**HR Data Analytics Project**

By Juventus Chakma Paritosh

This is a data analytics project in respect of HR (Human Resources). In this project firstly SQL (Structured Query Language) has been used then the data has been presented with fine dashboard in Tableau. For SQL, PostgreSQL has been used.

A database has been created at first then a table was inserted into the database. Different SQL functions like SELECT, INSERT, COPY, WHERE, GROUP BY, ORDER BY, HAVING, CROSSTAB(), CAST(), ROUND(),SUBQUERIES, DATA TYPES etc. have been used. The main motive of this project is to find out the job satisfaction rating of employees in different department. It also tried to answer attrition rate in different department and by gender.

The queries of SQL are given below with few screenshots of the selective query outputs. At the end a screenshot of the Tableau Dashboard has been provided along with the link to Tableau Public where the Viz is present.

**Queries:**

-- at first we need to create a table with exact columns based on csv file --

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

)

-- let us check the table that has been created --

select \*

from hrdata; -- so far we can only see the structure with columns and no data in it –

A screenshot of a computer

Description automatically generated

/\* let us add data into the table. simply right click on the table and import the csv file.

after importing again run the query "select \* from hrdata" \*/

-- writing a query to get the sum of employee count --

select sum (employee\_count)

from hrdata; -- result shows as "1470" --

A screenshot of a computer

Description automatically generated

-- writing query to get sum of employee count only for education = high school --

select sum (employee\_count)

from hrdata

where education = 'High School'; -- result shows as "170" --

-- writing query to get sum of employee count only for department = sales --

select sum (employee\_count)

from hrdata

where department = 'Sales'; -- result shows as "446" --

-- writing query to get sum of employee count only for department = R&D --

select sum (employee\_count)

from hrdata

where department = 'R&D'; -- result shows as "961" --

-- writing query to get sum of employee count only for education\_field = Medical --

select sum (employee\_count) as employee\_count

from hrdata

where education\_field = 'Medical'; -- result shows as "464" --

-- writing a query to get the count of attrition from hrdata but attrition = Yes --

select count (attrition)

from hrdata

where attrition = 'Yes'; -- result shows as "237" --

/\* writing a query to get the count of attrition from hrdata but attrition = Yes and

education = 'Doctoral Degree' \*/

select count (attrition)

from hrdata

where attrition = 'Yes' and education = 'Doctoral Degree'; -- result shows as "5" --

/\* writing a query to get the count of attrition from hrdata but attrition = Yes and

department = 'R&D' \*/

select count (attrition)

from hrdata

where attrition = 'Yes' and department = 'R&D'; -- result shows as "133"

/\* writing a query to get the count of attrition from hrdata but attrition = Yes and

department = 'R&D' and education\_field = 'Medical'\*/

select count (attrition)

from hrdata

where attrition = 'Yes' and department = 'R&D' and education\_field = 'Medical'; -- result shows as "47"

/\* writing a query to get the count of attrition from hrdata but attrition = Yes and

department = 'R&D' and education\_field = 'Medical' and education = 'High School' \*/

select count (attrition)

from hrdata

where attrition = 'Yes' and department = 'R&D' and education\_field = 'Medical'

and education = 'High School'; -- result shows as "9" --

-- writing a query to get attrition rate --

select round (((select count (attrition) from hrdata where attrition = 'Yes') /

sum (employee\_count)) \* 100, 2) from hrdata; -- result shows as "16.12" --

A screenshot of a computer

Description automatically generated

-- writing a query to get attrition rate and from department sales --

select round (((select count (attrition) from hrdata where attrition = 'Yes' and department = 'Sales') /

sum (employee\_count)) \* 100, 2) from hrdata

where department = 'Sales'; -- result shows as "20.63" --

A screenshot of a computer

Description automatically generated

-- writing query to get active employees --

select sum (employee\_count) - (select count (attrition) from hrdata where attrition = 'Yes')

from hrdata; -- result shows as "1233" --

-- writing query to get active employees but gender = male --

select sum (employee\_count) - (select count (attrition) from hrdata where attrition = 'Yes' and gender = 'Male')

from hrdata where gender = 'Male'; -- result shows as "732" --

-- writing query to get average age --

select round (avg (age),0) as Aveg\_age from hrdata; -- result shows as "37" --

-- writing query to get attrition by gender --

select gender, count (attrition) from hrdata

where attrition = 'Yes'

group by gender; -- result shows as "female = 87 and male = 150" --

select gender, count (attrition) from hrdata

where attrition = 'Yes'

group by gender

order by count (attrition) desc;

A screenshot of a computer

Description automatically generated

-- writing query to get attrition by gender and education = high school --

select gender, count (attrition) from hrdata

where attrition = 'Yes' and education = 'High School'

group by gender

order by count (attrition) desc; -- result shows as "female = 11 and male = 20" --

-- writing query to get attrition by department --

select department, count (attrition) from hrdata

where attrition = 'Yes'

group by department; -- result shows as "HR = 12, Sales = 92 and R&D = 133" --

select department, count (attrition) from hrdata

where attrition = 'Yes'

group by department

order by count (attrition) desc;

-- writing query to get percentage of attriton by department –

select department, count (attrition),

round ((cast (count (attrition) as numeric) / -- here changing the data type to numeric with "cast" --

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

from hrdata

where attrition = 'Yes'

group by department

order by count (attrition) desc; -- result shows as "HR = 5.06%, Sales = 38.82% and R&D = 56.12%" --

A screenshot of a computer

Description automatically generated

-- writing query to get percentage of attriton by department and by gender --

select department, count (attrition),

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

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

from hrdata

where attrition = 'Yes' and gender = 'Female'

group by department

order by count (attrition) desc; -- result shows as "HR = 6.90%, Sales = 43.68% and R&D = 49.43%" --

A screenshot of a computer

Description automatically generated

-- writing query to get number of employee by age group --

select age, sum (employee\_count)

from hrdata

group by age

order by age;

-- writing query to get number of employee by age group and department = R&D --

select age, sum (employee\_count)

from hrdata

where department = 'R&D'

group by age

order by age;

-- writing query to get atrrition by education\_field --

select education\_field, count (attrition)

from hrdata

where attrition = 'Yes'

group by education\_field

order by count (attrition) desc;

-- writing query to get atrrition by education\_field and department = 'Sales' --

select education\_field, count (attrition)

from hrdata

where attrition = 'Yes' and department = 'Sales'

group by education\_field

order by count (attrition) desc;

-- writing query to get attrition rate by different age group --

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

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, gender;

-- if crosstab extension is not existing --

create extension if not exists tablefunc;

-- writng query to get rating for job satisfaction --

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 ct (job\_role varchar (50), one numeric, two numeric, three numeric, four numeric)

order by job\_role;

A screenshot of a computer

Description automatically generated

**Tableau Dashboard**

**Link to Tableau Public:**

<https://public.tableau.com/views/DataAnalyticsHR/AnalyticsHR?:language=en-US&:display_count=n&:origin=viz_share_link>A screenshot of a graph

Description automatically generated