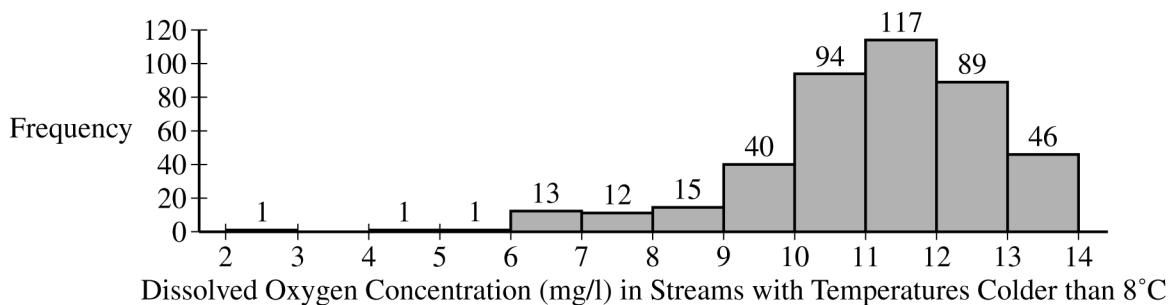


Begin your response to **QUESTION 1** on this page.

STATISTICS**SECTION II****Total Time—1 hour and 30 minutes****6 Questions****Part A****Suggested Time—1 hour and 5 minutes****5 Questions**

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

- As part of a study on the chemistry of Alaskan streams, researchers took water samples from many streams with temperatures colder than 8°C and from many streams with temperatures warmer than 8°C . For each sample, the researchers measured the dissolved oxygen concentration, in milligrams per liter (mg/l).



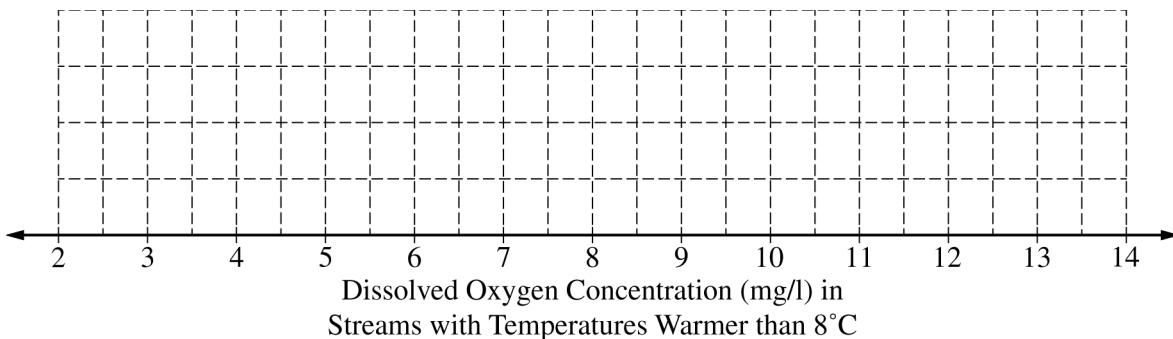
- The researchers constructed the histogram shown for the dissolved oxygen concentration in streams from the sample with water temperatures colder than 8°C . Based on the histogram, describe the distribution of dissolved oxygen concentration in streams with water temperatures colder than 8°C .

GO ON TO THE NEXT PAGE.

Continue your response to **QUESTION 1** on this page.

Min	Q1	Median	Q3	Max	Mean	Std. Dev.
2.10	4.39	5.43	6.12	13.45	5.54	1.64

- (b) The researchers computed the summary statistics shown in the table for the dissolved oxygen concentration in streams from the sample with water temperatures warmer than 8°C. Use the summary statistics to construct a box plot for the dissolved oxygen concentration in streams with water temperatures warmer than 8°C. Do not indicate outliers.



- (c) The researchers believe that streams with higher dissolved oxygen concentration are generally healthier for wildlife. Which streams are generally healthier for wildlife, those with water temperature colder than 8°C or those with water temperature warmer than 8°C? Using characteristics of the distribution of dissolved oxygen concentration for temperatures colder than 8°C and characteristics of the distribution of dissolved oxygen concentration for temperatures warmer than 8°C, justify your answer.

GO ON TO THE NEXT PAGE.

Begin your response to **QUESTION 2** on this page.

2. A developer wants to know whether adding fibers to concrete used in paving driveways will reduce the severity of cracking, because any driveway with severe cracks will have to be repaired by the developer. The developer conducts a completely randomized experiment with 60 new homes that need driveways. Thirty of the driveways will be randomly assigned to receive concrete that contains fibers, and the other 30 driveways will receive concrete that does not contain fibers. After one year, the developer will record the severity of cracks in each driveway on a scale of 0 to 10, with 0 representing not cracked at all and 10 representing severely cracked.

(a) Based on the information provided about the developer’s experiment, identify each of the following.

- Experimental units

- Treatments

- Response variable

(b) Describe an appropriate method the developer could use to randomly assign concrete that contains fibers and concrete that does not contain fibers to the 60 driveways.

GO ON TO THE NEXT PAGE.

Question 1: Focus on Exploring Data**4 points****General Scoring Notes**

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

Model Solution	Scoring
<p>(a) The histogram of dissolved oxygen concentration in Alaskan streams with water temperatures colder than 8°C is unimodal and skewed left with a median between 11 and 12 mg/l.</p> <p>The first quartile is in the bin from 10-11 mg/l and the third quartile is in the bin from 12-13 mg/l, so the IQR is approximately 2 mg/l.</p> <p>There do not appear to be any high outliers, but there are several potential low outliers because the values in the 2-3, 4-5, and 5-6 bins are all certainly more than 1.5 IQR below the first quartile.</p>	<p>Essentially correct (E) if the description of the distribution satisfies component 1 <i>AND</i> at least three of components 2-5:</p> <ol style="list-style-type: none"> Includes context of dissolved oxygen concentration Shape: The distribution is skewed left Center: The center of the distribution is between 11 mg/l and 12 mg/l Spread: Refers to at least one measure of variation (i.e., the range is between $14 - 2 = 12$ mg/l and $13 - 3 = 10$ mg/l; all values are between 2 mg/l and 14 mg/l or the IQR is approximately 2 mg/l) Unusual features: potential outliers or a gap between 3 mg/l and 4 mg/l <p>Partially correct (P) if the response satisfies component 1 and two components out of components 2-5 <i>OR</i> if the response satisfies at least three out of components 2-5 but does not satisfy component 1.</p> <p>Incorrect (I) if the response does not meet the criteria for E or P.</p>

Additional Notes:

- Context**
 - Component 1 can be satisfied with a reference to dissolved oxygen concentration, dissolved oxygen, amount of oxygen, or mg/l.

- **Shape:**
 - Component 2 cannot be satisfied if a response describes the histogram as normal or approximately normal.
- **Center:**
 - A response that addresses center using definitive language such as “the mean (median) of the distribution is 11.5” cannot satisfy component 3.
 - A response that addresses center using approximate language such as “the median of the distribution is approximately 11.5” must, for any single measure of center, specify a numeric value that is between 11 mg/l and 12 mg/l, inclusive, to satisfy component 3.
- **Spread:**
 - A response recognizing all values in the sample fall between 2 mg/l and 14 mg/l, satisfies component 4 only for these exact endpoints and need not appeal to a specific measure of spread such as the range.
 - A response that uses interval language must use it correctly. For example, “the observations range from 2 mg/l and 14 mg/l,” satisfies component 4 because it correctly indicates that all observations are between 2 mg/l and 14 mg/l, inclusive. However, a statement such as “the range is between 2 mg/l and 14 mg/l,” is incorrect because the range is a single number, i.e., $14 - 2 = 12$ mg/l.
 - A response that appeals to a specific measure of spread using approximate language, such as “the IQR is approximately 2,” must specify a numeric value within the bounds appropriate to that measure of spread shown in the following table.

Statistic	Min (mg/l)	Max (mg/l)
Range	10	12
IQR	1.5	2.5
Standard Deviation	1.7	1.8

- **Unusual Features:**
 - Component 5 cannot be satisfied if a response indicates that the distribution has an unusual feature other than potential outliers or a gap.
 - Definitive language such as “there is an outlier” does not satisfy component 5.