

# Midterm Exam Part I | Multiple Choice Questions

**Due Mar 12 at 11:59pm**

**Points 40**

**Questions 40**

**Available Jan 10 at 12am - May 9 at 11:59pm**

**Time Limit 60 Minutes**

## Instructions

Part I of the midterm exam is a closed book, closed note, and closed Google.

Part I is timed at 60 minutes or 1.5 minutes per multiple choice question. Do not exceed your time allotment without a detailed explanation.

Part I contains 40 multiple choice questions worth 1 point each. The questions are taken from Chapter 13 of the DB text (9 questions), Chapter 14 of the DB text (5 questions), Big Data PDF (5 questions), OLAP PDF (6 questions), Data Profiling & ODS PDFs (5 questions), MS Excel Manipulating a Pivot Table (3 questions), and Data Quality-Integrity-Ethics PDFs (7 questions).

Choose the response that best answers the question posed. Be sure to read all of the answers, as each question only has one best answer.

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	59 minutes	28 out of 40

!  Correct answers will be available Mar 12 at 11:59pm - May 9 at 12am.

Score for this quiz: **28** out of 40

Submitted Mar 11 at 6:28pm

This attempt took 59 minutes.

### Question 1

1 / 1 pts

With respect to this feature in a star schema, it can also be computed or derived at run time:

attributes

characteristics facts dimensions

Incorrect

**Question 2**

0 / 1 pts

Data marts are:

 any of these responses small, single-subject data warehouse subset created from data extracted from a larger data warehouse created for supporting faster data access

Incorrect

**Question 3**

0 / 1 pts

Which of the following are not a characteristic of high-quality information?

 Unique Consistent Complete Timely Accessible

**Question 4****1 / 1 pts**

Numeric measurements (values) that represent a specific business aspect or activity in a star schema are considered:

- dimensions
- characteristics
- attributes
- facts

**Question 5****1 / 1 pts**

Decision support data is typically housed in a data warehouse, which does not have the following characteristic:

- volatile
- subject-oriented
- integrated
- time-variant

**Question 6****1 / 1 pts**

Which of the following is not consider a business intelligence tool:

- OLAP

- data warehouse
- dashboards
- portals
- data-mining
- data visualization
- spreadsheet

**Question 7**

1 / 1 pts

Quantitative data describes:

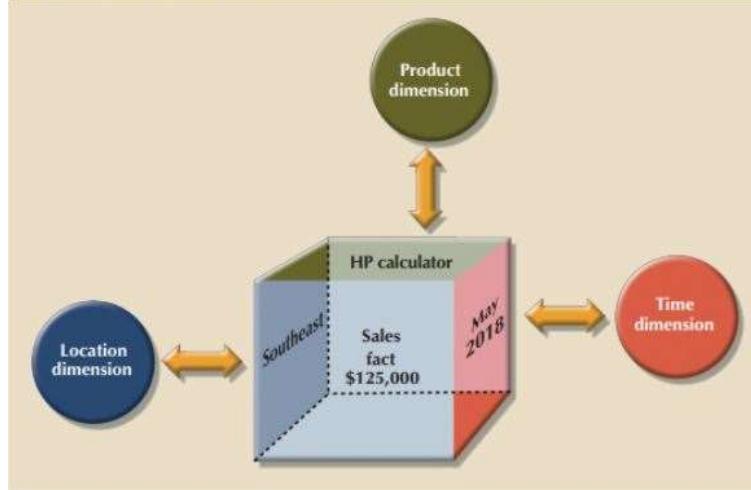
- ordinal data
- numeric facts or measures of data
- qualities of data
- nominal data

Incorrect

**Question 8**

0 / 1 pts

Given Figure 13.5, a month (e.g., January, February, ..., December) would be considered as a(an):

**FIGURE 13.5 SIMPLE STAR SCHEMA**

- characteristic
- dimension
- fact
- attribute

Incorrect

**Question 9**

0 / 1 pts

Qualifying characteristics that provide additional perspectives in a star schema.

- dimensions
- facts
- characteristics
- attributes

**Question 10****1 / 1 pts**

The six primary phases of the data-mining process does not include:

- data understanding
- deployment
- data preparation
- business understanding
- data testing
- data modeling
- evaluation

**Incorrect****Question 11****0 / 1 pts**

The following presents data, trends, and relationships in a easily understandable and eye appealing manner:

- infographic
- pie chart
- bar graph
- data visualization

**Question 12****1 / 1 pts**

Using data visualization, businesses can benefit by:

- highlighting all forms of data
- recognizing data patterns
- insight into marketing behavior
- redundant information processing

**Question 13****1 / 1 pts**

Which of the following is not an example of unstructured data?

- voice mail data
- satellite images
- social media data
- video data
- photographic data
- sensor data

**Incorrect****Question 14****0 / 1 pts**

Which of the following is not an example of structured data?

- point-of-sale data

- financial data
- accounting data
- click-stream data
- text messages
- Weblog data

**Question 15**

1 / 1 pts

With respect to OLAP cube operations, this operation takes a section out of the cube to analyze a specific category:

- slice
- dice
- drill-down
- roll-up
- pivot

**Question 16**

1 / 1 pts

With respect to OLAP cube operations, this operation makes dimensions more specific, making more instances.

- drill-down
- roll-up

- dice
- pivot
- slice

**Question 17**

1 / 1 pts

Uses relational databases and familiar relational query tools to store and analyze multi-dimensional data.

- ROLAP
- OLAP
- HOLAP
- MOLAP

**Question 18**

1 / 1 pts

In online analytical processing, this object represents multiple perspectives (i.e., multidimensionality) to data via dimensions.

- Pivot
- Drill-down
- Roll-up
- Cube
- Pivot

**Question 19****1 / 1 pts**

With respect to OLAP cubes, this operation switches the categories between two dimensions:

- drill-down
- roll-up
- slice
- pivot
- dice

**Question 20****1 / 1 pts**

With respect to OLAP cube operations, this operation makes a subcube with the same dimension but more specific instances.

- pivot
- dice
- drill-down
- slice
- roll-up

**Question 21****1 / 1 pts**

A characteristic of Big Data that describes the quantity of data to be stored is:

- Variety
- Volume
- Veracity
- Velocity

**Question 22****1 / 1 pts**

A characteristic of Big Data that describes the variations in the structure of data to be stored is:

- Velocity
- Veracity
- Variety
- Volume

**Question 23****1 / 1 pts**

A characteristic of Big Data that describes the speed at which data enters the system and must be processed is:

- Volume

Veracity Velocity Variety**Question 24**

1 / 1 pts

A new generation of database management systems that is not based on the traditional RDBMS is:

 Hadoop Cassandra NewSQL NoSQL**Question 25**

1 / 1 pts

A database model that attempts to provide ACID-compliant transactions across a highly distributed infrastructure is:

 Cassandra Hadoop NewSQL NoSQL

Incorrect

**Question 26****0 / 1 pts**

Which of the following characteristics **does not** describe an operational data store (ODS)?

- detailed-only
- current-valued
- non-volatile
- integrated
- subject-oriented

**Question 27****1 / 1 pts**

Which of the following **is not** an application of data profiling statistics?

- assess risk in integrating data for new applications, including the challenges of joins
- understanding data challenges early to avoid late project surprises, time delays, and cost overruns
- data quality metrics—does data meet standards?
- Insights extracted from data profiling can determine how easy or difficult it will be to use existing data for other purposes along with providing metrics on data quality.

- assess if metadata accurately describes actual data
- six sigma methodologies applied to enterprise data
- none of these responses

**Question 28**

1 / 1 pts

A process where available data in columns of an existing database table are examined, data statistics are collected, and insights extracted determine how easy or difficult it will be to use existing data for other purposes.

- data integrity
- data mining
- data quality
- data profiling

**Question 29**

1 / 1 pts

Which of the following is not accurate about an operational data store (ODS).

- contents of the ODS are consistently updated through the course of business operations



designed to quickly perform relatively simple queries on small amounts of data



designed for complex queries on large amounts of data



is a type of database often used as an interim area for a data warehouse

Incorrect

**Question 30**

0 / 1 pts

Which of the following is not a data profiling metric?

 domain pattern or standard format standard deviation (quantitative) median value (middle) dependencies modal value (mode or most frequent) mean value (if quantitative) attribute name variance data type minimum / maximum value number of records / number of null records

**Question 31****1 / 1 pts**

In a Microsoft Excel pivot table, which of the following is used to display fields on the side?

- report filter
- row labels
- column labels
- values

**Question 32****1 / 1 pts**

In a Microsoft Excel pivot table, which of the following is used to display summary numeric data?

- report filter
- values
- row labels
- column labels

**Incorrect****Question 33****0 / 1 pts**

In a Microsoft Excel pivot table, which of the following is used to display fields at the top ?

report filter column labels values row labels**Question 34**

1 / 1 pts

Government regulations on data collection primarily are concerned with which of the following?

 money laundering privacy and personal identifiable information privacy personal identifiable information

Incorrect

**Question 35**

0 / 1 pts

Lee & Strong (2003) empirically identified which of the following as an important prerequisite for producing high-quality data?

 understanding knowledge wisdom information

**Question 36****1 / 1 pts**

To ensure data integrity reflects the dynamic, global nature of real-world states, organizations need which of the following?

- referential integrity rules in the database
- user-defined integrity rules in the database
- data integrity rules in the database
- data integrity as an embedded process

**Question 37****1 / 1 pts**

Outliers are not data quality issues and outliers are determined as:

- outside  $1.5 \times$  the interquartile range (IQR) of data elements
- below the 25th percentile of data elements
- above the 75th percentile of the data elements
- inside  $1.5 \times$  the interquartile range (IQR) of data elements

**Incorrect****Question 38****0 / 1 pts**

Which of the following statements is most correct?

- Data integrity issues will lead to low data quality.
  - any of these responses
  - Data integrity may not correct poor data quality.
  - Poor data integrity and/or poor data quality yield flawed information.
- Poor data integrity and/or poor data quality data, which when used in decision making, yield flawed decisions.

Incorrect

**Question 39**

0 / 1 pts

Which of the following explains what is good and bad with respect to duty and obligation?

- legal authority
- privacy
- morals
- ethics

**Question 40**

1 / 1 pts

Which of the following refers to systemizing, defending, and recommending concepts of right and wrong conduct in relation to data, particularly personal data.

- information ethics

- privacy ethics
- data ethics
- personal identifiable information ethics

Quiz Score: **28** out of 40