YouTube Video Data Downloads

- As of 2021-12-20

```
In [1]:  # Imports
         import pandas as pd
         import numpy as np
         import math
         import pickle
         from api_keys import youtube_key
         from googleapiclient.discovery import build
```

Obtaining Playlist ID for Pulling Video Data

```
In [2]:  youtube = build('youtube','v3',developerKey=youtube_key)

         # Channel ID pulled from website
         channel_id = "UCZNTsLA6t6bRoj-5QRmqt_w"
```

```
In [3]:  request = youtube.channels().list(part='snippet,contentDetails,statistics',
                                            id=channel_id)
         response = request.execute()
         print(response)
```

{'kind': 'youtube#channelListResponse', 'etag': 'f13VZ1KkcgoXX2KKxLVyezzl0FE', 'pageInfo': {'totalResults': 1, 'resultsPerPage': 5}, 'items': [{'kind': 'youtube#channel', 'etag': '2czfmUQy4Z3ZmxOqw-5i1leCYNo', 'id': 'UCZNTsLA6t6bRoj-5QRmqt_w', 'snippet': {'title': 'LowkoTV', 'description': 'Hello everyone, my name is Lowko and welcome to my channel. I have a passion for strategy games and have been producing gaming video content for over 13 years! \n\nI usually upload videos Monday-Saturday on this channel, and daily on the MoreLowko YouTube channel.', 'customUrl': 'lowkotv', 'publishedAt': '2008-02-11T14:09:44Z', 'thumbnails': {'default': {'url': 'https://yt3.ggpht.com/RRBX1HmO7Ng4B2Gp9nHXtYM3n9Pui5LkUu7aix6SseaExKxXCsPckI_XCHFN6-AKpYMxoX7A=s88-c-k-c0x00ffffff-no-rj', 'width': 88, 'height': 88}, 'medium': {'url': 'https://yt3.ggpht.com/RRBX1HmO7Ng4B2Gp9nHXtYM3n9Pui5LkUu7aix6SseaExKxXCsPckI_XCHFN6-AKpYMxoX7A=s240-c-k-c0x00ffffff-no-rj', 'width': 240, 'height': 240}, 'high': {'url': 'https://yt3.ggpht.com/RRBX1HmO7Ng4B2Gp9nHXtYM3n9Pui5LkUu7aix6SseaExKxXCsPckI_XCHFN6-AKpYMxoX7A=s800-c-k-c0x00ffffff-no-rj', 'width': 800, 'height': 800}}, 'localized': {'title': 'LowkoTV', 'description': 'Hello everyone, my name is Lowko and welcome to my channel. I have a passion for strategy games and have been producing gaming video content for over 13 years! \n\nI usually upload videos Monday-Saturday on this channel, and daily on the MoreLowko YouTube channel.'}, 'country': 'NL'}, 'contentDetails': {'relatedPlaylists': {'likes': '', 'uploads': 'UUZNTsLA6t6bRoj-5QRmqt_w'}}, 'statistics': {'viewCount': '249103618', 'subscriberCount': '476000', 'hiddenSubscriberCount': False, 'videoCount': '3526'}}]}

```
In [4]:  playlist_id = response['items'][0]['contentDetails']['relatedPlaylists']['uploads']
         print(playlist_id)
```

```
UUZNTsLA6t6bRoj-5QRmqt_w
```

```
In [5]:  # Total # of videos for sanity check later
         num_videos = int(response['items'][0]['statistics']['videoCount'])
         num_videos
```

```
Out[5]:  3526
```

```
In [6]:  # Limit of 50 requests to api, grabbing number of loops in order to get all the data
         num_pages = math.ceil(num_videos/50)
         num_pages
```

```
Out[6]:  71
```

Obtaining Video IDs for All Videos On Channel

Starting with Video Responses containing ID for each video

```
In [7]:    # Container for Video Responses
           video_responses = []
```

```
In [8]:    # Looping through all pages since we're limited to 50 requests per call
           new_request = youtube.playlistItems().list(part='snippet,contentDetails',
                                                       playlistId=playlist_id,
                                                       maxResults=50)

           for page in range(num_pages):
               new_response = new_request.execute()
               video_responses.append(new_response)
               if page == (num_pages-1):
                   break
               else:
                   new_request = youtube.playlistItems().list(part='snippet,contentDetails',
                                                              playlistId=playlist_id,
                                                              maxResults=50,
                                                              pageToken=new_response['nextPageToken'])
```

```
In [9]:    # Checks
           len(video_responses)
```

```
Out[9]:    71
```

Video IDs

```
In [10]:   # Container for Video IDs
           video_ids = []
```

```
In [11]:   # Looping through the list of video dictionaries
           for response_index in range(len(video_responses)):
               for vid in video_responses[response_index]['items']:
                   video_ids.append(vid['contentDetails']['videoId'])
```

```
In [12]:   # Check if # of video IDs match the number of videos from summary
           print(len(video_ids))
           print(num_videos)
```

```
           3526
           3526
```

```
In [13]:   video_ids[:5]
```

```
Out[13]:   ['LHDQ8iSbdck', 'IM7OdVKFW4M', 'ReDXHIf6cuc', 'kWfuYGNq1C4', 'IUqQ_rVnSXg']
```

```
In [14]:   # Confirming no duplicates
           len(set(video_ids))
```

```
Out[14]:   3526
```

Obtaining Video Stats

```
In [15]:   # Grouping Video Ids into groups of 50 to pull max stats per request call
           combined_ids= [video_ids[i:i + 50] for i in range(0, len(video_ids), 50)]
```

```
len(combined_ids)
```

Out[15]: 71

In [16]:
```python
# Video stats data pull
video_stats=[]
for comb in combined_ids:
    stat_request = youtube.videos().list(part="snippet,contentDetails,statistics",
                                          id=comb)
    stat_response = stat_request.execute()
    video_stats.append(stat_response)
```

In [17]:
```python
len(video_stats)
```

Out[17]: 71

In [18]:
```python
len(video_stats[-1])
```

Out[18]: 4

In [19]:
```python
# Preview of 1 video stats
video_stats[0]['items'][0]
```

Out[19]:
{'kind': 'youtube#video',
 'etag': 'LDXfsBPHP_Uyp3WNbkUCxIb8f84',
 'id': 'LHDQ8iSbdck',
 'snippet': {'publishedAt': '2021-12-21T19:00:14Z',
  'channelId': 'UCZNTsLA6t6bRoj-5QRmqt_w',
  'title': 'StarCraft 2: EPIC COMEBACKS! (Scarlett vs ByuN)',
  'description': "Best-of-3 series of professional StarCraft 2 between Scarlett (Zerg) and ByuN (Terran). It's obvious that these two players are very familiar with each other as they play build orders and strategies that otherwise wouldn't make much sense in this series.\n\nSupport my work on Patreon: http://www.patreon.com/lowkotv\nBecome a YouTube member: https://lowko.tv/join\n\nMy second channel: http://lowko.tv/morelowko\nLowko Merch: http://lowko.tv/merch\n\nBe part of the community on Discord: http://discord.gg/lowkotv\nThe hardware setup I use: https://lowko.tv/setup/\n\nStarCraft 2 is a military science fiction real-time strategy video game developed and released by Blizzard Entertainment.\n\n#StarCraft2 #SC2 #RTS\n\n--\nhttp://lowko.tv\nhttp://twitter.com/lowkotv\nhttp://facebook.com/lowkotv\nhttp://instagram.com/lowkotv\nhttp://twitch.tv/lowkotv",
  'thumbnails': {'default': {'url': 'https://i.ytimg.com/vi/LHDQ8iSbdck/default.jpg',
    'width': 120,
    'height': 90},
   'medium': {'url': 'https://i.ytimg.com/vi/LHDQ8iSbdck/mqdefault.jpg',
    'width': 320,
    'height': 180},
   'high': {'url': 'https://i.ytimg.com/vi/LHDQ8iSbdck/hqdefault.jpg',
    'width': 480,
    'height': 360},
   'standard': {'url': 'https://i.ytimg.com/vi/LHDQ8iSbdck/sddefault.jpg',
    'width': 640,
    'height': 480},
   'maxres': {'url': 'https://i.ytimg.com/vi/LHDQ8iSbdck/maxresdefault.jpg',
    'width': 1280,
    'height': 720}},
  'channelTitle': 'LowkoTV',
  'tags': ['Lowko',
   'LowkoTV',
   'Simon Heijnen',
   'Tutorial',
   'Guide',
   'Commentary',
   'How-to',
   'StarCraft 2',
   'SC2',
```

```
    'StarCraft 2: Legacy of the Void',
    'Tournament',
    'Strategy',
    'StarCraft 2 cast',
    'Zerg',
    'Protoss',
    'Terran',
    'StarCraft 2021',
    'StarCraft 2 2021',
    'Scarlett',
    'ByuN'],
   'categoryId': '20',
   'liveBroadcastContent': 'none',
   'localized': {'title': 'StarCraft 2: EPIC COMEBACKS! (Scarlett vs ByuN)',
    'description': "Best-of-3 series of professional StarCraft 2 between Scarlett (Zerg) and ByuN (Terran). It's obvious that these two players are very familiar with eac
h other as they play build orders and strategies that otherwise wouldn't make much sense in this series.\n\nSupport my work on Patreon: http://www.patreon.com/lowkotv\nB
ecome a YouTube member: https://lowko.tv/join\n\nMy second channel: http://lowko.tv/morelowko\nLowko Merch: http://lowko.tv/merch\n\nBe part of the community on Discord:
http://discord.gg/lowkotv\nThe hardware setup I use: https://lowko.tv/setup/\n\nStarCraft 2 is a military science fiction real-time strategy video game developed and rel
eased by Blizzard Entertainment.\n\n#StarCraft2 #SC2 #RTS\n\n--\nhttp://lowko.tv\nhttp://twitter.com/lowkotv\nhttp://facebook.com/lowkotv\nhttp://instagram.com/lowkotv\n
http://twitch.tv/lowkotv"},
   'defaultAudioLanguage': 'en-GB'},
  'contentDetails': {'duration': 'PT51M',
   'dimension': '2d',
   'definition': 'hd',
   'caption': 'false',
   'licensedContent': True,
   'contentRating': {},
   'projection': 'rectangular'},
  'statistics': {'viewCount': '8088',
   'likeCount': '500',
   'favoriteCount': '0',
   'commentCount': '35'}}
```

Pulling out video stats into one main list

```python
In [20]: all_stats = []
         for response_index in range(len(video_stats)):
             for vid in video_stats[response_index]['items']:
                 all_stats.append(vid)
         len(all_stats)
```

Out[20]: 3526

Saving results into a pickle file

```python
In [21]: pickle.dump(all_stats, open( r"Data/all_stats.p", "wb" ) )
```