📊 HR Analysis SQL Query Documentation

🗃️ Database Setup

-- Create the HR_Analysis database

- CREATE DATABASE HR_Analysis;

-- View all records in the hr_analysis table

- SELECT * FROM hr_analysis;

-- Describe the structure of the hr_analysis table

- DESCRIBE hr_analysis;

🔧 Column Normalization

-- Convert startDate from string to proper DATE format

- UPDATE hr_analysis
  SET startDate = STR_TO_DATE(startDate, "%m/%d/%Y");

-- Modify StartDate column to DATE type

- ALTER TABLE hr_analysis
  MODIFY COLUMN StartDate DATE;

-- Rename and standardize column names for clarity and consistency

- ALTER TABLE hr_analysis

  CHANGE `Job Role` job_role VARCHAR(20),

  CHANGE `Employee ID` employee_id INT,

  CHANGE `Performance Score` performance_score VARCHAR(50),

  CHANGE `Engagement Score` engagement_score INT,

  CHANGE EmployeeClassificationType employee_type VARCHAR(50),

  CHANGE `current employee rating` current_employee_rating INT,

  CHANGE `satisfaction score` satisfaction_score INT,

  CHANGE `work-life balance score` work_life_balance_score INT;

- ALTER TABLE hr_analysis

  CHANGE `Training Program Name` training_program VARCHAR(50),

  CHANGE `Training Type` training_type VARCHAR(50),

  CHANGE `Training Outcome` training_outcome VARCHAR(50),

  CHANGE `Training Months` training_months VARCHAR(50),

  CHANGE `Training Cost` training_cost VARCHAR(50);

Employee Demographics

-- Total number of employees

- SELECT COUNT(*) FROM hr_analysis;

-- Gender-wise employee count

- SELECT COUNT(employee_id), GenderCode FROM hr_analysis
  GROUP BY GenderCode;

-- Marital status-wise employee count

- SELECT MaritalDesc, COUNT(employee_id) FROM hr_analysis
  GROUP BY MaritalDesc;

Job Role Distribution

-- Employee count per job role

- SELECT job_role, COUNT(employee_id) AS no_of_employees FROM hr_analysis
  GROUP BY job_role
  ORDER BY no_of_employees;

-- Top 5 job roles with highest number of employees

- SELECT job_role, COUNT(employee_id) AS no_of_employees FROM hr_analysis
  GROUP BY job_role
  ORDER BY no_of_employees DESC
  LIMIT 5;

Employment Status

-- Number of active employees

- SELECT COUNT(employeestatus) AS Employees_status FROM hr_analysis
  WHERE EmployeeStatus = "active";

-- Number of terminated employees

- SELECT COUNT(employeestatus) AS Employees_status FROM hr_analysis
  WHERE EmployeeStatus = "terminated";

Employee Type Breakdown

-- Total employees by type (temporary, part-time, full-time)

- SELECT COUNT(employee_id) AS total_employees, employee_type FROM hr_analysis
  GROUP BY employee_type;

-- Job role vs employee type distribution

- SELECT job_role, employee_type, COUNT(employee_id) AS total_employees
  FROM hr_analysis
  GROUP BY job_role, employee_type
  ORDER BY job_role, total_employees DESC;

📌 Status by Job Role

-- Active employees by job role (Top 10)

- SELECT job_role, COUNT(employeestatus) AS Employees_Active FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY job_role ORDER BY Employees_Active DESC LIMIT 10;

-- Terminated employees by job role

- SELECT job_role, COUNT(employeestatus) AS Employees_Terminated FROM hr_analysis
  WHERE EmployeeStatus = "terminated"
  GROUP BY job_role
  ORDER BY Employees_Terminated DESC;

🪪 Payzone Analysis

-- List of payzones

- SELECT payzone FROM hr_analysis
  GROUP BY payzone;

-- Total employees per payzone

- SELECT COUNT(employee_id) AS total_employees, payzone FROM hr_analysis
  GROUP BY payzone
  ORDER BY total_employees DESC;

-- Payzone A,B,C: breakdown by job role, division, department, and employee type

- SELECT job_role, division, DepartmentType, employee_type, COUNT(employee_id)
  AS total_employees, payzone FROM hr_analysis
  WHERE PayZone = "zone a"
  GROUP BY PayZone, job_role, division, DepartmentType, employee_type
  ORDER BY total_employees DESC;

-- Payzone distribution by employee type

- SELECT employee_type, payzone, COUNT(employee_id) AS total_employees FROM hr_analysis
  GROUP BY employee_type, PayZone
  ORDER BY PayZone, total_employees DESC;

-- Payzone A: distribution by marital status

- SELECT maritaldesc, payzone, COUNT(employee_id) AS total_employees FROM hr_analysis WHERE payzone = "zone a"
  GROUP BY maritaldesc, payzone;

-- Payzone distribution by state

- SELECT payzone, state, COUNT(employee_id) AS total_employees FROM hr_analysis
  GROUP BY state, payzone
  ORDER BY payzone, total_employees DESC;

-- Payzone distribution by age

- SELECT payzone, age, COUNT(employee_id) AS total_employees FROM hr_analysis
  GROUP BY payzone, age
  ORDER BY payzone, total_employees DESC;

☑ Performance Score Analysis

-- Total employees by performance score

- SELECT performance_score, COUNT(employee_id) AS total_employees FROM hr_analysis
  GROUP BY performance_score;

-- Performance score vs employee status

- SELECT COUNT(employee_id) AS total_employees, Performance_Score, EmployeeStatus FROM hr_analysis
  WHERE performance_score = "fully meets"
  GROUP BY Performance_Score, EmployeeStatus;

-- Active employees with "fully meets" score by job role, department, division

- SELECT COUNT(employee_id) AS total_employees, job_role, DepartmentType, Division, performance_score, EmployeeStatus FROM hr_analysis
  WHERE performance_score = "fully meets" AND EmployeeStatus = "active"
  GROUP BY job_role, DepartmentType, Division, performance_score, EmployeeStatus
  ORDER BY total_employees DESC;

-- Active employees with "fully meets" score by state

- SELECT COUNT(employee_id) AS total_employees, Performance_Score, EmployeeStatus, state FROM hr_analysis
  WHERE performance_score = "fully meets" AND EmployeeStatus = "active" GROUP BY Performance_Score, EmployeeStatus, state ORDER BY total_employees DESC;

📊 Engagement & Rating

-- Average engagement score

- SELECT AVG(engagement_score) AS average_engagement_score FROM hr_analysis;

-- Active employees by engagement score

- SELECT COUNT(employee_id) AS total_employees, engagement_score FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY engagement_score;

-- Average employee rating

- SELECT AVG(current_employee_rating) FROM hr_analysis;

-- Average rating by job role

- SELECT AVG(current_employee_rating) AS Average_rating, job_role FROM hr_analysis
  GROUP BY job_role
  ORDER BY Average_rating DESC;

-- Active employees by rating

- SELECT current_employee_rating, COUNT(employee_id) AS total_employees FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY current_employee_rating;

-- Active employees by rating and state

- SELECT COUNT(employee_id) AS total_employees, current_employee_rating, State FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY state, current_employee_rating
  ORDER BY current_employee_rating DESC;

-- Active male employees by age and rating

- SELECT COUNT(employee_id) AS total_employees, current_employee_rating, age, gendercode FROM hr_analysis
  WHERE EmployeeStatus = "active" AND GenderCode = "male"
  GROUP BY age, GenderCode, current_employee_rating
  ORDER BY current_employee_rating DESC;

😊 Satisfaction Score

-- Average satisfaction score

- SELECT AVG(satisfaction_score) FROM hr_analysis;

-- Average satisfaction score by job role

- SELECT AVG(satisfaction_score) AS Average_Satisfaction_score, job_role FROM hr_analysis
  GROUP BY job_role;

-- Total employees by satisfaction score

- SELECT COUNT(employee_id) AS total_employees, satisfaction_score FROM hr_analysis
  GROUP BY satisfaction_score;

-- Active employees by satisfaction score

- SELECT COUNT(employee_id) AS total_employees, satisfaction_score FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY satisfaction_score;

-- Active employees by satisfaction score and state

- SELECT COUNT(employee_id) AS total_employees, satisfaction_score, state FROM hr_analysis
  WHERE EmployeeStatus = "active"
  GROUP BY state, Satisfaction_Score
  ORDER BY total_employees DESC;

🏆 Employee Total Rank

-- Rank employees based on combined score of performance, engagement, satisfaction, and work-life balance

- WITH scores AS (SELECT employee_id, EmployeeStatus, (performance_score + engagement_score + satisfaction_score + work_life_balance_score) AS total_score FROM hr_analysis)
  SELECT employee_id, EmployeeStatus, total_score, RANK() OVER (ORDER BY total_score DESC) AS Rank_Position FROM scores ORDER BY total_score DESC;

🎓 Training Program Analysis

-- Training programs by job role

- SELECT job_role, training_program FROM hr_analysis
  ORDER BY job_role;

-- Total employees by training type

- SELECT COUNT(employee_id), training_type FROM hr_analysis
  GROUP BY training_type;

-- Total employees by training outcome

- SELECT training_outcome, COUNT(employee_id) AS total_employees FROM hr_analysis GROUP BY training_outcome
  ORDER BYtraining_type;

-- Total employees by training outcome

- SELECT training_outcome, COUNT(employee_id) AS total_employees
  FROM hr_analysis
  GROUP BY training_outcome
  ORDER BY total_employees DESC;

-- Average training duration in months

- SELECT AVG(training_months) AS average_training_months
  FROM hr_analysis;

-- Total training cost

- SELECT SUM(CAST(REPLACE(training_cost, '$', ' ') AS DEC