**DATA ANALYST PORTFOLIO SQL PROJECT**

**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

)

**Import Data in Table Using Query**

COPY hrdata FROM 'D:\hrdata.csv' DELIMITER ',' CSV HEADER;

**Employee Count:**

select sum(employee\_count) as Employee\_Count from hrdata;

**Attrition Count:**

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

**Attrition Rate:**

select

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

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

from hrdata;

**Active Employee:**

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

*OR*

select (select sum(employee\_count) from hrdata) - count(attrition) as active\_employee from hrdata

where attrition='Yes';

**Average Age:**

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

**Attrition by Gender**

select gender, count(attrition) as attrition\_count from hrdata

where attrition='Yes'

group by gender

order by count(attrition) desc;

**Department wise Attrition:**

select department, count(attrition), round((cast (count(attrition) as numeric) /

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

where attrition='Yes'

group by department

order by count(attrition) desc;

**No of Employee by Age Group**

SELECT age, sum(employee\_count) AS employee\_count FROM hrdata

GROUP BY age

order by age;

**Education Field wise Attrition:**

select education\_field, count(attrition) as attrition\_count from hrdata

where attrition='Yes'

group by education\_field

order by count(attrition) desc;

**Attrition Rate by Gender for different Age Group**

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

round((cast(count(attrition) as numeric) / (select count(attrition) from hrdata where attrition = 'Yes')) \* 100,2) as pct

from hrdata

where attrition = 'Yes'

group by age\_band, gender

order by age\_band, gender desc;

**Job Satisfaction Rating**

-Run this query first to activate the cosstab() function in postgres

CREATE EXTENSION IF NOT EXISTS tablefunc;

-Then run this to get o/p-

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;