

Total Score: ___89___ out of 100 points

Following Guidelines (10 points)	
Was the final project turned in on time?	
	Total: ___10___ out of 10 points

Project set up and Data Collection and Cleaning	
(5 points) Were explanatory and response variables identified? Was the type of each variable mentioned? Was a reference given for the data set? There were two questions asked, are they being tackled separately or together?	Yes
(5 points) Was the data set described appropriately? If only part of the data were used, is the procedure for selecting that part described (Both variables or observations)? Was missing data discussed and explained how it was handled.	<p>The description of your recoding is a little confusing..."</p> <ul style="list-style-type: none"> Down.Pmt.Present: This flag is set to FALSE if the Down Payment is missing, otherwise it is set to TRUE. Once this value is computed, the missing values for Down Payment variable are set to -1. <p>So you are creating a new categorical variable and I'm pretty sure I know why you converted the missing of the original variable to -1, but you don't explain why this is helpful. Is it because you are still going to use this variable in some capacity? The bullet points are explaining it a little, but an upfront explanation of what that is going to help you do would be helpful to articulate your model building process. Its allowing you to model the variable as continuous without throwing things out due to missing</p>
	Total: ___9___ out of 10 points

Data Analysis and Model Building	
(10 points) Were appropriate descriptive statistics shown for various aspects of the data (response variables, explanatory variables, assessment of correlations, multicollinearity assessments? Were obvious issues discussed, mentioned, and/or attempted to address.	<p>This report was a little light in regards to the exploratory analysis. Missing issues were well discussed but with the exception of figure 2, theres really know way of getting a feel for how these variables behave. This, to me, really hurts the report, because if that was done, it would be much easier to explain what your final regression model is saying about your variables and how they are in relationship to</p>

	getting funded.
(10 points) Was there a graphical display of the response variable and the explanatory variable or a subset of them? Was the display appropriate? Were any special or unusual features noted and explained?	Just wish there was more here.
(10 points) Discussion of feature selection approaches (both algorithm and common sense logic). Summary of the procedures and appropriate displays to illustrate.	<p>“For creating the model, we used the backwards/stepwise model selection algorithm using AIC and BIC with the following interaction terms:” Just exclude the word interaction. As not everything is interaction terms in the bullet point list.</p> <p>After saying that, I think your intent was to just discuss the interactions because the result in Figure 3 include things like ResubmittalTRUE. I think it just needs to be thoroughly explained what variables are going in AS IS, and which ones are involved in interactions for missing, and others that are straight up traditional interactions.</p> <p>Including the bullet points to describe whats being considered for the modeling I think was a good strategy. However with the stuff above, its hard to get a feel for all of explanatories that were being used at it makes it hard looking at Figure 3 since there are is only one interaction in there.</p>
	Total: __25__ out of 30 points
Inference (For the concise model)	
(10 points) Was the appropriate inferential procedure used? Were the null and alternative hypotheses given? Were the population and sample correctly identified? Were the results interpreted in context?	<p>Here is a question for everyone to think about. If you know it, then great. If not then go find out....For the outputs in Figure 3 and Figure 4 are these the tables from running a final logistic model using the features from feature selection on the entire data set, or are these generated from the cross validations (kfolds) themselves to pick the predictors, or is this the test set? I’m asking this not really as a criticism, but just to emphasize the importance of knowing what this train function is actually providing when you ask it to save predictions and do a summary of the object.</p> <p>This could be done a bunch of different ways, so its important to know which one is it.</p> <p>You explicitly state Figure 5 is from the test set, Figure 3 is estimates from the 10fold cross validation. But its not clear where Figure 4 is coming from. If its from the training, then you should have shown it for the test. R code looks like its from test but just say it in the report.</p> <p>Just give me a little more out of your interpretation and</p>

	<p>summary of the final model. How does a variable effect the odds? You don't have a huge complicated model so that is a lost opportunity.</p> <p>All of these things you talked about at the end, you could have provided figures or tables of these and put them in the EDA portion that I dinged you on earlier.</p> <p>There are quite a few insignificant predictors in Figure 3. Are they critical for predicting? ROC's look good if removed? Etc. Even if not shown, a discussion on that would have been insightful.</p> <p>In reporting of the final model, you guys just stripped some R code in there. No coefficients are present so it looks like you are saying all the coefficients are 1. Either be sound with the formulation "$\log(\text{odds of funding}) = B_0 + b_1x + \dots$" or just put the explanatory variables in words inside of a sentence</p>
(10 points) Were the assumptions checked with the appropriate statistics or graphics?	No these were not discussed but we know diagnostics of logistic are rather tough. The cv and the test is sufficient.
(10 points) Was the appropriate output from software shown in order to support conclusions of the inferential procedure?	ROC's and the Confusion matrix were reported.
	Total: __25__ out of 30 points

Conclusions (10 points)	
Were weaknesses in the study pointed out? Were the conclusions given in the context of the appropriate population from which the sample was drawn?	Yes
	Total: __10__ out of 10 points

Clarity (10 points)	
(5 points) Is the report clear and concise? Did the report have sections? Did paragraph transitions	General writing of the report was sound. No issues here.

flow smoothly from one paragraph to the next? Were words overused? Were there excessive spelling or typographical errors?	
(5 points) Did graphics have appropriate titles and labeling? Were figure captions appropriate?	Yes
	Total: __89__ out of 10 points

Additional Comments: