<project title>

GA Data Science - Final Project Outline

<name>

September 18, 2014

1. Problem to be solved

<summary / description of problem to be solved>

2. Description of dataset

* What dataset will you be using for your project?
* How will you access this data? API? Scraping? Other?
* What sort of information is in the dataset? In other words, what features are available?
* How will you turn that data into a training set? (If using a supervised approach)
* How do you anticipate processing that data to get it into a form to use for your modeling?

3. Hypothesis

What is your hypothesis? In other words, what do you hope to predict or otherwise learn as the outcome of your project?

What are some of the features you might use?

4. Statistical methods I plan to use and why

Think back to our 2x2 matrix slide: Is your problem a classification problem or a regression problem? Will your approach use supervised approaches or unsupervised approaches?

Which of the machine learning algorithms that we have learned do you plan to use for your final project and why? Which do you explicitly NOT plan to use and why?

5. Applications the finding may have

Once you have completed your project, what are some of the applications of your findings? In other words, how might those findings be applied? What is the “practical” value of the model you will have built?

Also, what will your deliverable be in addition to your code and data? Will you write a report in the style of CS229? Will you create a visualization? (NOTE: Do NOT attempt to learn D3 on top of everything else unless you are already a javascript ninja! Seriously.)

REFERENCES:

[Optional] links to relevant sources