8-2017

A data science pipeline for educational data : a case study using learning catalytics in the active learning classroom.

Asuman Cagla Acun Sener
University of Louisville

A DATA SCIENCE PIPELINE FOR EDUCATIONAL DATA: A CASE STUDY USING LEARNING CATALYTICS IN THE ACTIVE LEARNING CLASSROOM

By

Asuman Cagla Acun Sener
B.S., Hacettepe University, 2013

A Thesis
Submitted to the Faculty of the
J.B. Speed School of Engineering University of Louisville
In Partial Fulfillment of the Requirements
for the Degree of

Master of Science in Computer Science

Department of Computer Engineering and Computer Science
University of Louisville
Louisville, Kentucky

August 2017

A DATA SCIENCE PIPELINE FOR EDUCATIONAL DATA: A CASE STUDY USING LEARNING CATALYTICS IN THE ACTIVE LEARNING CLASSROOM

By

Asuman Cagla Acun Sener
B.S., Hacettepe University, 2013

A Thesis Approved on

August 7, 2017

by the following Thesis Committee:

____
Dr. Olfa Nasraoui, Thesis Director

____
Dr. Nihat Altiparmak

____
Dr. Jeff Hieb

DEDICATION

This thesis is dedicated to my husband

Samil Sener

to my son

Everest

and

every human who fights for justice.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Olfa Nasraoui for her direction, assistance, guidance and immense knowledge. Specifically, after being a new mother, she always encouraged me and I always felt her support. I would also like to thank the members of my thesis committee, Dr. Nihat Altiparmak for all his guidance and support after starting my master's degree and Dr. Jeff Hieb, an innovative educator, for providing his course data that was used in this thesis. I would also like to express my thanks to my husband, Samil, for his understanding and patience during those times when there was no light at the end of anything. He encouraged me, believed in me, and made me stick with it.

ABSTRACT

A DATA SCIENCE PIPELINE FOR EDUCATIONAL DATA: A CASE STUDY USING LEARNING CATALYTICS IN THE ACTIVE LEARNING CLASSROOM

Asuman Cagla Acun Sener

August 7, 2017

This thesis presents an applied data science methodology on a set of University of Louisville, Speed School of Engineering student data. We used data mining and classic statistical techniques to help educational researchers quickly see into the data trends and peculiarities. Our data includes scores and information about two Engineering Fundamental Class. The format of these classes is called an inverted classroom model or flipped class. The purpose of this study is to analyze the data in order to uncover potentially hidden information, tell interesting stories about the data, examine student learning behavior and learning performance in an active learning environment, including collaborative learning in a flipped classroom model.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Educational Data Mining (EDM) is defined by The Educational Data Mining community website, www.educationaldatamining.org "as an emerging discipline, concerned with developing methods for exploring the unique types of data that come from the educational setting, and using those methods to better understand students, and the settings which they learn in." EDM develops methods and applies techniques from statistics, machine learning, and data mining to analyze data collected during teaching and learning. EDM tests learning theories and informs educational practice [1]. Rather than the theory of learning, in this thesis, we focus on the computational aspects of educational data mining, namely designing the data science pipeline that can reveal patterns in education data.

Benjamin Franklin says "Tell me and I forget. Teach me and I remember. Involve me and I learn." That is the fundamental idea of the flipped classroom approach. Our thesis presents the results of applying exploratory educational data mining on data of student activities in a flipped classroom model.

1.1 Flipped Classroom

The flipped classroom is a pedagogical model in which the traditional lecture and assignments of a course are reversed [2]. Bishop and Verleger [3], conducted an extensive survey of the research on the flipped classroom and added more on the current definition of the flipped classroom. They define the flipped classroom as an educational technique that consists of two parts: interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom. A graphic representation of this definition is shown in Figure 1.1.

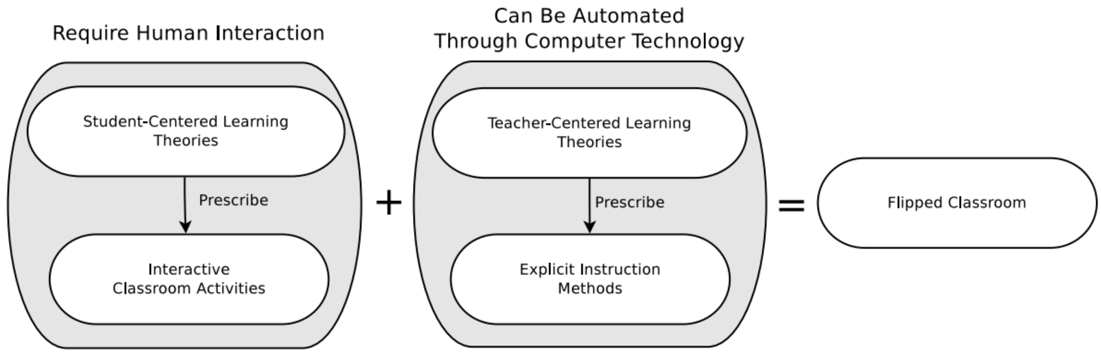


Figure 1.1: Flipped Classroom [3]

The courses that we are concerned with in this thesis are Introductory Engineering Fundamental Courses, whose course structure, based on the syllabus can be described as follows [4] [5]:

The material in each unit is divided into multiple lessons. Each lesson has a single corresponding assignment. Included in that lesson, are links to specific relevant sections in the textbook, links to video lectures, and these are followed by a few practice questions for the material in that lesson. Students are expected to read the sections in the textbook and watch the videos, making notes as they go through the material. These assignments have a due date, and students are expected to have read the chapter section and watched the videos, and attempted the practice questions by the due date. Completing these assignments means coming to class prepared, and class Readiness Assessment Test (RATs) expect that students have completed the unit lesson for that day.

Class meetings are centered on working problems in small groups. At the beginning of each class meeting, students take a short Readiness Assessment Test (RAT). The RAT includes basic questions. This is an individual work, and finishes in

5 minutes. After the RAT, the instructor quickly reviews that day's lesson material, and then the remainder of the time, students work in small groups solving more difficult problems related to that lesson or previous unit lessons.

1.2 Objectives of the Thesis

The main objective of this thesis is to analyze student activity data to uncover potentially hidden information that can help tell data stories and help understand student learning behavior and learning performance in an active learning environment and in collaborative groups within a flipped classroom model.

1.3 Contributions

We propose a data science pipeline methodology to analyze and visualize raw educational data, based on classical statistical methods such as factor analysis, visualization methods such as heat maps, and machine learning algorithms such as decision tree learning. Our biggest effort was on the data preparation phase which started with raw data. This phase required understanding the domain and how the data is related to its context. Many iterations were also required while generating visualizations in order to reveal useful information.

1.4 Organization of the Thesis

The remainder of this thesis is organized as follows. Chapter 2 provides a literature review of our applied methodology and related work. Chapter 3 continues with the methodology which are followed by the experimental results are presented in Chapter 4. Finally, Chapter 5 summarizes the results.

LITERATURE REVIEW

In this chapter, we present a brief review of the methods that are used in our data science pipeline.

2.1 Exploratory Factor Analysis

Exploratory factor analysis (EFA) is a statistical method used to uncover the underlying structure of a relatively large set of variables [6]. Factor analysis could be described as orderly simplification of interrelated measures. Traditionally, factor analysis has been used to explore the possible underlying structure of a set of interrelated variables without imposing any preconceived structure on the outcome [7].

To determine the number of factors, Cattell [8] introduced scree plots, which are visual tools used to help determine the number of important components or factors in multivariate settings, such as principal component analysis and factor analysis. The scree plot is examined for a natural break between the large eigenvalues and the remaining small eigenvalues.

After applying EFA, factor loadings need to be rotated to become interpretable [9]. There are two main factor rotation methods; orthogonal rotation and oblique rotation. An orthogonal rotation assumes that the factors are uncorrelated, while an oblique rotation assumes that factors are correlated [10].

2.2 Visual Data Analysis

Bar charts, histograms, scatter plots, social network graphs, stream graphs, tree maps, gratt charts, heat maps, and correlation plots are different techniques used for data visualization [11].

"Visual data analysis is a way of discovering and understanding patterns in large datasets via visual interpretation. It is used in the scientific analysis of complex processes. Visual data analysis is an emerging field, a blend of statistics, data mining, and visualization that promises to make it possible for anyone to sift through, display, and understand complex concepts and relationships" [1].

2.3 Decision Tree

Decision tree learning is a method for approximating discrete-valued target functions, in which the learned function is represented by a decision tree [12]. Most of the decision tree algorithms developed from ID3, which is developed by Ross Quinlan [13]. Decision tree J48, which we used in our research, is the implementation of algorithm ID3 developed by the WEKA project team [14].

In pseudo code, the general algorithm for building decision trees is [15]:

1. Check for the above base cases.
2. For each attribute a, find the normalized information gain ratio from splitting on a.
3. Let a_best be the attribute with the highest normalized information gain.
4. Create a decision node that splits on a_best.
5. Recur on the sub lists obtained by splitting on a_best, and add those nodes as children of node.

2.4 Related Work

Based on the meta-analysis research paper [16], the authors found that the most popular techniques for educational data mining (EDM) were: clustering, followed by classification, sequential pattern mining, prediction, and association rule analysis. Also, Baker [17] divides EDM research in the following general categories: prediction, clustering, relationship mining, discovery with models, and distillation of data for human judgment.

Specifically, for flipped classroom data analysis, several efforts have been reported [18] [19]. They are mainly focused on comparing student scores of flipped classroom and traditional class methods for the same department and same course, and they are mostly engineering departments [20]. Also, some of them use student feedback for data analysis [21] [22] [23].

2.5 Summary

In this chapter, we reviewed background on exploratory factor analysis, visualization and decision trees, because of their relevance to our work. We concluded with existing work in educational data mining. In the next chapter, we will present our data science pipeline on educational data.

METHODOLOGY

In this chapter, we present the different steps of our data science pipeline.

3.1 Data Science Pipeline

Figure 3.1: Data Science Pipeline depicts the general flow and stages of our methodology, which includes four major stages.

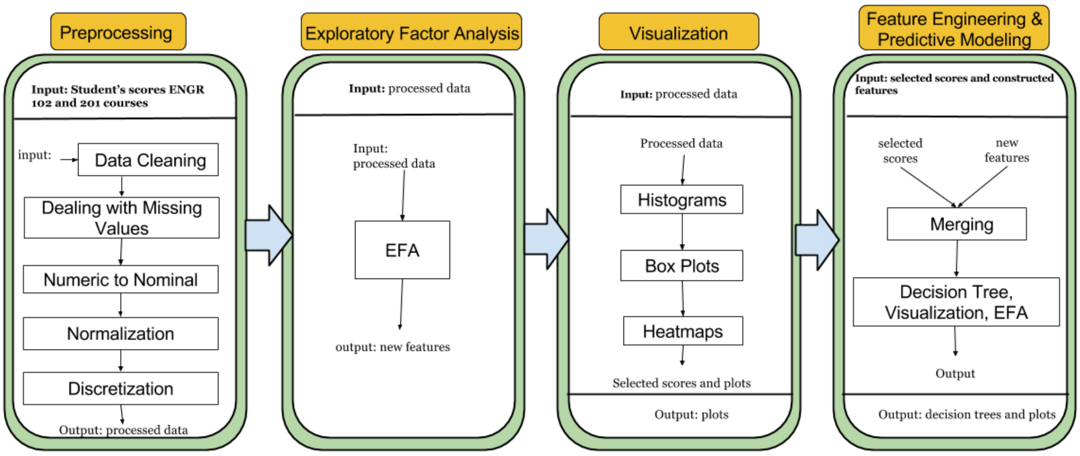


Figure 3.1: Data Science Pipeline

3.1.1 Preprocessing

Before we analyzed the data, we performed the following data preprocessing steps:

- Data Cleaning: We removed the features that we will not use in our analysis.
- Dealing with Missing Values: There are a small number of N/A values. We filled them with zero, which makes sense because if there is no score, this means that the student did not participate in the test.

- Numerical to Nominal: We converted attributes' numerical values to nominal values before building decision trees.
- Normalization: We experimented with centering our data to a zero mean (the mean for the entire class for one activity or exam). Normalized values allow the comparison between different scores in terms of how they are changing relative to each activity's class average.
- Discretization: We discretized values by mean to obtain more accurate results from the prediction model.

3.1.2 Exploratory Factor Analysis (EFA)

To apply EFA, we used the R language, because its libraries supporting EFA were preferable to Python. The R libraries that used are readxl, ggplot2, psych, corrplot, and GPArotation. We used the oblique rotation method to rotate factor loadings.

3.1.3 Visualization

To visualize our data, we used Python, which is a very popular programming language especially for data science [24]. Histograms, boxplots, and mainly heat maps were created in our study. The following libraries are used: pandas, matplotlib pyplot, ggplot, plotly, numpy, scipy stats, and seaborn.

3.1.4 Feature Engineering

Feature engineering is used when building predictive models where we clearly have an outcome to predict (a discrete class label or continuous variable). Feature engineering can also help in unsupervised learning and preliminaries exploratory analysis to allow us to dig stories that may be hidden within the data such as whether there are distinct groups, trends, or correlations. It can also help us build more meaningful visualizations. After exploratory factor analysis and visual data analysis, we constructed new features that we confirmed, and then built decision tree models to predict the final score.

3.2 Summary

In this chapter, we presented our methodology for the data science pipeline. In the next chapter, we will present our experimental results based on our education data sets.

EXPERIMENTAL RESULTS

4.1 IRB Statement

This research was approved by the Institutional Review Board at our institution. An independent evaluator monitored the research to ensure that students assigned to the control group received fair treatment, despite having spacing withheld from their instructional plan in Introductory Calculus for Engineers.

4.2 Data Sets

Our data sets include the following information about a set of students at University of Louisville, Speed School of Engineering.

- Key features of students [student id, gender and ACT math score],
- Scores of the Introductory Calculus for Engineers Course 1,
- Scores of the Introductory Calculus for Engineers Course 2.

4.3 Student Demographics

In our dataset, the total number of students is 190, including 43 females and 147 males. In our classes, the number of males is almost three times higher than the number of females. Also, 77% of the students are males and 23% of the students are females. This distribution is depicted in Figure 4.1: Histogram of gender; F is female, M is male

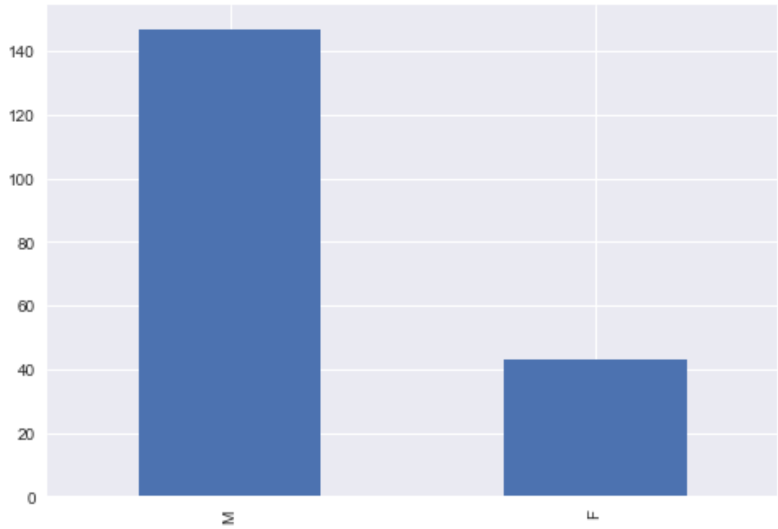


Figure 4.1: Histogram of gender; F is female, M is male

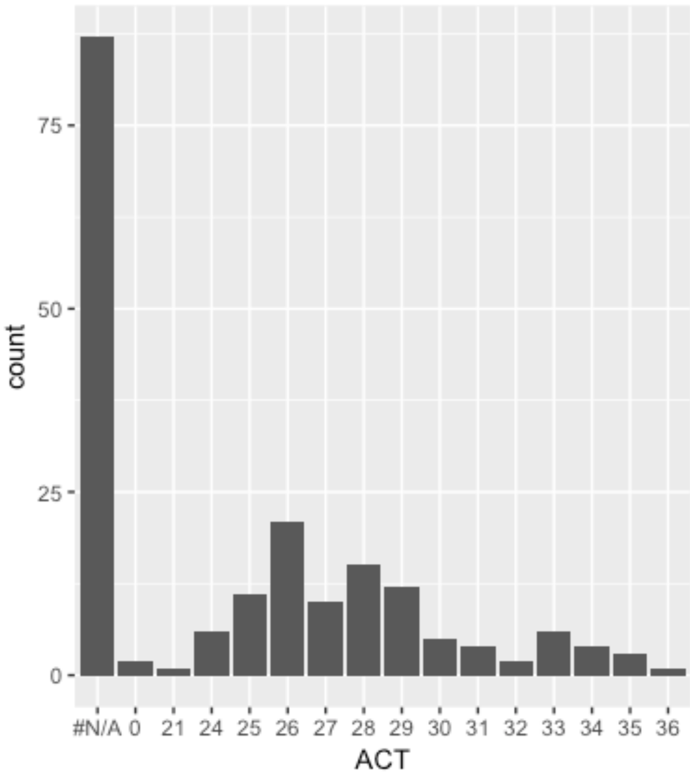


Figure 4.2: Histogram of ACT math scores

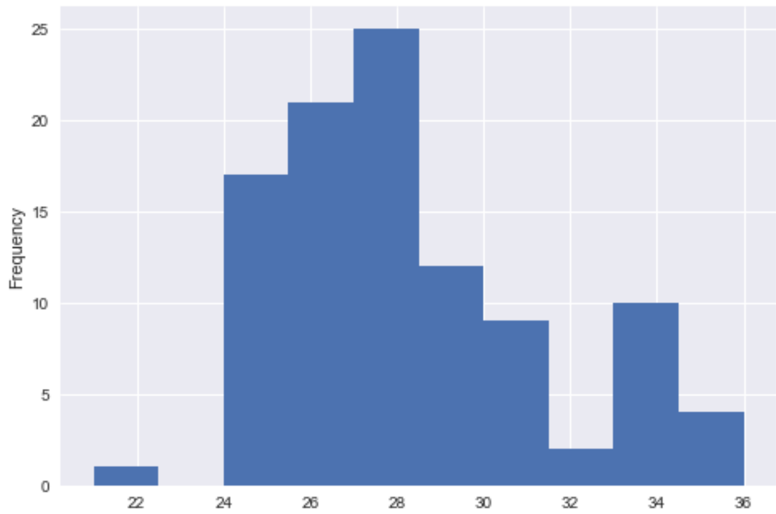


Figure 4.3: Histogram of ACT math score without missing scores (NA) and zeroes

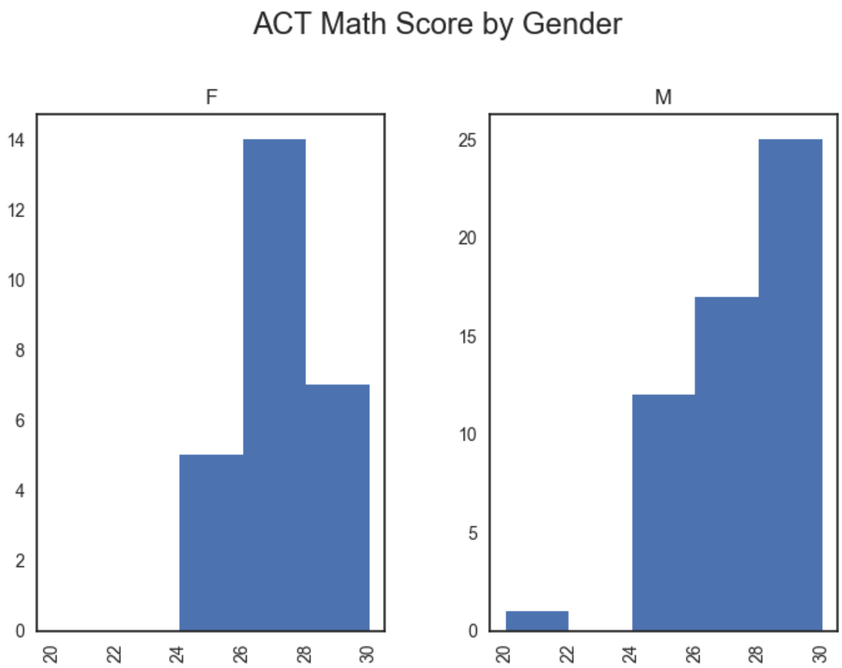


Figure 4.4: Histogram of ACT math scores grouped by gender

Figure 4.5: Violin plot of ACT math scores grouped by gender

We can see in Figure 4.2 that almost half of the student scores are not available. For the known values, Figure 4.3, shows scores that are mainly distributed in the range [24-30] and 14% of the student have the highest score range in [33-36]. In

Figure 4.4 and Figure 4.5, male students are seen to have higher average ACT math score than females. This visualization, called violin plot, is structured as follows: The thick black bar in the center represents the interquartile range, the thin black line extended from it represents the 95% confidence intervals, and the white dot is the median.

4.4 The Introductory Calculus for Engineers Course 1

The Course 1 dataset has [97 rows x 335 columns] corresponding to 91 students with 335 attributes of combined homework, class activities, and exams scores, including for some of the scores, their date and time spent.

4.4.1 EFA for Course 1

We divide this section into three parts as follows:

1. Homework and lesson assignments,
2. Class activities, and
3. Exams.

In the following subsections, we present our exploratory factor analysis (EFA) results.

4.4.1.1 Homework Assignments

In this section, we present our exploratory factor analysis results for homework assignments.

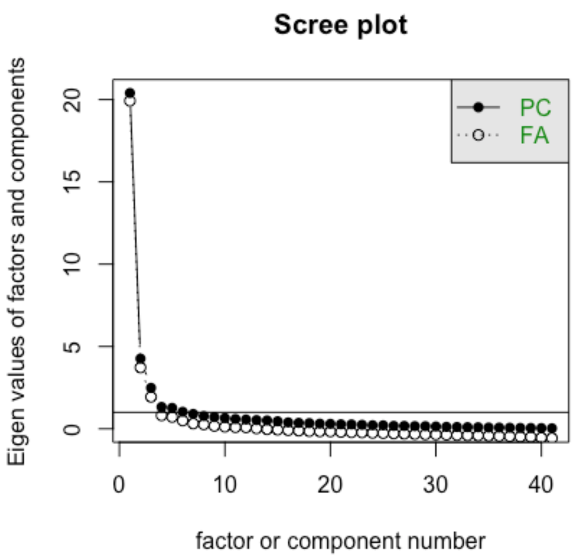


Figure 4.6: Scree Plot of Course 1 Homework

From the scree plot in Figure 4.6, we observe that there are three significant factors over 41 variables in homework scores.

```
Loadings:
View(x, title)      MR1    MR3    MR2
HWU1L1_S                          0.607
HWU1L2_S                          0.744
HWU1L3_S                          0.819
HWU2L1_S                          0.894
HWU2L2_S                          0.867
HWU2L3_S                          0.553
HWU2L3_Supdated                   0.671
HWU2_additional            0.417
HWU3L1_S                          0.629
HWU3L2_S                          0.578
HWU3L3_S                   0.491
HWU4L1_S                   0.685
HWU4L2_S                   0.887
HWU4L3_S                   0.741
HWU5L1_S                   0.721
HWU5L2_S                   0.629
HWU5L3_S                   0.705
HWU6L1_S                   0.735
HWU6L2_S                   0.812
HWU6L3_S                   0.782
HWU7L1_S                   0.824
HWU7L2_S                   0.661
HWU7L3_S                   0.606
HWU8L1_S            0.444  0.571
HWU8L2_S            0.426  0.548
HWU8L3_S            0.575
HWU9L1_S            0.482
HWU9L2_S            0.538
HWU9L3_S            0.535
HWU10L1_S           0.767
HWU10L2_S           0.813
HWU10L3_S           0.795
HWU11L3_S           0.876
HWU12L1_S           0.902
HWU13L3_S           0.694
HWU11L1_S           0.795
HWU11L2_S           0.799
HWU12L3_S           0.865
HWU12L2_S           0.895
HWU13L1_S           0.758
HWU13L2_S           0.621

                 MR1   MR3   MR2
SS loadings    9.398 8.344 4.992
```

Figure 4.7: Factor Loadings of Homework Assignment Scores of Course 1

Table 1: Factors of Homework Scores of Course 1

Factor 1	Homework Unit 1 to 3
Factor 2	Homework Unit 3 to 7
Factor 3	Homework Unit 7 to 13

4.4.1.2 Class Activities

In this section, we present our exploratory factor analysis results for in-class activities.

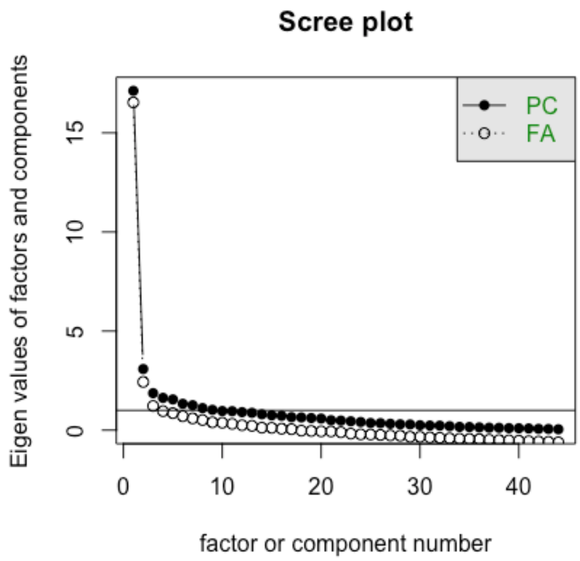


Figure 4.8: Scree Plot of Course 1 Class Activities

Table 2: Factors of Course 1 Class Activities

Factor 1	Class Activities Unit 8 to 13
Factor 2	Class Activities Unit 1 to 7

The scree plot, shown in Figure 4.8, reveals two factors. Table 2 shows that Factor 1 includes the last 6 units which range between Units 8-13, while Factor 2 includes the first 7 unit activities. If we look closely at the factor loadings, we can see that the most significant attributes are in factor 1 and they are the reviews of each unit.

Loadings:

View(x, title)

	MR1	MR2
U1L1_CA_8/25/2015		0.527
U1L2-CA 8/27/15		
U1L3-CA 8/28/2015		0.525
U1 Review-CA 8/31/2015		
U2L1-CA 9/2/2015		0.431
U2L2-CA 9/3/2015		0.487
U2L3-CA 9/4/2015		
U3L1-CA 9/9/2015		0.554
U3L2-CA 9/10/2015		0.625
U3L2-CA 9/14/2015		0.637
U4L1-CA 9/16/2015		0.519
U4L2-CA 9/17/2015		0.590
U4L3-CA 9/18/2015		0.626
U4 Review-CA 9/21/2015		
U4L1-CA 9/23/2015		0.870
U4L2-CA 9/24/2015		0.442
U4L3-CA 9/25/2015		0.595
U6L1-CA 9/30/2015	0.415	
U6L2-CA 10/01/2015		0.737
U6L3-CA 10/02/2015		0.716
U7L1-CA 10/12/2015		0.446
U7L2-CA 10/14/2015		0.498
U8L1-CA 10/16/2015	0.499	
U8L2-CA 10/19/2015	0.502	
U8L3-CA 10/21/2015	0.547	
U9L1-CA 10/23/2015	0.481	
U9L2-CA 10/26/2015	0.636	
U9L3-CA 10/28/2015	0.608	
U9 Review-CA 10/28/2015	0.816	
U10L1-CA 10/30/2015	0.611	
U10L2-CA 11/2/2015	0.426	
U10L3-CA 11/4/2015	0.490	
U11L1-CA 11/6/2015		
U11L2-CA 11/9/2015	0.661	
U11L3-CA 11/10/2015	0.552	
U12L1-CA 11/13/2015	0.754	
U12L2-CA 11/16/2015	0.670	
U11 Review-CA 11/11/2015	0.816	
U12L3-CA 11/17/2015	0.578	
U12 Review-CA 11/18/2015	0.847	
U13L1-CA 11/20/2015	0.847	
U13L2-CA 11/23/2015	0.749	
U13L3-CA 11/24/2015	0.642	
U13 Review-CA 11/30/2015	0.717	

Figure 4.9: Factor Loadings for Class Activities of Course 1

4.4.1.3 Exams

In this section, we present our exploratory factor analysis results for exam scores.

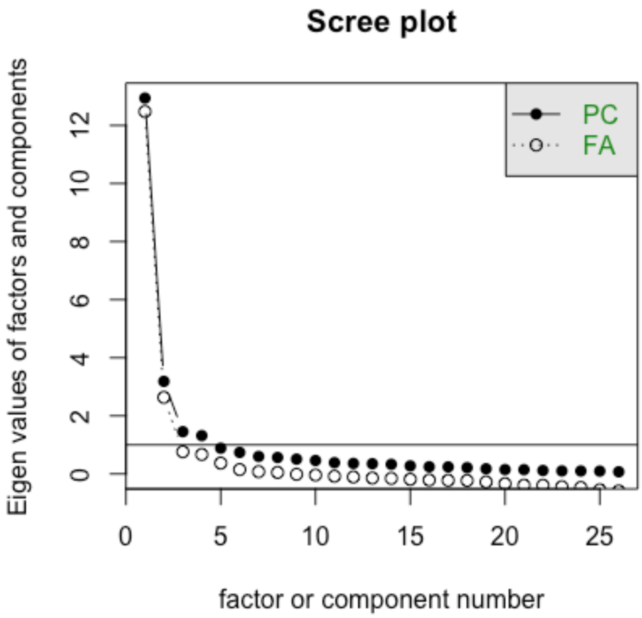


Figure 4.10: Scree Plot of Course 1's Homework

From the scree plot in Figure 4.10, we observe that there are two significant factors over 26 variables in the data set. Table 3 shows that these factors are: Units up to Unit 7 and units after Unit 7, respectively.

Table 3: Factors of Exams of Course 1

Factor 1	Exams Unit 1 to 7
Factor 2	Exams Unit 7 to 13

```
Loadings:
             MR1    MR2
Exam1P1_S
Exam1P2_S
Exam2P1_S           0.462
Exam2P2_S           0.695
Exam3P1_S           0.727
Exam3P2_S           0.607
Exam4P1_S           0.881
Exam4P2_S           0.851
Exam5P1_S           0.574
Exam5P2_S           0.735
Exam6P1_S           0.490
Exam6P2_S           0.763
Exam7P1_S    0.437  0.424
Exam7P2_S           0.557
Exam8P1_S    0.603
Exam8P2_S    0.439  0.452
Exam9P1_S    0.609
Exam9P2_S    0.634
Exam10P1_S   0.801
Exam10P2_S   0.569
Exam11P1_S   0.873
Exam11P2_S   0.946
Exam12P1_S   0.946
Exam12P2_S   0.937
Exam13_S     0.916
FinalExam_S  0.850

                 MR1   MR2
SS loadings    7.910 6.153
Proportion Var 0.304 0.237
Cumulative Var 0.304 0.541
```

Figure 4.11: Factor Loadings for Exams of Course 1

4.4.2 Visual Data Analysis for Course 1

We separated this section into three parts as follows:

1. Homework
2. Class activities
3. Exams

In the next subsections, we present all related visualizations.

4.4.2.1 Homework

In this section, we present visualizations related to homework and lesson assignments. To facilitate interpreting each visualization, we attempt to summarize its analysis within its own caption, rather than in the main text.

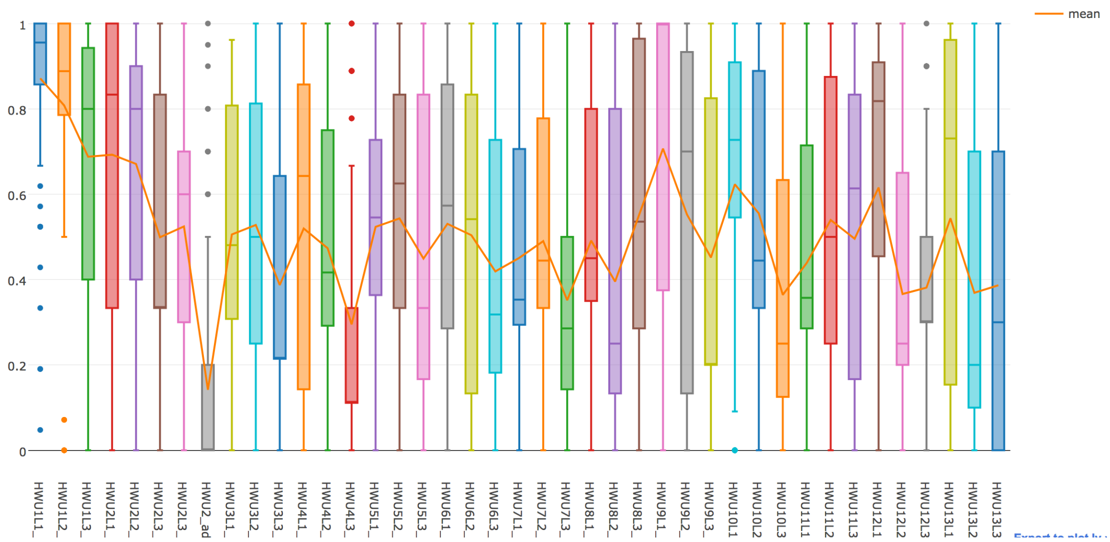
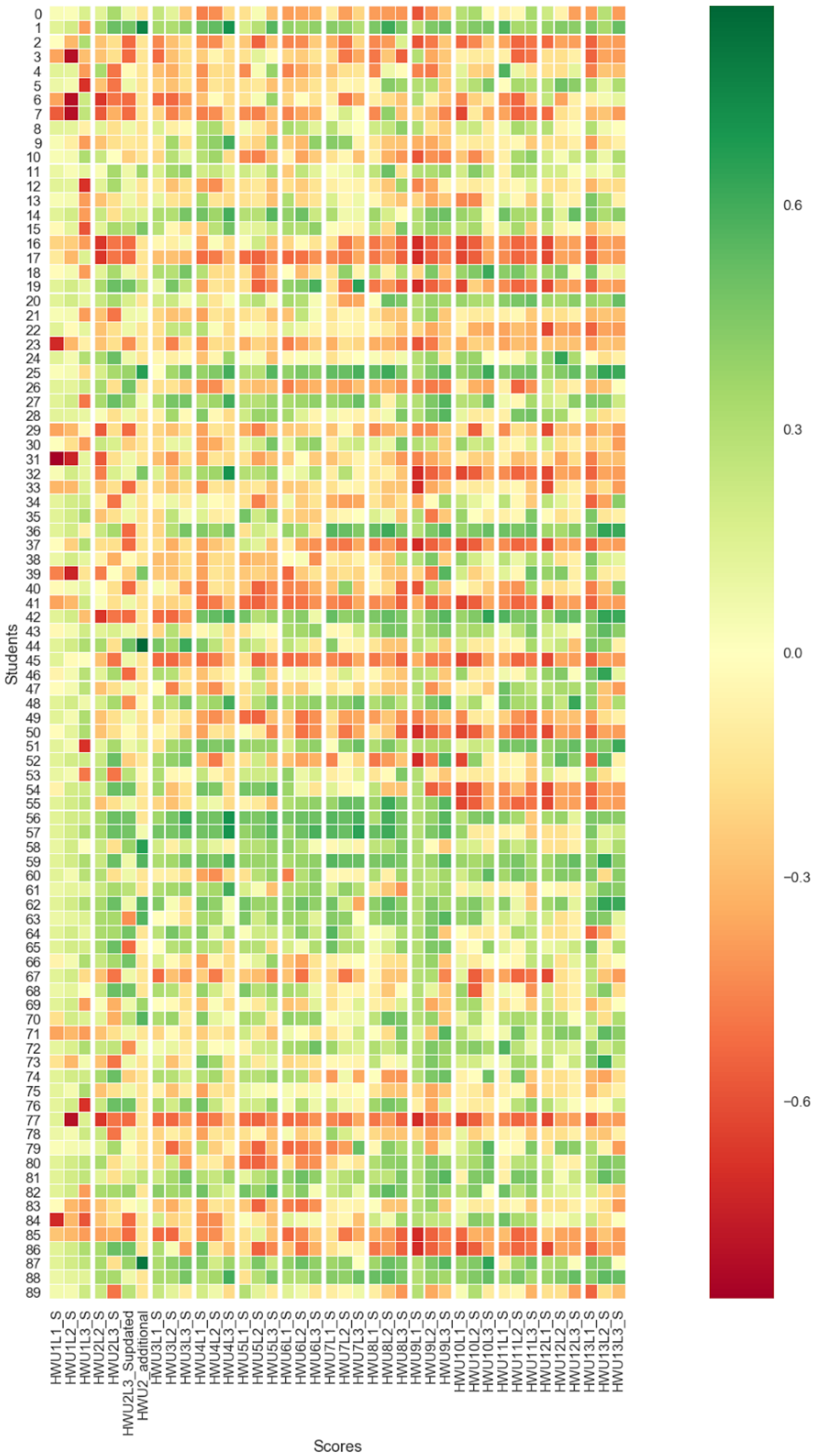


Figure 4.12: Course 1 Homework Scores of All Students.

In Figure 4.12, a downward trend can be observed within most units as content advances within the unit. In only one case, the downward trend continues to the consecutive unit (from Unit 1 to Unit 2). In all other cases, the trend is reversed with the next unit which restarts at a higher level (e.g., Unit 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13).



Students
Scores
HWU1L1_S
HWU1L2_S
HWU1L3_S
HWU2L2_S
HWU2L3_S
HWU2L3_Supdated
HWU2_additional
HWU3L1_S
HWU3L2_S
HWU3L3_S
HWU4L1_S
HWU4L2_S
HWU4L3_S
HWU5L1_S
HWU5L2_S
HWU5L3_S
HWU6L1_S
HWU6L2_S
HWU6L3_S
HWU7L1_S
HWU7L2_S
HWU7L3_S
HWU8L1_S
HWU8L2_S
HWU8L3_S
HWU9L1_S
HWU9L2_S
HWU9L3_S
HWU10L1_S
HWU10L2_S
HWU10L3_S
HWU11L1_S
HWU11L2_S
HWU11L3_S
HWU12L1_S
HWU12L2_S
HWU12L3_S
HWU13L1_S
HWU13L2_S
HWU13L3_S
0.6
0.3
0.0
−0.3
−0.6

Figure 4.13: Homework Scores of Course 1 in columns vs students in rows. Data is centered to a zero mean, which is shown in yellow color in the heat map. Thick white lines separate each unit. Column names: HW is homework; U X is for Unit X; L Y is Lesson Y. Raw scores range from 0 to 1. This visualization shows that there are three types of student performance levels based on homework scores; high level (scores above 0.6), average level (scores around zero), and low level (scores below -0.3). Students tend to maintain their performance level throughout the semester. If a student does well on homework, they keep up with the high level and vice versa, which is a different trend compared to class activity and exam scores.

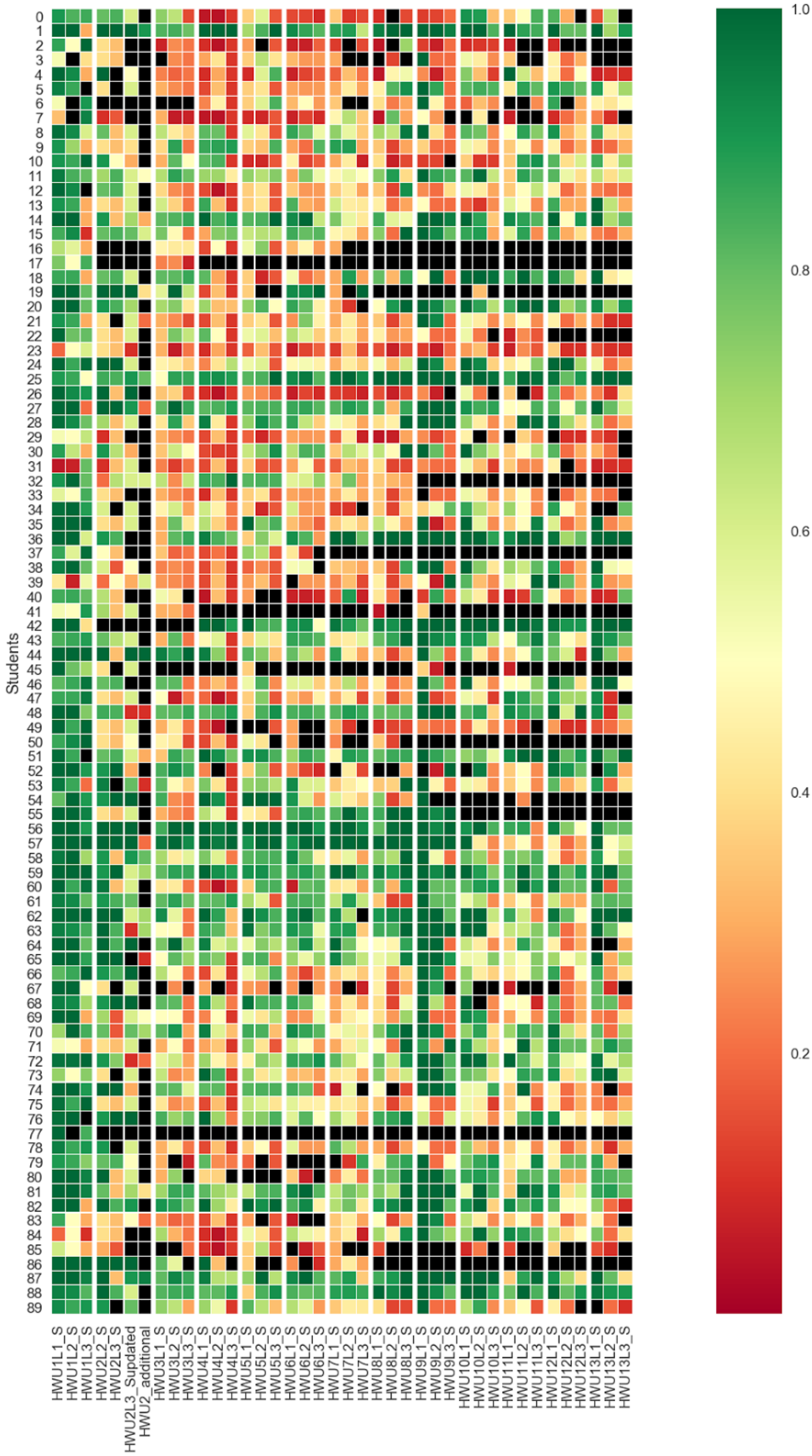

Students
Scores
HWU1L1_S
HWU1L2_S
HWU1L3_S
HWU2L2_S
HWU2L3_S
HWU2L3_Supdated
HWU2_additional
HWU3L1_S
HWU3L2_S
HWU3L3_S
HWU4L1_S
HWU4L2_S
HWU4L3_S
HWU5L1_S
HWU5L2_S
HWU5L3_S
HWU6L1_S
HWU6L2_S
HWU6L3_S
HWU7L1_S
HWU7L2_S
HWU7L3_S
HWU8L1_S
HWU8L2_S
HWU8L3_S
HWU9L1_S
HWU9L2_S
HWU9L3_S
HWU10L1_S
HWU10L2_S
HWU10L3_S
HWU11L1_S
HWU11L2_S
HWU11L3_S
HWU12L1_S
HWU12L2_S
HWU12L3_S
HWU13L1_S
HWU13L2_S
HWU13L3_S
1.0
0.8
0.6
0.4
0.2

Figure 4.14: Homework Scores of Course 1 in columns vs students in rows. Data is not normalized. Column names: HW is homework; U X is for Unit X; L Y is Lesson Y. Raw scores range from 0 to 1. Black color is a score of 0. This visualization helps distinguish between zero (shown in black) and very low scores. Students who do not attend the lessons fall into two different types: The first type do not do the homework, while the second type attempts/tries to do so, but still get very low scores.

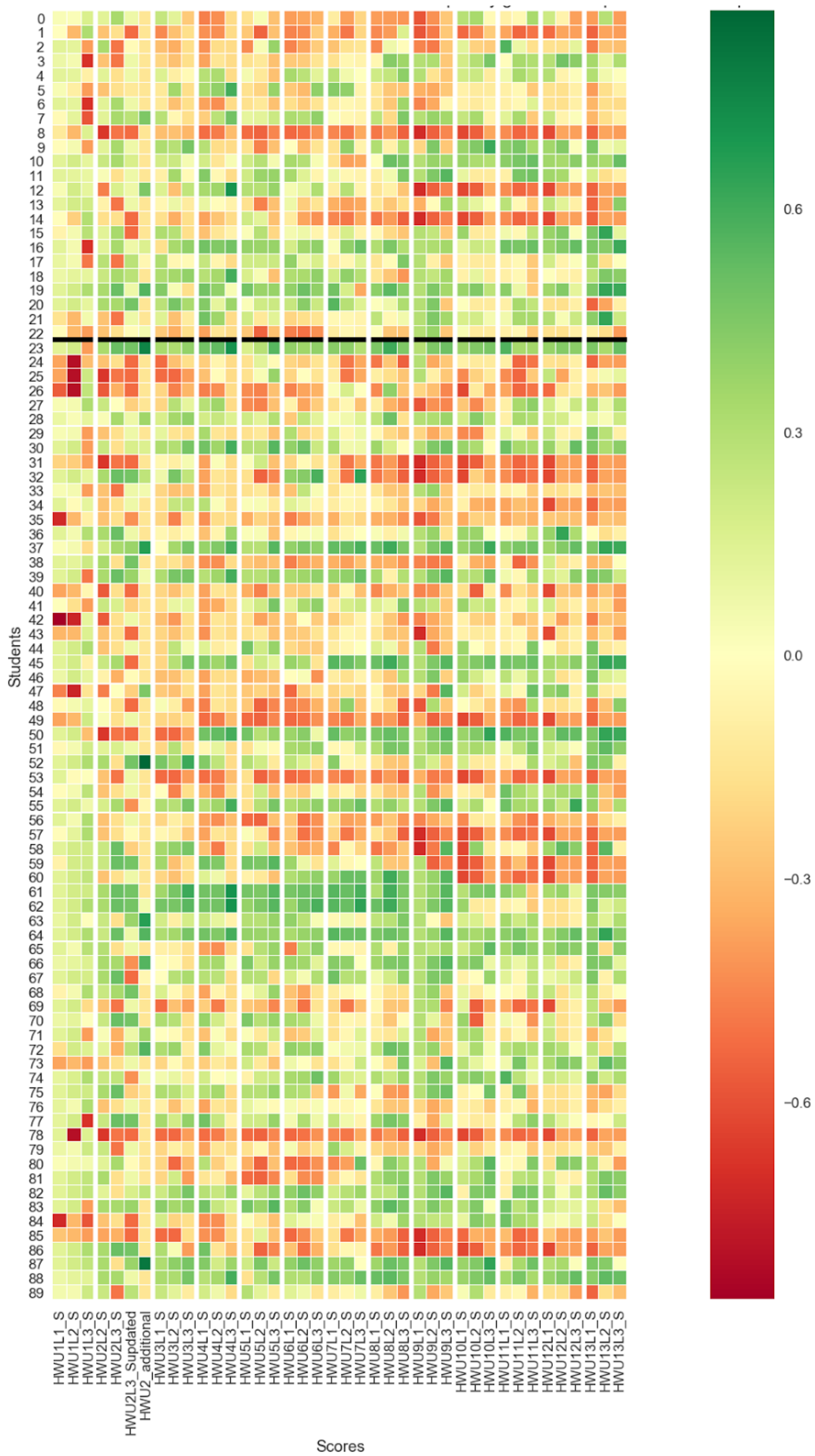

Students
Scores
HWU1L1_S
HWU1L2_S
HWU1L3_S
HWU2L2_S
HWU2L3_S
HWU2L3_Supdated
HWU2_additional
HWU3L1_S
HWU3L2_S
HWU3L3_S
HWU4L1_S
HWU4L2_S
HWU4L3_S
HWU5L1_S
HWU5L2_S
HWU5L3_S
HWU6L1_S
HWU6L2_S
HWU6L3_S
HWU7L1_S
HWU7L2_S
HWU7L3_S
HWU8L1_S
HWU8L2_S
HWU8L3_S
HWU9L1_S
HWU9L2_S
HWU9L3_S
HWU10L1_S
HWU10L2_S
HWU10L3_S
HWU11L1_S
HWU11L2_S
HWU11L3_S
HWU12L1_S
HWU12L2_S
HWU12L3_S
HWU13L1_S
HWU13L2_S
HWU13L3_S
0.6
0.3
0.0
−0.3
−0.6

Figure 4.15: Homework Scores of Course 1 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Column names: HW is homework; U X is for Unit X; L Y is Lesson Y. Raw scores range from 0 to 1. The black line separates gender; the first part is for female students and the second part is for male students. This visualization shows us that there is no significant difference in homework scores between female and male students.

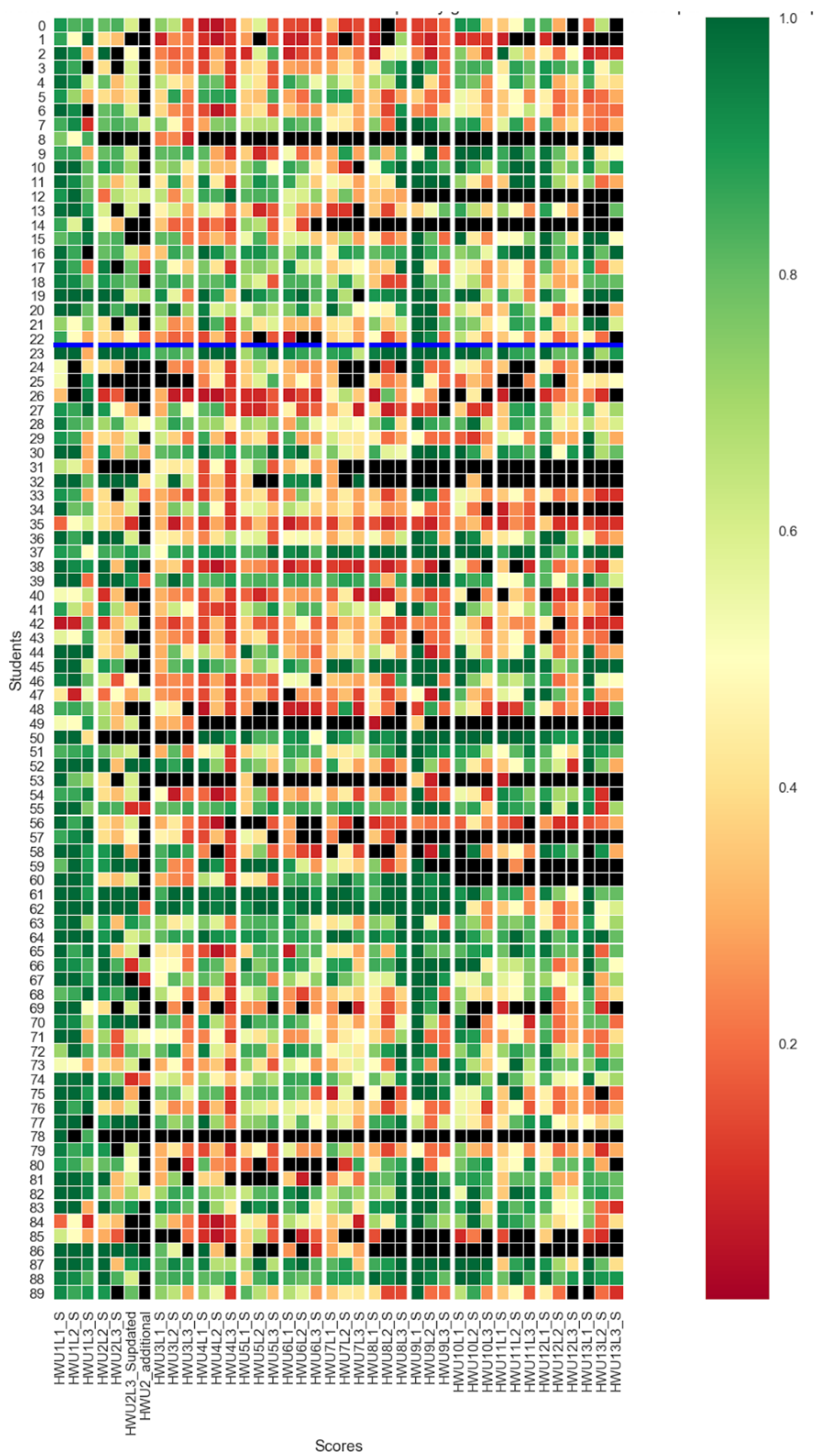


Figure 4.16: Homework Scores of Course 1 in columns vs students in rows. Data is not normalized. Column names: HW is homework; U X is for Unit X; L Y is Lesson Y. Raw scores range from 0 to 1. The blue line separates gender; the first part is for female students and the second part is for male students. This visualization helps

us distinguish between zero and very low scores. When considering zero, there is no significant difference between female and male students.

4.4.2.2 Class Activities

In this section, we present all visualizations related to class activities.

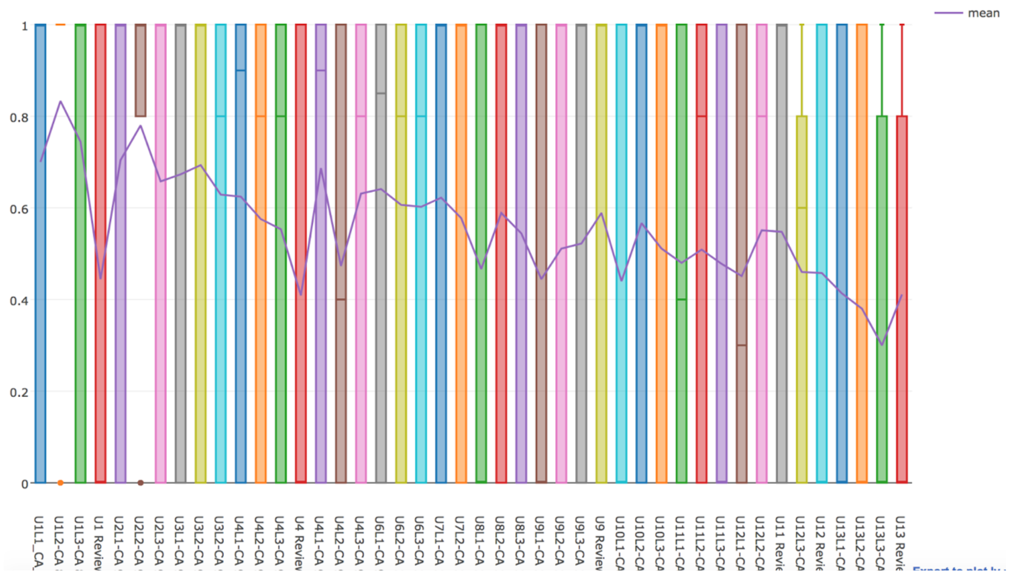


Figure 4.17: Course 1 Class Activity Scores of All Students. In some cases, (Unit 2, Unit 10) students start with a low activity score in a new unit then improve, and the get worse.

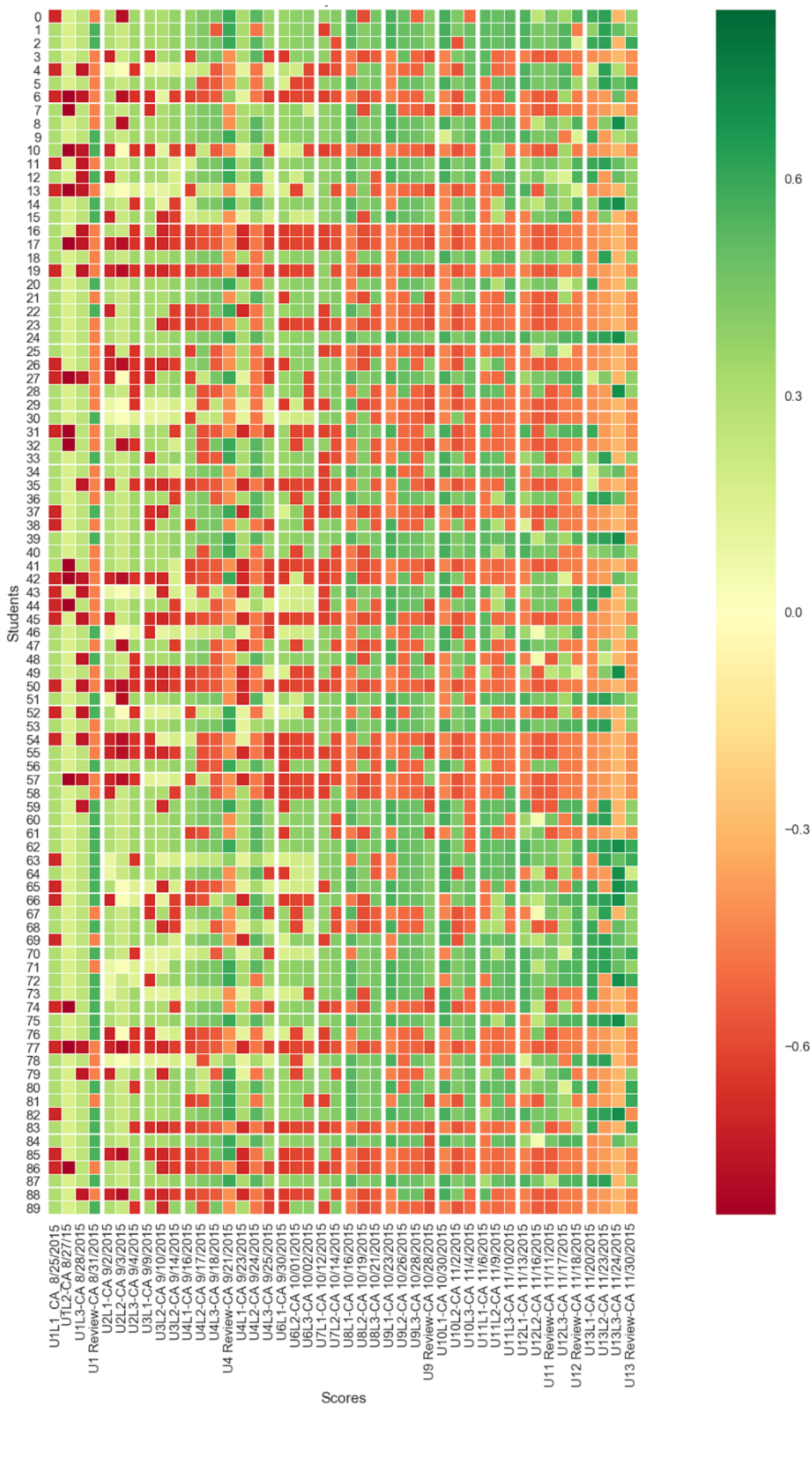
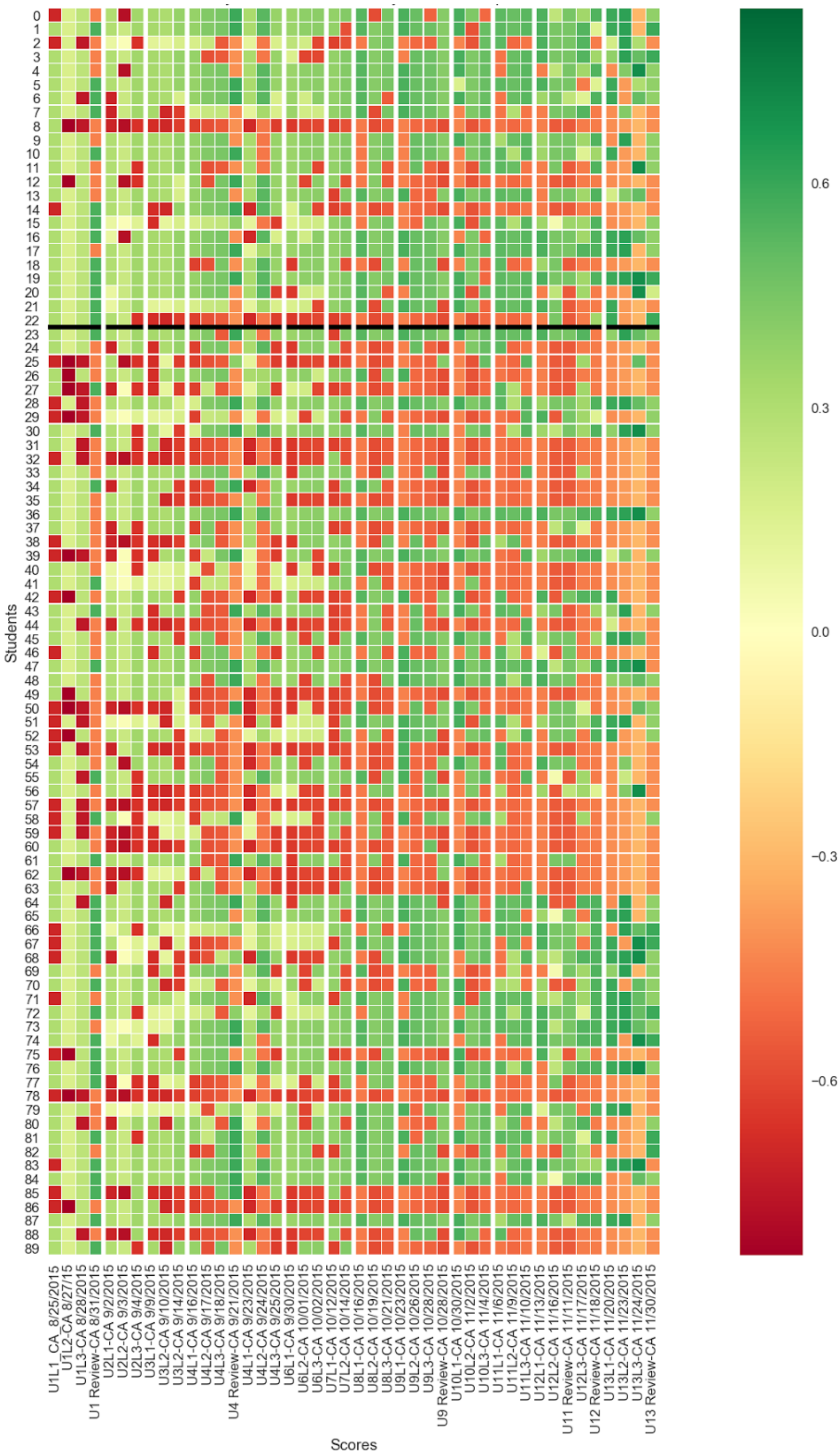

Students
Scores
U1L1_CA_8/25/2015
U1L2-CA 8/27/15
U1L3-CA 8/28/2015
U1 Review-CA 8/31/2015
U2L1-CA 9/2/2015
U2L2-CA 9/3/2015
U2L3-CA 9/4/2015
U3L1-CA 9/9/2015
U3L2-CA 9/10/2015
U3L2-CA 9/14/2015
U4L1-CA 9/16/2015
U4L2-CA 9/17/2015
U4L3-CA 9/18/2015
U4 Review-CA 9/21/2015
U4L1-CA 9/23/2015
U4L2-CA 9/24/2015
U4L3-CA 9/25/2015
U6L1-CA 9/30/2015
U6L2-CA 10/01/2015
U6L3-CA 10/02/2015
U7L1-CA 10/12/2015
U7L2-CA 10/14/2015
U8L1-CA 10/16/2015
U8L2-CA 10/19/2015
U8L3-CA 10/21/2015
U9L1-CA 10/23/2015
U9L2-CA 10/26/2015
U9L3-CA 10/28/2015
U9 Review-CA 10/28/2015
U10L1-CA 10/30/2015
U10L2-CA 11/2/2015
U10L3-CA 11/4/2015
U11L1-CA 11/6/2015
U11L2-CA 11/9/2015
U11L3-CA 11/10/2015
U12L1-CA 11/13/2015
U12L2-CA 11/16/2015
U11 Review-CA 11/11/2015
U12L3-CA 11/17/2015
U12 Review-CA 11/18/2015
U13L1-CA 11/20/2015
U13L2-CA 11/23/2015
U13L3-CA 11/24/2015
U13 Review-CA 11/30/2015
0.6
0.3
0.0
−0.3
−0.6

Figure 4.18: Class Activity Scores of Course 1in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Thick white lines separate each unit. U X is for Unit X; L Y is Lesson Y. Scores range from 0 to 1. In this visualization, we see that from unit 1 to 13 red color becomes darker to lighter color and greens are opposite of the red; they become lighter to darker. From this score change, we understand that class activities become more difficult than the previous unit and student performance reduced.



Students
Scores
U1L1_CA_8/25/2015
U1L2-CA 8/27/15
U1L3-CA 8/28/2015
U1 Review-CA 8/31/2015
U2L1-CA 9/2/2015
U2L2-CA 9/3/2015
U2L3-CA 9/4/2015
U3L1-CA 9/9/2015
U3L2-CA 9/10/2015
U3L2-CA 9/14/2015
U4L1-CA 9/16/2015
U4L2-CA 9/17/2015
U4L3-CA 9/18/2015
U4 Review-CA 9/21/2015
U4L1-CA 9/23/2015
U4L2-CA 9/24/2015
U4L3-CA 9/25/2015
U6L1-CA 9/30/2015
U6L2-CA 10/01/2015
U6L3-CA 10/02/2015
U7L1-CA 10/12/2015
U7L2-CA 10/14/2015
U8L1-CA 10/16/2015
U8L2-CA 10/19/2015
U8L3-CA 10/21/2015
U9L1-CA 10/23/2015
U9L2-CA 10/26/2015
U9L3-CA 10/28/2015
U9 Review-CA 10/28/2015
U10L1-CA 10/30/2015
U10L2-CA 11/2/2015
U10L3-CA 11/4/2015
U11L1-CA 11/6/2015
U11L2-CA 11/9/2015
U11L3-CA 11/10/2015
U12L1-CA 11/13/2015
U12L2-CA 11/16/2015
U11 Review-CA 11/11/2015
U12L3-CA 11/17/2015
U12 Review-CA 11/18/2015
U13L1-CA 11/20/2015
U13L2-CA 11/23/2015
U13L3-CA 11/24/2015
U13 Review-CA 11/30/2015
0.6
0.3
0.0
−0.3
−0.6

Figure 4.19: Class Activity Scores of Course 1 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Normalization makes the comparison meaningful mainly along one column. Comparison of one student's scores across different units is only meaningful for the student score evolution relative to the class average in each of those units, rather than an absolute comparison of the scores. Thick white lines separate each unit. The black line separates gender; the first part is for female students and the second part is for male students. U X is for Unit X; L Y is Lesson Y. Scores range from 0 to 1. Female students attend lessons more than male students. They get higher scores in class activity assignments. The scores (relative to class average) tend to improve in later units.

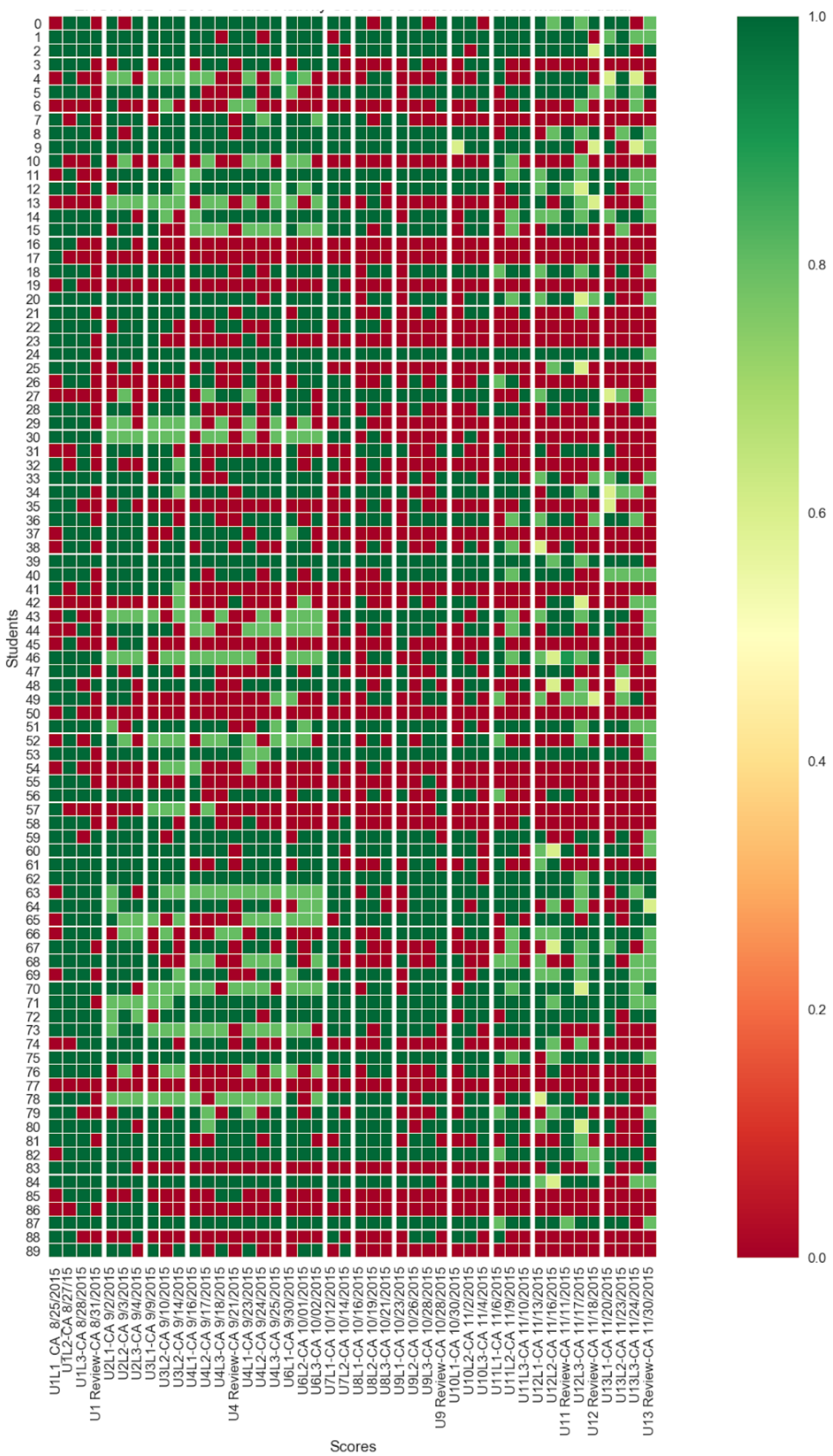
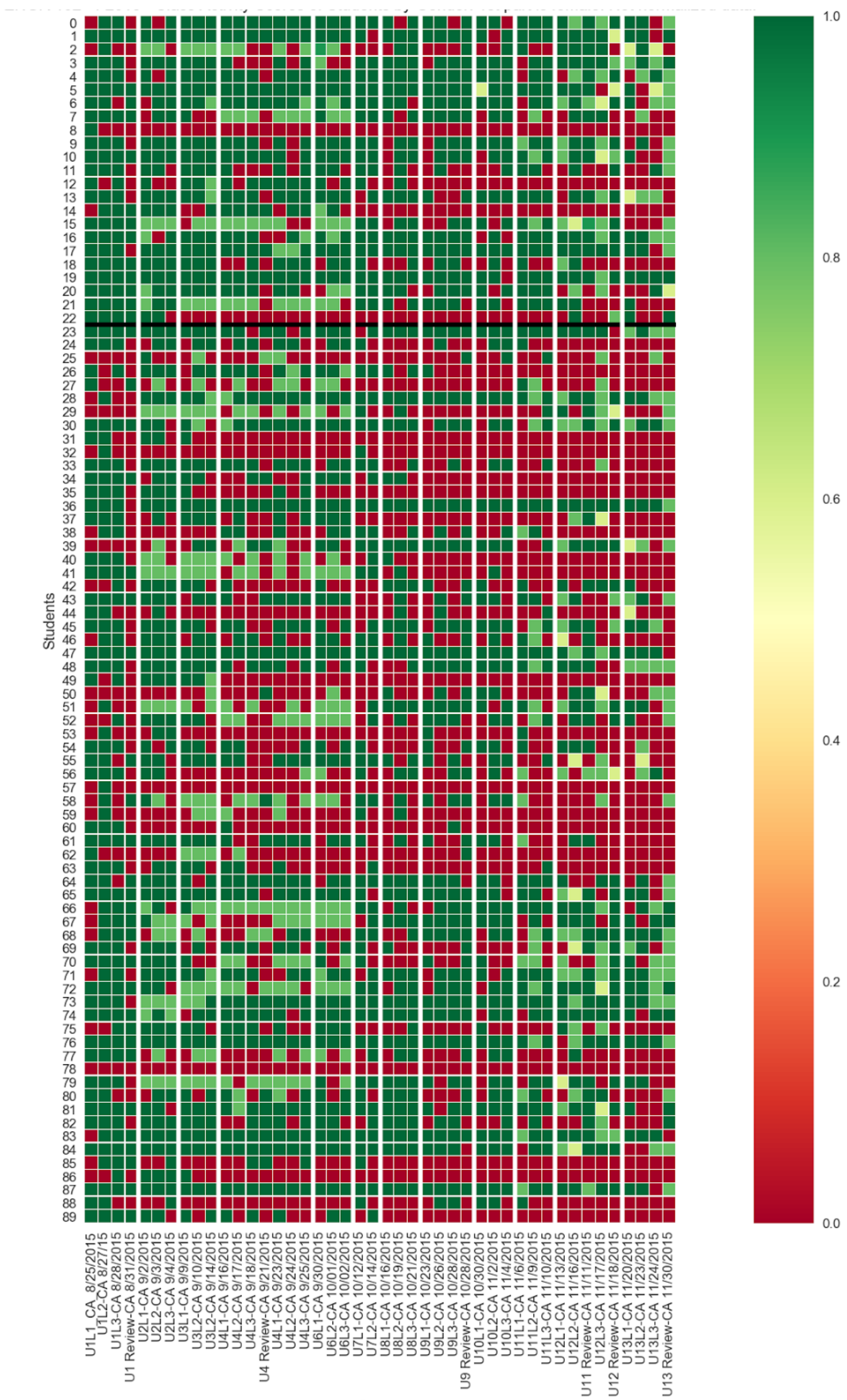

Students
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89
U1L1_CA_8/25/2015
U1L2-CA 8/27/15
U1L3-CA 8/28/2015
U1 Review-CA 8/31/2015
U2L1-CA 9/2/2015
U2L2-CA 9/3/2015
U2L3-CA 9/4/2015
U3L1-CA 9/9/2015
U3L2-CA 9/10/2015
U3L2-CA 9/14/2015
U4L1-CA 9/16/2015
U4L2-CA 9/17/2015
U4L3-CA 9/18/2015
U4 Review-CA 9/21/2015
U4L1-CA 9/23/2015
U4L2-CA 9/24/2015
U4L3-CA 9/25/2015
U6L1-CA 9/30/2015
U6L2-CA 10/01/2015
U6L3-CA 10/02/2015
U7L1-CA 10/12/2015
U7L2-CA 10/14/2015
U8L1-CA 10/16/2015
U8L2-CA 10/19/2015
U8L3-CA 10/21/2015
U9L1-CA 10/23/2015
U9L2-CA 10/26/2015
U9L3-CA 10/28/2015
U9 Review-CA 10/28/2015
U10L1-CA 10/30/2015
U10L2-CA 11/2/2015
U10L3-CA 11/4/2015
U11L1-CA 11/6/2015
U11L2-CA 11/9/2015
U11L3-CA 11/10/2015
U12L1-CA 11/13/2015
U12L2-CA 11/16/2015
U11 Review-CA 11/11/2015
U12L3-CA 11/17/2015
U12 Review-CA 11/18/2015
U13L1-CA 11/20/2015
U13L2-CA 11/23/2015
U13L3-CA 11/24/2015
U13 Review-CA 11/30/2015
Scores
1.0
0.8
0.6
0.4
0.2
0.0

Figure 4.20: Class Activity Scores of Course 1 in columns vs students in rows. Data is not normalized. Thick white lines separate each unit. U X is for Unit X; L Y is Lesson Y. Scores range from 0 to 1. Red color in the heat map represents a score of zero which means the student is absent. There is an advantage of not normalizing this data, since we can clearly see absences. Class Activity score can be only [0;0.6;0.8;1]. If the student attends lessons, even with low score in the activity, he/she gets mostly above 0.8, a score of 0.6 is rare. This visualization depicts the attendance of students. Almost half of the students do not attend the class regularly.



Students
Scores
U1L1_CA_8/25/2015
U1L2-CA 8/27/15
U1L3-CA 8/28/2015
U1 Review-CA 8/31/2015
U2L1-CA 9/2/2015
U2L2-CA 9/3/2015
U2L3-CA 9/4/2015
U3L1-CA 9/9/2015
U3L2-CA 9/10/2015
U3L2-CA 9/14/2015
U4L1-CA 9/16/2015
U4L2-CA 9/17/2015
U4L3-CA 9/18/2015
U4 Review-CA 9/21/2015
U4L1-CA 9/23/2015
U4L2-CA 9/24/2015
U4L3-CA 9/25/2015
U6L1-CA 9/30/2015
U6L2-CA 10/01/2015
U6L3-CA 10/02/2015
U7L1-CA 10/12/2015
U7L2-CA 10/14/2015
U8L1-CA 10/16/2015
U8L2-CA 10/19/2015
U8L3-CA 10/21/2015
U9L1-CA 10/23/2015
U9L2-CA 10/26/2015
U9L3-CA 10/28/2015
U9 Review-CA 10/28/2015
U10L1-CA 10/30/2015
U10L2-CA 11/2/2015
U10L3-CA 11/4/2015
U11L1-CA 11/6/2015
U11L2-CA 11/9/2015
U11L3-CA 11/10/2015
U12L1-CA 11/13/2015
U12L2-CA 11/16/2015
U11 Review-CA 11/11/2015
U12L3-CA 11/17/2015
U12 Review-CA 11/18/2015
U13L1-CA 11/20/2015
U13L2-CA 11/23/2015
U13L3-CA 11/24/2015
U13 Review-CA 11/30/2015
1.0
0.8
0.6
0.4
0.2
0.0

Figure 4.21: Class Activity Scores of Course 1 in columns vs students in rows. Data is not normalized. Thick white lines separate each unit. The black line separates gender; the first part is for female students and the second part is for male students. U X is for Unit X; L Y is Lesson Y. Scores range from 0 to 1. Red color in the heat map is equal to zero which means the student is absent. There is an advantage of not normalizing this data. Class Activity score can be only [0; 0.6; 0.8; 1]. If the student attends the class, even with low score in the activity, he/she gets mostly above 0.8, a score of 0.6 is rare. This visualization depicts the attendance of students. Absences increase after Unit 7, which is in the middle of the semester and is close to the last date to drop the course in the semester. Clearly, female students attend lessons more than male students. Male students do not attend lessons continuously and regularly.

4.4.2.3 Exams

In this section, we present all visualizations related to exam scores.

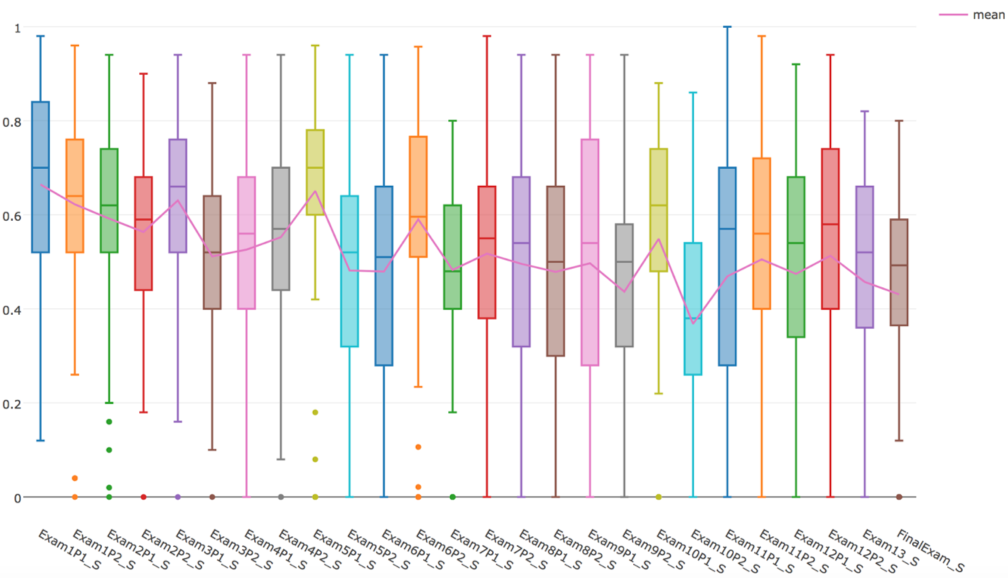


Figure 4.22: Course 1 Exam Scores. In general, exam scores decrease as the units advance.

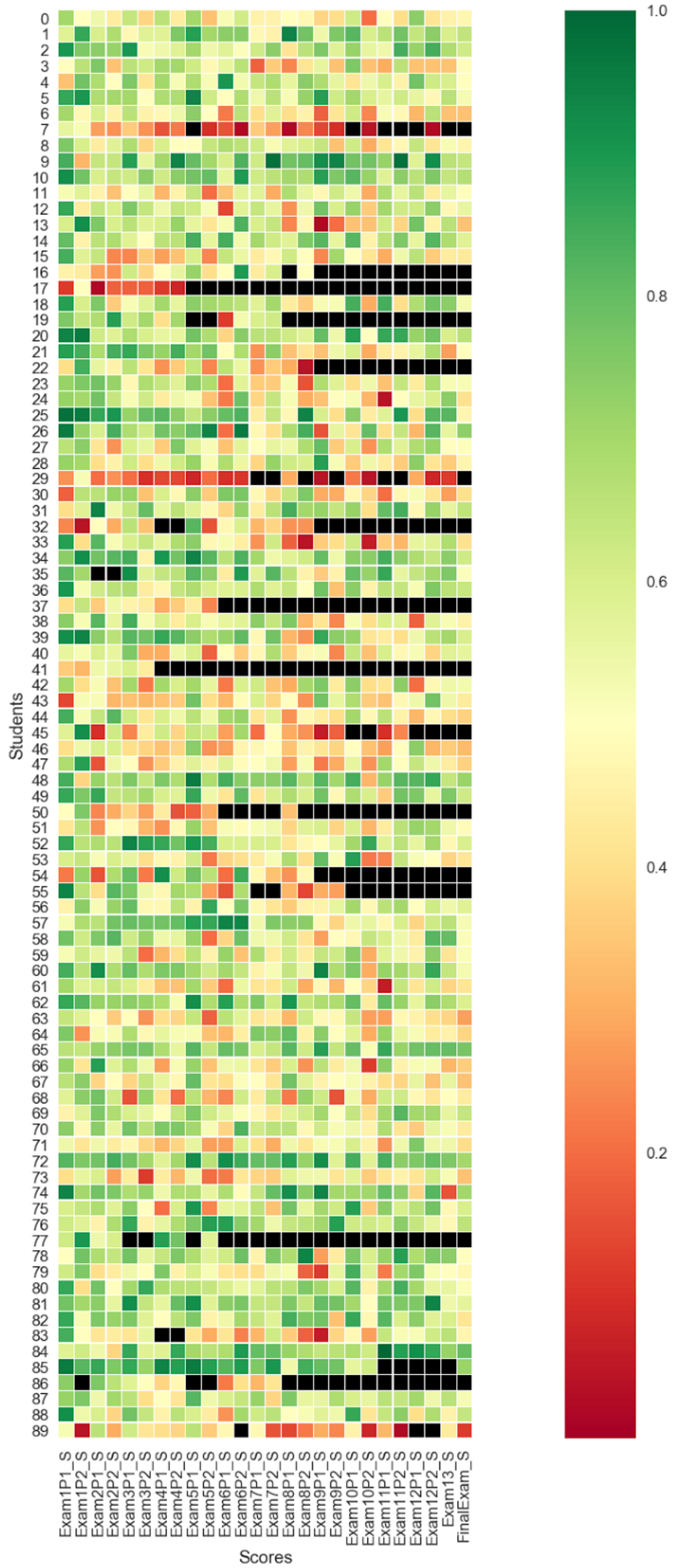


Figure 4.23: Exams Scores of Course 1 in columns vs students in rows. Data is not normalized. Exam X is for Unit X. There are two parts of exams in each unit; P1 is

part 1, P2 is part 2. Scores range from 0 to 1. The Black color is a score of 0 which means the student was absent. This visualization shows that students with consecutive absences drop out of the class; and he/she either does not attend or fails in the final exam. Overall drop out ratio is 1 in 5.58 students and 16% of the whole class. Most students drop after Unit 7. We can also see how in most cases; exam scores decrease after absences in previous exams.

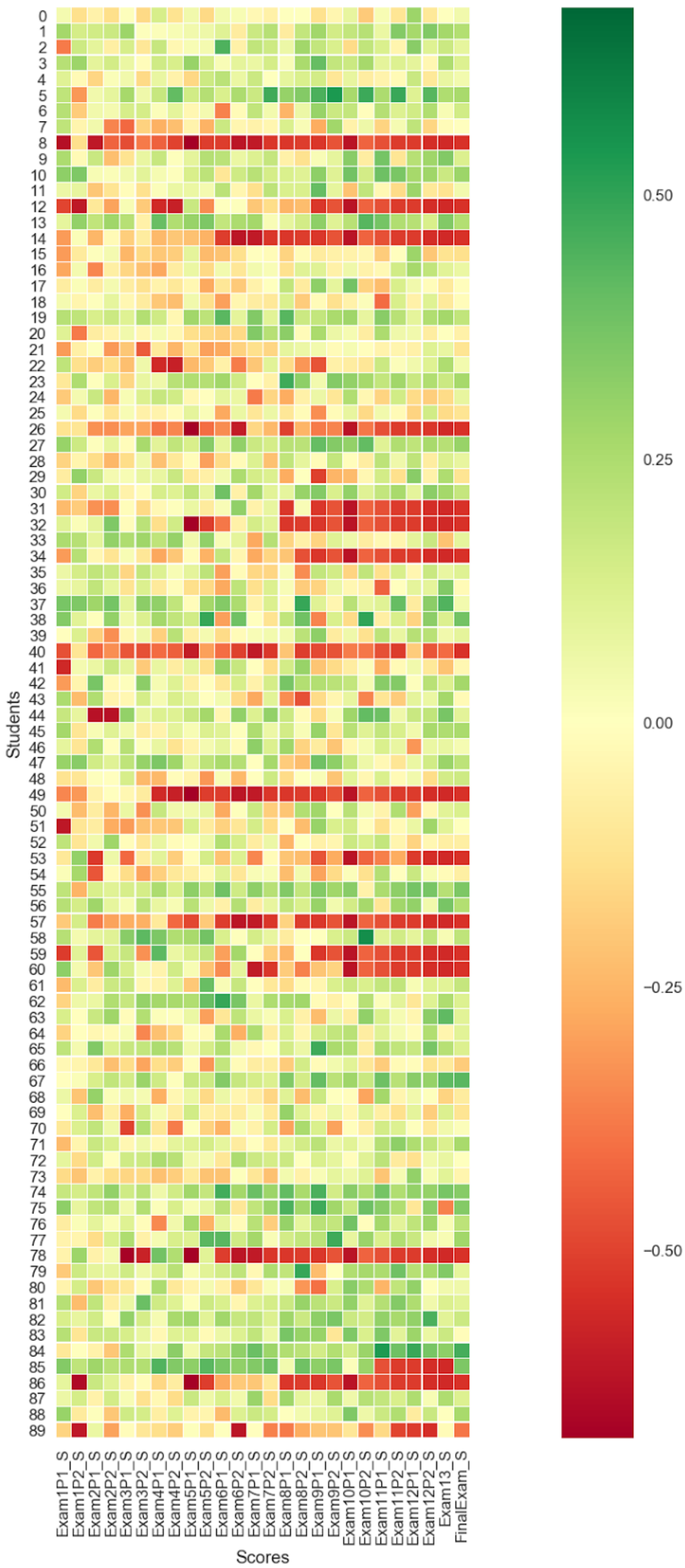


Figure 4.24: Exams Scores of Course 1 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Exam X is for Unit X.

There are two parts of exams in each unit; P1 is part 1, P2 is part 2. Scores range from 0 to 1. This visualization shows that students with scores below the average tend to get scores below the average in the final exam. By combining this plot with Figure 4.23, we can say that when we divide the semester into two parts, namely before and after Unit 7, there are three kinds of student behavior. The first type (Example: students 8, 40 and 78) perform below average in the first part of the semester, and get very low scores or drop the course. The second type (Example: students 1, 2, and 4) perform below average in the first part of the semester, but get better scores after Unit 7 when compared with the first part of the course. We can say that by the middle of the semester, student behavior may have changed and this has an impact on whether they pass the class or fail. Another student type gets a score above average and experiences continuous success; these correspond to the greener scores in the heat map.

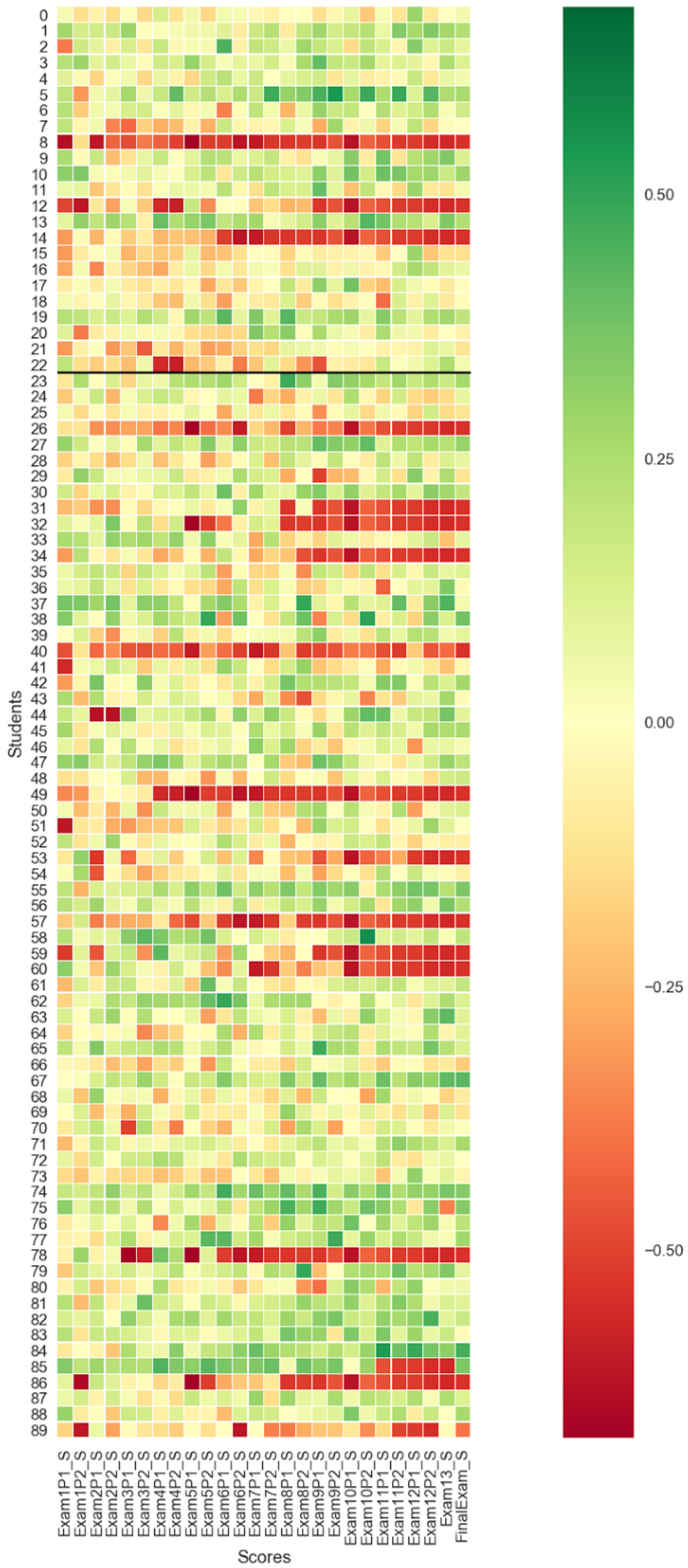


Figure 4.25: Exams Scores of Course 1 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Exam X is for Unit X.

There are two parts of exams in each unit; P1 is part 1, P2 is part 2. Scores range from 0 to 1. The black line separates gender; the first part is for female students and the second part is for male students. In addition to other plots, this visualization shows that male students tend to fail the class more than females. 1 in 7.6 female students drop the class; on the other hand, 1 in 5.5 male students drop this class. Overall drop out ratio is 1 in 5.58 students and 16% of the whole class. Also, with the exception of students who end up dropping the class, there is an improvement trend in scores towards later exams: The right side of the plot shows greener and less orange cells than the left side.

4.5 The Introductory Calculus for Engineers Course 2

The Introductory Calculus for Engineers Course 2 data, corresponds to 100 students with 297 attributes, consisting of combined homework, lesson assignments, class activity, and exam scores.

4.5.1 EFA for Course 2

In this section, we present the exploratory factor analysis results for all scores.

4.5.1.1 All Scores

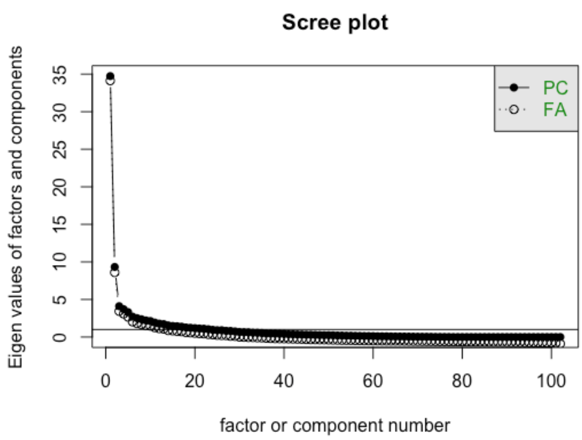


Figure 4.26: Scree Plot of Course 1

From the scree plot in Figure 4.26, we observe that there are two significant factors over 102 variables in the data set for students who took Course 1. Class activities and exams are grouped into factor 1; while homework assignments and lesson assignments are grouped into factor 2.

Table 4: Scree Plot of All Scores for Course 1

Factor 1	Class activities and exams
Factor 2	Homework and lesson assignments

Loadings:

	MR1	MR2
U1-L1&2-CA	0.580	
U1-L1-S		0.608
U1-L2-S		0.626
U1-L3-CA	0.480	
U1-L3-S		0.695
U1-L4-CA	0.514	
U1-L1-4-CA	0.585	
U1-L4-S		0.730
U1-L1-CA	0.558	
U1-R1-CA	0.591	
U1-L5-S		0.522
U1-HW1-S		0.600
U1-HW1-PM-S		0.529
U1-HW1-PM-S		0.580
Exam1		
U2-L2-CA	0.609	
U2-L2-S		0.527
U2-L3-CA	0.588	
U2-L3-S		0.655
U2-L4-CA	0.632	
U2-L4-S		0.698
U2-RD-CA	0.716	
U2-R1-CA	0.557	
U2-L5-CA	0.528	
U2-L5-S		0.769
U2-L6-CA	0.510	
U2-L6-S		0.484
U2-R2-CA	0.606	
U2-L7-CA	0.631	
U2-L7-S		0.722
U2-L9-CA	0.576	
U2-L9-S		0.618
U2-L6-8-CA	0.556	
U2-L8-S		0.641
Unit 2 HWK-1-S		0.586
U2-HWK2-S		0.496
U2-HWK3-S		0.529
Exam2	0.402	
U3-L1-CA	0.694	
U3-L1-4R-CA	0.729	
U3-L1-S		0.517
U3-L2-CA	0.577	
U3-L2-S		0.427
U3-L3-4-CA	0.799	
U3-L3-S		0.745
U3-L4-S		0.717
U3-R2-CA	0.721	
U3-R1-CA	0.737	
U3-HWK-P1(L1-L4)-S		0.525
U3-L5-CA	0.638	
U3-L5P2-CA	0.597	
U3-L5-S		0.628
U3-L6-CA	0.696	
U3-HWK-P2(L5)-S		0.629
U3-L6-S		0.784
U3-L7-CA	0.707	
U3-L7P2-CA	0.672	
U3-L7-S		0.668
U3-HWK-P3(L6-L7)-S		0.596
U3-L8-CA	0.578	
U3-L8-S		0.525
Exam 3	0.437	
U4-L1-2-CA	0.744	
U4-L1-CA	0.697	
U4-L1-S		0.573
U4-L2-CA	0.566	
U4-L2-S		0.577
U4-L3-CA	0.579	
U4-L3-S		0.632
U4-R1-CA	0.725	
U4-R2-CA	0.696	
U4-HWK-Part 1(L1-L2)-S	0.496	0.444
U4-HWK-P2(L3)-S	0.456	0.444
Exam 4	0.529	
MidTerm	0.502	
U5-D1-CA	0.695	
U5-L1-CA	0.683	
U5-L1-S		0.546
U5-L2-CA	0.645	
U5-L2-3P2-CA	0.613	
U5-L2-3-CA	0.622	
U5-L2-S		0.503
U5-L3-S		
U5-R2-CA	0.629	
U5-R1-CA	0.630	
U5-L5-CA	0.599	
U5-L4-S		0.553
U5-L5-S		0.432
U5-HWK-P1-S	0.547	0.423
U5-HWK-P2-S	0.568	
U5-HWK-P3-S	0.554	
Exam 5	0.615	
U6-L1-CA	0.690	
U6-R1-CA	0.721	
U6-L1-S	0.588	
U6-L2-CA	0.636	
U6-L2-S	0.515	
U6-HWK-S	0.572	
U6-L3-CA	0.644	
U6-L3-S	0.649	
Exam 6	0.666	
Final	0.539	

	MR1	MR2
SS loadings	25.453	15.884
Proportion Var	0.250	0.156
Cumulative Var	0.250	0.405

Figure 4.27: Factor Loadings All Scores of Course 1

4.5.2 Visual Data Analysis for Course 2

We separated this section into four parts as follows:

1. Homework and lesson assignments,
2. Class activities,
3. Exams, and
4. All scores.

In the following subsections, we present all related visualizations.

4.5.2.1 Homework and Lesson Assignments

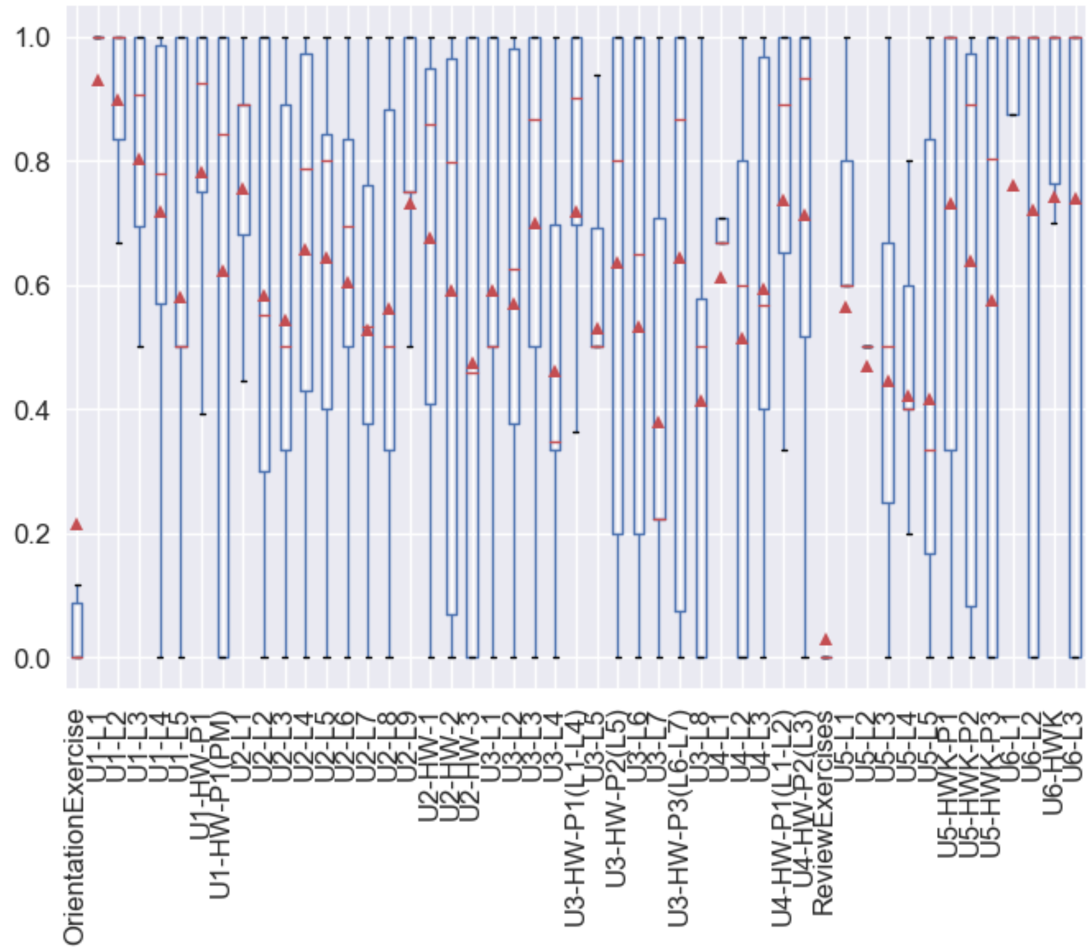


Figure 4.28: Boxplots of Course 2 Homework and Lesson Assignments of All Students. The red line is the median and the red triangle is mean. Scores range between 0 and 1. This boxplot supports all inferences that we made in the heat maps.

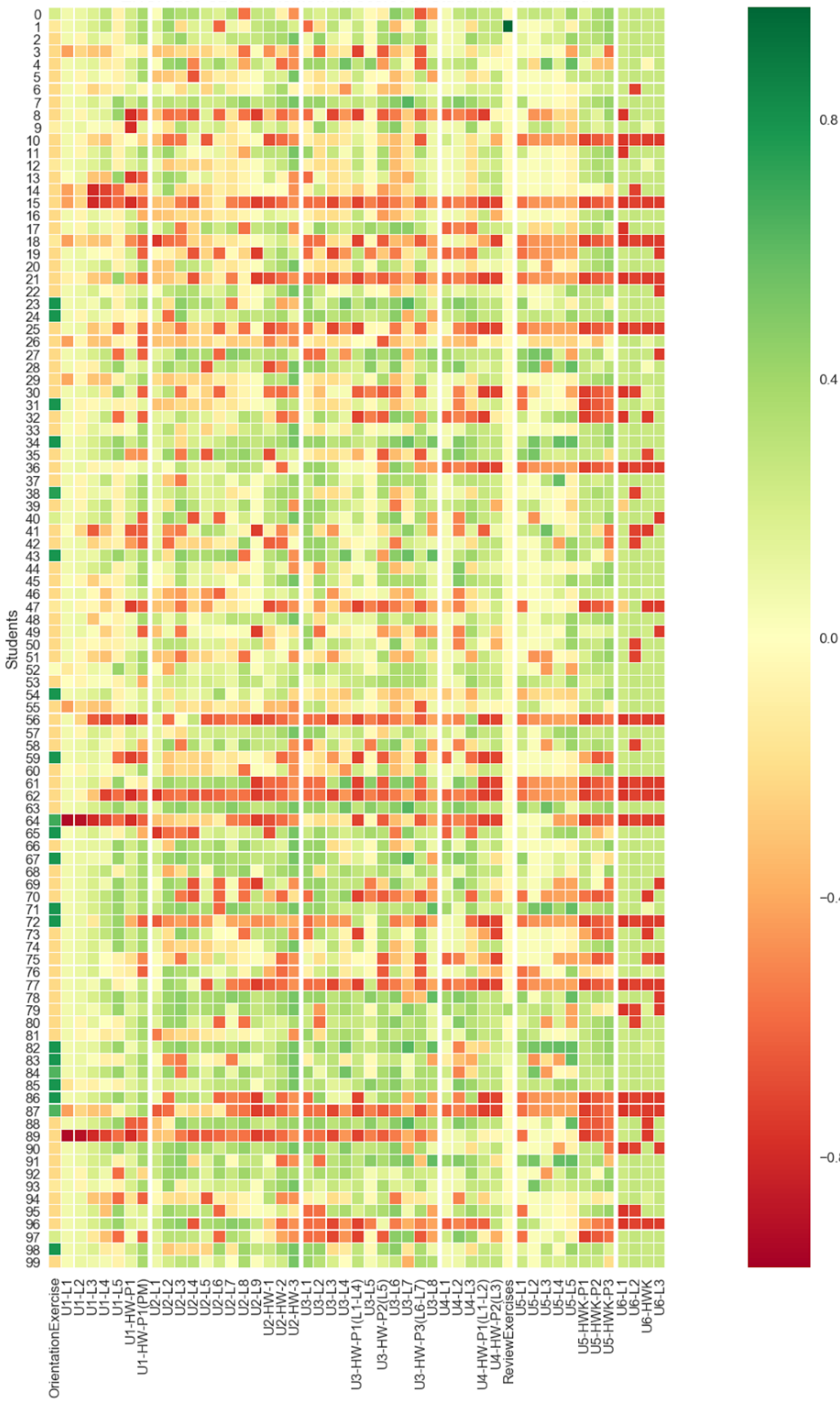

Students
Scores
0.8
0.4
0.0
−0.4
−0.8
OrientationExercise
U1-L1
U1-L2
U1-L3
U1-L4
U1-L5
U1-HW-P1
U1-HW-P1(PM)
U2-L1
U2-L2
U2-L3
U2-L4
U2-L5
U2-L6
U2-L7
U2-L8
U2-L9
U2-HW-1
U2-HW-2
U2-HW-3
U3-L1
U3-L2
U3-L3
U3-L4
U3-HW-P1(L1-L4)
U3-L5
U3-HW-P2(L5)
U3-L6
U3-L7
U3-HW-P3(L6-L7)
U3-L8
U4-L1
U4-L2
U4-L3
U4-HW-P1(L1-L2)
U4-HW-P2(L3)
ReviewExercises
U5-L1
U5-L2
U5-L3
U5-L4
U5-L5
U5-HWK-P1
U5-HWK-P2
U5-HWK-P3
U6-L1
U6-L2
U6-HWK
U6-L3

Figure 4.29: Heat Map of Homework and Lesson Assignments Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. U X is for Unit X; L Y is Lesson Y; HW/K is homework. Raw scores range from 0 to 1. Vertical white lines separate each unit. This visualization shows us that students tend to maintain their performance level. There is significant no change across units. However, we clearly observe that throughout Unit 1-6, there is a difference between homework scores and lesson assignments. From lesson assignments to homework, the color changes light orange to light green. This means that students have better performance in homework than the lesson assignments. Also, this shows that students have better performance on homework after working in class collaboratively on the same topic.

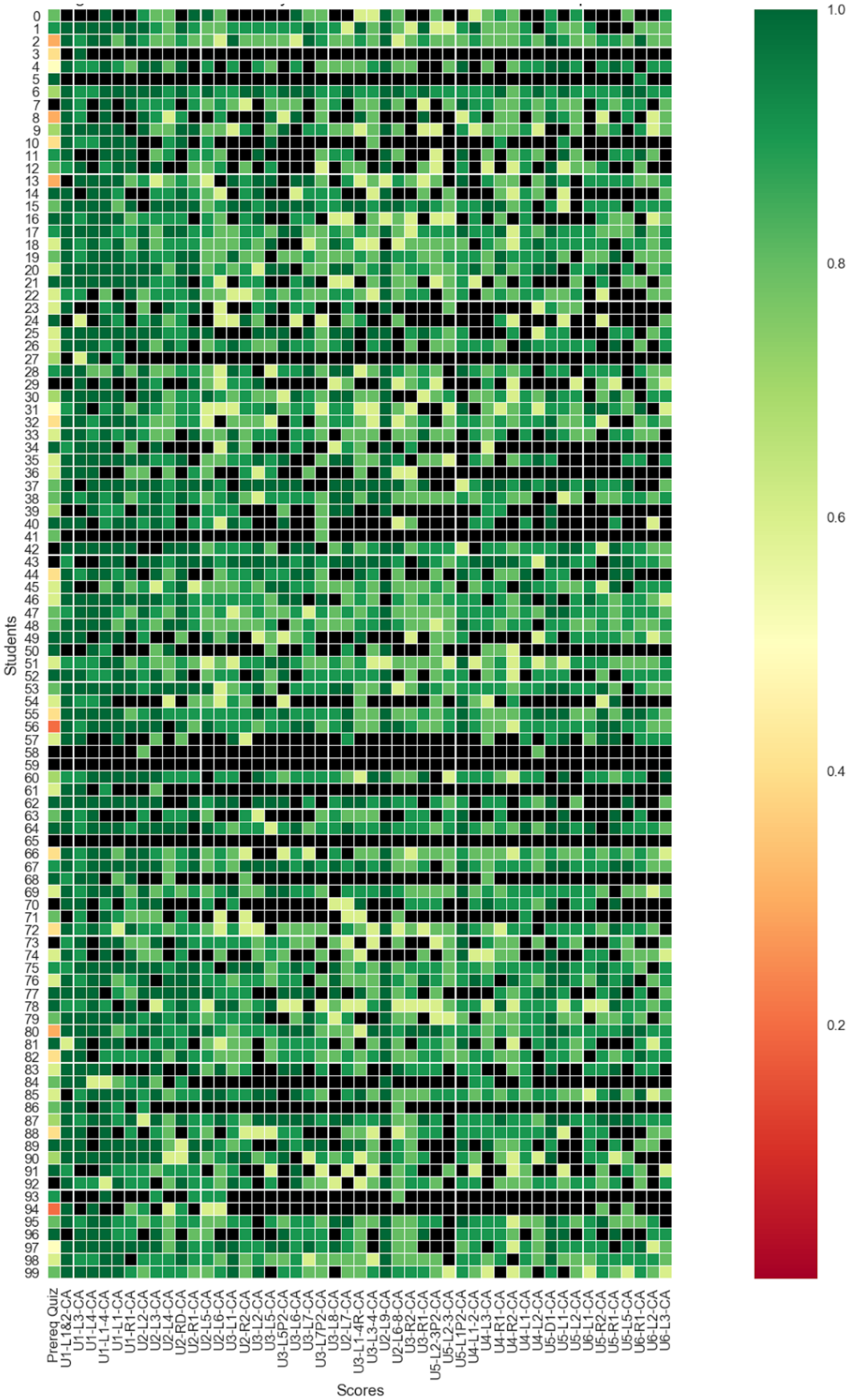

Students
Scores
Prereq Quiz
U1-L1&2-CA
U1-L3-CA
U1-L4-CA
U1-L1-4-CA
U1-L1-CA
U1-R1-CA
U2-L2-CA
U2-L3-CA
U2-L4-CA
U2-RD-CA
U2-R1-CA
U2-L5-CA
U2-L6-CA
U3-L1-CA
U2-R2-CA
U3-L2-CA
U3-L5-CA
U3-L5P2-CA
U3-L6-CA
U3-L7-CA
U3-L7P2-CA
U3-L8-CA
U2-L7-CA
U3-L1-4R-CA
U3-L3-4-CA
U2-L9-CA
U2-L6-8-CA
U3-R2-CA
U3-R1-CA
U5-L2-3P2-CA
U5-L2-3-CA
U5-L1P2-CA
U4-L1-2-CA
U4-L3-CA
U4-R1-CA
U4-R2-CA
U4-L1-CA
U4-L2-CA
U5-D1-CA
U5-L1-CA
U5-L2-CA
U6-L1-CA
U5-R2-CA
U5-R1-CA
U5-L5-CA
U6-R1-CA
U6-L2-CA
U6-L3-CA
1.0
0.8
0.6
0.4
0.2

Figure 4.30: Heat Map of Class Activity Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; CA is class activity. Scores range from 0 to 1. Black color in the heat map is equal to zero which means the student is absent. There is an advantage of not normalizing this data. Class Activity score can be only [0; 0.6; 0.8; 0.9; 1]. If the student attends the class, even with low score in the activity, he/she scores minimum 0.6. This visualization shows the attendance of students and shows clearly consecutive absences.

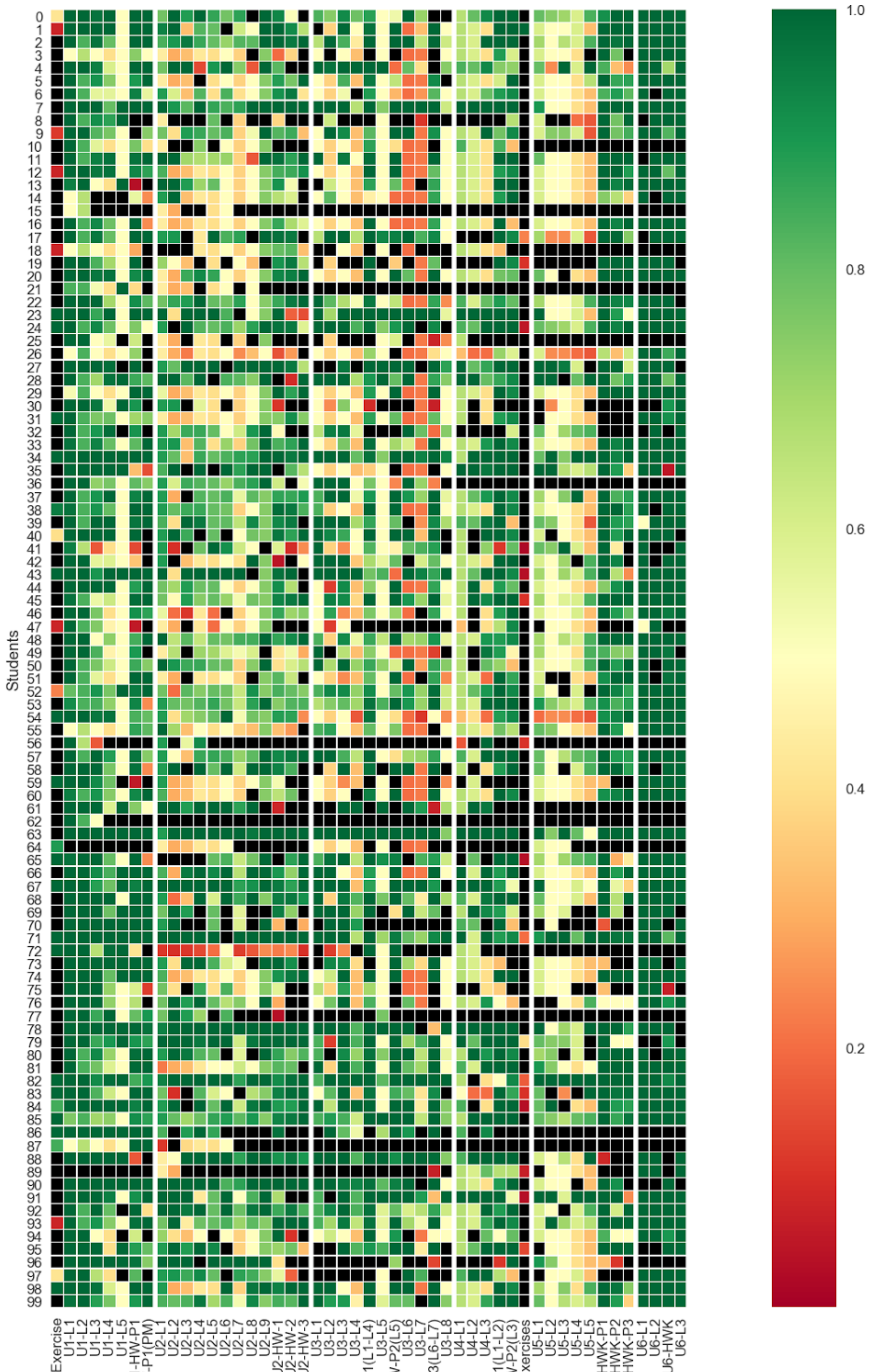

Students
Scores
OrientationExercise
U1-L1
U1-L2
U1-L3
U1-L4
U1-L5
U1-HW-P1
U1-HW-P1(PM)
U2-L1
U2-L2
U2-L3
U2-L4
U2-L5
U2-L6
U2-L7
U2-L8
U2-L9
U2-HW-1
U2-HW-2
U2-HW-3
U3-L1
U3-L2
U3-L3
U3-L4
U3-HW-P1(L1-L4)
U3-L5
U3-HW-P2(L5)
U3-L6
U3-L7
U3-HW-P3(L6-L7)
U3-L8
U4-L1
U4-L2
U4-L3
U4-HW-P1(L1-L2)
U4-HW-P2(L3)
ReviewExercises
U5-L1
U5-L2
U5-L3
U5-L4
U5-L5
U5-HWK-P1
U5-HWK-P2
U5-HWK-P3
U6-L1
U6-L2
U6-HWK
U6-L3
1.0
0.8
0.6
0.4
0.2

Figure 4.31: Heat Map of Homework and Lesson Assignment Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; HW/K is homework. Scores range from 0 to 1. Black cells represent a score of zero. Vertical white lines separate each unit. We see from the heat map that Unit 3 Lesson 6-7 and Unit 5 Lesson 2-5 are the hardest lessons for the majority of students. Also, students performed better in Unit 6, which is the last unit, than any other unit. We also observe how in several units, e.g. Unit 5, the homework scores are significantly better than the Lesson scores, and that the last lesson score before a homework score tends to be the lowest compared to preceding lesson scores in the same unit.

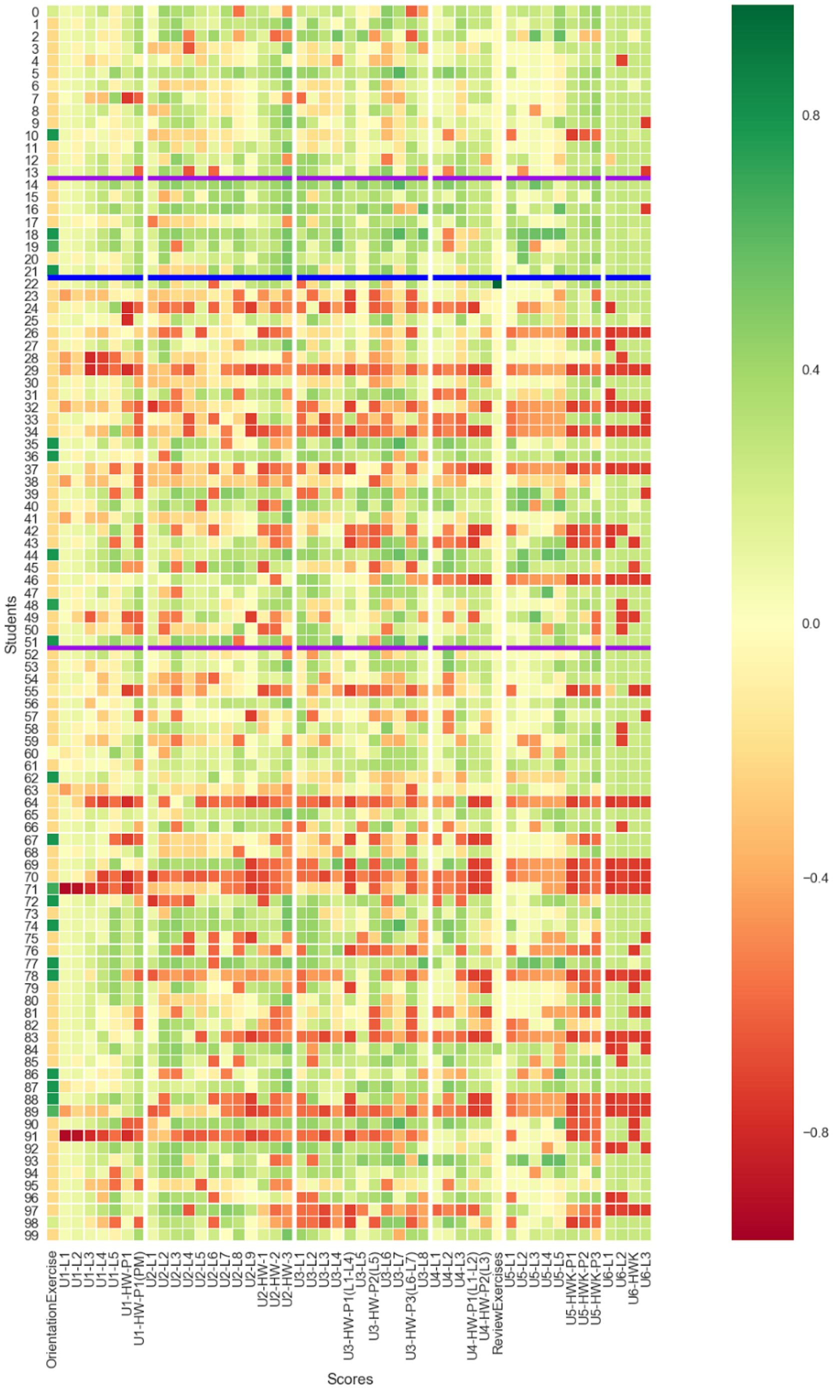


Figure 4.32: Heat Map of Homework and Lesson Assignments Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is shown in yellow color in the heat map. U X is for Unit X; L Y is Lesson Y; HW/K is homework. Vertical white lines separate each unit. The blue line separates gender; the first part is for female students and the second part is for male students.. Horizontal purple lines

separate students who took the flipped class [Course 1] before (within the same gender) from those who did not. The first part, above the purple line in the same gender group, are the students who took the flipped class before and the second part (below the purple line) did not take the flipped class prior to this class. This heat map clearly shows that female students have better performance than male students in homework. When we consider each gender separately, and compare scores depending on whether the flipped class had been taken before, we notice that for males, there is no difference; however, within the female group, those students who did not take the flipped class before, seem to have better performance than the group who did take the flipped class before. However, the sample size is too small for any meaningful conclusion. Also, students who did not take the flipped class before are actually retaking the class.

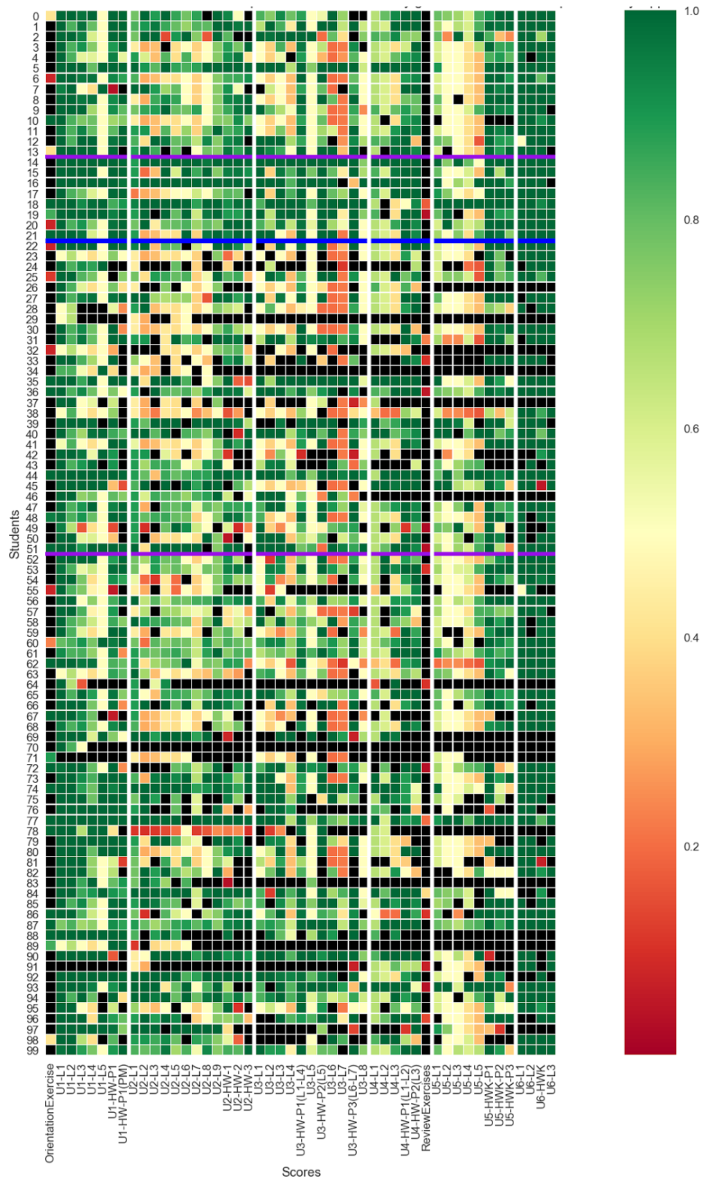


Figure 4.33: Heat Map of Homework and Lesson Assignments Scores Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; HW/K is homework. Black cells are equal to zero. Vertical white lines separate each unit. The blue line separates gender; the first part is for female students and the second part is for male students. The horizontal purple line separates those who took the flipped class [Course 1] before from those who did not within the same gender. The first part above the purple line took the flipped class and the second part did not take it. This heat map helps us to distinguish a zero from a low score. By not

considering the zero scores, in Unit 3 Lesson 6-7, most of the students have the lowest scores in lesson assignments compared to the other units. This shows that they may have had a hard time understanding these topics by themselves. When we check the homework score, which is due after the class meeting, they performed better compared to the lesson assignment. Also, we see that the same trend happened in Unit 5. This specific example may show the impact of collaborative learning.

4.5.2.2 Class Activities

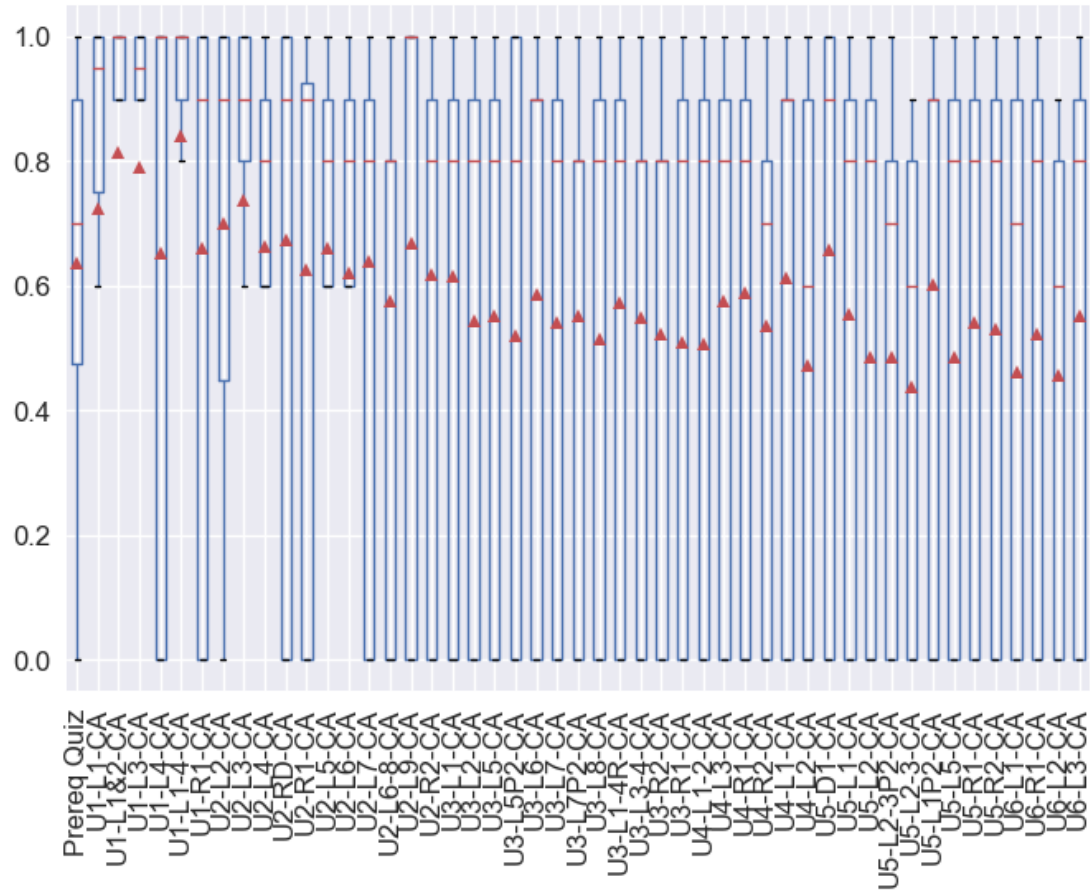


Figure 4.34: Boxplots of Course 2 Class Activity Scores of All Students. Unit 1-2 have higher average than unit 3-6. For all scores, the mean is lower than the median. The distribution is skewed to the left. There are major outliers in the left tail, which are absences.

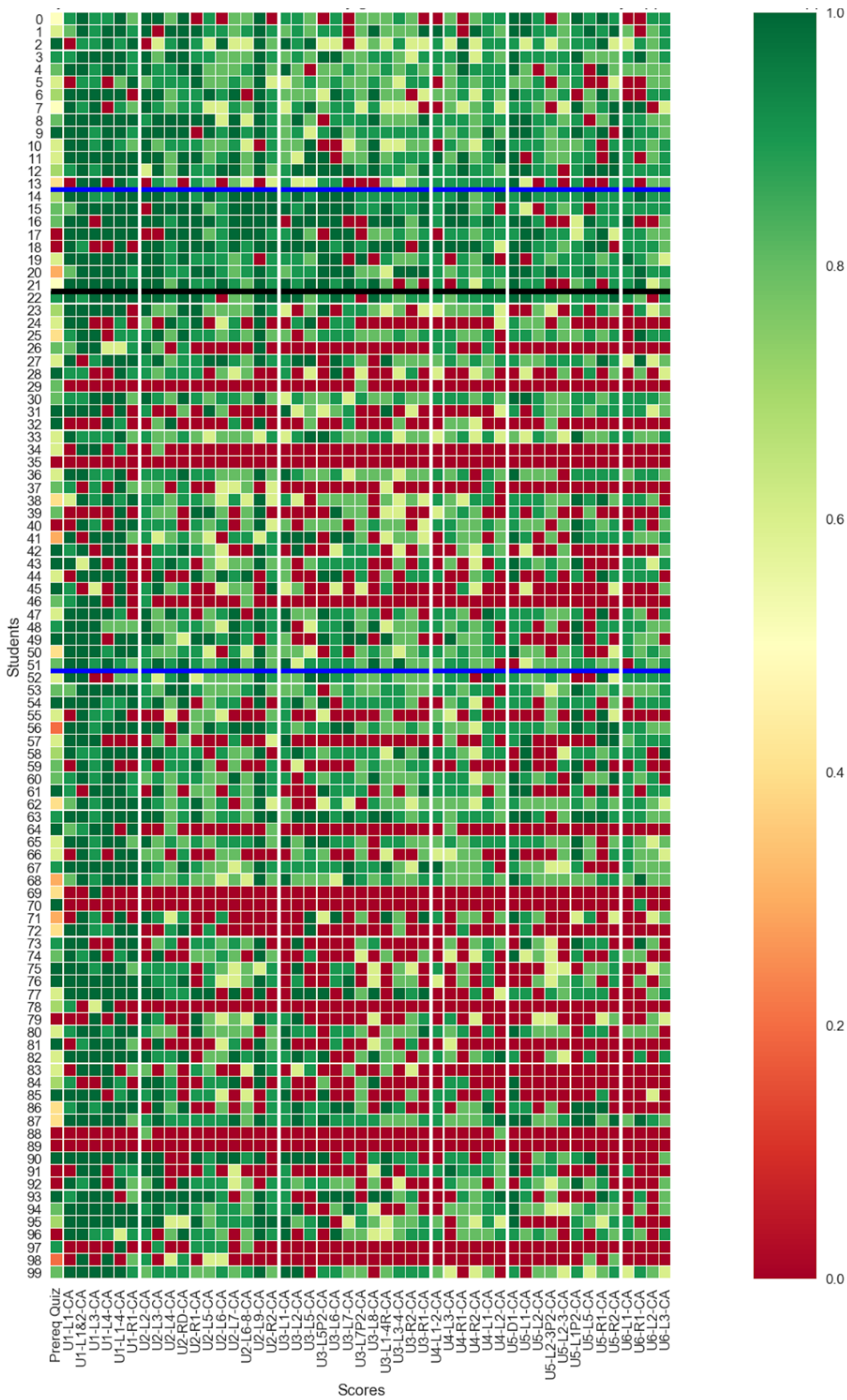


Figure 4.35: Class Activity Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; CA is class activity. Scores range from 0 to 1. Vertical white lines separate each unit. The black line separates gender; the first part is for female students and the second part is for male students. The horizontal blue line separates those who took the flipped class [Course 1] before from

those who did not within the same gender. Dark red color in the heat map is equal to zero which means the student is absent. There is an advantage of not normalizing this data. Class Activity score can be only [0; 0.6; 0.8; 0.9; 1]. If the student attends the class, even with low score in the activity, he/she scores a minimum of 0.6. This visualization shows the attendance pattern of students and shows clearly consecutive absences. Course 2 is a course that students took in their second year of school, so when we compare the students of Course 2 with the students of Course 1 by their class activity scores, we see that the students of Course 2 established their own pace and characteristics, which is different than how they did in the Course 1. There was no female student that made more than three consecutive absences and there were no withdrawed female students either. Female students have a higher participation rate compared to male students. Also, the statements that we made for homework apply for the class activities.

4.5.2.3 Exams

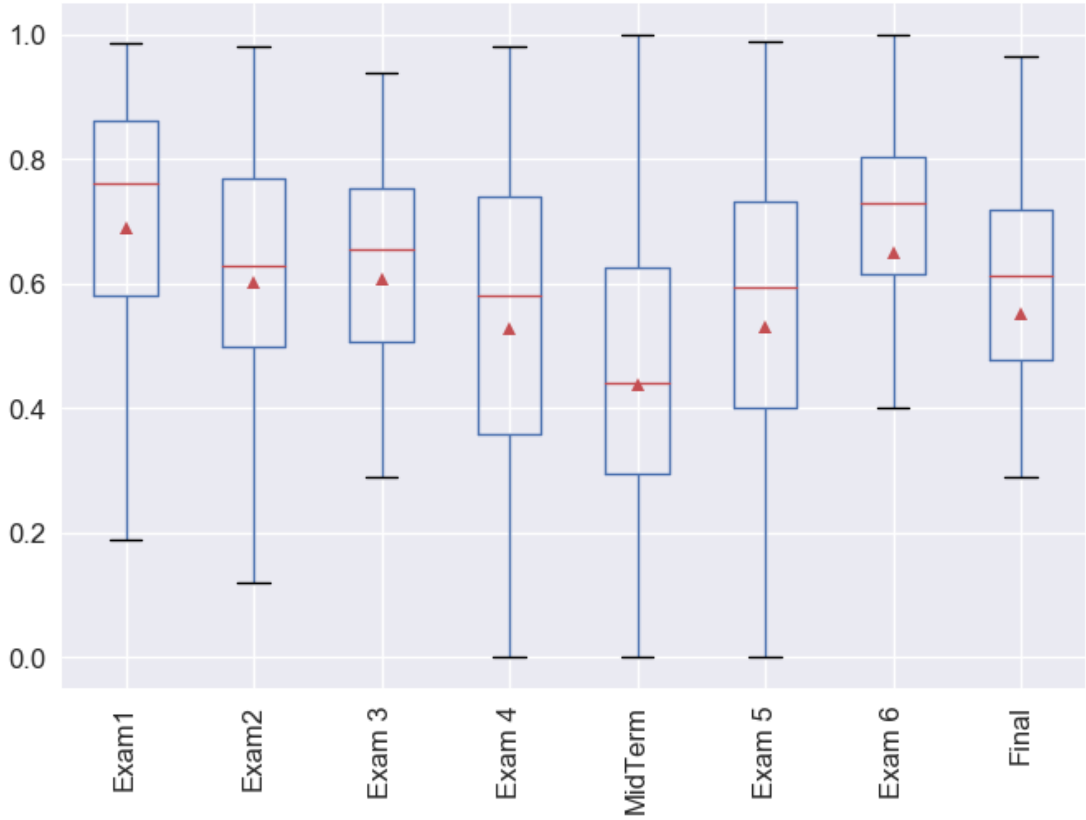
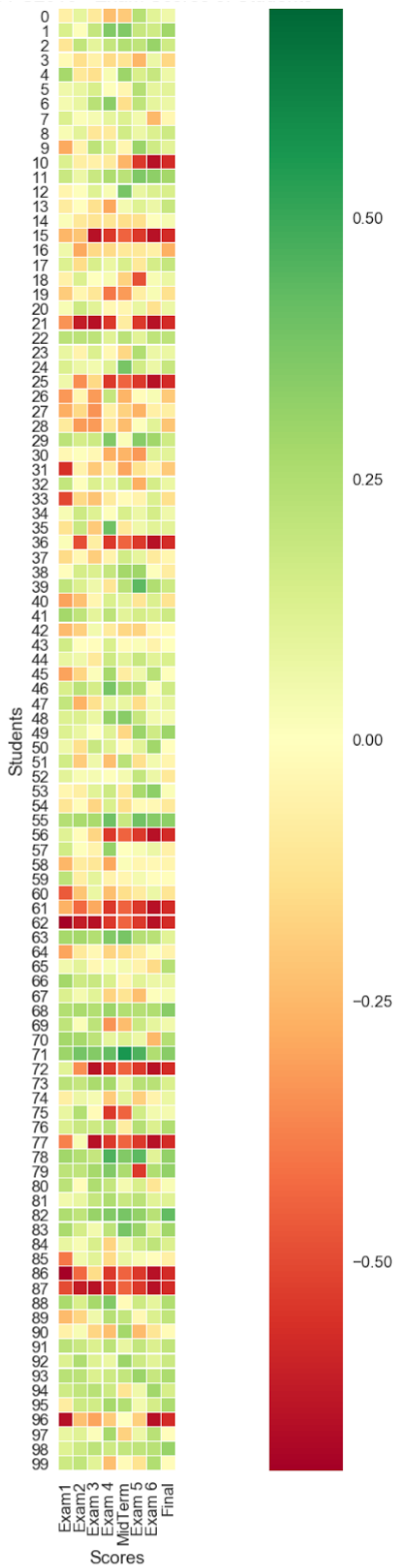


Figure 4.36: Boxplots of Course 2 Exam Scores of All Students. Red line is the median and the red triangle is the mean. Scores range between 0 to 1. This boxplot supports all inferences that we made in the heat maps. The lowest score exam is the midterm. The average exam scores range in [0.45 to 0.7].



Students
Scores
Exam1
Exam2
Exam 3
Exam 4
MidTerm
Exam 5
Exam 6
Final
0.50
0.25
0.00
−0.25
−0.50

Figure 4.37: Heat Map of Final Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map

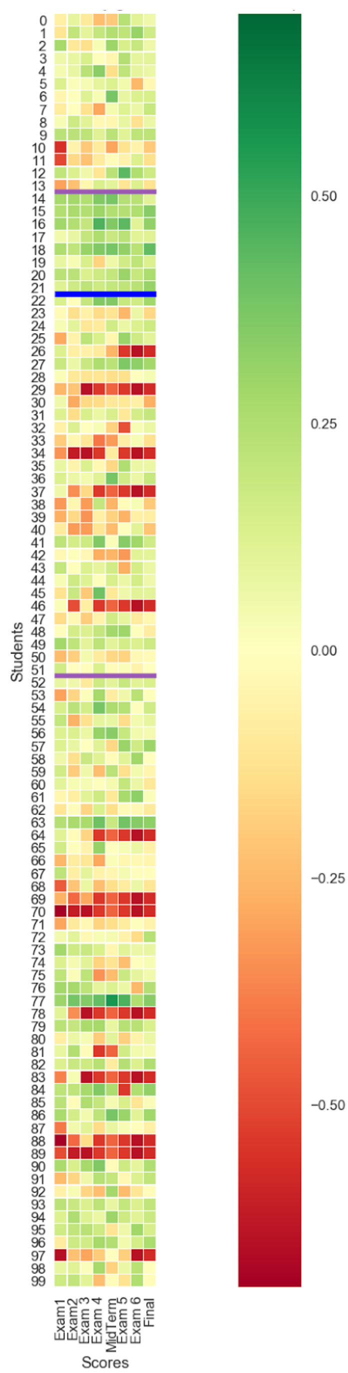


Figure 4.38: Heat Map of Final Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. The blue line separates gender; the first part is for female students and the second part is for male

students. The horizontal purple line separates those who took the flipped class [Course 1] before from those who did not within the same gender. This heat map shows that female students perform better that male students in exams. When we consider for each gender whether the flipped class is taken before or not; for males there is no difference, however the female group, who did not take the flipped class before, seem to do better than the group who took the flipped class. However, this may be due the small sample size and the association of retakes with not taking the flipped class before.

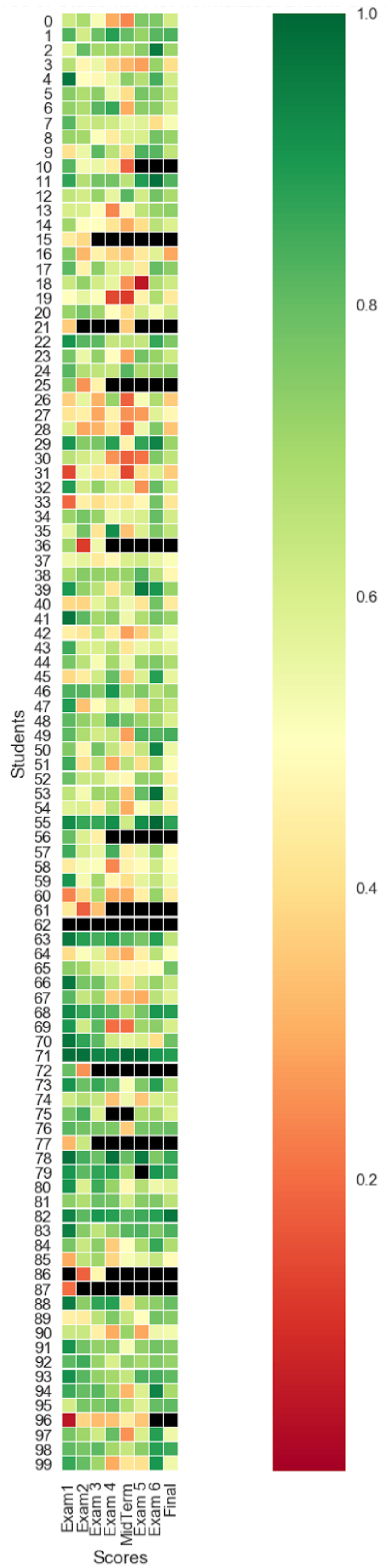


Figure 4.39: Heat Map of Final Scores of Course 2 in columns vs students in rows. Data is not normalized. Black cells are equal to zero. This heat map helps us distinguish a zero from a low score. In this visualization, we can see the drop outs. Overall, the dropout rate is 16%, which the same as Course 1.

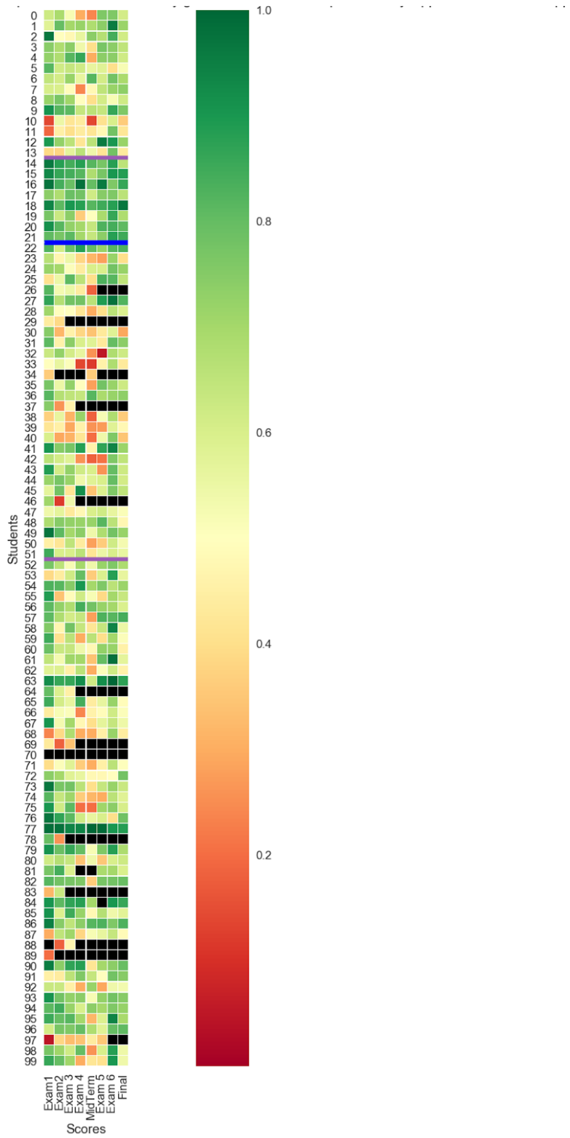


Figure 4.40: Heat Map of Final Scores of Course 2 in columns vs students in rows. Data is not normalized. Black cells are equal to zero. The blue line separates gender; the first part is for female students and the second part is for male students. The horizontal purple line separates those who took the flipped class [Course 1] before from

those who did not within the same gender. This heat map helps us distinguish a zero from a low score. In this visualization, we can also see drop outs. There is no drop out in females; in males it is 1 in 6 students. Also, the dropout rate does not change in both groups regardless of having taken the flipped class before.

4.5.2.4 All Scores

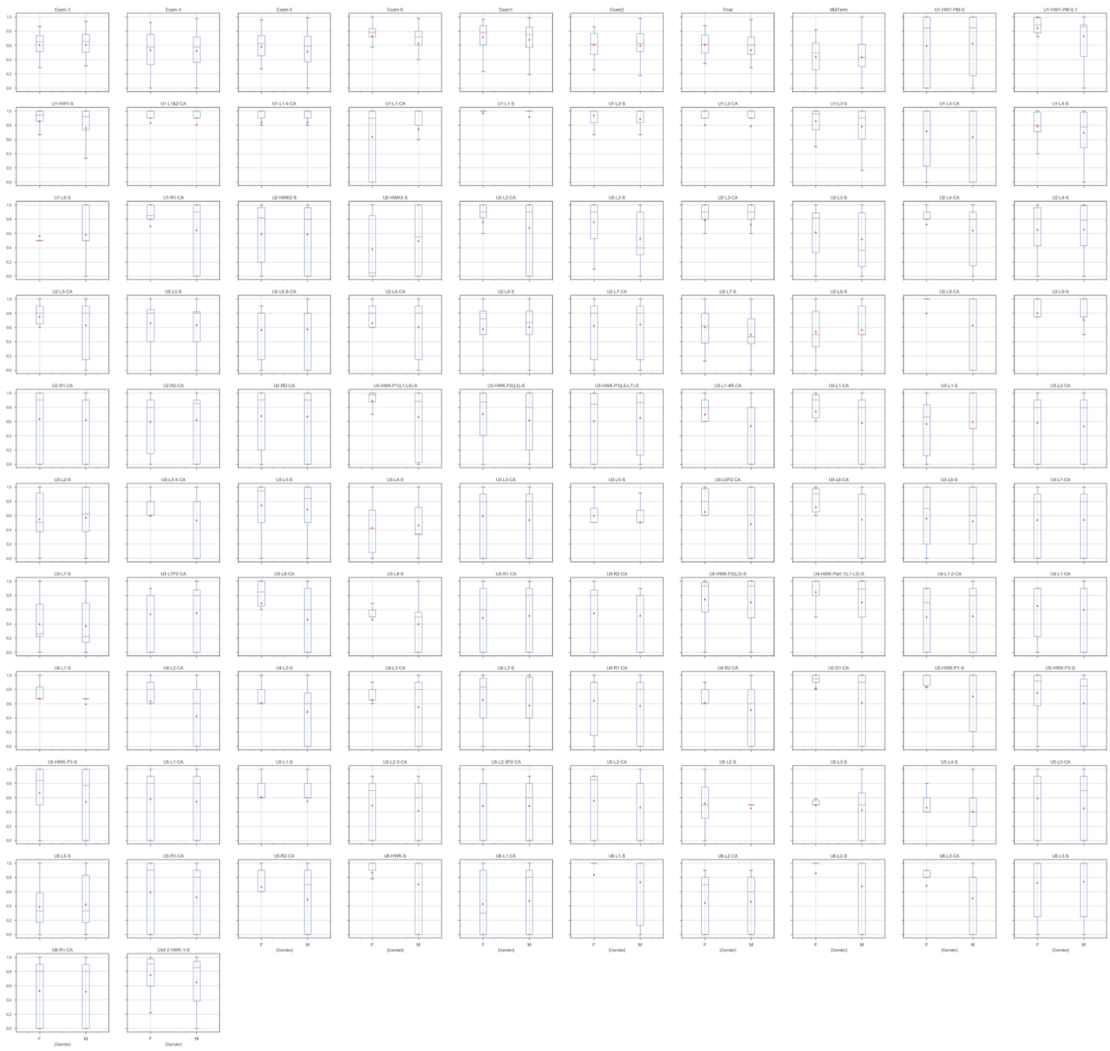


Figure 4.41: Boxplots of All Scores Grouped by Gender for Course 2. F is female; M is male. We use these boxplots to confirm findings in the heat maps.

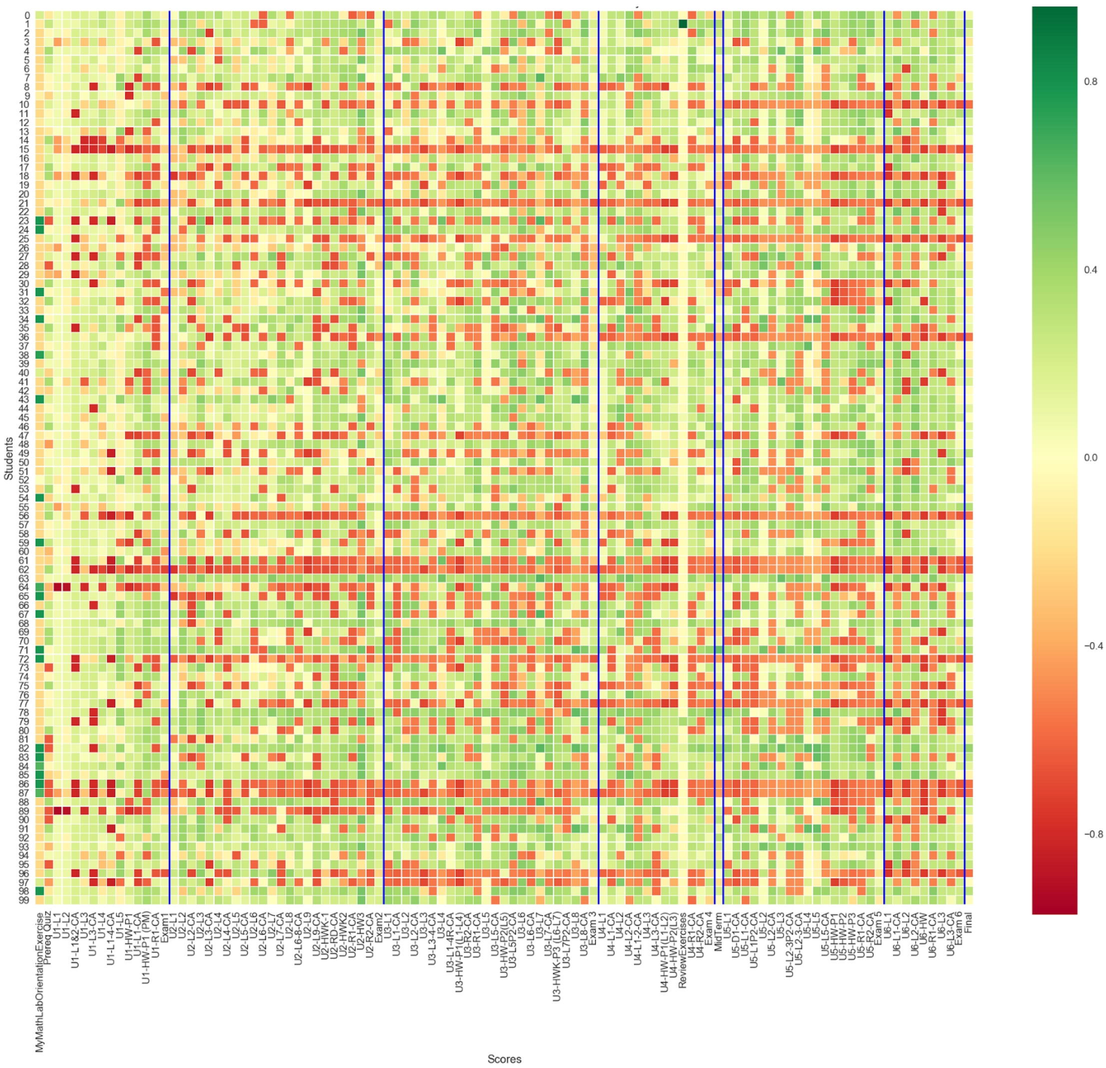


Figure 4.42: Heat Map of All Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. U X is for Unit X; L Y is Lesson Y; HW is homework; CA is class activities. Vertical blue lines separate each unit. In Unit 1 to 6, the overall colors change from light orange to light green. This means that students, except for dropouts, gain performance throughout the semester.

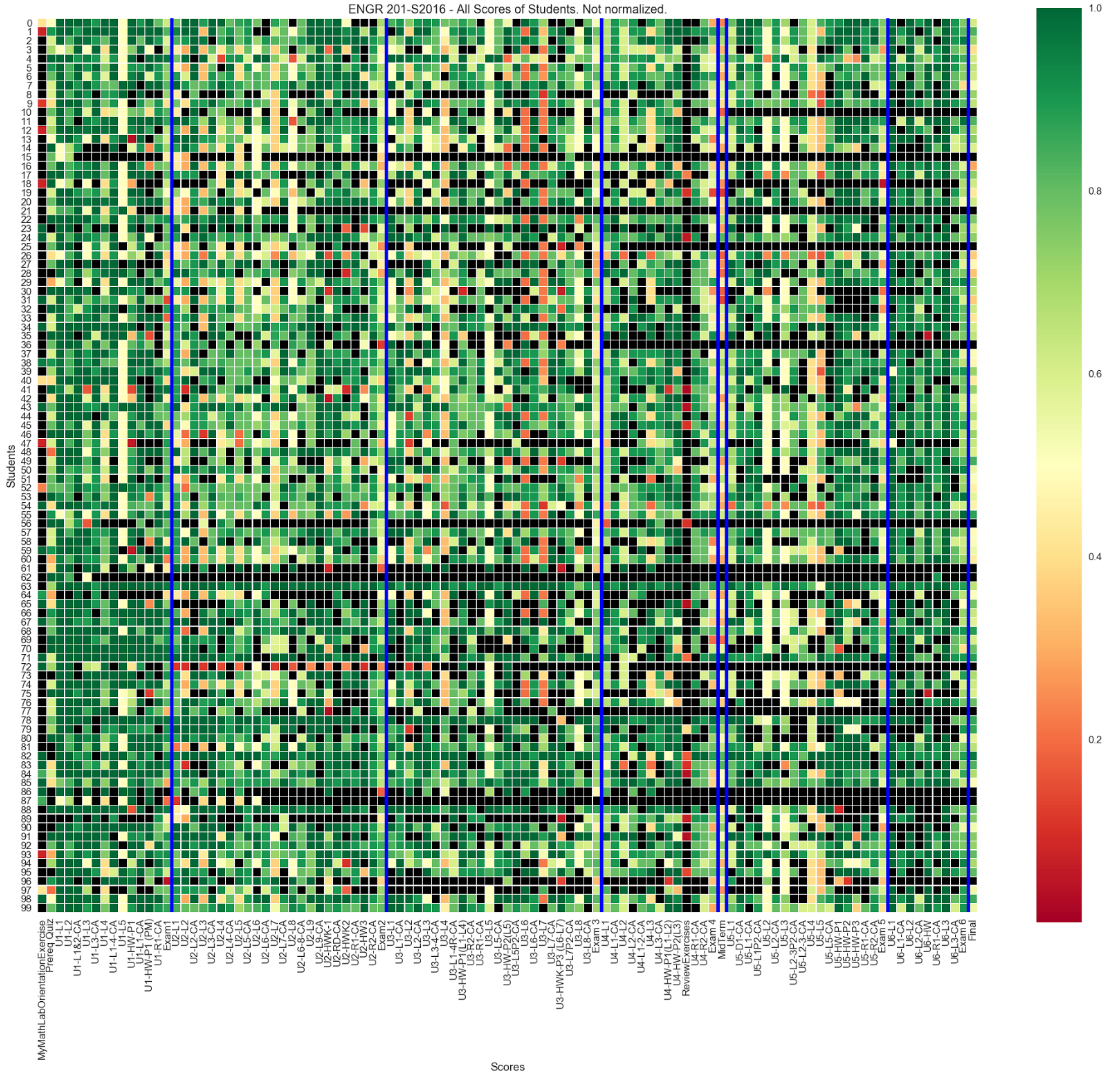


Figure 4.43: Heat Map of All Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; HW is homework; CA is class activities. Vertical blue lines separate each unit. This heat map helps us distinguish a zero from a low score.

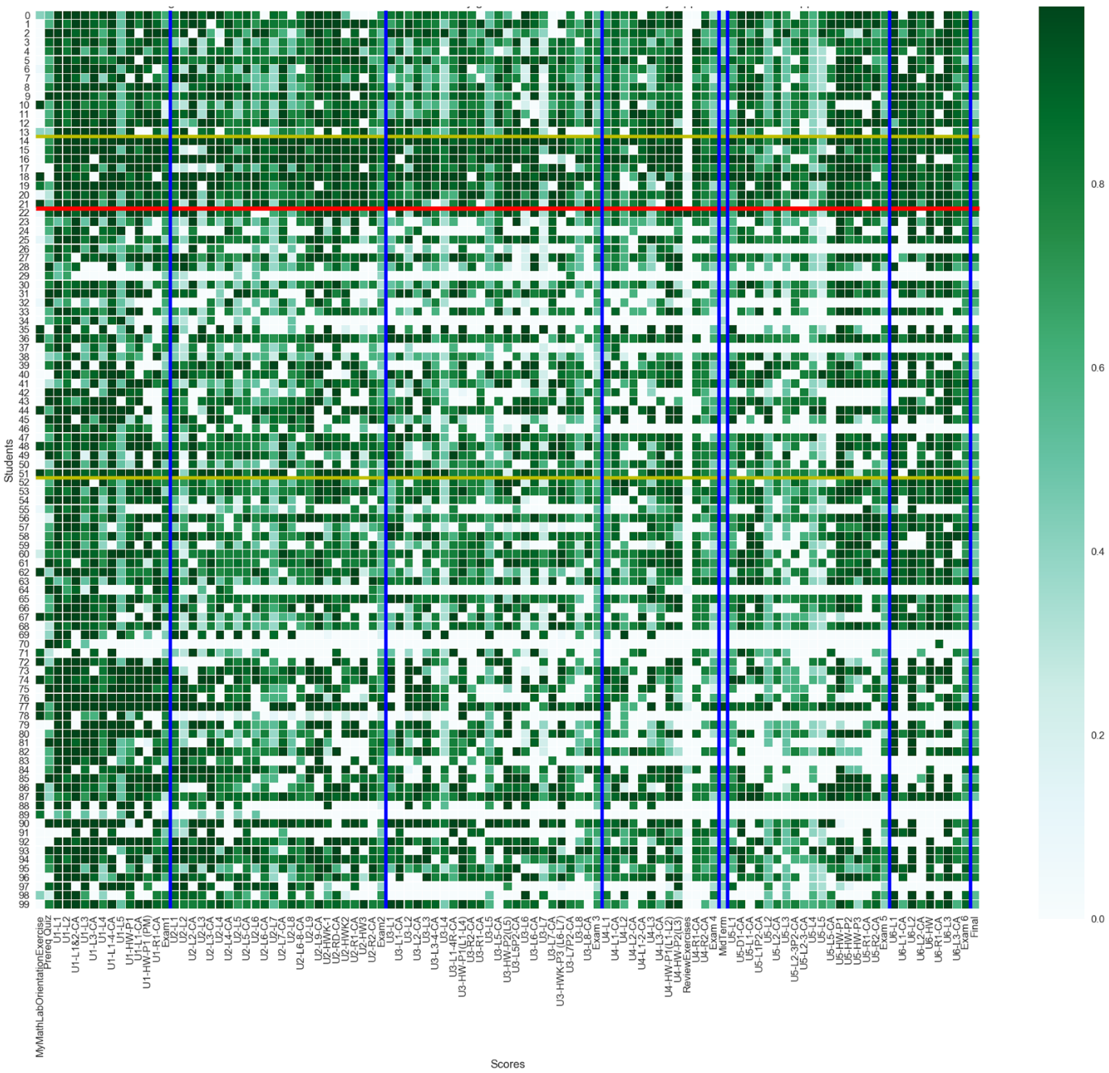


Figure 4.44: All Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; HW/K is homework; CA is class activities. Vertical blue lines separate each unit. The red line separates gender; the first part is for female students and the second part is for male students. The horizontal yellow line separates those who took the flipped class [Course 1] before from those who did not within the same gender. White color is equal to zero. This heat map helps us distinguish a zero from a low score.

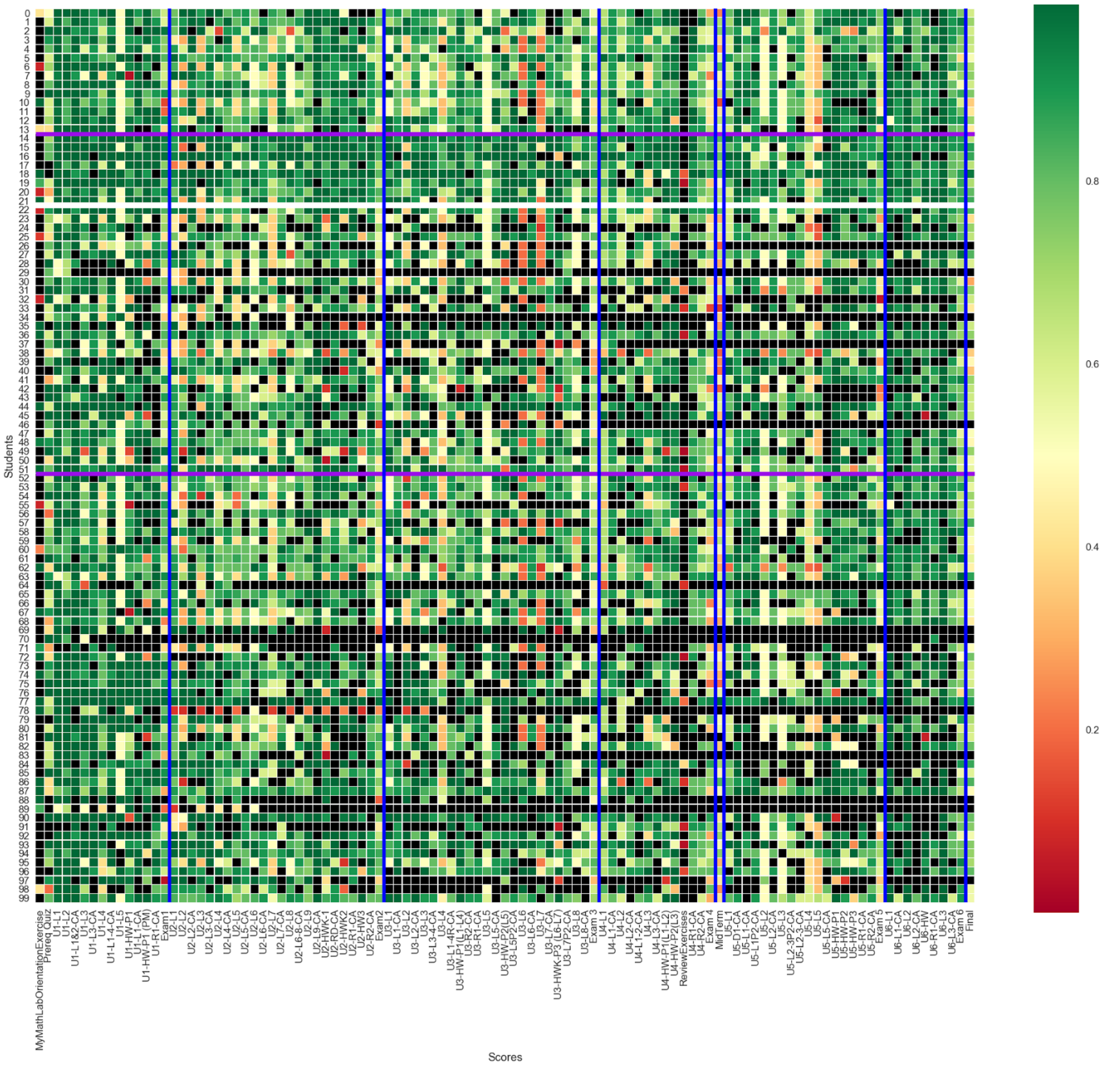


Figure 4.45: All Scores of Course 2 in columns vs students in rows. Data is not normalized. U X is for Unit X; L Y is Lesson Y; HW is homework; CA is class activities. Black cells are equal to zero. Vertical blue lines separate each unit. The white line separates gender; the first part is for female students and the second part is for male students. The horizontal purple line separates those who took the flipped class [Course 1] before from those who did not within the same gender. This heat map helps us distinguish a zero from a low score. In this visualization, we can see that among female students there is only one student with consecutive zero scores (which is 4). On the other hand, among male students, whether they took a flipped class before or not, they have many consecutive zeroes.

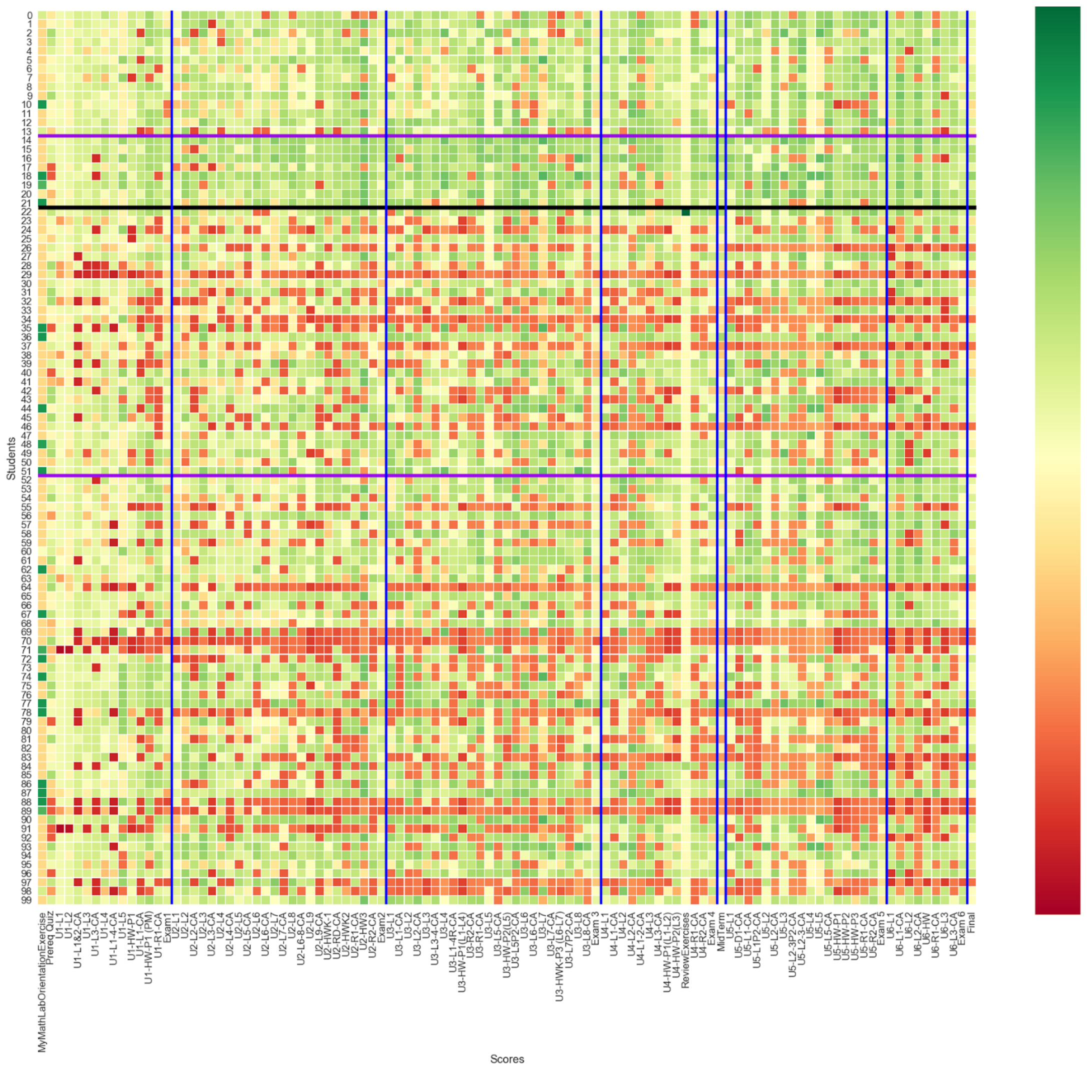


Figure 4.46: Heat map All Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. U X is for Unit X; L Y is Lesson Y; HW is homework. Vertical blue lines separate each unit. The black line separates gender; the first part is for female students and the second part is for male students. The horizontal purple line separates those who took the flipped class [Course 1] before from those who did not within the same gender. This heat map clearly shows us female students perform significantly better than male students. When we consider each gender depending on whether the flipped class was taken before or not; for males there is no difference, however the female group who did not take the flipped class before, perform better than the group who took the flipped class. Overall, female

students who did not take the flipped class before are the most successful group in the class. It is a very interesting implication. When we ask why, we cannot say that they are successful because they took a similar concept course before. We do not think that is by chance. By making an educated guess, we may say that this successful female group might be close friends. We checked their student IDs and they all started school in year while other female students started school a year before, which supports our idea.

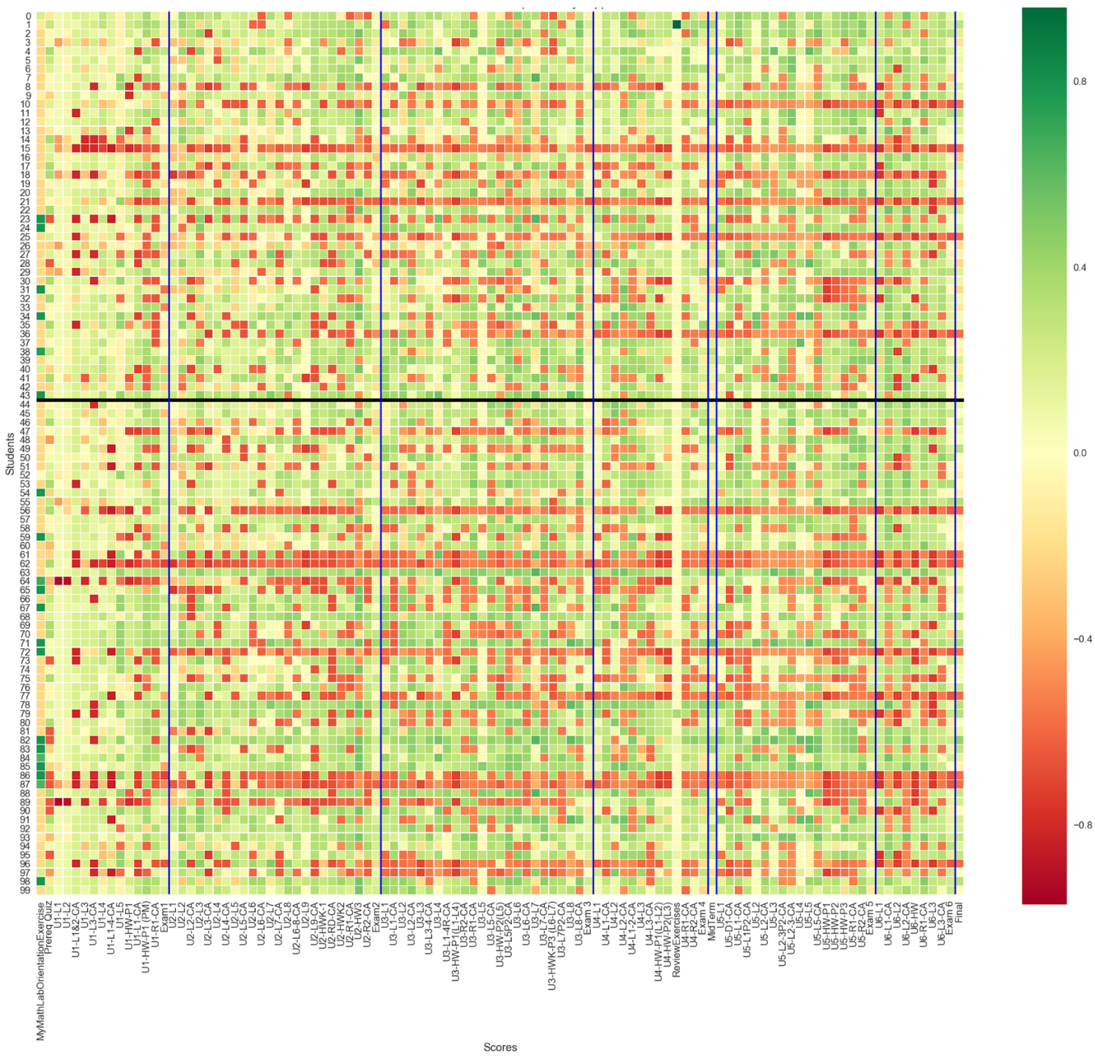


Figure 4.47: Heat Map of All Scores of Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Vertical blue lines separate each unit. Horizontal black line separates that is the flipped class [Course

1] taken before or not by gender. First part of the line took the flipped class and second part did not take it. This heat map shows us that there is no difference between students who took a flipped class before or not.

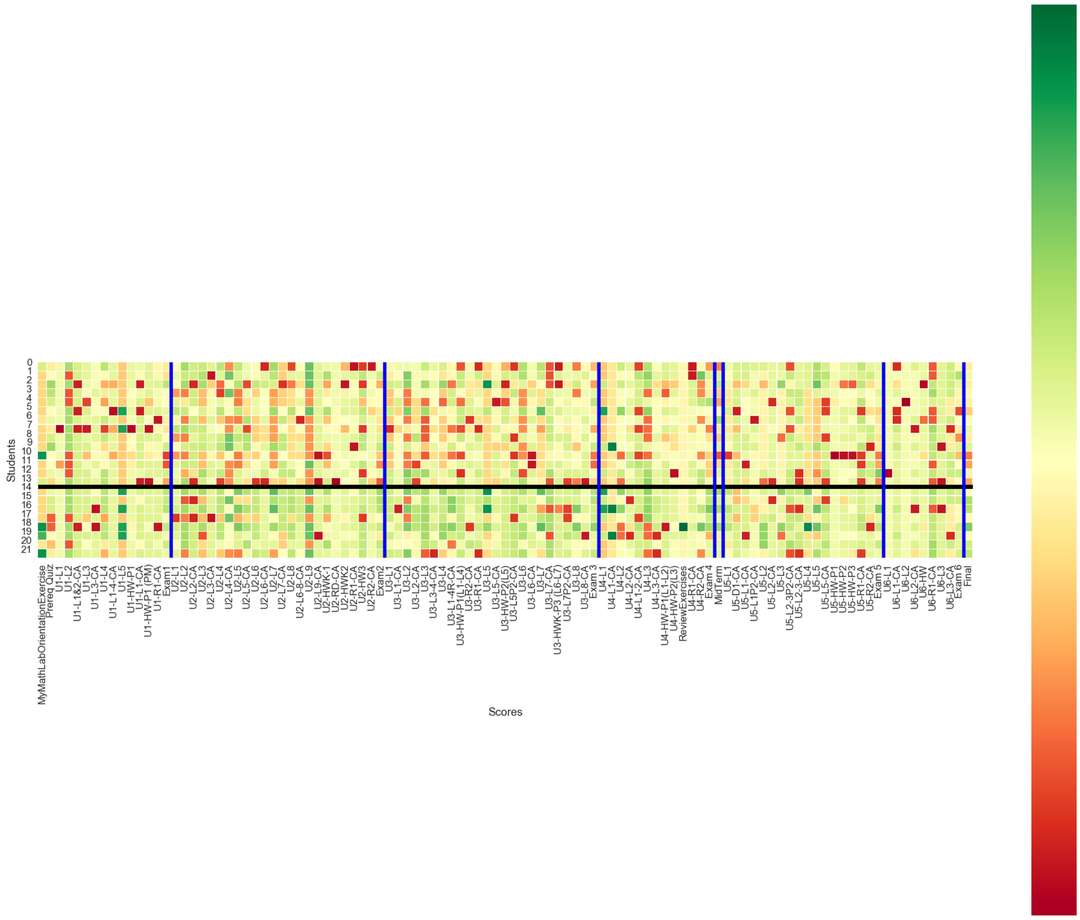


Figure 4.48: All Scores of Female Students in Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Vertical blue lines separate each unit. The horizontal black line separates those who took the flipped class [Course 1] before from those who did not. We normalized the scores just for females to clearly see the effect of the being familiar with the flipped class. We confirm that not taking a flipped class before does not affect student success negatively.

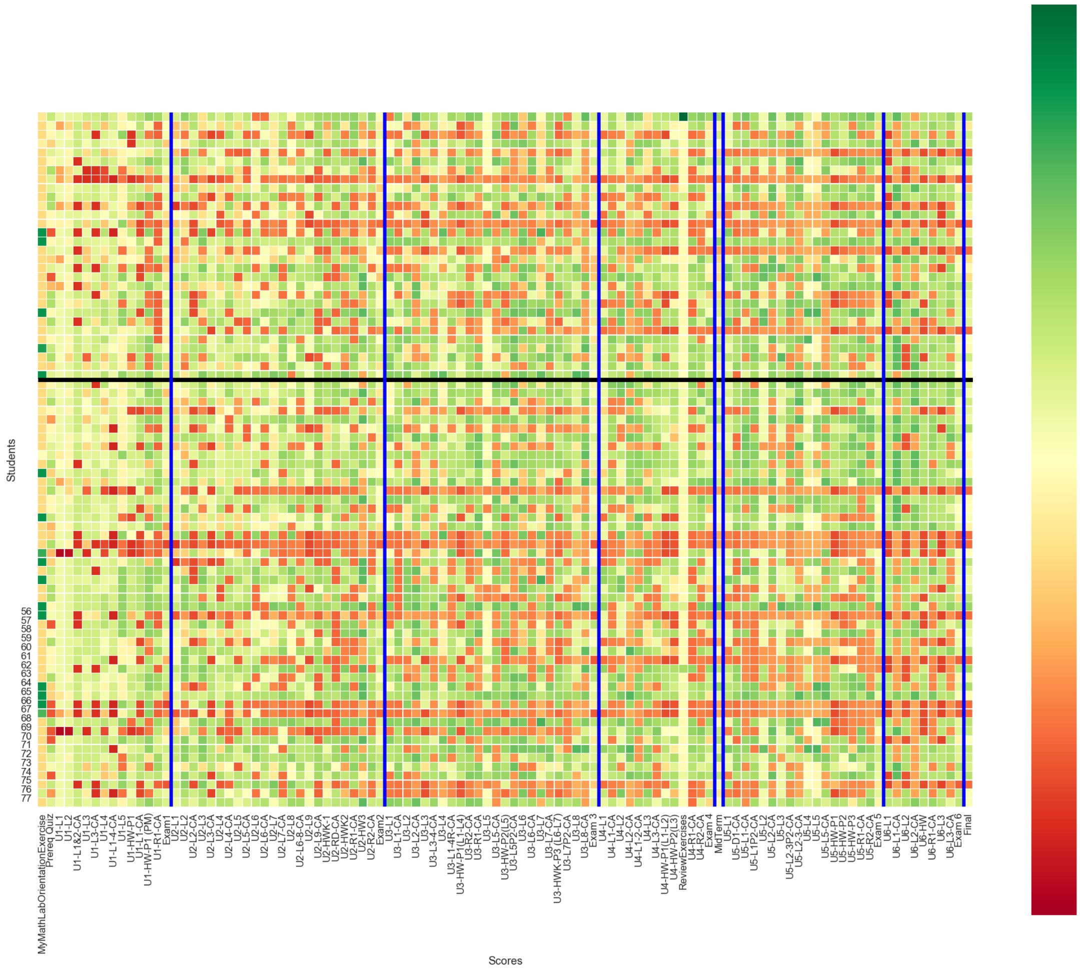


Figure 4.49: All Scores of Male Students in Course 2 in columns vs students in rows. Data is centered to a zero mean, which is yellow color in the heat map. Vertical blue lines separate each unit. The horizontal black line separates those who took the flipped class [Course 1] before from those who did not. The first part above the line took the flipped class and the second part did not take it. We normalized the scores just for male to clearly see the effect of being familiar with the flipped class model. We confirm that not taking a flipped class before does not affect student success.

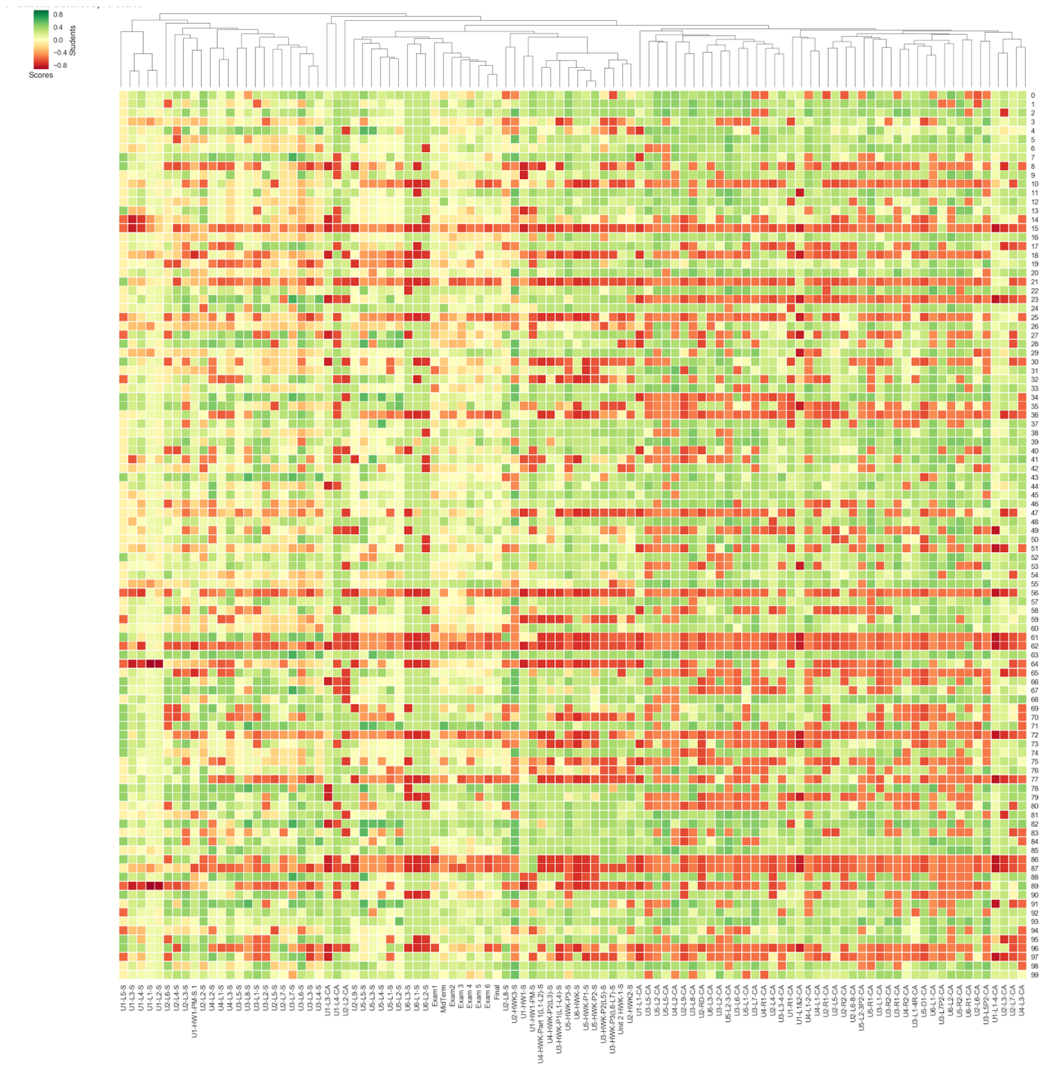


Figure 4.50: All Scores Clustered by Each Unit Activity. Normalized data.

Table 5: Clusters for k=3

Cluster 1	Lesson Assignments(LA) Unit 1 to 4
Cluster 2	LAs Unit 5-6, Exams and HW scores
Cluster 3	Class Activity scores

Table 6: Clusters for k=5

Cluster 1	Lesson Assignments(LA) Unit 1 to 4
Cluster 2	LAs Unit 5-6 and Exams
Cluster 3	HW scores

Cluster 4	Mixed Combination of Class Activity scores
Cluster 5	Mixed Combination of Class Activity scores

We clustered all scores for k=3 and k=5. Homework and exams are in the same cluster (Cluster 2, for K=3). That shows the significance of the homework. Also, LA Unit 1-4 and LA Unit 5-6 are in different clusters. We confirm this by the correlation plot.

4.5.3 Feature Engineering for Course 2

In this section, we constructed new features to predict the final score by building decision trees.

4.5.3.1 Constructed Dataset 1

In dataset 1; we calculated the mean of all scores that a student gets before each exam, then we used it as a variable. As seen in Figure 4.51, UnitX_mean is the average of all scores in Unit X, except the exam score. We applied this method to all units, and created a new dataset.

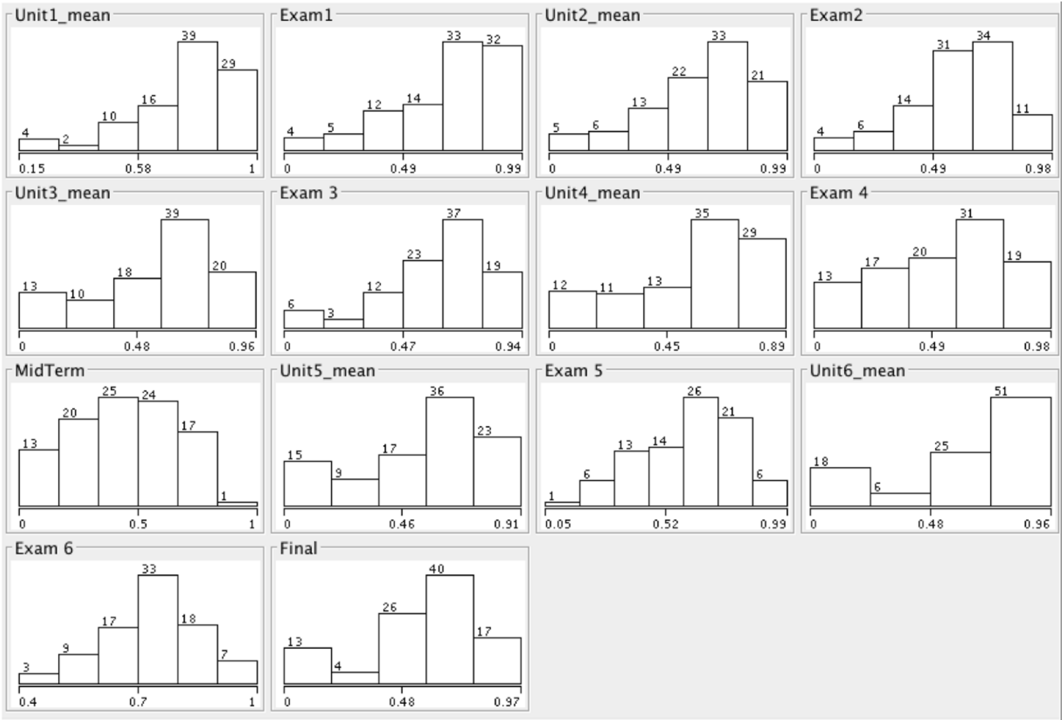


Figure 4.51: Histogram of Constructed Dataset 1

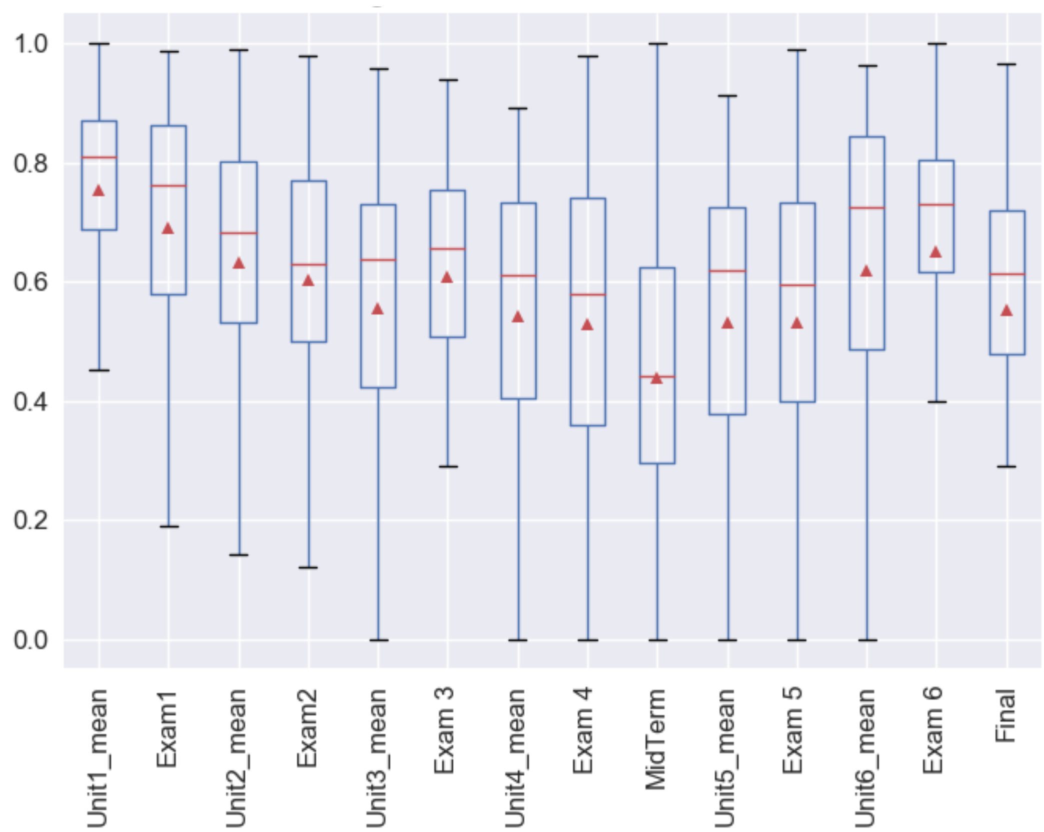


Figure 4.52: Box Plot of Constructed Dataset 1 for Course 2. Score means are very close to the exam score of that unit.

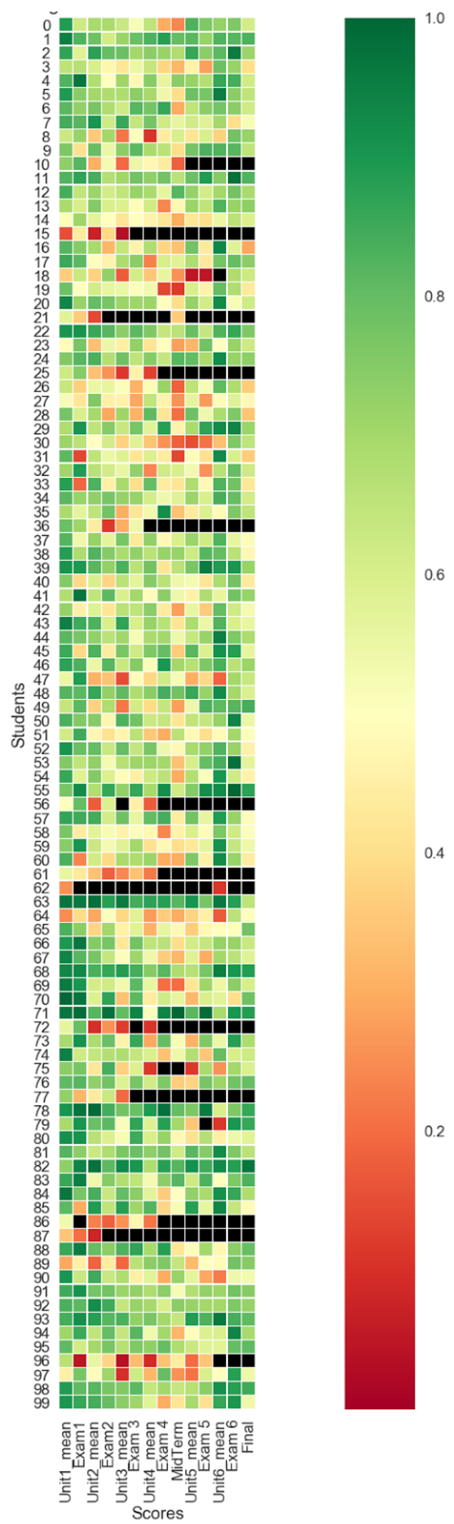


Figure 4.53: Heat Map of Constructed Dataset 1. Raw data. Black cells are equal to zero.

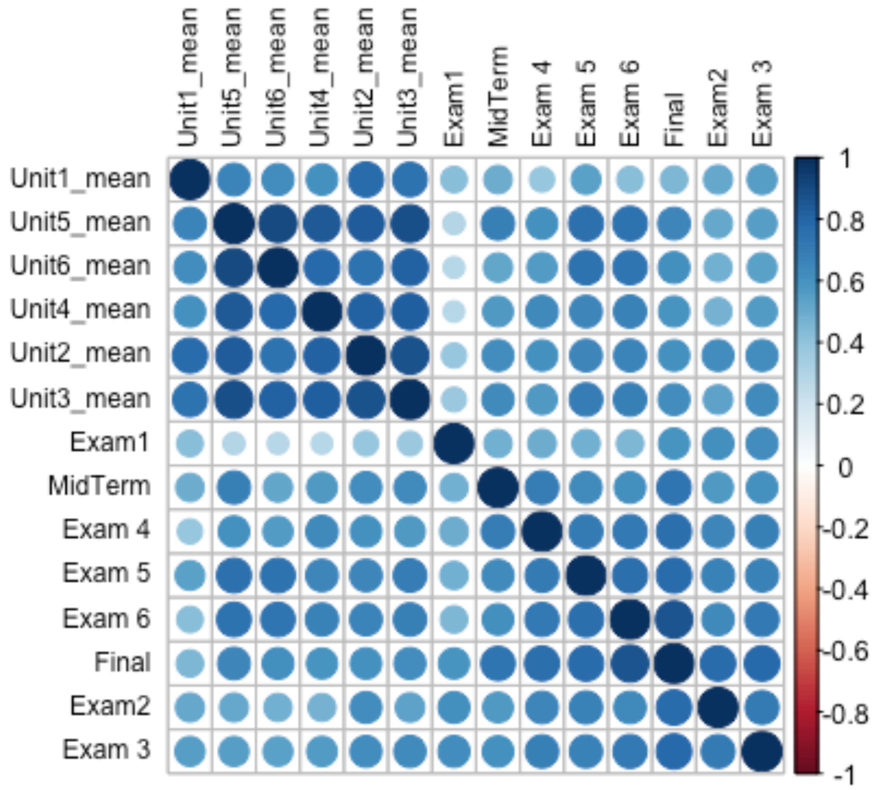


Figure 4.54: Correlation Plot of Constructed Features 1 for Course 2

In Figure 4.54, correlation plot of constructed dataset 1 shows that unit mean scores and exam scores are correlated separately in each. Also, exam 5 and 6 are more correlated with unit mean scores than other units.

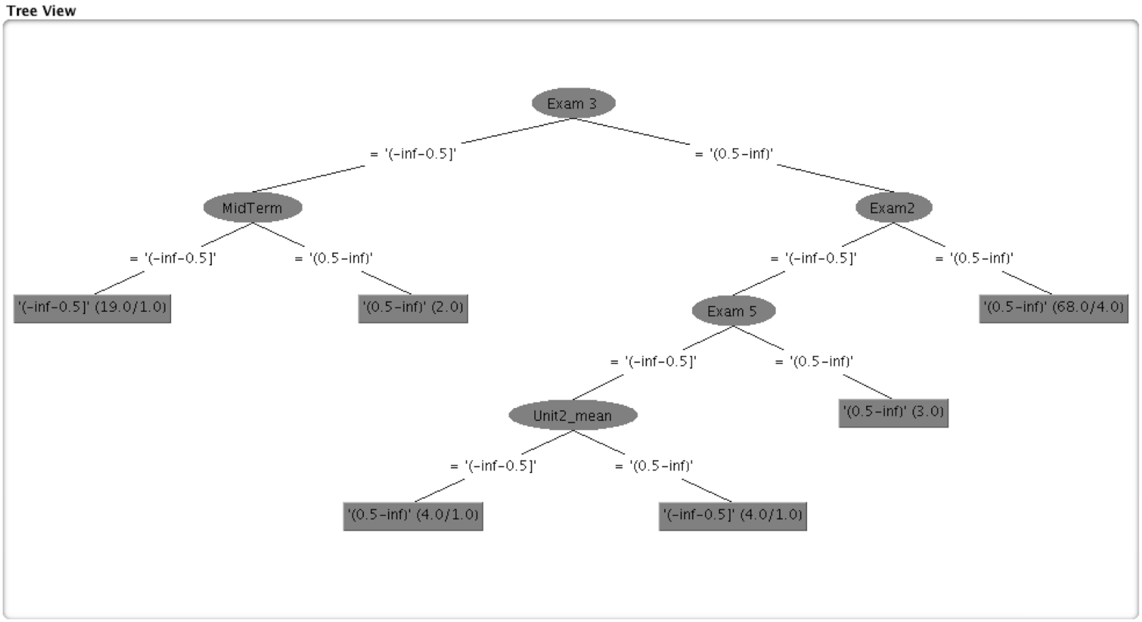


Figure 4.55: Decision Tree of Constructed Dataset 1 for Course 2 for the prediction of the final score.

Figure 4.55 is a pruned decision tree, built using the J48 algorithm [14]. It has an 85% accuracy in predicting the final score. We observe that the midterm is a strong predictor of the final score.

4.5.3.2 Constructed Dataset 2

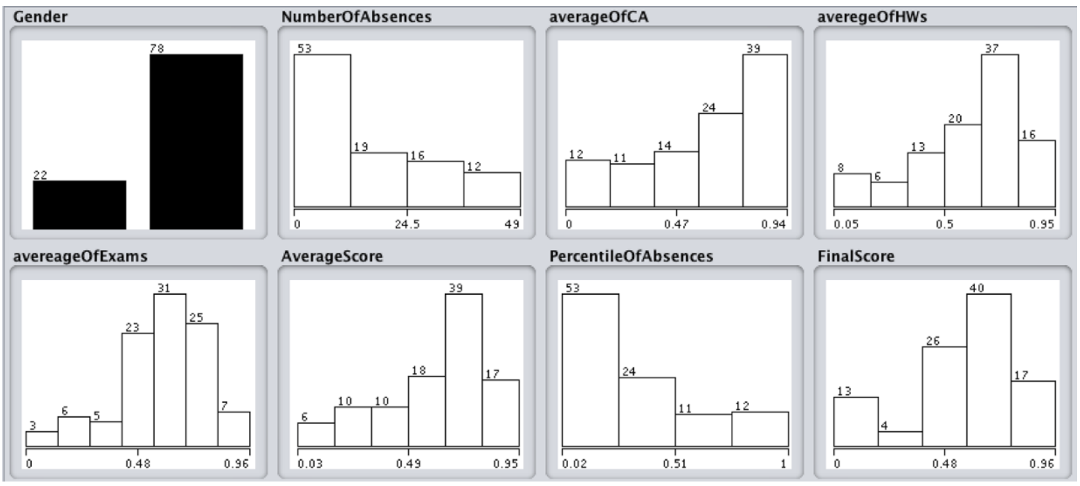


Figure 4.56: Histogram of Tree of Constructed Features 2 for Course 2

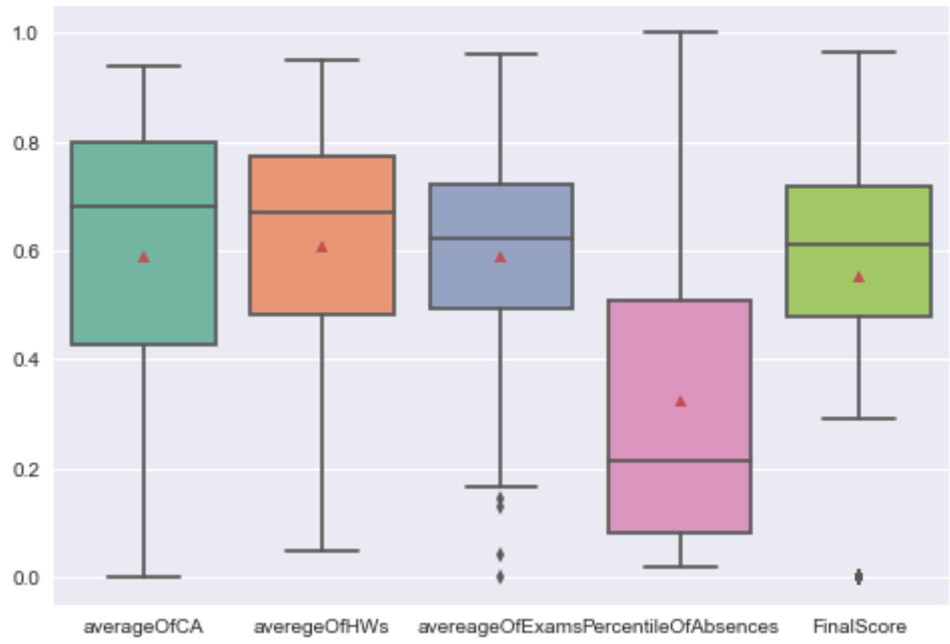


Figure 4.57: Box Plots of Tree of Constructed Features 2 for Course 2 shows that the averages of all features, except for the percentile of absences, are in the same range as the final score.

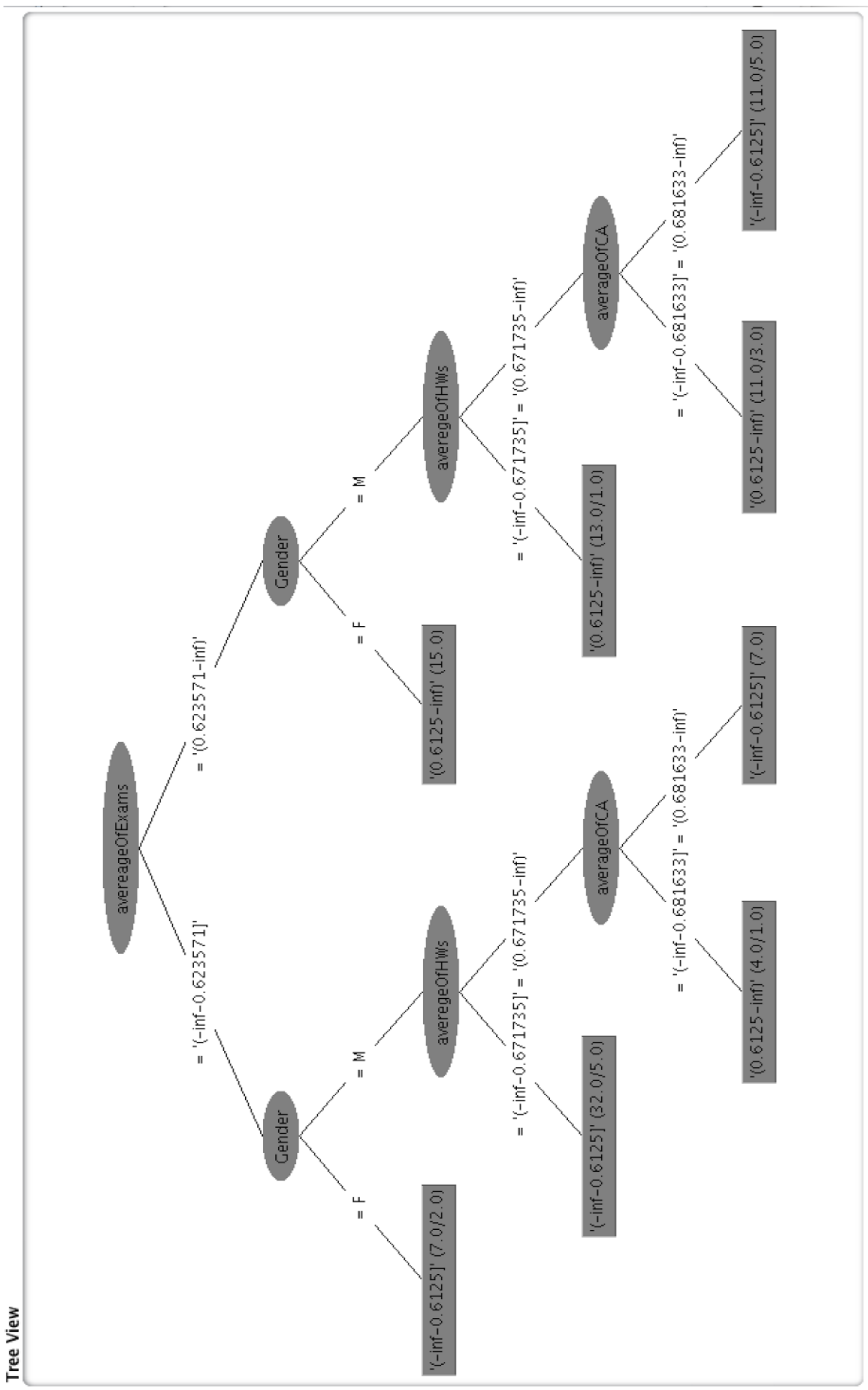


Figure 4.58: Decision Tree of Constructed Features 2 for Course 2 for the prediction of the final score. This is a pruned decision tree, build with the J48 algorithm.

It has 77% accuracy in predicting the final score. We observe that the gender feature has the highest impact on our prediction model.

OPINIONS OF STUDENTS WHO TOOK THE FLIPPED CLASSES

By applying a data science on the flipped class dataset, we extracted knowledge to understand how students did in a flipped classroom, and to predict their final score, etc. But how about student opinions about the flipped class approach?

Education is heavily involved with psychology, and the flipped classroom is an innovative pedagogical approach in the traditional education system. We believe that student opinion is important and may be a good complement to the quantitative data analysis in this study. Absent the evaluation reports, we looked at *ratemyprofessor.com* to see if there are any comments about Courses 1 and 2. In the page of Dr. Jeff Hieb[1], who is the instructor and implementer of the flipped classroom, we found 29 total ratings about this professor along with comments about Courses 1 and 2. We recognize that the online rating data can be unreliable, since in many cases, there is no guarantee of authenticity and that the data may suffer from selection bias (e.g., unhappy students may submit online ratings more than happy students) and other random sampling biases.

Overall, the quality ratings of Dr. Hieb is 4.8 out of 5, 100% of the raters say they would take the class again, and the level of difficulty is 3.2 out of 5. There are many great comments about the professor. Specifically, we looked for Course 1 and 2. The general student opinion is that these classes need much effort, however they give better understanding about the lecture. Several quotes of the students are included below:

---

[1] http://www.ratemyprofessors.com/ShowRatings.jsp?tid=1209438

- “It's a flipped classroom, its more work but you get a deeper knowledge of the material.”
- “The inverse classroom method he uses for the calc classes work very well.”
- “Engineering based calculus is tough but he structures his classroom in a way that makes it very do-able.”
- “His teaching style is way different from other calc professors he has a group teaching style vs a lecture hall which i find more helpful.”
- His teaching style is different than most engineering classes, but it are structured around the student. Engineering math is still hard, but Hieb is great.

As we can see, the student opinion is extremely positive about the flipped classroom and there is not even one negative comment.

CONCLUSION

In this thesis, we presented a data science pipeline to analyze the education data set consisting of scores in lessons, homework, exams, etc. in a flipped classroom model for J.B. Speed School of Engineering Students. We used a combination of classical statistical methods with computational visualization and machine learning. Some of the visualizations revealed trends in the increase of scores within and across units, as well as differences based on gender and having taken the flipped classroom before. To confirm some of our findings about gender and the flipped class factor, we applied the findings chi-square test of independence. For gender; the p-value was 0.004068 which is less than 0.05; the average score is thus dependent on the gender of students. For the flipped class; p-value was 0.6659 which is less than 0.05, the average score is independent of the flipped class factor of students. However, we emphasize that visualizations tend to be interpreted subjectively, while rigorous statistical tests remain the best way to verify certain conclusions. On the other hand, visualizations, especially on large data sets, can reveal certain patterns that we may not anticipate, and thus help generate hypotheses to be tested in a later stage.

While our objective was not to predict the final score, we did build machine learning models that can predict this score based on a variety of constructed features. The main goal of these models was to explore which features had the biggest impact on the final score, generally considered as a measure of overall student success in a class.

Future work involves improving and constructing new visualizations, as well as continuing some of the hypothesis generation and rigorous statistical testing and modeling. Other approaches such as sequential pattern mining are also needed to support some of the visual inspection of the heat maps. Other data can also be captured to support investigations that leverage data science, based on some of the conclusions we made and unanswered questions.

REFERENCES

[1] U.S. Department of Education Office of Educational Technology, "Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief," Center for Technology in Learning SRI International , Washington, D.C, 2012.

[2] Educause, "FLIPPED CLASSROOMS," 1 February 2012. [Online]. Available: https://net.educause.edu/ir/library/pdf/eli7081.pdf. [Accessed 15 July 2016].

[3] M. A. V. Jacob Bishop, "The Flipped Classroom: A Survey of the Research," in *2013 ASEE Annual Conference & Exposition*, Atlanta, 2013.

[4] D. J. Hieb, "Syllabus of The Introductory Calculus for Engineers Course 1," Blackbord, Louisville, 2016.

[5] J. Hieb, "Syllabus of The Introductory Calculus for Engineers Course 2," Blackboard, Louisville, 2015.

[6] Wikipedia, "Exploratory Factor Analysis," Wikipedia, -, 2017.

[7] D. Child, The essentials of factor analysis, second edition, London: Cassel Educational Limited, 1990.

[8] C. R. B, "The scree test for the number of factors," in *Multivariate Behav. Res.* , Urbana-Champaign, University of Illinois, Urbana-Champaign, ILl, 1966, pp. 1:245-76.

[9] J. W. Osborne, "What is Rotating in Exploratory Factor Analysis?," *Practical Assessment, Research & Evaluation,* vol. 20, no. 2, p. 7, 2015.

[10] J. D. Brown, "Choosing the Right Type of Rotation in PCA and EFA," *Shiken: JALT Testing & Evaluation SIG Newsletter,* vol. 13, no. 3, pp. 20-25, November 2009.

[11] Wikipedia, "Data Visualization," Wikipedia, Wikipedia, 2017.

[12] T. Mitchell, "Decision Tree Learning," in *Machine Learning*, McGraw Hill, 1997, pp. 52-80.

[13] R. Quinlan, C4.5: programs for machine learning, San Francisco: Morgan Kaufmann Publishers Inc, 1993.

[14] E. Frank, "Weka," WEKA, - - -. [Online]. Available: http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/J48.html. [Accessed 15 July 2017].

[15] Wikipedia, "C4.5 algorithm," Wikipedia, - - -. [Online]. Available: https://en.wikipedia.org/wiki/C4.5_algorithm#cite_note-4. [Accessed 15 July 2017].

[16] Z. T. Siti Khadijah Mohamada, "Educational data mining: A review," in *The 9th International Conference on Cognitive Science*, Malaysia, 2013.

[17] R. S. Baker, "Data Mining for Education," in *International Encyclopedia of Education (3rd edition)*, Oxford, 2012.

[18] C. G. Merrett, "Using Textbook Readings, YouTube Videos, and Case Studies for Flipped Classroom Instruction of Engineering Design," in *Proc. 2015 Canadian Engineering Education Association (CEEA15) Conf.*, Canada, 2015.

[19] D. N. A. G. M. S. Kenneth A. Connor, "Faculty Development and Patterns of Student Grouping in Flipped Classrooms Enabled by Personal Instrumentation," in *2017 ASEE Annual Conference & Exposition* , Columbus, 2017.

[20] H. G. S. Gabriela C. Weaver, "Design, Implementation, and Evaluation of a Flipped Format General Chemistry Course," *Chemical Education Research ,* vol. 9, no. 92, pp. 1437-1448, 2015.

[21] A.-M. K. ,. M. O.-L. Jonathan Verrett, "I flipped my tutorials: a case study of implementing active learning strategies in engineering," in *Proc. 2015 Canadian Engineering Education Association (CEEA15) Conf.*, Canada, 2015.

[22] F. A. Fei Geng, "Biotechnology labs reinvented through experiential learning: Enhancing student outcomes through the “flipped lab”," in *2015: Proceedings of the Canadian Engineering Education Association (CEEA) Conference*, Canada, 2015.

[23] C. L. H. L. A. R. Mary Lou Maher, "Flipped Classroom Strategies for CS Education," in *SIGCSE '15 Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, Kansas City, 2015.

[24] M. Theuwissen, "KDnuggets," DataCamp, - - -. [Online]. Available: http://www.kdnuggets.com/2015/05/r-vs-python-data-science.html. [Accessed 15 July 2017].

CURRICULUM VITAE

NAME: Asuman Cagla Acun Sener

ADDRESS: Department of Computer Engineering and Computer Science

University of Louisville

Louisville, KY 40292

EDUCATION: M.S. Computer Science

University of Louisville

2015-2017

B.S. Computer Engineering

Hacettepe University

2009-2013

EXPERIENCE: Software Developer

Innova IT Solutions, Ankara, Turkey

2013-2014

AWARDS: Study Abroad Scholarship

Ministry of Education, Turkey

2014