History G Feedback and Updates

Clara Seo and Graham Chickering

Feedback from Classmates

1)I incurred some errors in the Shiny app, and if believe it was because
minimum frequency was set to a higher number than the frequency of any word
in the selected chapter. If the user could be notified that this input is
illegal (or even better yet, not have the option to increase the slider past
a given chapter's maximum frequency for any word), that might alleviate some
user confusion as opposed to the current generic error.

RESPONSE: To fix this problem, we have added a checker that outputs a notification saying "Inputs not valid:Enter New Inputs" whenever an error would occur. This notification now prompts the user to either change the the minimum frequency or the number of words they want to see, and will not show the word cloud or the sentiment analysis graphs until this error has been corrected.

2) For all the great and thorough explanation and interpretation given with most of the project's
analysis, that given with the word cloud feels light in comparison. Perhaps a
sample interpretation of one or two of the figures (including an explanation of
how the parameters of the word cloud were chosen) would help guide the reader in
making sense of the many word clouds they could generate with the Shiny app.

RESPONSE: We ended up providing a couple sample interpretations of some of the word clouds in the report itself. We clarified how a word cloud is created and what the different sizes of the different words mean in the word cloud itself, with respect to how bigger words means they are more prevalent in the chapter. Our interpretations in the report itself should make it easier and more understandable for someone using our shiny app to understand the word cloud output themselves when they pick and change the inputs and the chapters they want to investigate.

3) Building on the wordcloud app, e.g. show the raw counts (quantitative)
to augment the qualitative (of the wordcloud), then maybe add inter-chapter
comparisons

RESPONSE: To build on the shiny app, we recognized that the word cloud by itself was not all that helpful. To build on what we already did, we decided to add two more panels that one can now observe sentiment analysis graphs or look at a table output of the sentiment analysis that shows the exact frequency and type of sentiment. By combining sentiment analysis with the word cloud itself, one can get a better understanding of the words and overall sentiment of each individual chapter. While the app only allows someone to look at one chapter (or the entire book) at a time, inter-chapter comparisons could still be made by changing back and forth between the chapters by changing the input selection.

4) Synthesising results from sentiment analysis in a more cohesive
way, e.g. in the discussion the result of each lexicon/approach is thrown out in
just chapter numbers, but how they are connected or even just giving the chapter
title in context would be helpful

RESPONSE: We made minor edits to make a more logical analysis. We changed the layout from bing to nrc to overall sentiment and gave chapter titles so that one can look at the different sentiments for the individual chapters, and added extra annotations for chapters that were outliers and stood out. We felt moving from bing to nrc to overall sentiment was a logical progression and highlighted some of the different sentiments that could be found in each of the chapters.

5) (This might be just me but) depending on the choice of input parameters in
the Shiny app, an error message might appear. Perhaps text could be printed to
say what the issue is.

RESPONSE: We have fixed this issue by rather than having an error message appear, instead having a popup notification appear saying "Inputs not valid:Enter New Inputs". This notification will continue popping up until the input is valid and then either the word cloud or sentiment analysis will appear.

6) At some occasional points throughout the paper, the structure of sentences can be convoluted and difficult to follow.

RESPONSE: We made a bunch of edits to both the executive summary and the analysis portion of the report to better clarify what exactly we are investigating and what our results were. We believe the edits we made addressed this issue and now much easier to understand and interpret the results from our project.

7) I would suggest printing out tables in your report using kable. This will
allow for an overall cleaner presentation of information, and I find it easier
to read a table when it's presented using kable. Also, with kable, you can
easily number the tables.

RESPONSE: We ended up changing all the tables in the report to use kable in order to make the output of our tables and the overall report look cleaner and easier to interpret.

8) Some of the words in the word clouds were arguably
too faint in color to see against the white background of the Shiny app. I would
suggest changing the color scheme of the word clouds so that smaller words are
still visible.

RESPONSE: While our original thought was that we wanted the color scheme to reflect Amherst's colors, we agree that some of the words can be very faint and hard to see so we adjusted the color scheme slightly so that that smaller words are still visible.

FEEDBACK FROM NICK

1) Lots of corrections and edits to make in the Executive Summary. This included adding full citations, fixing grammatical mistakes, and improving on the overall analysis of what we learned.

RESPONSE: To edit the executive summary, we ended up fixing the citations and turned them into full citations, as well as fixing some of the small grammatical mistakes that were in our original summary. We also clarified what tf-idf is a measure of and why it was important for our analysis. One of the other big pieces we made changes was in our original analysis we focused a lot on the limitations of our analysis, so to build off that we focused on insights that we found through our text analysis to compliment the original work that we did. We also made sure to reflect back to our goals we laid out in the beginning and address if we were successful or not. These were the major changes that we made for the Executive Summary for Biddy.

2) Clean tables and limit number of rows for tf-idf output

RESPONSE: We changed out table output to use kable instead to make the output of our tables look cleaner and easier to read and interpret.

3) Add annotations to sentiment analysis graphs

RESPONSE: We added the annotations to specific chapters on the graphs and also changed some of the labels of one of our axises.

4) Perhaps drop "amherst" and "college" since it's so common?

RESPONSE: We ended up deciding to keep the words "amherst" and "college" in our word clouds. While we recognize that it might seem obvious that a book about amherst college is going to contain a high frequency of those two words, we felt leaving those words in there better reflected what each individual chapter was about. That way we know the other words and what is being talked about is related to Amherst College specifically.

5) Shiny App fairly minimal and word clouds on their own are not particularly compelling.

RESPONSE: To make our shiny app more substantial we ended up building off the word cloud to now include graphs that show the sentiment of the top words in each chapter and a table that shows the output of the words being looked at on the graph. Each of these graphs is now contained in their own separate tab so that the user can determine what type of graph or table they want to look at. We also changed the overall theme of the app so that it has a much more polished and cleaned up look to it. We also added a popup- notification that makes sure that inputs that are selected will actually contain some output, and if not then they require the user to change their inputs until the criteria are met. Overall these additions will allow one to observe not just the top words in each chapter, but the sentiment of those words themselves in order to get an understanding of not just what is being talked about in the selected chapter, but how those topics are being talked about and if they are positive or negative.