I3–TD1
(Descriptive Statistics)

1. For statements (a)–(h), state whether descriptive or inferential statistics has been used.
   (a) By 2040 at least 3.5 billion people will run short of water (World Future Society).
   (b) In a sample of 100 on-the-job fatalities, 90% of the victims were men.
   (c) In a survey of 1000 adults, 34% said that they posted notes on social media websites (Source: AARP Survey).
   (d) In a poll of 3036 adults, 32% said that they got a flu shot at a retail clinic (Source: Harris Interactive Poll).
   (e) Allergy therapy makes bees go away (Source: Prevention).
   (f) Drinking decaffeinated coffee can raise cholesterol levels by 7% (Source: American Heart Association).
   (g) The average stay in a hospital for 2000 patients who had circulatory system problems was 4.7 days.
   (h) Experts say that mortgage rates may soon hit bottom (Source: USA TODAY).
2. For statements (a)–(i), classify each as nominal-level, ordinal-level, interval-level, or ratio-level measurement
   (a) Pages in the 25 best-selling mystery novels.
   (b) Rankings of golfers in a tournament.
   (c) Temperatures inside 10 pizza ovens.
   (d) Weights of selected cell phones.
   (e) Salaries of the coaches in the NFL.
   (f) Times required to complete a chess game.
   (g) Ratings of textbooks (poor, fair, good, excellent).
   (h) Number of amps delivered by battery chargers.
   (i) Ages of children in a day care center. Categories of magazines in a physician's office (sports, women's, health, men's, news).
3. For statements (a)-(h), classify each variable as qualitative or quantitative.
   (a) Marital status of nurses in a hospital.
   (b) Time it takes to run a marathon.
   (c) Weights of lobsters in a tank in a restaurant.
   (d) Colors of automobiles in a shopping center parking lot.
   (e) Ounces of ice cream in a large milkshake.
   (f) Capacity of the NFL football stadiums.
   (g) Ages of people living in a personal care home.
   (h) Different vitamins taken.
4. For statements (a)-(h), classify each variable as discrete or continuous.
   (a) Number of pizzas sold by Pizza Express each day.

(b) Relative humidity levels in operating rooms at local hospitals.

(c) Number of bananas in a bunch at several local super-markets.

(d) Lifetimes (in hours) of 15 iPod batteries.

(e) Weights of the backpacks of first-graders on a school bus.

(f) Number of students each day who make appointments with a math tutor at a local college.

(g) Blood pressures of runners in a marathon.

(h) Ages of children in a preschool.

5. **How People Get Their News** The Brunswick Research Organization surveyed 50 randomly selected individuals and asked them the primary way they received the daily news. Their choices were via newspaper (N), television (T), radio (R), or Internet (I). Construct a categorical frequency distribution for the data and interpret the results.

N	N	T	T	T	I	R	R	I	T
I	N	R	R	I	N	N	I	T	N
I	R	T	T	T	T	N	R	R	I
R	R	I	N	T	R	T	I	I	T
T	I	N	T	T	I	R	N	R	T

6. **College Completions** The percentage (rounded to the nearest whole percent) of persons from each state completing 4 years or more of college is listed below.

Percentage of persons completing 4 years of college

23	25	24	34	22	24	27	37	33	24
26	23	38	24	24	17	28	23	30	25
30	22	33	24	28	36	24	19	25	31
34	31	27	24	29	28	21	25	26	15
26	22	27	21	25	28	24	21	25	26

(a) Organize the data into a grouped frequency distribution with 5 classes.

(b) Find the relative frequency.

(c) Construct a histogram, frequency polygon, and ogive.

7. **Ages of the Vice Presidents at the Time of Their Death** The ages of the Vice Presidents of the United States at the time of their death are listed below.

90	83	80	73	70	51	68	79	70	71	
72	74	67	54	81	66	62	63	68	57	
66	96	78	55	60	66	57	71	60	85	
76	98	77	88	78	81	64	66	77	93	70

(a) Use the data to construct a frequency distribution with 6 classes.

(b) Find the relative frequency.

(c) Construct a histogram, frequency polygon, and ogive.

8. **Activities While Driving** A survey of 1200 drivers showed the percentage of respondents who did the following while driving. Construct a vertical bar graph and a horizontal bar graph for the data.

Drink beverage	80%
Talk on cell phone	73
Eat a meal	41
Experience road rage	23
Smoke	21

9. **Calories of Nuts** The data show the number of calories per ounce in selected types of nuts. Construct vertical and horizontal bar graphs for the data.

Types	Calories
Peanuts	160
Almonds	170
Macadamia	200
Pecans	190
Cashews	160

10. **Space Launches** The data show the number of U.S. space launches for the 10-year periods from 1960 to 2009. Construct a time series graph for the data and analyze the graph.

Year	60–69	70–79	80–89	90–99	100–109
Launches	614	247	199	300	206

11. **High School Dropout Rate** The data show the high school dropout rate for students for the years 2003 to 2009. Construct a time series graph and analyze the graph.

Year	2003	2004	2005	2006	2007	2008	2009
Percent	9.9	10.3	9.4	9.3	8.7	8.0	8.1

12. **Spending of College Freshmen** The average amounts spent by college freshmen for school items are shown. Construct a pie graph for the data.

Electronics/computers	$728
Dorm items	344
Clothing	141
Shoes	72

13. **Career Changes** A survey asked if people would like to spend the rest of their careers with their present employers. The results are shown. Construct a pie graph for the data and analyze the results.

Answer	Number of people
Yes	660
No	260
Undecided	80

14. **Peyton Manning's Colts Career** Peyton Manning played for the Indianapolis Colts for 14 years. (He did not play in 2011.) The data show the number of touch-downs he scored for the years 1998–2010. Construct a dotplot for the data and comment on the graph.

26	33	27	49	31	27	33
26	26	29	28	31	33	

15. **Songs on CDs** The data show the number of songs on each of 40 CDs from the author's collection. Construct a dotplot for the data and comment on the graph.

10	14	18	11
11	15	16	10
10	17	10	15
22	9	14	12
18	12	12	15
21	22	20	15
10	19	20	21
17	9	13	15
11	12	12	9
14	20	12	10

16. The traffic situation in X-City is getting worse, and it is high time a solution was offered. The company hired to work on the project took a survey of the estimated amount of vehicles that move on the road daily and for various intervals. The result of this survey is illustrated in the table below.

Time	Cars	Buses	Bikes
1-2pm	37	45	42
2-3pm	44	34	26
3-4pm	23	39	27
4-5pm	29	41	48

Construct a multiple line graph to visualize the data. Hence, determine the vehicle with the highest frequency and that with the lowest frequency.

17. Draw a multiple bar graph for the following data which represented agricultural production for the priod from 2010-2013.

Year	Food grains (tones)	Vegetables (tones)	Others (tones)
2010	100	30	10
2011	120	40	15
2012	130	45	25
2013	150	50	25

18. The heights (in cm) of a sample of the students in a class are shown:

50	52	70	72	65	52	60
75	51	64	65	55	67	70

Find the mean, mode, median, inter quartile range, midrange, variance, and standard deviation for the data.

19. **Households of Four Television Networks** A survey showed the number of viewers and number of households of four television networks. Find the average number of viewers, using the weighted mean.

Households	1.4	0.8	0.3	1.6
Viewers (in millions)	1.6	0.8	0.4	1.8

20. **Magazines in Bookstores** A survey of bookstores showed that the average number of magazines carried is 56, with a standard deviation of 12. The same survey showed that the average length of time each store had been in business was 6 years, with a standard deviation of 2.5 years. Which is more variable, the number of magazines or the number of years?

21. **Average Earnings of Workers** The average earnings of year-round full-time workers 25–34 years old with a bachelor's degree or higher were \$58,500 in 2003. If the standard deviation is \$11,200, what can you say about the percentage of these workers who earn.

    (a) Between \$47,300 and \$69,700?

    (b) More than \$80,900?

    (c) How likely is it that someone earns more than \$100,000?

22. **Costs to Train Employees** For a certain type of job, it costs a company an average of \$231 to train an employee to perform the task. The standard deviation is \$5. Find the minimum percentage of data values that will fall in the range of \$219 to \$243. Use Chebyshev's theorem.

23. **Exam Completion Time** The mean time it takes a group of students to complete a statistics final exam is 44 minutes, and the standard deviation is 9 minutes. Within what limits would you expect approximately 95% of the students to complete the exam? Assume the variable is approximately normally distributed.

24. **Exam Grades** Which of these exam grades has a better relative position?

    (a) A grade of 82 on a test with $\bar{x} = 85$ and $s = 6$.

    (b) A grade of 56 on a test with $\bar{x} = 60$ and $s = 5$.

25. Check each data set for outliers.

    (a) 506, 511, 517, 514, 400, 521

    (b) 3, 7, 9, 6, 8, 10, 14, 16, 20, 12

26. Check each data set for outliers.

    (a) 14, 18, 27, 26, 19, 13, 5, 25

    (b) 112, 157, 192, 116, 153, 129, 131

27. The following sample data are the midterm examination test scores for 30 students:

55	60	91	85	60	70	89	99	59	67
72	82	60	68	57	74	64	70	68	91
89	90	83	40	79	85	71	80	76	81

    a. Find the mean, mode, median, variance, standard deviation, Q_1, and Q_3 of the data.

    b. Construct a frequency table with 5 classes.

    c. Using the grouped data formula, find the mean, mode, median, variance, standard deviation,Q_1, and Q_3 for the table in part (b) and compare it to the results in part (a).

    d. Construct a histogram and comment on the shape of the distribution.

    e. Find the percentile values of 55, 60, and 74.

28. For the following data:

6.3	2.9	4.5	1.1	1.8	4.0	1.2	3.1	2.0	4.0
7.0	2.8	4.3	5.3	2.9	8.3	4.4	2.8	3.1	5.6
4.5	4.5	5.7	0.5	6.2	3.7	0.9	2.4	3.0	3.5

(a) Find the mean, mode, median, variance, standard deviation,Q_1, Q_3, and 90th percentile.

(b) Construct a frequency table with 5 classes.

(c) Using the grouped data formula, find the mean, mode, median, variance, standard deviation,Q_1, Q_3, and 90th percentile for the frequency table constructed in part (b) and compare it to the results in part (a).

(d) Construct a histogram, and comment on the shape of the data.

29. In recent years, due to low interest rates, many homeowners refinanced their home mortgages. Linda Lahey is a mortgage officer at Down River Federal Savings and Loan. Below is the amount refinanced for 20 loans she processed last week. The data are reported in thousands of dollars and arranged from smallest to largest.

59.2	59.5	61.6	65.5	66.6	72.9	74.8	77.3	79.2	83.7
85.6	85.8	86.6	87.0	87.1	90.2	93.3	98.6	100.2	100.7

a. Find the median, first quartile, and third quartile.

b. Find the 26th and 83rd percentiles.

c. Draw a box plot of the data and comment on the shape of the distribution.

30. **Hours Worked** The data shown here represent the number of hours that 12 part-time employees at a toy store worked during the weeks before and after Christmas. Construct two boxplots and compare the distributions.

Before	38	16	18	24	12	30	35	32	31	30	24	35
After	26	15	12	18	24	32	14	18	16	18	22	12

31. Many times in statistics it is necessary to see if a set of data values is approximately normally distributed. There are special techniques that can be used. One technique is to draw a histogram for the data and see if it is approximately bell-shaped. (Note: It does not have to be exactly symmetric to be bell-shaped.)
The numbers of branches of the 50 top libraries are shown.

67	84	80	77	97	59	62	37	33	42
36	54	18	12	19	33	49	24	25	22
24	29	9	21	21	24	31	17	15	21
13	19	19	22	22	30	41	22	18	20
26	33	14	14	16	22	26	10	16	24

1. Construct a frequency distribution for the data.
2. Construct a histogram for the data.
3. Describe the shape of the histogram.

4. Based on your answer to question 3, do you feel that the distribution is approximately normal?
5. Find the mean and standard deviation for the data.
6. What percent of the data values fall within 1 standard deviation of the mean?
7. What percent of the data values fall within 2 standard deviations of the mean?
8. What percent of the data values fall within 3 standard deviations of the mean?
9. Does your answer help support the conclusion you reached in question 4? Explain.