

# Quiz - Chapter 4

**Due** Feb 16 at 11:59pm

**Points** 100

**Questions** 21

**Available** Feb 4 at 12am - Feb 16 at 11:59pm 13 days

**Time Limit** 42 Minutes

This quiz was locked Feb 16 at 11:59pm.

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	20 minutes	100 out of 100

Score for this quiz: **100** out of 100

Submitted Feb 16 at 9:39pm

This attempt took 20 minutes.

### Question 1

4 / 4 pts

Converting continuous valued numerical variables to ranges and categories is referred to as discretization.

Correct!

☒ True

☐ False

### Question 2

4 / 4 pts

During classification in data mining, a false positive is an occurrence classified as true by the algorithm while being false in reality.

Correct!

☒ True

☐ False

### Question 3

4 / 4 pts

Using data mining on data about imports and exports can help to detect tax avoidance and money laundering.

Correct!

☒ True

☐ False

### Question 4

4 / 4 pts

Open-source data mining tools include applications such as IBM SPSS Modeler and Dell Statistica.

Correct!

☐ True

☒ False

### Question 5

4 / 4 pts

In the opening case, police detectives used data mining to identify possible new areas of inquiry.

Correct!

☐ True

☒ False

### Question 6

5 / 5 pts

Clustering partitions a collection of things into segments whose members share

Correct!

- ☒ similar characteristics.
- ☐ dissimilar characteristics.
- ☐ similar collection methods.
- ☐ dissimilar collection methods.

### Question 7

5 / 5 pts

What is the main reason parallel processing is sometimes used for data mining?

Correct!

- ☐ because the hardware exists in most organizations, and it is available to use
- ☐ because most of the algorithms used for data mining require it
- ☒ because of the massive data amounts and search efforts involved
- ☐ because any strategic application requires parallel processing

### Question 8

5 / 5 pts

What does the robustness of a data mining method refer to?

Correct!

☐

its ability to predict the outcome of a previously unknown data set accurately

☐

its speed of computation and computational costs in using the mode

☐

its ability to construct a prediction model efficiently given a large amount of data

☒

its ability to overcome noisy data to make somewhat accurate predictions

### Question 9

5 / 5 pts

What does the scalability of a data mining method refer to?

☐

its ability to predict the outcome of a previously unknown data set accurately

☐

its speed of computation and computational costs in using the mode

☒

its ability to construct a prediction model efficiently given a large amount of data

☐

its ability to overcome noisy data to make somewhat accurate predictions

Correct!

### Question 10

5 / 5 pts

In the Target case study, why did Target send a teen maternity ads?

**Correct!**

☐

Target's analytic model confused her with an older woman with a similar name.

☐

Target was sending ads to all women in a particular neighborhood.

☒

Target's analytic model suggested she was pregnant based on her buying habits.

☐

Target was using a special promotion that targeted all teens in her geographical area.

### Question 11

5 / 5 pts

Understanding customers better has helped Amazon and others become more successful. The understanding comes primarily from

☐

collecting data about customers and transactions.

☐

developing a philosophy that is data analytics-centric.

☒

analyzing the vast data amounts routinely collected.

☐

asking the customers what they want.

**Correct!**

### Question 12

5 / 5 pts

Which broad area of data mining applications partitions a collection of objects into natural groupings with similar features?

☐

associations

Correct!

- ☐ visualization
- ☐ classification
- ☒ clustering

### Question 13

5 / 5 pts

In estimating the accuracy of data mining (or other) classification models, the true positive rate is

Correct!

- ☒ the ratio of correctly classified positives divided by the total positive count.
- ☐ the ratio of correctly classified negatives divided by the total negative count.
- ☐ the ratio of correctly classified positives divided by the sum of correctly classified positives and incorrectly classified positives.
- ☐ the ratio of correctly classified positives divided by the sum of correctly classified positives and incorrectly classified negatives.

### Question 14

5 / 5 pts

Data preparation, the third step in the CRISP-DM data mining process, is more commonly known as \_\_\_\_\_.

Correct!

data preprocessing

orrect Answers

data preprocessing

### Question 15

5 / 5 pts

In the Influence Health case, the company was able to evaluate over \_\_\_\_\_ million records in only two days.

Correct!

195

orrect Answers

195

one hundred ninety-five

### Question 16

5 / 5 pts

\_\_\_\_\_ was proposed in the mid-1990s by a European consortium of companies to serve as a nonproprietary standard methodology for data mining.

Correct!

CRISP-DM

orrect Answers

CRISP-DM

### Question 17

5 / 5 pts

In the Dell case study, engineers working closely with marketing, used lean software development strategies and numerous technologies to create a highly scalable, singular \_\_\_\_\_.

**Correct!**

data mart

**Correct Answers**

data mart

### Question 18

5 / 5 pts

In \_\_\_\_\_, a classification method, the complete data set is randomly split into mutually exclusive subsets of approximately equal size and tested multiple times on each left-out subset, using the others as a training set.

**Correct!**

k-fold cross-validation

**Correct Answers**

k-fold cross-validation

k-fold cross-validation

### Question 19

5 / 5 pts

The data mining in cancer research case study explains that data mining methods are capable of extracting patterns and \_\_\_\_\_ hidden deep in large and complex medical databases.

**Correct!**

relationships

**Correct Answers**

relationships

### Question 20

5 / 5 pts



Patterns have been manually \_\_\_\_\_ from data by humans for centuries, but the increasing volume of data in modern times has created a need for more automatic approaches.

**Correct!**

extracted

**Correct Answers**

extracted

### Question 21

5 / 5 pts

Data are often buried deep within very large \_\_\_\_\_, which sometimes contain data from several years.

**Correct!**

databases

**Correct Answers**

databases

Quiz Score: **100** out of 100