

1. Get the businesses in San Jose, CA that are still "in business". Results should be sorted by review counts in descending order. Return for the top 7 businesses, the business id, name, review count, and average rating of the business.

**Solution:**

```
SELECT T.BID, BU.B_NAME, T.CNT, T.AVRG
FROM
(SELECT B.BUSN_ID BID, COUNT(B.BUSN_ID) CNT, AVG(R.RATING) AVRG
FROM business B, Reviews R
WHERE B.CITY = 'San Jose'
AND B.STATE_NM = 'CA'
AND B.OPEN='T'
AND R.BUSN_ID = B.BUSN_ID
GROUP BY B.BUSN_ID
ORDER BY CNT DESC
) T, BUSINESS BU
WHERE T.BID = BU.BUSN_ID
AND ROWNUM <= 7;
```

**Output:**

BID	B_NAME	CNT	AVRG
B1	Big Surf	6	2
B4	Safelite AutoGlass	6	2.3E+00
B23	The Table	6	4
B24	Top Nosh Café	6	4
B3	Pedley Richard Architect	6	2.2E+00
B5	SMITH THOMSON	6	4
B2	AirBorne COFFEE	6	3

7 rows selected

B6	Catch Your Big Break	6	3.5
----	----------------------	---	-----

2. Get the review counts for businesses in each business category. Results should be sorted by the Review counts in descending order.

**Solution:**

```
select distinct(b.bus_catg_id),b.busn_id ,count(r.rev_id)
from business b, reviews r
where r.busn_id= b.busn_id
group by b.busn_id, b.bus_catg_id
order by b.bus_catg_id;
```

**Output:**

BUS_CATG_I	BUSN_ID	COUNT(R.REV_ID)
-----		
BCT1	B1	6
BCT10	B10	2
BCT10	B12	1
BCT10	B13	1
BCT10	B14	1
BCT10	B15	1
BCT10	B16	1
BCT10	B17	1
BCT10	B18	1
BCT10	B19	1
BCT10	B20	1
BCT10	B21	1
BCT11	B11	1
BCT2	B2	6
BCT3	B3	6
BCT4	B4	6
BCT5	B22	5
BCT5	B23	6
BCT5	B24	6
BCT5	B25	6
BCT5	B26	6
BCT5	B27	5
BCT5	B28	5
BCT5	B5	6
BCT6	B6	6
BCT7	B7	2
BCT8	B8	2
BCT9	B9	2

28 rows selected

3. *Find the average rating across all reviews written by a particular user.*

**Solution:**

```
select Y.YELP_ID, nvl(T.avrg,0)
from
(select R.AUTHOR usr, nvl(round(avg(R.RATING)),0) avrg
from business b , reviews r
where B.BUSN_ID = R.BUSN_ID
```

```

group by R.AUTHOR) T
right outer join YELP_USER Y
on Y.YELP_ID = T.usr
order by Y.YELP_ID;

```

**OUTPUT:**

YELP_ID	NVL(T.AVRG,0)
Y1	3
Y10	3
Y11	0
Y12	0
Y13	5
Y2	4
Y3	4
Y4	4
Y5	4
Y6	4
Y7	3
Y8	3
Y9	3

13 rows selected

4. *Get the businesses in San Jose, CA that have been reviewed by more than 5 `elite' users. Users who have written more than 10 reviews are called `elite' users. Results should be ordered by the `elite' user count in descending order. Return for the top 5 businesses, the business id, business name, business review count, average rating, and the count of the `elite' users for the particular business.*

**Solution:**

```

select b.busn_id, count(r.rev_id) rev_cnt, round(avg(r.rating),0) avg_rtg, count(r.author)
usr_cnt
from business b, reviews r
where b.state_nm ='CA'

```

```

and b.city='San Jose'

and b.busn_id = r.busn_id

and r.author

in (select DISTINCT(r.author)

from YELP_USER y, reviews r

where r.author= y.yelp_id

having count(r.rating) >=10

group by r.author)

group by b.busn_id

having count(r.author) > 5

order by count(r.author) desc;

```

**Output:**

BUSN_ID	REV_CNT	AVG_RTG	USR_CNT
B1	6	2	6
B4	6	2	6
B23	6	4	6
B6	6	4	6
B24	6	4	6
B3	6	2	6
B25	6	4	6
B5	6	4	6
B2	6	3	6
B26	6	4	6

10 rows selected

5. *Get the businesses in San Jose, CA that have the highest percentage of reviews with rating of 5, and have been reviewed at least 5 times. Results should be ordered by the percentage in descending order. Return for the top 5 businesses, the business id, business name, review count, and percentage of reviews with rating of 5.*

**Solution:**

```

select t.busn_id, t.b_name, ((t.no_fives/count(r.rev_id))*100) percent_of_reviews,
count(r.rev_id) no_of_reviews, t.no_fives
from
(select r1.busn_id , b1.b_name, count(r1.rating) no_fives
from reviews r1, business b1
where r1.rating=5
and r1.busn_id = b1.busn_id

```

```

and r1.busn_id
in (
select r.busn_id
from reviews r , business b
where r.busn_id= b.busn_id
and b.city='San Jose'
and b.state_nm='CA'
group by r.busn_id
having count(r.rev_id)>=5
) group by r1.busn_id, b1.b_name
order by r1.busn_id) t, reviews r
where t.busn_id = r.busn_id
group by t.busn_id, t.b_name,t.no_fives ;

```

**Output:**

BUSN_ID	B_NAME	PERCENT_OF_REVIEWS	NO_OF_REVIEWS	NO_FIVES
-----				
B22	Hobee's	100	5	5
B27	Southern Kitchen	80	5	4
B28	Holder's Country Inn	100	5	5

6. Find top two businesses, which have highest average ratings in each business category. For each business, list business id, business name and business category name.

**Solution:**

```

SELECT b.BUSN_ID,b.B_NAME,b.BUS_CATG_ID
FROM BUSINESS b
WHERE b.BUSN_ID IN (
    SELECT temp.BUSN_ID
    FROM(SELECT b1.BUSN_ID,b1.BUS_CATG_ID
    FROM BUSINESS b1, REVIEWS r1
    WHERE b1.BUSN_ID = r1.BUSN_ID
    GROUP BY b1.BUSN_ID,b1.BUS_CATG_ID
    ORDER BY AVG(r1.Rating) DESC
    ) temp
    WHERE temp.BUS_CATG_ID = b.BUS_CATG_ID
    AND ROWNUM <= 2
);

```

**Output:**

BUSN_ID	B_NAME	BUS_CATG_I
-----		
B1	Big Surf	BCT1
B2	AirBorne COFFEE	BCT2

B3	Pedley Richard Architect	BCT3
B4	Safelite AutoGlass	BCT4
B6	Catch Your Big Break	BCT6
B7	Bay Area Coffee Shop	BCT7
B8	China Coffee Toffee	BCT8
B9	Hastings Water Works	BCT9
B11	In-N-Out Burger	BCT11
B12	Coffee Bar and Bistro	BCT10
B13	Renegades	BCT10
B22	Hobee's	BCT5
B28	Holder's Country Inn	BCT5

13 rows selected

7. *Find the oldest user who never reviewed any businesses.*

**Solution:**

```
select T.yid from
(select u.yelp_id yid, (months_between(current_date, u.dob )/12) age
from yelp_user u
where u.yelp_id
not in (
select r.author
from business b, reviews r
where R.BUSN_ID = B.BUSN_ID)
order by age desc) T
where rownum = 1;
```

**OUTPUT:**

YID

-----

Y11

8. *Find every business in CA that has the word "Coffee" in its name.*

**Solution:**

```
select B.BUSN_ID, B.B_NAME
from business b
```

```
where B.STATE_NM = 'CA'
and B.B_NAME like '%Coffee%';
```

**Output:**

BUSN_ID	B_NAME
-----	
B7	Bay Area Coffee Shop
B8	China Coffee Toffee
B12	Coffee Bar and Bistro

9. *List all 5-star businesses that have been reviewed by any users between the ages of 15 and 25. A 5-star business is a business which has an average review rating of 5.*

**Solution:**

```
select T.busn, r.author, round(months_between(current_date, u.dob )/12) age
from
(select bb.BUSN_ID busn from business bb, reviews rr
where bb.BUSN_ID = rr.BUSN_ID
group by bb.BUSN_ID
having avg(rr.rating) =5) T, reviews r, yelp_user u
where T.BUSN = r.BUSN_ID
and r.author = U.YELP_ID
and round(months_between(current_date, u.dob )/12)>15
and round(months_between(current_date, u.dob )/12)<25;
```

**Output:**

BUSN	AUTHOR	AGE
-----		
B15	Y13	20
B18	Y13	20
B16	Y13	20
B17	Y13	20
B13	Y13	20

B21	Y13	20
B14	Y13	20
B19	Y13	20
B22	Y1	23
B22	Y13	20
B20	Y13	20
B12	Y13	20

12 rows selected

10. List users that have been to at least 3 distinct states. Order by number of states traveled to (decreasing), break ties with user id (increasing). For each user, list his/her user id, name and list of states traveled to.

**Solution:**

```
select distinct(bb.state_nm),yelp_id, F_NAME from (SELECT yelp_id, F_NAME, count(*) AS
number_states
```

```
FROM
```

```
( SELECT u.yelp_id, u.F_NAME, b.state_nm, count(*)
```

```
FROM yelp_user u, reviews r, business b
```

```
WHERE u.yelp_id = r.author
```

```
AND b.busn_id = r.busn_id
```

```
GROUP BY b.state_nm, u.yelp_id, u.F_NAME
```

```
) T
```

```
GROUP BY yelp_id, F_NAME
```

```
HAVING count(*) >= 3
```

```
ORDER BY count(*) DESC, yelp_id ASC) T, REVIEWS rr, Business bb
```

```
where yelp_id = rr.AUTHOR
```

```
and bb.busn_id = rr.busn_id ;
```

**OUTPUT:**

ST	YELP_ID	F_NAME
CA	Y13	Kelley
VA	Y13	Kelley
OH	Y13	Kelley



AZ      Y13      Kelley