Q1: Choose the right answers: (20 Marks)

1. The types of relationship in star schema is .....

A. Many to many

B. one to one

C. one to many

D. many to one

2. A homogenous distributed database is which of the following?

A. The same DBMS is used at each location and data are not distributed across all nodes

B. The same DBMS is used at each location and data are distributed across all nodes

C. A different DBMS is used at each location and data are not distributed across all nodes

D. A different DBMS is used at each location and data are distributed across all nodes

3. K-means is an example of

A. Association rule

B. Clustering

C. Regression

D. Classification

4. A .....technique is to reduce the redundancies in data representation in order to decrease data storage requirements

A. Data Removal

B. Data Mining

C. Data Cleaning

D. Data Compression

5. Select non predictive data mining technique from below options

A. Summarization

B. Classification

C. Regression

D. Time Series Analysis

6. ..... describes a distribution model in which applications are hosted by a service provider and made available to user

A. Cloud service

B. PaaS

C. IaaS

D. SaaS

7. Which statement is true about web server?

A. A web server is a standard interface between Common Gateway Interface (CGI) and application server

B. A Web server can serve as a front end to a variety of information services.

C. The application server is working in conjunction with the web server, where application server displays and the web server interact.

D. All of the above

8. Language Bindings are

A. Object manipulation language

B. Mechanisms to invoke OQL from language

C. Procedures for operation on databases and transactions

D. All of above

9. Semi structured data is simply any information that uses a self-describing schema such as XML or JSON

A. True

B. False

10. Noise is

A. Component of a network

B. In the context of KDD and data mining, this refers to random errors in a database table

C. One of the defining aspects of a data warehouse

D. None of these

Q2 : Answer five of the following (Just five): what are the differences between? (20 Marks)

1. HTML & XML file
2. OODB & ORDB
3. Homogeneous & Heterogeneous
4. ROLAP & MOLAP
5. Predictive & Descriptive
6. Hard Clustering & Soft Clustering
7. Persistent objects & transient objects in OODB

Q3 : Data Warehouse (15 Marks)

Suppose that a data warehouse consists of the three dimensions *time*, *doctor*, and *patient*, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit.

A. Enumerate three classes of schemas that are popularly used for modeling data warehouses.

B. Draw a schema diagram for the above data warehouse using one of the schema classes listed in 1.

C. Starting with the base cuboid [day, doctor, patient], what specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2020?

Q4 : Data - Web Mining & Distributed Database (15 Marks)

A. What is descriptive and predictive data mining?

B. Given the training set in following table , build classification model and apply it to the test set following

Weekend	Weather	Parents	Money	Decision
W_1	Sunny	Yes	Rich	Cinema
W_2	Sunny	No	Rich	Tennis
W_3	Windy	Yes	Rich	Cinema
W_4	Rainy	Yes	Poor	Cinema
W_5	Rainy	No	Rich	Stay in
W_6	Rainy	Yes	Poor	Cinema
W_7	Windy	No	Poor	Cinema
W_8	Windy	No	Rich	Shopping
W_9	Windy	Yes	Rich	Cinema
W_{10}	Sunny	No	Rich	Tennis

C. Explain at least four differences between how transactions are handled in a distributed database vs. a centralized database

Q5 : XML Database (15 Marks)

A. What is parser in XML used for? And what are the common APIs for XML parsers?

B. Write an XML document which match the following XSD schema

```
<xs:element name="shipto">
 <xs:complexType>
  <xs:sequence>
   <xs:element name="name" type="xs:string"/>
   <xs:element name="address" type="xs:string"/>
   <xs:element name="city" type="xs:string"/>
   <xs:element name="country" type="xs:string"/>
  </xs:sequence>
 </xs:complexType>
</xs:element>
```

C. Write an XPath expression that returns root element

Q6 : Query processing & Web based Database (15 Marks)

A. What is the difference between a Web server and an application server? What functionality do typical application servers provide?

B. State the steps of query processing and indicate the purpose of each step.

C. Translate the following query to relational algebra and indicate which the best equivalent for best cost?

```
Select salary
From employee
Where salary < 1000
```

Good Luck