

Data Warehouse and Data Mining Lab Viva Questions (50 Q&A)

Q1: What is data warehousing?

Ans: Data warehousing is the process of collecting, storing, and managing large volumes of data from different sources in a central repository for analysis and reporting.

Q2: What is a data mart?

Ans: A data mart is a subset of a data warehouse focused on a specific business line or team.

Q3: Define OLAP.

Ans: OLAP (Online Analytical Processing) is a technology that enables users to analyze data from multiple dimensions interactively.

Q4: What is ETL in data warehousing?

Ans: ETL stands for Extract, Transform, Load. It is the process of extracting data from source systems, transforming it into a suitable format, and loading it into a data warehouse.

Q5: What is a star schema?

Ans: A star schema is a type of database schema that is used in data warehousing. It consists of a central fact table connected to dimension tables.

Q6: Difference between star schema and snowflake schema?

Ans: In a star schema, dimension tables are denormalized, while in a snowflake schema, dimension tables are normalized into multiple related tables.

Q7: What is data mining?

Ans: Data mining is the process of discovering patterns, correlations, and useful information from large sets of data using machine learning, statistics, and database systems.

Q8: What is classification in data mining?

Ans: Classification is a data mining technique used to predict categorical labels or classes of data

items based on training data.

Q9: What is clustering?

Ans: Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.

Q10: What is support and confidence in association rule mining?

Ans: Support indicates how frequently an itemset appears in the dataset, while confidence indicates how often a rule has been found to be true.

Q11: What is data warehousing?

Ans: Data warehousing is the process of collecting, storing, and managing large volumes of data from different sources in a central repository for analysis and reporting.

Q12: What is a data mart?

Ans: A data mart is a subset of a data warehouse focused on a specific business line or team.

Q13: Define OLAP.

Ans: OLAP (Online Analytical Processing) is a technology that enables users to analyze data from multiple dimensions interactively.

Q14: What is ETL in data warehousing?

Ans: ETL stands for Extract, Transform, Load. It is the process of extracting data from source systems, transforming it into a suitable format, and loading it into a data warehouse.

Q15: What is a star schema?

Ans: A star schema is a type of database schema that is used in data warehousing. It consists of a central fact table connected to dimension tables.

Q16: Difference between star schema and snowflake schema?

Ans: In a star schema, dimension tables are denormalized, while in a snowflake schema, dimension tables are normalized into multiple related tables.

Q17: What is data mining?

Ans: Data mining is the process of discovering patterns, correlations, and useful information from large sets of data using machine learning, statistics, and database systems.

Q18: What is classification in data mining?

Ans: Classification is a data mining technique used to predict categorical labels or classes of data items based on training data.

Q19: What is clustering?

Ans: Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.

Q20: What is support and confidence in association rule mining?

Ans: Support indicates how frequently an itemset appears in the dataset, while confidence indicates how often a rule has been found to be true.

Q21: What is data warehousing?

Ans: Data warehousing is the process of collecting, storing, and managing large volumes of data from different sources in a central repository for analysis and reporting.

Q22: What is a data mart?

Ans: A data mart is a subset of a data warehouse focused on a specific business line or team.

Q23: Define OLAP.

Ans: OLAP (Online Analytical Processing) is a technology that enables users to analyze data from multiple dimensions interactively.

Q24: What is ETL in data warehousing?

Ans: ETL stands for Extract, Transform, Load. It is the process of extracting data from source systems, transforming it into a suitable format, and loading it into a data warehouse.

Q25: What is a star schema?

Ans: A star schema is a type of database schema that is used in data warehousing. It consists of a central fact table connected to dimension tables.

Q26: Difference between star schema and snowflake schema?

Ans: In a star schema, dimension tables are denormalized, while in a snowflake schema, dimension tables are normalized into multiple related tables.

Q27: What is data mining?

Ans: Data mining is the process of discovering patterns, correlations, and useful information from large sets of data using machine learning, statistics, and database systems.

Q28: What is classification in data mining?

Ans: Classification is a data mining technique used to predict categorical labels or classes of data items based on training data.

Q29: What is clustering?

Ans: Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.

Q30: What is support and confidence in association rule mining?

Ans: Support indicates how frequently an itemset appears in the dataset, while confidence indicates how often a rule has been found to be true.

Q31: What is data warehousing?

Ans: Data warehousing is the process of collecting, storing, and managing large volumes of data from different sources in a central repository for analysis and reporting.

Q32: What is a data mart?

Ans: A data mart is a subset of a data warehouse focused on a specific business line or team.

Q33: Define OLAP.

Ans: OLAP (Online Analytical Processing) is a technology that enables users to analyze data from

multiple dimensions interactively.

Q34: What is ETL in data warehousing?

Ans: ETL stands for Extract, Transform, Load. It is the process of extracting data from source systems, transforming it into a suitable format, and loading it into a data warehouse.

Q35: What is a star schema?

Ans: A star schema is a type of database schema that is used in data warehousing. It consists of a central fact table connected to dimension tables.

Q36: Difference between star schema and snowflake schema?

Ans: In a star schema, dimension tables are denormalized, while in a snowflake schema, dimension tables are normalized into multiple related tables.

Q37: What is data mining?

Ans: Data mining is the process of discovering patterns, correlations, and useful information from large sets of data using machine learning, statistics, and database systems.

Q38: What is classification in data mining?

Ans: Classification is a data mining technique used to predict categorical labels or classes of data items based on training data.

Q39: What is clustering?

Ans: Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.

Q40: What is support and confidence in association rule mining?

Ans: Support indicates how frequently an itemset appears in the dataset, while confidence indicates how often a rule has been found to be true.

Q41: What is data warehousing?

Ans: Data warehousing is the process of collecting, storing, and managing large volumes of data

from different sources in a central repository for analysis and reporting.

Q42: What is a data mart?

Ans: A data mart is a subset of a data warehouse focused on a specific business line or team.

Q43: Define OLAP.

Ans: OLAP (Online Analytical Processing) is a technology that enables users to analyze data from multiple dimensions interactively.

Q44: What is ETL in data warehousing?

Ans: ETL stands for Extract, Transform, Load. It is the process of extracting data from source systems, transforming it into a suitable format, and loading it into a data warehouse.

Q45: What is a star schema?

Ans: A star schema is a type of database schema that is used in data warehousing. It consists of a central fact table connected to dimension tables.

Q46: Difference between star schema and snowflake schema?

Ans: In a star schema, dimension tables are denormalized, while in a snowflake schema, dimension tables are normalized into multiple related tables.

Q47: What is data mining?

Ans: Data mining is the process of discovering patterns, correlations, and useful information from large sets of data using machine learning, statistics, and database systems.

Q48: What is classification in data mining?

Ans: Classification is a data mining technique used to predict categorical labels or classes of data items based on training data.

Q49: What is clustering?

Ans: Clustering is the process of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.

Q50: What is support and confidence in association rule mining?

Ans: Support indicates how frequently an itemset appears in the dataset, while confidence indicates how often a rule has been found to be true.