

# DATA ANALYST PORTFOLIO SQL PROJECT FOR BEGINNERS

## TESTING TABLEAU/ POWER BI REPORTS IN SQL

### 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 crosstab() 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;
```