Practice on SQL Operators,
Filtering Data, Aggregate
and Sorting Functions

Relevel

by Unacademy

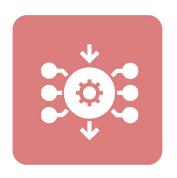


Additional Questions - Set 1

Instructions:

- We will use the mode.com for this set of questions.
- We will use three datasets with 10 questions each





We will use **tutorial.oscar_nominees dataset** for this segment

What is this table?

This is a table of oscar nominees over the years across multiple categories - columns are given below and

include

- Year of event
- Award category
- Name of the nominee
- Whether the nominee won or not
- ID column to make sure every entree has a unique identifier



year	category	nominee	movie	winner	id
2006	actress in a supporting role	Abigail Breslin	Little Miss Sunshine	false	1
1984	actor in a supporting role	Adolph Caesar	A Soldier's Story	false	2
2006	actress in a supporting role	Adriana Barraza	Babel	false	3
2002	actor in a leading role	Adrien Brody	The Pianist	true	4
1942	actress in a supporting role	Agnes Moorehe	The Magnificent Ambersons	false	5
1944	actress in a supporting role	Agnes Moorehe	Mrs. Parkington	false	6
1948	actress in a supporting role	Agnes Moorehe	Johnny Belinda	false	7
1964	actress in a supporting role	Agnes Moorehe	HushHush, Sweet Charlotte	false	8
1936	actor in a supporting role	Akim Tamiroff	The General Died at Dawn	false	9
1943	actor in a supporting role	Akim Tamiroff	For Whom the Bell Tolls	false	10
1972	actor in a supporting role	Al Pacino	The Godfather	false	11
1973	actor	Al Pacino	Serpico	false	12
1974	actor	Al Pacino	The Godfather Part II	false	13
1975	actor	Al Pacino	Dog Day Afternoon	false	14
1979	actor in a leading role	Al Pacino	And Justice for All	false	15
1990	actor in a supporting role	Al Pacino	Dick Tracy	false	16
1992	actor in a leading role	Al Pacino	Scent of a Woman	true	17
1992	actor in a supporting role	Al Pacino	Glengarry Glen Ross	false	18
2004	actor in a supporting role	Alan Alda	The Aviator	false	19
1966	actor	Alan Arkin	The Russians Are Coming The Russians Are Coming	false	20

Question-1:

Write a query to display all the records in the table tutorial.oscar_nominees



Answer-1:

SELECT * **FROM** tutorial.oscar_nominees



Question-2:

Write a query to find the distinct values in the 'year' column



Answer-2:

SELECT DISTINCT year FROM tutorial.oscar_nominees



Question-3:

Write a query to filter the records from year 1999 to year 2006



Answer-3:

SELECT * FROM tutorial.oscar_nominees WHERE year BETWEEN 1999 AND 2006



Question-4:

Write a query to filter the records for either year 1991 or 1998.



Answer-4:

SELECT * FROM tutorial.oscar_nominees WHERE year= 1991 or year =1998



Question-5:

Write a query to return the winner movie name for the year of 1997.



Answer-5:

SELECT movie FROM tutorial.oscar_nominees WHERE year= 1997 AND winner = True



Question-6:

Write a query to return the winner in the 'actor in a leading role' and 'actress in a leading role' category for the year of 1994,1980, and 2008.



Answer-6:

SELECT
nominee,
year,
category,
winner
FROM
tutorial.oscar_nominees
WHERE
category IN ('actor in a leading role','actress in a leading role')
AND year IN (1980,1994,2008)
AND winner = True



Question-7:

Write a query to return the name of the movie starting from letter 'a'?



Answer-7:

SELECT

movie

FROM

tutorial.oscar_nominees

WHERE

lower(movie) LIKE ('a%')



Question-8:

Write a query to return the name of movies containing the word 'the'.



Answer-8:

SELECT

movie

FROM

tutorial.oscar_nominees

WHERE

lower(movie) LIKE ('%the%')



Question-9:

Write a query to return all the records where the nominee name starts with "c" and ends with "r".



Answer-9:

SELECT

*

FROM

tutorial.oscar_nominees

WHERE

lower(nominee) LIKE ('c%') AND lower(nominee) LIKE ('%r')



Question-10:

Write a query to return all the records where the movie was released in 2005 and movie name does not start with 'a' and 'c' and nominee was a winner



Answer-10:

SELECT

*

FROM

tutorial.oscar_nominees

WHERE

(lower(movie) NOT LIKE ('a%') AND lower(movie) NOT LIKE ('c%'))

AND year = 2005

AND winner = True





We will use tutorial.kag_conversion_data dataset for this segment

What is this table?

This is the report of 3 advertising campaigns that were run on facebook- columns are given below and include

- Unique identifier for a user who was shown the advertisement (ad). This remains constant across different ad campaigns (universal for facebook)
- Unique identifier for a campaign this dataset has 3 campaigns
- Unique identifier used by facebook for mapping an ad ID within a campaign (valid only for this campaign)
- Age range
- Gender
- Amount of interest of a user will have to a certain campaign
- Impressions number of times a person has seen a campaign
- Number of times a person clicked on a campaign (side note clicks/impressions = click through rate)
- Spent amount spent by the advertiser to show an ad to a user
- Total conversion number of times a user attempted to purchase (in this case, attempted to purchase the advertised project)
- Approved conversion number of times the user actually purchased the product



bi_be	xyz_compoign_id	ffl_campaign_id.	age	gender	interest	impressions.	elioks	spent	total comenion	approved_conversion
708746	916	103916	30-34	M	19	7960	- 1	1.4300	2	1
798749	916	103917	30-34	M	16	17861	2	1.8200	2	0
708771	916	103920	30-54	M	30	663	0.	0	. 1	0
798875	916	103929	30-34	M	28	A218	1	1.25	1	0
706818	916	103928	30-34	36	28	4133	.1	1.3900	1	1
798820	916	103929	30-34	M	19	1915	0.	0	1	1
708889	910	103840	30-34	M	15	15615	3	4.7700	. 1	0
708865	916	103941	30-34	M	16	10901	1	1.2700	1	1
708953	916	103951	30-34	M	27	2365	1	1.5	1	0
708958	916	103952	30-31	M	20	9502	3	3.1600	1	0
708979	916	103995	30-34	M	31	1224	-0	0	1	0
709023	916	103962	30-54	M	7	735	0	0	- 1	0
700008	910	103965	30-34	3M	10	5117	0	0	- 1	0
709043	916	103965	30-34	M	16	5120		0	1	0
799058	916	103968	30-34	50	30	14669	7	10.2800	1	1
799106	916	103976	30-34	M	28	1241	0	0	1	Y
700115	916	103978	30-34	M	30	2308	1	0.5700	1	0
709124	916	103979	30:34	M	21	1004		0	1	
709179	916	103988	35-39	M	10	4627		1.6900	1	0
709183	916	103989	35-39	M	16	21026	4	4.6300	2	7.

Question-1:

Write a query to count the total number of records in the tutorial.kag_conversion_data dataset.



Solution: 1

SELECT
COUNT(*) AS total_records
FROM
tutorial.kag_conversion_data



Question-2:

Write a query to count the distinct number of fb_campaign_id.



Solution - 2

SELECT

COUNT(DISTINCT fb_campaign_id) AS total_records

FROM

Tutorial.kag_conversion_data



Question-3:

Write a query to find the maximum spent, average interest, minimum impressions for ad_id.



Solution-3

SELECT

MAX(spent) AS max_spent,

AVG(interest) AS avg_interest,

MIN(impressions) AS min_impression

FROM

Tutorial.kag_conversion_data

GROUP by ad_id



Question-4:

Write a query to create an additional column spent per impressions(spent/impressions)



Solution- 4

```
*,
spent/impressions AS spent_per_impression
FROM
Tutorial.kag_conversion_data
```



Question-5:

Write a query to count the ad_campaign for each age group.



Solution-5

SELECT
age,
COUNT(ad_id) AS num_campaign
FROM
tutorial.kag_conversion_data
GROUP BY
Age



Question-6:

Write a query to calculate the average spent on ads for each gender category.



Solution-6

SELECT

gender,

AVG(spent) AS avg_spent

FROM

tutorial.kag_conversion_data

GROUP BY

Gender



Question-7

Write a query to find the total approved conversion per xyz campaign id. Arrange the total conversion in descending order.



Solution-7:

SELECT

xyz_campaign_id,

SUM(approved_conversion) AS total_approved_conversion

FROM

tutorial.kag_conversion_data

GROUP BY

xyz_campaign_id

ORDER BY

total_approved_conversion DESC



Question-8:

Write a query to show the fb_campaign_id and total interest per fb_campaign_id. Only show the campaign which has more than 300 interests.



Solution - 8:

SELECT
fb_campaign_id,
SUM(interest) AS total_interest
FROM
tutorial.kag_conversion_data
GROUP BY
fb_campaign_id
HAVING SUM(interest) > 300
ORDER BY
total_interest DESC



Question-9:

Write a query to find the age and gender segment with maximum impression to interest ratio. Return three columns - age, gender, impression_to_interest.



Solution - 9:

```
SELECT
age,
gender,
SUM(impressions)/SUM(interest) AS impresssion_to_interest
FROM
tutorial.kag_conversion_data
GROUP BY
age,
gender
ORDER BY
impresssion_to_interest DESC
LIMIT 1
```



Question-10:

Write a query to find the top 2 xyz_campaign_id and gender segment with the maximum total_unapproved_conversion (total_conversion - approved_conversion).



Solution -10

```
SELECT

xyz_campaign_id,
gender,

SUM(total_conversion - approved_conversion) AS total_unapproved_conversion

FROM

tutorial.kag_conversion_data

GROUP BY

xyz_campaign_id,
gender

ORDER BY

total_unapproved_conversion DESC

LIMIT 2
```





Caselet-3:

We will use **tutorial.crunchbase_companies** dataset for this segment (you can use select * FROM tutorial.crunchbase_companies to show all the columns)

What is this table?

This contains data on companies across multiple countries, their history, current status, and the amount of funding they have received. Columns are given below and include

- Permalink this is a directory containing the name of a company as part of a directory of all companies (Read up more on permalinks online)
- Name of the company
- Website link
- Category that the company belongs to
- Total funding received
- Current operational status is it active, closed, or acquired by another company
- Country code country that the company was established in



- State code state that the company was established in only available for US
- Region that the company was established in
- City that the company was established in
- Number of rounds of funding that happened for the country
- Date of establishment of company
- Month of establishment of company
- Quarter of establishment of company
- Year of establishment of company
- First date of funding
- Last date of funding
- Last date when a milestone activity was recorded for the company
- Unique identifier to differentiate each entry



persist	mere.	to the page of	comment code.	Totaling STIE, OUT	mene .	- mercusia	PRINCIPAL COMPA	Pagino	- 670	South a receiv	Translation of	Brother, rooms	Security Security	Barried, year	tre, to disper-	Bet Serving at	(MC) THE REAL PROPERTY.	
by home with	DALL	http://www.debt.co			questing			printers.							Táriona	19/5/76		
School State Company State Com	1105(6)			200000	sample .	ide.	68	Print.	And it windows	1	17574	2000	2010/21	2016	11/10/02	BANKEY.		- 2
dismostraji kiris Aumaniy	SSS (brossing	National State of Concession	coPitated	Copiesed	COMPANY OF	174	11	Fremouskinking	Costineed Frank	+	5070/15	(80.00 - 10)	301106	2011	620014	9/90/64		
decreased by constitution	THE BRANCH	TREATMENT REPORTS	and the same	400,000	(freel)	40		Statement West	Strategy Worth						ACHORY	5/19/01	1000	- 4
(towards)): September 1	340 to grown	Note Daniel Delgal men bear	genteurs, y folipse	5790000	-			-							102011	DATE:		
principal public deserva-	80v	March September	ments.	NOWANA	trees(ing	. 959	NO.	Property.	Street,		4,500	(more last	2011.02	201	\$1500	P(1)16	14944	- 6
Surtamy's communications.	MV Dominion	Wash sense bear	parters yields	10000	meeting.	256	DN .	Lautegerm	List Ropette						67670	4000		
Section of Property	Street Co.	Marketermen		1.565404	displant.	046	87	Mary Tools	Stor Farb	-	A7574	access.	(F11.68	800	W/8078	WHO M.	Material	
conceptivion	36000	Morrecountry	140	20000	OWNERS.			ammore's			4/3/8/	(200-01)	190401	197	298	S7608		
Account to the	mary.	Mississipping	#00/F(0)	1.46400	CONTROL .	0.0	UK:	36.696	\$15,0960 (day)	,	4907	period t	Vernor	300	1/0/10	1000	1000	10
Josephal Alli debbi-	sala boytoki.	Neg. (1 Millione con con-	tools.	1797004	consing	100	80	tyrite	tyl file	1	171794	1004-01	1599-01	194	M9.54	1/2/11		- 11
company to 25 made	10:15 Months	the Dean St Month one.	and the same	1700000	countries.	100	667.	Designation St.	Dischline		1/1001	2001.01	2001-015	3561	M896	0.76%		- 10
American Contract of the	Milijakowanah	http://www.Edilphibitesprode	900		menting.	0014		General Dis-		-					MATERIAL ST	474.11		8.0
Deland Williams to	SECTION AND ADDRESS OF	1000	400	2000	minter:	99.	18	18 Tag	Day 3 persons	1	17076	DESCRIPT.	2011.00	.201	171799	018(9)	1011	- 11
Services Statement Services	2007 April 18 miles	Michely Williams into	40	(1964)771	terming.	6108	1905	Delite	0501.0		1/1/09	200001	VINE UT	246	ANCTO	10/14/01/01	67978	10
Surgery Willymone	TEST Meson	Mod November	940	1390	condition.	1980		Test	Peril	1	11/00/95	2000/11	STREET, ST.	300	100011	19/10/16	NO.	16
domests 100de to	empla to	Mindress Williams	Total field by	(270)164	meeting.			printed 6							11100	97 5/66		41
Acres (November 1999)	White	Missione Milator	wholes	Inmoor.	toward a	100		Bridge.	belled.	-					1/viss	101.00		16
Domest of Mileson	SECTIVE .	Stockers Olyhouse	protect	10000	projekter)	10.6	68	0.00	Sec - Carechard		6/78/71	2011 64	derri tide	gar-t	116/77	TVM	80719	110
Commo/Others	410440	TITLE AND STREET	spillment .		menting.	100	100	New York	New York		1,4000	2000.61	1965-24	290	2007	Meto	16990	200

Question-1

Find the top 5 countries(country code) with the highest number of operating companies. Ensure the country code is not null.



Answer-1

```
SELECT
country_code,
COUNT(name) as num_companies
FROM
tutorial.crunchbase_companies
WHERE
country_code IS NOT NULL AND status = 'operating'
GROUP BY
country_code
ORDER BY
num_companies DESC
LIMIT 5
```



Question-2:

How many companies have no country code available in the dataset



Answer-2:

SELECT

COUNT(name) as num_companies

FROM

tutorial.crunchbase_companies

WHERE

country_code IS NULL



Question-3:

Find the number of companies starting with letter 'g' founded in France(FRA) and still operational(status = operating)

Answer-3:

SELECT

COUNT(name) AS num_companies

FROM

tutorial.crunchbase_companies

WHERE

lower(name) LIKE 'g%'

AND country_code = 'FRA'

AND status ='operating'



Question-4:

How many advertising, founded after 2003, are acquired?



Answer-4:

SELECT

COUNT(name) AS num_companies

FROM

tutorial.crunchbase_companies

WHERE

founded_year > 2003

AND status = 'acquired'

AND category_code = 'advertising'



Question-5:

Calculated the average funding_total_usd per company for the companies founded in the software, education, and analytics category.



Solution-5:

```
SELECT
category_code,
SUM(funding_total_usd)/COUNT(category_code) AS average_funding_total_usd
FROM
tutorial.crunchbase_companies
WHERE
category_code IN ('software', 'education','analytics') AND funding_total_usd IS NOT NULL
GROUP BY
Category_code
```



Question-6:

Find the city having more than 50 closed companies. Return the city and number of companies closed.



Answer-6:

SELECT
city,
COUNT(name) AS num_closed_compnay
FROM
tutorial.crunchbase_companies
WHERE
status = 'closed'
AND city IS NOT NULL
GROUP BY
city
HAVING (COUNT(name)) > 50



Question-7:

Find the number of bio-tech companies who are founded after 2000 and either have more than 1Mn funding or have ipo and secured more than 1 round of funding.



Answer-7:

SELECT
COUNT(name) AS num_compnay
FROM
tutorial.crunchbase_companies
WHERE
founded_year > 2000
AND (funding_total_usd > 1000000 OR (status = 'ipo' AND funding_rounds >1))
AND category_code = 'biotech'



Question-8:

Find all number of all acquired companies founded between 1980 and 2005 and founded in the city ending with the word 'city'. Return the city name and number of acquired companies.



Answer-8:

```
city,
COUNT(name) AS num_acquired_compnay
FROM
tutorial.crunchbase_companies
WHERE
founded_year BETWEEN 1980 AND 2005
AND status = 'acquired'
AND lower(city) LIKE ('%city')
GROUP BY
City
```



Question-9:

Find the number of 'hardware' companies founded outside 'USA' and did not take any funding. Return the country code and number of hardware companies in descending order.



Answer-9:

SELECT
country_code,
COUNT(name) AS num_company
FROM
tutorial.crunchbase_companies
WHERE
category_code = 'hardware'
AND funding_total_usd IS NULL
AND country_code ◇'USA'
GROUP BY
country_code
ORDER BY
Num_company DESC



Question-10:

Find the 5 most popular company category(category with highest companies) across the city Singapore, Shanghai, and Bangalore. Return category code and number of companies



Answer-10:

```
SELECT
category_code,
COUNT(name) AS num_compnay
FROM
tutorial.crunchbase_companies
WHERE
city IN ('Bangalore', 'Shanghai', 'Singapore')
GROUP BY
category_code
ORDER BY
num_compnay DESC
LIMIT 5
```



THANK YOU