



SQL EMPLOYEES MANAGEMENT Project

WWW.TNLLIN.COM



BY DIVYANSH PATEL

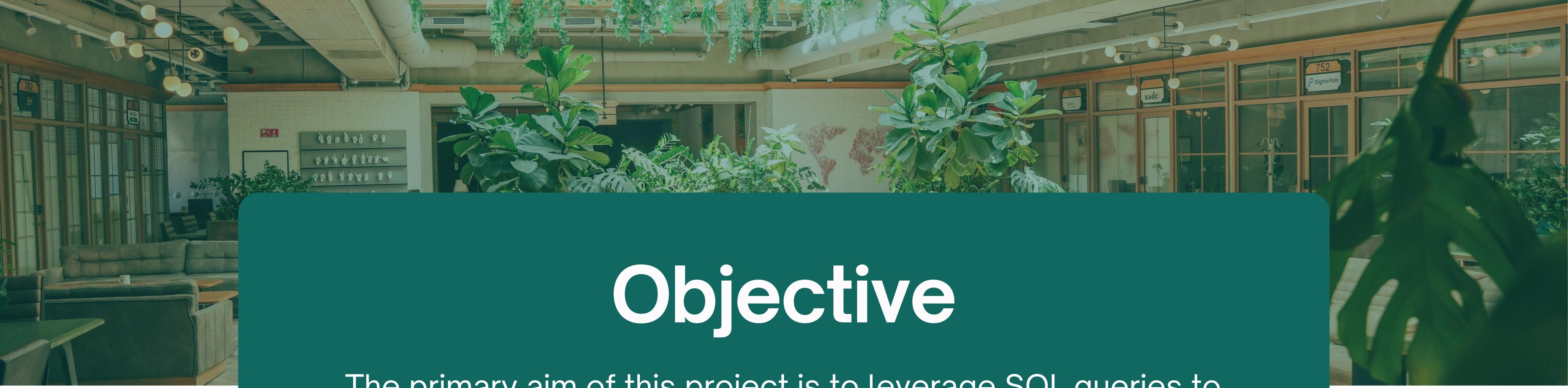
Overview

Project Title:

SQL Analysis of Employee Data

This SQL Analysis of Employee Data project aims to harness the power of SQL to explore and analyze various dimensions of an organization's employee data. By systematically querying the database, we seek to uncover actionable insights that can inform strategic decisions and drive organizational improvements. Here's a detailed overview of the project





Objective

The primary aim of this project is to leverage SQL queries to delve into the employee database and extract meaningful insights. By analyzing various aspects such as employee demographics, salary distribution, tenure, and performance metrics



About Me

My name is Divyansh Patel, and I am a passionate data analyst with a keen interest in leveraging data to derive actionable insights.

Background

High School:- ST JOSEPH SCHOOL PIPARIYA
COMMERCE WITH 92.2%

Education: BBA bia 3rd sem in
[Business Intelligence and Analytics]
from LNCT University BHOPAL .



*Skills: SQL, Data Analysis, Power BI ,
Python , Visualization, Database
Management*

Methodology

We will employ a variety of SQL queries

01

Aggregate Functions

To calculate averages, counts, and sums

02

JOIN Operations

To combine data from multiple tables

03

Subqueries

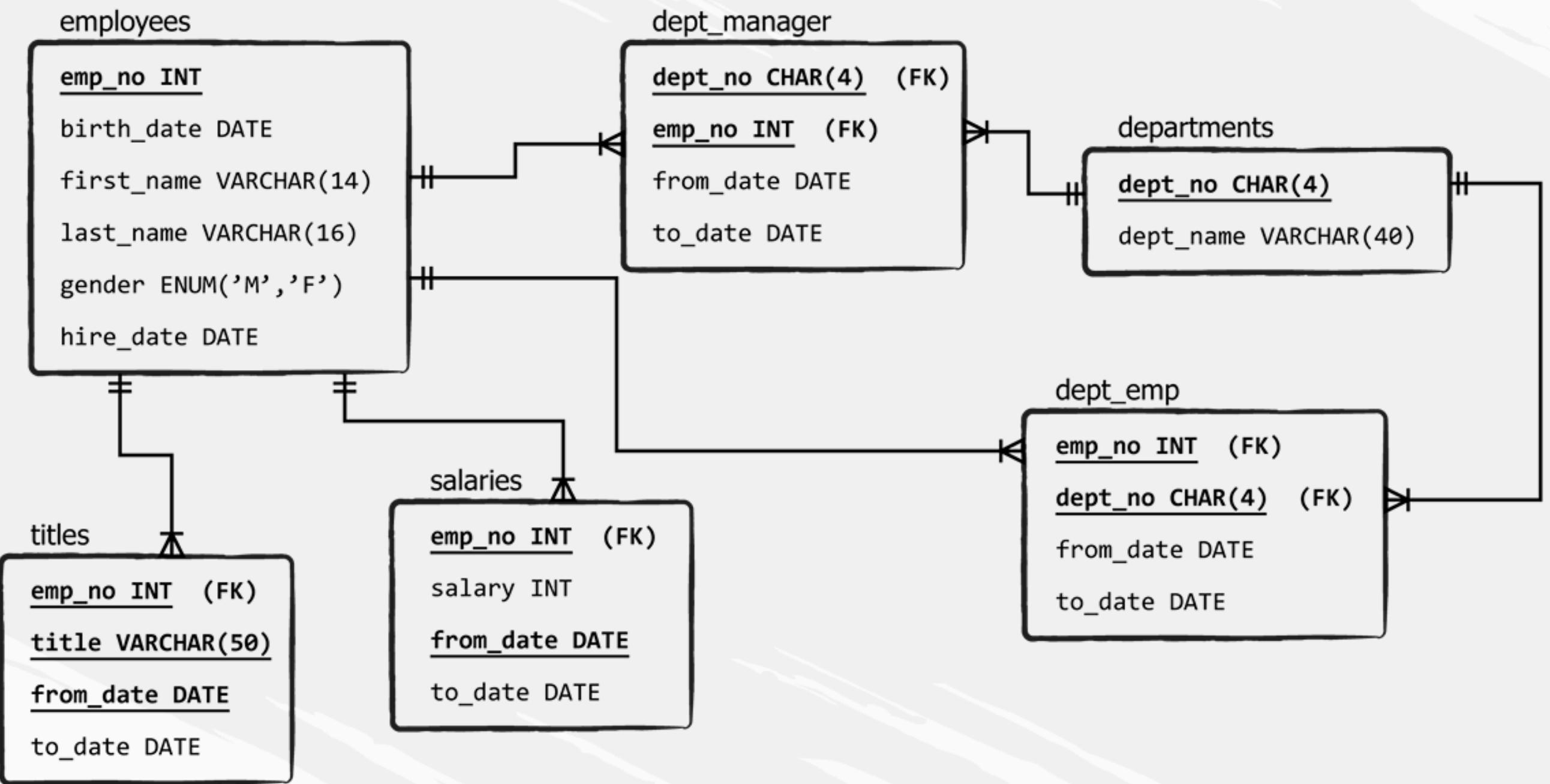
For complex data extraction and calculation.

04

Case Statements

To categorize and segment data effectively

database



ARE YOU READY

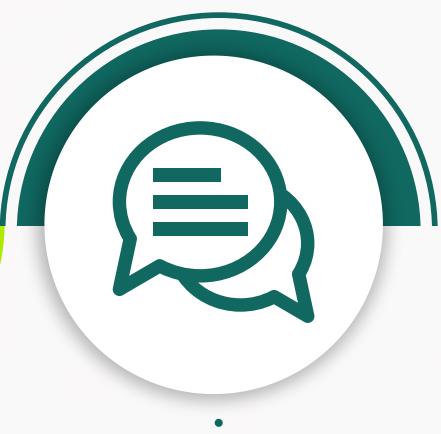
Basic
Queries



Intermediate
Queries



Advanced
Queries



Question: Write an SQL query to find the total number of distinct employees and the average salary for each gender. Group by gender

SELECT

```
e.gender, COUNT(DISTINCT e.emp_no) as total_employees ,  
ROUND(AVG(s.salary), 2) as avg_salary
```

FROM

```
employees e  
JOIN  
salaries s ON s.emp_no = e.emp_no
```

GROUP BY e.gender;

| | gender | total_employees | avg_salary |
|---|--------|-----------------|------------|
| ▶ | M | 60964 | 63755.91 |
| | F | 40832 | 63769.12 |

count total employees in each department

SELECT

```
d.dept_name, COUNT(e.emp_no) AS total_emp
```

FROM

```
dept_emp e
```

JOIN

```
departments d ON d.dept_no = e.dept_no
```

GROUP BY d.dept_name

ORDER BY total_emp DESC;

| | dept_name | total_emp |
|---|--------------------|-----------|
| ▶ | Development | 85707 |
| | Production | 73485 |
| | Sales | 52245 |
| | Customer Service | 23580 |
| | Research | 21126 |
| | Marketing | 20211 |
| | Quality Management | 20117 |
| | Human Resources | 17786 |
| | Finance | 17346 |

How can we determine which department has the lowest total salary expenditure, and what is the total salary expenditure for each department?

SELECT

```
d.dept_name, SUM(s.salary) AS total_expenditure
```

FROM

```
dept_emp e
```

JOIN

```
departments d ON d.dept_no = e.dept_no
```

JOIN

```
salaries s ON s.emp_no = e.emp_no
```

GROUP BY d.dept_name

ORDER BY total_expenditure DESC

;

| | dept_name | total_expenditure |
|---|--------------------|-------------------|
| ▶ | Development | 16466602819 |
| | Production | 14171898953 |
| | Sales | 13570391450 |
| | Marketing | 4624646965 |
| | Customer Service | 4473462983 |
| | Research | 4080782490 |
| | Finance | 3924369743 |
| | Quality Management | 3665941336 |
| | Human Resources | 3194340951 |

Intermediate Queries



Question: Write an SQL query to find the department number, average salary paid, and total number of total contract by each manager. Group by employee number, first name, and department number, and order by average salary paid in descending order

SELECT

```
m.dept_no,  
AVG(s.salary) AS avg_salary_paid,  
COUNT(e.emp_no) AS total_contract  
FROM  
dept_manager m  
JOIN  
employees e ON e.emp_no = m.emp_no  
JOIN  
salaries s ON s.emp_no = e.emp_no  
GROUP BY e.emp_no , e.first_name , m.dept_no  
ORDER BY avg_salary_paid DESC;
```

| | dept_no | avg_salary_paid | total_contract |
|---|---------|-----------------|----------------|
| ▶ | d001 | 89128.2778 | 18 |
| | d001 | 87570.5882 | 17 |
| | d007 | 87422.1250 | 16 |
| | d008 | 87217.2778 | 18 |
| | d007 | 84242.4444 | 18 |
| | d006 | 81301.2222 | 18 |
| | d002 | 72822.7778 | 18 |
| | d002 | 68809.0000 | 18 |
| | d006 | 67995.5000 | 16 |
| | d008 | 65916.6667 | 15 |
| | d004 | 63609.7059 | 17 |
| | d003 | 62990.2222 | 18 |
| | d009 | 62860.3636 | 11 |
| | d005 | 62568.1667 | 18 |

How can we calculate the average number of employees managed by each manager in different departments, and which department has the most efficient management ratio?

SELECT

```
d.dept_name,  
d.dept_no,  
COUNT(DISTINCT de.emp_no) AS total_emp,  
COUNT(DISTINCT m.emp_no) AS total_manager,  
ROUND((COUNT(DISTINCT de.emp_no) / COUNT(DISTINCT m.emp_no)),  
0) AS manager_per_emp
```

FROM

```
dept_emp de  
JOIN  
departments d ON d.dept_no = de.dept_no  
JOIN  
dept_manager m ON m.dept_no = de.dept_no  
GROUP BY d.dept_name , d.dept_no  
ORDER BY total_emp DESC;
```

| | dept_name | dept_no | total_emp | total_manager | manager_per_emp |
|---|--------------------|---------|-----------|---------------|-----------------|
| ▶ | Development | d005 | 85707 | 2 | 42854 |
| | Production | d004 | 73485 | 4 | 18371 |
| | Sales | d007 | 52245 | 2 | 26123 |
| | Customer Service | d009 | 23580 | 4 | 5895 |
| | Research | d008 | 21126 | 2 | 10563 |
| | Marketing | d001 | 20211 | 2 | 10106 |
| | Quality Management | d006 | 20117 | 4 | 5029 |
| | Human Resources | d003 | 17786 | 2 | 8893 |
| | Finance | d002 | 17346 | 2 | 8673 |

Question: Write an SQL query to create two temporary tables, `female_total` and `male_total`, which store the total number of female and male employees by department, respectively. Then, join these tables to list the department name, total number of male employees, and total number of female employees.

```
create temporary table female_total
select d.dept_name, ep.gender ,count(e.emp_no) as total_emp from dept_emp e
join
departments d on d.dept_no = e.dept_no  join employees ep on ep.emp_no = e.emp_no
where ep.gender ='F'
group by d.dept_name,ep.gender
order by total_emp desc ;

create temporary table male_total
select d.dept_name, ep.gender ,count(e.emp_no) as total_emp from dept_emp e
join
departments d on d.dept_no = e.dept_no  join employees ep on ep.emp_no = e.emp_no
where ep.gender ='M'
group by d.dept_name,ep.gender
order by total_emp desc;

SELECT
    male_total.dept_name,
    male_total.gender,
    male_total.total_emp,
    female_total.gender,
    female_total.total_emp
FROM
    male_total
    JOIN
    female_total ON male_total.dept_name = female_total.dept_name;
```

| | dept_name | gender | total_emp | gender | total_emp |
|---|--------------------|--------|-----------|--------|-----------|
| ▶ | Development | M | 51449 | F | 34258 |
| | Production | M | 43936 | F | 29549 |
| | Sales | M | 31391 | F | 20854 |
| | Customer Service | M | 14132 | F | 9448 |
| | Research | M | 12687 | F | 8439 |
| | Quality Management | M | 12039 | F | 8078 |
| | Marketing | M | 12174 | F | 8037 |
| | Human Resources | M | 10711 | F | 7075 |
| | Finance | M | 10331 | F | 7015 |

How can we determine which department has the highest average salary, and what is the count of employees and average salary for each department?

• SELECT

```
d.dept_name,  
COUNT(e.emp_no) AS total_emp,  
CONCAT(ROUND(AVG(s.salary), 2), '$') AS avg_salary_dept  
  
FROM  
dept_emp e  
JOIN  
departments d ON d.dept_no = e.dept_no  
JOIN  
salaries s ON s.emp_no = e.emp_no  
GROUP BY d.dept_name  
ORDER BY avg_salary_dept DESC;
```

| | dept_name | total_emp | avg_salary_dept |
|---|--------------------|-----------|-----------------|
| ▶ | Sales | 167999 | 80776.62\$ |
| | Marketing | 64319 | 71901.72\$ |
| | Finance | 55935 | 70159.47\$ |
| | Research | 68165 | 59866.24\$ |
| | Production | 238024 | 59539.79\$ |
| | Development | 276733 | 59503.58\$ |
| | Customer Service | 76137 | 58755.44\$ |
| | Quality Management | 63984 | 57294.66\$ |
| | Human Resources | 57708 | 55353.52\$ |

Question: Write an SQL query to find the department name, total employees, total salary expenditure (formatted with a dollar sign), average salary (formatted with a dollar sign), and department's share in total salary expenditure (as a percentage). Group by department name and order by total expenditure in descending order

SELECT

```
d.dept_name,  
COUNT(e.emp_no) AS total_emp,  
CONCAT(SUM(s.salary), '$') AS total_expenditure,  
CONCAT(ROUND(AVG(s.salary), 2), '$') AS avg_salary_dept,  
CONCAT(ROUND(SUM(s.salary)) / (SELECT  
    SUM(salary) AS total_exp_on_salary  
  FROM  
    salaries) * 100,  
        2),  
        '%' AS share_in_salary_exp  
FROM  
dept_emp e  
  JOIN  
departments d ON d.dept_no = e.dept_no  
  JOIN  
salaries s ON s.emp_no = e.emp_no  
GROUP BY d.dept_name  
ORDER BY total_expenditure DESC;
```

| dept_name | total_emp | total_expenditure | avg_salary_dept | share_in_salary_exp |
|--------------------|-----------|-------------------|-----------------|---------------------|
| Marketing | 64319 | 4624646965 \$ | 71901.72\$ | 7.50% |
| Customer Service | 76137 | 4473462983 \$ | 58755.44\$ | 7.25% |
| Research | 68165 | 4080782490 \$ | 59866.24\$ | 6.62% |
| Finance | 55935 | 3924369743 \$ | 70159.47\$ | 6.36% |
| Quality Management | 63984 | 3665941336 \$ | 57294.66\$ | 5.94% |
| Human Resources | 57708 | 3194340951 \$ | 55353.52\$ | 5.18% |
| Development | 276733 | 16466602819 \$ | 59503.58\$ | 26.70% |
| Production | 238024 | 14171898953 \$ | 59539.79\$ | 22.98% |
| Sales | 167999 | 13570391450 \$ | 80776.62\$ | 22.00% |

Advanced Queries



How can we determine the average salary paid and the number of salary records for each department manager, and which manager has the highest average salary?

SELECT

```
e.emp_no,  
e.first_name,  
m.dept_no,  
AVG(s.salary) AS avg_salary_paid,  
COUNT(e.emp_no) AS total_cont
```

FROM

```
dept_manager m  
JOIN  
employees e ON e.emp_no = m.emp_no  
JOIN  
salaries s ON s.emp_no = e.emp_no
```

GROUP BY e.emp_no , e.first_name , m.dept_no

ORDER BY avg_salary_paid DESC;

| | emp_no | first_name | dept_no | avg_salary_paid | total_cont |
|---|--------|-------------|---------|-----------------|------------|
| ▶ | 110022 | Margareta | d001 | 89128.28 | 18 |
| | 110039 | Vishwani | d001 | 87570.59 | 17 |
| | 111133 | Hauke | d007 | 87422.13 | 16 |
| | 111400 | Arie | d008 | 87217.28 | 18 |
| | 111035 | Przemyslawa | d007 | 84242.44 | 18 |
| | 110725 | Peternela | d006 | 81301.22 | 18 |
| | 110085 | Ebru | d002 | 72822.78 | 18 |
| | 110114 | Isamu | d002 | 68809.00 | 18 |
| | 110800 | Sanjoy | d006 | 67995.50 | 16 |
| | 111534 | Hilary | d008 | 65916.67 | 15 |
| | 110344 | Rosine | d004 | 63609.71 | 17 |
| | 110422 | Fiona | d005 | 62520.50 | 15 |

How can we determine the average salary and the count of distinct employees for each department, and which department has the highest average salary?

SELECT

```
d.dept_name,  
m.dept_no,  
AVG(s.salary) AS avg_salary_paid,  
COUNT(DISTINCT e.emp_no) AS total_cont  
  
FROM  
dept_manager m  
JOIN  
employees e ON e.emp_no = m.emp_no  
JOIN  
salaries s ON s.emp_no = e.emp_no  
JOIN  
departments d ON d.dept_no = m.dept_no  
GROUP BY d.dept_name , m.dept_no  
ORDER BY avg_salary_paid DESC;
```

| | dept_name | dept_no | avg_salary_paid | total_cont |
|---|--------------------|---------|-----------------|------------|
| ▶ | Marketing | d001 | 88371.6857 | 2 |
| | Sales | d007 | 85738.7647 | 2 |
| | Research | d008 | 77535.1818 | 2 |
| | Finance | d002 | 70815.8889 | 2 |
| | Quality Management | d006 | 67130.9355 | 4 |
| | Development | d005 | 59658.1176 | 2 |
| | Human Resources | d003 | 58286.0556 | 2 |
| | Production | d004 | 56233.4000 | 4 |
| | Customer Service | d009 | 54959.6724 | 4 |

How can we determine the total salary paid and the count of distinct employees for each department, and which department has the highest total salary paid?

SELECT

```
d.dept_name,  
m.dept_no,  
SUM(s.salary) AS avg_salary_paid,  
COUNT(DISTINCT e.emp_no) AS total_cont
```

FROM

```
dept_manager m  
JOIN  
employees e ON e.emp_no = m.emp_no  
JOIN  
salaries s ON s.emp_no = e.emp_no  
JOIN  
departments d ON d.dept_no = m.dept_no  
GROUP BY d.dept_name , m.dept_no  
ORDER BY avg_salary_paid DESC;
```

| | dept_name | dept_no | avg_salary_paid | total_cont |
|---|--------------------|---------|-----------------|------------|
| ▶ | Quality Management | d006 | 4162118 | 4 |
| | Production | d004 | 3374004 | 4 |
| | Customer Service | d009 | 3187661 | 4 |
| | Marketing | d001 | 3093009 | 2 |
| | Sales | d007 | 2915118 | 2 |
| | Research | d008 | 2558661 | 2 |
| | Finance | d002 | 2549372 | 2 |
| | Human Resources | d003 | 2098298 | 2 |
| | Development | d005 | 2028376 | 2 |

note : move to next question we need to create temporary table age_of_emp_comp

```
create temporary table age_of_emp_comp
SELECT
    e.emp_no AS employee_number,
    e.first_name as first_name ,
    e.gender AS gender,
    YEAR(e.birth_date) AS birth_year,
    YEAR(MIN(s.from_date)) AS start_year,
    (YEAR(MIN(s.from_date))-YEAR(e.birth_date)) as age_at_join,
    (YEAR('2002-01-01') - YEAR(MIN(s.from_date))) AS years_since_start,
    (YEAR('2002-01-01') - YEAR(e.birth_date)) AS age_in_2002,
    YEAR('2002-01-01') AS reference_year,
    YEAR(SYSDATE()) AS current_year,
    (YEAR(SYSDATE()) - YEAR(e.birth_date)) AS current_age
FROM
    employees e
    JOIN salaries s ON e.emp_no = s.emp_no
GROUP BY
    e.emp_no,
    e.gender;
```

Question: Write an SQL query to retrieve employee number, first name, gender, birth year, start year, age at joining, years since start (as of 2002), age in 2002, reference year (2002), current year, and current age. Group by employee number and gender

SELECT

```
e.emp_no AS employee_number,  
e.first_name AS first_name ,  
e.gender AS gender,  
YEAR(e.birth_date) AS birth_year,  
YEAR(MIN(s.from_date)) AS start_year,  
(YEAR(MIN(s.from_date))-YEAR(e.birth_date)) as age_at_join,  
(YEAR('2002-01-01') - YEAR(MIN(s.from_date))) AS years_since_start,  
(YEAR('2002-01-01') - YEAR(e.birth_date)) AS age_in_2002,  
YEAR('2002-01-01') AS reference_year,  
YEAR(SYSDATE()) AS current_year,  
(YEAR(SYSDATE()) - YEAR(e.birth_date)) AS current_age
```

FROM

```
employees e  
JOIN salaries s ON e.emp_no = s.emp_no
```

GROUP BY

```
e.emp_no,  
e.gender;
```

| | employee_number | first_name | gender | birth_year | start_year | age_at_join | years_since_start | age_in_2002 | reference_year | current_year | current_a |
|---|-----------------|------------|--------|------------|------------|-------------|-------------------|-------------|----------------|--------------|-----------|
| ▶ | 10001 | Georgi | M | 1953 | 1986 | 33 | 16 | 49 | 2002 | 2024 | 71 |
| | 10002 | Bezalel | F | 1964 | 1996 | 32 | 6 | 38 | 2002 | 2024 | 60 |
| | 10003 | Parto | M | 1959 | 1995 | 36 | 7 | 43 | 2002 | 2024 | 65 |
| | 10004 | Chirstian | M | 1954 | 1986 | 32 | 16 | 48 | 2002 | 2024 | 70 |
| | 10005 | Kyoichi | M | 1955 | 1989 | 34 | 13 | 47 | 2002 | 2024 | 69 |
| | 10006 | Anneke | F | 1953 | 1990 | 37 | 12 | 49 | 2002 | 2024 | 71 |
| | 10007 | Tzvetan | F | 1957 | 1989 | 32 | 13 | 45 | 2002 | 2024 | 67 |
| | 10008 | Saniya | M | 1958 | 1998 | 40 | 4 | 44 | 2002 | 2024 | 66 |
| | 10009 | Sumant | F | 1952 | 1985 | 33 | 17 | 50 | 2002 | 2024 | 72 |
| | 10010 | Duangkaew | F | 1963 | 1996 | 33 | 6 | 39 | 2002 | 2024 | 61 |

Question: Write an SQL query to group employees based on their age at joining into the categories: Rising Stars (20-30), Prime Movers (31-40), and Seasoned Experts (41-50). Calculate the total number of employees in each age group . and sum the employees accordingly. Group the results by age group

```
select age_group , sum(total_emp) from  
(select age_at_join , count(employee_number) as total_emp ,  
case  
when age_at_join between 20 and 30 then 'Rising Stars'  
when age_at_join between 31 and 40 then 'Prime Movers'  
when age_at_join between 41 and 50 then 'Seasoned Experts'  
end as age_group  
from age_of_emp_comp  
group by age_at_join) as tab11  
group by age_group;
```

| | age_group | sum(total_emp) |
|---|------------------|----------------|
| ▶ | Prime Movers | 58242 |
| | Rising Stars | 28878 |
| | Seasoned Experts | 14676 |

Question: Write an SQL query to classify employees based on their years of experience into the following categories: Emerging Talents (0-5 years), Proficient Professionals (5-10 years), Seasoned Specialists (10-15 years), and Veteran Virtuosos (15+ years). Calculate the total number of employees in each experience category and group the results accordingly

```
select age_of_experience , sum(total_emp) as total_employees from  
(select years_since_start , count(employee_number) as total_emp ,  
case  
when years_since_start between 0 and 5 then 'Emerging Talents'  
when years_since_start between 5 and 10 then 'Proficient Professionals'  
when years_since_start between 10 and 15 then 'Seasoned Specialists'  
else 'Veteran Virtuosos'  
end as age_of_experience  
from age_of_emp_comp  
group by years_since_start) as tab11  
group by age_of_experience ;
```

| | age_of_experience | total_employees |
|---|--------------------------|-----------------|
| ► | Veteran Virtuosos | 12836 |
| | Proficient Professionals | 34009 |
| | Seasoned Specialists | 34191 |
| | Emerging Talents | 20760 |

Write an SQL query to determine the total number of employees in each age group (Rising Stars, Prime Movers, Seasoned Experts) from the age_emp table, specifically where to_date is '9999-01-01'. Group the results by age group, and calculate the sum of employees in each age category

```
→ select age_group , sum(tot_emp_age) from (
  select age , count(emp_no) as tot_emp_age ,
  → case
    when age between 20 and 30 then 'Rising Stars'
    when age between 31 and 40 then 'Prime Movers'
    when age between 41 and 50 then 'Seasoned Experts'
  end as age_group
  from age_emp
  where to_date = '9999-01-01'
  → group by age) as b
  group by age_group ;
```

| | age_group | sum(tot_emp_age) |
|---|------------------|------------------|
| ▶ | Seasoned Experts | 62430 |
| ▶ | Prime Movers | 19205 |

Write an SQL query to determine the total number of employees in each age group (Rising Stars, Prime Movers, Seasoned Experts), further segmented by gender, from the age_emp table where the to_date is '9999-01-01'. Group the results by both age group and gender, and calculate the sum of employees in each category

```
select age_group , gender , sum(tot_emp_age) from  
  (select age , gender , emp_no, count(emp_no) as tot_emp_age ,  
    case  
      when age between 20 and 30 then 'Rising Stars'  
      when age between 31 and 40 then 'Prime Movers'  
      when age between 41 and 50 then 'Seasoned Experts'  
    end as age_group  
  from age_emp  
  group by age, gender,emp_no)as b  
group by age_group , gender;
```

| | age_group | gender | sum(tot_emp_age) |
|---|------------------|--------|------------------|
| ▶ | Prime Movers | M | 14429 |
| | Prime Movers | F | 9561 |
| | Seasoned Experts | M | 46535 |
| | Seasoned Experts | F | 31271 |

Conclusion

This SQL Analysis of Employee Data project has successfully demonstrated the power of structured queries to extract valuable insights from organizational data. Through a series of comprehensive analyses, we have been able to:

Understand Employee Demographics: By analyzing gender distribution and departmental demographics, we have a clearer picture of the workforce composition, which is crucial for planning diversity and inclusion initiatives.

Evaluate Salary Distribution: We've identified key trends in salary averages, top earners, and potential less gender pay gaps. This insight aids in ensuring equitable and competitive compensation strategies.

Analyze Tenure: By examining tenure, we can identify patterns in employee retention and turnover. This helps in strategizing HR practices to retain valuable talent and manage workforce stability.

Assess Performance Metrics: Correlating performance metrics with salary and tenure provides a deeper understanding of what drives high performance within the organization.



THANK YOU

For watching this presentation



divyanshpatel@751 gmail.com