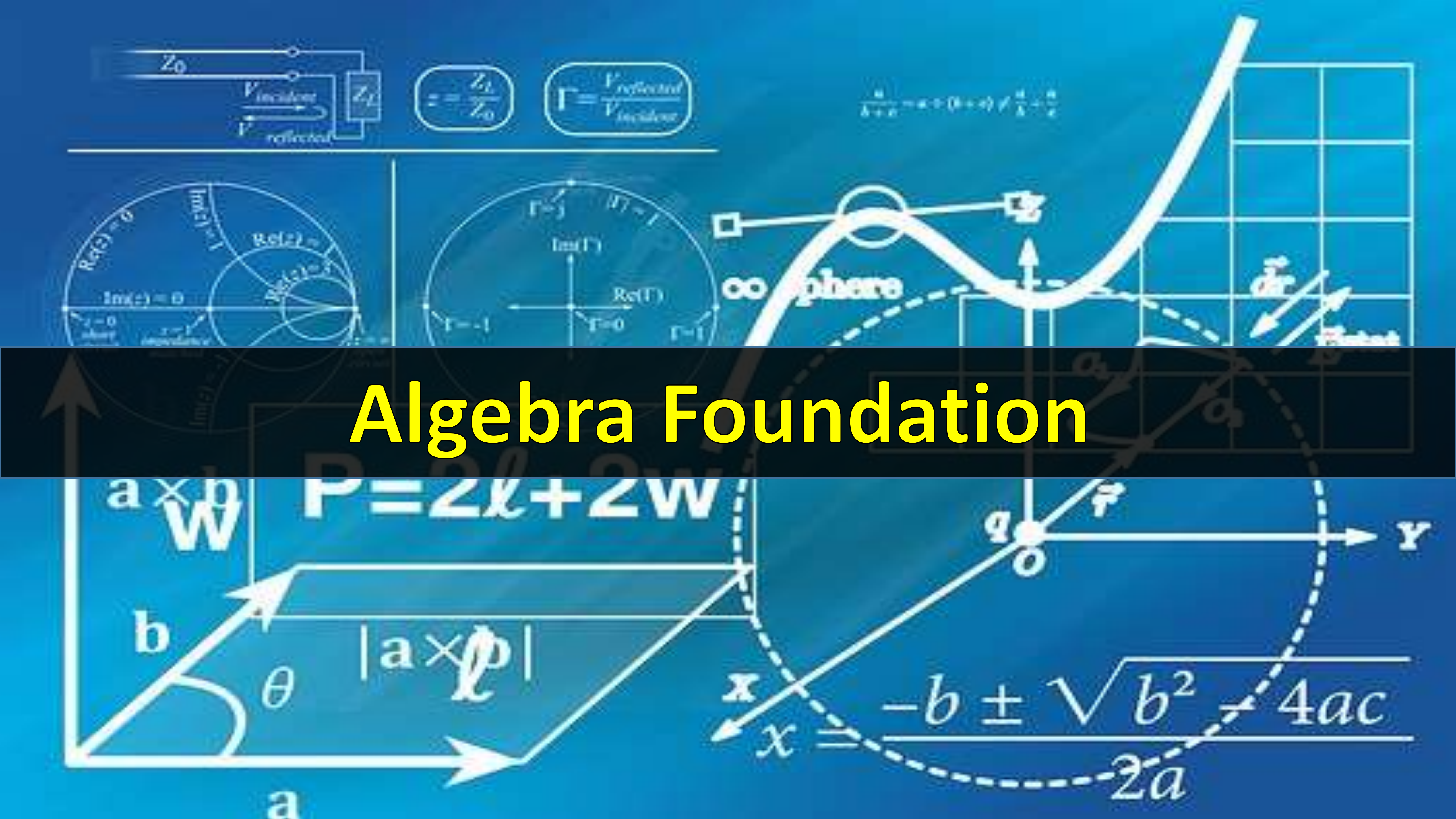
Math And Statistics For Data Science

By Eng. Mohammed Marwan Shahin

Outlines

- **Algebra Foundation** - Covers Algebraic Equations, Quadratic Equations and Functions
- **Calculus** - First order and second order derivatives, Partial Derivatives, Gradients
- **Linear Algebra** - Vectors and their Operations, Matrix and its operations, Vector Transformation using Matrices.
- **Probability** - Probability Basics, Conditional Probability, Random Variables and Random processes
- **Descriptive & Inferential Statistics**



Algebra Foundation

Balancing

Al-Kitāb al-mukhtaṣar fī hīsāb al-ğabr wa'l-muqābala

calculation

-al-Kwarizmi

820 AD

Algebra : Means Restoration or Completion

What is Algebra?

Variables

Constants

Equations

Left Side = Right Side

Equation

$$3X + 4 = 10$$

$$3X + 4 - 4 = 10 - 4$$

$$\frac{3X}{3} = \frac{6}{3}$$

$$X = 2$$

Equation

$$3X + 4 = 10$$

$$3(2) + 4 = 10$$

$$6 + 4 = 10$$

Distributive Property

$$3\,(X + 2) = 12$$

$$3X + 6 = 12$$

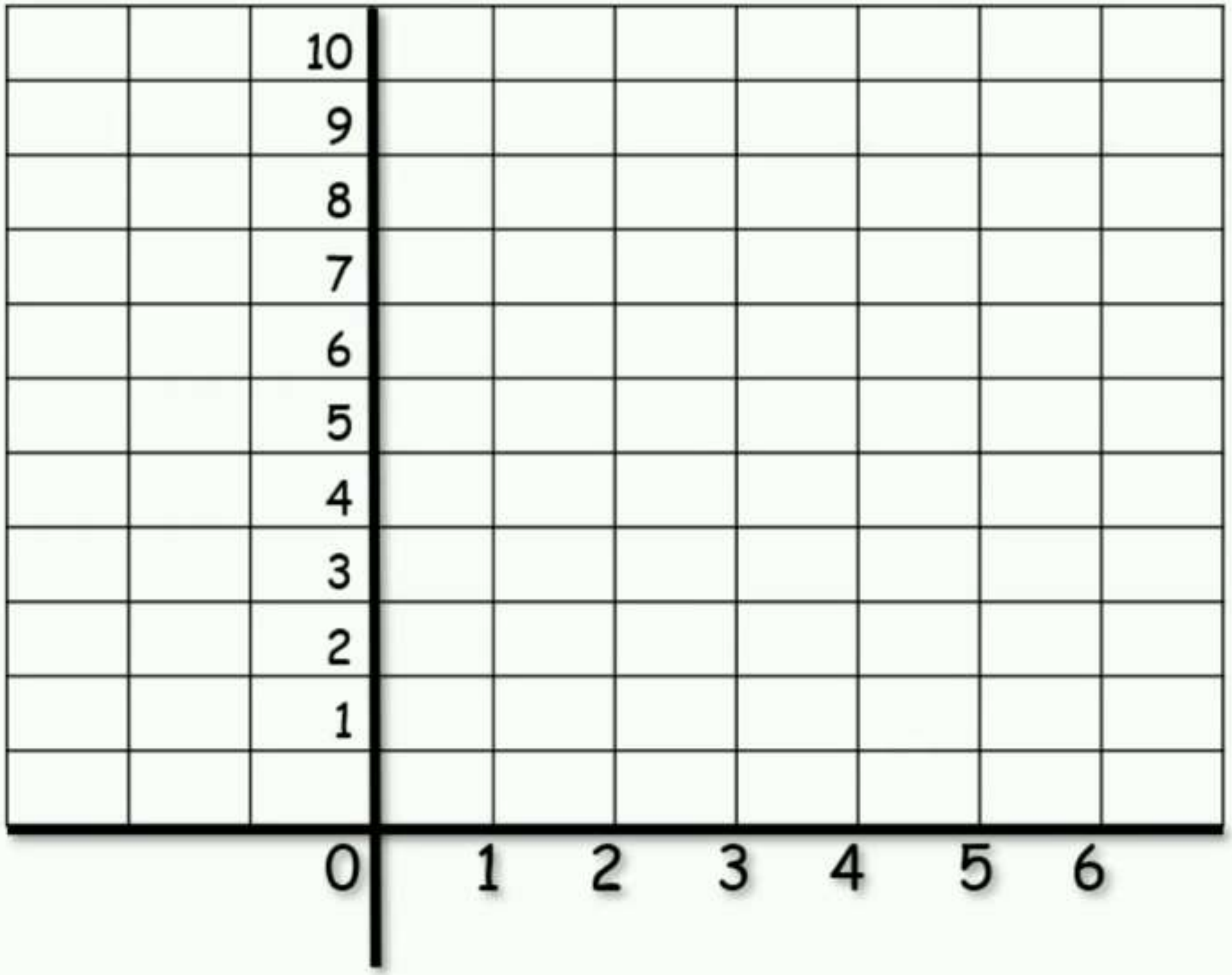
Variable

$$Age = Grade + 5$$

$$y = x + 5$$

Linear Equation

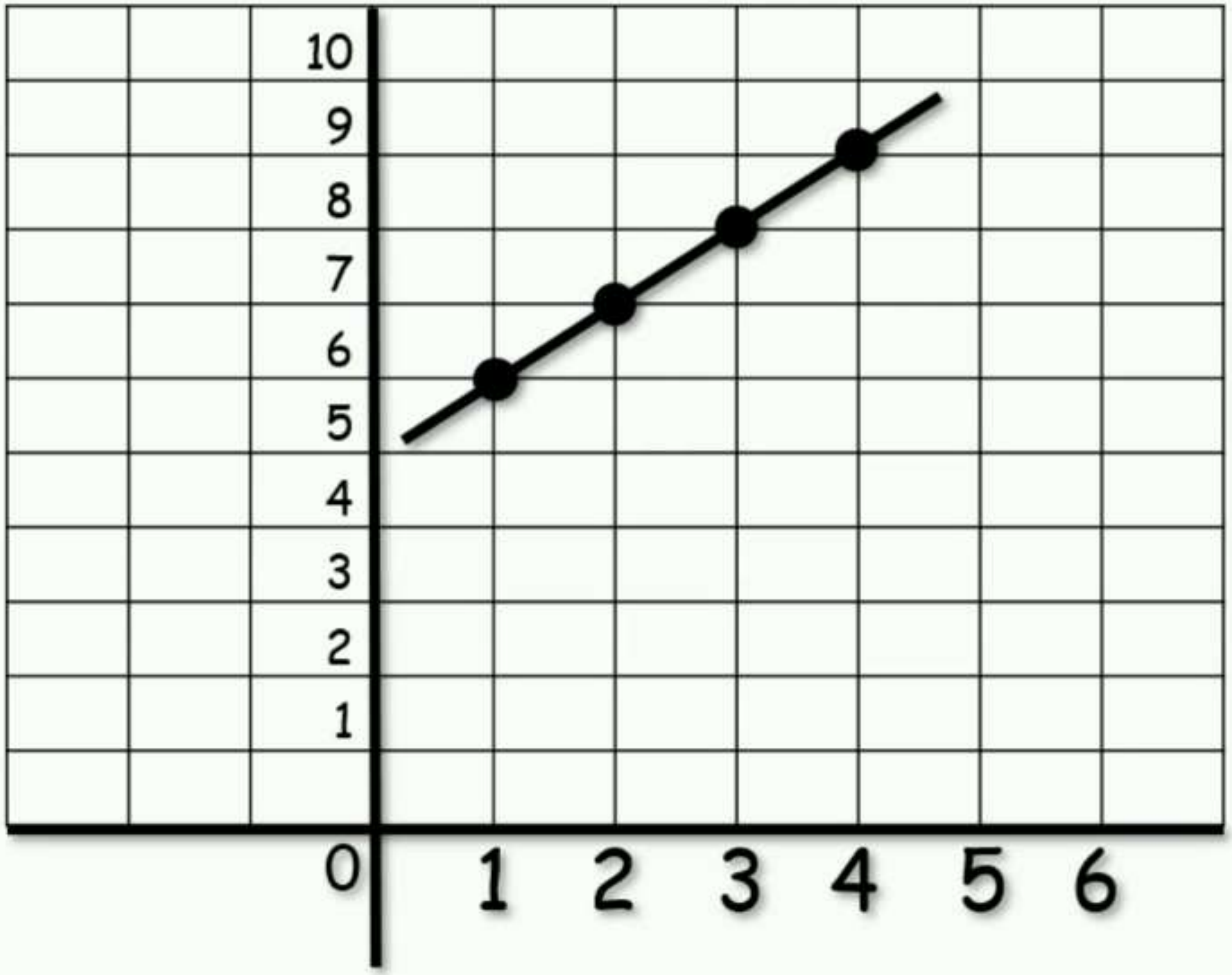
Linear Equation



$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

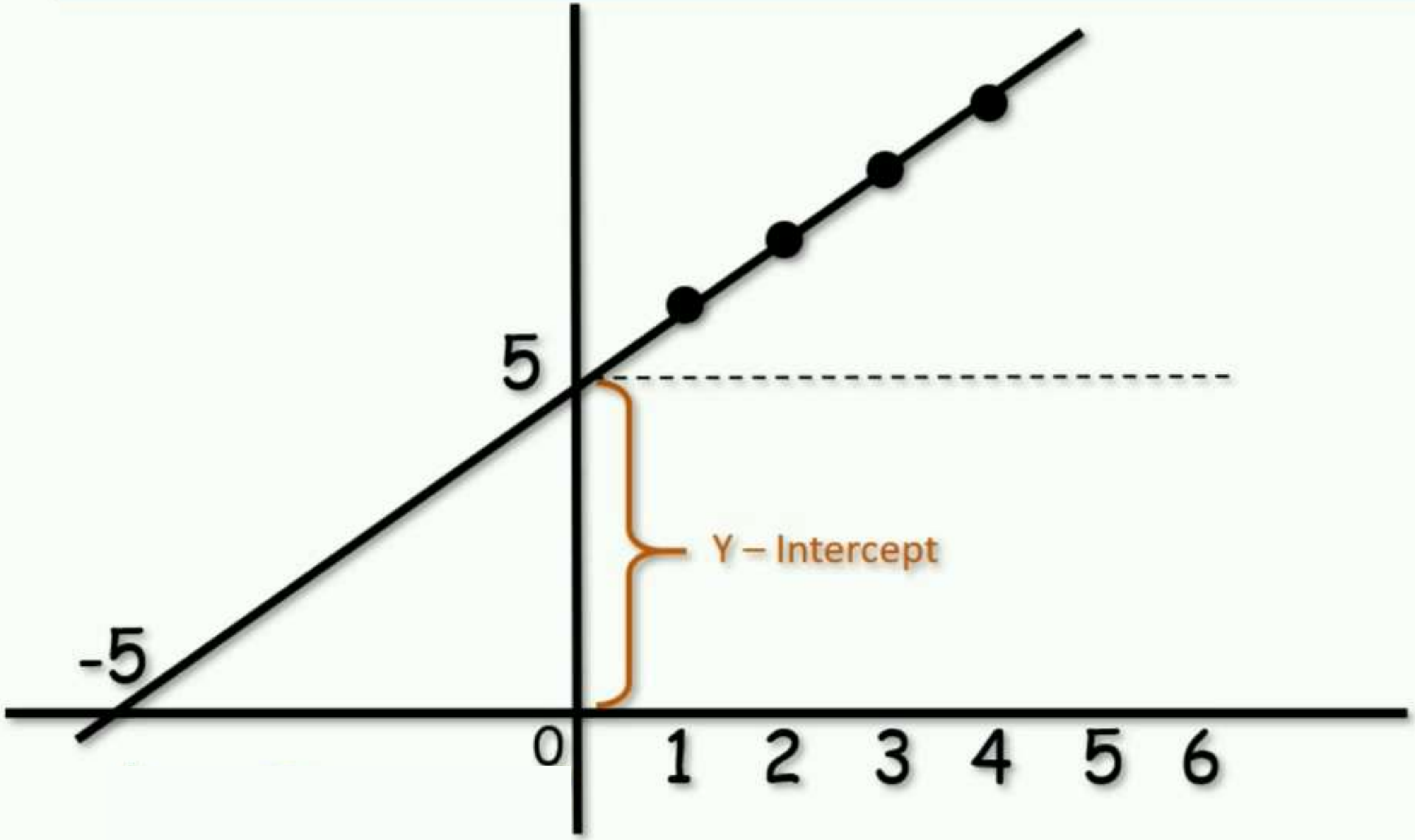
Linear Equation



$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

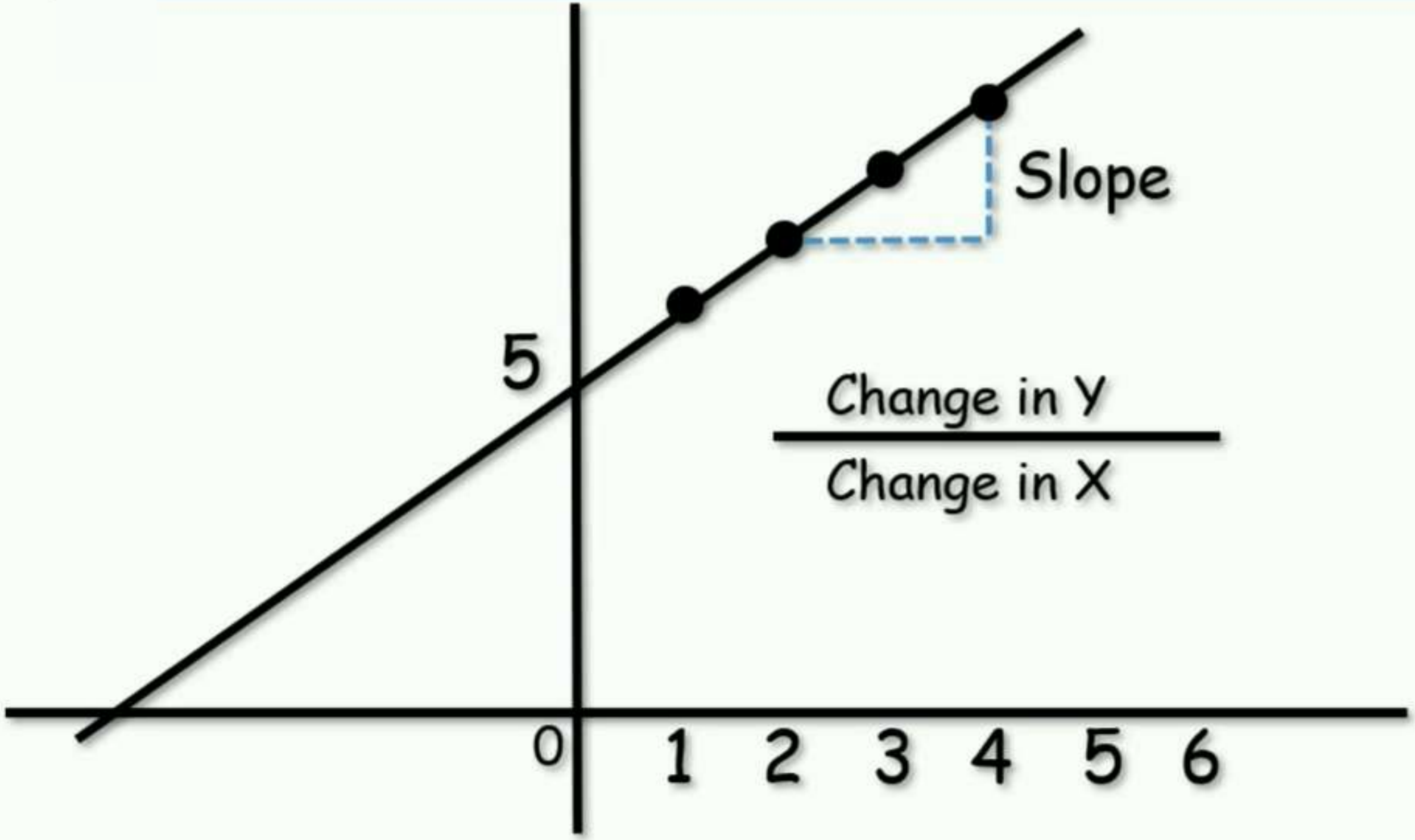
Linear Equation



$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

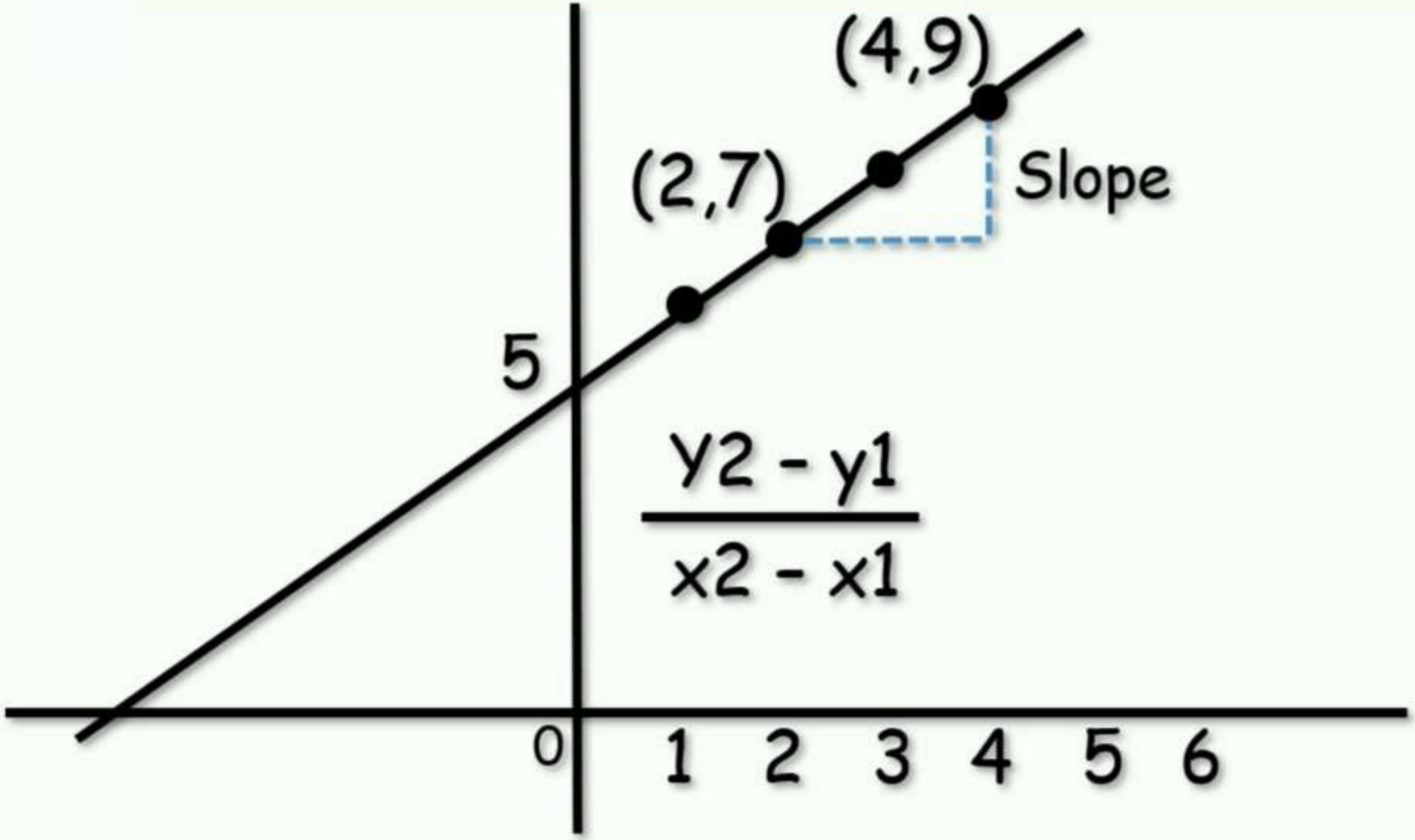
Linear Equation



$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

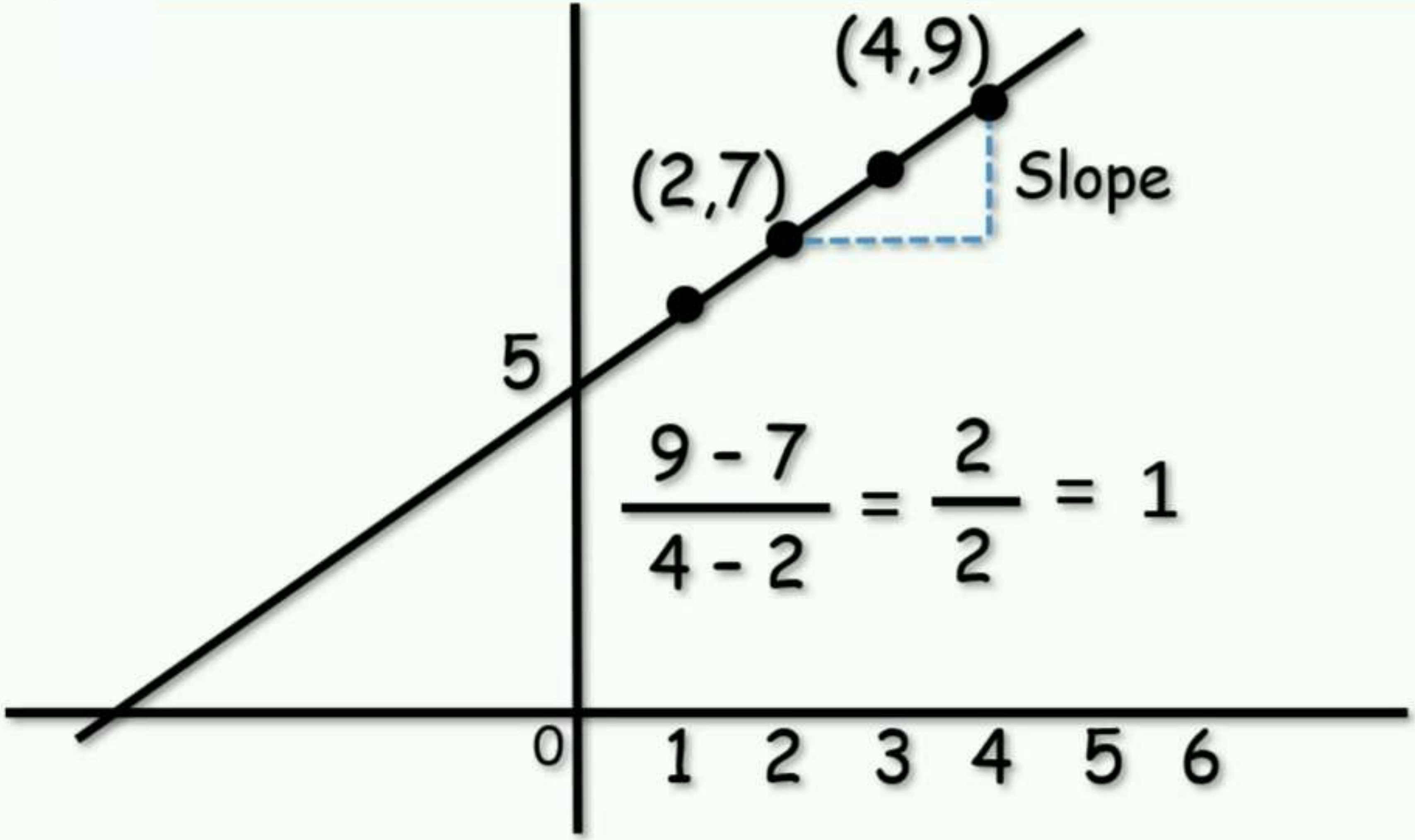
Linear Equation



$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

Linear Equation



$$\frac{9-7}{4-2} = \frac{2}{2} = 1$$

$$y = x + 5$$

x	y
1	6
2	7
3	8
4	9

Algebra Subjects

Exponents

Logarithm

Polynomial

Factoring

Quadratic Equations

Exponents

- How many times we should multiply a number by itself?

Index or exponent or power

$$4^2 = 4 \times 4$$

Base

$$4^3 = \underline{4} \times \underline{4} \times \underline{4}$$

Exponents

$4^{-3} =$

$4^{0} =$

Exponents

$$4^{-3} = 1 \div (4 \times 4 \times 4)$$

$$4^{0} = 1$$

Exponents Arithmetic

$$x^3 * x^2 = \boxed{}$$

Exponents Arithmetic

$$X^3 * X^2 = X^{3+2} = X^5$$

$$X * X * X \quad * \quad X * X$$

Exponents Arithmetic

$$X^3 \times X^2 = X^{3+2} = X^5$$

$$X^3 \div X^2 =$$

Exponents Arithmetic

$$X^3 \times X^2 = X^{3+2} = X^5$$

$$X^3 \div X^2 = X^{3-2} = X^1$$

$$\frac{X * X * X}{X * X}$$

In this lecture...

- Exponents
- Logarithm
- Polynomial
- Factoring
- Quadratic Equations

Logarithm

$$4^3 = 64$$

$$4^? = 64 \Rightarrow \log_4(64) = \boxed{}$$

Logarithm

$$4^3 = 64$$

$$4^? = 64 \quad \Rightarrow \quad \log_4(64) = 3$$

Logarithm

$$\log_2(4) = ?$$

$$2^{?} = 4$$

In this lecture...

- Exponents
- Logarithm
- Polynomial
- Factoring
- Quadratic Equations

Polynomial

- Poly → Many
- Nomial → Terms

$$2x^2 \quad + \quad 3x \quad + 5$$

exponents variables constants

terms

Polynomial Arithmetic

- Poly → Many
- Nomial → Terms

$$
\begin{array}{r}
2x^2 + 3x + 5 \\
+ \quad 3x^2 - 2x + 3 \\
\hline
\end{array}
$$

Polynomial Arithmetic

- Poly → Many
- Nomial → Terms

$$
\begin{array}{rrrr}
 & 2x^2 & +\ 3x & +\ 5 \\
+ & 3x^2 & -\ 2x & +\ 3 \\
\hline
 & 5x^2 & +\ 1x & +\ 8
\end{array}
$$

Polynomial Arithmetic

$$
\begin{array}{r}
2x^2 + 3x + 5 \\
\times \quad 3x^2 - 2x + 3 \\
\hline
\end{array}
$$

Polynomial Arithmetic

$$
\begin{array}{llllll}
 & 2x^2 & + 3x & + 5 & & \\
\times & 3x^2 & - 2x & + 3 & & \\
\hline
 & 6x^4 & - 4x^3 & + 6x^2 & & \\
 & & 9x^3 & - 6x^2 & + 9x & \\
 & & & 15x^2 & - 10x & + 15
\end{array}
$$

In this lecture...

- Exponents
- Logarithm
- Polynomial
- Factoring
- Quadratic Equations

Factoring

- What can I multiply with what to get the required equation or number?

$$5 \times 3 = 15$$

Factors

Factoring

- What can I multiply with what to get the required equation or number?

$$3x + 9 \Rightarrow 3\,(x + 3)$$

Factoring

- What can I multiply with what to get the required equation or number?

$$2x^3 - 8x \rightarrow \boxed{}$$

Factoring

- What can I multiply with what to get the required equation or number?

$$2x^3 - 8x \Rightarrow 2\ (x^3 - 4x) \Rightarrow 2x\ \boxed{(x^2 - 4)}$$

Difference of Squares

Difference of Squares

$$(a^2 - b^2) =$$

$$(x^2 - 4) =$$

Difference of Squares

$$(a^2 - b^2) = (a + b)(a - b)$$

$$(x^2 - 4) =$$

Difference of Squares

$$(a^2 - b^2) = (a + b)(a - b)$$

$$(x^2 - 4) = (x^2 - 2^2) = (x + 2)(x - 2)$$

Factoring

- What can I multiply with what to get the required equation or number?

$$2x^3 - 8x \rightarrow$$

Factoring

- What can I multiply with what to get the required equation or number?

$$2x^3 - 8x \rightarrow 2x(x+2)(x-2)$$

In this lecture...

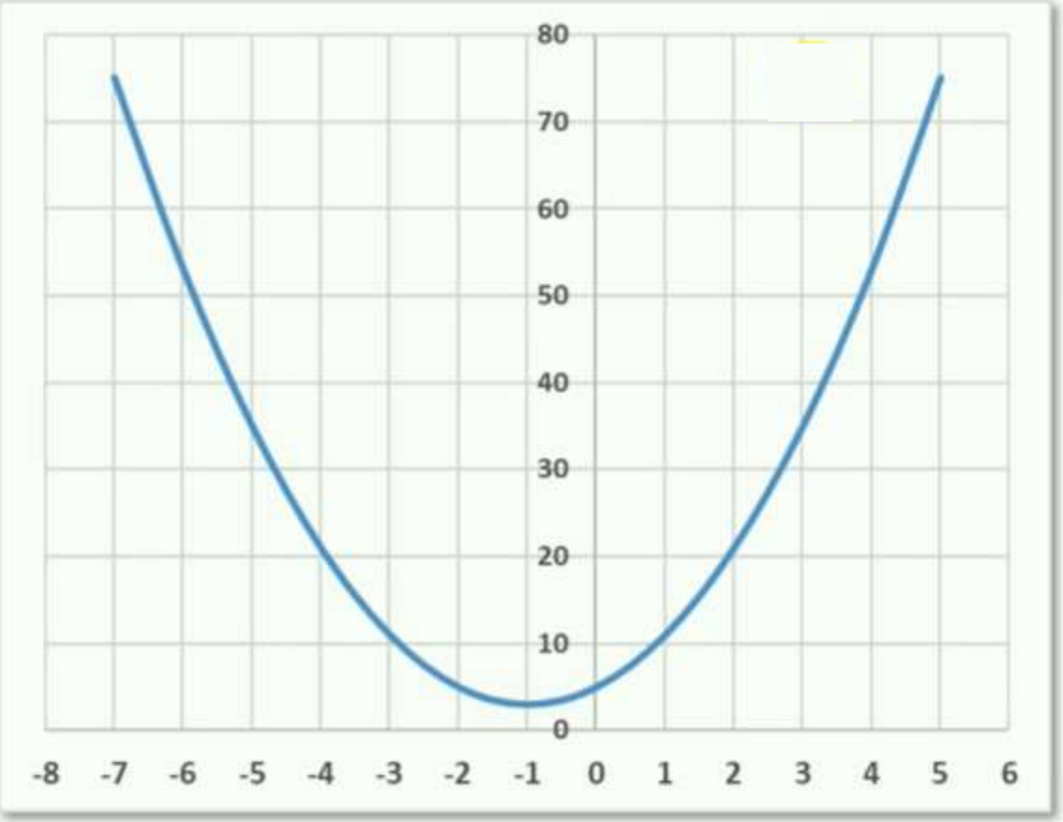
- Exponents
- Logarithm
- Polynomial
- Factoring
- **Quadratic Equations**

Quadratic Equation

Special type of polynomial with “Quad” or Square.

$$ax^2 + bx + c$$

$$2x^2 + 4x + 5$$



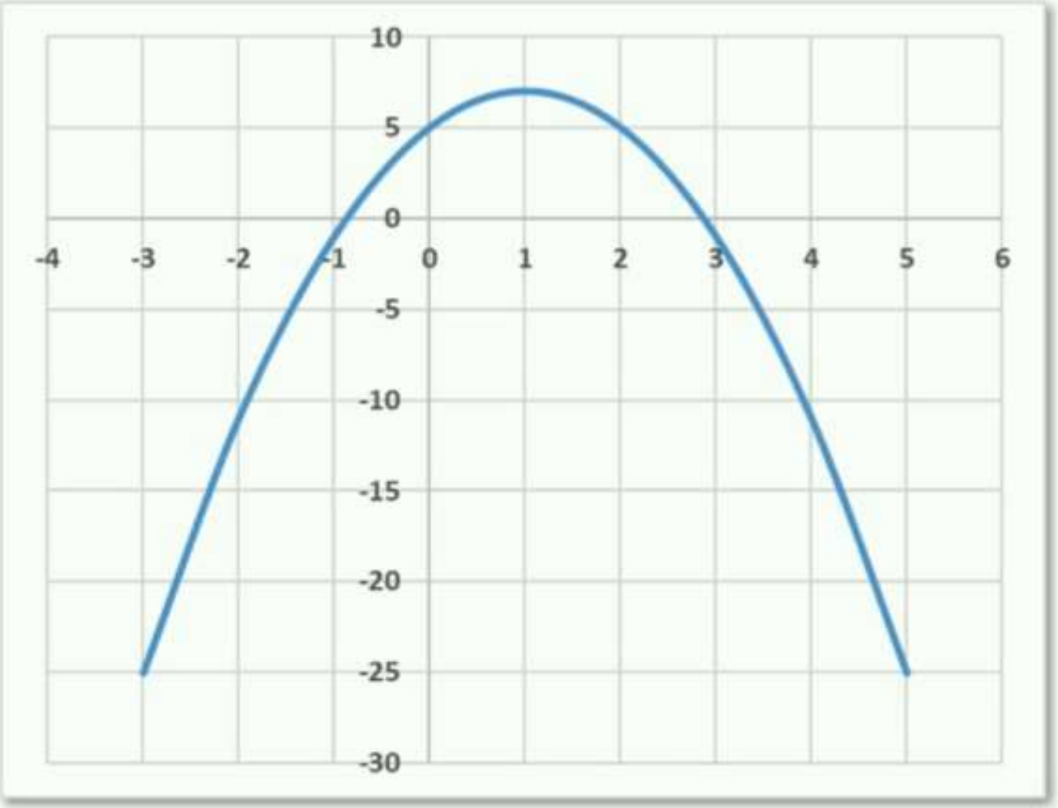
Quadratic equation

$-2x^2 + 4x + 5$

$2x^2 + 4x + 5$

Quadratic equation

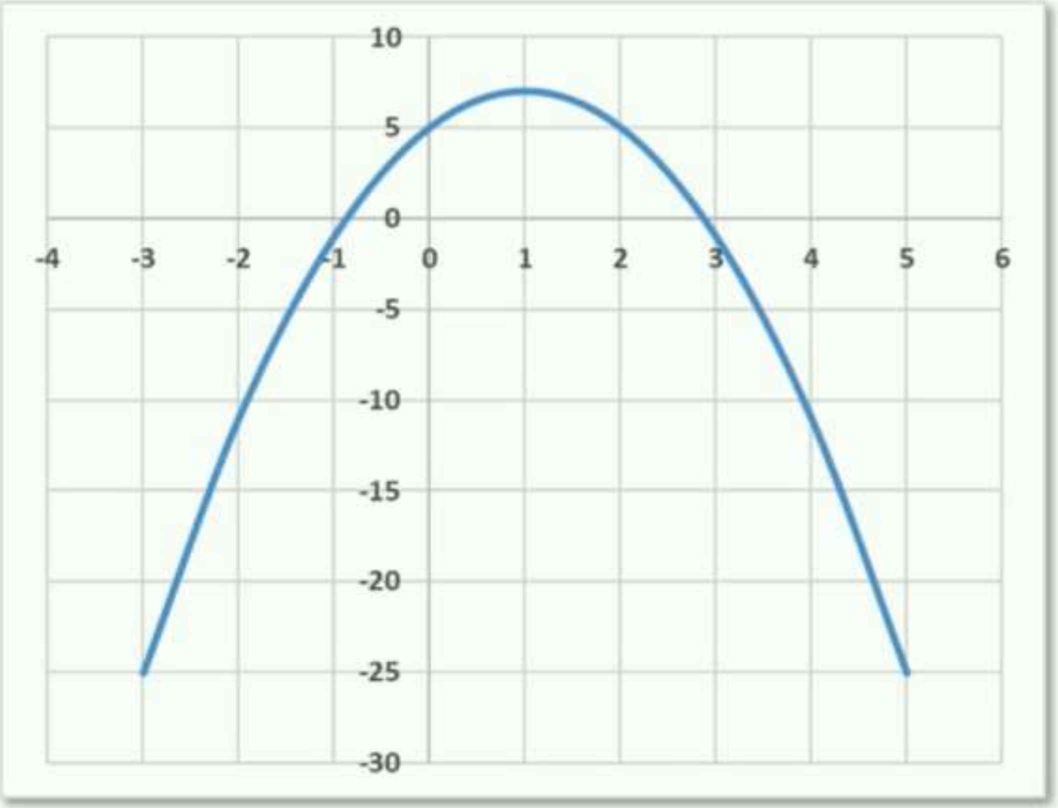
$-2x^2 + 4x + 5$



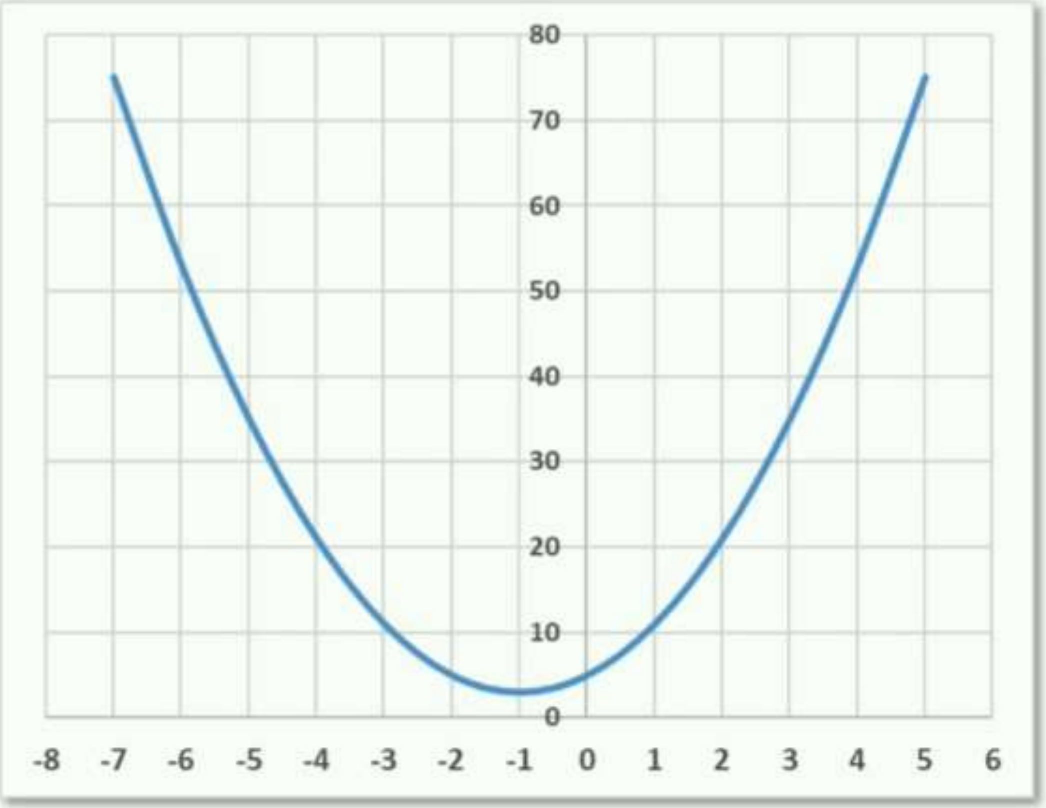
$2x^2 + 4x + 5$

Quadratic equation

$$-2x^2 + 4x + 5$$



$$2x^2 + 4x + 5$$



Common Equations

$$(x + a)^2 =$$

$$(x - a)^2 =$$

$$(x + a) * (x - a) =$$

Common Equations

$$(x + a)^2 = x^2 + 2xa + a^2$$

$$(x - a)^2 = x^2 - 2xa + a^2$$

$$(x + a) * (x - a) = x^2 - a^2$$

What is a Function?

Input → function → Output

Process the input

X → Square → X^2

$$f(x) = x^2$$

What is a Function?

$$f(x) = x^2$$

function input output

What is a Function?

$$f(x) = x^2$$

$$f(x) = x + 5$$

$$f(x) = 4x - 9$$

Age as a function of grade?

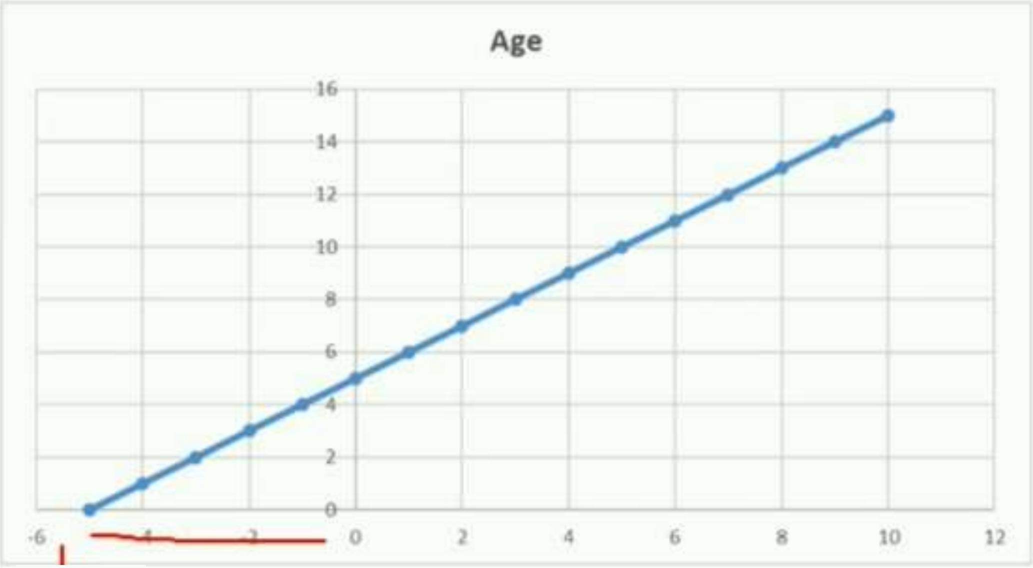
$$f(x) = x + 5$$

Grade	Age
1	1 + 5 = 6
2	2 + 5 = 7
3	3 + 5 = 8
4	4 + 5 = 9
5	5 + 5 = 10

Allowed values → 1 to 10

Age as a function of grade?

$$f(x) = x + 5$$

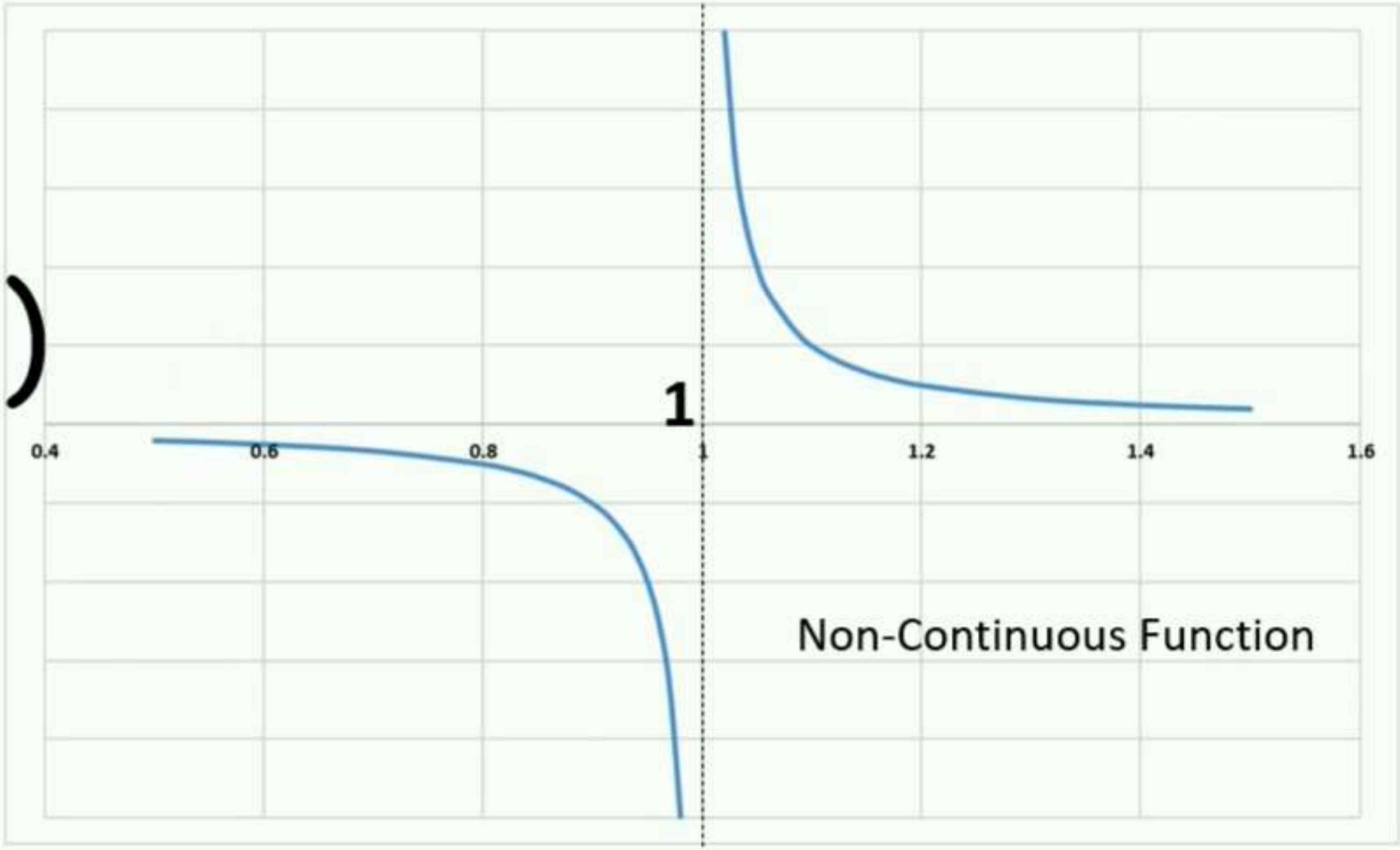


Grade	Age
1	1 + 5 = 6
2	2 + 5 = 7
3	3 + 5 = 8
4	4 + 5 = 9
5	5 + 5 = 10

Allowed values → 1 to 10

Continuous Functions

$$f(x) = 1/(x - 1)$$



It's not a continuous function : if it's continuous function we should draw it without lifting the pen.

? Quiz

Question 1:

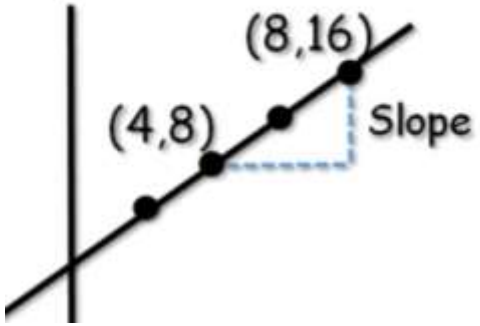
- **What will be the value of "x" when we solve the following equation?**

  $2x + 4 = 18$

- ○ 7
- ○ 11
- ○ 9
- ○ 6

Question 2:

- **What will be the slope of the line in this plot**



- ○ 4
- ○ 1
- ○ 2
- ○ 8

Question 3:

- **What will be the value of the following equation?**

$$\log_2 (8) =$$

- ○ 3
- ○ 4
- ○ 2
- ○ None of the above.

Question 4:

$$f(x) = 1/x$$

- **The following function is NOT a continuous function. True or False?**

True

False

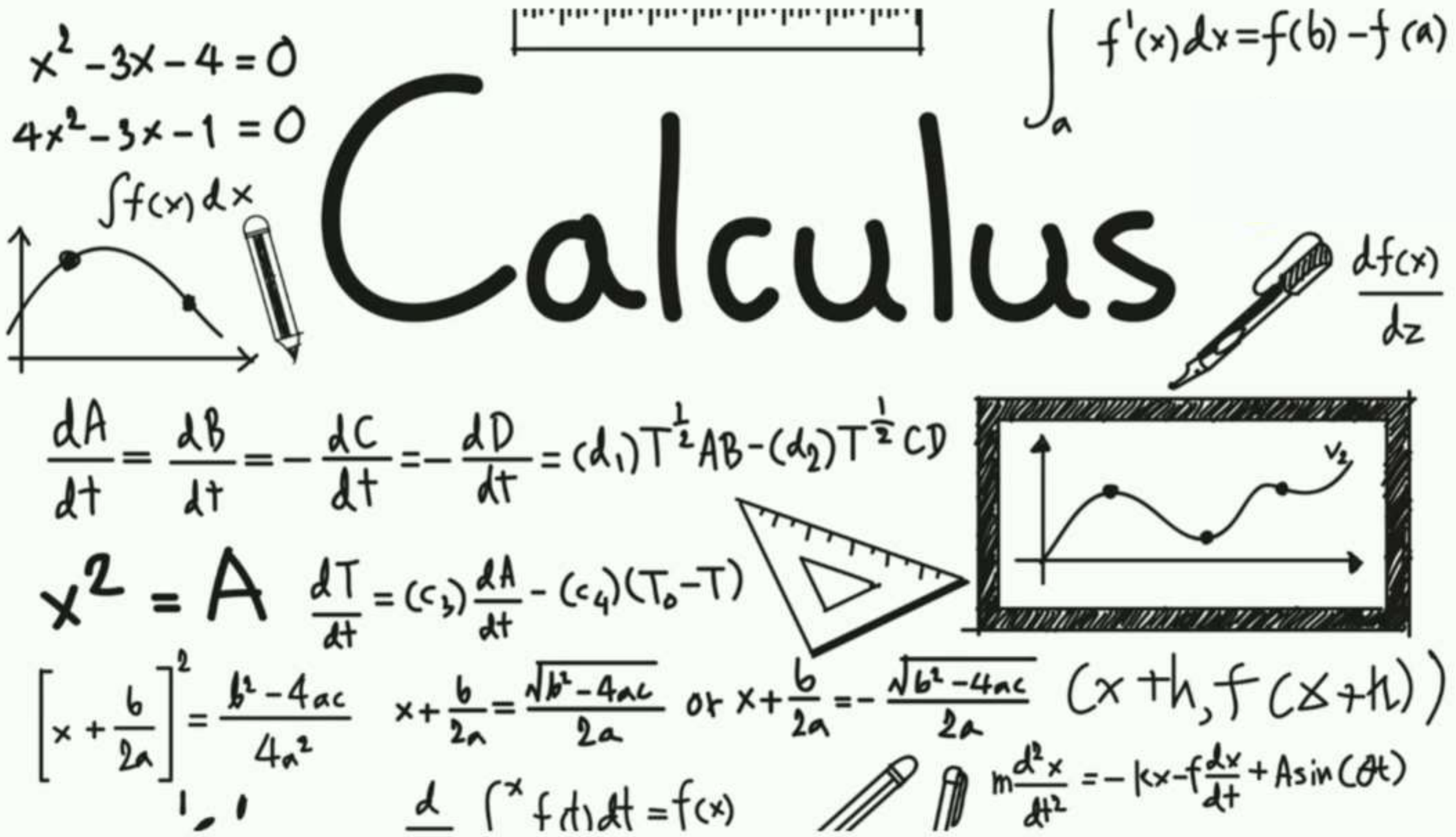
Question 4:

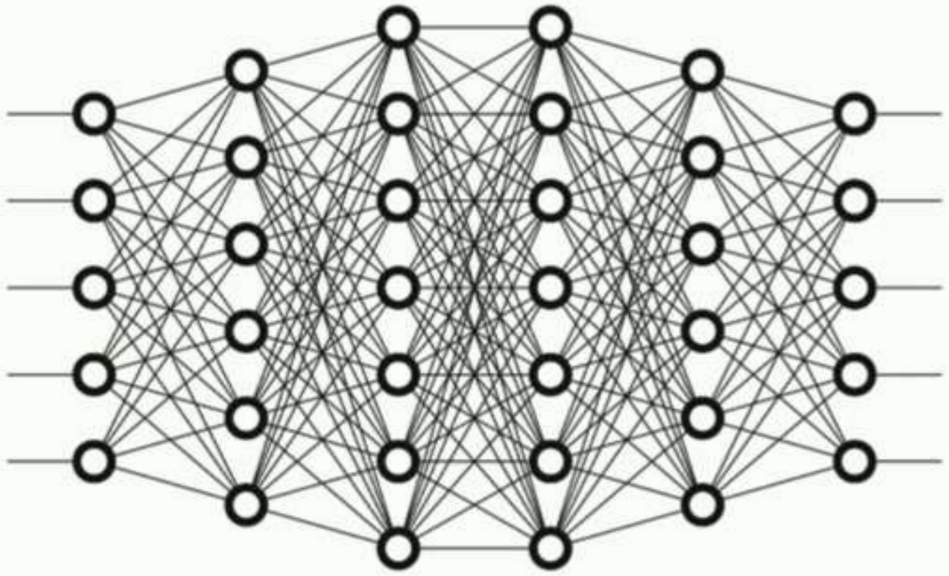
$$f(x) = 1/x$$

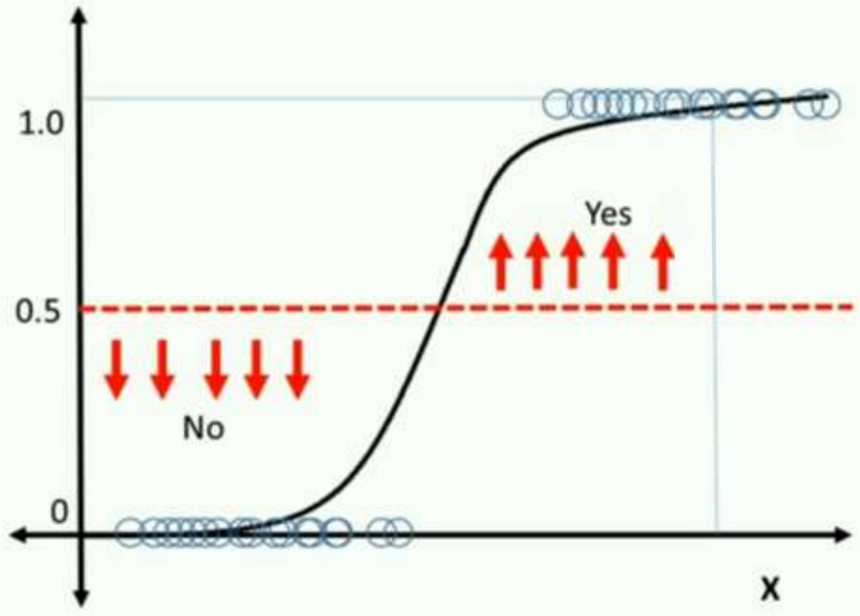
- **The following function is NOT a continuous function. True or False?**

True

False


Calculus
x² − 3x − 4 = 0
4x² − 3x − 1 = 0
∫f(x) dx
∫ₐ f'(x) dx = f(b) − f(a)
df(x)/dz
dA/dt = dB/dt = − dC/dt = − dD/dt = (d₁)T^½ AB − (d₂)T^½ CD
v₂
x² = A
dT/dt = (c₃) dA/dt − (c₄)(T₀ − T)
[x + b/2a]² = (b² − 4ac)/4a²
x + b/2a = √(b² − 4ac)/2a or x + b/2a = − √(b² − 4ac)/2a
(x + h, f(x + h))
m d²x/dt² = − kx − f dx/dt + A sin(θt)
d/ ∫ˣ f(t) dt = f(x)





$$\frac{dy}{dx} = 24x^3 - 6x^2 - 24x + 1$$

$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$

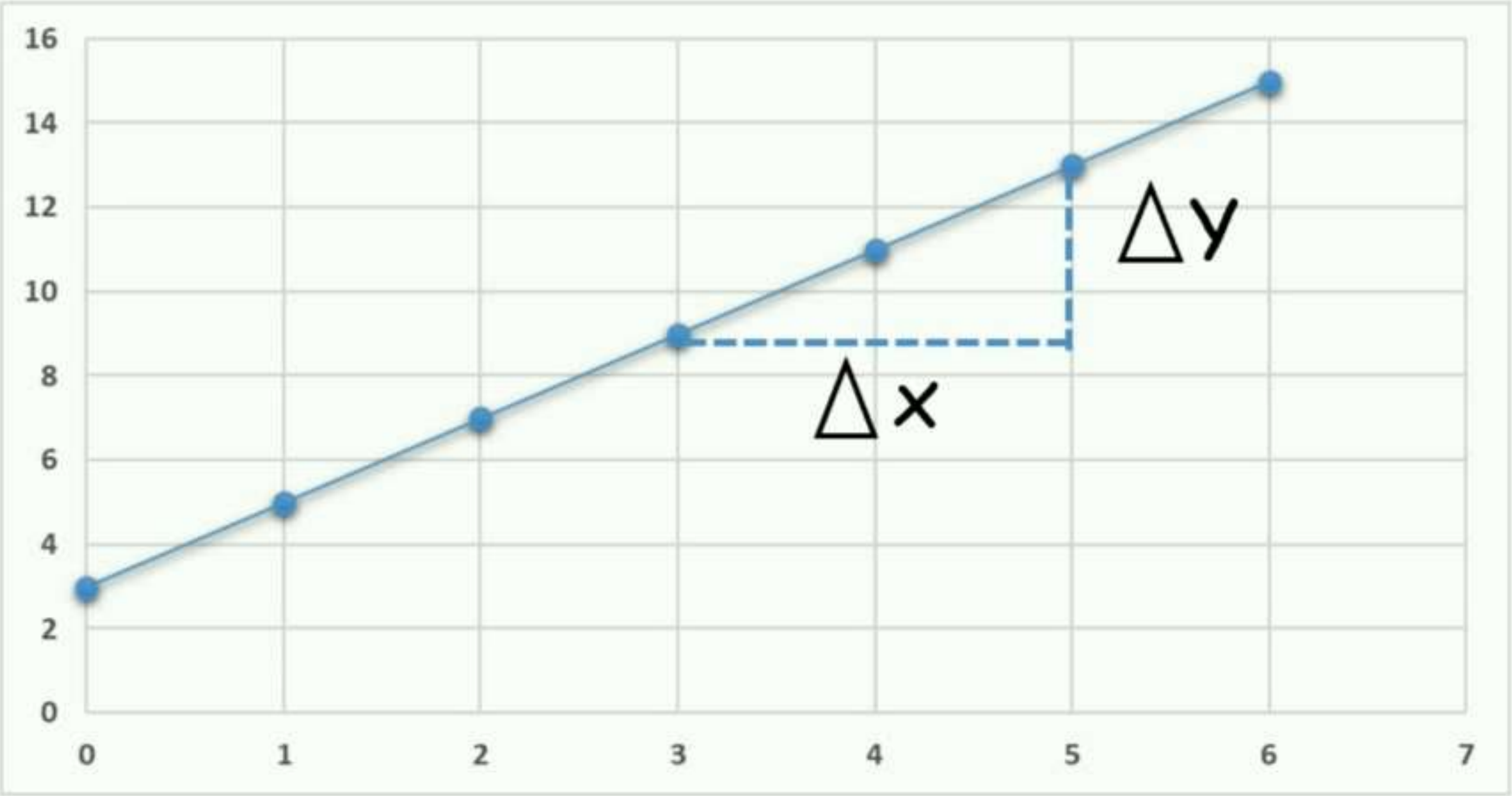
$$\frac{\partial(f(x,y))}{\partial x} = 2X$$

Rate of Change

$$y = 2x + 3$$

Rate of Change

$$= \frac{\Delta y}{\Delta x}$$



Rate of Change

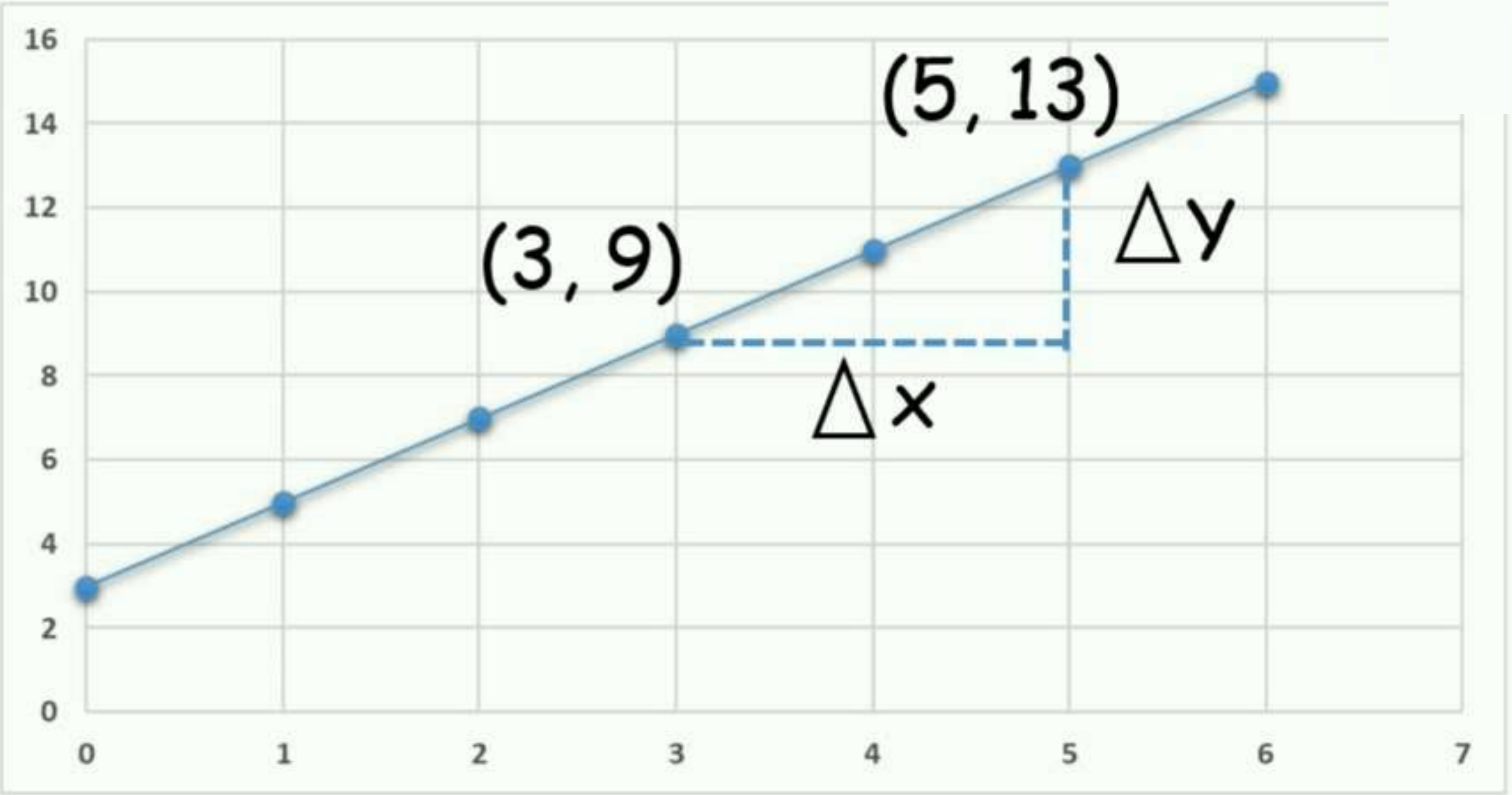
$$y = 2x + 3$$

Rate of Change

$$= \frac{\Delta y}{\Delta x}$$

$$= \frac{13 - 9}{5 - 3}$$

$$= 2$$

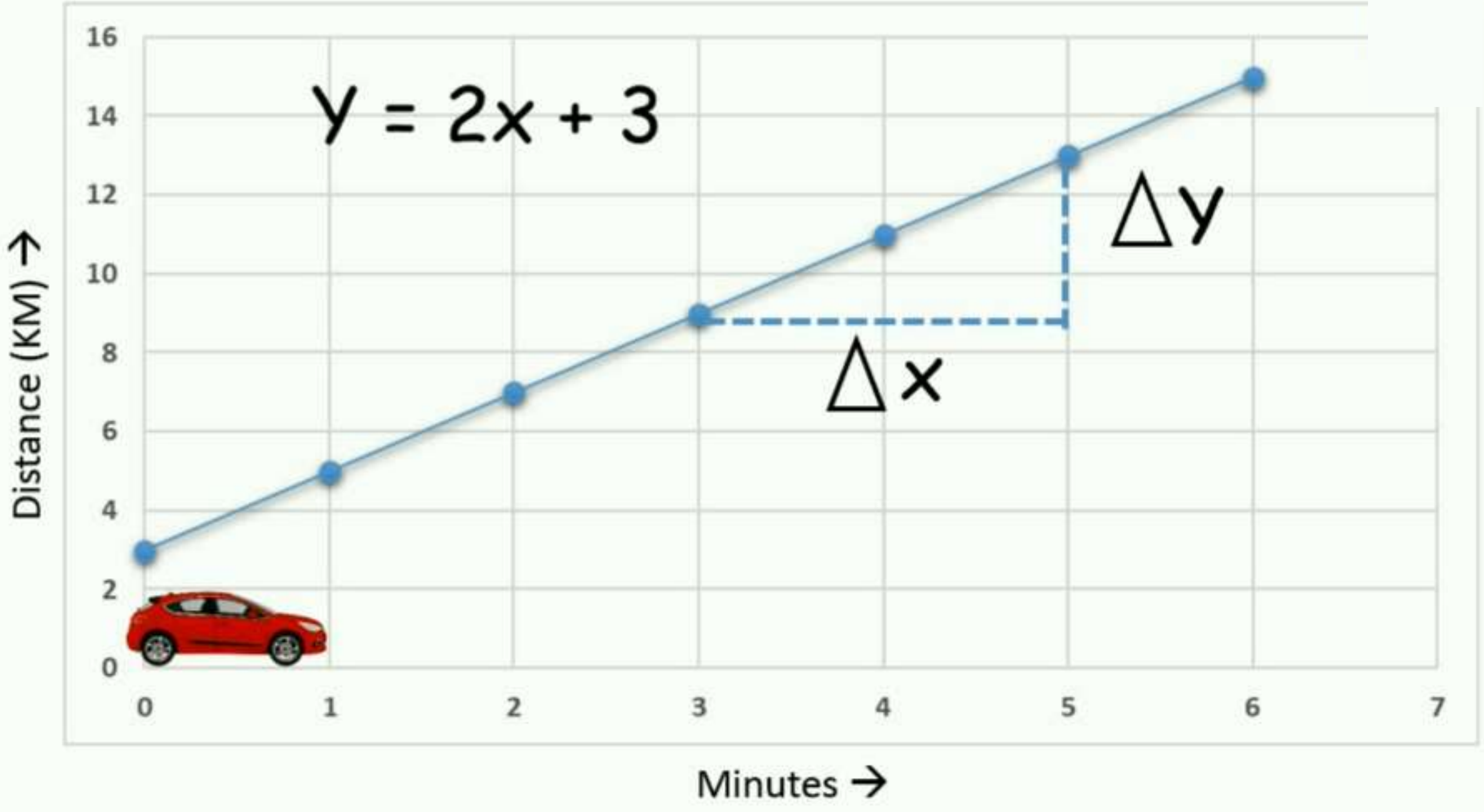


Rate of Change

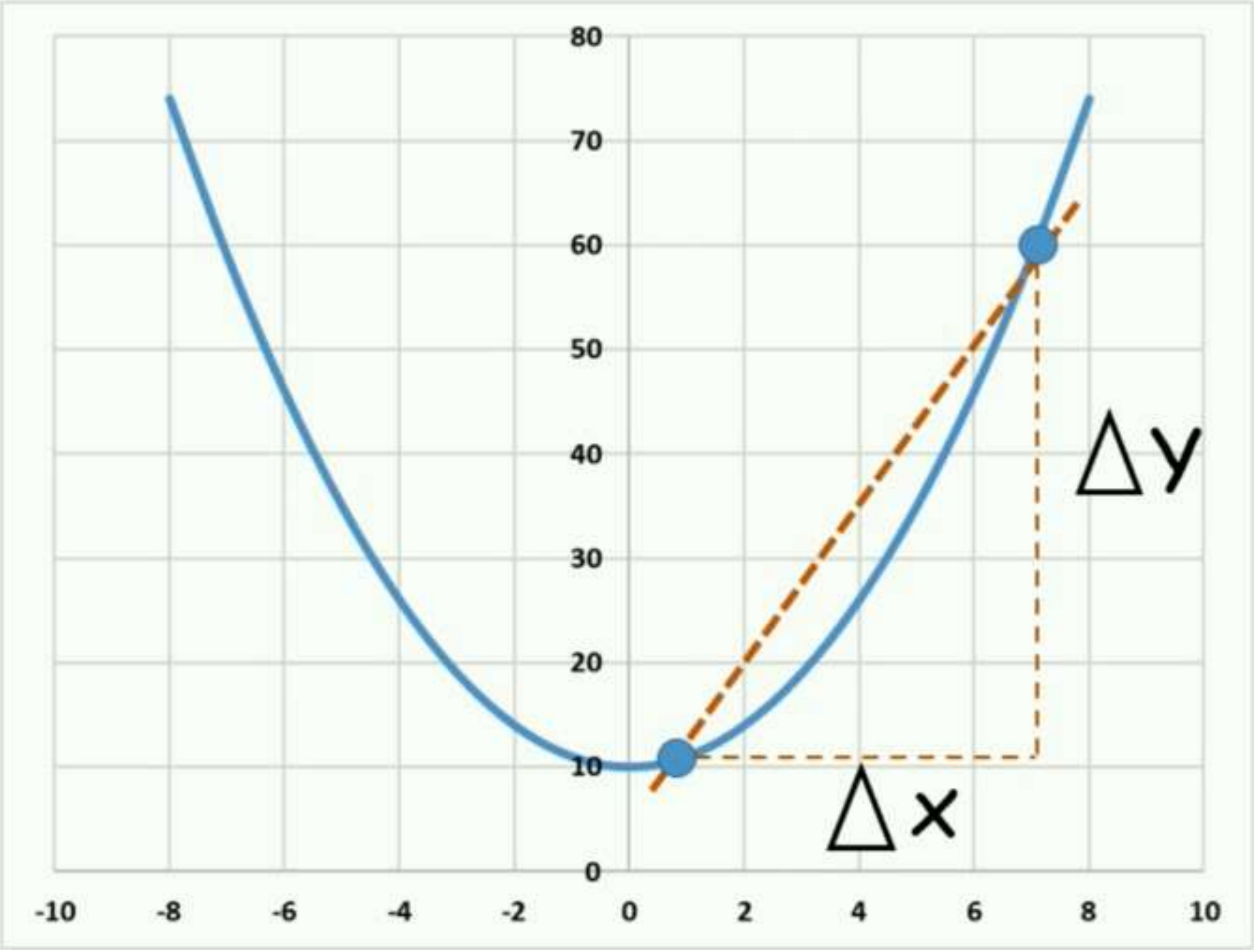
Rate of Change

$$\frac{\Delta \text{ Distance}}{\Delta \text{ Time}}$$

$= \quad 2\text{KM/minute}$



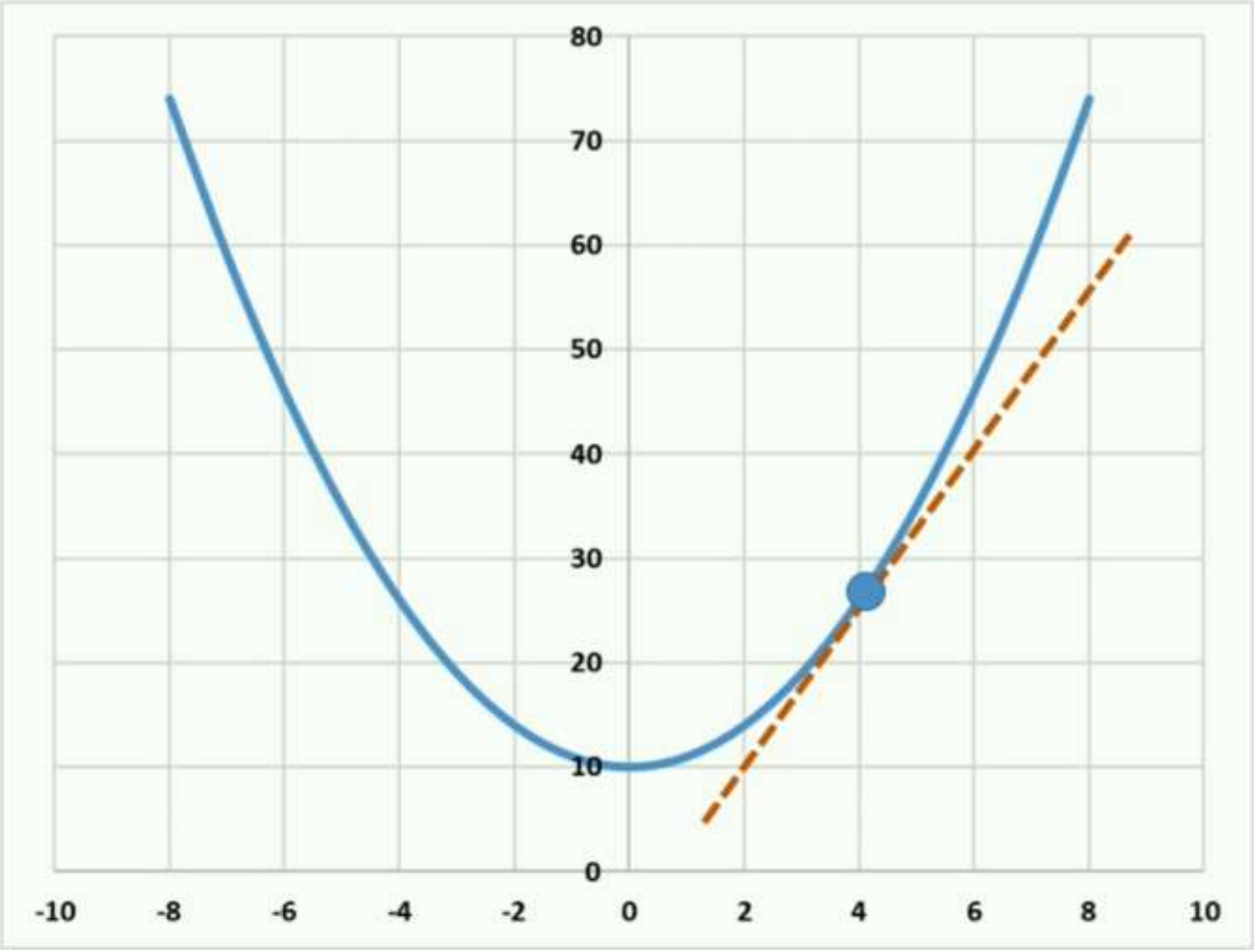
Slope between two points



Average Slope

$$= \frac{\Delta y}{\Delta x}$$

Slope at “The Point” or Gradient

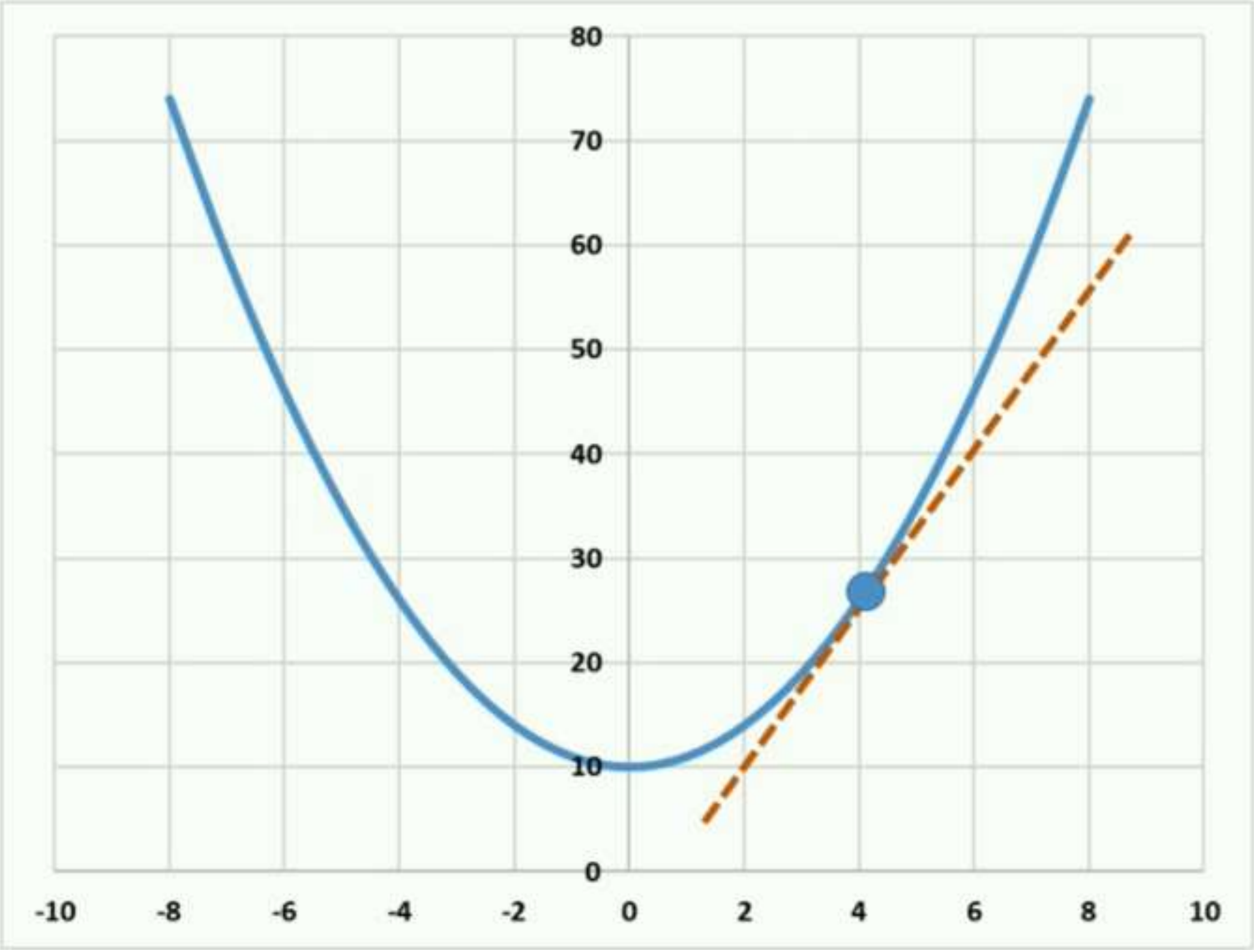


$$\Delta y = 0$$

$$\Delta X = 0$$

$$\text{Slope} = \frac{\Delta y}{\Delta x} = \text{undefined}$$

Slope at "The Point" or Gradient

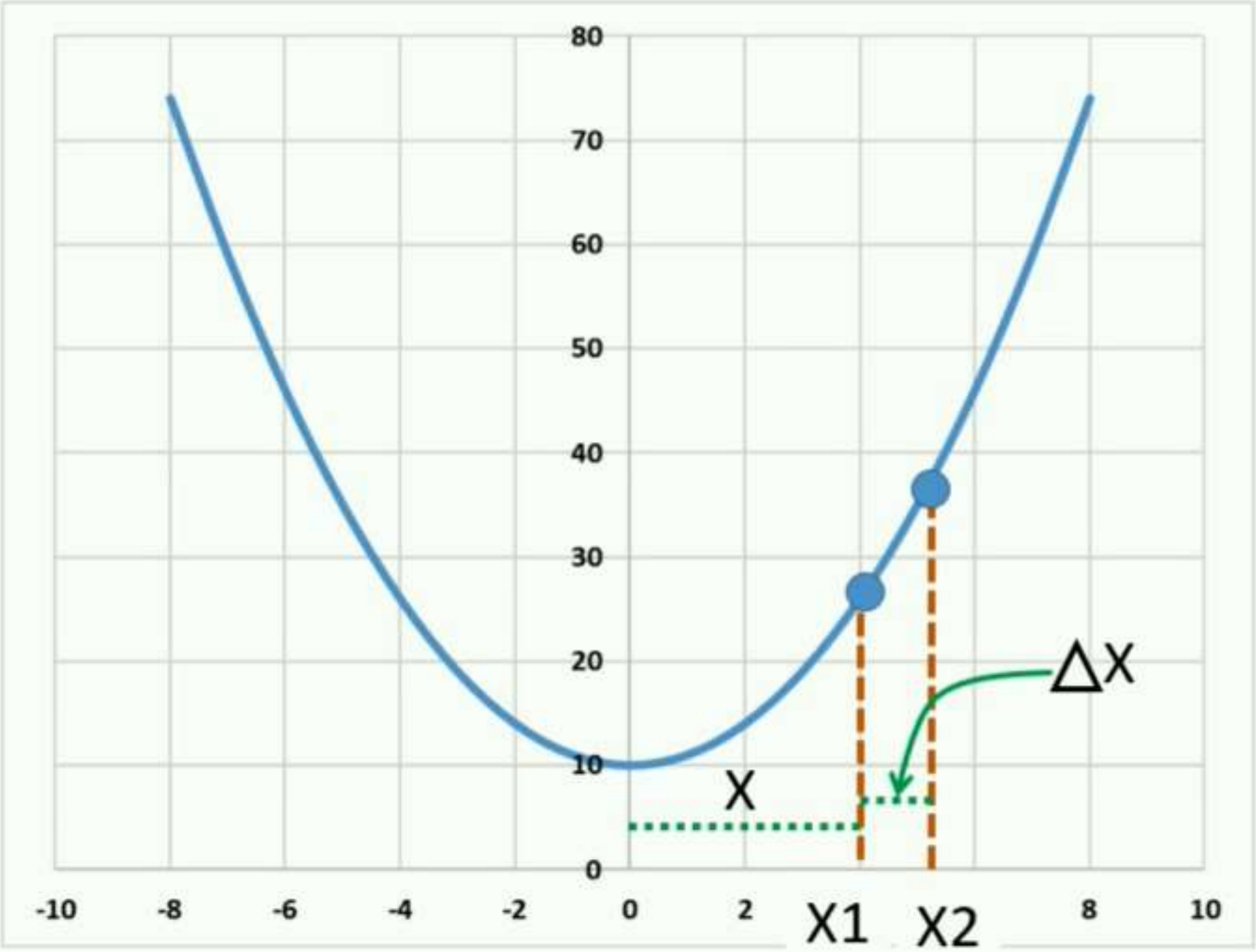


$\Delta y = 0$

$\Delta X = 0$

$$Y = f(x) = x^2 + 10$$

Slope at the point – Rate of Change



$$Y = f(x) = x^2 + 10$$

$$\Delta Y = y_2 - y_1 \qquad \Delta X = x_2 - x_1$$

$$x_1 = x; \qquad x_2 = x + \Delta x$$

$$Y_2 = f(x_2) = f(x + \Delta x)$$

$$= (x + \Delta x)^2 + 10$$

Slope at the point – Rate of Change

$$Y_2 = f(x_2) = f(x + \Delta x)$$

$$= (x + \Delta x)^2 + 10$$

$$= (x^2 + 2x^*\Delta x + \Delta x^2) + 10$$

Slope at the point – Rate of Change

$$Y_2 = f(x_2) = f(x + \Delta x)$$

$$= (x + \Delta x)^2 + 10$$

$$= (x^2 + 2x^*\Delta x + \Delta x^2) + 10$$

$$\Delta Y = y_2 - y_1 = \underbrace{\boxed{(x^2 + 2x^*\Delta x + \Delta x^2) + 10}}_{y_2} - \underbrace{\boxed{(x^2 + 10)}}_{y_1}$$

Slope at the point – Rate of Change

$$Y_2 = f(x_2) = f(x + \Delta x)$$

$$= (x + \Delta x)^2 + 10$$

$$= (x^2 + 2x * \Delta x + \Delta x^2) + 10$$

$$\Delta Y = y_2 - y_1 = \overset{y_2}{\boxed{(\cancel{x^2} + 2x * \Delta x + \Delta x^2) + \cancel{10}}} \overset{-}{-} \overset{y_1}{\boxed{(\cancel{x^2} + \cancel{10})}}$$

Slope at the point – Rate of Change

$$\Delta Y = y_2 - y_1 = \underbrace{\boxed{(\cancel{x^2} + 2x^*\Delta x + \Delta x^2) + \cancel{10}}}_{} - \boxed{(\cancel{x^2} + \cancel{10})}$$

(above the boxes: y_2 – y_1)

$$\Delta Y = 2x^*\Delta x + \Delta x^2$$

$$= \Delta x \ (2x + \Delta x)$$

Distributive Property

3 (X + 2) = 12

3X + 6 = 12

Slope at the point – Rate of Change

$$\Delta Y = \Delta x \ (2x + \Delta x)$$

$$\text{Average Slope} = \frac{\Delta y}{\Delta x} = \frac{\Delta x \ (2x + \Delta x)}{\Delta x}$$

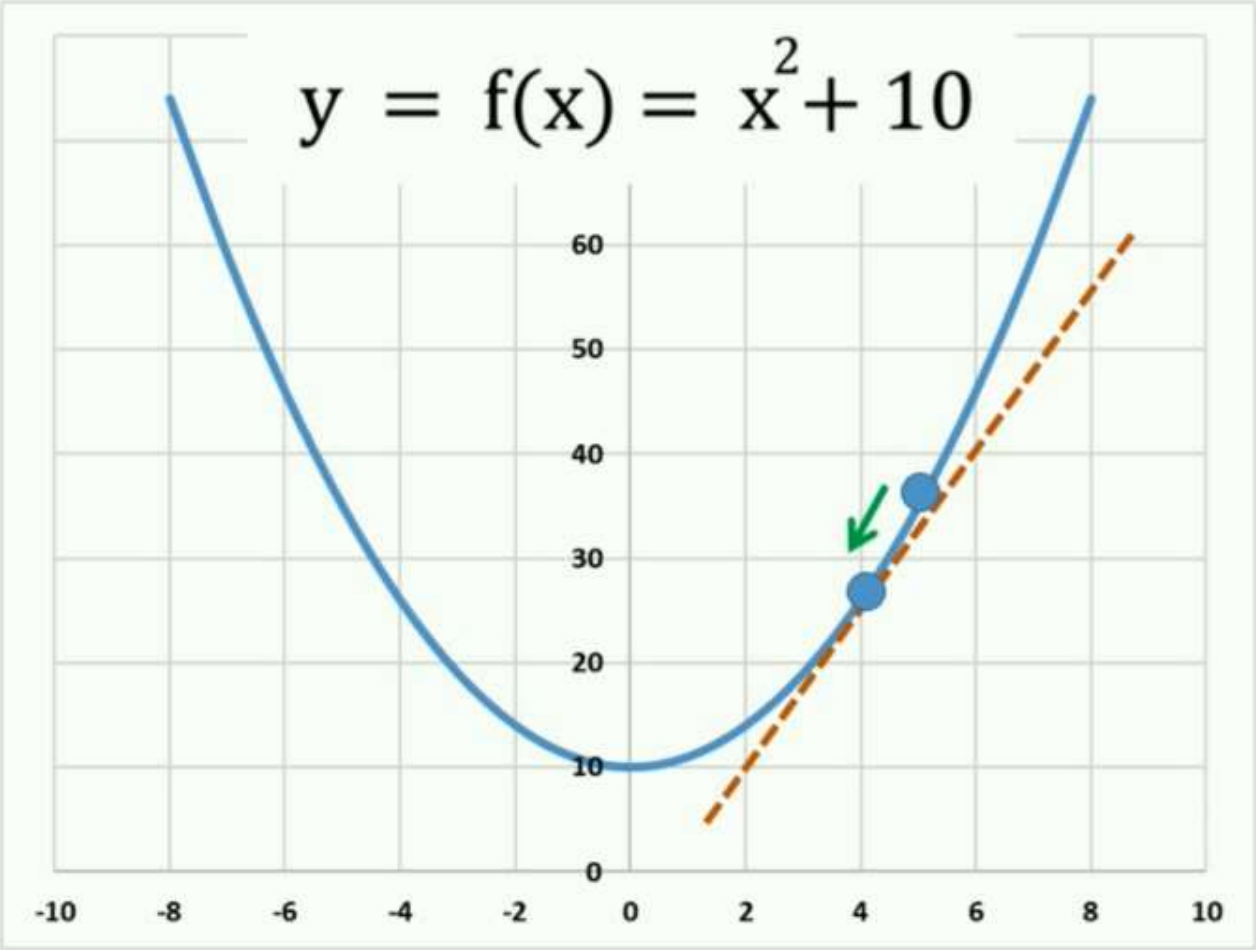
Slope at the point – Rate of Change

$$\Delta Y = \Delta x \, (2x + \Delta x)$$

$$\text{Average Slope} = \frac{\Delta y}{\Delta x} = \frac{\cancel{\Delta x}\,(2x + \Delta x)}{\cancel{\Delta x}}$$

$$\text{Average Slope} = 2x + \Delta x$$

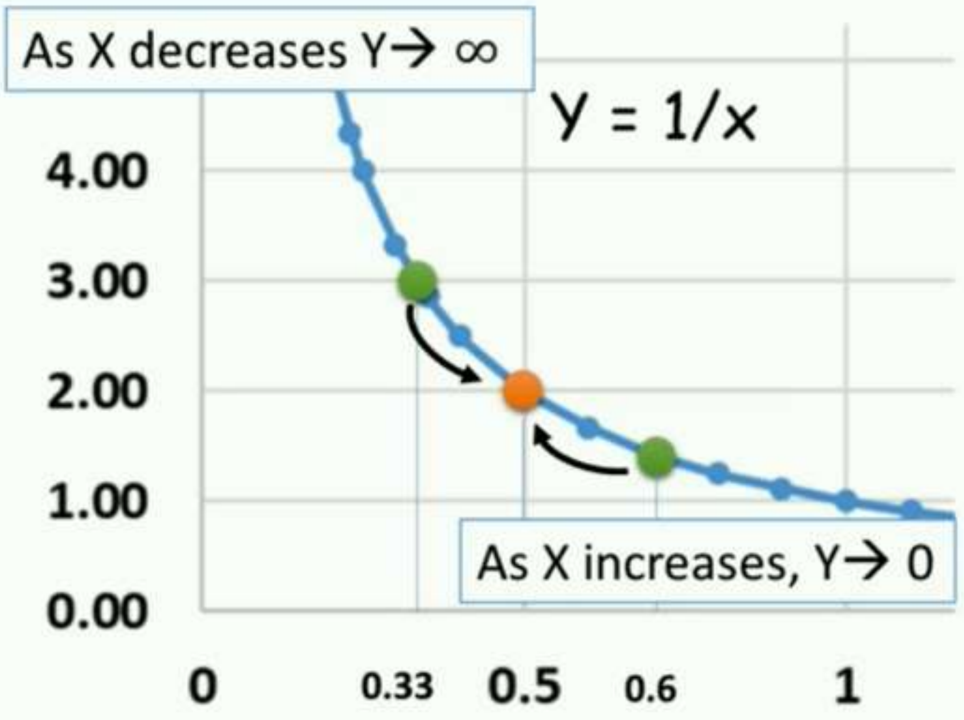
Derivative



Slope $= 2x + \Delta x$

$\Delta x \rightarrow 0$

Limits



$$\lim_{x \to 0.5} \frac{1}{x} = 2$$

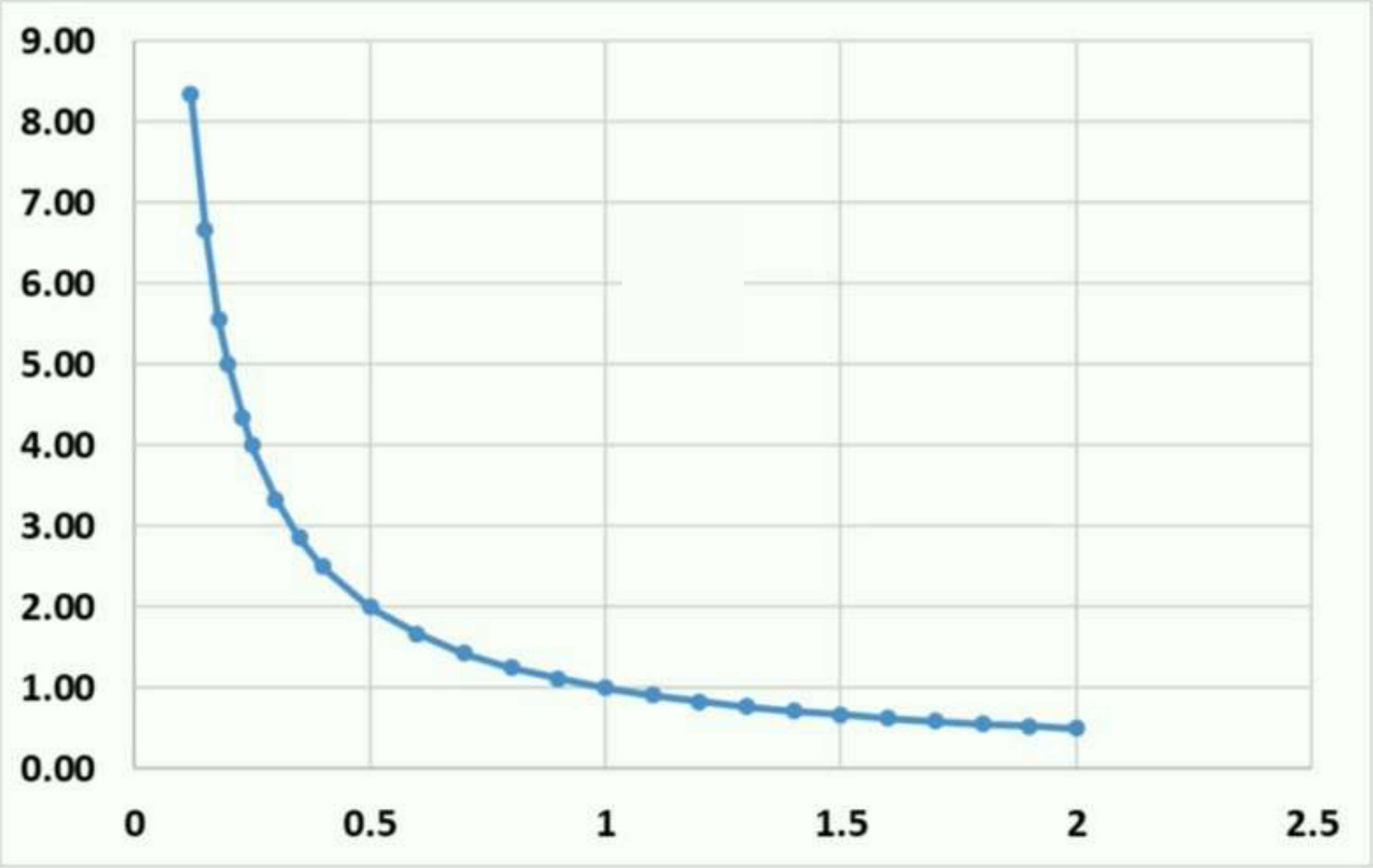
$$\lim_{x \to \infty} \frac{1}{x} = 0$$

$$\lim_{x \to 0} \frac{1}{x} = \infty$$

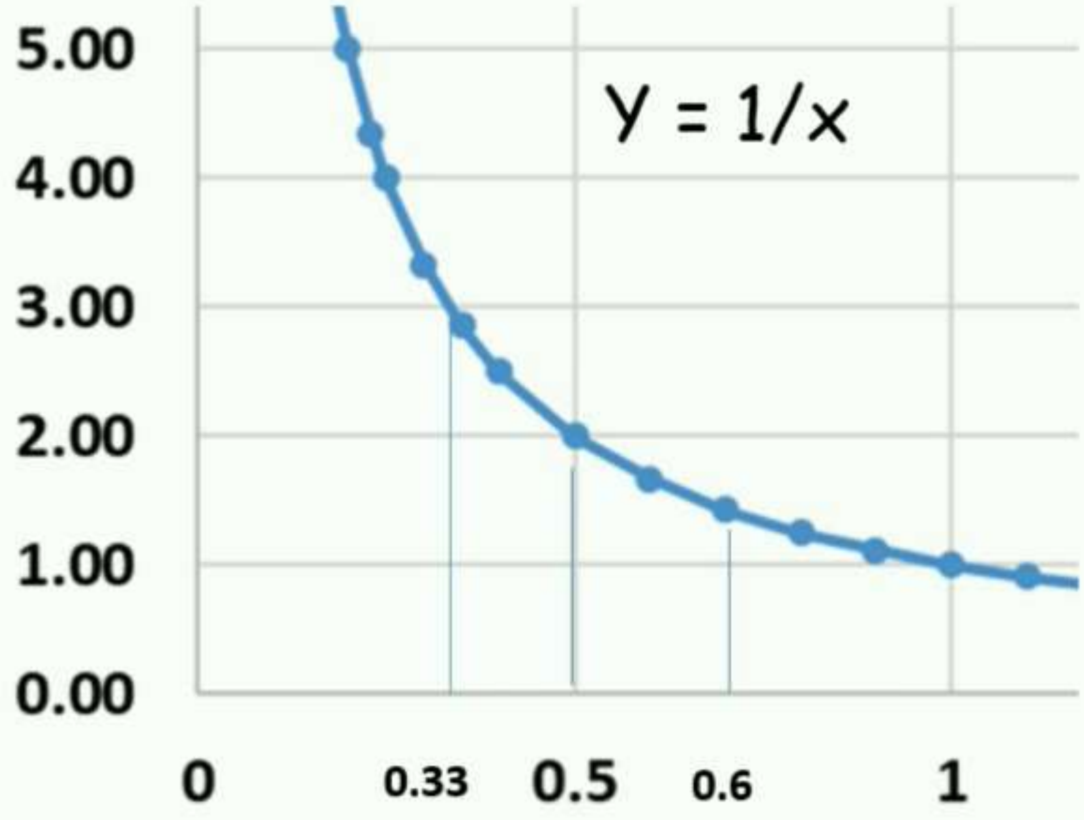
Limits

$$y = 1/x$$

$$x \neq 0$$

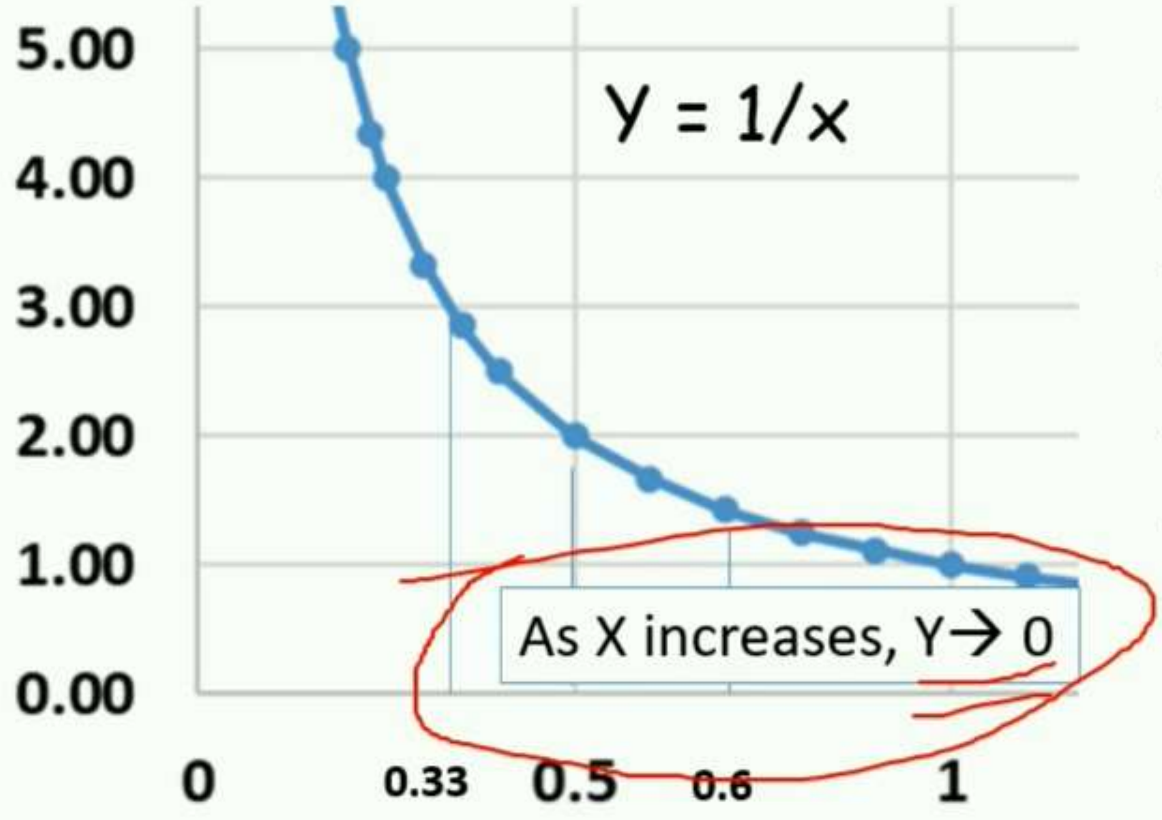


Limits



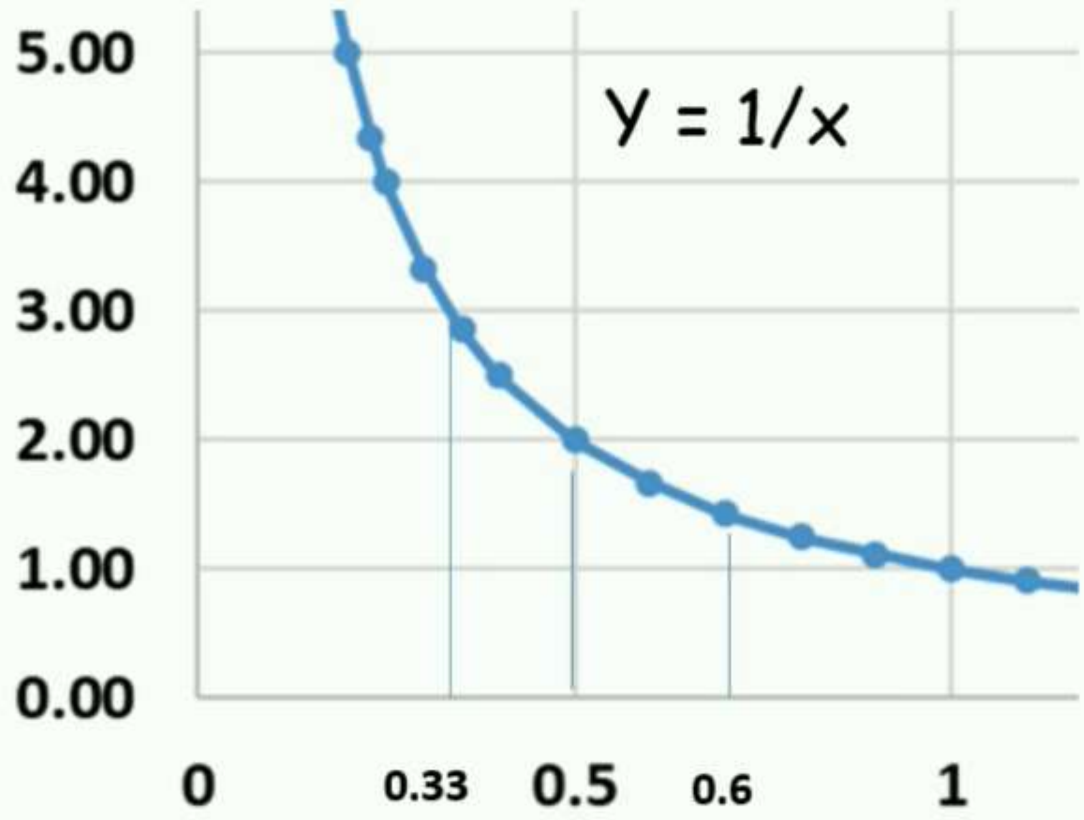
X	y
1	1
10	0.1
100	0.01
1000	0.001
10,000	0.0001
1,000,000	0.000001

Limits



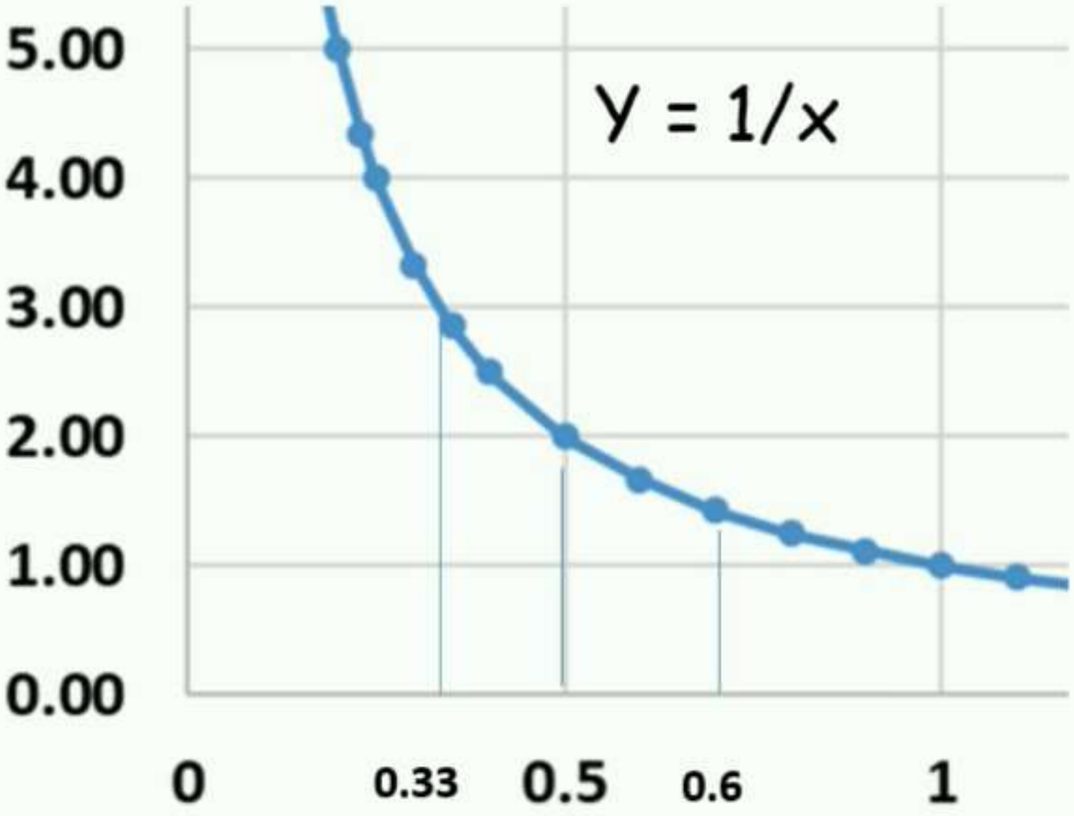
X	Y
1	1
10	0.1
100	0.01
1000	0.001
10,000	0.0001
1,000,000	0.000001

Limits



X	y
1	1
0.5	2
0.01	100

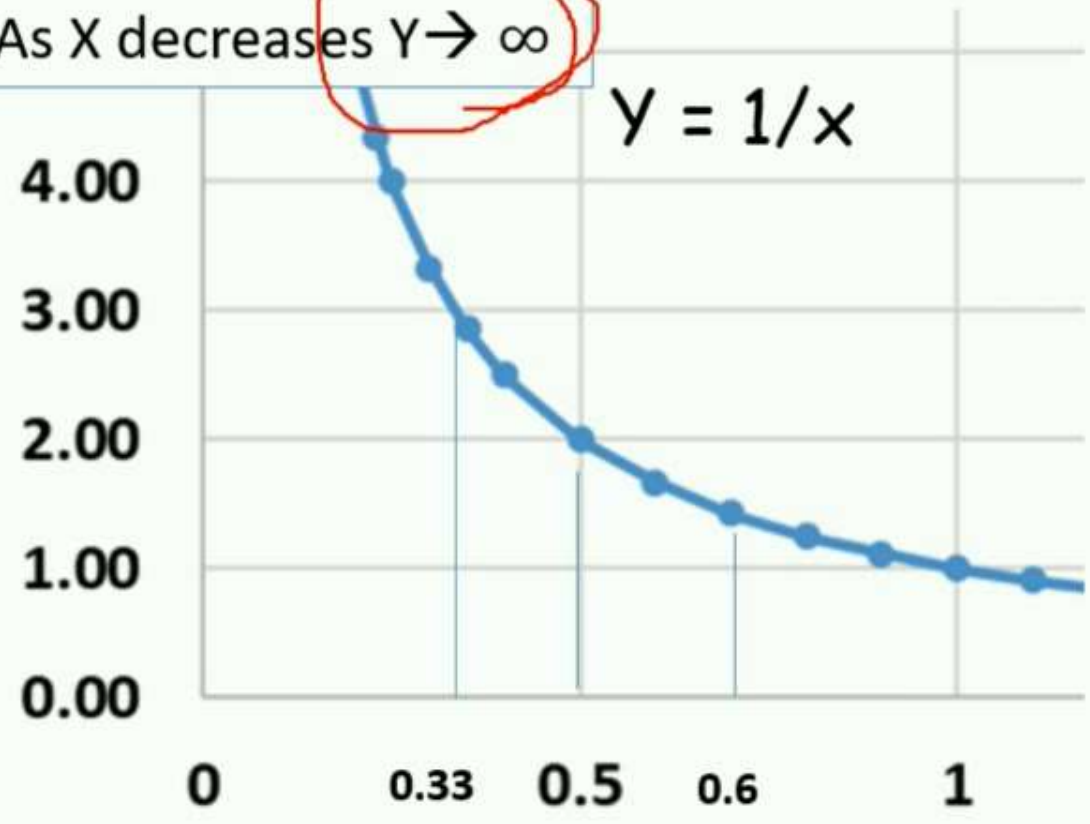
Limits



X	Y
1	1
0.5	2
0.01	100
0.001	1000
0.0001	10,000
0.000001	1,000,000

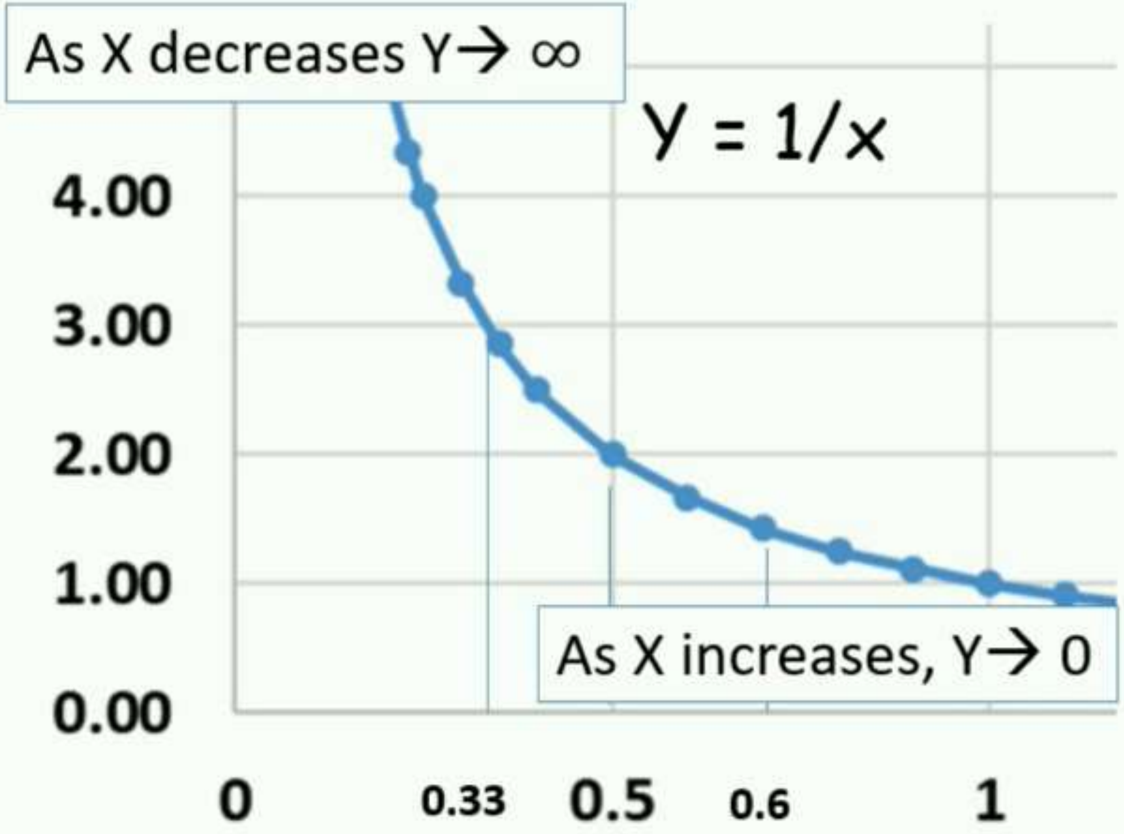
Limits

As X decreases Y→ ∞



X	Y
1	1
0.5	2
0.01	100
0.001	1000
0.0001	10,000
0.000001	1,000,000

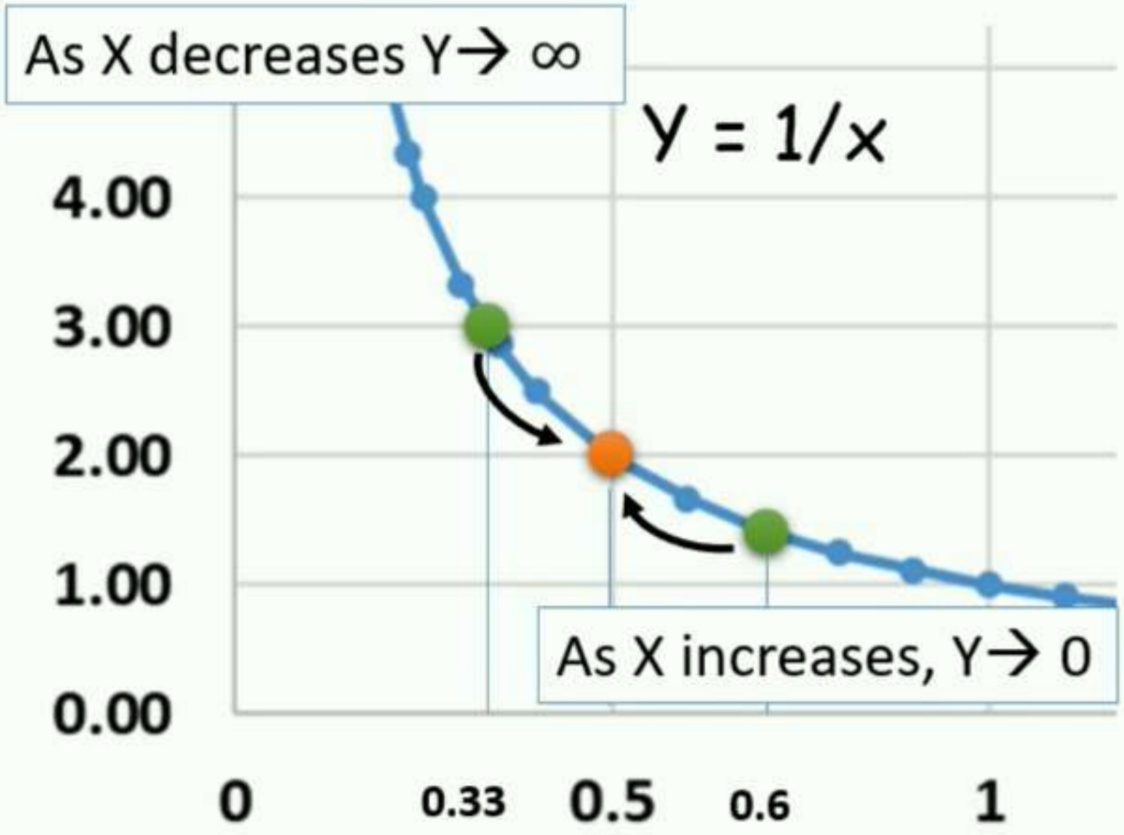
Limits



X	y
1	1
10	0.1
100	0.01
1000	0.001
10,000	0.0001
1,000,000	0.000001

X	y
1	1
0.5	2
0.01	100
0.001	1000
0.0001	10,000
0.000001	1,000,000

Limits

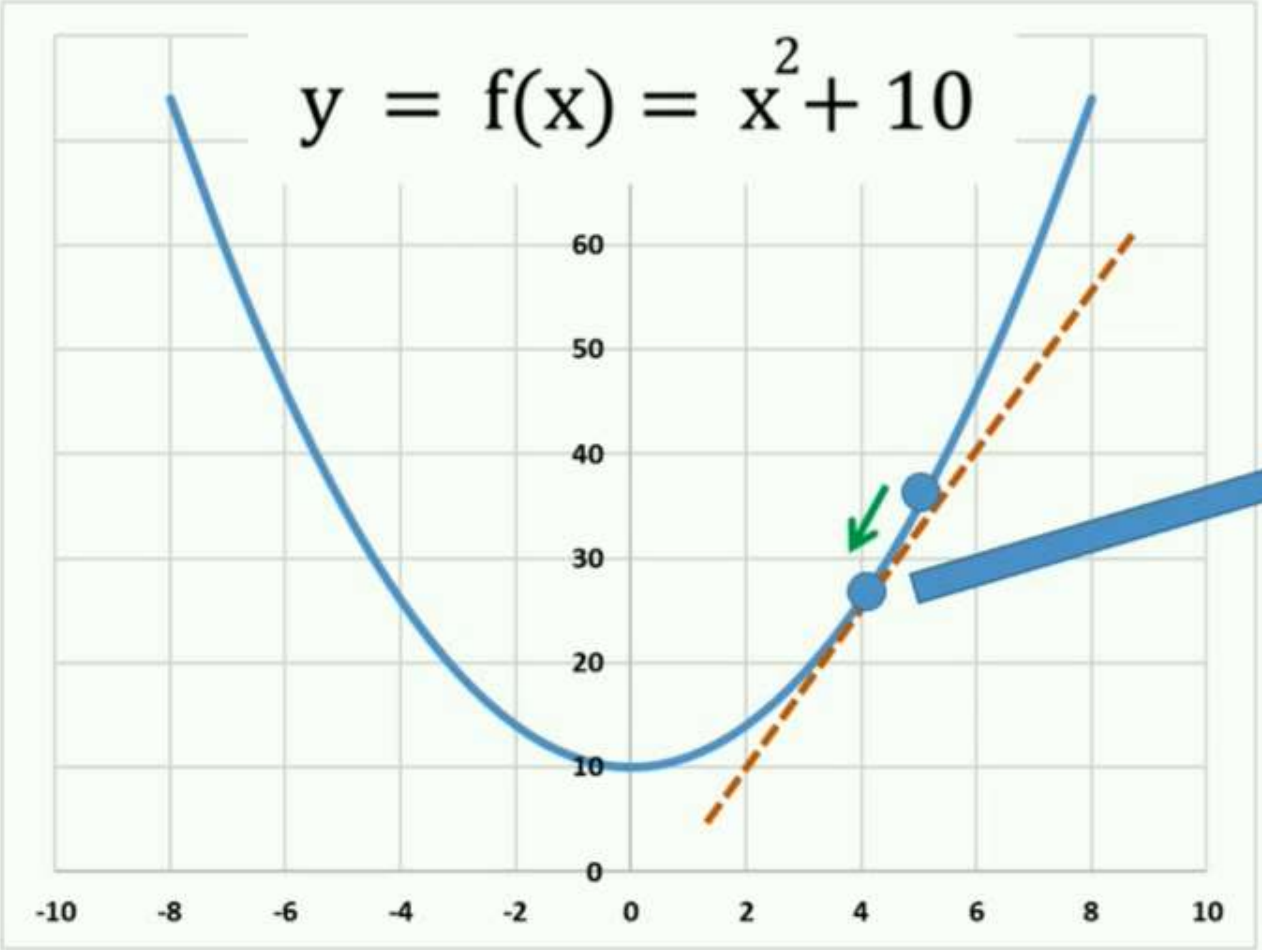


$$\lim_{x \to 0.5} \frac{1}{x} = 2$$

$$\lim_{x \to \infty} \frac{1}{x} = 0$$

$$\lim_{x \to 0} \frac{1}{x} = \infty$$

Derivative

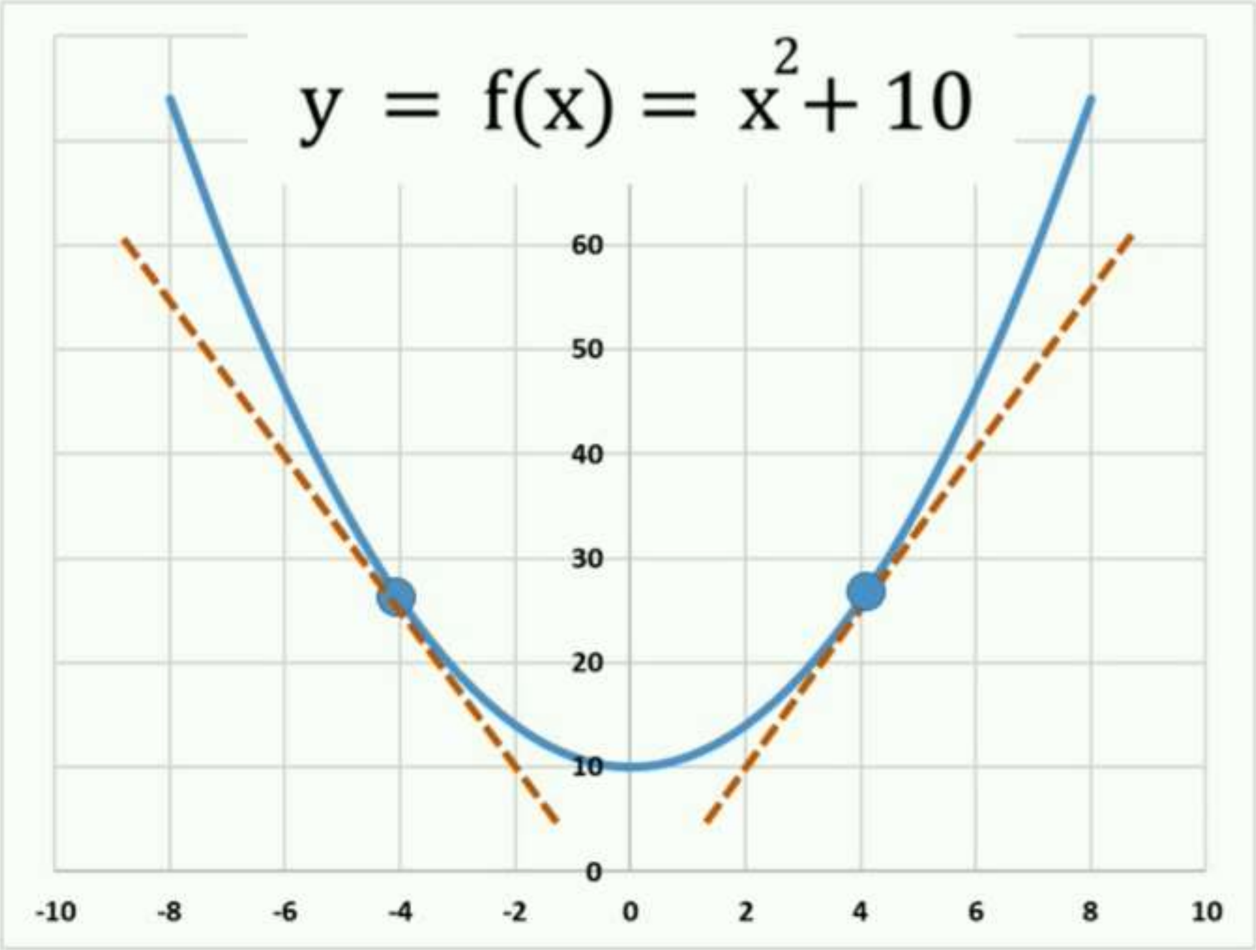


$$\text{Slope} = 2x + \Delta x$$

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} (2x + \underbrace{\Delta x}_{0})$$

$$\frac{dy}{dx} = 2x$$

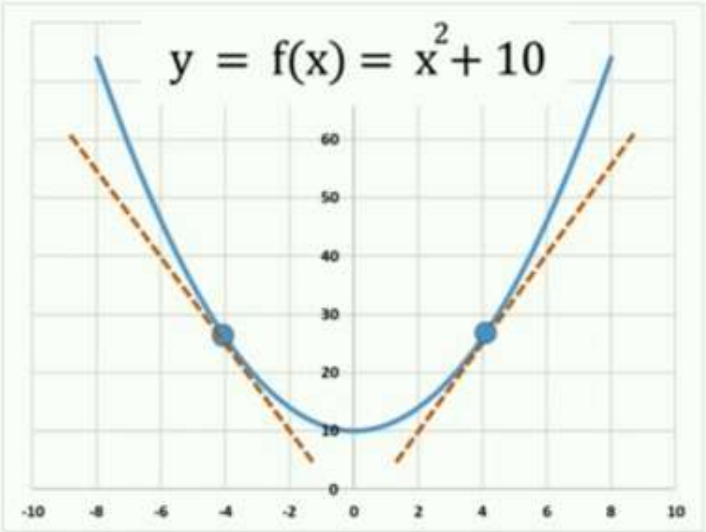
Derivative



$$\frac{dy}{dx} = 2x$$

1. $x = 4; \quad slope = 8$
2. $x = -4; \quad slope = -8$

Why to know the "Slope at the Point"?

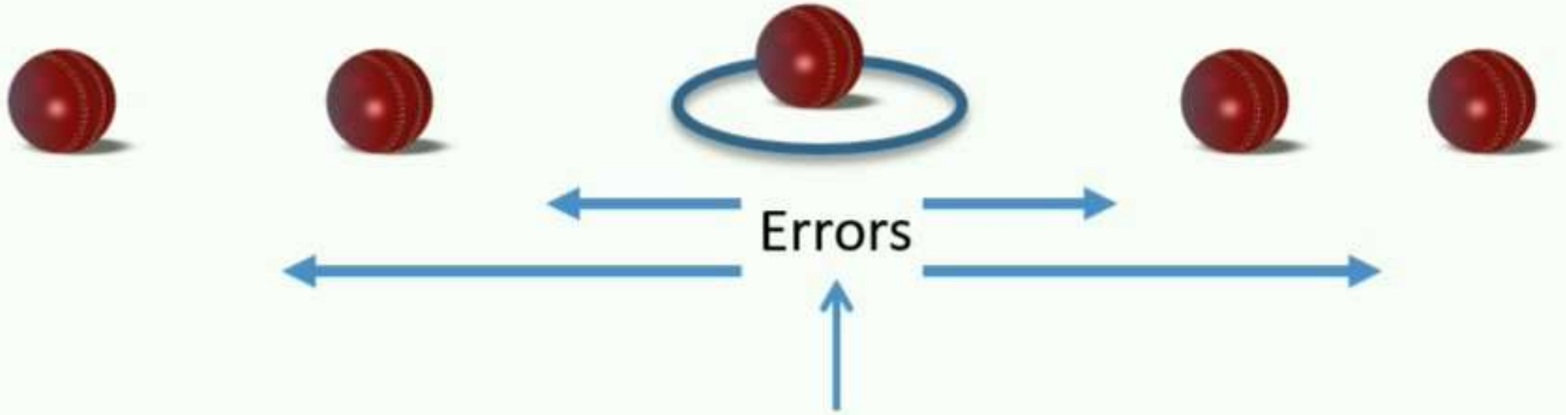


$$\frac{dy}{dx} = 2x$$

1. $x = 4$; slope $= 8$
2. $x = -4$; slope $= -8$

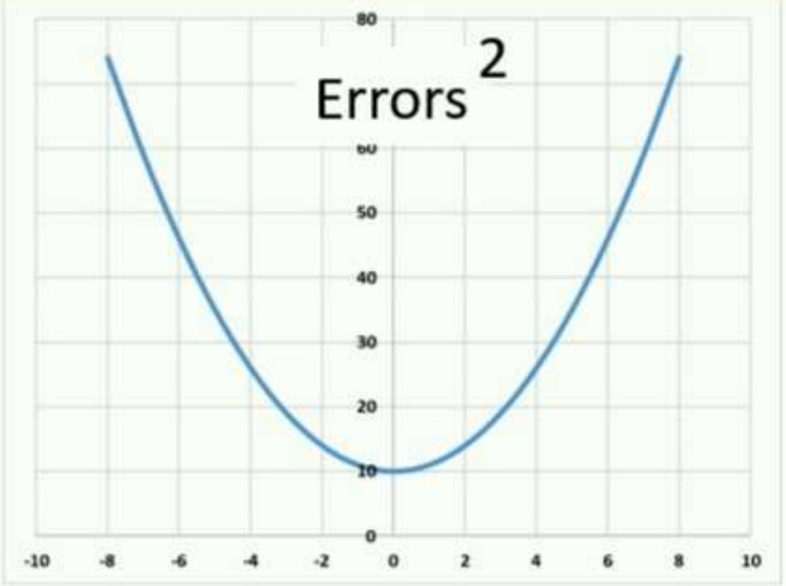
Why to know the “Slope at the Point”?

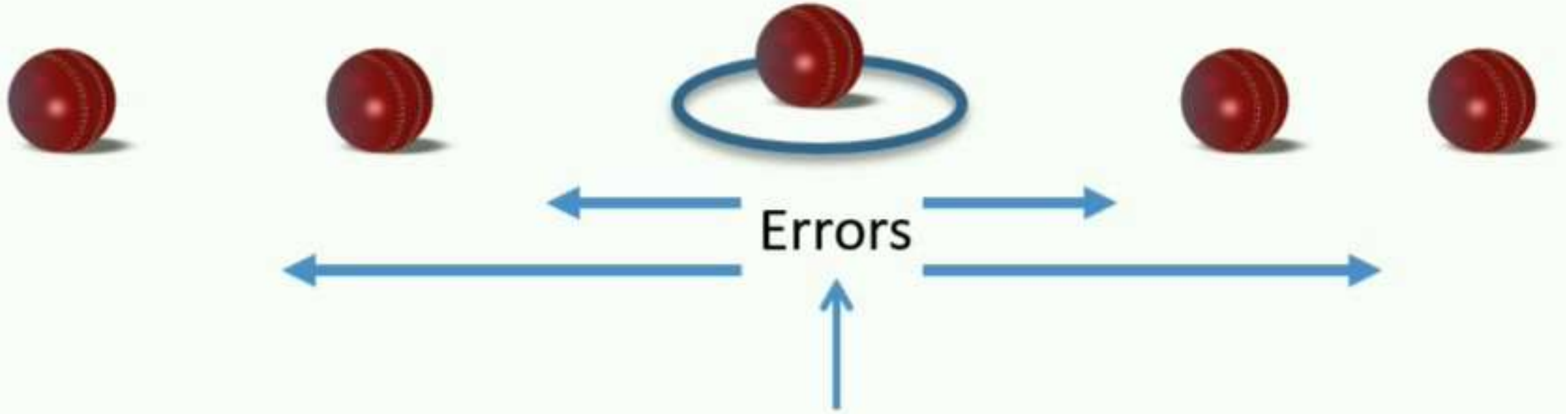
Throw



Why to know the "Slope at the Point"?

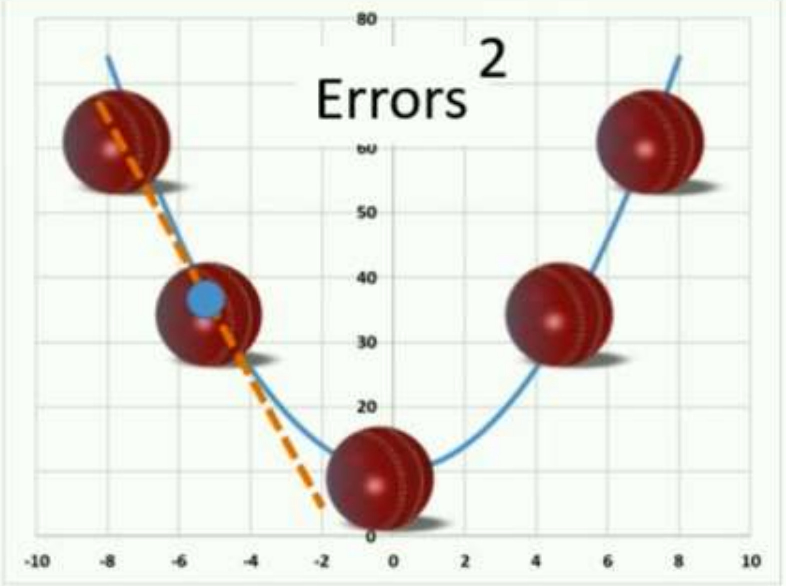
Throw

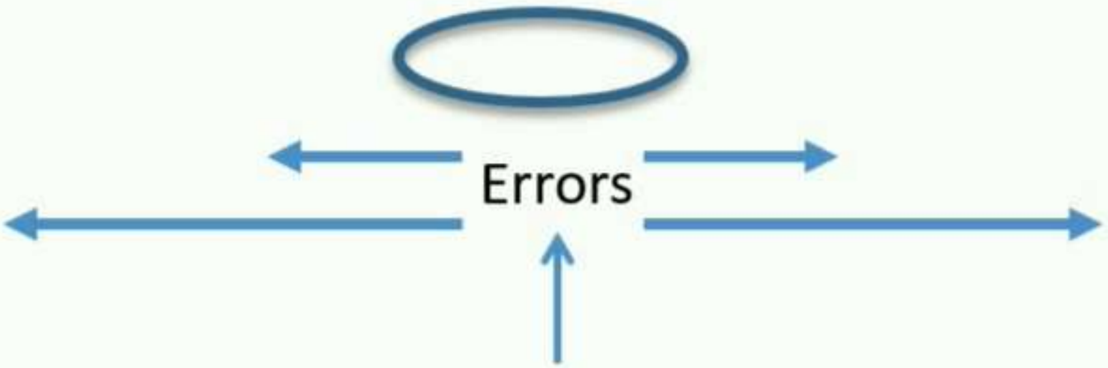




Why to know the “Slope at the Point”?

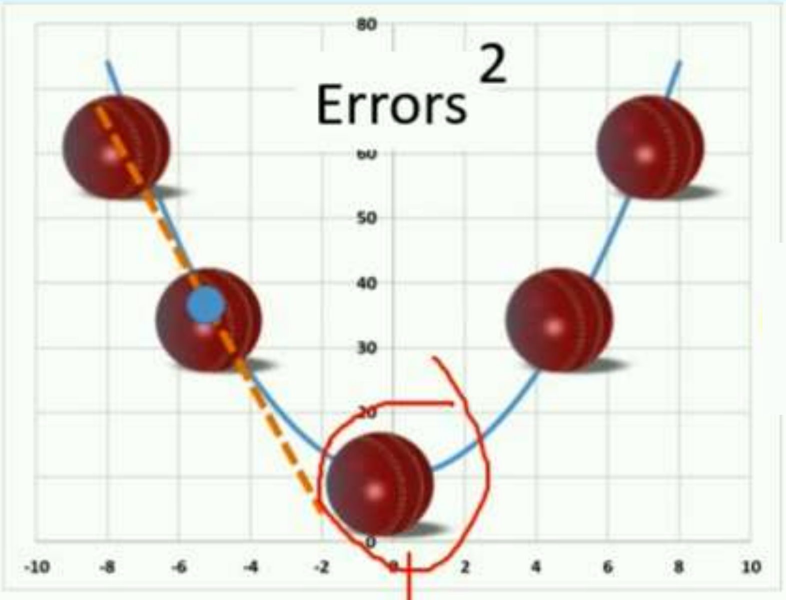
Throw



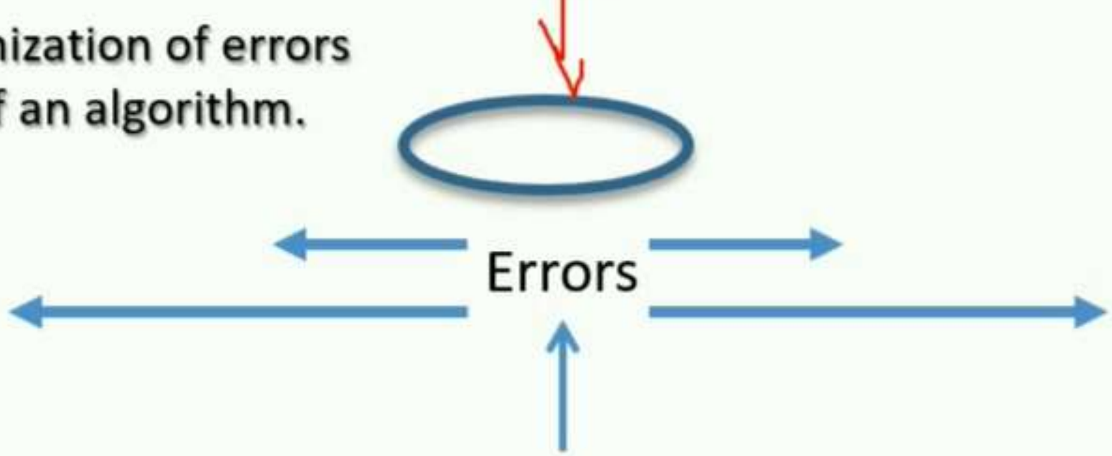


Why to know the “Slope at the Point”?

Throw



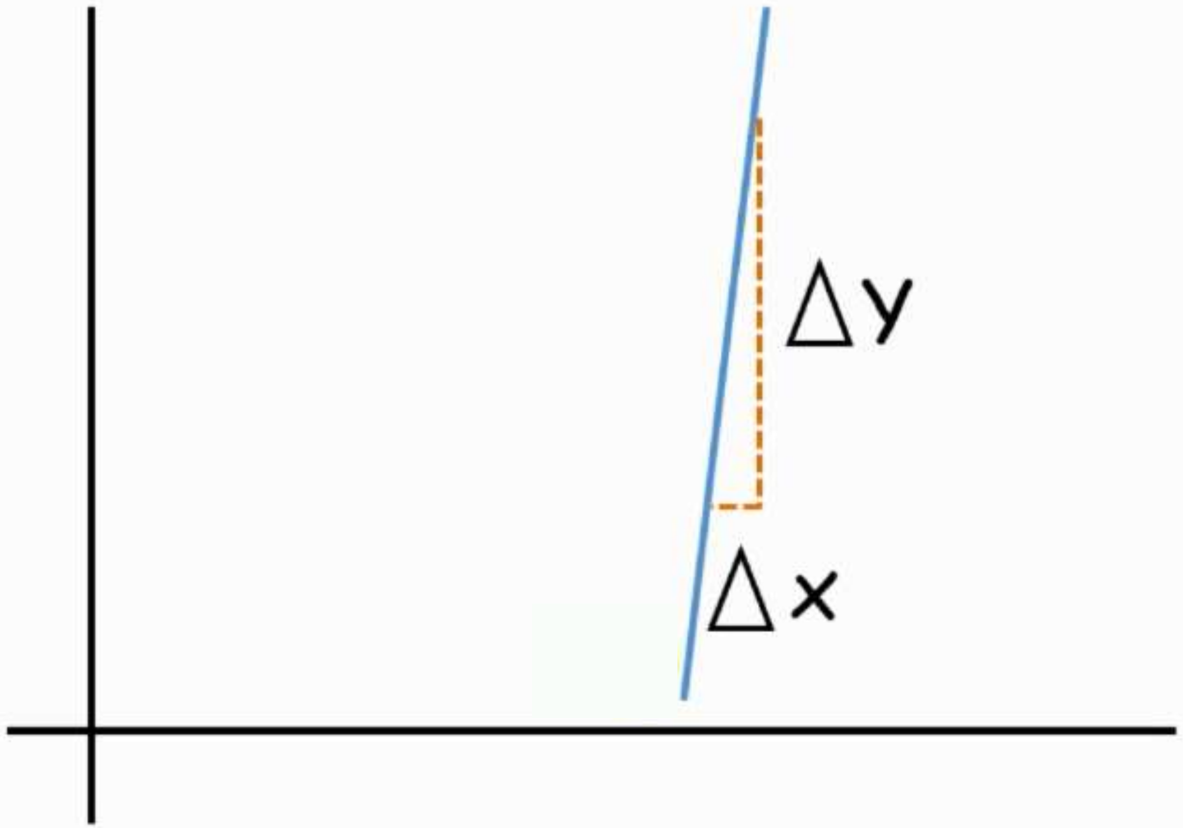
Optimization or Minimization of errors
to get best accuracy of an algorithm.



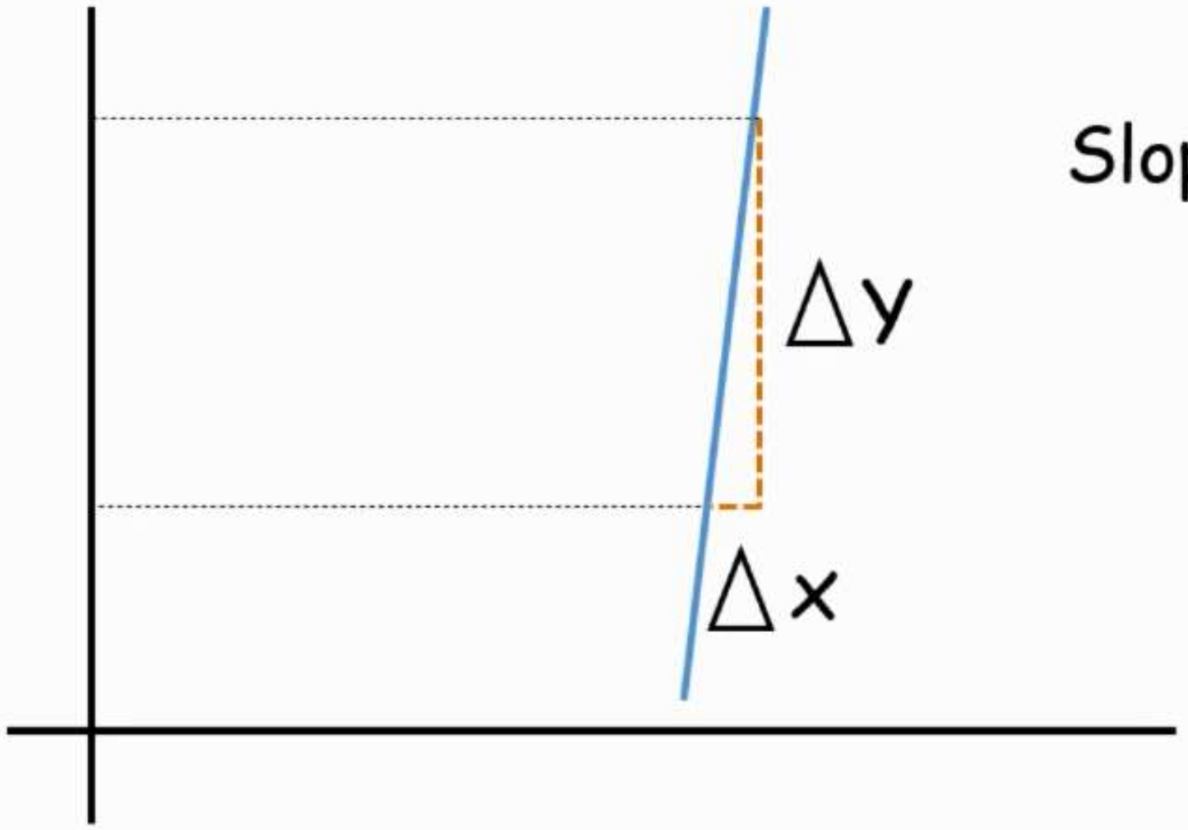
Derivative rules

- Derivative of a vertical line
- Derivative of a horizontal line
- Differentiability for various functions
- Power rule of derivative

Derivative Rules



Derivative Rules



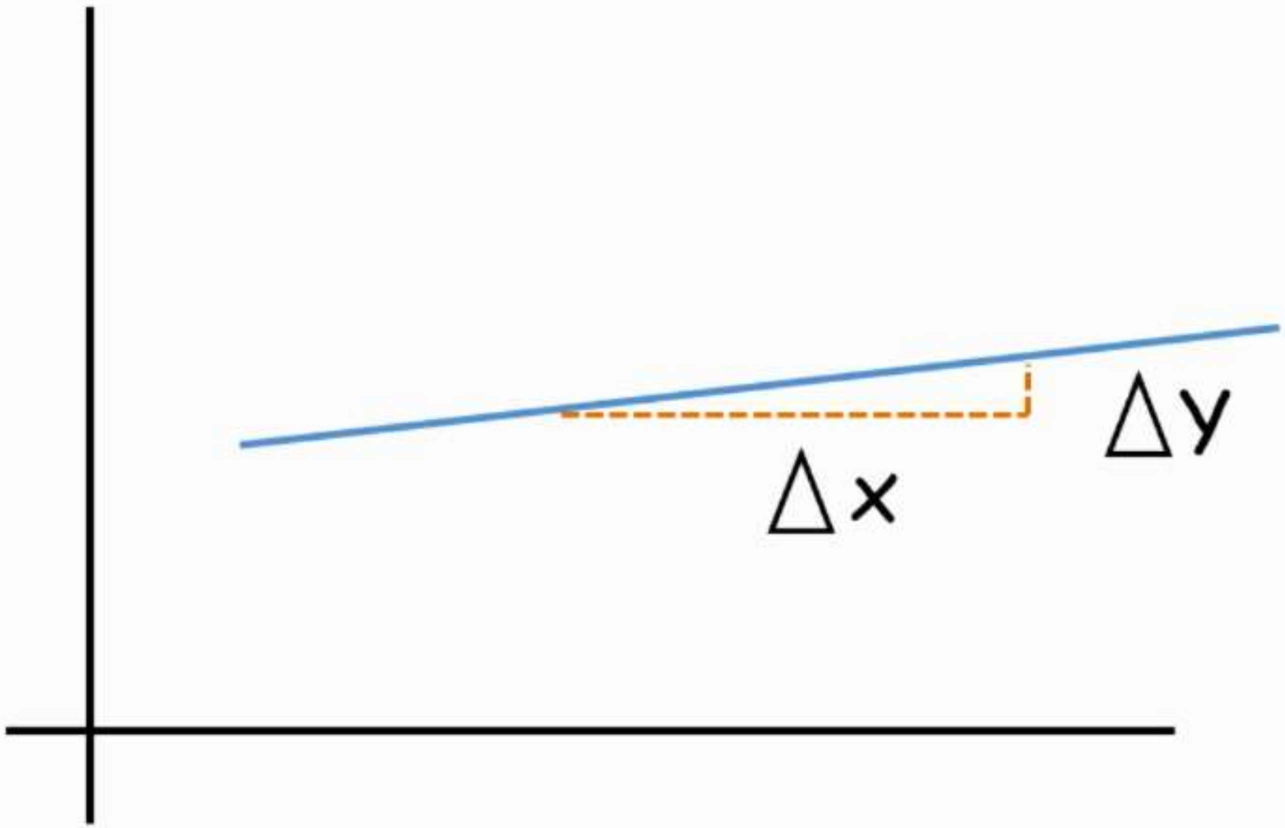
$$\text{Slope} = \frac{\Delta y}{\Delta x} \uparrow$$

Derivative Rules

$$\Delta x = 0$$

$$\text{Slope} = \frac{\Delta y}{\Delta x} = \infty \text{ or Undefined}$$

Derivative Rules

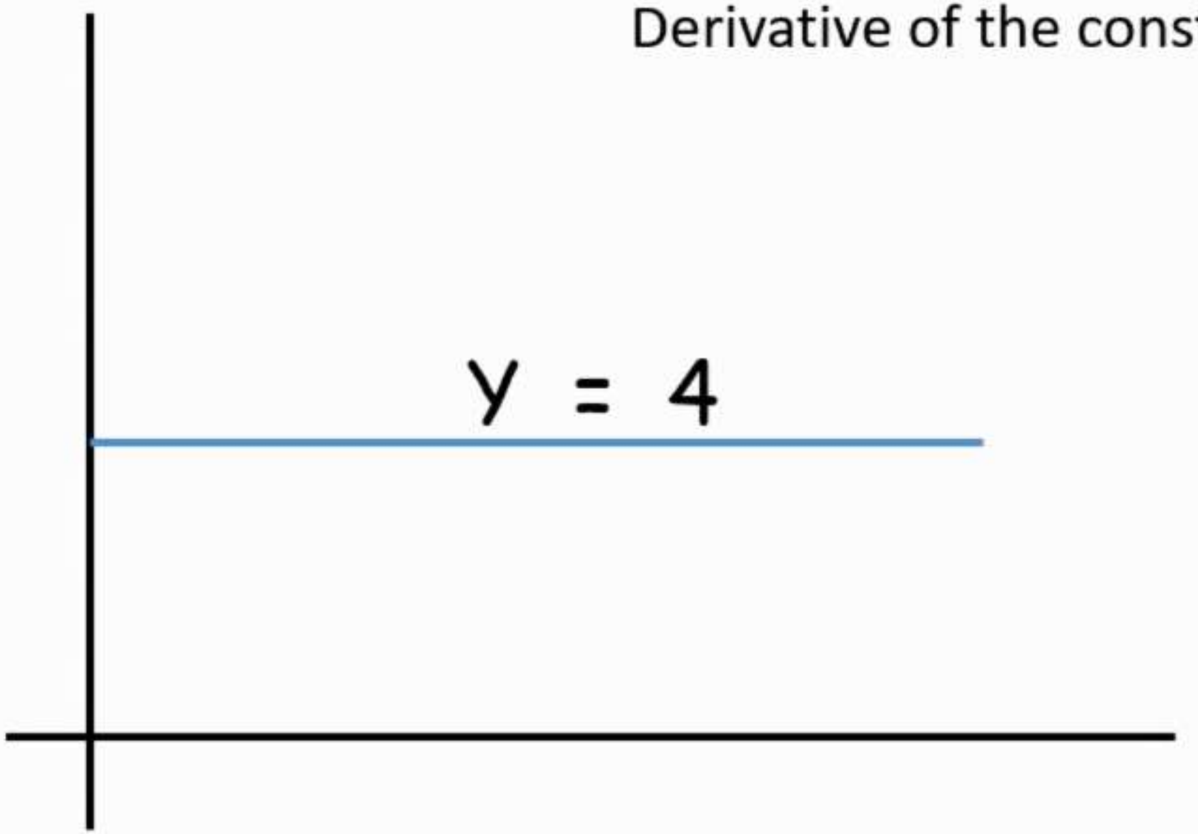


$$\text{Slope} = \frac{\Delta y}{\Delta x} \downarrow$$

$$\text{Slope} = \frac{\Delta y}{\Delta x} = 0$$

Derivative Rules – Constant

Derivative of the constant is ZERO.



$$\Delta y = 0$$

$$\frac{dy}{dx} = 0$$

$$\frac{d\,(4)}{dx} = 0$$

Power Rule of Derivative

$$y = f(x) = x^2 + 10 \Rightarrow \frac{dy}{dx} = 2x$$

$$y = f(x) = x^3 + 10 \Rightarrow \frac{dy}{dx} = 3x^2$$

Power Rule of Derivative

$$y = f(x) = x^2 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = \boxed{}$$

$$y = f(x) = x^3 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = \boxed{}$$

Power Rule of Derivative

$$y = f(x) = x^2 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = 2x$$

$$y = f(x) = x^3 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = 3x^2$$

Power Rule of Derivative

$$y = f(x) = ax^{n} \Rightarrow \frac{dy}{dx} = a*n\ x^{n-1}$$

Power Rule of Derivative

$$y = f(x) = x^2 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = 2x$$

(Original Index – 1)

$$y = f(x) = x^3 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = 3x^2$$

Remove constant

(Original Index * Original Coefficient)

Power Rule of Derivative

$$y = f(x) = x^3 + 10 \Rightarrow \frac{dy}{dx} = 3x^2$$

(Original Index – 1)

Remove constant

(Original Index * Original Coefficient)

$$y = f(x) = 2x^3 + 4x^2 - \underline{7x} + 9 \Rightarrow \frac{dy}{dx} = \boxed{}$$

Power Rule of Derivative

(Original Index – 1)

$$y = f(x) = x^3 + 10 \Rightarrow \frac{dy}{dx} = 3x^2$$

Remove constant

(Original Index * Original Coefficient)

$$y = f(x) = 2x^3 + 4x^2 - 7x + 9 \Rightarrow \frac{dy}{dx} = 6x^2 + 8x^1 - 7x^0$$

Second Order Derivative

$$\frac{d\left(\frac{dy}{dx}\right)}{dx} = \frac{d^2y}{dx^2}$$

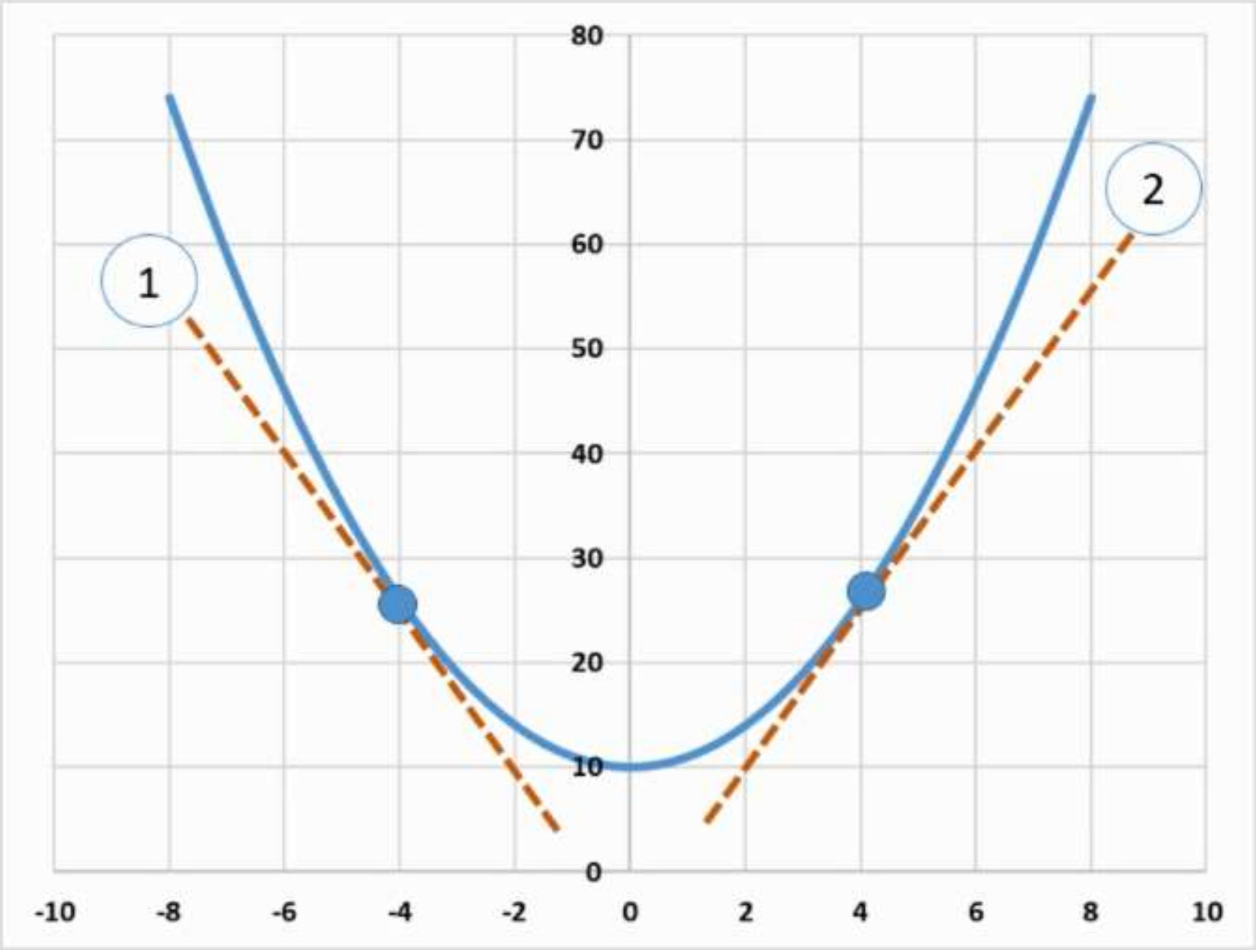
Derivative for directions

$$y = f(x) = x^2 + 10$$

$$\frac{dy}{dx} = 2x$$

1 $\frac{dy}{dx} = -8$

2 $\frac{dy}{dx} = +8$



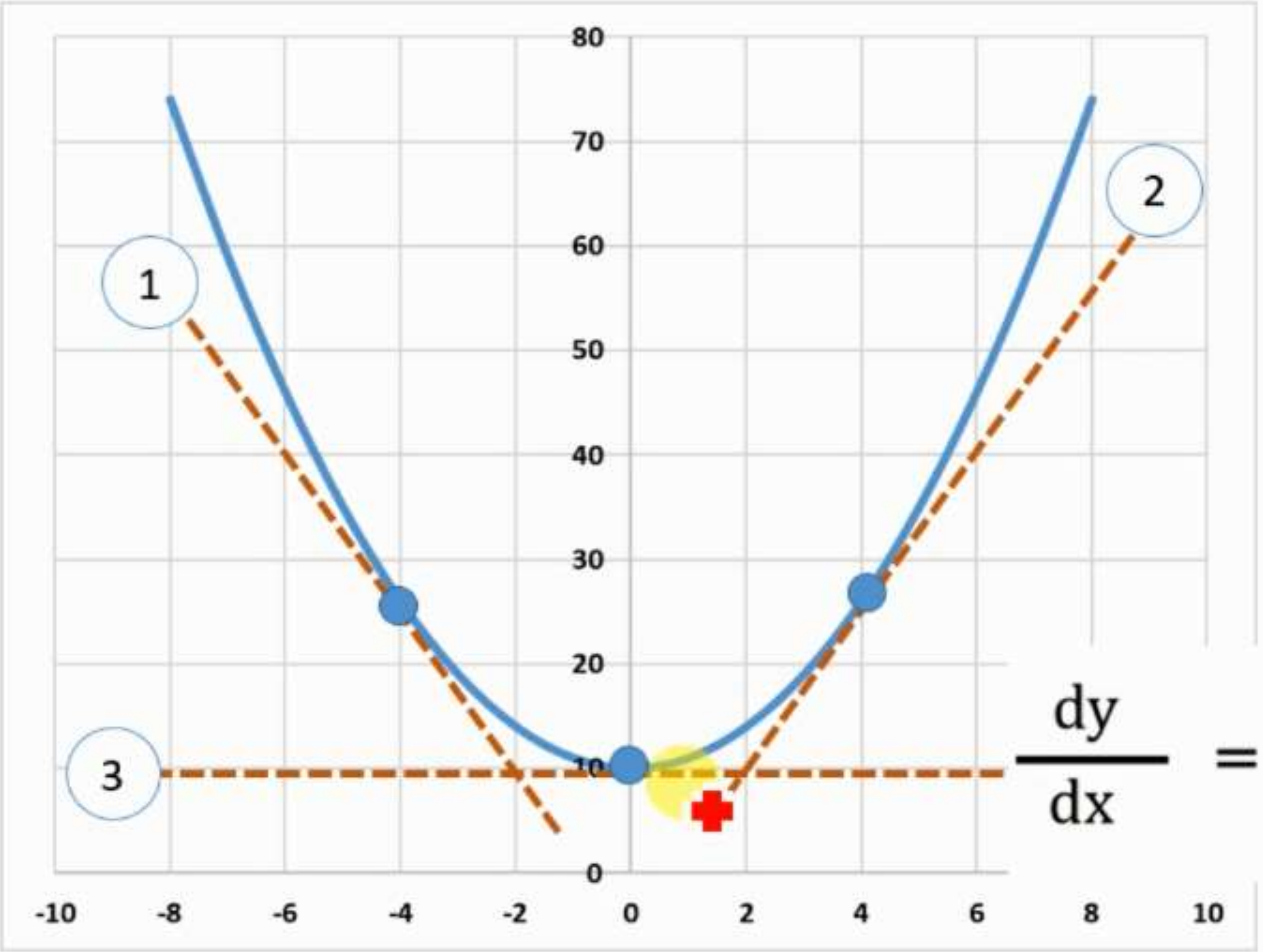
Derivative for directions

$$y = f(x) = x^2 + 10$$

$$\frac{dy}{dx} = 2x$$

1 $\dfrac{dy}{dx} = -8$

2 $\dfrac{dy}{dx} = +8$



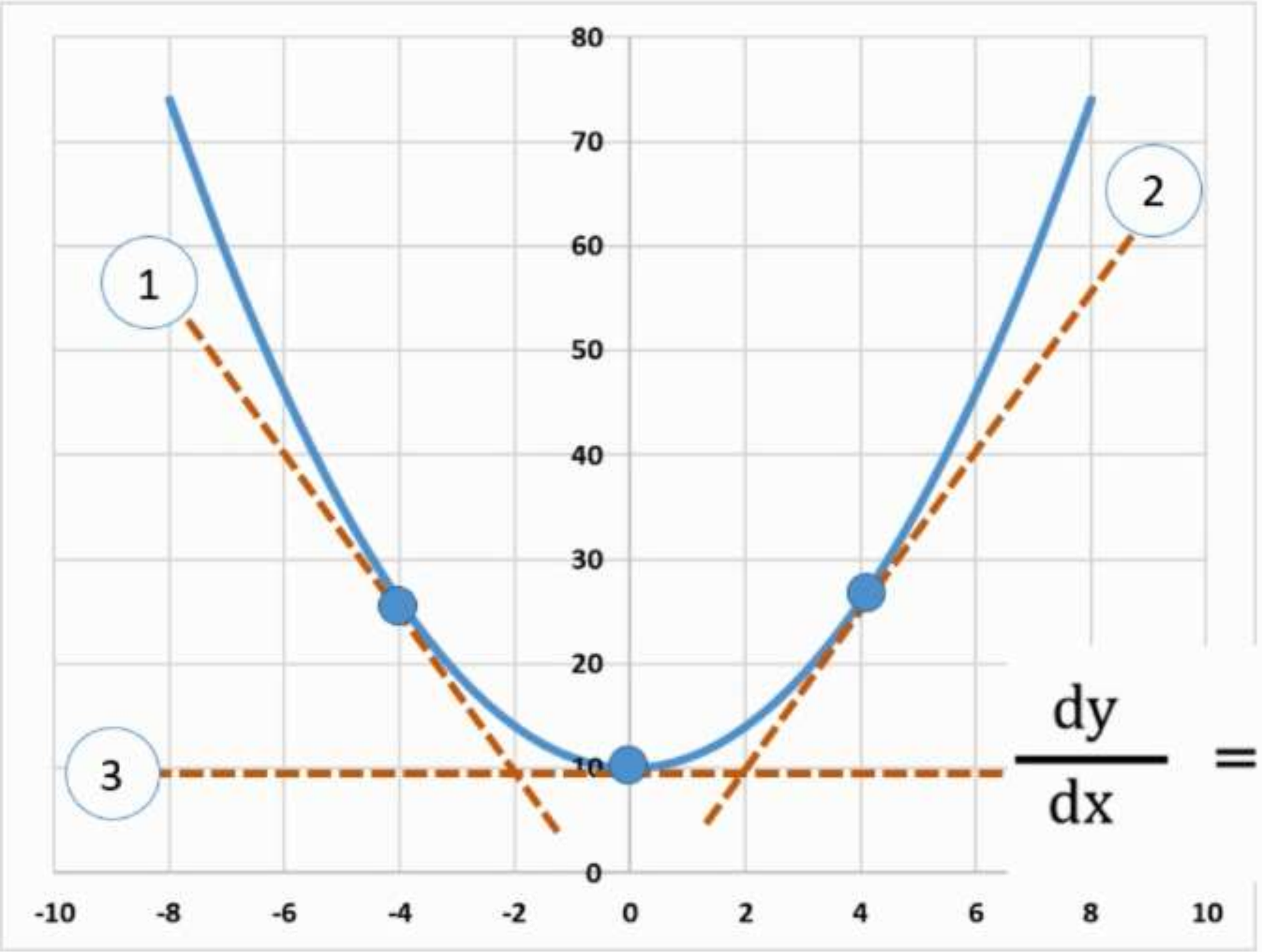
$$\frac{dy}{dx} = 0$$

Derivative Direction

$$y = f(x) = x^2 + 10$$

$$y = f(x) = 0 + 10$$

$$y = f(x) = 10$$



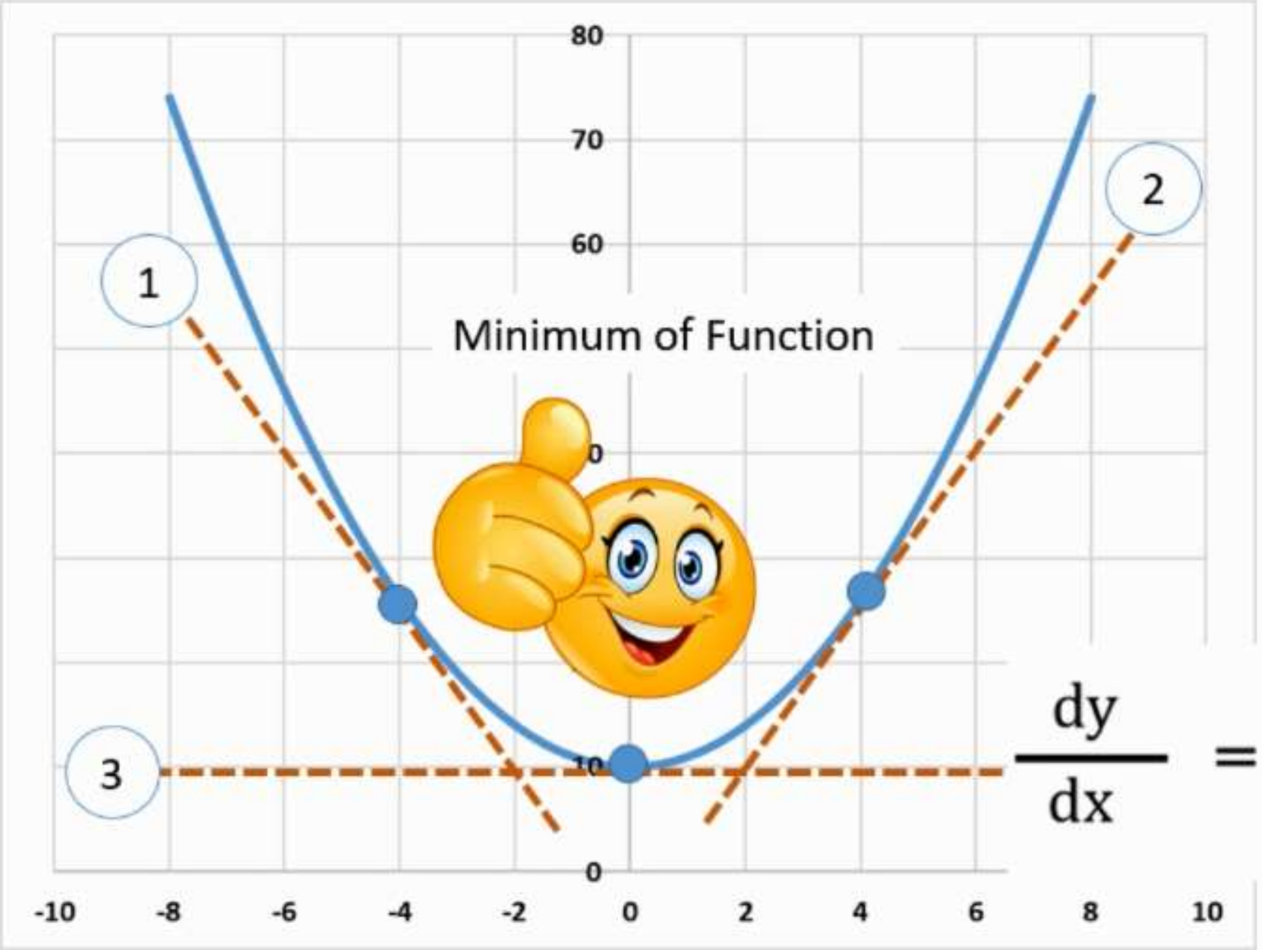
$$\frac{dy}{dx} = 0$$

Derivative Direction

$$y = f(x) = x^2 + 10$$

$$y = f(x) = 0 + 10$$

$$y = f(x) = 10$$



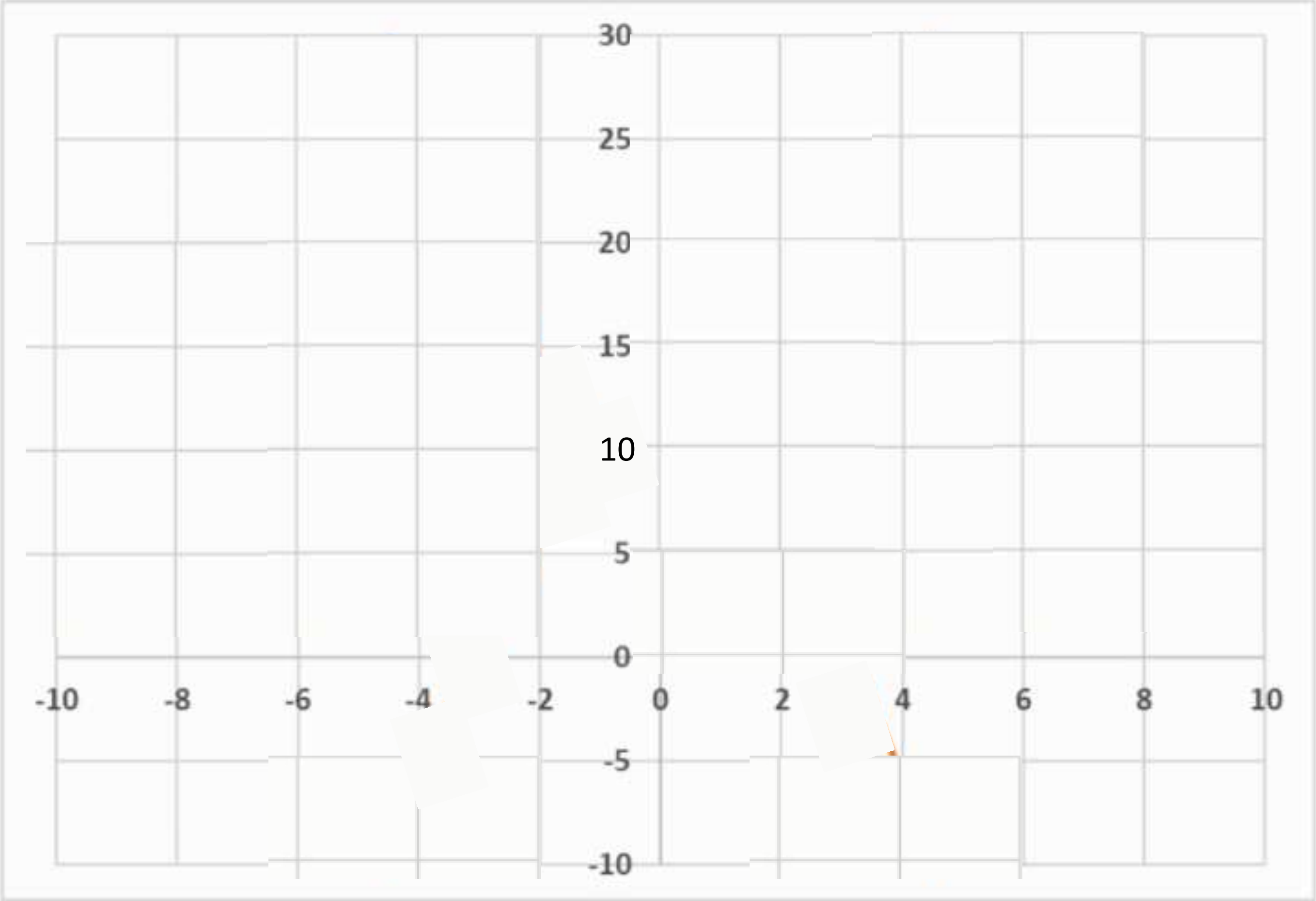
$$\frac{dy}{dx} = 0$$

Maxima or Minima?

$$y = f(x) = x^2 + 10$$

?

$$y = f(x) = -x^2 + 10$$

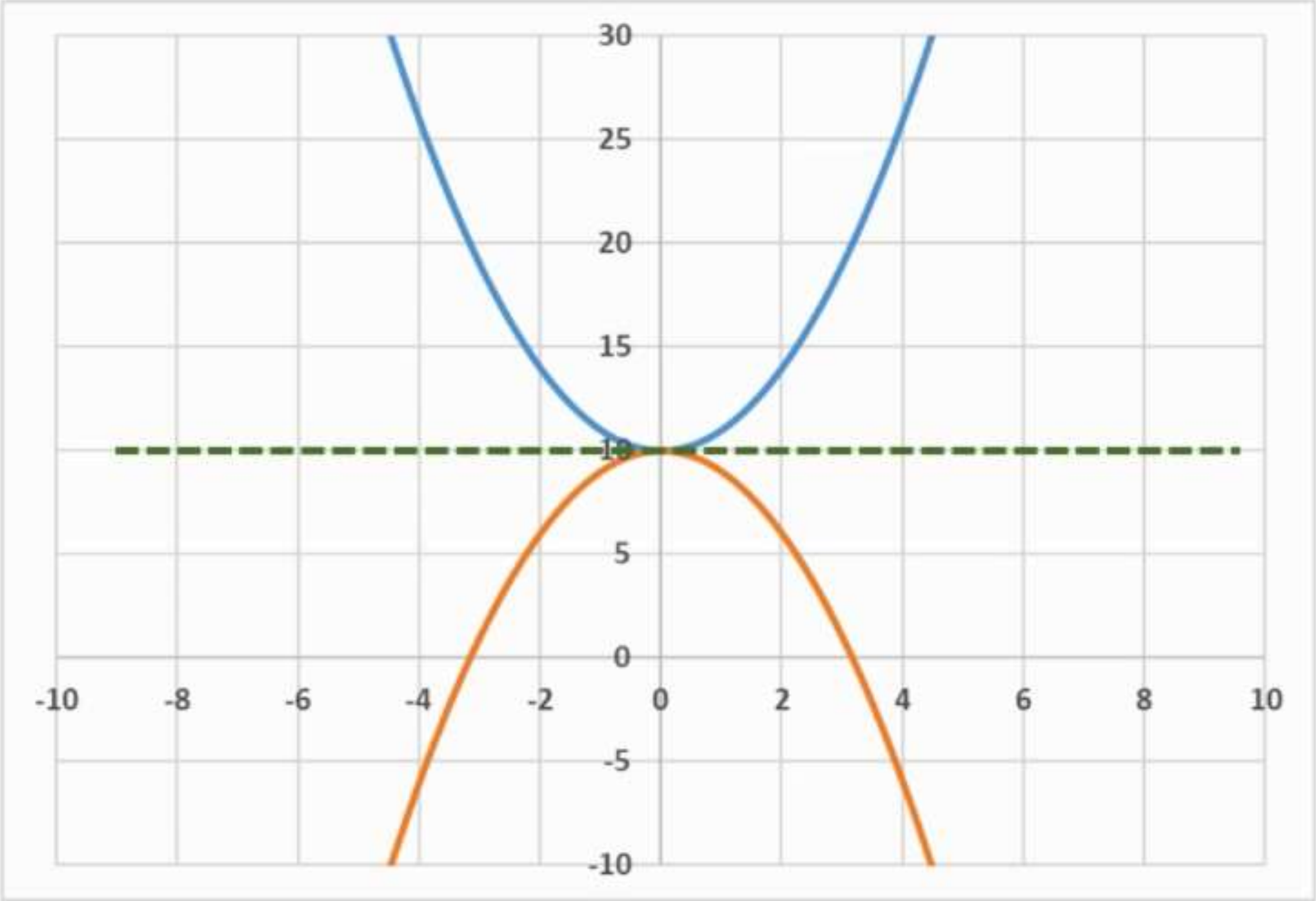


Maxima or Minima?

$$y = f(x) = x^2 + 10$$

?

$$y = f(x) = -x^2 + 10$$



Second Order Derivative

$$\frac{d\left(\frac{dy}{dx}\right)}{dx} = \frac{d^2y}{dx^2}$$

Second Order Derivative

$$y = f(x) = x^2 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = \quad \Rightarrow \quad \frac{d^2y}{dx^2} =$$

$$y = f(x) = -x^2 + 10 \quad \Rightarrow \quad \frac{dy}{dx} = \quad \Rightarrow \quad \frac{d^2y}{dx^2} =$$

Second Order Derivative

$$y = f(x) = x^2 + 10 \Rightarrow \frac{dy}{dx} = 2x \Rightarrow \frac{d^2y}{dx^2} = 2$$

$$y = f(x) = -x^2 + 10 \Rightarrow \frac{dy}{dx} = -2x \Rightarrow \frac{d^2y}{dx^2} = -2$$

Rules for Maxima and Minima

Second Derivative < 0 → Local Maxima

Second Derivative > 0 → Local Minima

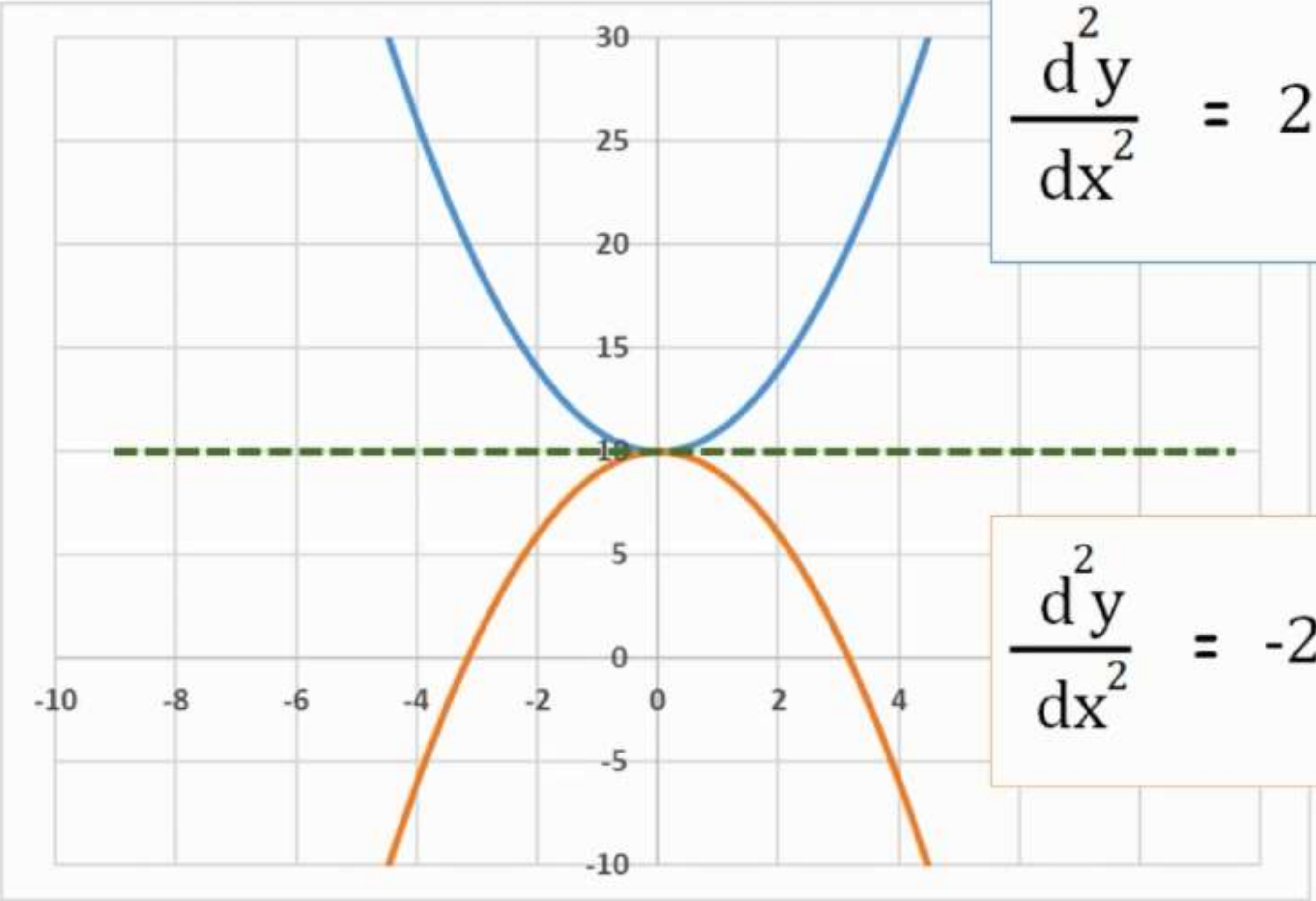
Maxima or Minima?

$$y = f(x) = x^2 + 10$$

Minima at y = 10

$$y = f(x) = -x^2 + 10$$

Maxima at y = 10



$$\frac{d^2y}{dx^2} = 2$$

$$\frac{d^2y}{dx^2} = -2$$

Maxima or Minima?

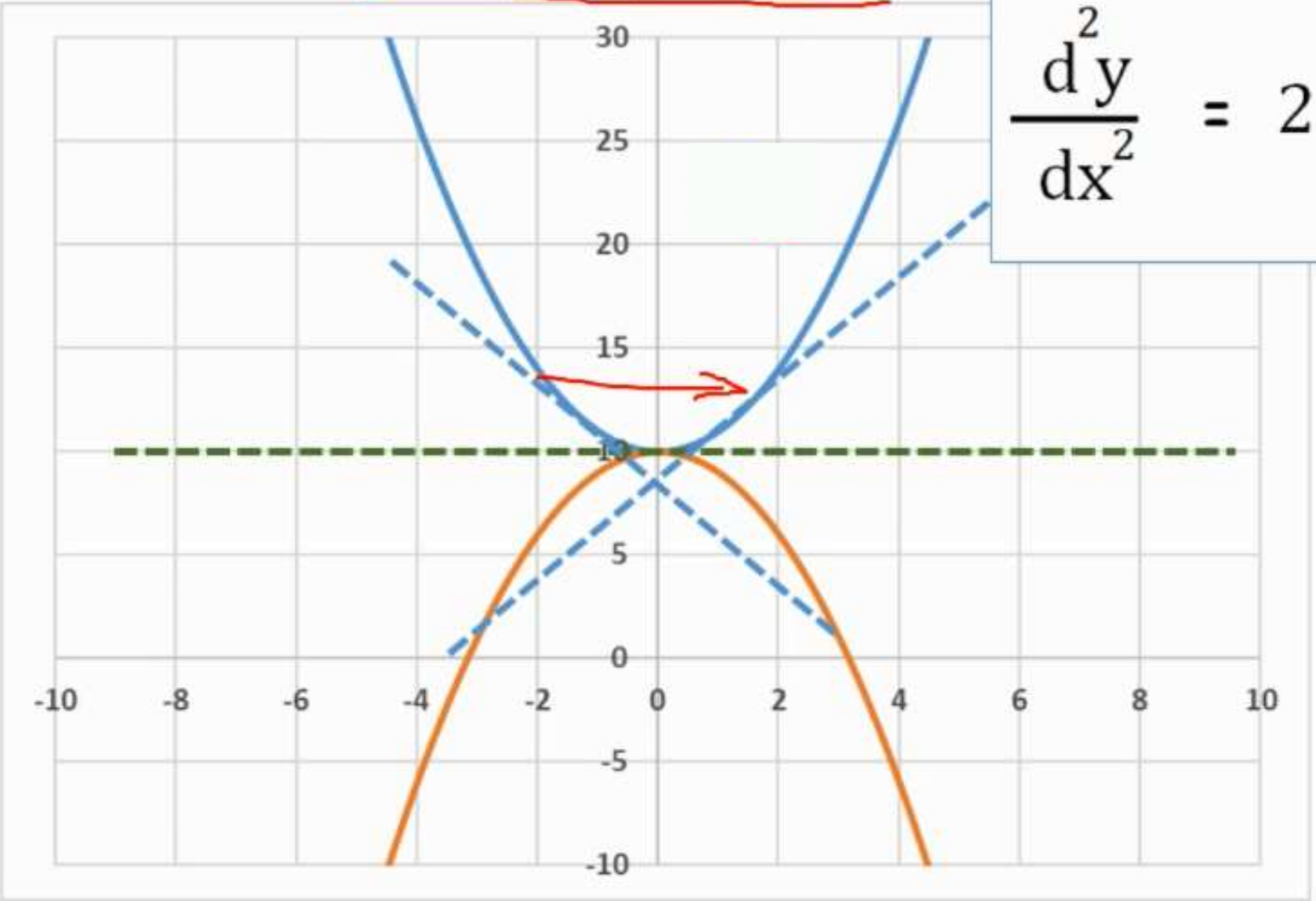
$$y = f(x) = x^2 + 10$$

Minima at y = 10

$$y = f(x) = -x^2 + 10$$

Maxima at y = 10

Shift from Negative slope to Positive

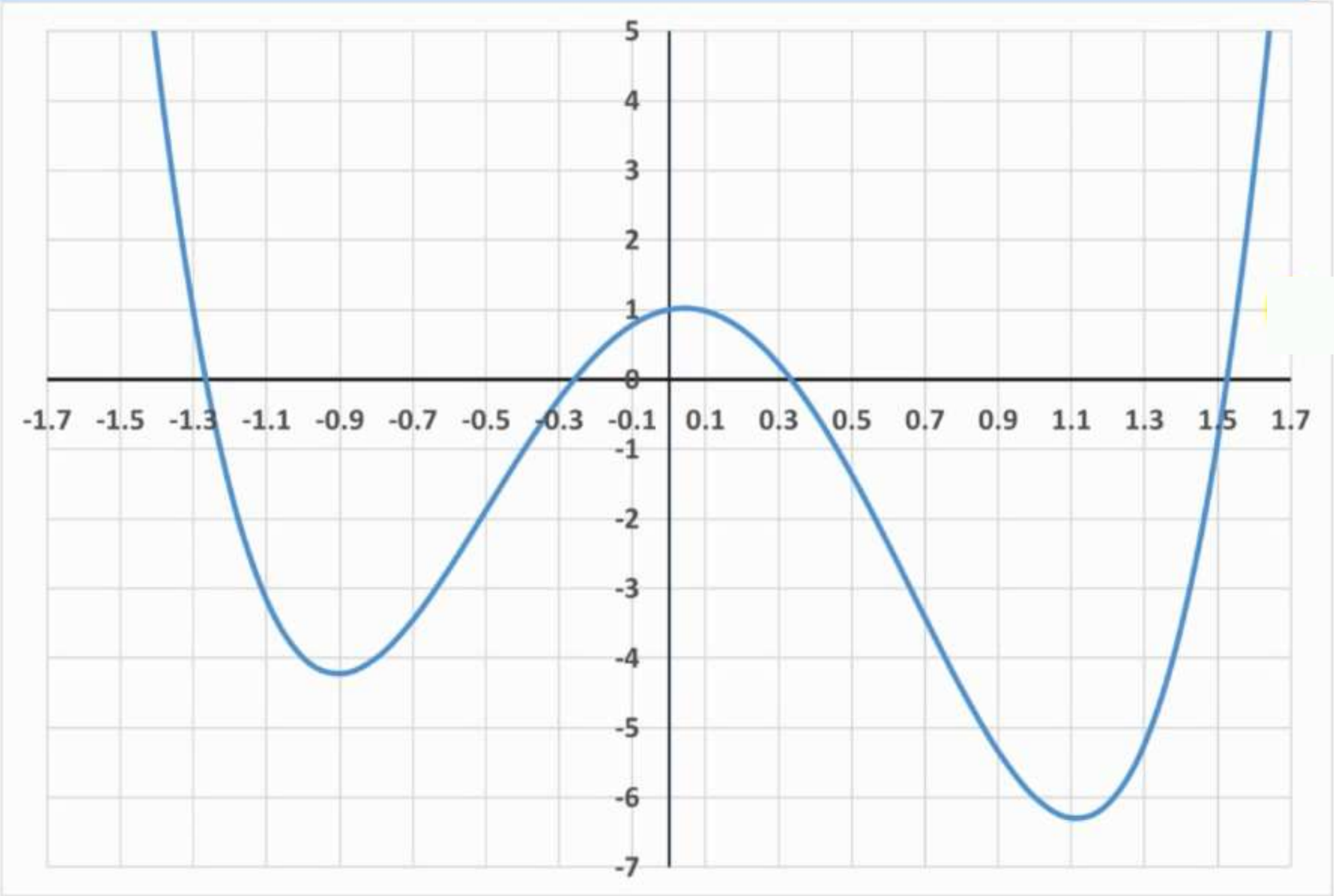


$$\frac{d^2y}{dx^2} = 2$$

Shift from Positive slope to Negative

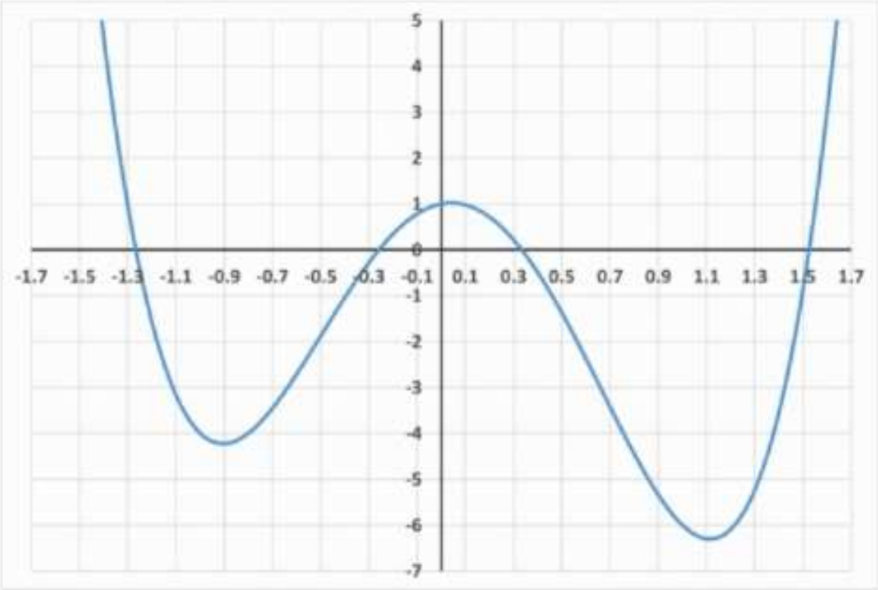
Derivative for Maxima and Minima

$$y = 6x^4 - 2x^3 - 12x^2 + x + 1$$



Derivative for Maxima and Minima

$$y = 6x^4 - 2x^3 - 12x^2 + x + 1$$



Step 1 – Get the first Derivative

Step 2 – Get the Second Derivative

Step 3 – Identify points where slope is zero

Step 4 – Get the second derivative when slope is zero

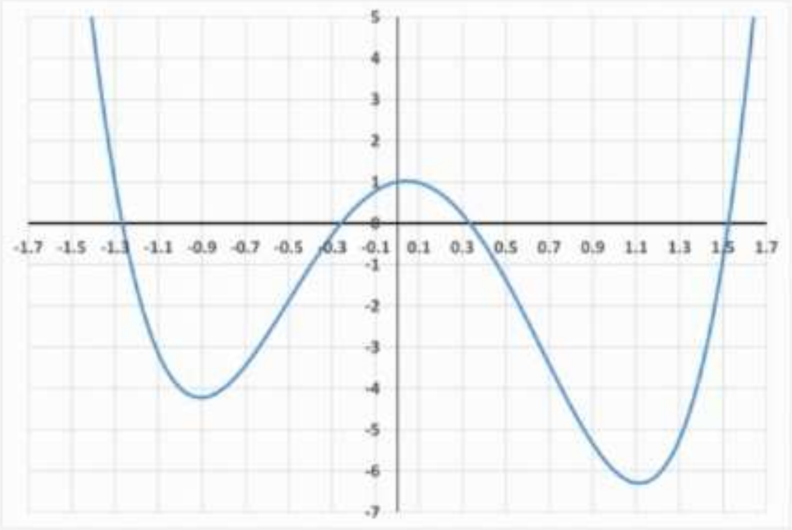
Step 5 – Apply the rules for maxima and minima

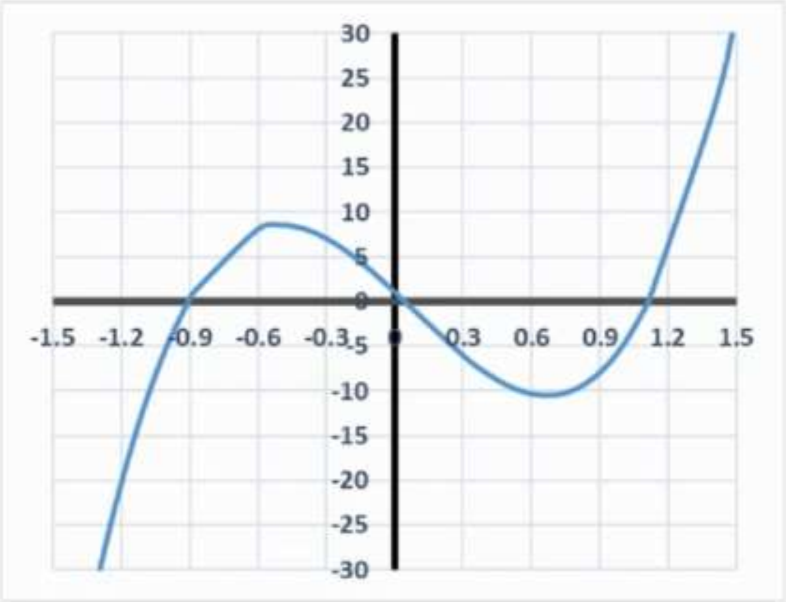
Derivative for Maxima and Minima

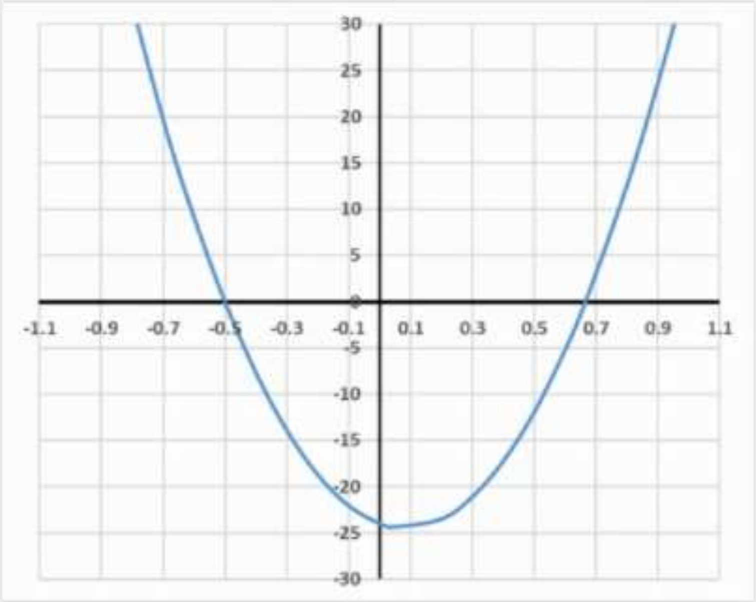
$$y = 6x^4 - 2x^3 - 12x^2 + x + 1$$

$$\frac{dy}{dx} = 24x^3 - 6x^2 - 24x + 1$$

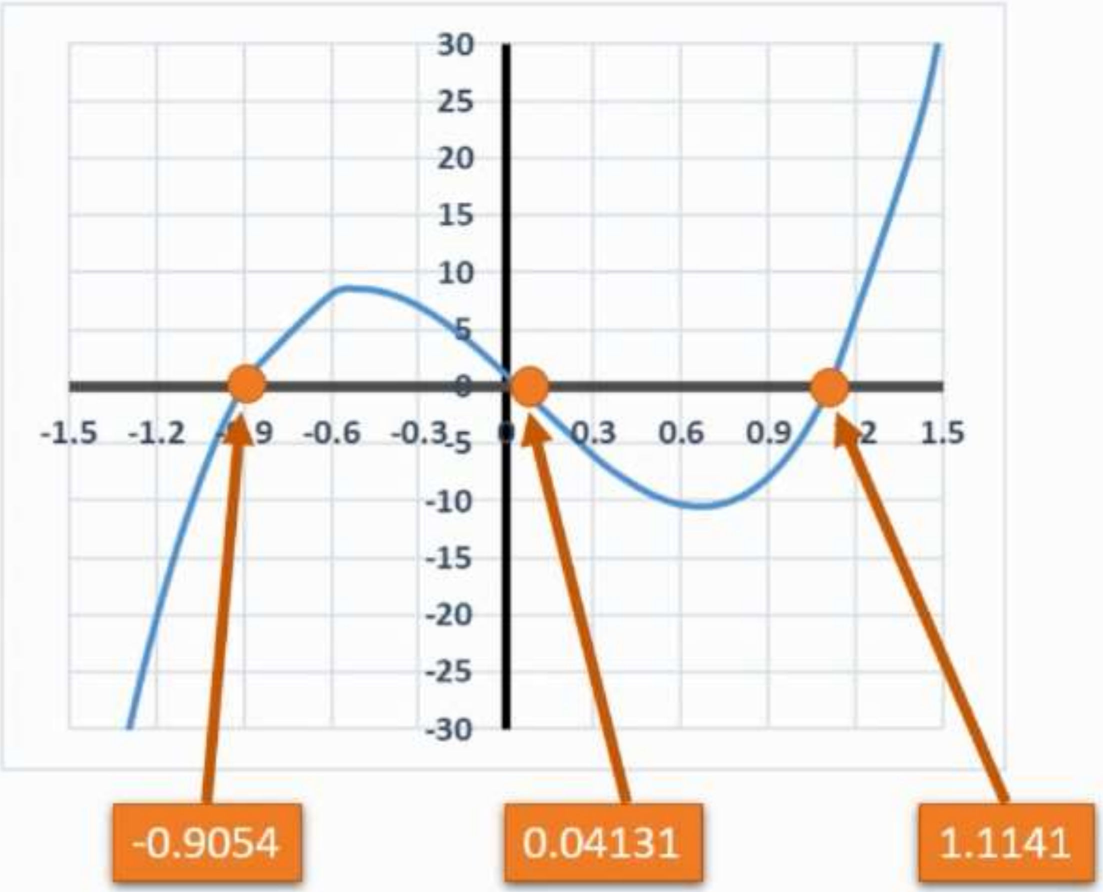
$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$







First Derivative for zero slope

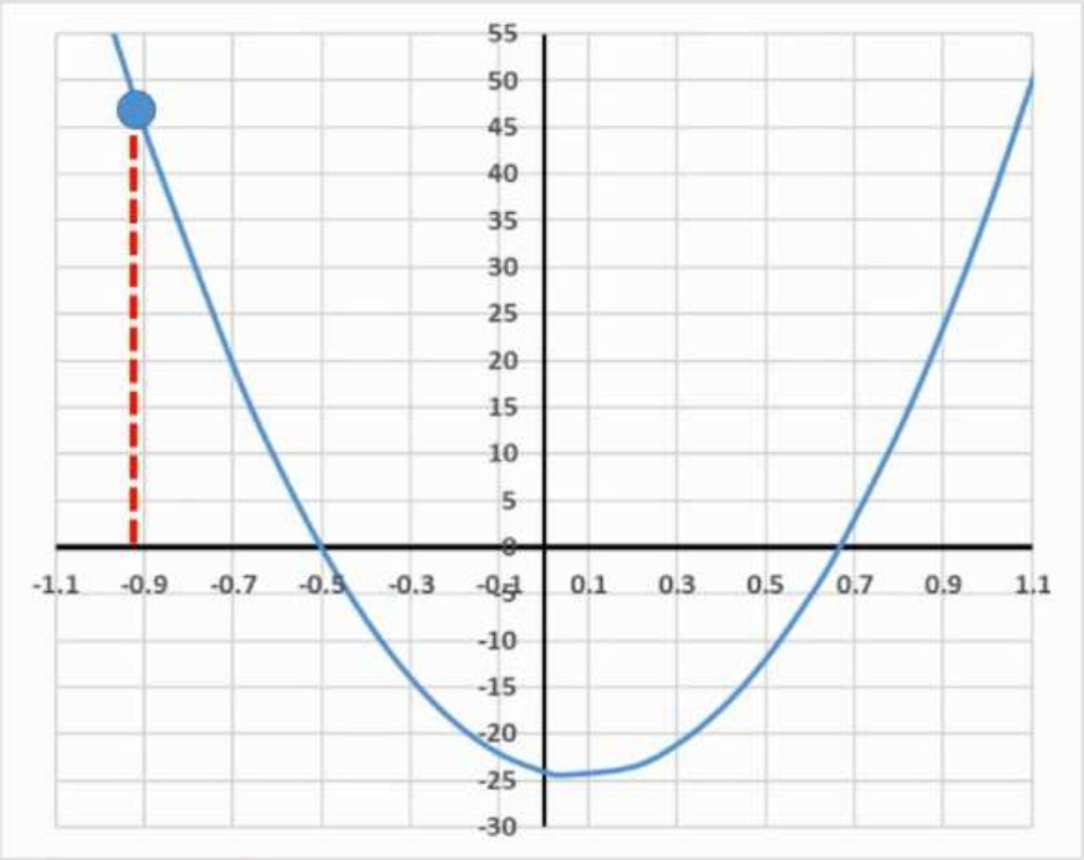


$$\frac{dy}{dx} = 24x^3 - 6x^2 - 24x + 1$$

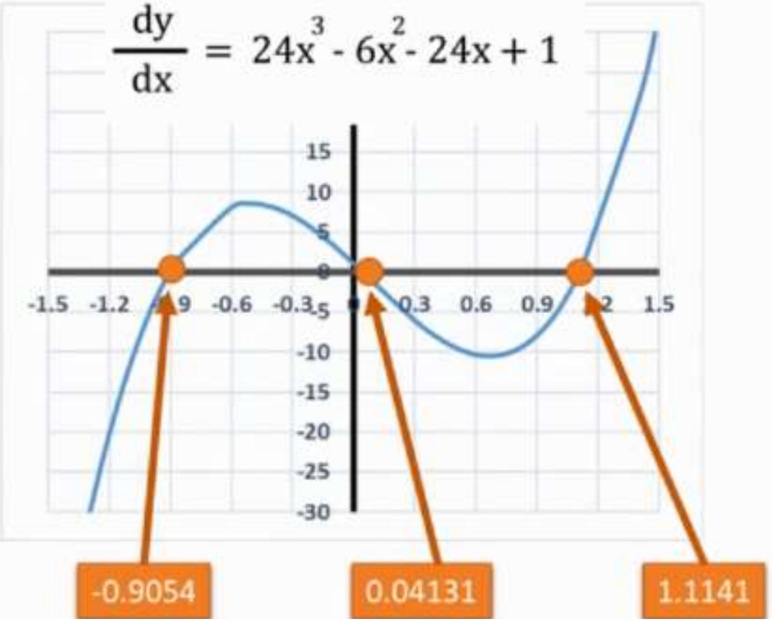
Second Derivative

$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$

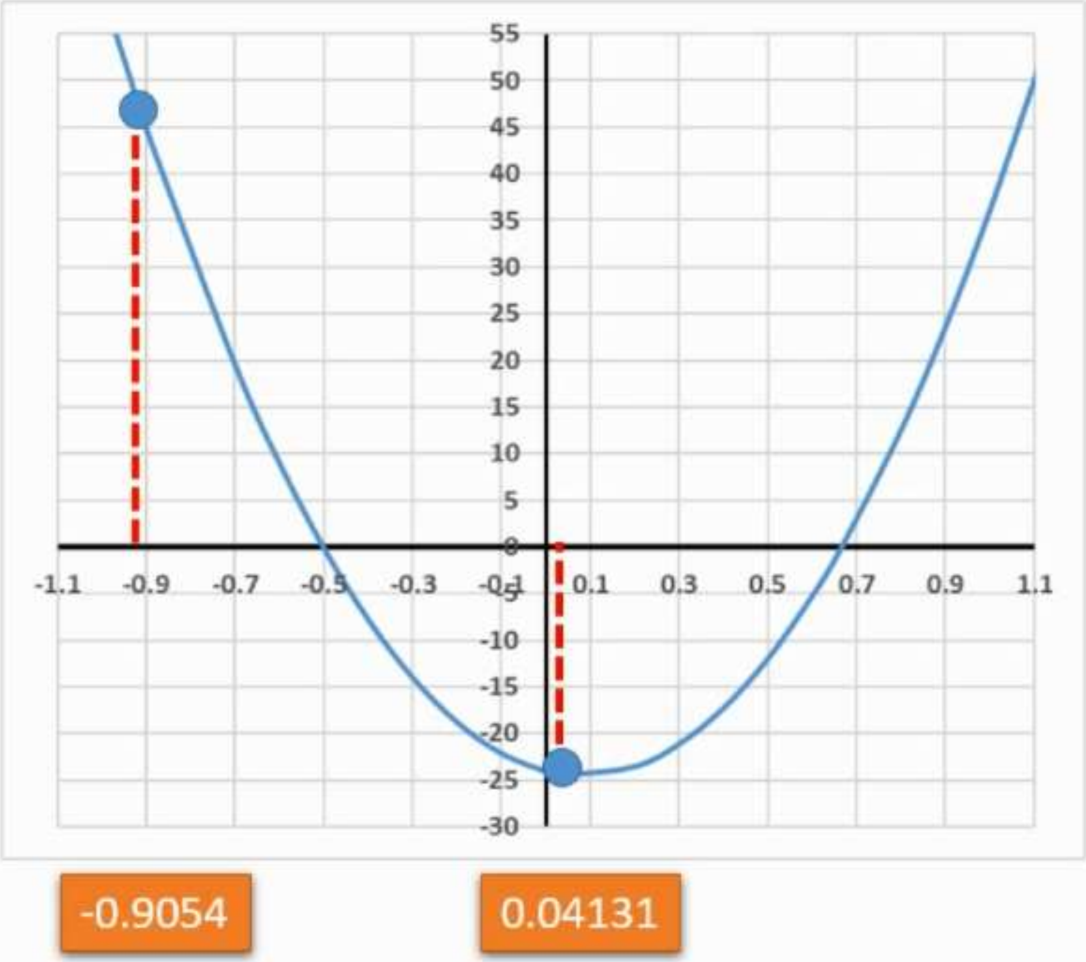
$$f(-0.9054) = 72x^2 - 12x - 24 = +45.88$$



-0.9054



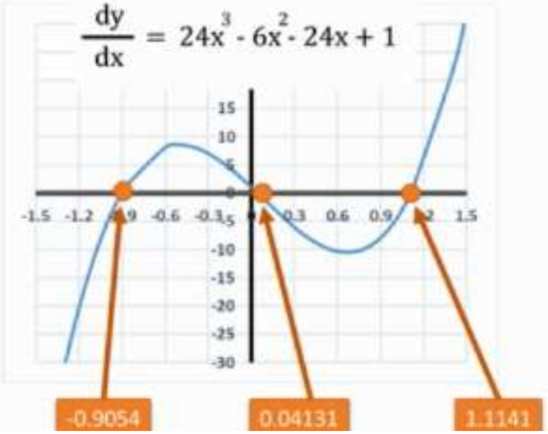
Second Derivative



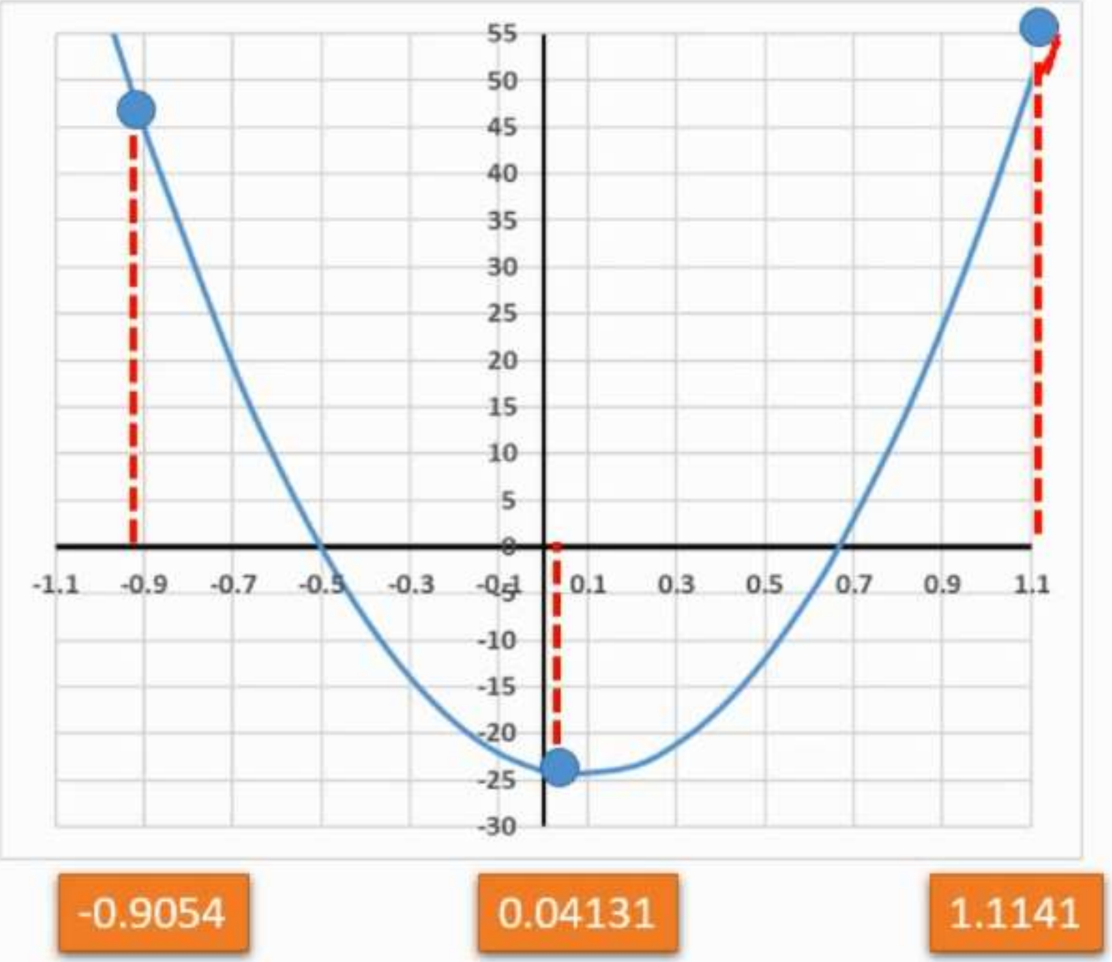
$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$

$$f(-0.9054) = 72x^2 - 12x - 24 = +45.88$$

$$f(0.0413) = 72x^2 - 12x - 24 = -24$$



Second Derivative



$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$

$$f(-0.9054) = 72x^2 - 12x - 24 = +45.88$$

$$f(0.0413) = 72x^2 - 12x - 24 = -24$$

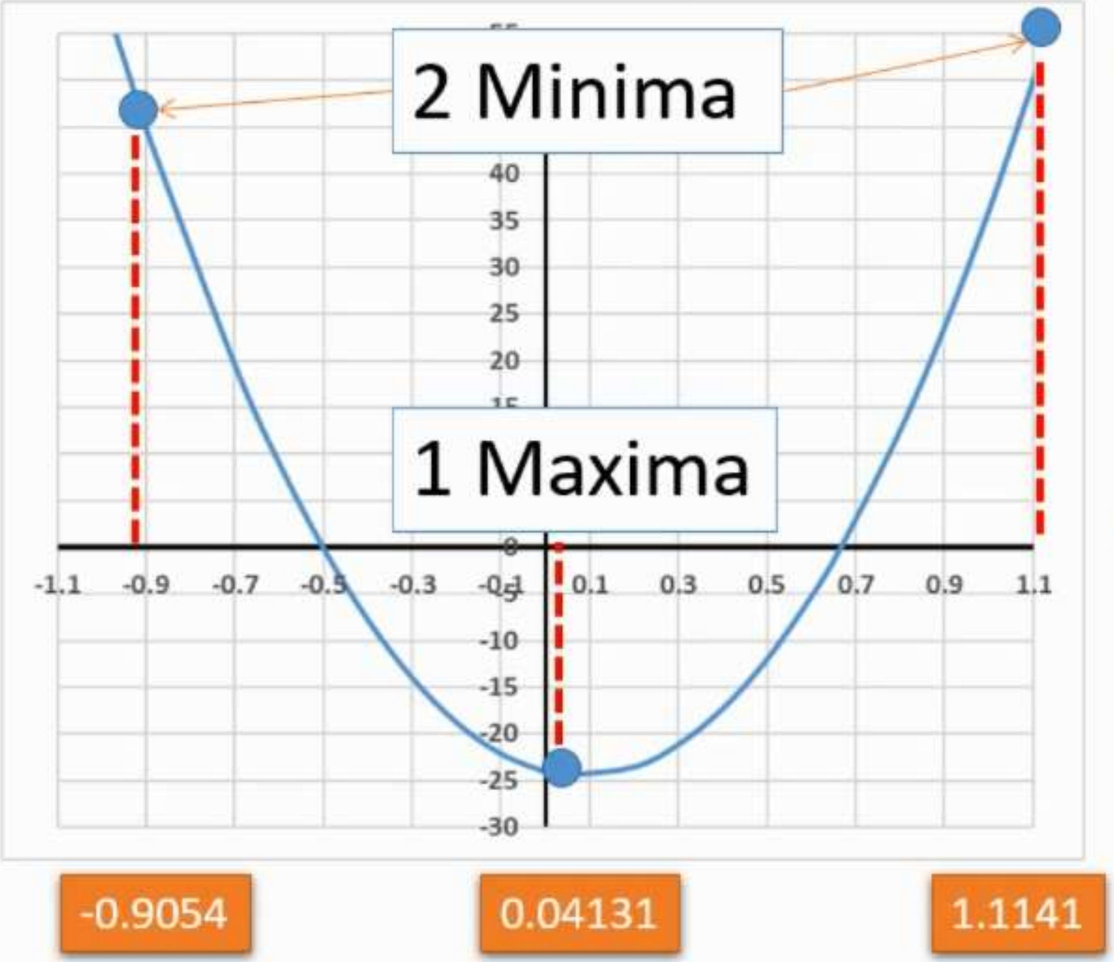
$$f(1.1141) = 72x^2 - 12x - 24 = +52$$

Rules for Maxima and Minima

Second Derivative < 0 → Local Maxima

Second Derivative > 0 → Local Minima

Get Maxima and Minima

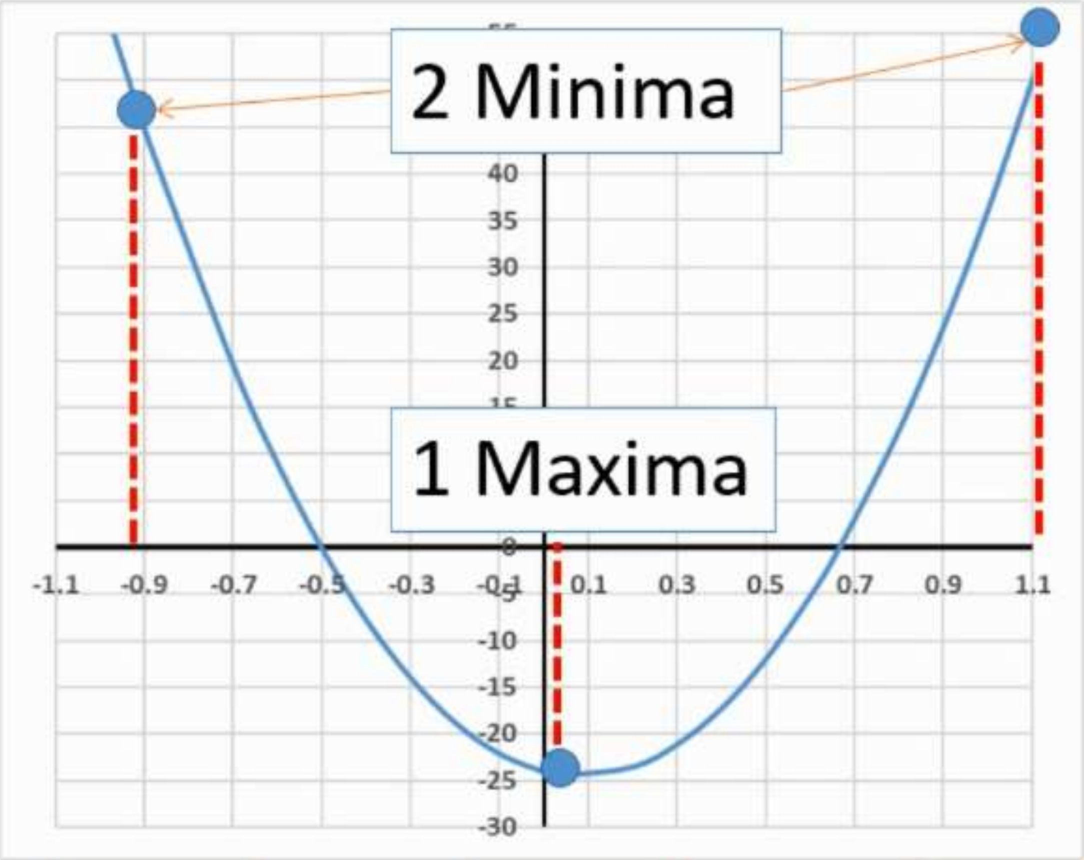


$$f(-0.9054) = 72x^2 - 12x - 24 = +45.88$$

$$f(-0.0413) = 72x^2 - 12x - 24 = -24$$

$$f(1.1141) = 72x^2 - 12x - 24 = +52$$

Get Maxima and Minima



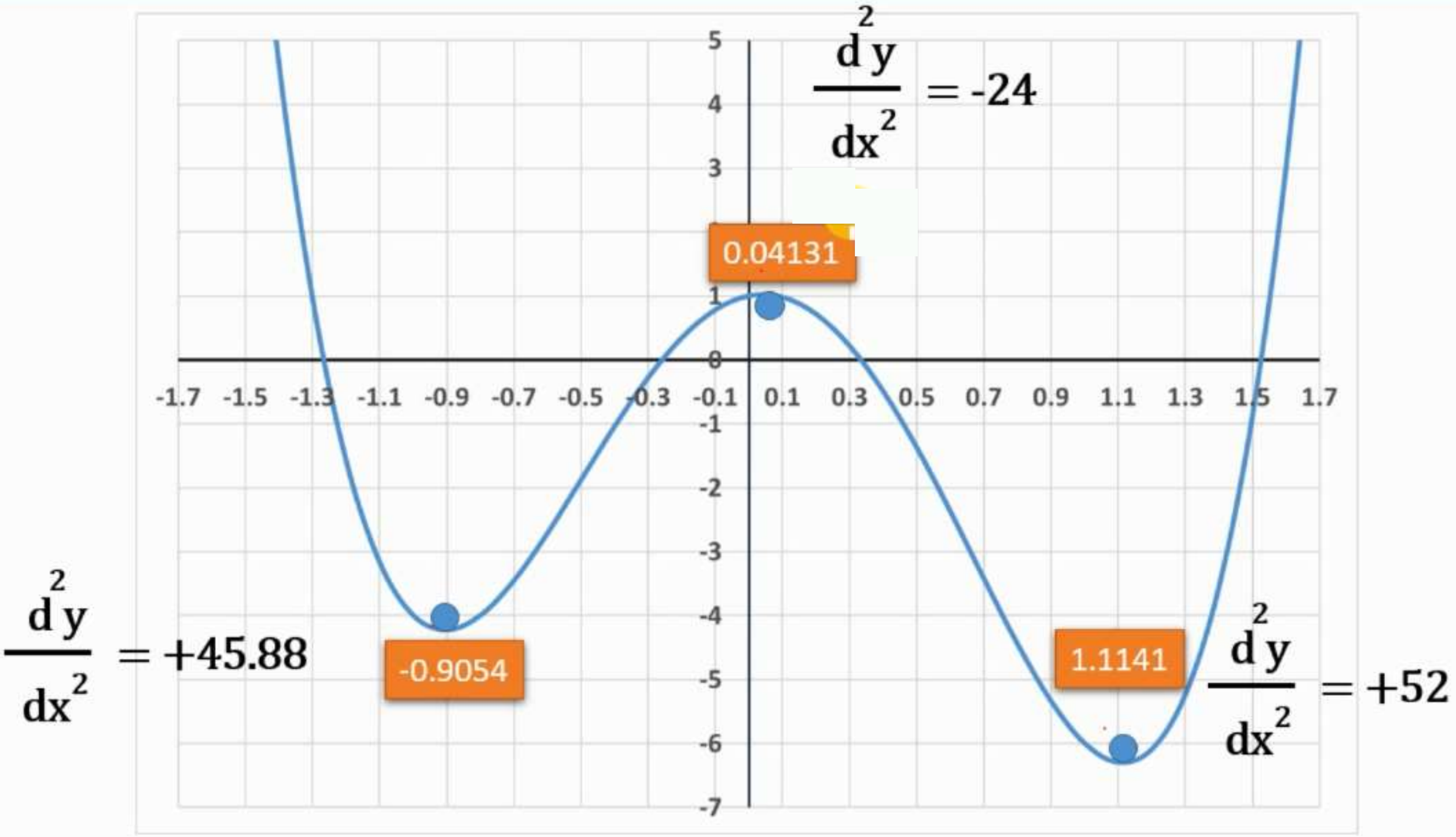
-0.9054 0.04131 1.1141

$$f(-0.9054) = 72x^2 - 12x - 24 = +45.88$$

$$f(-0.0413) = 72x^2 - 12x - 24 = -24$$

$$f(1.1141) = 72x^2 - 12x - 24 = +52$$

Actual Minima and Maxima

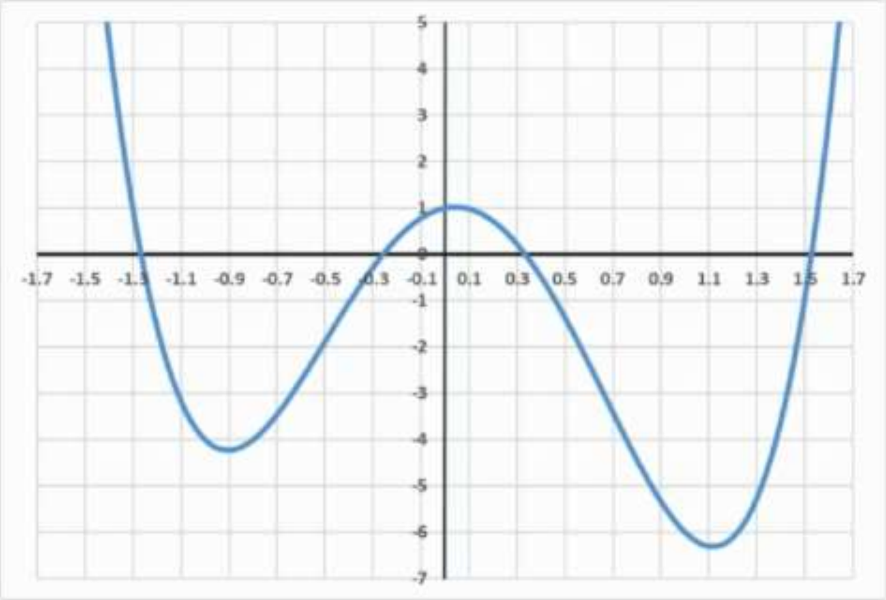


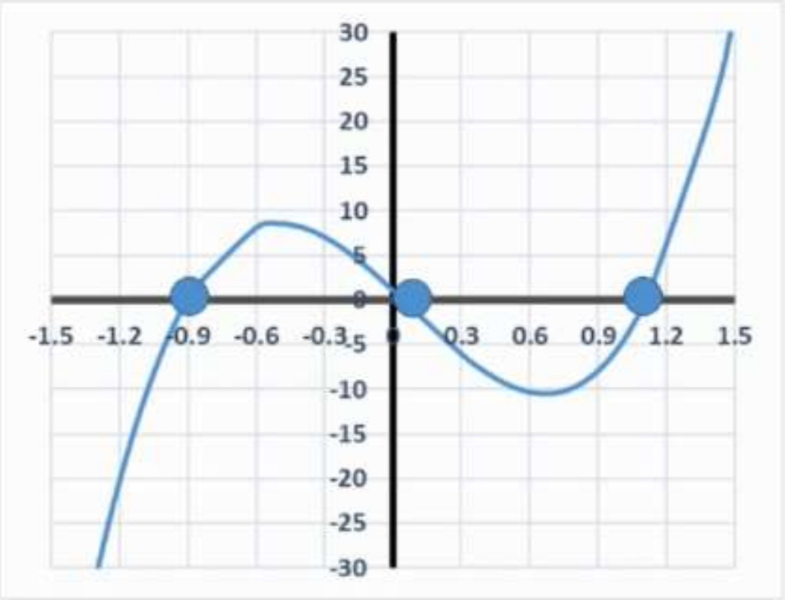
Derivative for Maxima and Minima

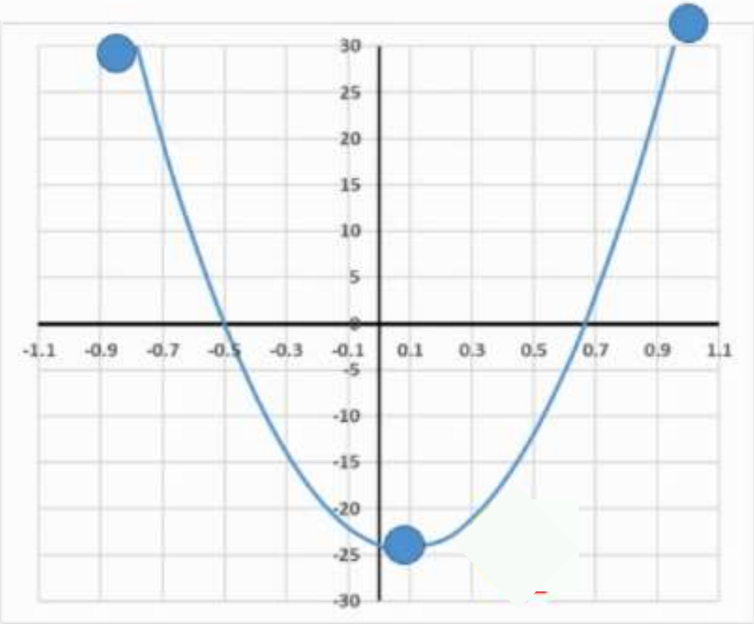
$$y = 6x^4 - 2x^3 - 12x^2 + x + 1$$

$$\frac{dy}{dx} = 24x^3 - 6x^2 - 24x + 1$$

$$\frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$





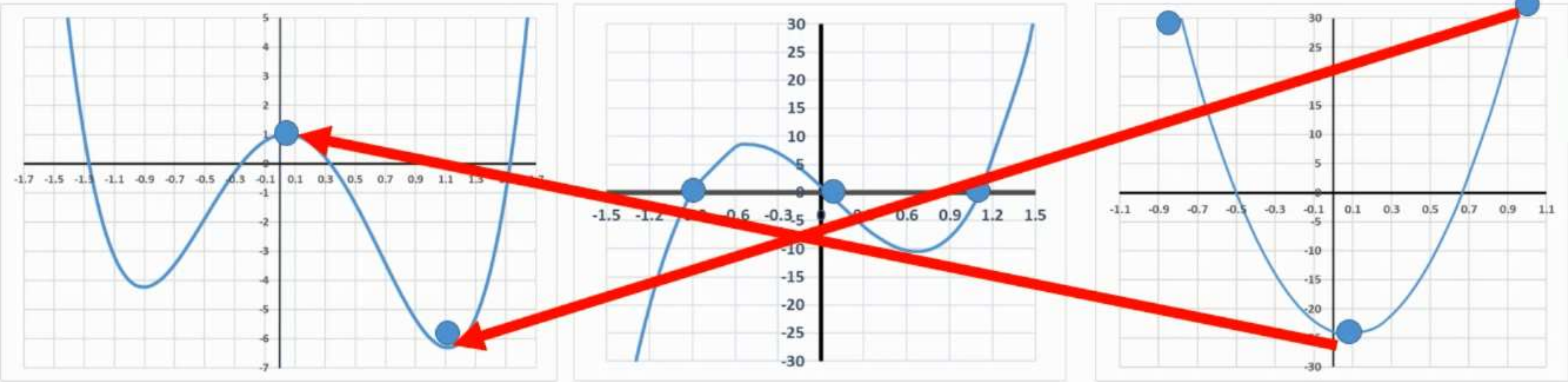


-0.9054 | 0.04131 | 1.1141

-0.9054 | 0.04131 | 1.1141

Derivative for Maxima and Minima

$$y = 6x^4 - 2x^3 - 12x^2 + x + 1 \qquad \frac{dy}{dx} = 24x^3 - 6x^2 - 24x + 1 \qquad \frac{d^2y}{dx^2} = 72x^2 - 12x - 24$$



-0.9054 | 0.04131 | 1.1141

-0.9054 | 0.04131 | 1.1141

Quiz

Question 1:

$y = 1/x$

What will be the answer of the following limit?

$$\lim_{x \to 2} \frac{1}{x} = ?$$

- ◯ 1
- ◯ 2
- ◯ Undefined
- ◯ 0.5

Question 2:

First Order Derivative of a horizontal line is zero. True or False?

- True
- False

Question 3:

$$y = 2x^3 + x^2 + 4x + 9$$

Using the chain rule of derivative, what will be the first order derivative of the y, in the following equation?

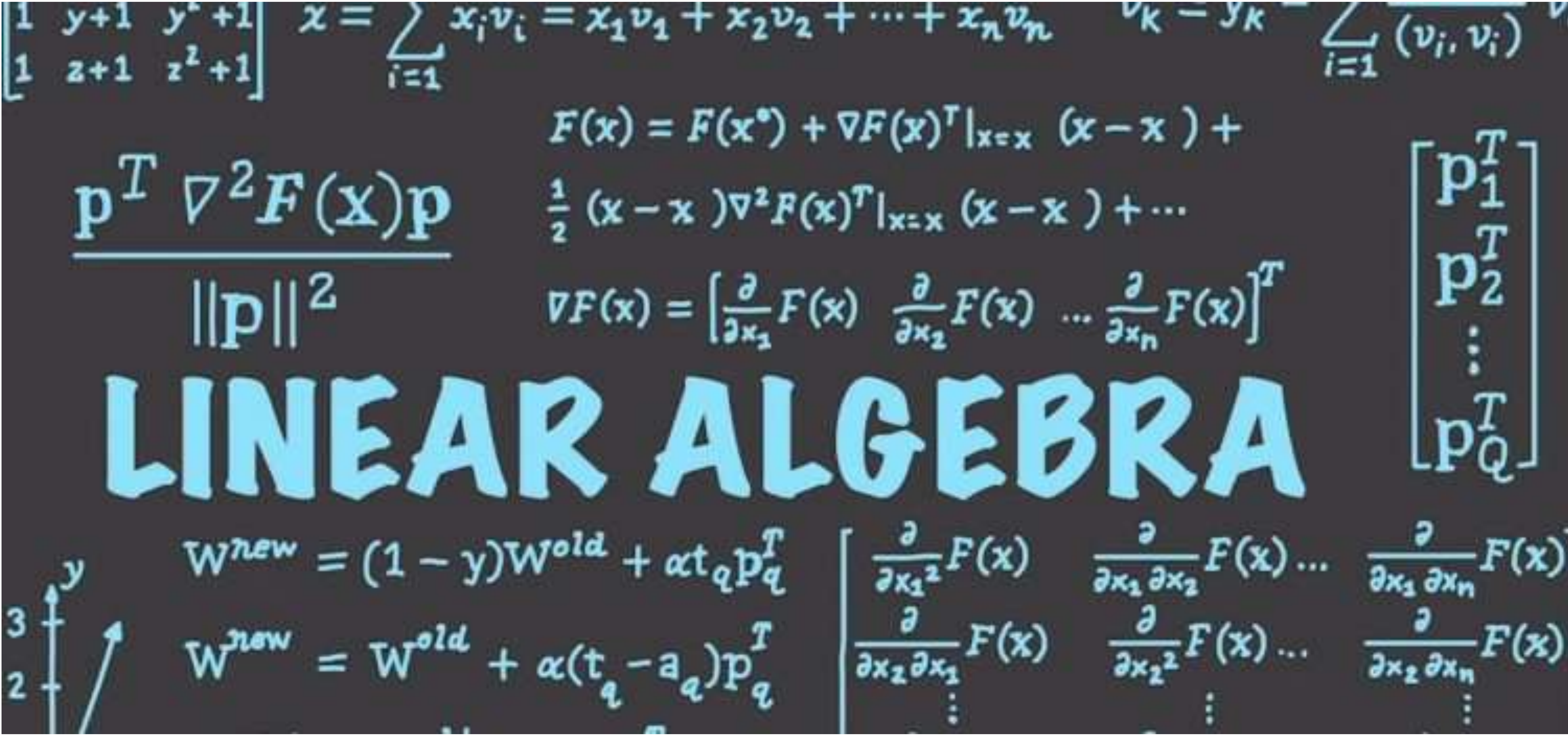
$$\frac{dy}{dx} = ?$$

- 3x^2 + 2x + 4
- 6x^2 + 2x + 4
- 2x^2 + x + 4
- None of the above

Question 4:

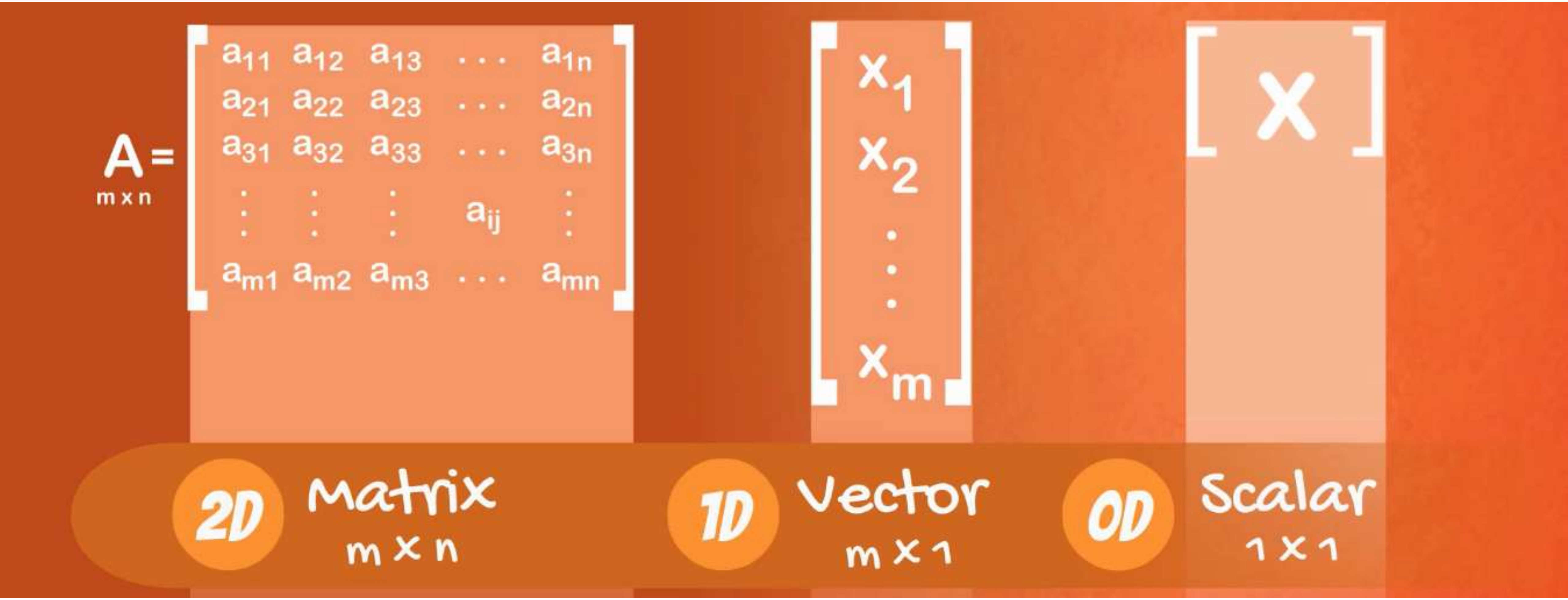
If the double derivative of the function is less than zero, we have got the minima of the function. True or False?

- **True**
- **False**



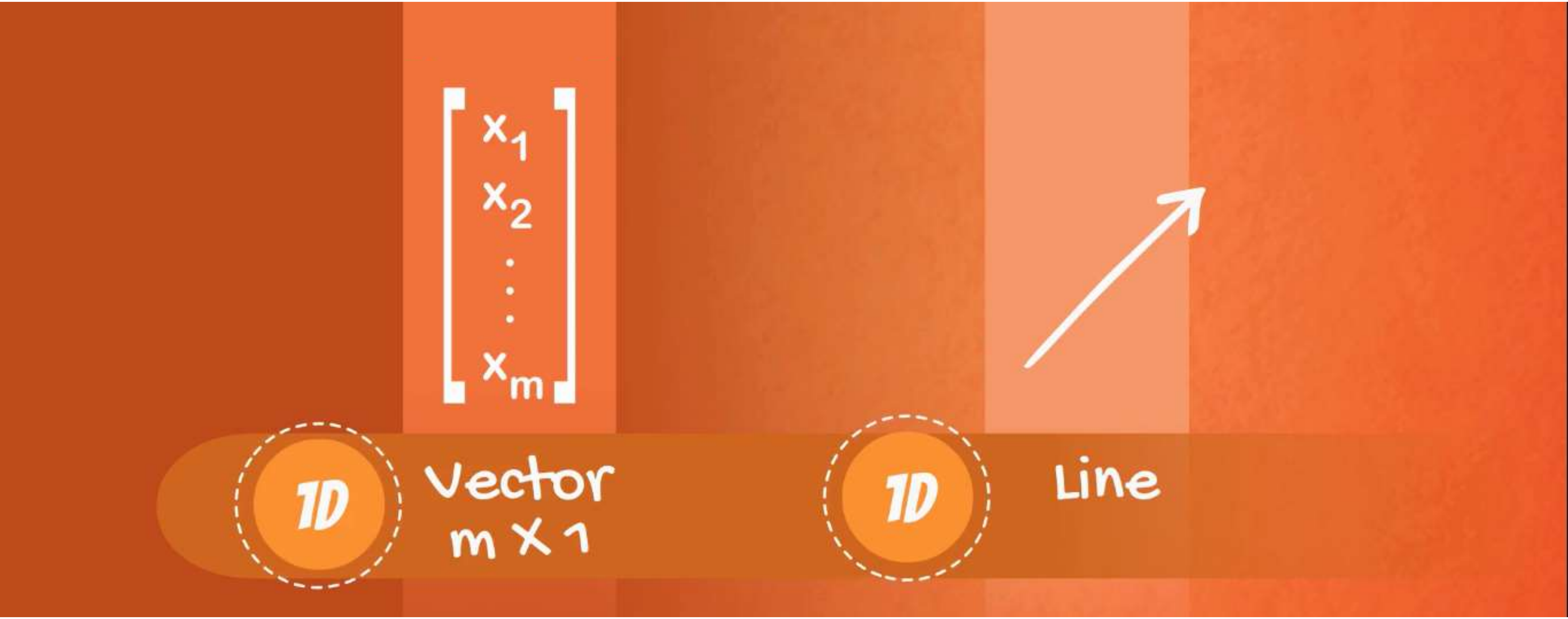
LINEAR ALGEBRA
$x = \sum_{i=1} x_i v_i = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$
$\frac{p^T \nabla^2 F(x) p}{\|p\|^2}$
$\nabla F(x) = \left[\frac{\partial}{\partial x_1} F(x) \ \frac{\partial}{\partial x_2} F(x) \ \ldots \ \frac{\partial}{\partial x_n} F(x)\right]^T$
$W^{new} = (1-y)W^{old} + \alpha t_q p_q^T$
$W^{new} = W^{old} + \alpha(t_q - a_q)p_q^T$

Matrix Vs. Vector Vs. Scalar



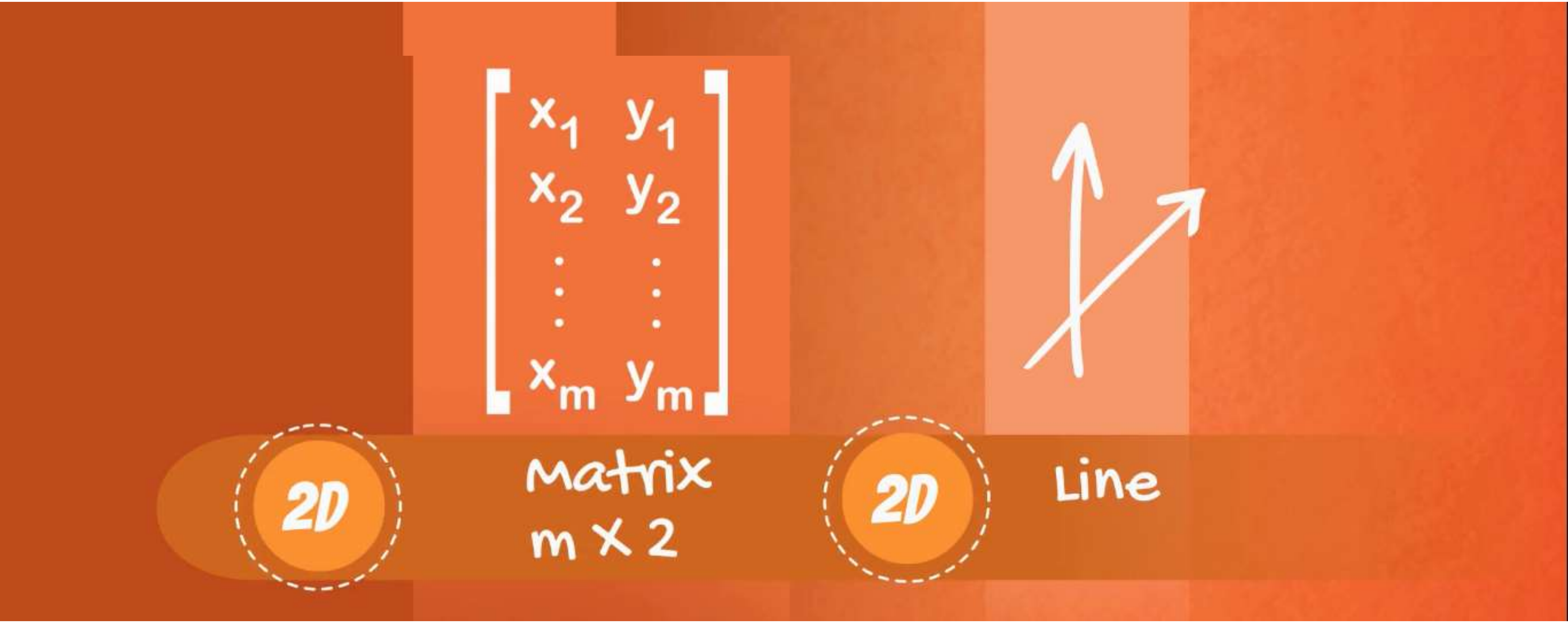
Scalar

Vector



$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$
1D
Vector
m x 1
1D
Line

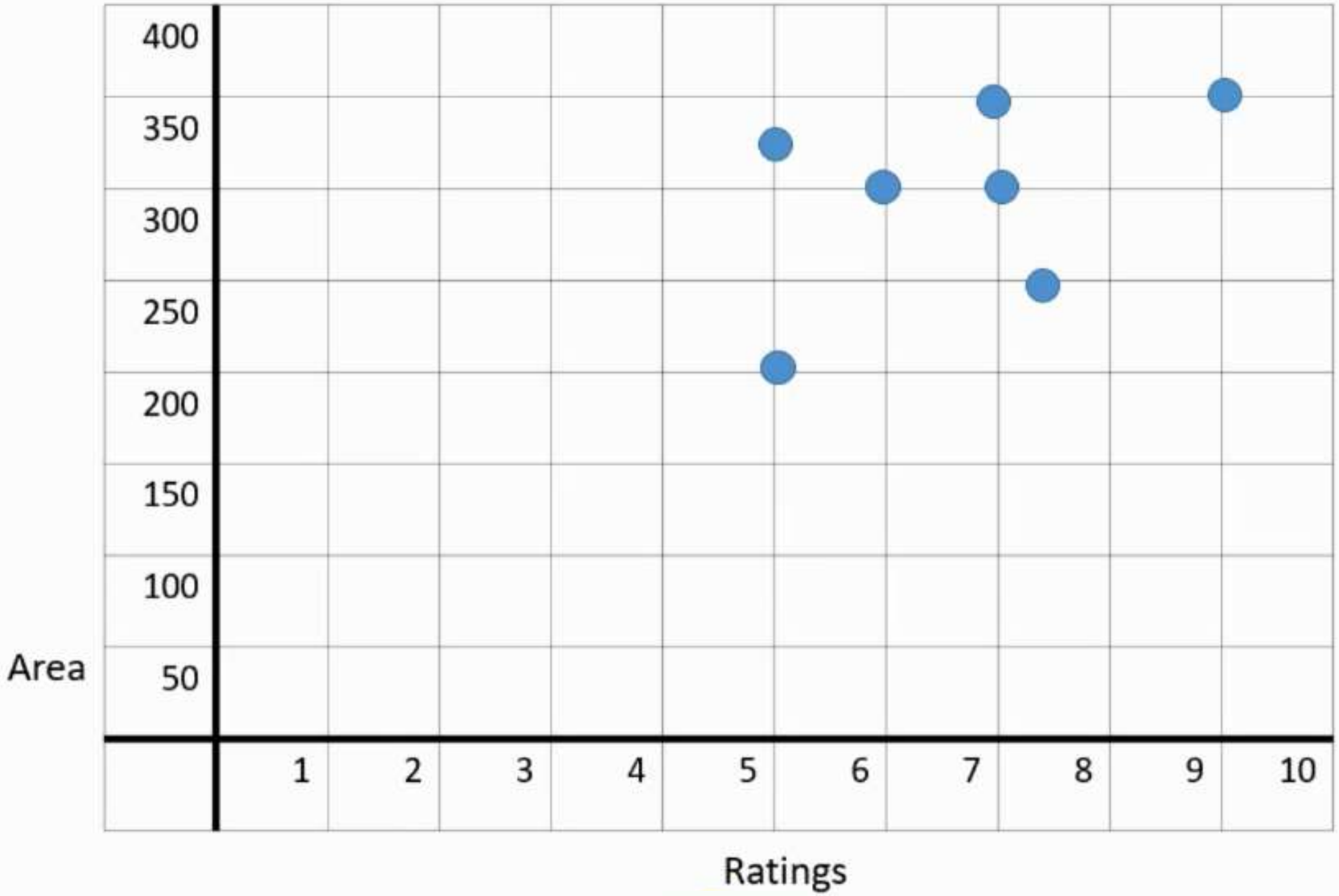
Matrix



$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_m & y_m \end{bmatrix}$$
2D
Matrix
m X 2
2D
Line

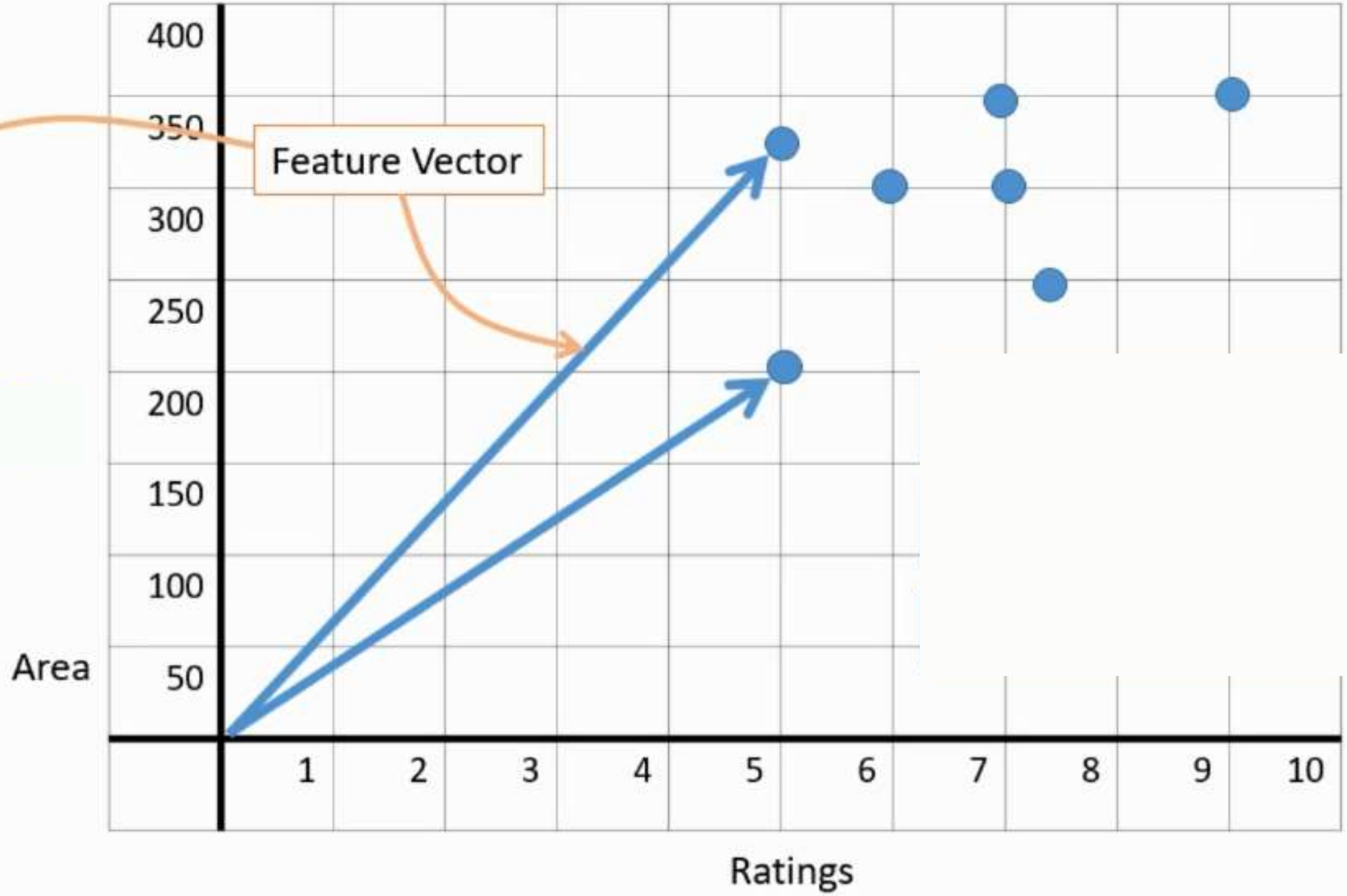
Vectors in Machine Learning

Rating	Area sq. mtr
5	200
7	300
5	325
8	250
6	300
7	350
7.5	250
9	350



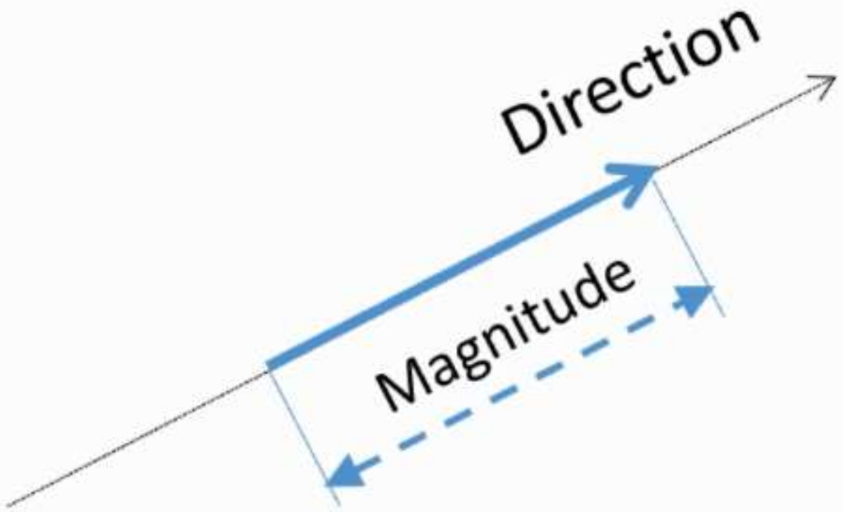
Vectors in Machine Learning

Rating	Area sq. mtr
5	200
7	300
5	325
8	250
6	300
7	350
7.5	250
9	350



What is a vector?

$$\vec{V}$$



What is a vector?

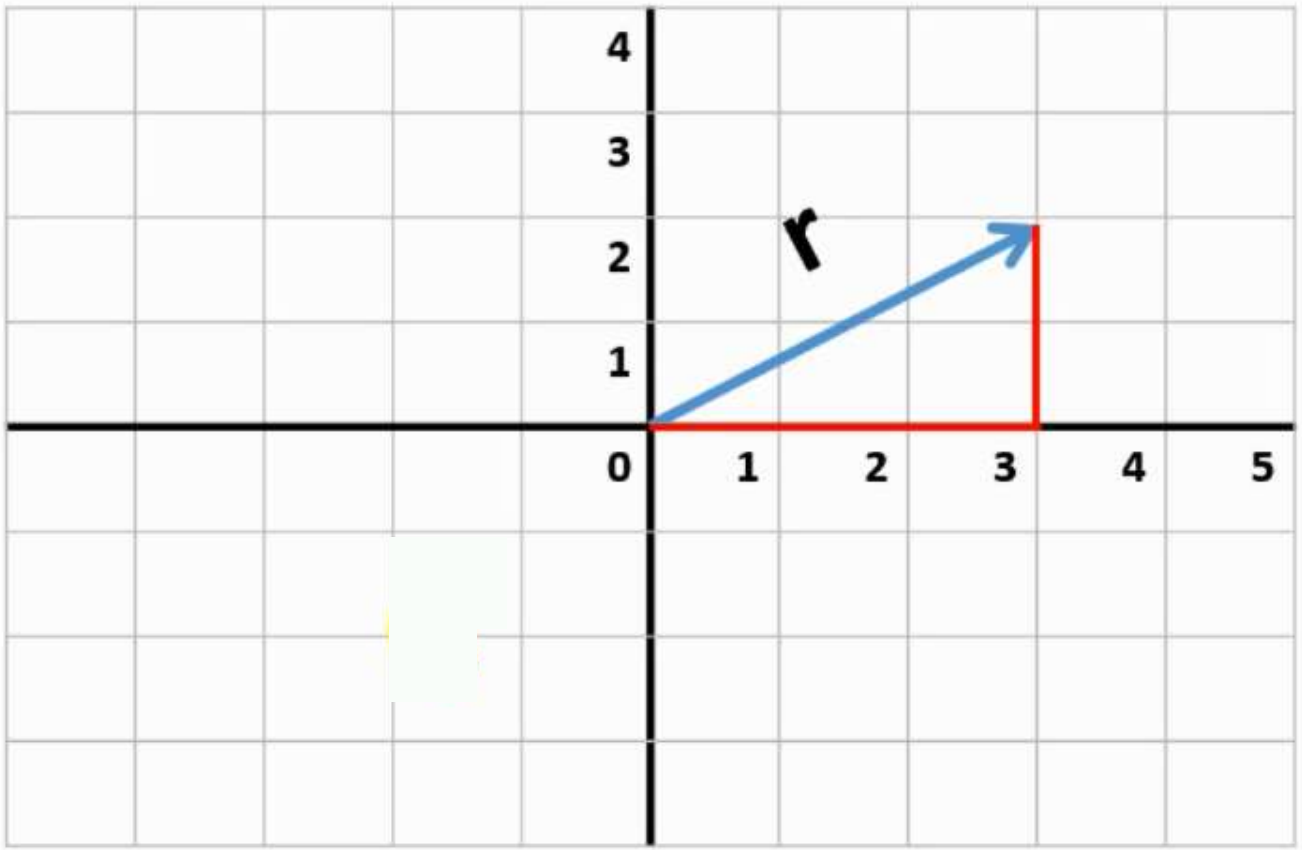
Cartesian:

$$\vec{V} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$r^2 = a^2 + b^2$$

$$= 9 + 4$$

$$= \sqrt{13}$$



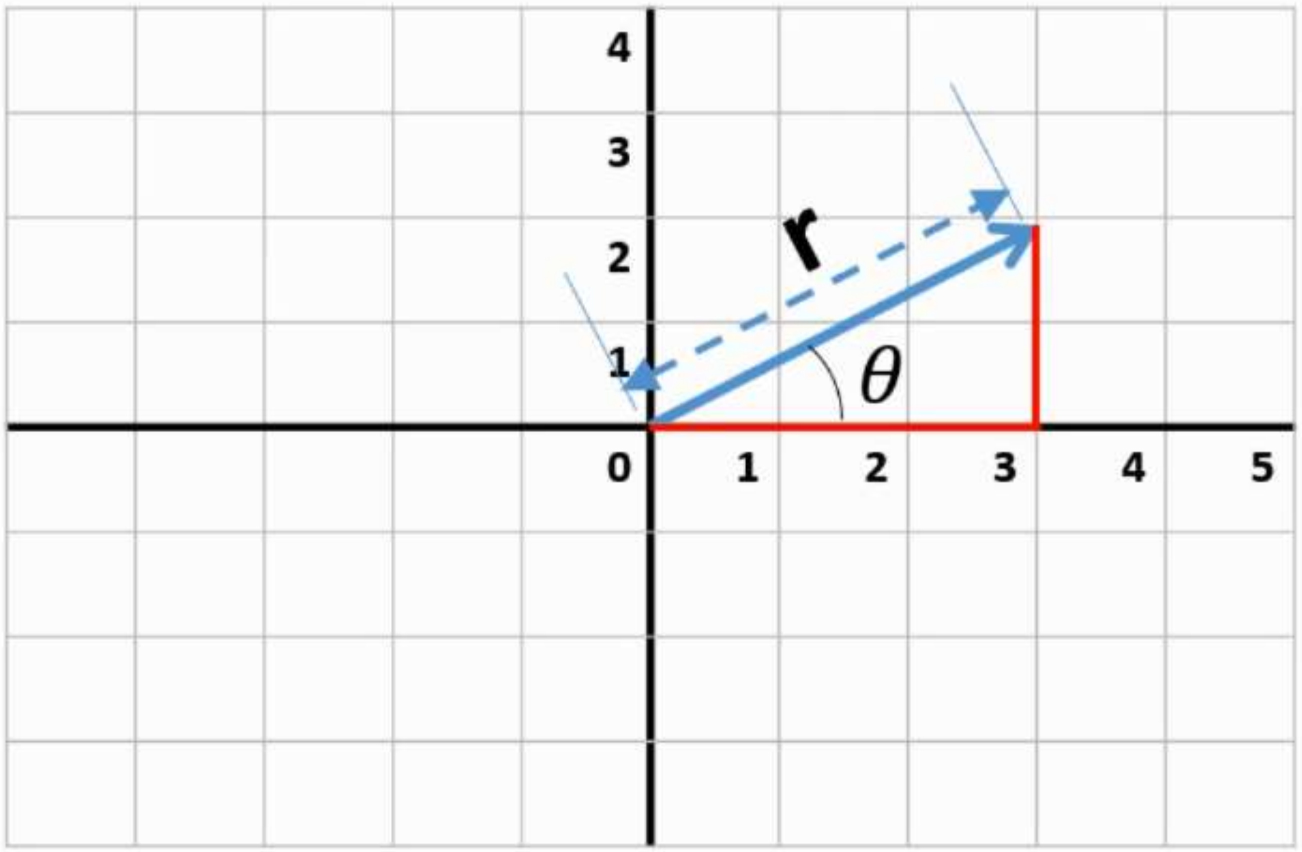
What is a vector?

Cartesian:

$$\vec{V} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

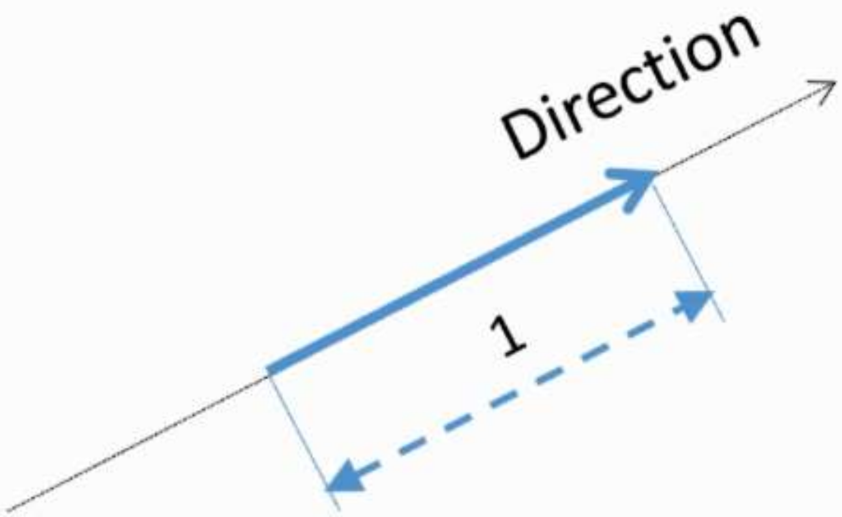
Polar:

$$\vec{V} = (r,\ \theta)$$

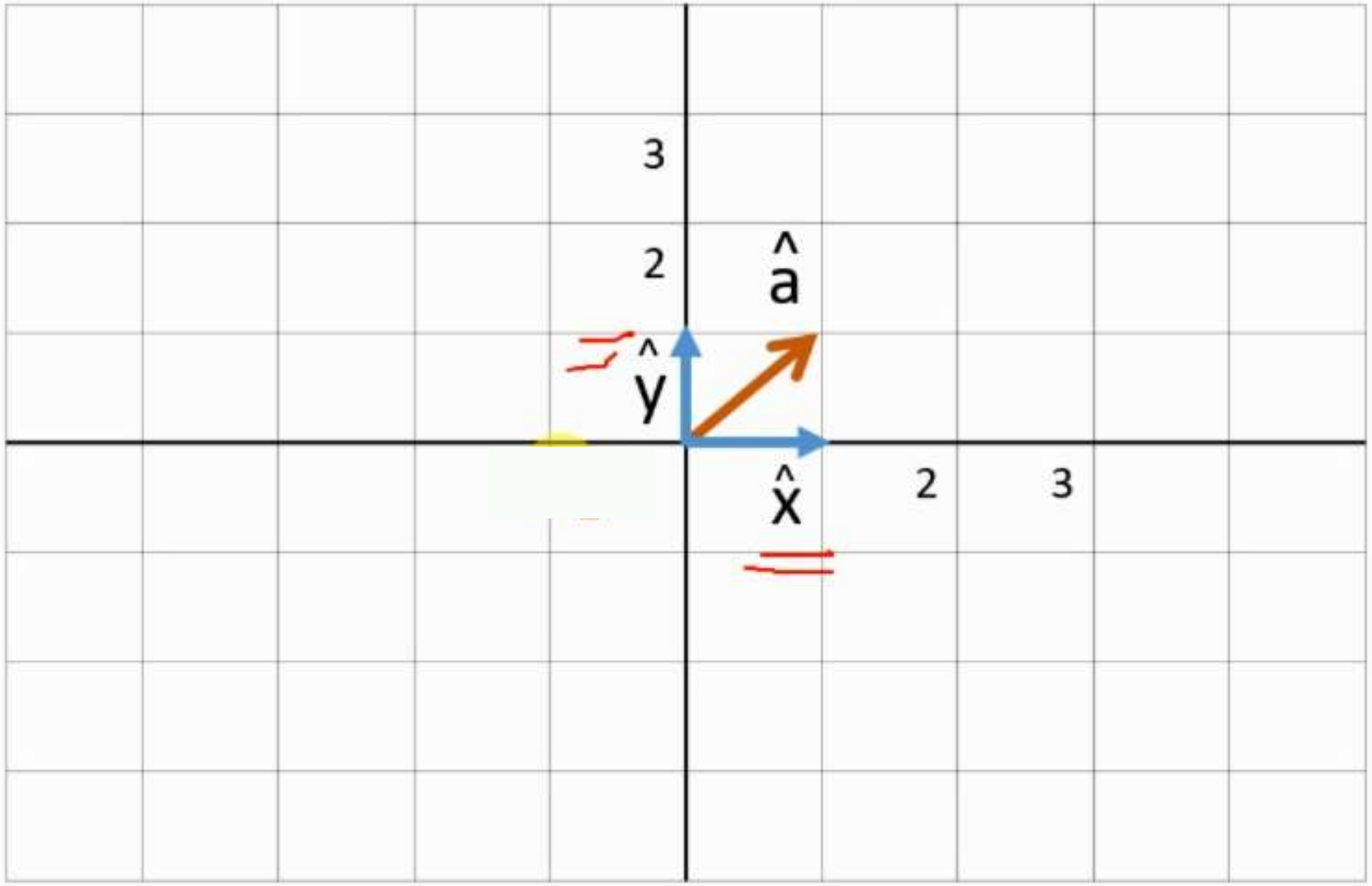


Unit Vector

$\hat{a}$



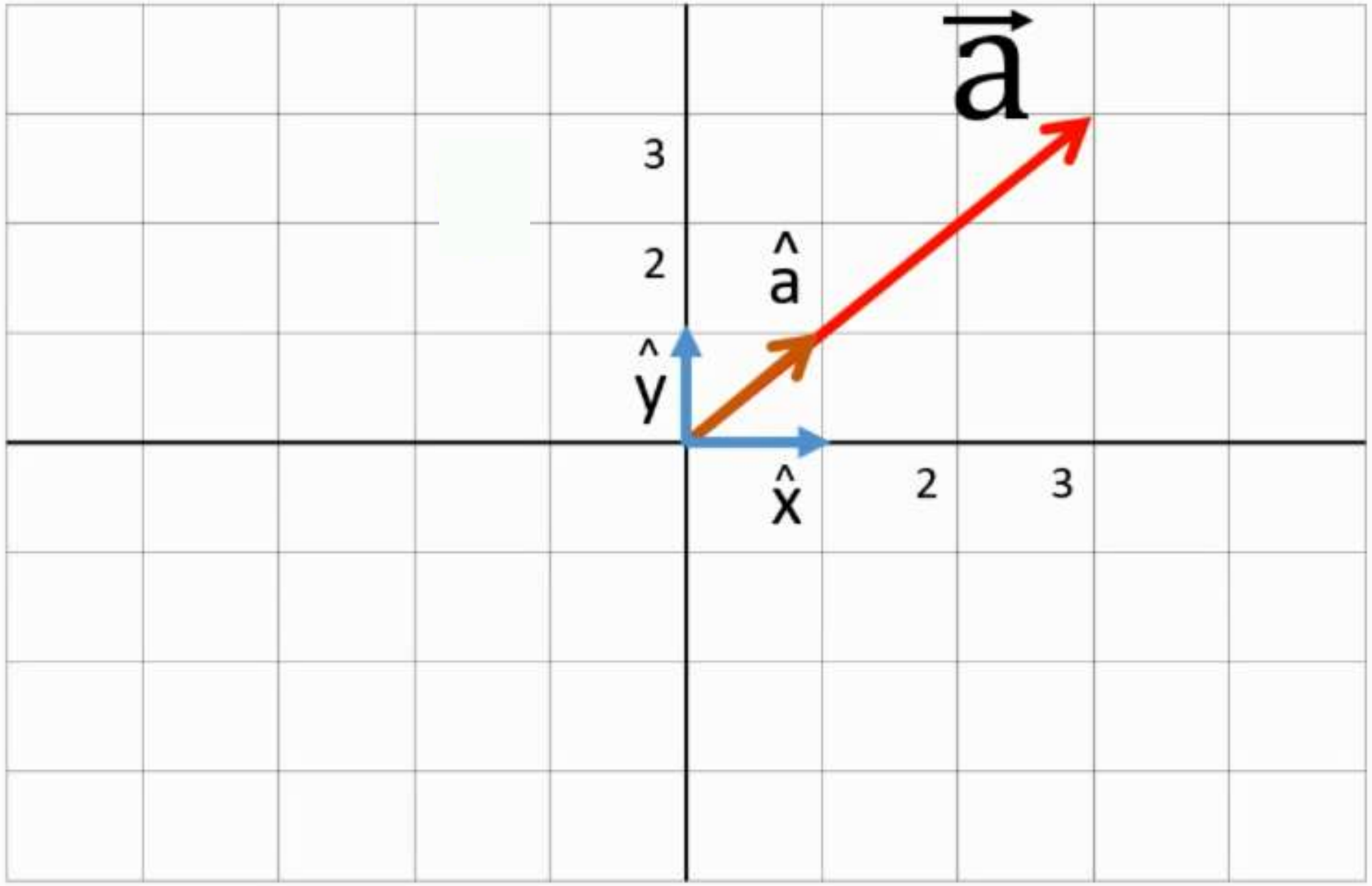
Unit Vector



Unit Vector

$$\vec{a} = 3 * \hat{a}$$

$$\vec{a} = 3 * \hat{x} + 3 * \hat{y}$$

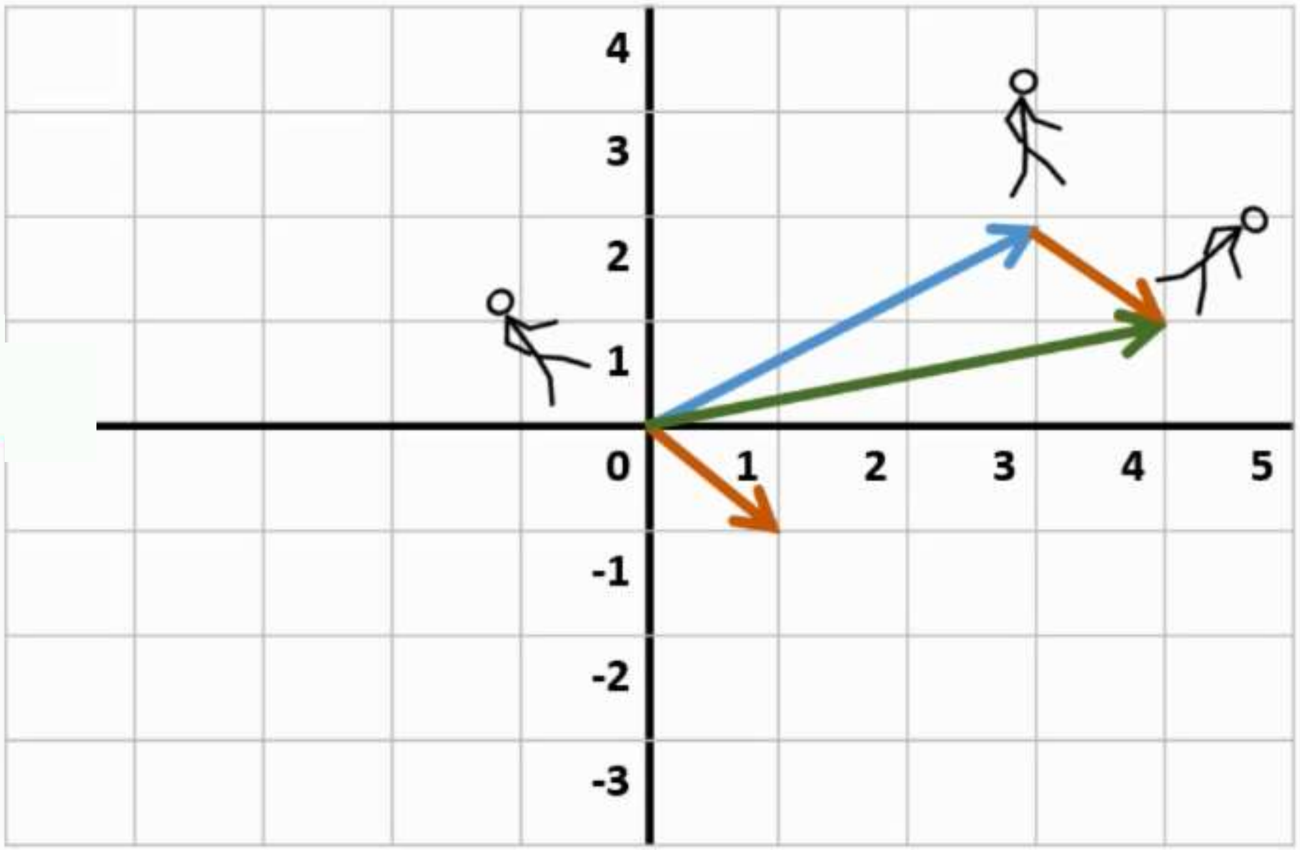


Vector Addition

$$\vec{V}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\vec{V}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\vec{V}_1 + \vec{V}_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

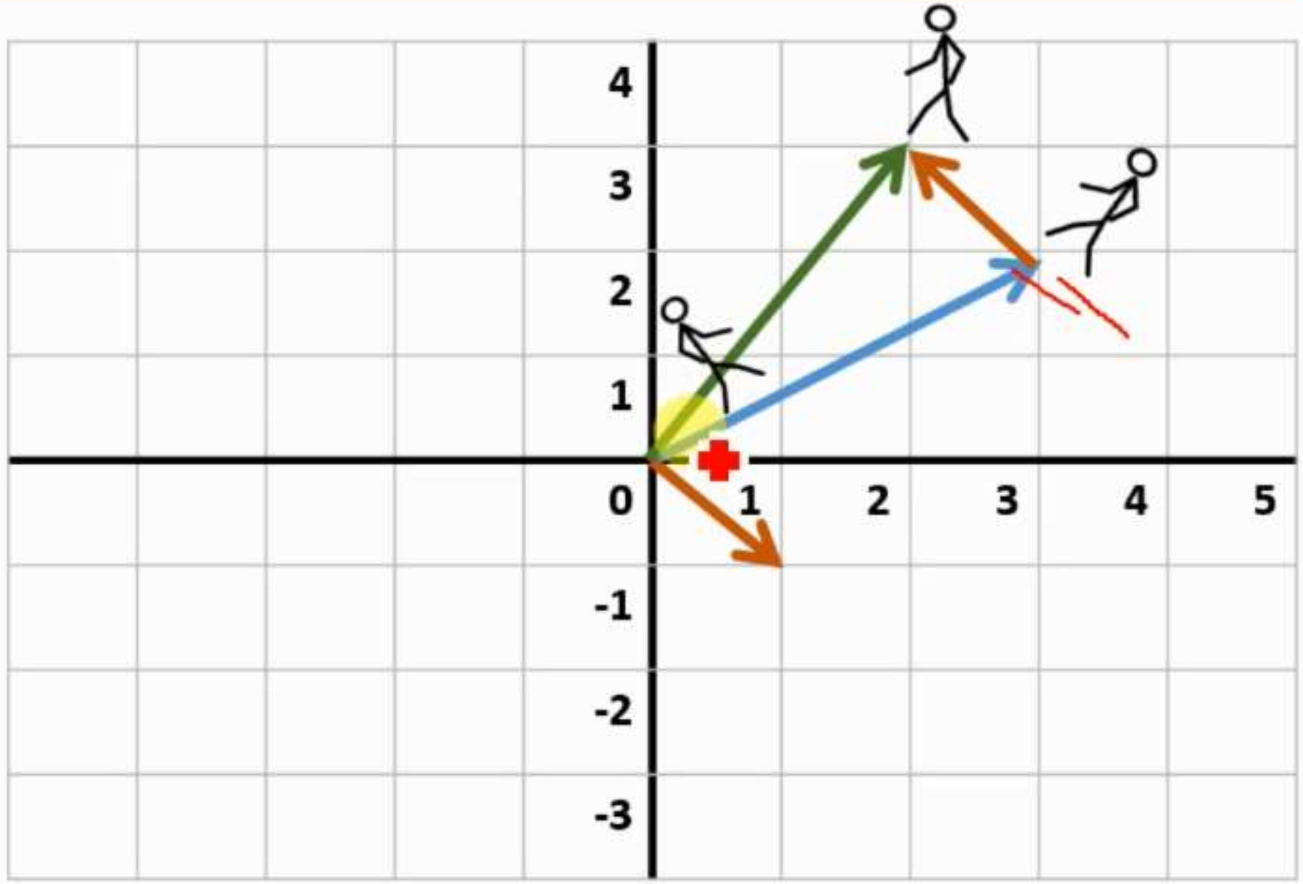


Vector Subtraction

$$\vec{V}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\vec{V}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\vec{V}_1 - \vec{V}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$



Why should we learn Matrices?

What is a Matrix?

$$X = \left.\underbrace{\begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}}_{\text{Columns}}\right\} \text{Rows}$$

Why should we learn Matrices?

Fixed Acidity	Volatile Acidity	Citric Acid	Residual Sugar	Chlorides	Free Sulfur Dioxide	Total Sulfur Dioxide	Density	pH	Sulphates	Alcohol	Quality
7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.8	0.88	0	2.6	0.098	25	67	0.9968	3.2	0.68	9.8	5
7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	5
11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	6
7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.4	0.66	0	1.8	0.075	13	40	0.9978	3.51	0.56	9.4	5
7.9	0.6	0.06	1.6	0.069	15	59	0.9964	3.3	0.46	9.4	6
7.3	0.65	0	1.2	0.065	15	21	0.9946	3.39	0.47	10	7
7.8	0.58	0.02	2	0.073	9	18	0.9968	3.36	0.57	9.5	7

Why should we learn Matrices?

7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.8	0.88	0	2.6	0.098	25	67	0.9968	3.2	0.68	9.8	5
7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	5
11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	6
7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.4	0.66	0	1.8	0.075	13	40	0.9978	3.51	0.56	9.4	5
7.9	0.6	0.06	1.6	0.069	15	59	0.9964	3.3	0.46	9.4	6
7.3	0.65	0	1.2	0.065	15	21	0.9946	3.39	0.47	10	7
7.8	0.58	0.02	2	0.073	9	18	0.9968	3.36	0.57	9.5	7

Why should we learn Matrices?

7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.8	0.88	0	2.6	0.098	25	67	0.9968	3.2	0.68	9.8	5
7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	5
11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	6
7.4	0.7	0	1.9	0.076	11	34	0.9978	3.51	0.56	9.4	5
7.4	0.66	0	1.8	0.075	13	40	0.9978	3.51	0.56	9.4	5
7.9	0.6	0.06	1.6	0.069	15	59	0.9964	3.3	0.46	9.4	6
7.3	0.65	0	1.2	0.065	15	21	0.9946	3.39	0.47	10	7
7.8	0.58	0.02	2	0.073	9	18	0.9968	3.36	0.57	9.5	7

Why should we learn Matrices?

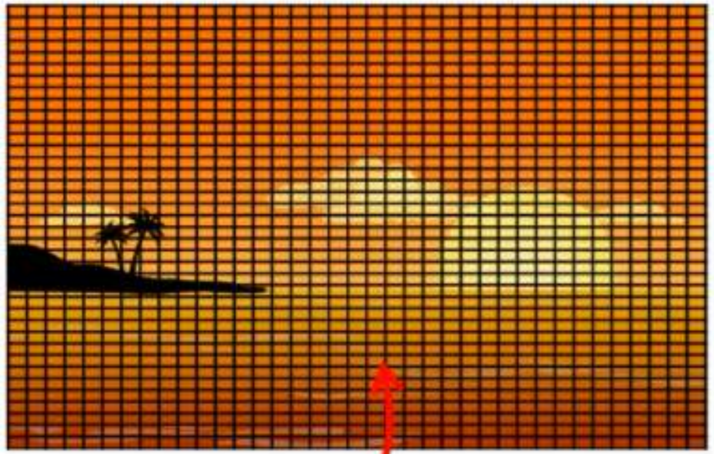
$$\begin{bmatrix}
7.4 & 0.7 & 0 & 1.9 & 0.076 & 11 & 34 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.8 & 0.88 & 0 & 2.6 & 0.098 & 25 & 67 & 0.9968 & 3.2 & 0.68 & 9.8 & 5 \\
7.8 & 0.76 & 0.04 & 2.3 & 0.092 & 15 & 54 & 0.997 & 3.26 & 0.65 & 9.8 & 5 \\
11.2 & 0.28 & 0.56 & 1.9 & 0.075 & 17 & 60 & 0.998 & 3.16 & 0.58 & 9.8 & 6 \\
7.4 & 0.7 & 0 & 1.9 & 0.076 & 11 & 34 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.4 & 0.66 & 0 & 1.8 & 0.075 & 13 & 40 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.9 & 0.6 & 0.06 & 1.6 & 0.069 & 15 & 59 & 0.9964 & 3.3 & 0.46 & 9.4 & 6 \\
7.3 & 0.65 & 0 & 1.2 & 0.065 & 15 & 21 & 0.9946 & 3.39 & 0.47 & 10 & 7 \\
7.8 & 0.58 & 0.02 & 2 & 0.073 & 9 & 18 & 0.9968 & 3.36 & 0.57 & 9.5 & 7
\end{bmatrix}$$

Why should we learn Matrices?

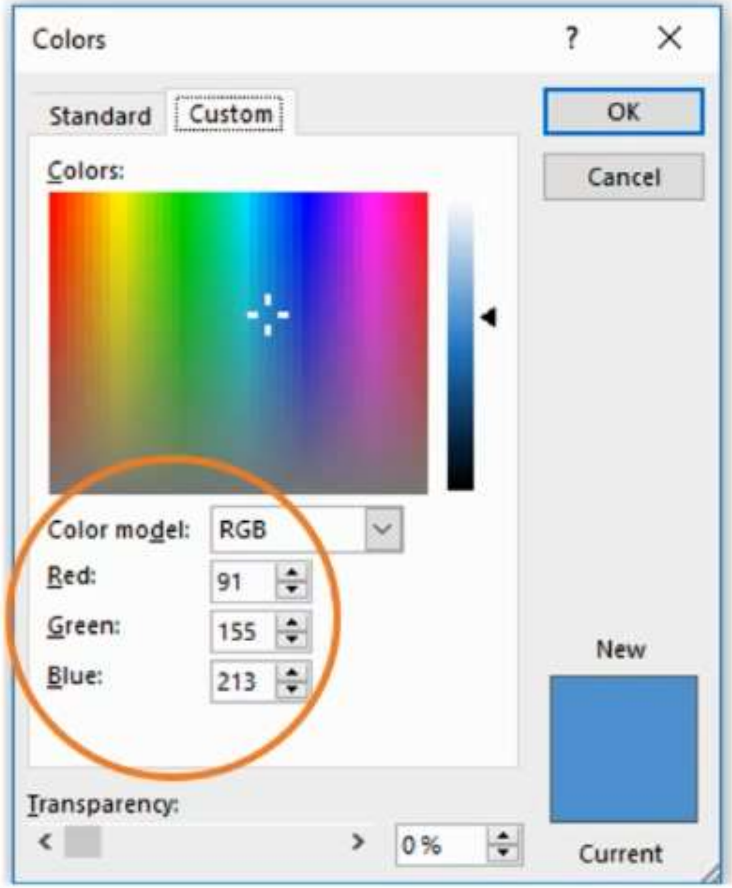
$$
\begin{bmatrix}
7.4 & 0.7 & 0 & 1.9 & 0.076 & 11 & 34 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.8 & 0.88 & 0 & 2.6 & 0.098 & 25 & 67 & 0.9968 & 3.2 & 0.68 & 9.8 & 5 \\
7.8 & 0.76 & 0.04 & 2.3 & 0.092 & 15 & 54 & 0.997 & 3.26 & 0.65 & 9.8 & 5 \\
11.2 & 0.28 & 0.56 & 1.9 & 0.075 & 17 & 60 & 0.998 & 3.16 & 0.58 & 9.8 & 6 \\
7.4 & 0.7 & 0 & 1.9 & 0.076 & 11 & 34 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.4 & 0.66 & 0 & 1.8 & 0.075 & 13 & 40 & 0.9978 & 3.51 & 0.56 & 9.4 & 5 \\
7.9 & 0.6 & 0.06 & 1.6 & 0.069 & 15 & 59 & 0.9964 & 3.3 & 0.46 & 9.4 & 6 \\
7.3 & 0.65 & 0 & 1.2 & 0.065 & 15 & 21 & 0.9946 & 3.39 & 0.47 & 10 & 7 \\
7.8 & 0.58 & 0.02 & 2 & 0.073 & 9 & 18 & 0.9968 & 3.36 & 0.57 & 9.5 & 7
\end{bmatrix}
$$

Why should we learn Matrices?

Matrix of Pixels



$$\begin{array}{c} R \\ G \\ B \end{array} \begin{bmatrix} 230 \\ 169 \\ 43 \end{bmatrix}$$



Matrix Addition

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 8 & -1 \\ 5 & -2 & -3 \end{bmatrix}$$

$$X + Y = \begin{bmatrix} \quad \end{bmatrix} = \begin{bmatrix} \quad \end{bmatrix}$$

Matrix Addition

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 8 & -1 \\ 5 & -2 & -3 \end{bmatrix}$$

$$X + Y = \begin{bmatrix} 2+1 & 3+8 & 4+(-1) \\ 1+5 & 6+(-2) & 7+(-3) \end{bmatrix} = \begin{bmatrix} 3 & 11 & 3 \\ 6 & 4 & 4 \end{bmatrix}$$

Matrix Subtraction

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 8 & -1 \\ 5 & -2 & -3 \end{bmatrix}$$

$$X - Y = \begin{bmatrix} \end{bmatrix} = \begin{bmatrix} \end{bmatrix}$$

Matrix Subtraction

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 8 & -1 \\ 5 & -2 & -3 \end{bmatrix}$$

$$X - Y = \begin{bmatrix} 2 - 1 & 3 - 8 & 4 - (-1) \\ 1 - 5 & 6 - (-2) & 7 - (-3) \end{bmatrix} = \begin{bmatrix} 1 & -5 & 5 \\ -4 & 8 & 10 \end{bmatrix}$$

Matrix Multiplication

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

Scalar Multiplication

$$2 * X = 2 * \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

Matrix Multiplication

$$X \,.\, A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

Matrix Multiplication – Scalar

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

$$2 * X = 2 * \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

Matrix Multiplication – Scalar

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

$$2 * X = 2 * \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & 6 & 8 \\ 2 & 12 & 14 \end{bmatrix}$$

Let's see an example

	Average Price
Sports Shoes	$ 40
Formal	$ 30
Sandals	$ 20

	2016	2017	2018
Sports Shoes	2	3	3
Formal	3	4	3
Sandals	6	8	9

	2016	2017	2018
Sports Shoes	2 * 40	3 * 40	3 * 40
Formal	3 * 30	4 * 30	3 * 30
Sandals	6 * 20	8 * 20	9 * 20

	2016	2017	2018
Sports Shoes	80	120	120
Formal	90	120	90
Sandals	120	160	180
Total	**290**	**400**	**390**

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

$$2 \times \boxed{\underline{3} \qquad\qquad \underline{3}} \times 2$$

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

$$X \, . \, A =$$

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

$$X \,.\, A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

$$X \,.\, A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} \quad \end{bmatrix}$$

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix}$$

$$X \,.\, A = \begin{bmatrix} 2 & 3 & 4 \\ \color{red}{1} & \color{red}{6} & \color{red}{7} \end{bmatrix} \begin{bmatrix} 2 & \color{red}{3} \\ 1 & \color{red}{6} \\ 2 & \color{red}{3} \end{bmatrix} \Rightarrow \begin{bmatrix} 15 & 36 \\ 22 & 60 \end{bmatrix}$$

$$(1*3) + (6*6) + (7*3) = 60$$

Matrix Multiplication – Dot product

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 2 & 3 \end{bmatrix} \qquad X \, . \, A = \begin{bmatrix} 15 & 36 \\ 22 & 60 \end{bmatrix}$$

$$2 \; X \; \boxed{3 \qquad\qquad 3} \; X \; 2 \qquad\qquad 2 \; X \; 2$$

Matrix Division

$$A = \begin{bmatrix} 3 & 4 & -1 \\ 7 & -3 & 2 \end{bmatrix}$$

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

$$\frac{A}{X} =$$

Matrix Division

$$A = \begin{bmatrix} 3 & 4 & -1 \\ 7 & -3 & 2 \end{bmatrix} \qquad X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

$$\frac{A}{X} = \cancel{\frac{\begin{bmatrix} 3 & 4 & -1 \\ 7 & -3 & 2 \end{bmatrix}}{\begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}}}$$

Matrix Division

$$A = \begin{bmatrix} 3 & 4 & -1 \\ 7 & -3 & 2 \end{bmatrix} \qquad X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix}$$

$$\frac{A}{X} = A \,.\, X^{-1}$$

Inverse of a Matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$1/A = \text{Inverse of } A = A^{-1}$$

$$A^{-1} =$$

Inverse of a Matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$1/A = \text{Inverse of } A = A^{-1}$$

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Determinant of a Matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\text{Determinant} = ad - bc$$

Identity Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 1 \\ 2 & 3 & 8 \end{bmatrix}$$

$$A * I = \boxed{}$$

Identity Matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 1 \\ 2 & 3 & 8 \end{bmatrix}$$

$$A * I = A$$

Transposing Vectors/Matrix

- The values are not changing or transforming only their position is.
- Transposing the same vector (object) twice yields the initial vector (Object).
- A 3X1 matrix transposed is a 1X3 matrix.

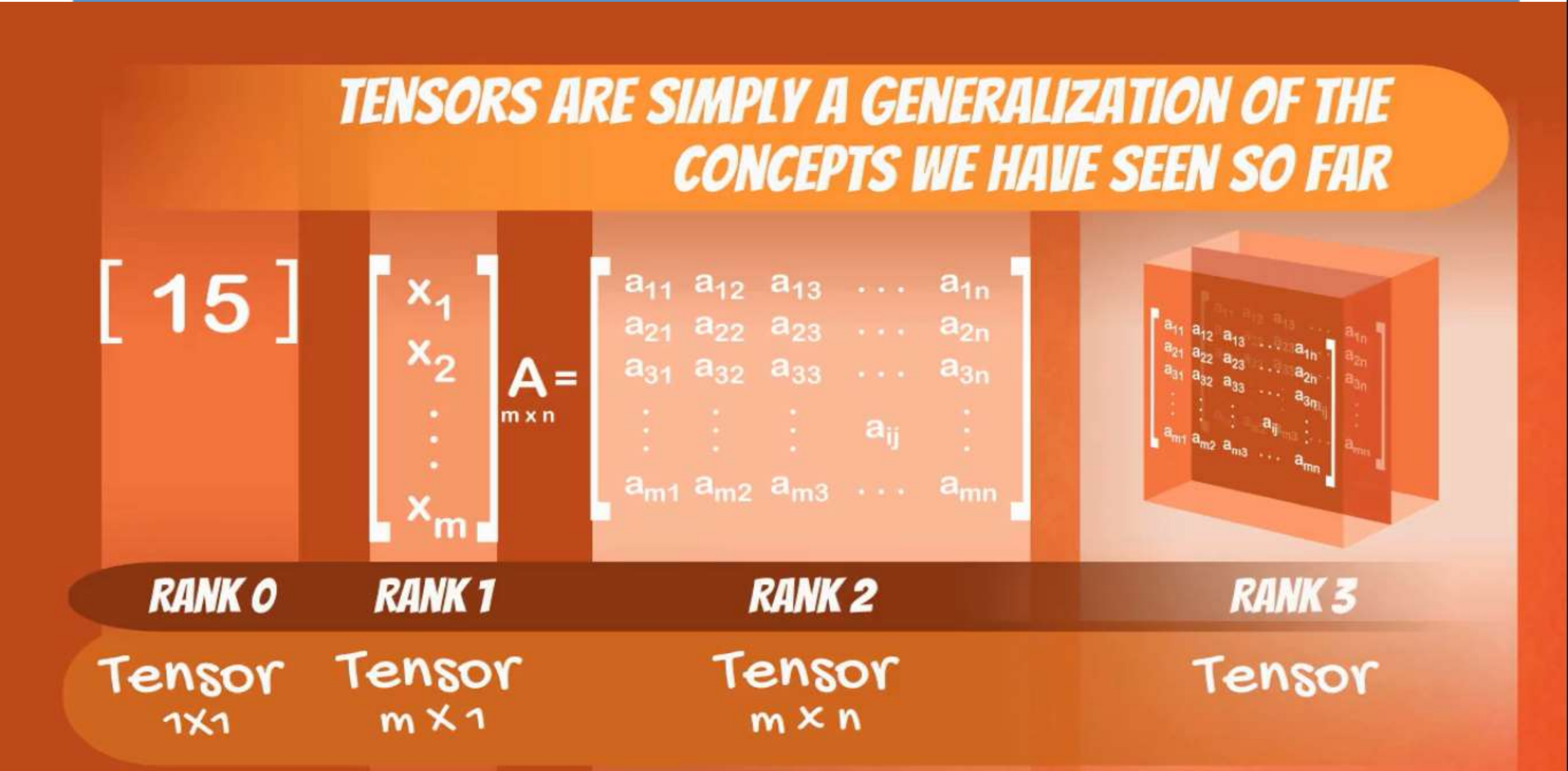
$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \longrightarrow X^T = \begin{bmatrix} \quad \end{bmatrix}$$

Transposing Vectors/Matrix

- The values are not changing or transforming only their position is.
- Transposing the same vector (object) twice yields the initial vector (Object).
- A 3X1 matrix transposed is a 1X3 matrix.

$$X = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 6 & 7 \end{bmatrix} \Longrightarrow X^T = \begin{bmatrix} 2 & 1 \\ 3 & 6 \\ 4 & 7 \end{bmatrix}$$

What is A Tensor ?



jupyter Tensors Last Checkpoint: a minute ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run Markdown

Import the relevant libraries

In [1]:
```
import numpy as np
```

Creating a tensor

In [2]:
```
m1 = np.array([[5,12,6],[-3,0,14]])
m1
```

Out[2]:
```
array([[ 5, 12,  6],
       [-3,  0, 14]])
```

In [3]:
```
m2 = np.array([[9,8,7],[1,3,-5]])
m2
```

Out[3]:
```
array([[ 9,  8,  7],
       [ 1,  3, -5]])
```

In [4]:
```
t = np.array([m1,m2])
```

In [5]:
```
t
```

Out[5]:
```
array([[[ 5, 12,  6],
        [-3,  0, 14]],

       [[ 9,  8,  7],
        [ 1,  3, -5]]])
```

Quiz

Question 1:

How many dimensions does a vector have?

- 0
- 1
- 2
- 3

Question 2:

What will be the result of adding the following two vectors?

$$\vec{V}_1 = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \qquad \vec{V}_2 = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

$$\vec{V}_1 + \vec{V}_2 = \begin{bmatrix} ? \\ ? \end{bmatrix}$$

- ○ -20
  6
- ○ -1
  1
- ○ 8
  2
- ○ 1
  5

Question 3:

What will be the result of multiplication or dot product of a 4 x 3 and 3 x 5 matrices?

- ○ 4 x 5
- ○ 4 x 3
- ○ 3 x 5
- ○ 3 x 3

Question 4:

The matrix division can be achieved using transpose of a matrix. True or False?

- True
- False

Question 5:

Multiplication of a matrix with the identity matrix is similar to multiplying a number by 1. True or False?

- True
- False

Probability & Statistics

Probability

- Probability is a numerical way of describing how likely something is going to happen.

Probability

- Probability is a numerical way of describing how likely something is going to happen.

What is the chance that it will be Head?

Probability

- Probability is a numerical way of describing how likely something is going to happen.

50% chance for both H or T

Probability

- Probability is a numerical way of describing how likely something is going to happen.

Toss it twice.

Probability

- Probability is a numerical way of describing how likely something is going to happen.

What is the chance of both being Heads?

Probability

- Probability is a numerical way of describing how likely something is going to happen.

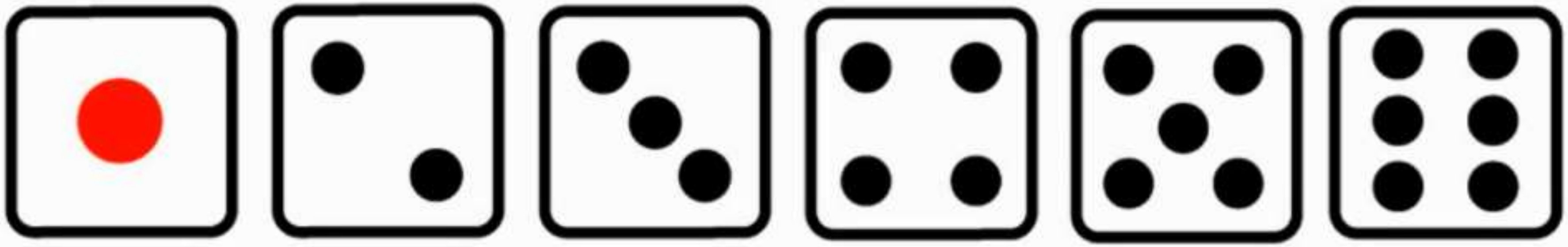
What is the chance of both being Heads?

1^{st}	2^{nd}
H	H
H	T
T	H
T	T

4

$$1/4 = 0.25 \text{ or } 25\%$$

Maximum and Minimum Probability

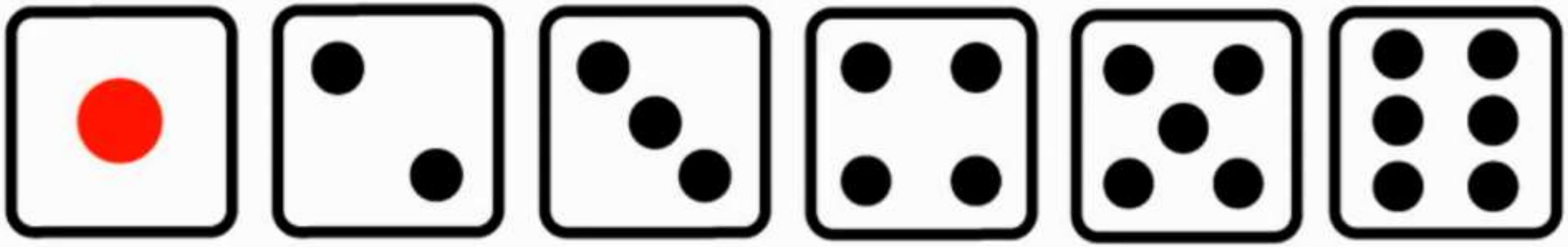


What is the probability of getting an 8?

1 2 3 4 5 6

P(8) = 0/6 = 0.0 or 0%

Maximum and Minimum Probability

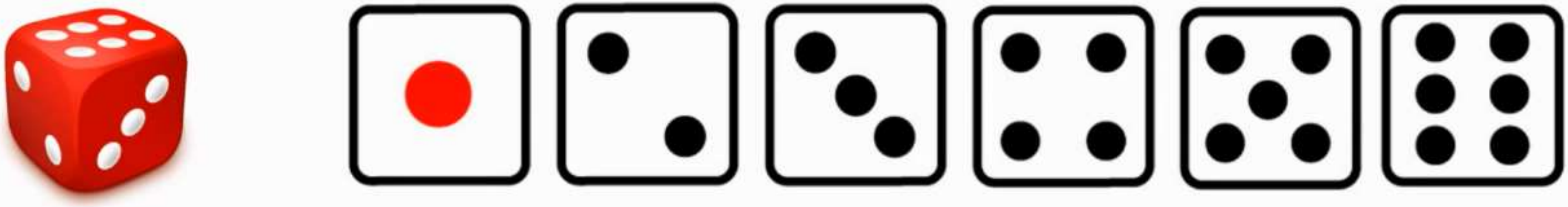


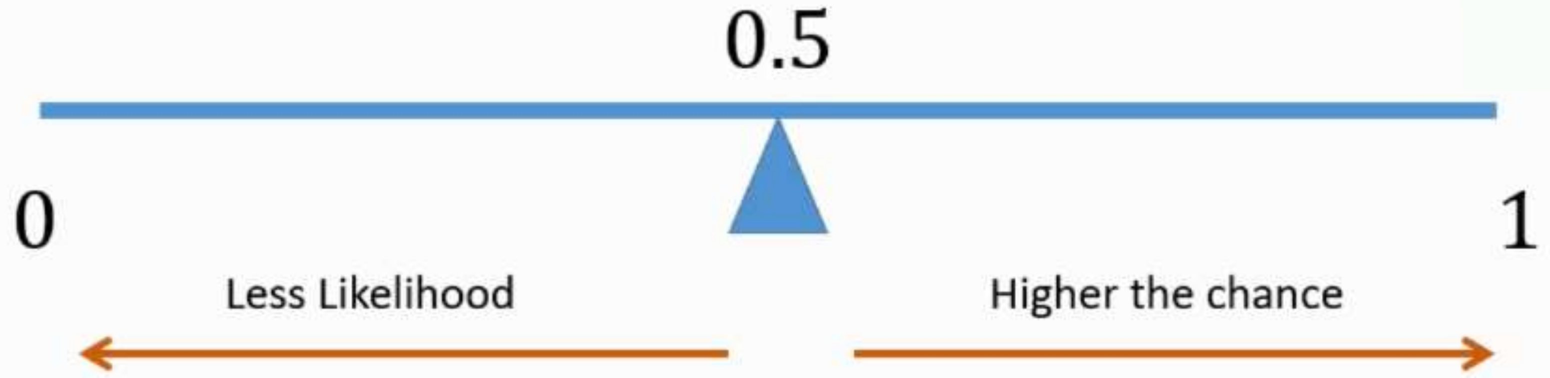
What is the probability of getting a number between 1 and 6?

1 2 3 4 5 6

P(between 1 and 6) = 6/6 = 1.0 or 100%

Maximum and Minimum Probability





Conditional Probability

A measure of the probability of an event (some particular situation occurring) given that another event has occurred.

-- Wikipedia

Conditional Probability in everyday

P(Rain| Cloudy)

P(Rain| Sunny)

Conditional Probability

Rain

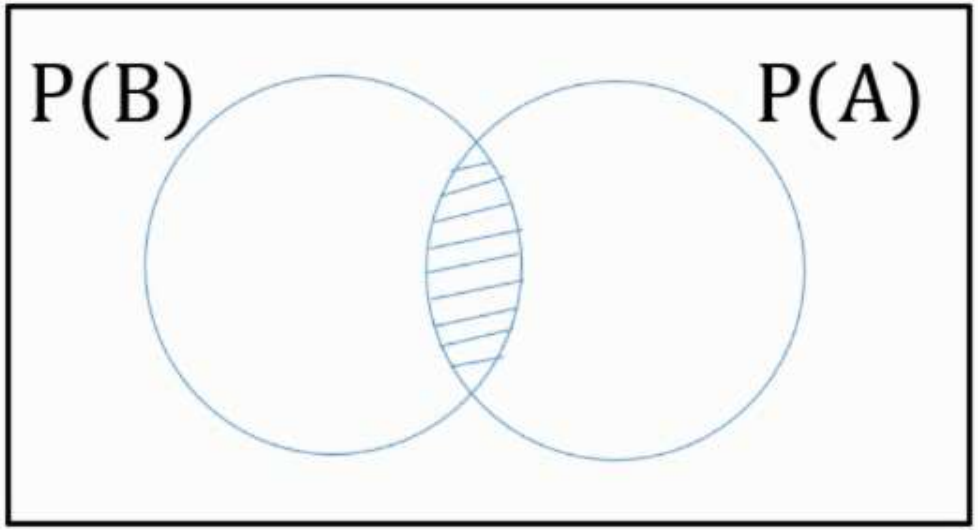
A measure of the probability of an event (some particular situation occurring) given that another event has occurred.

-- Wikipedia

Weather Condition

Conditional Probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$



A → Event whose probability we need to find
e.g. Will it rain?

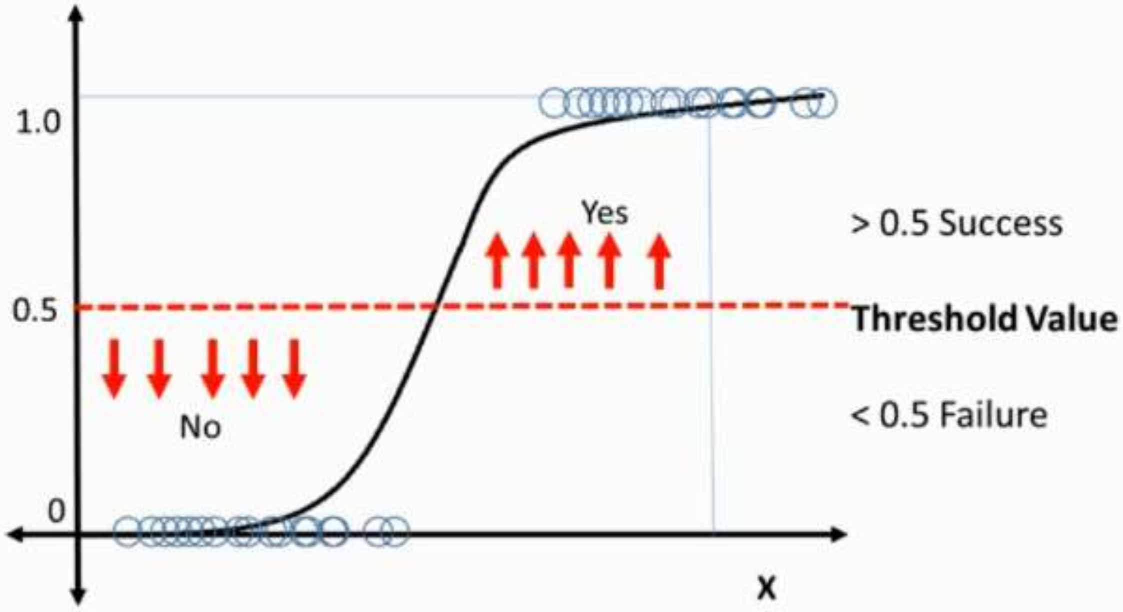
B → Even that has already occurred
e.g. It's already cloudy

How it is used in Data Science and Machine Learning?

Will this customer buy this product?

Will this customer default the loan?

Will the loan of this customer be approved?



Descriptive Statistics

Descriptive Statistics

- Descriptive statistics describes or summarizes a particular set of data.
- It gives an idea of the data in terms of:
  - The mean, mode, median, and range
  - Variance and standard deviation
  - Skewness
  - Count, maximum, and minimum

Inferential Statistics

Inferential statistics uses a sample of the population to make comparisons or predictions on the population.

Example: You have survey data on the model of phone a person would like to purchase. Using inferential statistics, you can estimate that around 70% of people prefer any type of android phone.

Understanding The Variables Using a Dataset

Loan_ID	Gender	Married	Dependents	Self_Employed	Income	LoanAmt	Term	CreditHistory	Property_Area	Status
LP001002	Male	No	0	No	$5,849.00		60	1	Urban	Y
LP001003	Male	Yes	1	No	$4,583.00	$128.00	120	1	Rural	N
LP001005	Male	Yes	0	Yes	$3,000.00	$66.00	60	1	Urban	Y
LP001006	Male	Yes	2	No	$2,583.00	$120.00	60	1	Urban	Y

Types of Variables

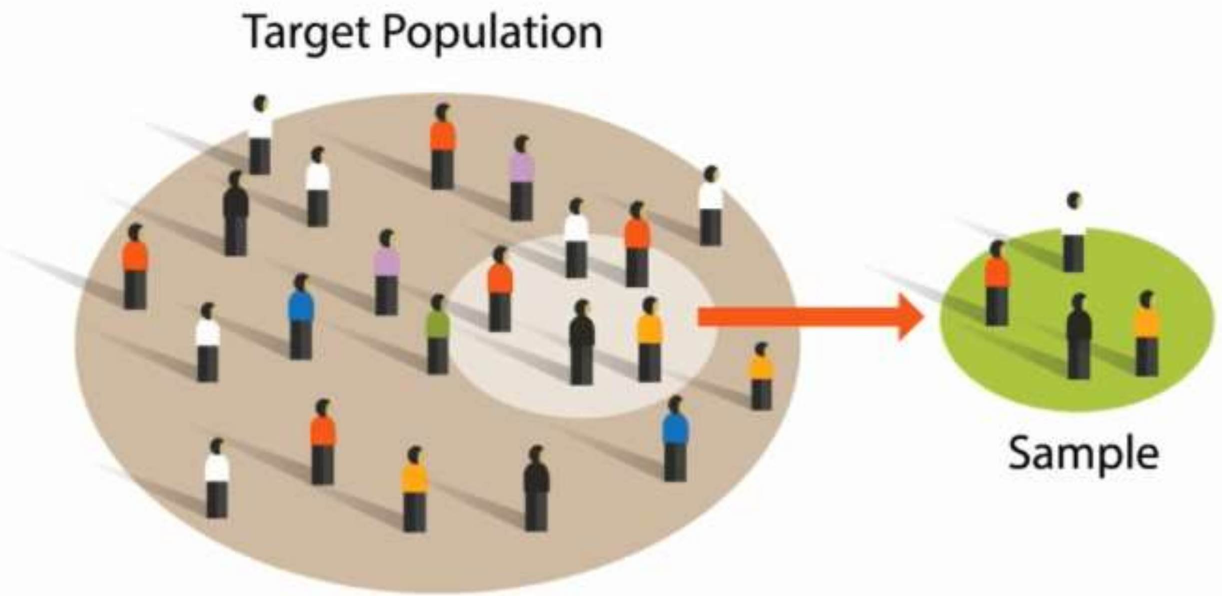
- **Predictor/Independent**
  - Gender
  - Married
  - Dependents
  - Self_Employed
  - Income
  - LoanAmt
  - Term
  - CreditHistory
  - PropertyArea
- **Target/Dependent**
  - Status

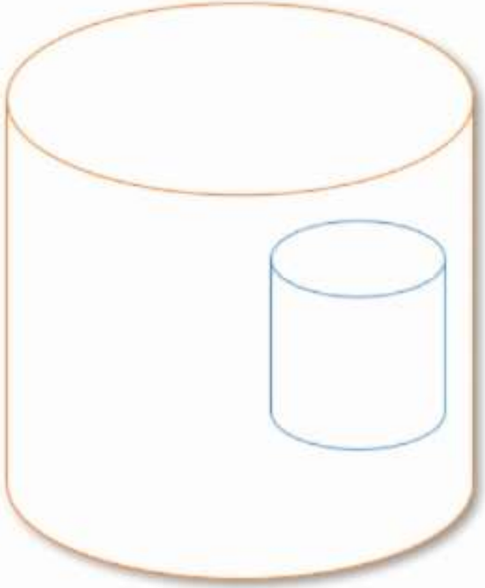
Data Type

- **Character/String**
  - Gender
  - Married
  - Self_Employed
  - Property_Area
  - Status
- **Numeric**
  - Dependents
  - Income
  - LoanAmt
  - Term
  - CreditHistory

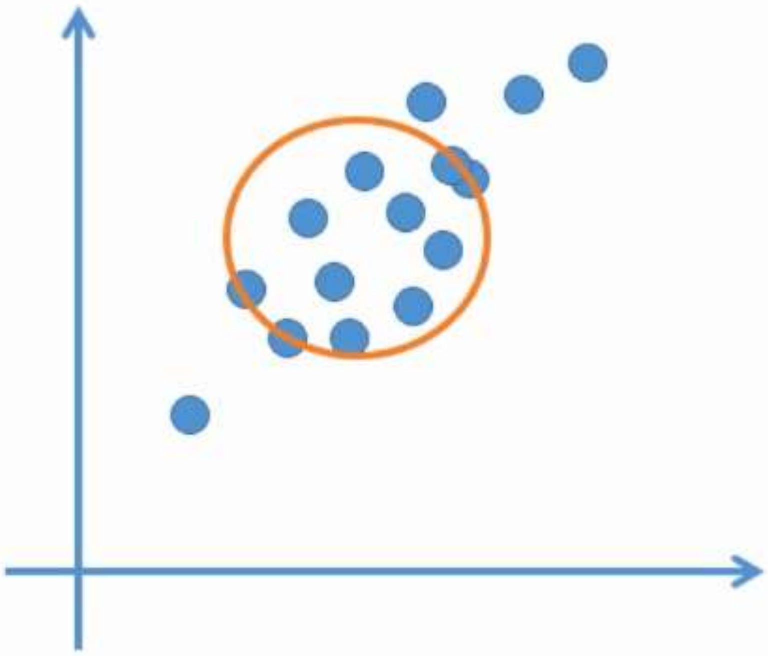
Population Vs. Sample

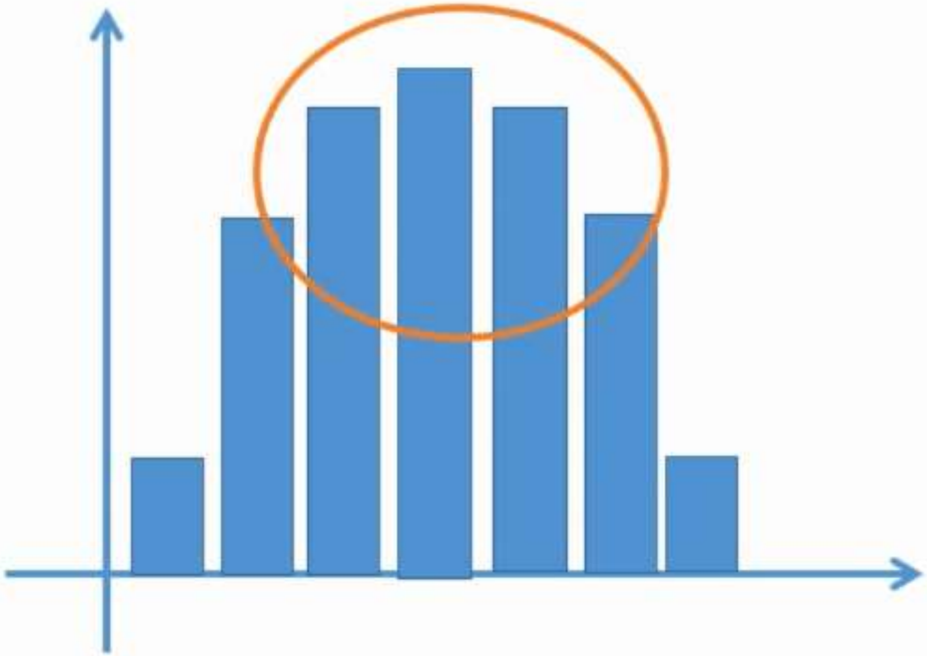
Population and Sample





Central Tendency of Data





Central Tendency

Single value that attempts to describe the whole data using a central point or central location of the data.

Central Tendency

- Mean
- Median
- Mode

Mean

Applicant	Loan Amount
Jitesh	$ 24,000
John	$ 18,000
Frans	$ 34, 000
Danny	$ 40,000
Cecile	$ 24,000
Scott	$ 16,000
Alex	$ 29,000

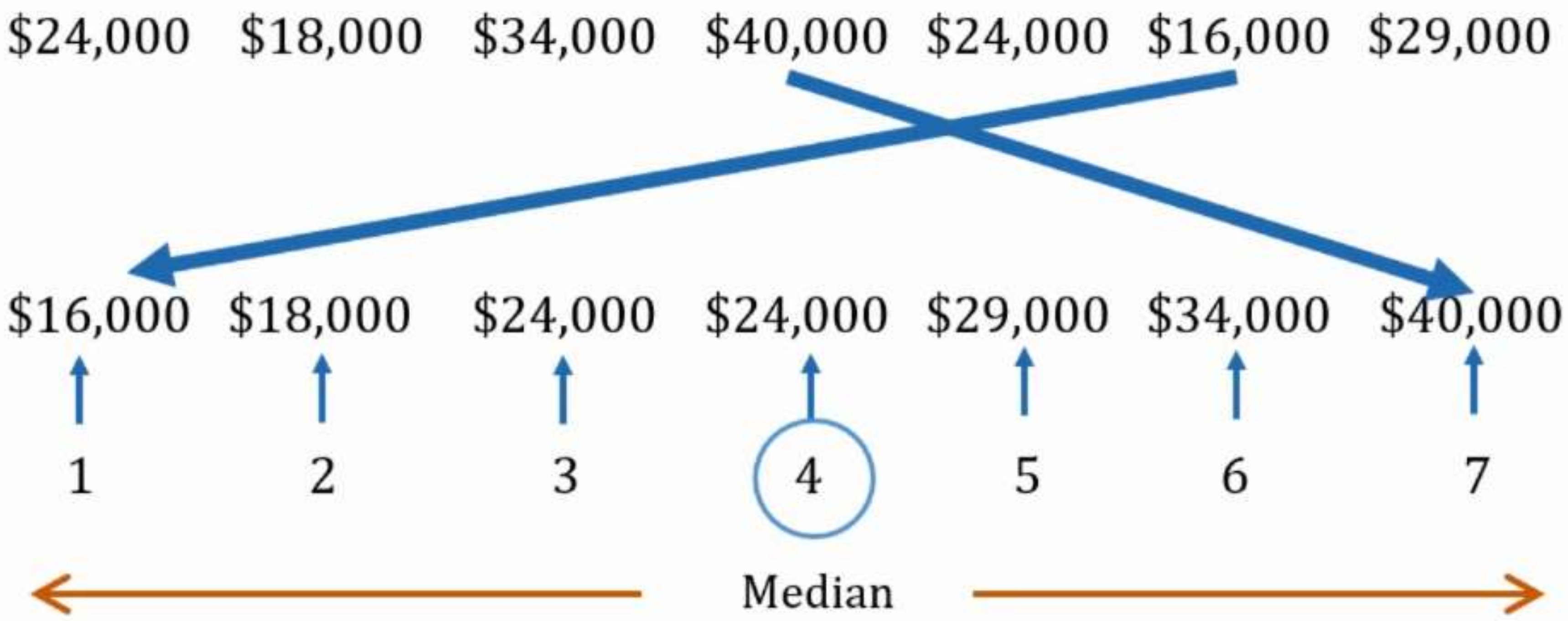
$$\text{Mean} = \frac{24000 + 18000 + 34000 + 40000 + 24000 + 16000 + 29000}{7}$$

$$= \frac{151000}{7}$$

$$\text{Mean} = \$25{,}167$$

Median

Applicant	Loan Amount
Jitesh	$ 24,000
John	$ 18,000
Frans	$ 34, 000
Danny	$ 40,000
Cecile	$ 24,000
Scott	$ 16,000
Alex	$ 29,000



Mode

Applicant	Loan Amount
Jitesh	$ 24,000
John	$ 18,000
Frans	$ 34, 000
Danny	$ 40,000
Cecile	$ 24,000
Scott	$ 16,000
Alex	$ 29,000

$$\text{Mode} = \$24{,}000$$

Outliers

Experience	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 19,000
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

Effect of Outliers

Experience	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 7,556
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

$ 6,915 ← Mean → $ 7,678

← Median →

Experience	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 19,000
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

Effect of Outliers

Experience	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 7,556
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

$ 6,915 ← Mean → $ 7,678

$ 7,200 ← Median → $ 7,200

Experience	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 19,000
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

Central Tendency

Spread in Data

Spread in Data

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20
Total	**140**

$\text{Mean} = 20$

$\text{Median} = 20$

Day	Temperature
1	22
2	23
3	21
4	18
5	19
6	17
7	20
Total	**140**

$\text{Mean} = 20$

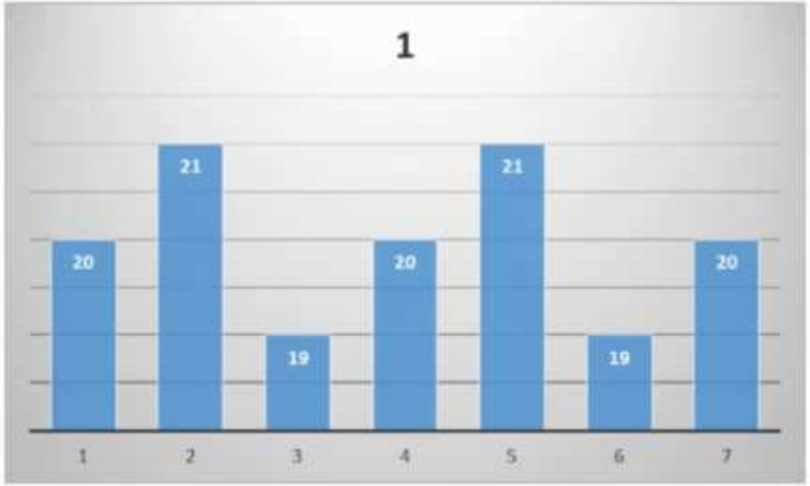
$\text{Median} = 20$

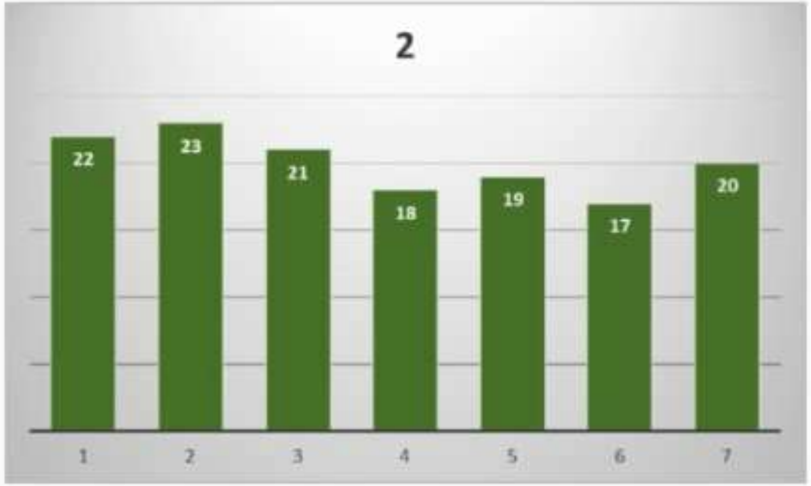
Day	Temperature
1	12
2	11
3	13
4	20
5	24
6	29
7	31
Total	**140**

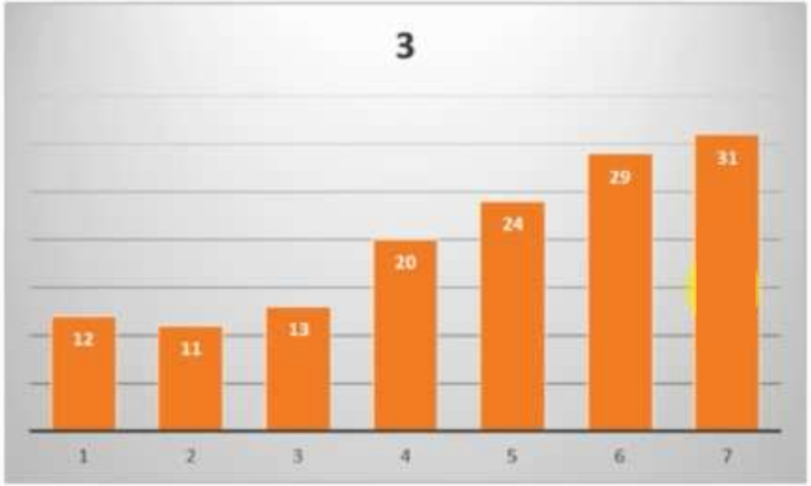
$\text{Mean} = 20$

$\text{Median} = 20$

Spread in Data







$\text{Mean} = 20$

$\text{Median} = 20$

$\text{Mean} = 20$

$\text{Median} = 20$

$\text{Mean} = 20$

$\text{Median} = 20$

Spread in Data

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20
Total	**140**

$$\text{Mean} = 20$$

$$\text{Median} = 20$$

Day	Temperature
1	22
2	23
3	21
4	18
5	19
6	17
7	20
Total	**140**

$$\text{Mean} = 20$$

$$\text{Median} = 20$$

Day	Temperature
1	12
2	11
3	13
4	20
5	24
6	29
7	31
Total	**140**

$$\text{Mean} = 20$$

$$\text{Median} = 20$$

Spread in Data

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20
Total	**140**

2

Day	Temperature
1	22
2	23
3	21
4	18
5	19
6	17
7	20
Total	**140**

6

Day	Temperature
1	12
2	11
3	13
4	20
5	24
6	29
7	31
Total	**140**

$\text{Mean} = 20$

$\text{Median} = 20$

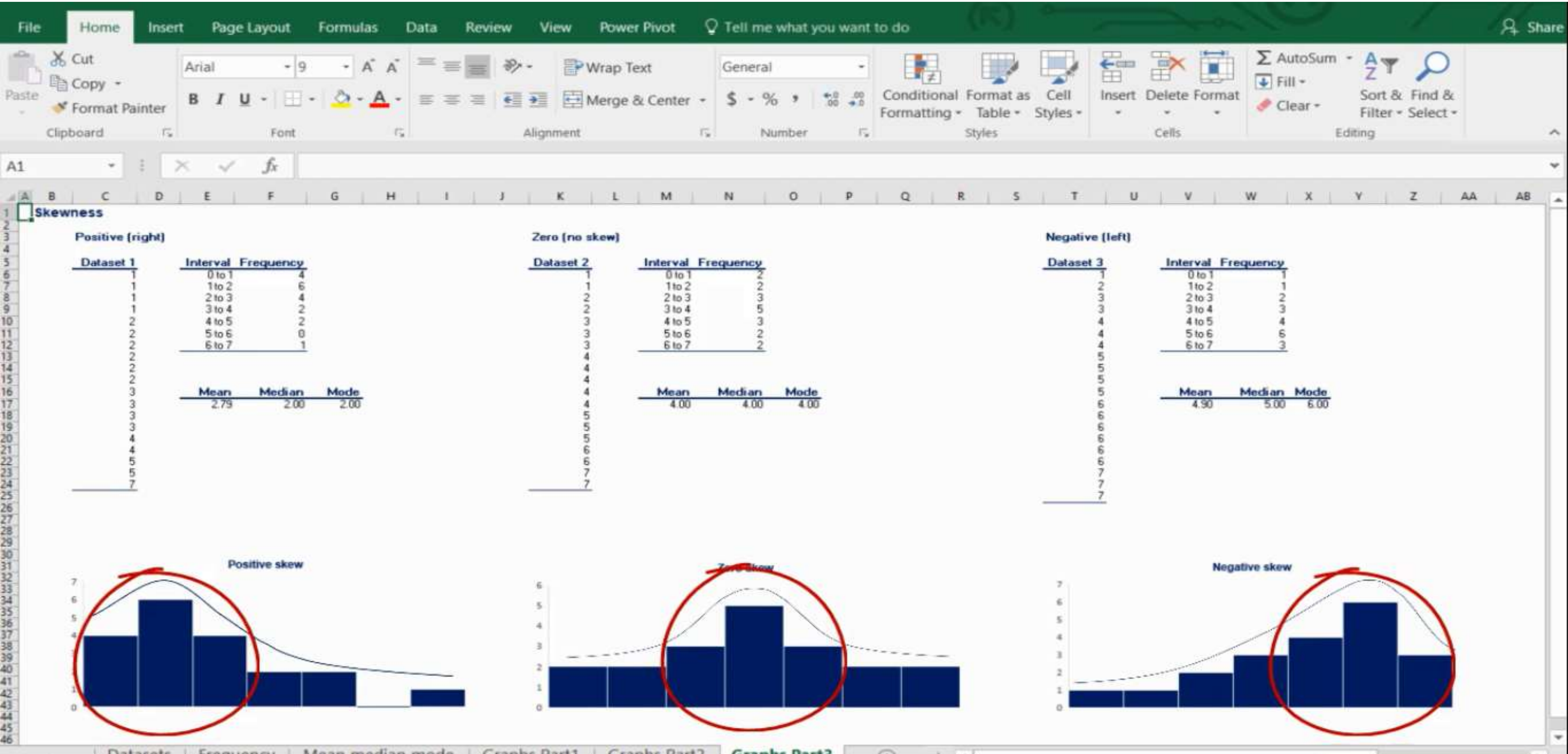
$\text{Mean} = 20$

$\text{Median} = 20$

$\text{Mean} = 20$

$\text{Median} = 20$

Measure of Asymmetry



Skewness

Positive (right)

Dataset 1
1
1
1
1
2
2
2
2
2
2
3
3
3
3
4
4
5
5
7

Interval	Frequency
0 to 1	4
1 to 2	6
2 to 3	4
3 to 4	2
4 to 5	2
5 to 6	0
6 to 7	1

Mean	Median	Mode
2.79	2.00	2.00

Zero (no skew)

Dataset 2
1
1
2
2
3
3
3
4
4
4
4
4
5
5
5
6
6
7
7

Interval	Frequency
0 to 1	2
1 to 2	2
2 to 3	3
3 to 4	5
4 to 5	3
5 to 6	2
6 to 7	2

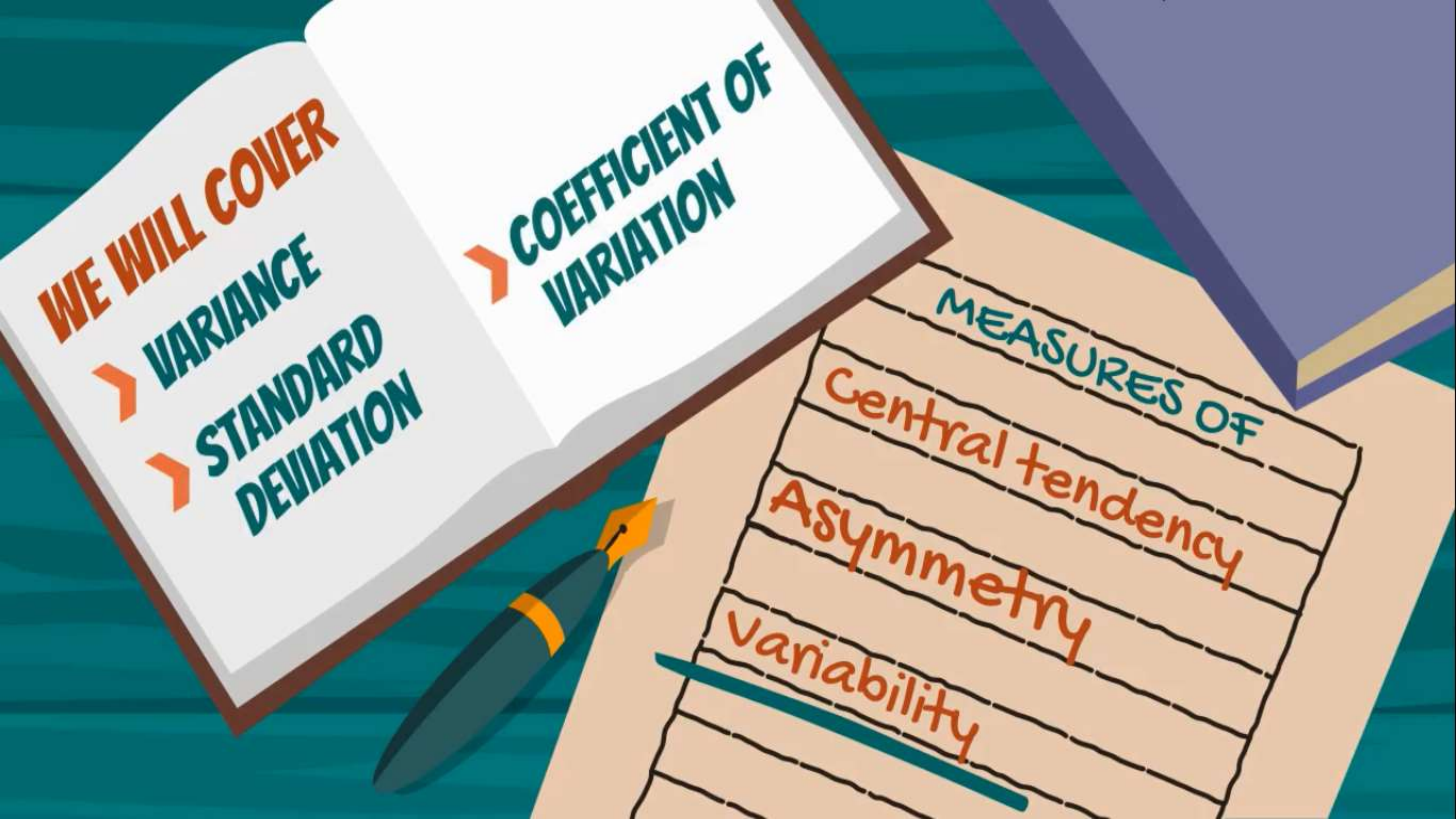
Mean	Median	Mode
4.00	4.00	4.00

Negative (left)

Dataset 3
1
2
3
3
4
4
4
5
5
5
5
6
6
6
6
6
6
7
7
7

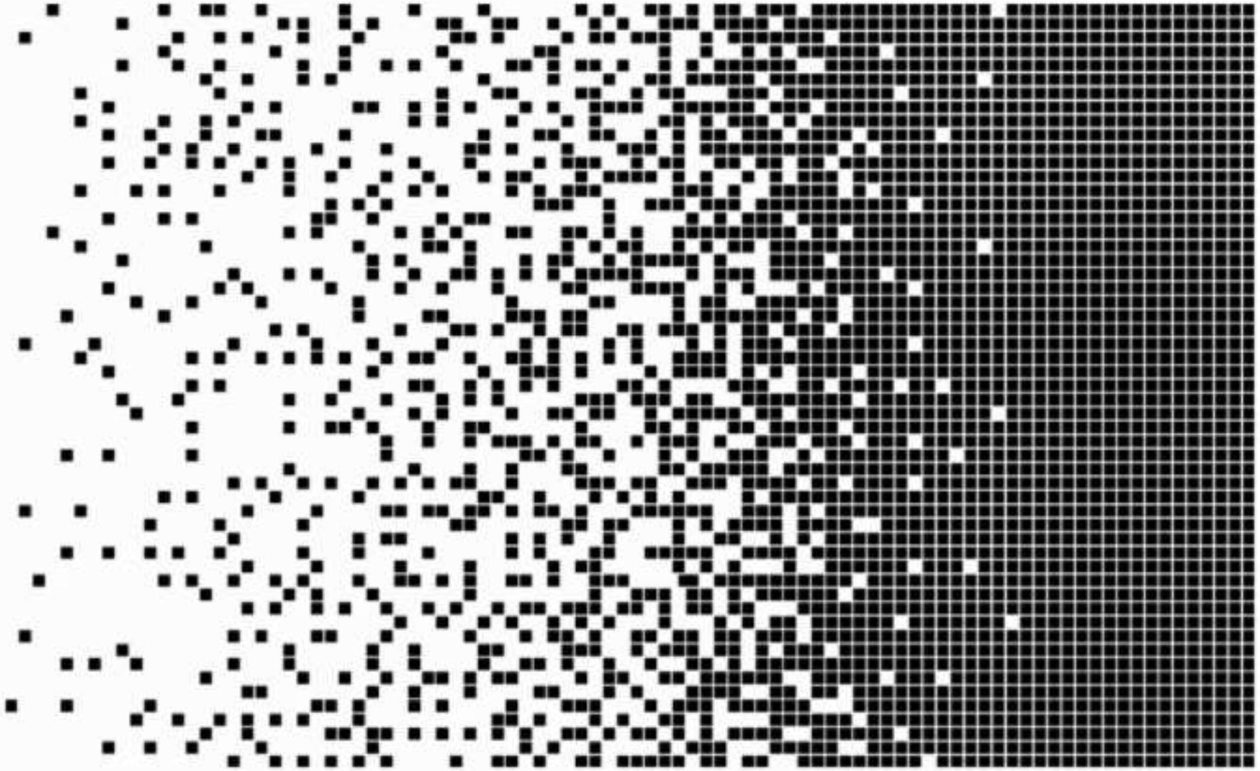
Interval	Frequency
0 to 1	1
1 to 2	1
2 to 3	2
3 to 4	3
4 to 5	4
5 to 6	6
6 to 7	3

Mean	Median	Mode
4.90	5.00	6.00


WE WILL COVER
VARIANCE
STANDARD DEVIATION
COEFFICIENT OF VARIATION
MEASURES OF
Central tendency
Asymmetry
Variability

Measure of Dispersion

- Variance
- Standard Deviation
- Percentile
- Range
- Interquartile range



Mean

$$\frac{\sum_{i=1}^{n} x_i}{n}$$

sample formula

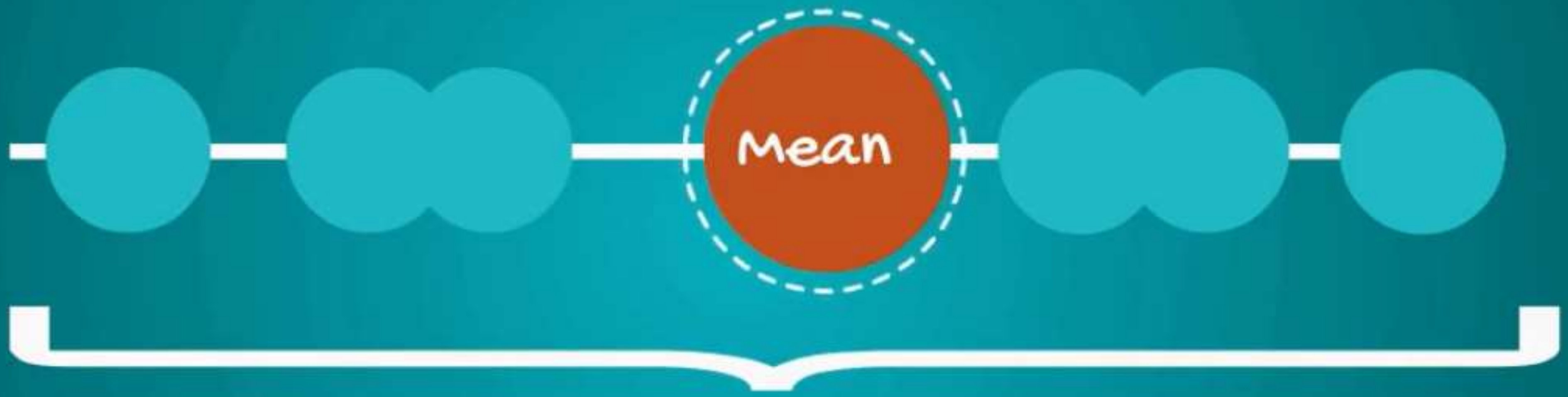
n is the size of the sample

$$\frac{\sum_{i=1}^{N} x_i}{N}$$

population formula

N is the size of the population

VARIANCE



Variance measures the dispersion of a set of data points around their mean

VARIANCE

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$$

population variance

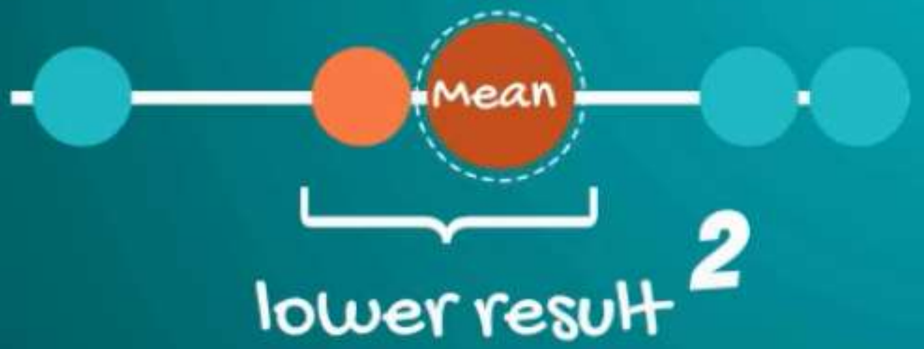
$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

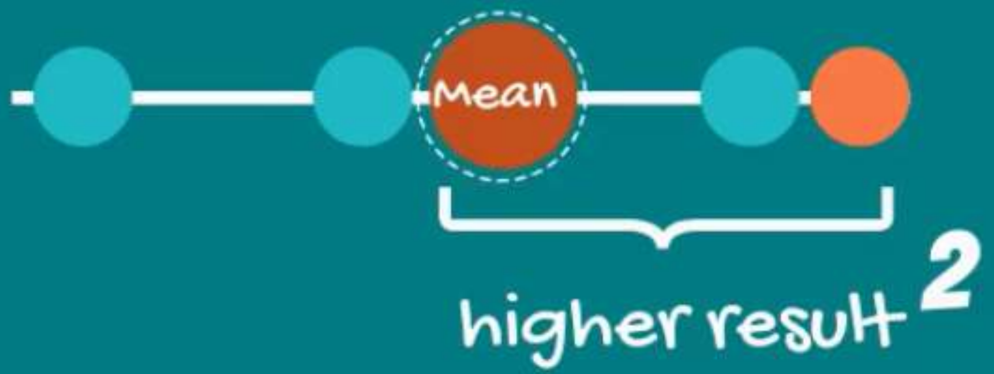
sample variance

VARIANCE

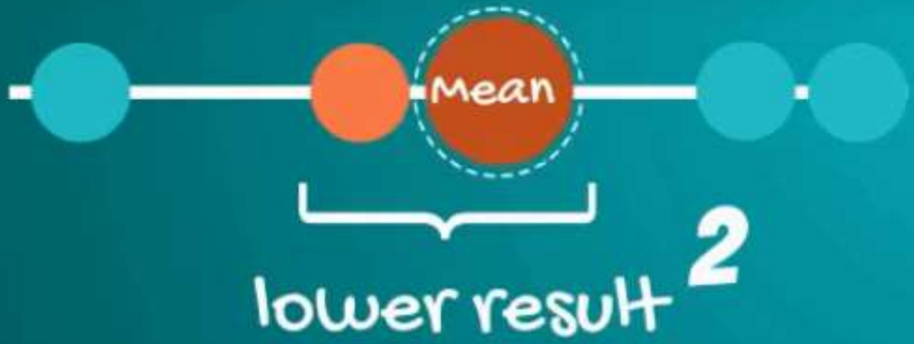
$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

population variance





- Dispersion is non-negative. Non-negative values don't cancel out
- Amplifies the effect of large differences

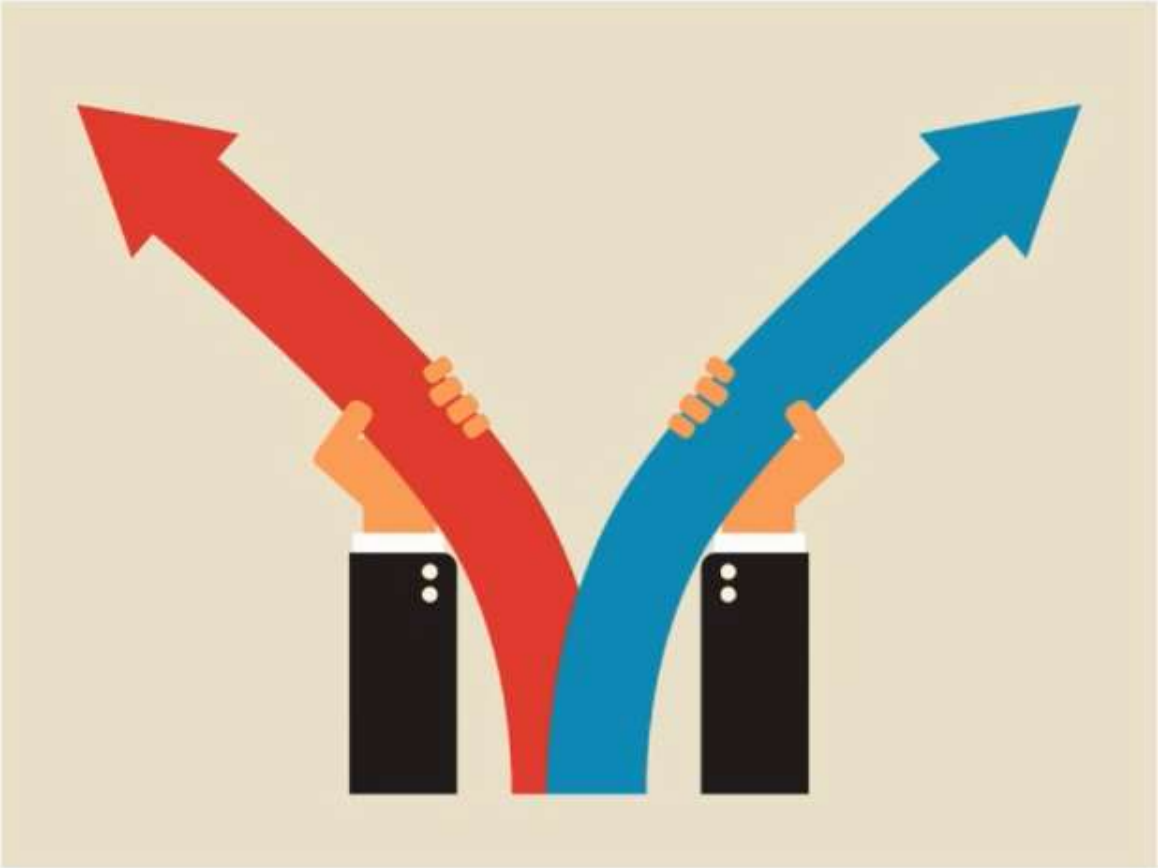


10 SQUARED IS 100

100 SQUARED IS 10,000!

Variance

Average of the **squared difference** of the data from the **Mean**.



Variance and Standard Deviation

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20

Difference
0
1
-1
0
1
-1
0

Difference Squared
0
1
1
0
1
1
0
Total = 4

$$\text{Average} = 4/7 = 0.57$$

Variance σ^2

$$\text{Mean} = 20$$

$$\sigma = 0.7559$$

σ → Standard Deviation

Variance and Standard Deviation

Day	X	$X - \bar{X}$	$(X - \bar{X})^2$
1	20	0	0
2	21	1	1
3	19	-1	1
4	20	0	0
5	21	1	1
6	19	-1	1
7	20	0	0

$$\text{Average} = 4/7 = 0.57$$

$$\text{Variance, } \sigma^2 = 0.57$$

$$\sigma = 0.7559$$

$$\text{Mean} = \bar{X} = 20$$

Variance and Standard Deviation

Day	X	$X - \bar{X}$	$(X - \bar{X})^2$
1	12	-8	64
2	11	-9	81
3	13	-7	49
4	20	0	0
5	24	4	16
6	29	9	81
7	31	11	121
			412

$$\text{Average} = 412/7 = 58.857$$

$$\text{Variance, } \sigma^2 = 58.857$$

$$\sigma = 7.67$$

$$\text{Mean} = \bar{X} = 20$$

Variance and Standard Deviation

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20

$\sigma = 0.7559$

$\text{Mean} = \overline{X} = 20$

Day	Temperature
1	12
2	11
3	13
4	20
5	24
6	29
7	31

$\sigma = 7.67$

$\text{Mean} = \overline{X} = 20$

Airline Arrival

$$\sigma = 10 \text{ Minutes}$$

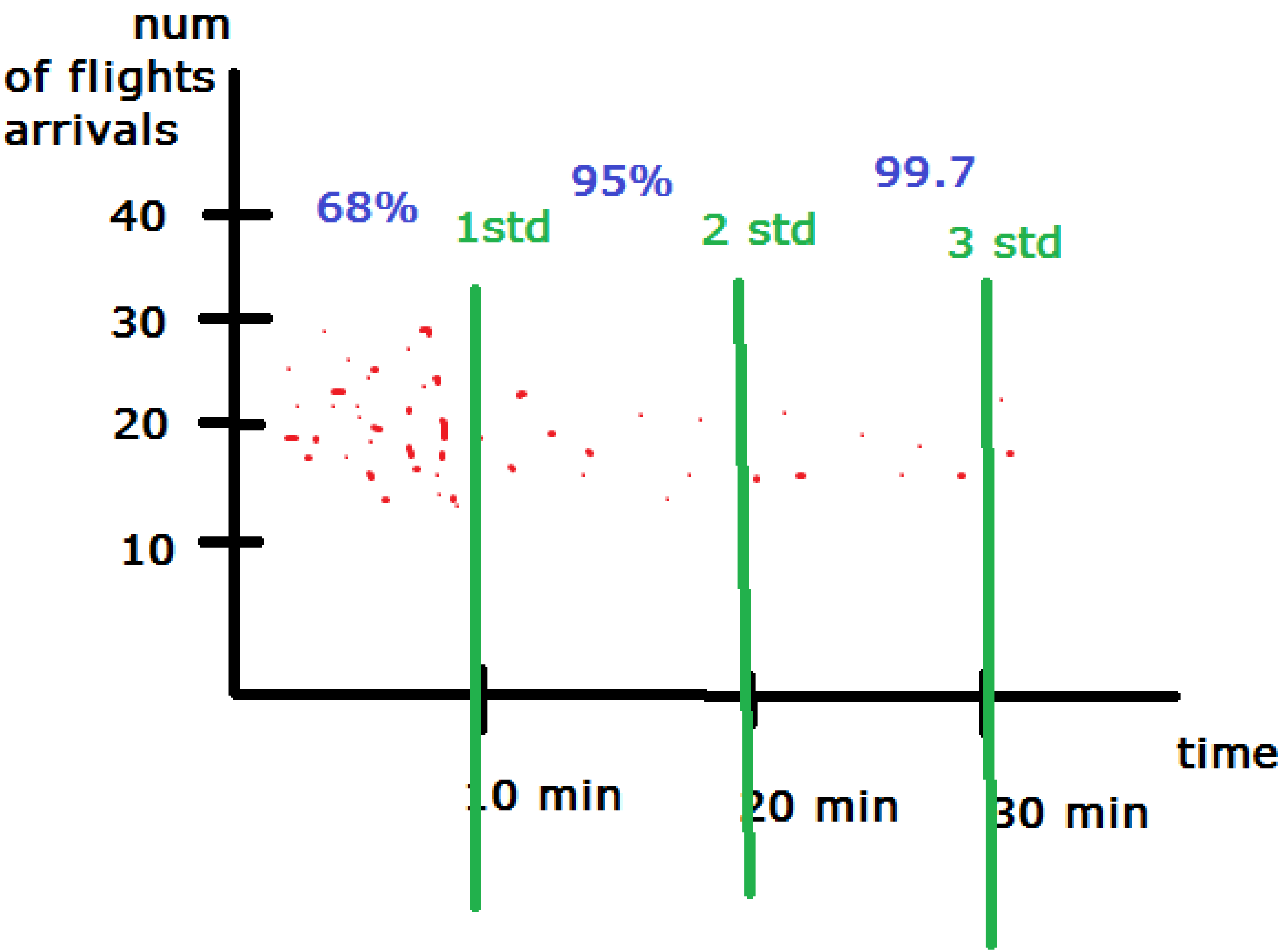
Majority of the flights will arrive in $\pm$ 10 minutes of scheduled time.

10 Minutes

20 Minutes

30 Minutes

Std Deviation from Mean	Data Represented
1σ	68%
2σ	95%
3σ	99.7%


num
of flights
arrivals
40
30
20
10
68%
1std
95%
2 std
99.7
3 std
time
10 min
20 min
30 min

```python
# Basic Descriptive Statistics
# ---------------------------------------

import statistics as st

data = [1, 2, 3, 3]

# Mean calculation
mean = st.mean(data)
median   = st.median(data)
mode     = st.mode(data)
std_dev  = st.stdev(data)
variance = st.variance(data)
```

Percentile


My GMAT percentile score is 91%.
That's a Great Score. You should try for top Universities..

What is Percentile?

The value below which a given percentage of observations in a group of observations falls...

– Wikipedia

Let’s understand it

Percentile

- Arrange the data in an order

- Calculate the percentage of observations or data points below a particular value.

Total Observations * 0.8

15 * 0.8 = 12

Row Number	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 7,556
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

Range

Difference between the highest and lowest value…

Range

Day	Temperature
1	20
2	21
3	19
4	20
5	21
6	19
7	20

Day	Temperature
1	22
2	23
3	21
4	18
5	19
6	17
7	20

Day	Temperature
1	12
2	11
3	13
4	20
5	24
6	29
7	31

$\text{Mean} = 20$
$\text{Range} = 2$

$\text{Mean} = 20$
$\text{Range} = 6$

$\text{Mean} = 20$
$\text{Range} = 20$

Inter Quartile Range (IQR)

1st Quartile

Median

3rd Quartile

Row Number	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 7,556
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

Q3 – Q1
Inter Quartile Range
IQR

$8,810 – $5,147 = $3,663

Inter Quartile Range (IQR)

1st Quartile

Median

3rd Quartile

Row Number	Salary
1	$ 3,725
2	$ 4,155
3	$ 4,627
4	$ 5,147
5	$ 5,718
6	$ 6,347
7	$ 7,039
8	$ 7,210
9	$ 7,423
10	$ 7,556
11	$ 8,369
12	$ 8,810
13	$ 8,940
14	$ 9,200
15	$ 9,458

IQR of 95%

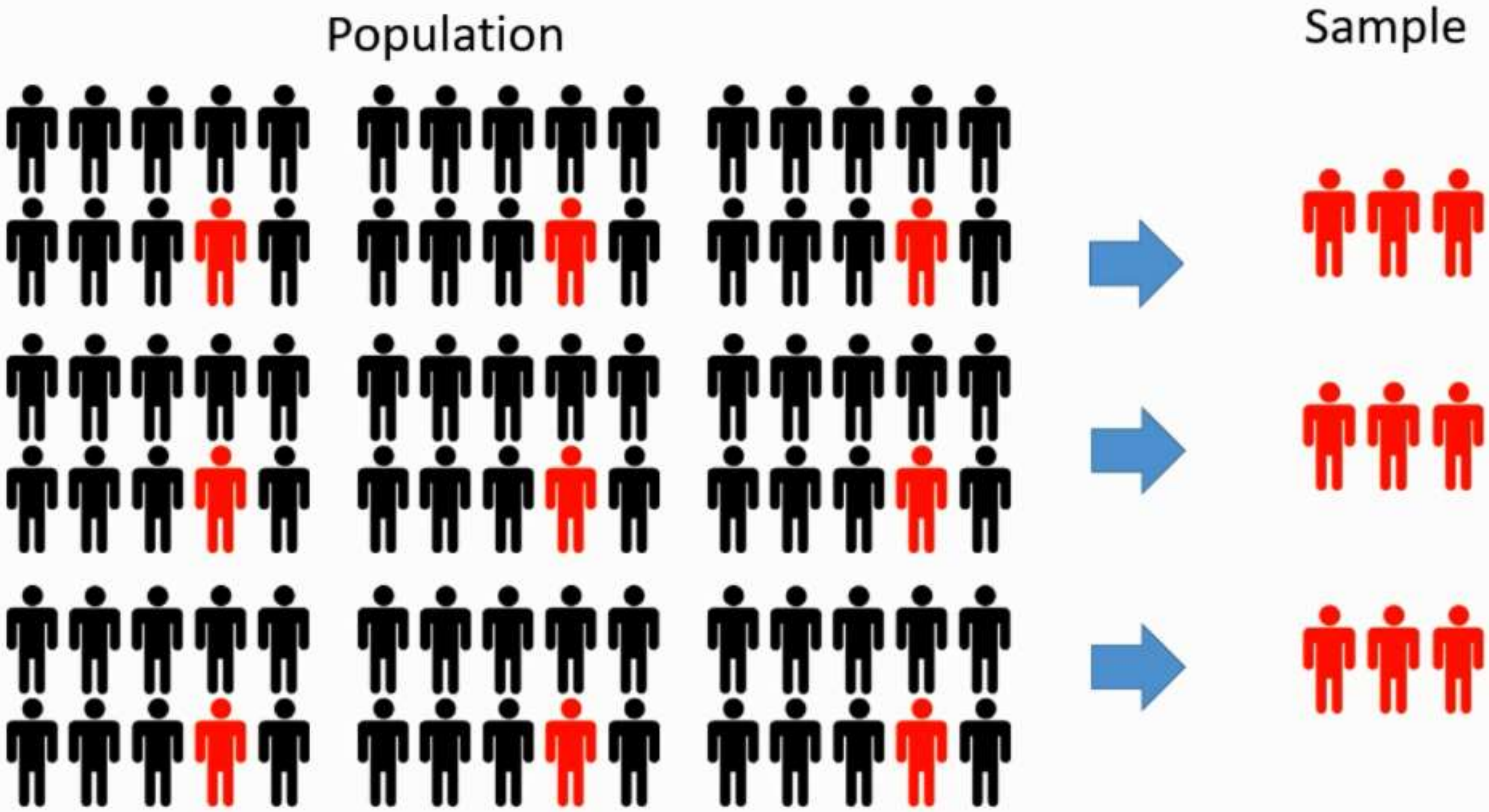

Inferential Statistics
Morris Charts
Line Chart
Area Chart
Bar Chart
Donut Chart
25%
Donut Chart
25%
50%
75%

Inferential Statistics

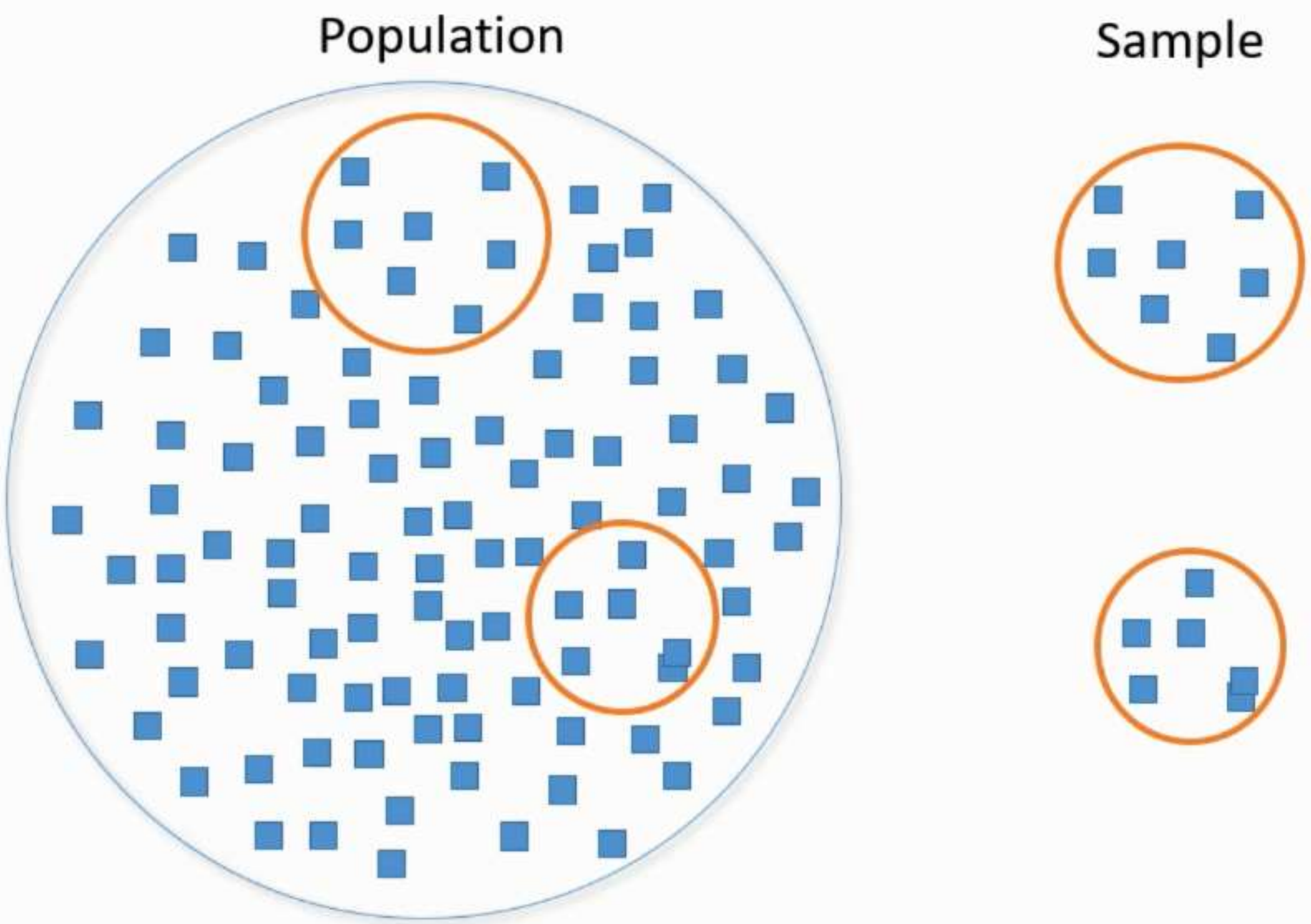
Analyse the <u>random sample</u> of the data from the given population and <u>infer/deduce the properties</u> of the population based on the <u>properties of the sample</u>.

Drawing Inference

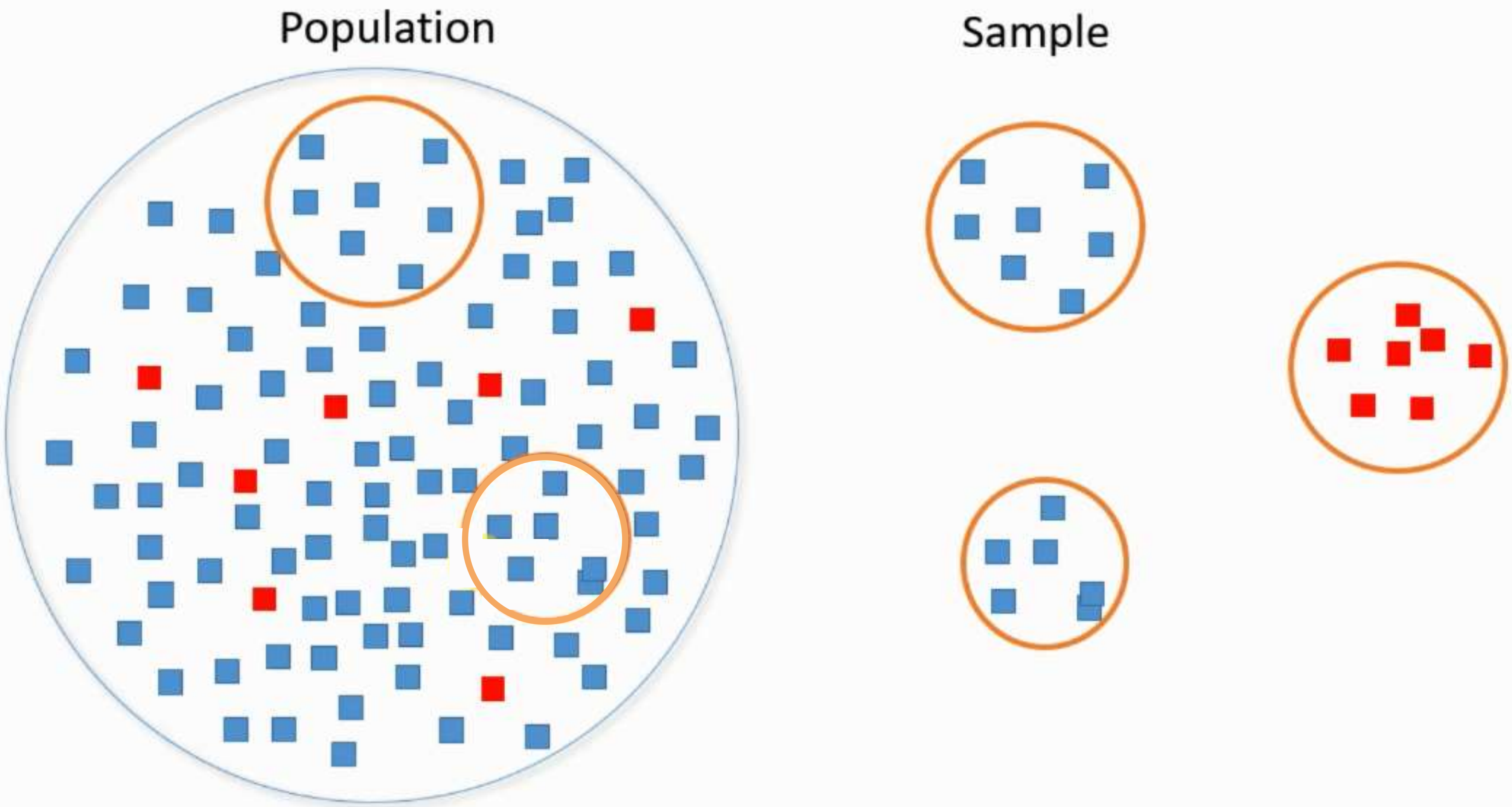
Population and Sample



Population and Sample



Population and Sample



Why to Sample?

Why to Sample?

- In real world situations, it's practically difficult to study the entire population.
- For example, a company that manufactures mobile phones cannot interact with every customer who purchased the product. In such cases, the company selects a sample of the population.
- In order to apply statistics and study the population, the sample must be random.
- A random sample is one in which each sample of a population has an equal chance or probability of being selected. It reduces the chances of human bias, making it highly representative of the population.

The process of selecting samples is called sampling. The units under study from a sample are called sampling units, and the number of units in a sample is called sample size.

Population and Sample

Find out the percentage of oil in the chips manufactured.

Find out the average diameter of the pencils manufactured in a day.

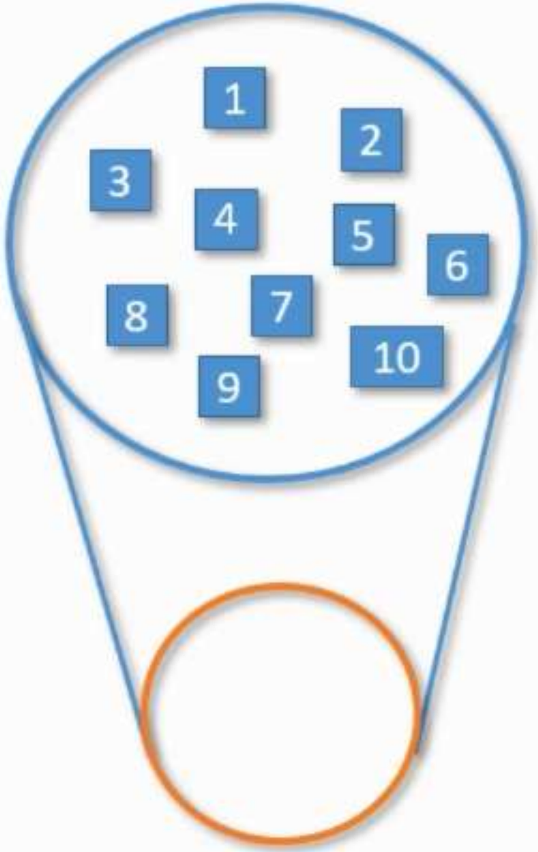
Sampling Techniques

- Simple Random Sampling
- Stratified Sampling

Simple Random Sampling

- Probability Based

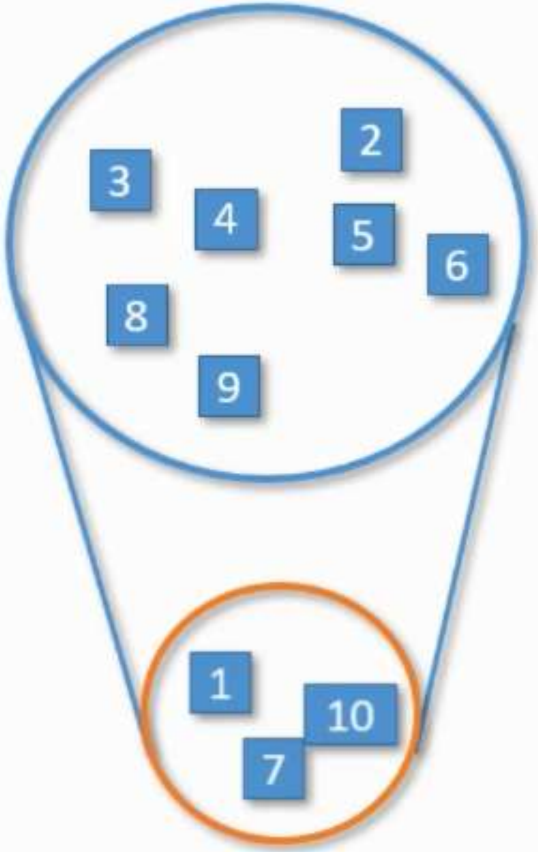
- Every record or item has an equal chance



$P(x) = 1/10$

Simple Random Sampling

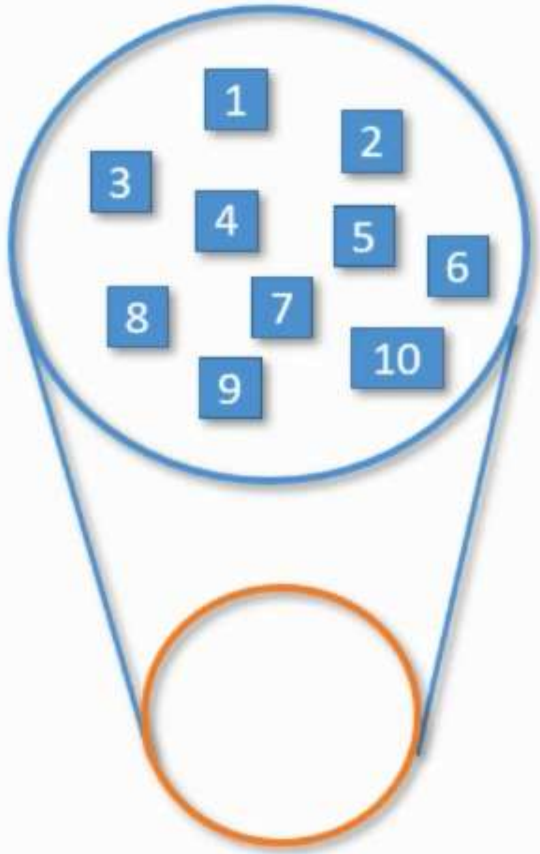
- Probability Based
- Every record or item has an equal chance
- Done without replacement



$P(x) = 1/10$

Stratified Sampling

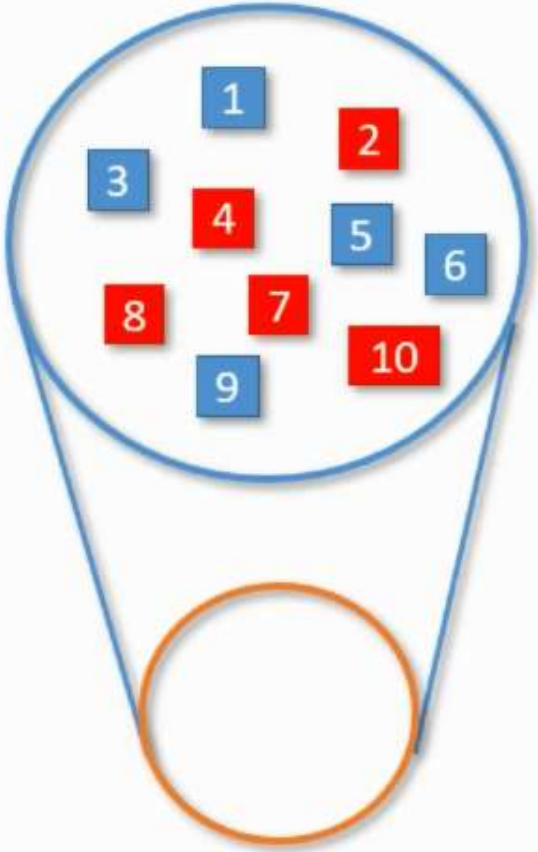
- Records are divided based on the characteristics or <u>Strata</u>



Stratified Sampling

- Records are divided based on the characteristics or <u>Strata</u>

- Records are chosen in equal proportions of Strata



Status

1. Approved
2. Not-Approved

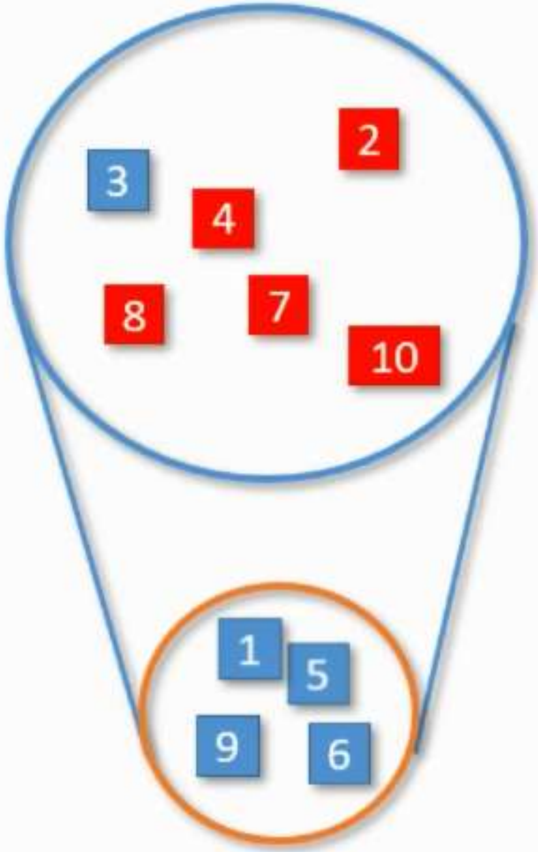
Sample of 40% of records.

Sample size of 4.

Stratified Sampling

- Records are divided based on the characteristics or <u>Strata</u>

- Records are chosen in equal proportions of Strata



Status

1. Approved
2. Not-Approved

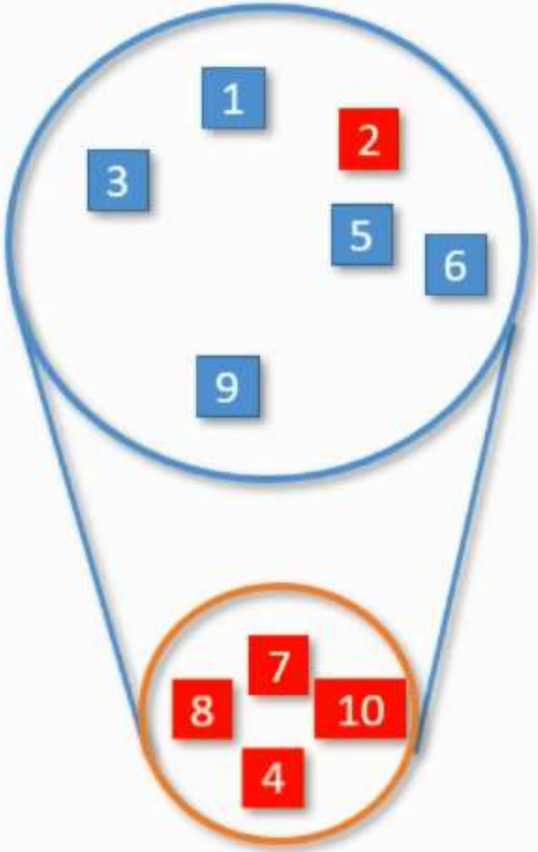
Sample of 40% of records.

Sample size of 4.

Stratified Sampling

- Records are divided based on the characteristics or <u>Strata</u>

- Records are chosen in equal proportions of Strata



Status

1. Approved
2. Not-Approved

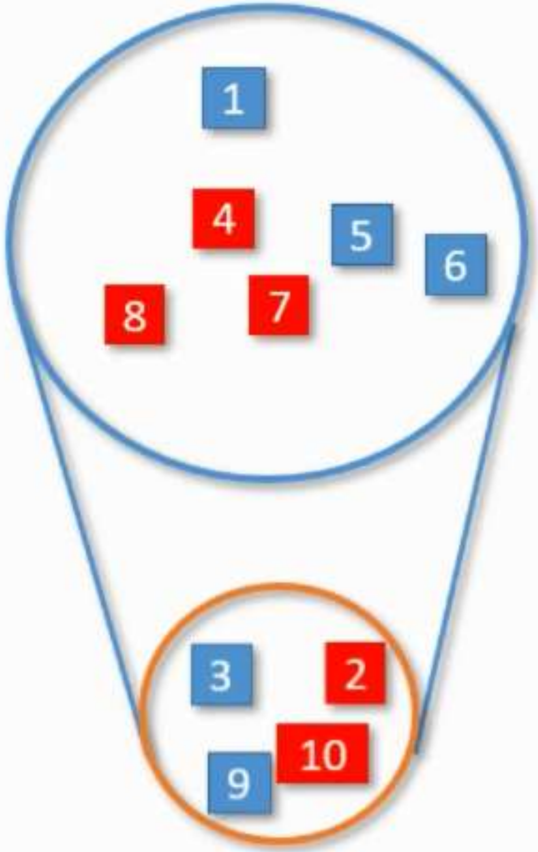
Sample of 40% of records.

Sample size of 4.

Stratified Sampling

- Records are divided based on the characteristics or <u>Strata</u>

- Records are chosen in equal proportions of Strata



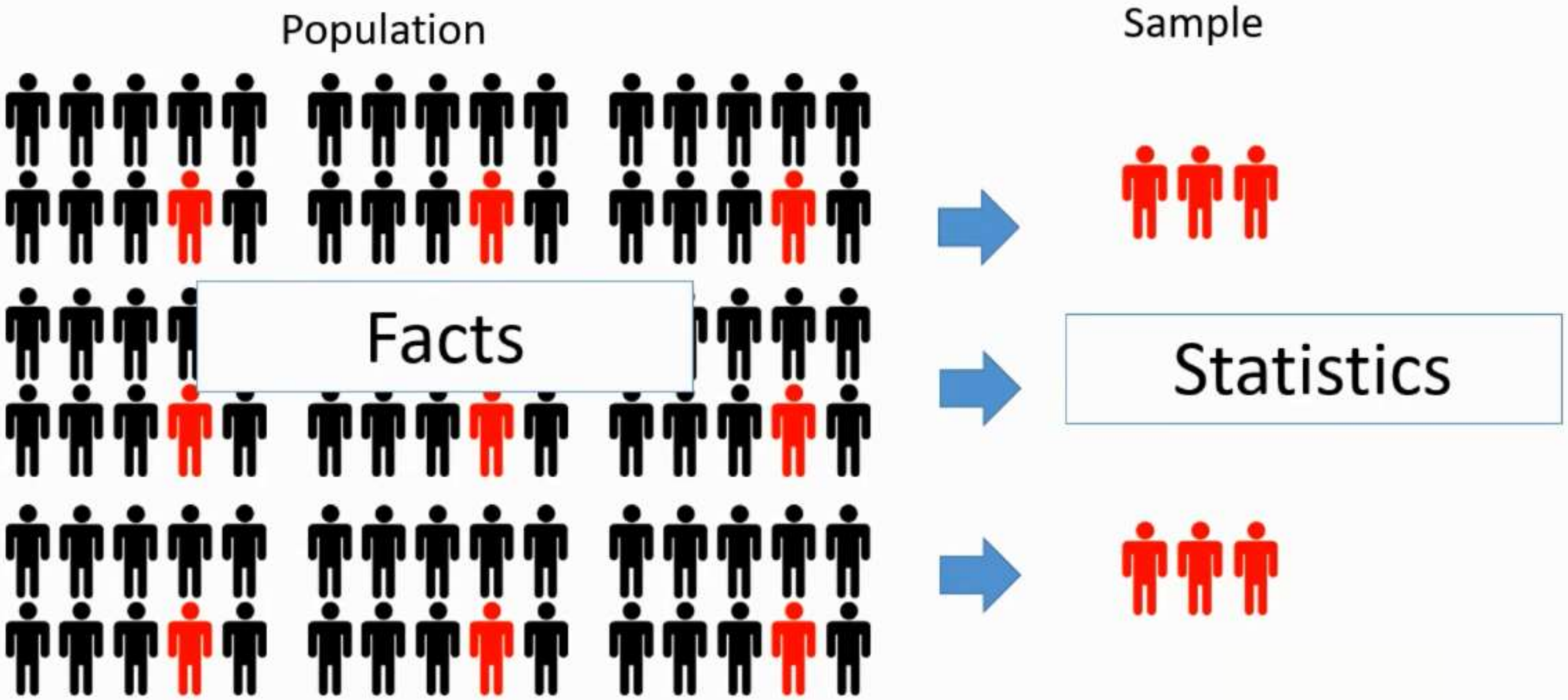
Status

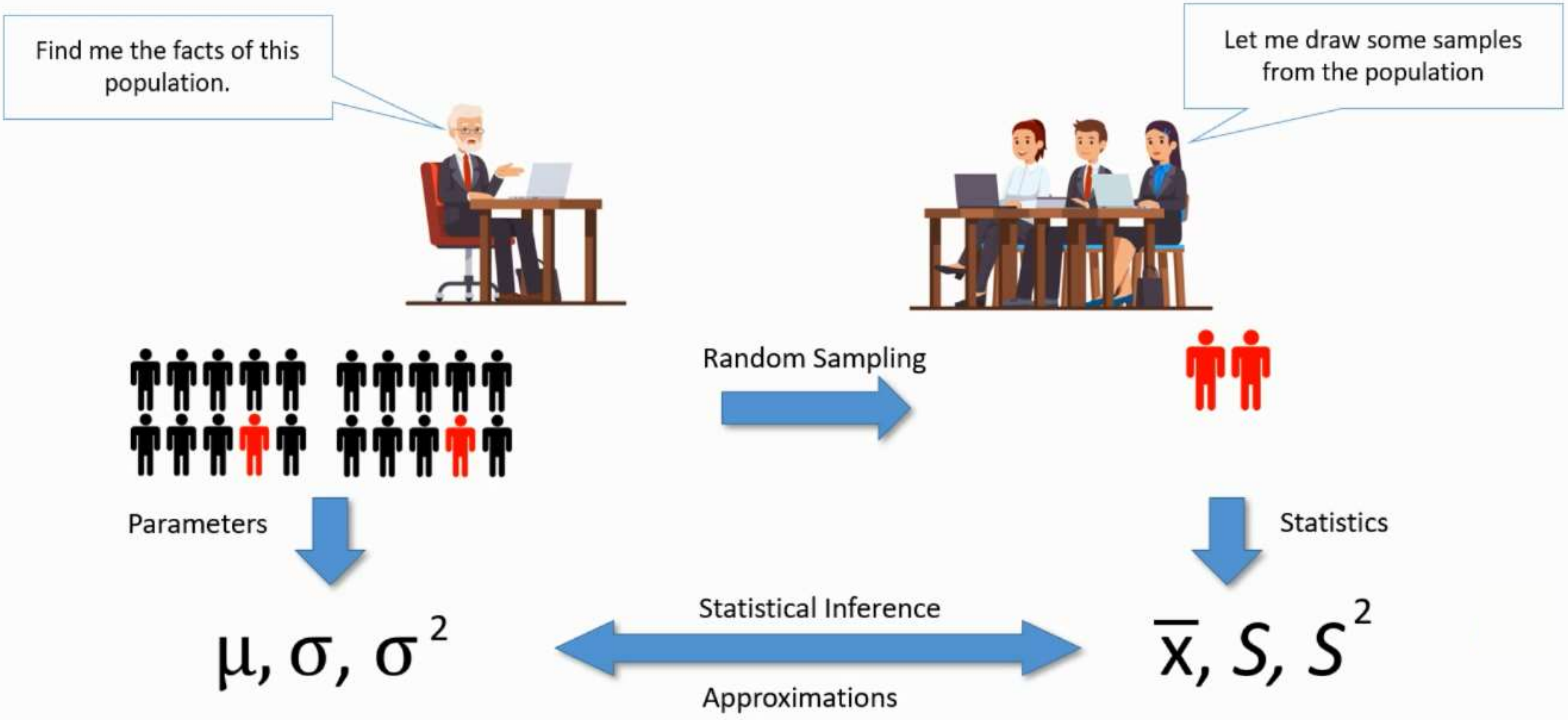
1. Approved
2. Not-Approved

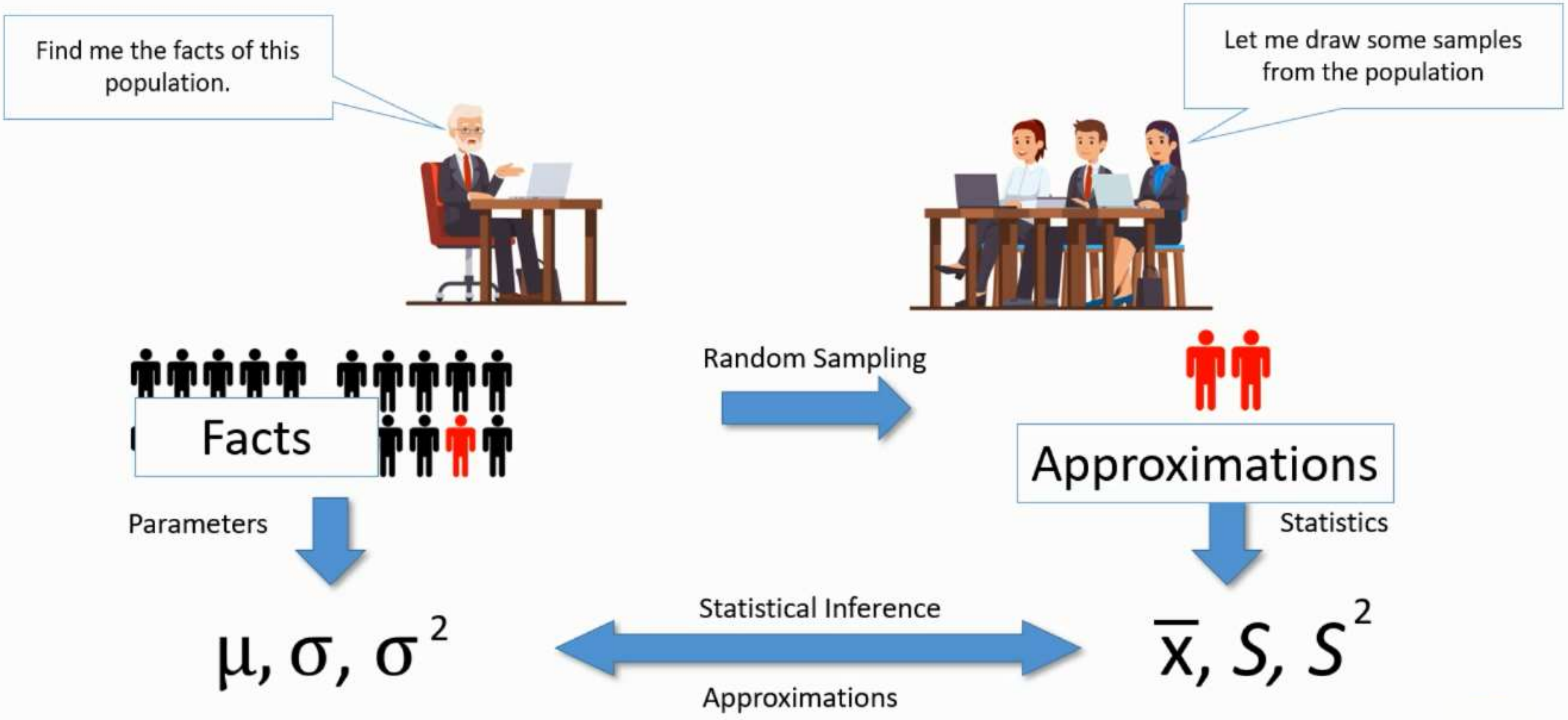
Sample of 40% of records.

Sample size of 4.

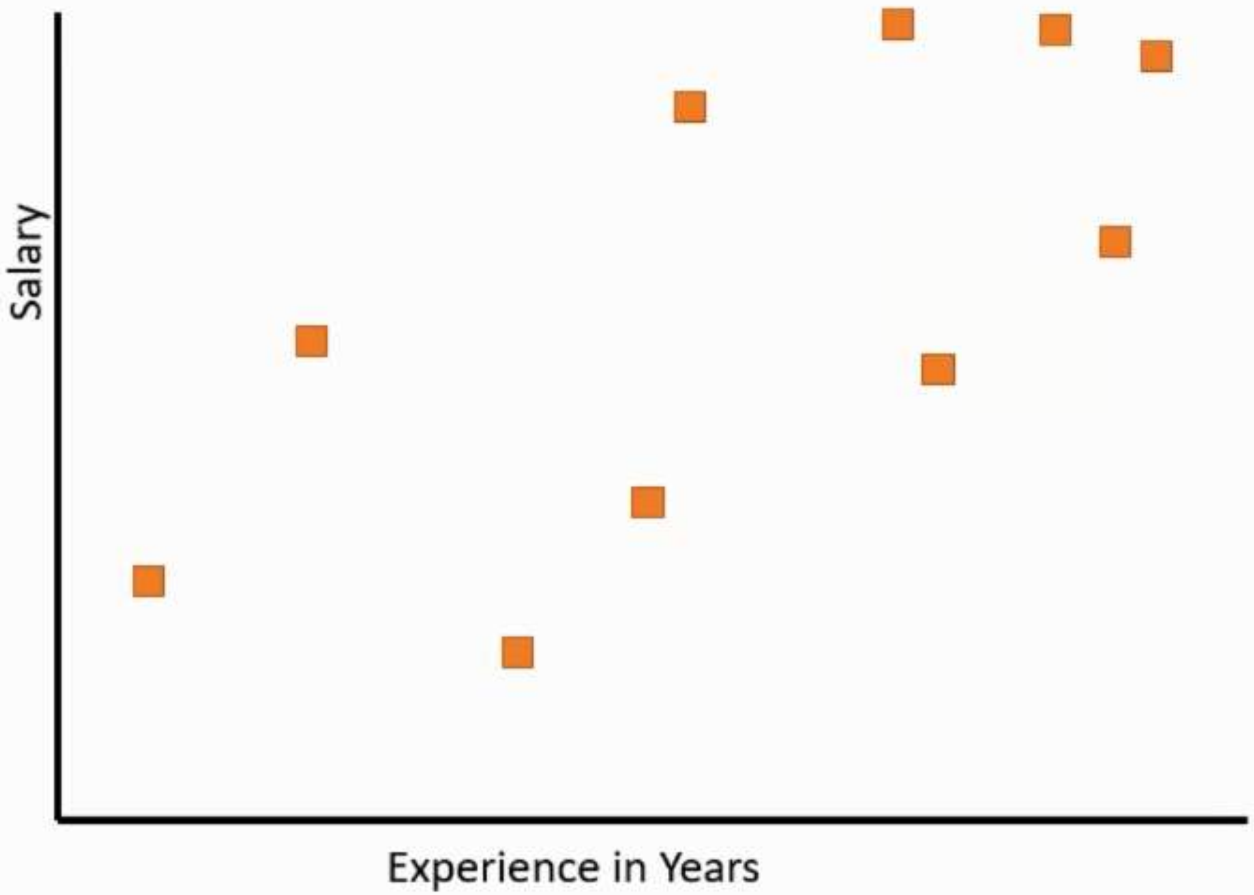
Population and Sample



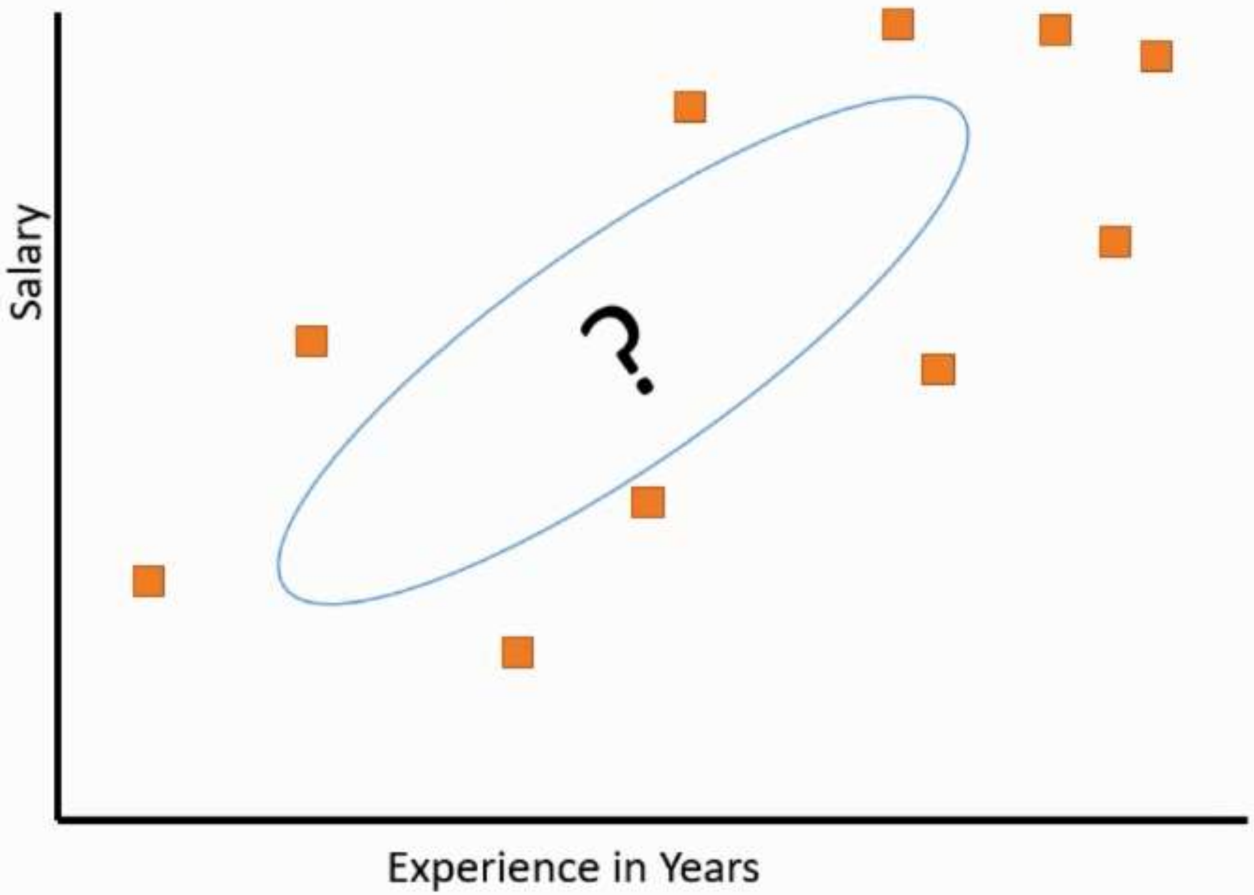

Find me the facts of this population.
Let me draw some samples from the population
Random Sampling
Parameters
Statistics
Statistical Inference
μ, σ, σ^2
$\bar{x}, S, S^2$
Approximations


Find me the facts of this population.
Let me draw some samples from the population
Random Sampling
Facts
Approximations
Parameters
Statistics
Statistical Inference
μ, σ, σ^2
Approximations
$\bar{x}, S, S^2$

Experience Vs Salary Sample

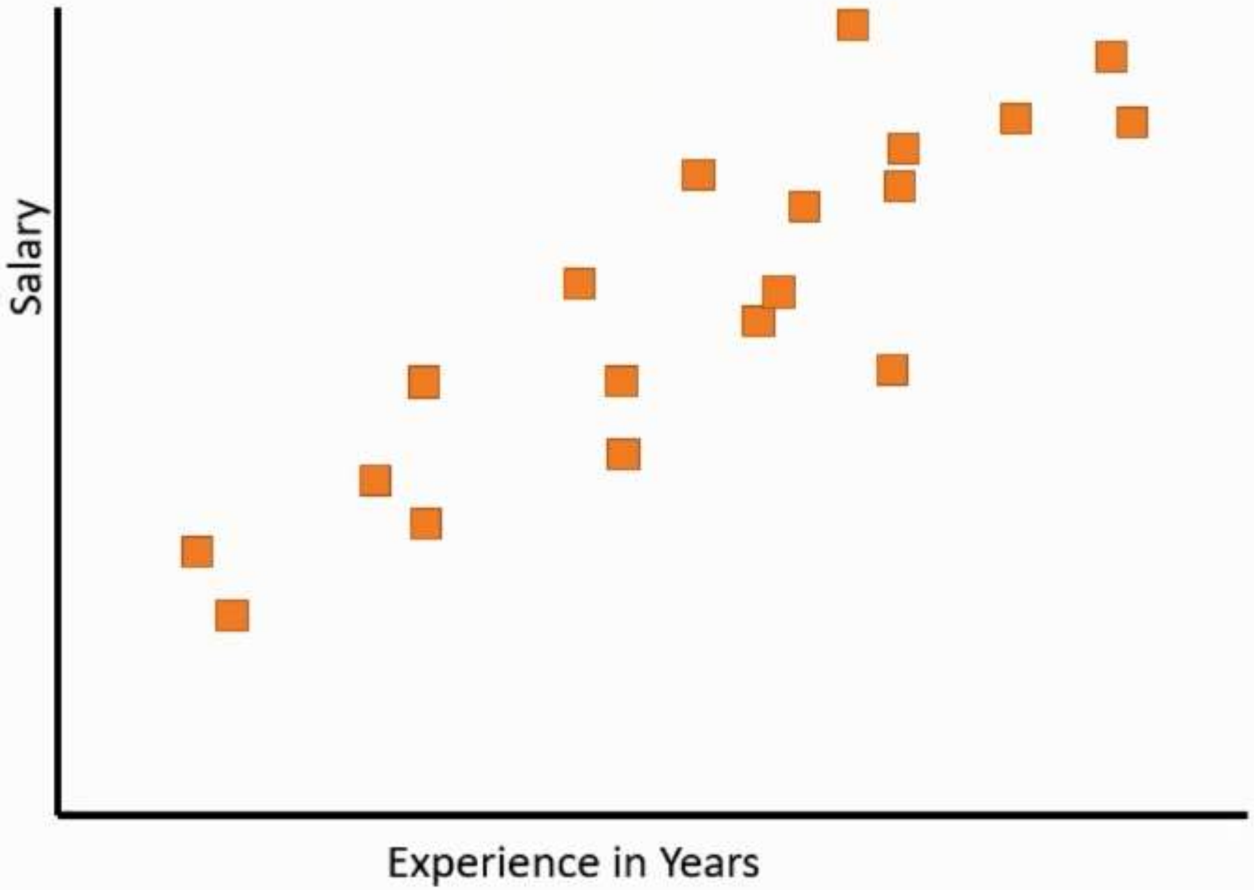


Experience Vs Salary Sample

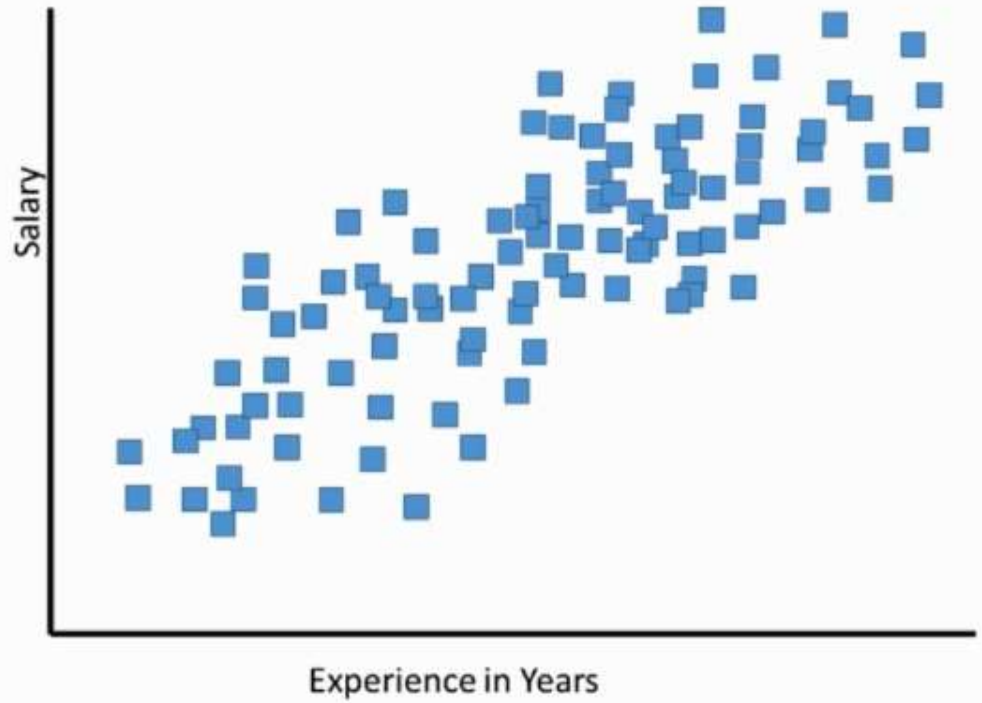


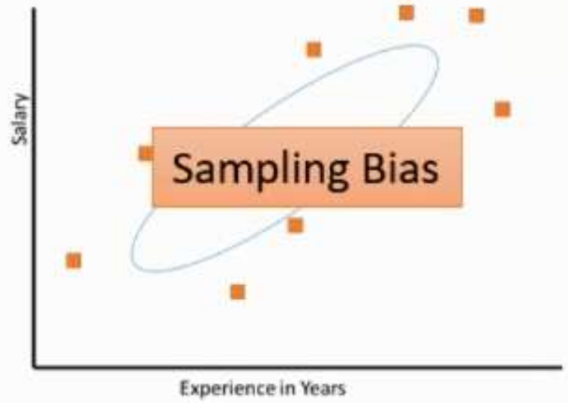
Experience Vs Salary Sample

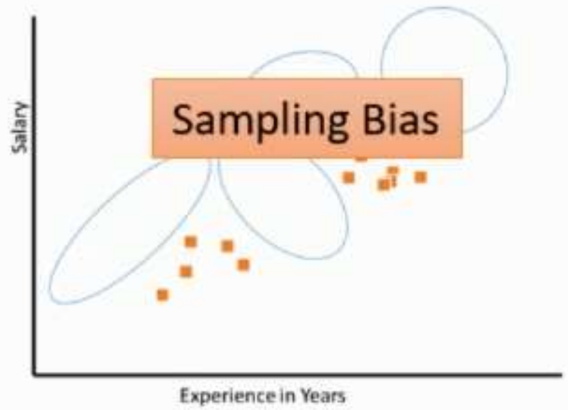
Experience Vs Salary Sample



Experience Vs Salary Sample







Incorrect Inferences

Sample bias

Sampling bias is a bias in which a sample is collected in such a way that some members of the intended population are less likely to be included than others.

-- Wikipedia

Classic Example - 1936 US Presidential Election

Alf Landon

Franklin Roosevelt

Classic Example - 1936 US Presidential Election

- Successful Presidential poll since 1916
- Correct predictions of outcomes for 1916, 1920, 1924, 1928 and 1932
- Most successful poll survey and a great marketing tool

Classic Example - 1936 US Presidential Election

57% of the vote

43% of the vote

- Sent surveys to 10 million people
- 2.4 million responses

Result – Winner with 60% votes

Classic Example - 1936 US Presidential Election

Classic Example - 1936 US Presidential Election

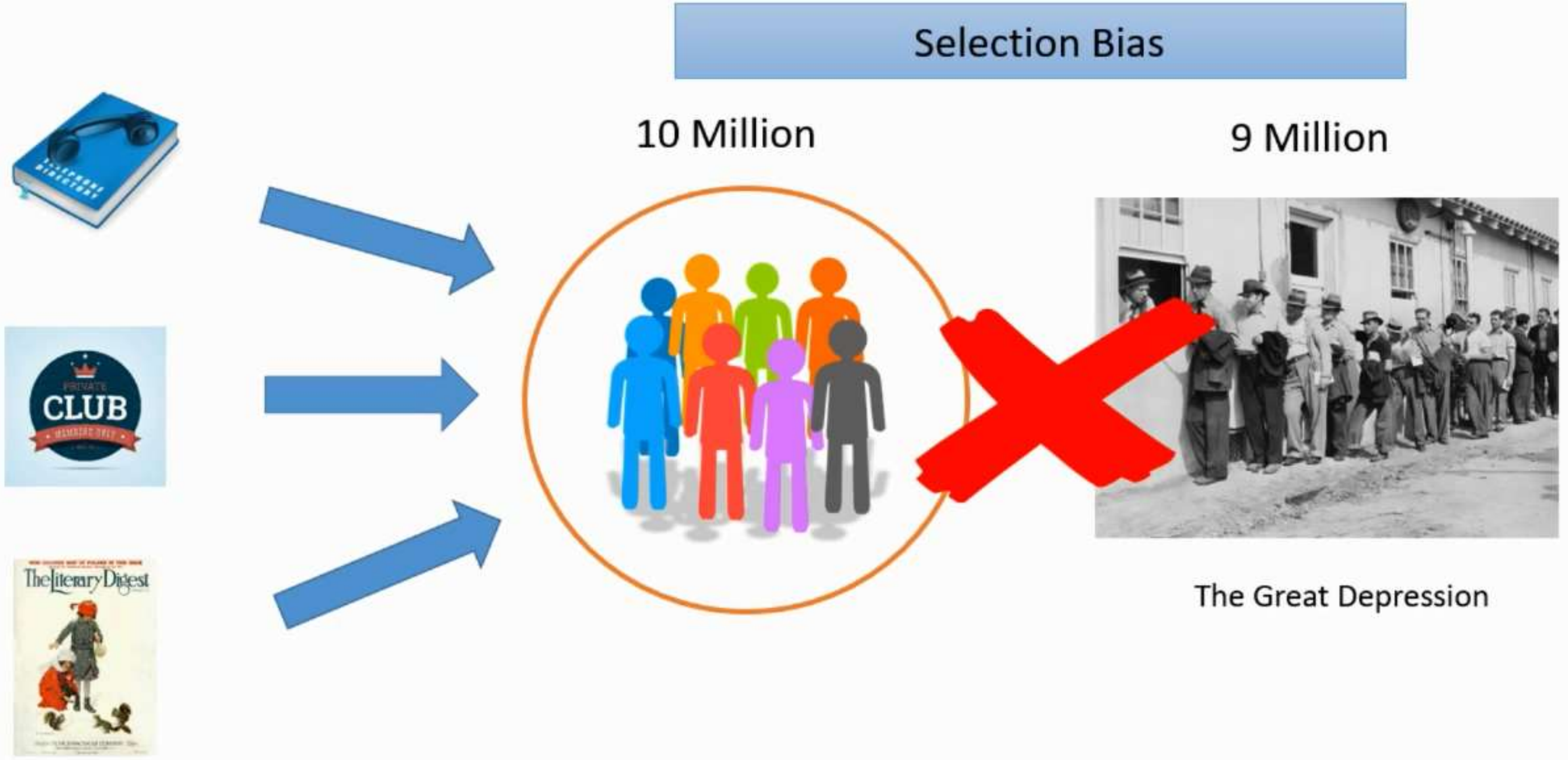
George Gallup

GALLUP

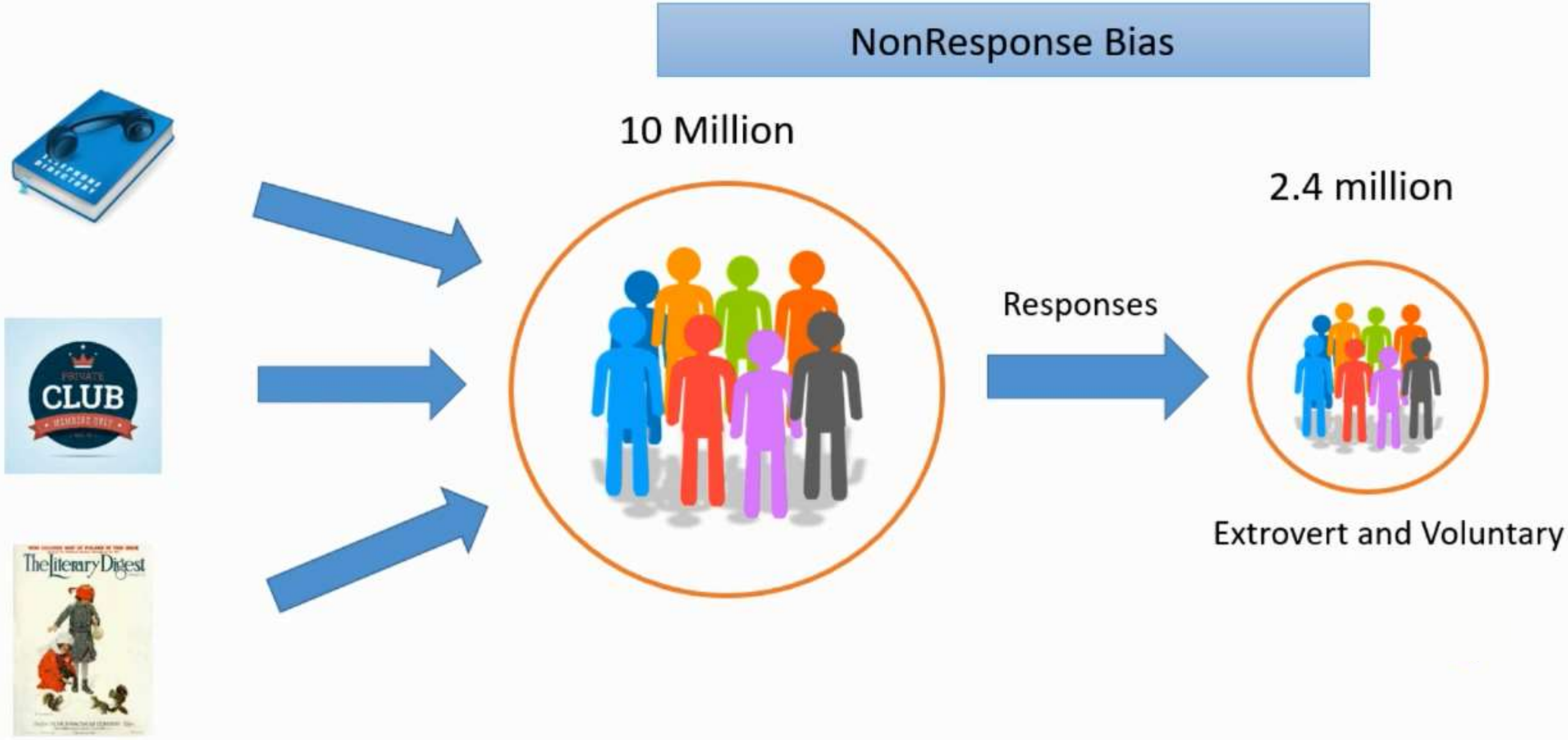
- Correct prediction
- Sample size of 50,000
- Recreated the incorrect result of The Literary Digest

Classic Example - 1936 US Presidential Election

Sampling Bias of The Literary Digest



Sampling Bias of The Literary Digest



Important Points for Sampling

- Data Size vs Data Quality

Completeness

Validity

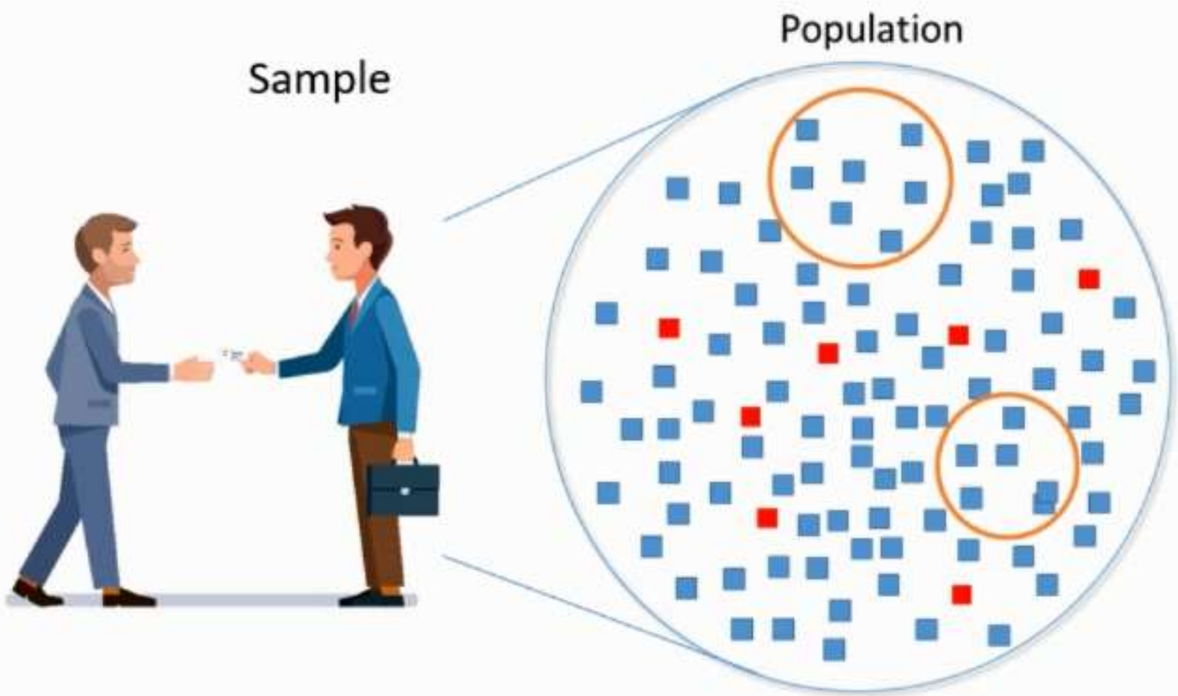
Accuracy

Availability

Timeliness

Important Points for Sampling

- Data Size vs Data Quality
- Sample is a Representation of Population



Important Points for Sampling

- Data Size vs Data Quality
- Sample is a Representation of Population
- Types of Sampling Errors

Important Points for Sampling

- Data Size vs Data Quality
- Sample is a Representation of Population
- Types of Sampling Errors

Population Specific Error

Selection Error

Non-Response Error

Sample Error

Sample Frame Error


in the next two lessons...
MEASURES OF RELATIONSHIP BETWEEN VARIABLES
OUR FOCUS:
Covariance
Linear correlation coefficient

Measurements of Single Variable Vs. Between Multiple Variables

Single Variable Measurements	Multiple Variable Measurements
- Mean / Mode / Median - Skewness - Variance - Standard deviation	- Covariance - Correlation coefficient

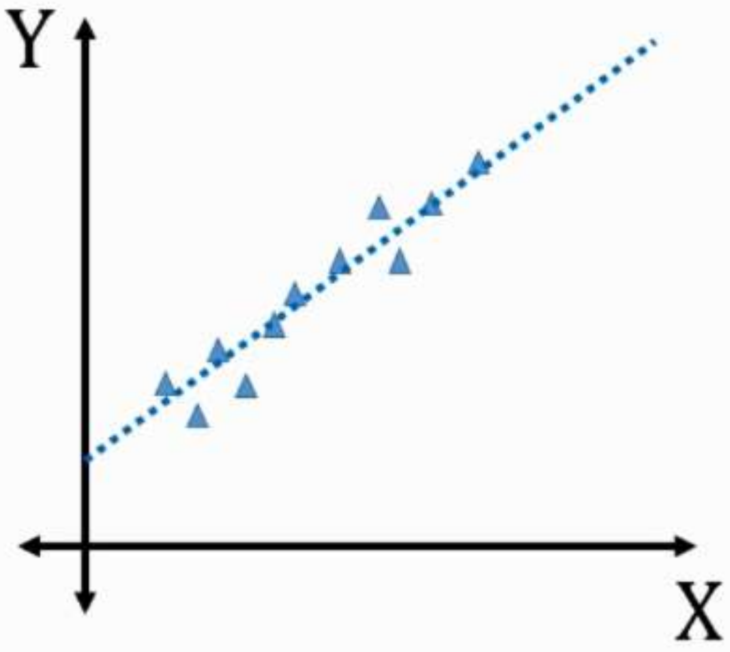
Correlation

Number of cigarettes smoked

Stress Level

Correlation

Statistically Correlated

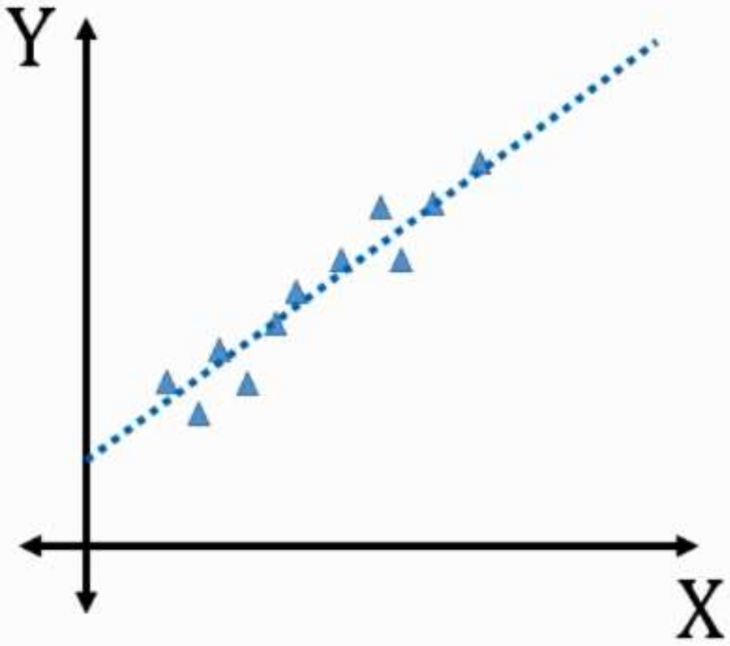


$X = \text{Height}$

$Y = \text{Weight}$

Statistically Correlated

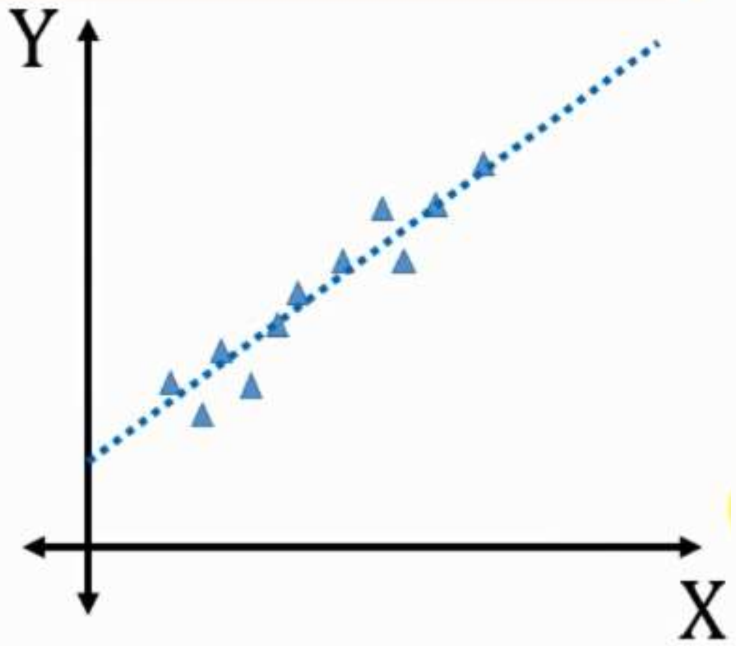
- Strength of the correlation – Coefficient of Correlation
- Direction of correlation – Sign of the Coefficient



Statistically Correlated

- Strength of the correlation – Coefficient of Correlation

- Direction of correlation – Sign of the Coefficient



Pearson Correlation Coefficient

$$r = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1) * \sigma_x * \sigma_y}$$

Correlation Coefficient

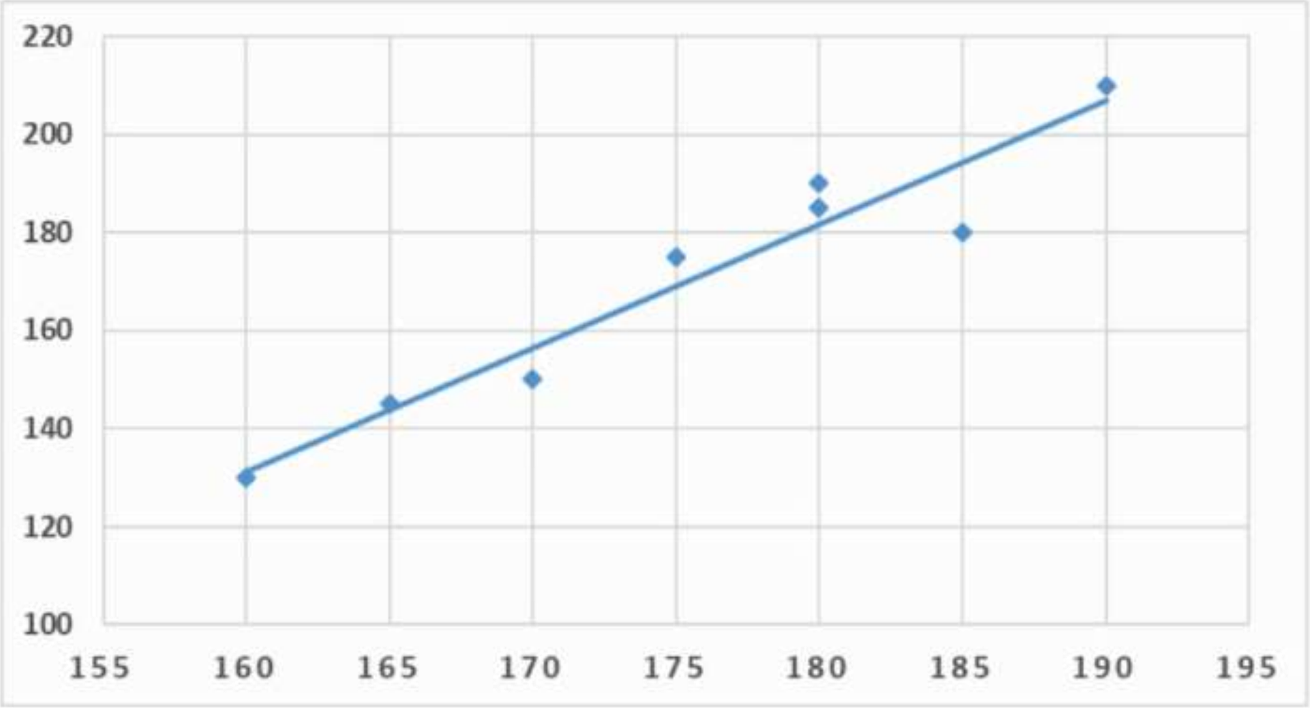
	Height X	Weight Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X}) * (Y - \bar{Y})$
	160	130	-15.625	-40.625	634.7656
	170	150	-5.625	-20.625	116.0156
	165	145	-10.625	-25.625	272.2656
	180	190	4.375	19.375	84.76563
	175	175	-0.625	4.375	-2.73438
	190	210	14.375	39.375	566.0156
	185	180	9.375	9.375	87.89063
	180	185	4.375	14.375	62.89063
Mean	**175.625**	**170.625**			**1821.875**
Std Dev	**10.155**	**25.651**			

$$r = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1) * \sigma_x * \sigma_y}$$

$$r = \frac{1821.875}{(8-1) * 10.155 * 25.651}$$

$$r = 0.96$$

Correlation Coefficient



Scatter Plot

$$r = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1) * \sigma_x * \sigma_y}$$

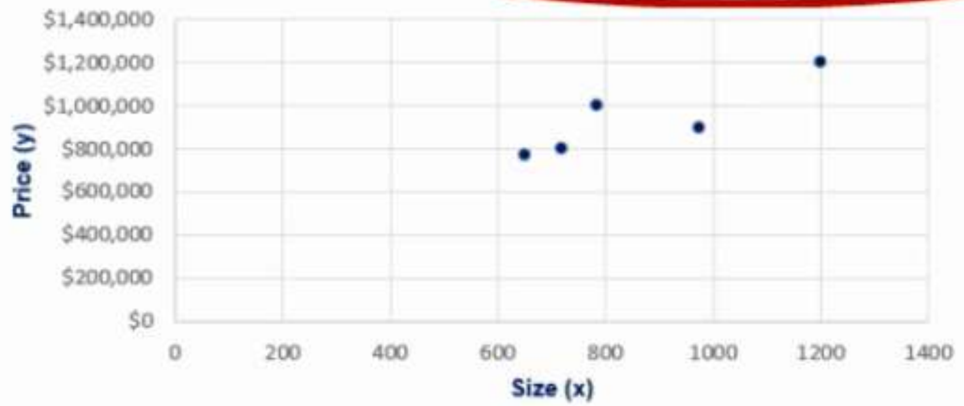
$$r = \frac{1821.875}{(8\text{-}1) * 10.155 * 25.651}$$

$$r = 0.96$$

Correlation coefficient

Housing data

	Size (ft.)	Price ($)		$(x-\bar{x})*(y-\bar{y})$
	650	772,000		34,776,000
	785	998,000		- 5,265,000
	1200	1,200,000		89,178,000
	720	800,000		19,418,000
	975	895,000		- 4,142,000
Mean	866	933,000	**Sum**	133,965,000
Standard dev.	222	173,615	**Sample size**	5
			Cov. sample	33,491,250
			Correlation coeff.	0.87



$$\frac{s_{xy}}{s_x s_y}$$

$-1 \leq$ correlation coefficient ≤ 1

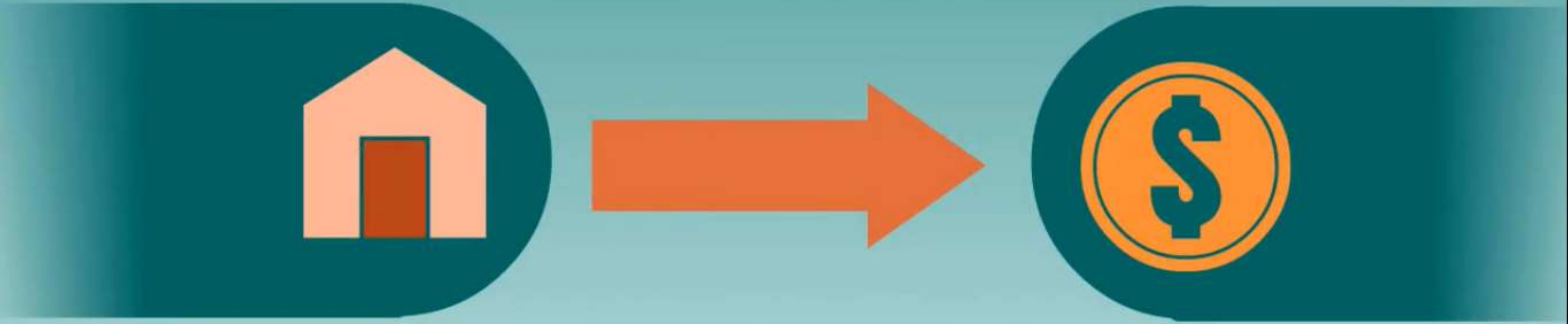
PERFECT POSITIVE CORRELATION

Correlaton coeff.= 1

the entire variability of one variable is explained by the other

RELATIONSHIP DIRECTION

Size determines price



CORRELATION OF 0

Absolutely independent variables

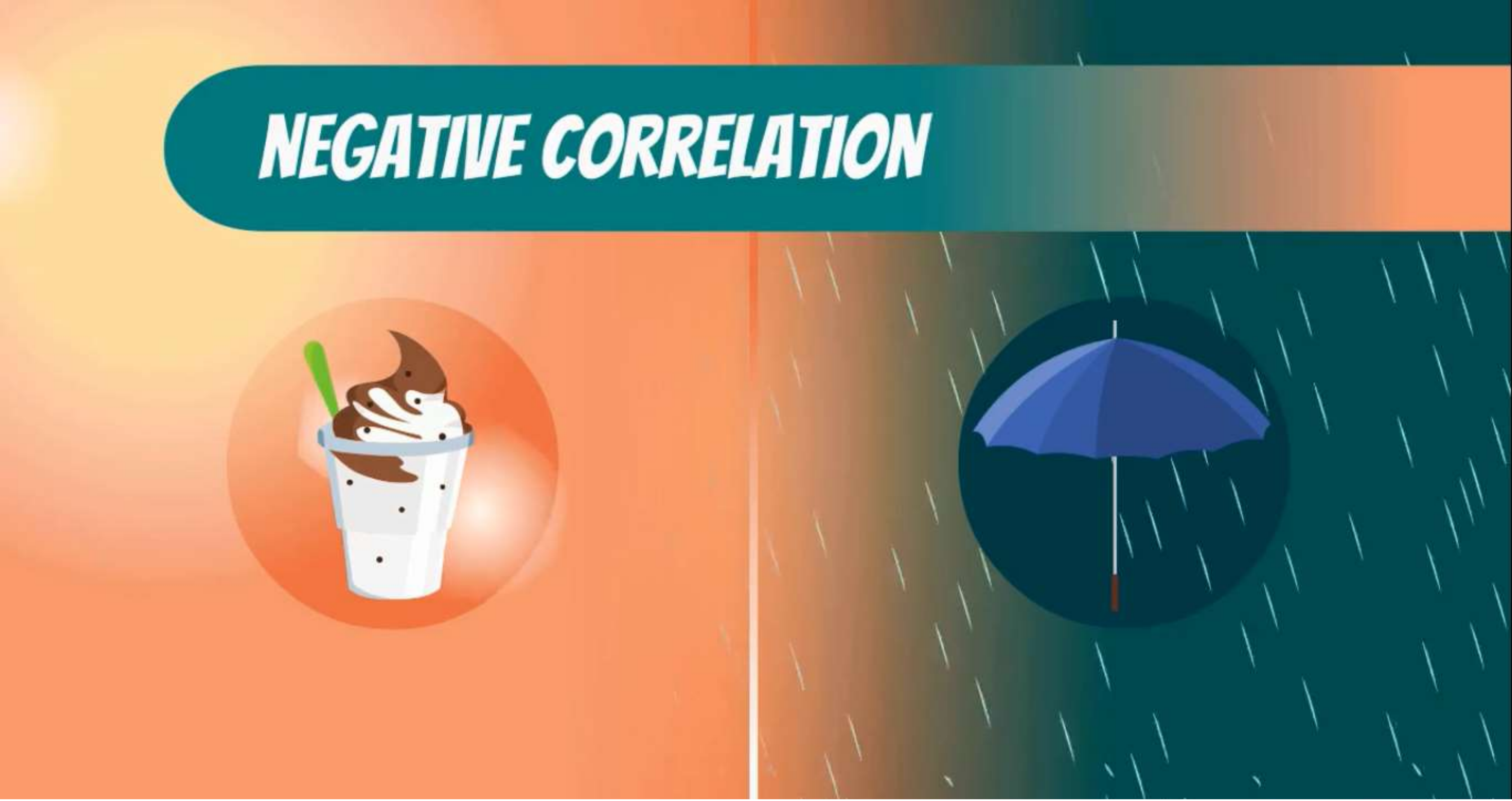
Coffee in Brazil

Houses in London

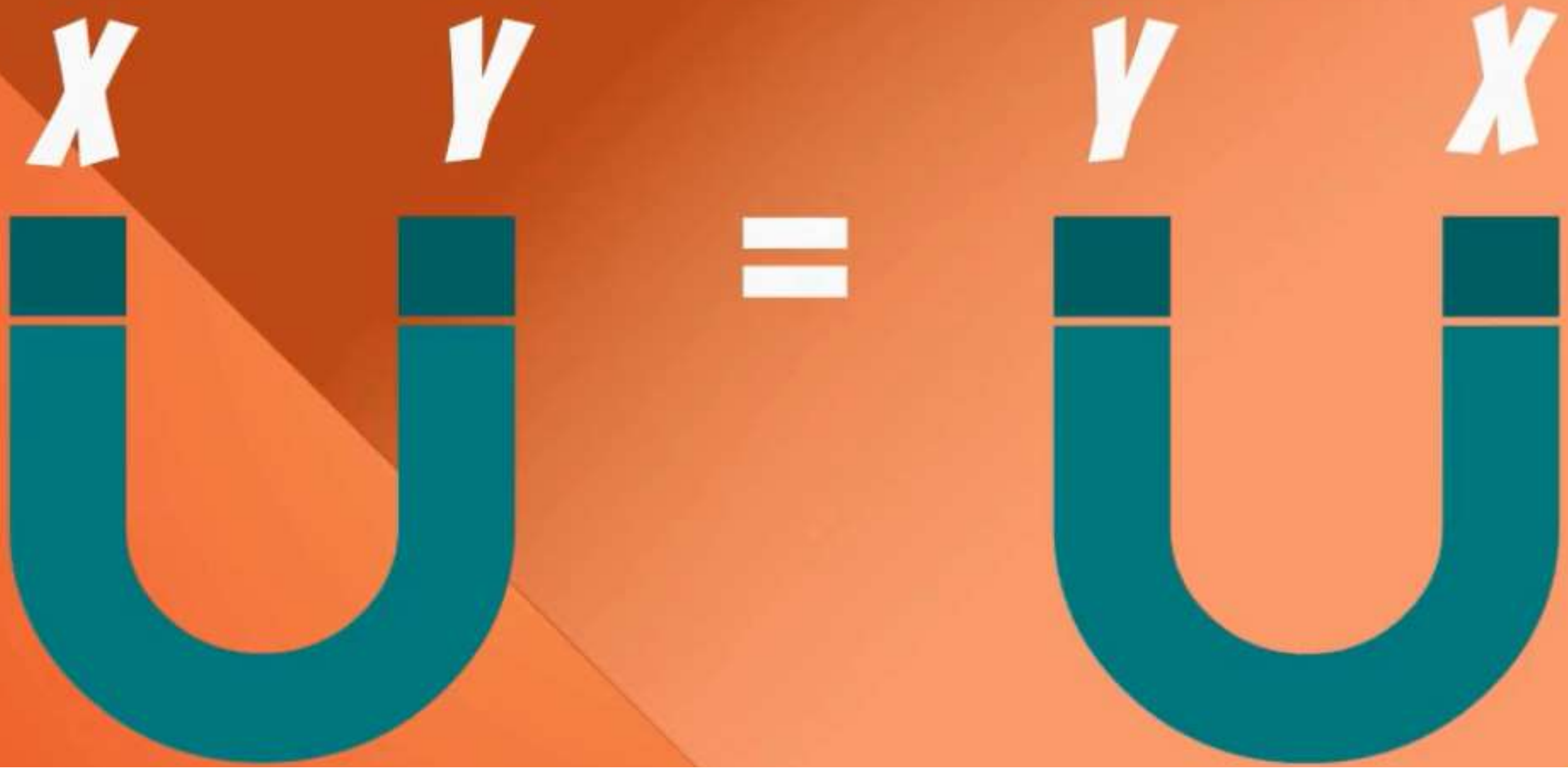
NEGATIVE CORRELATION

Perfect negative correlation of - 1

Imperfect negative correlation: (- 1,0)


NEGATIVE CORRELATION

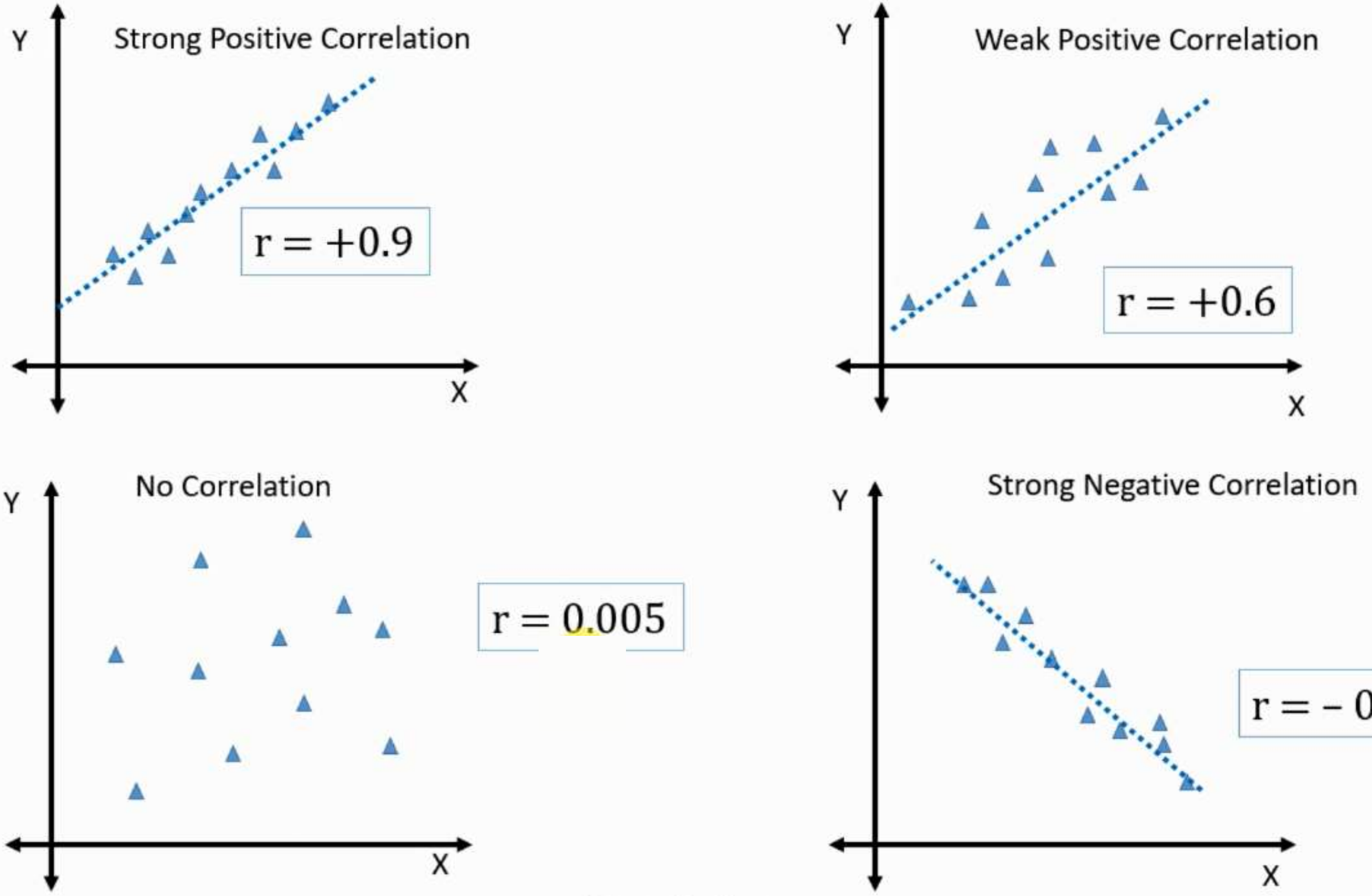
CORRELATION



CORRELATION

$$\frac{\mathrm{Cov}(x, y)}{\mathrm{Stdev}(x) * \mathrm{Stdev}(y)} = \frac{\mathrm{Cov}(y, x)}{\mathrm{Stdev}(y) * \mathrm{Stdev}(x)}$$

Symmetrical with respect to both variables

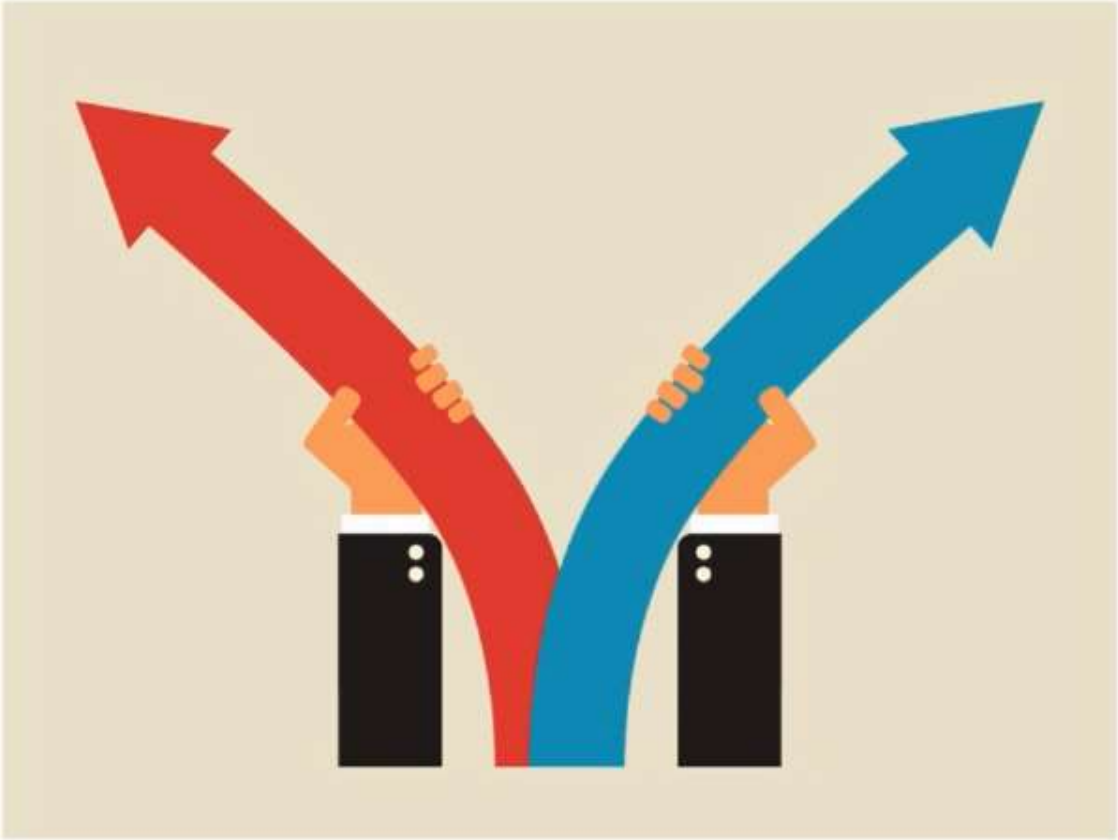


Strong Positive Correlation
Y
X
r = +0.9
Weak Positive Correlation
Y
X
r = +0.6
No Correlation
Y
X
r = 0.005
Strong Negative Correlation
Y
X
r = – 0.8

Variance

Average of the squared difference of the data from the Mean.

$$\text{Variance, } S_x^2 = \frac{\sum (x - \bar{x})^2}{(N - 1)}$$



Variance

Average of the squared difference of the data from the Mean.

$$\text{Variance, } S_x^2 = \frac{\sum (x - \bar{x}) * (x - \bar{x})}{(N - 1)}$$

Variance of X with respect to X.

Variance

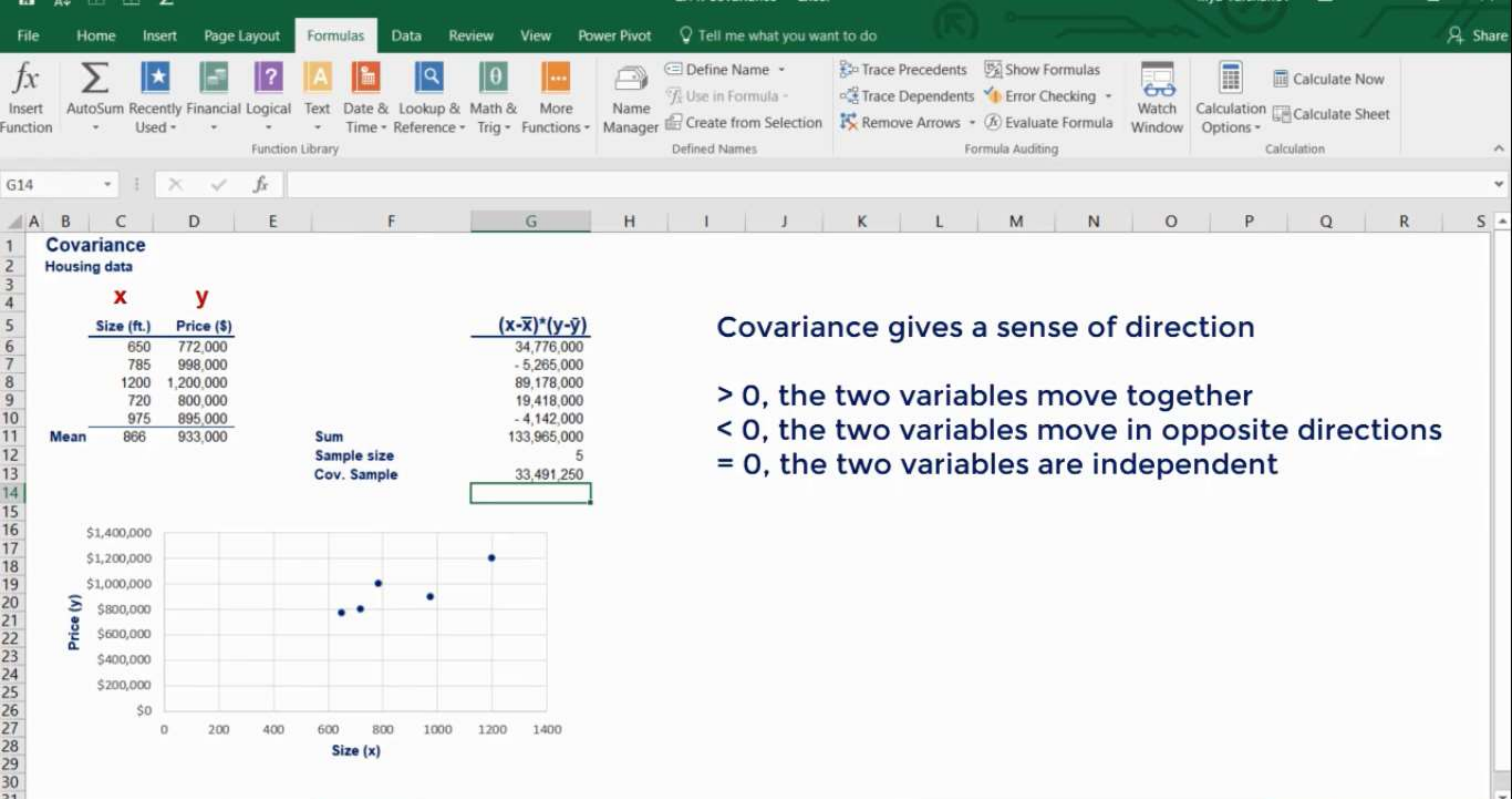
$$\text{Variance, } S_x^2 = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1)}$$

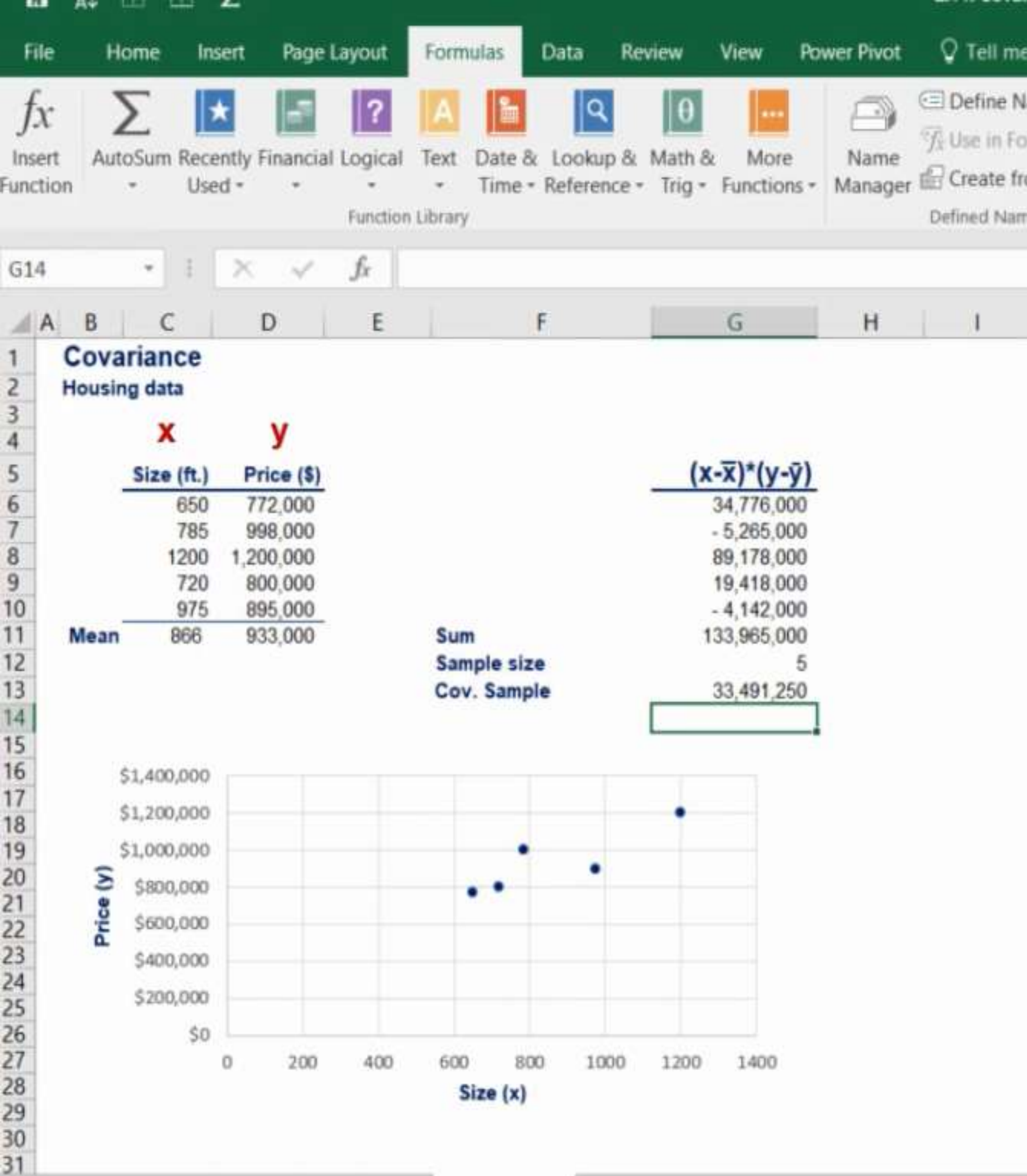
Variance of X with respect to Y.

Covariance

$$\text{Covariance, } S_{xy}^{2} = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1)}$$

Variance of X with respect to Y.


File Home Insert Page Layout Formulas Data Review View Power Pivot Tell me what you want to do Share
Insert Function AutoSum Recently Used Financial Logical Text Date & Time Lookup & Reference Math & Trig More Functions
Function Library
Name Manager Define Name Use in Formula Create from Selection
Defined Names
Trace Precedents Trace Dependents Remove Arrows Show Formulas Error Checking Evaluate Formula Watch Window
Formula Auditing
Calculation Options Calculate Now Calculate Sheet
Calculation
G14
Covariance
Housing data
x y
Size (ft.) Price ($)
650 772,000
785 998,000
1200 1,200,000
720 800,000
975 895,000
Mean 866 933,000
(x-x̄)*(y-ȳ)
34,776,000
-5,265,000
89,178,000
19,418,000
-4,142,000
Sum 133,965,000
Sample size 5
Cov. Sample 33,491,250
Covariance gives a sense of direction
> 0, the two variables move together
< 0, the two variables move in opposite directions
= 0, the two variables are independent
Price (y)
$1,400,000 $1,200,000 $1,000,000 $800,000 $600,000 $400,000 $200,000 $0
0 200 400 600 800 1000 1200 1400
Size (x)



5

50

0.0023456

33,491,250

Note: Result from covariance can be of any scale as mentioned above , that's why we use correlation coefficient to solve this issue and make the scale between (+1 to -1).

Covariance

Pearson Correlation Coefficient $$r = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1) * \sigma_x * \sigma_y}$$

$$\text{Covariance, } S_{xy}^2 = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1)}$$

Variance of X with respect to Y.

Covariance

Pearson Correlation Coefficient

$$r = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1) * \sigma_x * \sigma_y} = \frac{\textbf{Covar}\ (x, y)}{\sigma_x * \sigma_y}$$

$$\text{Covariance, } S_{xy}^2 = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1)}$$

Variance of X with respect to Y.

Covariance

	Height X	Weight Y	$X-\overline{X}$	$Y-\overline{Y}$	$(X-\overline{X})*(Y-\overline{Y})$
	160	130	-15.625	-40.625	634.7656
	170	150	-5.625	-20.625	116.0156
	165	145	-10.625	-25.625	272.2656
	180	190	4.375	19.375	84.76563
	175	175	-0.625	4.375	-2.73438
	190	210	14.375	39.375	566.0156
	185	180	9.375	9.375	87.89063
	180	185	4.375	14.375	62.89063
Mean	**175.625**	**170.625**			**1821.875**
Std Dev	**10.155**	**25.651**			

$$\text{Covariance, } S_{xy}^2 = \frac{\sum (x-\overline{x}) * (y-\overline{y})}{(N-1)}$$

$$\text{Covar}(x, y) = \frac{1821.875}{(8-1)}$$

$$\text{Covar}(x, y) = 260.27$$

Covariance

- Non-Standardised method of correlation
- Can be positive or negative

$$\text{Covariance, } S_{xy}^2 = \frac{\sum (x - \overline{x}) * (y - \overline{y})}{(N - 1)}$$

$$\text{Covar}(x, y) = \frac{1821.875}{(8\text{-}1)}$$

$$\text{Covar}(x, y) = 260.27$$

Covariance Matrix

	Height X	Weight Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X}) * (Y - \bar{Y})$
	160	130	-15.625	-40.625	634.7656
	170	150	-5.625	-20.625	116.0156
	165	145	-10.625	-25.625	272.2656
	180	190	4.375	19.375	84.76563
	175	175	-0.625	4.375	-2.73438
	190	210	14.375	39.375	566.0156
	185	180	9.375	9.375	87.89063
	180	185	4.375	14.375	62.89063
Mean	**175.625**	**170.625**			**1821.875**
Std Dev	**10.155**	**25.651**			

	x	y
x	Covariance (x, x)	Covariance(x, y)
y	Covariance (y, x)	Covariance (y, y)

	x	y	z
x			
y			
z			

Covariance Matrix

	Height X	Weight Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X}) * (Y - \bar{Y})$
	160	130	-15.625	-40.625	634.7656
	170	150	-5.625	-20.625	116.0156
	165	145	-10.625	-25.625	272.2656
	180	190	4.375	19.375	84.76563
	175	175	-0.625	4.375	-2.73438
	190	210	14.375	39.375	566.0156
	185	180	9.375	9.375	87.89063
	180	185	4.375	14.375	62.89063
Mean	**175.625**	**170.625**			**1821.875**
Std Dev	**10.155**	**25.651**			

	x	y
x	Variance(x)	Covariance(x, y)
y	Covariance (y, x)	Variance (y)

Variance – Covariance Matrix

$$\text{Covariance, } S_{xy}^{2} = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{(N - 1)}$$

Covariance Matrix

	Height X	Weight Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})*(Y-\bar{Y})$
	160	130	-15.625	-40.625	634.7656
	170	150	-5.625	-20.625	116.0156
	165	145	-10.625	-25.625	272.2656
	180	190	4.375	19.375	84.76563
	175	175	-0.625	4.375	-2.73438
	190	210	14.375	39.375	566.0156
	185	180	9.375	9.375	87.89063
	180	185	4.375	14.375	62.89063
Mean	**175.625**	**170.625**			**1821.875**
Std Dev	**10.155**	**25.651**			

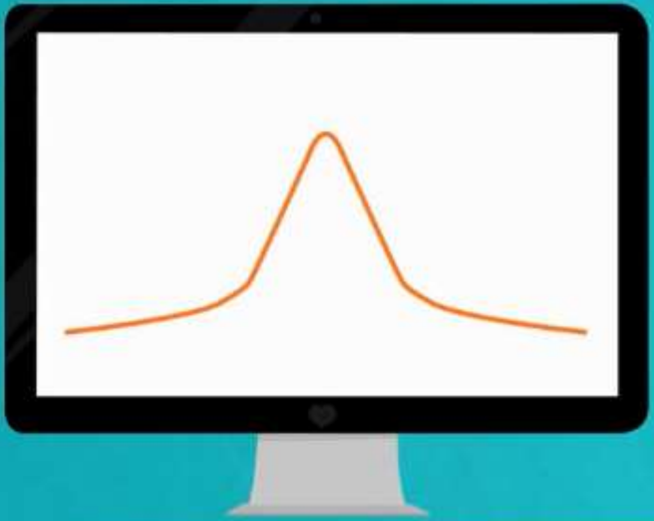
	x	y
x	103.125	260.27
y	260.27	710.26

Variance – Covariance Matrix

IN STATISTICS

DISTRIBUTION = PROBABILITY DISTRIBUTION

DISTRIBUTION = PROBABILITY DISTRIBUTION



NORMAL

BINOMIAL

UNIFORM

DEFINITION

A distribution is a function that shows the possible values for a variable and how often they occur.

ROLLING A DIE

<table>
<tr><th>OUTCOME</th><th>PROBABILITY</th></tr>
<tr><td>1</td><td rowspan="6">1/6</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td></tr>
<tr><td>6</td></tr>
</table>

SUM OF THE PROBABILITIES :

1 | 100%

=> we have exhausted all possibilities

ROLLING A DIE

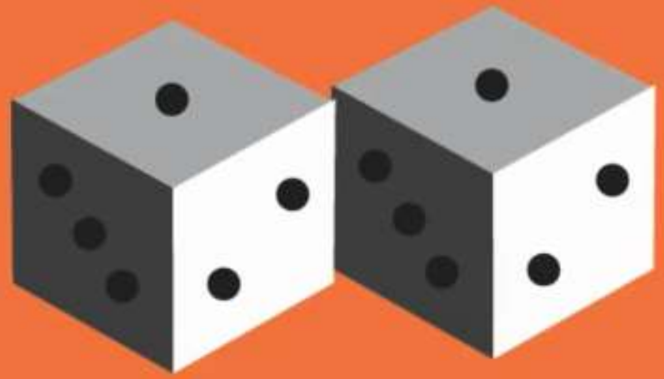
OUTCOME	PROBABILITY
1; 2; 3; 4; 5; 6	0.17
All else	0

DISCRETE UNIFORM DISTRIBUTION

ROLLING TWO DICE

Possibilities?

(1,1) (2,1) (1,2)



ROLLING TWO DICE

Possibilities? Total: 36

(1,1) (2,1) (3,1) (4,1) (5,1) (6,1)
(1,2) (2,2) (3,2) (4,2) (5,2) (6,2)
(1,3) (2,3) (3,3) (4,3) (5,3) (6,3)
(1,4) (2,4) (3,4) (4,4) (5,4) (6,4)
(1,5) (2,5) (3,5) (4,5) (5,5) (6,5)
(1,6) (2,6) (3,6) (4,6) (5,6) (6,6)

ROLLING TWO DICE

Probability of getting

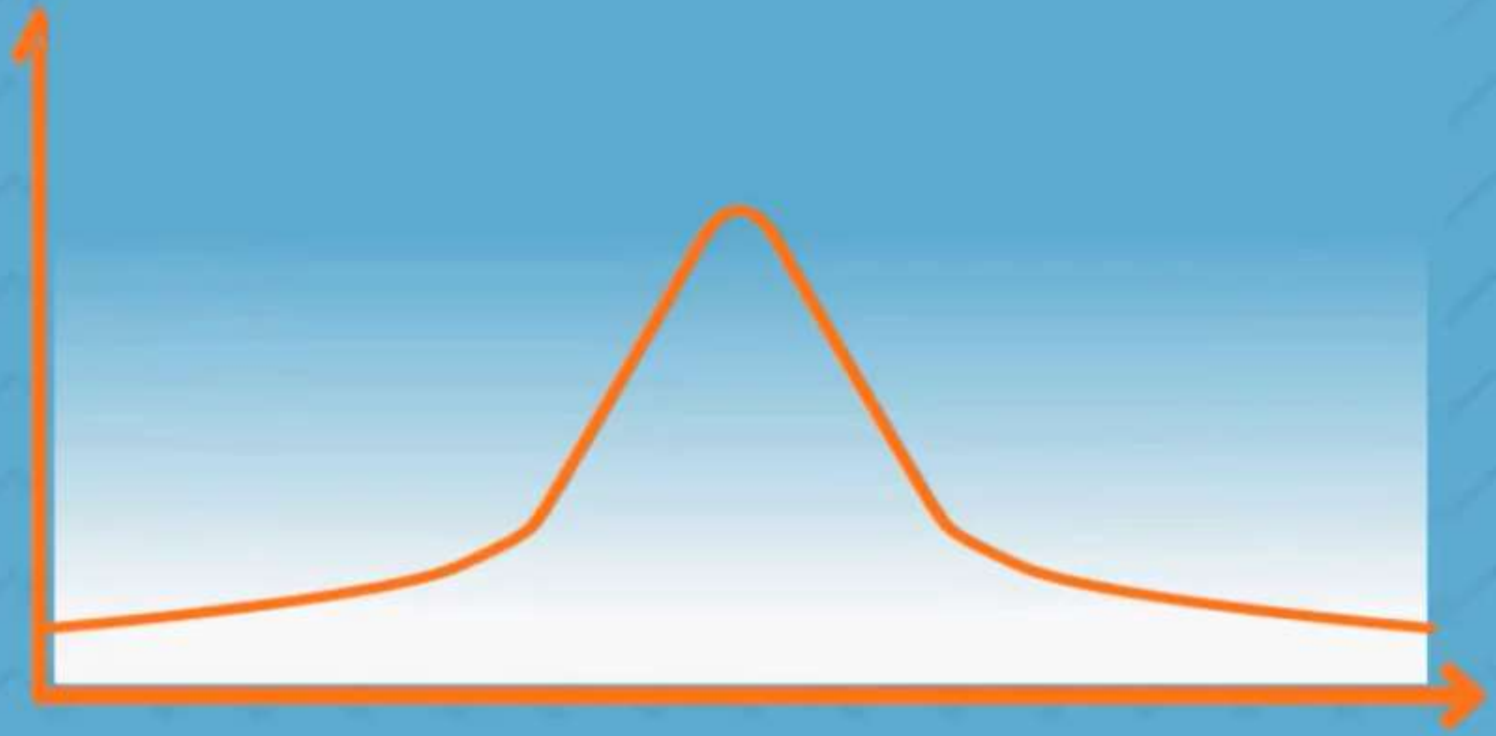
$\frac{2}{36}$ (1,2) and (2,1)

ROLLING TWO DICE

OUTCOME	PROBABILITY
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03
All else	0

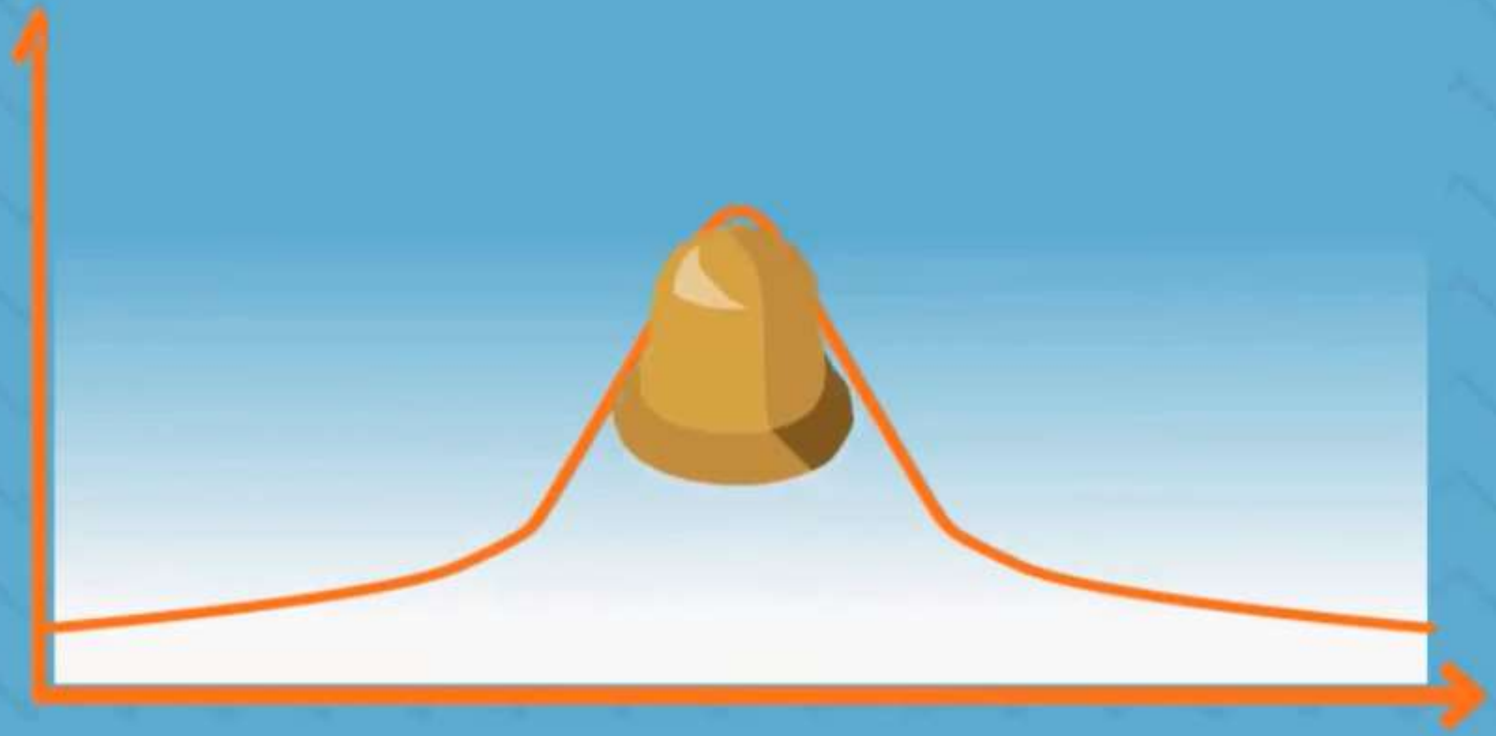

0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
2
3
4
5
6
7
8
9
10
11
12

NORMAL DISTRIBUTION

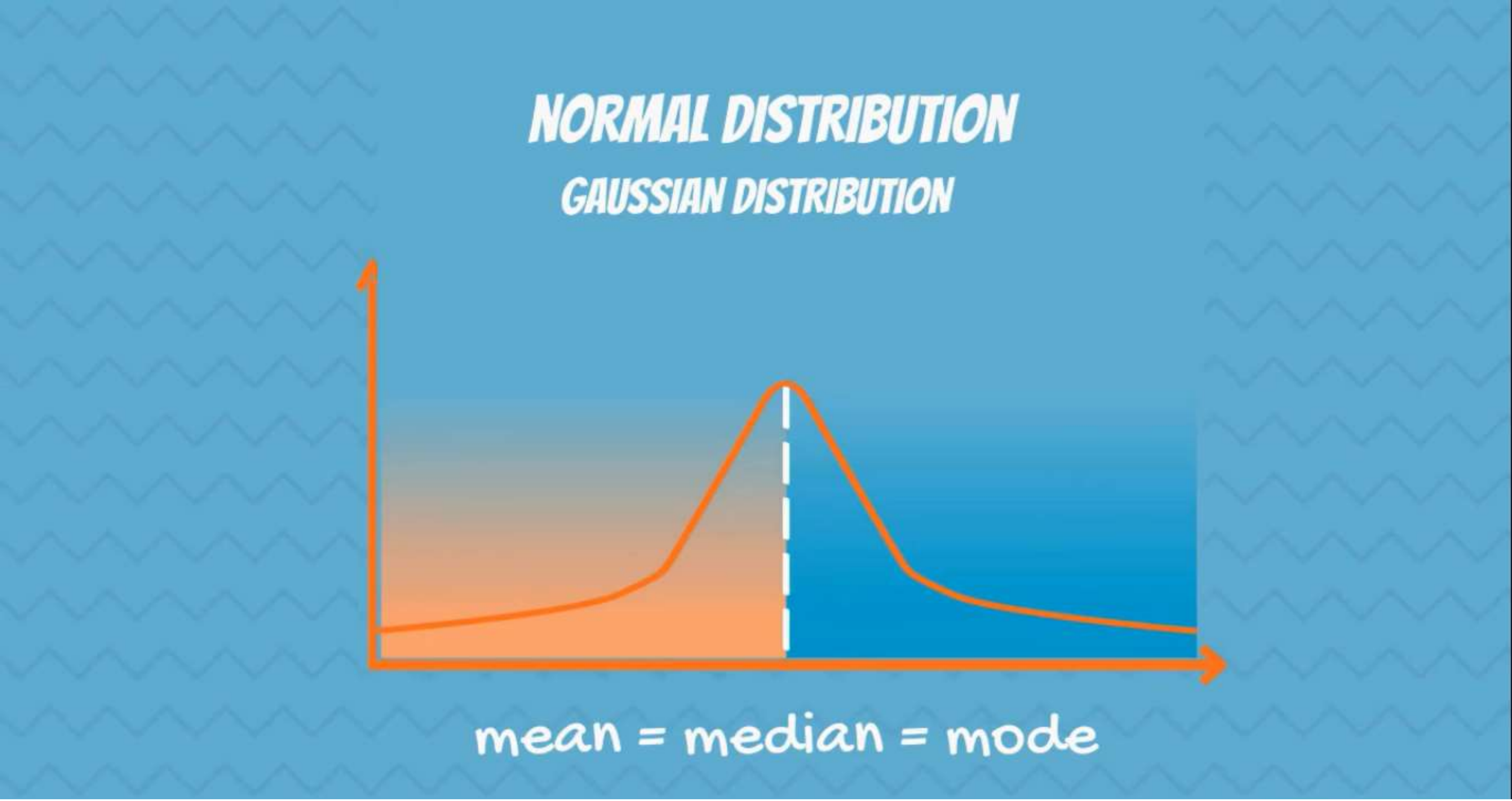


NORMAL DISTRIBUTION

GAUSSIAN DISTRIBUTION

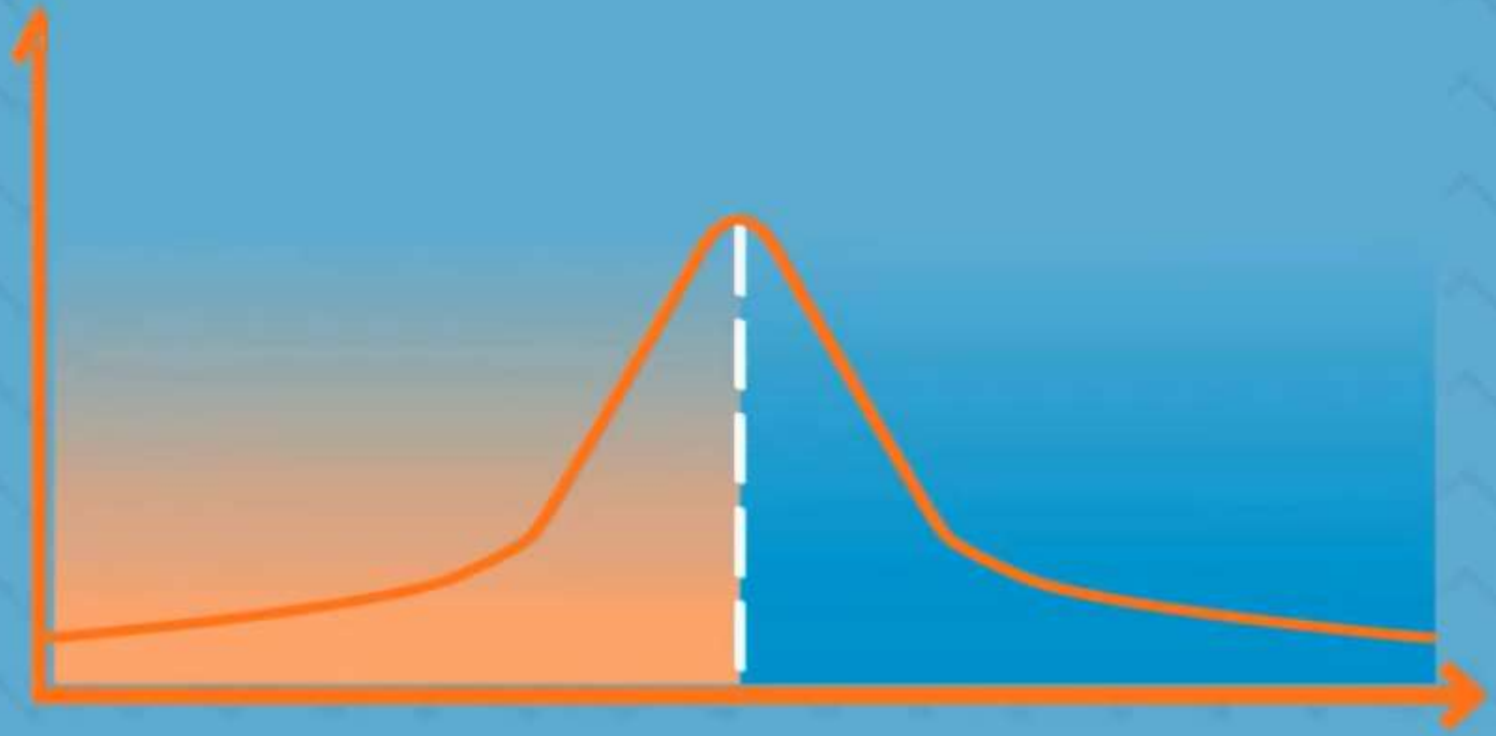


BELL CURVE


NORMAL DISTRIBUTION
GAUSSIAN DISTRIBUTION
mean = median = mode

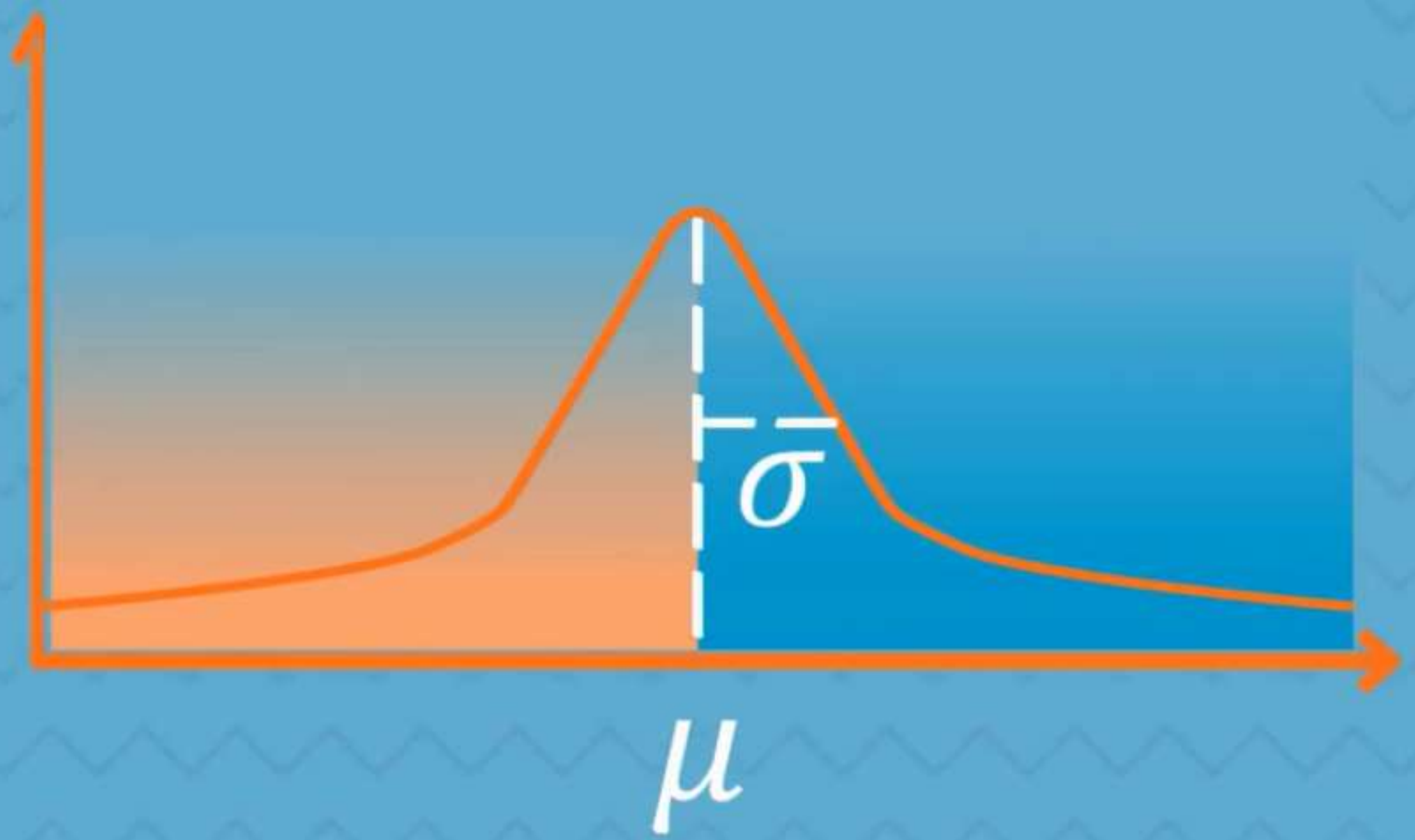
NORMAL DISTRIBUTION

GAUSSIAN DISTRIBUTION

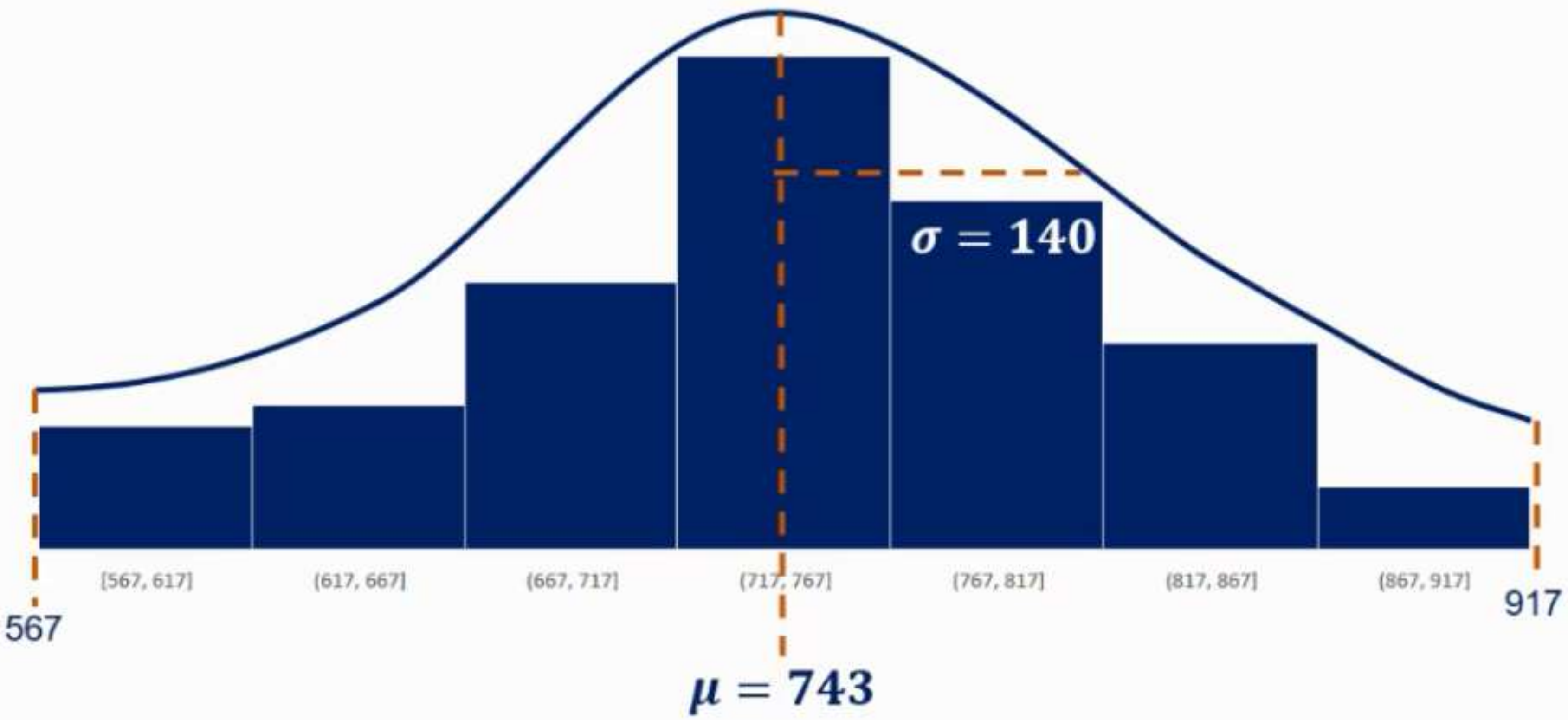


no skew

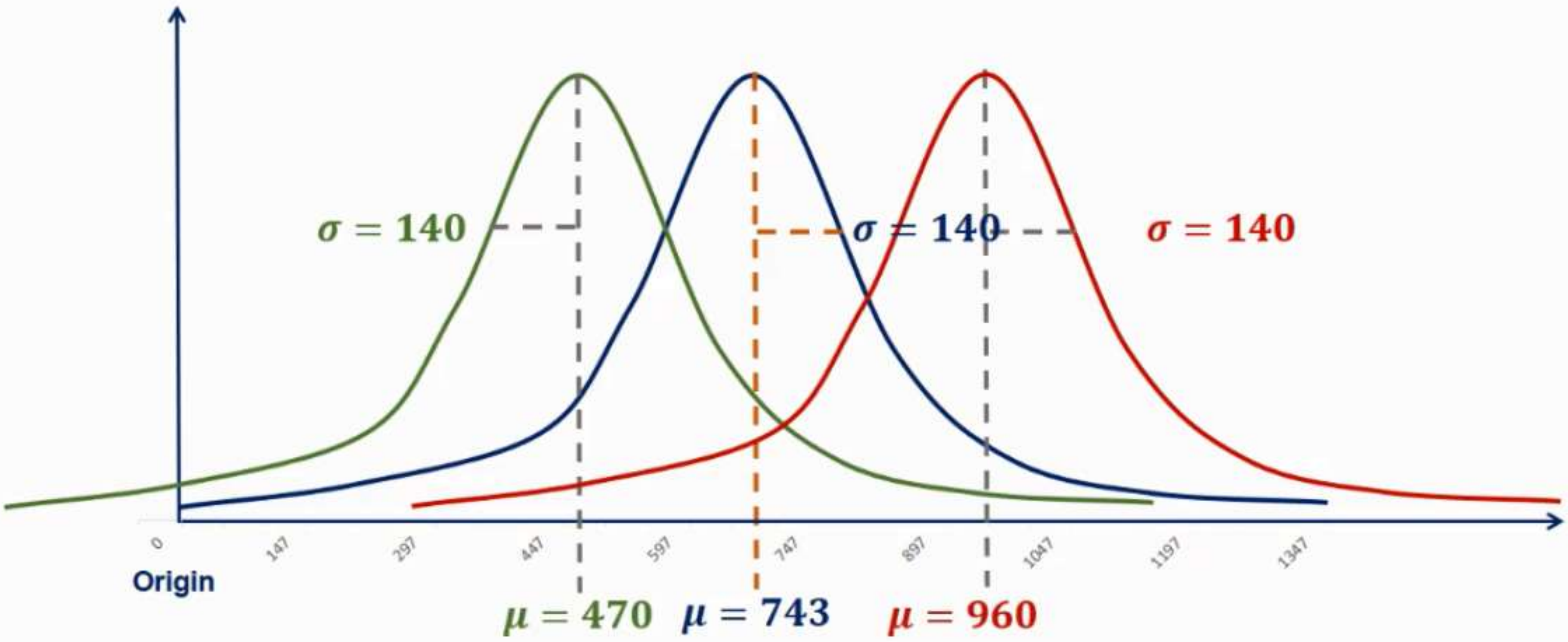
$$N \sim (\mu, \sigma^2)$$



Normal distribution

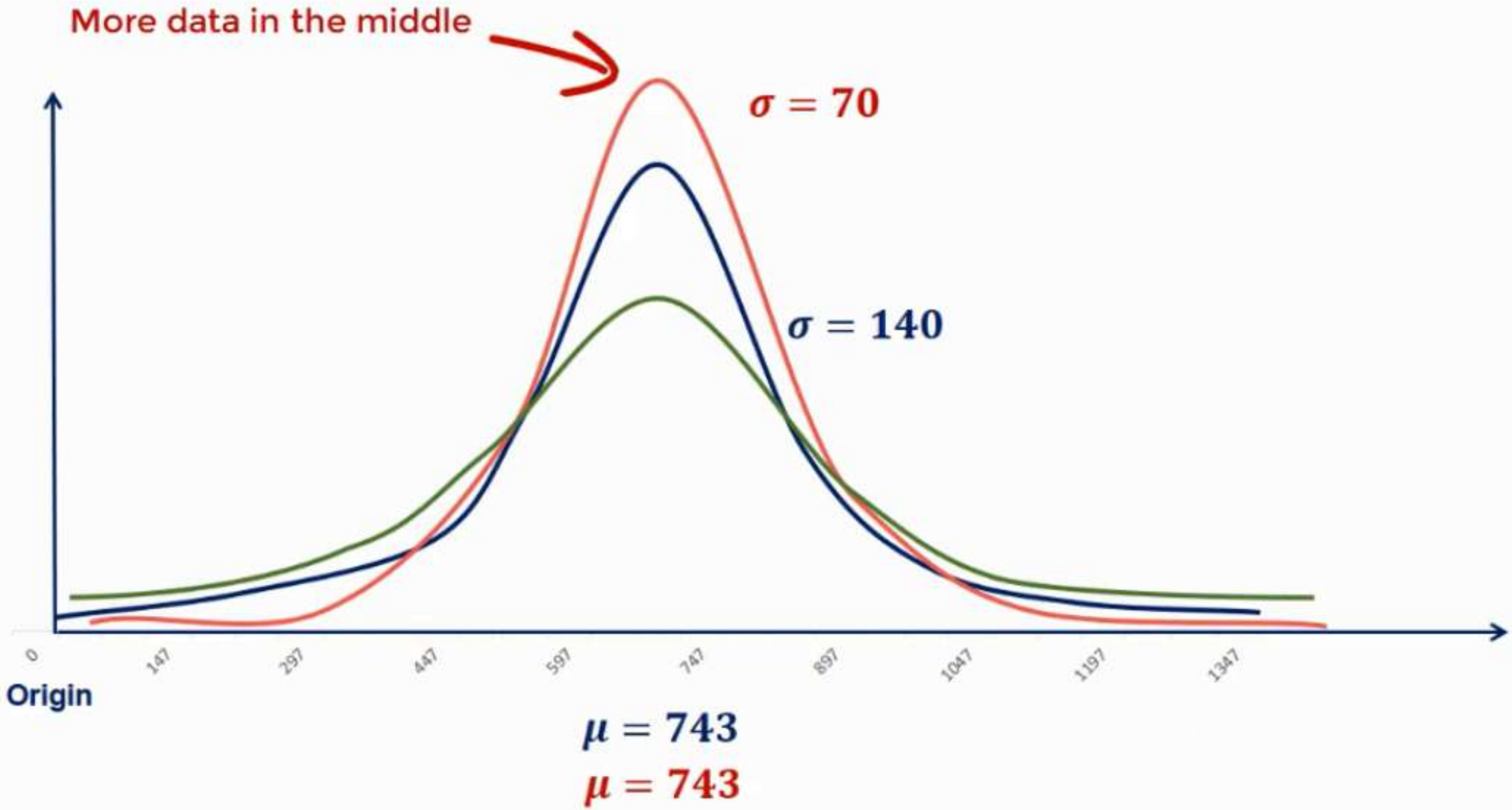


Normal distribution. Controlling for standard deviation



Note: When the mean is lower it will shift the distribution graph to the left, and when the mean value is higher it will shift the distribution graph to the right side.

Normal distribution. Controlling for the mean



Note: When the standard deviation is lower then more data will be concentrated on the middle and when standard deviation is higher means less data in the middle.

Discrete Variable

$$X = \{ 1, 2, 3, 4, 5, 6 \}$$

Experiment of dice with even Y.

Dice1 →

← Dice2

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$P(2) = 1/36$

$P(4) = 3/36$

$P(6) = 5/36$

$P(8) = 5/36$

$P(10) = 3/36$

$P(12) = 1/36$

$P(Y \text{ is even}) = 18/36 = 0.5$

Experiment of dice with even Y.

Dice1 →

↓ Dice2

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$P(2) = 1/36$ $P(3) = 2/36$

$P(4) = 3/36$ $P(5) = 4/36$

$P(6) = 5/36$ $P(7) = 6/36$

$P(8) = 5/36$ $P(9) = 4/36$

$P(10) = 3/36$ $P(11) = 2/36$

$P(12) = 1/36$

$$P(Y \text{ is even}) = 18/36 = 0.5$$

Distribution of Discrete Variables

Dice1 →

← Dice2

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$P(2) = 1/36$ $P(3) = 2/36$

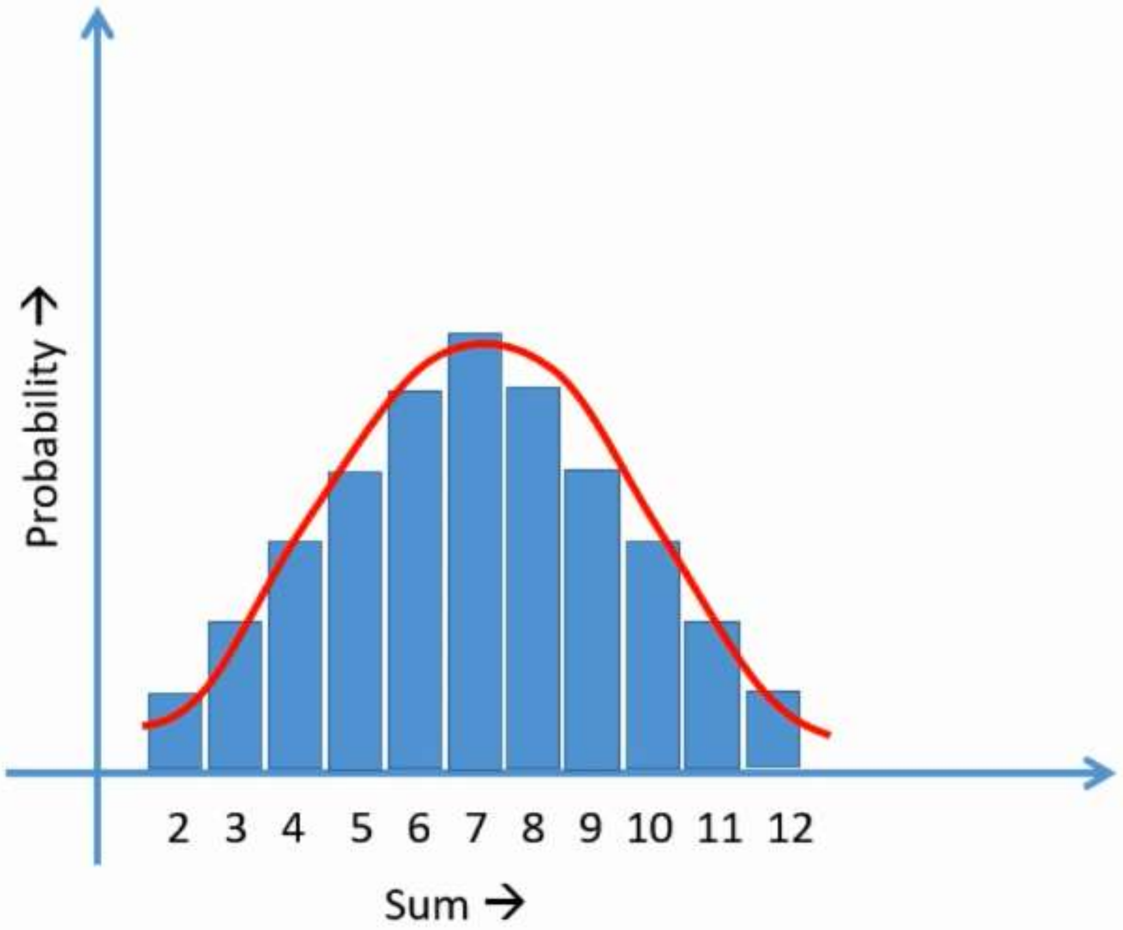
$P(4) = 3/36$ $P(5) = 4/36$

$P(6) = 5/36$ $P(7) = 6/36$

$P(8) = 5/36$ $P(9) = 4/36$

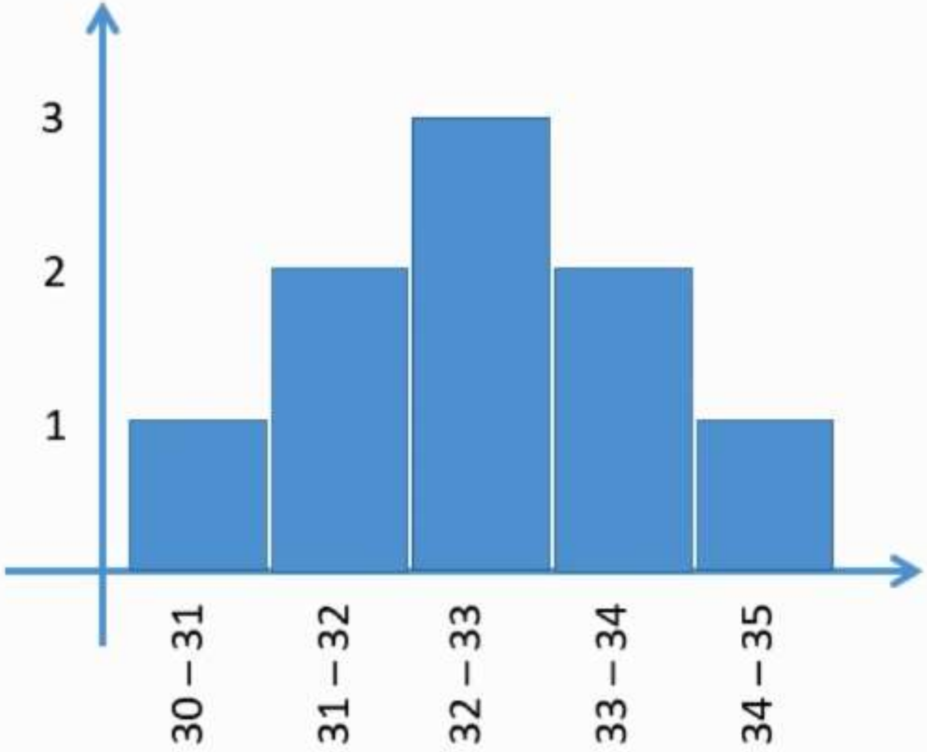
$P(10) = 3/36$ $P(11) = 2/36$

$P(12) = 1/36$



Distribution of Continuous Variable

Temperature
30.6
31.4
31.2
32.1
32.2
32.7
33.4
33.8
34.6



Frequency Distribution with Bins

What % of values are between

30-31 → 1/9 = 11.1%

31-32 → 2/9 = 22.2%

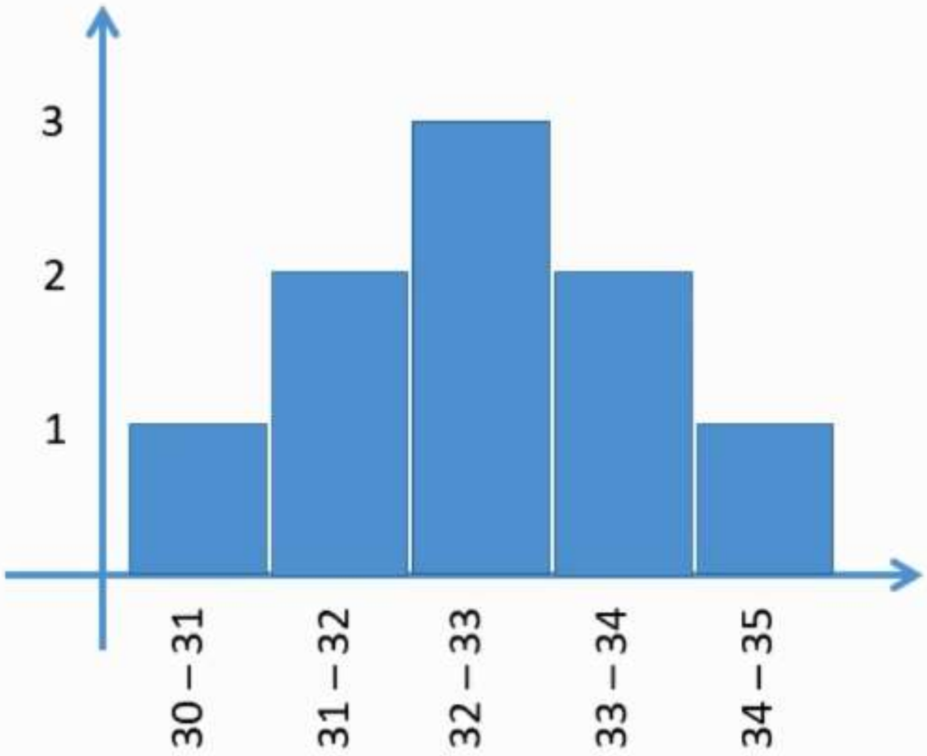
32-33 → 3/9 = 33.3%

33-34 → 2/9 = 22.2%

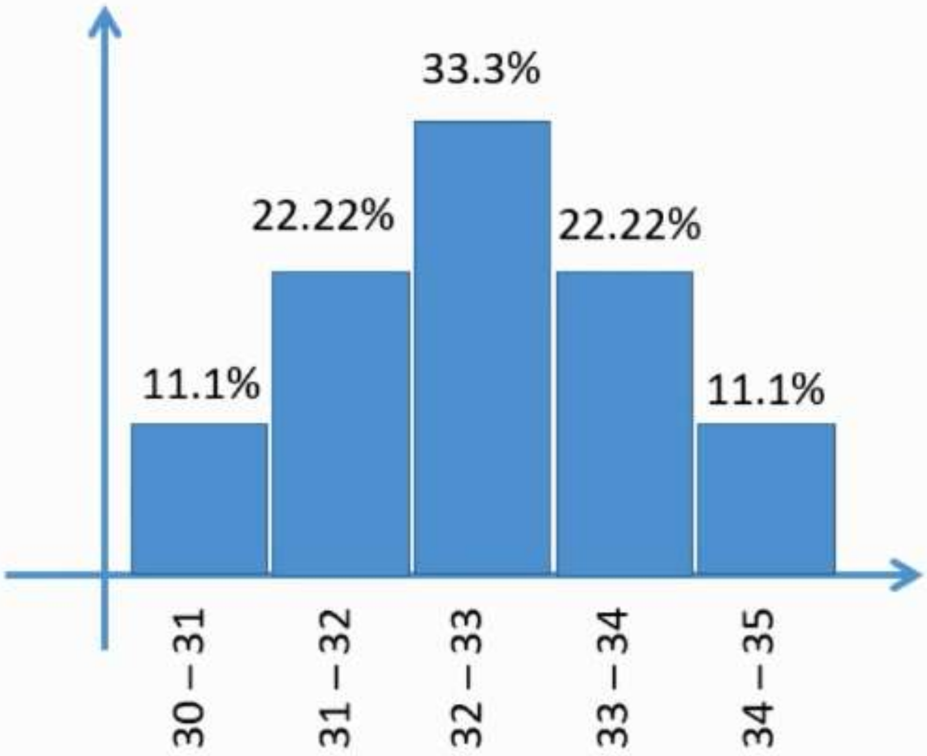
34-35 → 1/9 = 11.1%

Distribution of Continuous Variable

Temperature
30.6
31.4
31.2
32.1
32.2
32.7
33.4
33.8
34.6



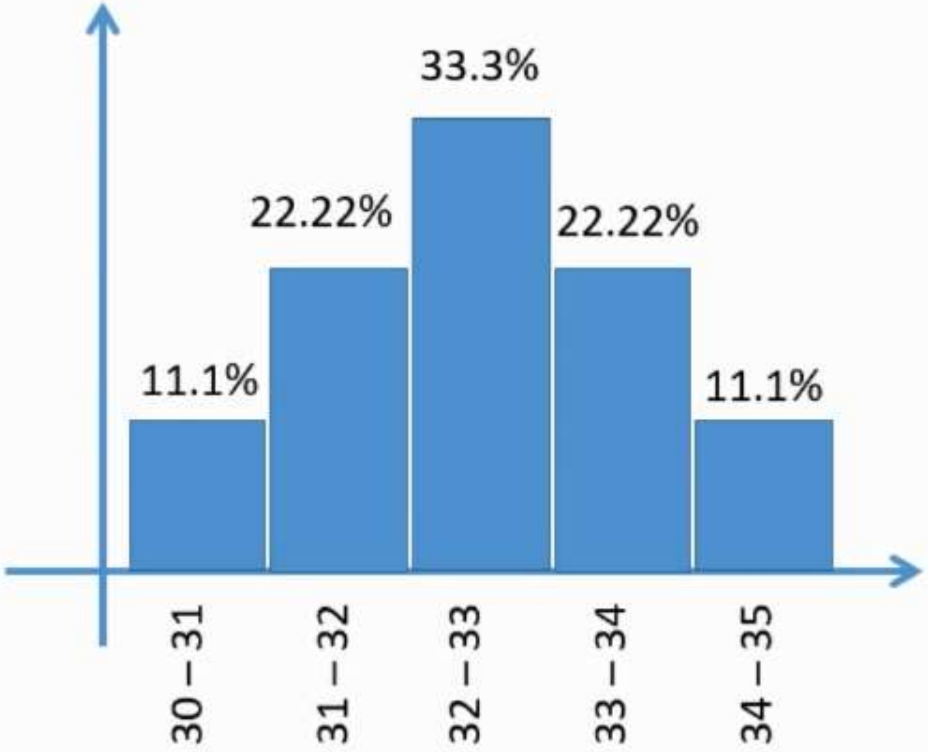
Frequency Distribution with Bins



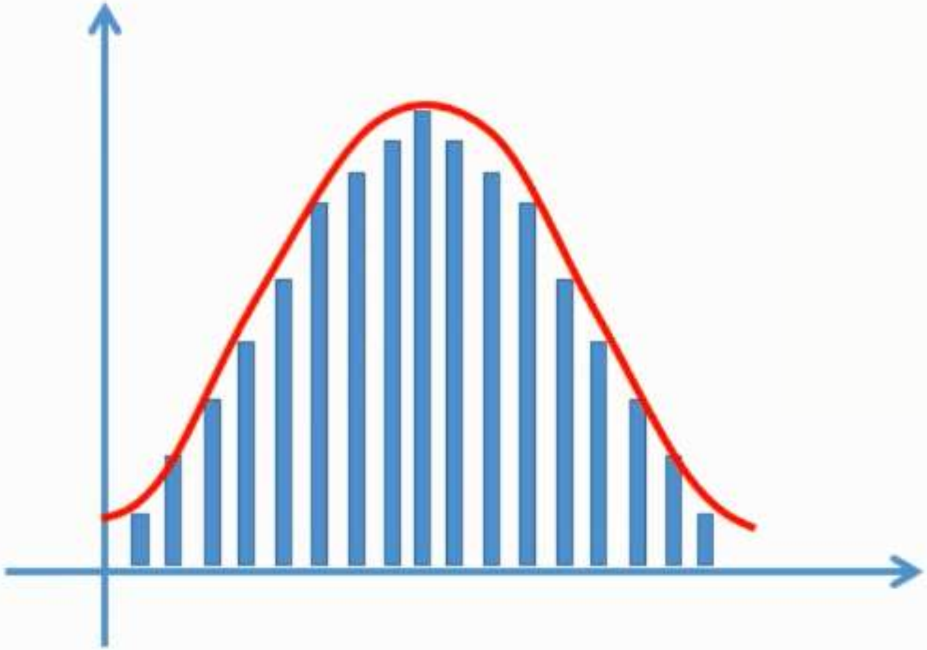
Probability of the Bins

Distribution of Continuous Variable

Temperature
30.6
31.4
31.2
32.1
32.2
32.7
33.4
33.8
34.6

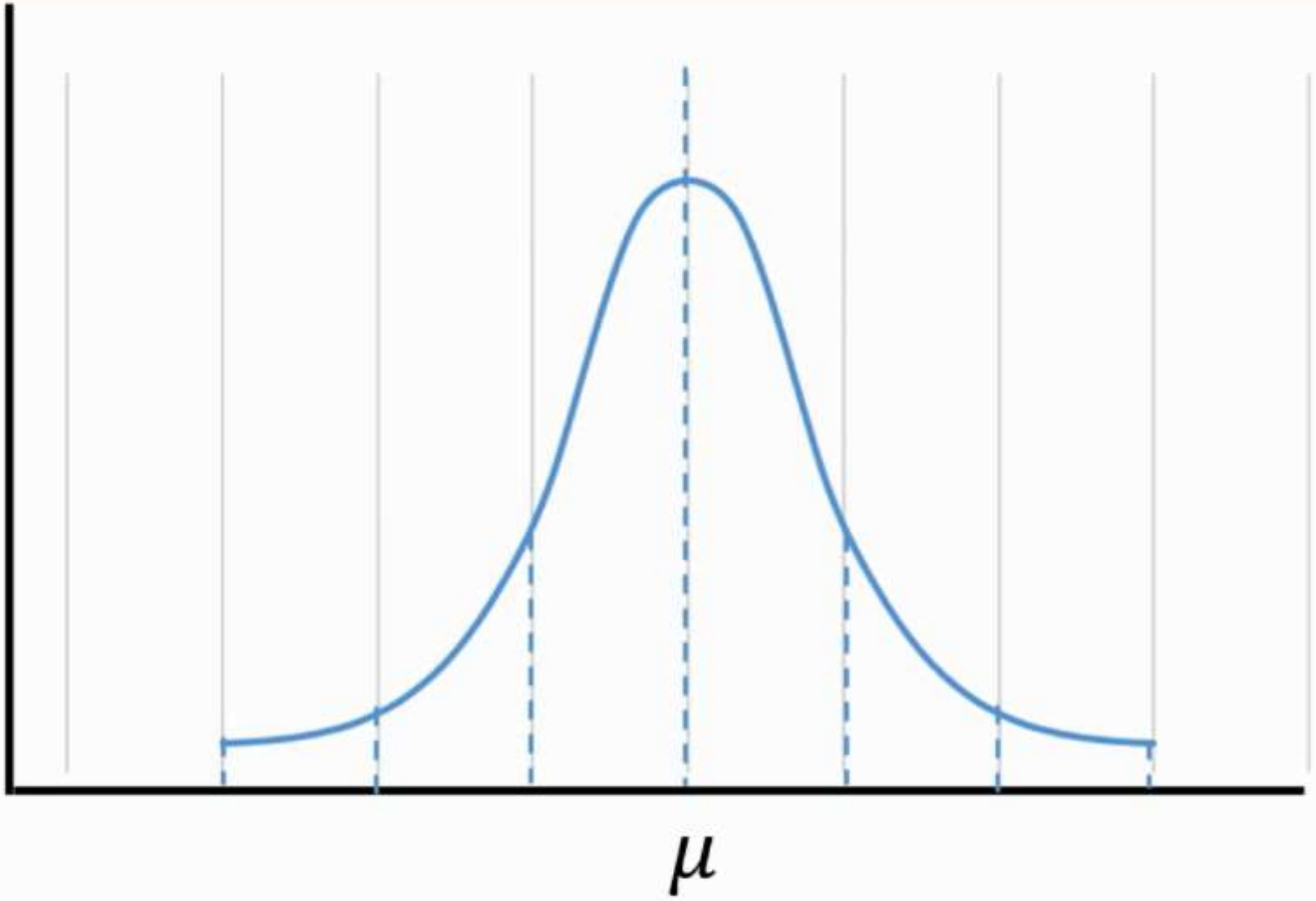


Probability of the Bins

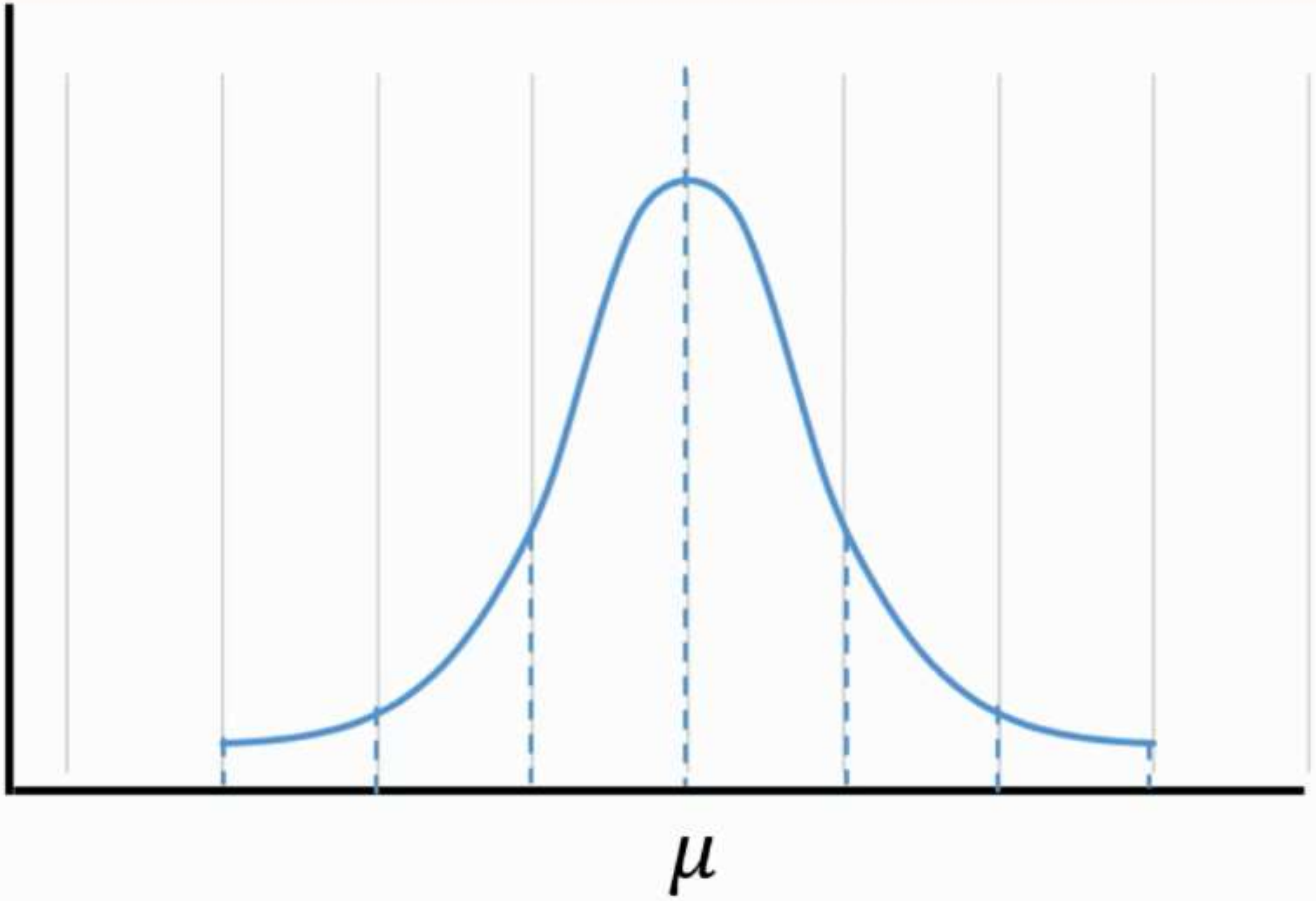


Probability Density

Normal Distribution – Bell Curve

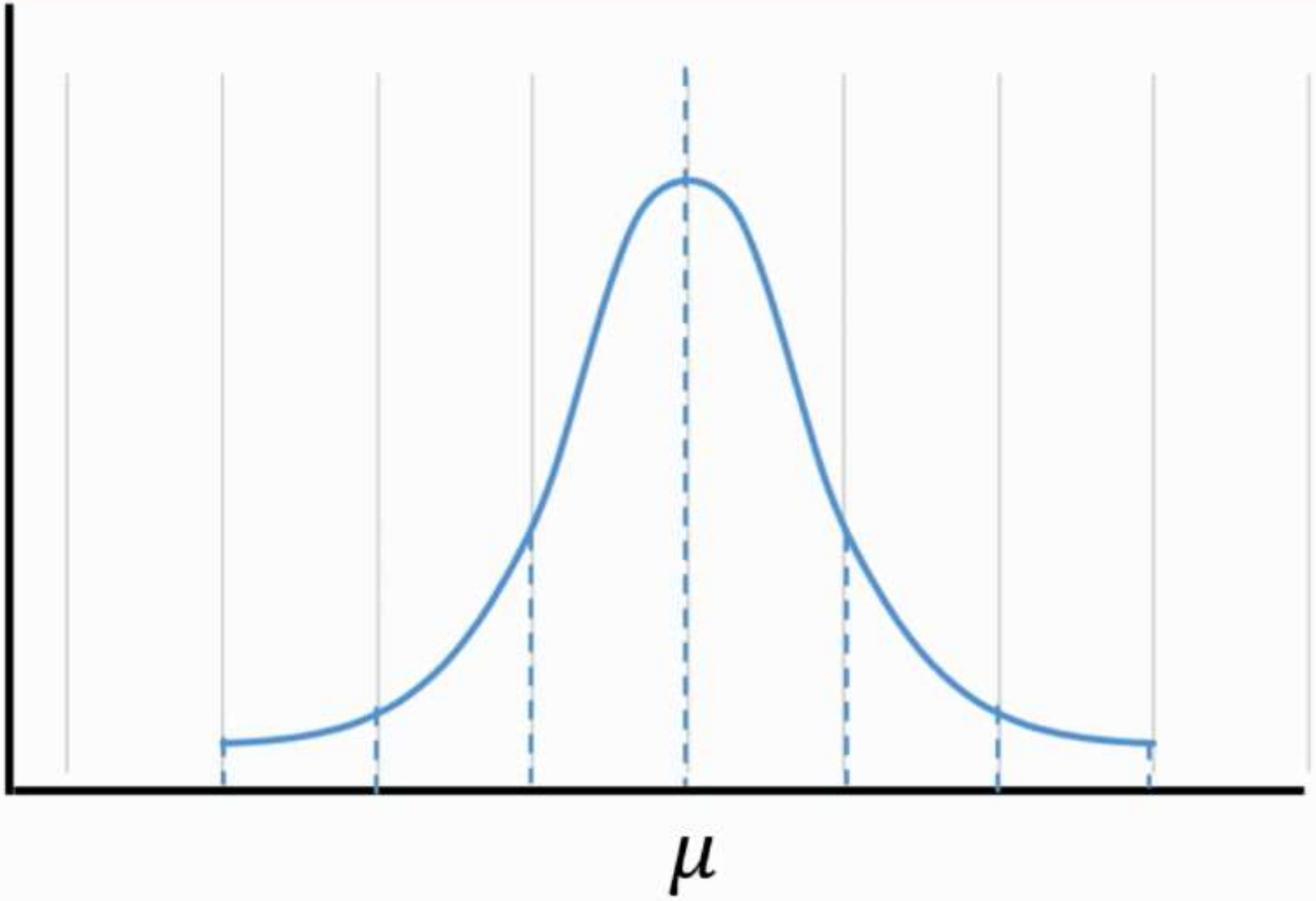


Normal Distribution – Bell Curve – Gaussian Distribution



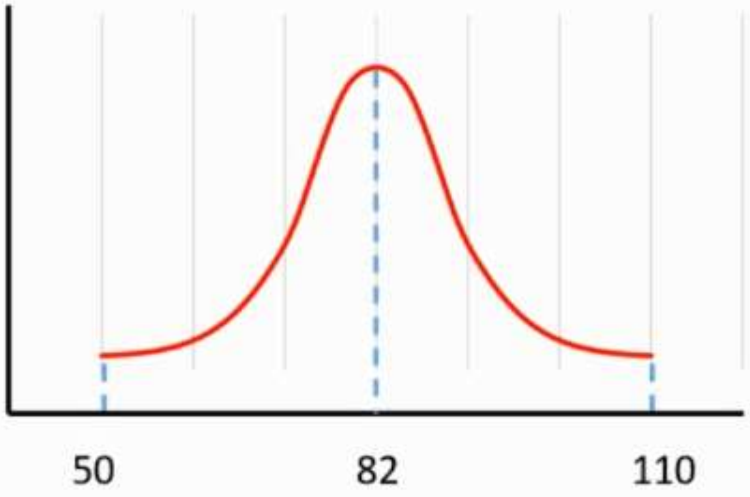
Carl Gauss

Normal Distribution – Bell Curve – Gaussian Distribution

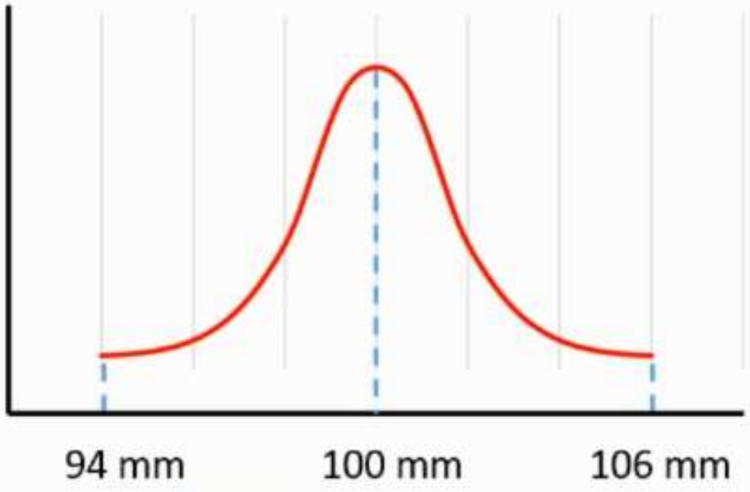


Examples of Normal Distribution

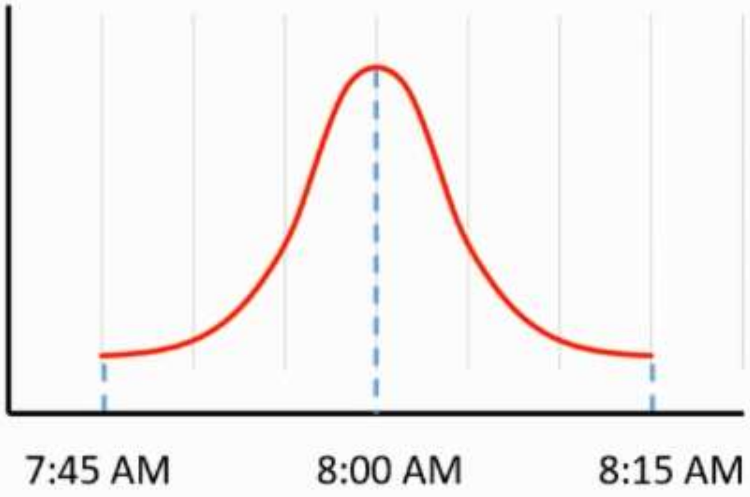
- Diastolic Blood Pressure



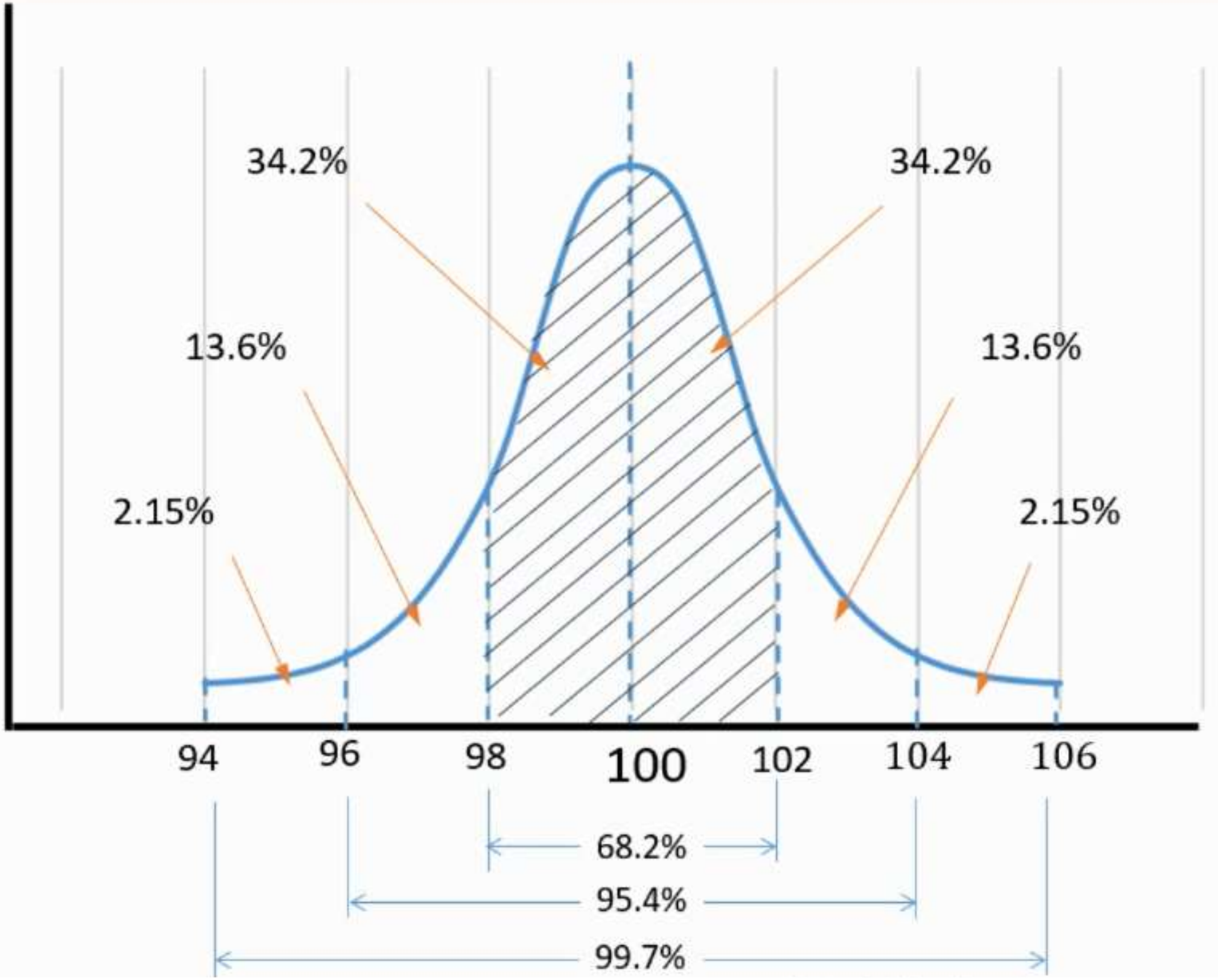
- Manufacturing



- Arrival Time at office



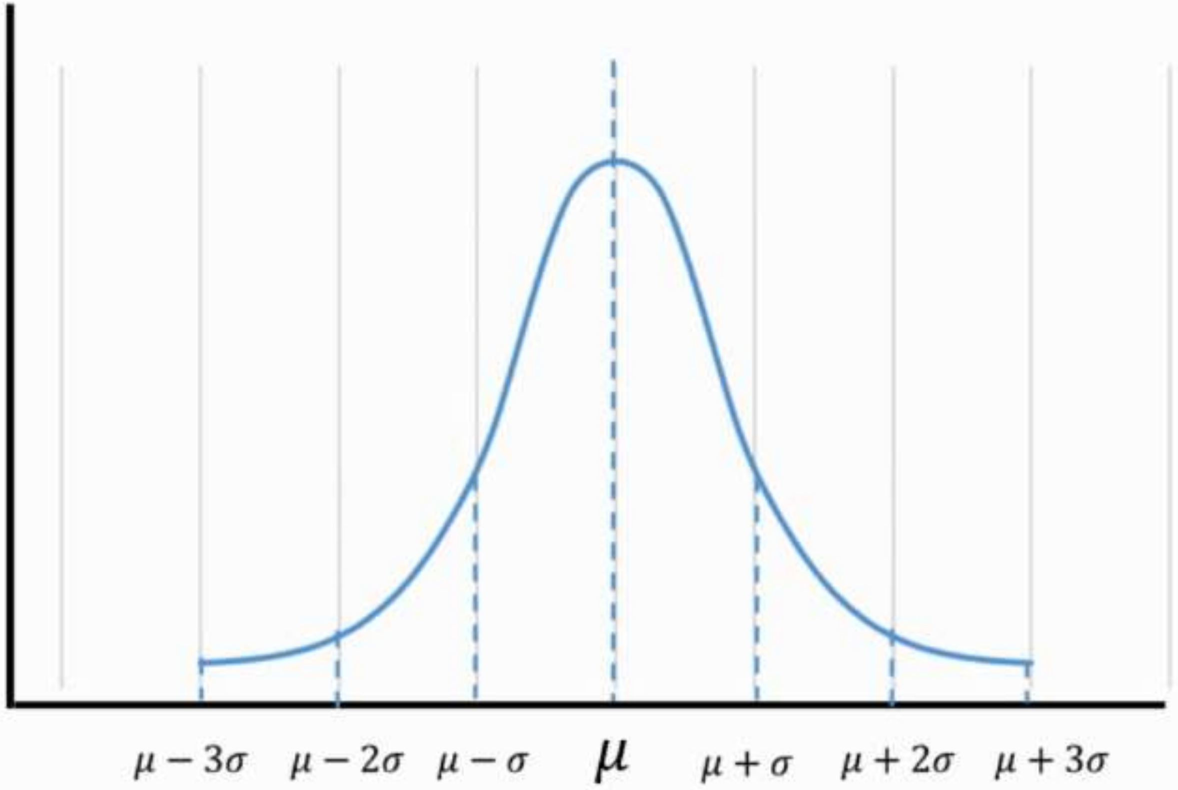
Normal Distribution – Bell Curve – Gaussian Distribution



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

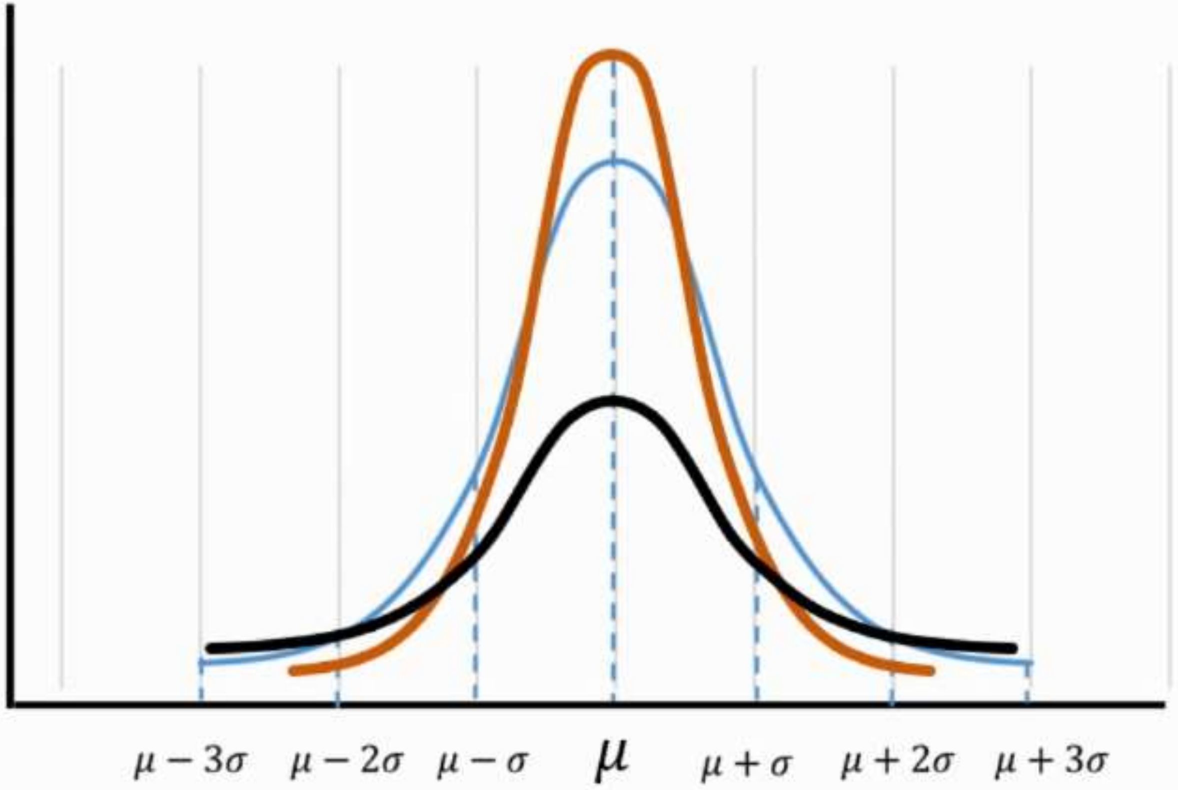
$\mu = 100$
$\sigma = 2$

Characteristics of Normal Distribution



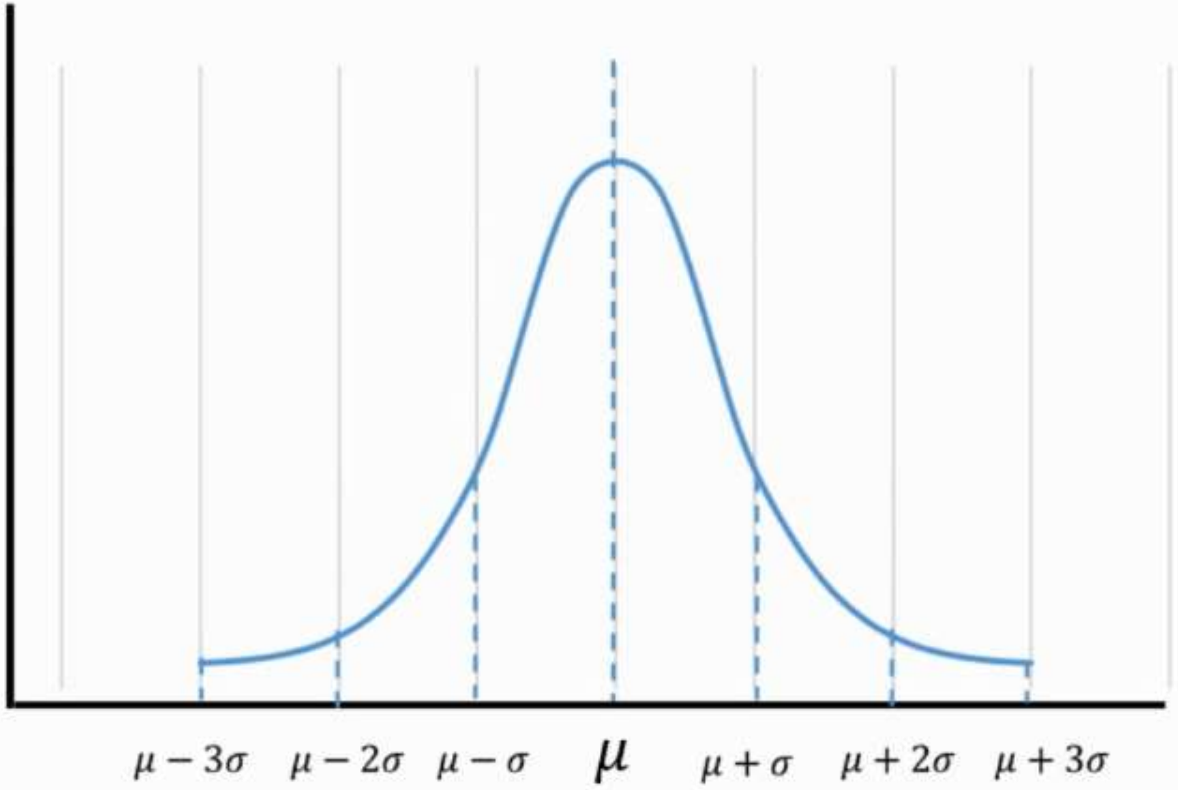
- Mean defines the centre of the graph
- Mean = Median = Mode
- Standard Deviation defines the width of the graph

Characteristics of Normal Distribution



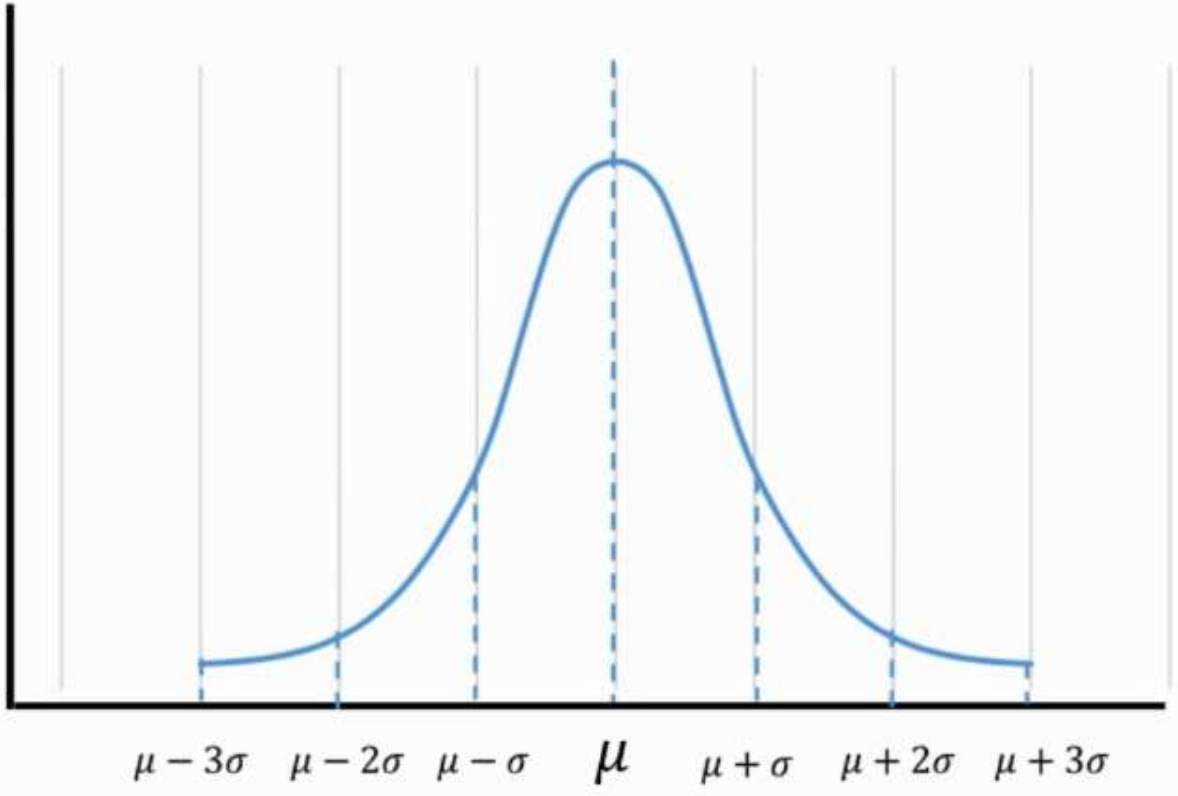
- Mean defines the centre of the graph
- Mean = Median = Mode
- Standard Deviation defines the width of the graph

Characteristics of Normal Distribution



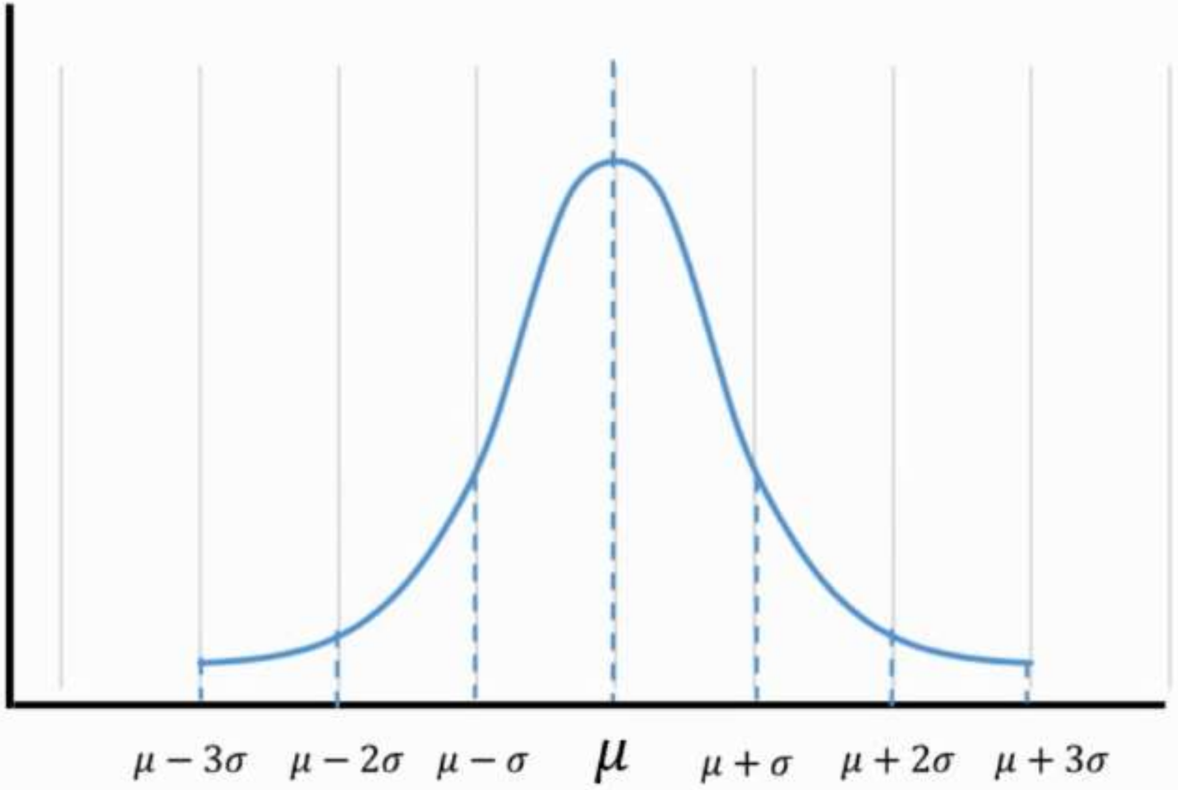
- Mean defines the centre of the graph
- Mean = Median = Mode
- Standard Deviation defines the width of the graph
- Entire distribution can be specified using mean and variance
- The total area under the curve is 1

Characteristics of Normal Distribution



- Mean defines the centre of the graph
- Mean = Median = Mode
- Standard Deviation defines the width of the graph
- Entire distribution can be specified using mean and variance
- The total area under the curve is 1
- Probability at a given point is zero

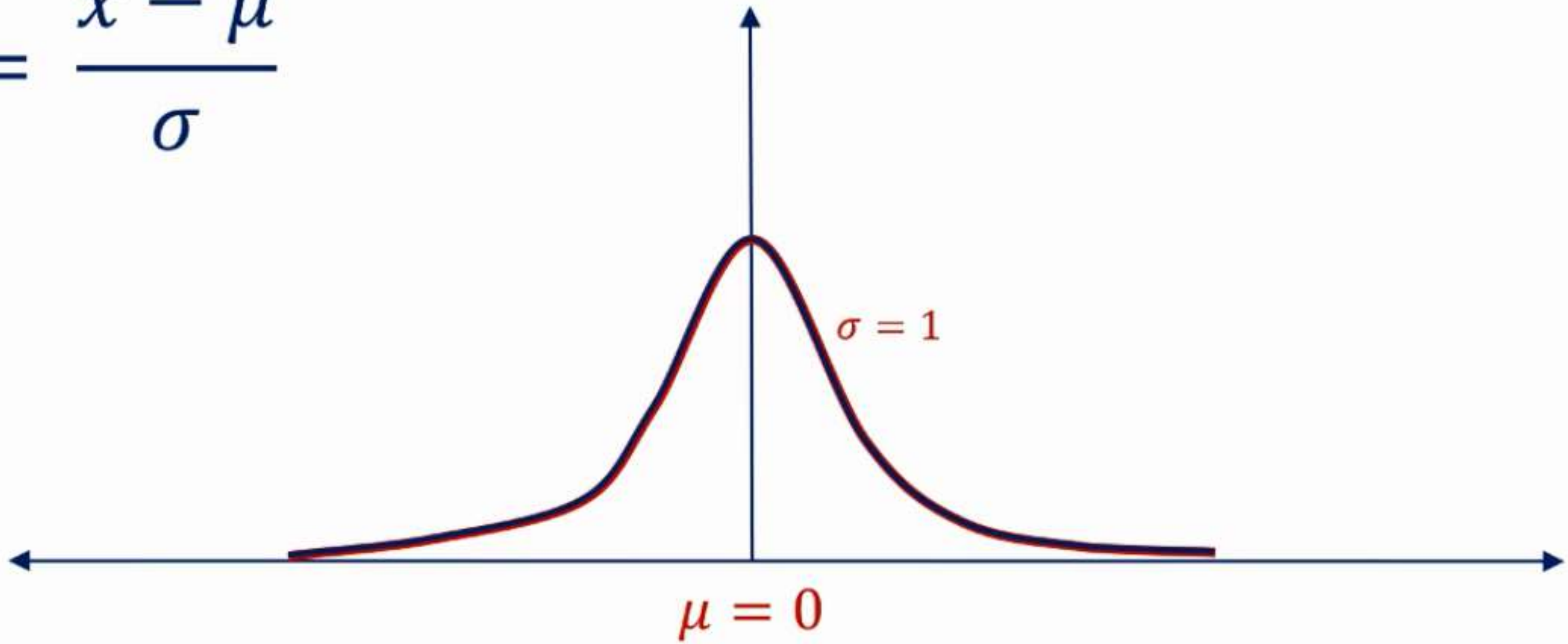
Characteristics of Normal Distribution



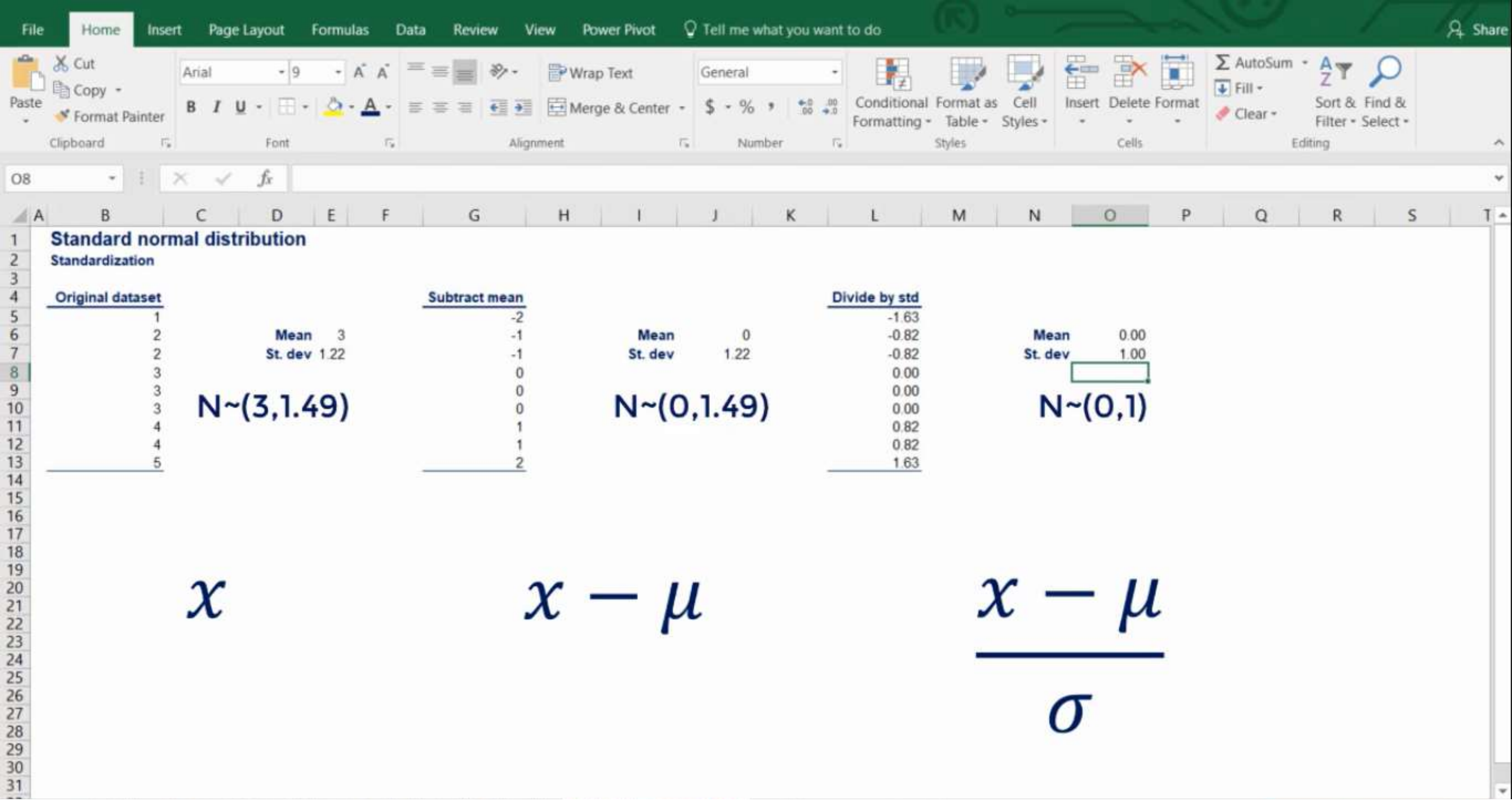
- Mean defines the centre of the graph
- Mean = Median = Mode
- Standard Deviation defines the width of the graph
- Entire distribution can be specified using mean and variance
- The total area under the curve is 1
- Probability at a given point is zero
- 68.2% of the area under the curve is within 1 σ of the mean
- 95.4% of the area under the curve is within 2 σ of the mean
- 99.7% of the area under the curve is within 3 σ of the mean

STANDARDIZATION

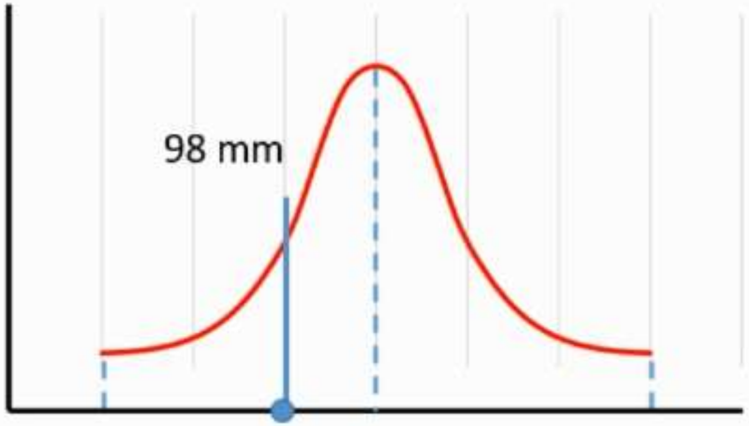
$$z = \frac{x - \mu}{\sigma}$$



Z


Standard normal distribution
Standardization
Original dataset
1
2
2
3
3
3
4
4
5
Mean 3
St. dev 1.22
N~(3,1.49)
Subtract mean
-2
-1
-1
0
0
0
1
1
2
Mean 0
St. dev 1.22
N~(0,1.49)
Divide by std
-1.63
-0.82
-0.82
0.00
0.00
0.00
0.82
0.82
1.63
Mean 0.00
St. dev 1.00
N~(0,1)
x
$x - \mu$
$\frac{x - \mu}{\sigma}$

Standard Normal Distribution



$$f(\mathrm{x}) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Area} = \lim_{\triangle \mathrm{x} \to 0} \sum_{i=1}^{n} \mathrm{f}(\mathrm{x}_i) * \triangle \mathrm{x}_i$$

$$\boxed{\int_{x1}^{x2} f(x)dx}$$

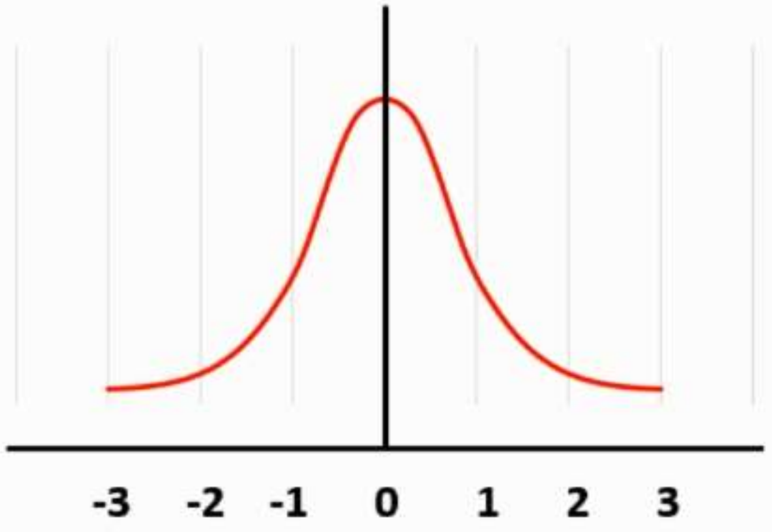
$$x2 = 98$$

$$x1 = -\infty$$

What is the probability that the diameter of the pipe will be below 98 mm?

Assuming the standard deviation is 2 mm and mean of 100 mm.

Z-Score



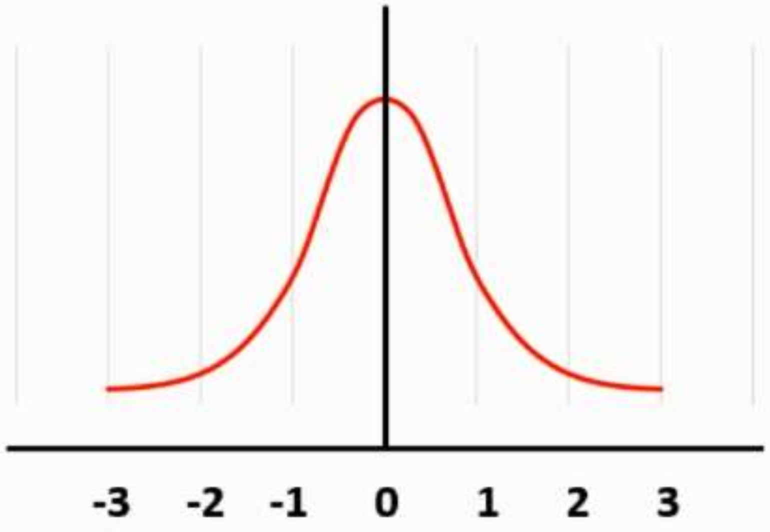
$$z = \frac{x - \mu}{\sigma}$$

For Mean,
σ = 2
x = 100 mm

$$z = \frac{100 - 100}{2} = 0$$

σ = 2 → 1
μ = 100 → 0

Z-Score



$\sigma = 2 \rightarrow 1$

$\mu = 100 \rightarrow 0$

$$z = \frac{x - \mu}{\sigma}$$

For Mean,

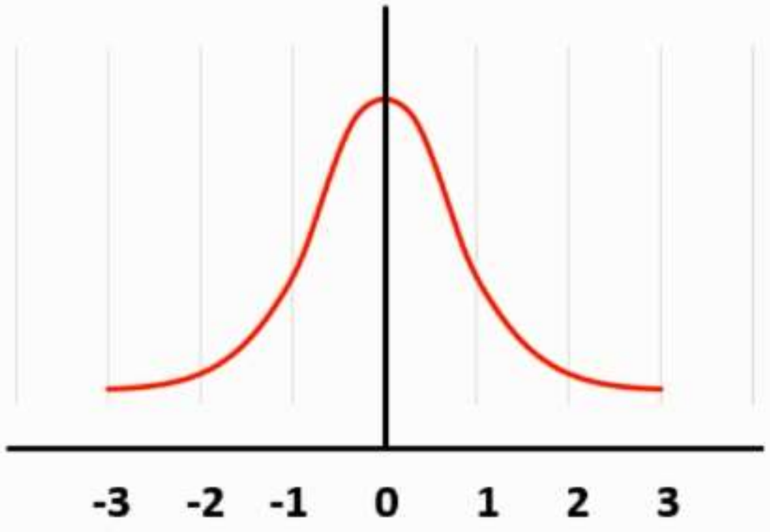
$\sigma = 2$

$x = 106$ mm

$$z = \frac{106 - 100}{2} = \underline{\underline{3}}$$

Z-Score

What is the probability that the diameter of the pipe will be below 98 mm?



$\sigma = 2 \rightarrow 1$

$\mu = 100 \rightarrow 0$

$$z = \frac{x - \mu}{\sigma}$$

For x,

$\sigma = 2$

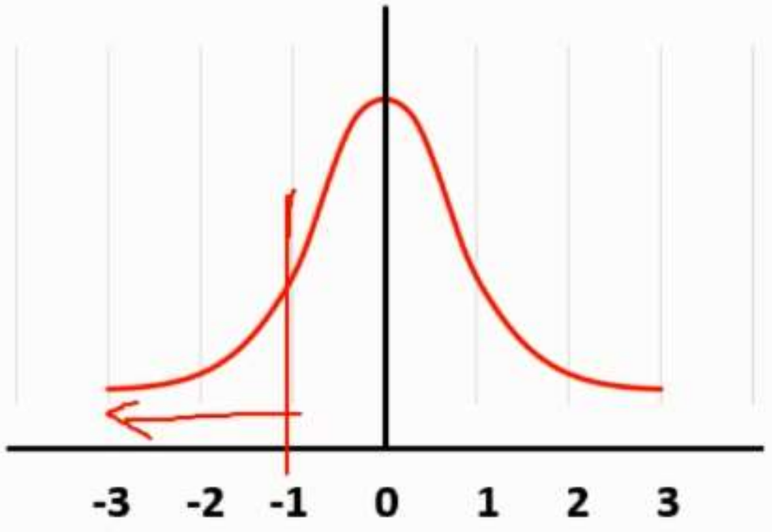
$x = 98$ mm

$$z = \frac{98 - 100}{2} = -2/2 = -1$$

Z-Score

What is the probability that the diameter of the pipe will be below 98 mm?

15.8%



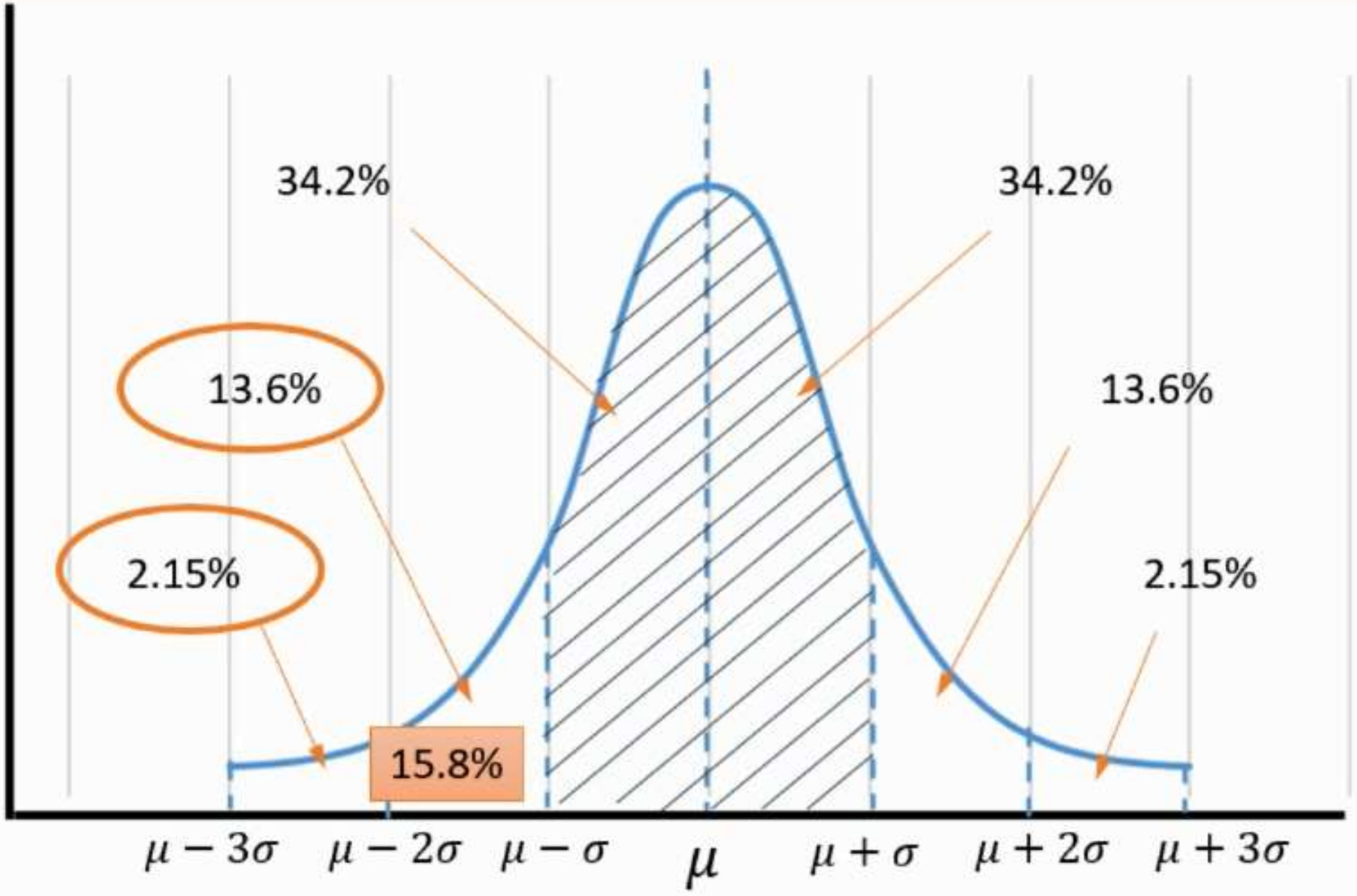
$$z = \frac{x - \mu}{\sigma}$$

For x,
σ = 2
x = 98 mm

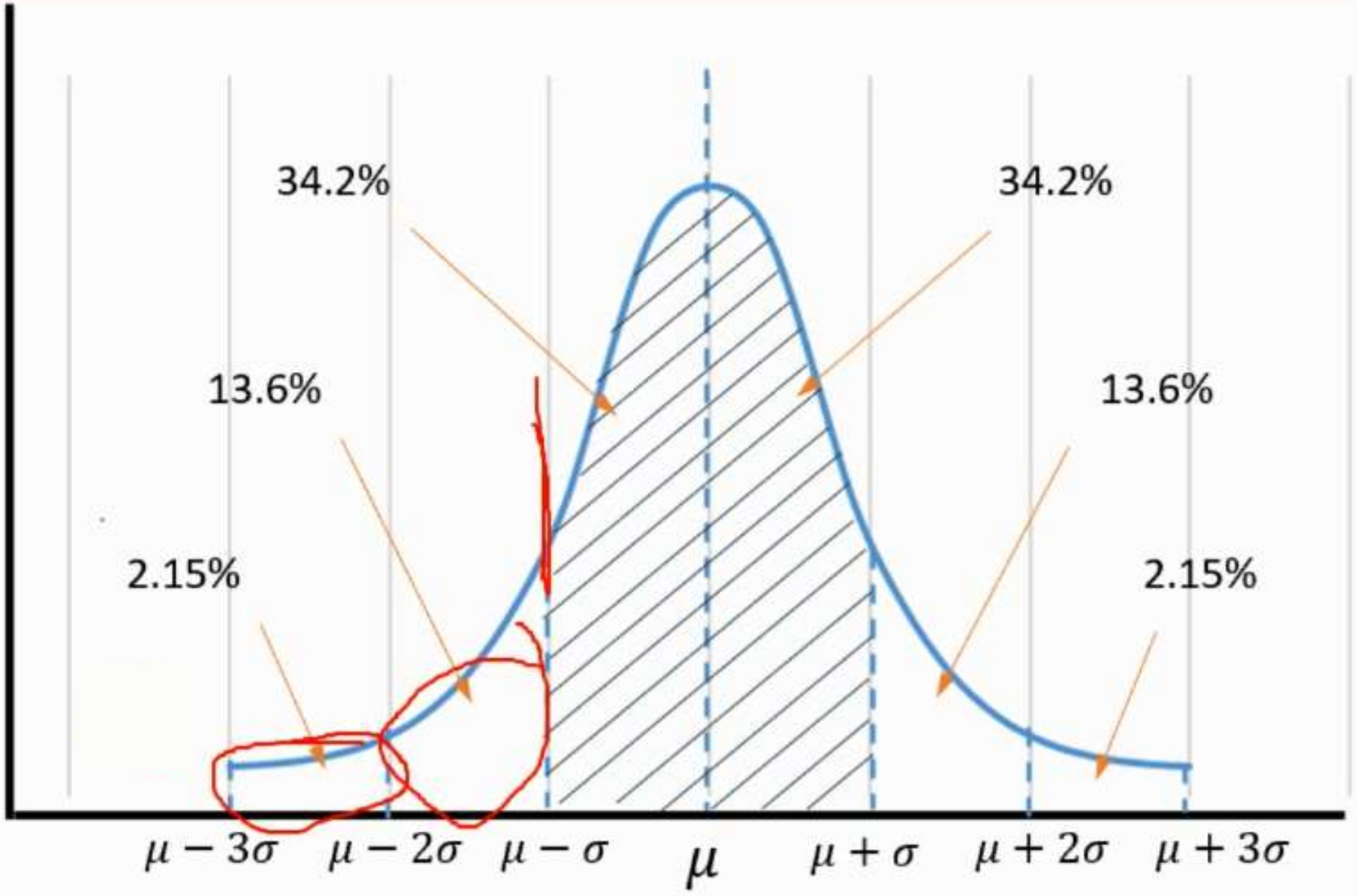
σ = 2 → 1
μ = 100 → 0

$$z = \frac{98 - 100}{2} = -2/2 = -1$$

That was easy....

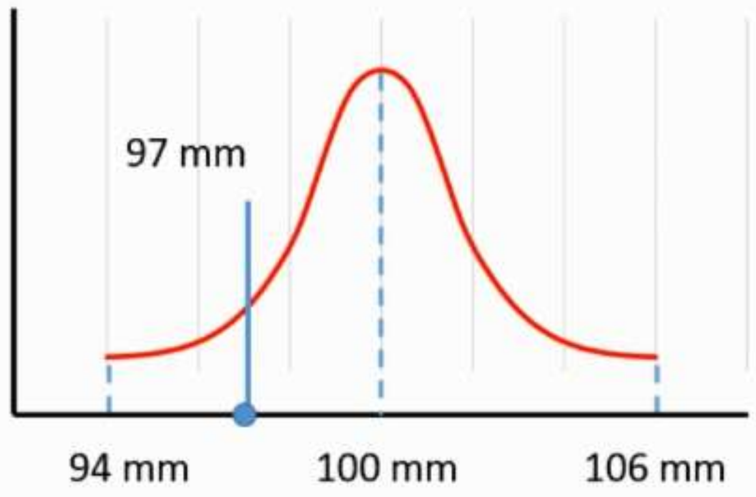


Z-Score



34.2%
34.2%
13.6%
13.6%
2.15%
2.15%
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$

Z-Score



96.5 mm

94.76 mm

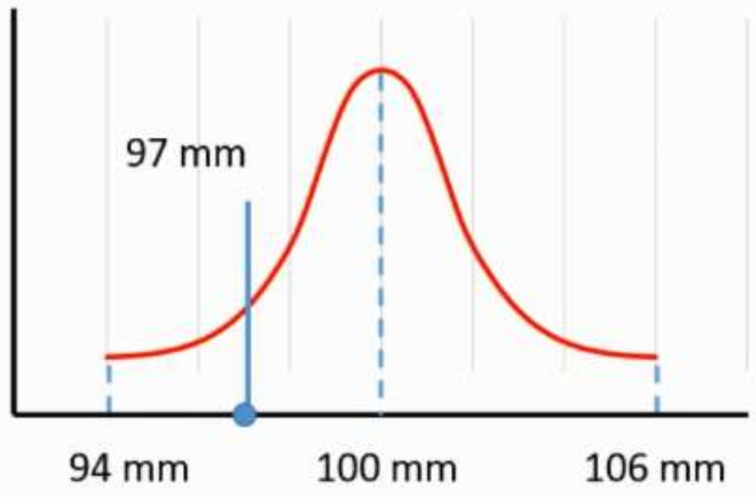
99.5 mm

101.4 mm

$\sigma = 2$
$\mu = 100$

What is the probability that the diameter of the pipe will be below 97 mm?

Z-Score



$$f(\mathrm{x}) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Area} = \lim_{\triangle x \to 0} \sum_{i=1}^{n} f(x_i) * \triangle x_i$$

$$\boxed{\int_{x1}^{x2} f(x)dx}$$

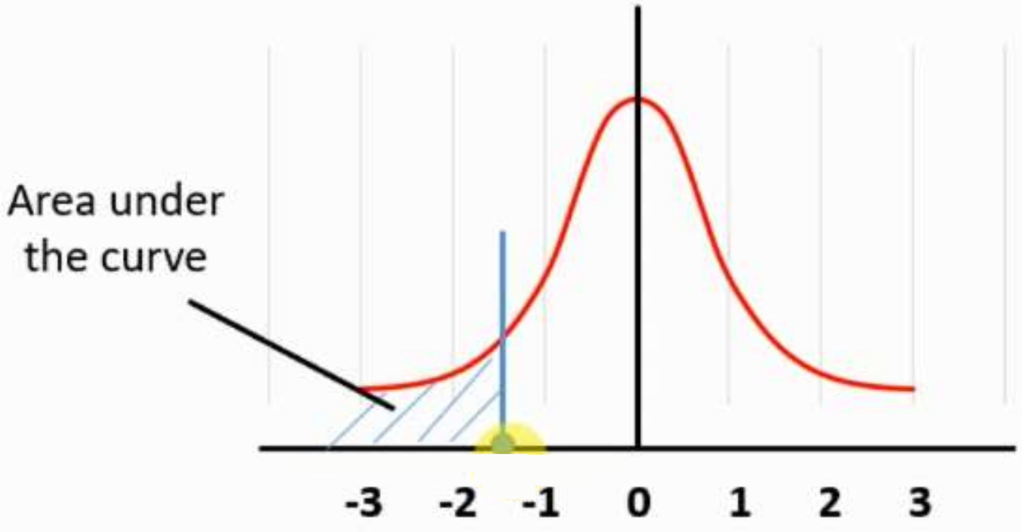
$x2 = 97$

$x1 = -\infty$

σ = 2
μ = 100

What is the probability that the diameter of the pipe will be below 97 mm?

Z-Score



$$z = \frac{x - \mu}{\sigma}$$

For X as 97,
σ = 2
x = 97 mm

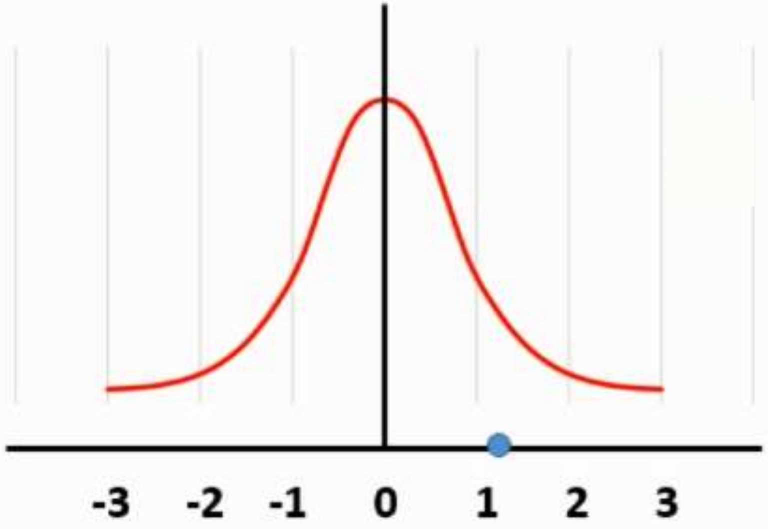
$$z = \frac{97 - 100}{2} = -3/2 = -1.5$$

σ = 2 → 1
μ = 100 → 0

What is the probability that the diameter of the pipe will be below 97 mm?

Z-Score Table

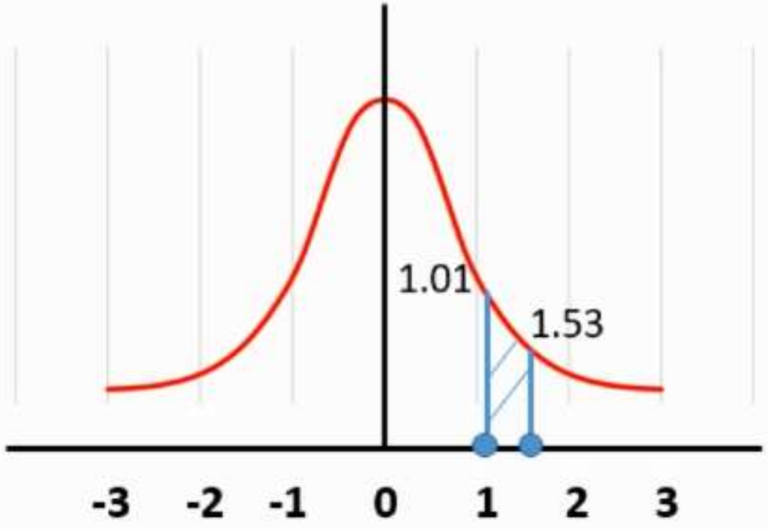
- Standard Normal Table
- Provides Cumulative Distribution Function Values



	0.00	0.01	0.02	0.03
1.00	0.841345	0.843752	0.846136	0.848495
1.10	0.864334	0.866500	0.868643	0.870762
1.20	0.884930	0.886861	0.888768	0.890651
1.30	0.903200	0.904902	0.906582	0.908241
1.40	0.919243	0.920730	0.922196	0.923641
1.50	0.933193	0.934478	0.935745	0.936992

Z-Score Table

- Standard Normal Table
- Provides Cumulative Distribution Function Values

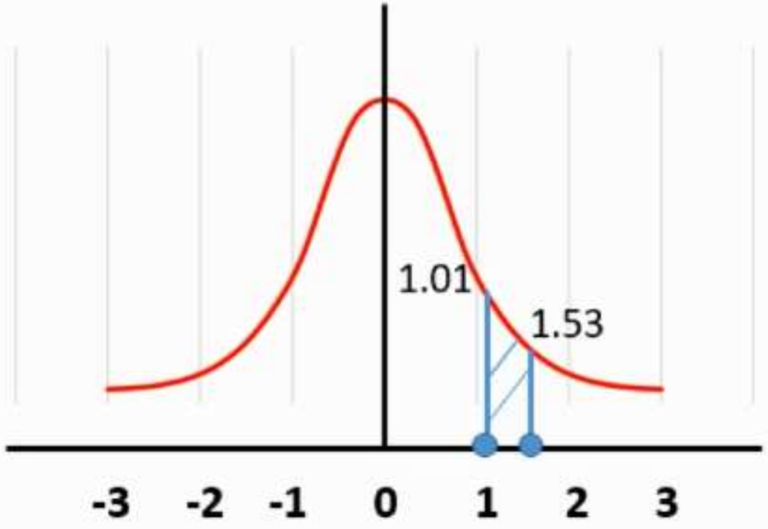


	0.00	0.01	0.02	0.03
1.00	0.841345	0.843752	0.846136	0.848495
1.10	0.864334	0.866500	0.868643	0.870762
1.20	0.884930	0.886861	0.888768	0.890651
1.30	0.903200	0.904902	0.906582	0.908241
1.40	0.919243	0.920730	0.922196	0.923641
1.50	0.933193	0.934478	0.935745	0.936992

Z-Score Table

- Standard Normal Table

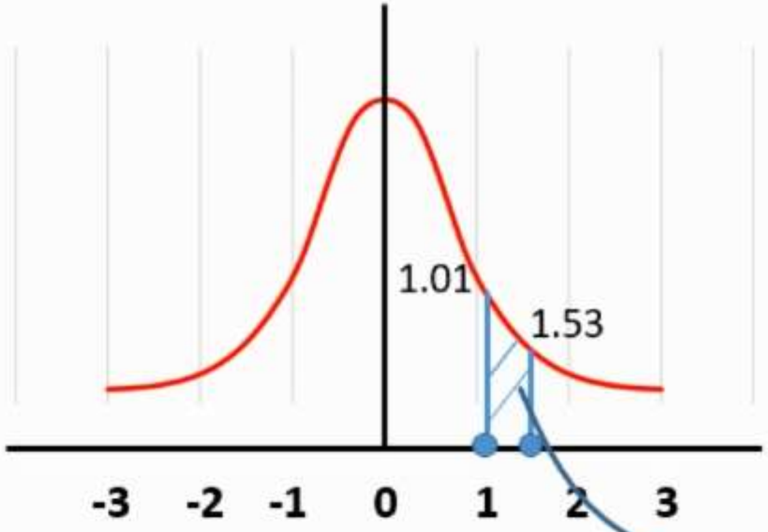
- Provides Cumulative Distribution Function Values



	0.00	0.01	0.02	0.03
1.00	0.841345	0.843752	0.846136	0.848495
1.10	0.864334	0.866500	0.868643	0.870762
1.20	0.884930	0.886861	0.888768	0.890651
1.30	0.903200	0.904902	0.906582	0.908241
1.40	0.919243	0.920730	0.922196	0.923641
1.50	0.933193	0.934478	0.935745	0.936992

Z-Score Table

- Standard Normal Table
- Provides Cumulative Distribution Function Values



	0.00	0.01	0.02	0.03
1.00	0.841345	0.843752	0.846136	0.848495
1.10	0.864334	0.866500	0.868643	0.870762
1.20	0.884930	0.886861	0.888768	0.890651
1.30	0.903200	0.904902	0.906582	0.908241
1.40	0.919243	0.920730	0.922196	0.923641
1.50	0.933193	0.934478	0.935745	0.936992

$$0.936992 - 0.843752 = 0.09324$$

Z-Score Table

What is the probability that the diameter of the pipe will be below 97 mm?

	0.00	0.01	0.02	0.03
-1.90	0.028717	0.029379	0.030054	0.030742
-1.80	0.035930	0.036727	0.037538	0.038364
-1.70	0.044565	0.045514	0.046479	0.047460
-1.60	0.054799	0.055917	0.057053	0.058208
-1.50	**0.066807**	0.068112	0.069437	0.070781
-1.40	0.080757	0.082264	0.083793	0.085343
-1.30	0.096800	0.098525	0.100273	0.102042
-1.20	0.115070	0.117023	0.119000	0.121000
-1.10	0.135666	0.137857	0.140071	0.142310
-1.00	0.158655	0.161087	0.163543	0.166023

Importance of Standard Normal Distribution and Z-Score

- Standardises the readings or scores
- Calculate the probability within the normal distribution
- Comparison of two records from different normal distribution at two different scale

Experience in years	Salary
1	$ 4,500
4	$ 7,200
4	$ 6,500
6	$ 8,500
7	$ 8,900

Quiz

Question 1:

- The high school principal asked you to conduct a survey on student satisfaction in the entire high school. You go and ask all your classmates about their opinion. Then you present the results to the principal. Was this the population or a sample drawn from it? How is the value that you presented called?

population, statistic

population, parameter

sample, statistic

sample, parameter

Question 2:

- You have a variable, representing the weight of a person. What type of data does it represent?

categorical, discrete

categorical, continuous

numerical, discrete

numerical, continuous

Question 3:

- Imagine we have a variable X, which follows a Normal Distribution with a mean of 4 and a variance of 9. We want to standardize Z, so what formula do we use for the transformation?

z = (x-4)/3

z = (4-x)/3

z = (4-x)/9

z = (4-x)/(-3)