

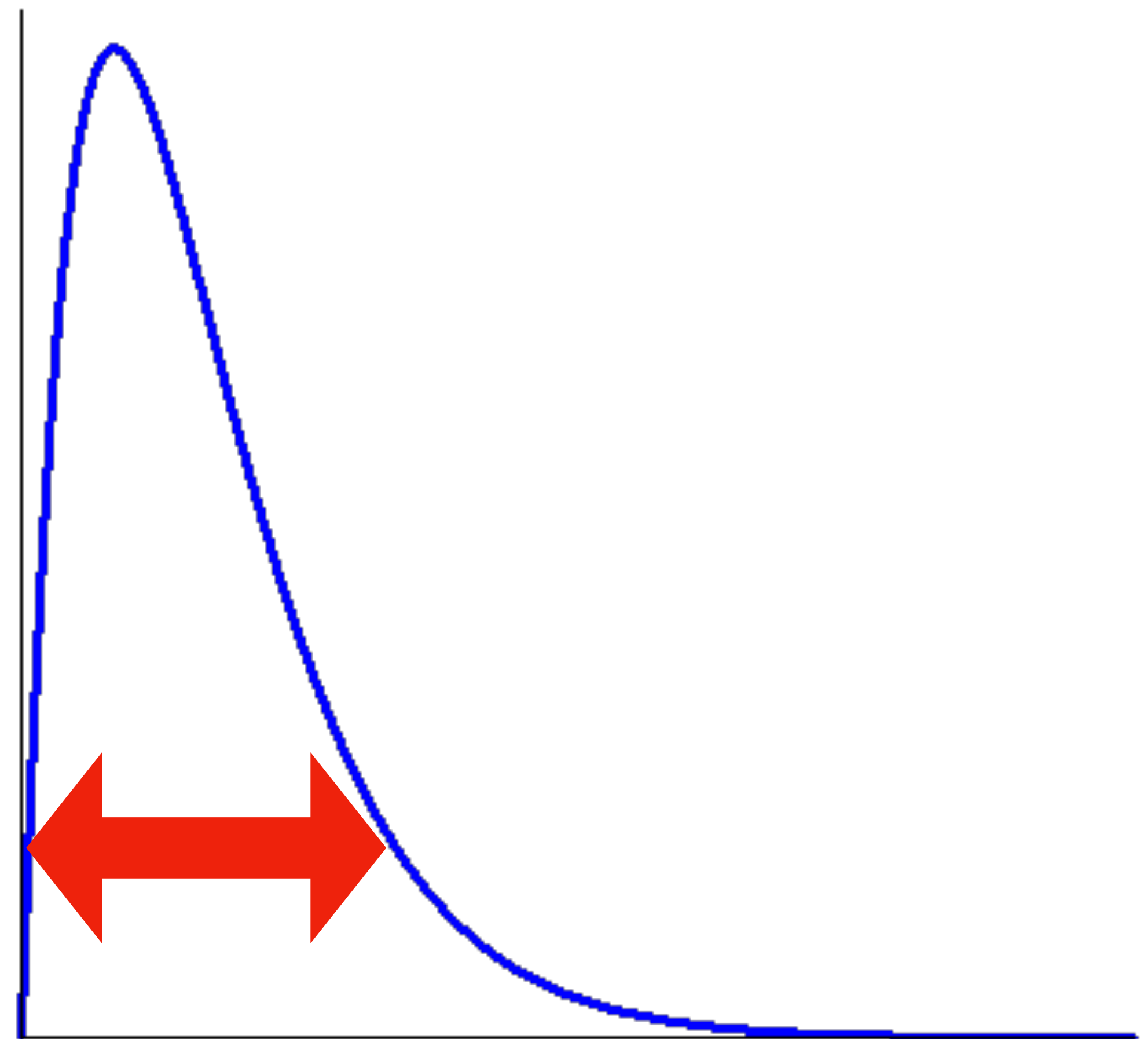
Variance:

- Measure of “spread” of a distribution
- Denoted by $\text{Var}[X]$ or σ^2
- Defined as

$$\text{Var}[X] = E[(X - \mu)^2]$$

where $\mu = E[X]$

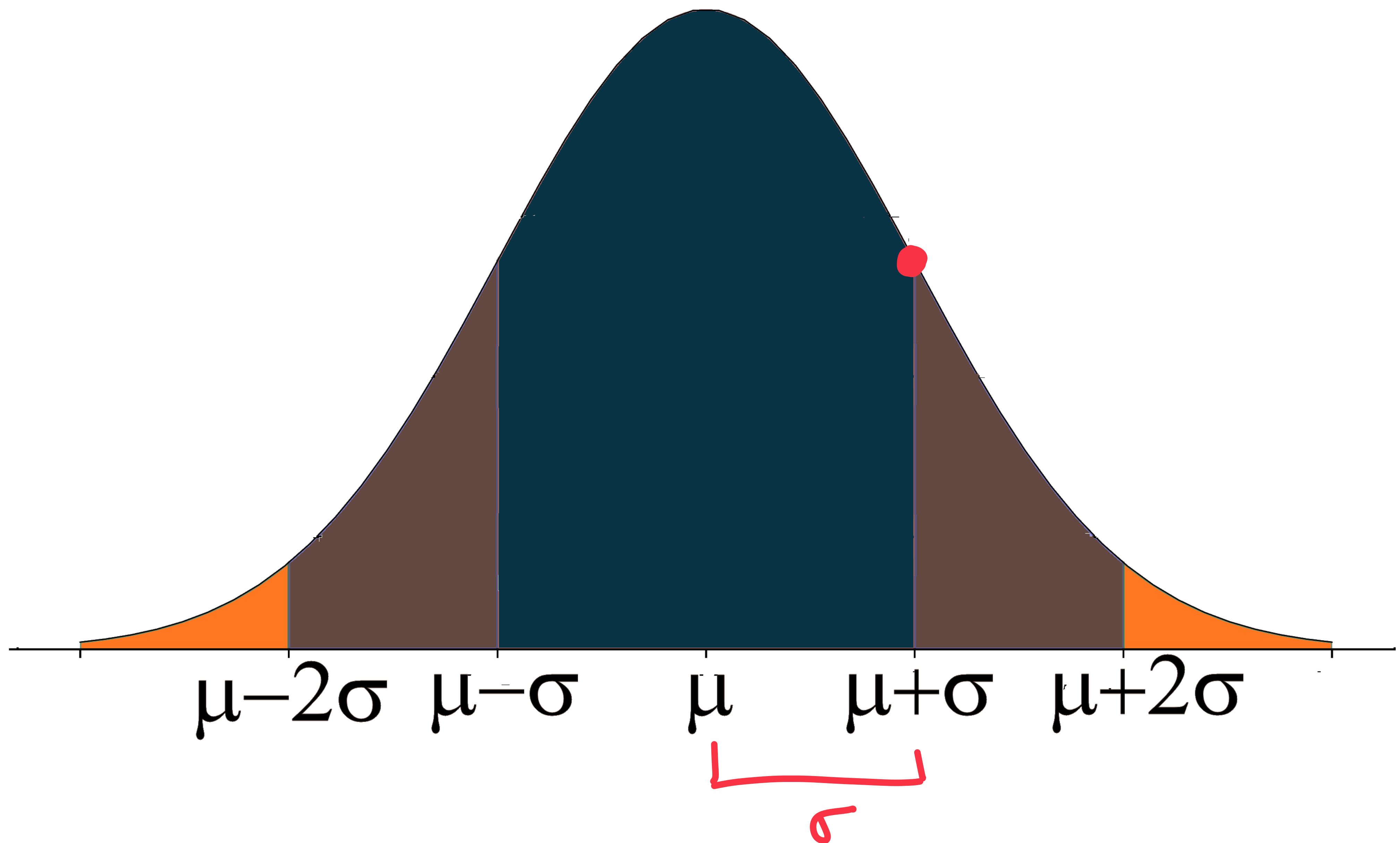
- Why not $E[X - \mu]$?
- ~~Why not~~ $E[|X - \mu|]$?



Standard Deviation:

$$\sigma = \sqrt{\sigma^2}$$

$$\text{Ex: } X \sim N(\mu, \sigma^2)$$



Another Way to Compute Variance:

$$\begin{aligned}\text{Var}[X] &= E[(X - \mu)^2] \\&= E[X^2 - 2\mu X + \mu^2] \\&= E[X^2] - 2\mu \underbrace{E[X]}_{\mu} + \mu^2 \\&= E[X^2] - 2\mu^2 + \mu^2 \\&= E[X^2] - \mu^2 \\&= E[X^2] - (E[X])^2\end{aligned}$$

Properties of Variance:

$$\text{Var}[aX] = ?$$

Let $Y = aX$.

Then

$$\mu_Y = E[Y] = E[aX] = aE[X] = a\mu_X.$$

$$\begin{aligned}\Rightarrow \text{Var}[aX] &= \text{Var}[Y] = E[(Y - \mu_Y)^2] \\ &= a^2 E[(X - \mu_X)^2] \\ &= a^2 \text{Var}[X]\end{aligned}$$

Properties of Variance:

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]?$$

Not necessarily!

- We will see that this is true if X and Y are independent.
- Need concept of “covariance”.

Covariance:

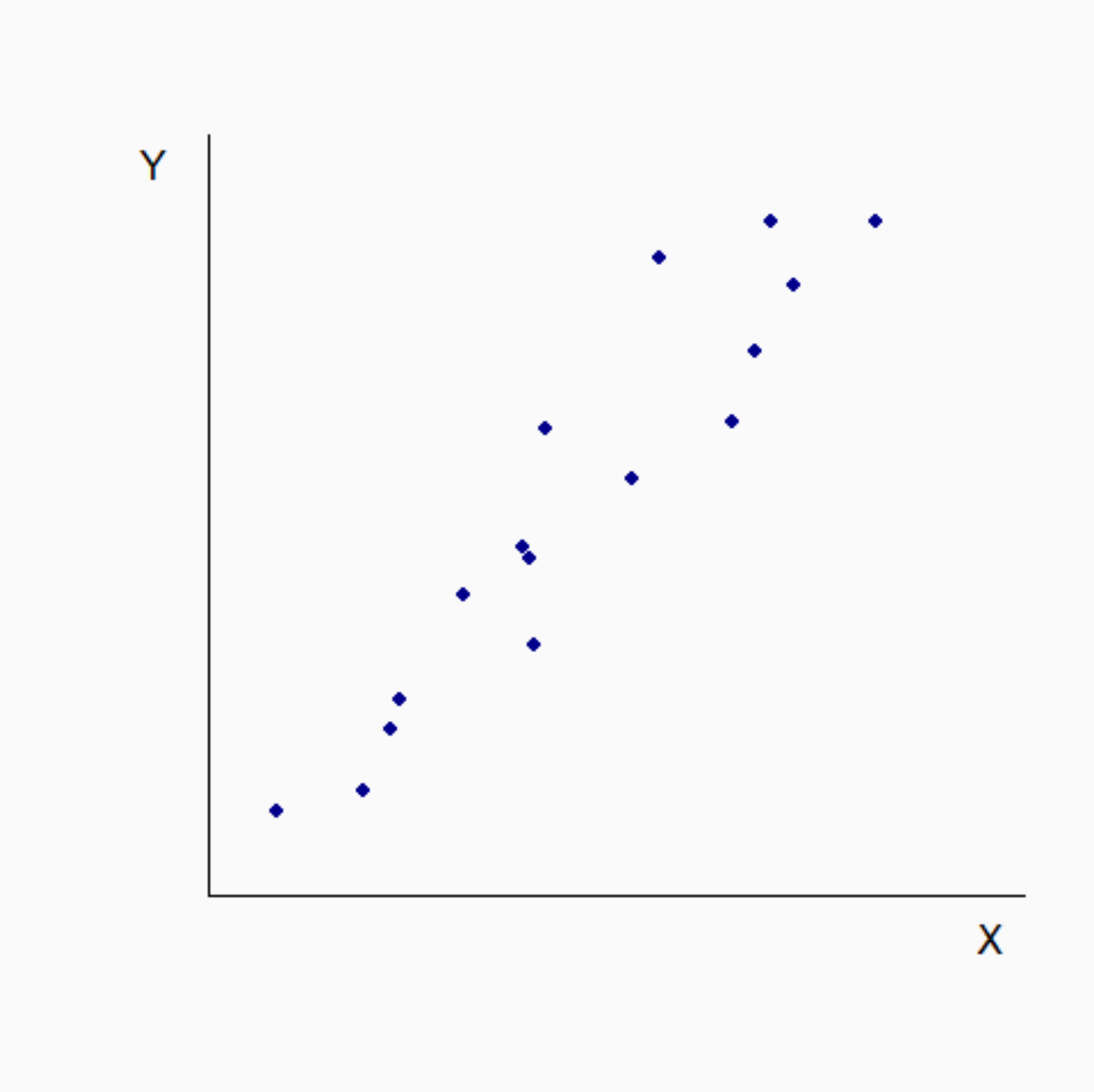
Let X and Y be random variables.

Define/denote the covariance as

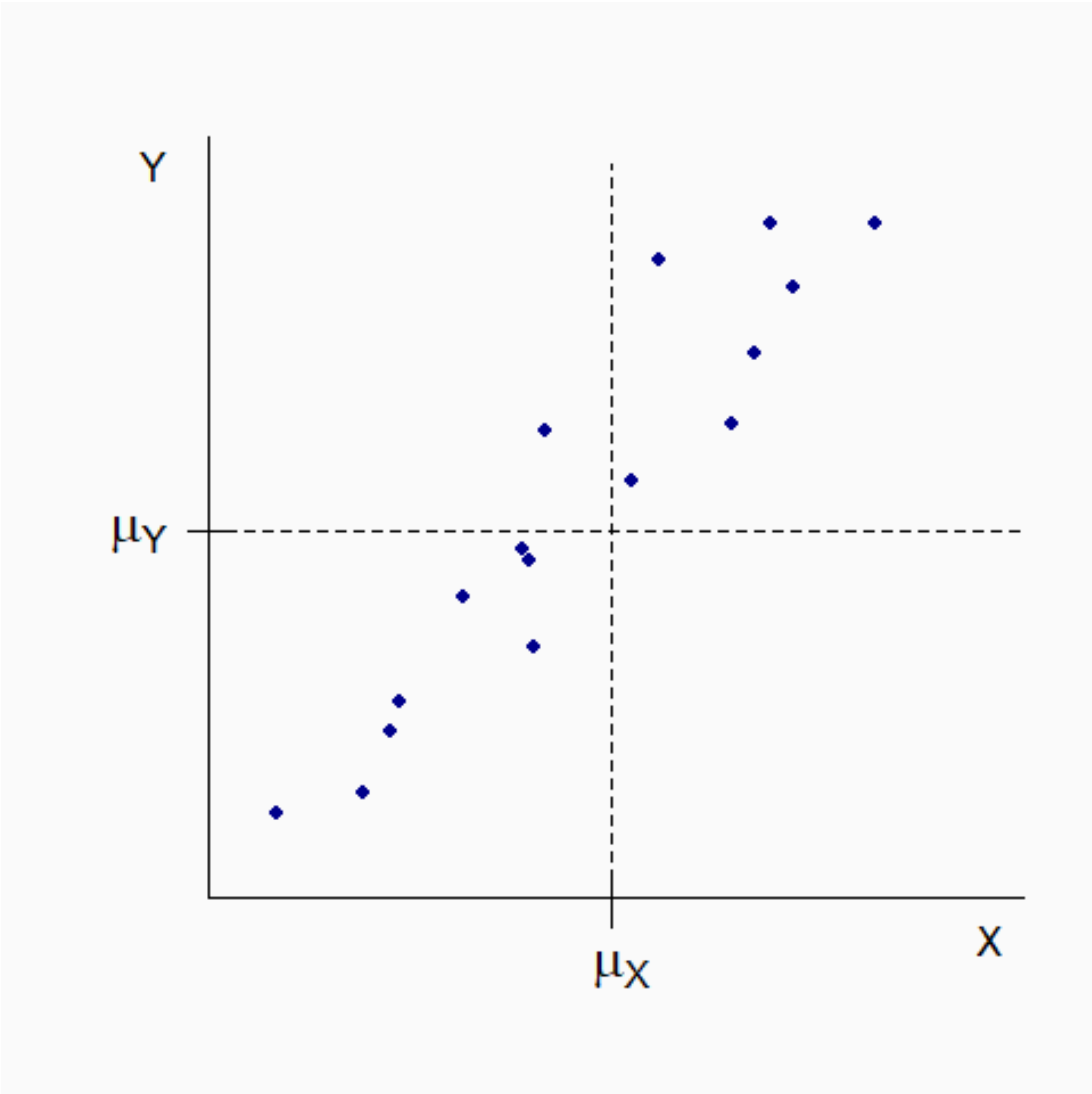
$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

- Notation: $\sigma_{X,Y}$
- Note that $\text{Cov}(X, X) = \text{Var}[X]$

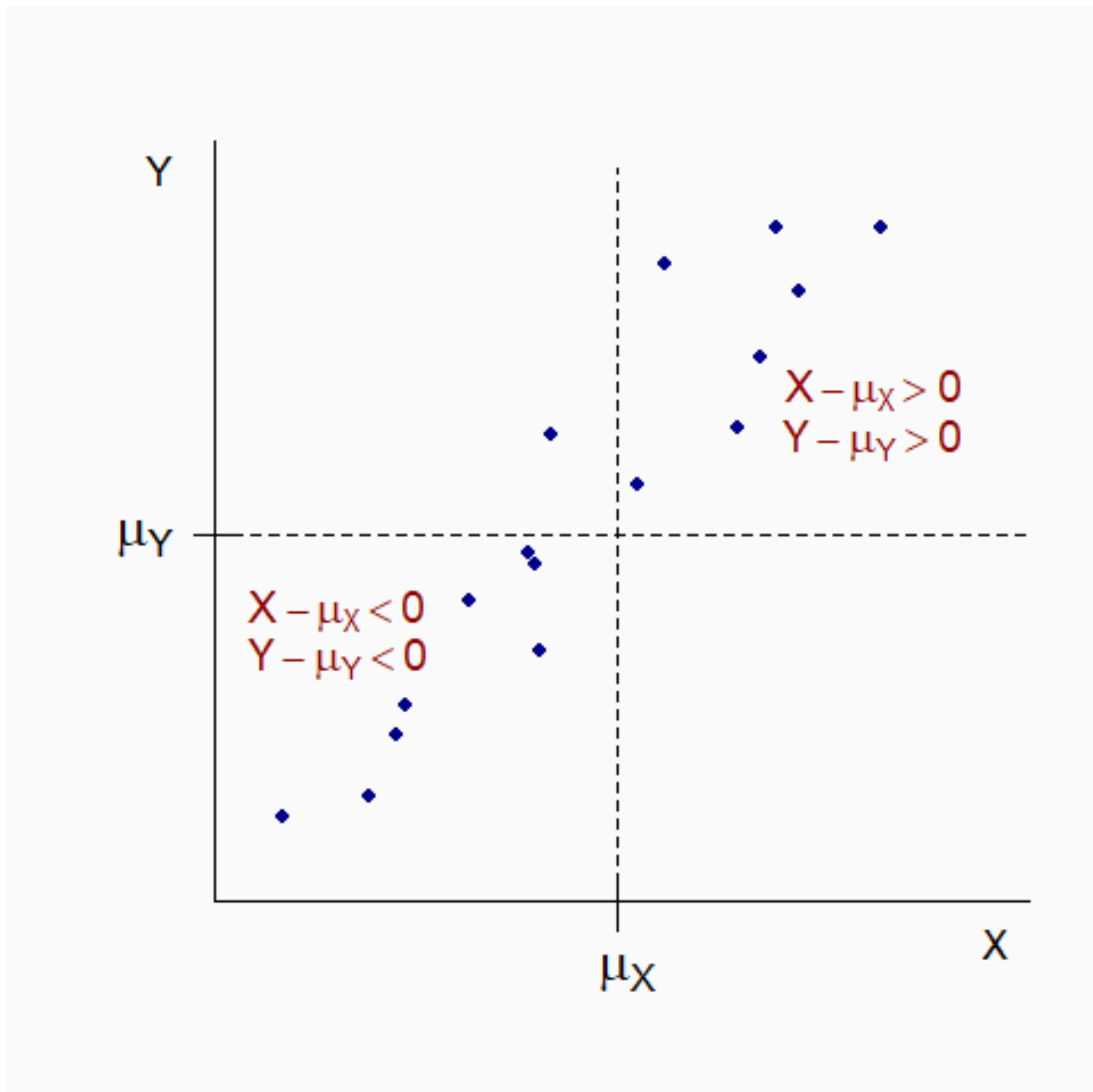
Example:



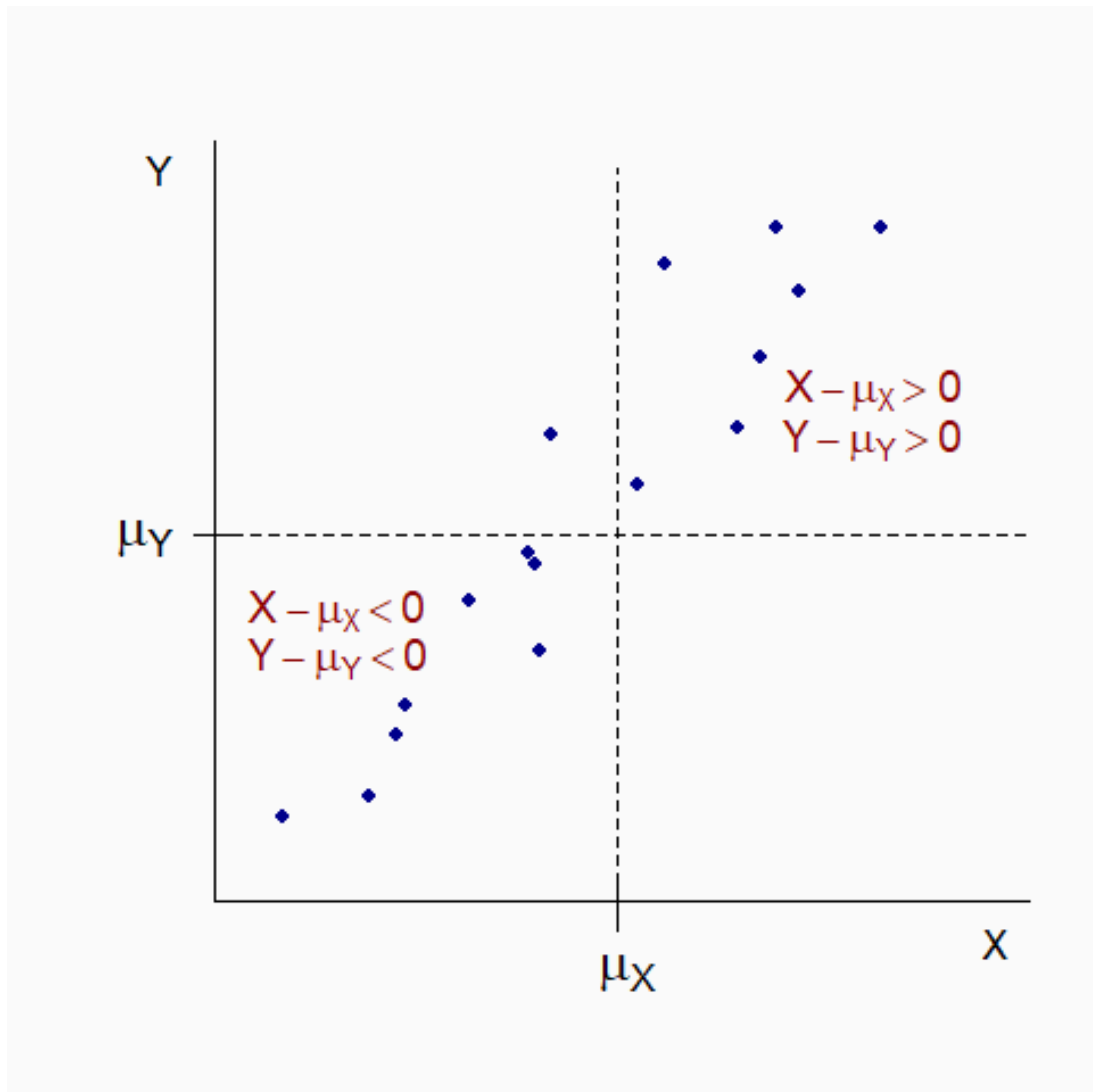
Example:



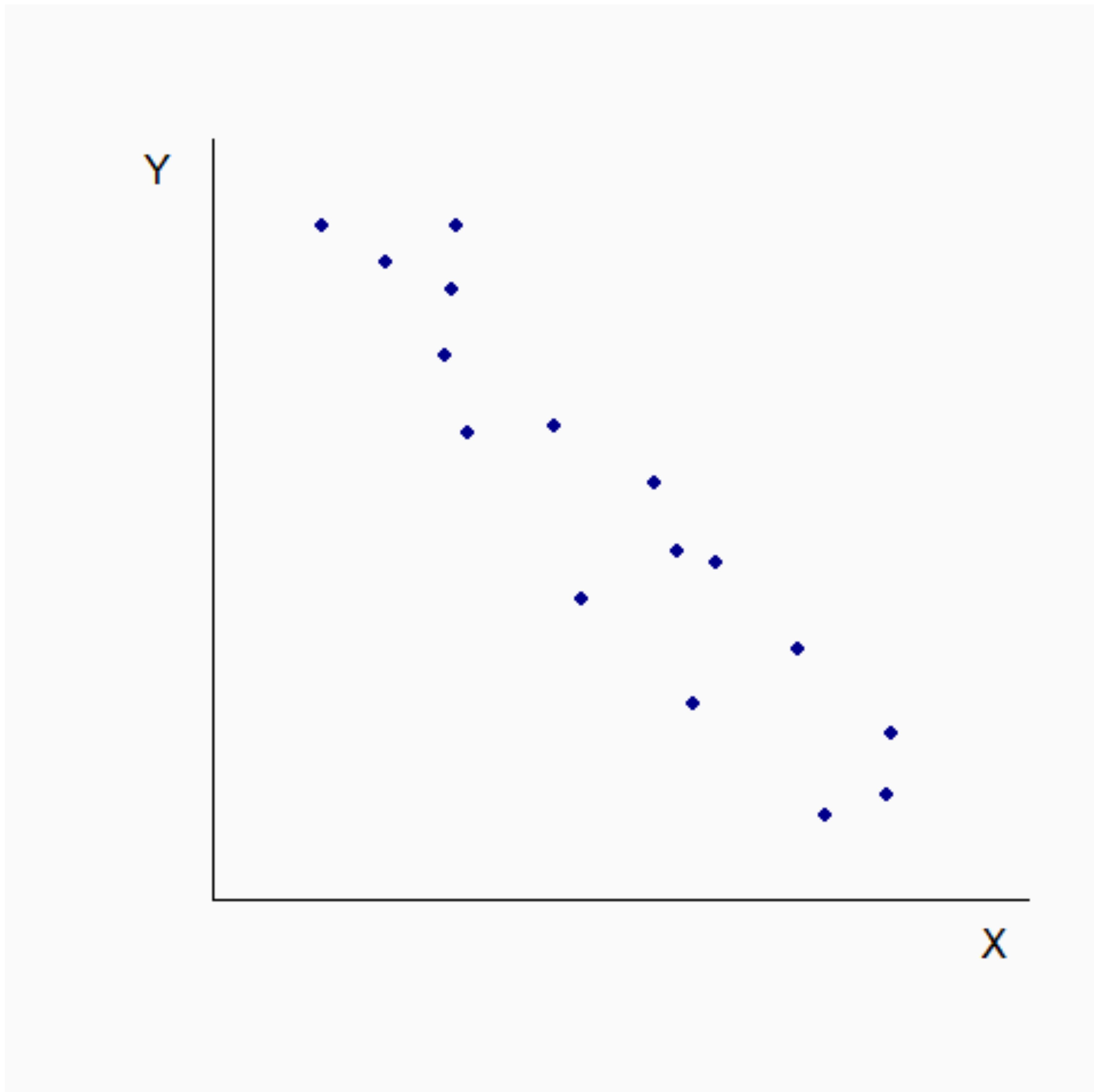
Example:



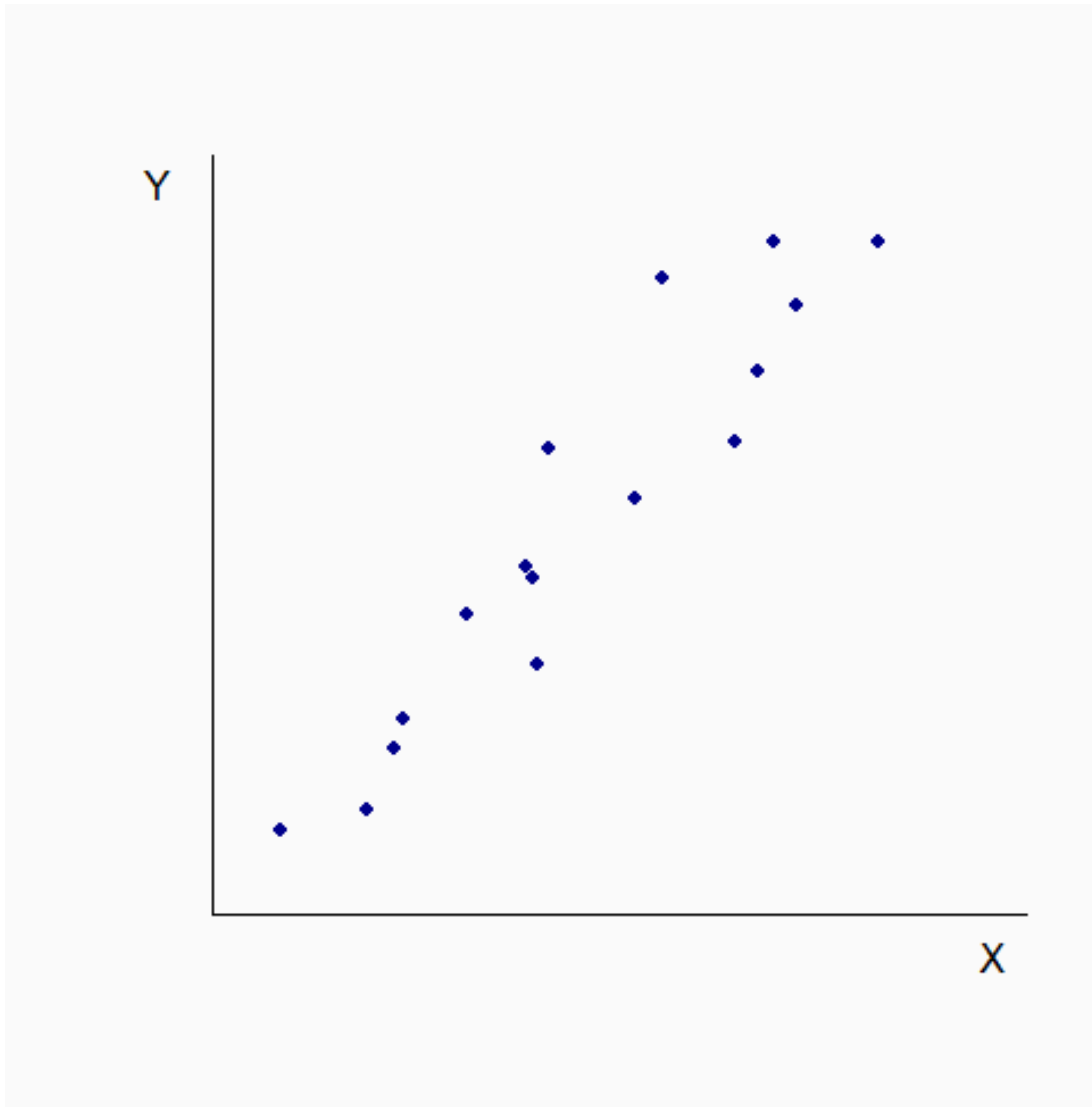
Example: Positive Covariance



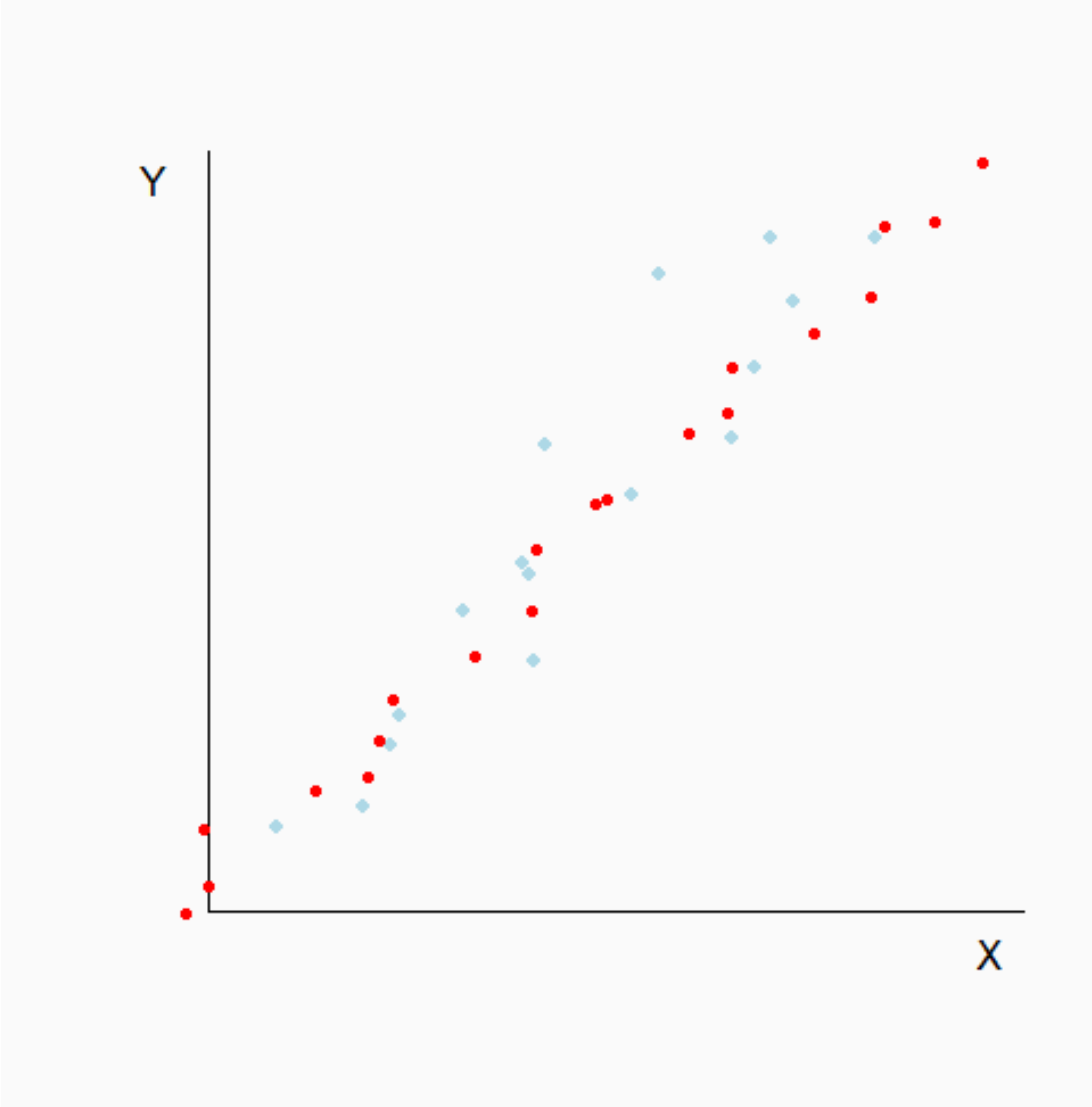
Example: Negative Covariance



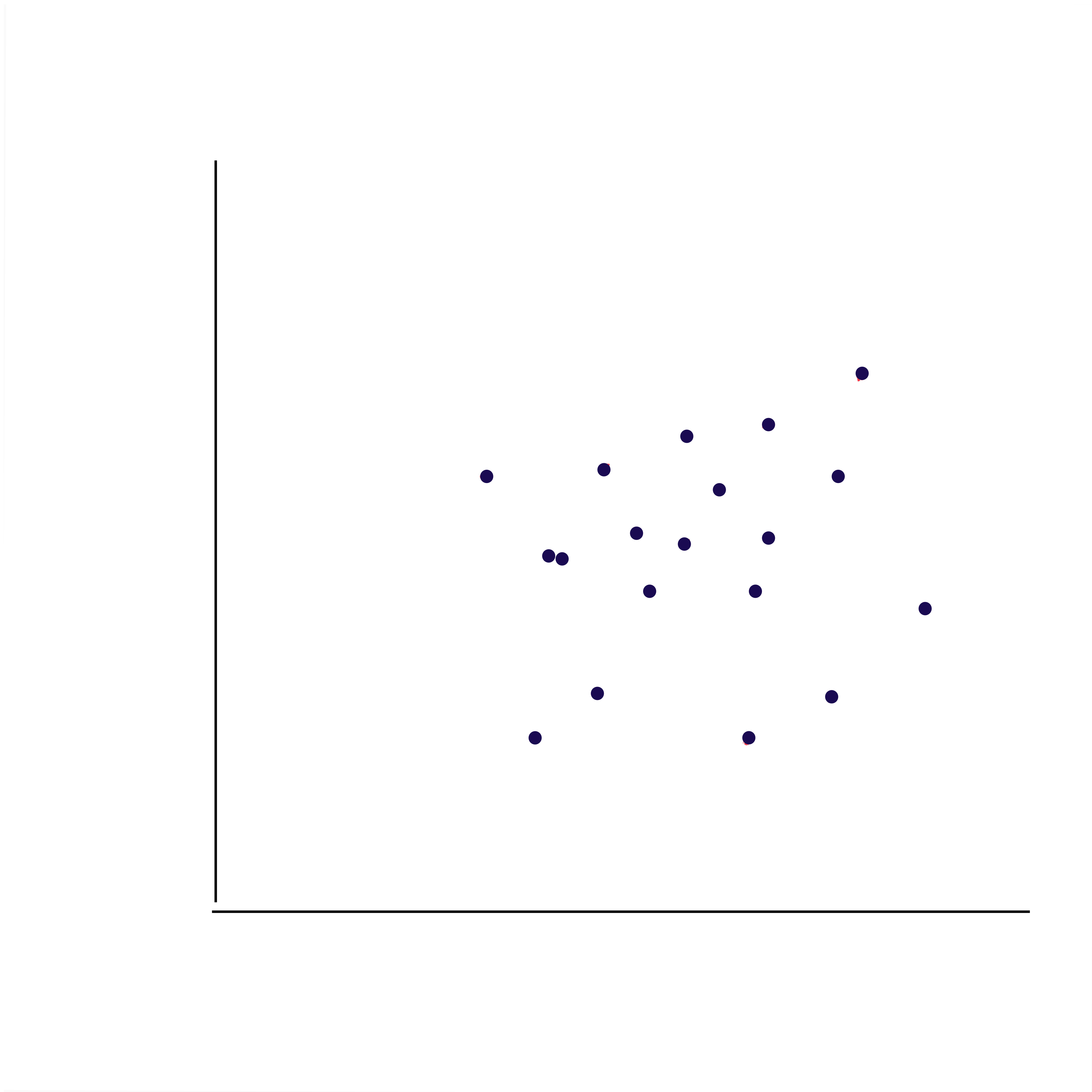
Example: Positive Covariance



Example: Larger Positive Covariance



Example: Covariance Close to Zero



- There is no upper/lower bound on covariance.
- Also, it is in weird units.
- Both problems can be addressed using **correlation**.

$$\text{Corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(x)\text{Var}(Y)}}$$

- Notation: $\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$
- Property: $-1 \leq \rho_{X,Y} \leq 1$

Covariance Alternate:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$= E[XY] - \mu_Y \underbrace{E[X]}_{\mu_X} - \mu_X \underbrace{E[Y]}_{\mu_Y} + \mu_X \mu_Y$$

$$= E[XY] - \mu_X \mu_Y$$

$$= E[XY] - E[X]E[Y]$$

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

- X and Y independent
 $\Rightarrow \text{Cov}(X, Y) = 0$
- $\text{Cov}(X, Y) = 0$ does not imply X and Y are independent.
- $\text{Cov}(X, Y) = 0$ if and only if
 $\text{Corr}(X, Y) = 0$
- If $\text{Corr}(X, Y) = 0$, we say that X and Y are **uncorrelated**.

independent \Rightarrow uncorrelated

uncorrelated \nRightarrow independent

Previous Example:

x	-1	0	1
P(X = x)	1/4	1/2	1/4

$$E[X] = 0 \quad \Rightarrow \quad E[X]E[Y] = 0$$

Define $Y = X^2$. Then $E[XY] = 0$

$$\text{Var}[X + Y] = ?$$

$$\text{Let } Z = X + Y.$$

Note that

$$\begin{aligned}\mu_Z &= E[Z] = E[X] + E[Y] \\ &= \mu_X + \mu_Y\end{aligned}$$

$$\begin{aligned}\text{Var}[X + Y] &= \text{Var}[Z] = E[(Z - \mu_Z)^2] \\ &= E[((X + Y) - (\mu_X + \mu_Y))^2] \\ &= E[((X - \mu_X) + (Y - \mu_Y))^2]\end{aligned}$$

$$\begin{aligned}\text{Var}[X + Y] &= E[((X - \mu_X) + (Y - \mu_Y))^2] \\&= E[(X - \mu_X)^2] + 2E[(X - \mu_X)(Y - \mu_Y)] \\&\quad + E[(Y - \mu_Y)^2] \\&= \text{Var}[X] + 2\text{Cov}(X, Y) + \text{Var}[Y]\end{aligned}$$

If X and Y are independent

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$