Derivation of the formulas

$$E = (Ax_1 + B - y_1)^2 + \cdots + (Ax_N + B - y_N)^2$$

$$\begin{aligned}
\frac{\partial E}{\partial A} &= 2(Ax_1 + B - y_1)\cdot x_1 + \cdots + 2(Ax_N + B - y_N)\cdot x_N \\
&= 2A(x_1^2 + \cdots + x_N^2) + 2B(x_1 + \cdots + x_N) - 2(x_1y_1 + \cdots + x_Ny_N) \\
&= 2A(\sum x^2) + 2B(\sum x) - 2(\sum xy)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial E}{\partial B} &= 2(Ax_1 + B - y_1)\cdot 1 + \cdots + 2(Ax_N + B - y_N)\cdot 1 \\
&= 2A(x_1 + \cdots + x_N) + 2B(1 + \cdots + 1) - 2(y_1 + \cdots + y_N) \\
&= 2A(\sum x) + 2B(\ N\ ) - 2(\sum y)
\end{aligned}$$

Derivation of the formulas

$$\frac{\partial E}{\partial A} = 2A(\sum x^2) + 2B(\sum x) - 2(\sum xy) = 0$$

$$\frac{\partial E}{\partial B} = 2A(\sum x) + 2B(N) - 2(\sum y) = 0$$

Solve for *A* and *B*

$$(\sum x^2)A+(\sum x)B=\sum xy$$

$$(\sum x\ \ )A+(\ N\ )B=\sum y$$

$$\begin{bmatrix}\sum x^2 & \sum x\\ \sum x & N\end{bmatrix}\begin{bmatrix}A\\ B\end{bmatrix}=\begin{bmatrix}\sum xy\\ \sum y\end{bmatrix}$$

$$\begin{bmatrix}\sum x^2 & \sum x\\ \sum x & N\end{bmatrix}^{-1}=\frac{1}{N(\sum x^2)-(\sum x)^2}\begin{bmatrix}N & -\sum x\\ -\sum x & \sum x^2\end{bmatrix}$$

so $$\begin{bmatrix}A\\ B\end{bmatrix}=\frac{1}{N(\sum x^2)-(\sum x)^2}\begin{bmatrix}N & -\sum x\\ -\sum x & \sum x^2\end{bmatrix}\begin{bmatrix}\sum xy\\ \sum y\end{bmatrix}$$

Solve for A and B

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{N(\sum x^2)-(\sum x)^2} \begin{bmatrix} N & -\sum x \\ -\sum x & \sum x^2 \end{bmatrix} \begin{bmatrix} \sum xy \\ \sum y \end{bmatrix}$$

$$= \frac{1}{N\cdot\sum x^2-(\sum x)^2} \begin{bmatrix} N(\sum xy)+(-\sum x)(\sum y) \\ (-\sum x)(\sum xy)+(\sum x^2)(\sum y) \end{bmatrix}$$

$$A = \frac{N(\sum xy)-(\sum x)(\sum y)}{N(\sum x^2)-(\sum x)^2}$$

$$B = \frac{(-\sum x)(\sum xy)-(\sum x^2)(\sum y)}{N(\sum x^2)-(\sum x)^2} = \frac{\sum y - A(\sum x)}{N}$$

Lung cancer deaths as a result of smoking

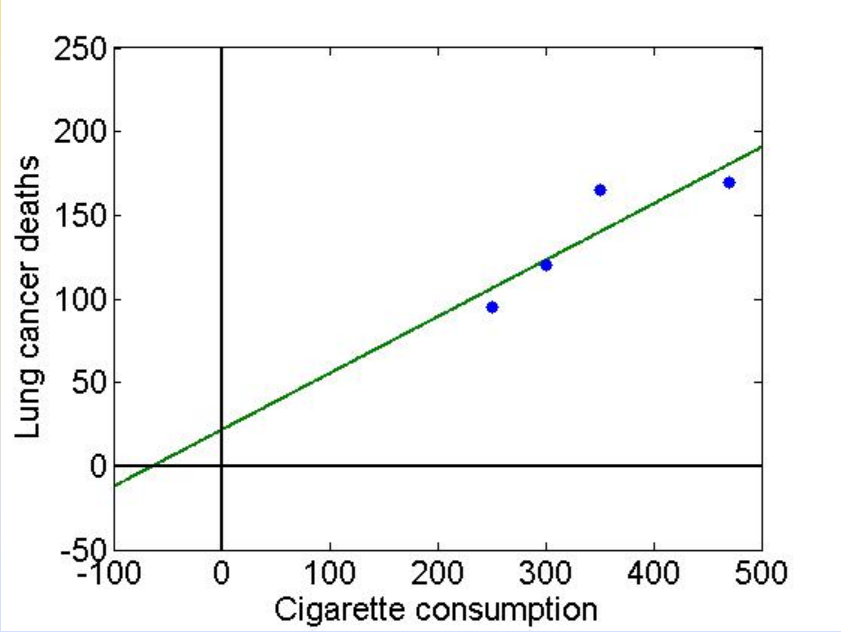
Country	Cigarettes	Lung Cancer Deaths
Norway	250	95
Sweden	300	120
Denmark	350	165
Australia	470	170

Best fit line: $y = .338x + 21.621$

↑ Deaths ↑ Cigs

What do .338 and 21.621 represent?

$y = .338x + 21.621$

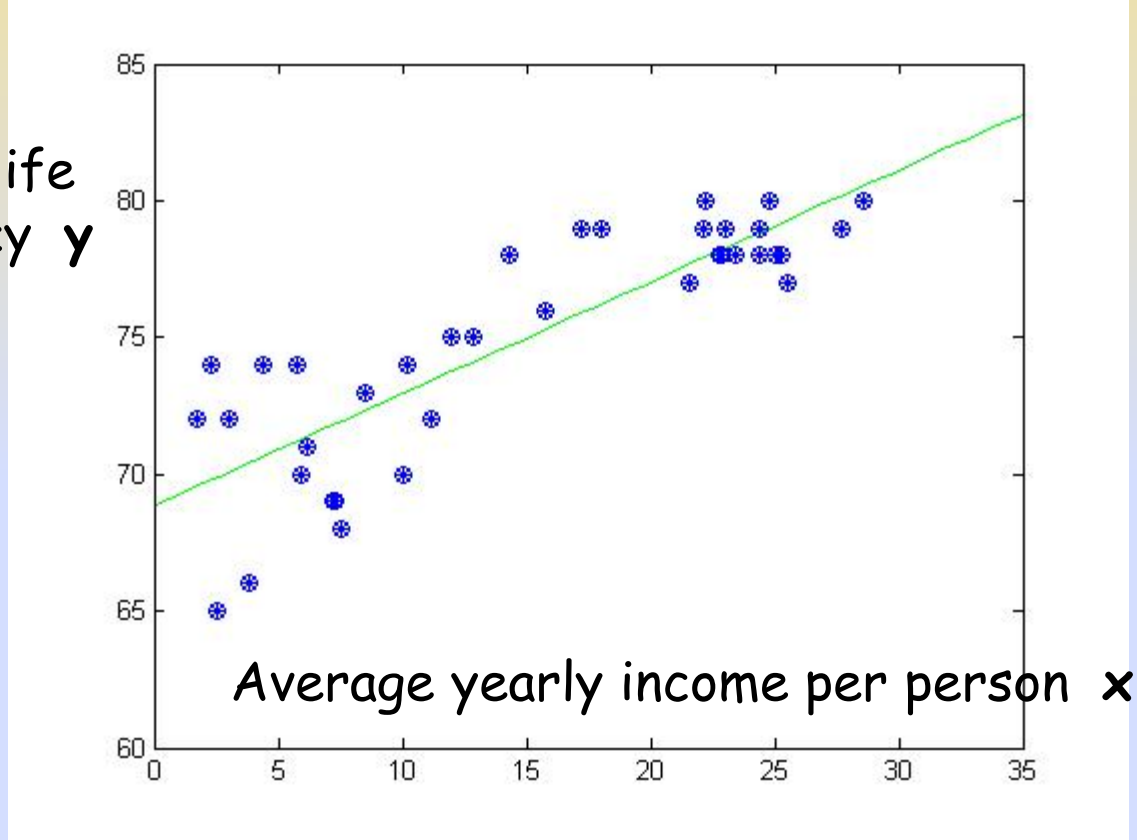


Suppose a country estimates that its cigarette consumption is 400 per person per year.
What lung cancer deaths can it expect?

Suppose a country wants to reduce its lung cancer death rate to 50 deaths (per million males per year).
What do they need to reduce the smoking rate to?

$$y = .4085x + 68.85$$

Age Income



What do .4085 and 68.85 represent?