The materials included in these files are intended for non-commercial use by AP teachers for course and exam preparation; permission for any other use must be sought from the Advanced Placement Program. Teachers may reproduce them, in whole or in part, in limited quantities, for face-to-face teaching purposes but may not mass distribute the materials, electronically or otherwise. These materials and any copies made of them may not be resold, and the copyright notices must be retained as they appear here. This permission does not apply to any third-party copyrights contained herein.
STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II grade—75

Directions: 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 of your results and explanation.

1. The summary statistics for the number of inches of rainfall in Los Angeles for 117 years, beginning in 1877, are shown below.

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN</th>
<th>MEDIAN</th>
<th>TRMEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>MIN</th>
<th>MAX</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
</table>

(a) Describe a procedure that uses these summary statistics to determine whether there are outliers.

To find outliers, take $1.5 \cdot IQR$ (Interquartile Range). Subtract this from the first quartile number, and add it to the third quartile number.

$$1.5 \cdot IQR = 1.5 \cdot (19.250 - 9.680) = 1.5 (9.57) = 14.355$$

$$9.680 - 14.355 = -4.675$$

$$19.250 + 14.355 = 33.605$$

Anything outside of these numbers (in this case, below -4.675 or above 33.605) is considered an outlier.
(b) Are there outliers in these data? Yes.
Justify your answer based on the procedure that you described in part (a).

Although it is impossible to have an outlier lower than -4.675 (because negative rainfall is impossible) we know there is an outlier on the upper end. The maximum value is 38.180, which lies above 33.605, thus making it an outlier. However, without seeing the actual data, we cannot be sure if there are others within the data besides 38.180.

(c) The news media reported that in a particular year, there were only 10 inches of rainfall. Use the information provided to comment on this reported statement.

10 inches of rainfall is not uncommon. Slightly more than 25% of all years have less rainfall than that, as displayed by the first quartile number. That year was by no means abnormal.
STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II grade—75

Directions: 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 of your results and explanation.

1. The summary statistics for the number of inches of rainfall in Los Angeles for 117 years, beginning in 1877, are shown below.

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN</th>
<th>MEDIAN</th>
<th>TRMEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>MIN</th>
<th>MAX</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
</table>

(a) Describe a procedure that uses these summary statistics to determine whether there are outliers.

Outliers are points lying outside plus or minus three standard deviations from the mean. One can find this range and see if the minimum and maximum are within this range. If either of these two points is outside the three standard deviations, they are outliers.
(b) Are there outliers in these data? Yes

Justify your answer based on the procedure that you described in part (a).

\[
\begin{align*}
\mu &= 30 \\
14.941 &\pm 3(6.747) \\
14.941 &\pm 20.241 \\
(-5.3, 35.182) \\
\text{Minimum} &= 4.880 \\
\text{within range} &\checkmark \\
\text{Maximum} &= 38.180 \\
\text{outside of range by} 2.978 &\Rightarrow \text{outlier}
\end{align*}
\]

(c) The news media reported that in a particular year, there were only 10 inches of rainfall. Use the information provided to comment on this reported statement.

\[
2 = \frac{x - \mu}{\sigma}
\]

\[
2 = \frac{10 - 14.941}{6.747} = \frac{-4.941}{6.747} = -0.7323
\]

The media reports this data as if it were rarely significantly below the average (mean) rainfall. However, after determining the z-score for that year, one can see that it is within one standard deviation from the mean and is therefore not rare by any means.
1. The summary statistics for the number of inches of rainfall in Los Angeles for 117 years, beginning in 1877, are shown below.

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN</th>
<th>MEDIAN</th>
<th>TRMEAN</th>
<th>STDEV</th>
<th>SE MEAN</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>MIN</th>
<th>MAX</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
</table>

(a) Describe a procedure that uses these summary statistics to determine whether there are outliers.

First, find the interquartile range $IQR$:

$$IQR = Q_3 - Q_1 = 19.250 - 9.680 = 9.57$$

An outlier is defined as any value lying $1.5IQR$ above $Q_3$ or below $Q_1$.

If $$x \leq 9.680 - 1.5(9.57) = -4.675$$ or $$x \geq 19.250 + 1.5(9.57) = 33.605$$

then $x$ is an outlier.

Since $x$ cannot be negative, if $x \geq 33.605$, then $x$ is an outlier.
(b) Are there outliers in these data? **Yes**

Justify your answer based on the procedure that you described in part (a).

There are no outliers below Q1, but there is at least one outlier above Q3. The maximum value is 38.190 which is greater than Q3 + 1.5IQR, or 33.605, so it is an outlier.

(c) The news media reported that in a particular year, there were only 10 inches of rainfall. Use the information provided to comment on this reported statement.

That is not an extreme observation. The probability \( p \) that there would be less than 10 inches of rain fall in a given year is given by

\[
p = N(z)
\]

where \( z = \frac{10 - \mu}{\sigma} = -0.73 \)

\( p \) is the area to the left of the normal curve:

\[
p = 23.27\%.
\]