**Question 1**

**5 / 5 pts**

Before diving into the data it is important to understand:

The desired decision of the stakeholders.

The origin of the raw material for the data.

How the data is stored and warehoused.

The context of the data related to the business.

**Question 2**

**5 / 5 pts**

Why do data analysists need to work closely with subject matter experts?

To determine changes in data status code.

To validate the assumptions upon which the analytics are based.

To reduce waste and increase effectiveness of the analytic processes.

To ensures the steps necessary for the analysis.

**Question 3**

**5 / 5 pts**

A major difference between quantitative and qualitative data is that quantitative data:

Is more difficult to statistically analyze.

is numeric and obtained from observations can be obtained from personal descriptions.

requires greater preparation before analysis.

**IncorrectQuestion 4**

**0 / 5 pts**

\_\_\_\_\_ data is used to document occurrences associated with adverse outcomes.

Variable.

Attribute.

Nonnumeric.

Count.

**Question 5**

**5 / 5 pts**

The major attributes of data stored in a field at the database level include all, except:

amount.

scale.

Size.

precision.

**IncorrectQuestion 6**

**0 / 5 pts**

Treatment of data according to a database attribute requires knowledge of the type of

integrity and management of the database.

field controls.

data precision and scale.

analyses must be performed.

**Question 7**

**5 / 5 pts**

Categorical data consist of values that are:

grouped in a random order.

mutually exclusive and do not follow a specific order

set with actual differences between integers.

dichotomous and orderly.

**Question 8**

**5 / 5 pts**

The Kelvin temperature scale represents \_\_\_\_\_\_ data.

categorical.

interval.

ratio.

ordinal.

**IncorrectQuestion 9**

**0 / 5 pts**

Frequency distributions can be used to display:

qualitative data.

clusters around a central value.

the spread of data.

a plot of quartile ranges.

**Question 10**

**5 / 5 pts**

The histogram is a distribution of data as it exists:

around the median.

over its range.

outside the mode.

Across its central tendency.

**Question 11**

**5 / 5 pts**

The most commonly used method to summarize data is the:

Mean.

range.

mode.

quartile.

**Question 12**

**5 / 5 pts**

Which of the follow statistical measures is most susceptible to outliers?

mode.

range.

Mean.

quartile.

**Question 13**

**5 / 5 pts**

Which of the following identified correlations between variables?

Scatter plots.

Histogram.

Frequency distribution.

Box-and-whisper plots (box plots).

**Question 14**

**5 / 5 pts**

Which of the following represents the core of data mining?

Columns contain multiple fields.

There are variations in data structure.

Cases are in rows and variables in columns.

Data are proportionally organized.

**Question 15**

**5 / 5 pts**

Database access can be inhibited due to each of the following, except:

Data mining may require format conversion.

Database is not organized to support data mining.

Unintentional monopolizing of business resources.

Legal and business obligations to privacy law.

**Question 16**

**5 / 5 pts**

A surveying the data should include at a minimum, a representation of:

outliers, mode, frequency, categorical variable values.

missing cases, minimum values, maximum values, outliers.

ratios, average, standard deviation, categorical variable values.

minimum values, maximum values, average, standard deviation.

**Question 17**

**5 / 5 pts**

Bar charts are an integral part of the data mining process because each bar represents:

visual summaries of continuous variables.

the mean value for each variable.

visual summaries of categorical data.

a range of values for each variable.

**Question 18**

**5 / 5 pts**

What is the initial step towards predictive value obtained from a scatterplot?

Calculation of mean and standard deviation.

Access of minimum and maximum values.

Creation of nominal structures for analytics.

Identification of relationships between variables.

**Question 19**

**5 / 5 pts**

The re-arrangement of data can best be done using the \_\_\_\_\_ control.

Drag and drop.

Selection.

Display.

Sort.

**Question 20**

**5 / 5 pts**

To merge datasets with different cases what analytic function must be used?

Append.

Count.

Identify.

Sort.