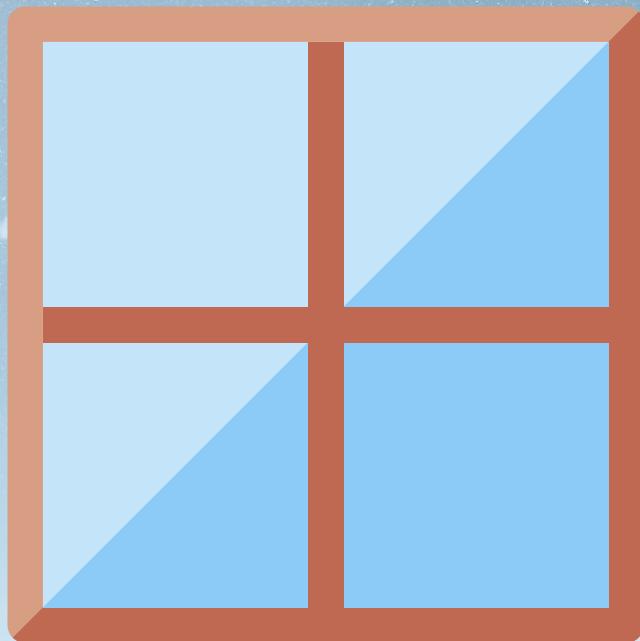
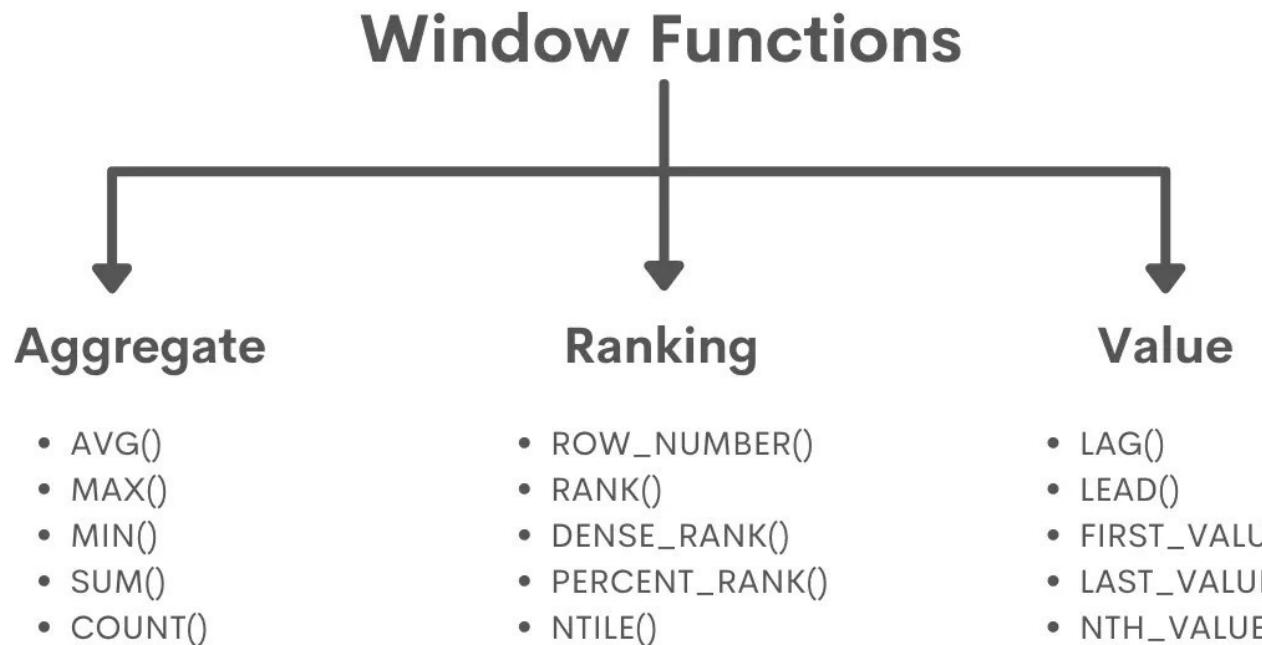


# SQL - Window Functions



# Types of Window Functions



## Rank(), Dense\_Rank(), Row\_Number()

Functions used - Rank(), Dense\_Rank(), Row\_Number()

Question - List the ten most valuable Asian startups outside China

```
Select Company, Industry, Country, Valuation_$Bn,  
Rank() Over(Order by Valuation_$Bn desc) as RK,  
Dense_Rank() Over(Order by Valuation_$Bn desc) as DRK  
Row_Number() Over(Order by Valuation_$Bn desc) as Row_No  
From unicorn_companies  
Where Continent = 'Asia' and Country <> 'China'  
Limit 10
```

Result -

#	Company	Industry	Country	Valutn(\$Bn)	RK	DRK	Row_No
	BYJUs	Edtech	India	22	1	1	1
	J&T Express	Supply Chain	Indonesia	20	2	2	2
	Swiggy	Supply Chain	India	11	3	3	3
	OYO Rooms	Travel	India	10	4	4	4
	Lalamove	Supply chain	Hong Kong	10	4	4	5
	Dunamu	Fintech	S_Korea	9	6	5	6
	Yanolja	Travel	S_Korea	9	6	5	7
	Dream11	Software	India	8	8	6	8
	Razorpay	Fintech	India	8	8	6	9
	NSE India	Fintech	India	7	10	7	10



## Percent\_Rank()

```
/* Percent_Rank() - This window function is used to calculate the percentile ranking of rows within a query result. It always returns values ranging from 0 to 1, with 1 being the highest(100) percentile */

-- Select the top 5 percentile of companies funded in each continent
-- Query
Select *
From (
Select Company, Industry, Continent, Funding_$Bn,
Percent_Rank() Over (Partition By Continent Order by
Funding_$Bn) as Funding_Percentile
From unicorn_companies)
Where Funding_Percentile >= 0.95

--Result
# Company      Industry    Continent   Funding_$Bn   Funding_Percentile
Opay          Fintech     Africa      0.57        1
SVOLT         Mobility    Asia        3            0.964
OYO Rooms    Travel      Asia        3            0.964
Zuoyebang    Edtech     Asia        3            0.964
Ola Cabs      Mobility    Asia        4            0.974
BYJUS         Edtech     Asia        4            0.974
Yuanfudao    Edtech     Asia        4            0.974
Swiggy        SCM         Asia        5            0.990
Epic Games   Gaming      N_America  7            0.996
SpaceX        Other       N_America  7            0.996
JUUL Labs    Consumer    N_America  14           1
Airwallex    Fintech    Oceania    0.8          1
C6 Bank       Fintech    S_America  2            0.95
Rappi         SCM         S_America  2            0.95
```



## Ntile()

```
/* Ntile() - This function is similar to the Row_Number() function.  
Difference being instead of assigning row_numbers, Ntile() divides rows  
into a given number of approximately equal groups then assigns  
numbers to these groups starting from 1 */  
  
-- Split South American start ups into 5 buckets based on valuation  
  
-- Query  
Select Company, Industry, Country, Funding_$Bn,  
Ntile(5) Over ( Order by Funding_$Bn Desc) as Funding_Group  
From unicorn_companies  
Where Continent = 'South America'  
  
--Result  


| # | Company   | Industry   | Country   | Funding_\$Bn | Funding_Group |
|---|-----------|------------|-----------|--------------|---------------|
|   | Rappi     | SCM        | Colombia  | 2            | 1             |
|   | C6 Bank   | Fintech    | Brazil    | 2            | 1             |
|   | Creditas  | Fintech    | Brazil    | 0.83         | 1             |
|   | Loft      | E-commerce | Brazil    | 0.79         | 1             |
|   | Neon      | Fintech    | Brazil    | 0.72         | 2             |
|   | Nuvemshop | E-commerce | Brazil    | 0.63         | 2             |
|   | iFood     | SCM        | Brazil    | 0.59         | 2             |
|   | Uala      | Fintech    | Argentina | 0.54         | 3             |
|   | Loggi     | SCM        | Brazil    | 0.51         | 3             |
|   | EBANX     | Fintech    | Brazil    | 0.46         | 3             |
|   | CargoX    | SCM        | Brazil    | 0.39         | 3             |
|   | NotCo     | AI         | Chile     | 0.37         | 4             |
|   | LifeMiles | Other      | Colombia  | 0.34         | 4             |
|   | Madeira   | E-commerce | Brazil    | 0.34         | 4             |
|   | Olist     | E-commerce | Brazil    | 0.32         | 5             |
|   | Unico     | AI         | Brazil    | 0.24         | 5             |
|   | Betterfly | AI         | Chile     | 0.21         | 5             |


```



## Lead() and Lag()

```
/* Lead() and Lag() window functions are used to access the preceding  
and succeeding value(s) of specified rows from the current row  
within its partition.
```

These functions are written a bit differently because they usually contain offset and default values in their syntax

```
-- Offset - Number of rows preceding or succeeding from the current row.  
It should be a positive integer value.  
If it is zero, the function evaluates the result for the current row.  
If we omitted, the function uses 1 by default.
```

```
-- Default_Value - Value that appears when there is no preceding  
or succeeding value.  
If omitted, the function returns a null value */
```

```
-- Syntax Example  
Lead 'or' Lag(Col_Name, Offset , Default_Value)  
Over (Partition By (Col_Name) Order By (Col_Name)
```



## Lead() Example

```
-- Lead() Example
-- Difference in Industry Funding compared to Next Industry

-- Query
Select Industry, Sum(Funding_$Bn) as Industry_Funding,
Lead(Sum(Funding_$Bn),1) Over(Order by Sum(Funding_$Bn)) as Next_Industry,
Lead(Sum(Funding_$Bn),1) Over(Order by Sum(Funding_$Bn)) -
SUM(Funding_$Bn) as Funding_Diff
From unicorn_companies
Group by 1
Order by 2

-- Result
# Industry      Industry_Funding  Next_Industry  Funding_Diff
Travel          12.64           15.14          2.5
Mobile          15.14           18.18          3.04
Data_Mgt        18.18           21.42          3.24
Hardware        21.42           21.88          0.46
Edtech          21.88           21.96          0.08
Cybersecurity   21.96           25.79          3.83
Consumer        25.79           33.98          8.19
Health          33.98           35.06          1.08
Automobile      35.06           40.89          5.83
Other           40.89           44.76          3.87
Supply chain    44.76           46.87          2.11
AI              46.87           73.50          26.63
Software        73.5            75.48          1.98
E-commerce      75.48           108.72         33.24
Fintech         108.72          Null           Null
```



## Lag() Example

```
-- Lag() Example
-- Calculate Year on Year change in No of New Unicorns

-- Query
Select Year(Joined) as Year_, Count(Year(Joined)) as No_of_Unicorns,
Lag(Count(Year(Joined)), 1, 0) Over(Order by Year(Joined)) as Previous_Year,
Count(Year(Joined)) - Lag(count(Year(Joined)), 1, 0)
Over(Order by (Year(Joined))) as YonY_Change
From unicorn_companies
Group by 1
Order by 1

-- Result
# Year_  No_of_Unicorns  Previous_Year  YonY_Change
  2007    1              0              1
  2011    2              1              1
  2012    4              2              2
  2013    3              4             -1
  2014   13              3              10
  2015   35             13              22
  2016   21             35             -14
  2017   44             21              23
  2018  103             44              59
  2019  104            103              1
  2020  108            104              4
  2021  519            108             411
  2022  116            519            -403
```



## First\_Value() and Last\_Value()

```
/* First_Value() and Last_Value() functions does exactly what their names says  
They extract the First/Last value obtained by the expression calculated  
in the window partition */
```

```
-- Companies with > 20X Returns showing highest and lowest return by Industry  
-- Query
```

```
Select Company, Industry, Country, (Valuation_$Bn/Funding_$Bn)as ROI,  
First_Value(Company) Over (Partition By Industry Order by  
Valuation_$Bn/Funding_$Bn Desc) as Highest_Return,  
Last_Value(Company) Over (Partition By Industry Order by  
Valuation_$Bn/Funding_$Bn Desc  
RANGE BETWEEN UNBOUNDED PRECEDING AND UNBOUNDED FOLLOWING) as Lowest_Return  
From unicorn_companies  
Where (Valuation_$Bn/Funding_$Bn) > 20  
Order by Industry
```

```
-- Result
```

#	Company	Industry	Country	ROI	Highest_Return	Lowest_Return
	Bytedance	AI	China	22.5	Bytedance	Bytedance
	ZongMu Tech	Auto	China	28.9	ZongMu Tech	ZongMu Tech
	L&P Cosmetic	Consumer	S_Korea	33.3	L&P Cosmetic	Genki Forest
	Genki Forest	Consumer	China	20.8	L&P Cosmetic	Genki Forest
	Il Makiage	E-commerce	USA	66.7	Il Makiage	eDaili
	SHEIN	E-commerce	China	50	Il Makiage	eDaili
	OpenSea	E-commerce	USA	30.2	Il Makiage	eDaili
	Xiaohongshu	E-commerce	China	21.7	Il Makiage	eDaili
	eDaili	E-commerce	China	20	Il Makiage	eDaili
	Dunamu	Fintech	S_Korea	128	Dunamu	Checkout.com
	CFGII	Fintech	USA	100	Dunamu	Checkout.com
	Upstox	Fintech	India	60	Dunamu	Checkout.com



## Nth\_Value()

```
/* Nth_Value() is a window function that extracts the
value from the
Nth row in an ordered set of rows */

-- Find the 2nd Funded City and No of companies per city
Megacities (cites more than 10bn Funding)
-- Query
Select City, Country, Count(Company) as No_of_Companies,
(Sum(Funding_$Bn),2) as Total_Funding,
Nth_Value(City,2) Over(Order by Sum(Funding_$Bn) Desc )
as 2nd_Funded_City
From unicorn_companies
Group by City
Having Total_Funding >= 10
Order by Total_Funding Desc

-- Result
City          Country    No_of_Companies  Funding   2nd_City
San Francisco USA        152             81.57    Null
Beijing       China      63              51.12    Beijing
New York     USA        103             38.52    Beijing
Bengaluru    India      29              26.9     Beijing
London        UK         34              26.3     Beijing
Shanghai      China      44              23.96    Beijing
Berlin        Germany    17              11.87    Beijing
Paris         France     19              10.78    Beijing
```