hich was created in the assignment 2 to NoSQL. The converted database contains all the domain related data from the assignmen t 2. The NoSQL database uses Mongodb. In []: # Importing libraries import pymongo from tweepy.streaming import StreamListener from tweepy import OAuthHandler from tweepy import Stream import pandas as pd import numpy as np import tweepy as tw from pandas import DataFrame In [50]: # Reading the CSV file from the previous assignment match = pd.read_csv('Match.csv', index_col=0) match.head() Out[50]: Team2 match_date Season_Year Venue_Name City_Name match_id Team1 Match_SK Royal Kolkata 335987 Challengers 4/18/2008 2008 Chinnaswamy 0 Knight Bangalore Bangalore Riders Stadium Punjab Chennai Cricket Kings XI 335988 4/19/2008 2008 Chandigarh Super Association Punjab Kings Stadium, Mohali Delhi Rajasthan Feroz Shah 4/19/2008 335989 2008 Delhi Daredevils Royals Kotla Wankhede Mumbai 4/20/2008 Challengers 2008 Mumbai 335990 Indians Stadium Bangalore Kolkata Deccan Eden 335991 Knight 4/20/2008 2008 Kolkata Gardens Chargers Riders In [33]: # Connecting to the mongodb database mongodb_host= 'localhost' mongodb_port = '27017' client = MongoClient(mongodb_host + ':' + mongodb_port) db = client['cricket'] print (db) Database(MongoClient(host=['localhost:27017'], document_class=dict, tz_a ware=False, connect=True), 'cricket') In [80]: # Using Twitter API to fetch details consumer_key = 'Ulg0a0EDEy09m90eQ1WjgY8ax' consumer_secret = 'hJWMDqHG6rAtNn5fV6oIgNwaVcdy2yTIY0WXNWoy0KDD8111UZ' access_token = '715063474020159488-TWPJ4GzyfYJ94Iy4ItvIu4c7Qd065nD' access_token_secret = '2mgrQXQ5m6yG1zmZcUBluykXg5LGrgw4KZYeNKPtDCicI' auth = tw.OAuthHandler(consumer_key, consumer_secret) auth.set_access_token(access_token, access_token_secret) api = tw.API(auth, wait_on_rate_limit=True) search_words = "%#cricket" $date_since = "2018-11-16"$ tweets = tw.Cursor(api.search, g=search words, lang="en", since=date_since).items(100) tweets Out[80]: <tweepy.cursor.ItemIterator at 0x2a63fb37f08> In [39]: # Reading the tweets tweets = [tweet.text **for** tweet **in** tweets] tweets Out[39]: ['RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & India\xa0should be p layed t...', 'Wisden has just depressed me by highlighting that Steve Smith scored e leven Test centuries against us in the 2010s... https://t.co/7X5hAeYCNg', 'RT @SunilAh64145529: Howzat #Cricket 🛭 https://t.co/3vnLCuQHNZ', 'RT @timesofindia: #Cricket #TheAshes 🛮 \n\n2005 Ashes was one of the a ll time great cricket series: @RickyPonting \n\nRead: https://t.co/ACofo C...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & India\xa0should be p 'RT @TheCricketWire: @windiescricket @Russell12A []\n\nBest Strike Rates in IPL History \n(30+ Batting Average) \n\n10 Andre Russell 186.41\n20 R "RT @Daniel86criket: Waqas, Apologies for spoiling your fame night budd y but the future of Pakistan isn't looking bright. #Cricket https://t...", 'RT @MarkDuffett: A Pair of Aces - The Caribbean Bromance of 1973 http s://t.co/rNm9yqRIZh #Cricket', 'Missing your #Sport ? No #cricket #football or #tennis - Support me at Redbubble #RBandME: https://t.co/AF3N096GlP... https://t.co/WbXskQ1jt2', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p layed t...', 'RT @daniel86cricket: Which cricket format is the best? #Cricket', 'RT @Cricketracker: Justify your answers. #Cricket https://t.co/yAbK7sm 'RT @Saj_PakPassion: Michael Clarke\xa0"Australia & mp; probably every other team sucked up to India & were too scared to sledge Kohli &am p; other Indi...', 'RT @Sai PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p layed t...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p 'RT @Saj_PakPassion: Michael Clarke\xa0"Australia & mp; probably every other team sucked up to India & amp; were too scared to sledge Kohli & am p; other Indi...', "RT @Kittysvinvault: Another of my favourite finds, this tiny tin conta ining the pocket cricket game 'Owzthat', sometimes known as 'Trench C...", 'RT @Karl_McDermott: Day 15 of groundsman isolation, finished with the roller for the moment, time to cut some grass!! #lords #lovelords #cr...', 'Follow my account i will be posting some thing special for cricket lov ers #cricket #special #fans', 'New Account..." \[\nMust Follow... \[\] \n#Twitter #Trending #FolloMe #foll obackforfolloback #follotrick... https://t.co/DBnWOQBxRB', '@windiescricket @Russell12A 🛮 \n\nBest Strike Rates in IPL History \n(3 0+ Batting Average) \n\n10 Andre Russell 186.41\n20 ... https://t.co/YTuYDs R9sa', 'RT @Woodwindtutor: Some of @brookeccnorwich ladies have been busy the last few days. I Us newbies have a bit of work to do! #StayAtHomeChal...', 'The #CPLT20 Party comes to #CEENTV! \nMondays and Thursday - CPL Rewin d HOWZAT!\n000\nsign Up today and Watch Free for... https://t.co/W9Uvgg82g 4', 'RT @PoppingCreaseSA: I simply cannot believe there are supposed Cricke t fans out there that still, to this day, do not appreciate or even,...', 'RT @PoppingCreaseSA: I simply cannot believe there are supposed Cricke t fans out there that still, to this day, do not appreciate or even,...', 'RT @PayntrCricket: 3 FOR 2 OFFER!\n\nTreat yourself to our new range o f accessories, a perfect addition to your exercise routine\n\n#payntr #p 'RT @capitalkidscric: Day2! Challenge 2! \n\nSomething a little less ti ring today, it's time for cricket to meet tennis! \n\nLongest rally win 'RT @CobrasCricket: (1/3) [It's official! The new members of the World Sports Betting Cape Cobras are ...\n\n#Cricket #Cobras #CobrasCricket h t...', "RT @kanishkaab: Here's my pick of 10 awesome #cricket documentaries yo u can watch online during the #lockdown .There are notable omissions..." 'RT @PoppingCreaseSA: I simply cannot believe there are supposed Cricke t fans out there that still, to this day, do not appreciate or even,...', 'Twenty20 World Cup Cricket 2020: Testing, isolating players on agenda in bid to keep tournament in Australia this y... https://t.co/ov3iCPMbWv', 'A Pair of Aces - The Caribbean Bromance of 1973 https://t.co/rNm9vqRIZ h #Cricket'. 'RT @downthegroundtw: Here's an interesting question for you all ↓\n\nW hat was the best Test innings of the last decade? It has to be an inn 'RT @downthegroundtw: Here's an interesting question for you all \$\n\nW hat was the best Test innings of the last decade? It has to be an inn 'RT @AhmerNajeeb: Another tough choice for cricket fans. Who will you c hoose between the both\n\n1: Inzimam-ul-haq\n2: Muhammad Yousaf\n\n#cric ket', 'Benjamin Andrew Stokes is an English international cricketer and curre nt vice-captain of the England Test team. Sto... https://t.co/AqFgJ6P0q9', 'RT @Saj PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', 'I simply cannot believe there are supposed Cricket fans out there that still, to this day, do not appreciate or eve... https://t.co/iBbFadJgVy', 'RT @KKRiders: Presenting, The Wisden Leading T20 Cricketer of 2019! [\n\n\ Choice @WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenC ri...', "It's not only the batsman that's out. #cricket #twunts #dumbfucks #lon donlockdown\n\nhttps://t.co/qWaoc9Tsi6", 'RT @KKRiders: Presenting, The Wisden Leading T20 Cricketer of 2019! \n\n\ Choice @WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenC "RT @Daniel86criket: Waqas, Apologies for spoiling your fame night budd y but the future of Pakistan isn't looking bright. #Cricket https://t...", 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p laved t...', "RT @KKRiders: Raced and paced his way to the top! []\n\nThe number 1 Te st bowler has also made his way into the elite list of 'Wisden Cricke t...", 'Bowlers love to see those LED lights up for sure ≶□\n\n#KheloAzaadiSe #Love #Cricket #Welovecricket... https://t.co/AlMHIn6ZTr', 'RT @Saj_PakPassion: Howzat #Cricket https://t.co/keoU8UcXfe', 'RT @Daniel86criket: Sad State of affairs. #Cricket https://t.co/8x8coT "RT @kanishkaab: Here's my pick of 10 awesome #cricket documentaries yo u can watch online during the #lockdown .There are notable omissions..." 'RT @Daniel86criket: One has to go. #Cricket https://t.co/QPOr5sgvDV' 'Our latest news articles is up! We explore a world where the COVID-19 outbreak didn't happen #cricket #cricketcoach... https://t.co/aaXMCrWox1' '@durfiteymun Great finish to a fantastic game of #Cricket', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p layed t...', 'RT @chand2579: So here it is @dizzy259 in a chat with me on his career highlights, @VVSLaxman281, injuries, @ImIshant, @sachin_rt and a lo...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', 'RT @guernseycricket: Back to cricket - A simple drill involving just a tennis ball for any close in fielders or wicket keepers, @lukewright...', 'RT @FarziCricketer: Pain in life was easy to deal when we were togethe r. #cricket', "RT @Daniel86criket: When your team is doing worse than Oman and Hong K ong so you pull out Babar's stats to cover it up. #Cricket https://t...", 'RT @tumbowl: @bebutch @abc730 @JoshFrydenberg Idealogical constraints from little johnny? What does this mean? Did little johnny have ideal...', 'Welcome to @cricclubonline ! Your number one stop for cricket news, co aching tips, updates and analysis! Stay tuned... https://t.co/kHRh9FLE4J', 'DSC Shoc Cricket Bat\n\nwats app for more details 0333 3426216\n\nhttp s://t.co/jxGS1MkH00\n\n#fearless #karachirain... https://t.co/ogeArhoYRj', "RT @KKRiders: Our guest for this week's #KnightsUnplugged is Eoin Morg an! [\n\n[Saturday, 11 April\n[9 PM\n[https://t.co/x0Gp9QfSsf\n\nRais e y...", 'RT @Daniel86criket: SL v IND series since 2015\n\nT20I - L L L L L\n ODI - L L $\ln - L = 12 \ln L = 12 \ln c$ \nBoltzm...', 'A million times better all round cricketer than that other bloke \n\n# ViratKohli \n#BenStokes \n#Cricket \n#number1 https://t.co/PaqF1h3JFt', "RT @Daniel86criket: Waqas, Apologies for spoiling your fame night budd v but the future of Pakistan isn't looking bright. #Cricket https://t...", 'RT @CobrasCricket: (2/3) 🛘 It's official! The new members of the World Sports Betting Cape Cobras are ...\n\n#Cricket #Cobras #CobrasCricket h 'RT @daniel86cricket: Which cricket format is the best? #Cricket', 'ShahiD Afridi Baseball Six @SAfridiOfficial #COVID\n#lockdownextension #IPL2020\n#Cricketnews #IPL #IPLT20... https://t.co/0VL03JFp3Z', 'Another great day of hard work from our youngster @GeorgeCharlesP1 som e great drills here for our fellow juniors. K... https://t.co/6vgCT1Jqwf', 'RT @Daniel86criket: One has to go. #Cricket https://t.co/QPOr5sgvDV', 'Geeky personalised gifts for sports fans. Any RTs much appreciated. Ta ke care.\n\nhttps://t.co/R08It6JJgC #ShopIndie... https://t.co/rNzuIkSrk 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p laved t...', 'RT @AkberAJaffri: Interviewing @faf1307 & @amlahash in #CapeTown l ast August, ahead of WorldXI tour of Pakistan, one of the highlights of 'RT @KKRiders: Catch him yet? Shake your head to see the knight! [\n\n# KorboLorboJeetbo #KKR #Cricket #ShakeWho #QuarantineActivities https:// 'RT @AkberAJaffri: PCB releases NOC policy for its players, all central ly contracted players may seek clearance for up to a maximum of four...', '@just1doctorwala @scroll_in PLEASE give us a break !\n\nCan't these pl ayers STOP peddling their wares at least during... https://t.co/B86Eu2qdB 'RT @AkberAJaffri: Interviewing @faf1307 & @amlahash in #CapeTown l ast August, ahead of WorldXI tour of Pakistan, one of the highlights of 'RT @nnis_sports: DHONI'S BIGGEST ASSET WAS HIS CALM UNDER PRESSURE\n@m sdhoni @PaddyUpton1 @ICHOfficial @BCCI @imVkohli #MSDhoni #Cricket ht t…', "RT @KKRiders: Raced and paced his way to the top! []\n\nThe number 1 Te st bowler has also made his way into the elite list of 'Wisden Cricke 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'Back to cricket - A simple drill involving just a tennis ball for any close in fielders or wicket keepers,... https://t.co/nSR3ZaVZZh', 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', "RT @KKRiders: Raced and paced his way to the top! []\n\nThe number 1 Te st bowler has also made his way into the elite list of 'Wisden Cricke t...", 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'RT @briancjeeves: #YellowSport Essex ace Harmer named as one of Wisde n's elite @EssexCricket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', 'RT @intouchcricket: @intouchcricket and @EpsomC_Cricket are producing a series of simple, informative and fun coaching drills. This straigh...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xaOshould be p layed t...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', '#Cricket\n\nThough he picked Gautam Gambhir as the best captain he has played under in IPL, Rajasthan Royals batsman... https://t.co/lPBLdORdXb', '#Cricket\n\nEven though Australian super batsman David Warner wished f or a challenge from Indian captain Virat Kohli... https://t.co/boAsQEEeo 'No suprise to see that Cricketers, Rugby Players and im sure other spo rtsmen and women have taken a Paycut. Footbal... https://t.co/dudwZEXDH1', 'RT @ghulamabbasshah: #Pakistan's legendary former pacer Shoaib Akhtar (@shoaib100mph) proposed a three-match Indo-Pak #Cricket series to r...', 'RT @Saj_PakPassion: Shoaib Akhtar\xa0has suggested that whenever possi ble a three-match ODI series between Pakistan & amp; India\xa0should be p layed t...', 'RT @SonySportsIndia: A musical one from @joybhattacharj to start thing s off for the day! <code>[][]</code> \n\n#SonySportsOuiz #SonySports #Ouiz #Music #Cr 'RT @SonySportsIndia: #Bingo if you get this []!\n\n#SonySportsQuiz #Son ySports #Sports #Quiz #ContestAlert #Cricket #Contest #India https://t. In [40]: # Splitting the tweets user_list = [tweet.split(':')[0] for tweet in tweets] tweets_list = [tweet.split(':')[-1] **for** tweet **in** tweets] tweets_list Out[40]: [' Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xa0should be played t...', '//t.co/7X5hAeYCNg', '//t.co/3vnLCuQHNZ', '//t.co/ACofoC...', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', @windiescricket @Russell12A []\n\nBest Strike Rates in IPL History \n (30+ Batting Average) \n\n10 Andre Russell 186.41\n20 Ris...', '//t...', '//t.co/rNm9yqRIZh #Cricket', '//t.co/WbXskQ1jt2', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xa0should be played t...', Which cricket format is the best? #Cricket', '//t.co/yAbK7smh0r', Michael Clarke\xa0"Australia & probably every other team sucked u p to India & amp; were too scared to sledge Kohli & amp; other Indi...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xaOshould be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xaOshould be played t...', Michael Clarke\xa0"Australia & mp; probably every other team sucked u p to India & amp; were too scared to sledge Kohli & amp; other Indi...', " Another of my favourite finds, this tiny tin containing the pocket cr icket game 'Owzthat', sometimes known as 'Trench C...", Day 15 of groundsman isolation, finished with the roller for the mome nt, time to cut some grass!! #lords #lovelords #cr...', 'Follow my account i will be posting some thing special for cricket lov ers #cricket #special #fans', '//t.co/DBnWOQBxRB', '//t.co/YTuYDsR9sa', ' Some of @brookeccnorwich ladies have been busy the last few days. 🛭 U s newbies have a bit of work to do! #StayAtHomeChal...', '//t.co/W9Uvqq82g4', ' I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', ' 3 FOR 2 OFFER!\n\nTreat yourself to our new range of accessories, a p erfect addition to your exercise routine\n\n#payntr #pe...', Day2! Challenge 2! \n\nSomething a little less tiring today, it's tim e for cricket to meet tennis! \n\nLongest rally wins!...', ' (1/3) [It's official! The new members of the World Sports Betting Ca pe Cobras are ...\n\n#Cricket #Cobras #CobrasCricket ht...', " Here's my pick of 10 awesome #cricket documentaries you can watch onl ine during the #lockdown .There are notable omissions...", ' I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', '//t.co/ov3iCPMbWv', '//t.co/rNm9yqRIZh #Cricket', ' Here's an interesting question for you all \$\n\nWhat was the best Tes t innings of the last decade? It has to be an inni...', Here's an interesting question for you all ↓\n\nWhat was the best Tes t innings of the last decade? It has to be an inni...', ' Muhammad Yousaf\n\n#cricket', '//t.co/AqFgJ6P0q9', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & amp; India\xa0should be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & amp; India\xa0should be played t...', '//t.co/iBbFadJgVy', ' Presenting, The Wisden Leading T20 Cricketer of 2019! [\\n\n] Choice @ WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenCri...', '//t.co/qWaoc9Tsi6', ' Presenting, The Wisden Leading T20 Cricketer of 2019! □\n\n□ Choice @ WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenCri...', '//t...', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & amp; India\xa0should be played t...' " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", '//t.co/AlMHIn6ZTr', '//t.co/keoU8UcXfe', '//t.co/8x8coTrsYc', " Here's my pick of 10 awesome #cricket documentaries you can watch onl ine during the #lockdown .There are notable omissions...", '//t.co/QPOr5sgvDV', '//t.co/aaXMCrWox1', '@durfiteymun Great finish to a fantastic game of #Cricket', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & mp; India\xa0should be played t...', So here it is @dizzy259 in a chat with me on his career highlights, @ VVSLaxman281, injuries, @ImIshant, @sachin_rt and a lo...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & amp; India\xa0should be played t...', Back to cricket - A simple drill involving just a tennis ball for any close in fielders or wicket keepers, @lukewright…', ' Pain in life was easy to deal when we were together. #cricket', '//t....', ' @bebutch @abc730 @JoshFrydenberg Idealogical constraints from little johnny? What does this mean? Did little johnny have ideal...', '//t.co/kHRh9FLE4J', '//t.co/ogeArhoYRj', '//t.co/x0Gp9QfSsf\n\nRaise y...', ' SL v IND series since 2015\n\nT20I - L L L L L\nODI - L L L\nTest -L L \n\nSL = 12L\ncancel L on both sides\nS = 12\n\nBoltzm...', '//t.co/PaqF1h3JFt', '//t...', ' (2/3) [It's official! The new members of the World Sports Betting Ca pe Cobras are ...\n\n#Cricket #Cobras #CobrasCricket ht...', Which cricket format is the best? #Cricket', '//t.co/0VL03JFp3Z', '//t.co/6vgCT1Jqwf', '//t.co/QPOr5sgvDV', '//t.co/rNzuIkSrk6', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', ' Interviewing @faf1307 & amp; @amlahash in #CapeTown last August, ahead of WorldXI tour of Pakistan, one of the highlights of m...', '//...', ' PCB releases NOC policy for its players, all centrally contracted pla yers may seek clearance for up to a maximum of four...', '//t.co/B86Eu2qdBQ', ' Interviewing @faf1307 & amp; @amlahash in #CapeTown last August, ahead of WorldXI tour of Pakistan, one of the highlights of m...', DHONI'S BIGGEST ASSET WAS HIS CALM UNDER PRESSURE\n@msdhoni @PaddyUpt on1 @ICHOfficial @BCCI @imVkohli #MSDhoni #Cricket htt...', " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', '//t.co/nSR3ZaVZZh', ' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & mp; India\xa0should be played t...', @intouchcricket and @EpsomC_Cricket are producing a series of simple, informative and fun coaching drills. This straigh...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & amp; India\xa0should be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & amp; India\xa0should be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & mp; India\xa0should be played t...', '//t.co/lPBLd0RdXb', '//t.co/boAsQEEeoN' '//t.co/dudwZEXDH1' ' #Pakistan's legendary former pacer Shoaib Akhtar (@shoaib100mph) pro posed a three-match Indo-Pak #Cricket series to r...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & amp; India\xa0should be played t...', A musical one from @joybhattacharj to start things off for the day! ['//t....'] In [41]: | # Reading the tweets tweets_list Out[41]: [' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', '//t.co/7X5hAeYCNg', '//t.co/3vnLCuQHNZ', '//t.co/ACofoC...', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', @windiescricket @Russell12A []\n\nBest Strike Rates in IPL History \n (30+ Batting Average) \n\n10 Andre Russell 186.41\n20 Ris...', '//t...', '//t.co/rNm9yqRIZh #Cricket', '//t.co/WbXskQ1jt2', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xa0should be played t...', Which cricket format is the best? #Cricket', '//t.co/yAbK7smh0r', ' Michael Clarke\xa0"Australia & probably every other team sucked u p to India & amp; were too scared to sledge Kohli & amp; other Indi...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xa0should be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & amp; India\xa0should be played t...', Michael Clarke\xa0"Australia & probably every other team sucked u p to India & were too scared to sledge Kohli & other Indi...', " Another of my favourite finds, this tiny tin containing the pocket cr icket game 'Owzthat', sometimes known as 'Trench C...", Day 15 of groundsman isolation, finished with the roller for the mome nt, time to cut some grass!! #lords #lovelords #cr...', 'Follow my account i will be posting some thing special for cricket lov ers #cricket #special #fans', '//t.co/DBnWOQBxRB', '//t.co/YTuYDsR9sa', ' Some of @brookeccnorwich ladies have been busy the last few days. 🛭 U s newbies have a bit of work to do! #StayAtHomeChal...', '//t.co/W9Uvqq82g4', ' I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', ' I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', ' 3 FOR 2 OFFER!\n\nTreat yourself to our new range of accessories, a p erfect addition to your exercise routine\n\n#payntr #pe...', Day2! Challenge 2! \n\nSomething a little less tiring today, it's tim e for cricket to meet tennis! \n\nLongest rally wins!...', (1/3) I It's official! The new members of the World Sports Betting Ca pe Cobras are ...\n\n#Cricket #Cobras #CobrasCricket ht...', " Here's my pick of 10 awesome #cricket documentaries you can watch onl ine during the #lockdown .There are notable omissions..." I simply cannot believe there are supposed Cricket fans out there tha t still, to this day, do not appreciate or even,...', '//t.co/ov3iCPMbWv' '//t.co/rNm9ygRIZh #Cricket', ' Here's an interesting question for you all \\n\nWhat was the best Tes t innings of the last decade? It has to be an inni...', ' Here's an interesting question for you all \$\n\nWhat was the best Tes t innings of the last decade? It has to be an inni...', Muhammad Yousaf\n\n#cricket', '//t.co/AqFqJ6P0q9', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xaOshould be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xaOshould be played t...', '//t.co/iBbFadJgVy', ' Presenting, The Wisden Leading T20 Cricketer of 2019! [\n\n] Choice @ WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenCri...', '//t.co/qWaoc9Tsi6', ' Presenting, The Wisden Leading T20 Cricketer of 2019! [\\n\n] Choice @ WisdenAlmanack\n\n#KorboLorboJeetbo #KKR #Cricket #WisdenCri...', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xaOshould be played t...', " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", '//t.co/AlMHIn6ZTr', '//t.co/keoU8UcXfe', '//t.co/8x8coTrsYc', " Here's my pick of 10 awesome #cricket documentaries you can watch onl ine during the #lockdown . There are notable omissions...", '//t.co/QPOr5sgvDV', '//t.co/aaXMCrWox1' '@durfitevmun Great finish to a fantastic game of #Cricket'. ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & amp; India\xaOshould be played t...', So here it is @dizzy259 in a chat with me on his career highlights, @ VVSLaxman281, injuries, @ImIshant, @sachin rt and a lo...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', Back to cricket - A simple drill involving just a tennis ball for any close in fielders or wicket keepers, @lukewright...', Pain in life was easy to deal when we were together. #cricket', @bebutch @abc730 @JoshFrydenberg Idealogical constraints from little johnny? What does this mean? Did little johnny have ideal...', '//t.co/kHRh9FLE4J', '//t.co/ogeArhoYRj', '//t.co/x0Gp9QfSsf\n\nRaise y...', 'SL v IND series since 2015\n\nT20I - L L L L L\nODI - L L L\nTest -L L \n\nSL = 12L\ncancel L on both sides\nS = 12\n\nBoltzm...', '//t.co/PaqF1h3JFt', '//t...', ' (2/3) 🛘 It's official! The new members of the World Sports Betting Ca pe Cobras are ...\n\n#Cricket #Cobras #CobrasCricket ht...', ' Which cricket format is the best? #Cricket', '//t.co/0VL03JFp3Z', '//t.co/6vgCT1Jqwf' '//t.co/QPOr5sgvDV' '//t.co/rNzuIkSrk6', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xa0should be played t...', Interviewing @faf1307 & @amlahash in #CapeTown last August, ahead of WorldXI tour of Pakistan, one of the highlights of m...', ' PCB releases NOC policy for its players, all centrally contracted pla yers may seek clearance for up to a maximum of four...', '//t.co/B86Eu2qdBQ', ' Interviewing @faf1307 & amp; @amlahash in #CapeTown last August, ahead of WorldXI tour of Pakistan, one of the highlights of m...', DHONI'S BIGGEST ASSET WAS HIS CALM UNDER PRESSURE\n@msdhoni @PaddyUpt on1 @ICHOfficial @BCCI @imVkohli #MSDhoni #Cricket htt...', " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', '//t.co/nSR3ZaVZZh', ' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon Harmer @WisdenCricket @WisdenAlmanack...', ' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', " Raced and paced his way to the top! [\n\nThe number 1 Test bowler has also made his way into the elite list of 'Wisden Cricket...", ' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', ' #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', #YellowSport Essex ace Harmer named as one of Wisden's elite @EssexCr icket @Simon_Harmer_ @WisdenCricket @WisdenAlmanack...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match O DI series between Pakistan & India\xa0should be played t...', @intouchcricket and @EpsomC_Cricket are producing a series of simple, informative and fun coaching drills. This straigh...', ' Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & amp; India\xa0should be played t...', '//t.co/lPBLd0RdXb', '//t.co/boAsQEEeoN', '//t.co/dudwZEXDH1' ' #Pakistan's legendary former pacer Shoaib Akhtar (@shoaib100mph) pro posed a three-match Indo-Pak #Cricket series to r...', Shoaib Akhtar\xa0has suggested that whenever possible a three-match 0 DI series between Pakistan & India\xaOshould be played t...', A musical one from @joybhattacharj to start things off for the day! ['//t....'] In [44]: # Putting the tweets in DataFrame df = DataFrame ({'user_ID':user_list,'tweets':tweets_list}) df.head() Out[44]: user_ID tweets 0 RT @Saj_PakPassion Shoaib Akhtar has suggested that whenever pos... 1 Wisden has just depressed me by highlighting t... //t.co/7X5hAeYCNg 2 RT @SunilAh64145529 //t.co/3vnLCuQHNZ RT @timesofindia 3 //t.co/ACofoC... RT @Saj_PakPassion Shoaib Akhtar has suggested that whenever pos... **Answering the questions** 1) What are tags are associated with a person, place or thing? # Tags associated with match In [54]: Tags_Match = match.loc[:,['Team1','Team2','match_date']] Tags_Match Out[54]: Team1 Team2 match date Match SK 0 Royal Challengers Bangalore Kolkata Knight Riders 4/18/2008 Chennai Super Kings 1 Kings XI Punjab 4/19/2008 2 Delhi Daredevils Rajasthan Royals 4/19/2008 3 Mumbai Indians Royal Challengers Bangalore 4/20/2008 Kolkata Knight Riders **Deccan Chargers** 4/20/2008 632 Delhi Daredevils Royal Challengers Bangalore 5/14/2017 633 Mumbai Indians Rising Pune Supergiants 5/16/2017 634 Sunrisers Hyderabad Kolkata Knight Riders 5/17/2017 635 Kolkata Knight Riders 5/19/2017 Mumbai Indians 636 Mumbai Indians Rising Pune Supergiants 5/21/2017 637 rows × 3 columns # Tags associated with venue Tags_venue = match.loc[:,['Venue_Name','City_Name','Country_Name']] Tags_venue Out[55]: Venue_Name City_Name Country_Name Match_SK 0 M Chinnaswamy Stadium India Bangalore 1 Punjab Cricket Association Stadium, Mohali Chandigarh India Delhi India Feroz Shah Kotla 3 Wankhede Stadium India Mumbai Kolkata **Eden Gardens** India ... 632 Delhi India Feroz Shah Kotla 633 Wankhede Stadium Mumbai India 634 M Chinnaswamy Stadium Bengaluru India 635 M Chinnaswamy Stadium India Bengaluru Rajiv Gandhi International Stadium Uppal Hyderabad (Deccan) 636 India 637 rows × 3 columns In [56]: # Tags associated with Result Tags_Result = match.loc[:,['match_winner','Win_Type','Outcome_Type','Win _Margin']] Tags_venue Out[56]: City_Name Country_Name Venue_Name Match_SK M Chinnaswamy Stadium Bangalore India 1 Punjab Cricket Association Stadium, Mohali Chandigarh India Feroz Shah Kotla Delhi India 3 Wankhede Stadium Mumbai India **Eden Gardens** Kolkata India 632 Feroz Shah Kotla Delhi India 633 Wankhede Stadium Mumbai India 634 M Chinnaswamy Stadium Bengaluru India 635 M Chinnaswamy Stadium Bengaluru India 636 Rajiv Gandhi International Stadium Uppal Hyderabad (Deccan) India 637 rows × 3 columns 3. What people, places or things are trending in your domain? In [86]: X = df.groupby('user_ID').agg({'tweets':np.size}).sort_values(['tweets'], ascending=**False**) X.head(5)Out[86]: tweets user_ID RT @Saj_PakPassion 19 RT @Daniel86criket RT @KKRiders RT @briancjeeves RT @PoppingCreaseSA It can be seen that the user_ID named @Saj_PakPassion has tweeted 19 times in our gathered. Hence being the most most trending person in our domain. 4. What people, places or things are popular in your domain? In [82]: y =match.groupby('City_Name').agg({'match_id':np.size}).sort_values(['ma tch_id'], ascending=False) y.head(5)Out[82]: match_id City_Name Mumbai 85 Kolkata 61 Delhi 60 Bangalore 58 Chennai 48 It can be seen that the city_name with Mumbai has hosted most number of matches. Hence,it is the most popular place in the domain **Design Explanations** In SQL databases we will storing values in the form of relational tables by linking each table with the primary key and foriegn key relationship. But in the NoSQL databases will be storing the data in the form of documents. In the SQL databases, the values are stored in the form of tables but in the NoSQL databases the values stored in the form of Collections. For Example, in our cricket database we will be having data for each match in a table and all its related data in the other table. But in the NoSQL

databases all data are stand alone document.

shtag similar to cricket.

The data used in this Assignment is downloaded from Kaggle. The d

Twitter API was used to fetch tweets and the user IDs with the ha

consumer_secret = 'hJWMDqHG6rAtNn5fV6oIgNwaVcdy2yTIY0WXNWoy0KDD81

ata is audited for null values. Also, Additionally an

We extracted data using the below Twitter API consumer_key = 'Ulg0a0EDEy09m90eQ1WjgY8ax'

Report

Web API

Assingnment 3 - IPL cricket database

The main aim of this assignment is to convert the SQL database w

Abstract