8th Grade Class Notes

Ethan Kent

September 16, 2024

Contents

1 Measures of Central Tendency and Spread

When analyzing a set of data, there are different ways to summarize both the center of the data and the spread of the data. Two common ways to describe the center of a dataset are the *median* and the *mean*.

Similarly, two ways to describe how spread out the data are include the *interquartile range* (IQR) and the *standard deviation*.

Each of these measures has its strengths, particularly when it comes to handling *outliers*—data points that are significantly higher or lower than the others.

1.1 Median and Interquartile Range

The median is the middle value in a set of data when the numbers are arranged in order. It splits the data into two equal halves: half of the values are below the median, and half are above.

The interquartile range (IQR) measures the spread of the middle 50% of the data. It is calculated as the difference between the third quartile (the value below which 75% of the data falls) and the first quartile (the value below which 25% of the data falls).

- The median and IQR are most useful when the data contains outliers. This is because the median is not affected by extremely large or small values, while the mean can be pulled significantly in one direction by outliers.

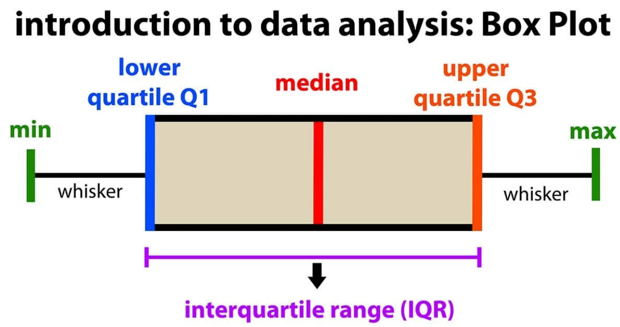


Figure 1: A box-and-whisker plot showing the distribution of data, highlighting the median, quartiles, and outliers.

- For example, consider the incomes of a small group of people. If one person in the group is a millionaire, the *mean* income would be much higher than most people's actual incomes, making it a misleading measure of the "typical" income. In this case, the *median* income would give a better idea of what most people in the group earn.

Example:

- Data set: $\{10, 12, 15, 18, 100\}$ (an outlier at 100).
- **Median**: The median is 15, which is a good representation of the middle of the data.
- **IQR**: The IQR tells us how spread out the middle 50% of the data is, ignoring the outlier at 100.

1.2 Mean and Standard Deviation

The mean is the average of the values in a data set. It is found by adding all the values together and dividing by the number of values. The standard deviation is a measure of how spread out the numbers are from the mean. A small standard deviation means that the values are close to the mean, while a large standard deviation indicates that the values are more spread out.

- The mean and standard deviation are best used when the data is *symmetric* and does not contain outliers. Because the mean is sensitive to extremely large or small values, outliers can skew the result and give a misleading picture of the data.
- When the data is fairly uniform, the mean provides a good measure of central tendency, and the standard deviation gives a clear indication of how much the data values vary around the mean.

Example:

- Data set: $\{10, 12, 15, 18, 20\}$.

- **Mean**: The mean is 15, which accurately reflects the central tendency of the data since there are no outliers.
- **Standard Deviation**
  - The standard deviation, σ (lowercase sigma), gives a sense of how closely clustered the data points are around the mean. In this population, $\sigma = 3.69$.
  - If the data set were $\{1, 8, 15, 22, 29\}$, the mean would still be 15, but $\sigma = 9.90$

1.3 Choosing Between Median and Mean

The key difference between using the median or the mean depends on whether your data contains **outliers**.

- If your data has extreme values (outliers), the *median* and *IQR* will give you a better understanding of the typical values in your data set.
- If your data does not have outliers and is relatively symmetric, the *mean* and *standard deviation* provide a more precise description of both the center and the spread of the data.

In summary:

- Use the *median* and *IQR* when there are *outliers* or a *skewed distribution*.
- Use the *mean* and *standard deviation* when the data is *symmetrically distributed* and *outlier-free*.

2 Plugging in Numbers for Equations

You may be asked to solve an equation when given a value for the variable. For example:

$$\begin{aligned}
d &= 3l + 2 && \text{Original equation} \\
d &= 3(4) + 2 && \text{Substitute } l = 4 \\
d &= 12 + 2 && \text{Multiply 3 by 4} \\
d &= 14 && \text{Simplify}
\end{aligned}$$

You might also be asked whether an ordered pair satisfies an equation. To solve this, remember that an ordered pair represents the x and y values, in that order. For example, given the equation: $y = 2x - 5$, you can determine if the ordered pair $(3, 1)$ satisfies the equation by substituting $x = 3$ and $y = 1$:

$$\begin{aligned}
y &= 2x - 5 && \text{Original equation} \\
1 &= 2(3) - 5 && \text{Substitute } x = 3 \text{ and } y = 1 \\
1 &= 6 - 5 && \text{Multiply 2 by 3} \\
1 &= 1 && \text{Simplify, the equation is true}
\end{aligned}$$

Since the equation is true when substituting the values, the ordered pair $(3, 1)$ does satisfy the equation. If the equation were false after substituting the values, then the ordered pair would not satisfy the equation.

3 Graphing

When asked to graph a linear equation on the coordinate plane, there are several approaches. First, we can make a table by choosing values for x and then solving for y. From this, we can make a table of x and y values, plot those points on the coordinate plane, and connect the points to form a line.

3.1 Creating a Table from an Equation

Let's take the equation $y = 2x + 1$. We can create a table by selecting values for x and calculating the corresponding y values.

x	y
-2	-3
0	1
2	5

Now we can plot the points $(-2, -3)$, $(0, 1)$, and $(2, 5)$ on the coordinate plane. By connecting these points, we form the line representing the equation $y = 2x + 1$.

3.2 Graphing Using Point-Slope Form

You can also take advantage of an equation in point-slope form by remembering how to find the slope and y-intercept. Recall that *slope* is $\frac{\text{rise}}{\text{run}}$, meaning how much the line goes up (or down) for every unit it moves horizontally. The slope, denoted as m, tells you how steep the line is, and the y-intercept, denoted as b, is the point where the line crosses the y-axis.

The general form of a linear equation is:

$$y = mx + b$$

where:

- m is the slope (rise/run)
- b is the y-intercept

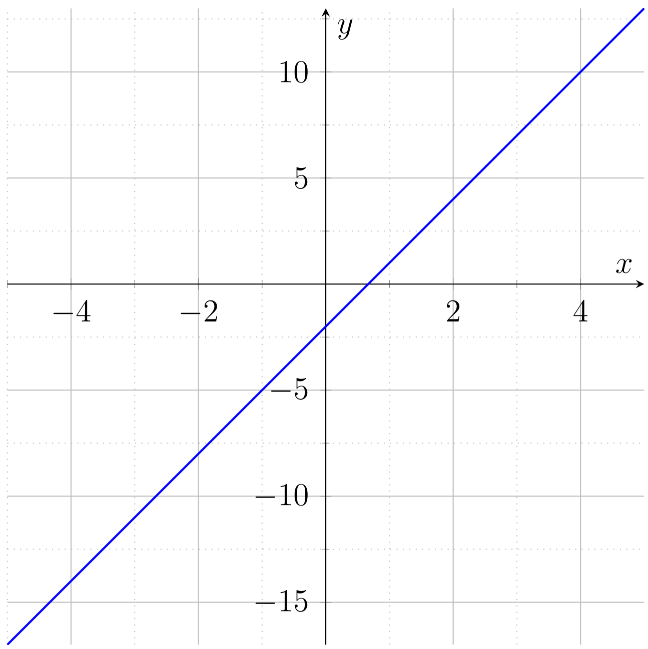
To graph an equation like $y = 3x - 2$, you can start by plotting the y-intercept, which is the point $(0, -2)$. From there, use the slope $m = 3$, which means for every 1 unit increase in x, y increases by 3 units (rise of 3 and run of 1). Plot a second point based on this slope, then draw a line through both points.

3.3 Example: Graphing Using Slope and Y-Intercept

Let's graph the equation $y = 3x - 2$. Here's how to do it:

1. The y-intercept is -2, so plot the point $(0, -2)$.
2. The slope is 3, which is $\frac{3}{1}$. This means from the y-intercept, move up 3 units and 1 unit to the right. Plot this second point at $(1, 1)$.
3. Now, connect the two points to form the line.

This is a quick and effective way to graph linear equations without needing to make a full table of values.



In this example, the line crosses the y-axis at -2 and follows a slope of 3.