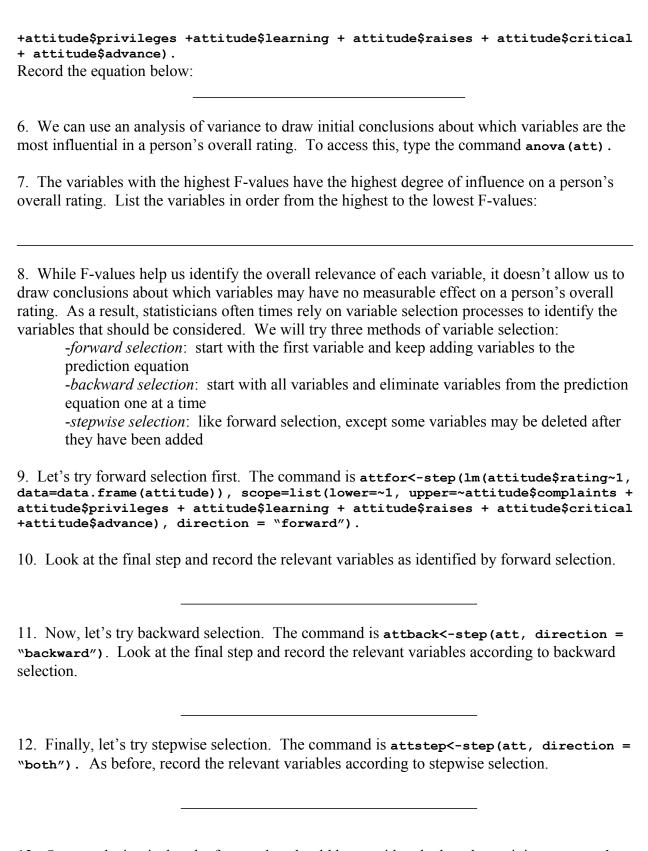
Lesson 2: Using R for Data Minin	Lesson 2	: Usir	ig R for	· Data	Mining
----------------------------------	----------	--------	----------	--------	--------

Name			

I. Many times, a situation depends on numerous factors. Someone's credit score, for example, isn't just determined by a single answer to a question on a loan application. List some of the
things you think a credit company asks for on a credit application.
Obviously, some pieces of information you listed above are more relevant than others in determining whether or not someone should get a loan. As a result, we need a process that can analyze these large amounts of data in order to determine which factors are the BEST predictors of someone's "credit-worthiness." One such process is called <i>data mining</i> .
II. Search the internet for five facts about data mining. Record them below:
1
2
3
4
5
III. We can use R to experiment with the data mining process. Let's first try out the process on a data set that is already in R.
1. Type attitude to see the data set we are going to use.
2. Type help(attitude) to see the details of the data set. It is always important to understand the source of data when trying to make predictions and draw conclusions!
3. List the variables presented in this data set below, and then circle the dependent variable:
4. Now return to your R console and load the data using the command data(attitude).

5. Let's find a linear model to relate the variables that impact a person's overall rating. Our command will be: att<-lm(attitude\$rating ~ attitude\$complaints



13. Our conclusion is that the factors that should be considered when determining someone's overall rating are:

14. 1	14. How could our conclusion be helpful to people studying the results of the survey?			
	Use the internet to research three uses of data mining. Make sure to explain how data ng is used in each of the three cases:			
3				