Data Appendix:

Information about the Analysis Data Files
Course_data_clean.csv: Our dataset, course_data_clean.csv, contains reviews at the course-review level. The reviews were entirely sourced from uwflow.com and the code that was used to web scrape the data is found at https://github.com/asusevski/course_reviews. We downloaded the csv file which contains the scraped data from https://www.kaggle.com/datasets/anthonysusevski/course-reviews-university-of-waterloo.

stopwords copy,txt: The custom stop words come from https://github.com/derekgreene/topic-model-tutorial/blob/master/stopwords.txt, a tutorial on topic modeling we leveraged to implement this project. We also added some additional stopwords that we deemed necessary from our EDA such as "course", "econ", and other words that do not provide meaning to our analysis.

Information about the variables in course_data_clean.csv

Column	Description	Potential Responses
course_code	Official code of the course established by the University of Waterloo	CS 115, ECON 101
course_title	Full title of the course	Introduction to Computer Science 1, Introduction to Microeconomics
num_ratings	Number of course reviews submitted by reviewers	2111, 1143
useful	Percentage of reviewers that stated they found the course useful	21%, 64%
easy	Percentage of reviewers that stated they found the course easy	10%, 70%
liked	Percentage of reviewers that stated they liked the course	23%, 45%
num_reviews	Number of reviewers that submitted a text review	86, 214
reviews	Course review submitted by the reviewers	● Not an "EASY" course at all. Anyone commentating that is probably a Math/Eng student who is already familiar with coding through other courses. Since the midterm/exam is on paper, you have to be able to imagine what your code is going to be able to do without actually testing it like you do in your assignments! Speaking of assignment, they're bloody hard! No joke, it might take you an entire day to problem solve your way through them and that might not even be enough - incredibly frustrating course and not recommended unless you

		really want to study coding in the future. DR. Racket is not even applicable to most organizations. If you want to learn coding, just pick up something off Amazon or get a pdf through a friend - don't spend money here unless you're actually planning on continuing studies in CS. ● Useful for my major? Nope. Useful for my general knowledge? Kinda. Easy? It's birdy for all profs except Smith.
course_rating	Whether the student liked or disliked the course overall	liked course, disliked course
course_rating_int	Whether or not the student liked the course displayed as an integer	1.0, 0.0