



HR DATASETS

MYSQL

Prepared by:
Reyan Azam Hashmi

Human Resource Data Analysis

This project analyzes employee data from a Human Resource database using MySQL to gain insights into workforce demographics and trends. It answers key HR-related questions such as gender and ethnicity breakdowns, age distribution, remote vs. headquarters staffing, average employment length, departmental turnover, job title distribution, and employee growth over time. The analysis provides valuable metrics to understand organizational structure, diversity, and retention patterns.

What is the gender breakdown of employees in the company?

```
select gender, count(*) from hr  
where age > 18 and termdate is null  
group by gender;
```

Result Grid			Filter Rows:
	gender	count(*)	
▶	Male	8259	
	Female	7464	
	Non-Conforming	448	

What is the race/ethnicity breakdown of employees in the company?

```
1 -- What is the race/ethnicity breakdown of employees in  
2 • select race, count(*) as Total from hr  
3 where age > 18 and termdate is null  
4 group by race;
```

Result Grid			Filter Rows:	Export:
	race	Total		
▶	Hispanic or Latino	1858		
	White	4602		
	Black or African American	2637		
	Asian	2563		
	Two or More Races	2645		
	American Indian or Alaska Native	976		
	Native Hawaiian or Other Pacific Islander	890		

What is the age distribution of employees in the company?

```
2 •      select
3      case
4      when age >=18 and age<=24 then '18-24'
5        when age >=25 and age<=34 then '24-34'
6        when age >=35 and age<=44 then '35-44'
7          when age >=45 and age<=54 then '45-54'
8            when age >=55 and age<=64 then '55-64'
9            else '65+'
10        end as age_group,
11        count(*)as count
12      from hr
13      WHERE age > 18
14      AND termdate IS NULL group by age_group;
```

▶	24-34	4910
	35-44	5071
	45-54	4853
	18-24	949
	55-64	388

How many employees work at headquarters versus remote locations?

```
2 • select location, count(*) as Total
3   from hr
4   where age > 18 and termdate is null
5   group by location;
```

Result Grid			Filter Rows:
	location	Total	
▶	Headquarters	12134	
	Remote	4037	




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

```
2 • SELECT
3     ROUND(AVG(TIMESTAMPDIFF(YEAR, hire_date, termdate)), 0) AS avg_employment_years
4 FROM hr
5 WHERE termdate <= CURDATE()
6     AND termdate IS NOT NULL
7     AND age >= 18;
8
```

Result Grid		Filter Rows:
	avg_employment_years	
▶	8	

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

```
select department,gender,count(*) from hr
where age>-18 and termdate is null
group by department,gender
order by department;
```

Result Grid   Filter Rows: <input type="text"/> Export: 			
	department	gender	count(*)
▶	Accounting	Female	1086
	Accounting	Male	1265
	Accounting	Non-Conforming	74
	Auditing	Female	18
	Auditing	Male	17
	Business Development	Female	545
	Business Development	Male	613
	Business Development	Non-Conforming	42
	Engineering	Female	2238
	Engineering	Male	2463



What is the distribution of job titles across the company?

```
select jobtitle, count(*) as total_employees  
from hr  
where age >= 18 and termdate is not null  
group by jobtitle;
```

	jobtitle	total_employees
▶	Business Analyst	126
	Business Systems Development Analyst	49
	Service Manager	46
	Senior Recruiter	25
	HR Manager	30
	Relationship Manager	36
	Developer III	17
	Desktop Support Technician	59
	Service Tech III	45
	Internal Auditor	12
	Administrative Officer	12
	Recruiter	36
	Training Manager	27
	Cost Accountant	47
	Software Test Engineer III	17

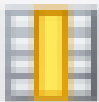

Which department has the highest termination rate?

```
2 • use human_resource;
3 • select department, Total_employees, terminated_employees,
4     terminated_employees/Total_employees as termination_rate
5     from
6     (select department, count(*) as Total_employees,
7         sum(case when termdate is not null and termdate <= curdate()
8             from hr
9             where age >= 18
10            group by department ) as sub_query
11     order by termination_rate desc;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
department	Total_employees	terminated_employees	termination_rate
Auditing	47	10	0.2128
Legal	280	44	0.1571
Training	1513	206	0.1362
Support	834	110	0.1319
Human Resources	1607	207	0.1288
Engineering	5874	747	0.1272
Sales	1612	205	0.1272
Research and Development	972	121	0.1245
Accounting	2956	366	0.1238
Product Management	578	71	0.1228
Services	1512	183	0.1210
Business Development	1447	163	0.1126
Marketing	447	48	0.1074

What is the distribution of employees across locations by state?

```
8 • select location_state,count(*)as count
9     from hr
10    where age>=18 and termdate is null
11    group by location_state
12    order by count desc;
```

Result Grid   Filter Rows		
	location_state	count
▶	Ohio	13100
	Pennsylvania	820
	Illinois	649
	Michigan	512
	Indiana	506
	Kentucky	308
	Wisconsin	276

What is the distribution of employees across locations by city?

```
2 • select location_city,count(*)as count
3 from hr
4 where age>=18 and termdate is null
5 group by location_city
6 order by count desc;
```

Result Grid			Filter Rows:
	location_city	count	
▶	Cleveland	12247	
	Chicago	265	
	Philadelphia	249	
	Cincinnati	205	
	Pittsburgh	203	
	Detroit	155	
	Dayton	154	
	Louisville	152	
	Indianapolis	141	
	Lexington	139	
	Columbus	126	
	Springfield	126	
	Milwaukee	126	
	Fort Wayne	110	
	Peoria	100	

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 <=curdate() and termdate is not null then 1 else 0 end)as terminations  
from hr  
where age>=18  
group by year(hire_date))as sub_query  
order by year asc;
```

	year	hires	terminations	net_change	net_change_percent
▶	2000	188	22	166	88.30
	2001	993	179	814	81.97
	2002	951	155	796	83.70
	2003	998	183	815	81.66
	2004	998	193	805	80.66
	2005	971	183	788	81.15
	2006	986	183	803	81.44
	2007	1001	154	847	84.62
	2008	978	144	834	85.28
	2009	1008	148	860	85.32
	2010	965	129	836	86.63
	2011	990	119	871	87.98
	2012	981	115	866	88.28
	2013	958	96	862	89.98
	2014	952	91	861	90.44

What is the tenure distribution for each department?

```
select department,round(avg(datediff(termdate,hire_date)/365),0)as average_tenure
from hr
where termdate<=curdate() and termdate is not null and age>=18
group by department ;
```

Result Grid			Filter Rows:
	department	average_tenure	
▶	Engineering	8	
	Services	9	
	Human Resources	8	
	Business Development	8	
	Sales	9	
	Support	8	
	Auditing	8	
	Training	8	
	Accounting	8	
	Research and Development	8	
	Product Management	7	
	Legal	7	
	Marketing	8	