

YouTube video player interface showing a video titled "CISD43 TextMining Twitter WebConf". The video content displays a Jupyter Notebook snippet with Python code for fetching tweets and user locations using the Tweepy library. The code is as follows:

```
In [15]: tweets = tweepy.Cursor(twitter_api.search,
                                q=new_search,
                                lang="en",
                                since=date_since).items(5)

        users_locs = [[tweet.user.screen_name, tweet.user.location] for tweet in tweets]
        users_locs

Out[15]: [['A_Melkishvili', 'Paris, France'],
          ['a_parrington', 'UK'],
          ['grantham_1c', 'Imperial College London'],
          ['changeobserved', ''],
          ['dougbrinkman', '']]
```

Below the code, the video explains how to create a Pandas DataFrame from a list of tweet data. The video player shows a progress bar at 21:46 / 1:05:17. The video has 175 views and was uploaded on May 5, 2021. The right sidebar features recommended videos, including "Step-By-Step Blogger Tutorial For Beginners", "Alberta lawyers share insights from the front lines of the...", "Python OOP Tutorial 1: Classes and Instances", and "Elon Musk talks Twitter, Tesla and how his brain works - live...".

YouTube video player interface showing a video titled "CISD43 TextMining Twitter WebConf". The video content displays a Jupyter Notebook snippet with Python code for fetching tweets and user locations using the Tweepy library. The code is as follows:

```
In [15]: tweets = tweepy.Cursor(twitter_api.search,
                                q=new_search,
                                lang="en",
                                since=date_since).items(5)

        users_locs = [[tweet.user.screen_name, tweet.user.location] for tweet in tweets]
        users_locs

Out[15]: [['A_Melkishvili', 'Paris, France'],
          ['a_parrington', 'UK'],
          ['grantham_1c', 'Imperial College London'],
          ['changeobserved', ''],
          ['dougbrinkman', '']]
```

Below the code, the video explains how to create a Pandas DataFrame from a list of tweet data. The video player shows a progress bar at 28:21 / 1:05:17. The video has 175 views and was uploaded on May 5, 2021. The right sidebar features recommended videos, including "Step-By-Step Blogger Tutorial For Beginners", "Alberta lawyers share insights from the front lines of the...", "Python OOP Tutorial 1: Classes and Instances", and "Elon Musk talks Twitter, Tesla and how his brain works - live...".

CISD43 TextMining Twitter V

youtube.com/watch?v=K65Et4Nvsk

Menu Services Foods Business Youtube Temp Finance Coding

YouTube Search

CISD43_TextMining_Twitter_Example2 - Adobe Acrobat Pro DC

Home Tools CISD43_TextMining_Twitter_Example2.pptx

6/18

4/29/2020 CISD43_TextMining_Twitter_Example2 - Jupyter Notebook

```
In [16]: 1 # Split the words from one tweet into unique elements
2 all_tweets_no_urls[0].lower().split()

Out[16]: ['what',
'a',
'pleasure',
'to',
'be',
'contributing',
'to',
'this',
'wonderful',
'tea',
'magazine',
'my',
'article',
'was',
'about',
'learning to fly']
```

CISD43_TextMining_Twitter_Example2.pptx

Date modified: 4/29/2020 11:19 AM
Size: 1.03 MB
Date created: 4/29/2020 11:19 AM

CISD43 TextMining Twitter WebConf

Unlisted

175 views · May 5, 2021

DISLIKE SHARE DOWNLOAD SAVE

Step-By-Step Step-By-Step Blogger Tutorial

CISD43 TextMining Twitter Lab Session May12

localhost:8888/notebooks/Documents/MALVAREZ/CISD43_TextMining_Twitter_part_1_MALVAREZ

jupyter CISD 43 Text Mining Twitter part 1 MALVAREZ Last Checkpoint: 28 minutes ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Python 3

```
RT @MoHFW_INDIA: #IndiaFightsCorona \n\nUnion Health Minister @drharsvardhan encouraged people to come forward and donate blood at the #Bloo...
RT @ANI: We are facing more problems as Maharashtra govt did not conduct the test of ppl who returned from Hazur Sahib in Nanded. If they h...
RT @SkyNews: Why is Britain now emerging as the country with the highest number of deaths from this pandemic? \n\nFormer Chief Scientific Ad...
RT @SuissePatiente: #cancer #patients particularly high risk group during #Covid19 #coronavirus pandemic, resulting in excess death #Study...
```

To Keep or Remove Retweets

```
In [5]: new_search = search_words + " -filter:retweets"
new_search

Out[5]: '#covid19 -filter:retweets'
```

```
In [6]: tweets = tweepy.Cursor(twitter_api.search,
q=new_search,
lang="en",
since=date_since).items(5)

[tweet.text for tweet in tweets]

Out[6]: ["We're now providing a range of safety equipment to help us all get back to work faster. As well as standard masks f... https://t.co/1QhuxFRASK",
'You cant drag the Worl Boss in a childs play, Wizkid should be given someone like Tommy Lee but not Vybz Kartel, t... https://t.co/BpPkDyNaxK',
'Of the 852 people who had close contact with confirmed Covid-19 cases — after the sixth day of their illness — no n... https://t.co/l8prL3V06C',
'Never forgetnDelay Blocking & lack of testing = Trump's fault\nLack of supplies to protect us from #COVID19 -T rump's... https://t.co/C5PAKVLPHK',
'Matt Hancock Just checking that someone has placed an order with Roche. Hopefully we are not late on this one! htt ps://t.co/hC4FtKfjXk']
```

6:39 / 29:21

Home | Mt. San Antonio College | Assignments: CISD-43-01-4045: | CISD43 Sentimental Analysis Q... | CISD43 Text Mining Twitter (Pro... | Real time Big Data Analytics Per... | CISD43 TextMining Twitter L... |

youtube.com/watch?v=ebshmj...fjl

Menu Services Foods Business Youtube Temp Finance Coding

YouTube Search

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Documents CISD43 Text Mining Twitter part 1 MALVAREZ Last Checkpoint: 32 minutes ago (autosaved)

Python 3

```
def remove_urls(text):
    """Remove URLs found in a text string with nothing
    (i.e. it will remove the URL from the string).

    Parameters
    -----
    text : string
        A text string that you want to parse and remove urls.

    Returns
    -----
    The same text string with url's removed.
    """
    return " ".join(re.sub("([0-9A-Za-z-]+)(?:http|https|ftp|mailto|tel|)://", "", text).split())

In [10]:
all_tweets_no_urls = [remove_urls(tweet) for tweet in all_tweets]
all_tweets_no_urls[1]

Out[11]:
['This is a house church where a service was being held and the CCP raided it coronavirus COVID19'.
'DOMR Security had threaten to use Anti Terrorist power to seize assets of demonstrators in tandem with CCP'.
'realdonaldtrump as well as many others are debating at the moment the culpability of the CCP in regard to'.
'AMCHSung 6 weeks ago the CCP had no problem accusing the US of bringing covid19 into China on purpose if you're'.
'wanted Origin of CoronavirusCOVID19 which killed 250,000 people world-wide/a/CCP #ThankYouCCP #China4JweBidde
a. https://a.co/0mndab9ier']

In [12]:
# Split the words from one tweet into unique elements
```

11:28 / 29:21

CISD43 TextMining Twitter Lab Session May12

Unlisted

72 views • May 12, 2021

Like Dislike Share Download Save

Recently uploaded Watched

[Exclusive] Elon Musk: A future worth nothing awaits about!

8:40 PM 4/20/2022

Home | Mt. San Antonio College | Assignments: CISD-43-01-4045: | CISD43 Sentimental Analysis Q... | CISD43 Text Mining Twitter (Pro... | Real time Big Data Analytics Per... | CISD43 TextMining Twitter L... |

youtube.com/watch?v=ebshmj...fjl

Menu Services Foods Business Youtube Temp Finance Coding

YouTube Search

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Documents CISD43 Text Mining Twitter part 1 MALVAREZ Last Checkpoint: 38 minutes ago (autosaved)

Python 3

```
Out[19]:
['house',
'church',
'service',
'held',
'ccp',
'raided',
'coronavirus',
'covid19']

In [20]:
# Let's redo the count of most common words
all_words_new = list(itertools.chain(*tweets_new))
counts_new = collections.Counter(all_words_new)
counts_new.most_common(15)

Out[20]:
[('ccp', 530),
('covid19', 398),
('china', 245),
('world', 86),
('virus', 80),
('chinese', 80),
('coronavirus', 76),
('amp', 74),
('us', 73),
('people', 59),
('wuhan', 57),
('pandemic', 38),
('communist', 37),
('chinas', 34),
('realdonaldtrump', 33)]
```

17:13 / 29:21

CISD43 TextMining Twitter Lab Session May12

Unlisted

72 views • May 12, 2021

Like Dislike Share Download Save

Recently uploaded Watched

[Exclusive] Elon Musk: A future worth nothing awaits about!

8:40 PM 4/20/2022

