Chapter 2: Graphical Methods for Describing Data Distributions
Concept Quiz

Name _____

The answers to the questions below frequently depend on remembering facts, understanding concepts, and knowing statistical vocabulary. Before answering, be sure to read carefully!

T F 1. A univariate data set must contain numerical, not categorical, data.

T F 2. A segmented bar chart uses rectangles rather than circles to display different categories in a set of data.

T F 3. A numerical variable is continuous if it's possible values correspond to isolated points on the number line.

T F 4. A time series plot is a graph of data that has been gathered at different points in time.

T F 5. A pie chart is most useful for numeric data.

T F 6. In a histogram, class intervals are required to be the same length.

T F 7. A bar chart is a graphical display of categorical data.

T F 8. If the upper tail of a distribution stretches out farther than the lower tail, the distribution is negatively skewed.

T F 9. In a scatter plot, both the horizontal and vertical axes must start at zero.

T F 10. An outlier is a data value that is unusually far from the rest of the values in a data set.

T F 11. When using histograms to compare groups of different sizes, one may use either frequencies or relative frequencies for the vertical axis and still be effective.

T F 12. A unimodal set of data is one that contains only one variable.

Chapter 2, Concept Quiz Answer Key

1. F
2. T
3. F
4. T
5. F
6. F
7. T
8. F
9. F
10. T
11. F
12. F

Chapter 2: Graphical Methods for Describing Data Distributions

Section 2.1-2.3

Name _____

1. The article "New Rules Would Protect Students" (*USA Today*, June 16, 2010) reported the percentage of students who received loans to attend college that had defaulted on the loan within 3 years of when the student was scheduled to begin repayment of the loan. Information was given for public colleges, private non-profit colleges, and for-profit colleges.

	Relative Frequency		
Loan Status	**Public Colleges**	**Private Non-profit Colleges**	**For-Profit Colleges**
Good Standing	.928	.953	.833
In Default	.072	.047	.167

a) Construct a comparative bar chart that would allow you to compare loan status for the three types of colleges.

b) The article states "those who attended for-profit schools were more likely to default than those who attended public or private non-profit schools." What aspect of the comparative bar chart supports this statement?

2. The stem-and-leaf plot below displays the weights (in ounces) for a random sample of tomatoes grown on a local farm

Weights of Tomatoes (oz)

```
2L|4
2H|
3L|2
3H|89          stem: ones
4L|13          leaf: tenths
4H|5589
5L|11122334
5H|668999
6L|223344
6H|5556
```

(a) Briefly describe the distribution of tomato weights.

(b) What is the weight of the heaviest tomato in the sample?

3. Knowledge of where animals forage for food is essential for effective wildlife management and conservation. The data below are foraging heights (in m) for a sample of Mountain Chickadees. The data are taken from two different species of fir trees: the Douglas fir and the White fir.

Douglas: 22, 15, 16, 16, 16, 17, 17, 17, 18, 19, 19, 10, 10, 13, 13, 14
6, 7, 7, 8, 8, 9, 3

White: 28, 25, 21, 21, 18, 18, 16, 14, 14, 15, 15, 15, 15, 12, 12, 13
13, 11, 11, 8, 8, 7, 5, 3

(a) Display these data using a comparative (back-to-back) stem-and-leaf plot. Use the stems shown below, and be sure to include all information needed to compare the two distributions.

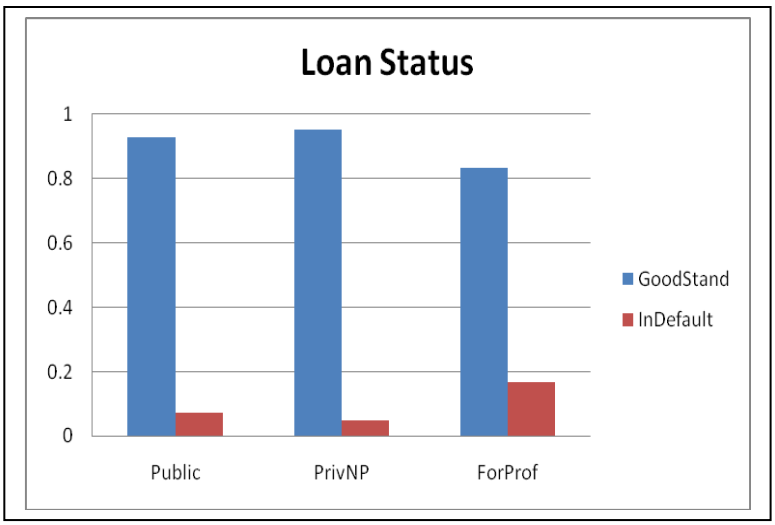
|0L|
|0H|
|1L|
|1H|
|2L|
|2H|

(b) Compare the distributions in part (a). What differences and similarities do you observe in the two distributions?

(c) When Mountain Chicadees are sampled to assess their health, they are captured in mist-nets, which are set to capture the birds at different heights. The mist nets are about 2 meters high, and directions must be given to the persons who set up the nests so that the mist-nets are positioned to capture the most birds. Based on the data from part (a) and your display in part (b), write a short paragraph describing the appropriate mist-net height in a forest region of Douglas firs and for a forest region of White firs.

Chapter 2, Quiz 1, Form A Answer Key

1. (a)



(b) The proportion of defaults for the for-profit schools is greatest, as indicated by the tallest height of the "default bar" for the for-profit schools.

.

2. a) This is moderately skewed in the negative direction, with most of the weight bunched in the higher 5 categories and fewer in the lower 5 weight categories. There is quite a bit of variability in weights.

b) The heaviest tomato is 6.6 ounces.

3. a) **Foraging heights of Mountain Chickadees (*m*)**

Douglas(m)		White(m)
3	0L	3
988776	0H	5788
43300	1L	11223344
9987776665	1H	5555688
2	2L	11
	2H	58

Stem: tens
Leaf: ones

b) The two distributions are similar in being reasonably symmetric. The distribution for Douglas firs is less widely spread out, and the values are generally a little lower than for the White firs.

c) To capture the most birds with a 2m high net, the 2m range where the data is most concentrated must be estimated. In a stand of Douglas pine the most dense part of the distribution is from about 0.5 to 1.5 meters off the ground. In a stand of White pine the most dense part of the distribution appears to be about 1.0 to 2.0 meters off the ground.

Chapter 2: Graphical Methods for Describing Data Distributions

Section 2.1-2.3

Name _____

1. One of the "media messages" sent by television is that appearance is very important. These messages are thought to spur young women to have unrealistic expectations about beauty, thus encouraging women to spend more money on cosmetics and toiletries. During a recent prime-time television season investigators kept track of "appearance" comments. Data on comments by TV males, classified by gender of the comment recipient and type of appearance comment, are given in the table below.

Insults, Compliments, and Neutral Appearance Comments (Males commenting)

Nature of comment	Males about males	Males about females
Compliments	52	106
Insults	41	15
Totals	93	121

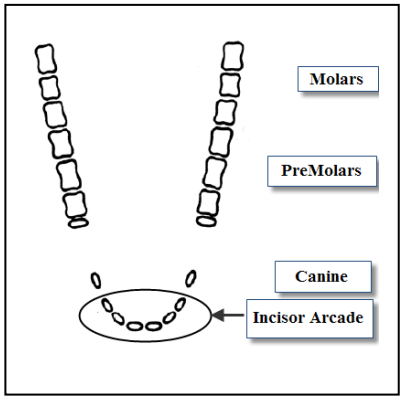
(a) Graphically present these data by constructing a comparative bar chart that would allow you to compare the comments about males and females.

(b) Using the information in the table, write a few sentences about how the comments differ for male recipients and female recipients.

2. Hoofed animals such as cattle and sheep are classified as browsers (leaf eaters) and grazers (grass eaters). Some biologists believe that the incisor arcade of the teeth, diagrammed below, is shaped differently for browsers and grazers, affecting their food selection. The "curvature" of an incisor arcade is a measure of its shape; large values indicate a longer, thinner mouth and small values a shorter, wider mouth. The back-to-back stem-and-leaf plot below compares the curvatures of incisor arcades for different species of browsers and grazers.

Curvature, Browsers vs. Grazers

Browser species $N = 18$		Grazer species $N = 54$
	1L	
	1H	9
	2L	1
	2H	889
4	3L	22244
99885	3H	55566667778999999
410	4L	022223
99987	4H	55556666677778889
321	5L	03
	5H	9
	6L	0
	6H	
	7L	
0	7H	



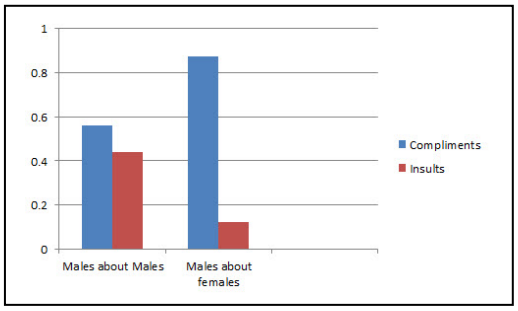
Generic mouth
(Exploded view)

(a) Briefly describe each distribution, noting any unusual features of the data.

(b) Some biologists believe browsers have thinner mouths because leafed plants frequently have thorns and spines. They think that browsers have evolved thinner mouths to get beyond the thorns and spines to the food part of the plant. Are the data in the back-to-back stem-and-leaf plot consistent with this theory? Justify your answer by appealing to specific aspect(s) of the plot.

Chapter 2, Quiz 1, Form B Answer Key

1. a) Comparative bar chart for male comments about males and females:



b) Males made more complimentary comments than insults for both genders. The proportion of compliments is significantly higher for comments that males made about females than for the comments that males made about males.

2. a) The distribution of browsers is centered in the low 40 range. Other than one prominent outlier at 70, the distribution is approximately symmetrical. The distribution of grazers is likewise centered in the low 40 range and is quite symmetrical, but more spread out than the distribution for browsers.

b) The evidence to support this theory is not very strong. The browsers may have a slightly higher average curvature, but the two distributions overlap quite a lot, indicating that the differences are modest at best.

Chapter 2: Graphical Methods for Describing Data Distributions

Section 2.4-2.5

Name _____

1. Artificial bird nests have been used extensively when studying predator behavior. Biologists have assumed predators will find and respond to artificial and natural nests in the same way. In a study to test this assumption the predation of 27 American Robin (*Turdus migratorius*) nests and 27 artificial wicker-basket nests were compared. Artificial nests were placed 30m from existing robin's nests in a randomly chosen direction. The natural nests and artificial nests were checked at 5-day intervals, and the number of nests (out of 27) that had been disturbed by predators in the interval was noted. The results are as follows:

Numbers of real and artificial Robin nests predated

Nest type	5-day check	10-day check	15-day check
Natural	12	17	19
Artificial	3	12	18

(a) Using a single set of axes, construct time series plots for (a) the number of natural nests disturbed by predators and (b) the number of artificial nests disturbed by predators.

(b) Does it appear that the artificial and natural nests have the same pattern of predation through time? Justify your answer by appealing to specific aspects of your plot in part (a).

2. A report in the *Des Moines Register* presented the average number of years of teaching experience for the teachers in elementary schools in Des Moines, Iowa. These data appear in the table below.

School	Average Years Exper.	School	Average Years Exper.	School	Average Years Exper.	School	Average Years Exper.
Adams	13.8	Hanawalt	12.7	McKee	17.5	Phillips	10.5
Capitol View	9.7	Hillis	17	McKinley	5.6	Pleasant Hill	12.8
Cattell	8.5	Howe	14.9	Mitchell	6.2	River Woods	5.9
Douglas	13.8	Hubbell	15.1	Monroe	16.2	Stowe	16.5
Downtown	7.0	Jackson	16.8	Moore	16.6	Studebaker	18.4
Edmunds	8.4	Jefferson	19.2	Morris	9.9	Wallace	4.8
Findley	14.9	King	7.0	Moulton	8.6	Willard	9.7
Garton	9.3	Longfellow	9.8	Oak Park	11.4	Windson	17.7
Granger	17.2	Lovejoy	9.9	Park Avenue	11.4	Woodlawn	17
Greenwood	14	Madison	20.8	Perkins	11.2	Wright	12.9

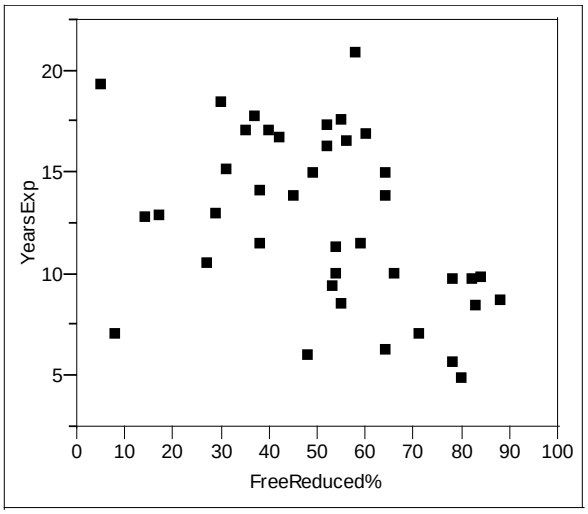
(a) Sketch a histogram for these data, using the following class intervals:

0 to < 4
4 to < 8, etc.

(b) Describe the distribution of the average number of years of teaching experience, and note any unusual features.

3. The *Des Moines Register* article in problem #2 also reported the percent of children eligible for Free and Reduced lunch to teacher experience. The percent of children eligible for Free/Reduced lunch is an indicator of the socio economic status of a neighborhood. The *Register* was concerned that less experienced teachers typically teach students from lower-income families. The accompanying scatter plot displays the data from the article.

YearsExp *vs.* FreeReduced%

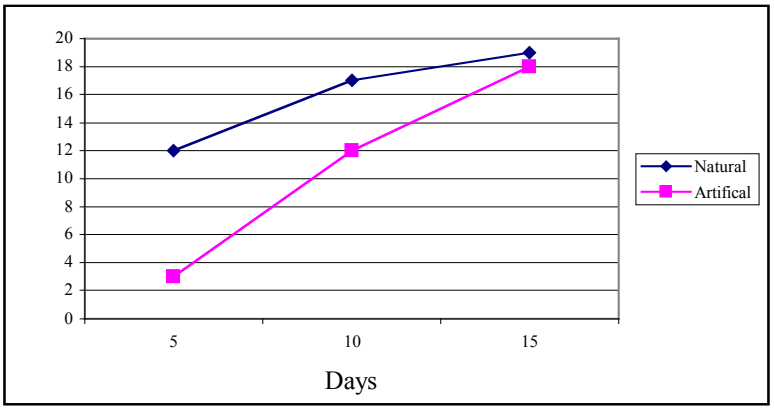


(a) If lesser experienced teachers typically teach students from low-income families, what sort of trend or pattern should appear in the scatter plot?

(b) Does this trend or pattern emerge in these data? Explain, referring to specific aspects of the scatterplot.

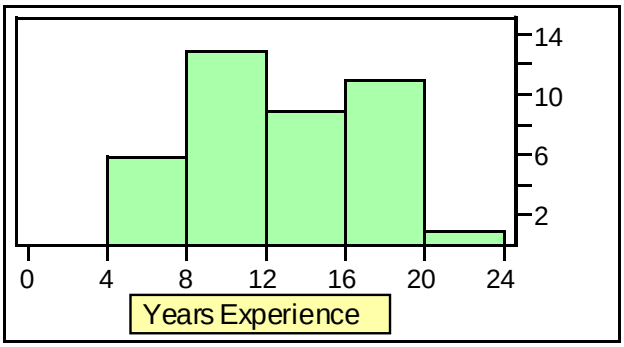
Chapter 2 Quiz 2, Form A Answer Key

1. a)



b) The artificial nests appear to be less likely to disturbed by predators. At 5 days far fewer have been disturbed, but over the course of the 15 days they almost catch up to the rate for the natural nests.

2. a)



b) This is a fairly symmetrical distribution with no apparent outliers. If anything it is perhaps slightly skewed toward the higher values.

3. a) If the newspaper is correct, then school with lower-income families would tend to have a higher rate of free lunches and a lower level of teacher experience. This would show in the graph as a decreasing trend: as the percentage of free lunches increases, the average years of experience should decrease.

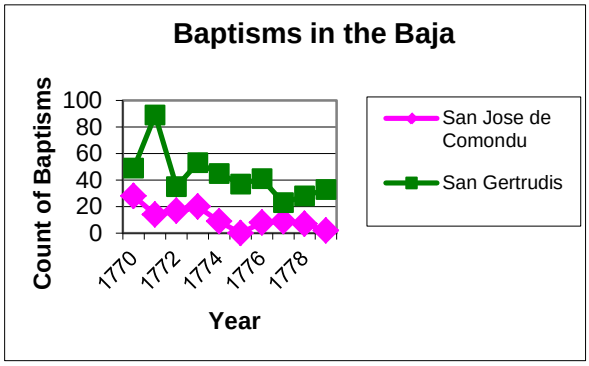
   b) This is shown in the graph. Those schools with more than 75% eligibility for free lunches have less experience (no more than 10 years), while the schools that have less than 75% eligibility have average experience values up to more than 20 years. However, this trend is only moderate at best.

Chapter 2: Graphical Methods for Describing Data Distributions

Section 2.4-2.5

Name _____

1. Information about the Spanish missions of Baja California exist in the form of baptisms and burial records. These records provide an opportunity to study Native American populations under Spanish influence. The graph at right displays two time series of the numbers of baptisms at two Central Baja missions, San Jose de Comondu and San Gertrudis.



In a short paragraph describe and compare the trends in the number of baptisms for the two missions in this time period, noting any interesting features apparent in the plots.

2. A report in the *Des Moines Register* presented on the percentages of students who qualify for Free/Reduced lunch programs in the elementary schools in Des Moines, Iowa. These data appear in the table below.

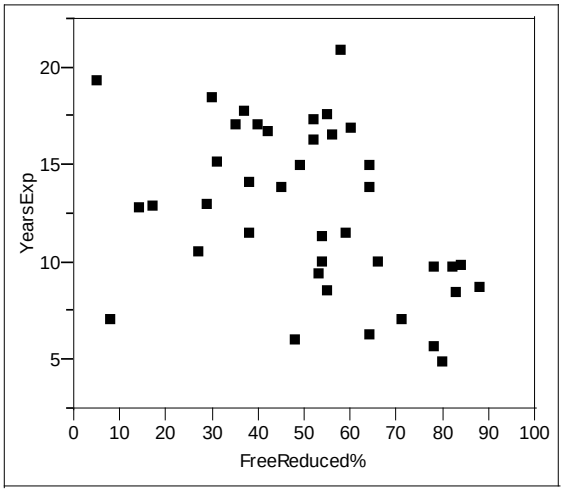
School	% F/R	School	% F/R	School	% F/R	School	% F/R
Adams	64	Hanawalt	14	McKee	55	Phillips	27
Capitol View	78	Hillis	35	McKinley	78	Pleasant Hill	17
Cattell	55	Howe	49	Mitchell	64	River Woods	48
Douglas	45	Hubbell	31	Monroe	52	Stowe	56
Downtown	8	Jackson	60	Moore	42	Studebaker	30
Edmunds	83	Jefferson	5	Morris	54	Wallace	80
Findley	64	King	71	Moulton	88	Willard	82
Garton	53	Longfellow	84	Oak Park	59	Windson	37
Granger	52	Lovejoy	66	Park Avenue	38	Woodlawn	40
Greenwood	38	Madison	58	Perkins	54	Wright	29

(a) Sketch a histogram for these data, using the following class intervals:

0 to < 20
20 to < 40, etc.

(b) Describe the distribution of Free/Reduced eligibility percents, and note any unusual features.

3. The *Des Moines Register* article in problem #2 also related the percent of children eligible for Free and Reduced lunch to teacher experience. The percent of children eligible for Free/Reduced lunch is an indicator of the socio economic status of a neighborhood. The *Register* was concerned that less experienced teachers typically teach students from low income families. The accompanying scatter plot displays the data from the article.



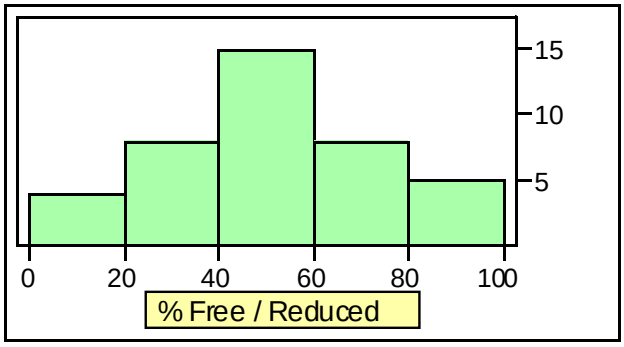
(a) If lesser experienced teachers typically teach students from low-income families, what sort of trend or pattern should appear in the scatter plot?

(b) Does this trend or pattern emerge in these data? Explain, referring to specific aspects of the scatterplot.

Chapter 2, Quiz 2, Form B Answer Key

1. Generally the trends seem to be parallel, except for an apparent greater variability at San Gertrudis in the 1770 – 1773 period. The trends shown in the two plots are roughly parallel except for those initial years of the decade.

2. a)



   b) This distribution is very symmetrical, with no outliers or any skewness. The distribution is centered at about 50 and there is quite a bit of variability in the percentages for the different schools.

3. a) If the newspaper is correct, then school with lower-income families would tend to have a higher rate of free lunches and a lower level of teacher experience. This would show in the graph as a decreasing trend: as the percentage of free lunches increases, the average years of experience should decrease.

   b) This is shown in the graph. Those schools with more than 75% eligibility for free lunches have less experience (no more than 10 years), while the schools that have less than 75% eligibility have average experience values up to more than 20 years. However, this trend is only moderate at best.

Chapter 2: Graphical Methods for Describing Data Distributions

Name _____

1. It is well known that the order of the questions in a survey can affect the responses. This is thought to occur because issues raised in one question are remembered when responding to later questions. In 1948 many people in the United States were concerned about the spread of communism as a political and economic philosophy. A survey given that year contained both of the following questions:

   (i) Do you think the United States should let Communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?

   (ii) Do you think a Communist country like Russia should let American newspaper reporters come in and send back to America the news as they see it?

   The following table contains the percentages of the Yes/No answers to question (i) given by respondents when asked these questions in different orders.

Responses to Communist reporter question in 1948

Question order	% Yes	% No
Question (i) asked first	37	63
Question (ii) asked first	73	27

   (a) Construct a comparative bar chart using these response percentages.

(b) Do you think the order of the questions made a difference in the responses? Explain why or why not in a few sentences.

2. Desert organisms must cope with extreme and variable conditions. Their coping strategies frequently involve retreating to a sub-surface refuge. Researchers in New Mexico, investigating the underground shelter choices by Gila Monsters (*Heloderma suspectum*), hypothesized that depth might be a factor. They reasoned that deeper shelters would be less susceptible to extremes of heat and cold and thus depth would be preferred in a shelter. Depth data for 20 chosen shelters and the 20 nearest available shelters not chosen are presented in the table at right.

Shelter Depths (cm)

Chosen	Nearest	Chosen	Nearest
94	54	33	32
89	6	15	14
32	57	27	7
76	50	27	32
65	46	85	28
30	41	52	52
68	43	80	50
84	59	33	55
10	55	79	69
58	53	71	55

a) Display these data using a comparative (back-to-back) stem-and-leaf plot. Use the stems shown below, and be sure to include all information needed to compare the two distributions.

|0|
|1|
|2|
|3|
|4|
|5|
|6|
|7|
|8|
|9|

(b) Describe the similarities and differences between the chosen and not chosen shelter depths.

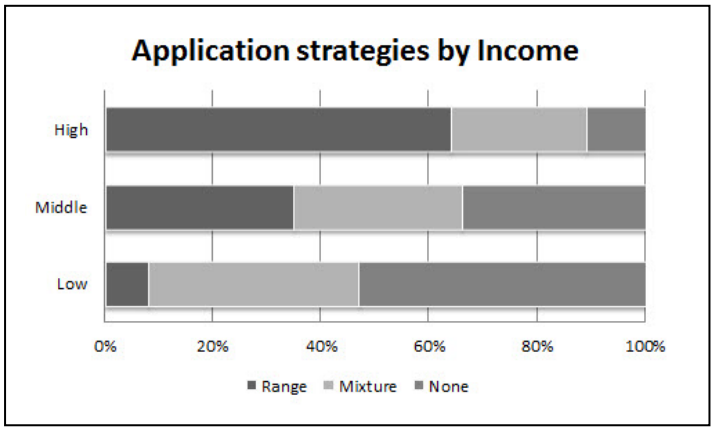
(c) Are these data consistent with the researchers' expectations? Explain, using specific aspect(s) of your plot in part (a).

3. A study by the National Bureau of Economic Research reports that college application strategies differ by income group. The three application strategies identified are: apply to a range of colleges, some selective some not; apply to no selective colleges; or, finally, use a mix of other strategies, such as applying to one selective college and one non-selective college. The percentages of students using these strategies, broken down by income, is shown in the table and the segmented bar chart.

In a few sentences describe how the college application strategies differ for the low, middle, and high income groups.

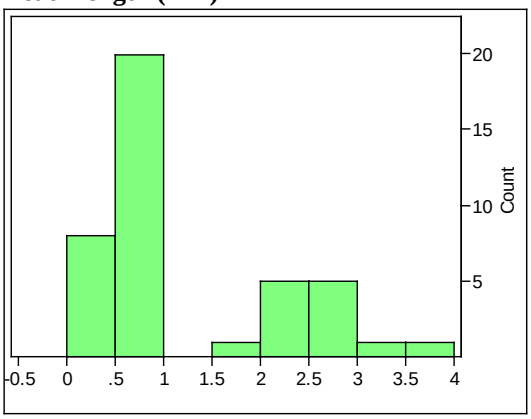
Application strategies by Income (%)

	Range	Mixture	None
Low	8	39	53
Middle	35	31	34
High	64	25	11

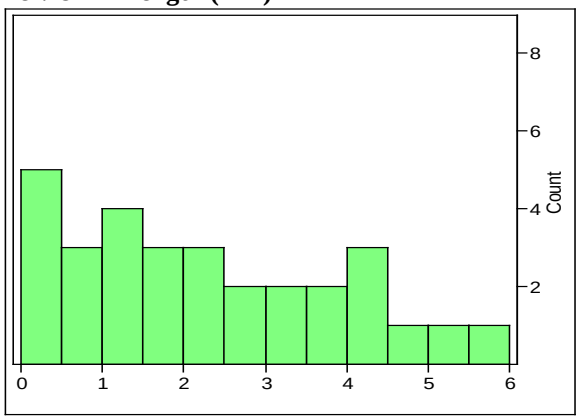


4. Marine biologists have reported the discovery of a new species of fish in the Western North Atlantic. Data on the head length (n=41) and pelvic fin length (n=30) are shown below.

Head Length (mm)



Pelvic Fin Length (mm)

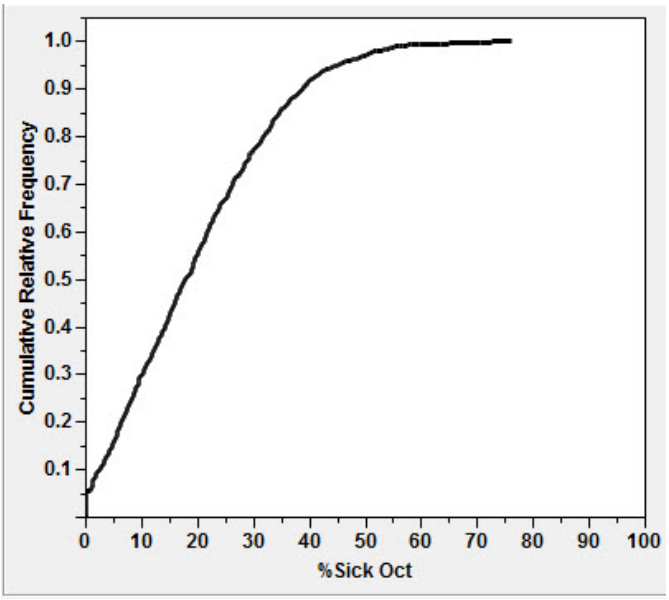


a) Describe the distributions of these two variables using appropriate statistical terms.

b) The distribution of head lengths appears to be rather strange. What do you think might account for the shape of this distribution?

5. During the first six months of the American Civil War, armies were amassed as never before in history. Large groups of men, transported far from home, lived in conditions that bred sickness. In a recent study involving Confederate muster rolls, a researcher constructed a cumulative relative frequency distribution of the percentages of Confederate soldiers reported sick for each company. A "Company" is a unit of soldiers, usually between about 70 and 120 in size. The cumulative relative frequency distribution for October, 1861, is shown below.

Cumulative Relative Frequency Distribution
Percentage reported sick by Company
October, 1861



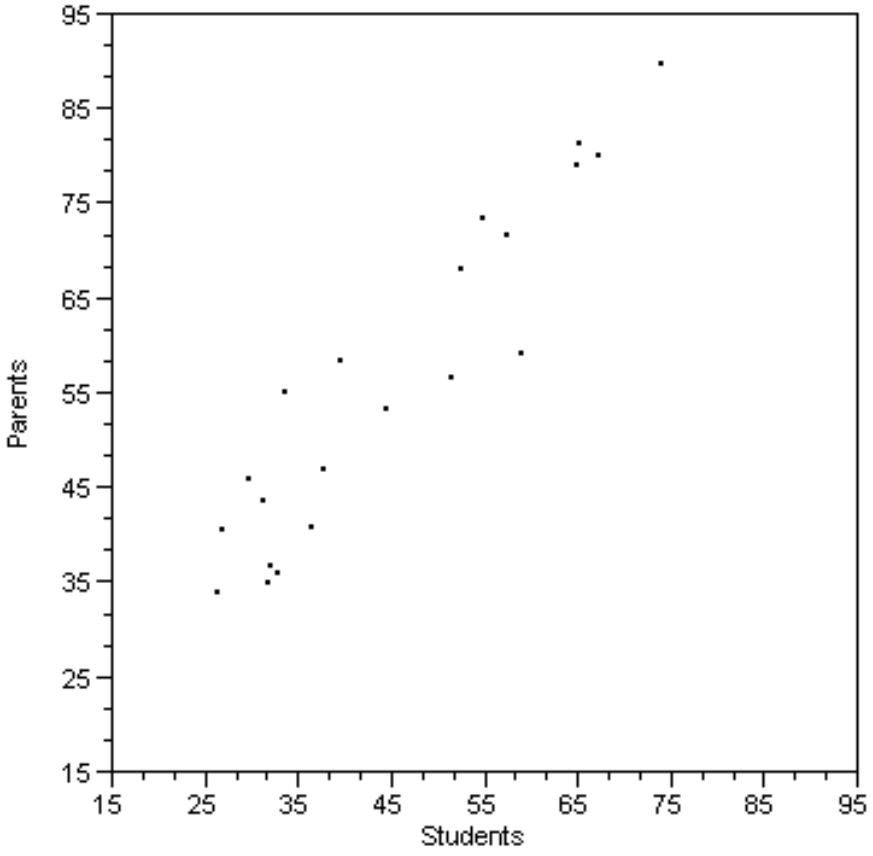
(a) Approximately what proportion of the Confederate companies had less than 10% sick?

(b) Approximately what proportion of the Confederate companies had more than 30% sick?

(c) Approximately 50% of the companies had less than what percentage of soldiers sick?

6. Investigators were interested in communication about sex and related topics between parents and 12-14 year olds. The investigators asked parents and students whether or not they had communicated with their child/parent about topics such as pregnancy, how to say "no," and methods of contraception. A scatter plot of their data is presented below. Each point represents one topic listed by the investigators. For example, 72.4% of parents and 57.6% of adolescents said the parents had discussed childbirth. This would be recorded as the point (57.6, 72.4).

Parents vs. Adolescents reporting of Discussion:
Sex Education Topics (% Yes)

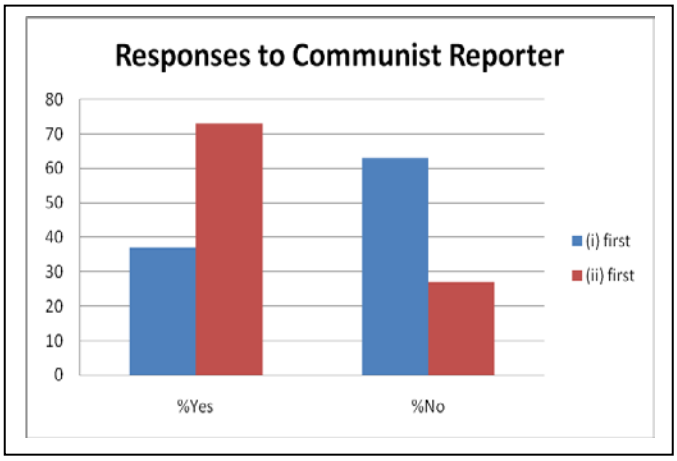


(a) On the graph above, sketch the line $y = x$, representing a line of "perfect agreement" between the parents and adolescents.

(b) What is a possible explanation of the placement of the points in the scatter plot relative to the line you sketched in part (a)?

Chapter 2, Test Form A Answer Key

1. a)



b) Yes, the order seems to have made a difference. The rate of "yes" responses was very nearly doubled if the question was asked after the one about American reporters in Russia.

2. a)

Depths (cm) for Chosen and Non chosen Shelters

```
Chosen          Non-chosen

    |0|67
  50|1|4
  77|2|8                    Stems: tens
3320|3|22                   Leafs: ones
    |4|136
  82|5|0023455579
  85|6|9
 961|7|
9540|8|
   4|9|
```

b) The chosen shelters are more variable in depth, and appear to be centered at about 50 cm. The non-chosen shelter depths are less variable, are skewed to the low side, and tend to not be as deep as the chosen shelters.

c) The researchers' expectations appear to be born out in these data. The shelters chosen by the Gila Monsters tend to be deeper than the potential shelters not chosen. As an example, 60 cm is a typical length of the chosen shelters, but is close to the maximum depth of the non chosen shelters.

3. The most striking difference is that a very low fraction of low income students use a range of strategies, whereas a very large fraction of high income students use a range of strategies. The middle income students are somewhere in between.

4. a) The distribution of head length is bimodal, with clumps centered at .5mm and 2.5mm, and a gap between these two clumps. The pelvic fin length is skewed to the right over a range from 0 to 6 mm.

   b) Generally one would not expect bimodal measurements unless perhaps there were a mixture of two subpopulations. In this case perhaps males are larger or smaller, or there may be two subspecies of fish that make up the identified species.

5. a) Approximately 25%

   b) Approximately 20%

   c) Approximately 20%

6. a) Draw diagonal line from (15,15) at lower left to (95,95) at upper right.

   b) All the points are above the line, indicating that for all topics, parents are generally saying that they have talked about the topics more so than indicated by the adolescents' response.

Chapter 2: Graphical Methods for Describing Data Distributions

Name _____

1. It is well known that the order of the questions in a survey can affect the responses. This is thought to occur because issues raised in one question are remembered when responding to later questions. In 1974 a survey asked the following questions about doctors and lawyers:

    1. Would you say that most doctors in this country are really interested in the public good, or are most doctors just out to make a lot of money?

    2. Would you say that most lawyers in this country are really interested in the public good, or are most lawyers just out to make a lot of money?

    The following table summarizes the responses to the lawyer question when these questions were asked in different orders.

Responses to lawyer question

Question order	% Public good	% Making money
Lawyer question asked first	26	74
Doctor question asked first	30	70

(a) Construct a comparative bar chart for these response percentages.

(b) Do you think the order of the questions made a difference in the responses? Explain.

2. Desert organisms must cope with extreme and variable conditions. Their coping strategies frequently involve retreating to a sub-surface refuge. Researchers in New Mexico, investigating the underground shelter choices by Gila Monsters (*Heloderma suspectum*), hypothesized that depth might be a factor. They reasoned that deeper shelters would be less susceptible to extremes of heat and cold and thus depth would be preferred in a shelter. Depth data for 20 chosen shelters and 20 available shelters not chosen are presented in the table at right.

Shelter Depths (cm)

Chosen	Nearest	Chosen	Nearest
94	54	33	32
89	6	15	14
32	57	27	7
76	50	27	32
65	46	85	28
30	41	52	52
68	43	80	50
84	59	33	55
10	55	79	69
58	53	71	55

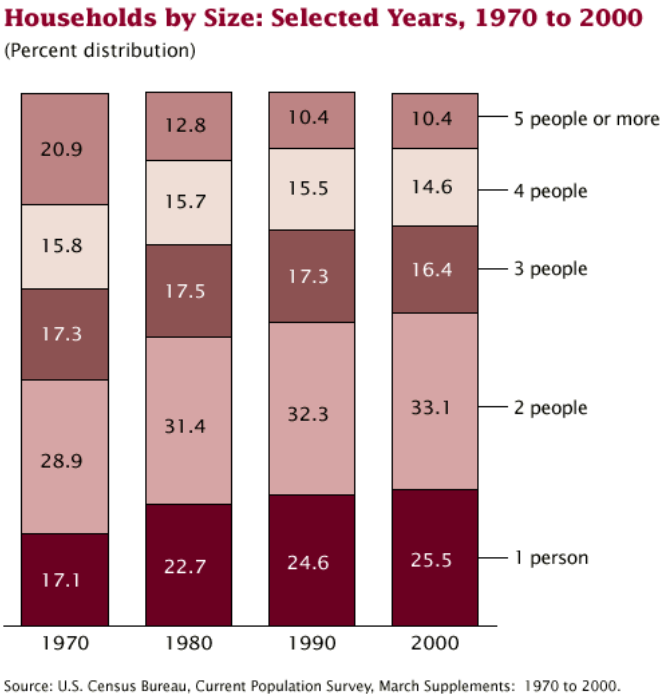
a) Display these data using a comparative (back-to-back) stem-and-leaf plot. Use the stems shown below, and be sure to include all information needed to compare the two distributions.

|0|
|1|
|2|
|3|
|4|
|5|
|6|
|7|
|8|
|9|

(b) Describe the similarities and differences between the chosen and not chosen shelter depths.

(c) Are these data consistent with the researchers' expectations? Explain, using specific aspect(s) of your plot in part (a).

3. As part of the United States Census, data is collected on the number of persons in each household. The census data for four decades are summarized below.



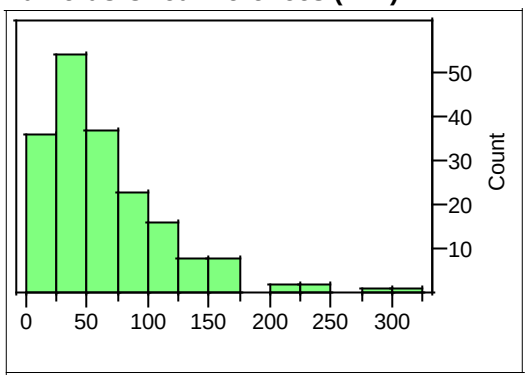
(a) In a few sentences describe how the proportion of households with 4 people has changed from 1970 to 2000.

(b) What size of household appears to have decreased the most from 1970 to 2000?

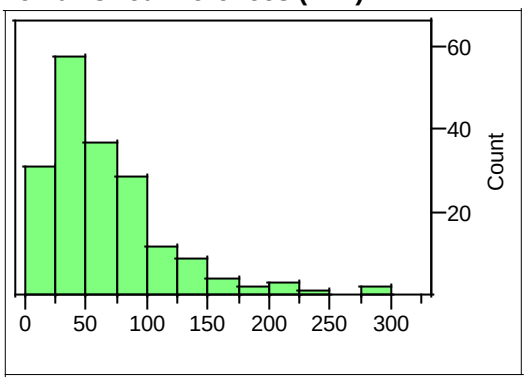
4. The average circumferences (distances around) in *mm* of the humerus, femur, and tibia of different species of mammals are displayed below. These long bones primarily support the body mass in mammals. Each data value is the mean circumference for the available measures for a single species.

(a) Describe the differences and similarities in the distributions of these different bones.

Humerus Circumferences (*mm*)

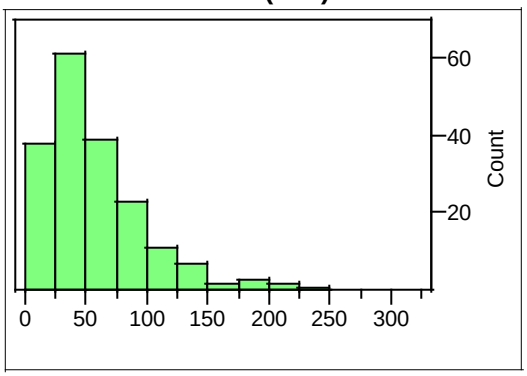


Femur Circumferences (*mm*)



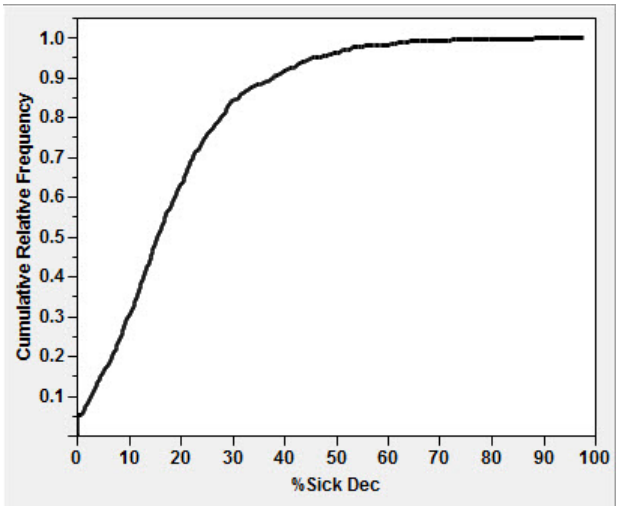
(b) Biologists speculate that greater weights might be supported by larger bones. Based on the histograms, if the biologists' speculations are correct, which of these bones (humerus, femur, or tibia) would seem to be the least able to support a large weight? Explain, referring to specific aspects of the histograms.

Tibia Circumference (*mm*)



5. During the first six months of the American Civil War, armies were amassed as never before in history. Large groups of men, transported far from home, lived in conditions that bred sickness. In a recent study involving Confederate muster rolls, a researcher constructed a cumulative relative frequency distribution of the percentages of Confederate soldiers reported sick for each company. A "Company" is a unit of soldiers, usually between about 70 and 120 in size. The cumulative relative frequency distribution for December, 1861, is shown below.

Cumulative Relative Frequency Distribution
Percentage reported sick by Company
December, 1861



(a) Approximately what proportion of the Confederate companies had less than 10% sick?

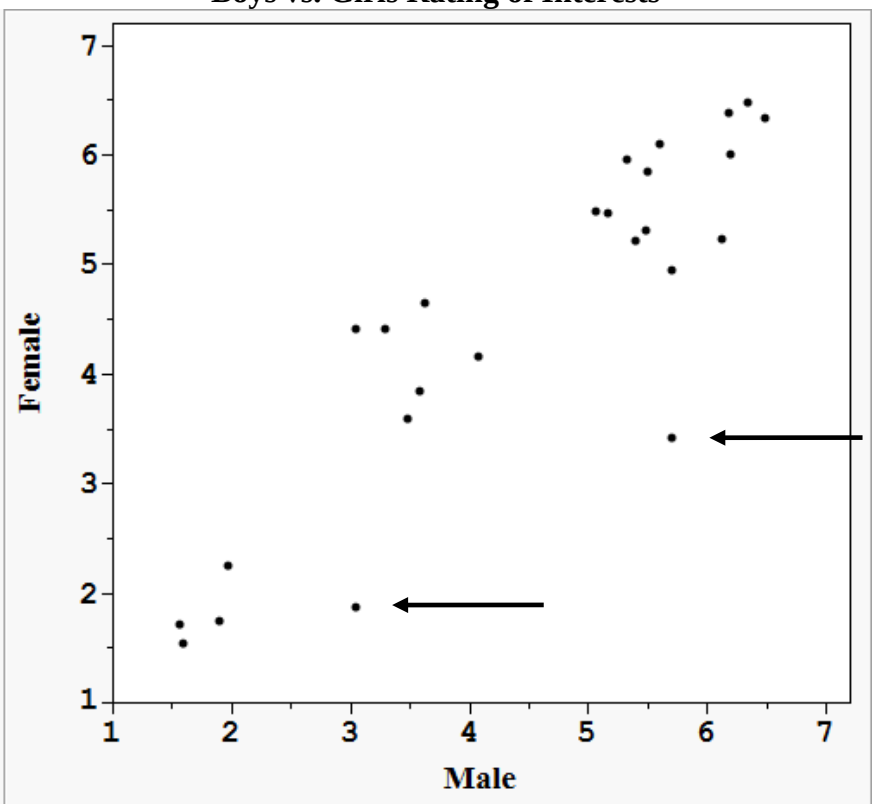
(b) Approximately what proportion of the Confederate companies had more than 30% sick?

(c) Approximately 50% of the companies had less than what percentage of soldiers sick?

6. Investigators studying the relationship between gender and the interests of early adolescents asked boys and girls in grades 5-9 to indicate on a 7-point scale their degree of interest in topics such as Life, Sports, and Cars.

The data in the scatter plot represent average levels of interest for boys and girls in grades 5 - 9. Each point represents one topic listed by the investigators. For example, the boys' average rating for "Opposite sex" was 6.2, while girls on average rated their interest in this topic as 6.0 on the 7 point interest scale. This is recorded as the point (6.2, 6.0).

Boys vs. Girls Rating of Interests



There are three interesting features of this scatter plot that you are to interpret below. What do you infer from each of these features of the scatterplot?

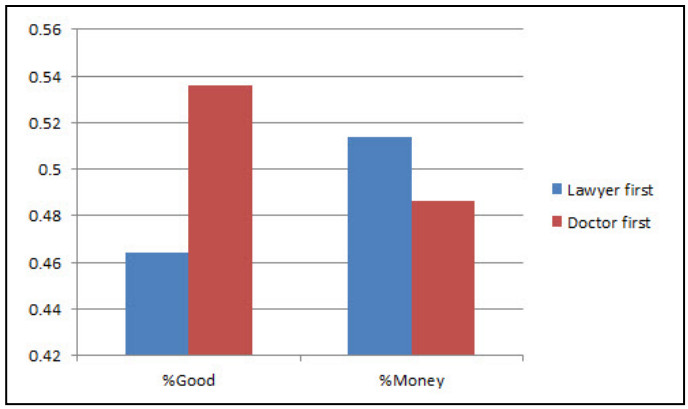
(a) The points generally seem to scatter around the line $y = x$.

(b) There seem to be three clusters of points, one in the lower left, one in the middle, and one in the upper right of the graph.

(c) There are two points (as indicated by the arrows) that do not seem to fit the overall pattern.

Chapter 2, Test Form B Answer Key

1. a)



b) There is only a small difference that is shown in this data, from 26% to 30%. The order of the questions has not made much of a difference in the responses to the question in this case.

2. a)

Depths (cm) for Chosen and Non chosen Shelters

```
Chosen          Non-chosen

    |0|67
  50|1|4
  77|2|8                    Stems: tens
3320|3|22                   Leafs: ones
    |4|136
  82|5|0023455579
  85|6|9
 961|7|
9540|8|
   4|9|
```

b) The chosen shelters are more variable in depth, and appear to be centered at about 50 cm. The non-chosen shelter depths are less variable, are skewed to the low side, and tend to not be as deep as the chosen shelters.

c) The researchers' expectations appear to be born out in these data. The shelters chosen by the Gila Monsters tend to be deeper than the potential shelters not chosen. As an example, 60 cm is a typical length of the chosen shelters, but is close to the maximum depth of the non chosen shelters.

3. a) The proportion of households with 4 people has stayed about the same, with only as slight decrease from 15.8% of households to 14.6%.

   b) The largest households, with 5 or more members decreased sharply, from 20.9% to 10.4% during the 30 years. This is a decrease of more than 50%. The other households either grew in size or only decreased slightly.

4. a) The shapes of these distributions are very similar, with a peak between 25 and 50 and positively skewed. The humerus and femur extend a little higher than the tibia, to about 300 vs. 200.

   b) The differences are not great, since the large parts of the distributions are very similar. Perhaps for the largest animals the humerus and femur support more of the weight, indicated by the fact that these distributions extend to higher values than does the tibia distribution.

5. (a) Approximately 30%

   (b) Approximately 20%

   (c) Approximately 15%

6. (a) This indicates that both boys and girls have similar ratings.

   (b) Answers will vary. One reasonable response is that the respondents tended to have either strong views or no opinion.

   (c) The two points (which happen to be motorcycles and war) are points that are associated with topics of much greater interest to boys than girls.