 What is the key idea for locating people and why do we expect it to work? Background and methods: - You will need to do some EDA - EDA - focus your discussion of the results on what it means for the model - Be selective with your EDA. Focus on results that have a relevance to the project. You can always mention that you tried other visualizations, but didn’t get much more out of it. If you add a graph, it should add to your analysis and cover something that cannot already be seen in other graphs. - What are the model types that you plan to build? - Stratification to have blue tarps represented in the training set - I guess you meant test set. - If I understand you correctly, you plan to repeat a 10-fold cross-validation 10 times / bootstraps), so setting both “v = 10, repeats = 10” in vfold\_cv. This means, 100 models for each model. This is a lot. It will be ok for part 1, but part 2 has models with a lot more tuning parameters, so your model building and validation time will be considerable. If after you heard my comment on this, you still want to set repeats=10, I suggest you take any of the methods and try it out. You can then see how much performance metrics vary between different repeats and assess if it is worth to increase the calculation times 10-fold. - Model training consists of these steps: - Model validation using 10-fold cross validation; this includes model tuning if applicable - what metric will you use? - Threshold selection - what metric will you use? - Are “penalized logistic regression” and “logistic regression” really two different models or just variations of the same? Results: - When you present metrics, talk about their relevance for the objective of this project - When you present your results, split it into results from model validation and results after threshold selection. e.g. ROC curves are only relevant for discussing model validation results and not threshold selection; reporting accuracy or any other threshold dependent metric prior to threshold selection is irrelevant. - If you visualize the same analysis for each model, combine them either in a single graph or a combination of graphs in one figure (using patchwork) - I wouldn’t call ROC a metric. The AUC under the ROC curve is one, but a ROC curve is a representation of all possible sensitivity/specificity values for different thresholds. Discussion: - Is accuracy really that important? You can increase the number of correctly predicted blue tarps, by just making everything a blue tarp, however at the same time, the effort increases. - If you look at model runtimes, differentiate between training and prediction times. You could also estimate how much we need to predict. A back of the envelope type estimate would be interesting. X images with YxZ pixels times the prediction time. Is it really a problem and it takes only a few minutes, or do we look into predictions times of days or even weeks? That actually would be something I haven’t seen. The holdout set from part 2 has location information for the pixels, so you could in principle estimate quite well, how many pixels are required to cover all of Haiti with images