**Hr Dashboard SQL Queries**

Import db into phpMyAdmin

**Data cleaning**

1. Change tablename

ALTER TABLE human\_resources RENAME hr;

1. Change id to employee id

ALTER TABLE hr

CHANGE COLUMN id emp\_id varchar(20) NULL;

1. Change birthdate format

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;

1. Change datatype of birthdate from varchar to date

[ALTER](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) [TABLE](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) hr MODIFY COLUMN birthdate [DATE](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/date-and-time-types.html);

1. Change hire\_date format

UPDATE IGNORE hr

SET hire\_date = CASE

WHEN hire\_date LIKE '%/%' THEN date\_format(str\_to\_date(birthdate, '%m/%d/%Y'), '%Y-%m-%d')

WHEN hire\_date LIKE '%-%' THEN date\_format(str\_to\_date(birthdate, '%m-%d-%Y'), '%Y-%m-%d')

ELSE NULL

END;

1. Change datatype of hire\_date from varchar to date

[ALTER](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) [TABLE](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) hr MODIFY COLUMN hire\_date [DATE](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/date-and-time-types.html)

1. Change format of termdate column

UPDATE hr

SET termdate = date(str\_to\_date(termdate,'%Y-%m-%d %H:%i:%s UTC'))

WHERE termdate IS NOT NULL AND termdate !='';

SELECT termdate FROM hr;

1. Change datatype of termdate

ALTER TABLE hr

MODIFY COLUMN termdate DATE;

1. Adding column age

[ALTER](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) [TABLE](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/alter-table.html) hr ADD COLUMN age [int](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/numeric-types.html" \t "mysql_doc)

1. Calculating age based on birthdate column

UPDATE hr

SET age = timestampdiff(YEAR,birthdate,CURRENT\_DATE());

SELECT \* FROM hr;

1. Youngest and eldest employee

SELECT

MIN(age) as youngest,

MAX(age) as oldest

FROM hr;

Count of ages less than 18

SELECT COUNT(\*) FROM hr

WHERE age < 18;

Output: 967

Therefore can be avoided or discarded (dataset contains over 22000 records)

**Data analysis:**

1. What is the gender breakdown of the employees in the company?

SELECT gender, COUNT(\*) as gender\_count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY gender;

2. Race/ ethnicity breakdown of employees in the company

SELECT race, COUNT(\*) AS race\_count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY race

ORDER BY race\_count DESC;

3. Age distribution of employees in the company

[SELECT](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/select.html) [MIN](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/group-by-functions.html#function_min)(age) as youngest, [MAX](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/group-by-functions.html#function_max)(age) AS oldest FROM hr WHERE age>=18 [AND](http://localhost/phpmyadmin/url.php?url=https://dev.mysql.com/doc/refman/5.5/en/logical-operators.html#operator_and) termdate = '0000-00-00'

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 age\_count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY age\_group;

1. Distribution by age\_group and gender

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 age\_count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY age\_group, gender;

1. How many employees work at headquarters versus remote locations?

SELECT location, COUNT(\*) as location\_count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY location;

6. Distribution of gender and departments across the company

SELECT gender, department, COUNT(\*) as count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY gender, department;

7. Distribution of job titles across the company

SELECT jobtitle, COUNT(\*) AS count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY jobtitle;

Turnover rate of the company

SELECT department,

total\_count,

terminated\_count,

terminated\_count/total\_count AS termination\_rate

FROM (

SELECT department,

COUNT(\*) as total\_count,

SUM(CASE WHEN termdate <> '0000-00-00' AND termdate<=CURRENT\_DATE() THEN 1 ELSE 0 END) as terminated\_count

FROM hr

WHERE age>=18

GROUP BY department

) AS subquery

ORDER BY termination\_rate

8. Distribution of employees across locations by city and state

SELECT location\_state, COUNT(\*) AS count

FROM hr

WHERE age>=18 AND termdate = '0000-00-00'

GROUP BY location\_state

ORDER BY count DESC;

9. How has the company’s employee count changed over time based on hire and term dates

SELECT

year,

hires,

terminations,

hires - terminations AS net\_change,

round((hires - terminations)/hires\*100,2) AS net\_change\_percent

FROM(

SELECT

YEAR(hire\_date) as year,

COUNT(\*) AS hires,

SUM(CASE WHEN termdate <> '0000-00-00' AND termdate <= CURRENT\_DATE() THEN 1 ELSE 0 END) as terminations

FROM hr

WHERE age>=18

GROUP by year

) as subquery

ORDER BY year ASC

10. Tenure distribution of each department

SELECT department, round(AVG(datediff(termdate,hire\_date)/365),0) AS avg\_tenure

FROM hr

WHERE termdate<=CURRENT\_DATE() AND termdate <> '0000-00-00' and age>=18

GROUP BY department;

1. Hires and terminations till now

SELECT

hires,

terminations

FROM(

SELECT

COUNT(\*) AS hires,

SUM(CASE WHEN termdate <> '0000-00-00' AND termdate <= CURRENT\_DATE() THEN 1 ELSE 0 END) as terminations

FROM hr

WHERE age>=18

) as subquery

Highest employees in department

SELECT department, COUNT(\*) AS max\_employees\_department

FROM hr

GROUP BY department

ORDER BY COUNT(\*) DESC

LIMIT 1;