



## Top 20 Data Warehouse Interview Questions and Answers

Share on your Social Media



# Top 20 Data Warehouse Interview Questions and Answers

Published On: June 3, 2024

## Data Warehouse Interview Questions and Answers

Businesses worldwide are starting to use data warehouses more frequently, and their popularity is rising. The **top 20 data warehouse interview questions and answers** will be covered in this article.

[Download Data Warehouse Interview Questions PDF](#)

## Data Warehouse Interview Questions and Answers for Freshers

## Featured Articles

Want to know more about becoming an expert in IT?

[Click Here to Get Started](#)

100% Placement Assurance

AUTHORISED CERTIFICATION PARTNER

IBI



Quick Enquiry

## Related Courses at SLA

- [Data Warehouse Training in OMR](#)
- [Data Warehousing Training in Chennai](#)

## Related Posts



**MEAN Stack Interview Questions and Answers**

## 1. Define Data Warehouse

The storage of data used for management decision support systems is called data warehousing (DW). A data warehouse is made up of many different types of data with high business conditions at one particular moment in time.

To put it succinctly, it is a collection of comprehensive data that is accessible for analysis and query.

## 2. Define a dimension table

A dimension table is a table that includes measurement attributes that are kept in fact tables. This table includes categories, logic, and hierarchies that can be traversed in nodes.

## 3. What is a fact table?

The fact table has foreign keys for the dimension tables in addition to the measurement of business processes.

For instance, if the company's operations involve producing bricks,

The average quantity of bricks produced by a single person or machine is an indicator of the company's operations.

## 4. What phases does data warehousing go through?

The four phases of data warehousing are as follows:

- Offline Operational Database
- Offline Data Warehouse
- Real-Time Data warehouse
- Integrated Data warehouse

## 5. Define OLTP

On-Line Transaction Processing, or OLTP for short, is a program that processes a lot of data at once and

## Answers

Published On: June 19, 2024

Introduction Since MEAN Stack combines several other applications as part of its functionality, it is...

## Top 15 Struts Interview Questions and Answers

Published On: June 18, 2024

Struts Interview Questions and Answers When it comes to developing Java web applications, Struts is...

## Top 20 C Sharp Interview Questions and Answers

Published On: June 17, 2024

C Sharp Interview Questions and Answers Microsoft created the general-purpose programming language C# together with...

## Top 20 VB.Net Interview Questions and Answers

Published On: June 17, 2024

changes it whenever it is received.

## [Data Warehousing Syllabus PDF](#)

VB.Net Interview Questions and Answers A wide range of applications, including desktop, web, and mobile...

### **6. Describe OLAP**

Online Analytical Processing, or OLAP for short, is a system designed to gather, organize, and handle multi-dimensional data for management and analysis needs.

### **7. Explain Data mining**

Data mining is defined as the process of examining data from several angles and dimensions and compiling the results into a meaningful report. able to query and obtain data in their preferred format from the database.

### **8. What is ETL?**

The acronym for Extract, Transform, and Load is ETL. ETL software is used to extract a desired subset of data by reading data from a defined data source.

It then uses lookup tables and rules to change the data to the desired state.

The resultant data is then loaded into the target database using the load function.

### **9. What is data warehousing in real-time?**

Business data is captured in real-time data warehousing as it happens. The data will be available in the flow and immediately usable upon completion of the business action.

### **10. Describe business intelligence**

DSS, or decision support system, is another name for business intelligence. It describes the methods, tools, and technologies used in the gathering, combining, and analyzing of data or information relevant to business. Seeing the data in the

information itself is far more helpful.

## **11. What is ODS?**

An operational data store, or ODS for short, is a store for operational data that is collected in real-time as opposed to long-term trend data.

## **12. What are factless fact tables?**

If there are no numerical fact columns in a fact table, it is said to be factless.

## **13. Explain aggregate tables.**

The tables that hold the current warehouse data that have been categorized according to a specific degree of dimension are known as aggregate tables. Compared to the original table, which includes more records, the aggregated tables make data retrieval easier.

This table improves query performance while lightening the database server's workload.

## **14. What is a confirmed fact?**

A table known as a "conformed fact" may be integrated with various fact tables and utilized in multiple data marts.

## **15. What is datamart?**

A data mart is a customized form of data warehousing that provides a quick overview of operational data to assist decision-makers based on historical trends and experiences.

## **16. How is the time dimension loaded?**

Time dimensions can be loaded using a program, and they are typically loaded through all potential dates in a year. One row every day might be used to symbolize 100 years in this case.

## **17. Non-additive facts: what are they?**

Facts that cannot be summed up for any of the

dimensions shown in the fact table are referred to as non-additive facts. The same data may be helpful if the dimensions alter.

## **18. What is SCD?**

SCD pertains to instances where records undergo gradual alterations and are characterized by gradually altering dimensions.

## **19. What are the types of SCD?**

The following are the three forms of SCD:

SCD 1: A fresh record is used in place of the old one.

SCD 2: The current customer dimension table gains a new entry in SCD 2.

SCD 3: A data set is updated with new information.

**[Data Warehousing Salary](#)**

## **Data Warehouse Interview Questions and Answers for Experienced**

### **20. What are the benefits and drawbacks of designing data warehouses using a top-down approach?**

The top-down method has the following benefits:

- Data marts have a constant-dimensional perspective since they are constructed from data warehouses.
- It is also believed that this model works best for business transformations. Large organizations decide to use this approach as a result.
- A data mart can be easily created from a data warehouse.

The top-down approach's drawback is that it requires a lot of money, time, and effort to create and maintain.

### **21. What are the benefits and drawbacks**

## **of designing data warehouses using a bottom-up approach?**

The benefits of the bottom-up method are as follows:

- Since the data marts are constructed initially, the reports are generated promptly.
- This gives us more room to accommodate data marts, which enables us to grow our data warehouse.
- Furthermore, building this model requires very little money or time.

One drawback of the bottom-up strategy is that it is not as strong as the top-down approach since the dimensional view of data marts is not as consistent as it is in the top-down approach.

## **22. What does data cleansing entail?**

The phrase “data purging” refers to methods of permanently deleting and eliminating data from a storage area. A range of steps and methods are involved in data purging, which is sometimes contrasted with data deletion.

- Deleted data is usually considered a temporary preference, whereas purging permanently removes data and frees up memory or storage space for new uses.
- In database management, one technique for data cleansing is the use of automatic data purging features.
- For instance, several Microsoft applications provide an automated purge method that employs a circular buffer system to remove outdated data to make space for newer information.
- In other situations, administrators will need to manually remove data from the database.

## **23. What do you mean by a data lake?**

A data lake is a sizable location where unstructured, semi-structured, and structured data can be

stored. There are no file or account size limitations, so you can save any kind of material there in its original format. It offers a substantial quantity of data for enhanced native integration and analytical performance.

A massive container that resembles a lake or river is called a data lake. A data lake has machine-to-machine connectivity, structured and unstructured data, and real-time log flow, much like a lake with multiple tributaries.

## **24. Distinguish between divisive and agglomerative hierarchical clustering.**

**Divisive Clustering:** This method also does away with the requirement to specify how many clusters to create in advance. It requires a way to divide clusters recursively until all the data is split into singletons, starting with breaking a cluster that includes all of the data.

**Agglomerative hierarchical clustering:** Flat clustering results in an unorganized collection of clusters. However, this arrangement provides more information. This clustering approach does not require us to provide the number of clusters ahead of time.

Each piece of data is initially treated as a singleton cluster by bottom-up algorithms, which then aggregate pairings of clusters until every cluster is combined into a single cluster containing all of the data.

The distinctions between the two are as follows:

- Divisive clustering is more complex than agglomerative clustering since it requires a flat clustering algorithm as a "subroutine" to separate each cluster until every data point has its own singleton cluster.
- Dividend clustering is more effective if we don't build a full hierarchy down to the individual data leaves.

Furthermore, a divided algorithm is more accurate. Agglomerative clustering makes decisions based on nearby points or local patterns without first looking at the global distribution of the data. These are set-in-stone choices. Divisive clustering considers the global distribution of data while producing top-level division decisions.

## Conclusion

We covered the most common **data warehousing interview questions and answers** in this article. Our **[data warehousing training in Chennai](#)** can help you become knowledgeable.

**[Data Warehousing Training](#)**

Share on your Social Media



## Softlogic Academy

## Softlogic Systems

### KK Nagar [Corporate Office]

No.10, PT Rajan Salai, K.K. Nagar, Chennai  
– 600 078.

**Landmark:** Karnataka Bank Building

**Phone:** [+91 86818 84318](tel:+918681884318)

**Email:** [enquiry@softlogicsys.in](mailto:enquiry@softlogicsys.in)

**Map:** [Google Maps Link](#)

**OMR**

## Navigation

---

- [About Us](#)
- [Blog Posts](#)
- [Careers](#)
- [Contact](#)
- [Placement Training](#)
- [Corporate Training](#)
- [Hire With Us](#)
- [Job Seekers](#)
- [SLA's Recently Placed Students](#)
- [Reviews](#)
- [Sitemap](#)

## Important Links

---

- [Disclaimer](#)
- [Privacy Policy](#)



No. E1-A10, RTS Food Street  
92, Rajiv Gandhi Salai (OMR),  
Navalur, Chennai - 600 130.

**Landmark:** Adj. to AGS Cinemas

**Phone:** [+91 89256 88858](tel:+918925688858)

**Email:** [info@softlogicsys.in](mailto:info@softlogicsys.in)

**Map:** [Google Maps Link](#)

[Terms and Conditions](#)

## Courses

---

Python

Software Testing

Full Stack Developer

Java

Power BI

Clinical SAS

Data Science

Embedded

Cloud Computing

Hardware and Networking

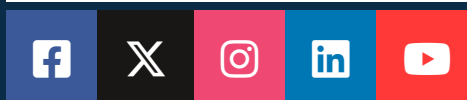
VBA Macros

Mobile App Development

DevOps

## Social Media Links

---



## Review Sources

---

Google

Trustpilot

Glassdoor

Mouthshut

Sulekha

Justdial

Ambitionbox

Indeed

Software Suggest

Sitejabber