Lauren Tahari

SI 206

Final Project

Goals:

My goals for this project was to get a better understanding of my gmail use and general Youtube use. I wanted to identify when I receive most of my emails (via Gmail) and when most Youtube videos are uploaded (based on subject content). I set out to extract Gmail information and content such as 100 received messages, their date, time window, day of week and message content. I also attempted to extract View Count from each uploaded YouTube video.

I wanted to create easy-to-read visualizations that would accurately depict these findings. Some of my other goals included that successfully using APIs, effectively caching data and storing information in a database that is easily retrievable. Initially, my goals also included using Facebook, Instagram and Github to extract data but I decided against using them.

Achieved or not?

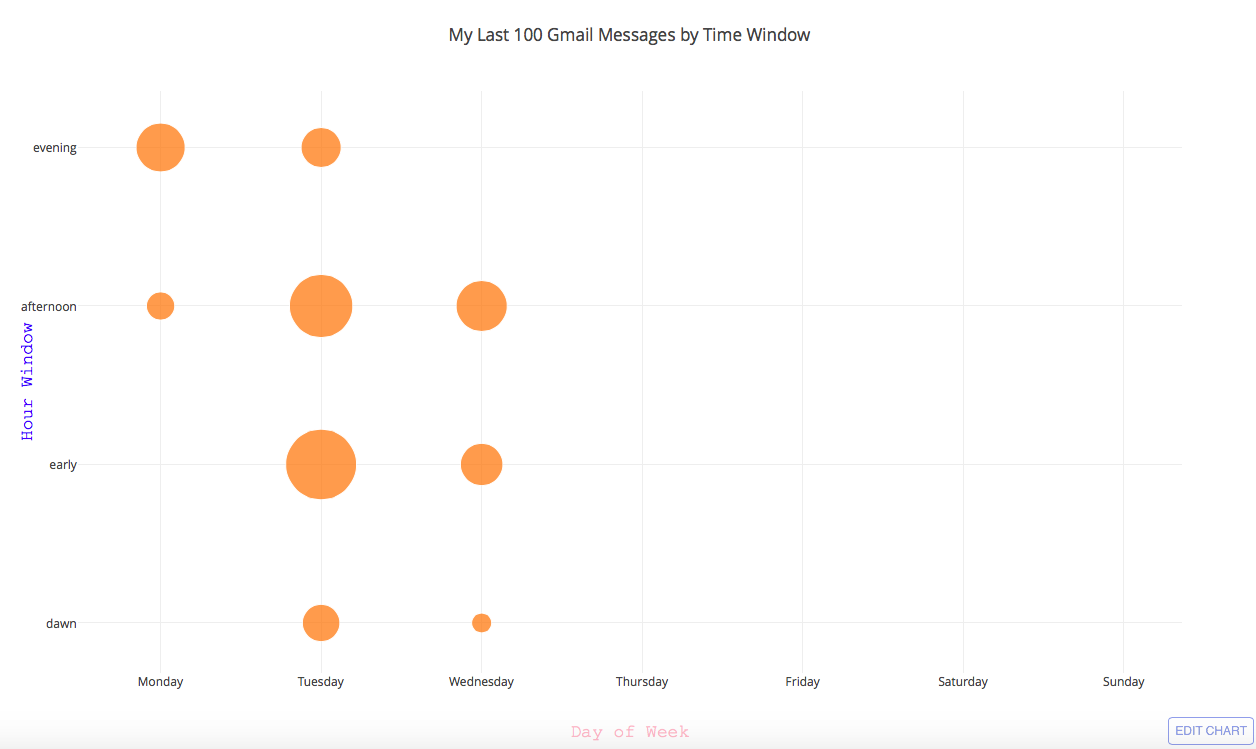
I successfully achieved these goals and gained a deeper understanding of my incoming email routine and at what time intervals I receive the most emails. Although I didn't extract view count, I was able to extract what day users uploaded videos most frequently to YouTube. I also created accurate visualizations depicting these findings. I achieved many of the goals that I set out to complete- however, I didn't attempt to use APIs to extract data from Facebook, Instagram and Github because I was content with the information I extracted from Gmail and YouTube.

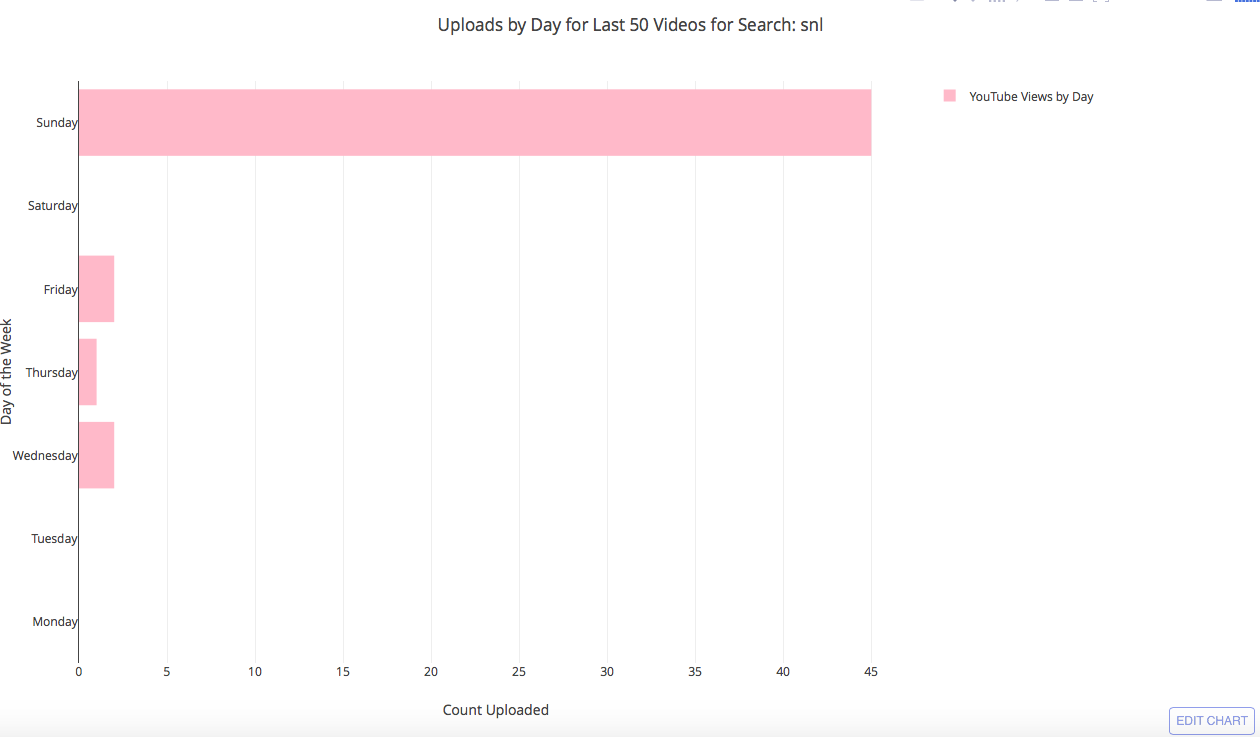
Problems

Overall, I had a bunch of de-bugging problems. I had difficulty converting the timestamp Gmail offered into understandable date and time format. Furthermore, I initially wanted to extract threads but I noticed that I haven’t had 100 threads in my Gmail account and therefore extracted emails. I also had difficulty identifying how to extract information from Youtube, but after using Youtube’s API help, I was able to successfully extract when videos were uploaded. Lastly, I found it quite difficult to manipulate visualizations on Plotly. I especially found it difficult to switch the bar graph from vertical to horizontal (which ended up being a very easy fix).

Social Media Report

Based on the visualizations below, it is evident that I receive most of my emails during the afternoon and early morning hours. Furthermore, it is clear that I receive approximately 100 emails every 3 days. Based on the second visualizations it is evident that most Saturday Night Live videos are uploaded on Sunday (with some being uploaded on Wednesday, Thursday and Friday)- however, it is important to note that these only accounts for 50 of the most recent uploads. Furthermore, the Youtube function can give users a deeper understanding of what day any type of video is uploaded (not just SNL subject related videos).





Instructions for running code

First, run the code. View a visualization of what days and what times I received my last 100 emails. You will be prompted to enter a search into Youtube. Search something you are interested in knowing what days those videos were uploaded on! Then-chose what color you want the bar graph to appear in.

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Issue Description** | **Location of Resource** | **Result**  **(did it solve the issue?** |
| 12/8 | How to set up Gmail API | <https://developers.google.com/api-client-library/python/start/get_started>  <https://developers.google.com/resources/api-libraries/documentation/gmail/v1/python/latest/gmail_v1.users.html>  <https://developers.google.com/resources/api-libraries/documentation/gmail/v1/python/latest/gmail_v1.users.messages.html> | Yes |
| 12/8 | How to Access Gmail Threads | <https://developers.google.com/resources/api-libraries/documentation/gmail/v1/python/latest/gmail_v1.users.messages.html#get> | No- no date time of treads available |
| 12/8 | Variable to messages | <https://developers.google.com/resources/api-libraries/documentation/gmail/v1/python/latest/gmail_v1.users.messages.html#get> | Yes |
| 12/8 | Extract Dates from Messages | <https://github.com/abhishekchhibber/Gmail-Api-through-Python/blob/master/gmail_read.py> | yes |
| 12/8 | HOW TO CONVERT FROM SECONDS (since January 1, 1970) to date format | <https://stackoverflow.com/questions/3682748/converting-unix-timestamp-string-to-readable-date-in-python> | yes |
| 12/8 | Identify which variable accesses each date format | <https://docs.python.org/2/library/time.html> | yes |
| 12/10 | DOUBLE CHECK IF CONVERSION IS CORRECT | <https://www.epochconverter.com> | yes |
| 12/10 | WRITE THE CSV FILE | <https://docs.python.org/2/library/csv.html> | yes |
| 12/10 | [TypeError: a bytes-like object is required, not 'str' in python and CSV](https://stackoverflow.com/questions/34283178/typeerror-a-bytes-like-object-is-required-not-str-in-python-and-csv) | <https://stackoverflow.com/questions/34283178/typeerror-a-bytes-like-object-is-required-not-str-in-python-and-csv> | yes |
| 12/10 | Used to check if Json was formatted correctly | <https://jsonformatter.curiousconcept.com> | yes |
| 12/10 | Youtube API how to begin | <https://developers.google.com/youtube/v3/getting-started> | yes |
| 12/10 | Retrieve video statistics | <https://developers.google.com/youtube/v3/docs/videos/list> | yes |
| 12/10 | How to extract upload date | <https://developers.google.com/youtube/v3/code_samples/python#search_by_location> | yes |
| 12/10 | how to change color of axis | <https://plot.ly/python/figure-labels/> | yes |
| 12/10 | HOW TO CHANGE TO HORIZONATAL GRAPH (didn't help) | <https://plot.ly/python/horizontal-bar-charts/> | Didn't help |