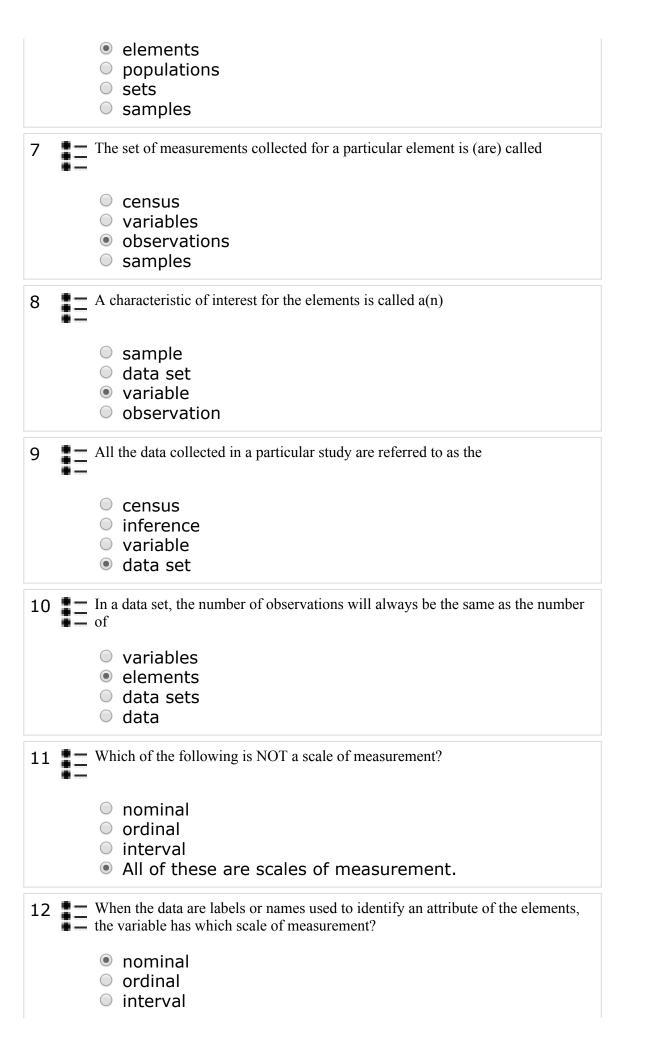
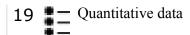
Chapter_1_Data_and_Statistics

Which of the following is an example of quantitative data? the player's number on a baseball uniform the serial number on a one-dollar bill the part number of an inventory item the number of people in a waiting line Which of the following is <u>not</u> an example of descriptive statistics? a histogram depicting the age distribution for 30 randomly selected students an estimate of the number of Alaska residents who have visited Canada a table summarizing the data collected in a sample of new-car buyers the proportion of mailed-out questionnaires that were returned Which of the following is an example of categorical data? social security number score on a multiple-choice exam height, in meters, of a diving board number of square feet of carpet The number observations in a complete data set having 10 elements and 5 variables is 5 • 10 **25 50** Facts and figures that are collected, analyzed and summarized for presentation and interpretation are data variables elements Both variables and elements are correct. The entities on which data are collected are



		o ratio
13	==	When the data are labels or names used to identify an attribute of the elements and the rank of the data is meaningful, the variable has which scale of measurement? onominal ordinal interval ratio
14		When the data have the properties of ordinal data and the interval between observations is expressed in terms of a fixed unit of measure, the variable has which scale of measurement? onominal ordinal interval ratio
15	: =	When the data have the properties of interval data and the ratio of two values is meaningful, the variable has which scale of measurement? onominal ordinal interval ratio
16	ŧ	Which scale of measurement can be either numeric or nonnumeric?
		 nominal ordinal interval nominal and ordinal
17	ŧ	Which of the following variables use the interval scale of measurement?
		 name of stock exchange time SAT scores Social Security number
18	: =	Which of the following variables use the ratio scale of measurement?
		 name of stock exchange time SAT scores Social Security number



- are always nonnumeric
- may be either numeric or nonnumeric
- are always numeric
- are always labels

20 __ Categorical data

- are always nonnumeric
- may be either numeric or nonnumeric
- are always numeric
- indicate either how much or how many

21 Tategorical data

- indicate either how much or how many
- can not be numeric
- are labels used to identify attributes of elements
- must be nonnumeric

22 — Ordinary arithmetic operations are meaningful

- only with categorical data
- only with quantitative data
- either with quantitative or categorical data
- with rank data

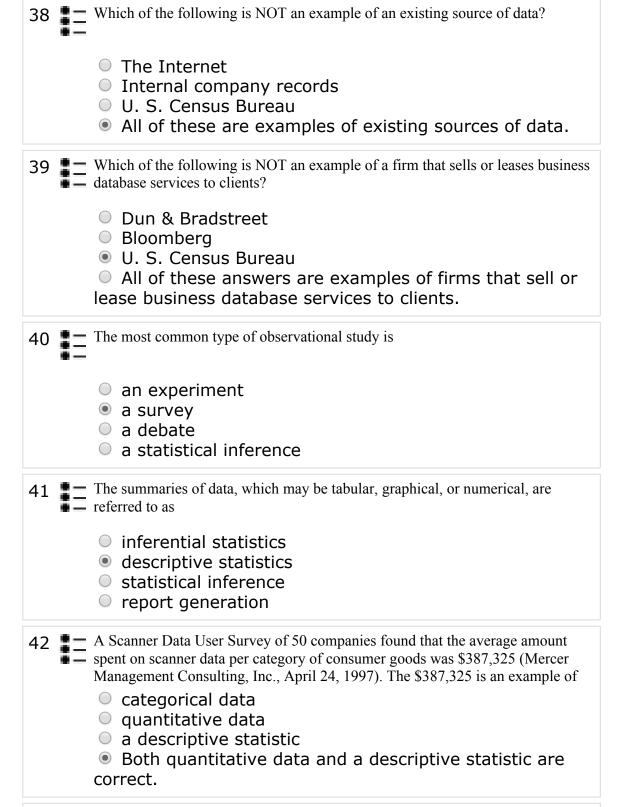
23 ___ Arithmetic operations are inappropriate for

- categorical data
- quantitative data
- both categorical and quantitative data
- large data sets
- 24 ___ In a questionnaire, respondents are asked to mark their gender as male or female.

 Gender is an example of a
 - categorical variable
 - quantitative variable
 - categorical or quantitative variable, depending on how the respondents answered the question
 - ratio variable
- In a questionnaire, respondents are asked to record their age in years. Age is an example of a

- categorical variable quantitative variable categorical or quantitative variable, depending on how the respondents answered the question ratio variable In an application for a credit card, potential customers are asked for their social security numbers. A social security number is an example of a categorical variable
 - - quantitative variable
 - categorical or quantitative variable, depending on how the respondents answered the question
 - ratio variable
- 27 Temperature is an example of
 - a categorical variable
 - a quantitative variable
 - either a quantitative or categorical variable
 - neither a quantitative nor categorical variable
- 28 🔭 For ease of data entry into a university database, 1 denotes that the student is an — undergraduate and 2 indicates that the student is a graduate student. In this case data are
 - categorical
 - quantitative
 - either categorical or quantitative
 - neither categorical nor quantitative
- 29 __ Income is an example of
 - categorical data
 - either categorical or quantitative data
 - currency data
 - quantitative data
- The birth weight of newborns, measured in grams, is an example of
 - categorical data
 - either categorical or quantitative data
 - neither categorical nor quantitative data
 - quantitative data
- 31 The weight of a ball bearing, measured in milligrams, is an example of
 - categorical data

			either categorical or quantitative data neither categorical nor quantitative data quantitative data
32	: =	The dolla	profits earned by a particular Fortune 500 company, measured in millions of ars, is an example of
		\bigcirc	categorical data either categorical or quantitative data neither categorical nor quantitative data quantitative data
33	: =	Data	collected at the same, or approximately the same, point in time are
		•	time series data static data cross-sectional data one-dimensional data
34	: =	Data	collected over several time periods are
		0	time series data time controlled data cross-sectional data time dependent data
35	: =	Stati	stical studies in which researchers do not control variables of interest are
		\bigcirc	experimental studies uncontrolled experimental studies not of any value observational studies
36	: =	Stati	stical studies in which researchers control variables of interest are
		\bigcirc	experimental studies control observational studies non experimental studies observational studies
37	: =		is an example of an existing source of data.
		•	An experimental study An uncontrolled experimental study U. S. Census Bureau An observational study



43 __ Statistical inference

- refers to the process of drawing inferences about the sample based on the characteristics of the population
- is the same as descriptive statistics
- is the process of drawing inferences about the population based on the information taken from the

	sample is the same as a census
44 📜	The collection of all elements of interest in a particular study is
	 the population the sampling statistical inference descriptive statistics
45	A portion of the population selected to represent the population is called
	 statistical inference descriptive statistics a census a sample
46	In a sample of 800 students in a university, 360, or 45%, live in the dormitories. The 45% is an example of
	 a sample a population statistical inference descriptive statistics
	In a sample of 800 students in a university, 160, or 20%, are Business majors. Based on the above information, the school's paper reported that "20% of all the students at the university are Business majors." This report is an example of a sample a population statistical inference descriptive statistics
48	Six hundred residents of a city are polled to obtain information on voting intentions in an upcoming city election. The six hundred residents in this study is an example of a(n) census sample observation population
49	A statistics professor asked students in a class their ages. On the basis of this information, the professor states that the average age of all the students in the university is 21 years. This is an example of a census descriptive statistics an experiment statistical inference

- The owner of a factory regularly requests a graphical summary of all employees' salaries. The graphical summary of salaries is an example of
 - a sample
 - descriptive statistics
 - statistical inference
 - an experiment
- The Department of Transportation of a city has noted that on the average there are 14 accidents per day. The average number of accidents is an example of
 - descriptive statistics
 - statistical inference
 - a sample
 - a population

52 **=** Exhibit 1-1

In a recent study based upon an inspection of 200 homes in Daisy City, 120 were found to violate one or more city codes.

Refer to Exhibit 1-1. The city manager released a statement that 60% of Daisy City's 3,000 homes are in violation of city codes. The manager's statement is an example of

- a census
- an experiment
- descriptive statistics
- statistical inference

53 **Exhibit 1-1**

In a recent study based upon an inspection of 200 homes in Daisy City, 120 were found to violate one or more city codes.

Refer to Exhibit 1-1. The Daisy City study described above is an example of the use of a

- census
- sample
- probability
- population

54 **Exhibit** 1-1

In a recent study based upon an inspection of 200 homes in Daisy City, 120 were found to violate one or more city codes.

Refer to Exhibit 1-1. The manager's statement that 60% of Daisy City's 3,000 homes are in violation of city codes is

an exactly correct statement

- only an approximation, since it is based upon sample information
- obviously wrong, since it is based upon a study of only 200 homes
- None of the other answers are correct.

55 **Exhibit 1-2**

In a sample of 3,200 registered voters, 1,440, or 45%, approve of the way the President is doing his job.

Refer to Exhibit 1-2. The 45% approval is an example of

- a sample
- descriptive statistics
- statistical inference
- a population

56 **Exhibit 1-2**

In a sample of 3,200 registered voters, 1,440, or 45%, approve of the way the President is doing his job.

Refer to Exhibit 1-2. A political pollster states, "Forty five percent of all voters approve of the President." This statement is an example of

- a sample
- descriptive statistics
- statistical inference
- a population
- 57 The process of analyzing sample data in order to draw conclusions about the characteristics of a population is called
 - descriptive statistics
 - statistical inference
 - data analysis
 - data summarization
- In a post office, the mailboxes are numbered from 1 to 5,000. These numbers represent
 - categorical data
 - time series data
 - either categorical or quantitative data
 - quantitative data
- The average age in a sample of 90 students at City College is 20. As a result of this sample, it can be concluded that the average age of all the students at City College
 - must be more than 20, since the population is always

larger than the sample

- must be less than 20, since the sample is only a part of the population
- could not be 20
- could be larger, smaller, or equal to 20
- 60 Since a sample is a subset of the population, the sample mean
 - is always smaller than the mean of the population
 - is always larger than the mean of the population
 - must be equal to the mean of the population
 - can be larger, smaller, or equal to the mean of the population
- 61 _ Since a sample is a subset of the population, a percentage that is calculated from the sample data
 - is always smaller than the corresponding percentage from the population
 - is always larger than the corresponding percentage from the population
 - must be equal to the corresponding percentage from the population
 - can be larger, smaller, or equal to the corresponding percentage from the population
- A sample of five Fortune 500 companies showed the following revenues (\$ millions): 7505.0, 2904.7, 7208.4, 6819.0, and 19500.0. Based on this information, which of the following statements is correct?
 - An estimate of the average revenue for all Fortune 500 companies is 8787.42 (\$\phi\text{millions}\).
 - The average revenue for all Fortune 500 companies is 8787.42 (\$ millions).
 - Over half of all Fortune 500 companies earn at least 7208.4 (\$ millions) in revenues.
 - If five other Fortune 500 companies were chosen, the average revenue would be 8787.42 (\$ millions).
- A sample of five Fortune 500 companies showed the following industry codes: banking, banking, finance, retail, and banking. Based on this information, which of the following statements is correct?
 - Sixty percent of the sample of five companies are banking industries.
 - Sixty percent of all Fortune 500 companies are banking industries.
 - Banking is the most common type of industry among all Fortune 500 companies.
 - If five other Fortune 500 companies were chosen, 60% of them would be banking industries.

64	: =	is a spreadsheet package.
		Statistical inference
		The bar chart
		The New York Stock ExchangeMicrosoft Excel
		MICIOSOIT EXCEI
65	: =	Dr. Kurt Thearling, a leading practitioner in the field, defines data mining as "the extraction of information from databases".
		thorough, insightful
		timely, accurate
		automated, predictive
		intentional, useful
		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
66	: =	The major applications of data mining have been made by companies with a strong focus.
		consumer
		manufacturing
		exporting
		research and development
57		Quantitative data that measure how many are; quantitative data that measure how much are
		interval; ratio
		o ratio; interval
		o continuous; discrete
		discrete; continuous
		4.55.55, 55.4.
58	: =	?Flight time from Cincinnati to Atlanta is an example of a variable and measurement.
		?discrete, interval
		• ?discrete, ratio
		?continuous, interval
		?continuous, ratio
		. continuous, rutio
59	: =	Which of the following is <u>not</u> a categorical variable??
		?your age on your last birthday
		• ?your cell phone area code
		?your accounting class start time
		?your high school graduation year
		· · · · · · · · · · · · · · · · · · ·
70	:=	Categorical data use either the or scale of measurement.?

- ?nominal, ordinal
- ?nominal, interval
- ?ordinal, interval
- ?interval, ratio
- 71 ___ The largest experimental statistical study ever conducted is believed to be for?
 - ?Cholera
 - Polio
 - ?Diphtheria
 - ?Malaria
- 72 The term *data warehousing* is used to refer to the process of doing all of the following except?
 - ?mining the data
 - ?capturing the data
 - ?storing the data
 - ?maintaining the data
- 73 The U.S. Census Bureau conducts a census every
 - ?4 years
 - ?5 years
 - ?8 years
 - ?10 years
- After the graduation ceremonies at a university, six graduates were asked whether they were in favor of (identified by 1) or against (identified by 0) abortion. Some information about these graduates is shown below.

♦ Graduate	<u>Sex</u>	<u>Age</u>	Abortion Issue	Class Rank
Nancy	F	22	1	�� 3
Michael	M	21	1	�� 2
Tammy	F	33	0	�� 1
John	M	38	0	20
Marlene	F	25	1	�� 4
David	M	19	0	�� 8

- a. How many elements are in the data set?
- b. How many variables are in the data set?
- c. How many observations are in the data set?
- d. Name the scale of measurement for each of the above (Sex, Age, Abortion Issue, Class rank).
- e. Which of the above (Sex, Age, Abortion Issue, Class rank) are categorical and which are quantitative variables?
- f. Are arithmetic operations appropriate for the variable "abortion issue"?

- a. 6
- b. 4
- c. 6
- d. Sex: nominal
- Age: ratio
- **♦** Abortion Issue: nominal
- Class rank: ordinal
- e. Sex: categorical
- Age: quantitative
- Abortion Issue: categorical
- Class Rank: categorical
- f. No

75	

An issue of Fortune Magazine reported that the following companies had the lowest sales per employee among the Fortune 500 companies.

•	Company	Sales per Employee (\$1000s)	Sales Rank
•	Seagate Technology	\$42.20	285
•	SSMC	42 .19	414
•	Russell	�� 41.99	480
•	Maxxam	♦♦ 40.88	485
•	Dibrell Brothers	\$2 2.56	470

- a. How many elements are in the above data set?
- b. How many variables are in the above data set?
- c. How many observations are in the above data set?
- d. Name the scale of measurement for each of the variables.

e. Name the variables and indicate whether they are categorical or quantitative.

Answer:

- a. 5
- b. 2
- c. 5
- d. Sales per employee: ratio; Sales rank: ordinal
- e. Sales per employee: quantitative; Sales rank: categorical

76

The following shows the temperatures (high, low) and weather conditions in a given Sunday for some selected world cities. For the weather conditions, the following notations are used: c = clear; cl = cloudy; sh = showers; pc = partly cloudy.

•	City	<u>Hi</u>	Lo	Condition	•
•	Acapulco	99	77	pc	•
•	Bangkok	92	78	pc	•
•	Mexico City	77	57	sh	•
•	Montreal	72	56	pc	•
•	Paris	77	58	�� c	•
•	Rome	88	68	cl	•
•	Toronto	78	61	�� c	•

- a. How many elements are in this data set?
- b. How many variables are in this data set?
- c. How many observations are in this data set?
- d. Name the scale of measurement for each of the variables.
- e. Name the variables and indicate whether they are categorical or quantitative.
- f. For which variables are arithmetic operations appropriate and for which are they not appropriate?

Answer:

- a. 7
- b. 3
- c. 7
- d. Hi: interval, Lo: interval, Condition: nominal
- e. Hi: quantitative, Lo: quantitative, Condition: categorical
- f. Hi: appropriate, Lo: appropriate, Condition: not appropriate

77	A magazine survey are s	e surveyed a sam shown below.	iple of its	s subscribers. Some of the	e responses from	n the
	*	•	•	Annual Household	•	

•	Subscriber ID	<u>Sex</u>	<u>Age</u>	Income (\$1000)	•
•	0006	F	22	45	•
•	4798	M	21	53	•
•	2291	F	33	82	•
•	4988	M	38	30	•
ŵ			٠		

♦	
a.	How many elements are in the data set?
b.	How many variables are in the data set?
c.	How many observations are in the data set?
d.	Name the scale of measurement for each of the variables

e. Which of the above (Sex, Age, Annual Household Income) are categorical and which are quantitative?

f. Are the data time series or cross-sectional?

Answer:

- a. 4
- b. 3
- c. 4
- d. Sex: nominal, Age: ratio, Annual Household Income: ratio
- e. Sex: categorical, Age: quantitative, Annual Household Income: quantitative
- f. cross-sectional

A magazine surveys a sample of its subscribers every year. Some of the responses are shown below.

•	•	Percent	Average	Average Annual
•	Year	<u>Female</u>	Age	Household Income (\$1000)
•	1996	5.8	35.2	40
•	1997	8.6	35.8	42

•	1998	7.3	33.9	41
•	1999	9.2	35.3	43
•	2000	9.8	36.6	46

- a. How many variables are in the data set?
- b. Name the scale of measurement for each variable.
- Which of the above (Year, Percent Female, Average Age, Average Annual
- Household Income) are categorical and which are quantitative?
- d. Are the data time series or cross-sectional?

- a. 4
- b. Year: interval, Percent Female: ratio, Average Age: ratio, Average Annual Household Income: ratio
- c. all are quantitative
- d. time series
- The following data shows the yearly income distribution of a sample of 200 employees at MNM, Inc.

•	Yearly Income (\$1000s)	Number of Employees	•
•	20 ? 24	�� 2	•
•	25 ? 29	48	•
•	30 ? 34	60	•
•	35 ? 39	80	•
•	40 ? 44	10	•

- a. What percentage of employees has yearly incomes of more than \$35,000?
- b. Is the figure (percentage) that you computed in Part a an example of statistical inference? If no, what kind of statistics does it represent?

 Based on this sample, the president of the company said that "45% of all our
- c. employees' yearly incomes are over \$35,000." The president's statement represents what kind of statistics?
- d. With the statement made in Part c, can we be assured that more than 45% of all employees' yearly incomes are at least \$35,000? Explain.

 What percentage of employees of the sample has yearly incomes of \$29,000

- e. or less?
- f. How many variables are presented in the above data set?
- g. The above data set represents the results of how many observations?

- a. 45%
- b. No, it is descriptive statistics.
- c. statistical inference
- d. No, this is simply an inference and approximation based on the sample information.
- e. 25%
- f. 2
- g. 200



A recent issue of a national magazine reported that in a national public opinion survey conducted among 2,000 individuals, 56% were in favor of gun control, 40% opposed gun control, and 4% had no opinion on the subject.

- a. What is the sample in this survey?
- b. Based on the sample, what percentage of the population would you think is in favor of gun control?
- Based on the sample, what percentage of the population would you think have no opinion on the subject?

Answer:

- a. The 2000 individuals who were approached
- b. 56%
- c. 4%



A pharmaceutical company is performing clinical trials on a new drug that is intended to relieve symptoms for allergy sufferers. Twelve percent of the 300 clinical trial participants experienced the side effect dry mouth.

- a. What is the population being studied?
- b. What is the sample being studied?
- Based on the sample, what percentage of the population do you think would suffer from dry mouth?

Answer:

- a. All allergy sufferers
- b. The 300 participants
- c. 12%

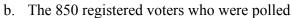


A polling organization conducts a telephone poll of 850 registered voters and asks which candidate they will vote for in the upcoming presidential election. Forty-three percent of the respondents prefer candidate A and 45% prefer candidate B.

- a. What is the population being studied?
- b. What is the sample being studied?
- c. Based on the sample, what percentage of the population do you think would vote for candidate B?

Answer:

a. All registered voters



c. 45%

Ω2	
$o_{\mathcal{I}}$	

The following table shows the starting salaries of a sample of recent business graduates.

•	Income (In \$1,000s)	Number of Graduates	•
•	15 ? 19	40	•
•	20 ? 24	60	•
•	25 ? 29	80	•
•	30 ? 34	18	•
•	35 ? 39	�� 2	•

- a. What percentage of graduates in the sample had starting salaries of at least \$30,000?
- b. Of the graduates in the sample, what percentage had starting salaries of less than \$25,000?
- c. Based on this sample, what percentage of all business graduates do you estimate to have starting salaries of at least \$20,000?

Answer:

- a. 10%
- b. 50%
- c. 80%



Michael, Inc., a manufacturer of electric guitars, is a small firm with 50 employees. The table below shows the hourly wage distribution of the employees.

•	Hourly Wages (In Dollars)	Number of Employees	•
•	10 ? 13	�� 8	•
•	14 ? 17	12	•
•	18 ? 21	20	•
Ŷ	22 ? 25	10	•

- a. How many employees receive hourly wages of at least \$18?
- b. What percentage of the employees has hourly wages of at least \$18?
- c. What percentage of the employees has hourly wages of less than \$14?

- a. 30
- b. 60%
- c. 16%

85



Laura Naples, Manager of Heritage Inn, periodically collects and tabulates information about a sample of the hotel's overnight guests. This information aids her in planning and scheduling decisions she must make. The table below lists data on ten randomly selected hotel registrants, collected as the registrants checked out. The data listed for each registrant are: number of people in the group; birth date of person registering; shuttle service used, yes or no; total telephone charges incurred; and reason for stay, business or personal.

Name of Registrant	People <u>In Group</u>	Birth Date (mm/dd/yy)	Shuttle <u>Used</u>	Telephone <u>Charges</u>	Reason For Stay
Adam Norquist	1	05/07/59	yes	\$0.00	personal
Mica Pepper	4	11/23/48	no	12.46	business
Claude Shepler	2	04/30/73	no	1.20	business
Amy Rodriquez	2	12/16/71	no	2.90	business
Tony DiMarco	1	05/09/39	yes	0.00	personal
Amy Franklin	3	09/14/69	yes	4.65	business
Tammy Roberts	2	04/22/66	no	9.35	personal
Ed Blackstone	5	10/28/54	yes	2.10	personal
Mary Silverman	1	11/12/49	no	1.85	business
Todd Atherton	2	01/30/62	no	5.80	business

- a. How many elements are there in the data set?
- b. How many variables are there in the data set?
- c. How many observations are there in the data set?
- d. What are the observations for the second element listed?
- e. What is the total number of measurements in the data set?
- f. Which variables are quantitative?
- g. Which variables are qualitative?
- h. What is the scale of measurement for each of the variables?
- i. Does the data set represent cross-sectional or times series data?
- j. Does the data set represent an experimental or an observational study?

- a. 10 elements
- b. 5 variables
- c. 10 observations
- d. 4, 11/23/48, no, 12.46, business
- e. 50
- f. people in group, telephone charges
- g. birth date, shuttle used, reason for stay
- h. people in group ratio scale, birth date ordinal scale, shuttle use nominal, telephone charge ratio scale, reason for stay nominal scale
- i. cross-sectional
- j. observational

86

Suppose the current weather report for your area contains the following information. Specify the measurement scale for each of the variables.

Temperature 84°

Wind Speed 10 mph

Wind Direction (from the) South

Sky Description Sunny Molds Level High

Answer:

Temperature - interval Wind Speed - ratio Wind Direction - nominal Sky Description - nominal Molds Level - ordinal



Molly Porter owns and operates two convenience stores, one on the East side of the city and the other on the South side. She has workforce-planning decisions to

make and has collected some recent sales data that are relevant to her decisions. Listed below are the monthly sales (\$000) at her two stores for the past six months.

<u>Store</u>	<u>March</u>	<u>April</u>	May	<u>June</u>	<u>July</u>	August
East	102	100	103	105	109	106
South	72	74	81	86	92	93

- a. Is the data set cross-sectional or time series data? Explain.
- b. Comment on any apparent patterns you see in the data.

Answer:

- a. Time series data for two variables: monthly sales for East store and monthly sales for South store.
- b. Both stores have been experiencing an overall rise in sales during the past six months. The South store's increase in sales (as a percentage of sales) has been greater than the East store's increase. The increases might be temporary, due to the seasonal nature of demand. It is also possible that the increases will continue.

88

Cortland Software, Inc. surveyed a sample of its programmers and recorded the following data on their: age, gender, years of programming experience, whether or not they have a master's degree, and annual salary.

?	?	?	Years of	Master's	?
<u>Programmer</u>	<u>Age</u>	<u>Gender</u>	<u>Experience</u>	<u>Degree</u>	Salary
1	32	Male	8	No	\$59,450
2	27	Male	2	Yes	62,750
3	31	Female	6	Yes	65,820
4	37	Males	5	No	54,620
5	29	Female	5	Yes	61,900
6	25	Female	4	No	48,370

?	
!	
a.	How many elements are in the data set?
b.	How many observations are in the data set?
c.	How many variables are in the data set?
d.	Name the scale of measurement for each variable.
e.	Are the data time series or cross-sectional?
?	
Ar	eswer:
?	
a.	6
b.	6
c.	5
d.	Age: ratio; MorF: nominal; Exper.: ratio; Degree: nominal; Salary: ratio
e.	Cross-sectional

?