Progress Reports - Social Media Comparison Team 1

Slick, Eric <ERIC.SLICK@UCDENVER.EDU>

Fri 10/1/2021 8:07 PM

To: Parcheta, Debra <Debra.Parcheta@ucdenver.edu>

Cc: Mansbacher, Peter <PETER.MANSBACHER@UCDENVER.EDU>; Dicochea, Andres <ANDRES.DICOCHEA@UCDENVER.EDU>;
Mortensen, James <JAMES.MORTENSEN@UCDENVER.EDU>

Hi Debi,

I hope you're doing well tonight, maybe enjoying a glass of wine! I wanted to send a consolidated email regarding the progress reports for each of our assigned social media outlets.

Peter: TikTok/Instagram

I'm thinking Instagram and tik Tok are going to be like YouTube. The process seems like it will involve crawling through each post to gather the info. Still need a lot more time to figure out that algorithm for both platforms. Also, Instagram has regular picture /video posts and reels. I'm thinking scope wise, just focusing on gathering data from the pictures/videos there. I'm not sure how incomplete of a picture that will create, data wise.

James: Youtube

Definitely possible with beautiful soup, just complex. To retrieve all videos released during users timeframe, we will have to use the brands' channel page (specifically— page with all video uploads) and we can create an array that has all the urls for the videos within the timeframe. Then, we will have to scrape all urls within that array.

Andres: Pinterest

- Initially I tried to run Beautiful Soup to web scrape, but it was not working because it would output the error of "Oh no! Pinterest doesn't work unless you turn on JavaScript"
- Had to use selenium
- This does make it, so you must place the browser web driver into the same folder as program
- Was able to get around the JavaScript error but the 'valuable' info is still hidden
- Example: Searching for comments in html code, you can find "Comments" but not the number

Myself: Facebook/Twitter

Facebook and Twitter both have been fruitful in returning lots of HTML code to me using BeautifulSoup. In terms of Facebook, I'm currently looking into evaluating brand handles based on user input and finding the correct Facebook page. For Twitter, I've had more trouble grabbing useful information but have identified a potential framework for use in: "Tweepy". Tweepy is also Python based and works extremely well at grabbing tweets and their related information.

Thanks, please let me know if you have any questions.
Eric