



# Predicting Student Adaptability Levels in Online Education

Kelly Blake | Rajeshwari Radhakrishna | Aisha Henderson |  
Sujatha Angajala

OCTOBER 2023

# Project Details

## Description:

This project utilizes supervised machine learning methods to develop a model that can predict the level of student adaptability in online learning, based on various demographic and socioeconomic features.

## What Are We Predicting?

What features influence student adaptability in an online learning setting? How can this information be leveraged to customize online learning programs and improve success for new digital learners?



# Dataset Information

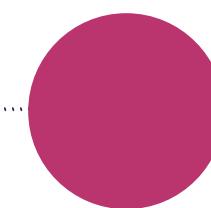
<sup>1</sup> Students Adaptability Level in Online Education

Column Name	Description	Type
Gender	Gender of Student	String
Age	Age range of the student	String
Education Level	Education Institution Level	String
Institution Type	Education Institution Type	String
IT Student	Studying as IT Student or Not	String
Location	Is Student Location in Town?	String
Load-Shedding	Level of load shedding	String
Financial Condition	Financial Condition of Family	String
Internet Type	Internet type used mostly in device	String
Class Duration	Daily class duration (in hours)	String
Self LMS	Institution's Own LMS Availability	String
Device	Device Used mostly in Class	String
Adaptability Level	Adaptability level of the student	String

- Data Source: Kaggle
- Total Rows: 1205 Rows
- The data collected was part of research conducted in Bangladesh after the onset of Covid, to understand and thus improve online learning for students across age ranges

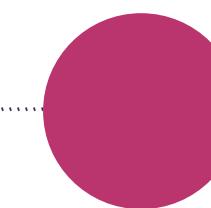
<sup>1</sup> <https://www.kaggle.com/datasets/mdmahmudulhasansuzan/students-adaptability-level-in-online-education>

# Coding and Machine Learning Tools



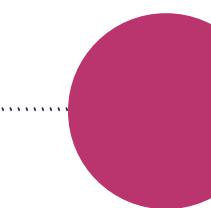
## **Postgres SQL**

Database used to house data



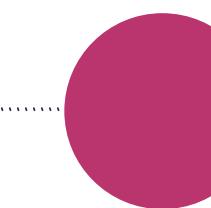
## **Jupyter Notebook**

Utilized for coding and python/machine learning libraries



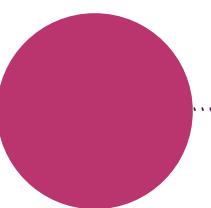
## **Python Pandas**

Used to read in and clean data from SQL



## **SciKit Learn - Random Forest**

Machine Learning Model used to train and predict dataset outcomes



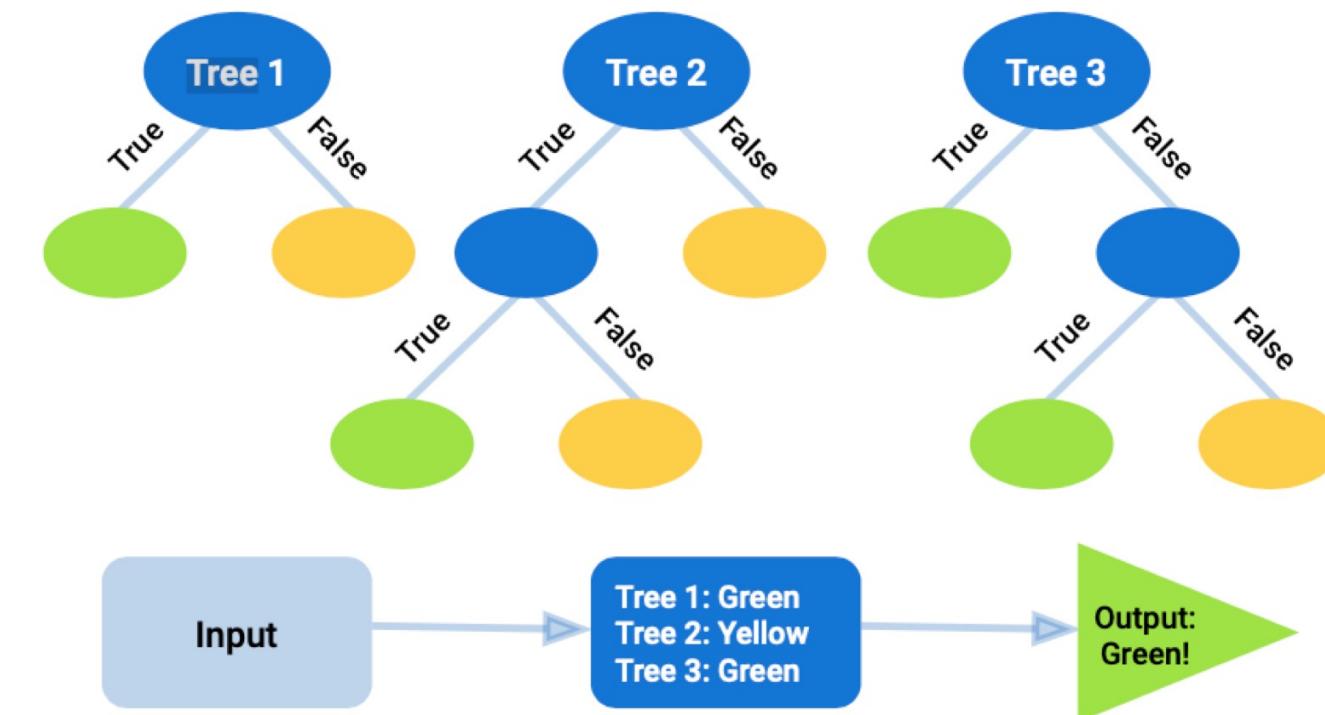
## **Tableau**

Used for Data Visualization

# Machine Learning Model

## Random Forest Algorithm

- Our model was trained using the Random Forest Algorithm to address a classification problem
  - A class label is determined based on the input data.
- The 'Adaptability' column used for the prediction contains multiple class labels:
  - High
  - Moderate
  - Low



# Machine Learning Model

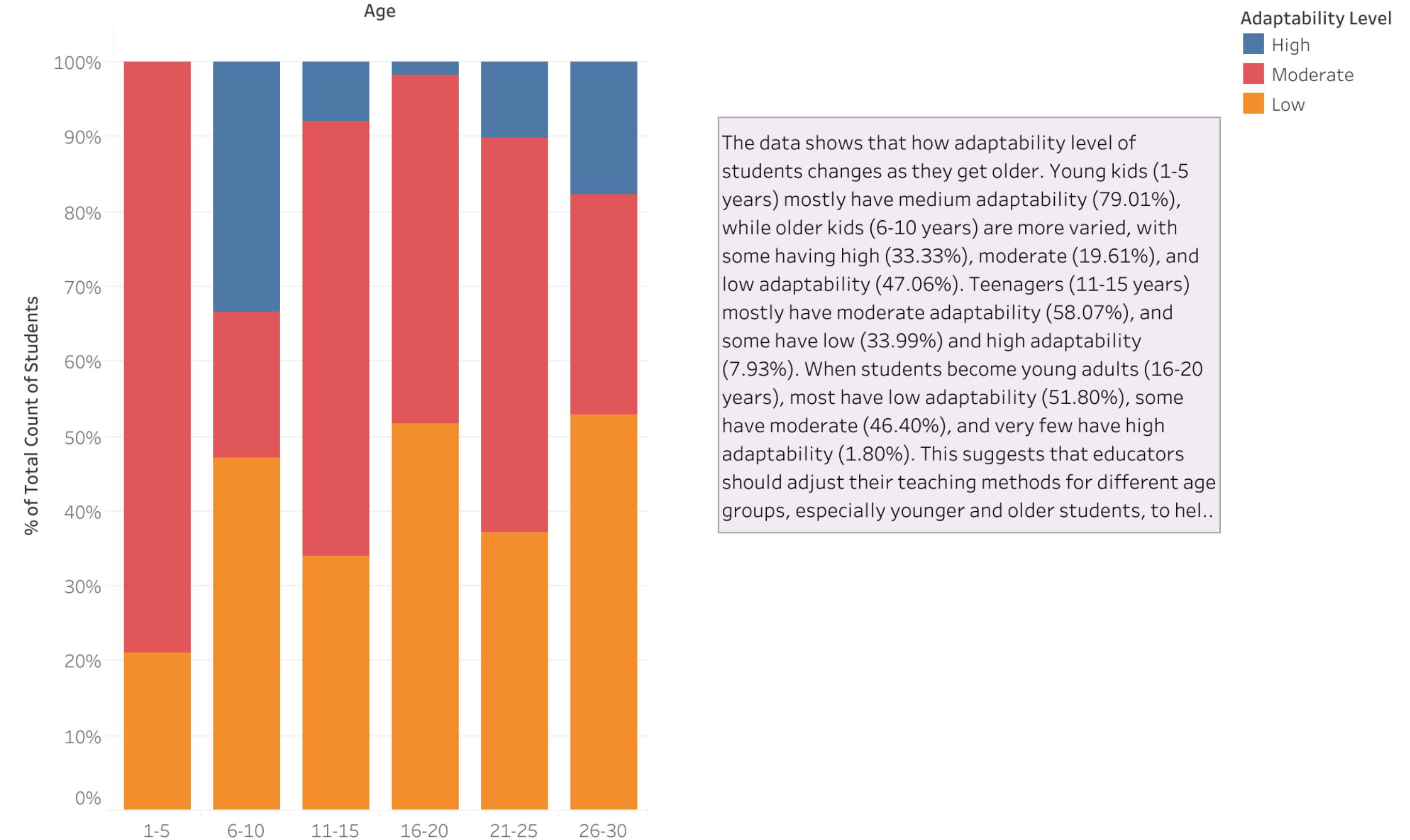
## Random Forest Algorithm

- For the dataset, the Random Forest Algorithm was used
- The problem we are solving for is a classification problem, in which a class label was anticipated based on the inputs of data
- The 'Adaptability' column used for the prediction contains multiple class labels:
  - High
  - Moderate
  - Low

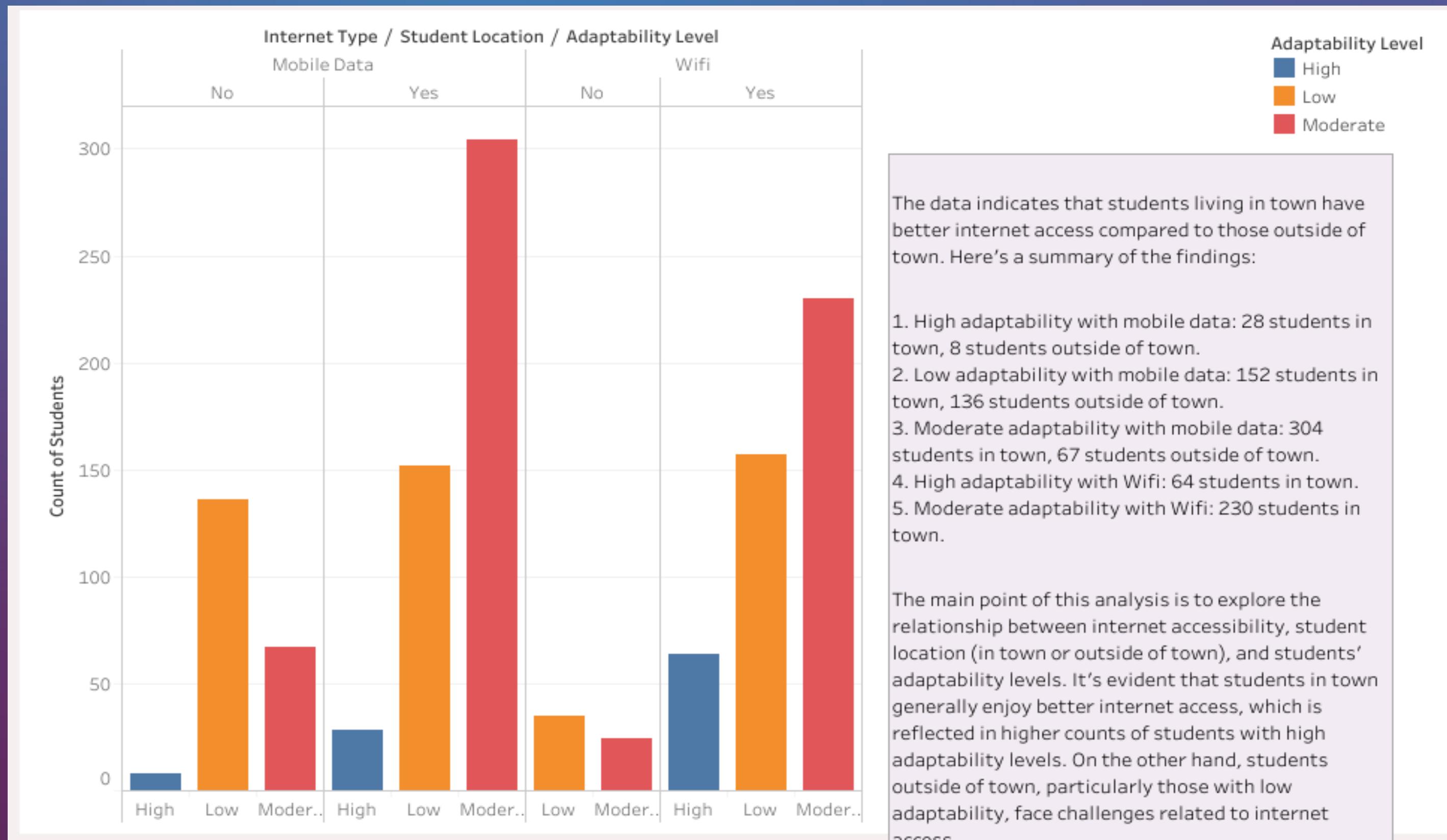
# Data Visualizations



# Adaptability Level by Student Age Range



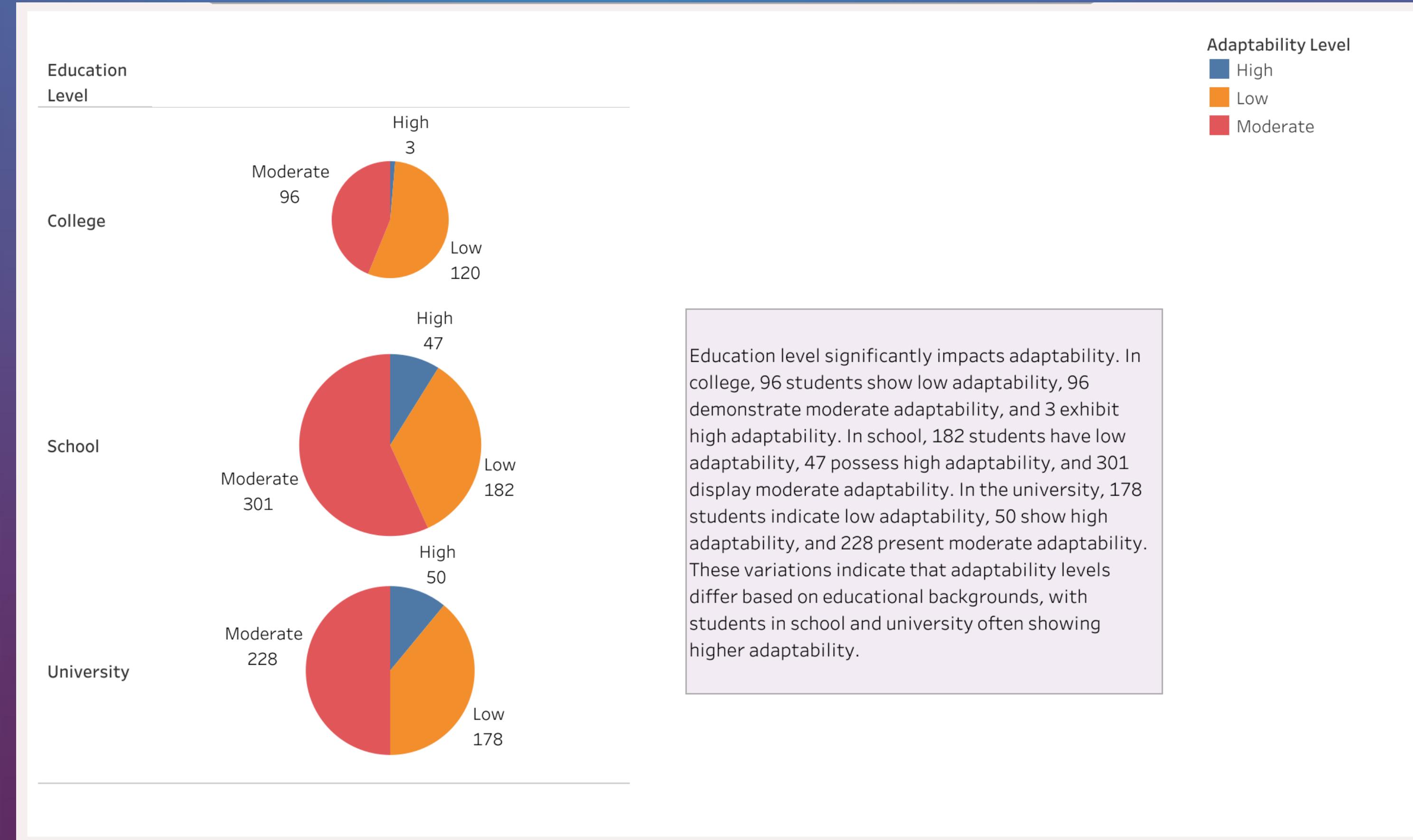
Was internet accessibility greater for students residing in town compared to those living outside of town?



# Can devices be effective tools for online learning?



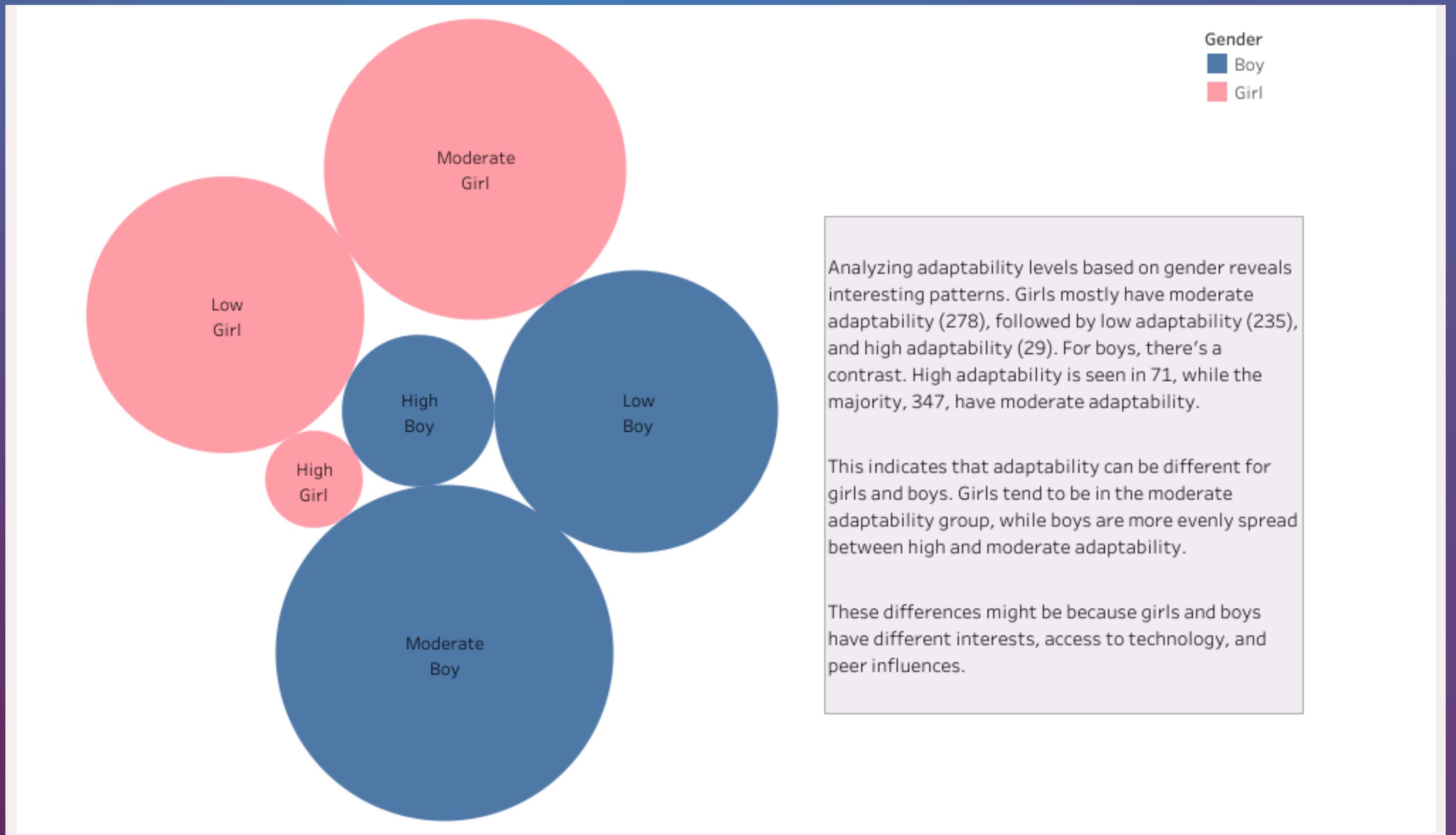
# How does the education level differ in adaptability?



# How does the Financial Condition of Students impact their online education adaptability?



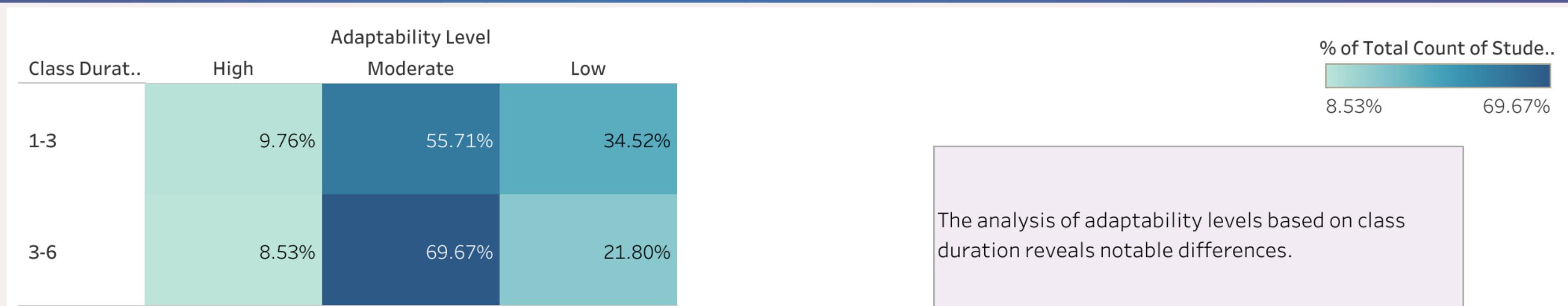
## Does adaptability differ based on student's gender?



# How does Load Shedding affect the adaptability of students in the context of online education?



# What is the relationship between Class Duration and student's adaptability?

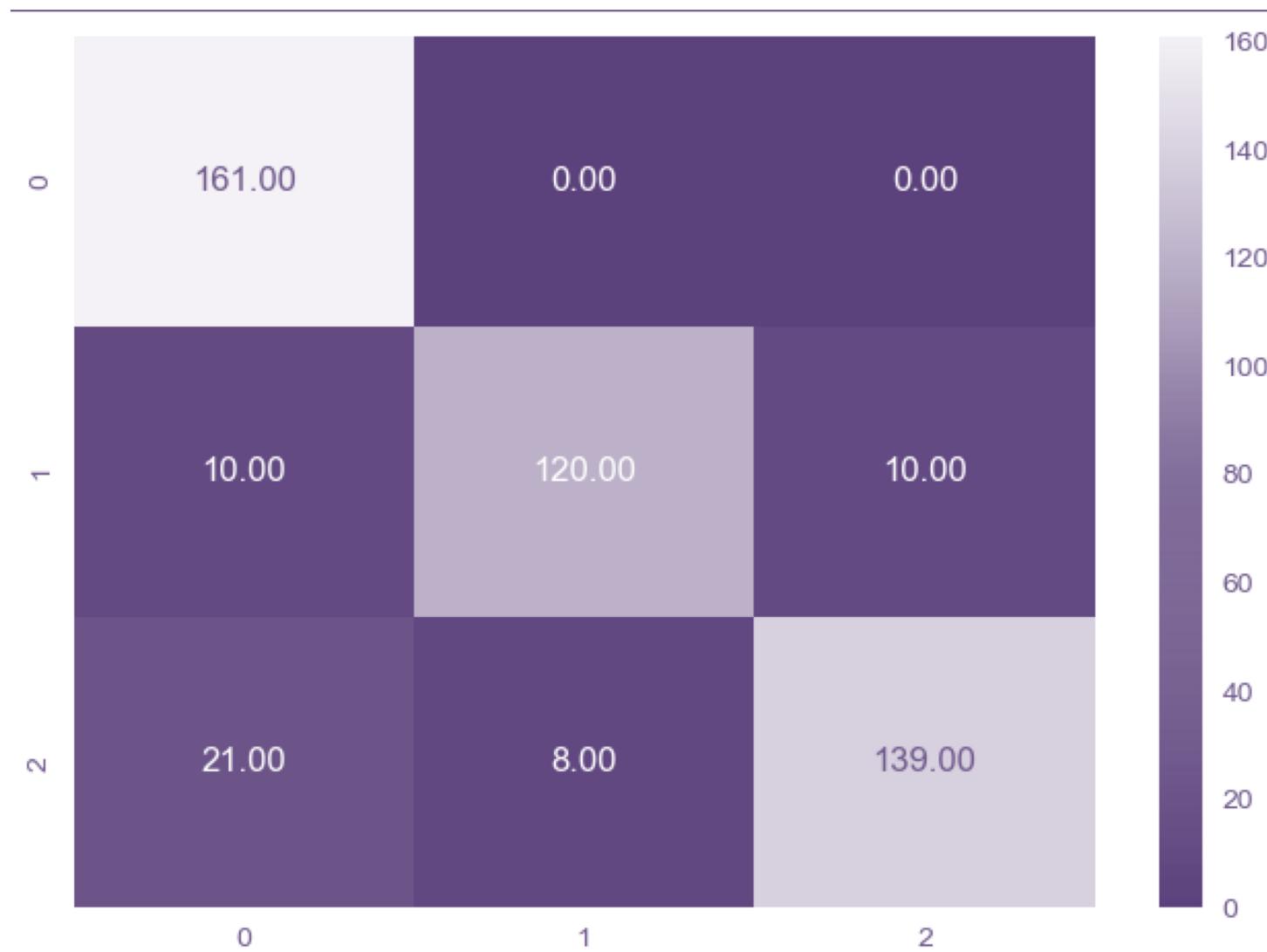


The analysis of adaptability levels based on class duration reveals notable differences.

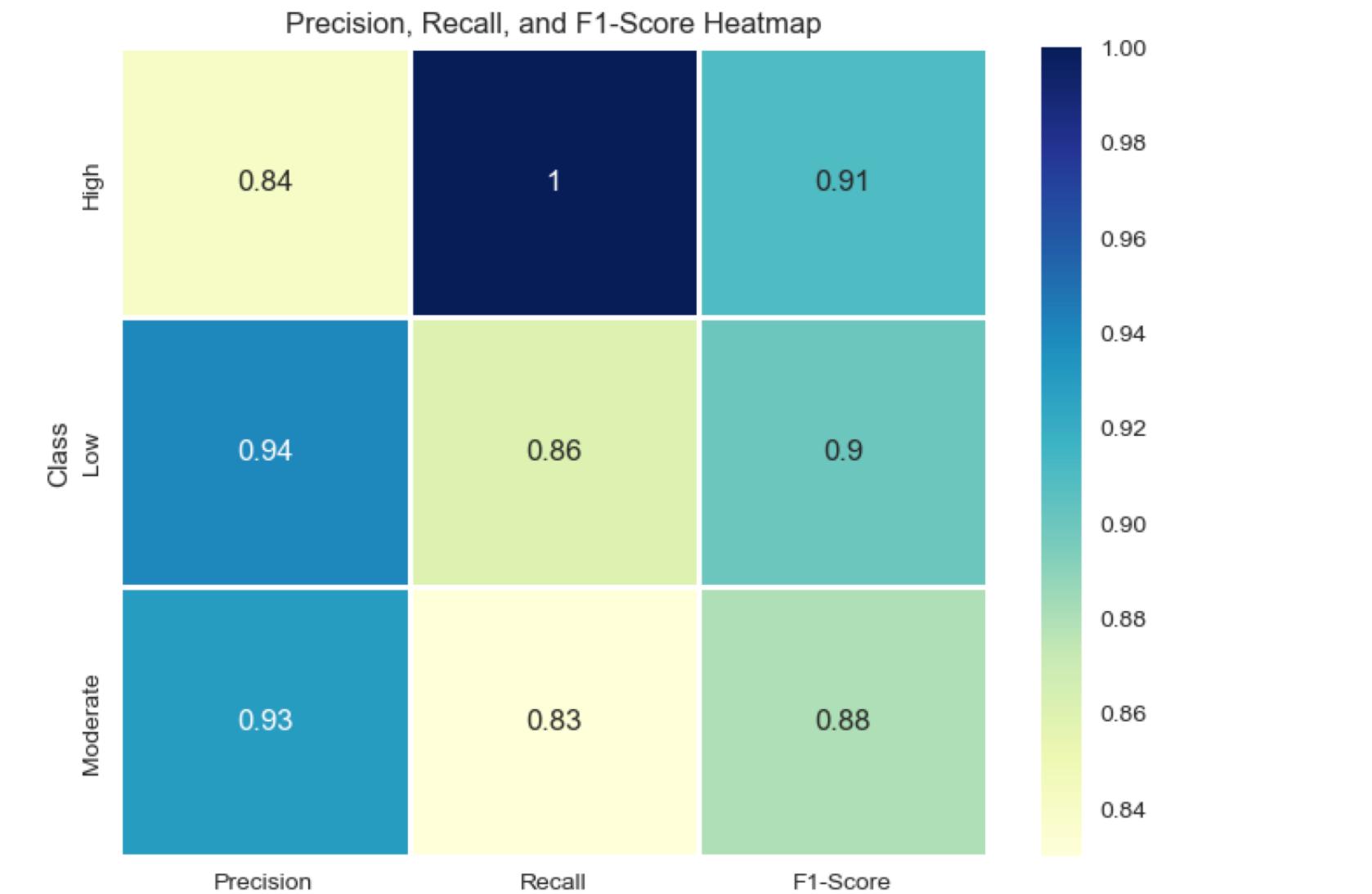
In the dataset, class duration is linked to students' adaptability in online education. Shorter classes (1-3 hours) tend to have a more even distribution of adaptability, while longer classes (3-6 hours) are associated with a higher prevalence of moderate adaptability. Educators may want to consider incorporating more interactive elements into longer classes to enhance students' adaptability.

# Final Analysis / Conclusion

## Confusion Matrix



## Classification Report



- The final confusion matrix, after balancing the data with RandomOverSampler reflects the model's effectiveness in classifying adaptability levels, with a notable ability to distinguish "High" adaptability, strong performance for "Low" adaptability, and accurate predictions for "Moderate" adaptability

- The final Classification Report reflects the model's effectiveness, with strong precision for predicting Moderate and Low adaptability levels, and accurate precision for High adaptability levels.
- Inversely, the recall scores indicate strong prediction for 'True Positives' for High adaptability, and accurate recall for the Low and Moderate classes
- Overall accuracy for the model indicates that 90% of the time, student adaptability levels were classified correctly

# THANK YOU!