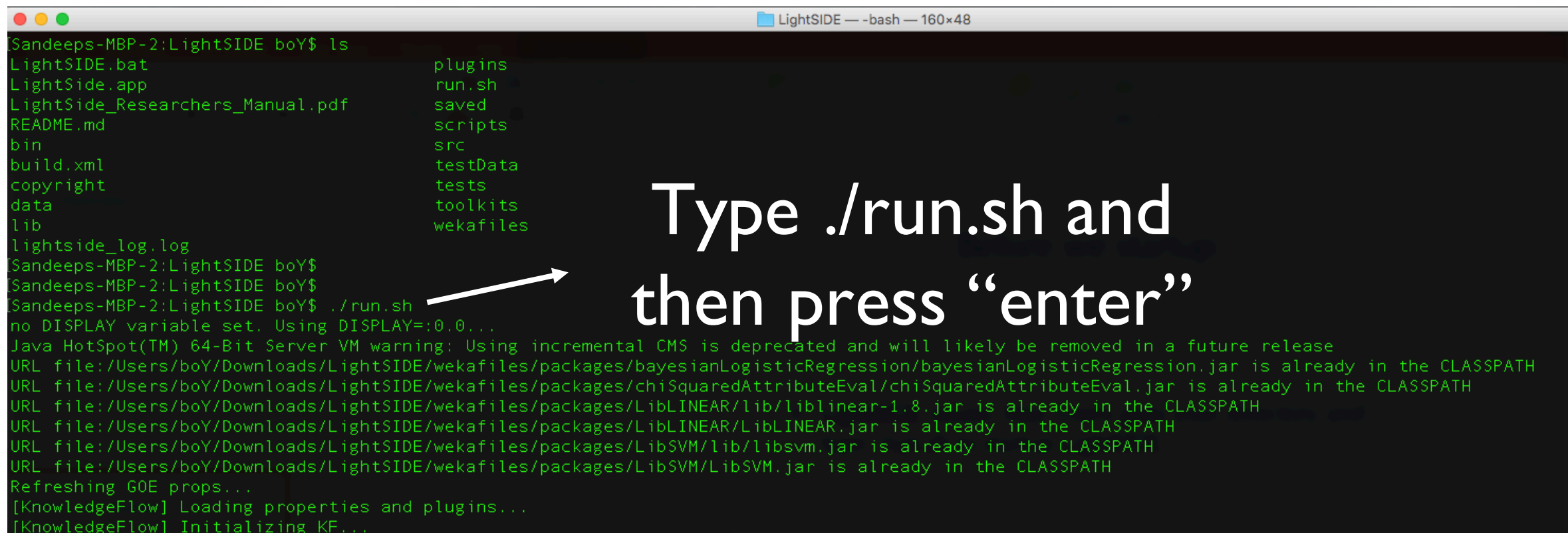# Linear Classifiers

Sandeep Avula
asandeep@unc.edu

Before we startup

- LightSIDE issues: Mac Users goto your directory and type ./run.sh to startup LightSIDE.

Type ./run.sh and then press "enter"

- HW2 out today. Deadline will be adjusted.

Overview

- Philosophical questions

- Derivatives: What are they good for?

- Linear regression

- *Multiple* linear regression

- Logistic regression

Philosophical Questions

- What would you do if …

- What does this have to do with linear classifiers?

Functions

$$x \longrightarrow \boxed{f(x)} \longrightarrow y$$

Derivatives

$y = 2x$

Derivatives

$$\frac{d}{x}\,(2x) = 2$$

Derivatives

$y = x^2$

Derivatives

$$\frac{d}{x}\left(x^2\right) = 2x$$
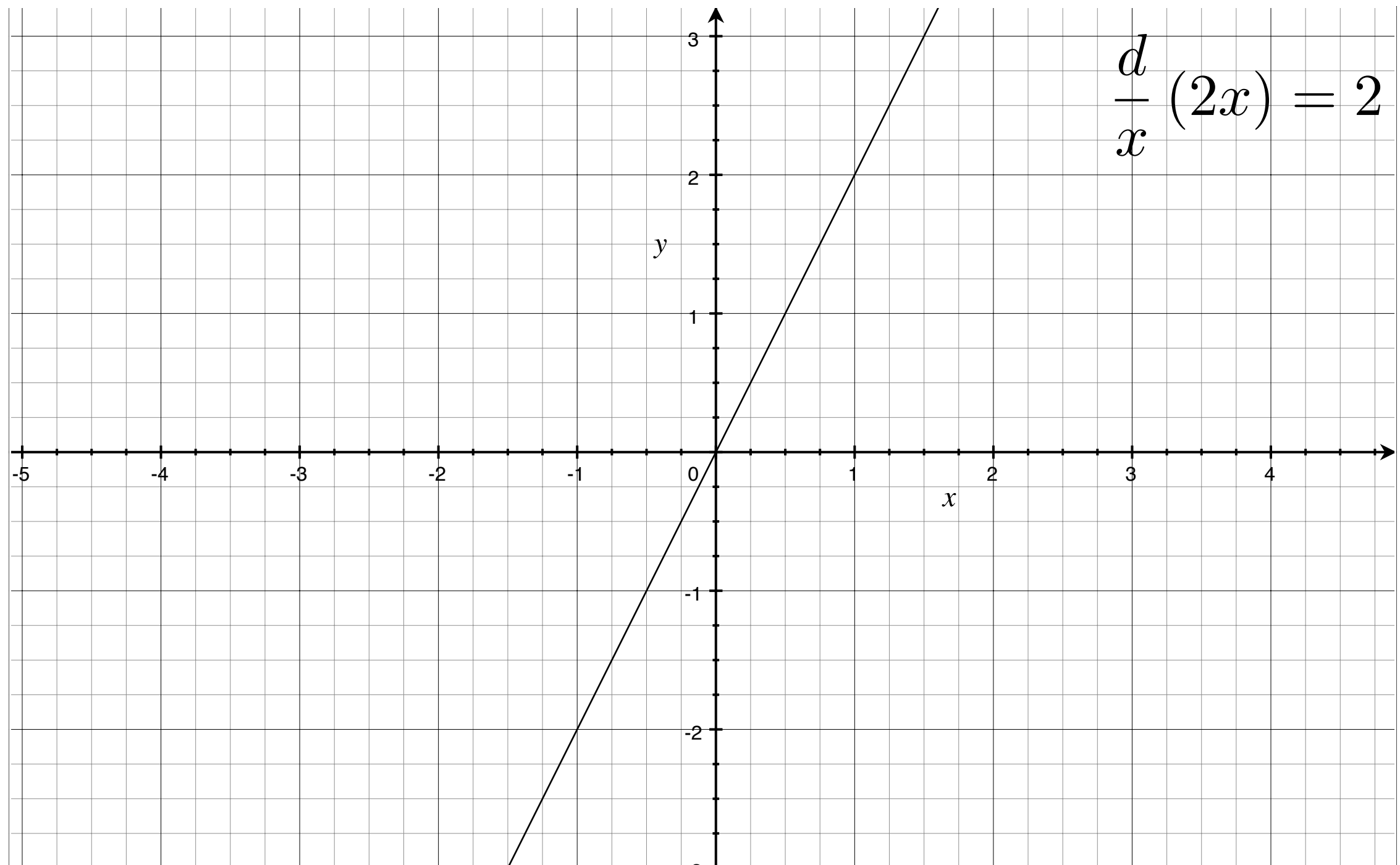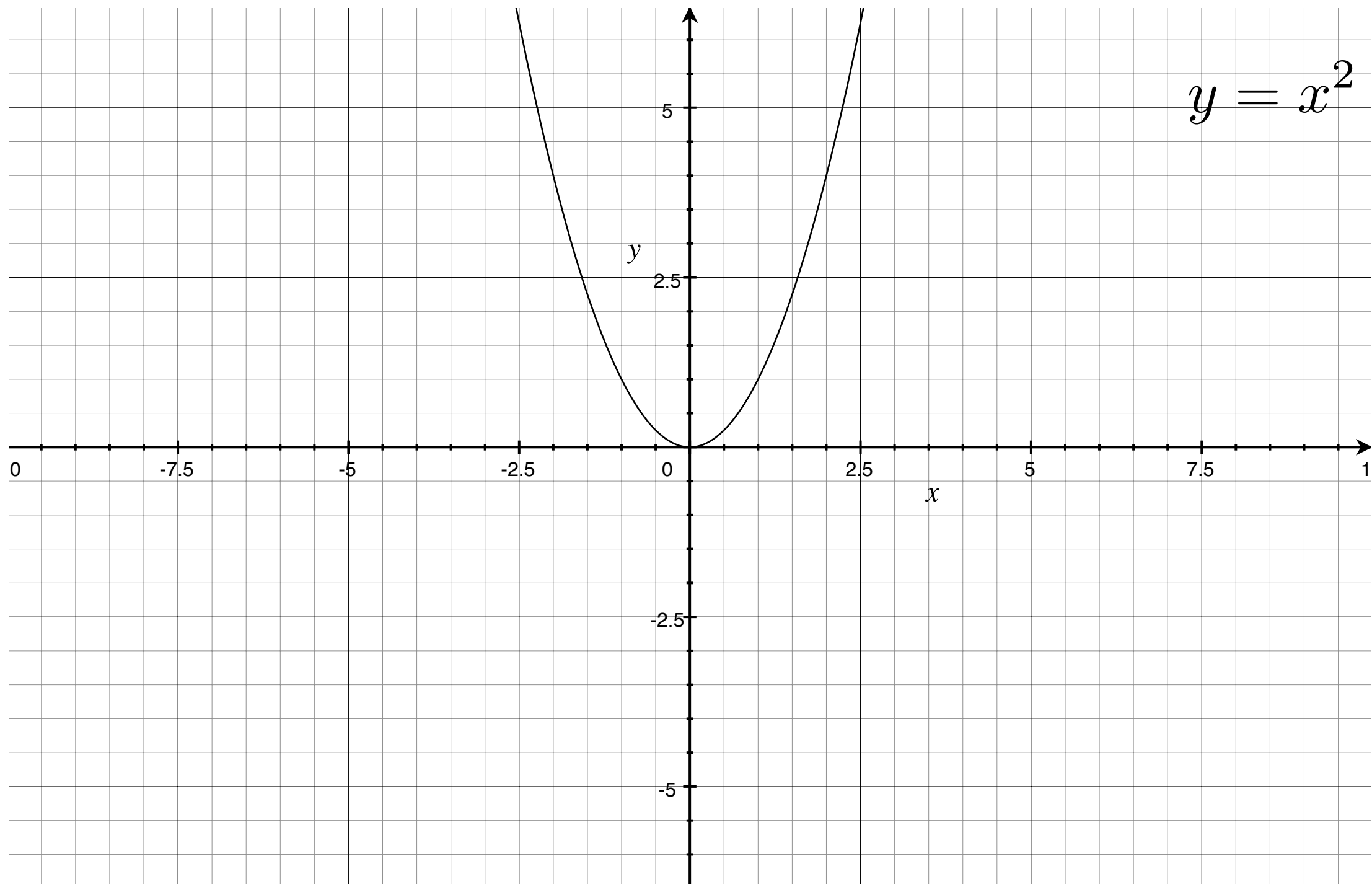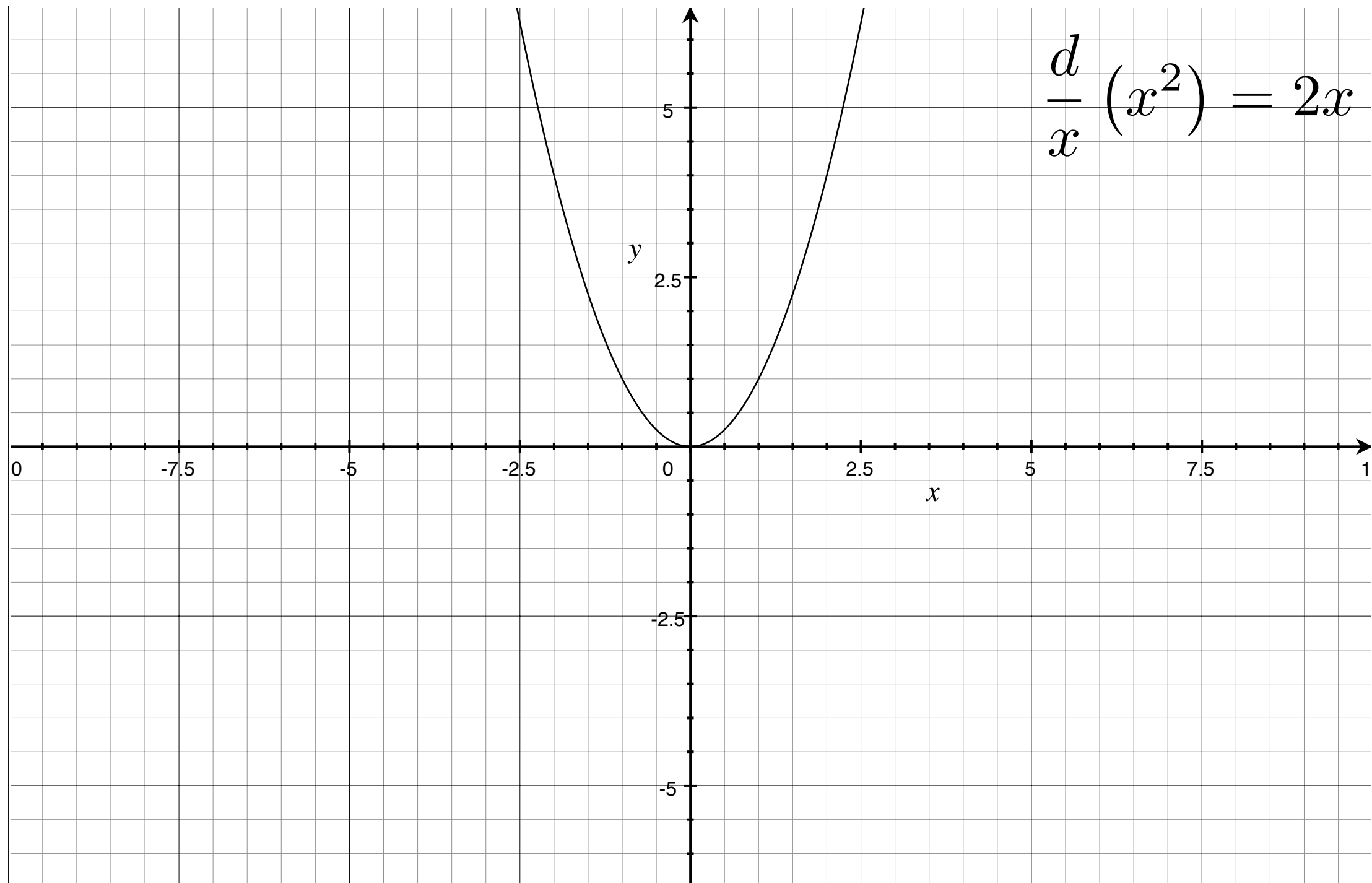
Derivatives: What are they good for?

- The derivative of f(x) outputs the slope of f(x) for a particular value of x

- A point of which the slope is zero is a point at which f(x) is at its highest or lowest value.

- What does this have to do with machine learning?

Derivatives

$$\frac{d}{x}\left(x^2\right) = 2x$$

$$2x = 0$$

$$x = 0$$
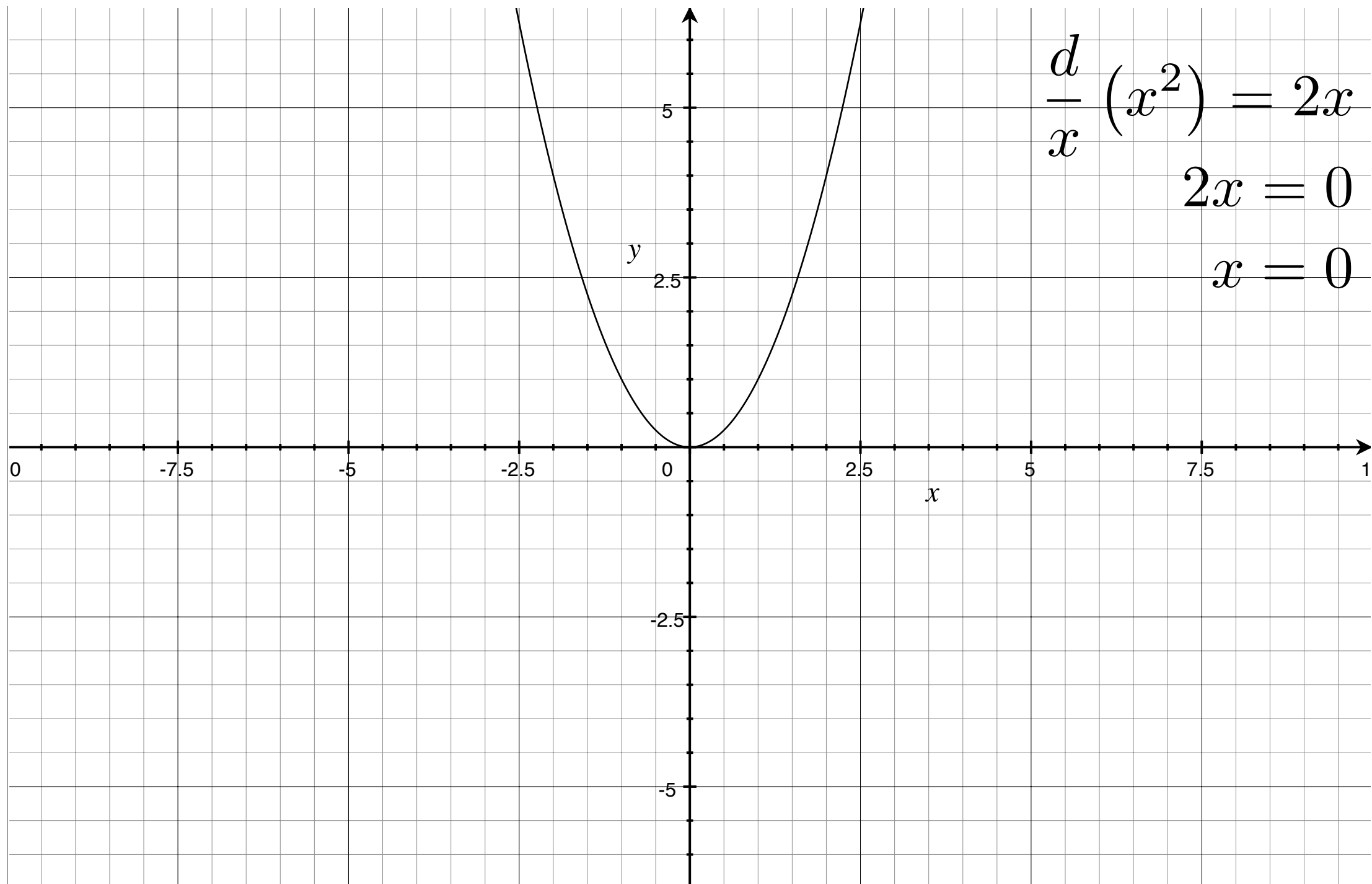
Computation Graphs

$$y = 3(a + bc)$$

Computation Graphs

$$y = 3(a + bc)$$

a

b

c

$u = (a + v)$

$v = bc$

$y = 3u$
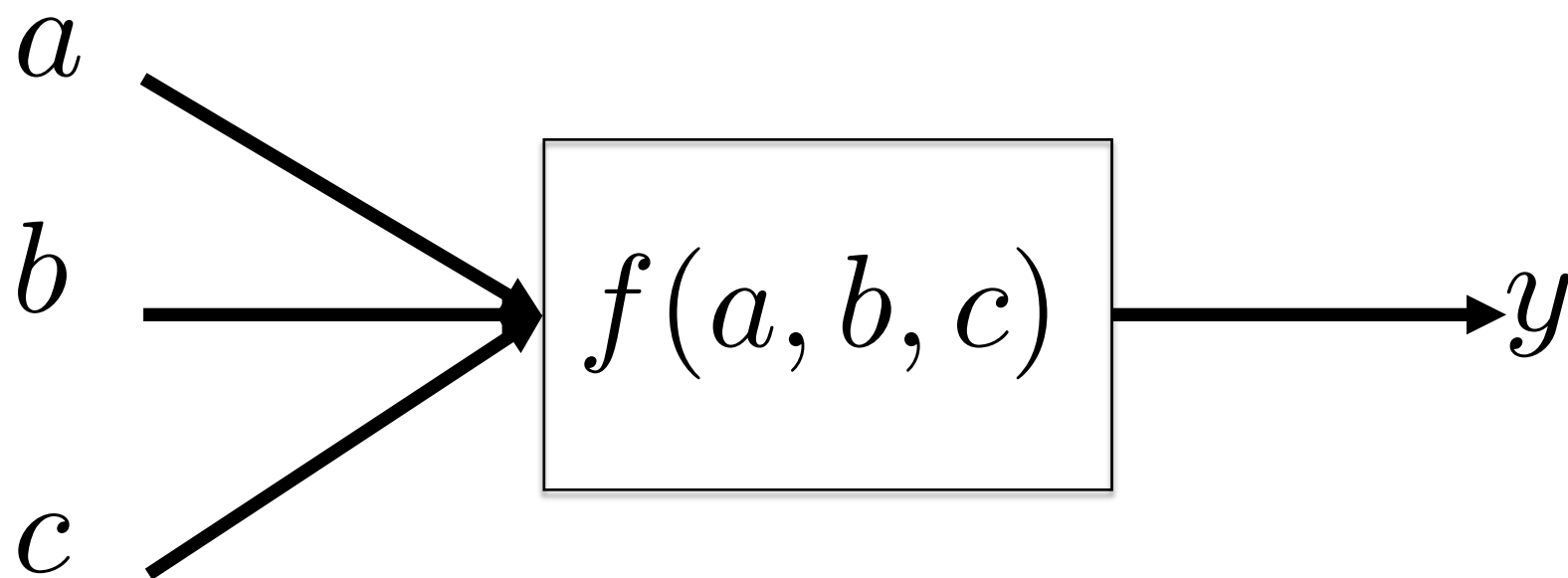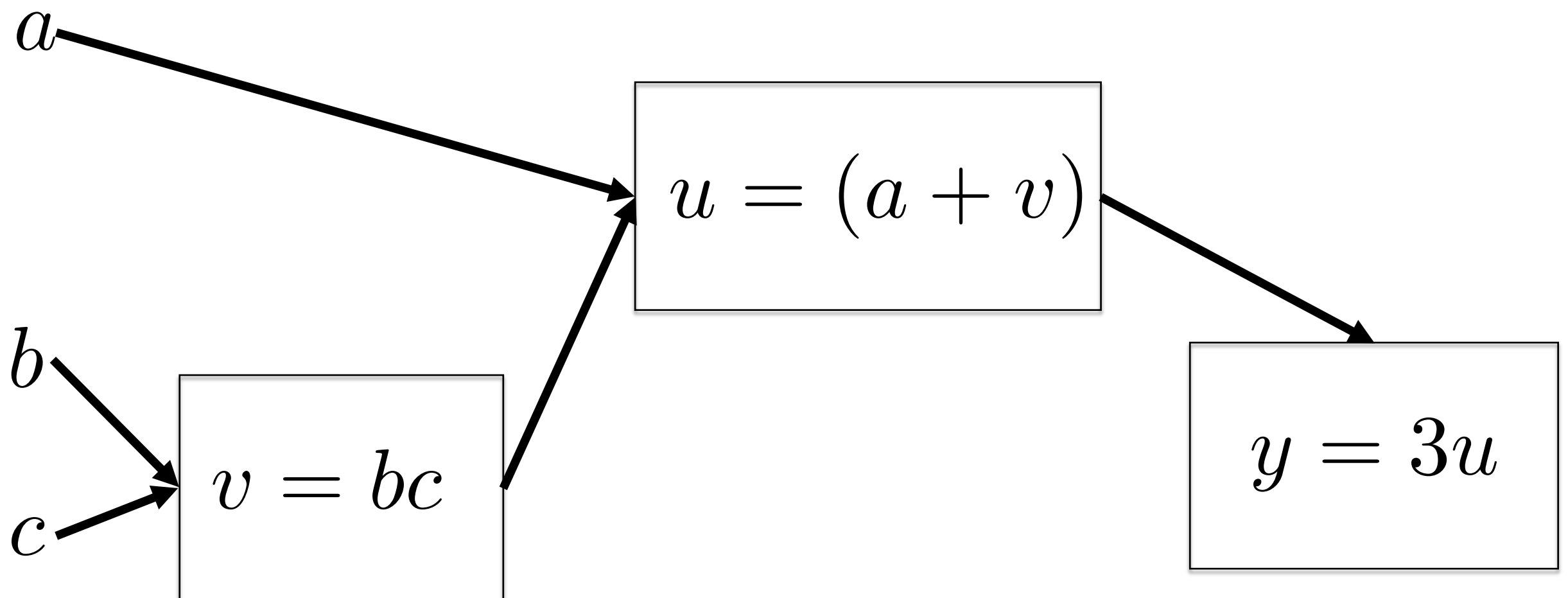
Derivatives: Chain Rule

$$y = 3(a + bc)$$

$$\frac{dy}{dc} = \frac{dv}{dc} \times \frac{du}{dv} \times \frac{dy}{du}$$

a

b

c

$u = (a + v)$

$v = bc$

$y = 3u$
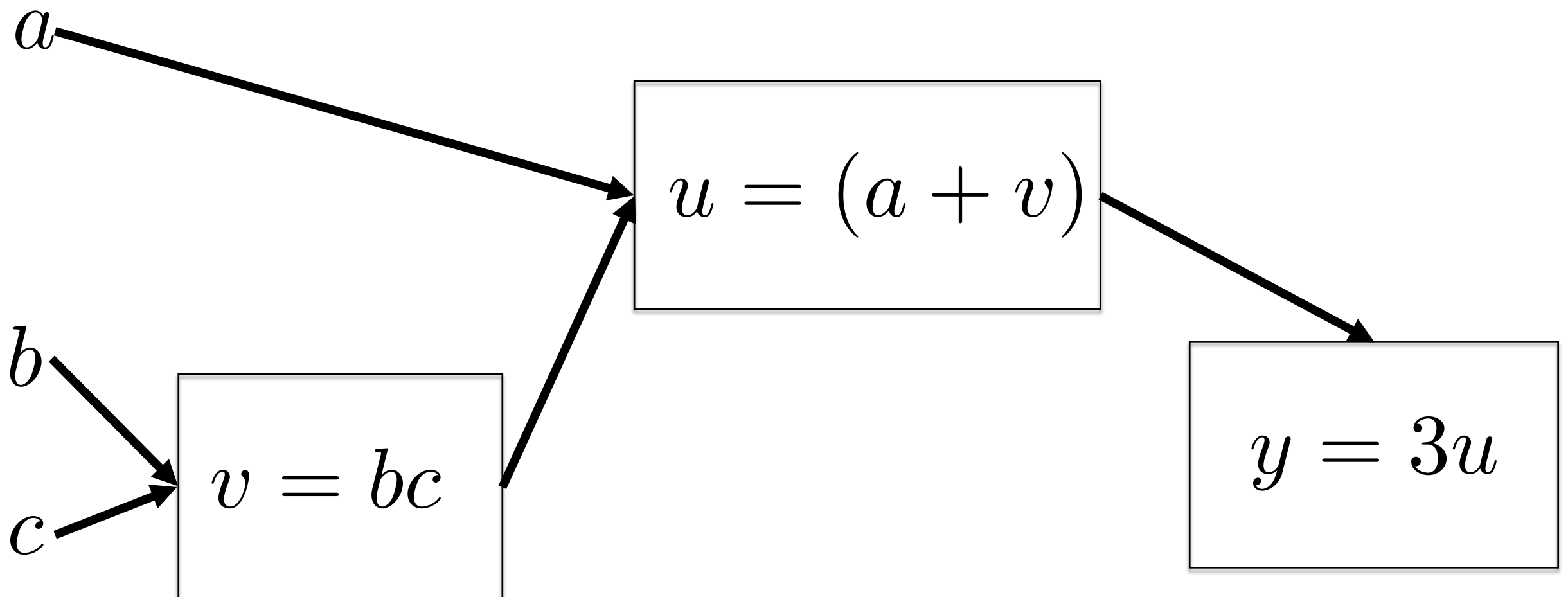
Derivatives: Chain Rule

$$y = 3(a + bc)$$

$$\frac{dy}{dc} = b \times 1 \times 3 = 3b$$
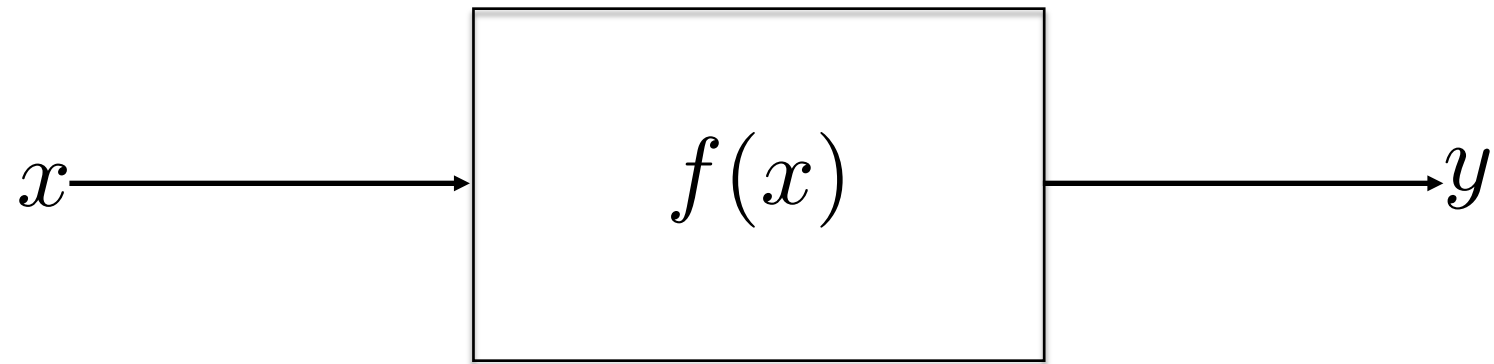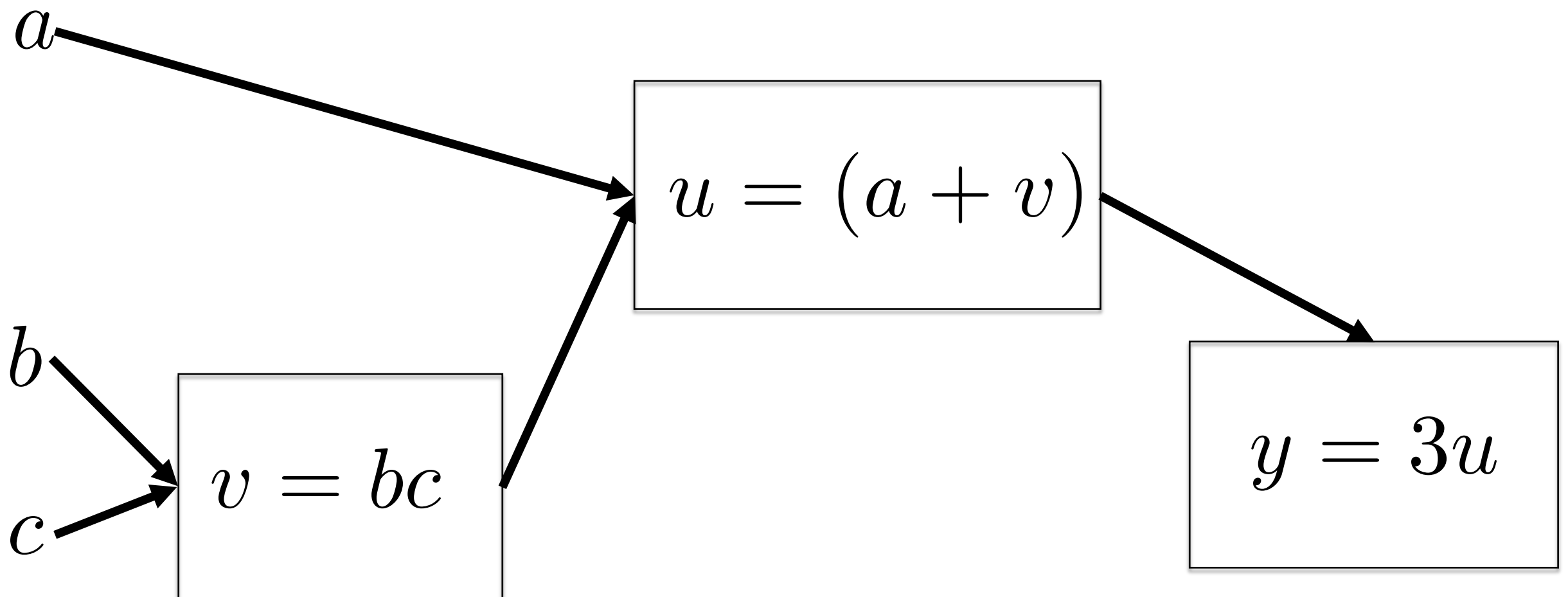
a

b

c

$$v = bc$$

$$u = (a + v)$$

$$y = 3u$$

Overview

- Philosophical questions

- Derivatives: What are they good for?

- **Linear regression**
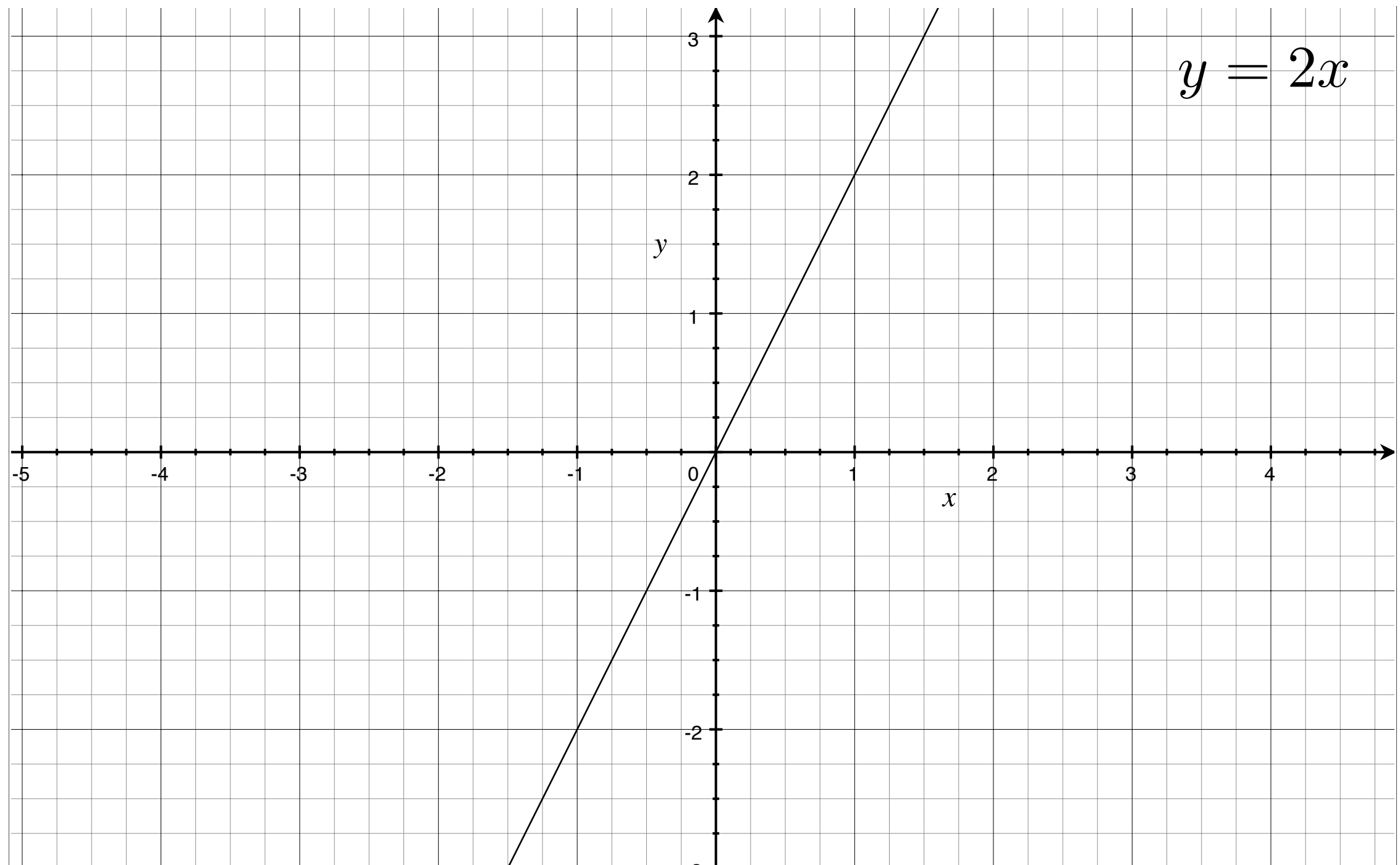
- Multiple linear regression

- Logistic regression

Linear Regression

$$y = wx + b$$

Linear Regression

yawns from students (y-axis)

Temperature in Rm 001 (x-axis)
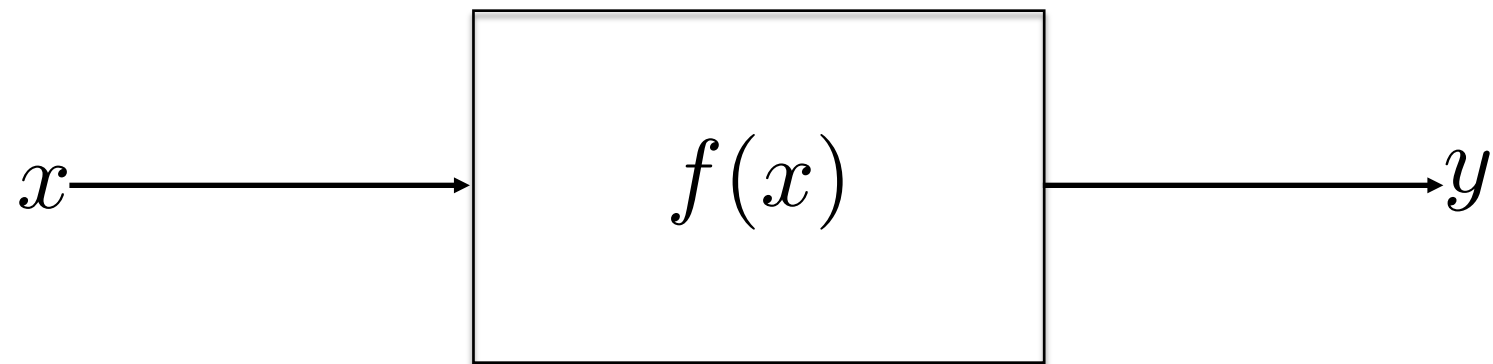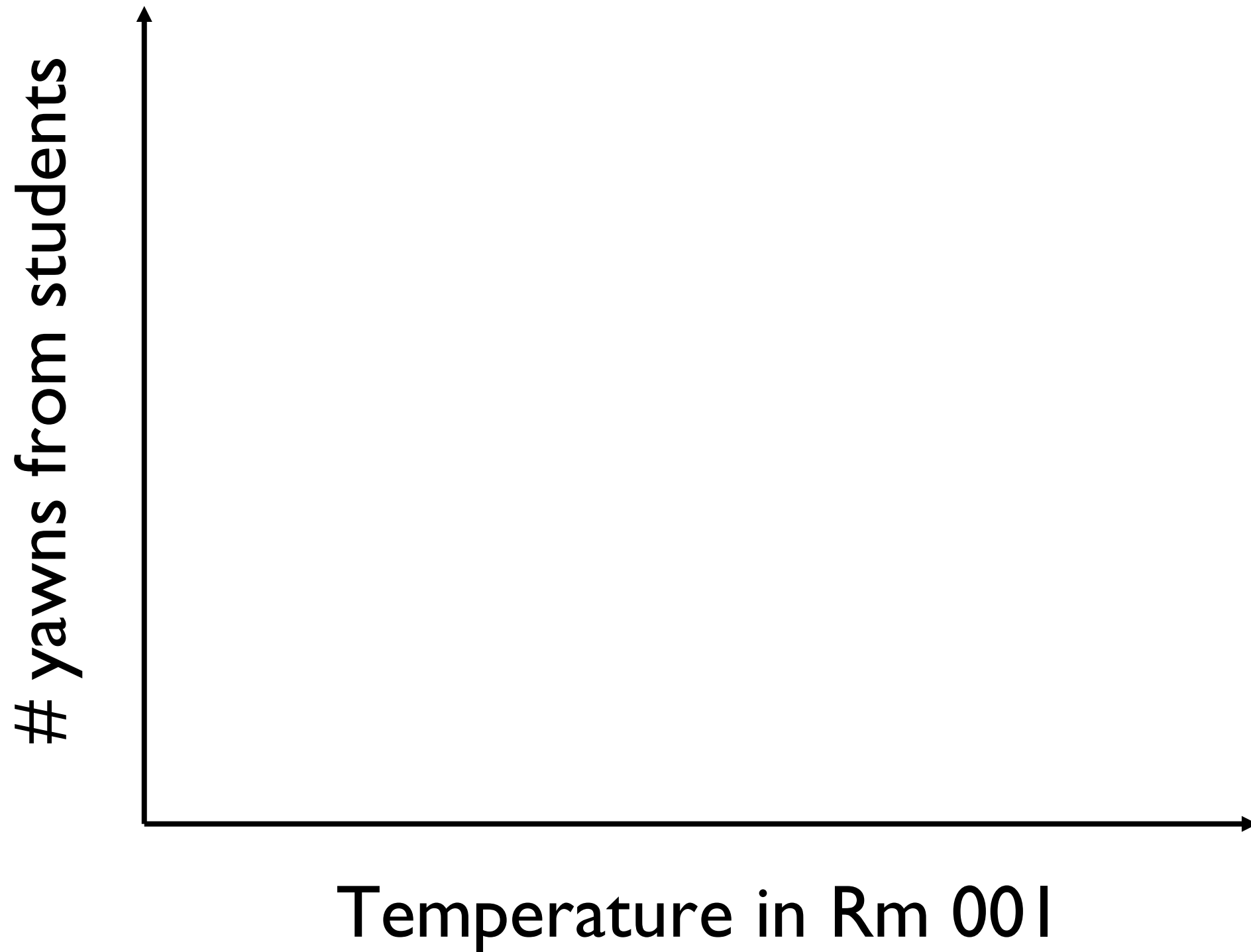
Linear Regression

Linear Regression

$$y = wx + b$$

yawns from students

Temperature in Rm 001
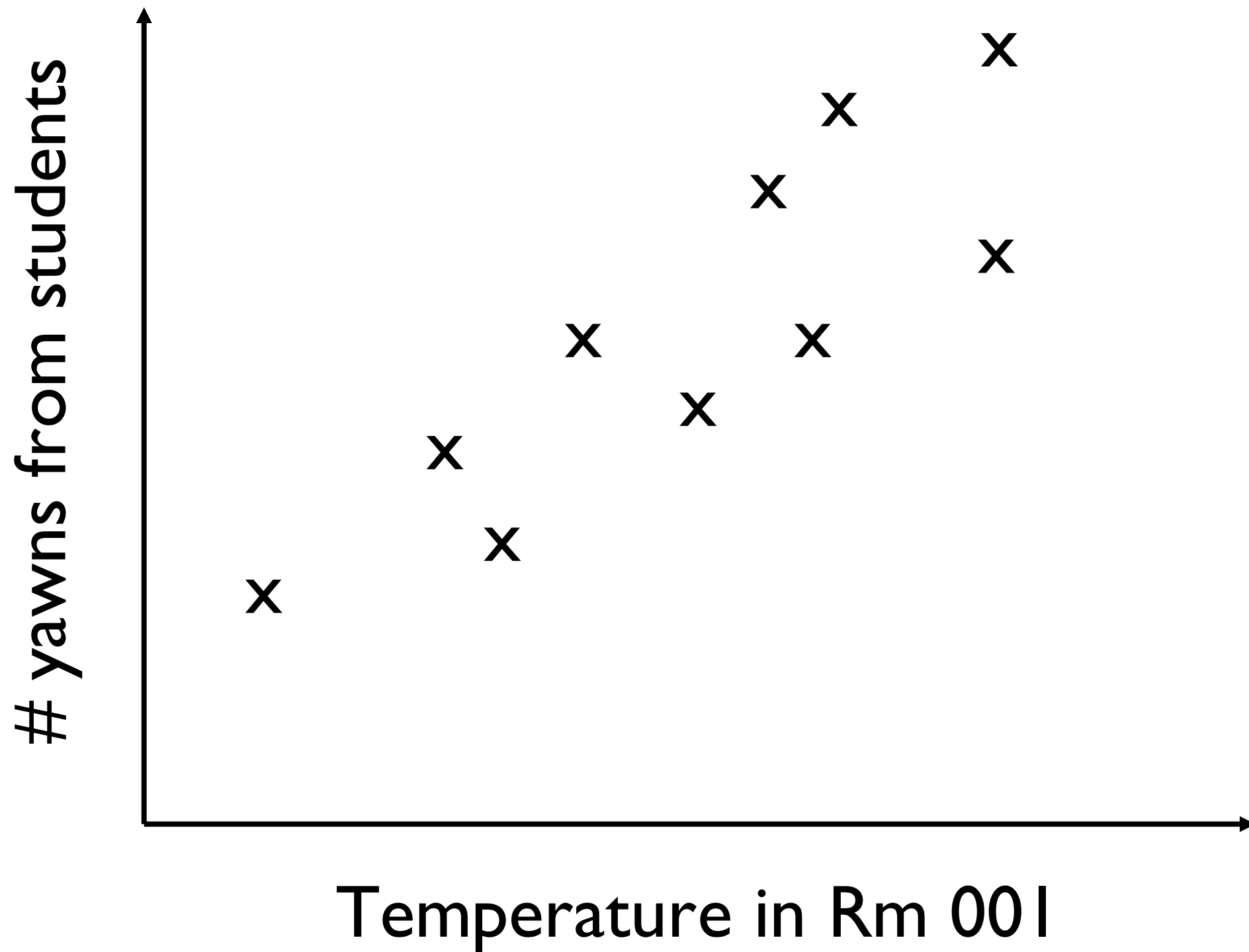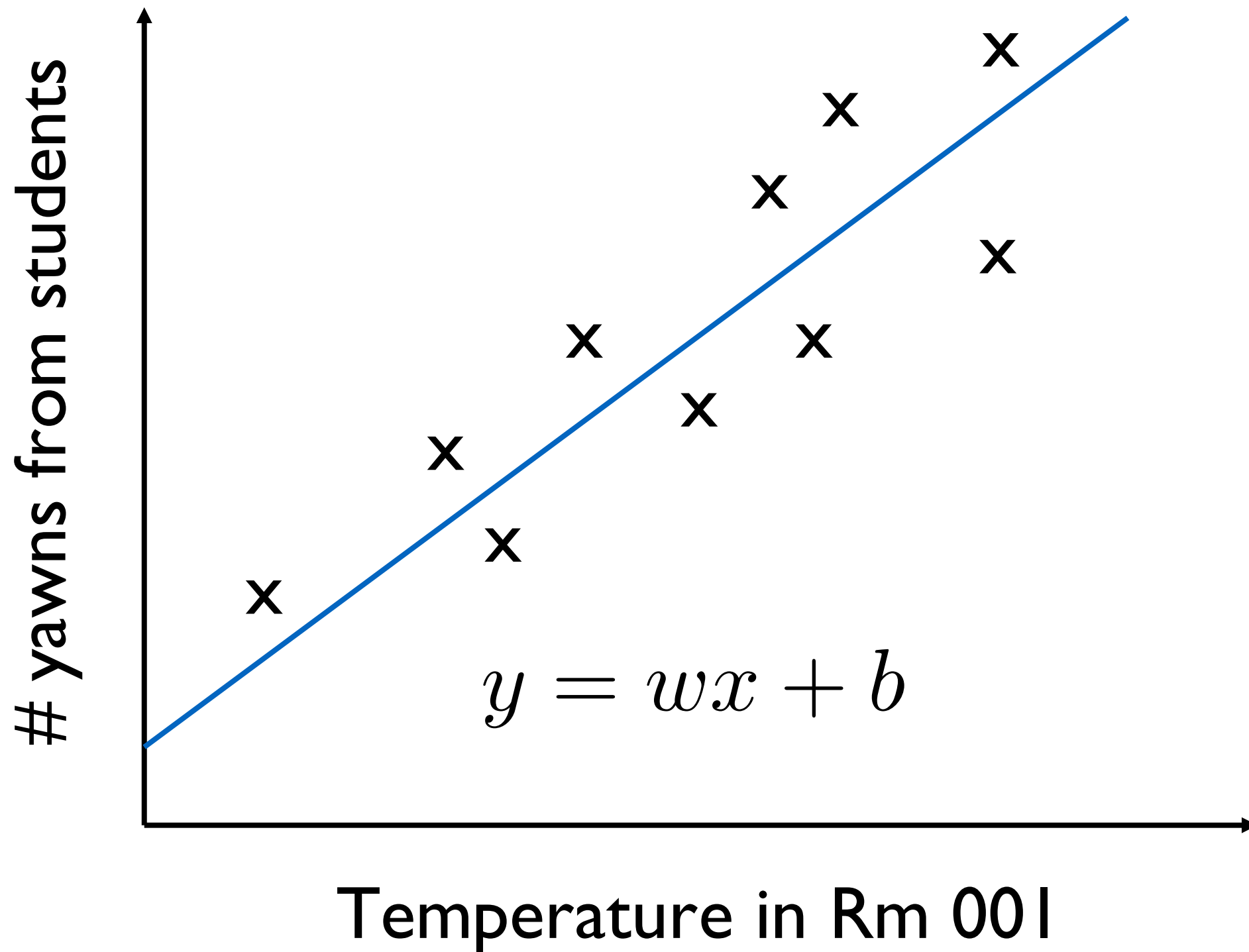
Linear Regression: Training

$$y = wx + b$$

- **Input:** set of *m* training examples (x,y)

- Find the value of *w* and *b* that minimize the error:

$$\sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)} \right)^2$$

Linear Regression: Training

$$y = wx + b$$

- Find the value of w and b that minimize the error:

$$\sum_{i=1}^{m} \left(y^{(i)} - \hat{y}^{(i)} \right)^2$$

$$\sum_{i=1}^{m} \left(y^{(i)} - wx^{(i)} - b \right)^2$$

Linear Regression: Training

$$y = wx + b$$

- Find the value of *w* and *b* that minimize the error:

$$\sum_{i=1}^{m} \left(y^{(i)} - wx^{(i)} - b \right)^2$$

- Take the derivative with respect to *w*, set it equal to 0, and solve for *w*.

- Take the derivative with respect to *b*, set it equal to 0, and solve for *b*.

Linear Regression: Training

- Find the value of *w* and *b* that minimize the error:

$$w = \frac{\frac{1}{m}\sum_{i=1}^{m}\left(x^{(i)} - \bar{x}\right)\left(y^{(i)} - \bar{y}\right)}{\sum_{i=1}^{m}\left(x^{(i)} - \bar{x}\right)^2}$$

$$b = \bar{y} - w\bar{x}$$

Linear Regression: Training

- Find the value of w and b that minimize the error:

$$w = \frac{\frac{1}{m} \sum_{i=1}^{m} \left(x^{(i)} - \bar{x}\right)\left(y^{(i)} - \bar{y}\right)}{\sum_{i=1}^{m} \left(x^{(i)} - \bar{x}\right)^2}$$

Always positive!

It depends!

$$b = \bar{y} - w\bar{x}$$

http://seismo.berkeley.edu/~kirchner/eps_120/Toolkits/Toolkit_10.pdf