**Subject- DWM TE SEM V(Academic year 2022-23)**

**(UNIVERSITY QUESTIONS ON UNIT 1, UNIT 2 ,UNIT 4)**

**Unit 1**

Q1 .Write short note on: Operational support system.

Q2. Illustrate the architecture of a typical DW system. **OR**

Explain Data Warehouse Architecture in detail

Q3. Differentiate Data Warehouse vs. Data Mart

Q4. Differentiate top-down and bottom-up approaches for building data

warehouse. Discuss the merits and limitations of each approach.

Q5. What is meant by Metadata in the context of Data warehouse? Explain the different types of metadata stored in a data warehouse. Illustrate with a suitable example. **OR** write short note on Metadata in Data warehouse

Q6. What is metadata?why do we need metadata when search engines like google seem so effective?

Q7. Write short note on: Factless Fact Table and Fact constellation.

Q8. Design the data warehouse for wholesale future Company. The data warehouse has to allow analyzing the company’s situation at least with respect to the Furniture, Customer and Time. More ever, the company needs to analyse. The furniture with respect to its type, category and material. The customers with respect to their spatial location, by considering at least cities, regions and states. The company is interested in learning the quantity, income and discount of its sales.

Q9. Describe the steps of ETL process. **OR** Discuss the process of Extraction, Transformation and Loading with a neat and labeled diagram.

Q10. Discuss various OLAP models and their architecture.

Q11. **All Design Star Schema case study problems**

**Unit 2**

Q1. Describe the steps involved in data mining when viewed as a process of knowledge Discovery

Q2. Explain types of attributes & Data visualization for data exploration.

Q3. Discuss different steps involved in data preprocessing.

Q4. Numerical on- i) Binning Method ii) 5 steps summary(mean, mode, medium, boxplot, histogram)

**Unit 4**

Q1. Explain K- means clustering algorithm? Apply K-means algorithms for the following data set with two clusters.

Data set={1,2,6,7,8,8,10,15,17,20}

Q2. Numerical on K-means