EX 1.3.1: Given the following sample of lifetimes of light bulbs (in years):

$$x: \ 8.5, \ 6.8, \ 7.7, \ 10.0, \ 11.3, \ 10.0, \ 9.9$$

a) Compute the sample mean, $\overline{x}$.

Observe that the sample size $n = (\#$ of data points in sample $x) = 7$

$$\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{7}\sum_{k=1}^{7} x_k = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7}{7} = \frac{8.5 + 6.8 + 7.7 + 10.0 + 11.3 + 10.0 + 9.9}{7} = \frac{64.2}{7} \approx \boxed{\mathbf{9.171}}$$

b) Compute the sample median, $\widetilde{x}$.

First, sort sample: $x: \ 6.8, \ 7.7, \ 8.5, \ 9.9, \ 10.0, \ 10.0, \ 11.3$

Then, since sample size $(n = 7)$ is odd, $\widetilde{x} = x_{([n+1]/2)} = x_{([7+1]/2)} = x_{(4)} = (4^{th}$ data point in sorted sample $x) = \boxed{\mathbf{9.9}}$

c) **Without visualizing the data**, identify the skewness of the sample.

Observe that $\overline{x} \approx 9.171 < 9.9 = \widetilde{x} \implies \overline{x} < \widetilde{x} \implies$ Sample is **negatively skewed**

d) Suppose the largest & smallest data points were removed from the sample. What trimming percentage achieves this?

Solve for p: $\frac{np}{100} = 1 \longleftarrow$ RHS is **one** since **one** extreme data point per side of sorted sample is to be removed.

$$\frac{np}{100} = 1 \implies \frac{7p}{100} = 1 \implies p = \frac{100}{7} \approx 14.29 \implies \text{Trimming percentage is } \boxed{\mathbf{14.29\%}}$$

e) Compute the trimmed mean using the trimming percentage found in the previous part.

First, remove the largest & smallest data points, then compute ordinary mean of remaining dataset:

(14.29% Trimmed x): $7.7, \ 8.5, \ 9.9, \ 10.0, \ 10.0$

$$\therefore \quad \overline{x}_{tr(14.29\%)} = \frac{7.7 + 8.5 + 9.9 + 10.0 + 10.0}{5} = \frac{46.1}{5} = \boxed{\mathbf{9.22}}$$

f) Compute the 33^{rd} percentile of the sample, $x_{0.33}$.

$x_{0.33} = x_{(\lceil 0.33n \rceil)} = x_{(\lceil 0.33(7) \rceil)} = x_{(\lceil 2.31 \rceil)} = x_{(3)} = (3^{rd}$ data point in sorted sample $x) = \boxed{\mathbf{8.5}}$

g) Which data point is the smallest such data point that is greater than or equal to 25% of the sample?

This by definition is the 25^{th} percentile, which is also the 1^{st} quartile:

$x_{Q1} = x_{0.25} = x_{(\lceil 0.25n \rceil)} = x_{(\lceil 0.25(7) \rceil)} = x_{(\lceil 1.75 \rceil)} = x_{(2)} = (2^{nd}$ data point in sorted sample $x) = \boxed{\mathbf{7.7}}$

h) Compute the lower hinge, x_{LH}, and upper hinge, x_{UH}, of the sample.

First, split the sorted sample into two halves using the sample median as guidance.

(Left Half of x): $6.8, \ 7.7, \ 8.5, \ 9.9 \implies x_{LH} = $ (Median of left half of x) $= \frac{7.7 + 8.5}{2} = \boxed{\mathbf{8.1}}$

(Right Half of x): $9.9, \ 10.0, \ 10.0, \ 11.3 \implies x_{UH} = $ (Median of right half of x) $= \frac{10.0 + 10.0}{2} = \boxed{\mathbf{10.0}}$

NOTE: The median $(\widetilde{x} = 9.9)$ is included in both halves since sample size $(n = 7)$ is odd.

LEGEND:

- RHS translates to "**R**ight **H**and **S**ide" (of an equation)
- $\implies$ translates to "which implies that"
- $\lceil x \rceil$ means **round** x **up** if not an integer: $\lceil 3 \rceil = 3, \ \lceil 3.1 \rceil = 4, \ \lceil 3.5 \rceil = 4, \ \lceil 3.9 \rceil = 4$

EX 1.3.2: Given the following sample of lifetimes of light bulbs (in years):

$$y: \;\; 8.5,\; 6.8,\; 7.7,\; 10.0,\; 11.3,\; 10.0,\; 9.9, 18.5$$

a) Compute the sample mean, $\overline{y}$.

Observe that the sample size $m = (\#$ of data points in sample $y) = 8$

$$\overline{y} = \frac{1}{m}\sum_{k=1}^{m} y_k = \frac{1}{8}\sum_{k=1}^{8} y_k = \frac{8.5+6.8+7.7+10.0+11.3+10.0+9.9+18.5}{8} = \frac{82.7}{8} = \boxed{\textbf{10.3375}}$$

b) Compute the sample median, $\widetilde{y}$.

First, sort sample: $\quad y: \;\; 6.8,\; 7.7,\; 8.5,\; 9.9,\; 10.0,\; 10.0,\; 11.3,\; 18.5$

Then, since sample size $(m = 8)$ is even, $\widetilde{y} = \frac{y_{(n/2)} + y_{(1+[n/2])}}{2} = \frac{y_{(4)} + y_{(5)}}{2} = \frac{9.9+10.0}{2} = \frac{19.9}{2} = \boxed{\textbf{9.95}}$

c) **Without visualizing the data**, identify the skewness of the sample.

Observe that $\overline{y} = 10.3375 > 9.95 = \widetilde{y} \implies \overline{y} > \widetilde{y} \implies$ Sample is $\boxed{\textbf{positively skewed}}$

d) Compute the 25% trimmed sample mean, $\overline{y}_{tr(25\%)}$.

First, determine how many extreme data points per side of sorted sample to trim:

$\frac{mp}{100} = \frac{(8)(25)}{100} = \frac{200}{100} = 2 \implies$ The 2 largest & 2 smallest data points should be trimmed from sample

Then, remove the two largest & two smallest data points, then compute ordinary mean of remaining dataset:

(25% Trimmed y) : $\;\; 8.5,\; 9.9,\; 10.0,\; 10.0$

$$\therefore \quad \overline{y}_{tr(25\%)} = \frac{8.5+9.9+10.0+10.0}{4} = \frac{38.4}{4} = \boxed{\textbf{9.6}}$$

e) Compute the 3^{rd} quartile of the sample, y_{Q3}.

$y_{Q3} = y_{0.75} = y_{(\lceil 0.75m \rceil)} = y_{(\lceil 0.75(8) \rceil)} = y_{(\lceil 6 \rceil)} = y_{(6)} = (6^{th}$ data point in sorted sample $y) = \boxed{\textbf{10.0}}$

f) Which data point is the largest such data point that is less than or equal to 87% of the sample?

This is equivalent to finding the smallest data point that's greater than or equal to $100\% - 87\% = 13\%$ of the sample.

$y_{0.13} = y_{(\lceil 0.13m \rceil)} = y_{(\lceil 0.13(8) \rceil)} = y_{(\lceil 1.04 \rceil)} = y_{(2)} = (2^{nd}$ data point in sorted sample $y) = \boxed{\textbf{7.7}}$

g) Compute the lower hinge, y_{LH}, and upper hinge, y_{UH}, of the sample.

First, split the sorted sample into two halves using the sample median as guidance.

(Left Half of y) : $\;\; 6.8,\; 7.7,\; 8.5,\; 9.9 \quad \implies y_{LH} =$ (Median of left half of y) $= \frac{7.7+8.5}{2} = \boxed{\textbf{8.1}}$

(Right Half of y) : $\;\; 10.0,\; 10.0,\; 11.3,\; 18.5 \quad \implies y_{UH} =$ (Median of right half of y) $= \frac{10.0+11.3}{2} = \boxed{\textbf{10.65}}$

NOTE: The median $(\widetilde{y} = 9.95)$ is not included in either half since sample size $(m = 8)$ is even.

LEGEND:

- $\implies$ translates to "which implies that"
- $\lceil x \rceil$ means **round** x **up** if not an integer: $\lceil 3 \rceil = 3, \;\; \lceil 3.1 \rceil = 4, \;\; \lceil 3.5 \rceil = 4, \;\; \lceil 3.9 \rceil = 4$