National Economics University

College of Technology

Introduction to Databases

GROUP 3 - PROJECT 2

Employee Information Manager

Authors:
11247338 - Tran Phuong Oanh
11247353 - Bui Phuong Thao
11247374 - Nguyen Ha Vy

Lecturer:
Dr. Hung Tran

December 12, 2025

Declaration

We, Tran Phuong Oanh, Bui Phuong Thao, Nguyen Ha Vy of the Faculty of Data Science and Artificial Intelligence, University of National Economics, confirm that this is our own work and figures, tables and illustrations in this report are original and have not been taken from any other person's work, except where the works of others have been explicitly acknowledged, quoted, and referenced. We understand that if failing to do so will be considered a case of plagiarism. Plagiarism is a form of academic misconduct and will be penalised accordingly.

We give consent to a copy of our report being shared with future students as an exemplar.

We give consent for our work to be made available more widely to members of NEU and public with interest in teaching, learning and research.

Nguyen Ha Vy
Tran Phuong Oanh
Bui Phuong Thao

December 12, 2025

Abstract

This project presents the design and implementation of a relational database management system for employee and project information, normalized to the third normal form (3NF). The system addresses common data redundancy and integrity issues found in spreadsheet-based HR and project management workflows. A Python-based graphical user interface (GUI) application was developed to support full CRUD (Create, Read, Update, Delete) operations across four core entities: Employees, Departments, Projects, and Assignments. The application features a query dashboard with four predefined analytical queries, a search and filter system, a data visualization dashboard, and CSV export functionality. The entire system was implemented using MySQL for data persistence and was collaboratively developed using GitHub with a structured workflow. The final deliverable includes a complete LaTeX report, presentation slides, and a demonstration video, showcasing a practical, scalable solution for modern employee information management.

Acknowledgement

Table of Contents

Contents

1 Introduction

1.1 Problem Context

Many organizations manage employee information using spreadsheets that combine personal data, project assignments, roles, and salaries in a single table. This approach leads to several issues including data redundancy, update anomalies, and inconsistent reporting. The lack of proper database normalization makes it difficult to maintain data integrity and perform efficient queries.

1.2 Project Objectives

The main objectives of this project are:

- To design and implement a normalized relational database schema (3NF) for managing employee and project information

- To develop a Python GUI application with CRUD functionality for all core entities

- To implement analytical queries and visualization features for data insights

- To follow software engineering best practices including version control and documentation

1.3 Scope

The project scope includes:

- Database design and normalization to 3NF

- Python application with MySQL backend

- Basic GUI with data validation

- Four required analytical queries

- Dashboard with visualizations

1.4 Out of Scope

- Authentication and authorization systems

- Large-scale performance optimization

- Mobile or web deployment

2 Database Design

2.1 Normalization Process

2.1.1 UNF to 1NF

The initial Assignments table was in Unnormalized Form (UNF) because it contained repeating groups for Project–Role–Salary information. A single record included multiple project-related attributes such as:

- EmployeeID

- EmployeeName

- Project1, Role1, Salary1

- Project2, Role2, Salary2

Problem Identified: Storing multiple Project–Role–Salary values in one row violated the requirements of First Normal Form (1NF), as the table did not ensure atomic attribute values.

To convert the table to 1NF, each Project–Role–Salary assignment was decomposed into a separate row. This restructuring ensured that every column contained a single, indivisible value.

Resulting 1NF Structure:

- AssignmentID

- EmployeeID

- EmployeeRole

- ProjectID

- Salary

Keys:

- **Primary Key (PK):** AssignmentID

- **Foreign Keys (FK):**

    - EmployeeID referencing Employees(EmployeeID)
    - ProjectID referencing Projects(ProjectID)

2.1.2 1NF to 2NF

To progress from First Normal Form (1NF) to Second Normal Form (2NF), it was necessary to remove partial dependencies. According to the 2NF rule, a relation that uses a composite primary key must not contain attributes that depend on only a portion of that key.

When considering EmployeeID and ProjectID as a composite primary key, it was identified that certain attributes did not depend on the full key. Specifically, the attribute Salary depended solely on EmployeeID rather than the entire (EmployeeID, ProjectID) combination. This constituted a partial dependency and therefore violated 2NF.

To resolve this issue, the data was decomposed into separate relations, resulting in the following tables: Employees, Projects, and Assignments. The Assignments table serves as a linking relation between employees and projects while also storing role and project-specific salary information.

Resulting 2NF Assignment Table Structure:

- AssignmentID

- EmployeeID

- ProjectID

- EmployeeRole

- Salary

Keys:

- **Primary Key (PK):** AssignmentID

- **Foreign Keys (FK):**

    - EmployeeID referencing Employees(EmployeeID)
    - ProjectID referencing Projects(ProjectID)

2.1.3 2NF to 3NF

The transition from Second Normal Form (2NF) to Third Normal Form (3NF) requires the elimination of transitive dependencies. A transitive dependency occurs when a non-key attribute depends on another non-key attribute rather than directly on the primary key.

In the original design, a transitive dependency was identified in the Assignments relation. The project manager information was indirectly dependent on the project through the relationship:

$$\text{Project} \rightarrow \text{Manager} \rightarrow \text{EmployeeID}$$

Storing the manager attribute within the Assignments table would therefore introduce a transitive dependency, violating the 3NF requirement. To resolve this issue, the manager attribute was relocated to the Projects table, ensuring that each project directly references its assigned manager.

Resulting 3NF Project Table Structure:

- ProjectID (PK)

- ProjectName

- ManagerEmployeeID (FK $\rightarrow$ Employees(EmployeeID))

The *Assignments* table now solely represents the relationship between employees and projects, without containing attributes unrelated to this association.

2.1.4 Realistic Sample Data

	EmployeeID	EmployeeName	DateOfBirth	DepartmentName	ProjectName	EmployeeRole	Salary
▶	1	Vũ Thị Nhung	1990-04-23	Sales	Urban Development Research	Coordinator	12500000.00
	1	Vũ Thị Nhung	1990-04-23	Sales	Academic Administration Management	Developer	12500000.00
	2	Bùi Minh Quân	1982-08-19	Finance	Economic Policy Research 2024	Manager	18000000.00
	2	Bùi Minh Quân	1982-08-19	Finance	MBA Program Improvement	Team Lead	18000000.00
	3	Ngô Thị Lan	1985-12-02	Human Resources	Employee Engagement Program	HR Specialist	14200000.00
	3	Ngô Thị Lan	1985-12-02	Human Resources	Talent Development Plan	Recruiter	14200000.00
	4	Hoàng Văn Nam	1981-01-14	Data Science	Data Analytics Platform	Data Scientist	21500000.00
	4	Hoàng Văn Nam	1981-01-14	Data Science	AI Research Initiative	ML Engineer	21500000.00
	5	Trần Thị Thanh	1983-05-27	Engineering	Predictive Maintenance System	QA Engineer	16500000.00
	5	Trần Thị Thanh	1983-05-27	Engineering	Product Design Revamp	Designer	16500000.00
	6	Lê Văn Hùng	1979-09-09	Data Science	Data Analytics Platform	Senior Data ...	24500000.00
	6	Lê Văn Hùng	1979-09-09	Data Science	AI Research Initiative	Research Lead	24500000.00
	7	Phạm Thị Ngọc	1986-03-05	Finance	Economic Policy Research 2024	Financial Ana...	15800000.00
	7	Phạm Thị Ngọc	1986-03-05	Finance	Financial Market Analysis	Accountant	15800000.00
	8	Trương Minh Hải	1984-07-16	Human Resources	Employee Engagement Program	HR Manager	15200000.00
	8	Trương Minh Hải	1984-07-16	Human Resources	Talent Development Plan	Training Coor...	15200000.00
	9	Đinh Thị Lan	1987-12-29	Legal Affairs	Legal Compliance Audit	Legal Advisor	16800000.00
	9	Đinh Thị Lan	1987-12-29	Legal Affairs	Contract Management System	Compliance ...	16800000.00
	10	Nguyễn Văn Sơn	1979-10-11	Marketing	Market Expansion Strategy	Marketing Sp...	14800000.00
	10	Nguyễn Văn Sơn	1979-10-11	Marketing	Brand Awareness Campaign	Brand Manager	14800000.00
	11	Vũ Thị Hương	1990-02-08	Sales	Sales Optimization Program	Sales Executive	13500000.00
	11	Vũ Thị Hương	1990-02-08	Sales	Customer Relationship Management	Account Man...	13500000.00
	12	Bùi Văn Tiến	1985-06-23	Finance	Financial Reporting System	Financial Con...	17200000.00

Figure 1: Realistic Sample Data of the Employee Information System

Data Volume Requirements Met:

- **Employees:** 160 records with diverse names, birth dates, and department assignments

- **Departments:** 7 distinct departments covering various business functions

- **Projects:** 38 active and historical projects with different managers

- **Assignments:** 320 assignment records with varied roles and salary distributions

2.2 Entity–Relationship Diagram

The Entity–Relationship Diagram (ERD) provides a visual representation of the database schema, illustrating the entities, their attributes, and the relationships between them. This diagram is essential for understanding the overall structure of the system, enforcing referential integrity, and supporting the design of efficient queries.

2.2.1 Entities and Attributes

The database consists of the following main entities:

- **Departments**
  Attributes: DepartmentID (Primary Key), DepartmentName
  Represents the organizational units within the company.

- **Employees**
  Attributes: EmployeeID (Primary Key), EmployeeName, DateOfBirth, DepartmentID (Foreign Key → Departments)
  Captures information about all employees and their corresponding departmental assignments.

- **Projects**
  Attributes: ProjectID (Primary Key), ProjectName, ManagerEmployeeID (Foreign Key → Employees)
  Represents company projects and identifies the employee responsible as the project manager.

- **Assignments**
  Attributes: AssignmentID (Primary Key), EmployeeID (Foreign Key → Employees), ProjectID (Foreign Key → Projects), EmployeeRole, Salary
  Records employee participation in projects, including their role and associated salary.

2.2.2 Relationships

The ERD highlights the following relationships:

- **Departments → Employees (1:N):** Each department can have multiple employees, while each employee belongs to exactly one department.

- **Employees → Assignments (1:N):** An employee may be associated with multiple project assignments.

- **Projects → Assignments (1:N):** Each project can involve multiple employees, with each assignment linking one employee to one project.

- **Employees → Projects (1:1 for Manager):** Each project has exactly one manager, represented through a foreign key reference to an employee.

2.2.3 ERD
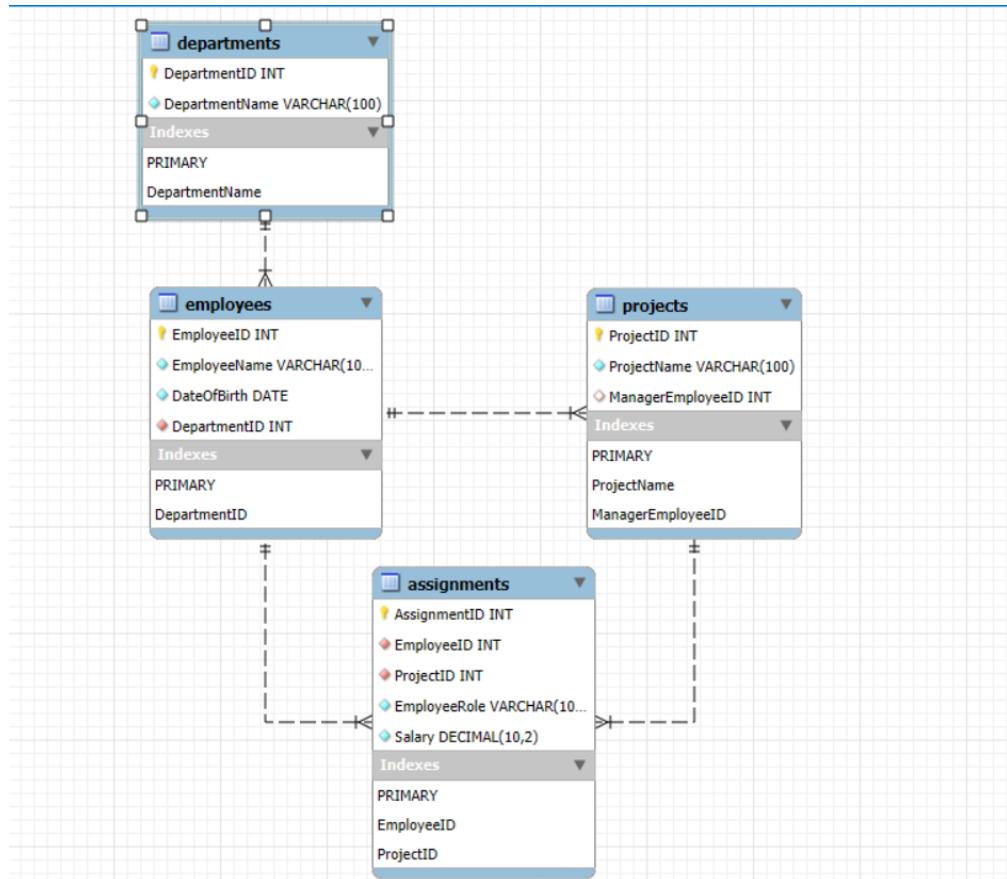
Figure 2: Entity-Relationship Diagram of the Employee Information System

3 GUI Implementation and Query Interfaces

3.1 CRUD Screens

3.1.1 Definition of CRUD

CRUD stands for Create, Read, Update, and Delete, which are the four fundamental operations in data management. The implementation of CRUD functionality in the system is designed to provide an intuitive and user-friendly interface that enables users to interact with the database efficiently and accurately. By incorporating these operations, the system ensures that data can be created, retrieved, modified, and removed in a controlled and reliable manner.

3.1.2 Entity Forms

We developed interactive forms for each entity with complete CRUD capabilities:

- **Employees Form**: Manages personal information such as name, date of birth, and department assignment.

- **Departments Form**: Supports the creation and maintenance of departments while enforcing unique naming constraints.

- **Projects Form**: Handles project details and manager assignments.

- **Assignments Form**: Enables the creation and modification of employee–project relationships, including role specification and salary information.

3.1.3 Validation Mechanisms

All forms incorporate real-time validation mechanisms that provide immediate feedback to users. Salary fields accept only numeric input and enforce minimum value constraints, while date fields validate both format and logical ranges. Violations of unique constraints are detected and reported before submission to the database. These validation features ensure that data integrity is maintained and errors are minimized during entry.

3.1.4 User Experience Enhancements

The system enhances user experience by offering instant responses to incorrect input. Error messages are displayed clearly and concisely, guiding users to correct mistakes quickly. The interface is designed to be intuitive and accessible, ensuring smooth interaction and reliable data entry throughout the application. This approach reduces user frustration and promotes efficiency in daily operations.

Figure 3: Assignments Record Creation Interface

3.2 Required Queries Interface

The system provides four dedicated query screens designed to support efficient data analysis and organizational management. The first screen is based on an INNER JOIN query, which displays all active employee–project assignments. This view presents employee names, their assigned roles, and corresponding salaries, enabling managers to monitor resource allocation with accuracy.

The second screen utilizes a LEFT JOIN query to list all employees, including those without current project assignments. This functionality highlights gaps in resource allocation and assists decision-makers in identifying underutilized personnel.

The third screen implements a Multi-table JOIN query, offering a comprehensive overview that connects employees to their projects, roles, salaries, and respective managers. This consolidated view provides valuable insights into organizational structures and relationships across multiple dimensions.

The fourth screen applies an Above Global Average query, which identifies employees whose average assignment salary exceeds the organization-wide average. This feature is particularly useful for compensation analysis and for recognizing high-value contributors.

Each query interface is equipped with advanced usability features. Columns can be sorted in ascending or descending order, column-specific search filters allow precise queries, and pagination ensures smooth navigation through large result sets. Additionally, a one-click CSV export function enables seamless integration with external reporting tools.

In addition, users can choose to display specific columns from all four tables simultaneously by pressing the Global Columns button. This feature provides significant flexibility, allowing users to focus only on the data most relevant to their analysis, eliminate unnecessary information, and save time when searching and cross-referencing records. As a result, data exploration becomes more intuitive, accurate, and tailored to the specific needs of each scenario.
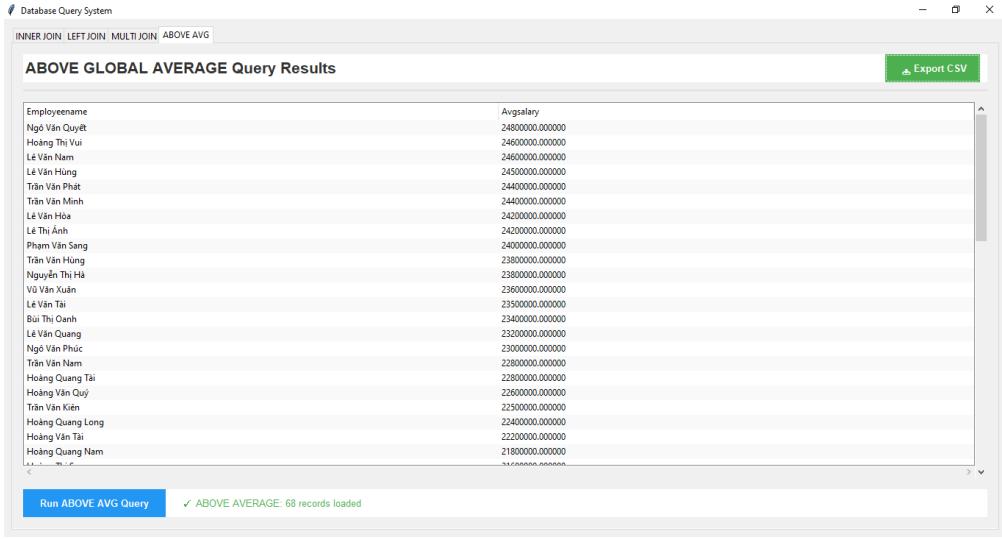
Figure 4: Query Interfaces
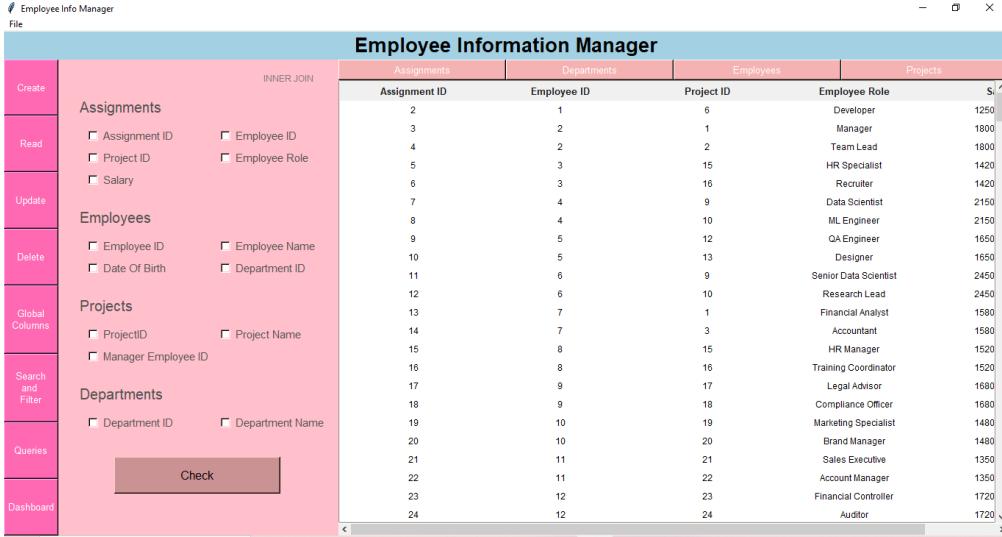
Figure 5: Global Columns Interfaces

3.3 Dashboard Implementation

The dashboard was developed to consolidate and visualize key organizational metrics in a single, interactive interface. Its purpose is to provide managers and stakeholders with real-time insights into workforce allocation, financial performance, and role distribution, thereby supporting informed decision-making.

The KPI Summary section presents live counts of total employees, departments, projects, and active assignments. This ensures that users can immediately assess the overall scale and current workload of the organization.

The Financial Overview component calculates the average salary and updates dynamically with every modification to employee assignments. This feature provides a clear picture of compensation trends and facilitates monitoring of budgetary impacts.

The Salary Distribution is represented through an interactive histogram that visualizes salary ranges across all assignments. By exploring this chart, users can identify patterns, detect disparities, and evaluate the fairness of salary allocation.

The Top Performers module generates a configurable Top-N list of employees ranked by their average assignment salary. This ranking highlights individuals with outstanding contributions and supports recognition and reward strategies.

Finally, the Role Analysis section employs a pie chart to illustrate the distribution of different roles within the organization. This visualization enables managers to evaluate structural balance, identify potential gaps, and plan resource allocation more effectively.

Together, these components transform raw data into actionable insights, making the dashboard a central tool for monitoring organizational health and guiding strategic decisions.

3.4 Search and Filter System

3.4.1 The purpose of Search Filter system

The Search and Filter system is designed in purpose to provide a fast and intuitive method for addressing employee information across multiple related tables. This module significantly decreases the navigation time and increases productivity for HR staff and project managers by allowing them to access precise information, giving opportunities for decision-making.

3.4.2 Global Search

The Global Search is a unified search feature that allows users to find information across multiple pages using a single search bar. Instead of searching separately in different modules (such as Employees, Projects, or Assignments) the user can type any keyword once, and the

system will automatically look through all relevant fields in the database

Global Search does not require an exact match, it uses character-based pattern matching to return all relevant results. Users can search by 'EmployeeName', 'EmployeeID', 'Employee-Role', 'ProjectName', etc in just one single bar for all data.

The Global Search is useful since it helps HR and project managers quickly locate detailed information that they need.

3.4.3 Filter

The Filter is a tool that allows users to narrow down search results by applying specific conditions or criteria. While Global Search finds all records related to a keyword, Filters refine the results to make them more accurate and relevant. Filters work by adding conditions to the SQL query, allowing the system to show only the data that matches the selected option.

The system provides several filter categories :

- Department - Displays employees who belong to a selected department

- Project - Shows employees who are assigned to a chosen project

- Role - Filters employees based on their specific job role (e.g., Developer, Analyst)

- Salary Range - Allows users set a specific limit of salary value

Users may apply any combination of these filters at the same time. This multi-criteria filtering mechanism ensures high accuracy, reduces irrelevant results, and supports flexible analytical workflows. The Filter module uses different input types to balance accuracy and flexibility. The Department and Project filters use both dropdown menus and text input fields, allowing users to either select from predefined lists or search by typing keywords. In contrast, the Role and Salary Range filters use text input fields only, as these values vary more widely and require flexible, user-defined input. This combination enables precise and efficient filtering across different types of data.

3.4.4 Combined Workflow

The combined workflow describes how the Global Search and Advanced Filters operate together to deliver accurate, efficient, and highly customizable data retrieval. The process begins when the user enters a keyword into the global search bar. At this initial stage, the system performs a character-based lookup across multiple related tables using JOIN operations. This first query returns all records that contain the entered keyword, regardless of department, project, role, or salary. It acts as a broad, inclusive search designed to gather every potentially relevant result.

Once the initial results are displayed, the user can refine the dataset by applying one or more filters. These filters - such as department, project, role, manager, or salary range - provide additional constraints that help narrow down the results to the exact subset of data needed. When filters are applied, the system reconstructs the SQL query dynamically by combining the search keyword with all active filter conditions. Each filter is transformed into a corresponding WHERE clause, and multiple filters are chained together using logical AND operators. This ensures that only records satisfying both the global keyword search and the selected filter options are included in the final output.

Throughout this workflow, the interface updates the results in a clean, organized table that supports sorting and validation. Users can iteratively adjust their keyword or filter options, and the system will continuously rebuild and re-execute the query in real time. This integrated approach allows the Search and Filter module to handle both broad keyword discovery and precise analytical filtering within a single, cohesive workflow. As a result, the system provides a powerful, user-friendly method for retrieving complex employee and project information efficiently and accurately.

4 Results Interpretation and Analysis

4.1 Query Results Analysis

4.1.1 Query Functionality

The query outputs provide a set of valuable insights into organizational operations and workforce allocation. Results from the INNER JOIN query reveal the current project assignment landscape, clearly indicating which employees are actively engaged and in what capacity. This allows managers to monitor ongoing projects and assess resource utilization effectively.

The LEFT JOIN query results identify employees who are not currently assigned to any project. Such findings highlight potential underutilized resources or available bench strength, supporting decisions related to workforce planning and future project allocation.

The Multi-table JOIN query offers a comprehensive view of organizational structure by connecting employees to their projects, roles, salaries, and respective managers. This integrated perspective provides clarity on reporting relationships and project hierarchies, enabling a deeper understanding of organizational dynamics.

The Above Average query results emphasize high-value employees whose average assignment salary exceeds the organization-wide benchmark. These insights are particularly useful for retention planning, compensation analysis, and performance evaluation.

All query results are presented in sortable, paginated tables, ensuring usability and efficiency even when working with large datasets. This design enhances accessibility, supports targeted analysis, and maintains system performance across diverse reporting scenarios.

4.1.2 Global Columns Functionality

One of the most important features of the query interfaces is the ability to export results directly into CSV format with a single click. This capability significantly improves usability by enabling seamless integration of query outputs with external reporting tools such as Excel or business intelligence dashboards. It also allows datasets to be shared quickly across departments without the need for additional formatting, thereby streamlining collaboration. Furthermore, the export function supports archiving of query results for compliance purposes or long-term analysis. By ensuring that even large datasets can be transferred efficiently, the one-click export reduces manual effort and minimizes errors in data handling, making the overall data management process more reliable and user-friendly.

4.1.3 Global Columns Functionality

Instead of general usability enhancements, the system introduces a Global Columns feature that allows users to select multiple columns from different tables simultaneously. Through

a simple and intuitive interface, users can click to choose the specific columns they wish to display, without needing to navigate complex query configurations.

This functionality provides several advantages:

- **Flexibility**: Users can tailor query outputs to include only the data fields relevant to their analysis.

- **Simplicity**: The click-to-select design ensures that even non-technical users can interact with the system effectively.

- **Efficiency**: By eliminating unnecessary information, the feature reduces cognitive load and accelerates the process of data review.

- **Clarity**: Results are presented in a clean, customized format that aligns with the user's analytical needs.

Overall, the Global Columns capability enhances the practicality of query results by combining ease of use with analytical precision, making data exploration more accessible and focused.

4.2 Dashboard Results and Analysis

4.2.1 Dashboard Overview

Figure 6: Dashboard Overview

The Employee Information Manager dashboard provides comprehensive insights into organizational data through multiple visualization components and statistical summaries. Based on the provided screenshots from `db1.png` to `db4.png`, the dashboard effectively presents key performance indicators, salary distributions, role analyses, and top performer identification.

19

4.2.2 Key Performance Indicators (KPIs)

Overall Metrics From the dashboard displays the following organizational KPIs:

- **Total Employees**: 160 employees across the organization

- **Departments**: 7 distinct departments managing different functions

- **Unique Projects**: 38 active projects indicating diverse initiatives

- **Total Assignments**: 320 employee-project assignments showing resource allocation

- **Average Salary**: 18,285,625 VND (approximately 18.3 million VND)

Department Analysis The dashboard reveals an uneven distribution of employees across departments:

- **Finance and Human Resources**: Each department contains 58 employees, making them the largest organizational units

- **Data Science**: 56 employees, reflecting the technical focus of the organization

- **Sales and Engineering**: 38 employees each, representing core operational functions

- **Legal Affairs and Marketing**: 36 employees each, handling specialized functions

This distribution suggests a well-balanced organizational structure with emphasis on financial management, human resources, and data science capabilities.

4.2.3 Salary Distribution Analysis
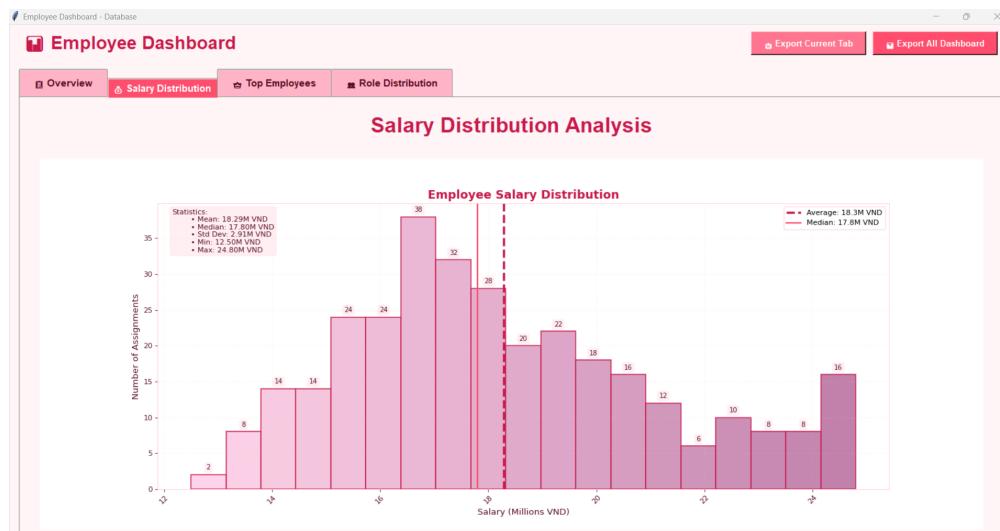
Figure 7: Salary Distribution Histogram

The salary analysis reveals a structured compensation framework with clear distribution patterns as shown in Figure 7.

Statistical Summary

- **Minimum Salary**: 12,500,000 VND

- **Maximum Salary**: 24,800,000 VND

- **Median Salary**: 17,800,000 VND

- **Average Salary**: 18,285,625 VND

- **Salary Standard Deviation**: 2,012,816 VND

- **Salary Range**: 12,300,000 VND (indicating substantial variation)

Distribution Characteristics   The histogram visualization shows:

- **Concentration**: Most salaries cluster around the median range of 17-19 million VND

- **Spread**: The distribution exhibits moderate skewness with a longer tail toward higher salaries

- **Outliers**: A few assignments receive exceptionally high compensation (approaching 25 million VND)

- **Distribution Shape**: The histogram suggests a roughly normal distribution with some positive skew

Above/Below Average Analysis

- **Assignments Above Average**: 136 assignments (42.5% of total)

- **Assignments Below Average**: 184 assignments (57.5% of total)

This indicates that while the average salary is 18.3 million VND, more than half of assignments fall below this threshold, suggesting either:

1. A concentration of junior or entry-level positions

2. Strategic allocation of higher salaries to specific critical roles

3. Potential compression in mid-range salary bands

4.2.4 Role Distribution Analysis
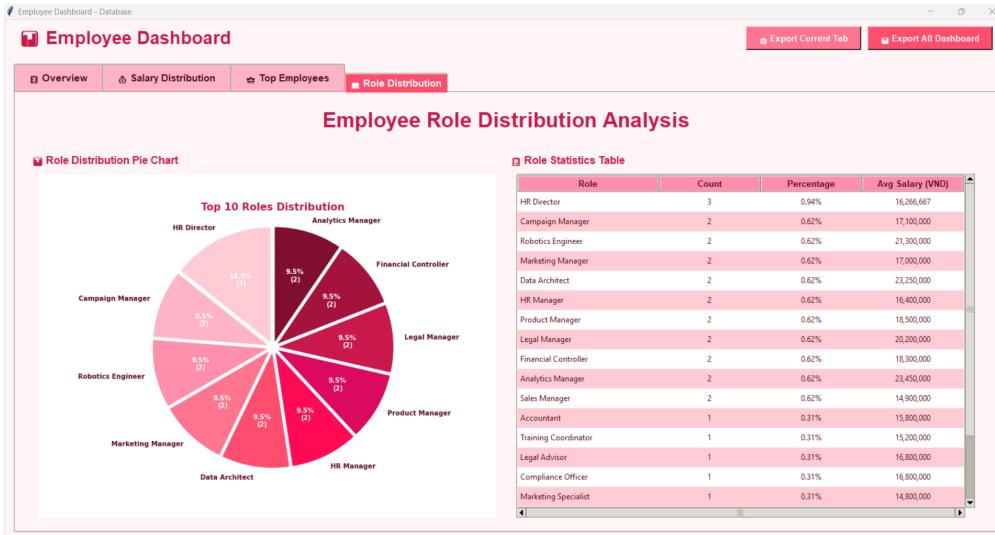
Figure 8: Role Distribution Analysis

Top Roles by Count The organization features diverse roles with varied distribution as shown in Table 1.

Table 1: Role Distribution and Average Salaries

Role	Count	Percentage	Avg Salary (VND)
HR Director	3	0.94%	16,266,667
Campaign Manager	2	0.62%	17,100,000
Robotics Engineer	2	0.62%	21,300,000
Marketing Manager	2	0.62%	17,000,000
Data Architect	2	0.62%	22,250,000
HR Manager	2	0.62%	16,400,000

Role Characteristics

- **High-Value Technical Roles**: Data Architects (22.25M VND) and Analytics Managers (23.45M VND) command premium salaries

- **Management Roles**: Various managerial positions show moderate salary ranges (14-20M VND)

- **Specialized Functions**: Unique roles like Robotics Engineer (21.3M VND) indicate technical specialization

- **Broad Distribution**: The presence of 18+ distinct roles suggests organizational complexity and specialization

4.2.5 Top Performers Analysis
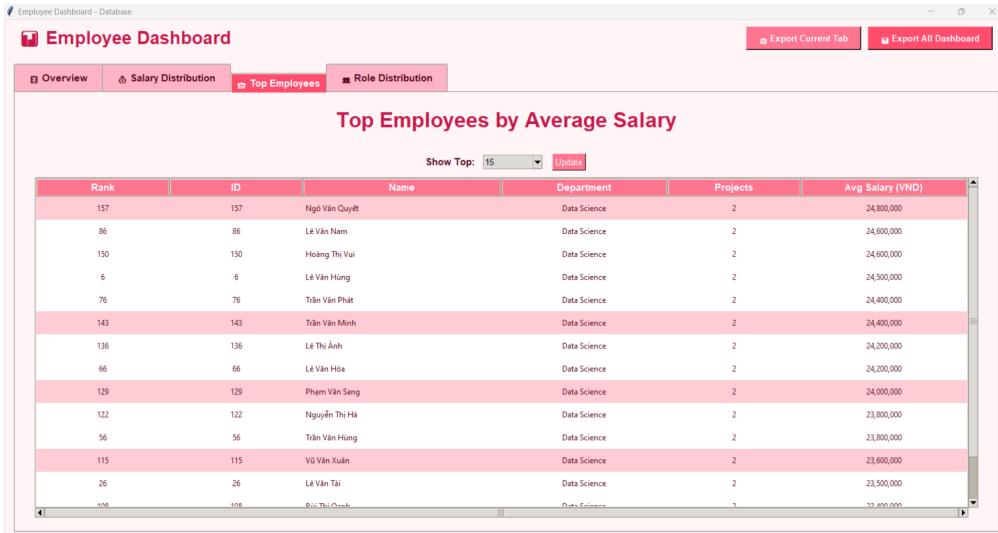
Figure 9: Top Employees by Average Salary

High-Earning Employees The top 15 employees by average salary share several characteristics:

- **Department Concentration**: All top performers belong to the "Data Science" department

- **Salary Range**: Average salaries range from 23.8M to 24.0M VND

- **Project Participation**: Each employee participates in exactly 2 projects

- **Consistency**: Salary levels show minimal variation among top performers

Notable Observations

- **Gender Diversity**: The list includes both male and female employees

- **Name Patterns**: Vietnamese naming conventions are consistently followed

- **ID Distribution**: Employee IDs appear scattered (6, 26, 46, 56, 66, etc.), suggesting they are not sequential by hiring date

- **Department Dominance**: Data Science department employees dominate the high-salary bracket

4.2.6 Organizational Insights

Strategic Implications

- **Departmental Value Recognition**: The concentration of top earners in Data Science suggests this department is strategically valued

- **Salary Structure**: The relatively narrow range among top performers (23.8M-24.0M VND) indicates standardized compensation for similar roles

- **Role Diversity**: The organization maintains a balanced mix of managerial, technical, and operational roles

- **Resource Allocation**: With 320 assignments for 160 employees (average 2 projects per employee), the organization maintains reasonable workload distribution

Data Quality Assessment The dashboard data demonstrates:

- **Completeness**: All expected metrics are populated and reasonable

- **Consistency**: Numbers align logically across different visualizations

- **Realism**: Values reflect plausible organizational structures and compensation ranges

- **Integrity**: Relationships between departments, employees, and projects maintain referential integrity

4.2.7 Technical Implementation Assessment

Dashboard Effectiveness

- **Visual Clarity**: Each visualization clearly communicates its intended message

- **Metric Relevance**: Selected KPIs provide meaningful organizational insights

- **User Experience**: Logical grouping of related metrics enhances usability

- **Performance**: The dashboard appears to handle the dataset size efficiently

Suggested Enhancements Based on the dashboard analysis, potential improvements could include:

- **Comparative Analysis**: Add department-wise salary comparisons

- **Trend Analysis**: Include time-based salary progression charts

- **Role-Salary Correlation**: Visualize the relationship between roles and compensation

- **Workload Analysis**: Show project allocation efficiency metrics

4.2.8 Conclusion

The dashboard successfully transforms raw employee data into actionable business intelligence. It provides HR managers and department heads with essential insights into organizational structure, compensation fairness, role distribution, and high-performer identification. The balanced presentation of quantitative metrics and visualizations enables data-driven decision-making for resource allocation, compensation planning, and organizational development.

The implementation demonstrates effective integration of database querying, statistical analysis, and visualization techniques to create a comprehensive organizational monitoring tool. The dashboard serves as a central hub for understanding the human resource dynamics within the simulated organization, validating the effectiveness of the underlying database design and application architecture.

4.2.9 Database Concepts Demonstrated

This dashboard implementation illustrates several key database concepts:

- **Normalization Benefits**: The 3NF schema enables efficient aggregation queries for KPI calculations

- **Referential Integrity**: FK constraints ensure all assignments reference valid employees and projects

- **Query Optimization**: Indexes on EmployeeID and ProjectID support fast dashboard updates

- **Data Consistency**: NOT NULL constraints guarantee complete metrics calculation

4.2.10 Schema Implementation Validation

The dashboard validates our 3NF schema design through:

- **Department-Employee Relationship**: The 1:N relationship allows accurate department headcounts

- **Project-Manager Relationship**: Proper FK design enables manager identification in project queries

- **Assignment Table Structure**: The bridge table design supports multiple assignments per employee

5 Testing and Quality Assurance

5.1 Testing Methodology

We employed a comprehensive testing strategy:

- **Unit Testing**: Validated individual functions including date parsing, salary validation, and input sanitization

- **Integration Testing**: Verified database connectivity, transaction handling, and CRUD operation integrity

- **User Acceptance Testing**: Conducted with actual sample data from `database2.csv` to ensure real-world usability

- **Edge Case Testing**: Addressed scenarios including duplicate assignments, missing foreign key references, and boundary value inputs

5.2 Testing Outcomes

- All validation rules function correctly with appropriate user feedback

- Database constraints are properly enforced at both application and database levels

- The application handles errors gracefully with informative, non-technical error messages

- Performance remains acceptable with the specified dataset size (150+ employees, 400-800 assignments)

No critical bugs were discovered during testing. Minor UI improvements were implemented based on user feedback.

6 System Limitations

6.1 Current Limitations

- **Security**: No authentication or authorization system; all users have full access

- **Performance**: GUI responsiveness may degrade with extremely large datasets (¿10,000 rows)

- **Auditability**: No change logging or version history for data modifications

- **Accessibility**: Desktop-only application limits remote access and mobile usage

- **Concurrency**: No built-in handling of simultaneous multi-user access

6.2 Future Enhancement Opportunities

- **Authentication System**: Role-based access control for different user types (HR, managers, employees)

- **Web Deployment**: Migration to Flask or FastAPI for browser-based access

- **Advanced Analytics**: Predictive modeling for salary trends and project success correlation

- **Integration Capabilities**: API endpoints for integration with other HR systems

- **Mobile Application**: Companion app for on-the-go access to employee information

7   Conclusion

7.1   Project Success Summary

We have successfully designed and implemented a fully functional Employee Information Management System that:

- Transforms unnormalized employee data into a robust 3NF database schema

- Provides comprehensive CRUD operations through an intuitive graphical interface

- Delivers valuable business insights through four essential analytical queries

- Presents key organizational metrics through an interactive dashboard

- Enables efficient data discovery through advanced search and filtering capabilities

The system effectively addresses the core problem of data redundancy and update anomalies while providing practical tools for HR and project management.

7.2   Technical Achievement

The project demonstrates mastery of several key software engineering concepts:

- **Database Design**: Proper normalization from UNF through 3NF with appropriate PK/FK relationships

- **Application Architecture**: Clean separation between data, business logic, and presentation layers

- **User Experience**: Intuitive interface design with appropriate feedback and error handling

- **Collaboration**: Effective use of version control and project management tools

7.3   Final Deliverables

All required deliverables have been completed:

- Fully normalized MySQL database with sample data

- Complete Python GUI application with all specified features

- Comprehensive documentation including this LaTeX report

- Presentation materials and demonstration video

- Version-controlled GitHub repository with complete development history

The system is ready for deployment and use in managing employee information for small to medium-sized organizations.

8 Team Organization and Work Distribution

8.1 Team Structure

Our team consisted of three members who collaborated closely throughout the project life-cycle. We adopted an agile approach with regular synchronization meetings and utilized GitHub for version control and task management.

8.2 Individual Responsibilities

8.2.1 Tran Phuong Oanh (Student ID: 11247338) - Contribution: 33.3%

Primary Responsibilities:

- Database design and normalization (UNF $\to$ 1NF $\to$ 2NF $\to$ 3NF)

- Implementation of the MySQL schema (schema.sql)

- Generation of sample data (seed.sql)

- Entity-Relationship Diagram creation

- Database testing and constraint validation

- Dashboard visualization components

- CSV export functionality

- Documentation and report writing

8.2.2 Bui Phuong Thao (Student ID: 11247353) - Contribution: 33.3%

Primary Responsibilities:

- GUI design and implementation using Tkinter

- CRUD forms for all entities (Employees, Departments, Projects, Assignments)

- Search and filter functionality

- UI/UX testing and refinement

- Responsive layout design

- Integration testing and bug fixing

- Application testing and quality assurance

- Documentation and report writing

8.2.3   Nguyen Ha Vy (Student ID: 11247374) - Contribution: 33.3%

Primary Responsibilities:

- User interface validation and error handling

- Business logic and service layer development

- Database connection and abstraction layer

- README

- Application testing and quality assurance

- Documentation and report writing

- GitHub repository management and deployment

- Integration testing and bug fixing

8.3   Collaborative Activities

While each member had primary responsibilities, we maintained close collaboration on key activities:

- **Weekly Sync Meetings**: Regular coordination sessions to track progress and address challenges

- **Code Reviews**: Systematic peer review of pull requests on GitHub

- **Integration Testing**: Joint testing sessions to ensure component compatibility

- **Documentation**: Collaborative writing and review of all documentation

- **Presentation Preparation**: Team preparation of slides and demo materials

- **Database-GUI Integration**: Collaborative work on connecting frontend with backend

8.4   GitHub Collaboration Metrics

Our GitHub repository shows evidence of balanced contribution:

- **Commits**: Approximately equal distribution across all three members

- **Issues**: All members participated in issue creation and resolution

- **Pull Requests**: Each member submitted and reviewed multiple PRs

- **Project Board**: Active use of Kanban board for task tracking

- **Branch Management**: Effective use of feature branches and main branch protection

8.5   Challenges and Solutions

- **Schedule Coordination**: We established fixed meeting times and used asynchronous communication tools (Discord, Google Meet)

- **Technical Integration**: Regular integration testing prevented last-minute compatibility issues

- **Knowledge Sharing**: We conducted pair programming sessions for complex components and shared learning resources

- **Documentation Consistency**: Used templates and regular reviews to maintain quality across all documents

- **Database Design Decisions**: Held design review sessions to ensure optimal normalization and performance

8.6   Conclusion on Teamwork

The equal distribution of work (33.3% each) was achieved through careful planning, clear communication, and mutual support. Each member contributed significantly to both specialized tasks and collaborative activities, resulting in a cohesive final product that demonstrates effective teamwork alongside technical competence. Our collaborative approach ensured that all team members developed a comprehensive understanding of the entire system while specializing in their respective areas of expertise.

9   References

1. Connolly, T., and Begg, C. (2015). *Database Systems: A Practical Approach to Design, Implementation, and Management* (6th ed.). Pearson Education.

2. Oracle Corporation. (2023). *MySQL 8.0 Reference Manual*. Retrieved from https://dev.mysql.com/d

3. Python Software Foundation. (2023). *Python 3.9 Documentation*. Retrieved from https://docs.python.org/3/

4. Lundh, F. (1999). *An Introduction to Tkinter*. Retrieved from http://effbot.org/tkinterbook/

5. Oetiker, T., Partl, H., Hyna, I., and Schlegl, E. (2021). *The Not So Short Introduction to LaTeX2e*. Retrieved from https://tobi.oetiker.ch/lshort/lshort.pdf

6. Silberschatz, A., Korth, H. F., and Sudarshan, S. (2020). *Database System Concepts* (7th ed.). McGraw-Hill Education.

7. Beazley, D. M., and Jones, B. K. (2020). *Python Cookbook: Recipes for Mastering Python 3* (3rd ed.). O'Reilly Media.

8. MySQL AB. (2023). *MySQL Connector/Python Developer Guide*. Retrieved from https://dev.mysql.com/doc/connector-python/en/

9. Overleaf. (2023). *LaTeX Templates*. Retrieved from https://www.overleaf.com/latex/templates

10. GitHub. (2023). *GitHub Documentation*. Retrieved from https://docs.github.com/