

# Data Warehouse Developer Interview Questions And Answers Guide.



**Global Guideline.**

<https://globalguideline.com/>



# Data Warehouse Developer Job Interview Preparation Guide.

## Question # 1

Explain me why is chameleon method used in data warehousing?

### Answer:-

Chameleon is a hierarchical clustering algorithm that overcomes the limitations of the existing models and the methods present in the data warehousing. This method operates on the sparse graph having nodes: that represent the data items, and edges: representing the weights of the data items.

This representation allows large dataset to be created and operated successfully. The method finds the clusters that are used in the dataset using two phase algorithm.

\* The first phase consists of the graph partitioning that allows the clustering of the data items into large number of sub-clusters.

\* Second phase uses an agglomerative hierarchical clustering algorithm to search for the clusters that are genuine and can be combined together with the sub-clusters that are produced.

[Read More Answers.](#)

## Question # 2

Tell us what is Hybrid SCD?

### Answer:-

Hybrid SCDs are a combination of both SCD 1 and SCD 2.

It may happen that in a table, some columns are important and we need to track changes for them i.e., capture the historical data for them whereas in some columns even if the data changes, we do not have to bother.

For such tables, we implement Hybrid SCDs, where in some columns are Type 1 and some are Type 2.

[Read More Answers.](#)

## Question # 3

Tell us what is the main difference between Inmon and Kimball philosophies of data warehousing?

### Answer:-

Both differ in the concept of building the data warehouse.

\* Kimball views data warehousing as a constituency of Data marts. Data marts are focused on delivering business objectives for departments in the organization. And the data warehouse is a conformed dimension of the data marts. Hence, a unified view of the enterprise can be obtained from the dimension modeling on a local departmental level.

\* Inmon explains in creating a data warehouse on a subject-by-subject area basis. Hence, the development of the data warehouse can start with data from the online store. Other subject areas can be added to the data warehouse as their needs arise. Point-of-sale (POS) data can be added later if management decides it is necessary.

\* Hence, Kimball-First Data Marts-Combined way -Data warehouse

\* Inmon-First Data warehouse-Later--Data marts

[Read More Answers.](#)

## Question # 4

Can you list the Schema that a data warehouse system can implements?

### Answer:-

A data Warehouse can implement star schema, snowflake schema, and fact constellation schema.

[Read More Answers.](#)

## Question # 5

Tell us what is Virtual Warehouse?

### Answer:-

The view over an operational data warehouse is known as virtual warehouse.

[Read More Answers.](#)

## Question # 6

Tell us what does the Query Manager responsible for?

**Answer:-**

Query Manager is responsible for directing the queries to the suitable tables.

[Read More Answers.](#)

**Question # 7**

Tell us out of star schema and snowflake schema, whose dimension table is normalized?

**Answer:-**

Snowflake schema uses the concept of normalization.

[Read More Answers.](#)

**Question # 8**

Explain me what kind of costs are involved in Data Marting?

**Answer:-**

Data Marting involves hardware & software cost, network access cost, and time cost.

[Read More Answers.](#)

**Question # 9**

Explain me what is the purpose of cluster analysis in Data Warehousing?

**Answer:-**

Cluster analysis is used to define the object without giving the class label. It analyzes all the data that is present in the data warehouse and compare the cluster with the cluster that is already running. It performs the task of assigning some set of objects into the groups also known as clusters. It is used to perform the data mining job using the technique like statistical data analysis. It includes all the information and knowledge around many fields like machine learning, pattern recognition, image analysis and bio-informatics. Cluster analysis performs the iterative process of knowledge discovery and includes trials and failures. It is used with the pre-processing and other parameters as a result to achieve the properties that are desired to be used.

Purpose of cluster analysis :-

- \* Scalability
- \* Ability to deal with different kinds of attributes
- \* Discovery of clusters with attribute shape
- \* High dimensionality
- \* Ability to deal with noisy
- \* Interpretability

[Read More Answers.](#)

**Question # 10**

Tell us what is snapshot with reference to data warehouse?

**Answer:-**

\* Snapshot refers to a complete visualization of data at the time of extraction. It occupies less space and can be used to back up and restore data quickly.

\* A snapshot is a process of knowing about the activities performed. It is stored in a report format from a specific catalog. The report is generated soon after the catalog is disconnected.

[Read More Answers.](#)

**Question # 11**

Explain me what is junk dimension?

**Answer:-**

\* In scenarios where certain data may not be appropriate to store in the schema, this data (or attributes) can be stored in a junk dimension. The nature of data of junk dimension is usually Boolean or flag values.

\* A single dimension is formed by lumping a number of small dimensions. This dimension is called a junk dimension. Junk dimension has unrelated attributes. The process of grouping random flags and text attributes in dimension by transmitting them to a distinguished sub dimension is related to junk dimension.

[Read More Answers.](#)

**Question # 12**

Please explain what is VLDB?

**Answer:-**

A very large database, or VLDB, is a database that contains an extremely large number of tuples (database rows), or occupies an extremely large physical file system storage space. A one terabyte database would normally be considered to be a VLDB.

[Read More Answers.](#)

**Question # 13**

Tell us what do OLAP and OLTP stand for?

**Answer:-**

OLAP is an acronym for Online Analytical Processing and OLTP is an acronym of Online Transactional Processing.

[Read More Answers.](#)

**Question # 14**

Can you define dimension?

**Answer:-**

The dimensions are the entities with respect to which an enterprise keeps the records.

[Read More Answers.](#)

**Question # 15**

Tell us what is Summary Information?

**Answer:-**

Summary Information is the area in data warehouse where the predefined aggregations are kept.

[Read More Answers.](#)

**Question # 16**

Tell me do you know what is Normalization?

**Answer:-**

Normalization splits up the data into additional tables.

[Read More Answers.](#)

**Question # 17**

Explain what are the reasons for partitioning?

**Answer:-**

Partitioning is done for various reasons such as easy management, to assist backup recovery, to enhance performance.

[Read More Answers.](#)

**Question # 18**

What is the types of OLAP server?

**Answer:-**

There are four types of OLAP servers, namely Relational OLAP, Multidimensional OLAP, Hybrid OLAP, and Specialized SQL Servers.

[Read More Answers.](#)

**Question # 19**

Can you tell us what is data mart?

**Answer:-**

Data mart contains the subset of organization-wide data. This subset of data is valuable to specific groups of an organization. In other words, we can say that a data mart contains data specific to a particular group.

[Read More Answers.](#)

**Question # 20**

Tell me the process that are involved in Data Warehousing?

**Answer:-**

Data Warehousing involves data cleaning, data integration and data consolidations.

[Read More Answers.](#)

**Question # 21**

Please explain what is conformed fact?

**Answer:-**

\* Conformed dimensions are the dimensions which can be used across multiple Data Marts in combination with multiple facts tables accordingly.

\* A conformed dimension is a dimension that has exactly the same meaning and content when being referred from different fact tables. A conformed dimension can refer to multiple tables in multiple data marts within the same organization.

[Read More Answers.](#)

**Question # 22**

Tell us what are the different types of SCD's used in data warehousing?

**Answer:-**

SCD (Slowly changing dimensions), are the dimensions in which the data changes slowly, rather than changing regularly on a time basis.

Three types of SCDs are used in data warehousing, which are defined as:

\* - SCD1: It is a record that is used to replace the original record even there is only one record existing in the database. The current data will be replaced and the new data will take its place.

\* - SCD2: It is the new record file that is added to the dimension table. This record exists in the database with the current data and previous data that is stored in the history.

\* - SCD3: This uses the original data that is modified to the new data. This consists of two records: one record that exist in the database and another record that will replace the old database record with the new information.

[Read More Answers.](#)

**Question # 23**



Tell us what is level of Granularity of a fact table?

**Answer:-**

A fact table is usually designed at a low level of Granularity. This means that we need to find the lowest level of information that can store in a fact table. e.g. Employee performance is a very high level of granularity. Employee\_performance\_daily, employee\_performace\_weekly can be considered lower levels of granularity.

The granularity is the lowest level of information stored in the fact table. The depth of data level is known as granularity. In date dimension, the level could be year, month, quarter, period, week, day of granularity.

The process consists of the following two steps:

\* Determining the dimensions that are to be included

\* Determining the location to locate the hierarchy of each dimension of information. The above factors of determination will be resented to the requirements.

[Read More Answers.](#)

**Question # 24**

Do you know what is Virtual Data Warehousing?

**Answer:-**

\* A virtual data warehouse provides a collective view of the completed data. A virtual data warehouse has no historic data. It can be considered as a logical data model of the containing metadata.

\* Virtual data warehousing is a 'de facto' information system strategy for supporting analytical decision making. It is one of the best ways for translating raw data and presenting it in the form that can be used by decision makers. It provides semantic map - which allows the end user for viewing as virtualized.

[Read More Answers.](#)

**Question # 25**

Tell us what language is the base of DMQL?

**Answer:-**

DMQL is based on Structured Query Language (SQL).

[Read More Answers.](#)

**Question # 26**

Tell us how many dimensions are selected in Slice operation?

**Answer:-**

Only one dimension is selected for the slice operation.

[Read More Answers.](#)

**Question # 27**

Do you know what is a warehouse manager?

**Answer:-**

Warehouse manager is responsible for the warehouse management process. The warehouse manager consist of third party system software, C programs and shell scripts. The size and complexity of warehouse manager varies between specific solutions.

[Read More Answers.](#)

**Question # 28**

Tell us how many fact tables are there in a star schema?

**Answer:-**

There is only one fact table in a star Schema.

[Read More Answers.](#)

**Question # 29**

What is the functions performed by OLAP?

**Answer:-**

OLAP performs functions such as roll-up, drill-down, slice, dice, and pivot.

[Read More Answers.](#)

**Question # 30**

Explain me which language is used for defining Schema Definition?

**Answer:-**

Data Mining Query Language (DMQL) is used for Schema Definition.

[Read More Answers.](#)

**Question # 31**

Can you tell me which one is faster, Multidimensional OLAP or Relational OLAP?

**Answer:-**

Multidimensional OLAP is faster than Relational OLAP.



[Read More Answers.](#)

### Question # 32

Tell us what is the benefit of normalization?

**Answer:-**

Normalization helps in reducing data redundancy.

[Read More Answers.](#)

### Question # 33

Can you explain me how does a Data Cube help?

**Answer:-**

Data cube helps us to represent the data in multiple dimensions. The data cube is defined by dimensions and facts.

[Read More Answers.](#)

### Question # 34

Explain me how many dimensions are selected in dice operation?

**Answer:-**

For dice operation two or more dimensions are selected for a given cube.

[Read More Answers.](#)

### Question # 35

Tell us what do you mean by Data Extraction?

**Answer:-**

Data extraction means gathering data from multiple heterogeneous sources.

[Read More Answers.](#)

### Question # 36

Do you know what is Data Warehousing?

**Answer:-**

Data Warehousing is the process of constructing and using the data warehouse.

[Read More Answers.](#)

### Question # 37

Explain me which one is faster, Multidimensional OLAP or Relational OLAP?

**Answer:-**

\* Multidimensional OLAP is faster than Relational OLAP.

\* MOLAP: Multi-dimensional OLAP

\* Data is stored in a multidimensional cube. The storage is not in the relational database, but in proprietary formats (one example is PowerOLAP's .olp file). MOLAP products can be compatible with Excel, which can make data interactions easy to learn.

\* ROLAP: Relational OLAP

\* ROLAP products access a relational database by using SQL (structured query language), which is the standard language that is used to define and manipulate data in an RDBMS. Subsequent processing may occur in the RDBMS or within a mid-tier server, which accepts requests from clients, translates them into SQL statements, and passes them on to the RDBMS.

[Read More Answers.](#)

### Question # 38

Explain me what is ODS?

**Answer:-**

\* An operational data store ("ODS") is a database designed to integrate data from multiple sources for additional operations on the data. Unlike a master data store, the data is not sent back to operational systems. It may be passed for further operations and to the data warehouse for reporting.

\* In ODS, data can be scrubbed, resolved for redundancy and checked for compliance with the corresponding business rules. This data store can be used for integrating disparate data from multiple sources so that business operations, analysis and reporting can be carried while business operations occur. This is the place where most of the data used in current operation is housed before it's transferred to the data warehouse for longer term storage or archiving.

\* An ODS is designed for relatively simple queries on small amounts of data (such as finding the status of a customer order), rather than the complex queries on large amounts of data typical of the data warehouse.

\* An ODS is similar to your short term memory where it only stores very recent information. On the contrary, the data warehouse is more like long term memory storing relatively permanent information.

[Read More Answers.](#)

### Question # 39

Tell us what is the difference between agglomerative and divisive Hierarchical Clustering?

**Answer:-**

\* Agglomerative Hierarchical clustering method allows the clusters to be read from bottom to top so that the program always reads from the sub-component first then moves to the parent whereas Divisive Hierarchical clustering uses top-bottom approach in which the parent is visited first than the child.



\* Agglomerative hierarchical method consists of objects in which each object creates its own clusters and these clusters are grouped together to create a large cluster. It defines a process of continuous merging until all the single clusters are merged together into a complete big cluster that will consist of all the objects of child clusters. However, in divisive clustering, the parent cluster is divided into smaller cluster and it keeps on dividing until each cluster has a single object to represent.

[Read More Answers.](#)

### Question # 40

Explain me what is active data warehousing?

#### Answer:-

\* An active data warehouse represents a single state of the business. Active data warehousing considers the analytic perspectives of customers and SUPPLIERS. It helps to deliver the updated data through reports.

\* A form of repository of captured transactional data is known as 'active data warehousing'. Using this concept, trends and patterns are found to be used for future decision making. Active data warehouse has a feature which can integrate the changes of data while scheduled cycles refresh. Enterprises utilize an active data warehouse in drawing the company's image in statistical manner.

[Read More Answers.](#)

### Question # 41

Explain me why do we override the execute method in struts?

#### Answer:-

As part of Struts Framework, we can develop the Action Servlet, ActionForm servlets (ActionServlet means what class extends the Action class, and ActionForm means what class extends the Action Form class) and other servlet classes.

In case of ActionForm class, we can develop validate() method. This method will return the ActionErrors object. In this method we can write the validation code. If this method returns null or ActionErrors with size=0, the web container will call execute() as part of the Action class.

If it returns size > 0, it will not call the execute() method. It will rather execute the jsp, servlet or html file as value for the input attribute as part of the attribute in struts-config.xml file.

[Read More Answers.](#)

### Question # 42

Tell us what is the very basic difference between data warehouse and operational databases?

#### Answer:-

A data warehouse contains historical information that is made available for analysis of the business whereas an operational database contains current information that is required to run the business.

[Read More Answers.](#)

### Question # 43

What are the functions of a load manager?

#### Answer:-

A load manager extracts data from the source system. Fast load the extracted data into temporary data store. Perform simple transformations into structure similar to the one in the data warehouse.

[Read More Answers.](#)

### Question # 44

Can you define metadata?

#### Answer:-

Metadata is simply defined as data about data. In other words, we can say that metadata is the summarized data that leads us to the detailed data.

[Read More Answers.](#)

### Question # 45

Do you know what is XMLA?

#### Answer:-

\* XMLA is XML for Analysis which can be considered as a standard for accessing data in OLAP, data mining or data sources on the internet. It is Simple Object Access Protocol. XMLA uses 'discover' and 'Execute' methods. Discover fetches information from the internet while Execute allows the applications to execute against the data sources.

\* XMLA is an industry standard for accessing data in analytical systems, such as OLAP. It is based on XML, SOAP and HTTP.

\* XMLA specifies MDXML as the query language. In the XMLA 1.1 version, the only construct in MDXML is an MDX statement enclosed in the tag

[Read More Answers.](#)

### Question # 46

Tell us what does subject-oriented data warehouse signify?

#### Answer:-

Subject oriented signifies that the data warehouse stores the information around a particular subject such as product, customer, sales, etc.

[Read More Answers.](#)

### Question # 47

Explain me what is the difference between view and materialized view?



### Answer:-

View:

- \* Tail raid data representation is provided by a view to access data from its table.
- \* It has logical structure that does not occupy space.
- \* Changes get affected in corresponding tables.

Materialized view:

- \* Pre-calculated data persists in materialized view.
- \* It has physical data space occupation.
- \* Changes will not get affected in corresponding tables.

[Read More Answers.](#)

### Question # 48

Do you know what does Metadata Respiratory contain?

### Answer:-

Metadata respiratory contains definition of data warehouse, business metadata, operational metadata, data for mapping from operational environment to data warehouse, and the algorithms for summarization.

[Read More Answers.](#)

### Question # 49

Explain me the functions of a warehouse manager?

### Answer:-

The warehouse manager performs consistency and referential integrity checks, creates the indexes, business views, partition views against the base data, transforms and merge the source data into the temporary store into the published data warehouse, backs up the data in the data warehouse, and archives the data that has reached the end of its captured life.

[Read More Answers.](#)

### Question # 50

Explain me the functions of data warehouse tools and utilities?

### Answer:-

The functions performed by Data warehouse tool and utilities are Data Extraction, Data Cleaning, Data Transformation, Data Loading and Refreshing.

[Read More Answers.](#)

### Question # 51

Tell us how do you load the time dimension?

### Answer:-

Time dimensions are usually loaded by a program that loops through all possible dates appearing in the data. It is not unusual for 100 years to be represented in a time dimension, with one row per day.

[Read More Answers.](#)

### Question # 52

Do you know what is load manager?

### Answer:-

A load manager performs the operations required to extract and load the process. The size and complexity of load manager varies between specific solutions from data warehouse to data warehouse.

[Read More Answers.](#)

### Question # 53

Can you list any five applications of data warehouse?

### Answer:-

Some applications include financial services, banking services, customer goods, retail sectors, controlled manufacturing.

[Read More Answers.](#)

### Question # 54

What is the phases involved in the data warehouse delivery process?

### Answer:-

The stages are IT strategy, Education, Business Case Analysis, technical Blueprint, Build the version, History Load, Ad hoc query, Requirement Evolution, Automation, and Extending Scope.

[Read More Answers.](#)

### Question # 55

Can you define data warehouse?

### Answer:-

Data warehouse is a subject oriented, integrated, time-variant, and nonvolatile collection of data that supports management's decision-making process.





[Read More Answers.](#)

Global Guideline - COM

## Data Warehouse Most Popular Interview Topics.

- 1 : [Data Warehousing Frequently Asked Interview Questions and Answers Guide.](#)
- 2 : [Ab Initio Frequently Asked Interview Questions and Answers Guide.](#)
- 3 : [Informatica Data Warehousing Frequently Asked Interview Questions and Answers Guide.](#)
- 4 : [ColdFusion Frequently Asked Interview Questions and Answers Guide.](#)
- 5 : [Cognos Data Warehouse Frequently Asked Interview Questions and Answers Guide.](#)
- 6 : [Business Objects Frequently Asked Interview Questions and Answers Guide.](#)
- 7 : [Micro Strategy Frequently Asked Interview Questions and Answers Guide.](#)
- 8 : [Warehouse Assistant Frequently Asked Interview Questions and Answers Guide.](#)
- 9 : [Extract, transform, load \(ETL\) Frequently Asked Interview Questions and Answers Guide.](#)
- 10 : [BO Designer 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](http://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](mailto:Info@globalguideline.com)