

Data Warehousing & Mining

**Unit 1**

---

1. What is data warehouse? (2017,2016,2014)
2. What are the advantages of data warehouse? (2017)
3. What is the future scope of data warehouse? (2017)
4. Define data cleaning. (2017)
5. Discuss star flake schema. (2017,2014,2013)
6. Explain capacity planning in data warehouse. (2017)
7. Explain data warehouse delivery methods. (2017)
8. Write short note on fact constellations. (2017)
9. Explain OLAP servers along with OLTP function and tools in details. (2017)
10. Write and explain the difference between database system and data warehouse. (2017)
11. Define OLTP. (2016)
12. What are data cubes? (2016)
13. Differentiate OLAP and OLTP. (2016,2015)
14. Explain clean and transform process with an illustrative example. (2016)
15. What do you mean by tuning the data warehouse? (2016)
16. Explain 3-tier data warehouse architecture. (2016)
17. Explain the security requirements in respect of data warehouse. Discuss the impacts of security on design and performance? Also explain the backup strategies and disaster recovery. (2016)
18. Suppose that a data warehouse consists of the 4-D , data, spectator, location, and game, and two measures, count and charge; where charge is the fare that a spectators may be students, adults, or senior, with each category having its own charge rate.
  - a. Draw a star schema diagram for the data warehouse.
  - b. Starting with the base cuboid [data, spectator, location, and game] what specific OLAP operations should one perform in order to list the total charge paid by student spectators at a particular place in 2015? (2016)
19. What is the purpose of load manager? (2015)
20. What is aggregation? (2015)
21. What is slice operation of OLTP? (2015)
22. What is snow flake schema? Explain it. (2015)
23. What is noisy data? Explain various data cleaning approaches. (2015)
24. Explain multidimensional schema with an example. (2015)
25. Explain horizontal and vertical partitioning. (2015)
26. Explain how we can tune the data warehouse? (2015)
27. Explain importance of data mining with business, technical and social context? (2015)
28. Explain backup strategies and disaster recovery techniques for data ware house. (2015)
29. Define fact. What is fact table? (2014,2013)
30. What do you understand by dimension? (2014)
31. Define data integrity. (2014)
32. What is the role of partition indexes? (2014)
33. Discuss in details the concept and application areas of data warehouse. (2014)
34. Write the details note on data warehousing architecture with explanation of staging area process and data marts. (2014)
35. What is the role of data warehouse administrator? Hence discuss the making of a data warehouse and list the other roles associated to it. (2014)
36. What do you understand by Business intelligence? (2013)

37. What is data flow diagram? (2013)
38. List the key benefits of data warehouse. (2013)
39. What is directional table? (2013)
40. Differentiate between operational data and data warehouse? (2013)
41. Discuss the role of DSS analyst. (2013)
42. Write a note on extract and load process. (2013)
43. Discuss the importance of security in data warehouse operations. (2013)
44. List and details the key milestones of a data warehouse project. What is the ROI of data warehouse project? (2013)
45. Discuss the designing approaches to data warehouse evolution. How does it helps in upgrading business decision? (2013)
46. Discuss the key abilities and responsibilities of data warehouse project manager. (2013)
47. Discuss the statement “close co-operation of the operational team and data warehouse team and data warehouse team is essential for success of data warehouse project”. (2013)

## Unit 2

---

1. What is data mining? (2017,2014,2013)
2. Name any three data mining technique. (2017,2014)
3. Define data cleaning. (2017)
4. Write short note on dimensionality reduction. (2017)
5. Explain the concept of data association and data generalization in detail alongb with analysis of attribute relevance. (2017)
6. Explain following:
  - a. Association rule mining
  - b. Statical measures in large database (2017)
7. Write and explain the concept of classification and prediction along with various issues incorporated with them. (2017)
8. Explain various category of clustering method. (2017)
9. Explain multidimensional data model in detail. (2017)
10. Explain density based method in cluster analysis. (2017)
11. What is fact constellation? (2016)
12. Define metadata. (2017,2016,2015,2014,2013)
13. Define decision tree. (2017,2016,2014)
14. What is gain ratio? (2016)
15. What is tree purning? (2016)
16. What is clustering? (2016)
17. What is graph mining? (2016)
18. What do you understand by data mart? Explain it. (2016)
19. What are data mining approaches? Name them. (2016)
20. What are the difference between the three main types of data warehouse usage; information processing analytical processing, and data mining? Discuss the motivational behind OLAP mining (OLAM). (2016)
21. What do you understand by cluster analysis? Define major clustering methods. Also explain the automatic cluster detection technique. (2016)
22. Discuss the applications of data mining in technical and social context. Explain the methodologies used for data mining. (2016)
23. What is summary table? (2015)
24. What do you mean by data extraction? (2015)
25. What types of costs are involved in data mining? (2015)

26. What is the difference between data warehouse and database? (2015)
27. What is staging? (2015)
28. What is ETL? Explain it. (2015)
29. What is hardware partitioning? (2015)
30. What is decision tree? Explain with example. (2015)
31. Explain the complexities in transformation of data . (2015)
32. Explain aggregation summary tables. (2015)
33. What do you mean by capacity planning? Explain it. (2015)
34. What do you mean by testing backup recovery? (2015)
35. Explain automatic cluster detection technique. (2015)
36. Compare and contrast the DLTP vs data warehouse approach. (2014)
37. Discuss the steps in extracted and load process. (2014)
38. Discuss the key activities associated with following system processes.
  - a. Extract and load process
  - b. Clean and transform process (2014)
39. What do you understand by data mining? Discuss the data mining approaches and methodologies. (2014)
40. What do you understand by dimensional modeling? Identify the business dimensions associated with an online bank. Create an illustrative star schema to demonstrate banking business facts and dimensions. (2014)
41. Define aggregation. (2013)
42. What is automatic cluster detection? (2013)
43. What are the key difference between operational data and data warehouse? (2013)
44. Write details notes on any three of following:-
  - a. Multi dimensional query management
  - b. Data transformation and grooming
  - c. Testing data warehouse
  - d. Business technical and social context of data mining
  - e. Data mining technique (2013)