PROJECT REPORT

SQL Analysis

QA Testing

SQL Analysis of the data is performed

Satyajeet Das Chowdhury
satyajeetdc@gmail.com

Objective

QA Testing using SQL

Create a HR Database in Postgres SQL and write SQL queries and create a test document to QA the HR ANALYTICS DASHBOARD developed in Microsoft Power BI.

- **Functional Validation** - Test each feature work as per the requirement. To verify all the filters and Action Filters on the report work as per the requirement.

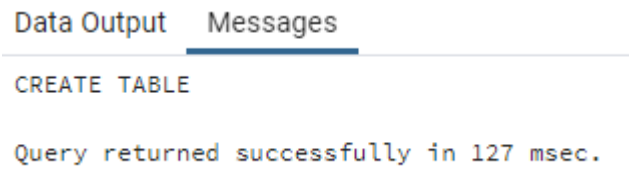
- **Data Validation** - Check accuracy and quality of data. To match the values in Tableau and Power BI report with SQL queries.

- **Test Document** - Create a Test document which will contain the screenshots and queries used to test the reports.

Process

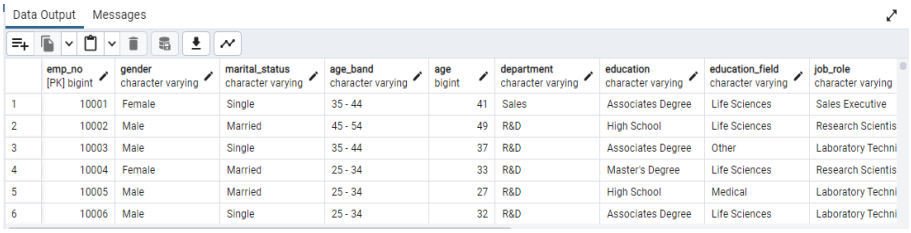
Creating the table schema:

```
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

SELECT * FROM hrdata

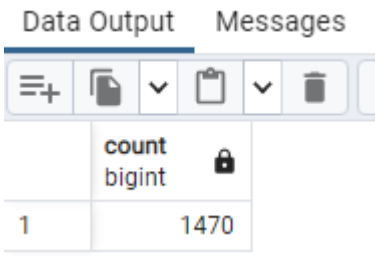


	emp_no [PK] bigint	gender character varying	marital_status character varying	age_band character varying	age bigint	department character varying	education character varying	education_field character varying	job_role character varying
1	10001	Female	Single	35 - 44	41	Sales	Associates Degree	Life Sciences	Sales Executive
2	10002	Male	Married	45 - 54	49	R&D	High School	Life Sciences	Research Scientis
3	10003	Male	Single	35 - 44	37	R&D	Associates Degree	Other	Laboratory Techni
4	10004	Female	Married	25 - 34	33	R&D	Master's Degree	Life Sciences	Research Scientis
5	10005	Male	Married	25 - 34	27	R&D	High School	Medical	Laboratory Techni
6	10006	Male	Single	25 - 34	32	R&D	Associates Degree	Life Sciences	Laboratory Techni

KPIs

Total Count of Employees

SELECT COUNT(emp_no) FROM hrdata

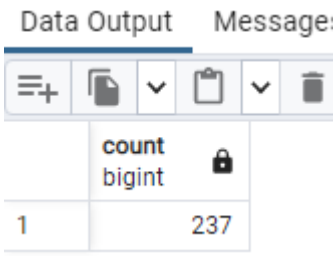


Data Output | Messages

	count bigint
1	1470

Attrition Count of employees

SELECT COUNT(attrition) FROM hrdata WHERE attrition = 'Yes'



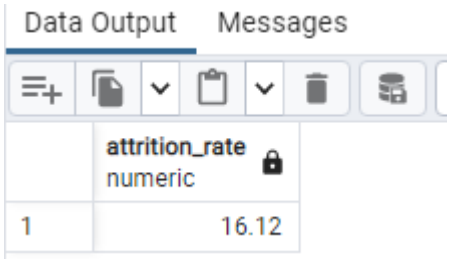
Data Output | Message

	count bigint
1	237

Attrition Rate of employees

SELECT ROUND((SELECT COUNT(attrition) FROM hrdata WHERE attrition = 'Yes')*100

/SUM(employee_count), 2) AS attrition_rate FROM hrdata

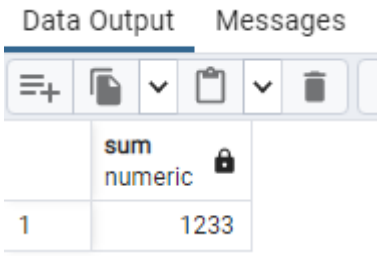


Data Output | Messages

	attrition_rate numeric
1	16.12

Active employees in the firm

SELECT SUM(active_employee) FROM hrdata

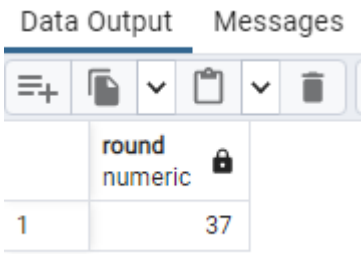


Data Output | Messages

	sum numeric
1	1233

Average age of the employees

SELECT ROUND(AVG(age),0) FROM hrdata



Data Output | Messages

	round numeric
1	37

Attrition by age

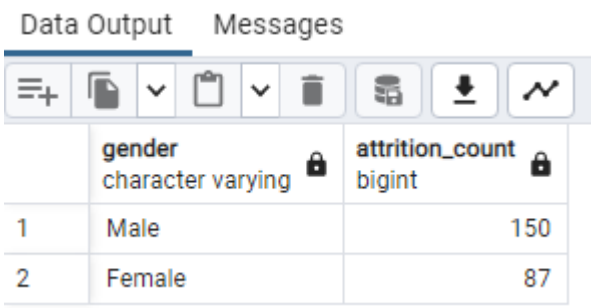
SELECT gender, COUNT(attrition) AS attrition_count

FROM hrdata

WHERE attrition = 'Yes'

GROUP BY gender

ORDER BY COUNT(attrition) DESC;



Data Output | Messages

	gender character varying	attrition_count bigint
1	Male	150
2	Female	87

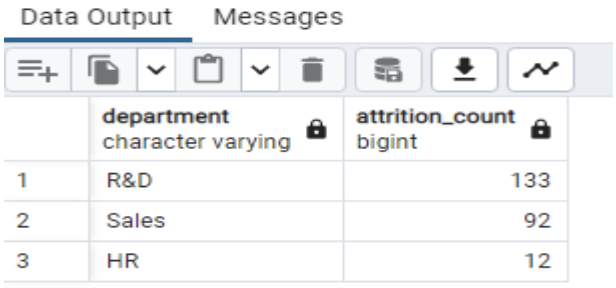
Other requirements

Attrition department wise

SELECT department, COUNT(attrition) AS attrition_count

FROM hrdata WHERE attrition = 'Yes'

GROUP BY department ORDER BY COUNT(attrition) DESC



Data Output | Messages

	department character varying	attrition_count bigint
1	R&D	133
2	Sales	92
3	HR	12

Attrition by education field

```
SELECT education_field, COUNT(attrition) AS attrition_count
FROM hrdata
WHERE attrition = 'Yes'
GROUP BY education_field
ORDER BY COUNT(attrition) DESC
```

Data Output Messages

	education_field character varying	attrition_count bigint
1	Life Sciences	89
2	Medical	63
3	Marketing	35
4	Technical Degree	32
5	Other	11
6	Human Resources	7

Attrition by age groups

```
SELECT age_band AS age_group, COUNT(attrition) AS attrition_count
FROM hrdata
WHERE attrition = 'Yes'
GROUP BY age_band
ORDER BY COUNT(attrition) DESC
```

Data Output Messages

	age_group character varying	attrition_count bigint
1	25 - 34	112
2	35 - 44	51
3	Under 25	38
4	45 - 54	25
5	Over 55	11

Count of employees by age group

```
SELECT age_band AS age_group, SUM(employee_count) AS employee_count
FROM hrdata
GROUP BY age_band
ORDER BY employee_count DESC
```

Data Output | Messages

	age_group character varying	employee_count numeric
1	25 - 34	554
2	35 - 44	505
3	45 - 54	245
4	Under 25	97
5	Over 55	69

Attrition rate by gender of different age groups

```
SELECT age_band AS age_group, gender, COUNT(attrition) AS 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 DESC;
```

Data Output | Messages

	age_group character varying	gender character varying	attrition bigint	percentage numeric
1	25 - 34	Male	69	29.11
2	25 - 34	Female	43	18.14
3	35 - 44	Male	37	15.61
4	35 - 44	Female	14	5.91
5	45 - 54	Male	16	6.75
6	45 - 54	Female	9	3.80
7	Over 55	Male	8	3.38
8	Over 55	Female	3	1.27
9	Under 25	Male	20	8.44
10	Under 25	Female	18	7.59

Attrition rate by gender of different departments

```
SELECT department, gender, COUNT(attrition) AS 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 department, gender
ORDER BY department, gender DESC;
```

Data Output | Messages

	department character varying	gender character varying	attrition bigint	percentage numeric
1	HR	Male	6	2.53
2	HR	Female	6	2.53
3	R&D	Male	90	37.97
4	R&D	Female	43	18.14
5	Sales	Male	54	22.78
6	Sales	Female	38	16.03

Job Satisfaction rating

```
-- activating the cosstab() function
CREATE EXTENSION IF NOT EXISTS tablefunc
```

Data Output | Messages

```
CREATE EXTENSION

Query returned successfully in 422 msec.
```

```
-- now the process

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
```

Data Output | Messages

	job_role character varying	one numeric	two numeric	three numeric	four numeric
1	Healthcare Representative	26	19	43	43
2	Human Resources	10	16	13	13
3	Laboratory Technician	56	48	75	80
4	Manager	21	21	27	33
5	Manufacturing Director	26	32	49	38
6	Research Director	15	16	27	22
7	Research Scientist	54	53	90	95
8	Sales Executive	69	54	91	112
9	Sales Representative	12	21	27	23

Result

All the details matched our Power BI analysis.

Total Tests : 13
Pass : 13
Fail : 0
Blocked : 0
Not Executed : 0