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

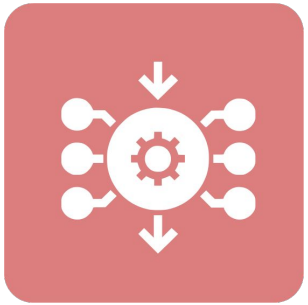
Relevel
by Unacademy

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

Caselet-1



Caselet-1

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

Caselet-1

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...	Hush...Hush, 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

Caselet-1

Question-1:

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

Caselet-1

Answer-1:

SELECT * FROM tutorial.oscar_nominees

Caselet-1

Question-2:

Write a query to find the distinct values in the ‘year’ column

Caselet-1

Answer-2:

SELECT DISTINCT year FROM tutorial.oscar_nominees

Caselet-1

Question-3:

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

Caselet-1

Answer-3:

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

Caselet-1

Question-4:

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

Caselet-1

Answer-4:

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

Caselet-1

Question-5:

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

Caselet-1

Answer-5:

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

Caselet-1

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.

Caselet-1

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
```

Caselet-1

Question-7:

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

Caselet-1

Answer-7:

```
SELECT
  movie
FROM
  tutorial.oscar_nominees
WHERE
  lower(movie) LIKE ('a%')
```

Caselet-1

Question-8:

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

Caselet-1

Answer-8:

```
SELECT
  movie
FROM
  tutorial.oscar_nominees
WHERE
  lower(movie) LIKE ('%the%')
```

Caselet-1

Question-9:

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

Caselet-1

Answer-9:

```
SELECT
  *
FROM
  tutorial.oscar_nominees
WHERE
  lower(nominee) LIKE ('c%') AND lower(nominee) LIKE ('%r')
```

Caselet-1

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

Caselet-1

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
```

Caselet-2

Caselet-2

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

Caselet-2

ad_id	xyz_campaign_id	fb_campaign_id	age	gender	interest	impressions	clicks	spent	total_conversion	approved_conversion
708746	916	103916	30-34	M	15	7350	1	1.4300	2	1
708749	916	103917	30-34	M	16	17861	2	1.8200	2	0
708771	916	103920	30-34	M	20	693	0	0	1	0
708815	916	103928	30-34	M	28	4259	1	1.25	1	0
708818	916	103928	30-34	M	28	4133	1	1.2900	1	1
708820	916	103929	30-34	M	29	1915	0	0	1	1
708889	916	103940	30-34	M	15	15615	3	4.7700	1	0
708895	916	103941	30-34	M	16	10951	1	1.2700	1	1
708953	916	103951	30-34	M	27	2355	1	1.5	1	0
708958	916	103952	30-34	M	28	9502	3	3.1600	1	0
708979	916	103955	30-34	M	31	1224	0	0	1	0
709023	916	103962	30-34	M	7	735	0	0	1	0
709038	916	103965	30-34	M	16	5117	0	0	1	0
709040	916	103965	30-34	M	16	5120	0	0	1	0
709058	916	103968	30-34	M	20	14669	7	10.2800	1	1
709105	916	103976	30-34	M	28	1241	0	0	1	1
709115	916	103978	30-34	M	30	2305	1	0.5700	1	0
709124	916	103979	30-34	M	31	1024	0	0	1	1
709179	916	103988	35-39	M	15	4627	1	1.6900	1	0
709183	916	103989	35-39	M	16	21026	4	4.6300	2	1

Caselet-2

Question-1:

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

Caselet-2

Solution: 1

```
SELECT
  COUNT(*) AS total_records
FROM
  tutorial.kag_conversion_data
```

Caselet-2

Question-2:

Write a query to count the distinct number of fb_campaign_id.

Caselet-2

Solution - 2

```
SELECT
  COUNT(DISTINCT fb_campaign_id) AS total_records
FROM
  Tutorial.kag_conversion_data
```

Caselet-2

Question-3:

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

Caselet-2

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
```

Caselet-2

Question-4:

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

Caselet-2

Solution- 4

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

Caselet-2

Question-5:

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

Caselet-2

Solution-5

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

Caselet-2

Question-6:

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

Caselet-2

Solution-6

```
SELECT
gender,
AVG(spent) AS avg_spent
FROM
  tutorial.kag_conversion_data
GROUP BY
  Gender
```

Caselet-2

Question-7

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

Caselet-2

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
```

Caselet-2

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.

Caselet-2

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
```

Caselet-2

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.

Caselet-2

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
```

Caselet-2

Question-10:

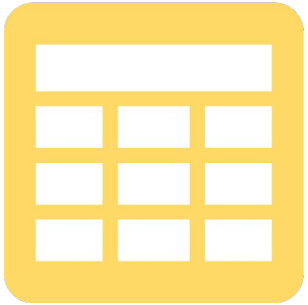
Write a query to find the top 2 xyz_campaign_id and gender segment with the maximum total_unapproved_conversion (total_conversion - approved_conversion).

Caselet-2

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



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?

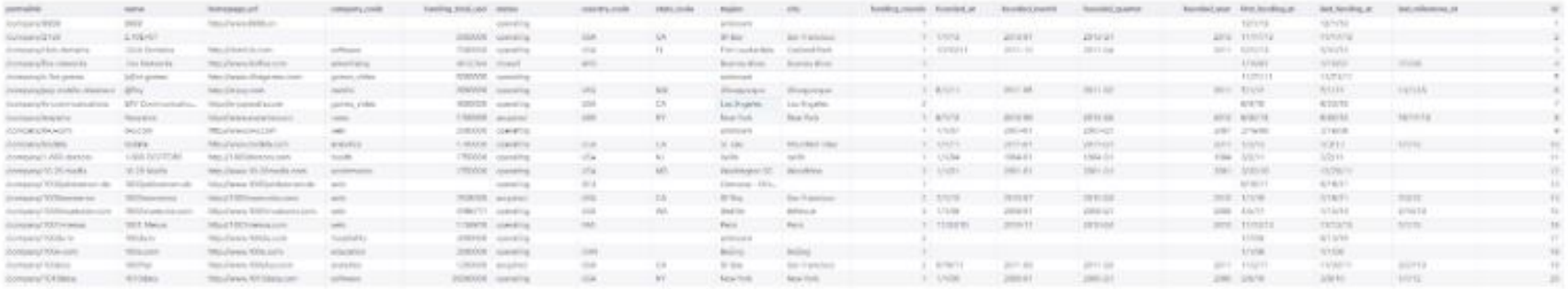
This contains data on companies across multiple countries, their history, current status, and the amount of funding they have received. Columns are given below and include

- 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

Caselet-3

- 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

Caselet-3



Caselet-3

Question-1

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

Caselet-3

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
```

Caselet-3

Question-2:

How many companies have no country code available in the dataset

Caselet-3

Answer-2:

```
SELECT
  COUNT(name) as num_companies
FROM
  tutorial.crunchbase_companies
WHERE
  country_code IS NULL
```

Caselet-3

Question-3:

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

Caselet-3

Answer-3:

```
SELECT
  COUNT(name) AS num_companies
FROM
  tutorial.crunchbase_companies
WHERE
  lower(name) LIKE 'g%'
  AND country_code = 'FRA'
  AND status ='operating'
```

Caselet-3

Question-4:

How many advertising, founded after 2003, are acquired?

Caselet-3

Answer-4:

```
SELECT
  COUNT(name) AS num_companies
FROM
  tutorial.crunchbase_companies
WHERE
  founded_year > 2003
  AND status = 'acquired'
  AND category_code = 'advertising'
```

Caselet-3

Question-5:

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

Caselet-3

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
```

Caselet-3

Question-6:

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

Caselet-3

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
```

Caselet-3

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.

Caselet-3

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'
```

Caselet-3

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.

Caselet-3

Answer-8:

```
SELECT
  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
```

Caselet-3

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.

Caselet-3

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
```

Caselet-3

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

Caselet-3

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