# *HR data analysis using postgresql*

First thing first let’s start by creating a table in our database to import our data into it

**Create Table**

create table hrdata

(

emp\_no int8 PRIMARY KEY,

gender varchar(50) NOT NULL,

marital\_status varchar(50),

age\_band varchar(50),

age int8,

department varchar(50),

education varchar(50),

education\_field varchar(50),

job\_role varchar(50),

business\_travel varchar(50),

employee\_count int8,

attrition varchar(50),

attrition\_label varchar(50),

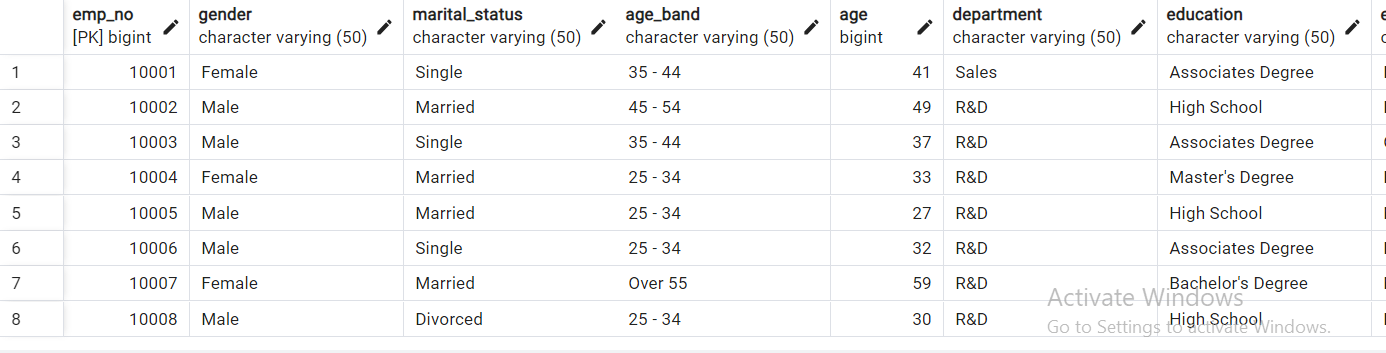
job\_satisfaction int8,

active\_employee int8 )

After creating the table , you can import your data into the table through many ways , the easiest one is through your application import wizard

Next let’s have a look at the data imported to our database

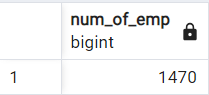
select \* from hrdata limit 8 ;



As you can see the data imported successfully.

## Calculate the total number of employees:

select count(employee\_count) as num\_of\_emp from hrdata ;



## Calculate the attrition cases :

We will access the attrition column and count the cases when the attrition column indicates to Yes

select count(attrition) as attrition\_cases from hrdata where attrition ='Yes' ;

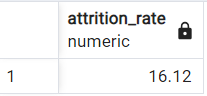


## Calculate the attrition rate :

We will simply calculate this by dividing the number of attrition cases on the total number of employees

\*we rounded the answer to 2 decimal points to be easy to read

select round((select count(attrition) from hrdata where attrition = 'Yes') / sum(employee\_count)\*100,2) as attrition\_rate from hrdata ;

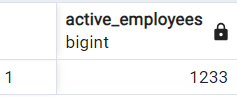


# Calculate the active employees

We will calculate the employees who are actively working at the moment by accessing the column ‘active employees’

select count(active\_employee) as active\_employees from hrdata

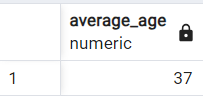
where active\_employee = 1 ;



# average age

\*Notice we rounded the number to the closest integer number so we get an average age

select round (avg (age),0) from hrdata ;

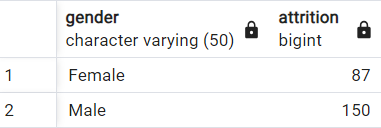


# attrition by gender :

check the attrition cases with respect to the gender

select gender , count(attrition) as Attrition from hrdata where attrition ='Yes'

group by gender;

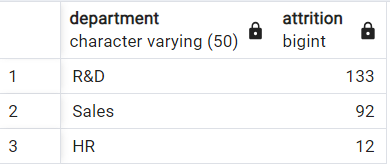


## attrition cases per department :

select department , count(attrition) as Attrition from hrdata where attrition ='Yes'

group by department

order by count(attrition) desc ;

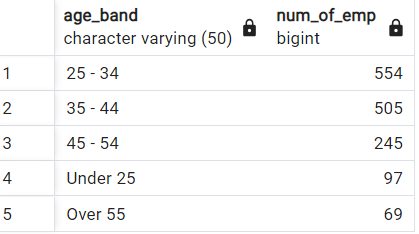


## number of employees by age group :

select age\_band , count(employee\_count) as num\_of\_emp from hrdata

group by age\_band

order by count(employee\_count) desc;

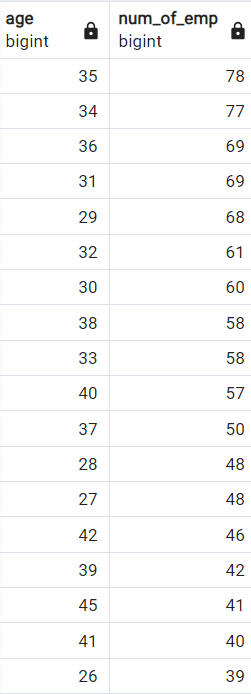


## number of employees by age :

select age , count(emp\_no) as num\_of\_emp from hrdata

group by age

order by count(emp\_no) desc;



## Attrition according to education:

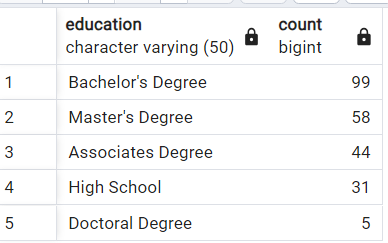
## 

select education , count(attrition) from hrdata

where attrition = 'Yes'

group by education

order by count(attrition) desc;



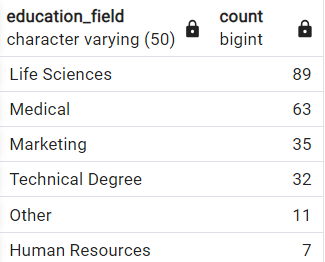
## Attrition according to education field :

select education\_field , count(attrition) from hrdata

where attrition = 'Yes'

group by education\_field

order by count(attrition) desc;



## Attrition by gender and different age group :

We select the age band and count of attrition for each age band , also the gender whether it’s a male or a female

Here comes the tricky part while calculating the percentage of attrition since we rounded the answer to 2 decimal places and castes the count of attrition to be a numeric value.

select age\_band , count(attrition) as Attrition , gender ,

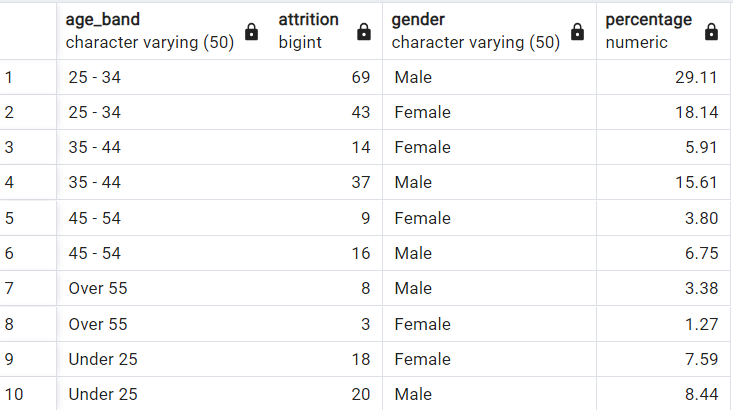
round((cast(count(attrition)as numeric)/

(select count(attrition) from hrdata where attrition ='Yes'))\*100 ,2) as percentage

from hrdata where attrition ='Yes'

group by age\_band , gender

order by age\_band ;



## Create a table containing the Job role and the satisfaction data for it :

Please note that we used the Crosstab function and you may face an error that you don’t have that function already so use that code

CREATE EXTENSION IF NOT EXISTS TABLEFUNC;

select \* from

crosstab ( 'select job\_role ,job\_satisfaction,sum(employee\_count)

from hrdata

group by job\_role,job\_satisfaction

order by job\_role,job\_satisfaction')

As crosstab (job\_role varchar(30),one\_star numeric , two\_stars numeric,three\_stars numeric , four\_stars numeric)

order by job\_role;

