Preliminary Data Analysis –Text

For MSHE 407

# Analyzing Text

## Wordle ([www.wordle.com](http://www.wordle.com))

* + 1. Click on the “create” tab at the top
    2. Wordle is a fun (though non-scientific) way to analyze themes in text. Wordle will generate an image of words where the most frequently used words are larger.
    3. Paste in text from interview transcripts, your lit review or any other text you’d like to “analyze” and hit “go”
       1. Note: You will need to have (or download) a Java plugin. Mac users may encounter an issue with viewing the Wordle but this issue can be resolved. See trouble shooting tips at http://www.wordle.net/faq#redx
    4. The Wordle will be generated in a random format which you can then modify.
       1. To modify the Worldle after it has been generated it is best to hit “open in window”
       2. You can modify the font, layout, the colors used, etc. using the dropdown menus at the top (I like “edit custom palette the best”
       3. Wordle automatically removes common words, but it can be interesting to see which common words are most frequently used. To do this go to “language” and select “do not remove common words”
    5. Once the Wordle looks how you want it to you can save it as a PDF and then insert it into your MP presentation etc.
  1. Microsoft Word
     1. One way to begin to quantify themes in an interview transcript is to use Word’s find and replace function
        1. Use the “list matches in sidebar” function
        2. To make sure that you’re not excluding variations of the word you’re searching for type only the first part of the word (e.g. “independ” instead of independent or independence) and choose “all word forms” from the first drop down menu in the side bar
        3. Word will highlight each instance of the word. You can then read the sentence surrounding each case to decide if it is evidence of a theme. If it is you can then highlight the text and add a code.
           1. It is a good idea to have your code reflect both the theme and the interviewee (for the “independ” code I may use “Ind 1” for mentions of independence in my first interview.
           2. The advantage of adding a code is that you don’t have to have the actual word “independence” to code it as such.
        4. After you’ve coded you can do the search process again, this time searching for your code. You can search for all codes “ind” to see how many total mentions and then for just “ind 1” to see how many of those were in the first interview.