Section 1.5 Math 141
The Method of Least Squares

Main ideas

Given data, we can find the line $y = Ax + B$ (I'm more used to notation $y = mx + b$, and maybe you are too, but we'll follow the book's notation) that best fits the data.

This line is called the <u>best fit</u> line or <u>least squares</u> line. The process of finding and using this line is sometimes called <u>linear regression</u>. We can also find all sorts of other types of functions (other than straight lines) to fit a given set of data, but that is beyond what we are doing in this class.

The slope A and $y-$intercept B give us useful information from our data points (x, y) about how y is related to x.

We can use the line $y = Ax + B$ to:
- Estimate the value of y that would result from a particular value x.
- Estimate the value of x that would be needed to produce a particular y.

In Class

1. Let's look at the formulas in the book for how to find the line that best fits given data. You don't need to know the formulas for the text, but you will for the homework.

2. You can also use technology to find the line. Let's use Excel to help us with the arithmetic. (You are welcome to do this on your homework.) Let's use Excel to simply find the line. (You should not do this on the homework. Show your own work.) As we saw, you can use technology to find many other models/functions to fit given data.

3. What do the slope and y-intercept of $y = .338 + 21.621$ mean? What do they tell us?

4. Suppose a country estimates its cigarette "consumption" is 400 per person per year. How many lung cancer deaths can it expect?

5. Suppose a country want to reduce its lung cancer death rate to 50 deaths (per million males per year). What smoking (cigarette consumption) rate do they need to reach?

Note: for HW 1.5.12, 13 and 14, the book suggests that you convert *years* to number of *years after the initial year given*. For example, think of the years 2012 – 2015 as 0 – 3 (the number of years after 2012). The book gives its solutions with this change.