

Project Report: Framework for Automated Data Injection and Dashboard Creation

Introduction and Project Objectives

Introduction

This project aimed to create a streamlined and efficient data management system for monitoring a digital literacy training program for students in grades 6th to 12th. The program focused on AI and robotics, fostering valuable skills for the future.

Project Objectives

The project objectives were to automate data ingestion and cleaning, create interactive dashboards for real-time monitoring and analysis, and generate insightful reports to track program progress and identify areas for improvement.

Data Sources and Ingestion Pipeline

1

The data sources included student demographics, attendance records, quiz scores, and feedback surveys. The data was generated using ChatGPT.

2

A Python script was developed to automate data ingestion from Excel files to MySQL Server.

3

The script ensures data integrity by validating data types and handling missing values, creating a clean and structured dataset for analysis.



Data Cleaning

The script employs a systematic approach to data cleaning by performing the following operations:

1. **Removing Duplicates:** Ensures the elimination of redundant records for streamlined data processing.
2. **Handling Missing Values:**
 - Fills missing categorical fields with default values (e.g., 'Unknown' for names, 'Absent' for attendance).
 - Imputes or marks missing numerical fields as necessary.
3. **Data Type Conversion:** Standardizes columns to appropriate data types, such as converting date columns to datetime format.
4. **Custom Cleaning Functions:**
 - Tailored cleaning processes based on file types (e.g., performance data, trainer data, attendance records, material assessments).

Dashboard Design and Deployment

1

Dashboard Design

Power BI was used to design a user-friendly and interactive dashboard that provides a comprehensive overview of the training program's performance.

2

Data Visualization

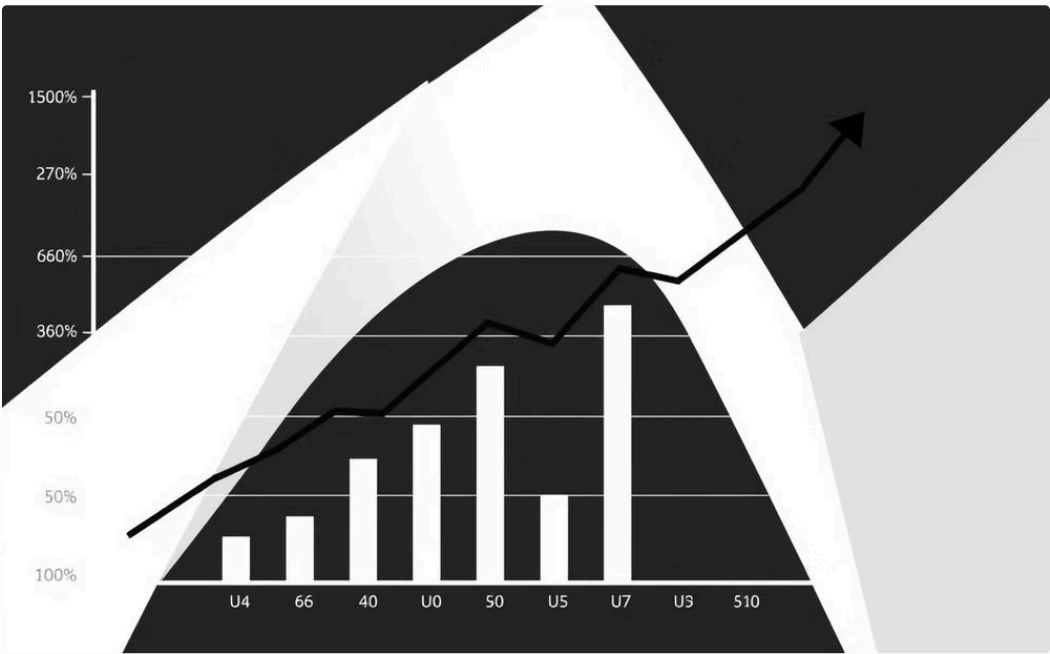
The dashboard features various visualizations, including charts, graphs, and tables, to present data in an accessible and meaningful way.

3

Dashboard Deployment

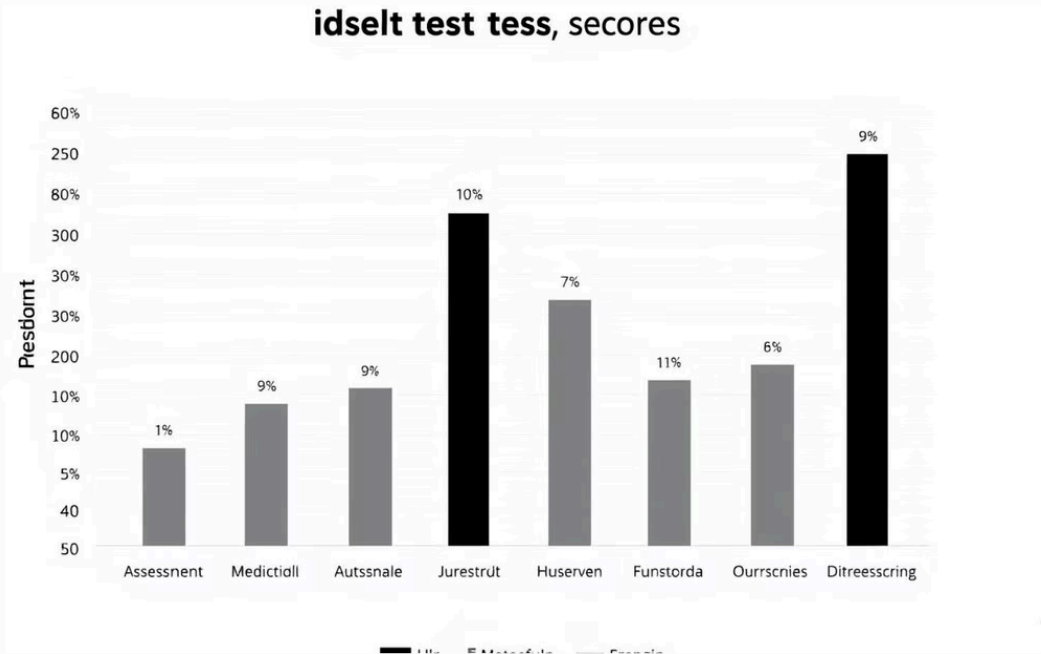
The dashboard was deployed on a secure server, allowing stakeholders to access real-time data and insights from anywhere with an internet connection.

Visualizations and KPI Tracking



Progress Tracking

The dashboard tracks key performance indicators (KPIs) such as student participation, quiz scores, and feedback ratings.



Performance Analysis

Visualizations help identify trends, patterns, and outliers in the data, providing valuable insights into program effectiveness and areas for improvement.

Real-Time Monitoring

10

Real-Time Updates

The dashboard provides real-time updates on data, allowing stakeholders to monitor program progress and make informed decisions.





Lessons Learned and Challenges Faced



Data Quality

Ensuring data quality was a key challenge, requiring careful data cleaning and validation processes.



Data Integration

Integrating data from multiple sources required a robust data ingestion pipeline and careful data mapping.



Visualization Best Practices

Choosing the right visualizations to communicate insights effectively was crucial for dashboard design.

Impact and Benefits of the Framework

1

Increased Efficiency

The framework significantly improved data management efficiency by automating data ingestion, cleaning, and injection, freeing up time for analysis and decision-making.

2

Enhanced Data Insights

Real-time monitoring and interactive dashboards provided deeper insights into program effectiveness and student performance, enabling informed decision-making.

3

Improved Program Outcomes

The framework contributed to positive program outcomes by identifying areas for improvement and facilitating data-driven decision-making.