	due in 6day 21h					
	Submission Phase					
	1. Do assignment ☐ (/introstats-001/human_grading/view/courses/970433/assessments/5/submissions)					
	Evaluation Phase					
	2. Evaluate peers வ (/introstats-001/human_grading/view/courses/970433/assessments/5/peerGradingSe	ets)				
	Results Phase					
	3. See results					
☐ In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.						
	Save draft Submit for grading					
	Results Phase  3. See results (/introstats-001/human_grading/view/courses/970433/assessments/5/results/mine)  In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.					

## Introduction to the Assignment

Hello everyone! Now that the videos and quizzes are running along smoothly, it's time for your first homework assignment. This assignment asks you to interpret some very special data -- data about you! Specifically, we will investigate data collected from students in this very MOOC, through the pre-course survey (available for viewing <a href="here">here</a> (https://spark-

<u>public.s3.amazonaws.com/introstats/Assignment1/PreCourseSurvey.pdf)</u>) and the background knowledge quiz. You'll have to look through various tables and figures about these data, and answer four different statistical questions about them, using the statistical ideas you have learned in Weeks 1 and 2 of this MOOC.

This assignment is not short, so please budget some extra time for it. (Yes, this is in ADDITION to the usual weekly load of new videos and quiz.) On the other hand, you are allowed just a limited number of words in your responses, so don't go overboard -- just concentrate on the most important points. Remember that your assignment will be viewed and graded by your peers next week. Your solutions do not have to be perfect or entirely complete, they just have to show that you have made a reasonable effort to think about the statistical issues involved. We hope that you will enjoy this assignment, and use it as a way of trying your hand at interpreting the results of a statistical analysis, without worrying TOO much

about the result as long as you make a sincere effort.

Note that this assignment has a HARD DEADLINE this coming Sunday (April 28) at 11:59 PM EDT, with no flexibility about that (sorry) because we need to start the peer assessment at that time. If you do not enter work before the submission deadline, you will not receive an evaluation for this assignment, nor will you be allowed to evaluate your classmates' work. So, please get started on this assignment right away, even before viewing the Week 4 videos and quiz (which are not required for this assignment, and which have a later Hard Deadline).

You can create and save multiple drafts any time during the submission process; your draft enters the system to be graded only when you click the "Submit for grading" button. (This means that if you haven't clicked "Submit for grading" by the deadline, your saved draft will not be evaluated, so please be sure to click it!) Even after submitting, you can return to your assignment, make changes and "Re-submit for grading" anytime up until the deadline. You will need to check the box "In accordance with the Honor Code..." to activate the Submit button.

Next week, we will continue with the peer assessment phase of this assignment. You will be asked to review five other students' assignments, and five other students will review your assignment too. Some general information about peer assessment is available <a href="https://example.com/heres/beauty-sep-align: review your assignment too.">https://example.com/heres/beauty-sep-align: review your assignment too.</a> Some general information about peer assessment is available <a href="https://example.com/heres/beauty-sep-align: review your assignment too.">https://example.com/heres/beauty-sep-align: review your assignment too.</a>

(http://help.coursera.org/customer/portal/articles/971175-how-do-peer-assessments-work-), and we will provide you with more details next week. Remember that you must participate fully in BOTH the assignment AND the peer assessment in order to receive credit for your work.

## **Background for the Assignment**

For this assignment, a simple random sample of 1,500 observations was selected from the 18,462 submissions of at least one of the pre-course survey and background knowledge quiz. Our data are the pre-course survey responses and answers to the background knowledge quiz for these 1,500 observations.

The key pdf document for this assignment is posted <a href="https://spark-public.s3.amazonaws.com/introstats/Assignment1/a1plotsandstats.pdf">here (https://spark-public.s3.amazonaws.com/introstats/Assignment1/a1plotsandstats.pdf</a>). That document gives more information about the data being considered, including five tables and four figures (plots). For this assignment, please use only the information provided in that document. (We won't be providing the data for now, partially because we don't want to disadvantage students who are not learning R, though we might post more details later.)

(As an aside, we note that the background survey also asked for what region of the world you live in. The possible response options were all continents, and several of you told us that where you live wasn't represented. We apologize for that; we simple used the standard Coursera regions, but now we see that they should be appropriately modified and expanded -- which we will do next time. Because of this concern, we decided not to use this question for the assignment. But for those of you who are curious about your fellow students: we have students enrolled from every continent (except Antarctica); 37% of respondents indicated they were from North America, 28% indicated Europe, 22% indicated Asia, 7% indicated South America, 4% indicated Africa, and 3% indicated Australia. And other students told us they were from Central America, New Zealand, and some Pacific islands, too.)

## The Actual Assignment

Your assignment is to use the information in the <u>pdf document (https://spark-public.s3.amazonaws.com/introstats/Assignment1/a1plotsandstats.pdf)</u> of tables and figures mentioned

state your answer.	
Assignment question 1	
(Categorical variables)	
Question 2 on the pre-course survey asked about the importance of some possible reasons for enrolling in the course. Consider the following two reasons: (1) "This class teaches skills that will help my job/career" and (2) "Because this course is offered by a prestigious university." Use Tables 2, 4, and 5 of	
the pdf document to answer the following two questions. Which of these two reasons seems to be most important to students in the class? And, do the reasons for enrolling differ depending on level of	
education?	
Format	
Words: 0 / 500	

above to answer the following four questions. Use the material and statistical knowledge that we have

covered in Weeks 1 and 2 of the MOOC; your answers are NOT expected to address material from Weeks 3 and 4 (nor anything we'll be covering later). Please focus on the most important features of the data that you observe, and be concise and to the point since you are allowed only a limited number of words to

## **Assignment question 2**

(One categorical and one quantitative variable)

For this question, we are interested in variables that are related to the total number of other MOOCs the student is enrolled in. Consider the following 2 variables: (1) level of education and (2) rating of the importance of "course is offered by a prestigious university". Use Table 3 and Figures 3 and 4 to draw conclusions about the relationships between each of these 2 variables and the total number of other MOOCs the student is enrolled in. Support your comments by referring to appropriate summary statistics and features of relevant plots.

Format			
	<u> </u>		

Assignment question 3		
Assignment question 3		
(Two quantitative variables)		
For this question, we are interested in whether the variables "t		
and "age" are related to the score on the background knowled to comment on the relationships you observe. Support your co	• .	
correlations and features of relevant plots.	g to appropria	-
Format		
Words: 0 / 500		
words. 07 300		
Assignment question 4		
To what group of people can you generalize any observations	you made about these data? Wh	u2
To what group of people call you generalize any observations	you made about these data: vvii	y :
Format		
Words: 0 / 500		
1		
☐ In accordance with the Honor Code, I certify that my ans	were here are my own work, and t	hat I
have appropriately acknowledged all external sources (if a		ιαι Ι
		a altimati
	Save draft Submit for gra	adıng

Words: 0 / 500