

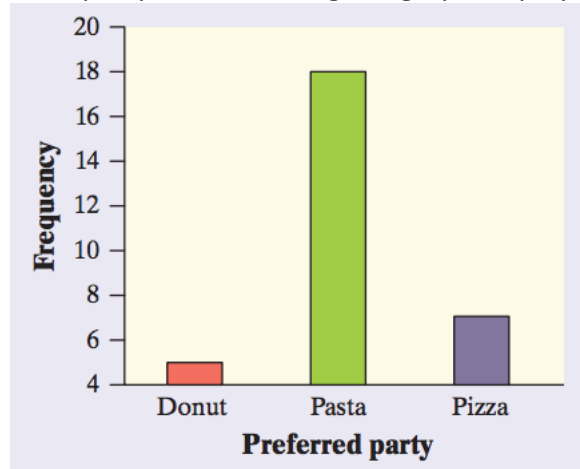


Whittier Tech

Grade 12:
Honors Statistics

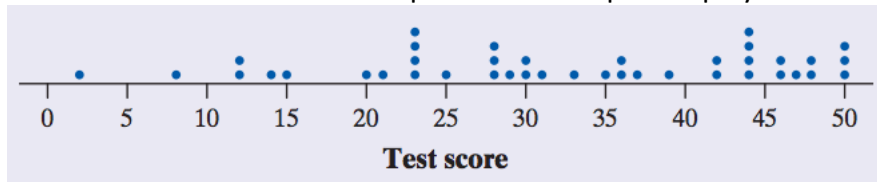
Section 1: Multiple Choice *Select the best answer for each question.*

1. You record the age, marital status, and earned income of a sample of 1463 women. The number and type of variables you have recorded are
 - a. 3 quantitative, 0 categorical.
 - b. 4 quantitative, 0 categorical.
 - c. 3 quantitative, 1 categorical.
 - d. 2 quantitative, 1 categorical.
 - e. 1 quantitative, 2 categorical.
2. The students in Mr. Tyson's high school statistics class were recently asked if they would prefer a pasta party, a pizza party, or a donut party. The following bar graph displays the data.



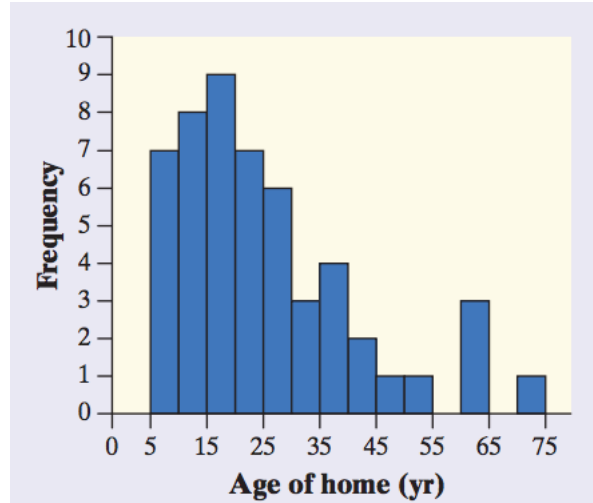
This graph is misleading because

- a. it should be a histogram, not a bar graph.
 - b. there should not be gaps between the bars.
 - c. the bars should be arranged in decreasing order by height.
 - d. the vertical axis scale should start at 0.
 - e. preferred party should be on the vertical axis and number of students should be on the horizontal axis.
3. Forty students took a statistics test worth 50 points. The dotplot displays the data. The third quartile is



- a. 45.
- b. 44.
- c. 43.
- d. 32.
- e. 23.

Questions 4-6 refer to the following setting. Realtors collect data in order to serve their clients more effectively. In a recent week, data on the age of all homes sold in a particular area were collected and displayed in this histogram.



4. Which of the following could be the median age?
 - a. 19 years
 - b. 24 years
 - c. 29 years
 - d. 34 years
 - e. 39 years

5. Which of the following is most likely true?
 - a. mean > median, range < IQR
 - b. mean < median, range < IQR
 - c. mean > median, range > IQR
 - d. mean < median, range > IQR
 - e. mean = median, range > IQR

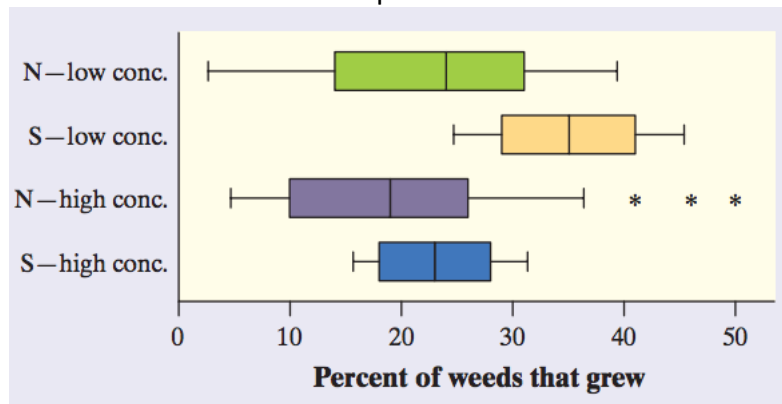
6. The standard deviation of the distribution of house age is about 16 years. Interpret this value.
 - a. The age of all houses in the sample is within 16 years of the mean.
 - b. The gap between the youngest and oldest house is 16 years.
 - c. The age of all the houses in the sample is 16 years from the mean.
 - d. The gap between the first quartile and the third quartile is 16 years.
 - e. The age of the houses in the sample typically varies by about 16 years from the mean age.

7. The mean salary of all female workers is \$35,000. The mean salary of all male workers is \$41,000. What must be true about the mean salary of all workers?
 - a. It must be \$38,000.
 - b. It must be larger than the median salary.
 - c. It could be any number between \$35,000 and \$41,000.
 - d. It must be larger than \$38,000.
 - e. It cannot be larger than \$40,000.

Questions 8 and 9 refer to the following setting. A survey was designed to study how business operations vary by size. Companies were classified as small, median, or large. Questionnaires were sent to 200 randomly selected businesses of each size. Because not all questionnaires are returned, researchers decided to investigate the relationship between the response rate and the size of the business. The data are given in the following two-way table.

		Business size		
		Small	Medium	Large
Response?	Yes	125	81	40
	No	75	119	160

8. What percent of all small companies receiving questionnaires responded?
- a. 12.5% b. 20.8% c. 33.3%
d. 50.8% e. 62.5%
9. Which of the following conclusions seems to be supported by the data?
- a. There are more small companies than large companies in the survey.
b. Small companies appear to have a higher response rate than medium or big companies.
c. Exactly the same number of companies responded as didn't respond.
d. Overall, more than half of companies responded to the survey.
e. If we combined the medium and large companies, then their response rate would be equal to that of the small companies.
10. An experiment was conducted to investigate the effect of a new weed killer to prevent weed growth in onion crops. Two chemicals were used: the standard weed killer (S) and the new chemical (N). Both chemicals were tested at high and low concentrations on 50 test plots. The percent of weeks that grew in each plot was recorded. Here are some boxplots of the results.



- Which of the following is *not* a correct statement about the results of this experiment?
- a. At both high and low concentrations, the new chemical results in better weed control than the standard weed killer.
b. For both chemicals, a smaller percentage of weeds typically grew at higher concentrations than at lower concentrations.
c. The results for the standard weed killer are less variable than those for the new chemical.
d. High and low concentrations of either chemical have approximately the same effects on weed growth.
e. Some of the results for the low concentration of weed killer show a smaller percentage of weeds growing than some of the results for the high concentration.

Section 2: Free Response Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

11. You are interested in how many contacts older adults have in their smartphones. Here are the data on the number of contacts for a random sample of 30 elderly adults with smartphones in a large city.

7	20	24	25	25	28	28	30	32	35
42	43	44	45	46	47	48	48	50	51
72	75	77	78	79	83	87	88	135	151

- Construct a histogram of these data.
 - Are there any outliers? Justify your answer.
 - Would it be better to use the mean and standard deviation or the median and IQR to describe the center and variability of this distribution? Why?
12. A study among the Pima Indians of Arizona investigated the relationship between a mother's diabetic status and the number of birth defects in her children. The results appear in the two-way table.

		Diabetic status		
		Nondiabetic	Prediabetic	Diabetic
Number of birth defects	None	754	362	38
	One or more	31	13	9

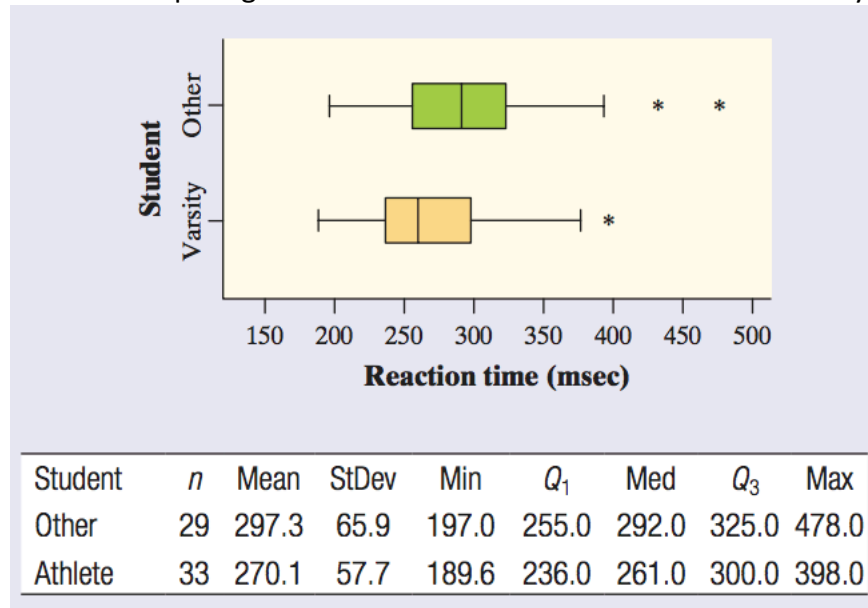
- What proportion of the women in this study had a child with one or more birth defects?
 - What percent of the women in this study were diabetic or prediabetic, and had a child with one or more birth defects?
 - Make a segmented bar graph to display the distribution of number of birth defects for the women with each of the three diabetic statuses.
 - Describe the nature of the association between mother's diabetic status and the number of birth defects for the women in this study.
13. The back-to-back stemplot shows the lifetimes of several Brand X and Brand Y batteries.

Brand X		Brand Y
	1	
	1	7
	2	2
	2	6
2 1 1 0	3	
9 9 7 7 5	3	
3 2 2 1	4	2 2 3 3 3 4
	4	5 6 8 8 9
4	5	0
5	5	

Key: 4|2 represents 420–429 hours.

- What is the longest that any battery lasted?
- Give a reason someone might prefer a Brand X battery.
- Give a reason someone might prefer a Brand Y battery.

14. Catherine and Ana suspect that athletes (i.e., students who have been on at least one varsity team) typically have a faster reaction time than other students. To test this theory, they have an online reflex test to 33 varsity athletes at their school and 20 other students. Here are parallel boxplots and numerical summaries of the data on reaction times (in milliseconds) for the two groups of students. Write a few sentences comparing the distribution of reaction time for the two types of students.

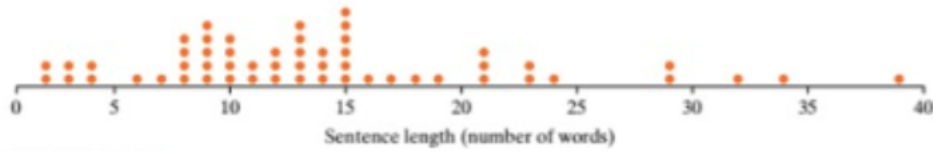


Section 3: Preliminary Exercises

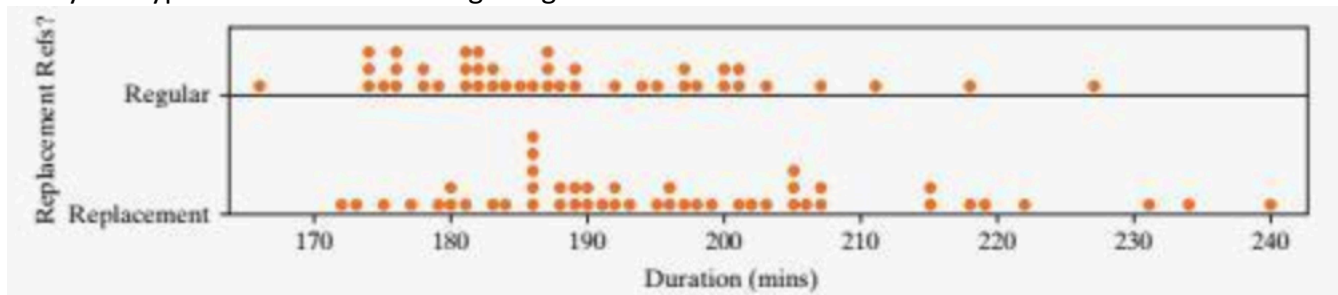
15. For each of the following research questions identify the observational units and variable(s).
- An article in a 2006 issue of the *Journal of Behavioral Decision Making* reports on a study involving 47 undergraduate students at Harvard. All of the participants were given \$50, but some (chosen at random) were told that this was a “tuition rebate,” while others were told that this was “bonus income.” After one week, the students were contacted again and asked how much of the \$50 they had spent and how much they had saved. Researchers wanted to know whether those receiving the “rebate” would tend to save more money than those receiving the “bonus.”
 - How much did a typical American consumer spend on Christmas presents in 2012?
 - Do college students who pull all-nighters tend to have lower grade point averages than those who do not pull all-nighters?
 - Is the residence situation of a college student (on-campus, off-campus with parents, off-campus without parents) related to how much alcohol the student consumes in a typical week?
 - Can you predict how far a cat can jump based on factors such as its length?
16. Refer back to #15. For each of the research questions (i-v), identify which variables are quantitative and which categorical. Also classify categorical variables according to whether they only have two possible outcomes (sometimes called a binary variable) or more than two possible outcomes.
17. Consider the students in your class as the observational units in a study.
- Identify three quantitative variables that could be recorded on these students.
 - Identify three categorical variables that could be recorded on these students.

18. A famous study titled “The Effect of Deliberative Practice in the Acquisition of Expert Performance,” published in *Psychological Review* (Ericsson, Krampe, and Tesch-Romer, 1993) led to the now-conventional wisdom that 10,000 hours of deliberate practice are necessary to achieve expert performance in skills such as music and sports. The researchers in this study asked violin students at the Music Academy of West Berlin to keep a diary indicating how they spent their time. The researchers also asked the students’ professors to indicate which were the top students with the potential for careers as international soloists, which were good violinists but not among the best, and which were studying to become music teachers. Researchers found that those in the top two groups devoted much more time per week to individual practice than did those in the group studying to become music teachers.
- Identify the observational units in this study.
 - Identify the quantitative variable in this study.
 - Identify the categorical variable in this study.
19. A recent study investigated self-reported anxiety levels immediately before a skydive among 11 first-time skydivers and 13 experienced skydivers (at least 30 jumps) recruited from a parachute center in northern England (Hare et al., 2013). The researchers found that anxiety levels were substantially higher (average anxiety score of 43, higher means more anxiety) among the first-time skydivers as compared to the experienced skydivers (average anxiety score of 27) on a standard test for anxiety (Spielberger State-Trait Anxiety Inventory). Identify the six steps of a statistical investigation in this study by answering the following questions about this study.
- Step 1: Ask a research question.** What is the research question being investigated in this study?
 - Step 2: Design a study and collect data.**
 - What are the observational units in the study?
 - How were the observational units obtained?
 - What are the variables in this study? Classify each as categorical or quantitative.
 - Step 3: Explore the data.**
 - What are the average anxiety scores in the two groups?
 - How many individuals are in each group?
 - Step 4: Draw inferences beyond the data.** What is the researchers’ conclusion about the difference in anxiety levels between the two groups?
 - Step 5: Formulate conclusions.** In order to make their conclusions more relevant, what are the researchers hoping about the 24 skydivers in this study compared to all skydivers in general?
 - Step 6: Look back and ahead.** The average age of novice skydivers was 21 and the average age of experienced skydivers was 28. Explain how this could impact the conclusions of the researchers. How could this be addressed in a future study?
20. Suppose that Nellie records the high temperature every day for one year in New York City, while Sandy does the same in San Diego.
- Who would you expect to have the larger *mean* of these temperatures? Explain.
 - Who would you expect to have the larger *standard deviation* of these temperatures? Explain.
21. Consider three students with the following distributions of 24 quiz scores:
- Amanda: 8,8
- Barney: 5, 6,6,6,6,7,7,7,7,7,7,8,8,8,8,8,8,9,9,9,9,10
- Charlene: 0,0,0,0,0,0,0,0,0,0,0,0,10,10,10,10,10,10,10,10,10,10,10,10
- Answer the following without bothering to do any calculations. Explain your choices.
- Which student has the *smallest* standard deviation of quiz scores?
 - Which student has the *largest* standard deviation of quiz scores?

22. The dotplot displays the distribution of sentence lengths (number of words in a sentence) for 55 sentences selected from John Grisham’s novel *The Confession*. Describe what this graph reveals, paying attention to the shape, center, variability, and unusual observations. Write as if you are describing this distribution to someone with no knowledge about sentence lengths.



23. In the 2012 National Football League (NFL) season, the first three week’s games were played with replacement referees because of a labor dispute between the NFL and its regular referees. Many fans and players were concerned with the quality of the replacement referees’ performance. We could examine whether data might reveal any differences between the three weeks’ games played with replacement referees and the next three week’s games that were played with regular referees. For example, did games generally take less or more time to play with replacement referees than with regular referees? The dotplots below display data about the duration of games (in minutes), separate by the type of referees officiating the game.



- What proportion of the 48 games officiated by replacement referees lasted for at least 3.5 hours (210 minutes)? What proportion of the 43 games officiated by regular referees lasted for this long?
 - What proportion of the 48 games officiated by replacement referees lasted for less than 3 hours (180 minutes)? What proportion of the 43 games officiated by regular referees lasted for this long?
 - Would you say that either type of referee tended to have longer games than the other on average? If so, which type of referee tended to have longer games and by about how much on average?
 - Would you say that either type of referee tended to have more variability in game durations? If so, which type of referee tended to have more variability?
 - Write a paragraph in which you compare the distributions of game durations between the two types of referees. Be sure to comment on shape, center, variability, and unusual observations.
24. It has been reported that the probability that a new business closes or changes owners within its first three years is about 0.60. Interpret what this probability means, beginning with: “About 60% of...”

25. An ice cream shop offers a discount price on a small ice cream cone- the price is obtained by rolling two dice and then taking the larger number followed by the smaller number to be the price (in cents). So, rolling a 3 and a 2 or a 2 and a 3 would result in a price of 32 cents. If I visit the ice cream shop with only 50 cents in my pocket, the probability that the ice cream cone costs no more than 50 cents is $\frac{4}{9}$, which is about 0.444. How would you interpret this probability? (Circle all that apply.)
- A. If I go to this ice cream shop a very large number of times with only 50 cents each time, then I will be able to afford the ice cream cone in about 44.4% of all such visits.
 - B. I will be able to afford the ice cream cone in four of my next nine visits to this shop if I always enter with two quarters.
 - C. Among all customers who enter this ice cream shop tomorrow, about $\frac{4}{9}$ will be able to afford the ice cream cone.
26. Which of the following numbers *cannot* represent a probability? Circle all that apply.
- a. 0.01
 - b. -.50
 - c. 1.05
 - d. 88%
 - e. 107%
27. Suppose that baseball Team A is better than baseball Team B. Team A is enough better that it has a $\frac{2}{3}$ probability of beating Team B in any one game, and this probability remains the same for each game, regardless of the outcomes of previous games.
- a. Explain what it means to say that Team A has a $\frac{2}{3}$ probability of beating Team B in any one game?
 - b. If Team A plays Team B for 3 games, is Team A guaranteed to win exactly twice?
 - c. If Team A plays Team B for 30 games, is Team A guaranteed to win exactly 20 times?
 - d. If Team A plays Team B for 30 games, do you think it's very unlikely that Team A will win exactly 20 times? Explain.