SQL CODES:

Creation of table for dataset1

```
create table dataset1(
        district varchar,
        state varchar,
        sex_ratio int,
        literacy float,
        growth float


        );
copy dataset1 FROM 'D:\census\dataset1.csv' DELIMITER ',' CSV
HEADER;
```

Dataset 1:

```
select * from dataset1;
```

Dataset 2:

```
select * from dataset2 ;
```

Q1) Number of rows in our dataset ?

```
select count(*) from dataset1;

select count(*) from dataset2 ;
```

Q2) A) Data only for Jharkhand and Bihar in dataset1 ?

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

Q2) B) Data only for Jharkhand and Bihar in dataset2 ?

```
select * from from dataset2 where state in ('Jharkhand' ,'Bihar');
```

Q3) Query total population of India ?

```
select sum(population) from dataset2;
```

Q4) Average growth rate ?

```
select avg(growth) as result from dataset1;
```

Q5) Average growth rate per state?

```
select state,avg(growth) as result
from dataset1
group by state
order by result desc ;
```

Q6) Average sex ratio by state (whole number) ?

```
select state,round(avg(sex_ratio),0) as result
from dataset1
group by state
order by result desc ;
```

Q7) Average literacy by state (whole number) ?

```
SELECT state, ROUND(CAST(AVG(Literacy) AS NUMERIC), 0) AS result
FROM dataset1
GROUP BY state
ORDER BY result DESC;
```

Q8) Query to sort the states and districts according to literacy rate in descending order?

```
select state,district,max(literacy ) as maximumliteracy
from dataset1
group by state,district
order by maximumliteracy desc ;
```

Q8) Average Literacy > 90 ?

```
SELECT state, ROUND(CAST(AVG(Literacy) AS NUMERIC), 0) AS result
FROM dataset1
GROUP BY state
having ROUND(CAST(AVG(Literacy) AS NUMERIC), 0) >90
ORDER BY result DESC;
```

Q9) States starting with letter A ?

```
Select distinct(state) from dataset1 where state like 'A%';
```

Q10) States ending with letter A ?

```
Select distinct(state) from dataset1 where state like '%a';
```

**Q 11) Return a list of states and districts from dataset1 ,
along with the total area from dataset2 for each combination of state and district**

```
select d.state,d.district,sum(c.area)
from dataset1 as d
join dataset2 as c
on d.state=c.state
group by d.state,d.district
order by sum(c.area) desc ;
```

Q 12) To find relation between literacy and area

```
SELECT d.state, d.district, d.literacy, c.area
FROM dataset1 AS d
JOIN dataset2 AS c
ON d.state = c.state
ORDER BY d.literacy DESC;
```

Q 13) To find relation between sex_ratio and area

```
select d.state,d.district,c.area,d.sex_ratio
from dataset1 as d
join dataset2 as c
on d.state=c.state
order by sex_ratio desc     ;
```

Q 14) To find relation between growth of population and area

```
select d.state,d.district,c.area,d.growth
from dataset1 as d
join dataset2 as c
on d.state=c.state
order by growth desc;
```