DataFest 2024

By Valtteri, Wendy & Sophie

Introduction

Statistics and Data Science Bachelor's Degrees have risen in popularity over the last decade (American Statistics Association, 2022), but Statistics remains as one of the least popular college majors as of 2024, perceivably due to the difficulty of the subject.

The aim of this project was to investigate student engagement data and make suggestions to help CourseKata improve the student experience of learning statistics. By doing so, we wish to find out which variables were the most impactful to student success and make learning of statistics easier for future students.

This analysis was based on CourseKata's collection of data on student engagement with three online Statistics and Data Science textbooks. To provide more focused feedback, we looked at the book:

College / Advanced Statistics and Data Science (ABCD).

Sanity Checks

Making sure the data is clean (e.g. all values are appropriate) before analysis.

Exploratory Data Analysis

Average Score of "Pulse" of Each Chapter			
Utility	**Expectancy**	**Intrinsic**	**Cost**
4.53 (1.08)	4.13 (1.17)	4.44 (1.12)	3.17 (1.31)

Pulse Questionnaire Placement

You're about to start a new chapter! Before you do, reflect on how it's going so far and rate your level of agreement with each of the following statements.

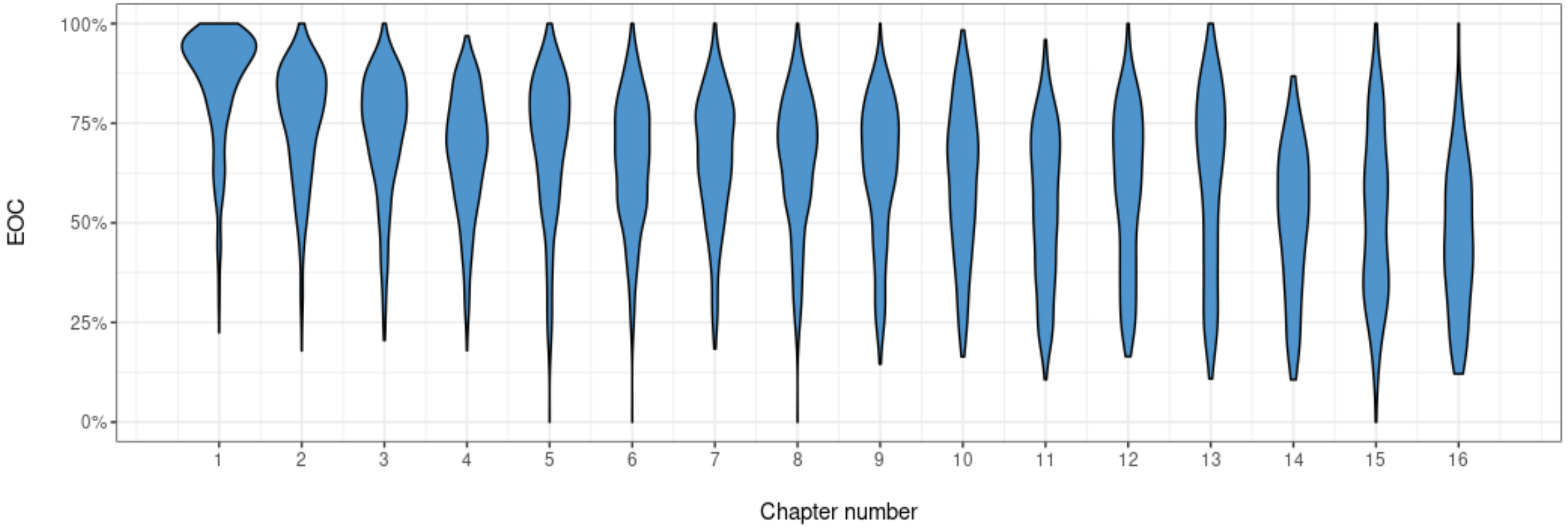
	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
I am confident in what I have learned in the previous chapter.	○	○	○	○	○	○
I think what I have learned in the previous chapter is useful.	○	○	○	○	○	○
My statistics ability is something about me that I can't change very much.	○	○	○	○	○	○
I was unable to put in the time needed to do well in the previous chapter.	○	○	○	○	○	○

Learnosity: Pulse_a2_1223

Chapter 2 - Understanding Data

2.1 Starting With a Bunch of Numbers

EOC (%) by Chapter in College / Advanced Statistics and Data Science (ABCD)



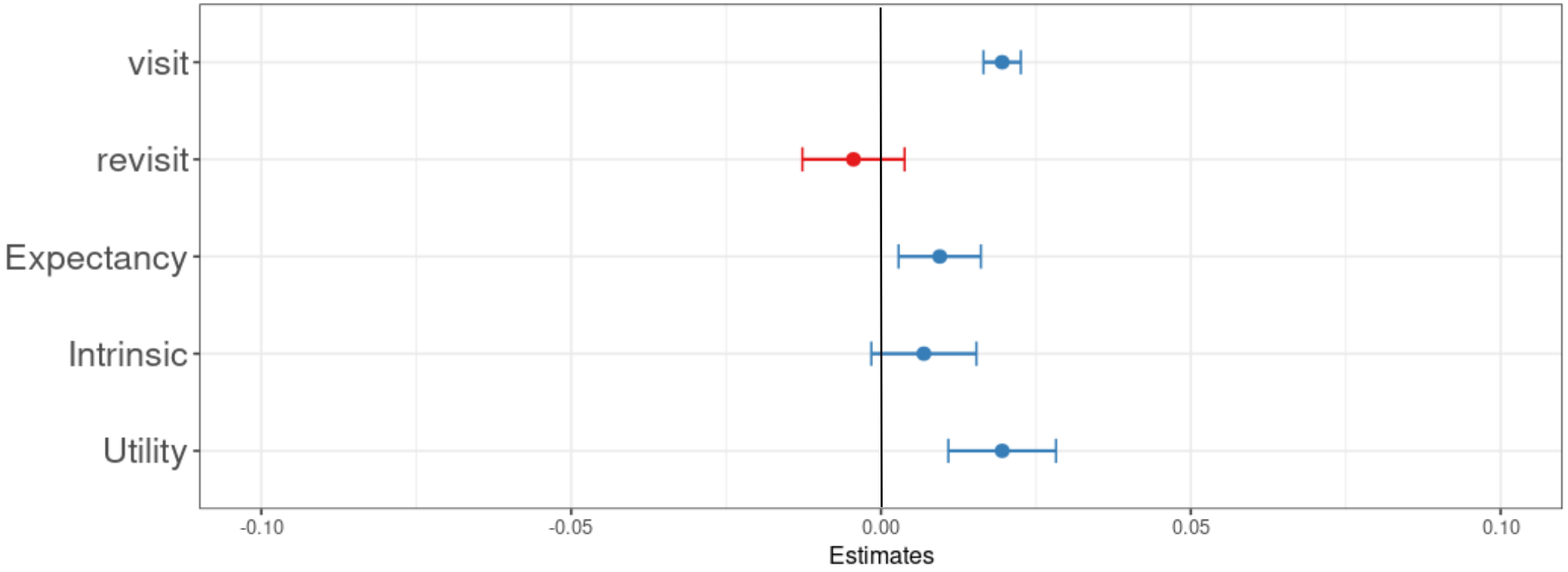
Bonferroni-corrected pairwise t-test: EOC (mostly) did not differ between chapters.

Motivation

How much influence does each variable have on the response variable (**End-of-Chapter Score)** while controlling for the individual differences between students and institutions?

Mixed Effects Models

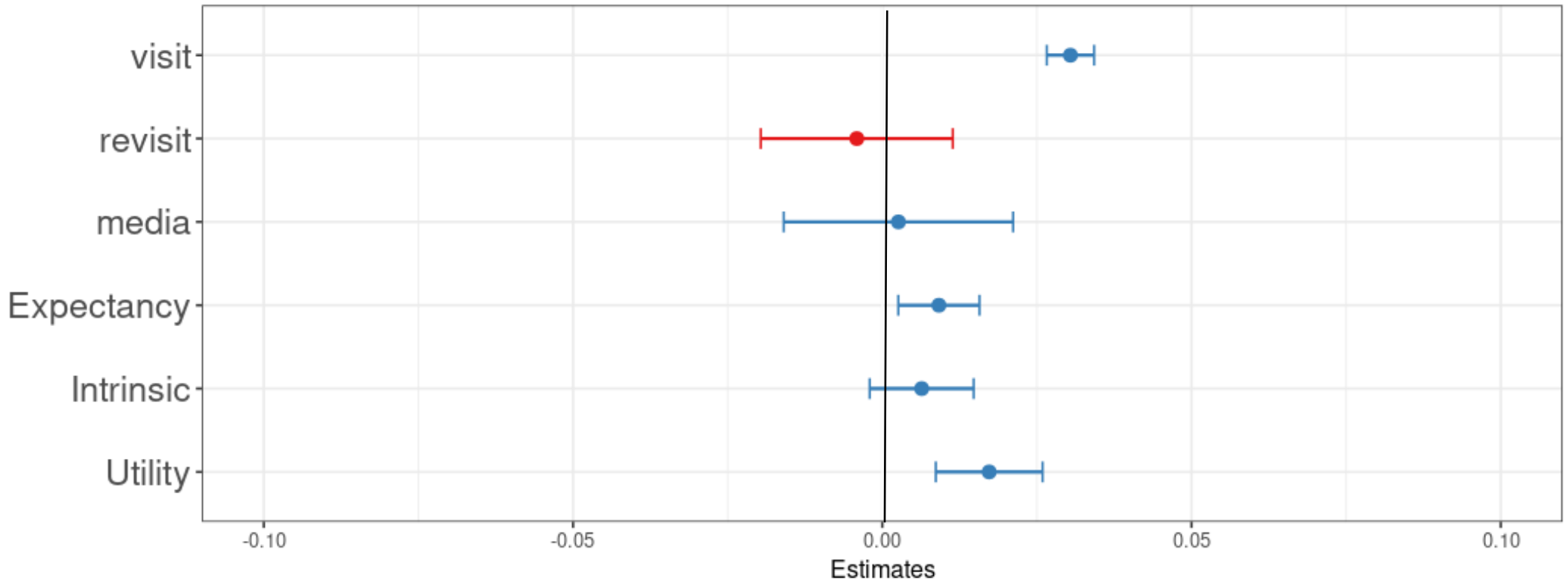
Mixed Effects Model without Media Engagement



Conditional R^2 = 0.69

Marginal R^2 = 0.07

Mixed Effects Model with Media Engagement



LCA

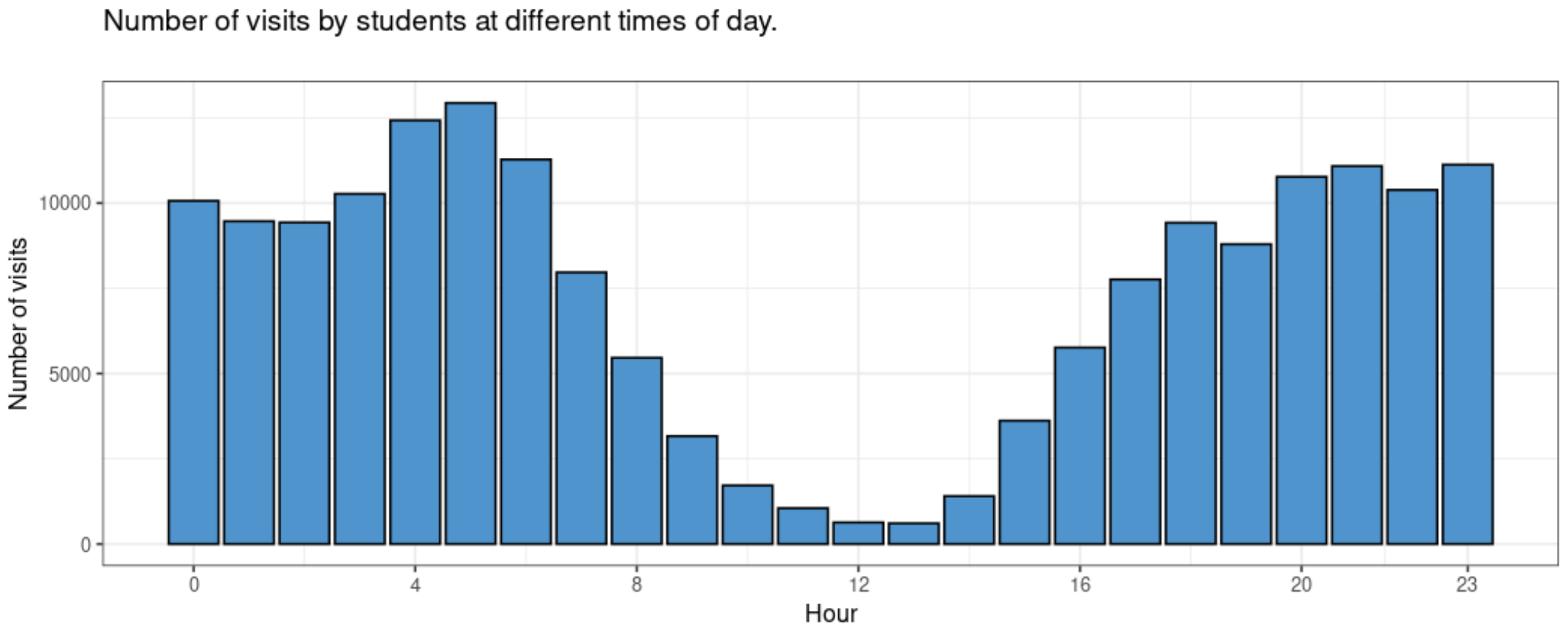
Conditional $R^2 = 0.70$

Marginal $R^2 = 0.09$

Further Analysis

Perform Latent Class Analysis – to further investigate the effect of different types of learning methods (interactive media, written textbooks, etc.) and how that influences student success

Time of Day Students Visit the Textbook



Remember to sleep!

Suggestions for CourseKata

- "Wellbeing Alert" to remind students to take a break when they are engaging with the textbook past midnight
- Students should maximize their engagement on their first visit to pages to improve their EOC performance
- Course organizers should ensure the chapters build off of one another
- Move the location of the Pulse Questionnaire to after the students take the end of chapter quiz, but before they receive their results