SQL Window Functions Cheat Sheet

https://www.demystify.cloud/

https://www.linkedin.com/in/sourav-banerjee-50b443106/

Aggregate Window Functions

FUNCTION	WHAT IT DOES
mean()	Returns an average value
max()	Returns a maximum value
min()	Returns a minimum value
sum()	Returns total value

USE CASE EXAMPLES

Average/minimum/maximum/total:

- sales per product and/or salesperson, store, country, etc.
- users' activity (post, like, comment) compared with users in other cities or countries
- streams per artist and/or month, time period (day, week, month, quarter, year), city, country, user, etc.
- salary by department, branch, city, country compared with the company's overall salary
- price per stock and time period
- orders by customer, time period, product, compared to other or all orders
- price paid for a ride between cities by the user, driver, time period, etc.

Aggregate Window Functions

FUNCTION	WHAT IT DOES
count()	Returns the number of times an element appears in a list or a string

USE CASE EXAMPLES

Count:

- the number of employees by department, years of experience, salary range, etc.
- the number of different items ordered by time period, customer, product, etc.
- the number of logins in an app by user, time period, location
- the number of likes, comments, posts by user, time period, location, etc.

Ranking Window Functions

FUNCTION	WHAT IT DOES
reset_indix()	Counts the number of rows across the entire data frame

USE CASE EXAMPLES

Rank:

- the employees or departments by salary
- the number of customers by time period
- the users by activity
- products by sales

Ranking Window Functions

FUNCTION	WHAT IT DOES
cumcount()	Counts the number of rows across the data groups
rank()	Ranks values based on a certain variable; ranking numbers may be skipped
rank(method = 'dense')	Ranks values based on a certain variable; ranking numbers are not skipped

USE CASE EXAMPLES

Rank:

- the employees by salary for every department
- the number of new customers for every time period, location, customer category, etc.
- the users' activity for every user category
- products by sales for every product category

Ranking Window Functions

FUNCTION	WHAT IT DOES
rank(pct= 'True')	Percentile representation of the ranks compared to the highest rank

USE CASE EXAMPLES

Find the:

- kth percentile of fraudulent insurance claims
- kth percentile of hours spent on an app
- kth percentile of orders
- kth percentile in every situation where the above ranking functions are used

Ranking Window Functions

FUNCTION	WHAT IT DOES
qcut()	Allows ranking based on quantiles beyond percentiles

USE CASE EXAMPLES

Find the:

- kth quantile of fraudulent insurance claims
- kth quantile of hours spent on an app
- kth quantile of orders
- kth quantile in every situation where the above ranking functions are used

Value Window Functions

FUNCTION	WHAT IT DOES
shift()	Represents value from another column but shifted by a single or multiple preceding or following rows

USE CASE EXAMPLES

Find the:

- rate of growth by comparing the current and previous period's number of hosts, users, clients, products sold, etc
- previous date of inspection, login, sales, hiring, etc.
- realized or budgeted costs, sales, number of employees for the current and the n future time periods

Value Window Functions

FUNCTION	WHAT IT DOES
nth()	Finds the first/last/nth value within the groups of a dataset

USE CASE EXAMPLES

Find the highest/lowest/nth:

- salary within a department or a company
- sales per product, salesperson, branch, time period, etc.
- budget or duration of a project
- distance driven
- date of sales, login, account creation, order, etc.