Question 1 (1 point) ✓ *Saved*

Total cost of ownership is never considered in an analytical model.

- ○ True
- ◉ False

Next Page

Page 1 of 22

Question 2 (3 points) ✓ *Saved*

What are the three ways to handle missing data values.

keep

delete

replace

Next Page

Page 2 of 22

Question 3 (1 point) ✓ *Saved*

Which of the following is not an advantage of using open source software for analytics

- ◉ It has been thoroughly engineered and extensively tested, validated and is completely documented
- ○ It can be used in combination with commercial software
- ○ It is available for free
- ○ A worldwide network of developers can work on it

Next Page

Page 3 of 22

Question 4 (1 point) ✓ *Saved*

What type anonymization would be grouping ages into 12-19, 20-30, 31-40, etc.

- ○ discretization
- ○ aggregation
- ◉ generalization
- ○ value distortion

Next Page

Page 4 of 22

Question 5 (1 point) ✓ *Saved*

Data is always normalized for analysis.

- ○ True
- ◉ False

Next Page

Page 5 of 22

Question 6 (1 point) ✓ *Saved*

Outliers can be treated as missing values.

- ◉ True
- ○ False

Next Page

Page 6 of 22

Question 7 (1 point) ✓ *Saved*

Which is not a critical success factor in analytic models?

- ○ Justifiability
- ○ Interpretability
- ● Stakeholder compliance
- ○ Operational efficiency

Next Page

Question 8 (1 point) ✓ *Saved*

Credit score is this type of risk analytics.

- ● risk
- ○ marketing
- ○ recommender
- ○ text

Next Page

Question 9 (1 point) ✓ *Saved*

The goal of linear regression is to find a 'best fit" line to represent the data.

- ● True
- ○ False

Next Page

Question 10 (1 point) ✓ *Saved*

This type of analytics states that data typically has a particular order.

Predictive	Sequence rules

Next Page

Question 11 (1 point) ✓ *Saved*

A typical data scientist is never required to program as part of their job.

- ○ True
- ● False

Next Page

Question 12 (4 points) ✓ *Saved*

What are the key differences between logistic regression and decision trees. Give examples of when you would prefer one over the other.

Paragraph | **B** | *I* | U | A | ≡ | ≡ | ▶ | 🔗 | + | ••• | ⛶

The key difference is that logistic regression uses a linear line for most of the data to predict. Decision tree splits up the data and find what's common. I would use a decision tree for predicting data for a pre-determined data like a person's income based on their age and other factor.

Next Page

Question 13 (1 point) *Saving...*

In social networks a node can be compared to an entity in the relational model, and an edge a relationship.

- (•) True
- ( ) False

Next Page

Question 14 (1 point) *Saved*

ROI only considers monetary gains.

- ( ) True
- (•) False

Next Page

Question 15 (1 point)

This is the process of taking a subset of the data to build a data model.

Sampling

Next Page

Question 16 (1 point) ✓ *Saved*

What type anonymization would be specifying city and state only, rather than a complete address?

- ○ Discretization
- ◉ Value distortion
- ○ Generalization
- ○ Aggregation

Next Page

Page 16 of 22

Question 17 (1 point)

This type on analytic groups items together.

Descriptive

Clustering

Next Page

Page 17 of 22

Question 18 (1 point) ✓ *Saved*

The same attribute must be considered as the same level of a decision tree.

- ◉ True
- ○ False

Next Page

Page 18 of 22

Question 19 (1 point) ✓ *Saved*

Data security ensures confidentiality, availability and integrity of data.

- (●) True
- ( ) False

Next Page

Question 20 (1 point) ✓ *Saved*

This type of anonymization involves adding a known value from a distribution to a value in order to hide the original value

- ( ) Aggregation
- ( ) Discretization
- (●) Value distortion
- ( ) Generalization

Next Page

Question 21 (1 point) ✓ *Saved*

The training set is used to test the analytic model.

- (●) True
- ( ) False

Next Page

Page 21 of 22

Question 22 (1 point)

This type of analytics finds relationships between data. Ex "when someone buys cereal they also buy milk"

Predictive

Association rules

Next Page

Page 22 of 22