Describing Distributions Numerically

Weight of contents of regular cola.

368, 367, 369, 370, 369, 370
366, 373, 365, 362, 378, 368

• What is a “typical” value?
• Look for the center of the distribution.
• What do we mean by “center”?

Describing Distributions Numerically

Weight of contents of regular cola.

36 |2
36*567899
37 |003
37*8

Describing Distributions Numerically

Weight of contents of regular cola.

36 |2
36*567899
37 |003
37*8
Measures of Center

- **Sample Midrange**
  - Average of the minimum and the maximum.
    \( \frac{362 + 378}{2} = 370 \text{ grams} \)
  - Greatly affected outliers.

Measures of Center

- **Sample Median**
  - A value that divides the data into a lower half and an upper half.
  - About half the data values are greater than the median about half are less than the median.

Sample Median

\[
\begin{array}{c}
36 & 2 \\
36* & 56788 \\
37 & 003 \\
37* & 8 \\
\end{array}
\]

Median = \( \frac{368 + 369}{2} \)
= 368.5 grams
Describing Distributions Numerically

• How much variation is there in the data?
• Look for the spread of the distribution.
• What do we mean by “spread”?

Measures of Spread

• Sample Range
  – The distance from the minimum and the maximum.
    \[(378 - 362) = 16 \text{ grams}\]
  – The length of the interval that contains 100% of the data.
  – Greatly affected outliers.

Quartiles

• Medians of the lower and upper halves of the data.
• Trying to split the data into fourths, quarters.
Quartiles

Lower quartile = (366+367)/2
= 366.5 grams

Upper quartile = (370+370)/2
= 370.0 grams

Measure of Spread

• InterQuartile Range (IQR)
  – The distance between the quartiles.
  IQR = 370 – 366.5 = 3.5 grams
  – The length of the interval that contains the central 50% of the data.

Five Number Summary

• Minimum 362 grams
• Lower Quartile 366.5 grams
• Median 368.5 grams
• Upper Quartile 370 grams
• Maximum 378 grams