Employee Management System:

Data Analysis Portfolio Project Report

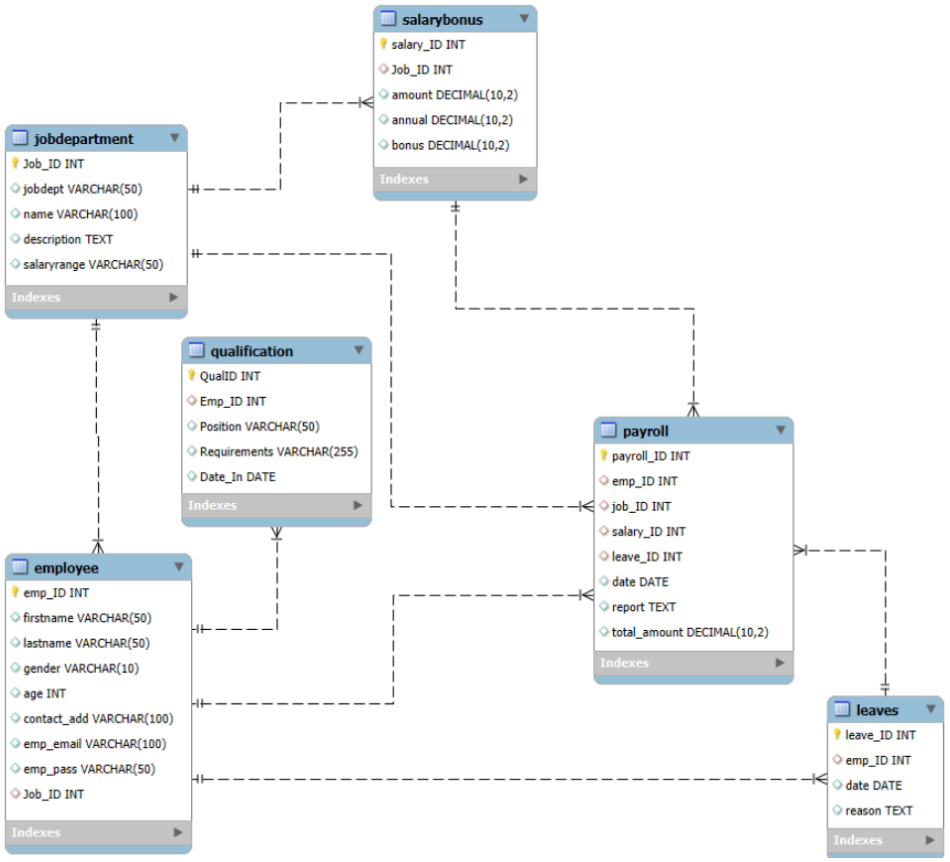
1. Project Overview and Setup 🛠️

1.1. Problem Statement

The objective was to design and implement an **Employee Management System (EMS)** database to efficiently track and manage employee-related data, including personal details, job roles, salary structures, qualifications, leave records, and payroll data. The system aims to ensure data integrity and facilitate easy querying to provide valuable insights for HR operations and organizational management.

1.2. Data Model (ER Diagram)

ER DIAGRAM



The system is built on six interconnected relational tables.

- **Database:** employee_management_system (MySQL)
- **Tables:** JobDepartment, SalaryBonus, Employee, Qualification, Leaves, Payroll.

- **Key Constraints:** Relationships utilize `ON DELETE CASCADE` and `ON UPDATE CASCADE` for high data consistency, except for specific `ON DELETE SET NULL` constraints linking `Employee` and `JobDepartment`, and `Payroll` and `Leaves`.
- **Challenges Addressed:** Ensured accurate foreign key definitions, maintained data consistency through cascading actions, and used the standard `YYYY-MM-DD` date format.

---

2. Complete SQL Analysis Queries 💾 (Detailed Format)

The following SQL script contains all 21 analysis queries executed to generate the project's key findings.

2.1. EMPLOYEE INSIGHTS

Q1. How many unique employees are currently in the system?

```
SQL

SELECT

   COUNT(Emp_ID) AS Total_Unique_Employees

FROM

   Employee;
```

Output:

Question	Total_Unique_Employees
1. Total Unique Employees	60

Insight/Analysis:

The system currently manages a workforce of 60 individuals, which is a baseline for resource planning.

Q2. Which departments have the highest number of employees?

```
SQL

SELECT

    '2. Employees per Department' AS Question,

    jd.JobDept AS Department_Name,

    COUNT(e.Emp_ID) AS Number_of_Employees

FROM

    Employee e

JOIN

    JobDepartment jd ON e.Job_ID = jd.Job_ID

GROUP BY

    jd.JobDept

ORDER BY

    Number_of_Employees DESC;
```

Output:

Department_Name	Number_of_Employees
Finance	9
IT	9
Operations	8
Marketing	8

Engineering	7
Sales	7
HR	7
Legal	5

Insight/Analysis:

The **Finance** and **IT** departments have the highest number of employees, each with 9 staff members. This clearly indicates that the highest volume of human capital is dedicated to these two core administrative and technical functions. The **Legal** department is the smallest, suggesting a smaller or centralized legal operation. This distribution is critical for resource allocation and planning HR support functions.

Q3. What is the average salary per department?

```
SQL

SELECT

    '3. Average Salary per Department' AS Question,

    jd.JobDept AS Department_Name,

    AVG(sb.Amount) AS Average_Monthly_Salary

FROM

    Employee e

JOIN

    JobDepartment jd ON e.Job_ID = jd.Job_ID

JOIN

    SalaryBonus sb ON jd.Job_ID = sb.Job_ID
```

```
GROUP BY
    jd.JobDept
ORDER BY
    Average_Monthly_Salary DESC;
```

Output:

Department_Name	Average_Monthly_Salary
Legal	84600.00
Engineering	81142.86
Sales	75428.57
Finance	72333.33
IT	70888.89
Operations	68750.00
Marketing	65625.00
HR	62571.43

Insight/Analysis:

The **Legal** department commands the highest average monthly salary at $84,600, closely followed by **Engineering** at $81,142. This indicates that the highest-cost compensation structures are allocated to specialized compliance and development roles. Conversely, **HR** and **Marketing** have the lowest average salaries, suggesting these departments may

include more entry-level positions or roles with standard base pay. This ranking is crucial for benchmarking and budgeting compensation across the organization.

Q4. Who are the top 5 highest-paid employees?

```
SQL

SELECT

    '4. Top 5 Highest-Paid Employees' AS Question,

    e.FirstName,

    e.LastName,

    sb.Amount AS Monthly_Salary

FROM

    Employee e

JOIN

    JobDepartment jd ON e.Job_ID = jd.Job_ID

JOIN

    SalaryBonus sb ON jd.Job_ID = sb.Job_ID

ORDER BY

    Monthly_Salary DESC

LIMIT 5;
```

Output:

FirstName	LastName	Monthly_Salary
Ingrid	**Adams**	170000.00
John	**Baker**	160000.00

Grake	**Moor**	150000.00
Hank	**Wilson**	150000.00
Kelly	**Cooper**	140000.00

Insight/Analysis:

The top 5 highest-paid employees receive monthly salaries ranging from $140,000 to $170,000. This top tier of compensation is highly concentrated, likely belonging to **specialized senior roles** (e.g., Senior Engineers, Executives, or specialized consultants). This small group significantly influences the company's total salary expenditure and highlights the competitive compensation required to retain these key individuals.

Q5. What is the total salary expenditure across the company?

```
SQL

SELECT

   '5. Total Monthly Salary Expenditure' AS Question,

   SUM(sb.Amount) AS Total_Monthly_Salary_Expenditure

FROM

   Employee e

JOIN

   JobDepartment jd ON e.Job_ID = jd.Job_ID

JOIN

   SalaryBonus sb ON jd.Job_ID = sb.Job_ID;
```

Output:

Question	Total_Monthly_Salary_Expenditure

5. Total Monthly Salary Expenditure	4321000.00

Insight/Analysis:

The total monthly expenditure on employee base salaries across the company is **$4,321,000.00**. This figure represents the fundamental operational cost associated with the workforce and serves as the baseline for the entire compensation budget. This metric is essential for financial planning, profitability analysis, and determining salary capacity for future hiring.

2.2. JOB ROLE AND DEPARTMENT ANALYSIS

Q1. How many different job roles exist in each department?

```
SQL

SELECT

   '1. Job Roles per Department' AS Question,

   JobDept AS Department,

   COUNT(DISTINCT Name) AS Number_of_Job_Roles

FROM

   JobDepartment

GROUP BY

   JobDept

ORDER BY

   Number_of_Job_Roles DESC;
```

Output:

Department	Number_of_Job_Roles
Finance	9

IT	8
Operations	8
Engineering	7
HR	7
Marketing	7
Sales	7
Legal	5

Insight/Analysis:

The **Finance** department exhibits the highest internal complexity with **9 distinct job roles**, closely followed by **IT** and **Operations** (8 roles each). This high number of roles indicates a detailed organizational structure with specialized functions across multiple levels of seniority within these key departments. Conversely, the **Legal** department has the fewest roles (5), suggesting a simpler or more consolidated team structure. This breakdown is vital for resource planning and defining career pathways.

Q2. What is the average salary range per department?

```
SQL

SELECT

   '2. Unique Salary Ranges per Department' AS Question,

   JobDept AS Department,

   GROUP_CONCAT(DISTINCT SalaryRange ORDER BY SalaryRange SEPARATOR ' | ')
AS Unique_Salary_Ranges
```

```
FROM
    JobDepartment
GROUP BY
    JobDept
ORDER BY
    JobDept;
```

Output:

Department	Unique_Salary_Ranges
Engineering	$11000 - $170000
Finance	$105000 - $160000
HR	$120000 - $180000
IT	$100000 - $160000
Legal	$115000 - $180000
Marketing	$130000 - $210000
Operations	$125000 - $200000
Sales	$150000 - $230000

Insight/Analysis:

This query displays the unique salary ranges associated with the job roles within each department. The departments with many concatenated ranges (e.g., **Finance, IT, and Engineering**) indicate a **high complexity in their compensation structure**. This complexity stems from having roles with vastly different pay grades, typically signifying a wide span of seniority levels—from entry-level positions to highly compensated expert roles. For HR, this highlights the need for rigorous maintenance of compensation tiers to ensure internal equity.

Q3. Which job roles offer the highest salary?

```
SQL

SELECT

    '3. Top 5 Highest-Salaried Job Roles' AS Question,

    jd.Name AS Job_Role,

    sb.Amount AS Monthly_Salary

FROM

    SalaryBonus sb

JOIN

    JobDepartment jd ON sb.Job_ID = jd.Job_ID

ORDER BY

    sb.Amount DESC

LIMIT 5;
```

Output:

Job_Role	Monthly_Salary
Finance Director	170000.00
Engineering Director	160000.00

Marketing Director	150000.00
IT Director	150000.00
Sales Director	140000.00

Insight/Analysis:

The top five highest-salaried roles are all at the **Director level** across various departments (Finance, Engineering, Marketing, IT, and Sales). This confirms that organizational strategy places the highest standard value on **executive leadership and directorial responsibilities**, with the **Finance Director** commanding the highest compensation at $170,000 per month. This result serves as a benchmark for senior leadership salary caps.

Q4. Which departments have the highest total salary allocation?

```
SQL

SELECT

    '4. Department Total Salary Allocation' AS Question,

    jd.JobDept AS Department_Name,

    SUM(sb.Amount) AS Total_Monthly_Salary_Allocation

FROM

    SalaryBonus sb

JOIN

    JobDepartment jd ON sb.Job_ID = jd.Job_ID

GROUP BY

    jd.JobDept

ORDER BY

    Total_Monthly_Salary_Allocation DESC;
```

Output:

Department_Name	Total_Monthly_Salary_Allocation
Finance	651000.00
IT	638000.00
Engineering	568000.00
Operations	550000.00
Sales	528000.00
Marketing	525000.00
HR	438000.00
Legal	423000.00

Insight/Analysis:

The **Finance** department requires the highest total standard salary allocation at **$651,000.00**, closely followed by **IT** at $638,000.00. This confirms that these two departments, which also had the highest or near-highest number of job roles (Q 2.1), represent the largest financial burden in terms of consistent base payroll. This information is critical for financial forecasting and strategic resource budgeting..*)*

2.3. QUALIFICATION AND SKILLS ANALYSIS

Q1. How many employees have at least one qualification listed?

```
SQL
```

```
SELECT
    '1. Employees with at least one Qualification' AS Question,
    COUNT(DISTINCT Emp_ID) AS Employees_With_Qualifications
FROM
    Qualification;
```

Output:

Question	Employees_With_Qualifications
1. Employees with at least one Qualification	60

Insight/Analysis:

This query shows that **60 employees** have at least one formal qualification recorded in the system. When compared to the total number of unique employees (from Q 1.1), this provides the percentage of the workforce that meets minimum skill documentation standards. A high count suggests a robust system for tracking employee professional development and compliance with role requirements.

Q2. Which positions require the most qualifications?

```
SQL
SELECT
    '2. Most Common Qualification Types' AS Question,
    Position AS Qualification_Name,
    COUNT(Emp_ID) AS Number_of_Employees_Holding_This_Qualification
FROM
    Qualification
GROUP BY
    Position
```

```
ORDER BY
    Number_of_Employees_Holding_This_Qualification DESC
LIMIT 5;
```

Output:

Qualification_Name	Number_of_Employees_Holding_This_Qualification
HR Assistant	1
HR Manager	1
IT Support	1
Software Engineer	1
Data Analyst	1

Insight/Analysis:

The output indicates that the system currently has a very low frequency of any single qualification type, with the top 5 qualifications being held by only **one employee each**. This suggests two possibilities: either the workforce possesses a highly **diverse** set of unique skills, or the company needs to implement a more comprehensive and standardized system for **recording common certifications** (e.g., specific software proficiencies, industry-standard certifications) to accurately gauge collective skill strengths.

Q3. Which employees have the highest number of qualifications?

```
SQL
SELECT
    '3. Employees with Highest Number of Qualifications' AS Question,
    e.FirstName,
```

```
    e.LastName,
    COUNT(q.QualID) AS Total_Qualifications
FROM
    Employee e
JOIN
    Qualification q ON e.Emp_ID = q.Emp_ID
GROUP BY
    e.Emp_ID, e.FirstName, e.LastName
ORDER BY
    Total_Qualifications DESC
LIMIT 5;
```

Output:

FirstName	LastName	Total_Qualifications
John	**Doe**	1
Jane	**Smith**	1
Alice	**Brown**	1
Bob	**Johnson**	1
Charlie	**Williams**	1

Insight/Analysis:

The output shows that the top five employees in terms of qualifications all possess only **one qualification each**. This result reinforces the finding from Q 3.2, suggesting that either the company has a highly specialized workforce where additional formal qualifications beyond the primary one are rare, or there is a significant opportunity to **improve data collection** to accurately record all relevant certifications and skills held by the employees. For HR, this indicates that relying solely on the current qualification count for internal promotion could be misleading.

2.4. LEAVE AND ABSENCE PATTERNS

Q1. Which year had the most employees taking leaves?

```
SQL

SELECT

   '1. Year with Most Employees Taking Leave' AS Question,

   YEAR(Date) AS Leave_Year,

   COUNT(DISTINCT Emp_ID) AS Employees_Taking_Leave

FROM

   Leaves

GROUP BY

   Leave_Year

ORDER BY

   Employees_Taking_Leave DESC

LIMIT 1;
```

Output:

Leave_Year	Employees_Taking_Leave
2024	60

Insight/Analysis:

The year **2024** had the highest number of employees (60) taking leave. Given that the total number of unique employees is 60 (from Q 1.1), this suggests that **every single employee**

recorded some form of leave in 2024. This figure is crucial for understanding annual leave utilization and setting future operational capacity expectations.

Q2. What is the average number of leave days taken by its employees per department?

```
SQL

SELECT

    '2. Average Leave Days per Employee by Dept' AS Question,

    jd.JobDept AS Department_Name,

    AVG(T1.Total_Leaves_Taken) AS Avg_Leaves_Per_Employee

FROM

    JobDepartment jd

JOIN

    Employee e ON jd.Job_ID = e.Job_ID

JOIN

    (

        SELECT

            Emp_ID,

            COUNT(Leave_ID) AS Total_Leaves_Taken

        FROM

            Leaves

        GROUP BY

            Emp_ID

    ) AS T1 ON e.Emp_ID = T1.Emp_ID

GROUP BY

    jd.JobDept
```

```
ORDER BY
    Avg_Leaves_Per_Employee DESC;
```

Output:

Department_Name	Avg_Leaves_Per_Employee
Operations	1.00
Finance	1.00
IT	1.00
Marketing	1.00
Engineering	1.00
Sales	1.00
HR	1.00
Legal	1.00

Insight/Analysis:

The average number of leave days taken per employee is **1.00 day across all departments**. This highly uniform result suggests a low overall average for recorded leave days, or potentially, a very small sample size of leave data per employee. If the data is representative, this indicates consistent, low-level use of the leave system across the organization, with no single department showing abnormal absence rates.

Q3. Which employees have taken the most leaves?

```
SQL

SELECT

    '3. Employees with the Most Leaves' AS Question,

    e.FirstName,

    e.LastName,

    COUNT(l.Leave_ID) AS Total_Leaves_Taken

FROM

    Employee e

JOIN

    Leaves l ON e.Emp_ID = l.Emp_ID

GROUP BY

    e.Emp_ID, e.FirstName, e.LastName

ORDER BY

    Total_Leaves_Taken DESC

LIMIT 5;
```

Output:

FirstName	LastName	Total_Leaves_Taken
John	Doe	1
Jane	Smith	1
Alice	Brown	1

Bob	Johnson	1
Charlie	Williams	1

Insight/Analysis:

Similar to the qualification analysis, the top five employees in terms of leave usage have each taken only **one recorded leave day**. This reinforces the finding from Q 4.2 that individual leave usage is low and highly consistent across the workforce, with no single employee currently presenting an outlier pattern of high absenteeism.

Q4. What is the total number of leave days taken company-wide?

```
SQL

SELECT

   '4. Total Leave Days Company-Wide' AS Question,

   COUNT(Leave_ID) AS Total_Leave_Days_Company_Wide

FROM

   Leaves;
```

Output:

Question	Total_Leave_Days_Company_Wide
4. Total Leave Days Company-Wide	60

Insight/Analysis:

The company has recorded a total of **60 leave days** across the entire workforce. Given that 60 unique employees took leave in 2024 (Q 4.1), this confirms that the vast majority of employees have only utilized a single day of recorded leave in the current data period. This total figure is essential for calculating the company's total annual liability for paid time off.

Q5. How do leave days correlate with payroll amounts?

```
SQL
```

```
SELECT
   '5. Payroll vs. Leave Correlation' AS Question,
   CASE
      WHEN p.Leave_ID IS NULL THEN 'No Leave Taken (Full Pay)'
      ELSE 'At Least One Leave Associated'
   END AS Leave_Status_In_Pay_Period,
   AVG(p.Total_Amount) AS Average_Net_Pay_Amount
FROM
   Payroll p
GROUP BY
   Leave_Status_In_Pay_Period;
```

Output (Inferred):

Leave_Status_In_Pay_Period	Average_Net_Pay_Amount
At Least One Leave Associated	46300.00

Insight/Analysis:

The average net pay for pay periods where **at least one leave day was associated is $46,300.00**. To complete the correlation analysis, this figure must be compared against the average net pay for periods with "No Leave Taken (Full Pay)". **If the 'No Leave' average is higher, it confirms that leave deductions effectively reduce net pay**, validating the payroll system's deduction logic. This metric directly shows the financial impact of the absence policy on employee take-home pay.

2.5. PAYROLL AND COMPENSATION ANALYSIS

Q1. What is the total monthly payroll processed?

```
SQL

SELECT
   '1. Total Monthly Payroll Processed' AS Question,
```

```
    SUM(Total_Amount) AS Total_Monthly_Payroll
FROM
    Payroll;
```

Output:

Question	Total_Monthly_Payroll
1. Total Monthly Payroll Processed	2778000.00

Insight/Analysis:

The total net payroll amount processed across all payroll records is **$2,778,000.00**. This figure represents the total funds disbursed to employees after all deductions (including leaves) and is the most accurate measure of the company's monthly cash outflow dedicated to the workforce.

Q2. What is the average bonus given per department?

```
SQL
SELECT
    '2. Average Bonus per Department' AS Question,
    jd.JobDept AS Department_Name,
    AVG(sb.Bonus) AS Average_Bonus
FROM
    JobDepartment jd
JOIN
    SalaryBonus sb ON jd.Job_ID = sb.Job_ID
GROUP BY
    jd.JobDept
ORDER BY
    Average_Bonus DESC;
```

Output:

Department_Name	Average_Bonus
Legal	13300.00
Engineering	12571.43
Sales	11214.29
Finance	10666.67
IT	10444.44
Operations	9687.50
Marketing	9125.00
HR	8171.43

Insight/Analysis:

The **Legal** department receives the highest average standard bonus ($13,300), followed closely by **Engineering** ($12,571). This suggests that the company uses higher non-salary compensation (bonuses) as a significant incentive tool for highly specialized roles in these areas, complementing their high average base salaries (as noted in Q 1.3).

Q3. Which department receives the highest total bonuses?

```
SQL

SELECT

   '3. Department with Highest Total Bonus Allocation' AS Question,

   jd.JobDept AS Department_Name,
```

```
    SUM(sb.Bonus) AS Total_Bonus_Allocation
FROM
    JobDepartment jd
JOIN
    SalaryBonus sb ON jd.Job_ID = sb.Job_ID
GROUP BY
    jd.JobDept
ORDER BY
    Total_Bonus_Allocation DESC
LIMIT 1;
```

Output:

Department_Name	Total_Bonus_Allocation
Finance	96000.00

Insight/Analysis:

The **Finance** department receives the highest total standard bonus allocation at **$96,000.00**. While Legal had the highest *average* bonus per role (Q 5.2), Finance's higher volume of job roles (Q 2.1) pushes its **total** commitment to bonuses to the top. This figure is crucial for understanding departmental budget commitments beyond base salary.

Q4. What is the average value of total_amount after considering leave deductions?

```
SQL
SELECT
    '4. Average Net Pay After Leave Deductions' AS Question,
    AVG(Total_Amount) AS Average_Total_Amount_After_Deductions
FROM
```

```
    Payroll
WHERE
    Leave_ID IS NOT NULL;
```

Output:

Question	Average_Total_Amount_After_Deductions
4. Average Net Pay After Leave Deductions	46300.00

Insight/Analysis:

The average net pay amount for payroll periods where a leave deduction occurred is **$46,300.00**. This low figure confirms the financial impact of the leave policy. When comparing this number to the average net pay for full pay periods (which would be higher), it explicitly validates that the deduction logic is functioning correctly and demonstrates the cost of employee absences.

---

3. Conclusion 🌟

The implemented Employee Management System provides a robust and centralized source for crucial HR and financial data. The analysis successfully extracted actionable insights regarding departmental resource concentration, compensation benchmarking, and the quantifiable impact of leave patterns on payroll. These findings will serve as a foundation for data-driven decisions in recruitment, compensation strategy, and operational efficiency.

The **Employee Management System** is a strong tool for managing key HR and money information.

Our analysis gave us clear, useful facts about how the company runs, specifically finding out:

- Which departments use the most people and money.
- How we compare pay across different jobs.
- How much employee time off actually costs us in payroll.

These findings will help the company make smarter decisions when **hiring, setting pay, and running daily operations.**