Mathematics For Data Science

IIT MADRAS

→ Natural Numbers: $N = \{0, 1, 2, \ldots\}$

So starting from 0 to positive ∞ we can sy these are natural numbers

→ If we want to include the negative numbers, then the range become → $\{\ldots, -3, -2, -1, 0, 1, 2, 3, 4 \ldots\}$

Together they called Integers!

Representation → $Z = \{\ldots, -3, -2, -1, 0, 1, 2 \ldots\}$

→ Important algorithm for prime numbers.

→ Sieve of Eratosthenes.

→ Every ~~m~~ number can be decomposed into prime factors.

→ $12 = 2 \cdot 2 \cdot 3 = 2^2 \cdot 3$

→ $126 = 2 \cdot 3 \cdot 3 \cdot 7 = 2 \cdot 3^2 \cdot 7$

→ Rational Number: (Fraction)

→ $\frac{p}{q}$ ; p, q are integers

→ Numerator (p)

→ Denominator (q)

→ We use Q to denote Rational Numbers

Greatest Common Divisor: GCD (18, 60) ?

→ $18 = 2 \times 3 \times 3$

→ $60 = 2 \times 2 \times 3 \times 5$

$GCD = 2 \times 3 = 6$

→ Rational Numbers are dense.

For example → between 1 to 2, there can be 1.4, 1.5 . . .

there is no gap

→ Integers are discrete. We can find gap there

For example, between 1 to 2 there are no middle elements we can find in integers.

→ Smallest number that is not a perfect square is $\sqrt{2}$.

$\sqrt{2}$ cannot be written as $\frac{p}{q}$

$\sqrt{2}$ is irrational

Similarly $\pi$, e, $\sqrt{3}$, $\sqrt{5}$, $\sqrt{7}$, $\sqrt{11}$ and many more numbers are irrational. (which can not be expressed as function $p/q$)

→ Real Numbers: $\mathbb{R}$ – all rational and irrational numbers.

→ Real numbers are also dense.

→ Every natural numbers is an integer

→ Every integer is a rational number

→ Every rational number is a real number

## SET Theory

→ Set is a collection of items.

For example → Days of the week = {Sun, Mon, Tue, Wed, Thu, Fri, Sat}

→ Factors of 24 = {1, 2, 3, 4, 6, 8, 12, 24}

→ Prime below 15 = {2, 3, 5, 7, 11, 13}

→ Sets may be infinite

→ They don't have a uniform type.

→ Sets are unordered. → {1, 2, 3}, {2, 1, 3} (same)

→ Duplicates in set doesn't mater → {1, 2, 3}, {1, 1, 2, 3} (same)

→ Cardinality: Number of items in a set.

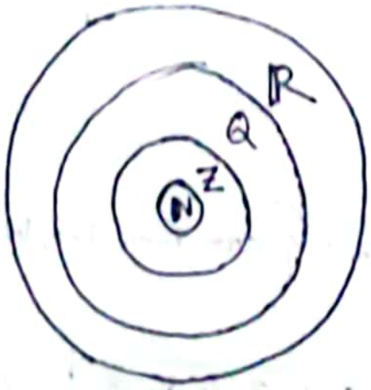
→ Elements: Items in a set are called elements.

→ Membership: $x \in X$, $x$ is an element of $X$.

**Subset:** $X \subseteq Y$ ( $\subseteq \rightarrow$ subset)

Example $\{2,3\} \subseteq \{1,2,3,4\}$

→ **Prime Numbers $\subseteq N$, $N \subseteq Z$, $Z \subseteq Q$, $Q \subseteq R$**




N = Natural Nums

Z = Integers

Q = Rational Nums

Z = Real nums

→ Every sets is a subset to **itself** → $X \subseteq X$

→ 2 sets are **equal** if they both are their subsets.

$X = Y$ if $X \subseteq Y$ and $Y \subseteq X$

→ **Proper Subset** Notation : $X \subset Y$

→ **Empty set** → $\emptyset$

→ **$\emptyset \subseteq X$** for every set $X$.

→ **Powersets:** set of subsets of a set

$X = \{a,b\}$

Powerset of $X = \{\emptyset, \{a\}, \{b\}, \{a,b\}\}$

$\{\emptyset\}$ → It is a set with one element which is an empty set.

$\emptyset$ → It is an empty set with no element.

→ Powerset of $\emptyset$ → $\{\emptyset\}$

→ Set with n element, we will have $2^n$ subsets.

## Construction of Subsets and Set operations.

**Set comprehension:**

→ The subset of even integers. $\{x \mid x \in \mathbb{Z}, \bmod 2 = 0\}$

Collect all the $x$

When $x$ belongs to Real numbers

and $x \% 2 == 0$,

→ Set of perfect squares: $\{m \mid m \in N, \sqrt{m} \in N\}$

→ Set of rationals in a reduced form

$\{p/q \mid p, q \in \mathbb{Z},\ \gcd(p,q) = 1\}$

→ Set of intervals

Integers from -6 to +6

$\{z \mid z \in \mathbb{Z},\ -6 \leq z \leq 6\}$

→ Closed Interval $[0,1] \rightarrow \{0, 0{\cdot}1, 0{\cdot}2, \ldots, 1\}$
$\{x \mid x \in \mathbb{R},\ 0 \leq x \leq 1\}$ ↳ include 0, 1

→ Open Interval $(0,1) \rightarrow \{0{\cdot}1, 0{\cdot}2, \ldots 0{\cdot}9\}$
$\{x \mid x \in \mathbb{R},\ 0 < x < 1$ ↳ Do not include 0, 1

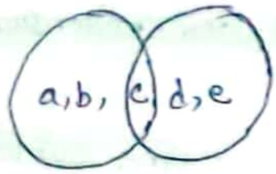
→ Left open $(0,1] \rightarrow \{x \mid x \in \mathbb{R},\ 0 < x \leq 1\}$

→ Right open $[0,1) \rightarrow \{x \mid x \in \mathbb{R},\ 0 < x < 1\}$

Operations

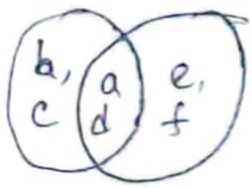
① Union → $X \cup Y$

$\{a,b,c\} \cup \{c,d,e\} = \{a,b,c,d,e\}$




② Intersection → $X \cap Y$

$\{a,b,c,d\} \cap \{a,d,e,f\} = \{a,d\}$



**Set difference:** Elements in X that are not in Y.

Represent as → $X \setminus Y$ or $X - Y$

$\{a,b,c,d\} \setminus \{a,d,e,f\} = \{b,c\}$

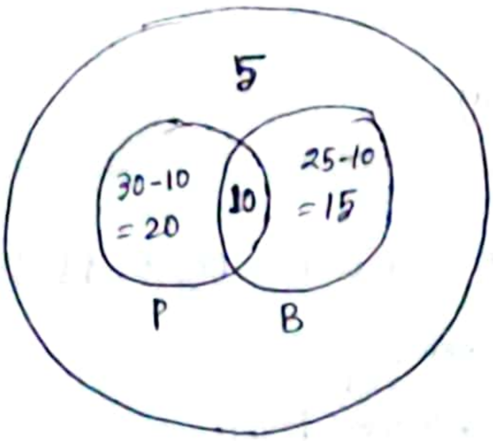
**Complement:** Suppose, $X = \{3, 9, 18, 12, 4\}$

Elements that are not in set X are complement of X

denotes by $\overline{X}$

**Questions on Set operations:**

① In a class 30 students took physics, 25 took Biology, 10 took Both and 5 took neither. How many students are there in the class?




$P = 30 - 10 = 20$

$B = 25 - 10 = 15$

$P \cap B = 10$

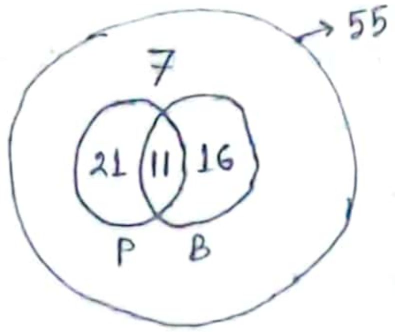
$\overline{P \cup B} = 5$

→ Total students $= 20 + 15 + 10 + 5$

$= 50$

2. In a class of 55 students, 32 students took Physics, 11 took Both Physics and biology. 7 students took neither subjects.

How many students took Biology but not Physics?




$P \cap B = 11$

$\text{Total} = 55$

$\overline{P \cup B} = 7$

$P = 32 - 11$
$= 21$

$B = 55 - (7 + 21 + 11)$
$= 16$

Also we can write like a equation in order to solve →

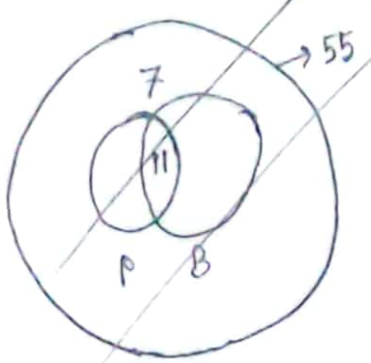
$7 + 21 + 11 + x = 55$

$\Rightarrow x = 55 - (7 + 21 + 11)$

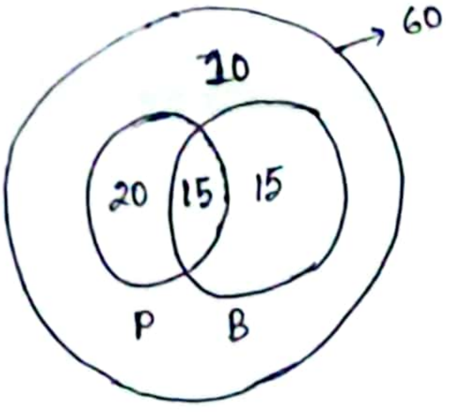
$= 16$

3. In a class of 55 students, 32 students took Physics, 11 took Both Physics and biology, and 7 took neither.

How many students took Both Biology ~~but not~~ and Physics?



(3) In a class of 60 students, 35 students took physics, 30 took Biology and 10 took neither. How many took Both physics and Biology?




$|Y|$ : Cardinality of $Y$ (Number of elements)

$|P| + |B| = 35 + 30 = 65$

$|P \cup B| = 60 - 10 = 50$

$\therefore |P \cap B| = 65 - 50$

$= 15$

Also we can solve it from the formula.

$\rightarrow |P \cup B| = |P| + |B| - |P \cap B|$

$\Rightarrow |P \cap B| = |P| + |B| - |P \cup B|$

$= 35 + 30 - 50$

$= 15$, So 15 students took both subjects.

**Cartesian Product** $\rightarrow$ all pairs $(a, b)$, $a \in A$ and $b \in B$

$A = \{1, 4, 7\}$ , $B = \{1, 16, 49\}$

$A \times B = \{(1,1), (1,16), (1,49), (4,1), (4,16), (4,49), (7,1), (7,16), (7,49)\}$

~~$B \times A = \{(1,1), (1,4), (1,7), 16,$~~

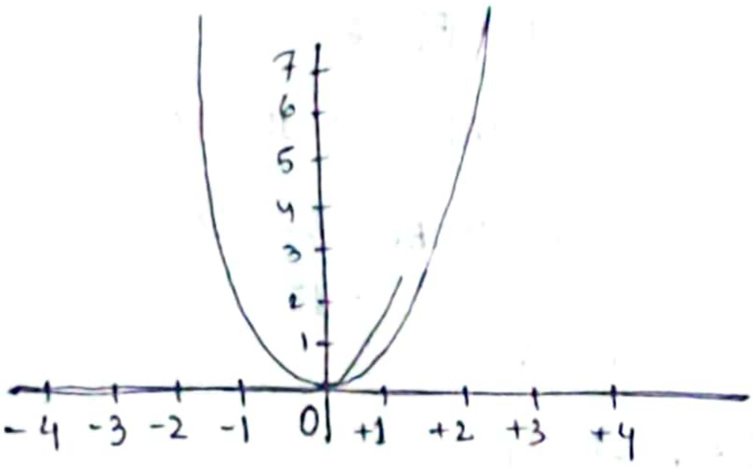
$B \times A = \{(1,1), (16,1), (49,1), (1,4), (16,4), (49,4), (1,7), (16,7), (49,7)\}$

$B \times B = \{(1,1), (1,16), (1,49), (16,1), (16,16), (16,49), (49,1), (49,16), (49,49)\}$

## Function Example:

$$y = F(x) = x^2$$

Associated relation $R_{sq} = \{(x,y) \mid x, y \in R, y = x^2\}$

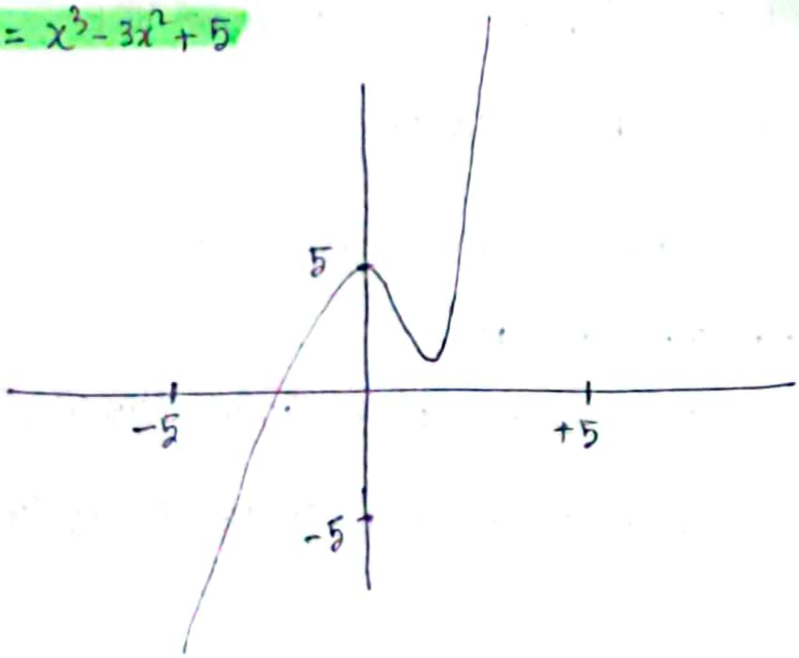



Range → $f(x) = x^2$ is always positive ∴ Range is 0 to $+\infty$

Minima → Minimum value = 0

Maxima → no maximum value.

② $y = f(x) = x^3 - 3x^2 + 5$

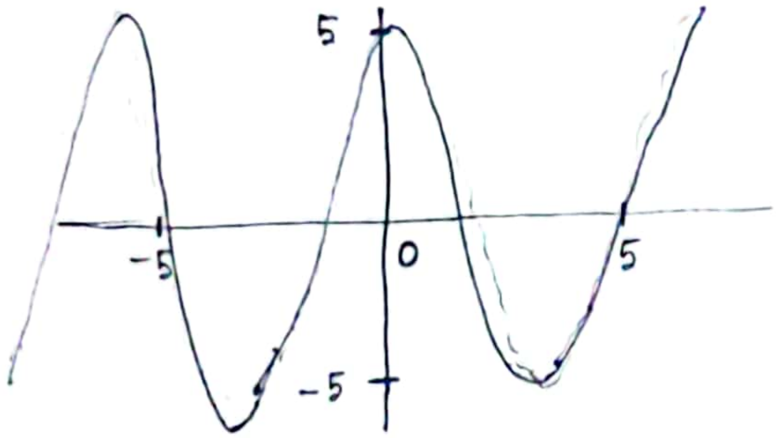


Range : $-\infty$ to $+\infty$

Minima : Has no global minima but local minima at 2

Maxima : Has no global maxima but local maxima at 0.

③ $F(x) = \cancel{\sin x}\ 5\sin(x)$




Range : bounded range from -5 to +5

Minima = Periodically attains minimum value -5, infinitely often

Maxima = Periodically attains maximum value +5, infinitely often.

Prime Numbers:

→ A prime Number P has exactly two Factors 1 and P.

→ Euclid proved, around in 300 BCE, that there cannot be a largest prime. Hence there must be infinitely many primes.