

Query:

```
use india_census
```

```
select * from dataset1
```

```
select * from dataset2
```

```
-- Total no.of data in both the dataset
```

```
select count(*) from dataset1
```

```
select count(*) from dataset2
```

```
-- Data of state jharkhand and bihar
```

```
select * from dataset1 where state in ("Jharkhand", "Bihar")
```

```
-- Total Population in india
```

```
select SUM(Population) as Total_Population from dataset2
```

```
-- Average Growth rate by states
```

```
select state,avg(growth)*100 as Avg_Growth from dataset1 group by state
```

```
-- Average sex ratio by different state
```

```
select state, round(avg(sex_ratio)) as avg_sex_ratio from dataset1 group by state order by  
avg_sex_ratio desc
```

```
-- All the state whose average literacy is greater than 80
```

```
select state, round(avg(literacy)) as avg_literacy from dataset1
```

```
group by state
```

```
having avg_literacy>80
```

```
order by avg_literacy desc
```

```
-- top 3 state having highest growth ratio
```

```
select state,avg(growth)*100 as Avg_Growth from dataset1 group by state
```

```
order by avg_growth desc limit 3
```

```
-- bottom 3 state having lowest growth ratio
```

```
select state,avg(growth)*100 as Avg_Growth from dataset1 group by state
```

```
order by avg_growth asc limit 3
```

```
-- top and bottom 3 states in literacy rate
```

```
select "Top 3 State in literacy" as category,
```

```
group_concat(state) as states from (select state, round(avg(literacy)) as avg_literacy from  
dataset1
```

```
group by state order by avg_literacy desc limit 3
```

```
) as top
```

```
union
```

```
select "Bottom 3 state in literacy" as category,
```

```
group_concat(state) as states from (select state, round(avg(literacy)) as avg_literacy from  
dataset1
```

```
group by state order by avg_literacy asc limit 3
```

```
) as bottom
```

```
-- state whose literacy greater than 80 and start with letter A
```

```

select * from (select state, round(avg(literacy)) as avg_literacy from dataset1
group by state
having avg_literacy>80
order by avg_literacy desc
) as bottom where state like "a%"

```

```

-- total males and females by state
select d.state,sum(d.males) total_males,sum(d.females) total_females from
(select c.district,c.state ,round(c.population/(c.sex_ratio+1),0) males,
round((c.population*c.sex_ratio)/(c.sex_ratio+1),0) females from
(select d1.district,d1.state,d1.sex_ratio/1000 sex_ratio,d2.population from dataset1 d1 inner
join dataset2 d2 on d1.district=d2.district ) c) d
group by d.state;

```

```

-- total literacy rate by state
select c.state,sum(literate_people) total_literate,sum(illiterate_people) total_lliterate from
(select d.district,d.state,round(d.literacy_ratio*d.population,0) literate_people,
round((1-d.literacy_ratio)* d.population,0) illiterate_people from
(select d1.district,d1.state,d1.literacy/100 literacy_ratio,d2.population from dataset1 d1
inner join dataset2 d2 on d1.district=d2.district) d) c
group by c.state

```

```

-- population in previous census
select sum(m.previous_census_population)
previous_census_population,sum(m.current_census_population)
current_census_population from(
select e.state,sum(e.previous_census_population)
previous_census_population,sum(e.current_census_population)
current_census_population from
(select d.district,d.state,round(d.population/(1+d.growth),0)
previous_census_population,d.population current_census_population from
(select a.district,a.state,a.growth growth,b.population from dataset1 a inner join dataset2 b
on a.district=b.district) d) e
group by e.state)m

```

```

-- population vs area
with cte as
(select q.*,r.total_area from (

```

```

select '1' as id,n.* from
(select sum(m.previous_census_population)
previous_census_population,sum(m.current_census_population)
current_census_population from(
select e.state,sum(e.previous_census_population)
previous_census_population,sum(e.current_census_population)
current_census_population from
(select d.district,d.state,round(d.population/(1+d.growth),0)
previous_census_population,d.population current_census_population from
(select a.district,a.state,a.growth growth,b.population from dataset1 a inner join dataset2 b
on a.district=b.district) d) e

```

```
group by e.state)m) n) q inner join (
```

```
select '1' as id,z.* from (  
select sum(area_km2) total_area from dataset2)z) r on q.id=r.id)
```

```
select (cte.total_area/cte.previous_census_population) as  
previous_census_population_vs_area, (cte.total_area/cte.current_census_population) as  
current_census_population_vs_area from cte
```

```
-- output top 3 districts from each state with highest literacy rate
```

```
select d1.* from  
(select district,state,literacy,rank() over(partition by state order by literacy desc) rnk from  
dataset1) d1  
where d1.rnk in (1,2,3) order by state
```