Psychoinformatics - Week 3 (Exercises) by 徐舒庭 (b11705018@ntu.edu.tw) 1 Analyze what videos go viral? (8 points) Please use YouTube APIs to carry out a data-driven or hypothesis-driven microstudy about the characteristics of viral videos. You need to present, here in this notebook, AT LEAST two statistical figures or tables as supporting evidence for your arguments. Each of these figures/tables deserves 4 points. Get Data from apiclient.discovery import build API_KEY = "AIzaSyBrKXvByiWJ_jN_lNzkm46oh03RB0V4gss" In []: import re def iso8601_duration_to_seconds(duration): # Regular expression to extract components (e.g., PT15M31S -> 15, 31) pattern = $r'P(?:T(?:(\d+)H)?(?:(\d+)M)?(?:(\d+)S)?)?'$ matches = re.match(pattern, duration) if matches: hours, minutes, seconds = [int(x) if x else 0 for x in matches.groups()] total_seconds = hours * 3600 + minutes * 60 + seconds return total_seconds else: return None def get_trending_videos(max_results=50): youtube = build('youtube', 'v3', developerKey=API_KEY) trending_response = youtube.videos().list(part='snippet, statistics, contentDetails', chart='mostPopular', regionCode='TW', maxResults=max_results).execute() title = [] channelId = []channelTitle = []categoryId = [] videoId = []viewCount = [] likeCount = [] commentCount = [] favoriteCount = [] tags = []duration = []for video in trending_response.get("items", []): title.append(video['snippet']['title']) videoId.append(video['id']) channelId.append(video['snippet']['channelId']) channelTitle.append(video['snippet']['channelTitle']) categoryId.append(video['snippet']['categoryId']) viewCount.append(int(video['statistics']['viewCount'])) if 'likeCount' in video['statistics'].keys(): likeCount.append(int(video['statistics']['likeCount'])) else: likeCount.append(0) commentCount.append(int(video['statistics']['commentCount'])) favoriteCount.append(video['statistics']['favoriteCount']) if 'tags' in video['snippet'].keys(): tags.append(video['snippet']['tags']) tags.append([]) duration.append(iso8601_duration_to_seconds(video['contentDetails']['duration'])) youtube_dict = { 'tags': tags, 'channelId': channelId, 'channelTitle': channelTitle, 'categoryId': categoryId, 'title': title, 'videoId': videoId, 'viewCount': viewCount, 'likeCount': likeCount, 'commentCount': commentCount, 'favoriteCount': favoriteCount, 'duration': duration return youtube_dict In []: **import** pandas **as** pd trend = get_trending_videos() df = pd.DataFrame(data=trend) Out[]: channelld channelTitle categoryId title videold viewCount likeCount commentCount favoriteCount duration tags 【中華隊滑輪溜冰】這就是對勝利的渴望!男子 愛爾達體育家族 0 UCCQvP4hsRW9emj0meGk15jg 17 GMOBUk4P_fg 1710346 24253 2382 399 0 接力隊不放棄衝線 0.01秒之差擊敗南韓摘金! ELTA Sports UCWAgfxHuoVWm4rmqYwoKvlw 杭州亞運 韓國 VS 台灣【看球閒聊直播】 [中華職棒, 大聯盟, mlb] 台南Josh 17 401013 2144 53 M9RihxdqHa4 0 10826 [Muse木棉花,新番,葬送的芙莉 葬送的芙莉蓮 首播100分鐘 第01-04話 | Muse木 13354 1054 2 蓮, 葬送のフリーレン, 芙莉蓮, 種 UCgdwtyqBunlRb-i-7PnCssQ Muse木棉花-TW 1H2cyhWYXrE 435848 0 6150 棉花 動畫 線上看 崎敦美,費倫,... [這群人, TGOP, 搞笑影片, This, 這群人 TGOP | 酒後失態的經典語錄 Classic UC6FcYHEm7SO1jpu5TKjNXEA 這群人TGOP 24 ZOI_fsLIAC8 444819 10758 422 0 920 Group, Of, People, 茵聲,... Quotations for Dum... [HYBE, HYBE LABELS, 하이브, 정국 (Jung Kook) '3D (feat. Jack Harlow)' mHNCM-10 2980348 203724 UC3IZKseVpdzPSBaWxBxundA **HYBE LABELS** 23542301 0 231 하이브레이블즈] YALSA [斗罗大陆, cos, 剧情, 唐三, 小舞, 【鬥羅大陸】唐三偶遇美女#cos #反轉 #鬥羅大 UC6ik7KPV5oL914e4LkpQMpA 梓俊与唐三 22 TLW9IK3NCFc 350351 11027 32 0 26 悬疑, 反转, 套路] 陸#shorts [年代,年代新聞,年代向錢看,陳凝 《#郭璽 指控 #馬文君 賣國! #阿扁當年想買8艘 6 觀,台灣,中國,美國,潛艦,海鯤號, 25 1466 UCBuHkb1AS_yRQ71meFNQ3VQ 年代向錢看 TSPp5XfDUtw 189702 0 0 2991 潛艦!》【年代向錢看】2023.10.0... 和平... [这不科学啊, 科学实验, 手工, 阿基 原来月饼會里的咸蛋黄并不都是从咸鸭蛋里抠出 这不科学啊 NnD9hbobOqQ UC3w1_57LtKg5ml1PHmPc7Eg 990782 44634 378 0 60 米吴,科学玩乐基地,科学, 来的!#fyp#fyptiktok#viral... Science, s... 【杭州亞運】沒到終點就在嗨!競速溜冰慘遭台 8 UCYkldEK001GxR884OZMFnRw 鏡週刊 j9TdOlo5xn0 553 [鏡週刊, 鏡傳媒, 新聞] 灣黃玉霖「絕頂左腿」大逆轉 南韓隊選手「樂 109987 120 0 70 極生悲」... 在英國開箱我在英國爆買的英國小物 [] UCgDQKFV2rMNzTE8Y3rHMVbQ 黃大謙 23 10657 549 0 KVqxvrEucVA 365060 953 [曼曼是老闆,美女老闆,嚴曼曼,美 22 故意叫老闆老婆,看她什麼反應#shorts QDcQGFRbx8U 女總裁,默默是我老闆,默默老闆, UCJ_dJtVyM-5PO0-3grBGtpw 曼曼是老板 321624 5965 114 0 48 霸氣老闆, 雨... [news, breaking news, live news, Baby seals cut loose from fishing net | ABC **ABC News** fQRI7FErNBI UCBi2mrWuNuyYy4gbM6fU18Q 25 6608491 155300 1552 0 60 daily news, w... News 窮小子送外賣被保安辱駡,沒想到窮小子是首富 [短劇, 男頻, 爽文, 都市, 劇情, 陸 UCyWIMLKD6XtCHpb_umtNCKg 王牌劇TV 繼承人,首富老爸跪求他回家繼承家產,保安當 L6V3mnq5CP4 143998 869 136 0 9723 12 劇, 女神, 美女, 主播, 贅婿, 窮小子, 場就失業... [别叫我狗哥,美食,搞笑,厨艺,红 烧排骨, 桓桓, 李牧桓, 探店, 生活, 别叫我狗哥 24 出差到成都, 瞧瞧我遇见了谁?【别叫我狗哥】 8351 56 59 UCSUdanhV0WKHiPyfAxHix9w 6gMwUrmgG4Y 271459 0 日常,... [巴黎, 酪梨女孩, 酪梨女孩巴黎, 小 【小吳】卡真的刷爆了... 🔥 『酪梨女孩in巴黎 吳, 小吳巴黎, 酪梨女孩刷卡, 刷卡, 見習網美小吳 iADGnJ_yGll UCAfcy122TZHqDQMMAwfbvBQ VLOG▮▮』價差超大!每間都買是什麼感覺...... 小吳刷卡,... [#camping, #survival, #bushcraft NOW you know how to do it! #camping 19 WEudznbEz4 22271978 485401 8749 UCRR3vpGoVKF3q_vZI7mCfjQ Marusya Outdoors 0 18 #outdoors] [健康2.0,健康,健康兩點靈,健康 譚敦慈 香蕉保存秘訣!不會變黑 可以慢慢吃【 20,健康2點靈,健康二點靈,鄭凱 UCAqzHNzLTC6Vm5UYE7xw_2A 健康2.0 dhsHa13m38M 137762 3200 50 59 挖健康】 @tvbshealth20 #s... [搞笑段子, 萌娃, 嘟嘟小可耐, 生活 嘟嘟把姊姊寫好的作業全撕了,還找爸爸來撐 22 腰,沒想到被姊姊暴打!#嘟嘟小可耐#萌娃日常 17 記錄, 萌娃神助攻, 專治不開心, 日 UCIpD8v1Qt7-DbSFUumF-ghg 嘟嘟小可耐 Kq6GPhi6n9g 198015 5847 0 31 常生活, 寶媽... ACS 冬季例行賽W3D2 x GCS 2023 夏季例行賽 18 UCeMjhoCCvujpObnt6yeZeNg Garena 傳說對決 zY5ki96html 274760 1421 110 0 36001 | W7D3- 2023/10... [三度, 忍笑, 海龜湯, 靈異, 笑話, 梗 【忍笑大挑戰51集】三度 VS 放火 三度C UCIIv2zQgmulRhEa1yCGLhpg 245291 11645 435 ZwcP5niZ7a4 0 941 圖, 梗文, 迷因, 搞笑, 爆笑, 挑戰] ft.@louislee0602 【DinTer】2023年BYG戰隊Q&A | 特哥經濟快撐 [Dinter, 丁特, IoI, 英雄聯盟, 特哥, UC0SkNQXPJ60hKEFubOz0fDA -xbGayZDaO4 9579 481 Dinter 1112 打野, jungle, BYG, 戰... 不住了?戰隊一年血虧500萬!有機會... 生下來就是雌雄同體?!擁有150世的記憶,用 UC2tQpW0dPiyWPebwBSksJ_g 脑洞乌托邦 意念控制物體,火遍全球的通靈"女巫",為何一 13379 21 MswMYjANxbs 539356 764 0 1576 夜消… 22 UCGPlcnciT5KXLgVjLumGQgQ 微疼 【微鬼畫】童年陰影|天橋下看到他就快閃 -gUK8UdzfGw 237025 6790 658 514 [短剧, 恋爱, 霸总, 宝妈, 纯爱, 豪 [ENG SUB]【完整版】《误嫁后,顾先生宠妻太 心动短剧社 23 门,甜宠,玛丽苏,土味剧,剧情,爱 UCgeTlp-3-qe_pxiUVSif6hQ 甚》和我的相亲对象闪婚了,什么?我的相亲 pUH1s6QPywM 1005280 8709 417 0 10511 Heartbeat [방탄소년단, BTS, BANGTAN, 알 정국 (Jung Kook) '3D (feat. Jack Harlow)' ${\sf UCLkAepWjdylmXSltofFvsYQ}$ **BANGTANTV** 10 Z3x9i7njCCo 7589036 1129773 34461 0 262 엠, RM, 슈가, SUGA, 제이홉, j... [科技, 3C, 開箱, 邦尼幫你, 評測, 「邦尼評測」必看!iPhone 15 最終到底推不 邦尼幫你 3C 科技 -UC2Zn5hlv7_hPsLlh_xwcH6g 25 28 45xzxF1fz88 162210 2611 538 0 1300 PTT, 科技評測, ptt, 2023, i... 推?7 大 關鍵優缺點 實測!iPhone... 開箱評測 [Joeman, 黃氏兄弟, 這群人, 眾量 級,阿神,老高,木曜,千千,賤葆, UCtSSBoYwD-yEkhvNdH0Zy3w 蔡桃貴 【蔡桃貴】兔寶第一次來我們家玩!(5Y1M22D) KkVW4L-nLuo 105485 2742 77 0 260 八年前我是他的大小姐, 他是我買來的窮小子前 [blackpink, aespa, aespa 1 男友, 八年后他成爲了赫赫有名的總裁, 我卻成 8LoHD31P1OM 27 UCmWYO7Ti6fp0dmCNZMWcPNw 阿星爱追剧 3645 248 4992 inkigayo, blackpink l... [民進黨,國民黨,2024,總統大選 揭!馬文君洩潛艦機密?證據在握? 火爆現場! 民視讚夯 Formosa 郭璽, 黄曙光, 馬文君, 賣國, 尹清 UCS3nuaXXw7nsiSzdJ5HJaoQ OCaRw-NGpZg 50115 1463 398 0 4043 郭璽.秦慧珠正面交鋒!【台灣最前線】2023.... TV Thumbs Up [晓涵哥, 晓涵哥来了, 外星人, ufo, 詹姆斯韋伯望遠鏡,最新照片發現宇宙Bug,宇 UCvoBl4rnVsetDKA_Tdk-jeA 曉涵哥來了 E7L1mBIPu3s 116209 4276 323 0 1143 探秘, 揭秘, 未解之謎, 探索宇宙, 宙大爆炸理論似乎存在問題「曉涵哥來了」 [美妝,都市生活,辦公,美女,黑絲, 對付綠茶就要快準狠?! #徐三柒#搞笑#變裝# **30** cp, 逆襲, 反轉, 閨蜜, 情感, 逆天化 徐三柒 25 UCBKNgbdWjD24nPZBLBoaTOg 4kDeYsYl1fg 155401 10912 0 58 短劇 [94要客訴,民進黨,國民黨,政治, 李正皓最新爆:馬文君粗暴凍結潛艇預算過程! 94要客訴 31 政論節目,台灣,中國,兩岸,蔡英 UC4twJx3tfIldmnC_zsuP45Q 25 牽涉聲紋的潛艇「俥葉」若被洩密:變海底鐵棺 rcOAy39_hm4 130094 3098 441 0 3641 材!名嘴.. 文,美... 滴謙CP才是官配!! 大謙會喜歡我排的英倫浪漫行 [學習, 英語, 英文, 阿滴] sYVO25oh3yw UCeo3JwE3HezUWFdVcehQk9Q 阿滴英文 232404 5507 238 0 962 程嗎? [棒球,中華職棒,中職, CPBL, 統 最強中華隊是哪一屆?盤點亞運奪牌史!業餘擊 33 一獅,中信兄弟,樂天桃猿,富邦悍 UCLEy8come0NWSxsZyw9M8Jg 427 0 Yahoo好棒棒 50839 52 718 HYa7cktxVQI 敗韓國全明星,韓總教練傻眼【Yahoo好棒棒】 將,台鋼.. [豆漿貓,娘娘,宅男,寵物,俊榮,阿 皓, 工程師, 療癒, 可愛, 萌寵, 豆 UCiXuTOccGliVKpcL1qQWMPg 豆漿- SoybeanMilk 【豆鬥榮 EP20】這兩姊弟...到底打夠了沒!? 5P77SOHQFwo 105799 6066 199 208 [三立新聞, 時事, 政治, 選情, 國民 公衛文不認罪又講謊話恐加偽證罪!王郁文明知 黨,民進黨,社會,快訊,最新消息, ${\sf UC9tnW2dpPDNc_aEOnhsRs4g}$ 前進新台灣 有浮報事實 關鍵對話曝庭上堅不認罪 律師黃帝 eNVXMJrr1f4 107458 1508 420 553 穎:瞎.. 前進新... [狈总,格局小了,老总,剧情,短视 频, 日常, 搞笑, 搞笑视频, 搞笑合 UCZH3J2P4RRW_ihkwbSCFcaw 狈总 23 華爾街之狼 #狈总 #shorts ozNBlabW4j8 141704 9514 77 0 61 集, 剧情... [為人處世,生活經驗,情感故事,退 男友請我參加親戚聚餐!他媽媽請了個大人物! 落日溫情 2198 54 休生活, 老年生活, 晚年生活, 子女 UCfcdbpglzm482Aj72lWu3lw 但他媽嫌我窮不讓上桌!吃飯時大人物卻給我敬 237199 4032 37 22 J5o0wFczUYs 0 养老,养老院... [방탄소년단, BTS, BANGTAN, 알 [CHOREOGRAPHY] 정국 (Jung Kook) '3D 10 477431 13637 ${\sf UCLkAepWjdylmXSltofFvsYQ}$ **BANGTANTV XpDEEnZQxNU** 2700933 0 173 엠, RM, 슈가, SUGA, 제이홉, j... (feat. Jack ... [美食,居家美食,情侣美食,情侣, 蝦滑肥牛響鈴捲一口下去滿滿的湯汁,太過癮 26 422694 12275 87 0 55 39 干饭, 韩式美食, 中式美食, 火锅, UCrUL9UWi-0XE0XasUzCMpvw 郭师傅与单大厨 hBOrCit7neE 了!#shorts 焖饭] 中職與美職的完美後援 古林睿煬劉致榮合力守成 愛爾達體育家族 UCCQvP4hsRW9emj0meGk15jg ai7bmF7PjzY 401634 4744 883 40 0 334 **ELTA Sports** | 杭州亞運 豬哥亮初次受訪 私密勁爆提問連發!《豬哥會 41 UCAdqIKWKOCzugzxdqPAh0zA 豬哥會社 24 342397 4429 24 0 40 54g-4Pwmq54 社》@FTV_ZhuGeClub 陸國慶人从众!國內旅遊爆發 | 亮:中好事不斷!倩: [中天新聞, 國際新聞, Cti Talk, 全 604 球大視野, 國際直球對決, 前進戰 全球大視野 25 美淪落日帝國 | 華為遙遙領先!蘋果坦承過熱 _6KdEHGyDns 455257 28456 0 7250 42 UCiwt1aanVMoPYUt_CQYCPQg 略高地,.. 【#... 【中華隊滑輪溜冰】戰術性提早衝刺放!女子接 愛爾達體育家族 17 43 UCCQvP4hsRW9emj0meGk15jg lud2_5nKOHo 430823 6310 307 0 352 **ELTA Sports** 力隊實力海放再拼一金! [养生, 中老年, 健康, 孝順, 真正的 103歲母親小跑看望93歲女兒 #真人真事 #正能 UCf2H_z2AHFDMlh_DkeYJ72Q 三味書屋 22 191345 6693 81 0 孝順, 笑了, 感動, 幸福, 晚年幸福 jvjPZX5vMAc 58 量 #人間溫暖 [Kpop, girl group, 1theK, UCYDmx2Sfpnaxg488yBpZIGg IVE 아이브 'Either Way' MV 14488554 45 starshipTV 10 _Hu4GYtye5U 567338 29206 0 238 Starshiptv, starship... [猫咪, 宠物, 动物, 萌宠, 小朋友, 宝 果果和十八人生第一张证件照get~₩#cat 15 宝,成长,互动,沉浸式,养猫,怀 十八的萌喵日记 navOUrPxti0 116292 5956 46 0 39 UCOxV6XArQpM34NFazf-sJZg #catandbaby #shorts 孕,... [最近紅什麼,達爾,小李,香香,丫 終於出國了?說走就走的旅行3【最近紅什麼】 23 47 丫,壘包,梗圖,戲劇,短劇,搞笑, UCmV89JP1Fx5jTAMYjmpJMQg 最近紅什麼 1LNtDhyznNo 92933 2584 245 0 2263 《Game》 綜藝.. [斗破苍穹, 斗破蒼穹, 動漫, 动漫 紫研成功出关!战斗力达斗皇级别!拥有穿越 动漫喵Meow 6995 198 48 漫畫,漫画,年番,劇集,剧集,番劇, UCZSDmkTnR847RfE799UUjkQ omRe87jkmtE 327593 0 1193 空间能力! | 斗破苍穹年番 Battle Thr... [5-minute crafts, 5mincraft, best Transforming a coin into a goldfish pendant 🐠 UCAhQHwhb7nb14Cn_mlkq2hg Handmade Heroes kqmQmsrm9NQ 4848270 86176 236 0 58 shorts, chal... **Hypothesis** 1: Videos with shorter durations are more likely to go viral. 2: Videos with more likes are more likely to go viral. 3: Videos with more comments are more likely to go viral. Data Analysis import pandas as pd import matplotlib.pyplot as plt import numpy as np # Plot the relationship between video duration and view count plt.figure(figsize=(10, 6)) plt.scatter(trend['duration'], trend['viewCount'], alpha=0.5) plt.title("Relationship between Video Duration and View Count") plt.xlabel("Video Duration (seconds)") plt.ylabel("View Count") # compute the correlation coefficient correlation = np.corrcoef(trend['duration'], trend['viewCount'])[0, 1] print(f"相關係數:{correlation}") plt.show() 相關係數:-0.12805774537087855 Relationship between Video Duration and View Count 1e7 2.0 1.5 View Count 1.0 0.5 0.0 10000 15000 20000 30000 5000 25000 35000 Video Duration (seconds) plt.figure(figsize=(10, 6)) plt.scatter(trend['likeCount'], trend['viewCount'], alpha=0.5) plt.title('Relationship between Like Count and View Count') plt.xlabel('Like Count') plt.ylabel('View Count') # compute the correlation coefficient correlation = np.corrcoef(trend['likeCount'], trend['viewCount'])[0, 1] print(f"相關係數: {correlation}") plt.show() 相關係數:0.7964278155062224 Relationship between Like Count and View Count 1e7 2.0 1.5

View Count 1.0 0.5 0.0 0.5 2.0 2.5 1.0 1.5 3.0 0.0 1e6 Like Count plt.figure(figsize=(10, 6)) plt.scatter(trend['commentCount'], trend['viewCount'], alpha=0.5) plt.title('Relationship between Comment Count and View Count') plt.xlabel('Comment Count') plt.ylabel('View Count') # compute the correlation coefficient correlation = np.corrcoef(trend['commentCount'], trend['viewCount'])[0, 1] print(f"相關係數: {correlation}") plt.show() 相關係數:0.7173996815262069 Relationship between Comment Count and View Count 1e7 2.0

1.5

View Count 1.0

0.5

0.0

Explanation

Hypothesis 1:

Hypothesis 2:

Hypothesis 3:

50000

Hence, we can't say that videos with shorter durations are more likely to go viral.

Hence, we can say that videos with more likes are more likely to go viral.

Hence, we can say that videos with more comments are more likely to go viral.

75000

25000

125000

100000

From the first figure, the correlation coefficient is -0.13, it shows that their is no significent relation between video duration and view count.

Please submit the PDF version of your notebook to NTU COOL before 10/6 (Friday).

From the second figure, the correlation coefficient is 0.8. Since it is bigger than 0.7, it shows that their is significent relation between like count and view count.

From the third figure, the correlation coefficient is 0.72. Since it is bigger than 0.7, it shows that their is significent relation between comment count and view count.

Comment Count

150000

175000

200000