1 Independent & Dependent Variables:

In the equation,

$$y = mx + c$$

we treat x as the independent variable.
Since, values of y depend on the values of x, y is treated as the dependent variable and we write it as $y = f(x)$.
The Sum Squared Error (SSE) is always calculated for the dependent variable.

2 Line Fitting:

To understand the concept of line fitting, let us take a question.

Table AT 3.1 shows the output values with respect to the function $f(t)$ for some input values t.

$f(t)$	1	4	3	8	7
t	1	2	3	4	5

Table AT 3.1

There are equations of four straight lines given as $f = t + 2$, $f = 2t + 2$, $f = 2t$, and $f = t$. Which equation is better fit for the given data?

Solution:
From the question and table it is clear that t is an independent variable and f is a dependent variable.
Figure AT 3.1 shows the relation between f and t.

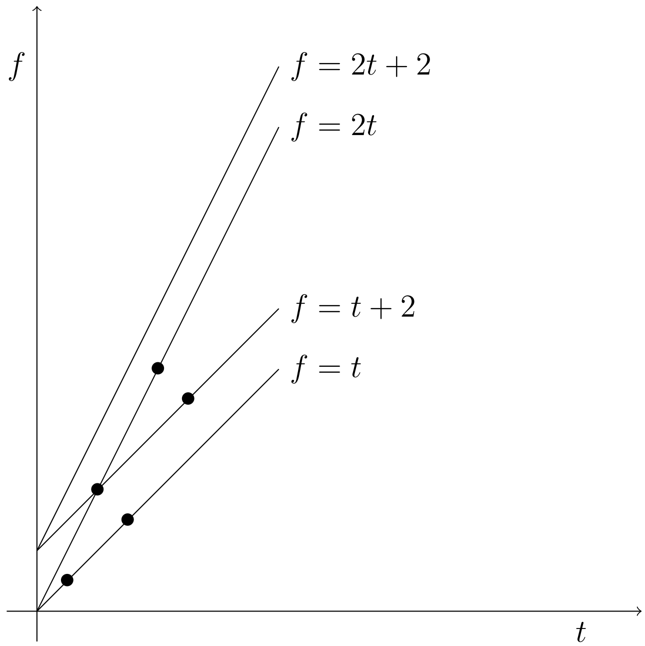

Figure AT 3.1

Here clearly all the lines except $f = 2t+2$ pass through two points. But we can not conclude which is the better fit here. To understand this, we need some algebraic approach and that can be done using sum squared error (SSE).

Table AT 3.2 shows the SSE calculated for one line $f = t + 2$.

	$f(t)$	1	4	3	8	7
	t	1	2	3	4	5
	$t + 2$	3	4	5	6	7
Errors :	$f(t) - t - 2$	-2	0	-2	2	0
Squared Errors:	$(f(t) - t - 2)^2$	4	0	4	4	0

Table AT 3.2

SSE = 4+0+4+4+0 = 12

After calculating SSE for all four equations:
SSS(t) = 24
SSS($t + 2$) = 12
SSS($2t$) = 19
SSS($2t + 2$) = 67

Therefore, $f = t + 2$ is the best fit among the given lines. This can be seen in figure too as it passes through two points and has the less distance from the other points comparatively.