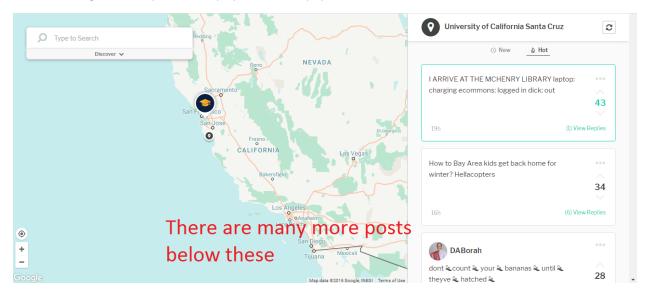
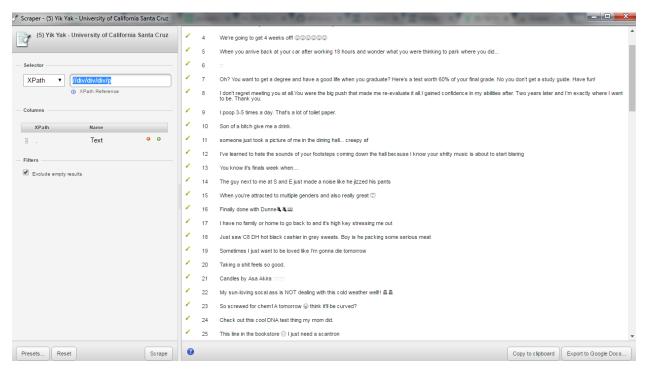
HOW WE COLLECTED OUR DATA

Step 1: Getting raw data from Yik-Yak website

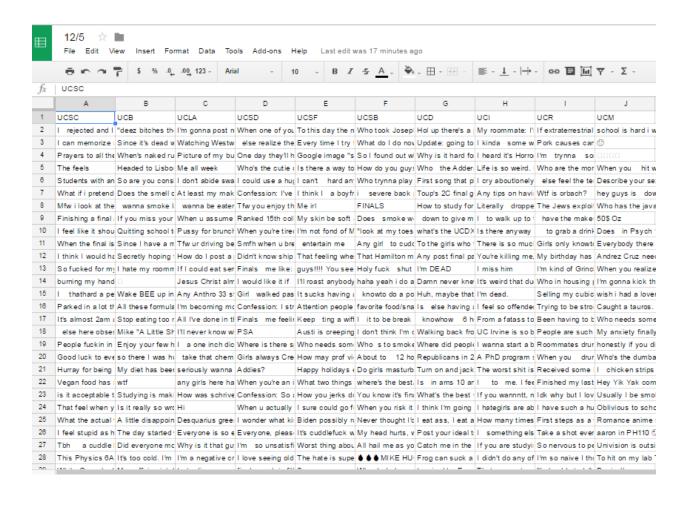
We first navigate to https://www.yikyak.com/map?place=(UC we wish to take data from)



We then use the chrome extension "Scraper" to grab the text of all the posts from the UC we are viewing, and copy all the text to clipboard



We do this for all UCs, and compile all data into a google sheet:



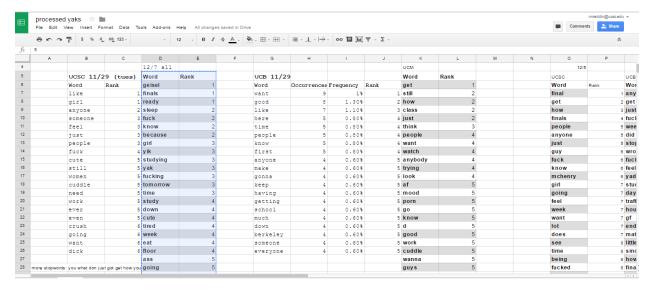
Step 2: Processing Data

We take a column of posts from a UC and put them into an online word processor found at wordcounter.com:

WORDCOUNTER

Here are your results. Nordcounter ranks the most frequently used words in any given body of text. Use this to see Word Frequency vhat words you overuse (is everything a "solution" for you?) or maybe just to find some final 83 reywords from a document. finals 50 Nordcounter is useful for writers, editors, students, and anyone who thinks that they might be fuck 43 speaking redundantly or repetitively -- and it's free! Eventually, I'm going to expand it so that people 38 35 ou can upload documents, but not yet. know 32 feel Enter the body of text here (to count & rank the word frequency): 32 one confession. I ve successfully found many of my crushes on fo, and feel like a scalker girl 30 Ifw you enjoy the views from Geisel but at night all you see when you look up from your study 29 laptop is your own reflection on the window time 29 Ranked 15th college in the world and we don't enough smart people to keep our fucking 28 take vifi going. Dammit Geisel When you're tired of studying or feeling down, remember that you're still better than me week 28 because you're not the idiot that drove away with your laptop on top of your car good 27 go 26 did 26 nclude Small Words ("the", "it", etc)? No -- exclude them ▼ 25 going 25 does Jse Only Roots (group variations together)? No 25 guys 24 how How Many Words should I list? 25 ▼ really 23 23 girls Go >> gonna 23 23 day because23 22 guy

Then we removed some more useless stopwords like go, did, plurals of words, etc.



Finally we converted these TSV into CSV because we were more comfortable working with CSV files.