APPENDIX 0 TABLE 1
STATISTICAL ANALYSIS COURSE SCHEDULE

Class	Topic	Assignments	Project
1	Introduction		
2	Review	R downloaded HW 1 Review Released	
3	Review		
4	Bootstrapping	HW 1 Review HW 2 Bootstrapping Released	
5	ANOVA		
6	ANOVA	HW 2 Bootstrapping HW 3 ANOVA Released	
7	ANOVA		
8	Exam Review	HW 3 ANOVA	
9	SLR[a] Introduction		
10	**Exam 1**	HW 4 SLR Released	
11	SLR[a] Inferences		
12	SLR[a] Estimation and Prediction	HW 4 SLR	
13	SLR[a] Residual Analysis and var trans	HW 5 SLR Released	
14	SLR[a] Examples		
15	MLR[b] Introduction	HW 5 SLR	
16	Exam 2 Review		
17	**Exam 2**		
18	Data Set Discussion		Dataset Selection Project
19	MLR[b] Inferences	HW 6 MLR Released	
20	MLR[b] Estimation and Prediction		Dataset final selection and descriptive stats
21	MLR[b] Examples and Lessons		
22	Exam Review	HW 6 MLR Due	
23	**Exam 3**		
24	Non-Parametric Testing	Exam 3 Due	
25	Bootstrapping Methods OR Categorical Data		
26	*p*-curve		
27	Project Presentations		Project Submissions
28	Project Presentations		

[a]Simple Linear Regression [b]Multiple Linear Regression

Some considerations about the course structure and time-line mentioned above: We recommend six homework assignments throughout the course. The homework is assigned based on topics taught. They are intended to build on the knowledge taught in class and bring light to the most important aspects. We recommend them to be shorter in length, due to the frequency of them, and such that the student consolidates their knowledge before a new topic is introduced. We recommend including problems that resort to the use of R, but also variations on the settings of the problem that would lead the student to understand the limitations of the methods, and ultimately, the theory behind the methods deployed.

The course is structured to have three exams. The first one is close notes, the second one is open notes in-class, and the third one is a take-home exam. The structure should be a mix of R and handwritten work, such that the student is able not only to run the code but also be able to interpret what is happening. The first exam would cover the review, bootstrap, and ANOVA. Exam 2 will cover simple linear regression, and the third exam will cover multiple linear regression. Further, we propose that there should be two classes between the last topic covered and the exams. This seeks to allow for enough time to provide feedback to the students in the different homework and give time to review proposed solutions and attend office hours.

Students will have to do a final project. This is an individual project that will be segmented in multiple submissions over the course. The final project will ensure that students will actively work on it throughout the semester and apply what is covered in class. The project will encompass all that is learned up to Multiple Linear Regression. The main goal of the project is to identify and implement a data analysis research project and use R as the primary analysis tool. The final project involves the submissions of a dataset to be chosen by the student subject to the instructor's approval, final project report, and a five-minute presentation. The dataset selection is due in the middle of the course to ensure that the students have started developing their final project, and a discussion is to be taken place in-class to address the next steps (i.e., the submission of descriptive statistics, main questions to be addressed, and dataset limitations). The final report will narrate the whole process that was conducted (i.e., describing the question of interest, how the dataset was used to analyze the question with detail on the steps used for the analysis). The findings and limitations of the study are the last sections of the paper. The project presentation – depending on the size of the class – could entail a peer-review group grading system involving three to five people in each group, who do a peer review of the presentations among them.