**SQL Project on HR**

**Objectives:**

1. **Understand Employee Demographics:**
2. **Assess Diversity and Inclusion:**
3. **Analyze Employee Age Distribution:**
4. **Evaluate Employee Location Distribution:**
5. **Study Termination Statistics:**
   * To calculate the average length of employment for terminated employees and assess patterns in employment duration before termination.
6. **Explore Gender Distribution in Departments and Job Titles:**
   * To analyze the gender distribution across departments and job titles, enabling a deeper understanding of gender representation in different roles.
7. **Assess Job Title Distribution:**
   * To determine the distribution of job titles across the company and identify the prevalence of each job role.
8. **Identify Department with High Turnover Rate:**
   * To identify the department with the highest turnover rate and explore potential factors contributing to turnover.
9. **Examine Employee Location Distribution:**
   * To understand the distribution of employees across locations by city and state and assess geographical representation.
10. **Track Employee Count Changes Over Time:**
    * To monitor changes in the company's employee count over time based on hire and termination dates, aiding in human resources planning.
11. **Analyze Tenure Distribution per Department:**
    * To assess the tenure distribution for each department and understand the average length of service for employees in various departments.

**Analysis**

-- What is the age distribution of employees in the company?

-- by 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 count

from hr

where age >= 18

group by age\_group,gender

order by age\_group,gender;

-- How many employees work at headquarters versus remote locations?

select location, count(\*) as count

from hr

where age>= 18

group by location;

|  |  |
| --- | --- |
| Headquarters | 15992 |
| Remote | 5255 |

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

select

round(avg(datediff(termdate,hire\_date))/365,0) as avg\_length\_employment

from hr

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

Avg length of employment = 8

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

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

from hr

where age>= 18

group by department,gender

order by department;

-- Which department has the 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 <= curdate() then 1 else 0 end) terminated\_count

from hr

where age >=18

group by department

) as subquery

order by termination\_rate desc;

|  |  |  |  |
| --- | --- | --- | --- |
| department | total\_count | terminated\_count | termination\_rate |
| Marketing | 480 | 453 | 0.9438 |
| Training | 1622 | 1526 | 0.9408 |
| Human Resources | 1727 | 1620 | 0.938 |
| Services | 1618 | 1508 | 0.932 |
| Accounting | 3192 | 2973 | 0.9314 |
| Engineering | 6387 | 5943 | 0.9305 |
| Product Management | 623 | 578 | 0.9278 |
| Business Development | 1569 | 1453 | 0.9261 |
| Sales | 1745 | 1614 | 0.9249 |
| Legal | 299 | 276 | 0.9231 |
| Auditing | 50 | 46 | 0.92 |
| Research and Development | 1032 | 944 | 0.9147 |
| Support | 903 | 824 | 0.9125 |

-- What is the distribution of employees across locations by city and state?

select location\_state, count(\*) as count

from hr

where age >=18

group by location\_state

order by count desc;

|  |  |
| --- | --- |
| **location\_state** | **count** |
| Ohio | 17252 |
| Pennsylvania | 1066 |
| Illinois | 833 |
| Indiana | 666 |
| Michigan | 650 |
| Kentucky | 417 |
| Wisconsin | 363 |

**Summary of Findings**

- There are more male employees

- White race is the most dominant while Native Hawaiian and American Indian are the least dominant.

- The youngest employee is 20 years old and the oldest is 57 years old

- 5 age groups were created (18-24, 25-34, 35-44, 45-54, 55-64). A large number of employees were between 25-34 followed by 35-44 while the smallest group was 55-64.

- A large number of employees work at the headquarters versus remotely.

- The average length of employment for terminated employees is around 7 years.

- The gender distribution across departments is fairly balanced but there are generally more male than female employees.

- The Marketing department has the highest turnover rate followed by Training. The least turn over rate are in the Research and development, Support and Legal departments.

- A large number of employees come from the state of Ohio.

- The net change in employees has increased over the years.

- The average tenure for each department is about 8 years with Legal and Auditing having the highest and Services, Sales and Marketing having the lowest.

**Limitations**

- Some records had negative ages and these were excluded during querying(967 records). Ages used were 18 years and above.

- Some termdates were far into the future and were not included in the analysis(1599 records). The only term dates used were those less than or equal to the current date.