Open Source Tools Lab Mini Project

YOUTUBE VIDEOS CRAWLER

Aim :

To build a tool using Python that will crawl over the videos of Youtube.

Problem Defination :

In this world surrounded by social media, we are stuck inside the world wide web. Youtube, one of the largest social media platform that has provided an opportunity for everyone to show their talent and earn for living.

The aim of this project is to display the recent 20 videos of a youtube channel by passing a link of that channel to this tool.

As the link is passed and crawl button is pressed, the tool will go to the desired link and crawl over the information of videos and store it in a CSV (Comma Separated Values) file. This file will be directly opened in Microsoft Excel containing the title of video, date of upload, views on the particular video, duration of the video and the link to that video.

Code for crawler

```
from tkinter import *
from bs4 import BeautifulSoup
import requests
import csv
import os
import sys

root = Tk()
root["bg"] = "black"
root.title('Youtube Videos Crawler')

def crawler(site_url):
    source = requests.get(site_url).text

    soup = BeautifulSoup(source, "lxml")

    csv_file = open('cms_scrape.csv', 'w')
    csv_writer = csv.writer(csv_file)
    csv_writer.writerow(['Title', 'Duration', 'Views', 'Upload Date', 'Link'])
    t_l = []
    d_l = []
    v_l = []
    ud_l = []
    l_l = []
    for ytd in soup.find_all('li', class_="channels-content-item yt-shelf-grid-item"):

        head = ytd.find('h3', class_= "yt-lockup-title ").a.text
        title = head
        t_l.append(title)

        duration = ytd.find('h3', class_= "yt-lockup-title ").span.text
        dur = duration[13:-1]
        d_l.append(dur)

        li = []
        li = ytd.find_all('ul', class_="yt-lockup-meta-info")[0].text.split('ws')
        view = li[0]+'ws'
        views = view
        v_l.append(views)

        date = li[1]
        up_date = date
        ud_l.append(up_date)

        vid_id = ytd.find('div', class_="yt-lockup clearfix yt-lockup-video yt-lockup-grid vve-check")['data-
context-item-id']
        vid_link = f'https://youtube.com/watch?v={vid_id}'
        link = vid_link
        l_l.append(link)

        try:
            csv_writer.writerow([title, dur, views, up_date, link])
        except Exception:
            continue
```

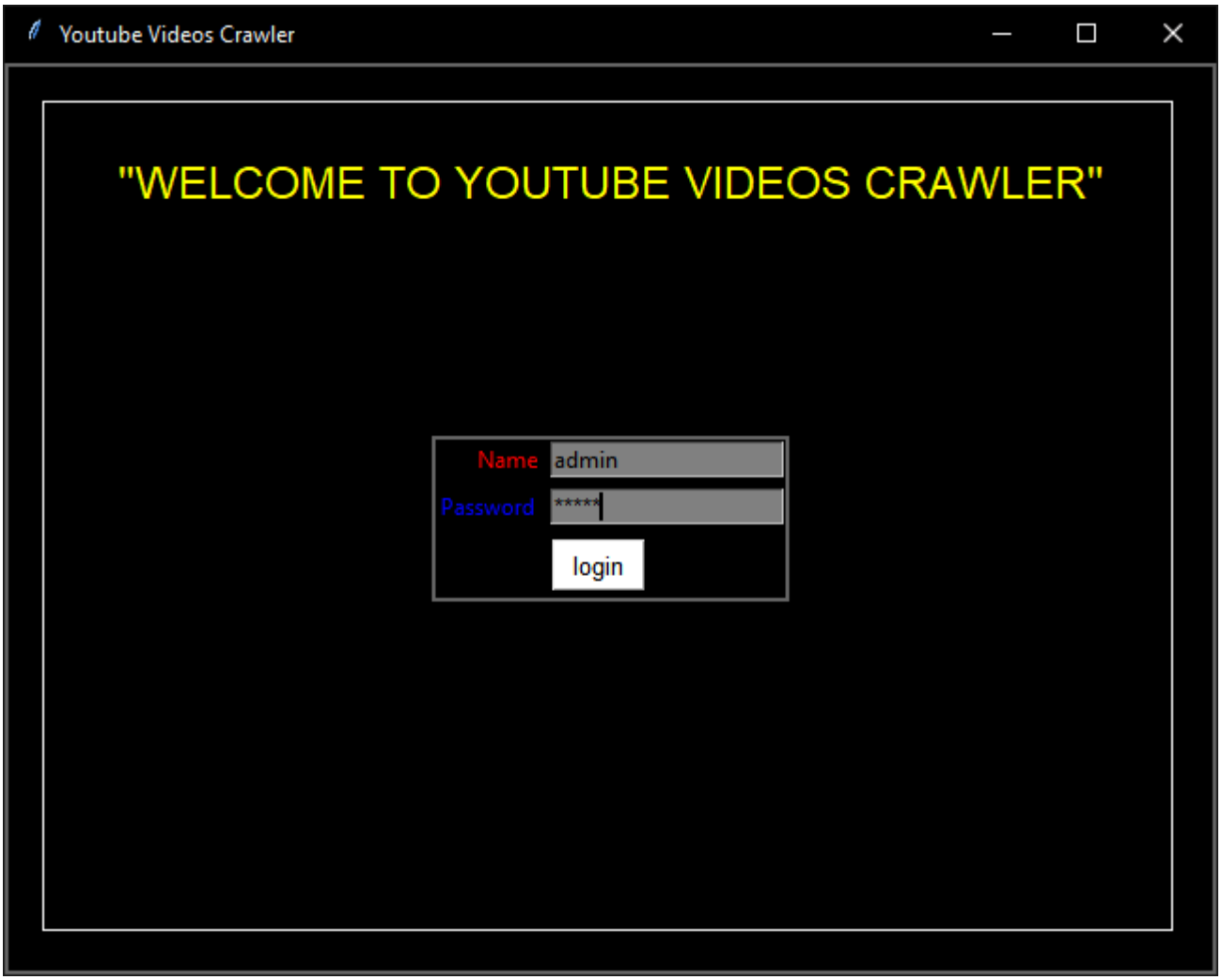
```
    csv_file.close()
    return 1

def actperformed(event):

    def call_crawler(event):
        get_link = enter_link_entry.get()
        flag = crawler(get_link)
        if flag is 1:
            nw.destroy()

            def open_file(event):
                x = sys.argv[0].split('/')
                pth = ''
                for i in x[:-1]:
                    pth = pth + i + '/'
                path = pth + 'cms_scrape.csv'
                os.startfile(path)
                root.destroy()

            root = Tk()
            root["bg"] = "black"
            root.title('Youtube Videos Crawler')
            root.geometry('320x240')

            c1 = Canvas(root, width=320, height=240, bg='black')
            button = Button(c1, text='Open File in Excel', width=18, height=3, bg='green', font=("Arial", 15))
            c1.pack()
            button.bind('<Button-1>', open_file)
            button.place(relx=0.2, rely=0.25)

            root.mainloop()

    get_name = name_entry.get()
    get_pass = name_entry.get()

    if(get_name == 'admin' and get_pass == 'admin'):

        root.destroy()
        nw = Tk()
        nw.title('Youtube Videos Crawler')
        nw['bg'] = 'black'
        nw.geometry("320x240")

        nc2 = Canvas(nw, width=320, height=240, bg='black')
        nc = Canvas(nc2, width=320, height=240, bg='black')

        new_label1 = Label(nc2, text='login successful', anchor='n', fg='green', bg='black', width=20)
        new_label1.config(font=("Courier", 17))

        enter_link = Label(nc, text='Enter the URL:', fg='white', bg='black', width=12)
        enter_link_entry = Entry(nc, bg='grey', width=20)
        crawl = Button(nc, bg='grey', activebackground='green', fg='black', text='crawl', width=6, height=1)

        new_label1.place(relx=0.5, rely=0.1, anchor='n')
```

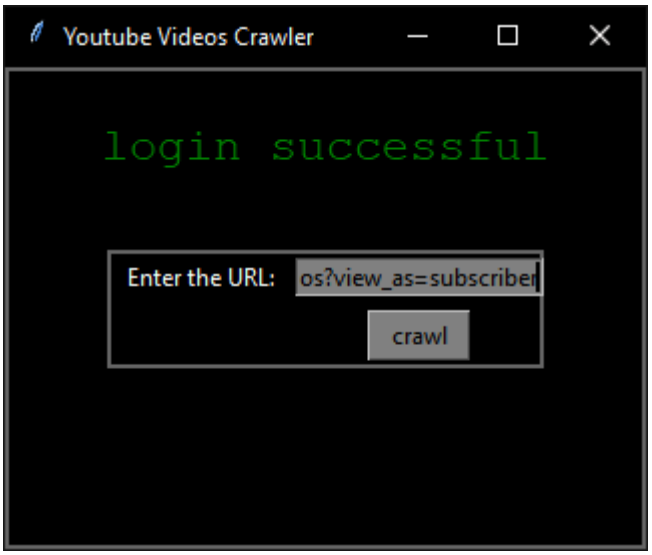
```
        enter_link.grid(row=2, column=1, padx=2, pady=3)
        enter_link_entry.grid(row=2, column=2, pady=3)
        crawl.bind('<Button-1>', call_crawler)
        crawl.grid(row=3, column=2, padx=1, pady=3)
        nc2.pack()
        nc.place(relx=0.5, rely=0.5, anchor=CENTER)

c3 = Canvas(root, width=640, height=480 ,bg='black')
c3.pack()
c2 = Canvas(c3, width=640, height=480, bg='black')
c2.pack()
c1 = Canvas(c2, width=640, height=480, bg='black')
c1.pack()

title = Label(c2, text='"WELCOME TO YOUTUBE VIDEOS CRAWLER"', anchor='n', fg='yellow', bg='black',
width=40)
title.config(font=("Arial", 18))
name = Label(c1, text='Name', bg='black', fg='red')
password = Label(c1, text='Password', bg='black', fg='blue')
name_entry = Entry(c1, bg='grey')
pass_entry = Entry(c1, bg='grey', show='*')
button = Button(c1, bg='white', activebackground='green', fg='black', text='login', width=5, height=1,
font=("Arial", 10))
border = c2.create_rectangle(20, 20, 620, 460, fill='black', outline='white')

title.place(relx=0.5, rely=0.1, anchor='n')
name.grid(row=2, column=2, sticky=E, padx=1, pady=2)
password.grid(row=3, column=2, sticky=E, padx=2, pady=2)
name_entry.grid(row=2, column=3, padx=3, pady=2)
pass_entry.grid(row=3, column=3, padx=3, pady=2)
button.grid(row=4, column=3, sticky=W, padx=4, pady=5)
button.bind("<Button-1>", actperformed)
c1.place(relx=0.5, rely=0.5, anchor=CENTER)

root.mainloop()
```

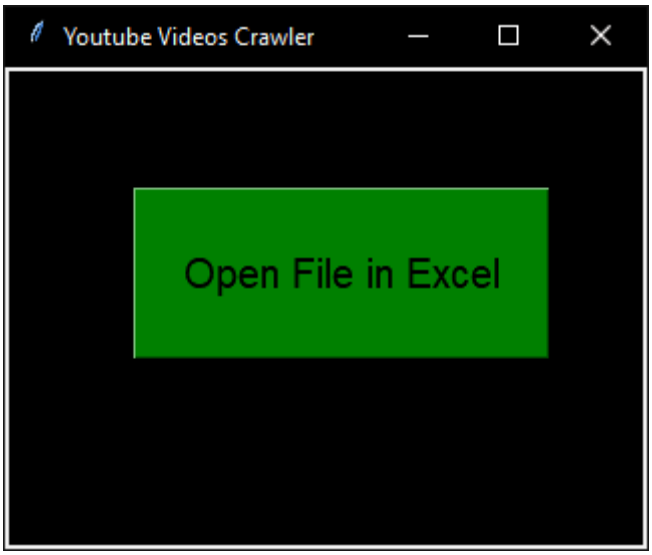
Code to convert .py to .exe

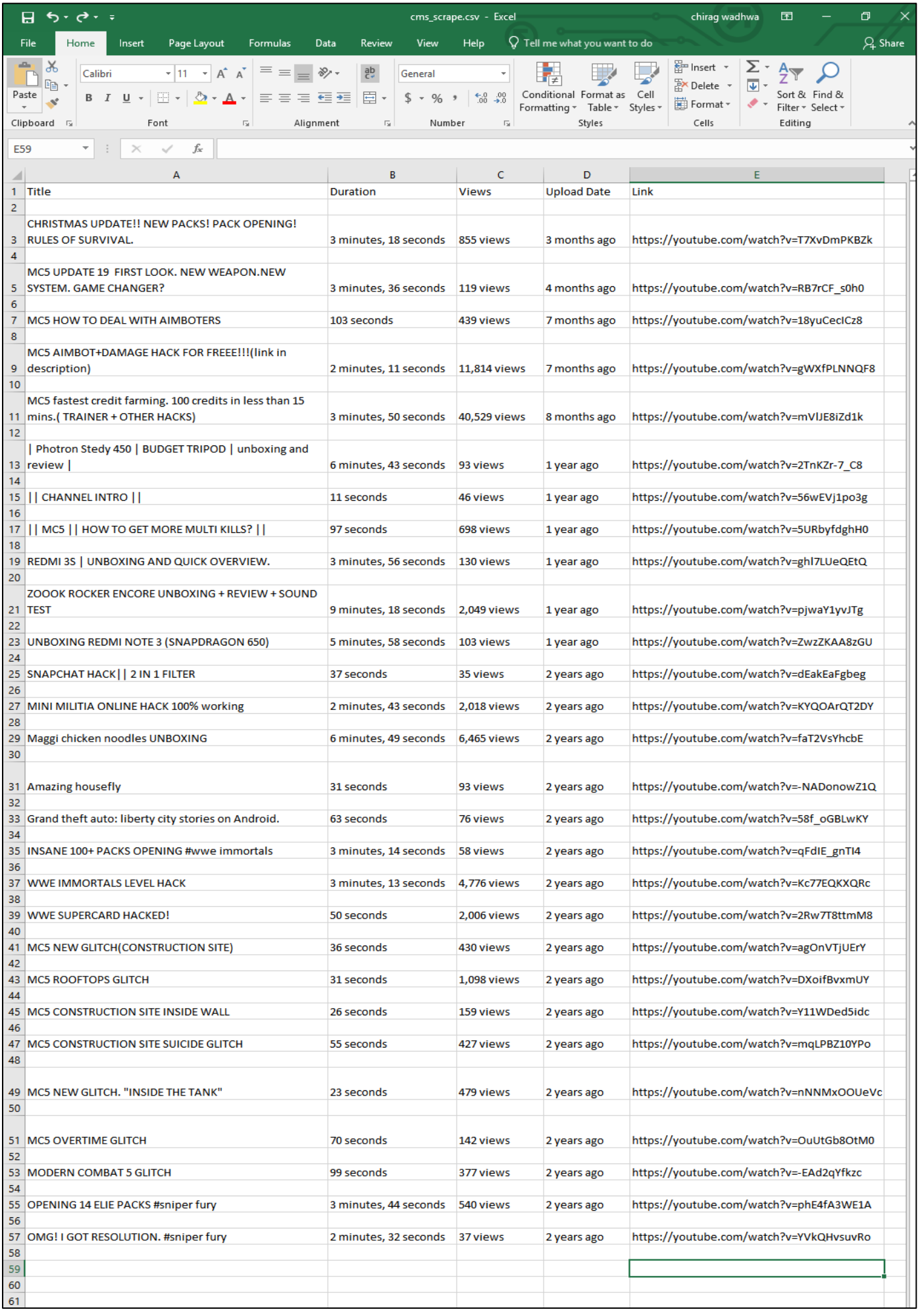
```
import sys
from cx_Freeze import setup, Executable
import matplotlib

base = None
if sys.platform == "win32":
   base = 'WIN32GUI'

setup(
   name = 'Youtube Videos Crawler',
   version = '1.0',
   options = {'build_exe':{'packages':['bs4', 'requests', 'matplotlib']}},
   executables = [Executable('login.py', base = base)])
```

Screenshots









	A	B	C	D	E
1	Title	Duration	Views	Upload Date	Link
2					
3	CHRISTMAS UPDATE!! NEW PACKS! PACK OPENING! RULES OF SURVIVAL.	3 minutes, 18 seconds	855 views	3 months ago	https://youtube.com/watch?v=T7XvDmPKBZk
4					
5	MC5 UPDATE 19 FIRST LOOK. NEW WEAPON.NEW SYSTEM. GAME CHANGER?	3 minutes, 36 seconds	119 views	4 months ago	https://youtube.com/watch?v=RB7rCF_s0h0
6					
7	MC5 HOW TO DEAL WITH AIMBOTERS	103 seconds	439 views	7 months ago	https://youtube.com/watch?v=18yuCecICz8
8					
9	MC5 AIMBOT+DAMAGE HACK FOR FREEE!!!(link in description)	2 minutes, 11 seconds	11,814 views	7 months ago	https://youtube.com/watch?v=gWXfPLNNQF8
10					
11	MC5 fastest credit farming. 100 credits in less than 15 mins.(TRAINER + OTHER HACKS)	3 minutes, 50 seconds	40,529 views	8 months ago	https://youtube.com/watch?v=mVlJE8iZd1k
12					
13	| Photron Stedy 450 | BUDGET TRIPOD | unboxing and review |	6 minutes, 43 seconds	93 views	1 year ago	https://youtube.com/watch?v=2TnKZr-7_C8
14					
15	|| CHANNEL INTRO ||	11 seconds	46 views	1 year ago	https://youtube.com/watch?v=56wEVj1po3g
16					
17	|| MC5 || HOW TO GET MORE MULTI KILLS? ||	97 seconds	698 views	1 year ago	https://youtube.com/watch?v=5URbyfdghH0
18					
19	REDMI 3S | UNBOXING AND QUICK OVERVIEW.	3 minutes, 56 seconds	130 views	1 year ago	https://youtube.com/watch?v=ghI7LUeQEtQ
20					
21	ZOOOK ROCKER ENCORE UNBOXING + REVIEW + SOUND TEST	9 minutes, 18 seconds	2,049 views	1 year ago	https://youtube.com/watch?v=pjwaY1yvJTg
22					
23	UNBOXING REDMI NOTE 3 (SNAPDRAGON 650)	5 minutes, 58 seconds	103 views	1 year ago	https://youtube.com/watch?v=ZwzZKAA8zGU
24					
25	SNAPCHAT HACK|| 2 IN 1 FILTER	37 seconds	35 views	2 years ago	https://youtube.com/watch?v=dEakEaFgbeg
26					
27	MINI MILITIA ONLINE HACK 100% working	2 minutes, 43 seconds	2,018 views	2 years ago	https://youtube.com/watch?v=KYQOArQT2DY
28					
29	Maggi chicken noodles UNBOXING	6 minutes, 49 seconds	6,465 views	2 years ago	https://youtube.com/watch?v=faT2VsYhcbE
30					
31	Amazing housefly	31 seconds	93 views	2 years ago	https://youtube.com/watch?v=-NADonowZ1Q
32					
33	Grand theft auto: liberty city stories on Android.	63 seconds	76 views	2 years ago	https://youtube.com/watch?v=58f_oGBLwKY
34					
35	INSANE 100+ PACKS OPENING #wwe immortals	3 minutes, 14 seconds	58 views	2 years ago	https://youtube.com/watch?v=qFdIE_gnTI4
36					
37	WWE IMMORTALS LEVEL HACK	3 minutes, 13 seconds	4,776 views	2 years ago	https://youtube.com/watch?v=Kc77EQKXQRc
38					
39	WWE SUPERCARD HACKED!	50 seconds	2,006 views	2 years ago	https://youtube.com/watch?v=2Rw7T8ttmM8
40					
41	MC5 NEW GLITCH(CONSTRUCTION SITE)	36 seconds	430 views	2 years ago	https://youtube.com/watch?v=agOnVTjUErY
42					
43	MC5 ROOFTOPS GLITCH	31 seconds	1,098 views	2 years ago	https://youtube.com/watch?v=DXoifBvxmUY
44					
45	MC5 CONSTRUCTION SITE INSIDE WALL	26 seconds	159 views	2 years ago	https://youtube.com/watch?v=Y11WDed5idc
46					
47	MC5 CONSTRUCTION SITE SUICIDE GLITCH	55 seconds	427 views	2 years ago	https://youtube.com/watch?v=mqLPBZ10YPo
48					
49	MC5 NEW GLITCH. "INSIDE THE TANK"	23 seconds	479 views	2 years ago	https://youtube.com/watch?v=nNNMxOOUeVc
50					
51	MC5 OVERTIME GLITCH	70 seconds	142 views	2 years ago	https://youtube.com/watch?v=OuUtGb8OtM0
52					
53	MODERN COMBAT 5 GLITCH	99 seconds	377 views	2 years ago	https://youtube.com/watch?v=-EAd2qYfkzc
54					
55	OPENING 14 ELIE PACKS #sniper fury	3 minutes, 44 seconds	540 views	2 years ago	https://youtube.com/watch?v=phE4fA3WE1A
56					
57	OMG! I GOT RESOLUTION. #sniper fury	2 minutes, 32 seconds	37 views	2 years ago	https://youtube.com/watch?v=YVkQHvsuvRo
58					
59					
60					
61					