

String Functions

Padding

- Used `LPAD()` and `RPAD()`
- Helps to add characters to left or right
- Mainly used for formatting output

Substring

- `SUBSTRING()` used to get part of a string
- Example: first few letters of movie title

Concatenation

- `CONCAT()` used to join strings
- Example: first name + last name with space

String Case & Changes

- `UPPER()` → converts to uppercase
- `LOWER()` → converts to lowercase
- `REVERSE()` → reverses the string
- `LENGTH()` → gives length of string

LOCATE & SUBSTRING_INDEX

- `LOCATE()` finds position of a character
- Used with `SUBSTRING()` for dynamic extraction
- Example:
 - extracting domain from email using `@`
- `SUBSTRING_INDEX()` used when delimiter is known
- Useful when string length is not fixed

Wildcards & LIKE

- `%` wildcard used with `LIKE`
- Used to match:
 - starting pattern
 - ending pattern
 - contains pattern

Regular Expressions (REGEX)

- REGEX is more powerful than LIKE
- Used `REGEXP` keyword

Examples

- Find titles with vowels
- Titles ending with specific letters

- Names with consecutive vowels
- `^` means cap starts with (not ! symbol)

Real-time Example

- using OCR + regex
- Used regex to extract data from bank statement PDFs

Math Functions

- `POWER()` for power calculation
- `FLOOR()` rounds down
- `CEIL()` rounds up
- `ROUND()` normal rounding
- `MOD()` gives remainder
- `RAND()` generates random number

Date & Time Functions

- `DATEDIFF()` to find number of days between dates
- Extracted parts of date:
 - YEAR
 - MONTH

- DAY
- Used month name and day name functions
- Used `NOW()` with `INTERVAL`
- Example: last 24 hours data

Data Manipulation

- Used `ALTER TABLE` to add new columns
- Used `UPDATE` to fill calculated values
- Important point:
 - Never update original/source data
 - Always create new column for transformations

CASE Statement

- `CASE` works like if-else
- Used for conditional grouping
- Example:
 - grouping customers based on name range

Casting / Data Type Conversion

- Casting didn't work as expected
- Asked us to research:

- decimal \rightarrow integer
- integer \rightarrow string
- Try different syntaxes and test