HR- EMPLOYEE DISTRIBUTION DASHBOARD

SQL QUERIES

1. DATA CLEANING
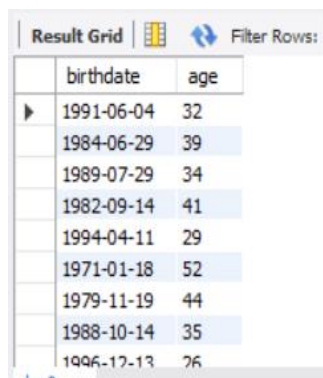
a. alter table hr
   change column ï»¿id emp_id varchar(20) null;

b. describe hr ;

c. set sql_safe_updates = 0;
   update hr
   set birthdate = case
       when birthdate like '%/%' then date_format(str_to_date(birthdate, '%m/%d/%Y'), '%Y-%m-%d')
       when birthdate like '%-%' then date_format(str_to_date(birthdate, '%m-%d-%Y'), '%Y-%m-%d')
       else null
   end;

d. alter table hr
   modify column birthdate date ;

e. update hr
   set hire_date = case
       when hire_date like '%/%' then date_format(str_to_date(hire_date, '%m/%d/%Y'), '%Y-%m-%d')
       when hire_date like '%-%' then date_format(str_to_date(hire_date, '%m-%d-%Y'), '%Y-%m-%d')
       else null
   end;

f. alter table hr
   modify column hire_date date ;

g. update hr
   set termdate = date(str_to_date(termdate, '%Y-%m-%d %H:%i:%s UTC'))

where termdate is not null and termdate <> '' ;

h. alter table hr
    modify column termdate date;

i. alter table hr
    add column age int;

j. update hr
    set age = timestampdiff(year, birthdate , curdate());

k. select birthdate, age from hr ;

birthdate	age
1991-06-04	32
1984-06-29	39
1989-07-29	34
1982-09-14	41
1994-04-11	29
1971-01-18	52
1979-11-19	44
1988-10-14	35
1996-12-13	26

l. select
        min(age) as youngest,
        max(age) as oldest
    from hr;

m. select count(*)
    from hr
    where age < 18;

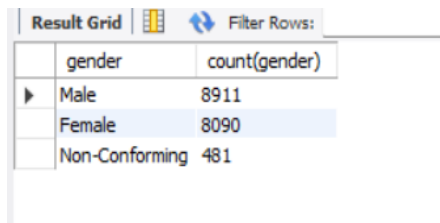## 2. DATA ANALYSIS

-- QUESTIONS

-- 1. What is the gender breakdown of employess in a company?

select gender , count(gender)

from hr

where age >= 18 and termdate = "

group by gender;

gender	count(gender)
Male	8911
Female	8090
Non-Conforming	481

-- 2. What is the race/ethinicity breakdown of employees in a company?
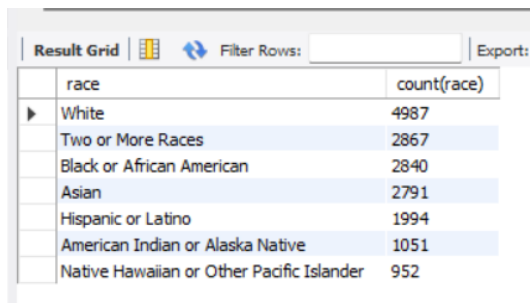
select race , count(race)

from hr

where age >= 18 and termdate = "

group by race

order by count(race) desc;

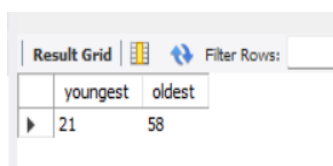| race	count(race)
White	4987
Two or More Races	2867
Black or African American	2840
Asian	2791
Hispanic or Latino	1994
American Indian or Alaska Native	1051
Native Hawaiian or Other Pacific Islander	952

-- 3. What is the age distribution of employees in a company?

select

    min(age) as youngest,

    max(age) as oldest

from hr

where age >= 18 and termdate = ";

youngest	oldest
21	58

```sql
select
    case
        when age >= 18 and age <= 24 then '18-24'
        when age >= 25 and age <= 34 then '25-34'
        when age >= 35 and age <= 44 then '35-44'
        when age >= 45 and age <= 54 then '45-54'
        when age >= 55 and age <= 64 then '55-64'
        else '65+'
        end as age_group,
    count(*) as count
    from hr
    where age >= 18 and termdate = ''
    group by age_group
    order by age_group;
```

Result Grid | Filter Row

age_group	count
18-24	1878
25-34	4964
35-44	4980
45-54	4716
55-64	944

-- GENDER DISTRIBUTIONS IN AGE GROUP

```sql
select
    case
        when age >= 18 and age <= 24 then '18-24'
        when age >= 25 and age <= 34 then '25-34'
        when age >= 35 and age <= 44 then '35-44'
        when age >= 45 and age <= 54 then '45-54'
        when age >= 55 and age <= 64 then '55-64'
        else '65+'
        end as age_group, gender,
```

count(*) as count

from hr

where age >= 18 and termdate = ''

group by age_group,gender

order by age_group, gender;

age_group	gender	count
18-24	Female	853
18-24	Male	979
18-24	Non-Conforming	46
25-34	Female	2350
25-34	Male	2477
25-34	Non-Conforming	137
35-44	Female	2229
35-44	Male	2612
35-44	Non-Conforming	139

-- 4. How many employees work at headquaters vs remote locations

select location, count(location)

from hr

where age >= 18 and termdate = ''

group by location

order by count(location) desc;

location	count(location)
Headquarters	13107
Remote	4375

-- 5. What is the average length of employment of employees who have been terminated?

select round(avg(datediff(termdate,hire_date))/365,0) as avg_employment_length

from hr

where termdate <= curdate() and age >= 18 and termdate <> '';

avg_employment_length
8

-- 6. How does the gender distribution vary across departments and job titles?

select gender, department, count(*) as department_count

from hr

where age >= 18 and termdate = ''

group by gender, department

order by department;

gender	department	department_count
Male	Accounting	1375
Female	Accounting	1175
Non-Conforming	Accounting	76
Female	Auditing	19
Male	Auditing	19
Male	Business Development	672
Female	Business Development	593
Non-Conforming	Business Development	42
Male	Engineering	2671

-- 7. What is the distribution of job titles across the comapny?

select jobtitle, count(*) as jobtitle_count

from hr

where age >= 18 and termdate = ''

group by jobtitle

order by count(*) desc ;

jobtitle	jobtitle_count
Research Assistant II	608
Business Analyst	552
Human Resources Analyst II	477
Research Assistant I	408
Account Executive	386
Staff Accountant I	364
Data Visualization Specialist	346
Human Resources Analyst	324

-- 8. Which department has highest turnover rate?

select department,

    total_count,

    terminated_count,

    terminated_count/total_count as termination_rate

    from(

select department, count(*) as total_count,

sum(case when termdate <> '' and termdate <= curdate() then 1 else 0 end) as terminated_count

from hr

where age >= 18

group by department) as subquery

order by termination_rate desc;

department	total_count	terminated_count	termination_rate
Auditing	50	8	0.1600
Legal	299	39	0.1304
Training	1622	193	0.1190
Research and Development	1032	120	0.1163
Human Resources	1727	196	0.1135
Accounting	3192	359	0.1125
Sales	1745	195	0.1117
Engineering	6387	707	0.1107

-- 9. What is the distribution of employess across different location by city and state?

select location_state, count(location_state) as count_state

 from hr

 where age >= 18 and termdate = ''

 group by location_state

 order by count_state desc;

location_state	count_state
Ohio	14144
Pennsylvania	892
Illinois	698
Michigan	550
Indiana	545
Kentucky	347
Wisconsin	306

-- 10. How has the company's employee count change over time based on hire date and term date?

select

  year,

  hires,

  terminates,

  hires - terminates as net_change,

round((hires-terminates)/hires * 100,2)as net_change_percent

from(

    select year(hire_date) as year,

    count(*) as hires,

    sum(case when termdate <> '' and termdate <= curdate() then 1 else 0 end) as terminates

    from hr

    where age <= 18

    group by year(hire_date)

) as subquery

    order by year asc;

year	hires	terminates	net_change	net_change_percent
2000	9	5	4	44.44
2001	40	6	34	85.00
2002	55	12	43	78.18
2003	54	5	49	90.74
2004	48	5	43	89.58
2005	59	7	52	88.14
2006	49	8	41	83.67
2007	32	6	26	81.25

-- 11. What is the tenure distribution for each department

    select department, round(avg(datediff(termdate, hire_date)/365),0) as avg_tenure

    from hr

    where termdate <= curdate() and termdate <> '' and age >= 18

    group by department

    order by avg_tenure desc;

department	avg_tenure
Sales	9
Engineering	8
Services	8
Auditing	8
Accounting	8
Research and Development	8
Marketing	8
Human Resources	7
Business Development	7