

Code: AT19

Subject: DATA WAREHOUSING AND DATA M

**AMIETE – IT (OLD SCHEME)**

Time: 3 Hours

**DECEMBER 2011**

Max. Marks: 100

**NOTE: There are 9 Questions in all.**

- Please write your Roll No. at the space provided on each page immediately after receiving the Question Paper.
- Question 1 is compulsory and carries 20 marks. Answer to Q.1 must be written in the space provided for it in the answer book supplied and nowhere else.
- The answer sheet for the Q.1 will be collected by the invigilator after 45 Minutes of the commencement of the examination.
- Out of the remaining EIGHT Questions answer any FIVE Questions. Each question carries 16 marks.
- Any required data not explicitly given, may be suitably assumed and stated.

**Q.1 Choose the correct or the best alternative in the following: (2×10)**

a. Data Mining includes

- (A) Analyzing large volumes of data to discover interesting associations or patterns
- (B) Querying a large data warehouse to uncover undiscovered facts
- (C) Very complex SQL query operations
- (D) Slicing and dicing until you uncover interesting details

b. Which schema is not used in data warehousing?

- (A) Star
- (B) Fact constellation
- (C) Snowflake
- (D) Hybrid schema

c. Find odd one out

- (A) ROLAP
- (B) TOLAP
- (C) MOLAP
- (D) HOLAP

d. Multiple data sources are combined in

- (A) Data cleaning
- (B) Data presentation
- (C) Data transforming
- (D) Knowledge representation

e. SKAT stands for:

- (A) Systematic Knot adding Technology
- (B) Symbolic Knowledge Acquisition Technology
- (C) Systematic Knowledge Acquisition Technology
- (D) System Knowledge and Technology

Code: AT19

Subject: DATA WAREHOUSING AND DATA M

- f. Which one of the following is a technique for data smoothing usually applied for data cleaning and sometimes for data discretization?
- (A) Histogram analysis. (B) Segmentation.  
(C) Binning. (D) None of the above
- g. EIS means
- (A) External Information System (B) External Internal System  
(C) Executive Internal System (D) Executive Information System
- h. OLTP stands for
- (A) online transaction processing systems  
(B) offline transaction processing systems  
(C) online transaction systems  
(D) online table processing
- i. During decision tree induction, pre-pruning leads to
- (A) construction of full grown trees (B) small and accurate trees  
(C) small and in-accurate trees (D) oversimplified and accurate trees
- j. Query and reporting tools are most appropriate for
- (A) Controlled predictable query environments  
(B) Ad hoc reporting requirements  
(C) Complex multifaceted business query applications  
(D) Discovery mode applications

---

**Answer any FIVE Questions out of EIGHT Questions.  
Each question carries 16 marks.**

---

- Q.2** a. What is a data warehouse? Explain the characteristics of a data warehouse? (8)  
b. Elaborate the major operating components, which are regularly monitored in data warehouse environment. (8)
- Q.3** a. Briefly explain Metadata ? List the items the Metadata store tracks. (8)  
b. Explain different levels of data modeling? Explain midlevel data modeling.(8)
- Q.4** a. Define data Integration and data Reduction. List strategies for data reduction. (8)  
b. List the challenges in the naturally evolving architecture. Describe any one of them. (8)

Code: AT19

Subject: DATA WAREHOUSING AND DATA M

- Q.5** a. Why does every structure in the data warehouse contain the time element? Explain (5)
- b. Distinguish between data warehouses and data marts. (5)
- c. Enumerate the building blocks of a data warehouse. (6)
- Q.6** a. Explain role of a EIS analyst. How does a data warehouse assist EIS analyst? (8)
- b. Explain the knowledge discovery process. (8)
- Q.7** a. What is external data? Why should it be compared to internal data over a period of time? Explain how this comparison is done. (8)
- b. What is the criterion for classification of Association rules? Explain. (8)
- Q.8** a. Discuss two distinguishing characteristics of a data driven methodology. (8)
- b. Explain the migration plan in building the data warehouse? (8)
- Q.9** Write a short note on (Any **FOUR**):
- (i) Data Cleaning
  - (ii) Event Mapping
  - (iii) Concept Hierarchy
  - (iv) Feedback loop
  - (v) Drill Down Analysis (4×4)