| Capstone  project |  |
| --- | --- |

**Stage 1**

**By: Group 6**

**Table Of Contents**

[**MySQL 3**](#_heading=h.vdbqd8o1z3dn)

[**SQL 3**](#_heading=h.cv1xed70e728)

[**Python 4**](#_heading=h.bfpz9oa51zi3)

[**PowerBI 4**](#_heading=h.76fc4zmbz13u)

[**Excel 5**](#_heading=h.cnbx3jea7fg)

| MySQL |  |
| --- | --- |

**Why was MySQL used?**

* To store and organize data in a structured and queryable format.
* To merge data from multiple sources such as CSV and TSV files.
* To create relational databases that efficiently manage student and course data

**Key Operations Performed:**

* Data Cleaning: Used DISTINCT queries to remove duplicate records.
* Data Formatting: Applied DATE\_FORMAT to standardize date formats.
* Data Merging: Combined multiple data sources using UNION and JOIN.

| SQL |  |
| --- | --- |

**Why was SQL used?**

* To extract data directly from MySQL and create analytical tables.
* To generate customized queries for filtering and analyzing data based on project requirements.
* To provide Power BI and Tableau with structured, ready-to-analyze datasets.

**Key Operations Performed:**

* Data Filtering and Analysis Queries: Used GROUP BY and ORDER BY for meaningful insights.
* Performance Optimization: Implemented indexes to speed up queries and improve database efficiency.

| Python |  |
| --- | --- |

**Why was Python used?**

* Automate data processing before uploading it to MySQL and Power

BI.

* To handle attendance records by merging multiple log files into a single dataset.
* To apply data cleaning techniques and remove missing or incorrect values.

**Key Operations Performed:**

* Used Pandas library for data manipulation and analysis.
* Cleaned missing data and replaced or left them empty as needed.
* Processed large datasets and converted them into an optimized format before loading them into the database.

| PowerBI |  |
| --- | --- |

**Why was Power BI used?**

* To create interactive dashboards and analyze student and course data.
* To display key metrics such as attendance rates, completion rates, and withdrawal trends.
* To connect with MySQL and generate dynamic reports for decision-makers.

**Key Operations Performed:**

* Established relationships between tables using student\_id as the primary key.
* Analyzed data using DAX (Data Analysis Expressions) to calculate metrics like attendance and withdrawal rates.
* Created interactive reports for easy visualization and trend identification

| Excel |  |
| --- | --- |

**Why was Excel used?**

* To process data and perform initial analyses before loading it into Power BI or MySQL.
* To use Pivot Tables and Charts for quick insights into student and course data.
* To create preliminary reports and review data visually without writing SQL queries.

**Key Operations Performed:**

* Data Cleaning using features like Remove Duplicates and Find & Replace.
* Initial Data Analysis using Pivot Tables to calculate attendance, dropout, and completion rates.
* Chart Creation to visualize trends and patterns in student data.
* Exporting Data in CSV or Excel format for further processing in Power BI or MySQL.

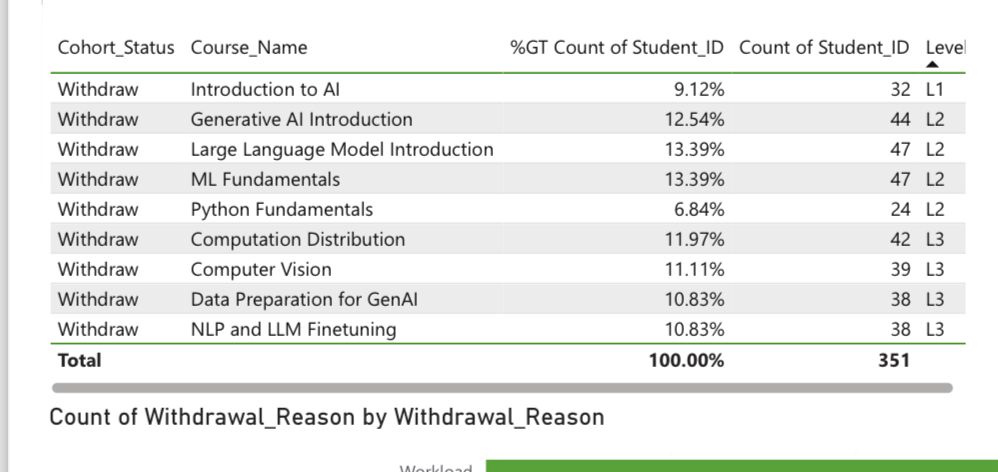


Table of Withdrawn Students Count

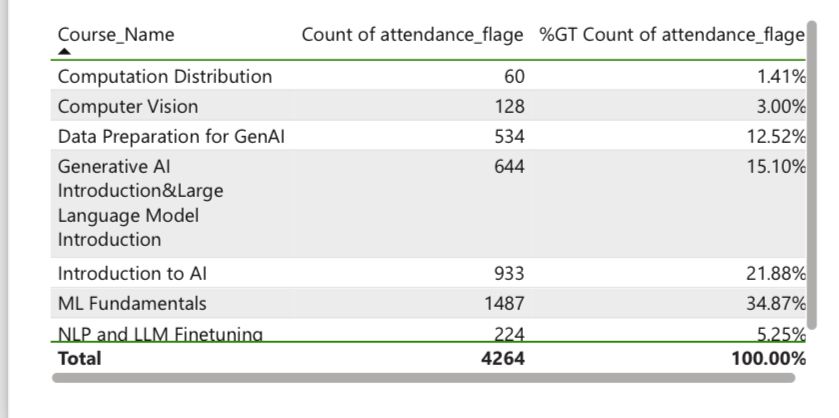


Table of Student Attendance Count

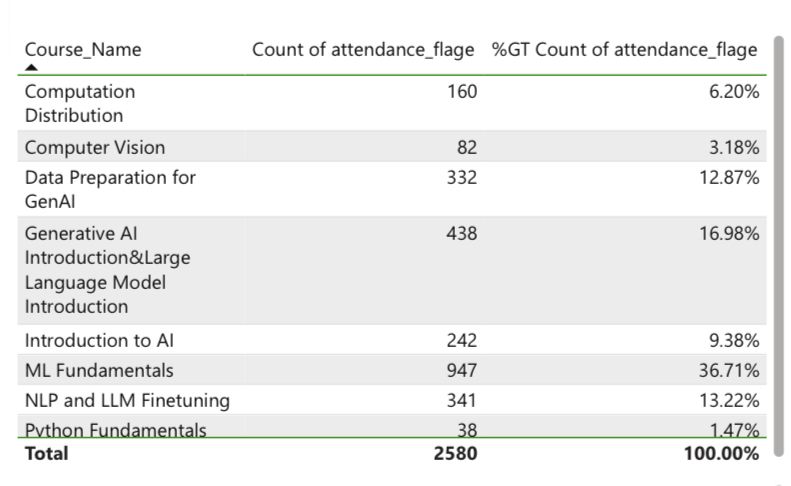
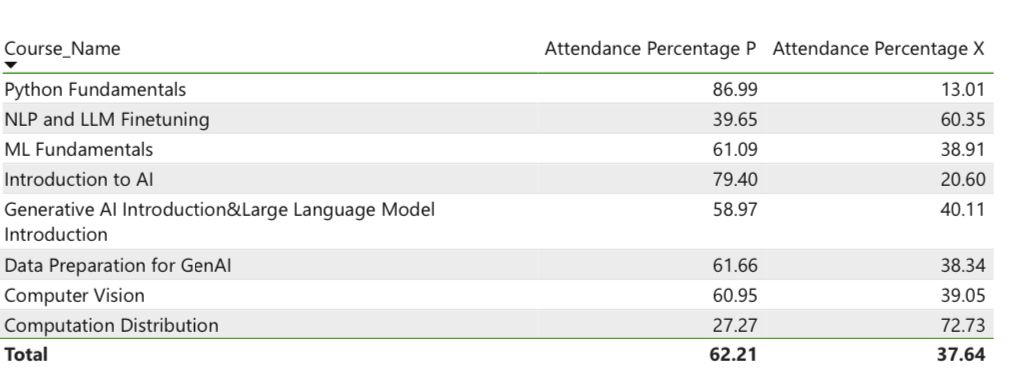
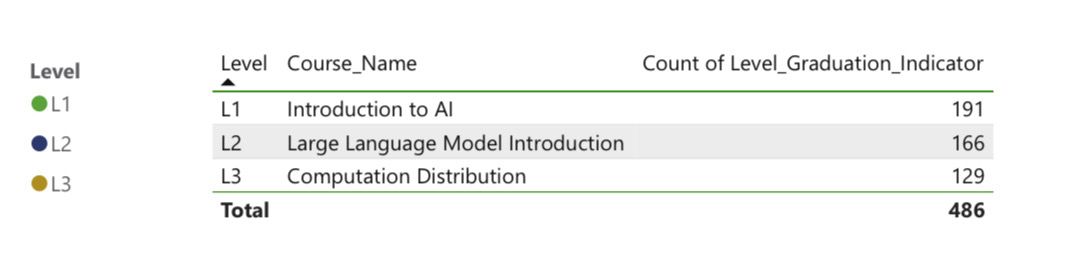


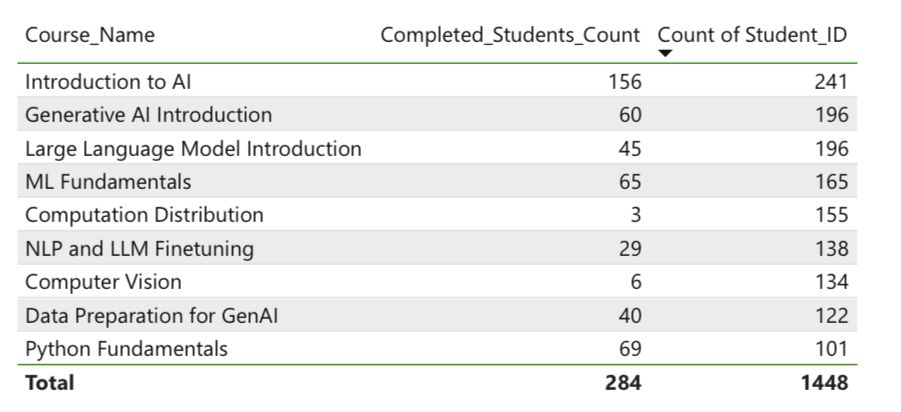
Table of Student Absence Count



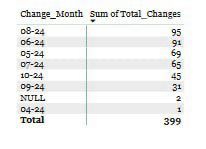
Percentage of Attendance and Absence



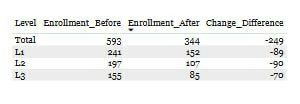
Number of Graduated Students

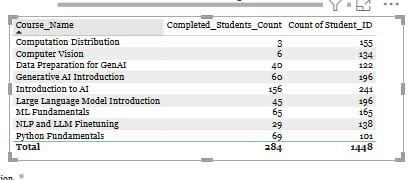


Number of Continuing Students for Each Course



The Month with the Most Change Requests





Difference in Enrollment Count Before and After Change Request by Level