Database Analyst Interview Questions And Answers Guide.

Database Analyst Job Interview Preparation Guide.

Question # 1

Tell me what is the IN operator?

Answer:-

IN is a conditional operator used in a WHERE clause and is shorthand for multiple OR conditional statements. It tests the expression that precedes it against a list of values that are passed in to the operator, which can either be comma-separated values or a subquery that returns a list of values. If the expression that precedes IN matches any of the elements in the list, the resulting value is TRUE, or 1; otherwise, the value is FALSE, or 0.

Read More Answers.

Question # 2

Please explain what is the difference between UNION and UNION ALL?

Answer:-

UNION will omit duplicate records, whereas UNION ALL will include duplicate records. UNION requires the server to do the additional work of removing any duplicates.

Read More Answers.

Question # 3

Explain me about a time when you could not meet a deadline?

Answer:-

This question gets into how well candidates handle stressful situations. You're looking for a data analyst who can anticipate when a deadline is not going to work and who can find a solution. Past behavior is a good predictor of future behavior.
What to look for in an answer:
* Ability to see big picture
* Decisiveness and being proactive
* Answers that do not blame others

Read More Answers.

Question # 4

Tell us why did you go into data analysis?

Answer:-

This query is a good way to get to know candidates as people. It can serve as an icebreaker at the beginning of an interview or, if it comes at the end, as a gentle way to bring your question portion to a close.
What to look for in an answer:
* Focused replies
* Personality
* Specifics

Read More Answers.

Question # 5

Tell us when might someone denormalize their data?

Answer:-

Typically done for performance reasons, to reduce the number of table joins. This is not a good idea in a transactional environment as there are inherent data integrity risks or performance risks due to excessive locking to maintain data integrity.
Questions related to the Unified Modeling Language (UML) or Entity-Relationship Diagrams (ERDs) may also be asked here.

Read More Answers.

Question # 6

Can you explain validation?

Answer:-

In this step, the model provided by the client and the model developed by the data analyst are validated against each other to find out if the developed model will meet the business requirements.

Question # 7

Can you explain me what is SQL?

Answer:-

SQL is short for Structured Query Language and is used to communicate with relational databases. It is the standard language used to retrieve, update, insert, and delete data when working with relational databases.

Question # 8

Tell me what is a primary key?

Answer:-

A primary key is a unique identifier for a particular record in a table. The primary key can't be NULL. A primary key can be a single column or a combination of columns in a table. Each table can contain only one primary key.

Question # 9

Tell me what kind of data analysis software experience do you possess?

Answer:-

I have advanced data analysis software experience. A few examples include creating PivotTables in Excel, producing databases from scratch in Access, and developing data mining algorithms in ELKI. Also in my previous role, I was tasked with upgrading the database to meet the demands of the market and the company to ensure it ran smoothly.

Question # 10

What is implementation of the Model and Tracking?

Answer:-

This is the final step of the data analysis process wherein the model is implemented in production and is tested for accuracy and efficiency.

Question # 11

Explain me some common problems that data analysts encounter during analysis?

Answer:-

Having a poor formatted data file. For instance, having CSV data with un-escaped newlines and commas in columns.
Having inconsistent and incomplete data can be frustrating.
Common Misspelling and Duplicate entries are a common data quality problem that most of the data analysts face.
Having different value representations and misclassified data.

Question # 12

Tell us what are the usual challenges a data analyst normally encounter?

Answer:-

Amongst the interview questions for data analyst, challenges faced is a sure-shot question put up by the interviewer. Here are a few challenges:
* Illegal values
* Duplicate entries
* Trying to identify data that is overlapping
* Regular misspelling
* Irregular value misrepresentation
Data analytics interview questions can come in various manners. There are data analytics questions for freshers and data analytics interview questions for experienced. Whichever ones apply to your present situation, make sure you are fully prepared.

Question # 13

Tell us how will you create a classification to identify key customer trends in unstructured data?

Answer:-

A model does not hold any value if it cannot produce actionable results, an experienced data analyst will have a varying strategy based on the type of data being analysed. For example, if a customer complain was retweeted then should that data be included or not. Also, any sensitive data of the customer needs to be protected, so it is also advisable to consult with the stakeholder to ensure that you are following all the compliance regulations of the organization and disclosure laws, if any.
You can answer this question by stating that you would first consult with the stakeholder of the business to understand the objective of classifying this data. Then, you would use an iterative process by pulling new data samples and modifying the model accordingly and evaluating it for accuracy. You can mention that you would follow a basic process of mapping the data, creating an algorithm, mining the data, visualizing it and so on. However, you would accomplish this in multiple segments by considering the feedback from stakeholders to ensure that you develop an enriching model that can produce actionable results.

Question # 14

Tell me which SQL command is used to add a row?

Answer:-

INSERT

Question # 15

Can you take a few minutes to explain how you would estimate how many shoes could potentially be sold in New York City each June?

Answer:-

Many interviewers pose questions that let them see an analyst's thought process without the aid of computers and data sets. After all, technology is only as good and reliable as the people behind it.
What to look for in an answer:
* Ability to identify variables/data segments
* Ability to communicate thought process
* Creativity

Question # 16

Please explain what are aggregate functions?

Answer:-

Aggregate functions perform calculations on a set of values and return a single value. The common aggregate functions are:
* COUNT (counts the number of rows in the table)
* SUM (returns the sum of all values of a numeric column)
* AVG (returns the average of all values of a numeric column)
* MIN (returns the lowest value of a numeric column)
* MAX (returns the highest value of a numeric column).
Aggregate functions are frequently used in combination with the GROUP BY statement.

Question # 17

Tell us what are the different types of subqueries?

Answer:-

There are two types of subqueries: correlated and uncorrelated.
An uncorrelated subquery is a independent query whose output is substituted into the main query. A correlated subquery, on the other hand, uses values from the outer query and therefore depends on the outer query. Such a subquery executes repeatedly, once for each row that is selected by the outer query.

Question # 18

As you said you have just been assigned a new analytics project. Where do you begin and what are the steps that follow?

Answer:-

The very first thing I would do is clearly define the problem or objective so I have a solid direction. Second, I would explore the data and become more familiar with it. This is extremely critical especially if I am working with a new set of data. Next, I would prepare the data for modeling. This entails data validation, detecting outliers, treating missing values, etc. With those steps completed, I would begin modeling the data until I discover the most significant or valuable results. Lastly, I would implement the model and track my results. As I'm sure you are aware, this process could vary slightly based upon the type of problem and the data and tools available.

Question # 19

Explain me what is data cleansing? Mention few best practices that you have followed while data cleansing?

Answer:-

From a given dataset for analysis, it is extremely important to sort the information required for data analysis. Data cleaning is a crucial step in the analysis process wherein data is inspected to find any anomalies, remove repetitive data, eliminate any incorrect information, etc. Data cleansing does not involve deleting any existing information from the database, it just enhances the quality of data so that it can be used for analysis.
Some of the best practices for data cleansing include -
* Developing a data quality plan to identify where maximum data quality errors occur so that you can assess the root cause and design the plan according to that.
* Follow a standard process of verifying the important data before it is entered into the database.
* Identify any duplicates and validate the accuracy of the data as this will save lot of time during analysis.
* Tracking all the cleaning operations performed on the data is very important so that you repeat or remove any operations as necessary.

Question # 20

Explain few of the best tools useful for data analytics?

Answer:-

Some of the best tools useful for data analytics are: KNIME, Tableau, OpenRefine, io, NodeXL, Solver, etc.

Question # 21

Tell us what do you use to get non-repeated values?

Answer:-

The DISTINCT keyword is used in the SELECT statement to eliminate repetition of identical data. It is also used in aggregate functions. When DISTINCT is used with only one column or expression, the query will strictly return the unique values for that particular column or expression. Similarly, when DISTINCT is used with multiple columns or expressions, the query will return only the unique combinations of those columns or expressions. Note that the DISTINCT keyword doesn't ignore the NULL value when sifting through data.

Question # 22

Tell me which data analysis software are you well-versed in?

Answer:-

This question lets you assess if candidates have the hard skills you need and can tell you what areas they might need training in. It is also another way to ensure basic competency.
What to look for in an answer:
* Software the job ad emphasized
* Experience with the software
* Ability to speak with familiarity

Question # 23

Tell me the typical data analysis process?

Answer:-

Data analysis deals with collecting, inspecting, cleansing, transforming and modelling data to glean valuable insights and support better decision making in an organization. The various steps involved in the data analysis process include -

Question # 24

Tell us how will you handle the QA process when developing a predictive model to forecast customer churn?

Answer:-

Data analysts require inputs from the business owners and a collaborative environment to operationalize analytics. To create and deploy predictive models in production there should be an effective, efficient and repeatable process. Without taking feedback from the business owner, the model will just be a one-and-done model.
The best way to answer this question would be to say that you would first partition the data into 3 different sets Training, Testing and Validation. You would then show the results of the validation set to the business owner by eliminating biases from the first 2 sets. The input from the business owner or the client will give you an idea on whether you model predicts customer churn with accuracy and provides desired results.

Question # 25

Can you write the command to remove all employees named John from the EMPLOYEE table?

Answer:-

DELETE from EMPLOYEE WHERE firstName = 'ABCDEF'

Question # 26

Tell me what does SQL stand for?

Answer:-

Structured Query Language

Question # 27

Tell me what is CROSS JOIN?

Answer:-

CROSS JOIN defines a Cartesian product on the sets of records from two or more joined tables, where the number of rows in the first table is multiplied by the number of rows in the second table.

Question # 28

Tell us what are the requirements needed for becoming a data analyst?

Answer:-

This is one of the most commonly asked data analyst interview questions. Listed below are the requirements needed for becoming a data analyst:
* Sound knowledge of statistical packages used in analyzing big datasets like Excel, SAS, SPSS and many others

* Very good knowledge of the programming language (Javascript, ETL frameworks or XML), reporting packages (Business Objects) and databases.
* Strong technical knowledge in areas like data models, segmentation techniques, data mining and database design
* Good skills in knowing how to run analysis, organization, collection and dissemination of big data accurately

Read More Answers.

Question # 29

Suppose you are assigned a new data anlytics project. How will you begin with and what are the steps you will follow?

Answer:-

The purpose of asking this question is that the interviewer wants to understand how you approach a given data problem and what is the though process you follow to ensure that you are organized. You can start answering this question by saying that you will start with finding the objective of the given problem and defining it so that there is solid direction on what need to be done. The next step would be to do data exploration and familiarise myself with the entire dataset which is very important when working with a new dataset.The next step would be to prepare the data for modelling which would including finding outliers, handling missing values and validating the data. Having validated the data, I will start data modelling untill I discover any meaningfuk insights. After this the final step would be to implement the model and track the output results.

This is the generic data analysis process that we have explained in this answer, however, the answer to your question might slightly change based on the kind of data problem and the tools available at hand.

Read More Answers.

Question # 30

Can you tell us what has been your most difficult analysis to date?

Answer:-

My biggest challenge was making prediction sales during the recession period and estimating financial losses for the upcoming quarter. Interpreting the information was a seamless process. However, it was slightly difficult to forecast future trends when the market fluctuates frequently. Usually I analyze and report on data that has already occurred. In this case, I had to research how receding economic conditions impacted varying income groups and then make an inference on the purchasing capacity of each group.

Read More Answers.

Question # 31

Tell me what do data analysts do?

Answer:-

This question is basic but serves an essential function. It weeds out the candidates who lack a rudimentary understanding of data analysis. It also lets you compare how well various candidates understand data analysis.
What to look for in an answer:
* Coverage of each step
* Mention of soft skills, such as communication
* Discussion of how data analysts benefit a company

Read More Answers.

Question # 32

Tell me in what situations should you use WHERE and HAVING in a statement?

Answer:-

Though both WHERE and HAVING are used to filter records, there is a subtle difference between the two. The WHERE clause is used to filter records from a result, whereas HAVING is used to filter groups. If the two are used together, the WHERE clause is applied first to filter rows; only after grouping finishes is the HAVING clause applied.

Read More Answers.

Question # 33

Tell me the difference between data profiling and data mining?

Answer:-

The difference between data profiling and data mining is:
Data Profiling is aimed at individual attributes' analysis. Information on different attributes like discrete values, value ranges and their data type, frequency, length are gotten from it. Data mining, on the other hand, targets unusual records detection, cluster analysis, sequence discovery and others.

Read More Answers.

Question # 34

Tell me what steps are in an analytics project?

Answer:-

The steps involved in an analysis project can be listed as:
* Problem identification
* Exploration of data
* Preparation of data
* Modeling
* Data Validation
* Implementation and tracking

Read More Answers.

Question # 35

Tell me what are your communication strengths?

Answer:-

My greatest communication strength would have to be my ability to relay information. I'm good at speaking in a simple, yet effective manner so that even people who aren't familiar with the terms can grasp the overall concepts. I think communication is extremely valuable in a role like this, specifically when presenting my findings. This is even more important when those findings could be beneficial or detrimental to other departments within the company and you need to make sure that everyone understands the overall message.

Read More Answers.

Question # 36

Tell us what is your process when you start a new project?

Answer:-

This question lets you measure candidates' organizational skills and how well they anticipate. It also gives you an opportunity to see if candidates' leadership or work styles are compatible with your company culture.
What to look for in an answer:
* Clear steps
* Deliberate process
* Consideration of deadline

Read More Answers.

Question # 37

Explain what is a view?

Answer:-

A view is a virtual table that consists of a subset of data from a table. The content of a view is defined by the query. A view takes up little space because it doesn't copy all data from the given table but only a subset of data as defined by the view. Note that a view can also display a combination of data from one or more tables. Views allow you to hide the complexity of large data and instead narrow in on areas of interest.

Read More Answers.

Question # 38

Tell us what is the GROUP BY statement used for?

Answer:-

GROUP BY is a statement that divides the result for a particular query into groups of rows. It is often used with aggregate functions such as SUM, AVG, MAX, MIN, and COUNT, which calculate information about each group. The SELECT statement returns one row for each group.

Read More Answers.

Question # 39

Tell us what are the best practices for data cleaning?

Answer:-

* Separate data depending on their attributes
* In the case of massive datasets, do a stepwise cleansing and improve on the data on every step until the data quality is good.
* For common data cleansing, you need to generate a set of scripts which include blanking out every value not matching a regex.
* Do analysis on the statistic for every column.
* Stay up to date with all cleaning operations, so changes could make when necessary.

Read More Answers.

Question # 40

Explain me what are the responsibilities of a Data Analyst?

Answer:-

To answer this question, you need to know that such responsibilities include:
* Interpret data and analyze results by using techniques of statistics and give reports.
* Look out for new areas or processes to improve opportunities.
* Get data from various sources (primary and secondary) and keep the systems running.
* Filter data from various sources and go through computer reports.
* Make sure all data analysis gets support and makes sure customers and staff relate well

Read More Answers.

Question # 41

Do you know what Is INNER JOIN?

Answer:-

INNER JOIN is the most common type of join. It returns all rows that are shared by two tables. Visually, it's analogous to identifying the overlap, or intersection, between two sets of data.

Read More Answers.

Question # 42

Tell us what was your most difficult data analyst project?

Answer:-

With a question like this, you glean insight into how candidates approach and solve problems. It also gives you a better idea of the type of work they have done.
What to look for in an answer:
* Explanation of how challenge(s) were overcome
* Lack of blaming others
* Discussion of why the project was difficult

Read More Answers.

Question # 43

Tell us what is the difference between Data Mining and Data Profiling?

Answer:-

Data Profiling, also referred to as Data Archeology is the process of assessing the data values in a given dataset for uniqueness, consistency and logic. Data profiling cannot identify any incorrect or inaccurate data but can detect only business rules violations or anomalies. The main purpose of data profiling is to find out if the existing data can be used for various other purposes.
Data Mining refers to the analysis of datasets to find relationships that have not been discovered earlier. It focusses on sequenced discoveries or identifying dependencies, bulk analysis, finding various types of attributes, etc.

Read More Answers.

Question # 44

Tell me what is Data Cleansing?

Answer:-

When answering this question, you should know that the definition of data cleansing is:
Data cleansing (also known as data cleaning) involves a data analyst discovering and eliminating errors and irregularities from the database to enhance data quality.

Read More Answers.

Question # 45

Tell me how often should you retrain a data model?

Answer:-

A good data analyst is the one who understands how changing business dynamics will affect the efficiency of a predictive model. You must be a valuable consultant who can use analytical skills and business acumen to find the root cause of business problems.
The best way to answer this question would be to say that you would work with the client to define a time period in advance. However, I would refresh or retrain a model when the company enters a new market, consummate an acquisition or is facing emerging competition. As a data analyst, I would retrain the model as quick as possible to adjust with the changing behaviour of customers or change in market conditions.

Read More Answers.

Question # 46

Do you know what is a subquery?

Answer:-

A subquery is, as the name suggests, a query within another query. The outer query is called the main query, while the inner query is the subquery. You can think of this in terms of composition of functions. A subquery is always executed first, and its result is then passed on to the main query.

Read More Answers.

Question # 47

Explain me what is self JOIN?

Answer:-

Self JOIN is a query that joins a table with itself. This is used to compare the values of a particular column with other values in the same column of the same table. Self JOIN uses aliases to name the original and duplicate tables.

Read More Answers.

Question # 48

Tell me what are the differences between primary and foreign keys?

Answer:-

The primary key is the column or set of columns used to uniquely identify the items in a table. A foreign key is used to uniquely identify the items in a different table, allowing join operations to happen.

Read More Answers.

Question # 49

Tell me what is the criteria to say whether a developed data model is good or not?

Answer:-

* The developed model should have predictable performance.
* A good data model can adapt easily to any changes in business requirements.
* Any major data changes in a good data model should be scalable.
* A good data model is one that can be easily consumed for actionable results.

Read More Answers.

Question # 50

Explain data Modelling?

Answer:-

The modelling step begins once the data has been prepared. Modelling is an iterative process wherein the model is run repeatedly for improvements. Data modelling ensures that the best possible result is found for a given business problem.

Read More Answers.

Question # 51

Tell us what is involved in typical data analysis?

Answer:-

Typical data analysis involves the collection and organization of data. Then, finding correlations between that analyzed data and the rest of the company's and industry's data. It also entails the ability to spot problems and initiate preventative measures or problem solve creatively.

Read More Answers.

Question # 52

Tell me what is the difference between LEFT JOIN and RIGHT JOIN?

Answer:-

A LEFT JOIN returns all records from the left table, even when they do not match in the right table. Missing values become NULL. In a similar manner, a RIGHT JOIN returns all records from the right table, even when they do match those in the left table. Missing values become NULL.

Read More Answers.

Question # 53

Tell us what are the important steps in data validation process?

Answer:-

Data Validation is performed in 2 different steps-
* Data Screening - In this step various algorithms are used to screen the entire data to find any erroneous or questionable values. Such values need to be examined and should be handled.
* Data Verification- In this step each suspect value is evaluated on case by case basis and a decision is to be made if the values have to be accepted as valid or if the values have to be rejected as invalid or if they have to be replaced with some redundant values.

Read More Answers.

Question # 54

What is Logic Regression?

Answer:-

Logic Regression can be defined as:
This is a statistical method of examining a dataset having one or more variables that are independent defining an outcome.

Read More Answers.

Question # 55

Tell me what is a database transaction?

Answer:-

A transaction is a single logical (atomic) unit of work, in which a sequence of operations (or none) must be executed. A transaction has a defined beginning and end. You can commit or roll back a transaction.

Read More Answers.

Question # 56

Tell us what do you know about interquartile range as data analyst?

Answer:-

A measure of the dispersion of data that is shown in a box plot is referred to as the interquartile range. It is the difference between the upper and the lower quartile.

Read More Answers.

Question # 57

Can you describe the differences in the first through fifth normalization forms?

Answer:-

Database candidates should be familiar with most if not all of these without needing to lookup definitions. Some of the other normalization forms are less commonly known/used, but could theoretically be asked. Knowing the differences between second and third is probably a good idea.

Read More Answers.

Question # 58

Do you know what is cardinality?

Answer:-

Thinking mathematically, it is the number of elements in a set. Thinking in the database world, cardinality has to do with the counts in a relationship, one-to-one, one-to-many, or many-to-many.

Database Analyst Interview Questions And Answers

Question # 59
Explain data Preparation?

Answer:-

This is the most crucial step of the data analysis process wherein any data anomalies (like missing values or detecting outliers) with the data have to be modelled in the right direction.

Question # 60
Tell me what are the elements of an ERD?

Answer:-

The three elements include the entities for which someone is seeking information, the attributes of those entities, and the relationships between the entities.

Question # 61
Explain data Exploration?

Answer:-

Having identified the business problem, a data analyst has to go through the data provided by the client to analyse the root cause of the problem.

Databases Programming Most Popular Interview Topics.

1 : SQL and PL/SQL Frequently Asked Interview Questions and Answers Guide.

2 : MySQL Programming Frequently Asked Interview Questions and Answers Guide.

3 : MS SQL Server Frequently Asked Interview Questions and Answers Guide.

4 : Database Administrator (DBA) Frequently Asked Interview Questions and Answers Guide.

5 : Data Structures Frequently Asked Interview Questions and Answers Guide.

6 : SQL Frequently Asked Interview Questions and Answers Guide.

7 : Data Modeling Frequently Asked Interview Questions and Answers Guide.

8 : RDBMS Frequently Asked Interview Questions and Answers Guide.

9 : Stored Procedure Frequently Asked Interview Questions and Answers Guide.

10 : PostgreSQL Frequently Asked Interview Questions and Answers Guide.

About Global Guideline.

Global Guideline is a platform to develop your own skills with thousands of job interview questions and web tutorials for fresher's and experienced candidates. These interview questions and web tutorials will help you strengthen your technical skills, prepare for the interviews and quickly revise the concepts. Global Guideline invite you to unlock your potentials with thousands of **Interview Questions with Answers** and much more. Learn the most common technologies at Global Guideline. We will help you to explore the resources of the World Wide Web and develop your own skills from the basics to the advanced. Here you will learn anything quite easily and you will really enjoy while learning. Global Guideline will help you to become a professional and Expert, well prepared for the future.

* This PDF was generated from https://GlobalGuideline.com at **November 29th, 2023**

* If any answer or question is incorrect or inappropriate or you have correct answer or you found any problem in this document then don't hesitate feel free and e-mail us we will fix it.

You can follow us on FaceBook for latest Jobs, Updates and other interviews material.
 www.facebook.com/InterviewQuestionsAnswers

Follow us on Twitter for latest Jobs and interview preparation guides
https://twitter.com/InterviewGuide

Best Of Luck.

Global Guideline Team
https://GlobalGuideline.com
Info@globalguideline.com