

## HR

### DATA CLEANING/MANIPUALTION

To convert the data type from var char to date in the date columns

```
ALTER TABLE hr
```

```
ALTER COLUMN birthdate TYPE DATE using(birthdate::DATE),
```

```
ALTER COLUMN hire_date TYPE DATE using(hire_date::DATE),
```

```
ALTER COLUMN termdate TYPE DATE using(termdate::DATE);
```

Using the CAST function can also be used to change the data type in the respective columns

```
UPDATE hr
```

```
SET birthdate = CAST(birthdate AS DATE),
```

```
    hire_date = CAST(hire_date AS DATE),
```

```
    termdate = CAST (termdate AS DATE);
```

To add an age column and get the age from the birthdate

```
ALTER TABLE hr
```

```
ADD COLUMN Age INT;
```

```
UPDATE hr
```

```
SET Age = EXTRACT (YEAR FROM AGE(CURRENT_DATE, birthdate));
```

OR

```
Set Age = DATE_PART ('year', CURRENT_DATE) - DATE_PART ('year', birthdate)
```

### QUESTIONS

Assuming the age needed to be employed is 18 and above, the following questions will be answered.


Total employee count

```
SELECT COUNT (id) AS total_number_employees
```

```
FROM hr
```

```
WHERE Age >= 18
```

```
AND termdate IS NULL;
```

	total_number_employees 
1	17482

Employees by gender

```
SELECT gender, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY gender;
```

	gender text	total_number_employees bigint
1	Male	8911
2	Female	8090
3	Non-Conforming	481

Employee count based on race

```
SELECT race, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY race
ORDER BY total_number_employees DESC;
```

	race character varying (50)	total_number_employees bigint
1	White	4987
2	Two or More Races	2867
3	Black or African American	2840
4	Asian	2791
5	Hispanic or Latino	1994
6	American Indian or Alaska Native	1051
7	Native Hawaiian or Other Pacific Islander	952

Age distribution in the company

```
SELECT
CASE
WHEN age BETWEEN 18 AND 30 THEN '<30'
WHEN age BETWEEN 31 AND 40 THEN '31-40'
WHEN age BETWEEN 41 AND 50 THEN '41-50'
ELSE '>50'
END AS Age_category,
COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY Age_category
```

ORDER BY total\_number\_employees DESC;

	age_category text	total_number_employees bigint
1	31-40	5078
2	41-50	4838
3	<30	4798
4	>50	2768

Breakdown of the genders according to age distribution.

SELECT

CASE

WHEN age BETWEEN 18 AND 30 THEN '<30'

WHEN age BETWEEN 31 AND 40 THEN '31-40'

WHEN age BETWEEN 41 AND 50 THEN '41-50'

ELSE '>50'

END AS Age\_category,

gender,

COUNT(id) AS total\_number\_employees

FROM hr

WHERE Age >= 18

AND termdate IS NULL

GROUP BY Age\_category, gender

ORDER BY total\_number\_employees DESC;

	age_category text	gender text	total_number_employees bigint
1	31-40	Male	2617
2	41-50	Male	2475
3	<30	Male	2429
4	31-40	Female	2312
5	<30	Female	2248
6	41-50	Female	2224
7	>50	Male	1390
8	>50	Female	1306
9	31-40	Non-Conforming	149
10	41-50	Non-Conforming	139
11	<30	Non-Conforming	121
12	>50	Non-Conforming	72

```

SELECT location, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY location
ORDER BY total_number_employees DESC;

```

	location character varying (50) 🔒	total_number_employees bigint 🔒
1	Headquarters	13107
2	Remote	4375

```

SELECT department, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY department
ORDER BY total_number_employees DESC;

```

	department character varying (100) 🔒	total_number_employees bigint 🔒
1	Engineering	5259
2	Accounting	2626
3	Human Resources	1430
4	Sales	1426
5	Training	1338
6	Services	1337
7	Business Development	1307
8	Research and Development	830
9	Support	732
10	Product Management	512
11	Marketing	410
12	Legal	237
13	Auditing	38

```

SELECT jobtitle, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY jobtitle
ORDER BY total_number_employees DESC
LIMIT 10;

```

	jobtitle character varying (100) 🔒	total_number_employees bigint 🔒
1	Research Assistant II	608
2	Business Analyst	552
3	Human Resources Analyst II	477
4	Research Assistant I	408
5	Account Executive	386
6	Staff Accountant I	364
7	Data Visualization Specialist	346
8	Human Resources Analyst	324
9	Software Engineer I	308
10	Systems Administrator I	302

```

SELECT location_city AS city, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY location_city
ORDER BY total_number_employees DESC
LIMIT 10;

```

	city character varying (100) 🔒	total_number_employees bigint 🔒
1	Cleveland	13233
2	Chicago	283
3	Philadelphia	268
4	Pittsburgh	229
5	Cincinnati	221
6	Louisville	174
7	Detroit	165
8	Dayton	161
9	Lexington	154
10	Indianapolis	154

```

SELECT location_state AS state, COUNT (id) AS total_number_employees
FROM hr
WHERE Age >= 18
AND termdate IS NULL
GROUP BY location_state
ORDER BY total_number_employees DESC
LIMIT 10;

```

	state character varying (100)	total_number_employees bigint
1	Ohio	14144
2	Pennsylvania	892
3	Illinois	698
4	Michigan	550
5	Indiana	545
6	Kentucky	347
7	Wisconsin	306

Information about employee duration

```

SELECT
    ROUND (AVG (DATE_PART ('year', termdate) - DATE_PART ('year',
hire_date)))
    AS average_length_before_termination_in_years,
    MAX (DATE_PART ('year', termdate) - DATE_PART ('year', hire_date))
    AS maximum_stay_before_termination_in_years,
    ABS (MIN (DATE_PART ('month', termdate) - DATE_PART ('month',
hire_date)))
    AS minimum_stay_before_termination_in_months
FROM hr
WHERE Age >=18
AND termdate IS NOT NULL
AND termdate <= CURRENT_DATE

```

	average_length_before_termination_in_years double precision	maximum_stay_before_termination_in_years double precision	minimum_stay_before_termination_in_months double precision
1	8	21	11

## INSIGHTS

This is exploratory data that shows the employee profile to provide HR with the demographic patterns in the company to assist in making several decisions as regards hiring. Combining this with productivity trends should give clearer information across the various categories.