Project Milestone 2
Group Name-Team COL362

Members

1)Ankit Raushan 2020CS10324
2)Sibasish Rout 2020CS10386
3)Keshav 2020CS10352

Database size

Relation name	No of rows	Size of relation
condition	1727	116.5Kb
country_crop_production	20567	790.9Kb
sub_groups_to_groups	125	3.3Kb
crops	1063	17.2Kb
generic_food	927	37.2Kb
india_crop_production	276	8.4Kb
nutrition	87	5Kb
state_conditions	35	859bytes
state_crop_production	246092	12.6Mb
Total size	**270899**	**13.6Mb**

Queries

```
--1--crops with maximum area of production in india

with table1(id,crop_name,total_area)as

    (select crop_id,crop_name,sum(area)

    from india_crop_production

    group by crop_id,crop_name),


    num(num_year)as

    (select count(distinct year_id)

    from india_crop_production)


select id,crop_name,total_area/num_year as average_production_area

from table1,num
```

```
order by average_production_area desc
limit 10;
```

--2--crops with maximum amount production in india

```
with table1(id,crop_name,total_production)as
    (select crop_id,crop_name,sum(production)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
    from india_crop_production)

select id,crop_name,total_production/num_year as
average_production
from table1,num
order by average_production desc
limit 10;
```

--3--crops with maximum yield in India

```
with table1(id,crop_name,total_area)as
    (select crop_id,crop_name,sum(area)
    from india_crop_production
    group by crop_id,crop_name),

    table2(id,crop_name,total_production)as
    (select crop_id,crop_name,sum(production)
```

```
    from india_crop_production
    group by crop_id,crop_name)

select t2.crop_name as crop_name,total_production/total_area as
average_yield
from table1 t1, table2 t2
where t1.id=t2.id
order by average_yield desc
limit 10;
```

--4--crop with average production area >=x having minimum yield

```
with table1(id,crop_name,total_area)as
    (select crop_id,crop_name,sum(area)
    from india_crop_production
    group by crop_id,crop_name),

    table2(id,crop_name,total_production)as
    (select crop_id,crop_name,sum(production)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
    from india_crop_production)

select t2.crop_name as crop_name, t2.total_production/num_year as
average_production, t1.total_area/num_year as
```

```
average_production_area,total_production/total_area as
average_yield
from table1 t1, table2 t2, num
where t1.id=t2.id and total_area/num_year >= 150
order by average_yield asc
limit 10;
```

--5--crop with average production area >=x having maximum yield

```
with table1(id,crop_name,total_area)as
    (select crop_id,crop_name,sum(area)
    from india_crop_production
    group by crop_id,crop_name),

    table2(id,crop_name,total_production)as
    (select crop_id,crop_name,sum(production)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
    from india_crop_production)

select t2.crop_name as crop_name, t2.total_production/num_year as
average_production, t1.total_area/num_year as
average_production_area,total_production/total_area as
average_yield
```

```
from table1 t1, table2 t2, num
where t1.id=t2.id and total_area/num_year >= 150
order by average_yield desc
limit 10;
```

--6--crop having maximum deviation in production

```
with table1(id,crop_name,total_production,appearance)as
    (select crop_id,crop_name,sum(production),count(crop_id)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
    from india_crop_production),

    table2(id,crop_name,mean_production,appearance)as
    (select id,crop_name,total_production/appearance,appearance
    from table1),

    table3(id,crop_name,temp,appearance)as
    (select id,i.crop_name,(production-
mean_production)*(production-mean_production),appearance
    from india_crop_production i,table2 t
    where crop_id=id),

    table4(id,crop_name,temp_dev,appearance)as
    (select id,crop_name,sum(temp),appearance
```

```
    from table3
    group by id,crop_name,appearance)

select id,crop_name,|/(temp_dev/appearance) as
deviation_in_production
from table4,num
order by deviation_in_production desc
limit 10;
```

--7--crop having minimum deviation in production

```
with table1(id,crop_name,total_production,appearance)as
    (select crop_id,crop_name,sum(production),count(crop_id)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
    from india_crop_production),

    table2(id,crop_name,mean_production,appearance)as
    (select id,crop_name,total_production/appearance,appearance
    from table1),

    table3(id,crop_name,temp,appearance)as
    (select id,i.crop_name,(production-
mean_production)*(production-mean_production),appearance
    from india_crop_production i,table2 t
```

```
    where crop_id=id),

    table4(id,crop_name,temp_dev,appearance)as
    (select id,crop_name,sum(temp),appearance
    from table3
    group by id,crop_name,appearance)

select id,crop_name,|/(temp_dev/appearance) as
deviation_in_production
from table4,num
order by deviation_in_production desc
limit 10;
```

--8--crop having maximum yield grown on area >=x requiring

rainfall <=y and rainfall >=z (avg temp=25 and avg rainfall=99.3)

```
with table1(id,crop_name,total_area,appearance)as
    (select crop_id,crop_name,sum(area),count(crop_id)
    from india_crop_production
    group by crop_id,crop_name),

    table2(id,crop_name,total_production)as
    (select crop_id,crop_name,sum(production)
    from india_crop_production
    group by crop_id,crop_name),

    num(num_year)as
    (select count(distinct year_id)
```

```
    from india_crop_production),

    table3(id, crop_name, average_yield,
average_production_area)as
    (select
t1.id,t1.crop_name,total_production/total_area,total_area/appearan
ce
    from table1 t1, table2 t2, num
    where t1.id=t2.id and total_area/appearance >= 0),

    cond(id,rain,temp,appearance)as
    (select crop_id,sum(rainfall),sum(temperature),count(crop_id)
    from condition
    group by crop_id)

select t.id as id,crop_name,temp/appearance as
avg_temp,rain/appearance as avg_rainfall,average_yield,
average_production_area
from table3 t,cond c
where t.id=c.id and temp/appearance>=0 and rain/appearance>=0
order by average_yield desc limit 10;
```

--9--country with maximum average wheat production

```
with
table1(country_name,id,crop_name,total_production,appearance)as
    (select
country_name,crop_id,crop_name,sum(production),count(crop_id)
```

```
    from country_crop_production
    where crop_name='wheat'
    group by country_name,crop_id,crop_name)

select country_name,id,crop_name,total_production/appearance as
average_production
from table1
order by average_production desc
limit 11;
```

--10--India's place in world in crop x

```
with
table1(country_name,id,crop_name,total_production,appearance)as
    (select
country_name,crop_id,crop_name,sum(production),count(crop_id)
    from country_crop_production
    where crop_name='maize'
    group by country_name,crop_id,crop_name),
    ind_prd(ind_production)as
    (select total_production/appearance
    from table1
    where country_name='IND')

select crop_name,count(country_name) as indias_position
from table1 ,ind_prd
where total_production/appearance >= ind_production
group by crop_name;
```

--11--country with maximum average crop production in a year

```
with table1(country_name,year_id,total_production)as
    (select country_name,year_id,sum(production)
    from country_crop_production
    group by country_name,year_id),
    table2(country_name,total_production,appearance)as
    (select country_name,sum(total_production),count(year_id)
    from table1
    group by country_name)

select country_name,total_production/appearance as
average_production
from table2
order by average_production desc
limit 10;
```

--12-- Maximum nutrition values queries--

```
with max_nutrition(crop_name,nutrition_type) as
(
    (
        select c.crop_name,'fat'
        from crops c,nutrition n
        where n.fat=
        (
            select max(n1.fat) from nutrition n1
        )
```

```
    and n.crop_id=c.crop_id
)
union all
(
    select c.crop_name,'calories'
    from crops c,nutrition n
    where n.calories=
    (
        select max(n1.calories) from nutrition n1
    )
    and n.crop_id=c.crop_id
)
union all
(
    select c.crop_name,'sodium'
    from crops c,nutrition n
    where n.sodium=
    (
        select max(n1.sodium) from nutrition n1
    )
    and n.crop_id=c.crop_id
)
union all
(
    select c.crop_name,'potassium'
    from crops c,nutrition n
    where n.potassium=
```

```
    (
        select max(n1.potassium) from nutrition n1
    )
    and n.crop_id=c.crop_id
)
union all
(
    select c.crop_name,'carbohydrates'
    from crops c,nutrition n
    where n.carbo_hydrate=
    (
        select max(n1.carbo_hydrate) from nutrition n1
    )
    and n.crop_id=c.crop_id
)
union all
(
    select c.crop_name,'sugars'
    from crops c,nutrition n
    where n.sugars=
    (
        select max(n1.sugars) from nutrition n1
    )
    and n.crop_id=c.crop_id
)
union all
(
```

```
        select c.crop_name,'protien'
        from crops c,nutrition n
        where n.protein=
        (
            select max(n1.protein) from nutrition n1
        )
        and n.crop_id=c.crop_id
    )
)
select crop_name,nutrition_type from max_nutrition;
```

--13--State queries--

```
with npk(cid,s) as
(
    select c.crop_id,(c.N+c.P+c.K) as sum_npk
    from condition c where (c.N is not null) and (c.P is not null)
and (c.K is not null)
    order by sum_npk asc
    limit 1
),
scmax(cid,prod) as
(
    select scp1.crop_id,max(scp1.production) from
state_crop_production scp1,npk n where scp1.crop_id=n.cid  group
by scp1.crop_id
),
state_max(cid,st) as
```

```
(
    select c.crop_name,scp.state_name
    from npk n,state_crop_production scp,scmax sc,crops c
    where n.cid=scp.crop_id and n.cid=c.crop_id
    and scp.production=sc.prod
)
select cid,st from state_max;
```

--14--Best crop according to climatic condition queries--

```
with temp1(cid) as
(
    select distinct scp.crop_id from state_crop_production scp
where scp.state_name  like 'punjab'
),
temp2(cid) as
(
    select distinct c.crop_id from crops c ,temp1 t1 ,condition
cd, state_conditions sc
    where c.crop_id not in (select * from temp1)
    and c.crop_id=cd.crop_id
    and abs(sc.temperature-cd.temperature)<2
    and abs(sc.rainfall-cd.rainfall)<10
    and abs(sc.ph-cd.ph)<0.5
),
temp3(cid) as
(
    select t2.cid,(n.protein+n.carbo_hydrate) as nv
```

```
    from temp2 t2,nutrition n where n.crop_id=t2.cid and
n.protein+n.carbo_hydrate>0
    order by nv desc limit 5
),
temp4(cid) as
(
    select t3.cid,(n.P+n.N+n.K) as tf
    from temp3 t3,condition n where n.crop_id=t3.cid and
n.P+n.N+n.K>0
    order by tf asc limit 1
)
select c.crop_name from crops c,temp4 where c.crop_id=temp4.cid;
```

Performance analysis

Query no	Estimated cost	Time of execution
1	18.15..18.18	3.733s
2	18.15..18.18	2.625s
3	22.09..22.11	1.259s
4	27.99..28.02	1.914s
5	27.99..28.02	3.482s
6	53.70..53.72	0.931s
7	53.70..53.72	0.875s
8	89.02..89.02	1.780s
9	517.27..517.30	1.323s
10	520.28..521.08	0.030s
11	638.51..638.53	1.827s
12	204.20..204.54	0.016s
13	12185.91..12186.23	0.070s
14	25777.41..25799.02	0.056s

Index Choices

Most of the queries that we have designed use the following attributes for join multiple times:

i)crop_id present in crops, state_crop_production, condition,nutrtion, india_crop_production etc. tables
ii)state_name present in state_condition and state_crop_production table
iii)crop_name present in crops table
Thus we have designed index for all these choices to improve our performance.