Reddit account for questions: atxProgramming

Andy's goal > track the below:

total subscriber count

Video view count

Video upload date

+ratio

Recent upload date (within 1-30 days) Video views (on a single video, not total video views for the channel): 25000+ Total subscriber count: 5000

Task#1

1. Python scrapes html from: <https://www.youtube.com/feed/trending>
2. Builds list of all users <a href="[/user/hyperkingames](https://www.youtube.com/user/hyperkingames) > <https://www.youtube.com/user/hyperkingames/videos>
3. Pull html from their videos and calculate ratio of [#views/#subs > 5]

subscribers">4,013</span>

1. If > 5 adds that video to a list
2. Email that list out

 --------------------------------------------------------------------------------------------------------------------------------

# Python snippets

From youtube trending all the subscribers:

>>> print(rawList)

{'/user/usjTV', '/user/KatyPerryMusic', '/user/VolbeatVEVO', '/user/channel4', '/user/Channel2ActionNews', '/user/hyperkingames', '/user/HellthyJunkFood', '/user/cmtvideo', '/user/ProZD', '/user/pcgamer', '/user/FamilyFeud', '/user/enyay', '/user/DaveHax', '/user/CBSBostonWBZ', '/user/CBS', '/user/TheLateLateShow', '/user/RealTime', '/user/ajw61185', '/user/MMAFightingonSBN', '/user/fishboy242', '/user/NBCNews', '/user/LateNightSeth', '/user/HollyscoopTV', '/user/corycotton', '/user/metrovideogame', '/user/AnonymousBelieberGuy', '/user/bbcintroducing', '/user/CBSNewsOnline', '/user/bgfilms', '/user/Mot898', '/user/MrMattyPlays', '/user/tldtoday', '/user/TEDEducation', '/user/coldfustion', '/user/JorjaSmithVEVO', '/user/dustinbones', '/user/TheAlexJonesChannel', '/user/ABCTheView', '/user/isleofmantt', '/user/PBSNewsHour', '/user/xbox', '/user/odeszamusic', '/user/lifenoggin', '/user/LastWeekTonight', '/user/rhettandlink2', '/user/ChrisStuckmann', '/user/TheFineBros', '/user/nigahiga', '/user/breakfastclubpowerfm', '/user/osnapitzari', '/user/AmericasGotTalent', '/user/Cbsfacethenation1', '/user/EEntertainment', '/user/Variety', '/user/MARVEL', '/user/WolfAliceVEVO', '/user/rudymancuso', '/user/CNN'}

Take user list and build new link list. Generated html list:

<https://www.youtube.com/user/hyperkingames/videos>

...

<https://www.youtube.com/user/usjTV/videos>

<https://www.youtube.com/user/KatyPerryMusic/videos>

<https://www.youtube.com/user/VolbeatVEVO/videos>

<https://www.youtube.com/user/Channel2ActionNews/videos>

<https://www.youtube.com/user/ProZD/videos>

<https://www.youtube.com/user/DaveHax/videos>

<https://www.youtube.com/user/CBSBostonWBZ/videos>

<https://www.youtube.com/user/TheLateLateShow/videos>

<https://www.youtube.com/user/ajw61185/videos>

<https://www.youtube.com/user/LateNightSeth/videos>

<https://www.youtube.com/user/HollyscoopTV/videos>

<https://www.youtube.com/user/corycotton/videos>

<https://www.youtube.com/user/CBSNewsOnline/videos>

<https://www.youtube.com/user/bgfilms/videos>

<https://www.youtube.com/user/Mot898/videos>

<https://www.youtube.com/user/TEDEducation/videos>

<https://www.youtube.com/user/JorjaSmithVEVO/videos>

<https://www.youtube.com/user/isleofmantt/videos>

<https://www.youtube.com/user/ABCTheView/videos>

<https://www.youtube.com/user/xbox/videos>

<https://www.youtube.com/user/LastWeekTonight/videos>

<https://www.youtube.com/user/rhettandlink2/videos>

<https://www.youtube.com/user/ChrisStuckmann/videos>

<https://www.youtube.com/user/TheFineBros/videos>

<https://www.youtube.com/user/AmericasGotTalent/videos>

<https://www.youtube.com/user/EEntertainment/videos>

<https://www.youtube.com/user/Variety/videos>

<https://www.youtube.com/user/MARVEL/videos>

<https://www.youtube.com/user/CNN/videos>

<https://www.youtube.com/user/channel4/videos>

<https://www.youtube.com/user/hyperkingames/videos>

<https://www.youtube.com/user/HellthyJunkFood/videos>

<https://www.youtube.com/user/cmtvideo/videos>

<https://www.youtube.com/user/pcgamer/videos>

<https://www.youtube.com/user/FamilyFeud/videos>

<https://www.youtube.com/user/enyay/videos>

<https://www.youtube.com/user/CBS/videos>

<https://www.youtube.com/user/RealTime/videos>

<https://www.youtube.com/user/MMAFightingonSBN/videos>

<https://www.youtube.com/user/fishboy242/videos>

<https://www.youtube.com/user/NBCNews/videos>

<https://www.youtube.com/user/metrovideogame/videos>

<https://www.youtube.com/user/AnonymousBelieberGuy/videos>

<https://www.youtube.com/user/bbcintroducing/videos>

<https://www.youtube.com/user/MrMattyPlays/videos>

<https://www.youtube.com/user/tldtoday/videos>

<https://www.youtube.com/user/coldfustion/videos>

<https://www.youtube.com/user/dustinbones/videos>

<https://www.youtube.com/user/TheAlexJonesChannel/videos>

<https://www.youtube.com/user/PBSNewsHour/videos>

<https://www.youtube.com/user/odeszamusic/videos>

<https://www.youtube.com/user/lifenoggin/videos>

<https://www.youtube.com/user/nigahiga/videos>

<https://www.youtube.com/user/breakfastclubpowerfm/videos>

<https://www.youtube.com/user/WolfAliceVEVO/videos>

<https://www.youtube.com/user/Cbsfacethenation1/videos>

<https://www.youtube.com/user/osnapitzari/videos>

<https://www.youtube.com/user/rudymancuso/videos>

Each of those links can then be scraped. At this point I hardcoded for the hyperkingames, just to proof of concept this. I’ll add code to loop tomorrow.

Output from python:

>>> spans.contents

['4,022'] <- subscribers

>>> viewsList

[<div class="yt-lockup-meta">

<ul class="yt-lockup-meta-info"><li>61,811 views</li><li>1 day ago</li></ul> </div>, <div class="yt-lockup-meta">

<ul class="yt-lockup-meta-info"><li>351 views</li><li>3 months ago</li></ul> </div>, <div class="yt-lockup-meta">

<ul class="yt-lockup-meta-info"><li>550 views</li><li>3 months ago</li></ul> </div>, <div class="yt-lockup-meta">

<ul class="yt-lockup-meta-info"><li>393 views</li><li>3 months ago</li></ul> </div>, <div class="yt-lockup-meta">

<ul class="yt-lockup-meta-info"><li>115 views</li><li>3 months ago</li></ul> </div>, <div class="yt-lockup-meta">

….. cut the list short, but this will provide 30 vids on the users videos page.

Haven’t done the ratio, but I’ll add that tomorrow. Shouldn’t be hard to calculate since I already have the views and sub count.

\*\*\*At this scale I think python on one machine is doable. If we are going to blow this up with more users and videos I think we could start looking into setting up some servers through Amazon AWS(fuckin bezos….) would be cool to try out\*\*\*