-- SQL Project: HR Analytics

-- Description: This project focuses on HR data analysis using SQL queries.

-- The queries cover employee count, attrition analysis, job satisfaction, and more.

-- ========================================

-- 1. Create the 'hrdata' Table

-- ========================================

CREATE TABLE hrdata (

emp\_no INT PRIMARY KEY, -- Unique employee number

gender VARCHAR(50) NOT NULL, -- Gender of the employee

marital\_status VARCHAR(50), -- Marital status of the employee

age\_band VARCHAR(50), -- Age group of the employee (e.g., 20-30, 30-40)

age INT, -- Actual age of the employee

department VARCHAR(50), -- Department the employee belongs to

education VARCHAR(50), -- Highest education level of the employee

education\_field VARCHAR(50), -- Field of education (e.g., Engineering, Business)

job\_role VARCHAR(50), -- Job role (e.g., Manager, Developer)

business\_travel VARCHAR(50), -- Business travel frequency

employee\_count INT, -- Employee count (useful for aggregations)

attrition VARCHAR(50), -- Attrition status (Yes/No)

attrition\_label VARCHAR(50), -- Attrition label (e.g., voluntary/involuntary)

job\_satisfaction INT, -- Job satisfaction rating (1-4)

active\_employee INT -- Indicates if the employee is currently active

);

-- ========================================

-- 2. Import Data into the 'hrdata' Table

-- ========================================

-- This query imports data from a CSV file into the 'hrdata' table.

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

-- ========================================

-- 3. Indexing for Performance Improvement

-- ========================================

-- Creating indexes on columns frequently used in WHERE clauses for faster querying.

CREATE INDEX idx\_attrition ON hrdata(attrition);

CREATE INDEX idx\_gender ON hrdata(gender);

CREATE INDEX idx\_department ON hrdata(department);

-- ========================================

-- 4. Basic Queries for HR Analysis

-- ========================================

-- 4.1. Employee Count

-- Retrieves the total number of employees.

SELECT SUM(employee\_count) AS Employee\_Count FROM hrdata;

-- 4.2. Attrition Count

-- Retrieves the total number of employees who have left (attrition = 'Yes').

SELECT COUNT(attrition) AS Attrition\_Count FROM hrdata WHERE attrition = 'Yes';

-- 4.3. Attrition Rate

-- Calculates the percentage of employees who have left.

SELECT ROUND(((SELECT COUNT(attrition) FROM hrdata WHERE attrition = 'Yes') / SUM(employee\_count)) \* 100, 2) AS Attrition\_Rate FROM hrdata;

-- 4.4. Active Employee Count

-- Retrieves the count of employees who are currently active (i.e., not part of attrition).

SELECT SUM(employee\_count) - (SELECT COUNT(attrition) FROM hrdata WHERE attrition = 'Yes') AS Active\_Employee FROM hrdata;

-- 4.5. Average Age of Employees

-- Retrieves the average age of employees.

SELECT ROUND(AVG(age), 0) AS Average\_Age FROM hrdata;

-- ========================================

-- 5. Detailed HR Analysis Queries

-- ========================================

-- 5.1. Attrition by Gender

-- Retrieves the number of employees who have left, grouped by gender.

SELECT gender, COUNT(attrition) AS Attrition\_Count FROM hrdata

WHERE attrition = 'Yes'

GROUP BY gender

ORDER BY COUNT(attrition) DESC;

-- 5.2. Department-wise Attrition

-- Retrieves attrition count and percentage, grouped by department.

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;

-- 5.3. Number of Employees by Age Group

-- Retrieves the number of employees, grouped by age.

SELECT age, SUM(employee\_count) AS Employee\_Count FROM hrdata

GROUP BY age

ORDER BY age;

-- 5.4. Education Field-wise Attrition

-- Retrieves the attrition count, grouped 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;

-- 5.5. Attrition Rate by Gender for Different Age Groups

-- Retrieves the attrition count and percentage, grouped by gender and age group.

SELECT age\_band, gender, COUNT(attrition) AS Attrition\_Count,

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;

-- ========================================

-- 6. Advanced Query: Job Satisfaction Rating

-- ========================================

-- Note: This query is specific to PostgreSQL.

-- 6.1. Enable Crosstab Function in PostgreSQL

-- Activate the crosstab() function for generating pivot-like tables.

CREATE EXTENSION IF NOT EXISTS tablefunc;

-- 6.2. Retrieve Job Satisfaction Ratings by Job Role

-- This query creates a crosstab to show job satisfaction levels across different job roles.

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;