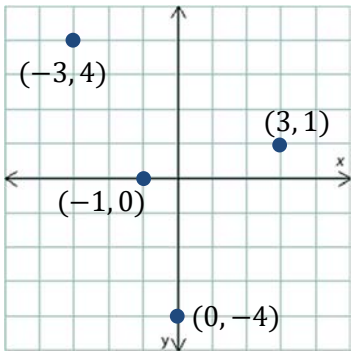
Review of Linear Equations

Graphing points and lines:

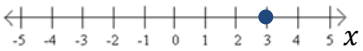
One way to think about graphing horizontal and vertical lines, is that they are just a generalization of going from one dimension to two-dimensions. On the coordinate plane, there are two axes: the x-axis, which is horizontal (right/left), and the y-axis which is vertical (up/down):



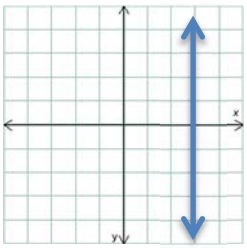
Points on a plane:

Points on a coordinate plane give the corresponding x and y-values for each spot on the plane. This involves looking at BOTH the corresponding values of x on the horizontal x-axis and y on the vertical y-axis that identify the individual point. We write this as (x, y).

We can think of this as being generated by placing a number line that represents the x on the page and sweeping it infinitely up and down through space, to generate a two-dimensional plane. If we do this, any dot that represents a value of x gets swept along with the number line axis that represents x. So, for example, if we had $x = 3$ represented by the following number line:



Then sweeping this number line infinitely up and down would generate the following blue line in the two dimensional plane, if each square on the plane represents one unit (e.g., this line is three squares to the right of zero on the x-axis):

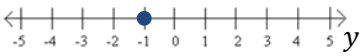


This line represents ALL the points on the plane where x is equal to 3. The value of y here (location corresponding to the vertical y-axis) varies as we move up and down the line.

So, for example, the points $(3, -6)$, $(3, 0)$, and $(3, 1000)$ are all on this line.

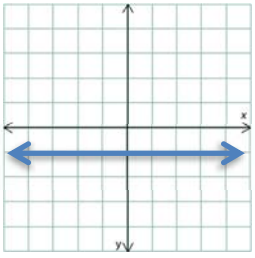
Similarly, we can think of the two-dimensional plane as being generated by placing the y-axis in a vertical position and then sweeping it infinitely left and right. For example, if we look at the number line representation for $y = -1$, we get:

Original number line y-axis



Number line (y-axis) placed vertically

Number line (y-axis) placed vertically and swept infinitely left and right through space to generate a two dimensional plane



This line represents ALL the points on the plane where y is equal to -1. The value of x here (location corresponding to the horizontal x-axis) varies as we move left and right on the line.

Here, for example, the points $(-8, -1)$, $(0, -1)$, and $(10, -1)$ are all on this line.

Slope of a line:

The slope of a line describes the change in the y-values over (divided by) the change in the x-values as we move from one point on the line to another. All we have to do is to identify two points on the line, and then count the change in the y-value from one point to the next (the vertical distance, keeping track of whether it is in the positive or negative direction), and then count the change in the x-value from one point to the next (the horizontal distance, keeping track of whether it is in the positive or negative direction).

Slope is a very important idea in linear regression and other forms of linear modeling. It depicts the **rate of change**, or the average **proportional relationship between x and y**. In other words, the slope depicts a relationship between x and y in which a **single unit change** in x always results in the **same fixed change** in y (we note that in linear regression, this is an average change—the slope of the fit line, not the slope literally as we move from one single data point to the next, which will vary quite a lot).

- For example, every time x increases by 1, y might increase by 2; this would be a slope of $\frac{2}{1}$.
- Or, every time x increases by 1, y might increase by $\frac{1}{2}$; this would be a slope of $\frac{1}{2}$ (every time x increases by 2, y increases by 1; so every time x increases by 1, y increases by $\frac{1}{2}$; we can also represent this mathematically as $\frac{1}{2} = \frac{\frac{1}{2}}{1}$).
- Or, every time x increases by 1, y might decrease by 4; this would be a slope of $\frac{-4}{1}$ or $-\frac{4}{1}$ (remember that $-\frac{4}{1} = \frac{-4}{1} = \frac{4}{-1}$ because a negative divided by a positive or vice versa always yields a negative answer).
- Or, every time x increases by 1, y might decrease by 0.25; this would be a slope of $\frac{-1}{4}$ or $-\frac{1}{4}$ (every time x increases by 4, y decreases by 1; so every time x increases by 1 y decreases by $\frac{1}{4}$; we can also represent this mathematically as $\frac{-1}{4} = \frac{-0.25}{1}$).

Here is a more formal definition of slope, if you find it helpful:

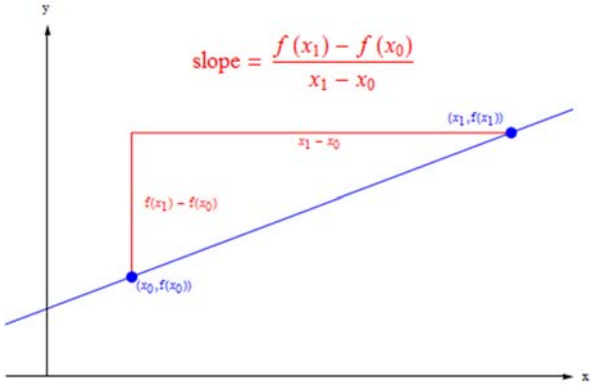
Slope (formal definition): Here is one formal way of representing the calculation as we have described it above, that you may have seen in earlier classes:

For two points (x_0, y_0) and (x_1, y_1), the slope can be calculated as:

$\frac{y_0 - y_1}{x_0 - x_1}$ or equivalently, $\frac{y_1 - y_0}{x_1 - x_0}$

(The main issue is simply that both y-values have to be at the top, both x-values have to be at the bottom, and the x- and y-values that correspond with one another [are in the same point] have to line up vertically to ensure that we are keeping track of the directions consistently.)

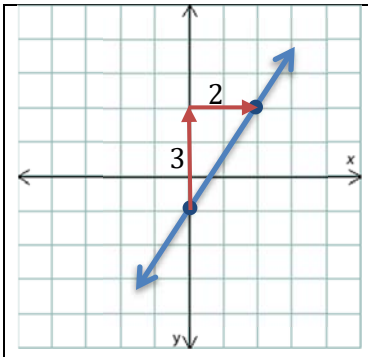
Graphically, this looks like this:



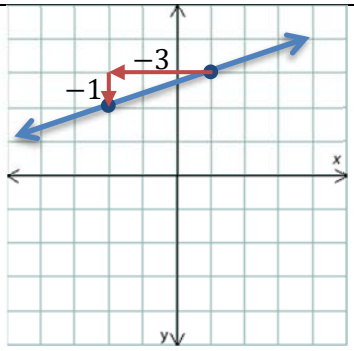
where the sign on the y-distance indicates whether one has to go up (positive) or down (negative) to go from (x_0, y_0) to (x_1, y_1), and the sign on the x-distance indicates whether one has to go to the right (positive) or to the left (negative) to go

from (x_0, y_0) to (x_1, y_1). (We note here that in this picture, $f(x_0)$ is just a fancy way of writing y_0 and $f(x_1)$ is just a fancy way of writing y_1—it is used to stress the fact that y_0 is a function of x_0 or that it is an output from plugging in x_0 in as an input into an equation).

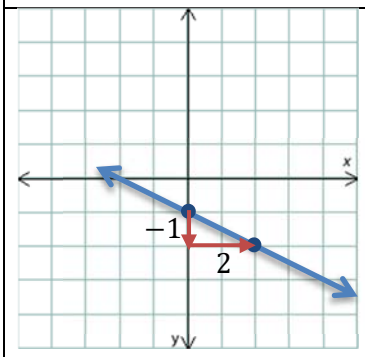
Let's look at some examples:



slope = $\frac{3}{2}$

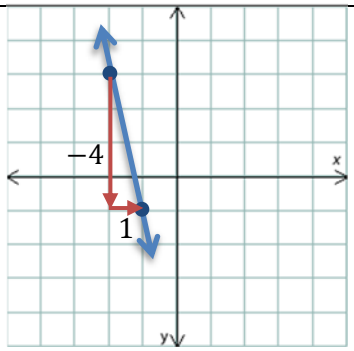


slope = $\frac{-1}{-3} = \frac{1}{3}$

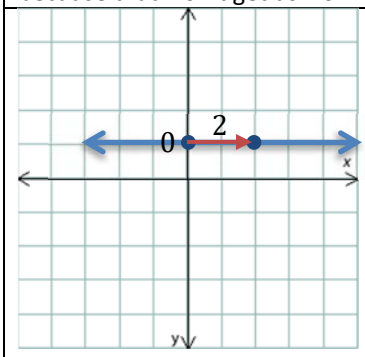


slope = $\frac{-1}{2} = -\frac{1}{2}$

Note: Be careful about the signs! We can start from the point $(0\ , -1)$ at the left and count toward the point $(2, -2)$, which is what we have done here, or we can start from the point $(2, -2)$ on the right and count toward the point $(0\ , -1)$ (in which case we would get slope = $\frac{1}{-2} = -\frac{1}{2}$ instead, which is equivalent since both $\frac{-1}{2}$ and $\frac{1}{-2}$ are equal to $-\frac{1}{2}$). They key thing here is that we have to keep track of the direction for both parts—either we have to go down (negative y direction) and then to the right (positive x direction), or we have to go up (positive y direction) and then to the left (negative x direction). It doesn't matter which one we choose, but we can't go up and to the right, or down and to the left, because that won't get us from one point on the line to another.

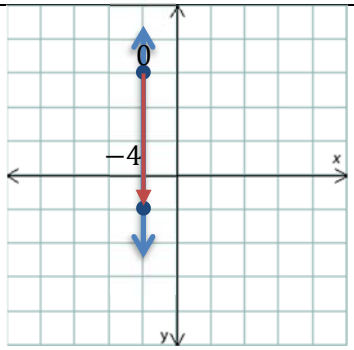


slope = $\frac{-4}{1} = -4$



slope = $\frac{0}{2} = 0$

Note: Horizonal lines have a slope of zero. This makes sense because if we pick any two points on the line, the distance between the two y-values will always be zero (because all y-values on the graph are equal to 1 here), but the distance between the two x-values will always be something that is non-zero. Because zero divided by any non-zero number is zero, we can see that the slope must be zero.



slope = $\frac{-4}{0} \rightarrow undefined$!

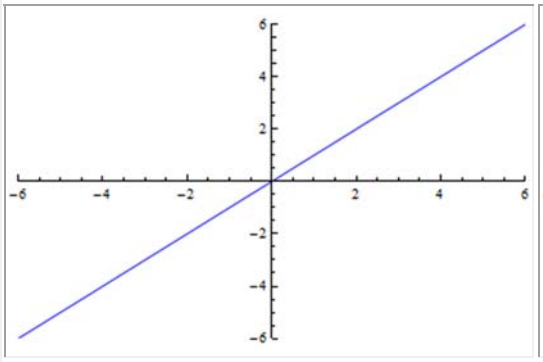
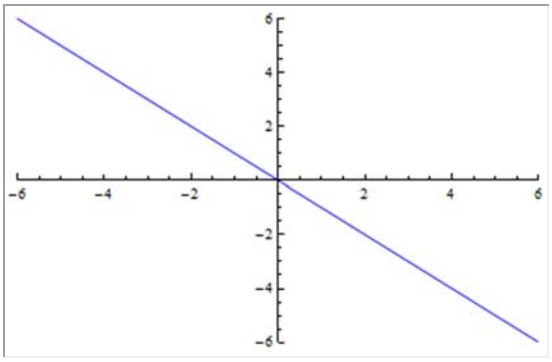
Note: Vertical lines have a slope that is undefined. This makes sense because if we pick any two points on the line, the distance between the two x-values will always be zero (because all x-values on the graph are equal to -1 here). Because anything divided by zero will be undefined (because we can't take some amount and break it up into zero pieces), we can see that the slope must be undefined.

Some Important properties of slope:

We notice that a slope has two key pieces of information:

1) Sign of the slope (positive/negative)

If the line is *increasing* from left to right, it has a *positive* slope (because as we go up, in the positive y-direction, we also go to the right, in the positive x-direction, and dividing a positive by a positive gives us a positive). **If the line is *decreasing* from left to right, it has a *negative* slope** (because as we go down, in the negative y-direction, we also go to the right, in the positive x-direction, and dividing a negative by a positive gives us a negative).

So this line has a positive slope:	And this line has a negative slope:
	

A line with a **positive** slope depicts a **direct** relationship in which y goes in the **same** direction as x (y always increases as x increases, or decreases as x decreases).

> (For example, as years of parental education (x) go up, a student's likelihood of completing a college degree also on average goes up (y). This relationship would be depicted by a positive slope on the fit line.)

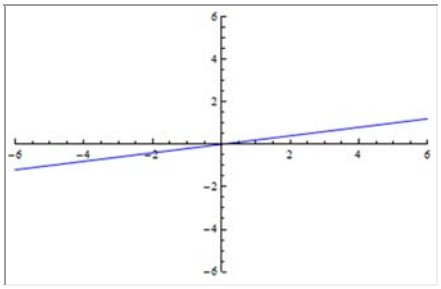
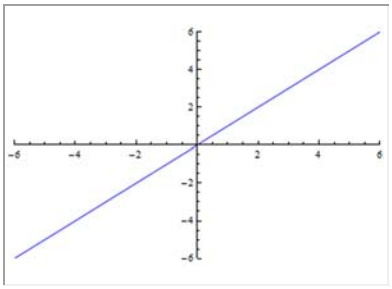
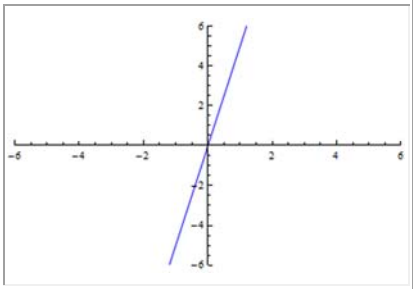
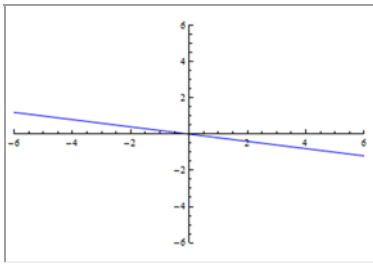
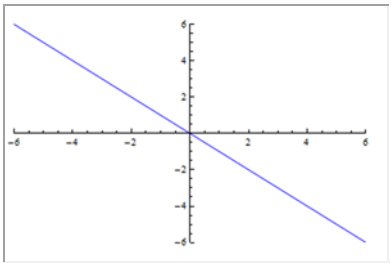
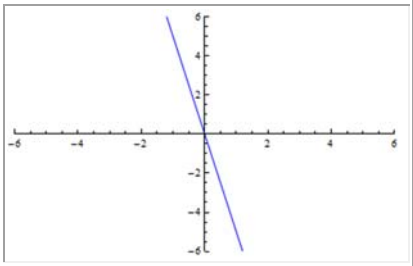
A line with a **negative** slope depicts an **inverse** relationship in which y goes in the **opposite** direction of x (y always decreases when x increases, or increases when x decreases).

> (For example, as the number of hours of paid and unpaid work outside of academics that a student has (x) go up, a student's likelihood of completing a college degree goes down on average (y). This relationship would be depicted by a negative slope on the fit line.)

2) "Steepness", or the **magnitude** of the slope:

The steeper the line, the greater the magnitude of the slope, and the stronger the relationship between x and y. For example, in the table below, all the lines in the first row have positive slope, and all the lines in the second row have negative slope; but in both the first row and the second row, the slope of the lines get steeper as we move from left to right, because the magnitude of the slope in both cases gets larger (the magnitude is just the size of the absolute value, or the distance from zero). Thus, the relationship between x and y gets **stronger** as we move from left to right.

But be careful not to confuse size with magnitude!: While the lines in the second row get steeper as we go from left to right, the slopes actually get smaller because a negative number with a larger magnitude is actually smaller. For example, -1 has a smaller magnitude than -5 because -1 is closer to zero, but $-5 < -1$ because -5 is to the left of -1 on the number line. (To avoid this problem, it is best to talk about stronger or weaker relationships, larger or smaller magnitudes, or more or less steep slopes, rather than talking about "bigger" or "smaller" numbers.)

	both lines have a slope that is not very steep (slope magnitude is less than 1)	**both lines have a slope that is steeper than the line to the left (slope magnitude is equal to 1)**	**both lines have a slope that is steeper than either line to the left (slope magnitude is greater than 1)**
lines with positive slope			
lines with negative slope			

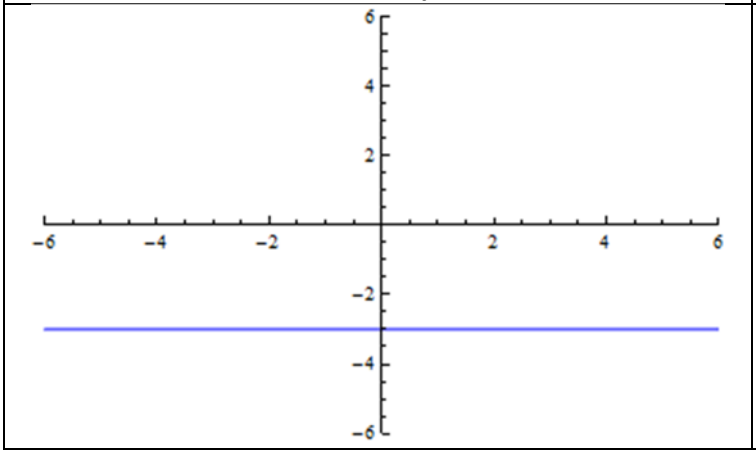
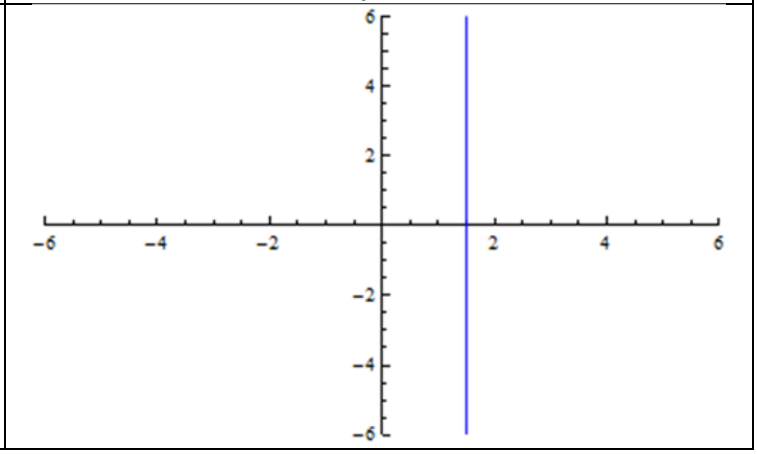
There are also a few other examples of what can happen with the "steepness" or magnitude of a slope:

In the table below, the **horizonal line** to the left has a **slope of zero**.

- This makes sense because if we pick any two points on the line, the distance between the two y-values will always be zero (because all y-values on the graph are equal to -3), but the distance between the two x-values will always be something that is non-zero. Because zero divided by any non-zero number is zero, we can see that the slope must be zero.
- This happens when y is always the same, no matter what the value of x is. In statistics, this would happen if there is NO relationship between x and y; in other words, y does not change on average when we vary values of x (in reality, y may have many different values, but when we model them statistically, the overall relationship is one in which these values do not change *on average* based on the value of x). For example, if student ID numbers are assigned randomly, then a student's likelihood of completing a college degree (y) will have NO relationship to their student ID number (x), and the slope of any line fitted to this data will on average be zero.

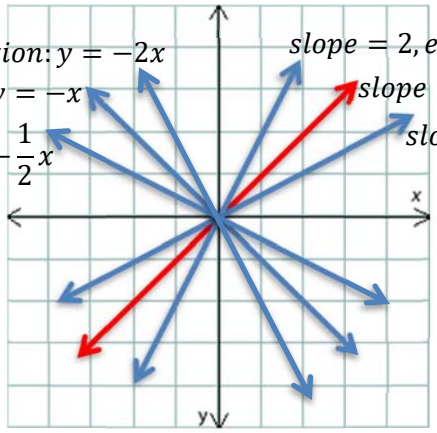
In the table below, the **vertical line** to the right has a slope that is **undefined**.

- This makes sense because if we pick any two points on the line, the distance between the two x-values will always be zero (because all x-values on the graph are equal to 2.5). Because anything divided by zero will be undefined (because we can't break something up into zero groups, or groups of size zero), we can see that the slope must be undefined.
- This is not typically a case we encounter in statistics with any reasonable type of dataset, but it can still be important to understand what an undefined slope would look like and what it means. A line with an undefined slope would depict a case in which values of x never vary, but where y takes on many different values. For example, if we collected data on students where we wanted to analyze the relationship between a student's grade level (x) and their score on a set of math items assessing a particular mathematical conception (y), but we only collected data from students in a single grade level, our attempts to graph the relationship between x and y would look like the graph on the right. It becomes obvious why we cannot analyze the relationship between x and y in that case: we don't have any variation in the value of x, so any attempt to model the relationship with a line with yield something with undefined slope, and no ability to say anything about the relationship between x and y.

this line has a slope of zero	**this line has a slope that is undefined**
	

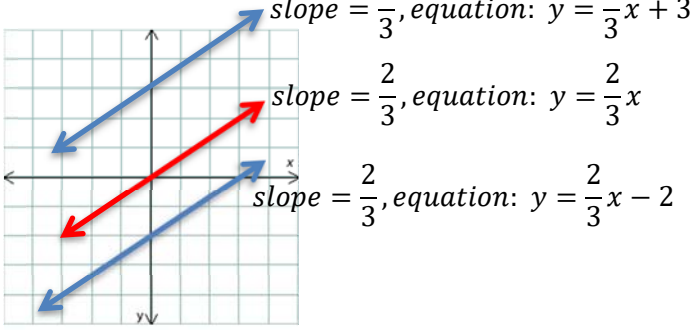
Intercepts and equations of lines

We now want to write the equations for a line, which will describe the relationship between the x and y values for all points on the line. Consider the red line below. We can see that for every value of x, y has exactly the same value. So this line can be described by the equation $y = x$, because for all points on the line, this equation will always be true. If we look at a similar line but with slope of 2, we can see that for every value of x for every point on the line, the value of y will be twice the value of x, so the equation that describes the points that line is $y = 2x$. Similarly, for the line with slope $\frac{1}{2}$ depicted here, every y value is actually half of the x value, for each point on the line, so the equation that explains the relationship between the x and y values for all points on the line is $y = \frac{1}{2}x$. We can generate similar equations for lines with negative slope.



But we notice that all of these lines go through the origin (what we call the point $(0, 0)$). But what about lines that are somewhere else in the plane? How do we describe them?

We know how to find the slope of any line if we have two points on the line—just by counting the y-distance and the x-distance from one point to the next. But this is not enough information for us to identify a unique line in the plane. For example, all of the following lines have the same slope ($slope = \frac{2}{3}$):



How can we generate the equation for each of these different lines? One thing we notice is that they are simply shifted up or down on the y-axis in each case. So the red line in the middle is represented by the equation $y = \frac{2}{3}x$, because for every value of x, to get y we have to take two-thirds of x to get the corresponding y-value for that point on the line. But for the line above that red line, we have to shift all of our y-values up by three in order for them to correspond to the right x value: on the red line, $x = 3$ gives us $y = 2$, but for the line above, for $x = 3$ we need to have $y = 2 + 3 = 5$. If we look at every point on the line $y = \frac{2}{3}x$, we have to add 3 to the result of $\frac{2}{3}x$ in order to get the correct y value, since the line is shifted up by 3. So the equation of the top line is $y = \frac{2}{3}x + 3$.

Similarly, the line below the red line is shifted down 2, and so each value of y has to be two less than what we get after multiplying $\frac{2}{3}$ by x, so the equation of that line is $y = \frac{2}{3}x - 2$.

This helps us to find the equation of a line:

- Every line with a defined slope (e.g. not vertical) can be described by an equation of the form: $y = mx + b$ where m represents the slope of the line and b represents the amount the line has been shifted up or down from the origin. Here b is also referred to as the y-intercept, or the y-value at the place where the line crosses the y-axis. Sometimes we write the y-intercept as a point, like this: $(0, b)$. One way to find the y-intercept from the equation of any line (even if it is not in the form $y = mx + b$) is simply to substitute in 0 for x and see what value we get for y: because that will always describe where the line crosses the y-axis, because the y-axis is the one place in the plane where x always equals zero.

  The Intercept:

  In statistics, the ***y*-intercept** on a fit line describes the average value of y when x is zero. So, if x represents the number of hours that a student spends each week on paid and unpaid work (not counting their schoolwork), and y represents the number of credits that a student earns in college that semester, then the y-intercept in this case describes how many credits a student would earn on average if they had **no** paid/unpaid work commitments outside of their schoolwork (although we note that the intercept may not actually have real-life interpretability if NO data points fell at or near zero—in that case, the intercept is really just a tool for generating the fit line, and not interpretable in and of itself).

- (It doesn't help us with equations of lines, and is not often used directly in statistical analysis, but we can also find what is called the x-intercept, or where the line crosses the x-axis.)

Linear equations in more than two-dimensions:

All of the descriptions above have been for two-dimensional representations of lines with only two variables. In reality, when we use linear regression we may have multiple dependent variables, for example like this:

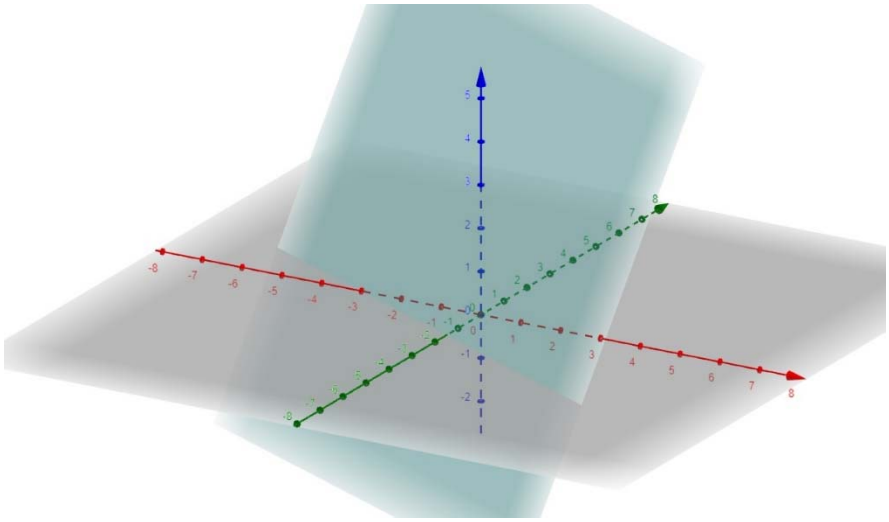
$$y = a_1x_1 + a_2x_2 + b$$

Or even more generally, like this (for any possible positive whole number n):

$$y = a_1x_1 + a_2x_2 + \cdots + a_nx_n + b$$

There are other readings on multiple linear regression, but here we discuss just the basic mathematical understanding needed in order to understand lines in more than two dimensions insofar as this will be necessary to interpreting multivariate linear regression.

Firstly, the line $y = a_1x_1 + a_2x_2 + b$ represents a two-dimensional plane in space. For example, the equation $y = x_1 + 2x_2 + 3$ is depicted in the image below (the blue tilted plane), where the vertical blue axis is the y-axis, the red axis from left to right is the x_1-axis, and the green axis which could be visualized as going back to forward is the x_2-axis.



However, in reality, we rarely need to think about these models in more than two dimensions. Why? Because we tend to interpret coefficients one at a time. In that case, if we want to think about the relationship between y and x_1, we simply hold x_2 constant and discuss what the conditional relationship is between y and x_1 when x_2 remains the same. In this case, we can pick any value of x_2 (called the reference value, but really any value will work, because it will not change the slope a_1) and consider the resulting two-dimensional line. For example, if we let $x_2 = 0$, then we get the line $y = x_1 + 3$ (which we note is one line on the plane above—it is the line along the blue plane that goes intersects the plane formed by the x_1-axis [where $x_2 = 0$] and the y-axes). The coefficient of x_1 is still 1 in both the full equation $y = x_1 + 2x_2 + 3$ and in this simplified equation when $x_2 = 0$, $y = x_1 + 3$. This means that as long as x_2 is held constant, the value of y always goes up by 1 every time x_1 goes up by 1 (on average, if this is the equation of a fit line). When using simple linear models like this, this is true no matter what value we pick as the reference value for x_2: for example, if we had chosen $x_2 = 10$, the resulting simplified line would have been $y = x_1 + 13$ (this is again a line on the plane, just the one that intersects the plane formed by the line $x_2 = 10$ and the y-axis). The slope is still 1, and so again, as long as x_2 is held constant, the value of y always goes up by 1 every time x_1 goes up by 1 (on average, if this is the equation of a fit line). The key point is that we can only compare the relationship between x_1 and y if we hold x_2 constant. So, we can only describe the relationship between the dependent variable and one independent variable at a time. But this allows us to ignore all the complications of the plane, and just focus on two-dimensional lines.

We can also think of this as controlling for x_2: this tells us what the relationship is between x_1 and y after we have controlled for variation in x_2. So, for example, if y represents the number of credits that a student earns in college, x_1 represents the number of hours that a student spends studying each week, and x_2 represents the student's GPA, looking at a_1 would tell us what the relationship is between the number of hours that a student spends studying each week and the number of credits they earn, when comparing students with the *same GPA*. Similarly, looking at a_2 will tell us the relationship between GPA and the number of credits that a student earns on average, when comparing students who spend the *same* amount of time studying each week. Because of this, in practice we can think of the lines separately when we are making inferences, something like this:

- The average relationship between credits earned and hours spent studying weekly, when controlling for GPA (or holding GPA constant) is $y = a_1x_1 + C$ (where C represents the value of $a_2x_2 + b$ for some fixed value of x_2). The value of C isn't really important to understanding what a_1 is or how to interpret it.
- Similarly, the average relationship between credits earned and GPA, when controlling for the number of hours spent studying each week (or holding hours spent studying constant) is $y = a_2x_2 + C$ (where C represents the value of $a_1x_1 + b$ for some fixed value of x_1).

So it is good enough to be able to visualize two dimensional lines in a plane when interpreting regression coefficients. We don't really need to visual 3D planes, or n-dimensional spaces (although it is useful to realize that the representation of the *whole* regression equation in multi-level regression is more complex than a simple two-dimensional line).