Models - Preventing harm through predictive analysis.
  + Social Impact - Considering broader societal consequences.

**Statistics Questions:**

* **Handling missing values in a dataset**
  + Listwise Deletion
    - In the listwise deletion method, an entire record is excluded from analysis if any single value is missing.
  + Average Imputation
    - Take the average value of the other participants' responses and fill in the missing value.
  + Regression Substitution
    - You can use multiple-regression analyses to estimate a missing value.
  + Multiple Imputations
    - It creates plausible values based on the correlations for the missing data and then averages the simulated datasets by incorporating random errors in your predictions.
* **Explanation of Normal Distribution**
  + Normal Distribution refers to a continuous probability distribution that is symmetric about the mean. In a graph, normal distribution will appear as a bell curve.
  + The mean, median, and mode are equal
  + All of them are located in the center of the distribution
  + 68% of the data falls within one standard deviation of the mean
  + 95% of the data lies between two standard deviations of the mean
  + 99.7% of the data lies between three standard deviations of the mean (3 sigma)
* **Time Series analysis**
  + Time Series Analysis is a statistical technique used to analyze data points collected or recorded at specific time intervals. It helps identify patterns, trends, and seasonal variations in data over time. Time series analysis is commonly used for forecasting future values based on historical data, such as stock prices, weather patterns, sales data, or economic indicators. Key components of time series analysis include:
    - Trend - Long-term movement in the data.
    - Seasonality - Regular and predictable fluctuations.
    - Cyclic Patterns - Irregular, long-term fluctuations.
    - Noise - Random variations in the data.
  + This analysis is vital for predicting future trends and making data-driven decisions.
* **Differences Between Overfitting and Underfitting**
  + Overfitting happens when a model learns the training data too closely, capturing noise and small fluctuations, which makes it perform well on training data but poorly on new, unseen data. Overfitted models are too complex, often with excessive parameters, causing them to lose generalizability.
  + Underfitting occurs when a model is too simple to capture the underlying patterns in the data, resulting in poor performance on both training and test data. Underfitted models have high bias and fail to capture relationships, leading to inaccurate predictions.
* **How to Treat Outliers in a Dataset**
  + Removal: If outliers are due to data entry errors or are irrelevant, they can be removed. However, this should be done cautiously to avoid losing meaningful data.
  + Transformation: Applying transformations like logarithmic or square root functions can reduce the impact of outliers, especially when data is skewed.
  + Capping/Flooring: Setting a maximum (cap) or minimum (floor) threshold for outliers to bring extreme values closer to the distribution’s range.
  + Imputation: Replacing outliers with median or mean values, especially useful if the outliers are within a reasonable range but still skew analysis.
  + Binning: Grouping values into bins can help in managing outliers without removing them, preserving the underlying data distribution.
  + Using Robust Models: Some algorithms are less sensitive to outliers, such as decision trees or models that use median-based statistics instead of mean-based.
* **Types of Hypothesis Testing**
  + Null and alternative hypothesis
* **Explanation of Type I and Type II Errors in Statistics**
  + In Hypothesis testing, a Type I error occurs when the null hypothesis is rejected even if it is true. It is also known as a false positive.
  + A Type II error occurs when the null hypothesis is not rejected, even if it is false. It is also known as a false negative.
* **Numeric Value Treated as Text in Microsoft Excel**
  + Using Apostrophe (`)

**SQL Questions:**

* What is SQL?
* What is a database?
* Differentiation between Power BI and Tableau
* Does SQL support programming language features?
* What is the difference between CHAR and VARCHAR2 datatype in SQL?
* What do you mean by data definition language?
* What do you mean by data manipulation language?
  + INSERT; UPDATE; DELETE
* What is the view in SQL?
* What do you mean by foreign key?
* What are table and Field?
* What is the primary key?
* What is a Default constraint?
* What is normalization?
  + Normalization in SQL is the process of organizing data within a database to reduce redundancy and improve data integrity. It involves dividing large tables into smaller, related tables and defining relationships between them to ensure efficient data storage and retrieval. This process helps minimize data anomalies and promotes consistency by following a set of rules called "normal forms."
  + The common normal forms include:
    - Ensuring each column contains atomic (indivisible) values, with each row unique.
    - Ensure that all non-key columns are fully dependent on the primary key.
  + Normalization aims to:
    - Eliminate redundant data.
    - Enhance data consistency and integrity.
    - Optimize the efficiency of database storage.
* What is Denormalization?
  + Denormalization is the process of intentionally adding redundancy to a database to improve read performance by reducing the need for complex joints and queries across multiple tables. It essentially reverses some aspects of normalization by combining tables, which can speed up data retrieval in applications with heavy reading requirements, such as reporting systems or analytics databases.
  + Key Points of Denormalization:
    - Improves Query Performance: It reduces the number of joins needed to retrieve data, which can be beneficial in read-heavy environments.
    - Increases Redundancy: Denormalization often duplicates data across tables, which can lead to storage inefficiencies and increase the risk of data anomalies.
    - Use Cases: Denormalization is typically used in OLAP (Online Analytical Processing) systems and for building data warehouses, where faster data retrieval outweighs the need for strict data integrity.
    - Example: Storing customer and order information in a single table instead of separate tables to avoid joining during retrieval.
  + Trade-offs
    - While denormalization can enhance performance, it also increases the need for data maintenance and can lead to data inconsistency if updates are not managed carefully.
* What is a query?
* What is a subquery?
* What are the different operators available in SQL?
  + Arithmetic Operators
  + Logical Operators
  + Comparison Operators
* What is a Constraint?
* What is MySQL collation?
  + MySQL collation is a set of rules that determines how string comparisons are made within the database. It defines the character set and the sorting rules (including case sensitivity) for storing and retrieving text. Collations are used to manage the sorting and comparison behavior of textual data based on specific languages, alphabets, and cultural norms.
  + Key Concepts of MySQL Collation:
    - Character Set: A character set defines the available characters in a collation. For example, UTF-8 or Latin1.
    - Sorting and Comparison Rules: Collation specifies how characters are ordered and compared. For instance, it can dictate whether comparisons are case-sensitive or case-insensitive.
    - Collation Names: A collation name usually follows the format: charset\_collation\_type. For example, utf8\_general\_ci (UTF-8 character set with general collation and case insensitivity).
    - Case Sensitivity:
    - Case-Insensitive (ci): ci in a collation name indicates case-insensitive comparisons (e.g., utf8\_general\_ci).
    - Case-Sensitive (cs): cs in a collation name indicates case-sensitive comparisons (e.g., utf8\_bin).
    - Default Collation: Each MySQL character set has a default collation that can be overridden at the table or column level.
  + Example:
    - If a table is set to use utf8\_general\_ci, a WHERE clause comparing "ABC" and "abc" would consider them equal. If set to utf8\_bin, the same comparison would be case-sensitive, and "ABC" and "abc" would be treated as distinct.
  + Changing Collation:
    - Collation can be changed for a database, table, or column, which affects how MySQL treats and sorts text data.
* What is user-defined functions (UDF)?
* What are all types of user-defined functions?
  + User-defined functions (UDFs) in SQL are custom functions created by users to perform specific operations that may not be covered by SQL’s built-in functions. They allow for modular, reusable code blocks that can be called within SQL queries to perform calculations, transformations, or other operations. UDFs can return a single value or a table, depending on the function type.
  + Types of User-Defined Functions
    - Scalar Functions: Return a single value, such as a numeric, text, or date value. These are used in calculations and transformations.
      * Example: A function to calculate the age based on a birth date.
    - Table-Valued Functions: Return a table result set, which can be queried like a regular table. Useful for generating a subset of data.
      * Example: A function that returns a list of employees within a specific department.
    - Inline Table-Valued Functions: Like table-valued functions, but they contain a single SELECT statement and don’t include BEGIN/END blocks. These are efficient and often faster than multi-statement functions.
      * Example: A function that retrieves customer data based on specific criteria.
  + Benefits of User-Defined Functions
    - Modularity: Encapsulate logic in one place, promoting code reusability.
    - Simplified Queries: Abstract complex operations, making queries easier to read and maintain.
    - Error Reduction: Reduces code duplication, which can help prevent errors in logic.
  + Example of a Scalar UDF
    - Here's a simple user-defined function to calculate the square of a number:
    - CREATE FUNCTION dbo.Square (@Number INT)
    - RETURNS INT
    - AS
    - BEGIN
    - RETURN @Number \* @Number;
    - END;
* What is a stored procedure?
  + A stored procedure in SQL is a precompiled set of one or more SQL statements that can be stored in the database and executed as a single unit. Stored procedures are used to perform complex operations like data manipulation, business logic, or administrative tasks on the database in a structured and reusable way. They help improve performance, reduce redundancy, and enhance security.
  + Key Benefits of Stored Procedures
    - Performance Improvement: Since stored procedures are precompiled, they execute faster than separate SQL queries.
    - Reusability: Once created, stored procedures can be called multiple times, which reduces code duplication.
    - Security: Permissions can be set to allow specific users to execute the procedure without giving them direct access to underlying tables.
    - Reduced Network Traffic: Sending a stored procedure call from an application is generally more efficient than sending multiple SQL queries.
  + Syntax for Creating a Stored Procedure
    - Here’s a basic structure for creating a stored procedure:
      * CREATE PROCEDURE GetEmployeesByDepartment
      * @DepartmentID INT
      * AS
      * BEGIN
      * SELECT EmployeeID, EmployeeName, JobTitle
      * FROM Employees
      * WHERE DepartmentID = @DepartmentID;
      * END;
      * EXEC GetEmployeesByDepartment @DepartmentID = 1;
* What are aggregate and scalar functions?
  + Aggregate Functions
    - Aggregate functions operate on a set of values and return a single summary value. They are commonly used with the GROUP BY clause to perform calculations on groups of data.
    - SUM(): COUNT(): AVG(): MAX(): MIN():
  + Scalar Functions
    - Scalar functions perform operations on individual values and return a single value for each input. They are often used to modify data or perform calculations at the row level.
    - UCASE() / UPPER(): LCASE() / LOWER(): LEN(): ROUND(): NOW(): CONCAT():
  + Key Differences Between Aggregate and Scalar Functions
    - Aggregate functions work on a group of values and return a single result, while scalar functions operate on a single value and return a single value.
    - Aggregate functions are often used with GROUP BY, whereas scalar functions are usually applied directly in SELECT queries or in conditions.
* What is an ALIAS command?
* What are Union, minus, and Interact commands?
* What is a T-SQL?
  + T-SQL, or Transact-SQL, is an extension of SQL (Structured Query Language) used primarily with Microsoft SQL Server. It adds programming constructs like variables, control-of-flow statements, and error handling to standard SQL, enabling more complex database operations and management tasks. Here are some key features of T-SQL:
    - **Procedural Programming Capabilities**: T-SQL includes functions such as IF...ELSE, WHILE, and error handling through TRY...CATCH, enabling procedural logic.
    - **Variables and Functions**: It allows the use of variables and offers built-in functions for string manipulation, date handling, and mathematical calculations.
    - **Transactions**: T-SQL supports explicit transactions, allowing users to commit or roll back changes as a unit of work.
    - **Advanced Querying**: With T-SQL, you can perform complex queries using window functions (ROW\_NUMBER(), RANK()), common table expressions (CTEs), and recursive queries.
    - **Triggers and Stored Procedures**: It supports triggers for automated responses to data changes and stored procedures for reusable code, optimizing performance and ensuring consistency.
* What is ETL in SQL?
  + **ETL is a process in Data Warehousing, and it stands for Extract, Transform, and Load. It is a process in which an ETL tool extracts the data from various data source systems, transforms it in the staging area, and then finally, loads it into the Data Warehouse system. These are three database functions that are incorporated into one tool to pull data out from one database and put data into another database.**
* How to copy tables in SQL?
  + CREATE TABLE new\_table AS
  + SELECT \* FROM original\_table;
* What is SQL injection?
  + **SQL Injection is a code injection technique where an attacker exploits vulnerabilities in an application’s SQL query handling to manipulate or access the database in unintended ways. By inserting or "injecting" malicious SQL code into a query, attackers can bypass authentication, retrieve sensitive data, modify or delete records, and even gain administrative control over the database.**
  + **How SQL Injection Works**
    - **SQL Injection typically occurs when user inputs (like login credentials or search terms) are not properly validated or sanitized. If a SQL query directly incorporates user inputs without validation, attackers can alter the query’s structure and influence database operations.**
  + **Example of SQL Injection**
    - SELECT \* FROM Users WHERE username = 'input\_username' AND password = 'input\_password';
    - SELECT \* FROM Users WHERE username = 'admin' -- ' AND password = 'anything';
    - **The -- starts a comment, which ignores the password condition, allowing access without a valid password.**
  + **Types of SQL Injection**
    - **Classic SQL Injection: Manipulating queries to gain unauthorized access.**
    - **Blind SQL Injection: Exploiting errors to retrieve data without direct feedback.**
    - **Union-Based Injection: Using UNION to retrieve data from other tables.**
    - **Error-Based Injection: Leveraging database error messages to gain information.**
  + **Preventing SQL Injection**
    - **Parameterized Queries/Prepared Statements: Ensure that SQL queries treat user inputs as data rather than executable code.**
    - **Input Validation and Sanitization: Verify and sanitize all user inputs to prevent malicious characters.**
    - **Use of ORM (Object-Relational Mapping): Helps generate safe SQL code automatically.**
    - **Limit Database Permissions: Restrict user privileges to minimize potential damage.**
    - **Regular Security Audits: Review code and conduct penetration testing to catch vulnerabilities.**
  + **SQL Injection can cause serious data breaches, so securing input handling is essential in web application development.**
* Can we disable a trigger? If yes, how?
* What are the differences between SQL and PL/SQL?

|  |  |
| --- | --- |
| SQL | PL/SQL (Procedural Language extensions to the Structured Query Language) |
| **We can embed SQL in PL/SQL** | **We cannot embed PL/SQL in SQL** |
| **SQL tells databases, what to do?** | **PL/SQL tells databases how to do.** |
| **We can execute one statement at a time in SQL** | **We can execute blocks of statements in PL/SQL** |
| **It is used for manipulating data.** | **It is used for creating applications.** |
| **SQL is very declarative in nature.** | **PL/SQL has a procedural nature.** |
| **SQL is a data-oriented language.** | **PL/SQL is a procedural language** |
| **SQL is a query execution or commanding language** | **PL/SQL is a complete programming language** |

* What is the difference between BETWEEN and IN operators in SQL?
* Write an SQL query to find the names of employees starting with 'A'
* Difference between Primary Key and Unique Constraints
* What is a Join in SQL? What are the Types of Joins?
* What is an Index?
* What is the On Delete Cascade Constraint?
* Explain WITH Clause in SQL.
* Different Attributes of Indexes
* What is a Cursor?
  + **A cursor in SQL is a database object that allows for the retrieval and manipulation of a result set row by row. It provides a way to iterate through a set of records returned by a query, enabling developers to process individual rows in a controlled manner.**
* Various Types of Relationships in SQL
  + **One-to-One Relationship.**
  + **One to Many Relationships.**
  + **Many to One Relationship.**
  + **Self-Referencing Relationship.**
* What is a Trigger?
* Difference between SQL DELETE and SQL TRUNCATE Commands
* Difference between Cluster and Non-Cluster Index
* What is a Live Lock?
* What is Case WHEN in SQL?
* Name different types of case manipulation functions available in SQL.
* What are local and global variables and their differences?
* Name the function which is used to remove spaces at the end of a string?
* What is the difference between TRUNCATE and DROP statements?
* Which operator is used in queries for matching patterns?
* Define SQL Order by the statement?
* Explain SQL Having statement?
* Explain SQL AND OR statement with an example?
* Define BETWEEN statements in SQL?
* Why do we use Commit and Rollback commands?
* What are ACID properties?
  + **ACID properties are a set of principles that guarantee reliable processing of database transactions. ACID stands for Atomicity, Consistency, Isolation, and Durability. These properties ensure that database transactions are processed reliably and help maintain the integrity of the database.**
    - **Atomicity**: Ensures that all parts of a transaction succeed or none do.
    - **Consistency**: Guarantees that a transaction takes the database from one valid state to another.
    - **Isolation**: Keeps transactions independent from one another, avoiding interference.
    - **Durability**: Guarantees that committed transactions are permanently recorded.
  + **Importance of ACID Properties**
    - **Maintaining data integrity in applications.**
    - **Ensuring reliability in multi-user environments.**
    - **Supporting complex transactions where consistency and correctness are paramount.**
* Are NULL values the same as zero or a blank space?
* What is the need for group functions in SQL?
* What is the need for a MERGE statement?
  + **The MERGE statement in SQL, also known as an "upsert," is used to perform conditional INSERT, UPDATE, or DELETE operations on a target table based on the results of a join with a source table. It allows for efficient handling of situations where you need to synchronize data between two tables, typically for scenarios involving data integration, ETL processes, or maintaining a data warehouse.**
  + **Key Benefits of the MERGE Statement:**
    - **Combines Multiple Operations: MERGE allows you to execute INSERT, UPDATE, and DELETE operations in a single statement, reducing the need for multiple SQL queries.**
    - **Improves Performance: By consolidating operations, MERGE can enhance performance, especially when dealing with large datasets, since it reduces the overhead associated with executing multiple queries.**
    - **Simplicity and Readability: The MERGE statement simplifies the logic required to perform complex operations, making the SQL code easier to read and maintain.**
    - **Data Synchronization: It's useful for synchronizing data between tables, ensuring that the target table reflects the latest state of the source table.**
    - **Conditional Logic: MERGE allows for complex conditional logic, enabling you to specify different actions based on the existence or absence of matching rows.**
      * MERGE INTO Customers AS c
      * USING NewCustomers AS nc
      * ON c.CustomerID = nc.CustomerID
      * WHEN MATCHED THEN
      * UPDATE SET c.CustomerName = nc.CustomerName
      * WHEN NOT MATCHED THEN
      * INSERT (CustomerID, CustomerName) VALUES (nc.CustomerID, nc.CustomerName)
      * WHEN NOT MATCHED BY SOURCE THEN
      * DELETE;
* How can you fetch common records from two tables?
* What are the advantages of PL/SQL functions?
* What is the SQL query to display the current date?
* What are Nested Triggers?
* How to find the available constraint information in the table?
* How do we avoid getting duplicate entries in a query without using the distinct keyword?
* The difference between NVL and NVL2 functions?
* What is the difference between COALESCE() & ISNULL()?
* Name the operator which is used in the query for appending two strings?
* Common SQL commands:
  + **DDL: CREATE, ALTER TABLE, DROP, TRUNCATE, and ADD COLUMN**
  + **DML: UPDATE, DELETE, and INSERT**
  + **DCL: GRANT and REVOKE**
  + **TCL: COMMIT, SET TRANSACTION, ROLLBACK, and SAVEPOINT**
  + **DQL: SELECT**

Link:

* [Top 70+ SQL Interview Questions and Answers for 2024](https://www.geeksforgeeks.org/sql-interview-questions/)

**Power BI Questions:**

* What is Power Bi?
* Features of Power BI.
* Differentiation between Power BI and Tableau
* Explain the difference between Power BI Desktop and Power BI Service.
* How do you import data into Power BI?
* What is the purpose of the Query Editor in Power BI?
* What is the role of visuals in a Power BI report?
* What is a filter in Power BI, and why is it useful?
* Major components of Power BI.
* Common chart in Power BI.
* DAX and benefits from using variables in DAX.
* Three fundamental concepts of DAX.
  + Syntax
  + Function
  + Context
* Most common DAX functions.
  + Aggregation Functions: SUM, MIN, MAX, AVG, COUNTROWS, DISTINCTCOUNT
  + Information Functions: ISBLANK, ISFILTERED, ISCROSSFILTERED
  + Statistical Functions: GEOMEAN, MEDIAN
  + Logical Functions: IF, AND, OR, SWITCH
  + Date & Time Functions: DATEDIFF, DATEVALUE
  + Filter Functions: VALUES, ALL, FILTER, CALCULATE, TOPN
  + Other Functions: UNION, INTERSECT, EXCEPT, NATURALINNERJOIN, NATURALLEFTEROUTERJOIN,
  + SUMMARIZECOLUMNS, ISEMPTY, VAR
* Calculate function in DAX
* Custom visualization.
* Different connectivity modes in power bi.
* Data source that we can connect to power bi.
* Views in power bi.
* Type of filters.
* Can you explain the basic steps for data cleaning and transformation in Power Query?
  + Data load > removed unwanted columns and rows > filter rows > rename the column and tables > correct the data type > handle missing data > standardize data format > merge and append the data > apply transformation (create calculated columns, pivot/unpivot data, split columns or performing some aggregation) > remove duplicates > group the data > sort it > and final review and then load it.
* What is a slicer in Power BI, and how is it used for data filtering?
* What are the benefits of using Power BI for data visualization and analysis?
  + Easy Integration with Data Sources
  + Interactive Visualizations
  + User-Friendly Interface
  + Powerful Data Modeling Capabilities
  + Advanced Analytics with DAX
  + Real-Time Dashboards and Monitoring
  + Collaboration and Sharing
  + Mobile Accessibility
  + Security and Compliance
  + Scalability for Large Data Volumes
  + AI and Machine Learning Integration
  + Cost-Effectiveness
* Explain various types of users who can use Power BI.
  + Business Analysts
  + Data Analysts and Data Scientists
  + Report Consumers (End Users)
  + Managers and Executives
  + Data Engineers
  + Developers and IT Professionals
  + Sales and Marketing Teams
  + Finance and Operations Teams
  + Customer Support Teams
  + Power BI Administrators
* Star Schema and its benefits.

Links:  
- [Top 30 Power BI Interview Questions and Answers](https://www.geeksforgeeks.org/power-bi-interview-questions-and-answers/)

* [25 Must-Know Power BI Interview Questions and Answers (2023)](https://www.dataquest.io/blog/power-bi-interview-questions-and-answers/)

**Tableau Questions:**

* What is Tableau?
* What are the different data connection options available in Tableau?
* How can you create a calculated field in Tableau?
* What is the difference between a dimension and a measure in Tableau?
* What is data blending in Tableau?
* What is the purpose of a parameter in Tableau?
* How can you perform data aggregation in Tableau?
* What Are the Data Types Supported in Tableau?
* What is Meant by ‘discrete’ and ‘continuous’ in Tableau?
* \*\* What Are the Filters? Name the Different Filters in Tableau.
  + Extract filters
  + Context filters
  + Data source filters
  + Filters on measures
  + Filters on dimensions
  + Table calculation filter
* What Are the Different Joins in Tableau?
* What is the Difference Between Joining and Blending?
* What is the Difference Between a Live Connection and an Extract?
* Is There a Difference Between Sets and Groups in Tableau?
* What is the Difference Between Treemaps and Heat Maps?
* What is the Difference Between .twbx And .twb?
* What Do You Understand by the Blended Axis?
* What is the Use of Dual-axis? How Do You Create One?
* What is the Level of Detail (LOD) Expression?
* How Can You Optimize the Performance of a Dashboard?
  + Maximize the number of fields and records. You can exclude unused fields from your visualization or use extract filters.
  + Limit the number of filters used, by avoiding quick filters and using action and parameter filters instead. These filters reduce query loads.
  + use Min/Max instead of Average because average functions require more processing time than Min/Max
  + Use boolean or numerical calculations more than string calculations. Computers can process integers and boolean much faster than strings.
* In Tableau, what are aggregation and disaggregation?
* What is KPI in Tableau?
* What are Bins in Tableau?
* What is a Quick Filter in Tableau?
* How do you calculate percentage in Tableau?
* What are the limitations of Tableau?
* What are cascading filters in Tableau?

**Python Questions:**