Cairo University	Cairo University – Institute Of Statistical Studies And Researches	
	Department: Information Systems & Technology	
	Academic Year: 2016 Semester: Second	
	Date: 29/5 Level: Master	

Course Title: Data Warehousing	Course code: IT608	Time: 3 Hours	Exam marks: 75	# Exam. Sheets: 5 pages

Exam. Instructions : Answer all Questions

Question One: True or False (15 Marks)

a. Creating Data warehouse involves only extracting data form operational environment and entering it into data warehouse.

b. The CLDS or spiral methodology can be used in the development of data warehouse.

c. The larger and more mature the organization the worse the problems of naturally evolving spider web architecture become.

d. Partitioning data warehouse data at application level is much easier.

e. Most of the DSS analytical processing occurs at the departmental level because it is easier and cheaper to get the data needed for processing there.

f. Data warehouse is constructed in a heuristic manner, where one phase of development depends entirely on the results attained in the previous phase.

g. Data warehouse contains corporate data that can be used for many different purposes, including sitting and waiting for future requirements which are unknown today.

h. Loading data on an ongoing basis presents the largest challenge to the data architect.

i. The process model applies to both the operational environment and the data warehouse environment.

j. Many development tools, such as CASE tools, can be applied to the data warehouse environment.

uestion Two : Choose (30 Marks)

a. The data warehouse requires an architecture that begins by looking at the …….. and then works down to the ……..

a) Whole, particulars.
b) Particulars, Whole.
c) System, Design

b. The …….. of data holds application-oriented primitive data only and primarily serves the TP community

a) operational level
b) data warehouse level
c) data mart level

c. Once the data warehouse built, two components must be monitored the …….. and……

a) data, activity
b) Growth, response time
c) Granularity, Partitioning

d. Operational databases contain ……..data

a) current-value

b) Archived

c) Summarized

e. The development of the data warehouse operates under ……..life cycle.

a) CLDS

b) DSLC

c) SDLC

f. Data in warehouse is stored in forms of……

a) Fact tables

b) Dimension tables

c) None of the above

g. Data mart is found at the level of:

a) Lightly summarized data

b) Highly summarized data

c) None of the above

h. ETL stands for:

a) Execute, Transport, Load

b) Extract, Transfer, Load

c) Exclude, Transact, Load

i. Data warehouse is:

a) Subject oriented

b) Application oriented

c) Transaction oriented

j. The older data in the data warehouse resides on:

a) Direct Access Storage Devices

b) Magnetic Tape

c) Flash memories

k. The key structure of the data warehouse must contain:

a) Current level data

b) Element of time

c) Primary data

l. ..... refers to the level of detail or summarization of the units of data in the data warehouse.

a) Partitioning

b) Granularity

c) Homogeneity

m. Elements that affect input output operations when accessing data warehouse

a) Normalization of data

b) Updating data

c) Distribution of data

n. The meta data of data warehouse contains the following

a) The structure of data as known by manager

b) The structure of data as known by DSS analyst

c) The structure of data of the operational environment

p. When we use data in data warehouse indirectly we have to consider the following

a) Formatting of data after it is retrieved

b) Analysis program and periodic refreshment

c) The compatibility of technologies between data warehouse environment and operational environment

Question Three : (30 Marks)

a. What is a data warehousing? And why we need to build it?

b. List the differences between primitive data and derived data?

c. In your opinion does data warehouse need monitoring? And why?

d. how does the level of granularity affect the design of data warehouse?

e. "It is always been argued that archival data exists only in data warehouse" are you agree with this point of view? Why?

f. Compare between operational environment and data warehousing environment concerning response time?

g. What is a spider web? Explain the problem associated with it?

h. What is the scenario for direct operational access of data warehouse?

i. Why integration of applications is not considered a simple task?

j. what are the techniques used to limit the amount of operational data scanned since the last refresh?

k. Compare between data model and process model concerning definition and usage.

l. What happened when different iterations for data warehouse development are done without a unifying data model?

m. What are the basic components of snapshot? and when the snapshot is created?

n. Why it is so important to have a metadata for data warehouse? explain its advantages?

o. Define Data Cyclicity? When it should be made and why?