40-item quiz on Data Mining and Data Warehousing, complete with answers. The questions are a mix of multiple-choice (MCQ), true/false, and short answer types for variety.

DATA MINING AND DATA WAREHOUSING QUIZ

Part A: Multiple Choice Questions (MCQs) – (20 items)

1. What is the purpose of data mining?

A. Data storage

B. Data cleaning

C. Discovering patterns and knowledge from large amounts of data

D. Data deletion

2. A data warehouse is primarily used for:

A. Online Transaction Processing

B. Storing passwords

C. Data analytics and decision support

D. File compression

3. Which of the following is NOT a data mining task?

A. Classification

B. Clustering

C. Summarization

D. Deleting redundant data

4. ETL stands for:

A. Extract, Transform, Load

B. Evaluate, Transform, Load

C. Extract, Transfer, Locate

D. Export, Transform, Load

5. OLAP stands for:

A. Online Analytical Processing

B. Offline Analytical Programming

C. Online Application Program

D. Object-Level Analysis Protocol

6. Which algorithm is used for classification in data mining?

A. K-Means

B. Apriori

C. Decision Trees

D. DBSCAN

7. Which one is a clustering algorithm?

A. K-Means

B. Naive Bayes

C. Decision Tree

D. FP-Growth

8. The star schema in a data warehouse consists of:

A. Only dimension tables

B. Fact and dimension tables

C. Only fact tables

D. None of the above

9. Which of the following is a type of OLAP?

A. MOLAP

B. DOLAP

C. HOLAP

D. All of the above

10. Association rule mining is used to find:

A. Time series

B. Correlations in data

C. Data summaries

D. Data normalization

11. Which metric measures the strength of association rules?

A. Accuracy

B. Confidence

C. Entropy

D. ROC

12. What is a snowflake schema?

A. A normalized star schema

B. A schema with no dimensions

C. An unstructured data schema

D. A schema used in operational databases

13. Which method reduces data volume but produces the same or similar analytical results?

A. Data cleansing

B. Data transformation

C. Data reduction

D. Data discretization

14. Outlier detection is used to:

A. Find missing values

B. Discover anomalies in data

C. Normalize data

D. Identify clusters

15. Which of the following is a predictive data mining task?

A. Clustering

B. Classification

C. Summarization

D. Association

16. Dimensionality reduction can be achieved by:

A. KNN

B. PCA

C. K-Means

D. Decision Tree

17. Which of these is a data preprocessing step?

A. Data warehousing

B. Data mining

C. Data cleaning

D. OLAP

18. Which data mining technique is best for customer segmentation?

A. Classification

B. Clustering

C. Regression

D. Association

19. Metadata in a data warehouse defines:

A. Data structure and constraints

B. Facts and dimensions

C. Query performance

D. Data loading schedule

20.Which is the last step in the KDD (Knowledge Discovery in Databases) process?

A. Data cleaning

B. Data selection

C. Data mining

D. Interpretation/Evaluation

Part B: True or False – (10 items)

21. Data mining and data warehousing are the same.

22. Data cubes are used in OLAP to represent multidimensional data.

23. Fact tables in a data warehouse store descriptive information.

24. Classification is a supervised learning technique.

25. ETL processes are not necessary in data warehousing.

26. Apriori algorithm is used for association rule mining.

27. Noise in data mining refers to valuable patterns.

28. Data normalization transforms data into a standard scale.

29. A data mart is smaller and more focused than a data warehouse.

30. Regression analysis can only be used for classification problems.

Part C: Short Answer – (10 items)

31. Define data warehousing.

32. What is data cleaning?

33. Name two common data mining algorithms.

34. What is the difference between OLTP and OLAP?

35. What is the role of a fact table in a star schema?

36. What does clustering do in data mining?

37. Define metadata in data warehousing.

38. List one use case for association rule mining.

39. What is a dimension table?

40. Why is data preprocessing important?