

Student Performance Dashboard

Interactive Analytics for Educational Insights

Group 4 (DEPI Data Project)

November 2025

Team Members:

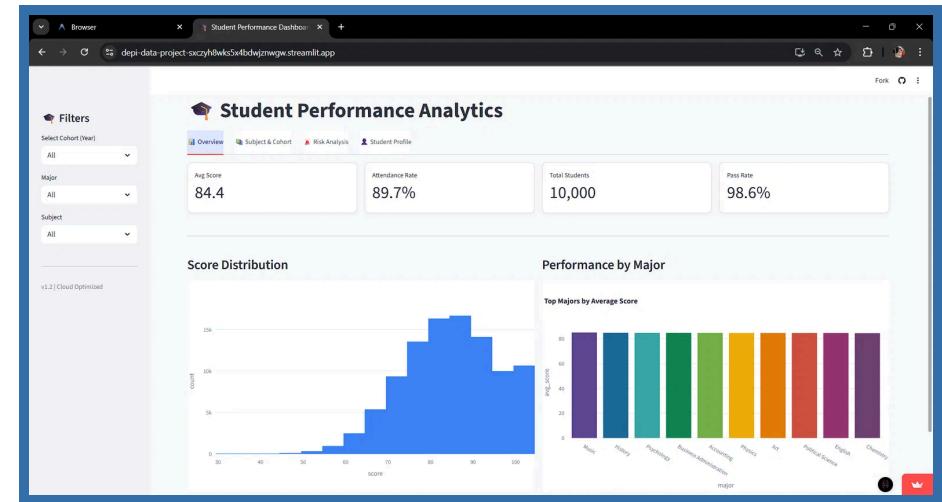
Abdulrahman Mustafa (Team Leader)

Malak Samir

Eman Wassem

Khaled Ehab

Ahmed Mohamed Galal Shoeib



Problem Statement: Bridging the Data-to-Insight Gap

⚠ The Challenge

- Educational institutions collect massive amounts of data.
- Teachers and admins struggle to find actionable insights in raw spreadsheets.
- Identifying "at-risk" students happens too late, hindering timely intervention.

💡 The Solution

- A centralized, interactive dashboard built on Streamlit.
- Real-time visualization of grades, attendance, and performance trends.
- **UVP: Early warning system** for students needing support, driven by a high-performance data stack.

Core Project Objectives: Setting the Performance Benchmark



1. Scalability

Process and analyze **1 Million+** student records efficiently.



2. Performance

Deliver insights in **< 5 seconds** using high-speed, in-memory analytics.



3. Usability

Create an intuitive interface for non-technical users, primarily teachers and administrators.

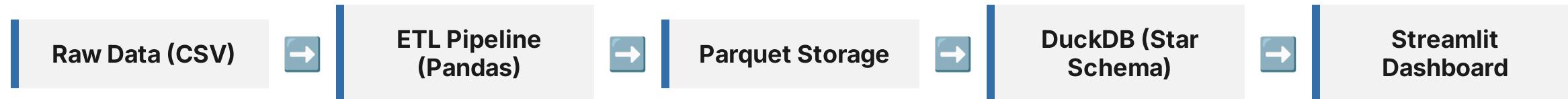


4. Accessibility

Ensure data is accessible via a robust, web-based platform (Streamlit Cloud).

Architecture & Tech Stack: Building for Scale and Speed

Data Workflow Pipeline



</> Core Languages & Databases

- **Python**– Core logic and scripting
- **DuckDB**– High-performance in-memory SQL

cube Frameworks & Tools

- **Streamlit**– Interactive web application
- **Plotly**– Dynamic and interactive charts
- **Pandas**– ETL pipeline processing

key DevOps & Deployment

- **GitHub Actions**– CI/CD for automated deployment
- **Streamlit Cloud**– Web-based hosting platform

Data Pipeline: From Raw Data to Analytics-Ready Storage



1. Extraction

Combined real IPEDS institutional data with synthetic student records.

Total Volume: **1,000,000 Students**



2. Transformation

Cleaned missing values and standardized formats. Modeled into a Star Schema:

Fact Table: Student Performance

Dimensions: Student, University, Course, Date

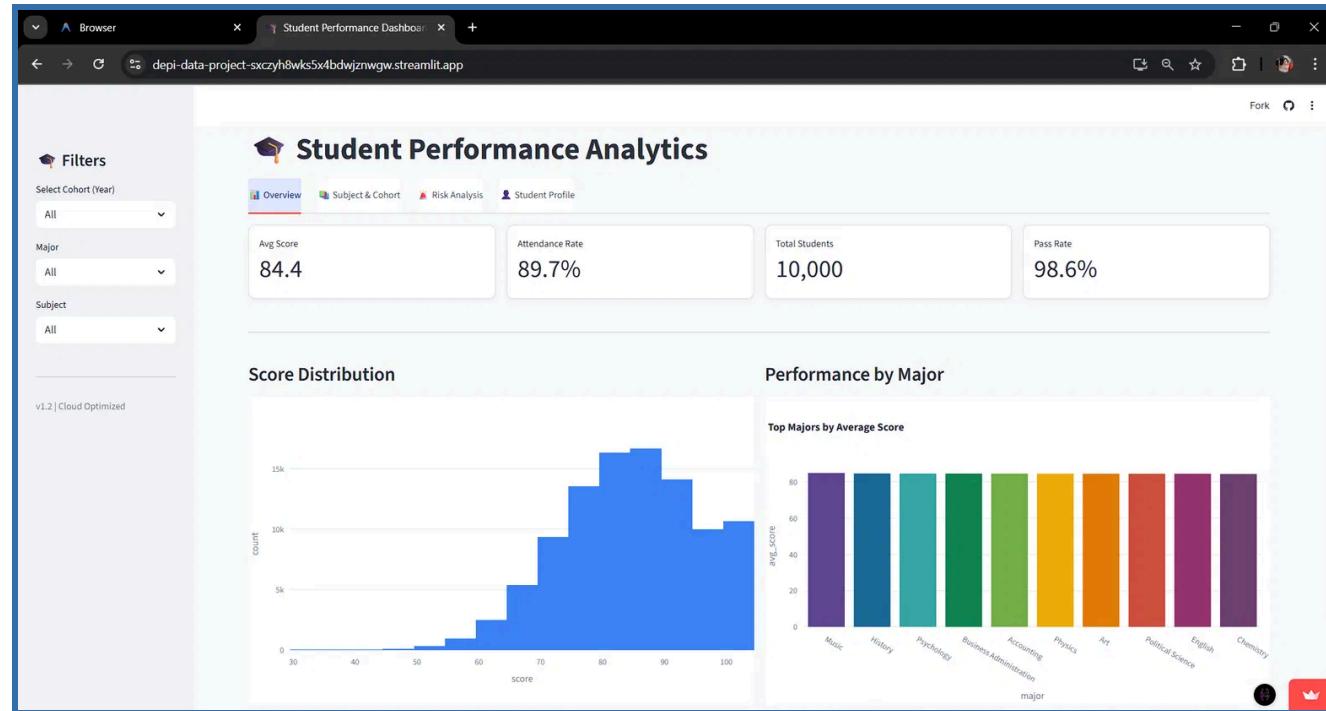


3. Loading

Saved as optimized **Parquet** files for efficient storage and querying.

Size Reduction: **90% smaller than CSV**

Key Features - Overview Dashboard



KPIs at a Glance

Instant view of **Average Score (84.4)**, **Attendance Rate (89.7%)**, and **Pass Rate (98.6%)**.



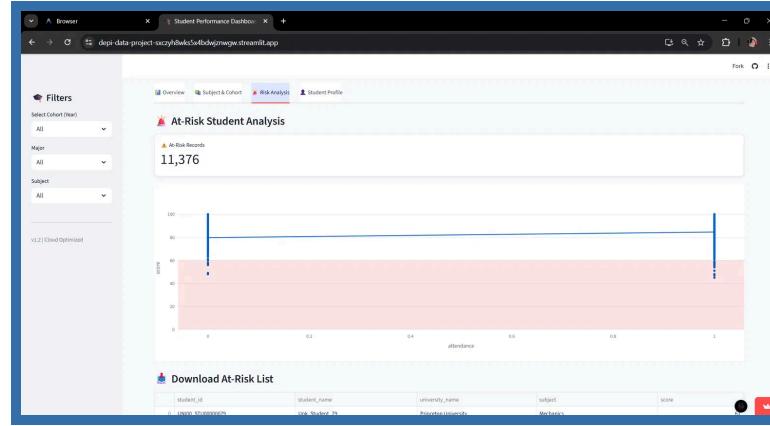
Distribution Analysis

Histograms showing the spread of grades across the entire student cohort for deeper insights.



Smart Filtering

Slice data instantly by Year, Major, or Subject to focus on specific student populations.



! Risk Analysis

- **Correlation Analysis:** Scatter plot showing the strong link between Attendance and Scores.
- **At-Risk Identification:** Automatically flags students meeting risk criteria.
- **Actionable Insights:** Teachers can download the at-risk list as CSV for immediate intervention.

At-Risk Criteria:

- Score < 60
- Attendance < 70%

Student Profile: 360° Academic View

🔍 Individual Lookup

Search for any student by ID to access their complete academic record instantly.

⌚ 360° View

See a student's full academic history, grade trends over time, and specific subject performance in one comprehensive dashboard.

🎓 Use Case

Parent-teacher conferences or academic advising sessions made efficient and data-driven.

The screenshot displays a web-based dashboard titled "Student Performance Dashboard". On the left, there is a sidebar with "Filters" for "Select Cohort (Year)", "Major", and "Subject", all currently set to "All". Below the filters is a note "v1.2 | Cloud Optimized". The main content area has tabs at the top: "Overview", "Subject & Cohort", "Risk Analysis", and "Student Profile", with "Student Profile" being the active tab. In the "Student Lookup" section, a search bar shows "77" and a status message "Student #77: Unk_Student_301048 | Major: Biology". Below this is an "Academic Summary" section showing "Avg Score: 88.1", "Attendance: 100.0%", and "Total Courses: 11". The "Course History" section contains a table with 8 rows of course data:

year	semester	subject	score	grade	attendance_flag
0	2013 Fall	Environmental Science	86	B+	
1	2013 Fall	Physiology	87	A-	
2	2013 Fall	Ecology	83	B+	
3	2013 Fall	Physics	88	A-	
4	2013 Fall	Molecular Biology	80	B	
5	2013 Fall	Philosophy	100	A+	
6	2013 Spring	Chemistry	61	D+	
7	2013 Spring	Evolution	96	A	

Challenges & Solutions: Engineering Resilience at Scale

 Challenge	 Solution
 Big Data (1M Rows)	Used DuckDB and Parquet for fast, columnar processing and efficient memory management.
 GitHub Limits (100MB)	Split large files into parts (`.part1`, `.part2`) and stitched them automatically during the ETL pipeline execution.
 Cloud Memory Constraints	Created a lightweight 10K Sample specifically optimized for the live demo on Streamlit Cloud.

Next Steps & Project Deliverables

Future Work

Predictive Models

Add Machine Learning algorithms to predict future student grades and performance trends.

Real-Time Data

Connect to live Learning Management Systems (LMS) for real-time data synchronization.

Mobile App

Optimize the dashboard layout and functionality for mobile devices and tablets.

Deliverables

Live Application

Fully functional Student Performance Dashboard deployed on Streamlit Cloud.

Source Code Repository

Complete GitHub repository with all code, configurations, and version control history.

Technical Documentation

Comprehensive README, code comments, and architecture documentation for maintainability.

Thank You!

We're excited to share our Student Performance Dashboard with you.

 Live Demo

depi-data-project-sxczyh8wks5x4bdwjznwgw.streamlit.app



Questions & Feedback?