

**Reg. No. :**

--	--	--	--	--	--	--	--	--	--

**Question Paper Code: 31255**

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2016

Fifth Semester

Computer Science and Engineering

01UCS505- DATA WAREHOUSING AND DATA MINING

(Regulation 2013)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions.

PART A - (10 x 2 = 20 Marks)

1. Define Data mart.
2. List the components of star schema.
3. Summarize the categories of reporting tools.
4. Give the formula for mean, median and mode.
5. List out the major tasks involved in Data preprocessing.
6. Differentiate operational database and informational database.
7. Distinguish Eager learners and Lazy learners.
8. Why tree pruning useful in decision tree induction?
9. Identify the advantages of STING over other clustering methods.
10. What is outlier analysis?

PART - B (5 x 16 = 80 Marks)

11. (a) Describe the data warehouse architecture with a neat diagram. (16)

Or

(b) Discuss two vendor specific Dynamic transformation tools in detail. (16)

12. (a) Describe the various Reporting and Query tools and their application. (16)

Or

(b) Explain reporting using Impromptu tool in detail. (16)

13. (a) Describe the architecture of typical data mining system with neat sketch. (16)

Or

(b) Describe about the various Data reduction techniques in detail. (16)

14. (a) What is classification? Explain the process of classification in Naïve-Bayesian classifier with a suitable example. (16)

Or

(b) Infer the process of classification using support vector machines. (16)

15. (a) Explain the K-means clustering with an example. (16)

Or

(b) Illustrate the strategies involved in finding various outliers. (16)