

Student name: _____

TRUE/FALSE - Write 'T' if the statement is true and 'F' if the statement is false.

A population is a set that includes all elements about which we wish to draw a conclusion.

true

false

If we examine some of the population measurements, we are conducting a census of the population.

true

false

A random sample is selected so that every element in the population has the same chance of being included in the sample.

true

false

An example of a quantitative variable is the manufacturer of a car.

true

false

An example of a qualitative variable is the mileage of a car.

true

false

Statistical inference is the science of using a sample of measurements to make generalizations about the important aspects of a population of measurements.

true

false

Time series data are data collected at the same time period.

true

false

Cross-sectional data are data collected at the same or approximately the same point in time.

true

false

Daily high temperature in a local community collected over a 30-day time period is an example of cross-sectional data.

true

false

The number of sick days taken by employees in 2008 for the top 10 technology companies is an example of time series data.

true

false

The number of sick days per month taken by employees for the last 10 years at Apex Co. is an example of time series data.

true

false

A quantitative variable can also be referred to as a categorical variable.

true

false

In a data set of information on college business students, an example of an element is their cumulative GPA.

true

false

In an observational study, the variable of interest is called a response variable.

true

false

In an experimental study, the aim is to manipulate or set the value of the response variable.

true

false

The science of describing the important aspects of a set of measures is called statistical inference.

true

false

It is possible to use a random sample from a population to make statistical inferences about the entire population.

true

false

Processes produce outputs over time.

true

false

Selecting many different samples and running many different tests can eventually produce a result that makes a desired conclusion be true.

true

false

Using a nonrandom sample procedure in order to support a desired conclusion is an example of an unethical statistical procedure.

true

false

An individual collecting data directly through planned experimentation is obtaining primary data.

true

false

Secondary data are data taken from an existing source.

true

false

Data warehousing is defined as a process of centralized data management and retrieval.

true

false

The term *big data* refers to the use of survey data by big business.

true

false

In order to select a stratified random sample, we divide the population into overlapping groups of similar elements.

true

false

If we sample without replacement, we do not place the unit chosen on a particular selection back into the population.

true

false

By taking a systematic sample in which we select every 100th shopper arriving at a specific store, we are approximating a random sample of shoppers.

true

false

A common practice in selecting a sample from a large geographic area is multistage cluster sampling.

true

false

Stratification can at times be combined with multistage cluster sampling to develop an appropriate sample.

true

false

In systematic sampling, the first element is randomly selected from the first (N/n) elements.

true

false

A sampling error can occur because of incomplete information.

true

false

The target population is the result of sampling from the original population that is of interest to the researcher.

true

false

Errors of non-observation occur when data values are recorded incorrectly.

true

false

A recording error is an error of observation.

true

false

A low response rate has no effect on the validity of a survey's findings.

true

false

Sampling error occurs because a characteristic of a random sample may not exactly equal the population characteristic that we are attempting to estimate.

true

false

Convenience sampling is a type of probability sampling in which we select elements to sample because we believe they have the highest probability of responding.

true

false

Judgment sampling is an example of convenience sampling.

true

false

Judgment sampling occurs when a person who is extremely knowledgeable about the population under consideration selects the population elements that they feel are most representative of the population.

true

false

Business analytics is a new field that does not use traditional statistics to analyze big data.

true

false

Prescriptive analytics involve methods used to find anomalies, patterns, and associations in data sets with the purpose of predicting future outcomes.

true

false

MULTIPLE CHOICE - Choose the one alternative that best completes the statement or answers the question.

A population that consists of all the customers who will use the drive-thru of the local fast food restaurant is called a(n) _____.

infinite population

random sample population

statistical population

finite population

In _____ we select elements because they are easy to sample.

random sampling

convenience sampling

judgment sampling

probability sampling

_____ sampling is where we know the chance that each element will be included in the sample, which allows us to make statistical inferences about the sample population.

Convenience

Voluntary

Probability

Judgment

Which of the following is not a method of numerical descriptive analytics?

factor analysis

cluster analysis

bullet graphs

association learning

_____ uses traditional or newer graphics to present visual summaries of business information.

Nonparametric predictive analytics

Parametric predictive analytics

Prescriptive analytics

Graphical descriptive analytics

Which of the following is not a supervised learning technique in predictive analytics?

linear regression

factor analysis

decision trees

neural networks

Transactional data are now used by businesses as part of
survey analysis.

big data.

descriptive statistics.

experimental studies.

_____ consists of a set of concepts and techniques that are used to describe populations and samples and to make statistical inferences about populations by using samples.

Traditional statistics

Random sampling

Data mining

Time series analysis

When we are choosing a random sample and we do not place chosen units back into the population, we are

sampling with replacement.

sampling without replacement.

using a systematic sample.

using a voluntary response sample.

Which of the following is a type of question used in survey research?

dichotomous

open-ended

multiple-choice

All of the other answers are correct.

Methods for obtaining a sample are called

sample surveys.

probability sampling.

random sampling.

sampling designs.

A _____ is a list of all the units in a population.

sample

frame

census

variable

Nonoverlapping groups of similar elements in a population are called
clusters.

frames.

strata.

stages.

A *Yes* or *No* question is _____.

dichotomous

evaluative

open-ended

systematic

_____ occurs when some population elements are excluded from the process of selecting the sample.

Nonresponse

Error of observation

Undercoverage

Sample frame

_____ is the difference between a numerical descriptor of the population and the corresponding descriptor of the sample.

Sampling error

Nonobservation error

Observation error

Nonresponse

Data that are collected by an individual through personally planned experimentation or observation are _____.

secondary data

quantitative data

primary data

variables

A ratio variable has the following characteristic.

qualitative

inherently defined zero value

categorical in nature

predictable

Which of the following is a quantitative variable?

the manufacturer of a cell phone

a person's gender

mileage of a car

whether a person is a college graduate

whether a person has a charge account

Which of the following is a categorical variable?

air temperature

bank account balance

daily sales in a store

whether a person has a traffic violation

value of company stock

Measurements from a population are called

elements.

observations.

variables.

processes.

The two types of quantitative variables are

ordinal and ratio.

interval and ordinal.

nominative and ordinal.

interval and ratio.

nominative and interval.

Temperature (in degrees Fahrenheit) is an example of a(n) _____ variable.

nominative

ordinal

interval

ratio

Jersey numbers of soccer players is an example of a(n) _____ variable.

nominative

ordinal

interval

ratio

The weight of a chemical compound used in an experiment that is obtained using a well-adjusted scale represents a(n) _____ level of measurement.

nominative

ordinal

interval

ratio

An identification of police officers by rank would represent a(n) _____ level of measurement.

nominative

ordinal

interval

ratio

_____ is a necessary component of a runs plot.

Observation over time

Qualitative variable

Random sampling of the data

Cross-sectional data

_____ is the science of using a sample to make generalizations about the important aspects of a population.

Time series analysis

Descriptive statistics

Random sample

Statistical inference

College entrance exam scores, such as SAT scores, are an example of a(n) _____ variable.

ordinal

ratio

nominative

interval

The number of miles a truck is driven before it is overhauled is an example of a(n) _____ variable.

nominative

ordinal

interval

ratio

A(n) _____ variable is a qualitative variable such that there is no meaningful ordering or ranking of the categories.

ratio

ordinal

nominative

interval

A person's telephone area code is an example of a(n) _____ variable.

nominative

ordinal

interval

ratio

Any characteristic of a population unit is a(n)

measurement.

sample.

observation.

variable.

Examining all population measurements is called _____.

census

frame

sample

variable

Any characteristic of an element is called a _____.

set

process

variable

census

The process of assigning a value of a variable to each element in a data set is called _____.

sampling

measurement

experimental analysis

observational analysis

A _____ is a display of individual measurements versus time.

runs plot

statistical analysis

random sample

measurement

Statistical _____ refers to using a sample of measurements and making generalizations about the important aspects of a population.

sampling

process

analysis

inference

A _____ is a subset of the units in a population.

census

process

sample

variable

A _____ variable takes on values that are numbers on the real number line.

qualitative

quantitative

categorical

nominative

A sequence of operations that takes inputs and turns them into outputs is a _____.

process

statistical inference

runs plot

random sampling

A(n) _____ variable can have values that indicate into which of several categories of a population it belongs.

qualitative

quantitative

ratio

interval

A set of all elements we wish to study is called a _____.

sample

process

census

population

_____ refers to describing the important aspects of a set of measurements.

Cross-sectional analysis

Runs plot

Descriptive statistics

Time series analysis

The change in the daily price of a stock is what type of variable?

qualitative

ordinal

random

quantitative

Data collected for a particular study are referred to as a data _____.

variable

measurement

set

element

A data set provides information about some group of individual _____.

variables

elements

statistics

measurements

When the data being studied are gathered from a published source, this is referred to as a(n) _____.

existing data source

observational data source

experimental data source

cross-sectional data source

One method of being sure a sample being studied can be used to make statistical inferences about the population is to select a

judgment sample.

voluntary response sample.

convenience sample.

probability sample.

Which of the following is *not* an example of unethical statistical practices?

inappropriate interpretation of statistical results

using graphs to make statistical inferences

improper sampling

descriptive measures that mislead the user

None of the other answers is correct.

If we collect data on the number of wins each team in the NFL had during the 2011-12 season, we have _____ data.

cross-sectional

time series

non-historical

survey

If we collect data on the number of wins the Dallas Cowboys earned each of the past 10 years, we have _____ data.

cross-sectional

time series

non-historical

survey

In a dataset about employees, which of the following would be considered a variable?

the specific employee "John Smith"

the company's total number of employees

the average salary across all departments

each employee's annual salary

Which of the following is an example of a qualitative variable?

the weight of a package in grams

the temperature in degrees Celsius

the number of pages in a book

the brand of a car

A dataset containing the average income of different countries in 2023 would be an example of time series data.

cross-sectional data.

both time series and cross-sectional data.

neither time series nor cross-sectional data.

A time series plot is primarily used to

compare different categories at a single point in time.

visualize the distribution of a single variable.

show the relationship between two variables.

track changes in a variable over time.

A study where researchers randomly assign participants to either a treatment group or a control group is an example of

an observational study.

an experimental study.

an existing data source study.

a survey study.

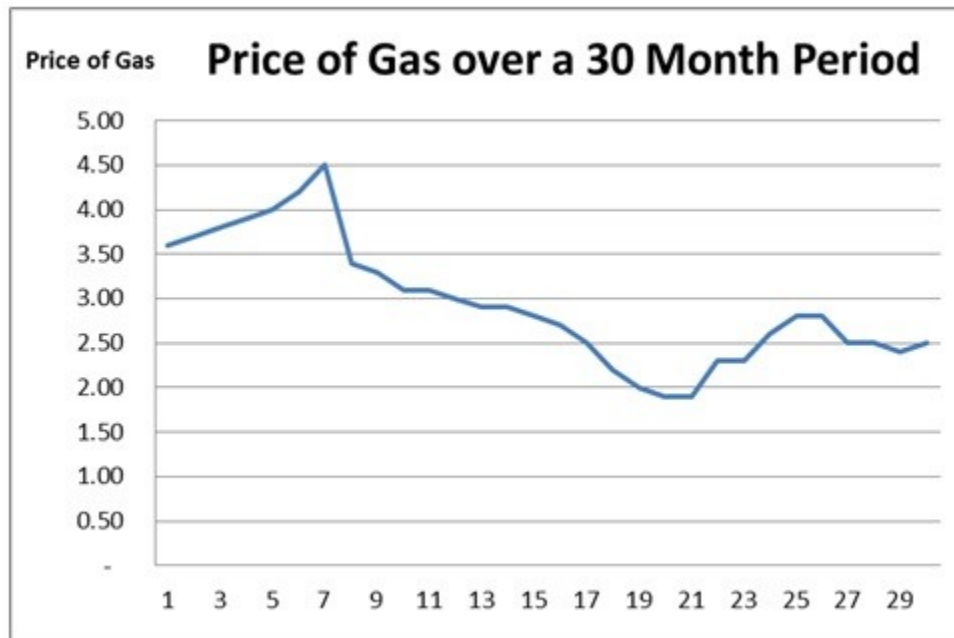
ESSAY. Write your answer in the space provided or on a separate sheet of paper.

A study is being conducted on the effect of gas price on the number of miles driven in a given month. Residents in two cities, one on the East Coast and one on the West Coast, are randomly selected and asked to complete a questionnaire on the type of car they drive, the number of miles they live from work, the number of children under 18 in their household, their monthly income, and the number of miles they have driven over the past 30 days. List the response variable(s).

A study is being conducted on the effect of gas price on the number of miles driven in a given month. Residents in two cities, one on the East Coast and one on the West Coast, are randomly selected and asked to complete a questionnaire on the type of car they drive, the number of miles they live from work, the number of children under 18 in their household, their monthly income, and the number of miles they have driven over the past 30 days. Is this an experimental or observational study?

A study is being conducted on the effect of gas price on the number of miles driven in a given month. Residents in two cities, one on the East Coast and one on the West Coast, are randomly selected and asked to complete a questionnaire on the type of car they drive, the number of miles they live from work, the number of children under 18 in their household, their monthly income, and the number of miles they have driven over the past 30 days. List the factor(s).

Looking at the runs plot of gasoline prices over the past 30 months, describe what it tells us about the price of gas during these 30 months.



Using the following data table of the average hours per week spent on Internet activities by 15- to 18-year-olds for the years 1999 to 2008, construct the runs plot and interpret.

1999	13.5
2000	15.0
2001	16.5
2002	17.7
2003	18.2
2004	19.3
2005	19.5
2006	19.9
2007	20.1
2008	20.4

Answer Key

Test name: CH1 Business Statistics E10

TRUE

A data set provides information about some group of individual elements, which may be people, object, events, or other entities.

FALSE

A census is defined as examining all of the population measurements.

TRUE

On each random selection, we give every element in the population the same chance of being chosen.

FALSE

This is an example of a qualitative or categorical variable.

FALSE

This is an example of a quantitative variable.

TRUE

Statistical inference is the process of using randomly selected data to make generalizations or deductions about the population that the random sample was drawn from originally.

FALSE

Time series data are collected over different time periods.

TRUE

Cross-sectional data are collected at the same point in time an example would be a specific month for a cell phone bill for several employees or sales during a specific time period.

FALSE

Cross-sectional data are collected at the same point in time. This is an example of time series data.

FALSE

This is an example of cross-sectional data. Time series data are collected at different time periods.

TRUE

This data is collected over different time periods; specifically 10 different years at Apex Co.

FALSE

Qualitative variables are also known as categorical variables.

FALSE

The college business students are the elements of the data set. The cumulative GPA is an example of a variable, which is a characteristic of an element (i.e., a college business student) in the data set.

TRUE

When initiating a study, we first define our variable of interest, or response variable. Other variables are called factors.

FALSE

In experimental studies, the aim is to manipulate the factor(s), which may be related to the response variable.

FALSE

This is the definition of descriptive statistics. Statistical inference is the science of using a sample of measurements to make generalizations about the population of measurements.

TRUE

Statistical inference is the science of using a sample of measurements to make generalizations about the important aspects of a population of measurements.

TRUE

A process is a sequence of operations that takes inputs (labor, materials, etc.) and turns them into outputs (products and services).

FALSE

Using different samples and tests to produce a desired conclusion does not make the conclusion true.

TRUE

This would be purposely selecting a biased sample that overrepresents population elements supporting a desired conclusion and would be unethical.

TRUE

Primary data are data collected by an individual or business directly through planned experimentation or observation.

TRUE

Secondary data is not collected by the individual who is conducting the study.

TRUE

Data warehousing has as its centralized ideal the creation and maintenance of a central repository for all of an organization's data.

FALSE

Big data is a term that arose from the huge capacity of data warehouses that contain massive amounts of data.

FALSE

A stratified random sample is created by dividing the population into non-overlapping groups.

TRUE

This type of sampling guarantees that all the elements in a sample will be different and thus we will have the fullest possible look at the population.

TRUE

This is a random sample that is not chosen with bias.

TRUE

This type of sampling is advantageous when selecting a sample from a frame or single list is not possible.

TRUE

Stratification can make a sample easier to select and while a whole frame may not exist a frame may exist in each small stratum.

TRUE

This is when you want to systematically select a sample of *n* lowercase *n* elements without replacement from a frame of *N* uppercase *N* elements.

TRUE

This is when we observe only a portion of the population included in the sample while the remainder is obscured.

FALSE

Target population is the entire population of interest.

FALSE

Errors of non-observation relate to population elements that are not observed.

TRUE

This is when either the respondent or interviewer incorrectly marks the answer and thus is an error of observation when data values are recorded incorrectly.

FALSE

Low response rates do affect the validity of a survey's results.

TRUE

Sampling error is the difference between a numerical descriptor of the population and the corresponding descriptor of the sample. The sample does not reflect the population.

FALSE

Convenience sampling is not probability sampling. Convenience sampling is a type of sampling in which we select elements because they are easy or convenient to sample.

FALSE

Judgment sampling has an extremely knowledgeable individual select the sample. Voluntary sampling occurs when participants self-select, which is a form of convenience sampling, where elements are selected because they are easy or convenient to sample.

TRUE

This type of sample depends upon the judgement of the person selecting the sample and thus is a biased sample.

FALSE

Business analytics is an extension of traditional statistics.

FALSE

This is the definition of predictive analytics. Prescriptive analytics uses results from predictive analytics to recommend courses of action within the business.

D

It is a finite population because only a finite number of customers will use the drive-thru. An infinite population would be defined as the theoretical potential number of customers.

B

Random sampling, judgment sampling, and probability sampling are methods of sampling in which the selected elements may not be convenient to sample.

C

Convenience, voluntary, and judgment sampling should not be used to make valid statistical inferences about a population.

C

Bullet graphs are a method of graphical descriptive analytics.

D

Predictive analytics (whether parametric or nonparametric) are methods used to predict values of a response variable on the basis of one or more predictor variables. Prescriptive analytics are techniques that combine external and internal constraints with results from descriptive or predictive analytics to recommend an optimal course of action.

B

Factor analysis is an unsupervised learning technique because there is no specific response variable involved, which is a requirement for a supervised learning technique.

B

By definition, big data are collected by business for effective decision making.

A

Definition of traditional statistics.

B

Sampling with replacement occurs when a selected element is replaced before another sample is taken; systematic and voluntary response samples are not random.

D

All three of the listed question types can be used in survey design.

D

Sample surveys are the result of sampling designs. Random sampling, stratified random sampling, cluster sampling, and systematic sampling are sampling designs which are types of probability sampling.

B

A sample can be only a part of a population; a census is the examination of the population and variable is a characteristic of an element of the population.

C

Strata are groups within a population sample which do not overlap.

A

Dichotomous questions consist of only two possible responses.

C

Exclusion of population elements in selection is not a result of nonresponse or error of observation because this occurs during the survey itself. Sampling error is a result of the survey process.

A

Nonresponse, nonobservation and observation error occur during the survey process. Sampling error is a result of the survey process.

C

By definition, primary data are collected while secondary data are from an existing source.

B

By definition, ratio variables are quantitative and have an absolute zero value.

C

A quantitative variable is measurable and noncategorical.

D

A categorical variable is qualitative, not measured.

B

By definition, elements are the members of the population and variables are characteristics of elements; a measurement (or observation) assigns a value to a variable for an element of the population. A process is a sequence of operations that takes inputs and turns them into outputs.

D

Nominative and ordinal are types of qualitative variables.

C

Temperature is quantitative (excludes nominative and ordinal), and the ratio of two temperatures is not meaningful.

A

Interval and ratio are quantitative variables; jersey numbers have no logical order.

D

Nominative and ordinal are qualitative variables; weight creates logical ratios: 60 lb is twice as heavy as 30 lb.

B

Interval and ratio are quantitative variables, nominative is only a naming category, and police rank has order.

A

A runs plot is a graphical display of time series data.

D

By definition, a time series is a study of data over time; descriptive statistics is the study of the measurements of population variables; a random sample is a data set.

D

Nominative and ordinal are qualitative variables; college entrance exam scores have no meaningful ratio and no inherently defined zero value.

D

Nominative and ordinal are qualitative variables; miles driven can have a meaningful ratio.

C

Ratio and interval are quantitative variables; ordinal implies order or rank.

A

This is a qualitative variable without order; therefore, a nominative variable.

D

Measurement and observation are methods attached to a variable; a sample is a subset of the units in a population.

A

By definition, a census looks at the entire population.

C

A process is a sequence of operations; a census looks at the entire population; set is related to population.

B

By definition, sampling is taking a portion of the population to measure; experimental and observational analysis are methods of obtaining data.

A

A runs plot is a graphical display of data over time.

D

By definition, inference is taking a sample of data and its measurements and relating those measurements to the population as a whole.

C

By definition, a census looks at an entire population; a variable is a characteristic of an element within the population; a process is a sequence of operations that produces elements of a population.

B

Qualitative, categorical, and nominative variables are non-quantitative variables.

A

By definition, a runs plot is a graphical display; random sampling is a method of selecting a portion of a population; statistical inference is the science of using a sample of measurements to infer about the entire population.

A

Quantitative, ratio, and interval all have similar definitions.

D

By definition, a census is the examination of all population measurements; a process is a sequence of operations; a sample is a subset of a population.

C

A runs plot and time series analysis both look at data over time; cross-sectional analysis looks at data collected at the same point in time.

D

Qualitative and ordinal have similar definitions; random variables are all characteristics of a population element.

C

By definition, a variable is a characteristic of an element; a measurement assigns a value to a variable; an element is one unit of a population.

B

By definition, measurements assign values to a variable of an element; statistics is the science of describing aspects of a set of measurements; variables are characteristics of elements in a population.

A

By definition, an experimental data source is a collection of data where one is able to manipulate values; an observational data source is a collection of data where one is unable to control factors. Cross-sectional is not a defined data source but rather a way of analyzing or displaying the data that have been collected.

D

Runs plots are a way of looking at processes over time, which can then be used to make inferences about a population. Simply looking at descriptive statistics (of which, proportion and cross-sectional analysis are methods or procedures) is not sufficient to make inferences.

B

It is unethical to use methods or procedures designed to mislead the audience that is viewing the findings.

A

A time series is a collection of data taken over time, while a cross-section is a collection of data taken at the same point in time.

B

A time series is a collection of data taken over time, while a cross-section is a collection of data taken at the same point in time.

D

D

B

D

B

Essay

The response variable in this study is the number of miles driven over the past 30 days.

Response variables are defined as the variable of interest in a study.

Essay

Observational study

An observational study occurs when analysts are unable to control the factors of interest. An experimental study occurs when values of factors that are related to the variable of interest can be set or manipulated.

Essay

Factors in this study are location of residence, type of car, number of miles from work, number of children under 18, and monthly income.

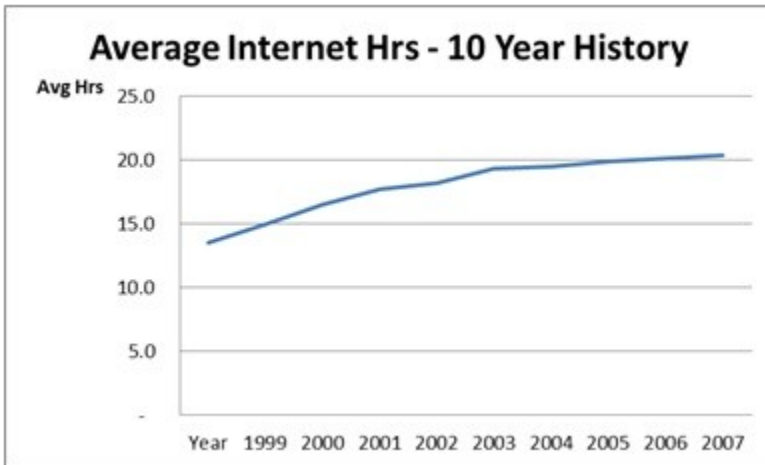
Factors are related to the variable of interest.

Essay

The price of gas peaked in the seventh month. The lowest price is observed around 20 to 21 months from the start of the data collection. At the end of the 30 months, gas price is beginning to show stability.

Observing the rise and fall of a time series or runs plot.

Essay



Displaying the average hours spent on Internet activities graphically results in a time series or runs plot. An increase over time in the amount of time can be observed through either the graph or data.