1. [3 Points] 3.53 (Survey) p. 201

(a) The population for the survey is the fans with season tickets.
(b) The sample for the survey is the 98 season ticket holders who responded to the survey.
(c) The response rate is $\frac{98}{150} \approx 0.653$.
(d) The nonresponse rate is $\frac{52}{150} \approx 0.347$.
(e) The response rate in the future might be improved by offering incentives to return the survey. For example, one might offer a drawing to meet players on the team. Alternatively, one might try to conduct this survey at the game. In this case, it might be easier to offer smaller incentives, such as concessions.

Responses to part (e) will vary. Acceptable responses for all parts earn 3 points. Acceptable responses for more than half, but not all parts, earns 2 points. Acceptable responses for fewer than half of the parts earns 1 point, and no correct answers earns 0 points.

2. [3 Points] 3.78 (Survey Questions) p. 205

(a) The question is clearly asking whether or not cell phones should come with warnings detailing dangers of their use. The statement that “some cell phone users have developed brain cancer” clearly slants the response towards agreeing that cell phones should require warnings.
(b) The question is somewhat clear in asking if a national health insurance system should be implemented. The question is slanted in favor of nationalized health insurance since it suggests that it would provide insurance for everyone and reduce administrative costs.
(c) The question is asking whether or not economic incentives should be put in place for some recycling. It’s clear that the question is slanted toward economic incentives for recycling.

Responses to all parts will vary. The questions may be construed as unclear, but all of them are slanted. Acceptable responses for all parts earn 3 points. Acceptable responses for more than half, but not all parts, earns 2 points. Acceptable responses for fewer than half of the parts earns 1 point, and no correct answers earns 0 points.

3. [3 Points] 3.125 (Type of Study) p. 229

(a) In a sample survey, you could ask questions that owners can easily answer. For example, what foods do the pets eat? Do the pets go outside?
(b) In an observational study, we might be interested in something that an owner wouldn’t be able to answer easily. This could include interactions between pets and owners. For example, do owners feed their pets when they make noise? How do pets respond to owners who work night jobs?
(c) In an experiment, we need to be able to alter some aspect of life between pets and owners. For example, what happens when the pet food is changed? How does the pet respond to having someone different take care of it for a period of time?

Responses to all parts will vary. If the student gives an answer for one of the options (survey, observational study, experiment) that would clearly be better as one of the other
options, then the answer is wrong. Acceptable responses for all parts earn 3 points. Acceptable responses for more than half, but not all parts, earns 2 points. Acceptable responses for fewer than half of the parts earns 1 point, and no correct answers earns 0 points.

4. [3 Points] 1.39 (Vehicle Colors) p. 27

(a) The bar chart for North America.

![Percentage of Car Colors, North America](chart1)

Figure 1 The bar chart for North American car percentages. White and black cars are the most common.

(b) The bar chart for Europe.

![Percentage of Car Colors, Europe](chart2)

Figure 2 The bar chart for European car percentages. Black and white cars are still the most common.

(c) The juxtaposed bar chart.

To get 3 points, all of the charts need to be correct, although we will not require labels for this assignment.

Award 2 points if the bar charts themselves contain minor errors.
Figure 3 The bar chart for North American and European car percentages by color. The proportions are similar, although white cars and silver cars are slightly more popular in North America, while black and gray are more popular in Europe.

To get 1 point, there need to be substantial flaws in the charts.
To get 0 points, the bar charts need to have not been produced correctly at all.

5. [3 Points] 1.71 (Potatoes) p. 50

(a) The mean and standard deviation of the potato weights are 6.424 and 1.399786 respectively. A histogram of their weights can be seen in Figure 4. The mean and standard deviation are good general measures for understanding the weight of the average potato taken from the bag and the variability of the weight. The histogram gives a good idea of the shape of the distribution of the potato weights.

(b) The histogram shows that there seem to be two distinct types of potato. There are high weight potatoes that exhibit low variance in their weights, and there are low
weight potatoes that exhibit high variance. The numerical summaries fail to capture this; so the histogram is useful in this regard.

(c) We divide the potatoes into two groups. The high weight group includes all potatoes of weight greater than 7 ounces, and the low weight group consists of all potatoes of at most 7 ounces. The mean and standard deviation for the first group are 7.81 and 0.152 respectively. The mean and the standard deviation for the second group are 5.5 and 1.028 respectively.

This might be preferable to just the mean and standard deviation of all the data together in that it gives more detailed information. The low weight potatoes have weights that vary more, and the weights of the heavy potatoes do not vary as much. Alternatively, this might not be preferable if we don’t need as granular an understanding of the distributions of the weights. In particular, we don’t know what the average potato weight is unless we also report the relative frequency of low and high weight potatoes.

To get all 3 points, the mean, standard deviation, and histogram all need to be correct in part (a). The histogram might look slightly different depending on the number of bins the students use. Box plots, pie charts, and bar charts are not acceptable. The five number summary is optional.

One thing to watch out for is a slightly reduced average (6.375 in particular). The reason for this is that if the data is read into R incorrectly, the students will leave out the first data point, which is 7.6, thus dropping the mean. Note if the student gets this wrong, but do not deduct points.

Otherwise, if any part of (a) is wrong, then deduct a point.

In part (b), the students need to say the numerical summaries are insufficient. Deduct a point if they don’t.

In part (c), the potatoes need to be divided into two groups. The 7 ounce mark forms a good cutoff. They could conceivably use a cutoff of up to 7.6 instead and still get the same partitioning. There is an equally sized gap from 5.3 to 6. This gap, or any other reasoned cutoff, is also fine.

Students can prefer the more detailed breakdown or the original statistics, but they have to offer a coherent reason. Otherwise, deduct a point (if there are any left).

6. [3 Points] 1.73 (Papua New Guinea) p. 50

(a) The five number summary for the CRP data can be found in Table 1.

<table>
<thead>
<tr>
<th>Min.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>3rd Qu.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>5.085</td>
<td>9.420</td>
<td>73.20</td>
</tr>
</tbody>
</table>

Table 1 The five number summary of the CRP values.
(b) A boxplot of the CRP data can be found in Figure 5.

![Boxplot of CRP levels in children](image1)

**Figure 5** The boxplot for CRP levels in children from a random sample.

(c) A histogram of the CRP data can be found in Figure 6.

![Histogram of CRP levels in children](image2)

**Figure 6** The histogram for CRP levels in children from a random sample.

(d) The distribution of CRP levels is very right-skewed. While most children have very little CRP, a few have very high levels. Since the boxplot has a harder time of showing just how pronounced the effect is, the histogram is preferable.

All the numbers and plots need to be correct to get all 3 points. In particular, note that the median will be 5.36 in one version of reading the data incorrectly. The quartiles might be slightly off depending on the function used. If they are, e.g. third quartile as 9.47, don’t take off points.

The score should be 2 if more than half of the sections are done properly.
The score should be 1 if no more than half of the sections are done properly. The score should be 0 if no sections are done properly.

7. [3 Points]

(a) In an observational study, women would simply report their hormone use (among other demographic factors of interest). Additionally, there would be records of whether or not they had a heart attack.

In an experiment, some women would be given hormones, and other women would not be given hormones. The groups would be randomized as well. The scientist then observes the outcomes for each group.

(b) Women who took hormones might also have been health-conscious in other ways. For example, they may have exercised more or eaten healthier.

If both sections are correct, award 3 points. There is some flexibility in the answers for each. For part (a), the essential point is that the assignment is not randomized in an observational study, but it is in an experiment. In (b), anything reasonable suffices.

If one section is good, award 2 points.

If neither section is good, award 1 point.

Award 0 points if the answers aren’t close at all.

8. [3 Points]

Let \( r_i \) be the weight of rower \( i \) in pounds. The first statement is that

\[
\frac{1}{7} \sum_{i=1}^{7} r_i = \frac{r_1 + r_2 + r_3 + r_4 + r_5 + r_6 + r_7}{7} = 158.6.
\]

The requirement that the average weight of the 8 rowers is at most 160 is

\[
\frac{1}{8} \sum_{i=1}^{8} r_i = \frac{r_1 + r_2 + r_3 + r_4 + r_5 + r_6 + r_7 + r_8}{8} \leq 160.
\]

From these two, we can see

\[
r_8 \leq 8 \cdot 160 - 7 \cdot 158.6 = 169.8.
\]

Thus the maximum weight of the last rower is 169.8 lbs.

Award 3 points for a correct solution.

Award 2 points if the answer is right but there’s little to no explanation.

Award 1 point if there is explanation but the answer is incorrect.

Award 0 points if the answer is wrong and there’s no explanation.

9. [3 Points]

(a) The mean salary is

\[
\frac{18,000,000 + (11) \cdot 455,000}{12} = 1,917,083
\]

The median salary is $455,000. This is easy to see since the average of the 6th and 7th highest paid players is the same as the minimum salary.
(b) All 11 rookies earn less than the mean, but no player earns less than the median.

(c) If the star’s salary rises to $28,000,000, the mean increases to

\[
\frac{28,000,000 + (11) \cdot 455,000}{12} = 2,750,417.
\]

However, the median salary is unchanged.

Award 3 points if all sections are correct.
Award 2 points if 2 sections are correct.
Award 1 point if 1 section is correct.
Award 0 points if no section is correct.

10. [3 Points]

The largest possible standard deviation occurs when each of the four rolls is as far as possible from the mean of the rolls. This means that we would roll 1 twice and 6 twice. The standard deviation in this case is

\[
\sigma = \sqrt{\frac{1}{4-1} [(1-3.5)^2 + (1-3.5)^2 + (6-3.5)^2 + (6-3.5)^2]} = \sqrt{\frac{4}{3} \cdot 2.5^2} = 2.886751.
\]

The smallest standard deviation occurs when the dice rolls are no different from the mean. This occurs when all four rolls yield the same number. This gives a standard deviation of 0.

Award 3 points for getting both the maximum and the minimum correct.
Award 2 points if only one of the two is correct.
Award 1 point if the numbers are correct but there is no explanation. Award 1 point if the numbers are incorrect but the explanation is correct.
Award 0 points if neither the explanations nor the results are correct.