MTH1W – Unit 9: Statistics

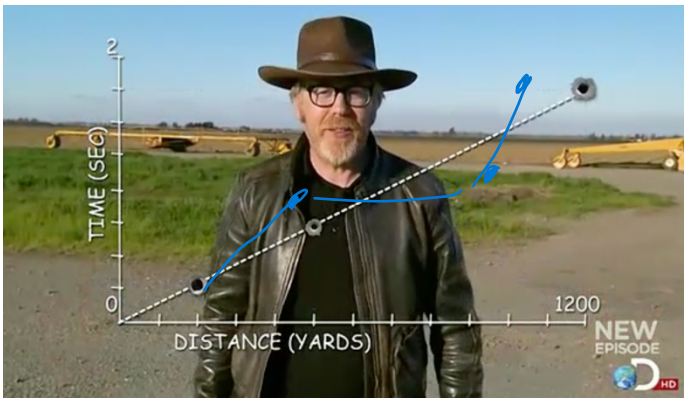
Name: Mr Hagen

Lesson #9.4: Scatter Plots and Lines of Best Fit

Date: May 25, 2023

Learning Goal: We are learning to create a line of best fit from a scatter plot.

Remember watching Adam and Jamie gather data points on the length of time it took for a bullet to travel to a target? After they had three data points, Adam was able to draw a graph known as a scatter plot. He was extremely fortunate that the three data points sat perfectly in a line. Taking this one step farther, we could also determine the equation of this line, thus allowing us to calculate any time or distance.



$b = -mx + y$

Today, we will take data from a table and plot it on a scatter plot. Then, we will determine the ***line of best fit***. This is a line in $y = mx + b$ form. We will also utilize [redacted] to create that line. The line of best fit best describes the relationship between the data points. This, like standard deviation, is not difficult, but it is a long and tedious task. However, on small data sets, it is completely possible to do by hand.

This process is called statistical modelling. We are only looking at creating a line, but you can do this with curves and make models of the trajectory of a rocket launch, the growth of a virus, the housing market, the sales for your company, and so much more! This is IMMENSLEY useful.

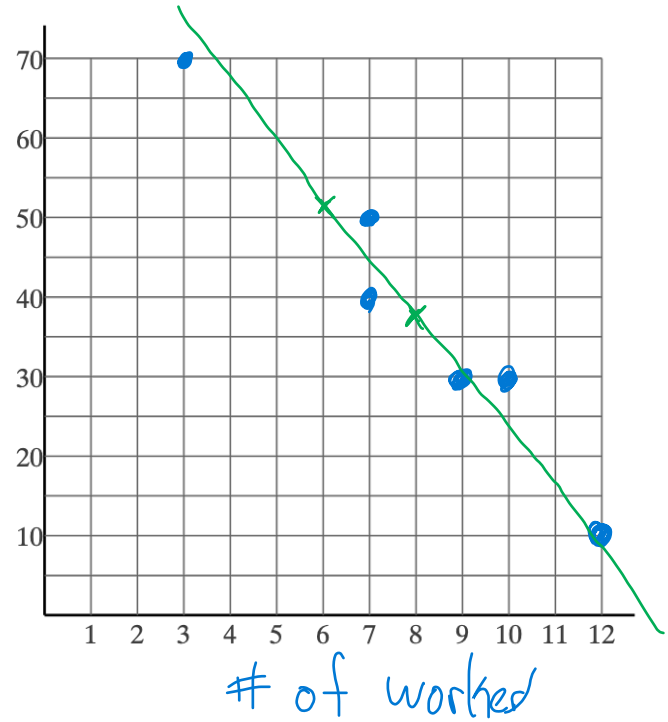
Example: Plot the points, calculate the line of best fit, then draw the line.

1. Plot the points.
2. Calculate the averages for the x and y coordinates.
3. Fill in the table (next page)
4. Calculate slope.
5. Determine $y = mx + b$.
6. Graph the line.
7. Answer any questions!

# of workers

hours to complete a job

X	Y
3	70
7	40
7	50
9	30
10	30
12	10



6. $x_a = \frac{48}{6} = 8$ ← average

$y_a = \frac{230}{6} = 38.3$

(3)

x	y	$x - x_a$	$(x - x_a)^2$	$y - y_a$	$(x - x_a)(y - y_a)$
3	70	$3-8=-5$	$(-5)=25$	31.7	$(-5)(31.7)=-158.5$
7	40	$7-8=-1$	1	1.7	$(-1)(1.7)=-1.7$
7	50	$7-8=-1$	1	11.7	-11.7
9	30	$9-8=1$	1	-8.3	-8.3
10	30	$10-8=2$	4	-8.3	-16.6
12	10	$12-8=4$	16	-28.3	-113.2

8 — 38.3

Run = add

Rise = add column

#6 use $x=6$, $y=?$

$y = -6.45(6) + 90$

$y = 51$ (6, 51)

use $x=8$, $y=?$

$y = -6.45(8) + 90$

$y = 38$ (8, 38)

(4) Slope

$m = \frac{\text{Rise}}{\text{Run}} = \frac{-310}{48} = -6.45$ (use 2 decimals)

(5) need b, using x_a and y_a

$b = -m x_a + y_a$

$b = -(-6.45)(8) + 38.3$

$b = 51.6 + 38.3 = 89.9 = 90$

$\therefore y = -6.45x + 90$

Now what? The best part is that we can now answer questions!

a) If we had one worker, how many hours would it take?

$x = 1$

$y = -6.45(1) + 90$

$y = 83.55$ hours

b) I need a job done in 20 hours, so how many workers should I hire?

$y = 20$

$20 = -6.45(x) + 90$

$-90 \quad -90$

$-70 = -6.45(x)$

$-6.45 \quad -6.45$

$10.85 = x$ $\therefore$ 11 workers.

Success Criteria:

- I can plot points on a scatter plot
- I can calculate the line of best fit