



Predicting Student Attrition Through Housing Data

Team -

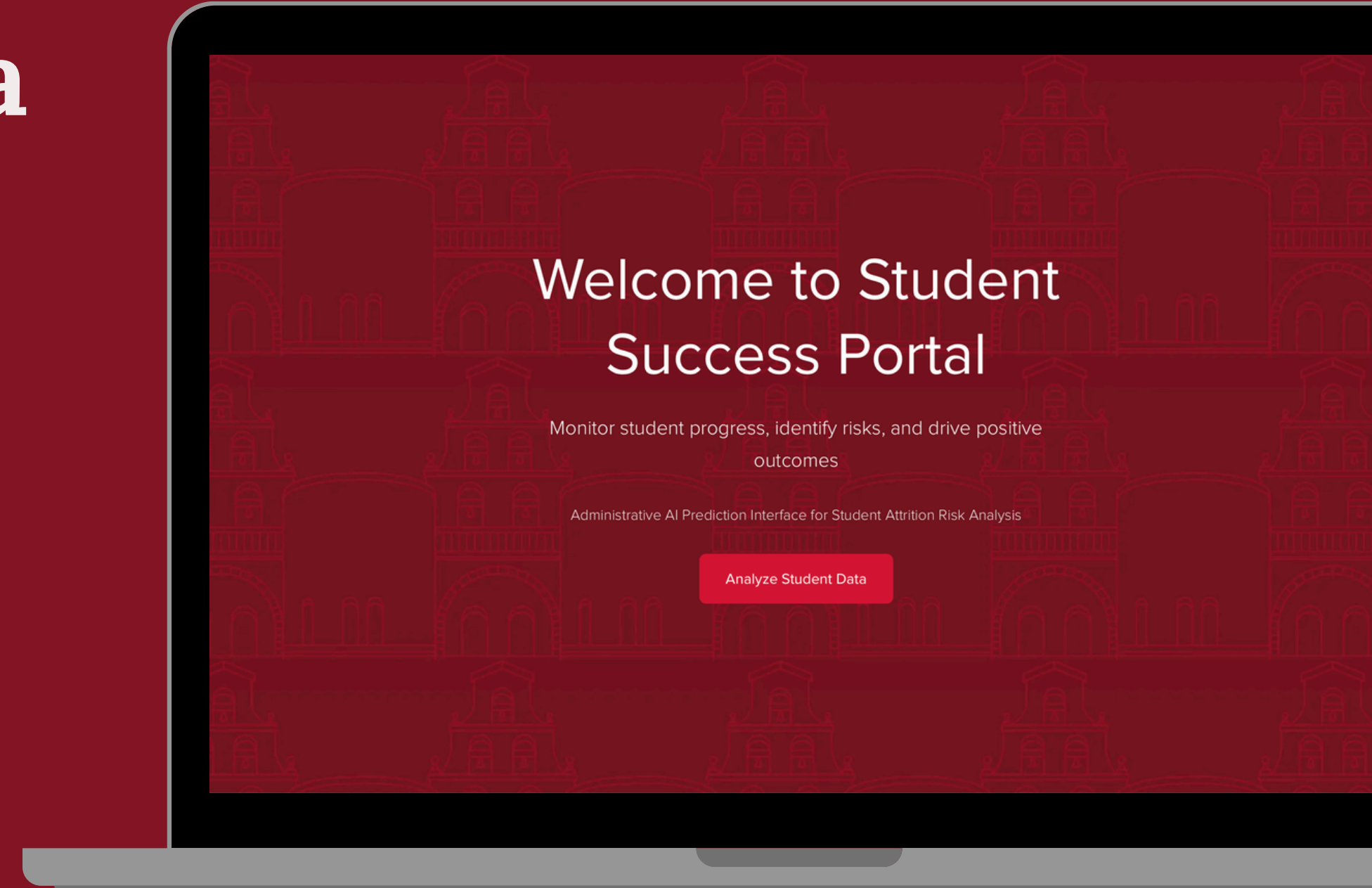
Mohith Kanthamneni

Kenny Garcia

Viridiana Delgado

Jess Alencaster

Julianna Arias



Meet the Team



Mohith Kanthamneni

Kenny Garcia



Viridiana Delgado

Jess Alencaster



Julianna Arias



The Problem

\$20,000+ 75%

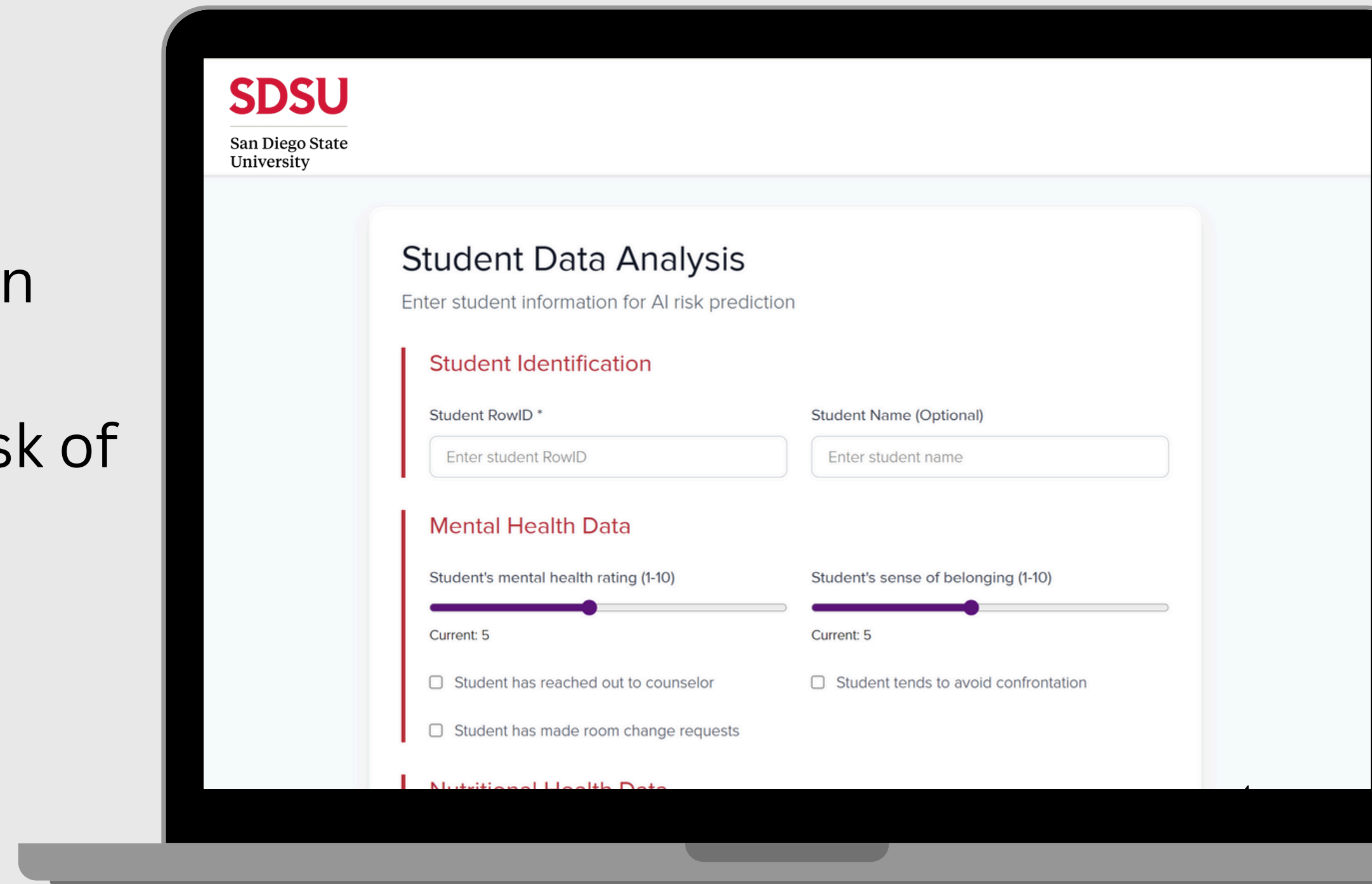
**Annual Housing Revenue
lost for every freshman
student who drops out**

**of SDSU students live off-
campus**



Dashboard Features

- Dashboard structured around customer needs: **housing administration**.
- Includes **main portal** and **student information form** based on chosen data columns.
- Submission results in student's risk of dropping out and contributing factors.



The screenshot shows a laptop displaying the SDSU (San Diego State University) Student Data Analysis dashboard. The dashboard is titled "Student Data Analysis" and includes the instruction "Enter student information for AI risk prediction". It features three main sections: "Student Identification", "Mental Health Data", and "Nutritional Health Data".

Student Identification

Student RowID * Student Name (Optional)

Mental Health Data

Student's mental health rating (1-10) Current: 5

☐ Student has reached out to counselor

☐ Student has made room change requests

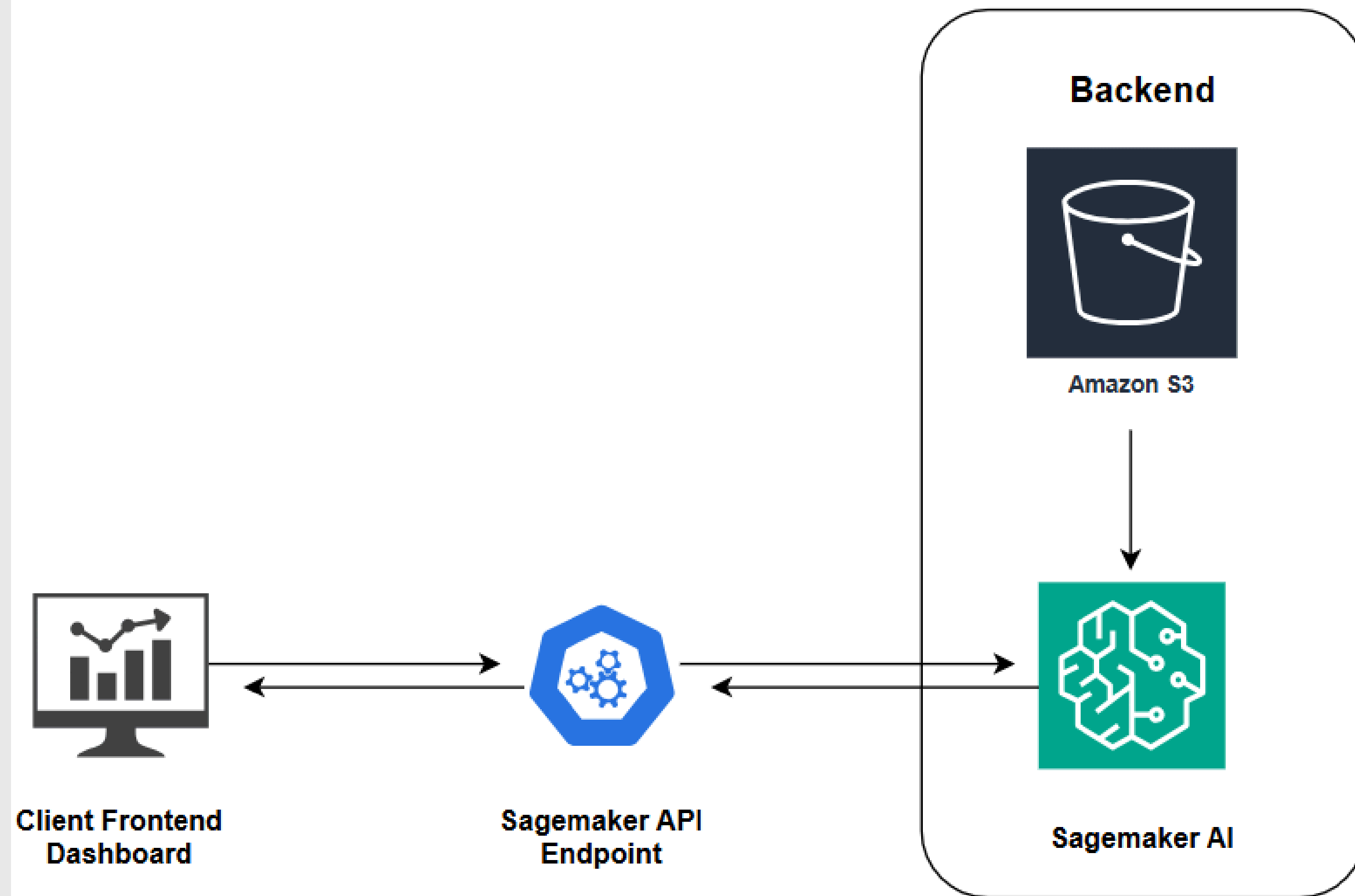
Student's sense of belonging (1-10) Current: 5

☐ Student tends to avoid confrontation

Nutritional Health Data



Project Structure



System Architecture

Project Design

Dataset

StarRez Housing Data (**16 out of 23 CSU campuses** use this system), Student Preferences survey, Housing assignment data, Learning communities and Academic Performance

Data Processing

Tools utilized: Jupyter Notebook, Copilot, Pandas

1. Created an EC2 server and virtual environment to host a Jupyter notebook.
2. Imported Pandas libraries and csv file data
3. Explored an AI predictive analysis model
4. Performed data transformation in the form of encodings.

Model Training

Cleaned data was utilized for model analysis:

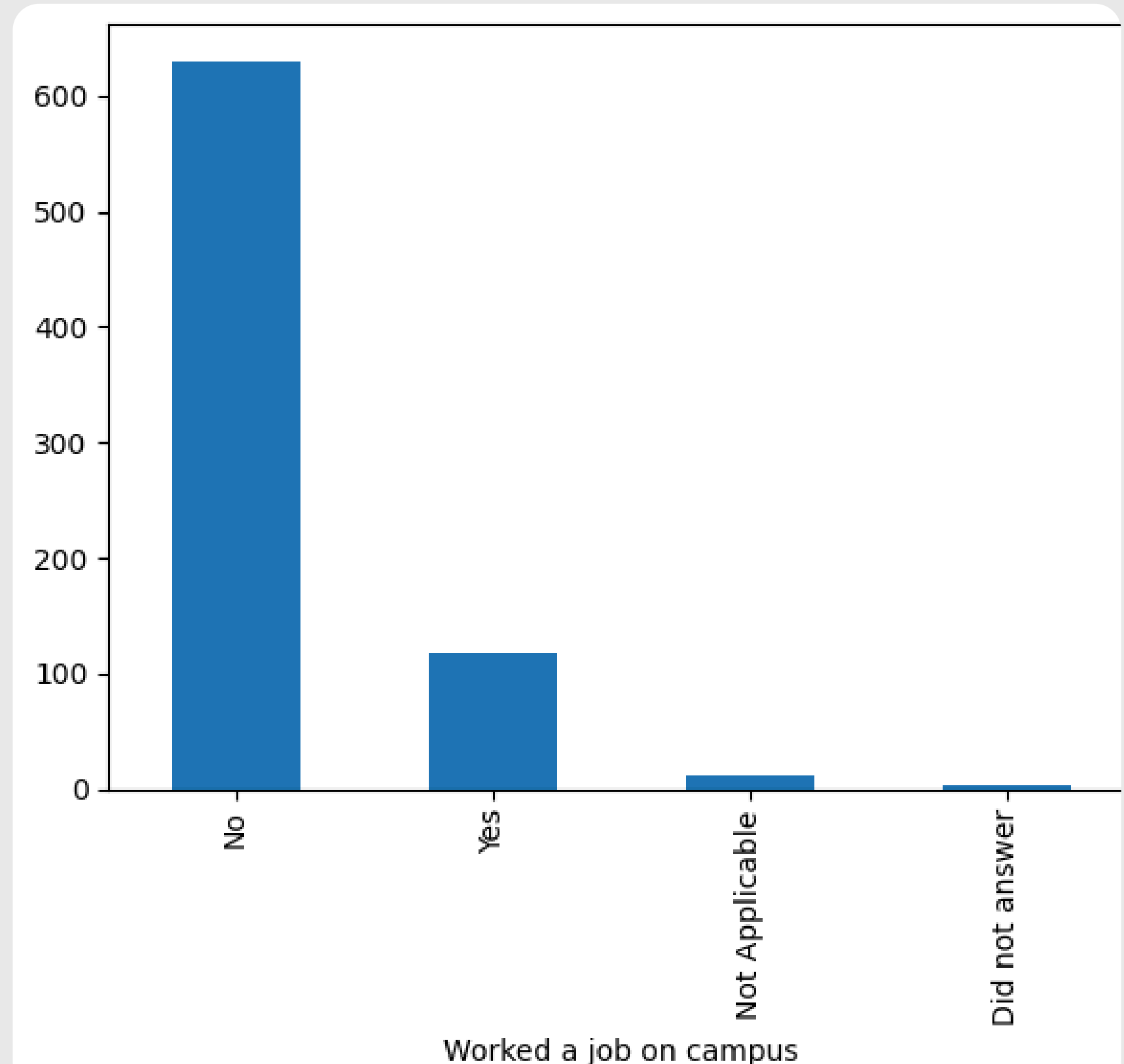
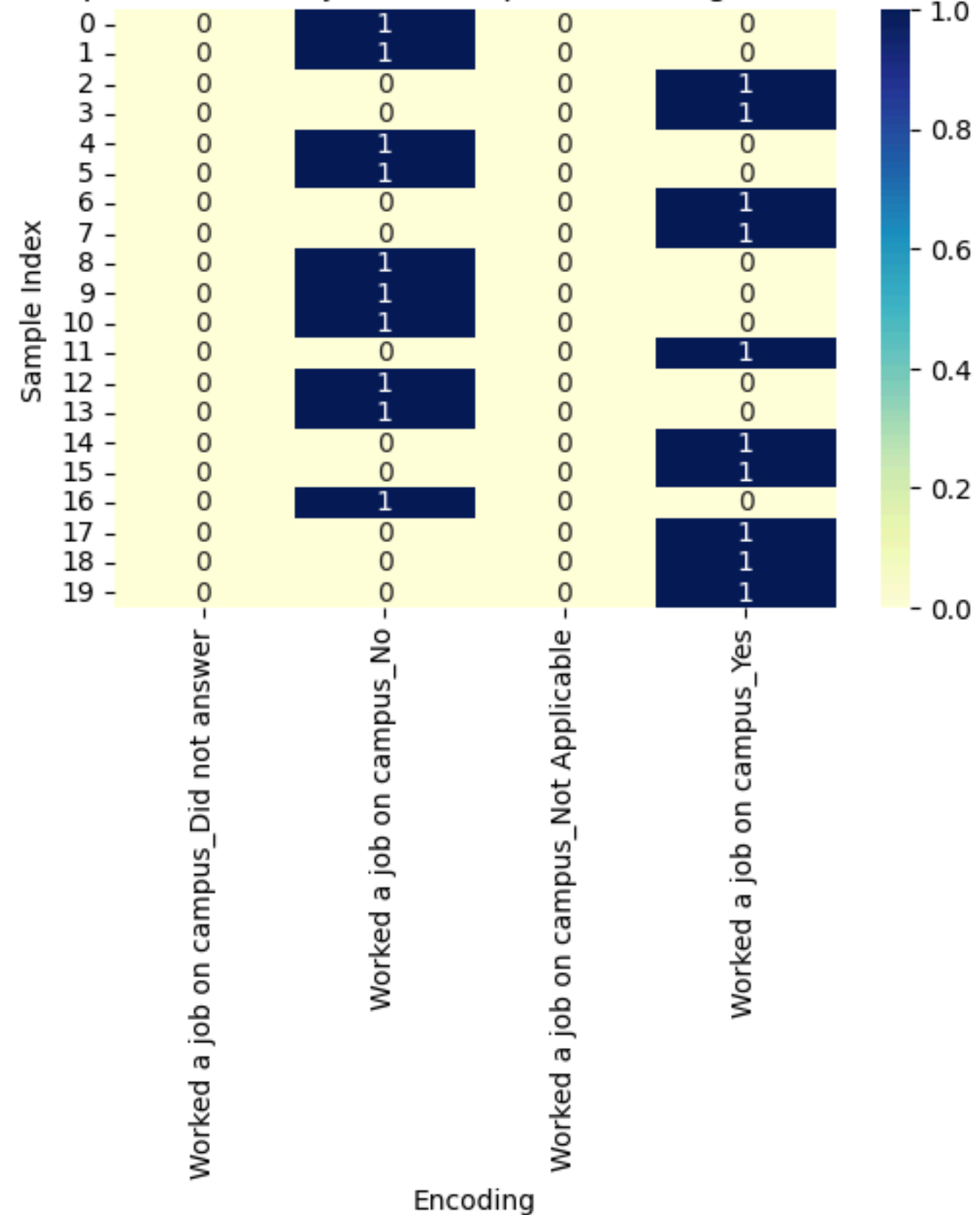
- **Logistical Regression**
- **Random Forest**
- **Sigmoid Neuron**

Sigmoid f1: 0.86



Data Visualization and Heat Maps

Heatmap of 'Worked a job on campus' Encodings (First 20 Rows)





Think Retention.

Predict the Future.

Empower Every Student.



Also, thank you to our mentor, Spandan!

Thank you!

Any questions?

