

# General Datawarehousing Job Interview Questions And Answers



**Interview Questions Answers**

<https://interviewquestionsanswers.org/>

## About Interview Questions Answers

**Interview Questions Answers . ORG** is an interview preparation guide of thousands of Job Interview Questions And Answers, Job Interviews are always stressful even for job seekers who have gone on countless interviews. The best way to reduce the stress is to be prepared for your job interview. Take the time to review the standard interview questions you will most likely be asked. These interview questions and answers on General Datawarehousing will help you strengthen your technical skills, prepare for the interviews and quickly revise the concepts.

If you find any **question or answer** is incorrect or incomplete then you can **submit your question or answer** directly with out any registration or login at our website. You just need to visit [General Datawarehousing Interview Questions And Answers](#) to add your answer click on the *Submit Your Answer* links on the website; with each question to post your answer, if you want to ask any question then you will have a link *Submit Your Question*; that's will add your question in General Datawarehousing category. To ensure quality, each submission is checked by our team, before it becomes live. This [General Datawarehousing Interview preparation PDF](#) was generated at **Wednesday 29th November, 2023**

You can follow us on FaceBook for latest Jobs, Updates and other interviews material.  
[www.facebook.com/InterviewQuestionsAnswers.Org](http://www.facebook.com/InterviewQuestionsAnswers.Org)

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

If you need any further assistance or have queries regarding this document or its material or any of other inquiry, please do not hesitate to contact us.

Best Of Luck.

**Interview Questions Answers.ORG Team**  
<https://InterviewQuestionsAnswers.ORG/>  
[Support@InterviewQuestionsAnswers.ORG](mailto:Support@InterviewQuestionsAnswers.ORG)



# General Datawarehousing Interview Questions And Answers Guide.

## Question - 1:

Explain me what are conformed dimensions?

### Ans:

Conformed dimensions are dimensions which are common to the cubes.(cubes are the schemas contains facts and dimension tables)  
Consider Cube-1 contains F1,D1,D2,D3 and Cube-2 contains F2,D1,D2,D4 are the Facts and Dimensions  
here D1,D2 are the Conformed Dimensions one dimension can share with more fact tables through primary key and foreign key relationship.

[View All Answers](#)

## Question - 2:

How to load the time dimension?

### Ans:

Time dimension are used to represent the datas or measures over a certain period of time.The server time dimension is the most widley used one by which we can represent the datas in hierachal manner such as quarter->year->months->week wise representations.

[View All Answers](#)

## Question - 3:

Explain Why are OLTP database designs not generally a good idea for a Data Warehouse?

### Ans:

Since in OLTP,tables are normalised and hence query response will be slow for end user and OLTP doesnot contain years of data and hence cannot be analysed.

[View All Answers](#)

## Question - 4:

Explain What are conformed dimensions?

### Ans:

Conformed dimensions mean the exact same thing with every possible fact table to which they are joined  
Ex:Date Dimensions is connected all facts like Sales facts,Inventory facts..etc

[View All Answers](#)

## Question - 5:

Explain What does level of Granularity of a fact table signify?

### Ans:

Granularity  
The first step in designing a fact table is to determine the granularity of the fact table. By granularity, we mean the lowest level of information that will be stored in the fact table. This constitutes two steps:  
Determine which dimensions will be included.  
Determine where along the hierarchy of each dimension the information will be kept.  
The determining factors usually goes back to the requirements

[View All Answers](#)

## Question - 6:



What is a Fact table?

**Ans:**

Fact Table contains the measurements or metrics or facts of business process. If your business process is "Sales", then a measurement of this business process such as "monthly sales number" is captured in the Fact table. Fact table also contains the foreign keys for the dimension tables.

Facts are organized in a table is called Fact table.

A Fact is a numeric values or a Business measure.

Every numeric is not a fact. a numeric which occupied a key performance indicator is called Facts

A Fact table contains a Facts at lower granularity level

[View All Answers](#)

**Question - 7:**

What is a data warehousing?

**Ans:**

Data Warehouse is a repository of integrated information, available for queries and analysis. Data and information are extracted from heterogeneous sources as they are generated....This makes it much easier and more efficient to run queries over data that originally came from different sources.

Typical relational databases are designed for on-line transactional processing (OLTP) and do not meet the requirements for effective on-line analytical processing (OLAP). As a result, data warehouses are designed differently than traditional relational databases.

A Data Warehousing is defined in 2 ways by 2 authors named "Ralph Kimball" and "W.H.Inman"

According to Ralph Kimball, . A D.W.H is a relational database which is specially design for business analysis but not for running the business.

. An enterprise D.W.H is design to make decision process. Hence it is called Decision Support System.

. A Data Warehouse is design to only read operations required for business analysis but not for transactional process. Hence it is called Read Only Database.

According to W.H.Inman, A Data Warehouse is a,

- 1) Time variant Database
- 2) Non-Volatile Database
- 3) Integrated Database
- 4) Subject oriented Database

and a Data Warehouse is a historical database

[View All Answers](#)

**Question - 8:**

Explain What are non-additive facts?

**Ans:**

Non-Additive: Non-additive facts are facts that cannot be summed up for any of the dimensions present in the fact table.

[View All Answers](#)

**Question - 9:**

Explain some of modeling tools available in the Market?

**Ans:**

These tools are used for Data/dimension modeling

1. Oracle Designer
2. ERWin (Entity Relationship for windows)
3. Informatica (Cubes/Dimensions)
4. Embarcadero
5. Power Designer Sybase

[View All Answers](#)

**Question - 10:**

What is an ER Diagram?

**Ans:**

The Entity-Relationship (ER) model was originally proposed by Peter in 1976 [Chen76] as a way to unify the network and relational database views.

Simply stated the ER model is a conceptual data model that views the real world as entities and relationships. A basic component of the model is the Entity-Relationship diagram which is used to visually represents data objects.

Since Chen wrote his paper the model has been extended and today it is commonly used for database design For the database designer, the utility of the ER model is:

it maps well to the relational model. The constructs used in the ER model can easily be transformed into relational tables.

it is simple and easy to understand with a minimum of training. Therefore, the model can be used by the database designer to communicate the design to the end user.

In addition, the model can be used as a design plan by the database developer to implement a data model in a specific database management software.

[View All Answers](#)

**Question - 11:**

Explain Why should you put your data warehouse on a different system than your OLTP system?

**Ans:**

A OLTP system is basically " data oriented " (ER model) and not " Subject oriented "(Dimensional Model) .That is why we design a separate system that will have a subject oriented OLAP system...

Moreover if a complex query is fired on a OLTP system will cause a heavy overhead on the OLTP server that will affect the daytoday business directly.

[View All Answers](#)

**Question - 12:**



List down some of the real time data-warehousing tools?

**Ans:**

ETL:

Informatica,  
Abinitio,  
Datastage etc.,

OLAP:

Business objects  
Cognos,  
Micro strategy,  
Hyperion etc.,

DW:

Oracle,  
DB2,  
Terradata,  
Sybase,  
Greenplum etc.,

[View All Answers](#)

**Question - 13:**

What is a general purpose scheduling tool?

**Ans:**

The basic purpose of the scheduling tool in a DW Application is to stream line the flow of data from Source To Target at specific time or based on some condition.

[View All Answers](#)

**Question - 14:**

Explain How are the Dimension tables designed?

**Ans:**

Most dimension tables are designed using Normalization principles upto 2NF. In some instances they are further normalized to 3NF.

Find where data for this dimension are located.

Figure out how to extract this data.

Determine how to maintain changes to this dimension (see more on this in the next section).

Change fact table and DW population routines.

[View All Answers](#)

**Question - 15:**

What is a dimension table?

**Ans:**

A dimensional table is a collection of hierarchies and categories along which the user can drill down and drill up. it contains only the textual attributes.

[View All Answers](#)

**Question - 16:**

Explain Which columns go to the fact table and which columns go the dimension table?

**Ans:**

The Primary Key columns of the Tables(Entities) go to the Dimension Tables as Foreign Keys.

The Primary Key columns of the Dimension Tables go to the Fact Tables as Foreign Keys.

[View All Answers](#)

**Question - 17:**

What is a Star Schema?

**Ans:**

Star schema is a type of organising the tables such that we can retrieve the result from the database easily and fastly in the warehouse environment.Usually a star schema consists of one or more dimension tables around a fact table which looks like a star,so that it got its name.

[View All Answers](#)

**Question - 18:**

Explain What are slowly changing dimensions?

**Ans:**

SCD stands for Slowly changing dimensions. Slowly changing dimensions are of three types

SCD1: only maintained updated values.

Ex: a customer address modified we update existing record with new address.

SCD2: maintaining historical information and current information by using

A) Effective Date

B) Versions

C) Flags

or combination of these

scd3: by adding new columns to target table we maintain historical information and current information



[View All Answers](#)

### Question - 19:

What is data mining?

#### Ans:

Data mining is a process of extracting hidden trends within a datawarehouse. For example an insurance dataware house can be used to mine data for the most high risk people to insure in a certain geographical area.

[View All Answers](#)

### Question - 20:

What are modeling tools available in the Market?

#### Ans:

There are a number of data modeling tools  
Tool Name Company Name  
Erwin Computer Associates  
Embarcadero Embarcadero Technologies  
Rational Rose IBM Corporation  
Power Designer Sybase Corporation  
Oracle Designer Oracle Corporation

[View All Answers](#)

### Question - 21:

What are the advantages of RAID 1, 1/0, and 5. What type of RAID setup would you put your TX logs?

#### Ans:

Transaction logs write sequentially and don't need to be read at all. The ideal is to have each on RAID 1/0 because it has much better write performance than RAID 5. RAID 1 is also better for TX logs and costs less than 1/0 to implement. It has a tad less reliability and performance is a little worse generally speaking. RAID 5 is best for data generally because of cost and the fact it provides great read capability.

[View All Answers](#)

### Question - 22:

Explain What type of Indexing mechanism do we need to use for a typical datawarehouse?

#### Ans:

On the fact table it is best to use bitmap indexes. Dimension tables can use bitmap and/or the other types of clustered/non-clustered, unique/non-unique indexes. To my knowledge, SQLServer does not support bitmap indexes. Only Oracle supports bitmaps.

[View All Answers](#)

### Question - 23:

What is pre-emptive and non-pre-emptive?

#### Ans:

Preemptive means taken as a measure against something possible, anticipated, or feared; preventive; deterrent: a preemptive tactic against a ruthless business rival. Non Pre-emptive is the exact opposite to Pre-emptive.No such preventive measures has been taken.

[View All Answers](#)

### Question - 24:

What are various ETL tools in the Market?

#### Ans:

Various ETL tools used in market are:  
1. Informatica  
2. Data Stage  
3. MS-SQL DTS(Integrated Services 2005)  
4. Abinitio  
5. SQL Loader  
6. Sunopsis  
7. Oracle Warehouse Bulider  
8. Data Junction

[View All Answers](#)

### Question - 25:

Explain What is real time data-warehousing?

#### Ans:

Real-time data warehousing is a combination of two things: 1) real-time activity and 2) data warehousing. Real-time activity is activity that is happening right now. The activity could be anything such as the sale of widgets. Once the activity is complete, there is data about it. Data warehousing captures business activity data. Real-time data warehousing captures business activity data as it occurs. As soon as the business activity is complete and there is data about it, the completed activity data flows into the data warehouse and becomes available instantly. In other words, real-time data warehousing is a framework for deriving information from data as the data becomes available.

[View All Answers](#)



### Question - 26:

What is the Differences between star and snowflake schemas?

#### Ans:

Star schema - all dimensions will be linked directly with a fact table.

Snow schema - dimensions may be interlinked or may have one-to-many relationship with other tables.

[View All Answers](#)

### Question - 27:

What is a lookup table?

#### Ans:

A lookup table is the one which is used when updating a warehouse. When the lookup is placed on the target table (fact table / warehouse) based upon the primary key of the target, it just updates the table by allowing only new records or updated records based on the lookup condition.

[View All Answers](#)

### Question - 28:

What is Snow Flake Schema?

#### Ans:

Snowflake Schema, each dimension has a primary dimension table, to which one or more additional dimensions can join. The primary dimension table is the only table that can join to the fact table.

[View All Answers](#)

### Question - 29:

Explain the Difference between OLTP and OLAP?

#### Ans:

Main Differences between OLTP and OLAP are:-

1. User and System Orientation

OLTP: customer-oriented, used for data analysis and querying by clerks, clients and IT professionals.

OLAP: market-oriented, used for data analysis by knowledge workers (managers, executives, analysts).

2. Data Contents

OLTP: manages current data, very detail-oriented.

OLAP: manages large amounts of historical data, provides facilities for summarization and aggregation, stores information at different levels of granularity to support decision making process.

3. Database Design

OLTP: adopts an entity relationship (ER) model and an application-oriented database design.

OLAP: adopts star, snowflake or fact constellation model and a subject-oriented database design.

4. View

OLTP: focuses on the current data within an enterprise or department.

OLAP: spans multiple versions of a database schema due to the evolutionary process of an organization; integrates information from many organizational locations and data stores

[View All Answers](#)

### Question - 30:

Explain What are the various Reporting tools in the Market?

#### Ans:

1. MS-Excel

2. Business Objects (Crystal Reports)

3. Cognos (Impromptu, Power Play)

4. Microstrategy

5. MS reporting services

6. Informatica Power Analyzer

7. Actuate

8. Hyperion (BRIO)

9. Oracle Express OLAP

10. Proclarity

[View All Answers](#)

### Question - 31:

Explain ssl?

#### Ans:

The Secure Sockets Layer (SSL) is a commonly-used protocol for managing the security of a message transmission on the Internet. SSL has recently been succeeded by Transport Layer Security (TLS), which is based on SSL. SSL uses a program layer located between the Internet's Hypertext Transfer Protocol (HTTP) and Transport Control Protocol (TCP) layers. SSL is included as part of both the Microsoft and Netscape browsers and most Web server products. Developed by Netscape, SSL also gained the support of Microsoft and other Internet client/server developers as well and became the de facto standard until evolving into Transport Layer Security. The "sockets" part of the term refers to the sockets method of passing data back and forth between a client and a server program in a network or between program layers in the same computer. SSL uses the public-and-private key encryption system from RSA, which also includes the use of a digital certificate.

TLS and SSL are an integral part of most Web browsers (clients) and Web servers. If a Web site is on a server that supports SSL, SSL can be enabled and specific Web pages can be identified as requiring SSL access. Any Web server can be enabled by using Netscape's SSLRef program library which can be downloaded for noncommercial use or licensed for commercial use.

TLS and SSL are not interoperable. However, a message sent with TLS can be handled by a client that handles SSL but not TLS.



[View All Answers](#)

### Question - 32:

Explain What is What are Semi-additive and factless facts and in which scenario will you use such kinds of fact tables?

#### Ans:

Snapshot facts are semi-additive, while we maintain aggregated facts we go for semi-additive.

EX: Average daily balance

A fact table without numeric fact columns is called factless fact table.

Ex: Promotion Facts

While maintain the promotion values of the transaction (ex: product samples) because this table doesn't contain any measures.

[View All Answers](#)

### Question - 33:

What is SCD1 , SCD2 , SCD3?

#### Ans:

SCD Stands for Slowly changing dimensions.

SCD1: only maintained updated values.

Ex: a customer address modified we update existing record with new address.

SCD2: maintaining historical information and current information by using

A) Effective Date

B) Versions

C) Flags

or combination of these

SCD3: by adding new columns to target table we maintain historical information and current information.

[View All Answers](#)

### Question - 34:

Explain me what is VLDB?

#### Ans:

VLDB stands for Very Large DataBase.

It is an environment or storage space managed by a relational database management system (RDBMS) consisting of vast quantities of information.

[View All Answers](#)

### Question - 35:

What is a level of Granularity of a fact table?

#### Ans:

Level of granularity means level of detail that you put into the fact table in a data warehouse. For example: Based on design you can decide to put the sales data in each transaction. Now, level of granularity would mean what detail are you willing to put for each transactional fact. Product sales with respect to each minute or you want to aggregate it upto minute and put that data.

[View All Answers](#)

### Question - 36:

Explain me What are Data Marts?

#### Ans:

Data Marts are designed to help manager make strategic decisions about their business.

Data Marts are subset of the corporate-wide data that is of value to a specific group of users.

There are two types of Data Marts:

1.Independent data marts ? sources from data captured form OLTP system, external providers or from data generated locally within a particular department or geographic area.

2.Dependent data mart ? sources directly form enterprise data warehouses.

[View All Answers](#)

### Question - 37:

What is ODS?

#### Ans:

1. ODS means Operational Data Store.

Submitted by Francis C. ( xxchen74 @ hotmail . com )

2. A collection of operation or bases data that is extracted from operation databases and standardized, cleansed, consolidated, transformed, and loaded into an enterprise data architecture. An ODS is used to support data mining of operational data, or as the store for base data that is summarized for a data warehouse. The ODS may also be used to audit the data warehouse to assure summarized and derived data is calculated properly. The ODS may further become the enterprise shared operational database, allowing operational systems that are being reengineered to use the ODS as there operation databases.

[View All Answers](#)

### Question - 38:

Explain What are the Different methods of loading Dimension tables?

#### Ans:



Conventional Load:

Before loading the data, all the Table constraints will be checked against the data.

Direct load:(Faster Loading)

All the Constraints will be disabled. Data will be loaded directly.Later the data will be checked against the table constraints and the bad data won't be indexed.

[View All Answers](#)

### Question - 39:

Explain piconet?

**Ans:**

The original Piconet was a USB-style expansion port on RM Nimbus computers.

These days, a piconet is an ad-hoc computer network linking a user group of devices using Bluetooth technology protocols to allow one master device to interconnect with up to seven active slave devices (because a three-bit MAC address is used). Up to 255 further slave devices can be inactive, or parked, which the master device can bring into active status at any time.

A piconet typically has a range of about 10 m and a transfer rate between about 400 and 700 kbit/s, depending on whether synchronous or asynchronous connection is used.

All Parked Slaves have 8 bit parked member address (PMA) and all the active slaves have 3 bit active member address (AMA). The AMA is used by the master to send packets to a specific slave and to identify that the slave has sent a response packet.

[View All Answers](#)

### Question - 40:

What is Normalization, First Normal Form, Second Normal Form, Third Normal Form?

**Ans:**

1.Normalization is process for assigning attributes to entities?Reducesdata redundancies?Helps eliminate data anomalies?Produces controlledredundancies to link tables

2.Normalization is the analysis offunctional dependency between attributes / data items of userviewsδ€•It reduces a complex user view to a set of small andstable subgroups of fields / relations

1NF:Repeating groups must beeliminated, Dependencies can be identified, All key attributesdefined.No repeating groups in table

2NF: The Table is already in1NF,Includes no partial dependencies?No attribute dependent on a portionof primary key, Still possible to exhibit transitivedependency,Attributes may be functionally dependent on non-keyattributes

3NF: The Table is already in 2NF, Contains no transitivedependencies

[View All Answers](#)

# Data Warehouse Most Popular & Related Interview Guides

- 1 : [Warehouse Assistant Interview Questions and Answers.](#)
- 2 : [Ab Initio Interview Questions and Answers.](#)
- 3 : [Data Stage Interview Questions and Answers.](#)
- 4 : [ColdFusion Interview Questions and Answers.](#)
- 5 : [Data Warehouse Supervisor Interview Questions and Answers.](#)
- 6 : [Informatica Data Warehousing Interview Questions and Answers.](#)
- 7 : [Data Warehousing Interview Questions and Answers.](#)
- 8 : [Data Warehouse BI Interview Questions and Answers.](#)
- 9 : [ETL \(Extract, transform, load\) Interview Questions and Answers.](#)
- 10 : [Micro Strategy Interview Questions and Answers.](#)

Follow us on FaceBook

[www.facebook.com/InterviewQuestionsAnswers.Org](http://www.facebook.com/InterviewQuestionsAnswers.Org)

Follow us on Twitter

<https://twitter.com/InterviewQA>

For any inquiry please do not hesitate to contact us.

Interview Questions Answers.ORG Team

[https://InterviewQuestionsAnswers.ORG/  
support@InterviewQuestionsAnswers.ORG](https://InterviewQuestionsAnswers.ORG/support@InterviewQuestionsAnswers.ORG)