This is a slightly longer (3-5 page) draft that should have the following:

* An introduction to the problem (based on points 1-2 your proposal above).
  + The Lending Club is an online, peer to peer lending platform. They aggregate requests for loans and investors in order to provide an alternative venue to the more common fundraising platforms. Investors participating in the platform face credit risk, the possibility that the loan they made will default. The lending club attempts to provide clarity by assigning a grade to loans that rates their likelihood of default and assigns interest rates based on said loans. I am exploring the possibility that it may be possible to predict the rate of default for the platform’s loans with greater accuracy, given the data they already collect, than the current rating system.
* A deeper dive into the data set:
  + What important fields and information does the data set have?
  + There are several
  + What are its limitations i.e. what are some questions that you cannot answer with this data set?
  + What kind of cleaning and wrangling did you need to do?
  + Are there other datasets you can find, use and combine with, to answer the questions that matter?
* Any preliminary exploration you’ve performed and your initial findings. Test the hypotheses one at a time. Often, the data story emerges as a result of a sequence of testing hypothesis e.g. You first tested if X was true, and because it wasn't, you tried Y, which turned out to be true.
* Based on these findings, what approach are you going to take? How has your approach changed from what you initially proposed, if applicable?

Add your code and milestone report to the github repository. As before, once your mentor has approved your milestone document, please share the github repository URL on the community and ask the community for feedback.