1. This question consists of two unrelated parts.

(8) (a) A survey of a large college class asked the following questions:

1. Are you female or male? (In the data, male = 0, female = 1)
2. Are you right-handed or left-handed? (In the data, right = 0, left = 1)
3. What is your height in inches?
4. How many minutes do you study on a typical weeknight?

The figure below shows histograms of the student responses, in scrambled order and without scale markings. Which histogram goes with each variable? Give a short explanation. *(Hint: 90 to 93 percent of the adult population is right-handed)*

(a) ![Histogram a]
(b) ![Histogram b]
(c) ![Histogram c]
(d) ![Histogram d]

3. For a large class, height should be approximately normally distributed.
2. Most people are right-handed.
4. Has to be either (b) or (c) and (b) is (2).
4. The only one left

(5) (b) Which five statistics (numbers) can be used to summarize a scatter diagram?

\[ \text{Avg}_x, \text{Std}_x, \text{Avg}_y, \text{Std}_y, r \]
(8) 2. In a large class, the average score on the final was 50 out of 100, and the SD was 20. The scores followed the normal curve.

Two brothers took the final. One placed at the 70th percentile and the other was at the 80th percentile. How many points separated them?

70th percentile:

By normal table: \( z \approx 0.525 \)

80th percentile:

By normal table: \( z \approx 0.85 \)

In original units (\( x = \text{Avg} + z \times \text{SD} \))

\[
50 + 0.525 \times 20 = 60.5 \\
50 + 0.85 \times 20 = 67
\]

The difference: 67 - 60.5 = 6.5

The answer: 6.5 points separated the two brothers.
3. We expect a car’s highway gas mileage to be related to its city gas mileage. Data for 1198 vehicles in the government’s 2008 Fuel Economy Guide give the regression line

\[
\text{highway mpg} = 4.62 + (1.109 \times \text{city mpg})
\]

for predicting highway mileage from city mileage.

(a) What is the slope of this line? Say in words what the numerical value of the slope tells you.

\[
\text{Slope} = 1.109
\]

Value of the slope is larger than 0. This indicates that there is a positive correlation between city mpg and highway mpg. A change of 1 in \(x\) will give a change of 1.109 in \(y\).

(b) What is the intercept? Explain why the value of the intercept is not statistically meaningful.

\[
\text{Intercept} = 4.62
\]

The value of the intercept is the predicted value of \(y\) when \(x = 0\), or for a car which has city mileage of 0. That is not statistically meaningful.
(4) 4. For men age 18-24 in the HANES study, the relationship between height and weight can be summarized as follows:

\[ y \quad \text{average height} = 70 \text{ inches, SD} = 3 \text{ inches} \]

\[ x \quad \text{average weight} = 162 \text{ pounds, SD} = 30 \text{ pounds}, \quad r = 0.47 \]

Estimate the average height of men that weight 117 pounds.

**Method 1:**

\[ 2_x = \frac{x - \text{Avg}_x}{\text{SD}_x} = \frac{117 - 162}{30} = -1.5 \]

\[ 2_y = r \times 2_x = 0.47 \times (-1.5) = -0.705 \]

\[ y = \text{Avg}_y + 2_y \times \text{SD}_y = 70 - 0.705 \times 3 = 67.885 \]

**Method 2:**

\[ \text{Slope} = r \times \frac{\text{SD}_y}{\text{SD}_x} = 0.47 \times \frac{3}{30} = 0.047 \]

\[ \text{intercept} = \text{Avg}_y - \text{Slope} \times \text{Avg}_x = 70 - 0.047 \times 162 = 62.386 \]

\[ y = 0.047 \times x + 62.386 \]

\[ = 0.047 \times 117 + 62.386 \]

\[ = 67.885 \]

We estimate that the average height of men that weight 117 pounds is 67.9 inches.

(4) 5. (a) The following is taken from the textbook. Fill in the first blank and use one of the following options for the second blank: (i) understate, (ii) overstate.

Ecological correlations ________ are based on rates and averages. They are often used in political science and sociology. And they tend to ________ the strength of an association.
(4) (b) Investigators are investigating the effect of long-term smoking. For this, they compare a treatment group consisting of 30 individuals who have smoked for over 40 years to a control group consisting of 30 individuals who have never smoked in the past 40 years. Is this a randomized controlled experiment or an observational study? Give a short explanation with your answer.

This is an observational study. The investigators can’t force someone to smoke/not smoke for 40 years. The subjects therefore assign themselves to treatment or control based on their smoking habits and the investigators just observe the difference between the groups (if any).

(5) 6. A personality test is administered to a large group of subjects. Four scores are shown below, in original units and in standard units. Fill in the blanks.

<table>
<thead>
<tr>
<th>original units</th>
<th>79</th>
<th>64</th>
<th>52</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>standard units</td>
<td>1.8</td>
<td>0.8</td>
<td>0.</td>
<td>-1.4</td>
</tr>
</tbody>
</table>

In standard units, the difference between the first two values is 1. The difference in original units is thus equal to 1 SD:

\[ SD = 79 - 64 = 15 \]

We know that

\[ x = \text{Avg} \pm 2 \times SD \]

or

\[ \text{Avg} = x - 2 \times SD = 79 - 1.8 \times 15 = 52 \]

52 is the average, which is 0 in standard units. For the other blank, we get

\[ x = \text{Avg} \pm 2 \times SD = 52 \pm (-1.4) \times 15 = 31 \]