Overview

Who We Are

We are students of Cohort 6 in the MIT Emerging Talent program. Through this program, we develop our understanding of computer science and data science concepts, as well as build our coding expertise. Currently, we are engaged in the initial phase of the Experiential Learning Opportunity (ELO) stage. This phase includes a Collaborative Data Science Project (CDSP), where we examine the complete data science lifecycle. Throughout this project, we apply our knowledge across key milestones, from problem identification to communicating results.

What We Do

As students from war-affected regions including Palestine, Ukraine, Yemen, and others, we are particularly motivated to support fellow students impacted by war and conflict. Our objective is to examine the effects of active conflict and war on students' academic performance, especially following the shift to online learning due to the closure or destruction of educational institutions. In this study, we are primarily focusing on Palestinian college students, with particular emphasis on the ongoing war in Gaza.

Problem Statement

In regions impacted by war, the educational system is often among the first societal structures to deteriorate. Students who previously demonstrated high levels of academic achievement and consistent attendance frequently experience significant declines in both areas once conflict begins. Contributing factors include violence, displacement, psychological trauma, and the breakdown of essential infrastructure such as internet connectivity and school availability. This project seeks to examine how armed conflict affects students' academic performance and attendance over time. By comparing academic records from before and during the conflict, we aim to measure the extent of disruption and identify patterns that can guide targeted interventions.

Research Question

How do disruptions caused by war including violence, displacement, and infrastructure breakdown affect students' academic performance and attendance in online learning environments among college students in war-affected regions?

Data Access and Current Limitations

We are in the process of getting data from the Islamic University of Gaza (IUG). This includes Moodle logs capturing various student interactions, with fields such as:

- User ID (can be anonymized for data privacy)
- Timestamp
- Event type (e.g., view, submit, login, forum post)
- Course information: title, description, students' level.
- Module
- Session information :Duration on page, clicks and downloads for materials and videos.
- Grade items
- Cohort or department identifiers
- Gender
- Level of study (when linked via user profiles)
- Place of residence (City or address)

What we are asking for

The data currently available to us only records students' performance in online learning after the onset of the war in Gaza. Due to technical issues and extensive damage to educational infrastructure, we lack data from before the conflict. To address this limitation, we propose using proxy data specifically, data from students in a nearby region not directly affected by war. We plan to request similar data from a university located in the West Bank. This will enable us to compare students' performance and attendance between Gaza and the West Bank. Our goal is to establish stronger causal inferences about the impact of war on academic

outcomes, rather than relying solely on correlations. We believe this comparison will significantly enhance the rigor and value of our study.

Sample Structure:

Date: Starting from the date of the first class held during the war.

Targeted Faculties: Medicine, Engineering, Information Technology.

Number of Students: Around 1,500 students of each gender.

Number of Courses per Student: 3–4 of the most important specialization courses.

Student Distribution: From all academic levels (first, second, third, fourth—graduate). All levels included for the Faculty of Medicine.

Type of Data:

- Detailed activity logs of students on the online learning platform (sessions, clicks, assignment submissions, quizzes and exams, grades), and any other details relevant to the study.
- Summary and information about each course within the target sample.
- Summary of each semester: start date, end date, and general academic calendar.

Sampling Strategy:

We aim for appropriate representation from each faculty and department, taking into account the following:

- Total number of students
- Courses registered online (remotely)
- Level of activity and engagement on Moodle
- Stratification: Faculty → Department → Year of study

If data is available for 3,000 students, a sample can be drawn to reflect the distribution of students across the three faculties, taking into consideration gender (male and female), academic levels, geographic distribution (residence), current

displacement location (if applicable), and academic performance (GPA or academic standing).

We are aware that the proportion of medical students will be lower than that of IT students.

Example:
If 60% of students are from the Faculty of Information Technology, 25% from the Faculty of Engineering, and 15% from the Faculty of Medicine, then in a sample of 3,000 students:

- 1,800 from the Faculty of Information Technology → distributed across its departments.
- 750 from the Faculty of Engineering → distributed across its departments.
- 450 from the Faculty of Medicine.