Cover

| Created → Jun 16, 2025 | Last update → Jul 4, 2025 |
| --- | --- |
| Examination System  **Under the Supervision of**  Eng. Rami Nagi | |
| **Team Members**  Nada Hosny Ibrahim  Mohamed Ahmed Ammar  Mahmoud Saied Elbaal  Sabry Yahia Ahmed  **Power BI Track | Tanta Branch**  **Intake 24 -25** | |

Intro and objectives

| Examination System | |
| --- | --- |

| 01 |
| --- |

### Introduction

#### **Purpose**

The **Examination System** project is designed to provide a unified platform for managing student academic records, assessments, certifications, and employment readiness. The roadmap outlines each critical step required to build a robust, scalable system for educational institutions and HR departments alike, ensuring data consistency, analytical capability, and real-time insights across the entire student lifecycle.

This roadmap covers the entire Smart Examination System: **OLTP design**, **Data Warehouse & HR Data Mart**, **SSRS reports**, **Power BI dashboards**, and a **Streamlit-based BI chatbot** for querying the data warehouse using natural language in English and Arabic.

It serves as a comprehensive guide for all stakeholders involved — from database architects to business analysts and Power BI developers — offering a clear path from raw data collection to actionable dashboards.

#### **Scope**

**This roadmap covers:**

* Designing the conceptual and logical data model (ERD)
* Creating a fully normalized transactional database using SQL Server
* Populating the database with realistic sample data
* Building a **Data Warehouse** using dimensional modeling
* Creating an **HR-specific Data Mart**
* Implementing **ETL processes using SSIS**
* Designing **SSRS reports** using stored procedures and deploying them in Power BI Desktop.
* Creating **Power BI dashboards** for Student, Instructor, Admin, and HR users with advanced DAX analytics for each.
* A **Streamlit-based BI chatbot** for querying the data warehouse using natural language.

#### **Importance**

This roadmap is central to achieving:

* **Data consistency**: Clean, validated, and integrated data across all departments
* **Performance insights**: From student readiness to hiring metrics
* **Data-driven decisions**: For academic improvements and HR planning
* **Scalability**: Easily adaptable to new departments, tracks, or recruitment partners

This system enables institutions to analyze student performance, manage certifications, assess job readiness, and respond to employer feedback. The inclusion of an **Arabic-English chatbot improves accessibility for non-technical users.**

By building a full analytical system on top of a stable OLTP foundation, the organization ensures not only operational efficiency but also long-term strategic visibility into student success and market alignment.

| 02 | | |
| --- | --- | --- |

### Vision & objectives

#### Vision

To empower education providers and HR teams with an integrated examination and employment intelligence platform that seamlessly bridges academic outcomes with real-world assessment and hiring needs, driven by data and automation.

#### Long-term objectives

#### **Objective 1:**

Build a centralized database that maintains student academic, course enrollment, exam, and employment records with complete referential integrity.

#### **Objective 2:**

Establish a scalable data warehouse and dedicated HR Data Mart that transforms raw data into meaningful dimensions and metrics.

#### **Objective 3:**

Deliver multi-dimensional dashboards and reports that support academic decision-making, HR recruitment alignment, and certification monitoring.

#### **Objective 4:**

A multilingual BI chatbot enhances data accessibility for non-technical users by allowing decision-makers to interact with dashboards via natural questions without needing SQL or BI tool experience.

#### Roadmap Alignment

This roadmap lays out the exact steps to reach the vision and long-term goals by:

* Designing a comprehensive ERD that captures the full academic and career lifecycle
* Constructing normalized and dimensional databases with integrity constraints
* Populating those systems with representative sample data for real testing
* Building a DWH and HR-centric data mart using SSIS to extract and transform data efficiently
* Producing visual analytics and automated reports for real-time HR and admin insights.
* Implementing structured data pipelines, designing a robust data warehouse, developing insightful dashboards, and integrating an NLP-based chatbot for conversational analytics.

It ensures that every technical decision is strategically aligned with business outcomes.

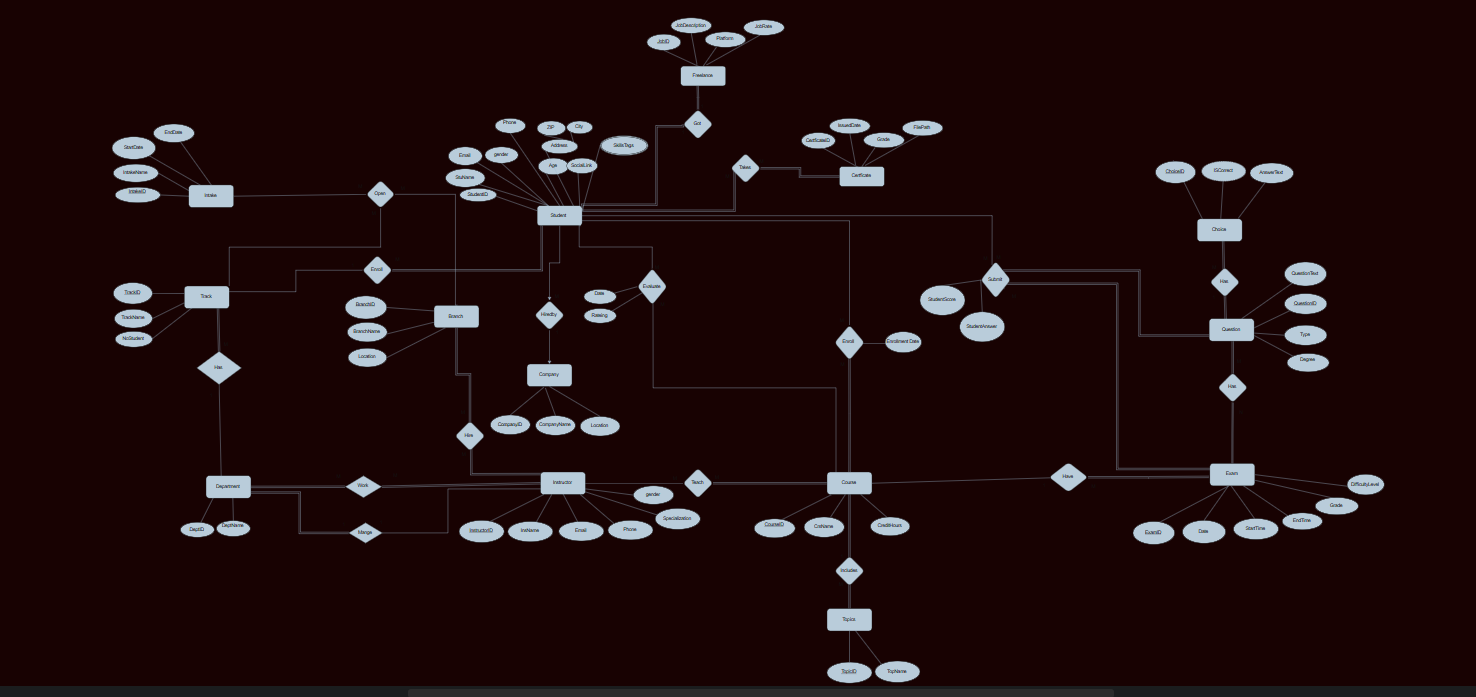
Key Initiatives

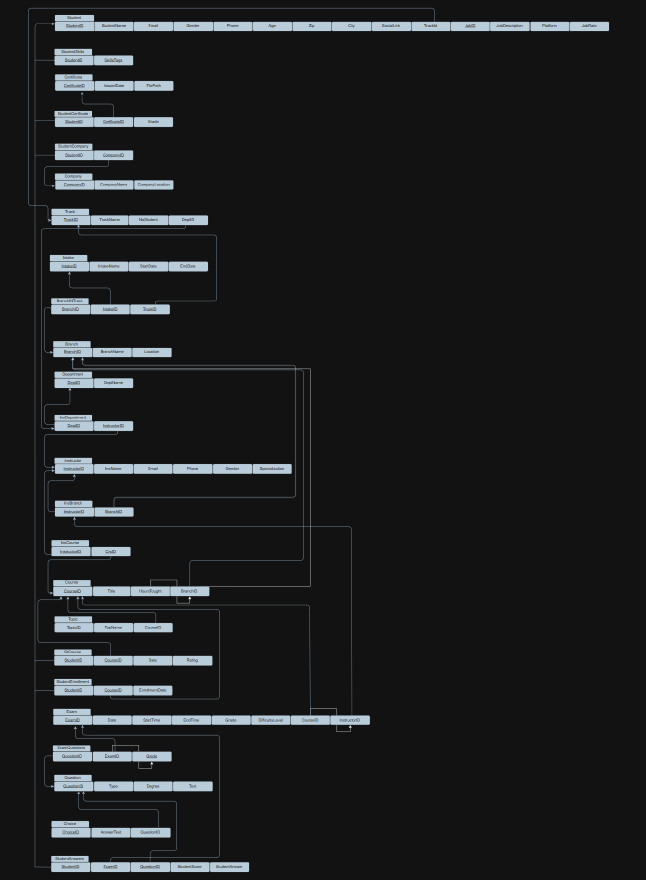
| Examination System | |
| --- | --- |

| 03 | |
| --- | --- |
| Key initiatives | |

#### **ERD**

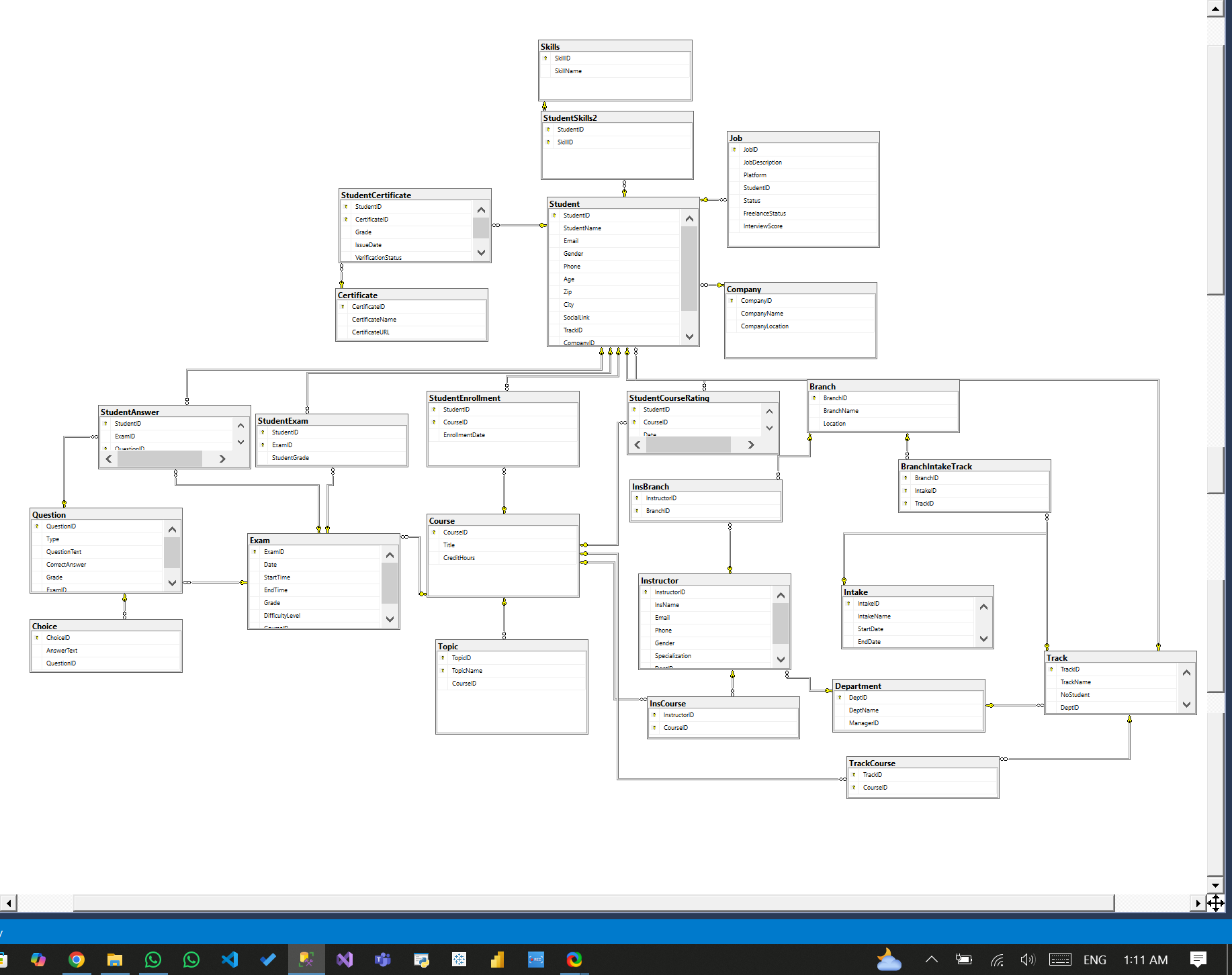
|  | |
| --- | --- |
| Description | Expected outcomes |
| * Design the logical structure of the system that includes all academic and career entities needed for student management, exam performance, instructors' performance, skill readiness, certification tracking, job placement, and Admin evaluation. | * An ERD that enables building a normalized **OLTP schema** suitable for transactional operations and later transformation into dimensional models. |
| 🛠️ Tools Used: | * **draw.io**: For creating the Entity Relationship Diagram (ERD) |
| Entity | **Description** |
| Student | Stores student info (Name, DOB, Gender, Contact, Branch, Track) |
| Track | Represents the academic track the student is enrolled in |
| Branch | Represents a physical or virtual branch location |
| Instructor | Manages instructor profile |
| Course | Represents course details (title, hours, branch) |
| Topic | Hold topic details like TopicID and Name. |
| Intake | Tracks intake session with date ranges |
| Department | Top-level organization unit over tracks |
| Exam | Represents an exam instance like ExamID, Start, and End Time |
| Question | Holds exam questions with difficulty level and grades |
| Choice | Stores student answers with correctness flags |
| Freelancing | Stores freelancing tasks and interview status |
| Company | Employer information |
| Certification | Stores certification metadata (title, topic, URL) |
| Skills | Stores student resume-related skills |
| Relationships: | One Student can have many Answers, Certifications, Skills, Placements, and skillsThe exam is linked to the Course and the InstructorA course can have many ExamsExams have many questionsThe course can be tough with many instructors  * Instructors can have a managerial level in branches.  A branch can have many intakesBranch depends on the Track |





#### **Mapping & Schema Design**

|  | |
| --- | --- |
| Description | Outcomes |
| * Define all database tables, their relationships, constraints, and logic used to implement the **Examination System** — starting from the normalized schema all the way to the HR Data Mart. | * A fully normalized, modular SQL schema is implemented. * Referential integrity is ensured across schemas using FKs and constraints. * Reusable views and joins form the basis for data warehouse loading. |
| Tools Used: | * **SQL Server Management studio** |
| Logical Relationships & Constraints | **All core relationships follow 1:N (one-to-many) or M:N (many-to-many) patterns.**   * Student → Answer → Question → Exam → Course → Track * Track → Branches → Intakes * Student ↔ Company via JobPlacement * Student ↔ Certification via StCertificationProgress * Instructor ↔ Department, Instructor ↔ Course * Branch, Intake, and Track form a junction via BranchInTrack   **All FKs were defined with:**   * ON UPDATE CASCADE * ON DELETE SET NULL or CASCADE based on use case |
| Data Validation Constraints | Examples:Gender in Student: Must be ('Male', 'Female')Score in Result: Must be between 0 and 100IsCorrect in Answer: Must be 0 or 1Difficulty Level in Question: Range between 1 and 5  * Interview score: Range between 1 and 5  JobPlacement.Status: ENUM style — ('Applied', 'Interviewed', 'Hired', 'Rejected')  * Freelance Status: ENUM style — ('Available', 'Not Available') |



## Data Population

|  | |
| --- | --- |
| Description | Outcomes |
| * Created realistic sample data to populate all the normalized tables across the scheme. This data simulates the full student journey — from intake and enrollment to exams, certifications, and job placement — enabling complete testing of the ETL process, SSRS reports, Power BI dashboards, and BI chatbot. | * All normalized tables are populated with consistent mock data. * Data supports full end-to-end testing of dashboards and reports. * Diverse patterns ensure that filters, KPIs, and drilldowns can be tested. |
| Tools Used: | * SQL Server Management studio * Python * Chat GPT |
| Data Population Strategy | * Manually insert seed data for lookup tables (Departments, Branches, Tracks) * Insert **students**, ensuring every student links to a valid branch and track * Populate **courses**, **instructors**, and course-topic relations * Generate **exams per course** with relevant questions and choices * Simulate **certification enrollments** and outcomes * Simulate **interview scores**, **job placements**, and **feedback** * Add mock **feedback**, **skills tags**, and weak **topic** analytics |
| Special Logic-Driven Inserts | Some tables required dynamic logic:Answer.IsCorrect was calculated by comparing the answer with the correct choice Result.TotalGrade was derived from SUM(Question.Grade) for all correct answers Certification.Score and CompletionStatus were varied across students to simulate real-world data Interview scores ranged from 1–5, and feedback was randomized using templates |
| Testing Coverage | The data was built to support:Over 100 studentsMultiple branches, intakes, tracks, coursesDiverse certification and job placement outcomesMock data for SSRS reports and Power BI dashboards covering:  * Student user Dashboards * Instructor User Dashboards * Admin user Dashboards * HR user Dashboards |



## Stored Procedures(SPs)

|  | |
| --- | --- |
| Description | Outcomes |
| * These procedures allow inserting**, updating, deleting, and selecting** students, enabling front-end or SSRS users to manage data easily. * Exam ops enable automating the Exam systems by generating the exam, getting its answer, and finally correcting it. * Stored procedures related to SSRS enable admins to create different reports that enable tracking of the overall operation. | * **CRUD** * InsertStudent * UpdateStudent * DeleteStudent * SelectAllStudents * **Exam Ops** * GenerateExam * GetExamAnswers * CorrectExam * **SSRS Reports** * sp\_GetGradesByStudent * sp\_GetExamQuestionsWithChoices * sp\_GetGradesByStudent * sp\_GetInstructorCoursesSummary * sp\_GetCourseTopics * sp\_GetStudentAnswersWithCorrection * **Admin** * sp\_DeleteAllData |
| Tools Used: | * SQL Server Management studio |
| CRUD Operations | * **SelectAllStudents**      * **InsertStudents**      * **UpdateStudent**      * **DeleteStudent** |
| Exam Automation | * **Generate Exam**      * **Get Exam Answer**      * **Correct Exam**      * **Insert Correction Data** |
| Report SPs | * **Getting Students' data through DepartmentNo**      * **Getting Student grades by ID**      * **Getting the instructor’s course and student count by InstructorID**      * **Getting Topics by CourseID**      * **Getting Exam Questions with Choices**      * **Getting Student Answers by Exam** |

## 

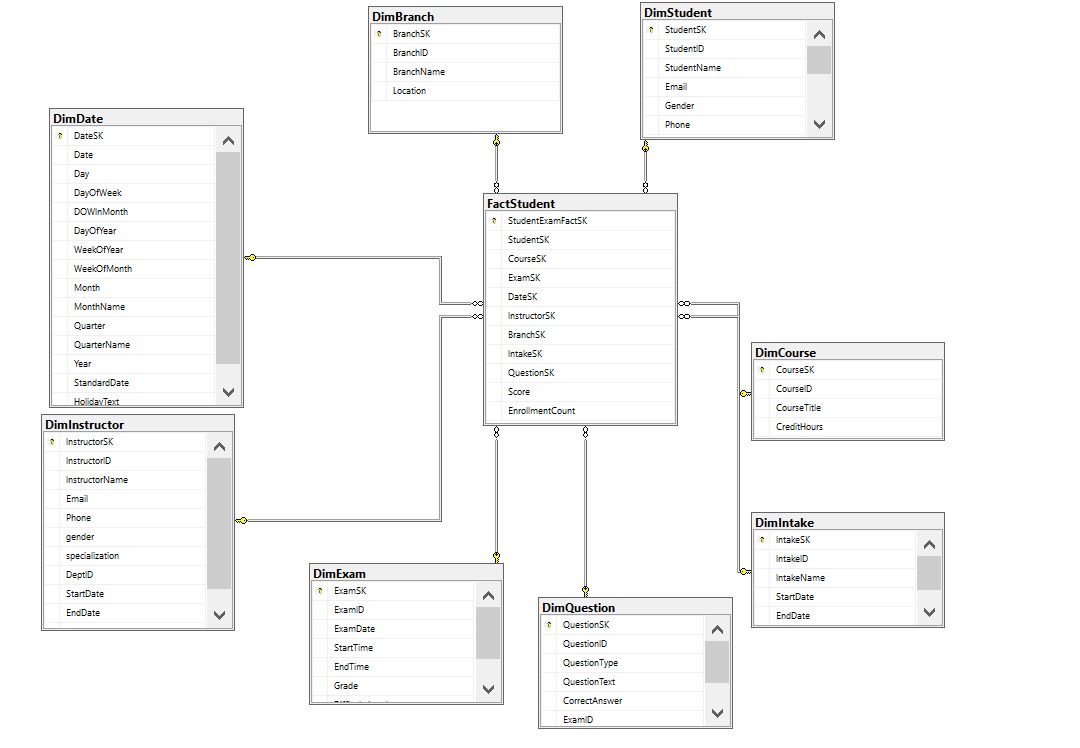
## DWH

##### Why We Built a Data Warehouse

Our transactional **Examination System** (OLTP) is highly normalized and optimized for data entry, but it's **not efficient for analytics** due to:

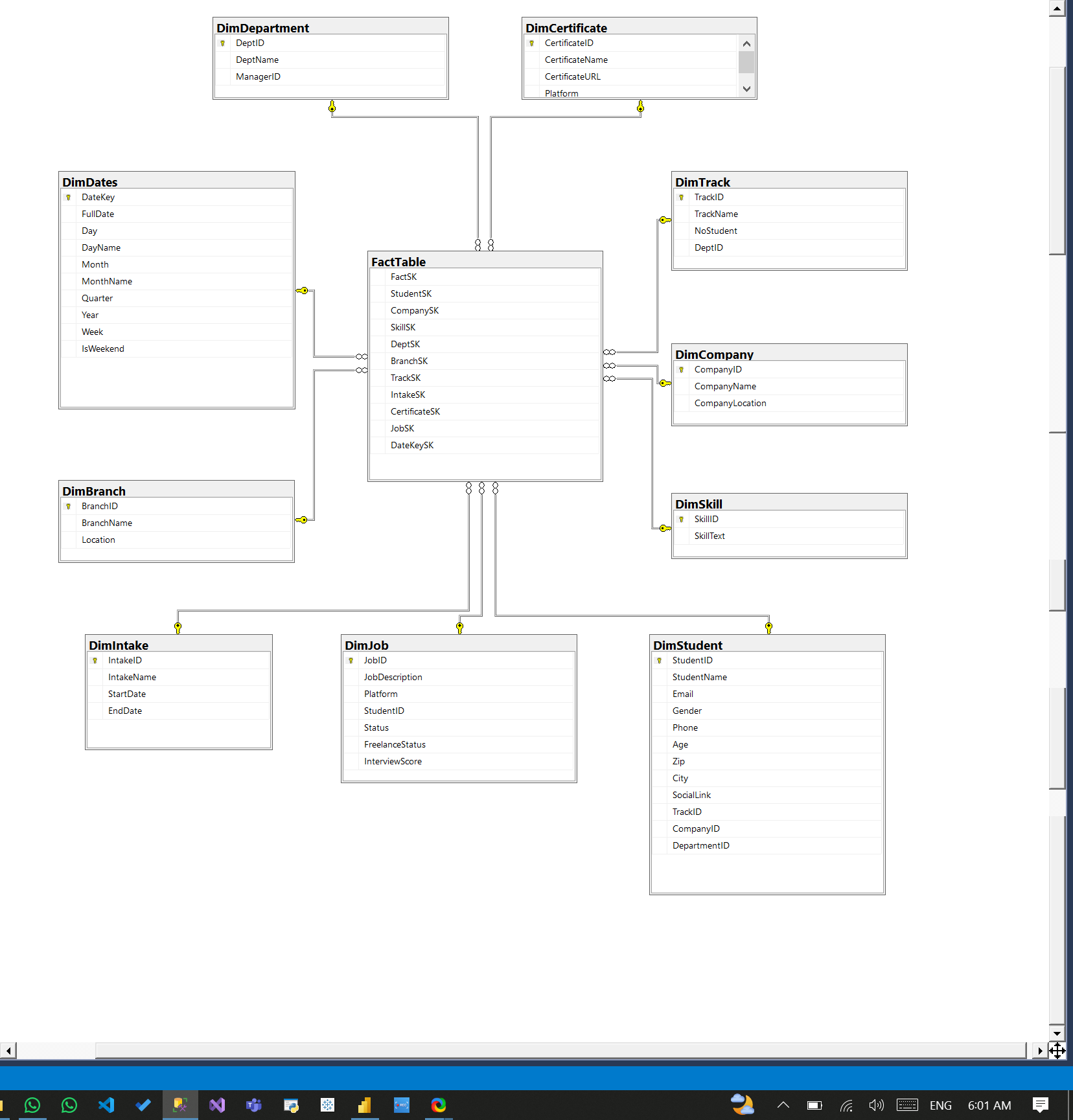
* Too many joins
* Business keys instead of surrogate keys
* No history tracking
* No aggregated metrics

|  | |
| --- | --- |
| Description | Outcomes |
| * This **Academic Data Warehouse (DW)** is designed to centralize and simplify the analysis of student performance, exam metrics, question difficulty, instructor involvement, and course-level outcomes. * It uses a **Star Schema** structure to support Power BI, BI chatbot, and SSRS with high-performance querying, intuitive data relationships, and minimized joins. | * Fully functioning star schema for **academic analytics** * Clean surrogate keys and dimensional separation * Perfectly suited for Power BI model, chatbot, and SSRS reporting * Ready for slicing by date, branch, course, or instructor |
| Tools Used | * SQL Server Management studio |
| FactStudent | **StudentFact** → SK (Primary Key) **StudentSK**  → FK to DimStudent **CourseSK**  → FK to DimCourse **ExamSK** → FK to DimExam **DateSK** → FK to DimDate  **InstructorSK** → FK to DimInstructor **BranchSK** → FK to DimBranch  **IntakeSK**  → FK to DimIntake  **QuestionSK** → FK to DimQuestion **Score** → score for question/Exam  **Enrollment Count** → How many students took this Course |
| Dimensions | **DimStudent →** Stores anonymized student info**DimCourse →** Stores course info**DimInstructor →** Stores instructor info**DimExam →** Stores exam info**DimQuestions →** Stores question info**DimIntake →** Stores intake info**DimBranch →** Stores branch info**DimDate →** Used as a universal date table to support time-based slicing in Power BI |
| Relationships | All dimension tables are linked **1-to-many** to the central **FactStudent** table via **surrogate keys**:  DimStudent.StudentSK = FactStudent.StudentSK  DimCourse.CourseSK = FactStudent.CourseSK  **This ensures:**   * No circular relationships * Clean Power BI model * Fast star-join performance |
| Typical Use Cases | **This data warehouse supports:**   * Track student exam performance per intake → Student + DimIntake * Compare course difficulty vs grades → DimCourse + DimExam + Score * Analyze instructor effectiveness → DimInstructor + Score * Time-based trends in performance → DimDate + Score * Question-level analysis → DimQuestion + CorrectAnswer |



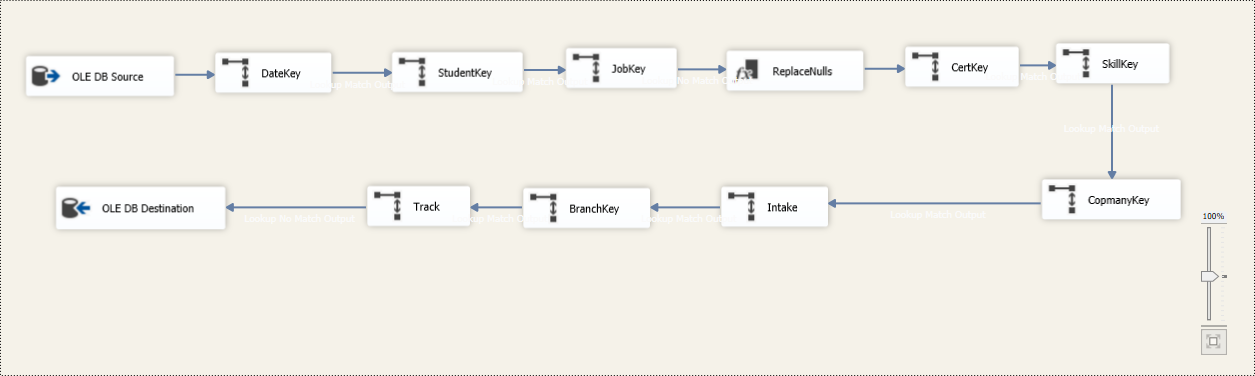
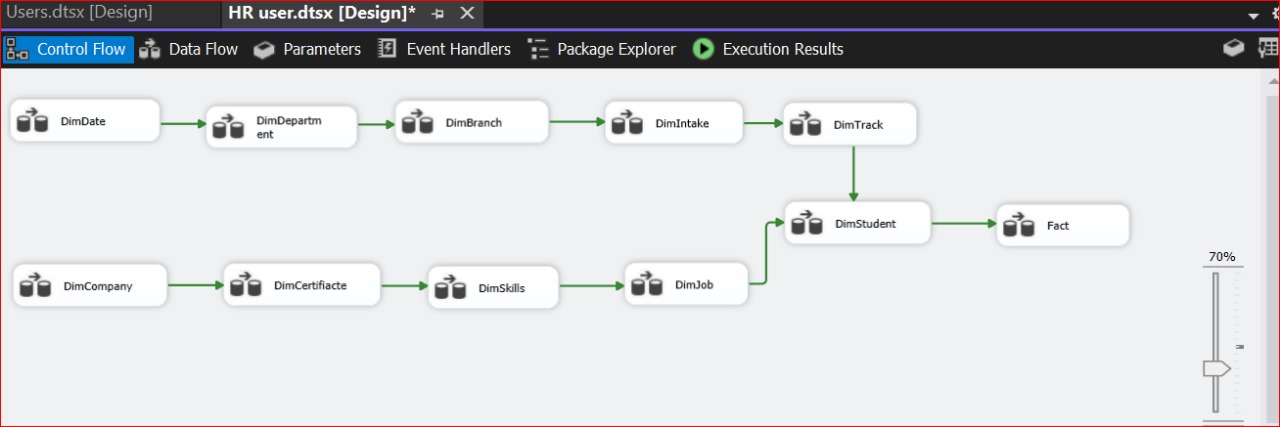
## HR Data Mart

|  | |
| --- | --- |
| Description | Outcomes |
| The **HR Data Mart** is a filtered, purpose-built subset of the main data warehouse focused specifically on **career outcomes**, **employability analytics**, and **talent tracking**. It provides HR users with the KPIs and dimensions needed to:   * Analyze student **readiness for hiring** * Track **certification completion** * Visualize **employer satisfaction** * Monitor **interview and resume performance** | * Create a centralized, denormalized **star schema** that allows fast, intuitive, and insightful reporting for: * **Performance analytics** * **Skill readiness** * **Certification tracking** * **Job placement** * **Interviewer feedback** |
| Tools Used: | * SQL Server Management studio |
| FactTable | **StudentFact** → SK (Primary Key) **StudentSK**  → FK to DimStudent **CertificateSK**  → FK to DimCertificate **SkillSK** → FK to DimSkill  **CompanySK** → FK to DimCompany  **JobSK** → FK to DimJob **BranchSK** → FK to DimBranch  **IntakeSK**  → FK to DimIntake  **TrackSK** → FK to DimTrack **DepartmentSK** → FK to DimDepartment  **DateSK** → FK to DimDate |
| Dimensions | **DimStudent →** Stores anonymized student info**DimCertificate →** Stores certificate Progressinfo**DimCompany →** Stores interview Progressinfo**DimJob →** Stores Freelance Progress info**DimSkill →** Stores Skills info**DimIntake →** Stores intake info**DimBranch →** Stores branch info**DimTrack →** Stores track info**DimDepartment →** Stores Department info**DimDate →** Used as a universal date table to support time-based slicing in Power BI |
| Relations | All dimension tables are linked **1-to-many** to the central **FactStudent** table via **surrogate keys**:  DimStudent.StudentSK = FactStudent.StudentSK  DimSkill.SkillSK =  FactStudent.SkillSK  **This ensures:**   * No circular relationships * Clean Power BI model * Fast star-join performance |
| Typical Use Cases | **This data Mart supports:**   * Student interaction rate → FactJobPlacement + DimCompany + DimJob * Top certification scores → Fact+CertificationProgress * Skill readiness score → Fact+StudentSkills * Resume keyword strength → Fact+SkillsTags * Interview weaknesses → FactJob + DimJob * Employer satisfaction → FactFeedback + DimStudent |



## SSIS

| Description | Outcomes |
| --- | --- |
| SSIS (SQL Server Integration Services) was used to build a **complete ETL process** for this project, transferring data from the transactional **Examination System** (OLTP) into both:   * The **Academic Data Warehouse** (FactStudent + related dimensions) * The **HR Data Mart** (job, skills, certs, feedback)   SSIS enabled:   * Controlled **data cleansing** * **Surrogate key** assignment * Handling **missing/malformed data** * Reusable and repeatable **data pipelines** | * SSIS provides a robust, scalable ETL pipeline * Data in the DW and HR mart is clean, consistent, and ready for Power BI * Surrogate keys and default fallbacks prevent model breaks * Logical precedence ensures no FK violations |
| Tools Used: | * Visual Studio 2022 (**SSIS**) |
| ETL Structure Overview | **Extract** → Use OLE DB Source to query raw OLTP tables  **Transform** → Use Lookup, Derived Column, Replace Nulls, Conditional Split  **Load**  → Use OLE DB Destination to insert into DW or DM tables |
| Load Order & Precedence Constraints | **To maintain referential integrity, data was loaded in this specific order:** 1. DimDate  2. DimDepartment  3. DimBranch  4. DimIntake  5. DimTrack  6. DimStudent  7. DimInstructor  8. DimCourse  9. DimExam  10. DimQuestion  11. DimCompany  12. DimCertificate  13. DimSkill  14. DimJob Each of these was implemented as a separate Data Flow Task in SSIS. |
| Key Components Used in SSIS | * **OLE DB Source**   Extracts data from: Student, exam.Exam, Job, etc.   * **Lookup** * Used to replace business keys with surrogate keys from DW dimensions. * **Example:** Replace StudentID with StudentSK from DimStudent * Handles unmatched lookups by redirecting to error outputs or setting to default (-1 or NULL). * **Replace Nulls / Derived Column** * Default values for missing gender, city, etc. * Conditional logic for derived metrics like Score%, CompletionStatus * **Data Conversion** * Used when mismatched data types between OLTP and DW exist * Especially useful for DATE ↔ DT\_DBTIMESTAMP * **OLE DB Destination**   Final step: Load into dimension or fact tables in DW |



## SSRS

| Description | Outcomes |
| --- | --- |
| SQL Server Reporting Services (SSRS) was used to deliver tabular and free-form reports to end users, especially ITI staff and educators.  These reports allow direct querying of the OLTP system and include parameters, expressions, and stored procedures for dynamic interaction.  Once complete, the reports were published to the Power BI Report Server, giving stakeholders a centralized place to access all reports. | * Over 6 SSRS reports developed using parameterized stored procedures. * All reports were deployed to Power BI Desktop for stakeholder access. * Reports support academic, administrative, and HR-related analysis. |
| Tools Used: | * Visual Studio 2022 (**SSRS**) * Report builder * Power BI |
| Reports Developed | * **Student Info by Department:**   Takes a DepartmentID parameter and shows the student's profile data |
| * **Student Grades by ID:**   Shows grades across all courses per selected StudentID |
| * **Instructor Courses Summary:**   Takes InstructorID and lists their courses + student count |
| * **Course Topics:**   Takes CourseID and returns all topics linked to that course |
| * **Exam Questions & Choices:**   Parameterized by ExamID; free-form layout showing questions and options |
| * **Exam Answers with Corrections:**   Requires both ExamID and StudentID; returns the actual answer and student's response |
| Techniques Used | **Stored Procedures:** Ensured optimized, secure, and filterable queries**Parameters:** Interactive filters for instructor, student, course, etc.**Expressions:** Conditional formatting, dynamic labels**Freeform Layouts:** Used for Exam + Answers layout with question groups**Sorting/Grouping:** Enabled in all table-based reports**Interactive Filters:** Used in dropdowns (e.g., department selector) |
| Deployment to Power BI Desktop | After testing in **Visual Studio 2022**, all SSRS **.rdl** files were imported into Power BI Desktop. |

## Dashboards

| Description | Outcomes |
| --- | --- |
| Power BI was used to create interactive dashboards for both:   * Academic analytics (e.g., student scores, exam trends, instructor performance) * HR insights (e.g., hiring, certifications, interview performance) | * +20 Dashboards completed from the student, instructor, Admin, and HR view * All DAX measures were optimized * Reusable, responsive, clean layout for business users * Highly interactive: slicing by date, intake, Branch, or track * Consistent visual identity across dashboards |
| Tools Used | Power BI Desktop |
| **Power BI Techniques Used** | * Slicer Panels: Filters by Track, Intake, Platform * Bookmarks: Show/hide Charts * Tooltip Pages: Hover details for skills, scores * DAX Measures: All KPIs were calculated using optimized DAX * Relationships: 1:M from Dimension to Fact, with single-direction filter flow |

### Student Dashboard

| Description | Report |
| --- | --- |
| **Student Performance Overview**  This dashboard presents key performance metrics for the ITI program, including total exams taken, average scores, and course ratings. It also breaks down performance by city, gender, and monthly trends. |  |
| Course Performance This dashboard highlights course-level performance and participation at ITI, showing average scores and enrollments by course and track. It also compares average scores across quarters to reveal trends over time. |  |
| Student Score This dashboard displays exam performance insights at ITI, including total and average scores, difficulty level impact, and branch-wise comparisons. It also tracks exam results over time and skill-based performance trends. |  |
| Student branch and course insights This dashboard provides a clear snapshot of student performance across ITI branches, highlighting total exams, scores, and pass rates. It also visualizes branch-wise participation and top-performing students for quick insights. |  |
| Exam Metric Overview This dashboard breaks down exam performance by course, difficulty level, and question type across 2,100 enrollments. It highlights average scores, pass rates, and course-specific insights to identify strengths and gaps in learning. |  |
| Exam Participation This dashboard tracks exam distribution and student progress across different intakes and courses. It shows total exams taken, top-performing students, and how exam participation varies by course and year. |  |

### Instructor Dashboard

| Description | Report |
| --- | --- |
| **Instructor Overview**  This report is designed to highlight key performance and demographic insights about instructors, including:   * The **total number of instructors** currently in the system * The **number of students assigned to each instructor** helping to understand teaching load distribution * A **gender breakdown of instructors** to assess diversity * The **average student score per instructor**, reflecting overall teaching effectiveness * Summary figures for **total male and female students**, as well as **total courses available** |  |
| Student Performance by Instructor This report is supposed to highlight key academic performance indicators across instructors and specializations, including:   * **Total number of participating students** * The **highest and lowest average student scores** * **Average score by specialization** (e.g., Web Development, Big Data) * **Instructor-wise performance** across different training branche * **Score trends over time** to monitor academic progress or decline |  |
| Instructor Activity This report is supposed to highlight the engagement and contribution of instructors across branches and intakes, including:   * Total number of **active vs. inactive instructors** * **Instructor distribution** by branch and specialization * Number of **training intakes covered per instructor** * **Monthly timeline** showing instructor start counts * Detailed metrics such as **total students**, **average scores**, and **fail/pass rates** by intake |  |
| Exam Overview This report is supposed to highlight exam-related insights across the training programs, including:   * Total numbers of **passed and failed students** * Number of **exams conducted** and **questions used** * Distribution of **question types per exam** * Breakdown of **exams by course title** * **Instructor pass rates** for comparative performance analysis |  |
| Instructor Deep Dive This report is supposed to highlight detailed teaching performance and student outcomes per instructor, including:   * Total **courses taught** by each instructor * **Pass rate** and **average student score** * Breakdown of **passed vs. failed students per instructor** * Instructor-specific teaching load and specialization * Timeline showing **months of instruction** and **student coverage** |  |

### Admin Dashboard

| Description | Report |
| --- | --- |
| **Exam Overview**  This report highlights essential metrics such as the number of exams taken, average scores, and students achieving full marks.  It provides a centralized view of exam activity across courses, helping the admin monitor scheduling, performance, and participation trends. |  |
| Instructor Performance Summary This report provides an at-a-glance evaluation of instructor effectiveness.  By displaying metrics like average student scores, pass rates, and teaching load, the dashboard helps identify top-performing instructors and areas needing development or support. |  |
| Course & Exam Health Summary This report assesses the overall structure and readiness of the course catalog.  It shows how many courses are active, how many are missing exams, and distributes courses by credit hours, supporting academic planning and curriculum quality monitoring. |  |
| Student Performance Snapshot This report summarizes student outcomes to give the admin insight into academic achievement levels.  It includes total student counts, score distributions, and averages to help identify high performers, struggling students, and general performance patterns. |  |
| Branch Comparison This report compares branches using custom metrics such as student enrollment, instructor quality, and course activity.  It helps the admin identify regional trends, resource allocation gaps, and opportunities for localized improvements. |  |

### Recruiters' “HR” Dashboard

| Description | Report |
| --- | --- |
| **Overview & Demographics**  This report is supposed to highlight key metrics like student counts and readiness to evaluate program performance.  It further breaks down student distribution by city, intake, and gender, aiding in understanding diversity and informing targeted strategies. |  |
| Certification Completion Tracker This report is supposed to track which students have completed, are in progress, or are lagging in certifications to support learning reinforcement and follow-ups. |  |
| Top Talent Pipeline This report is supposed to highlight high-performing students by skill coverage and certification progress to help HR identify best-fit candidates quickly. |  |
| Alumni Tracker This report is supposed to focus on interactions within the HR Tracking System.  It highlights interactions per company and platform to help evaluate engagement and identify effective recruitment channels. |  |
| Competence Coverage This report is supposed to highlight details of student skill distribution and competency over time, alongside average certification scores for various certificates.  It aims to highlight strengths and areas for development in student skill sets and program effectiveness. |  |
| Job Readiness &Drop Off This report is supposed to reveal where students drop out in the readiness funnel (from enrollment to placement) to help recover at-risk talent. |  |
| Student Resume Snapshot This report is supposed to provide a quick overview of a student's certifications, skills, and readiness score to help HR evaluate candidate profiles at a glance. |  |
| Student Profile With the ability to drill through to this specific profile for in-depth analysis.  This "Student Profile" report provides a detailed view of an individual student and key metrics. |  |

## BI Chatbot Assistant

| Description | Outcomes |
| --- | --- |
| This feature provides a multilingual natural language interface to query the Data Warehouse using Arabic or English.  It bridges the gap between non-technical HR managers and deep data insights. | * Transforms the DWH into a conversational BI experience * Empowers non-technical users to gain insight without dashboards or SQL * Lays the groundwork for voice or WhatsApp integration in the future |
| **Features & Techniques used** | * **LLM-Powered Queries:** Uses Google Gemini via LangChain to generate SQL * **Multilingual Support**: Accepts Arabic and English * **Schema-Aware Logic:** Validates user questions against the DWH schema * **Streamlit Chat UI:** Clean web-based interface for interaction * **Tech Stack:** Python + LangChain + SQLAlchemy + PyODBC + Streamlit |
| **Schema Support** | **Works with the main star schema:**   * FactStudent → holds scores, course, student, instructor, exam keys * DimStudent → student demographics and profile * DimExam → exam metadata and schedule * DimCourse → course info * DimInstructor → instructor information (name, gender, specialization, etc.) * DimBranch → training branch location and name * DimDate → calendar dates and breakdown (day, month, year) * DimQuestion → exam questions and correct answers * DimIntake → intake batch name and duration   These tables were structured with clean naming conventions to support LLM query generation. |
| **Sample Questions** | * **Arabic ⇒ ما هو متوسط درجات الطلاب في كل كورس؟** * **Arabic ⇒ كم عدد الطلاب في كل فرع؟** * **English ⇒ What is the average score by course?** * **English ⇒ How many students took exams in 2024?** |
| **Snippets** |  |

Timeline

| Examination System | |
| --- | --- |

| 04 |
| --- |

### Roadmap timeline

| Milestone |  | Initiative |  | Completion |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | | | | |
| Milestone 1 |  | ERD |  | Done |  |  |
| Milestone 1 |  | Mapping |  | Done |  |  |
| Milestone 2 |  | Data Population |  | Done |  |  |
| Milestone 3 |  | Data Warehouse |  | Done |  |  |
| Milestone 4 |  | Data Mart |  | Done |  |  |
| Milestone 5 |  | SSIS |  | Done |  |  |
| Milestone 6 |  | SSRS |  | Done |  |  |
| Milestone 7 |  | Dashboards |  | Done |  |  |
| Milestone 8 |  | BI Chatbot Assistant |  | Done |  |  |