1. Don’t even open this until you are told to do so.

2. Remember to turn your phone off now.

3. Please turn your hats around backwards or take them off.

4. Please put your backpack and other things along the walls or at the front of the room.

5. You need a gray, 81/2 × 11” scantron, pencil, calculator and you may have 5 sheets of notes.

6. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers clearly. Multiple marks will be counted wrong.

7. You will have 60 minutes to finish this exam.

8. If you have questions, please write out what you are thinking on the back of the page so that we can discuss it after I return it to you.

9. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of zero on the exam. You must work alone.

10. When you are finished please make sure you have filled in your name and marked your FORM (A, B, C or D) and 20 answers, then turn in JUST your scantron.

11. Good luck!
1. If we plotted the distance to each student’s hometown (College Station to home in miles), the distribution would be

A. bell-shaped — a few students live here in CS, a few in other states, most here in Texas
B. skewed right – only a very few live really far away
C. skewed left – only a few live in CS
D. skewed right – the few living in CS would be outliers
E. bimodal – a peak for those in Texas and a peak for those outside of Texas

2. Which of the following is true about the plot above?
There are 50 observations.

A. The median is between 15 and 17.5 (the tallest bin).
B. The minimum is 0 (the starting point of the horizontal axis).
C. The mean is smaller than the median because the distribution is skewed to the right.
D. Two of the above are true.
E. None of the above are totally true.

3. In College Station, most days we get 0 inches of rain, some days we get a moderate amount of rain, and very rarely do we get a lot of rain. If we wanted to report the center of College Station’s rain distribution, which statistic should we use?

A. The mean, because the distribution of rainfall values is symmetric.
B. The median, because the distribution of rainfall values is symmetric.
C. The mean, because it is in the middle of the distribution of rainfall values.
D. The median, because it is in the middle of the distribution of rainfall values.
E. The mean, because it takes into account all the values of rainfall.

4. Students often tell me that they marked the correct answer and then changed it. Is this good evidence that you should never change your answers?

A. Yes, if you’re just guessing, your first guess is always the best.
B. Yes, past experience is a good indicator of future happenings.
C. No, this is just anecdotal evidence.
D. No, you don’t know how good those students were, so you could be different.
E. No, you should always trust your own judgement.

5. Joe is in STAT303501. He made an 82 on his exam that had an average of 75 with a standard deviation of 6. Jane is in STAT303502 and made a 79. Her class average was 68 with a standard deviation of 8. Who did better relative to the rest of their class?

A. Joe because 82 is better than 79.
B. Joe because his class had a better average.
C. Jane because her z-score is higher.
D. Jane because she has more points above her class mean (11 vs. 7).
E. We can’t compare them because we don’t know the shape of either distribution of grades.

6. What would happen if the two classes had the same standard deviation?

A. Jane would be better only if the standard deviation was 8.
B. Jane would be better only if the standard deviation was 6.
C. We need to know what the standard deviation is to determine who would be better.
D. Again, we can’t compare the two.
E. It doesn’t matter what the standard deviation is, Jane would be better.

7. A random sample of Americans were asked how many hours they slept per night. The standard deviation was 2.2. Interpret this standard deviation in context.

A. The typical difference between the number of hours a person slept and the average is about 2.2.
B. About 2.2% of Americans had a number of hours of sleep that was more than 3 standard deviations above or below average.
C. The difference between the third and first quartiles of the number of hours a person slept is 2.2.
D. The average number of hours Americans slept is about 2.2.
E. The difference between the largest amount of sleep in the sample and the smallest is 2.2.
8. The General Social Survey asked randomly selected Americans, “Engineers don’t get as much fun out of life as other people do. Do you strongly agree, agree, disagree, or strongly disagree?” The table below classifies responses by gender.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>11</td>
<td>29</td>
<td>146</td>
<td>22</td>
<td>208</td>
</tr>
<tr>
<td>Female</td>
<td>4</td>
<td>33</td>
<td>147</td>
<td>24</td>
<td>208</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>62</td>
<td>293</td>
<td>46</td>
<td>416</td>
</tr>
</tbody>
</table>

How likely is a male who disagrees?

A. 146/147
B. 146/208
C. 146/293
D. 146/416
E. 208/293

9. The likelihood of an American agreeing is

A. 62/416
B. 208/416
C. $29/62 + 33/62$
D. 62/208
E. There's not a row for Americans, so we can't answer.

10. In one newspaper article, the author discussed having met someone who didn’t actually enjoy the taste of sugar in any form, candy or otherwise. Then the author asked readers to write in to talk about whether they liked sugar. A surprising 40% of 1200 readers who responded said that they also didn’t like sugar. Why is it not safe to infer that 40% of the population doesn’t like sugar?

A. This was only a single study. We should not use the results of one survey to make inferences to the general population.
B. Only 1200 people wrote in. In order to infer something about all Americans, they need a bigger sample size.
C. Perhaps those who also don’t like sugar were more likely to write in than those who do.
D. People who read this article are apparently less likely to like sugar than the general population.
E. People who read the newspaper might be wealthier than the general population.

11. Why would it be better if the distribution of scores on this exam was skewed left?

A. more than half the class would be above the mean
B. few people would do poorly
C. lots of people would make an A
D. All of the above are true.
E. Only two of the above are true.

12. Match the 5-Number-Summary to the correct boxplot.

<table>
<thead>
<tr>
<th>Min</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>15.8</td>
<td>17.5</td>
<td>17.9</td>
<td>18.9</td>
</tr>
<tr>
<td>B</td>
<td>11.8</td>
<td>14.5</td>
<td>15.05</td>
<td>16.7</td>
</tr>
<tr>
<td>C</td>
<td>9.9</td>
<td>15.7</td>
<td>17.3</td>
<td>18.5</td>
</tr>
<tr>
<td>D</td>
<td>15.5</td>
<td>18.4</td>
<td>19.2</td>
<td>19.7</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>15</td>
<td>16.8</td>
<td>17.8</td>
</tr>
</tbody>
</table>

A. A=4,B=2,C=1,D=3,E=5
B. A=3,B=2,C=1,D=4,E=5
C. A=3,B=2,C=5,D=4,E=1
D. A=3,B=5,C=1,D=4,E=1
E. A=4,B=2,C=5,D=3,E=1

13. A study published in the British Journal of Health Psychology found that dog owners are physically healthier than cat owners. Name a potential lurking variable that could explain why we shouldn’t conclude from this study that owning a dog causes better health.

A. physical activity level
B. how often the dog went to the vet
C. whether someone owned a dog
D. the health of a person
E. whether someone owned both

14. I have data on how often a person exercises (never, twice a week, every day, etc.) and whether they eat healthy (yes, no, sometimes). What would be the best way to display this data?

A. in a sorted list
B. with side-by-side boxplots
C. in a two-way table
D. in two pie charts
E. in multiple histograms
15. If the 5-Number-Summary for a dataset is: 65, 113, 134, 157, 209, what is the shape?

A. uniform  
B. normal  
C. skewed left  
D. skewed right  
E. skewed right with outliers

16. When a polling organization asked, “Should laws be passed to eliminate all possibilities of special interests giving huge sums of money to candidates?” 80% of the sample answered yes. When they posed the question, “Should laws be passed to prohibit interest groups from contributing to campaigns, or do groups have a right to contribute to the candidate they support?” only 40% said yes. This is an illustration of problems caused by:

A. lurking variables  
B. undercoverage  
C. voluntary response bias  
D. response bias from question wording  
E. nonresponse

17. Suppose I have 20 numbers ranging from 1 to 50. If I multiply all 20 by 5, which of the following would be true?

A. Each value in the 5-Number-Summary would be 5 times larger.  
B. All measures of spread would be 5 times larger.  
C. The mean and standard deviation of the set of z-scores (the distribution of z-scores) would be 5 times larger.  
D. All of the above would be true.  
E. Only two of the above would be true.

18. When are bar charts preferred over pie charts?

A. when there are a lot of categories, so there would be too many ‘slices’ in the pie  
B. when we want to compare variables, like voting by gender  
C. when the proportions don’t add up to 100%  
D. all of the above  
E. Never, either graph can be used for plotting categorical data?

19. What can be said about the three distributions above? There are the same number of observations in all three.

A. All values in group 3 are smaller than those in the other two groups.  
B. Group 2 has the largest standard deviation.  
C. Group 3 is the only one whose mean is approximately equal to its median.  
D. All of the above are true.  
E. Only two of the above are true.

20. The mayor of Houston is interested in finding out whether Houston residents support raising taxes to fund building more public libraries. Of 523 respondents, 256 said they would support the tax increase. What is the population of interest in this study?

A. the proportion of respondents who said they supported the tax increase  
B. the proportion of all Houston residents who would support the tax increase  
C. all Houston residents  
D. the 523 respondents  
E. the 256 respondents who said they supported the tax increase