Learning Task: If the Shoe Fits

Welcome to CSI at School! Over the weekend, a student entered the school grounds without permission. Even though it appears that the culprit was just looking for a quiet place to study, undisturbed by friends, school administrators are anxious to identify the offender and have asked for your help. The only available evidence is a suspicious footprint outside the library door.

After the incident, school administrators arranged for the data in the table below to be obtained from a random sample of this high school's students. The table shows the shoe print length (in cm), height (in inches), and gender for each individual in the sample.

<table>
<thead>
<tr>
<th>Shoe Print Length</th>
<th>Height</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>71</td>
<td>F</td>
</tr>
<tr>
<td>27</td>
<td>65</td>
<td>F</td>
</tr>
<tr>
<td>26</td>
<td>64</td>
<td>F</td>
</tr>
<tr>
<td>25.5</td>
<td>64</td>
<td>F</td>
</tr>
<tr>
<td>25</td>
<td>63</td>
<td>F</td>
</tr>
<tr>
<td>25.5</td>
<td>64</td>
<td>F</td>
</tr>
<tr>
<td>27</td>
<td>67</td>
<td>F</td>
</tr>
<tr>
<td>26</td>
<td>64</td>
<td>F</td>
</tr>
<tr>
<td>27</td>
<td>67</td>
<td>F</td>
</tr>
<tr>
<td>26.5</td>
<td>64</td>
<td>F</td>
</tr>
<tr>
<td>22.5</td>
<td>61</td>
<td>F</td>
</tr>
<tr>
<td>32</td>
<td>74</td>
<td>M</td>
</tr>
<tr>
<td>30</td>
<td>65</td>
<td>M</td>
</tr>
<tr>
<td>31</td>
<td>71</td>
<td>M</td>
</tr>
<tr>
<td>29.5</td>
<td>67</td>
<td>M</td>
</tr>
<tr>
<td>29</td>
<td>72</td>
<td>M</td>
</tr>
<tr>
<td>31</td>
<td>72</td>
<td>M</td>
</tr>
<tr>
<td>31</td>
<td>69</td>
<td>M</td>
</tr>
<tr>
<td>30</td>
<td>69</td>
<td>M</td>
</tr>
</tbody>
</table>

1. Explain why this study was an observational study and not an experiment.

2. Why do you think the school's administrators chose to collect data on a random sample of students from the school? What benefit might a random sample offer?
3. Suggest a graph that might be used to use to compare the shoe print length data distributions for females and males.

4. Describe one advantage of using comparative box plots instead of comparative dot plots to display this data.

5. For each gender calculate the five-number summary for the shoe print lengths. Additionally, for each gender, determine if there are any outlying shoe print length values.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Quartile 1 (Q1)</th>
<th>Median (Q2)</th>
<th>Quartile 3 (Q3)</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>29</td>
<td>29.75</td>
<td>30.5</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>Female</td>
<td>22.5</td>
<td>25</td>
<td>24.0</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>

6. Construct comparative box plots for the shoe print lengths of males and females. Discuss the similarities and differences in the shoe print length distributions for the males and females in this sample.

7. For each gender calculate the mean shoe print length. What information does the mean shoe print length provide?

\[
\text{Male } \bar{X} = \frac{243.5}{8} = 30.44 \\
\text{Female } \bar{X} = \frac{282}{11} = 25.44
\]
8. The mean will give us an indication of a typical shoe print length. In addition to knowing a typical length we would also like to know how much variability to expect around this length. For each gender calculate the **Range**, **Interquartile Range**, and **Mean Absolute Deviation** of the shoe print lengths.

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>IQR</th>
<th>M.A.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male</strong></td>
<td>8</td>
<td>1.25</td>
<td>0.81</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>4.5</td>
<td>2</td>
<td>1.03</td>
</tr>
</tbody>
</table>

9. If the length of a student's shoe print was 32 cm...

   A. Would you think that the print was made by a male or a female?

   B. How sure are you that you are correct? Explain your reasoning. Use results from Questions 5 through 8 in your explanation.

10. How would you answer Question 9 if the suspect's shoe print length was 27 cm?
1. There was not a hypothesis stated that needed to be tested. There was just a random sample taken to hopefully find the match to the suspicious footprint.

2. A random sample would be a representative sample of the school’s population. The benefit would be a “good” sample instead of a biased sample.

3. A box plot would be the best graph to compare the shoe print lengths of females and males.

4. One advantage of using box plots to compare data instead of dot plots is in how easy box plots are to make compared to dot plots. While dot plots show every piece of data, they are not easy to use in order to compare data. Also, the box plot gives you 5 data values for comparing the distributions and they are easy to see from the graph.
7. The average shoe print length for a male is almost 5 cm longer than that of the average female.

9A. Male

9B. I am very confident that a shoe print with a length of 32 cm is a male because of the following:

Mean: Female – 25.64     Male – 30.44
Median: Female – 26     Male – 30.5

The maximum length of a female shoe print is only 27 cm. The maximum length of a male shoe print is 32 cm.

10. If the shoe print length was 27 cm then I would feel confident that the suspect was a female using the same data listed in #9.

The minimum shoe print length of a male is only 29 cm.
5. Male
29, 29.5, 30, 30, 31, 31, 31, 32

Female
22.5, 24, 25, 25.5, 25.5, 26, 26, 26.5, 27, 27, 27

8. Male
range = 32 - 29 = 3
IQR = 31 - 29.75 = 1.25

Female
range = 27 - 22.5 = 4.5
IQR = 27 - 25 = 2
male

MAD = 0.81

\[ \bar{x} = 30.44 \]

\[ 29 - 30.44 = -1.44 \]
\[ 29.5 - 30.44 = -0.94 \]
\[ 30 - 30.44 = -0.44 \]
\[ 30 - 30.44 = -0.44 \]
\[ 31 - 30.44 = 0.56 \]
\[ 31 - 30.44 = 0.56 \]
\[ 31 - 30.44 = 0.56 \]
\[ 32 - 30.44 = 1.56 \]

\[ (1.44 + 0.94 + 0.44 + 0.44 + 0.56 + 0.56 + 0.56 + 1.56) / 8 \]

\[ = \frac{6.5}{8} = 0.8125 \]
Female

\[ \text{MAD} = 1.03 \]

\[ \bar{x} = 25.64 \]

22.5 - 25.64 = -3.14
24 - 25.64 = -1.64
25 - 25.64 = -0.64
25.5 - 25.64 = -0.14
25.5 - 25.64 = -0.14
26 - 25.64 = 0.36
26 - 25.64 = 0.36
26.5 - 25.64 = 0.86
27 - 25.64 = 1.36
27 - 25.64 = 1.36
27 - 25.64 = 1.36

\[
\frac{(3.14 + 1.64 + 0.64 + 0.14 + 0.14 + 0.36 + 0.36 + 0.86 + 1.36 + 1.36 + 1.36)}{11} = 1.03
\]