**1--To change to the required date format (yyyy-mm-dd)**

**Changing everything to yyyy-mm—dd fromat**

UPDATE hr\_data

SET birthdate = TO\_CHAR(TO\_DATE(birthdate, 'MM/DD/YYYY'), 'YYYY-MM-DD');

---------------------------------------------------------------------------------------------------------

**Changing 00 s to respective values (0071 to 1971, 0001 to 2001)**

UPDATE hr\_data

SET birthdate =

CASE

WHEN SUBSTRING(birthdate, 3, 2)::integer < 10 THEN CONCAT('20', SUBSTRING(hire\_date, 3))

ELSE CONCAT('19', SUBSTRING(birthdate, 3))

END

WHERE birthdate LIKE '00%';

(same was done for hire\_date) excluding integer < 25

**2--changing the column to date type (from varchar(50))**

ALTER TABLE hr\_data

ALTER COLUMN birthdate TYPE DATE

USING TO\_DATE(birthdate, 'YYYY-MM-DD');

**3--modifying the termdate colum**

UPDATE hr\_data

SET termdate = TO\_CHAR(TO\_DATE(left(termdate, 19), 'YYYY-MM-DD HH24:MI:SS'), 'YYYY-MM-DD');

**4—calculating age of employees**

alter table hr\_data

add column age varchar(50)

UPDATE hr\_data

SET age = EXTRACT(YEAR FROM AGE(current\_date, birthdate));

**QUESTIONS**

**1--What's the age distribution in the company?**

select max(age) as oldest,

min(age) as youngest

from hr\_data

{alter table hr\_data

alter column age

type int using age::integer}

-- changing age column to integer

--------------------------------------------------------------------------------------------

--age group

select age\_group, count(\*) as count from

(select case

when age >= 21 and age<=30 then '21 to 30'

when age >= 31 and age<=40 then '31 to 40'

when age >= 41 and age<=50 then '41 to 50'

else'50+'

end as age\_group

from hr\_data

) as subquery

group by age\_group

order by age\_group

---------------------------------------------------------------------------------------------------------

--age group by gender

select age\_group,gender, count(\*) as count from

(select case

when age >= 21 and age<=30 then '21 to 30'

when age >= 31 and age<=40 then '31 to 40'

when age >= 41 and age<=50 then '41 to 50'

else'50+'

end as age\_group, gender

from hr\_data

) as subquery

group by age\_group, gender

order by age\_group

**2--What's the gender breakdown in the company?**

select gender, count(\*) as count

from hr\_data

group by gender

order by gender

**3--How does gender vary across departments and job titles?**

-- across departments

select department,gender, count(\*) as count

from hr\_data

group by department,gender

order by department,gender

--------------------------------------------------------------------------------------------

--across job title

select jobtitle,gender, count(\*) as count

from hr\_data

group by jobtitle,gender

order by jobtitle,gender

**4--What's the race distribution in the company?**

select race, count(\*) as count

from hr\_data

group by race

order by count desc

**5--What's the average length of employment in the company?**

SELECT round(AVG(DATE\_PART('year', CAST(termdate AS DATE)) - DATE\_PART('year', CAST(hire\_date AS DATE)))::numeric,2) AS tenure

from hr\_data

WHERE termdate IS NOT NULL AND cast(termdate as date) <= CURRENT\_DATE;

**6--Which department has the highest turnover rate?**

select department, total\_count, terminated\_count

, round(cast(terminated\_count as numeric)/cast(total\_count as numeric),2) \* 100 as turnover\_rate from

(select department, count(\*) as total\_count,

sum(case

when termdate is not null and cast(termdate as date) <= current\_date then 1 else 0

end

) as terminated\_count

from hr\_data

group by department)as subquery

order by turnover\_rate desc

**7--What is the tenure distribution for each department?**

SELECT department,round(AVG(DATE\_PART('year', CAST(termdate AS DATE)) - DATE\_PART('year', CAST(hire\_date AS DATE)))::numeric,2) AS tenure

from hr\_data

WHERE termdate IS NOT NULL AND cast(termdate as date) <= CURRENT\_DATE

group by department

order by tenure desc;

**8--How many employees work remotely for each department?**

select location, count(\*)

from hr\_data

group by location

**9--What's the distribution of employees across different states?**

select location\_state, count(\*)

from hr\_data

group by location\_state

order by count desc

**10--How are job titles distributed in the company?**

select jobtitle, count(\*)

from hr\_data

group by jobtitle

order by count desc

**11--How have employee hire counts varied over time?**

select hire\_year, hires, hires - terminations as net\_change,

round((cast(hires as numeric)-cast(terminations as numeric))/cast(hires as numeric)\*100::numeric,2) as percent\_hire\_change from

(select extract(year from hire\_date) as hire\_year, count(\*) as hires,

sum(case when termdate is not null and cast(termdate as date) <= current\_date then 1

else 0 end )as terminations

from hr\_data

group by extract(year from hire\_date)

order by extract(year from hire\_date) ) as subquery

order by percent\_hire\_change desc