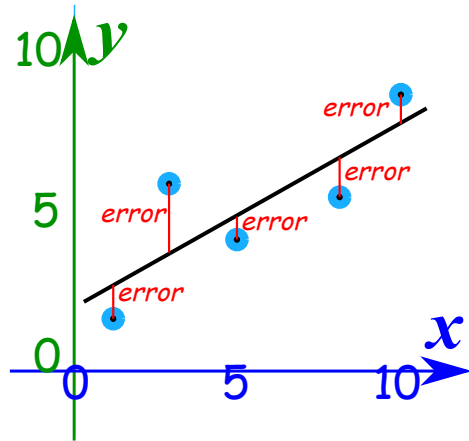


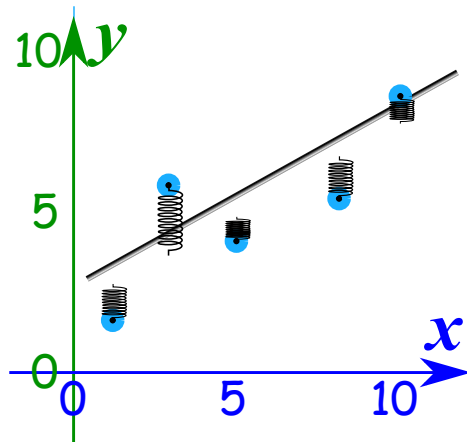
How does it work?

It works by making the total of the **square of the errors** as small as possible (that is why it is called "least squares"):



The straight line minimizes the sum of squared errors

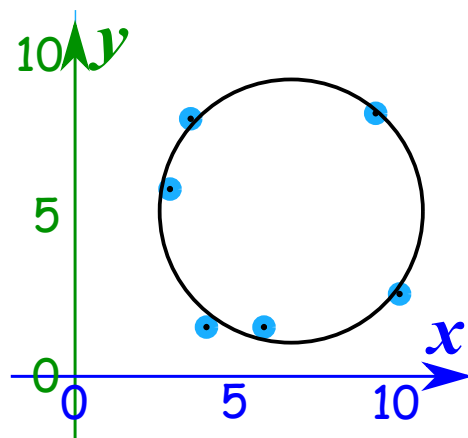
So, when we square each of those errors and add them all up, the total is as small as possible. You can **imagine** (but not accurately) each data point connected to a straight bar by springs:



Boing!

Outliers Be careful! Least squares is sensitive to **outliers**. A strange value will pull the line toward it.

Not Just For Lines This **idea** can be used in many other areas, not just lines.



A "circle of best fit"

But the formulas (and the steps taken) will be very different!