

# EMERALD TECHNOLOGIES

## EMPLOYEE PROMOTION DATA



# Employee Promotion Data

**Emerald Technologies** is a Nigerian multinational mobile telecommunications company, operating in many African and Asian countries. Its head office is in Lagos.

As of December 2022, **Emerald Technologies** recorded 1 Billion subscribers, making it the 3rd largest mobile network operator in the world, and the largest in Africa/Asia. They currently have issues around promotion and performance rating and need an experienced Data Scientist to help.

**You have been contacted as a Data Scientist Consultant to provide an in-depth analysis of what is happening in Emerald Technologies Using SQL**



# Exploratory Data Analysis

**Using the data answer the following questions:**

1. How many employees do we have in the organization and what is the maximum length of service?
2. How many employees are there in each department?
3. What is the proportion of male to female employees?
4. Group Employee age into 5 categories (20 – 29, 30 – 39, 40-49, 50-59, >60). What age group has the highest and lowest employee?
5. Who works in the Finance department?
6. Who has the highest average training score among all employees?
7. Which regions have the highest number of departures (employees who have left), and what are the corresponding departments?
8. Which department has the most employees, and which department has the fewest employees?
9. Who are the top 5 highest-earning employees in the 'Technology' department?
10. Who are the employees with awards in departments with more than 10 employees, and what are their department names?



# Tailored Data Analytics

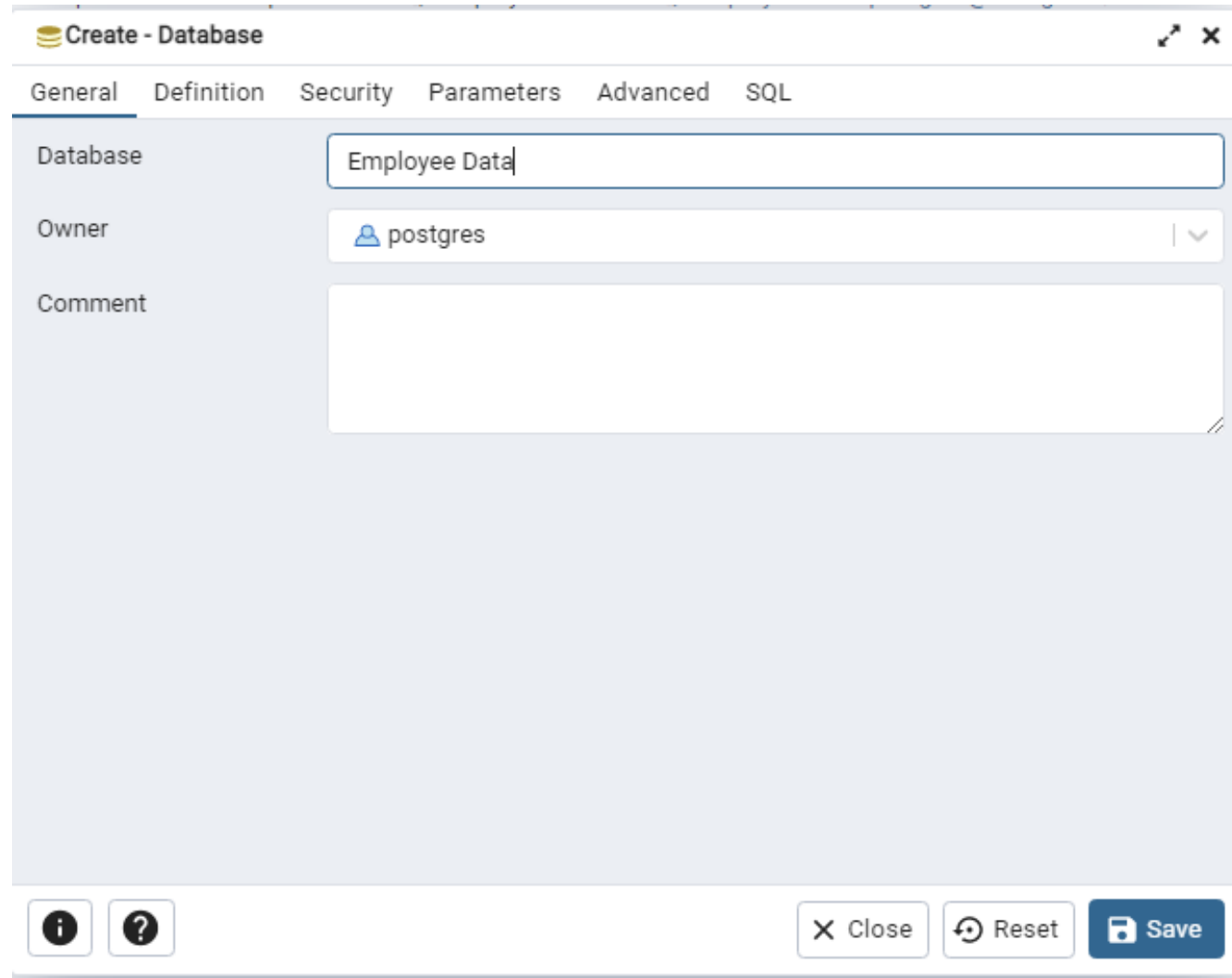
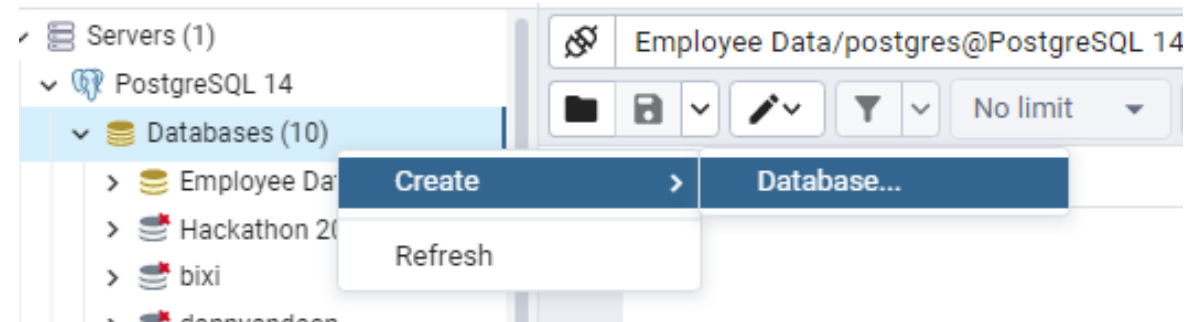


**Using the data answer the following questions:**

1. What is the average training score of employees in each department
2. What is the average previous year rating by department?
3. What is the average training score of employees by education type?
4. Group Average training score into grades (A,B,C,D,E,F) and what grade had the highest and lowest number of employees
5. Which three departments have the highest average job satisfaction among employees with a Bachelor's degree?
6. What is the average previous year rating by recruitment channel?
7. What is the split of gender by the previous year rating?
8. Based on the age group created what is the average previous year rating and average training score.
9. What is the average age of male and female employees, and how many employees are there for each gender?
10. Who are the top 5 highest-earning employees with a JobLevel of 3 or higher?

# How to Import the dataset (csv.)

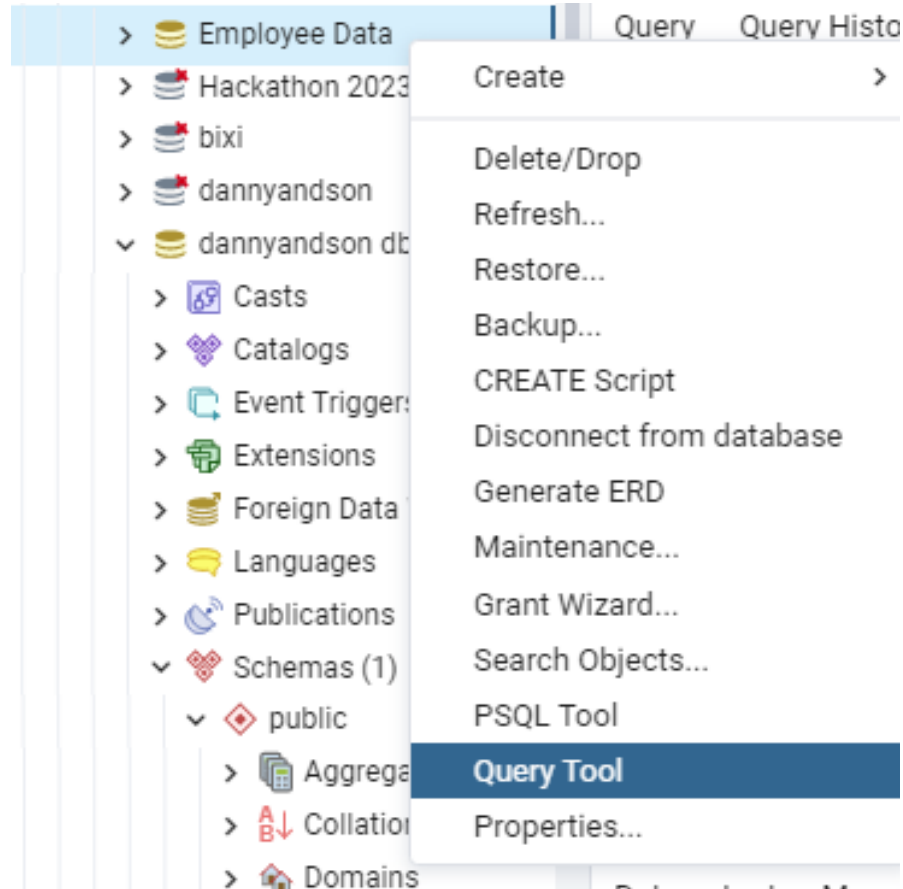
## Step 1: Create a database “Employee Data”





# How to Import the dataset (csv.)

**Step 2: Right-click on “Employee Data” and select Query Tool database**



# How to Import the dataset (csv.)

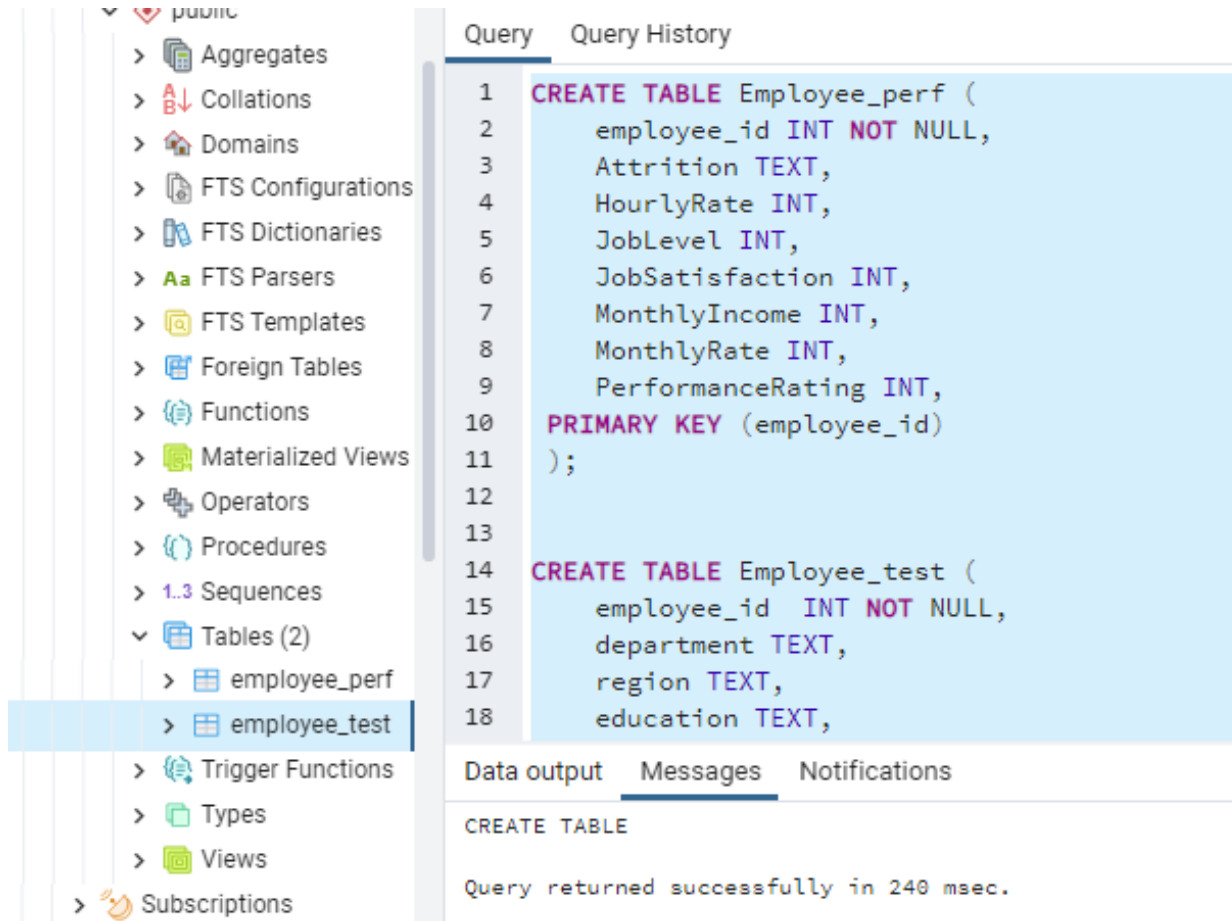
**Step 3: Paste the below query in the Query Editor and Execute.**

```
CREATE TABLE Employee_perf (  
  employee_id INT NOT NULL,  
  Attrition TEXT,  
  HourlyRate INT,  
  JobLevel INT,  
  JobSatisfaction INT,  
  MonthlyIncome INT,  
  MonthlyRate INT,  
  PerformanceRating INT,  
  PRIMARY KEY (employee_id)  
);
```

```
CREATE TABLE Employee_test (  
  employee_id INT NOT NULL,  
  department TEXT,  
  region TEXT,  
  education TEXT,  
  gender TEXT,  
  recruitment_channel TEXT,  
  no_of_trainings INT,  
  age INT,  
  previous_year_rating INT,  
  length_of_service INT,  
  awards_won INT,  
  avg_training_score INT,  
  PRIMARY KEY (employee_id)  
);
```

# How to Import the dataset (csv.)

You should get a successful query result and the table created under “Table” as shown in the screenshot below.



The screenshot displays a database management interface. On the left, a tree view shows the database structure under the 'public' schema. The 'Tables (2)' folder is expanded, showing 'employee\_perf' and 'employee\_test'. The 'employee\_test' table is selected. The main panel shows the SQL query used to create the table:

```
1 CREATE TABLE Employee_perf (  
2     employee_id INT NOT NULL,  
3     Attrition TEXT,  
4     HourlyRate INT,  
5     JobLevel INT,  
6     JobSatisfaction INT,  
7     MonthlyIncome INT,  
8     MonthlyRate INT,  
9     PerformanceRating INT,  
10    PRIMARY KEY (employee_id)  
11 );  
12  
13  
14 CREATE TABLE Employee_test (  
15     employee_id INT NOT NULL,  
16     department TEXT,  
17     region TEXT,  
18     education TEXT,
```

Below the query, the 'Data output' tab is active, showing the message: 'Query returned successfully in 240 msec.'



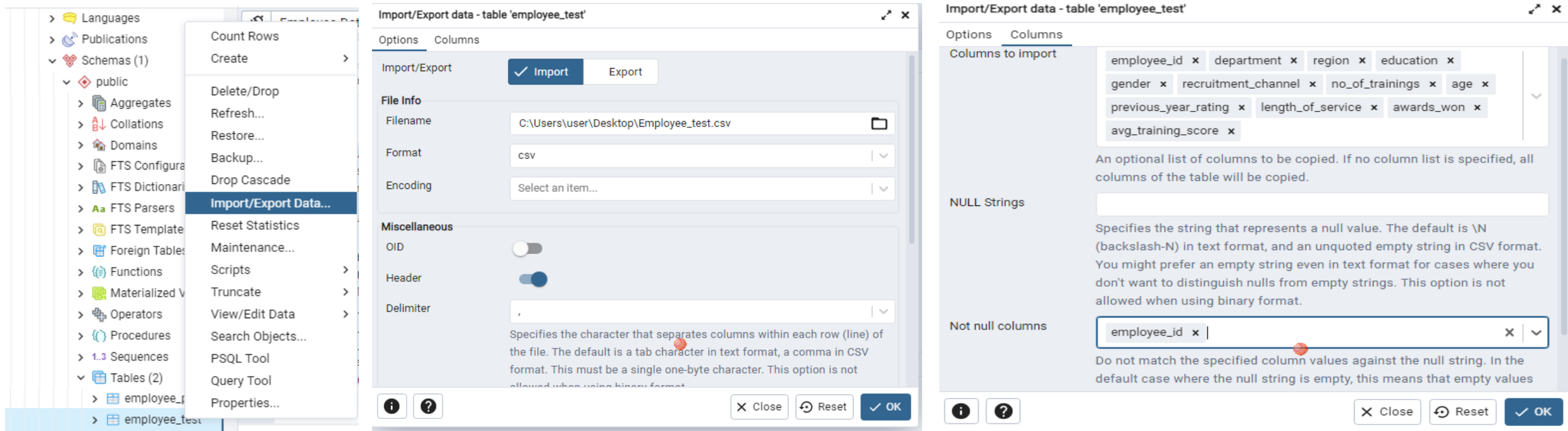
# How to Import the dataset (csv.)

## Step 4: Import the dataset employee\_performance.csv and employee\_test.csv:

### [Dataset](#)

- Right-click on the created table “employee\_test” and select “Import/Export Data”
- Select “import” and bring in the dataset from your computer files
- Enable “Header” and select comma as “Delimiter”
- Click on “Column” displayed at the top and select the columns to be imported in the “Columns to import” field
- Click OK or Save.

***Repeat steps for the second table***



The screenshot displays the 10Alytics interface with a right-click context menu open over the 'employee\_test' table. The menu includes options like 'Count Rows', 'Create', 'Delete/Drop', 'Refresh...', 'Restore...', 'Backup...', 'Drop Cascade', 'Import/Export Data...', 'Reset Statistics', 'Maintenance...', 'Scripts', 'Truncate', 'View/Edit Data', 'Search Objects...', 'PSQL Tool', 'Query Tool', and 'Properties...'. The 'Import/Export Data...' option is highlighted.

The 'Import/Export data - table 'employee\_test'' dialog is shown with the 'Options' tab selected. The 'Import/Export' section has 'Import' checked. The 'File Info' section shows the filename 'C:\Users\user\Desktop\Employee\_test.csv', format 'csv', and encoding 'Select an item...'. The 'Miscellaneous' section has 'OID' disabled, 'Header' enabled, and 'Delimiter' set to comma. A description for the delimiter states: 'Specifies the character that separates columns within each row (line) of the file. The default is a tab character in text format, a comma in CSV format. This must be a single one-byte character. This option is not allowed when using binary format.'

The 'Columns' tab is also shown, with 'Columns to import' listed as: employee\_id, department, region, education, gender, recruitment\_channel, no\_of\_trainings, age, previous\_year\_rating, length\_of\_service, awards\_won, and avg\_training\_score. The 'NULL Strings' section has a text input field. The 'Not null columns' section has a list box containing 'employee\_id'. A description for the not null columns states: 'Do not match the specified column values against the null string. In the default case where the null string is empty, this means that empty values'.