## Name: Prasanna Kshatri ID:2226435

First check that the tweets have been imported: show

collections

```
my_db> show collections
collectionName
dept
emp
weather
```

The Twitter data is stored in the weather collection. We can run db.weather.count() to count the number of documents: db.weather.count()

my\_db> db.weather.count()
DeprecationWarning: Collection.count() is deprecated. Use countDocuments or estimatedDocumentCount.
250

We can examine the contents of one of the documents by running: db.weather.findOne()

```
_id: ObjectId('65f261268c238a184dea5bdc'),
contributors: null,
truncated: false, text: 'I really might have seasonal depression because the sun is out?, the meather is great, and I feel more motivated than ever.',
is_quote_status: false,
in_reply_to_status_id: null,
id: Long('1224421887255185537'),
favorite_count: 0,
entities: { symbols: [], user_mentions: [], hashtags: [], urls: [] },
coordinates: null,
in_reply_to_screen_name: null,
in_reply_to_user_id: null,
retweet_count: 0,
id_str: '1224421087255105537',
favorited: false,
user: {
    follow_request_sent: false,
   has_extended_profile: false,
profile_use_background_image: true,
   default_profile_image: false,
   profile_background_image_url_https: 'https://abs.twimg.com/images/themes/theme17/bg.gif',
    verified: false
   translator_type:
   translator_type: none ;
profile_text_color: '3333333',
profile_image_url_https: 'https://pbs.twimg.com/profile_images/1176003516676091984/7WyNrPkk_normal.jpg',
profile_sidebar_fill_color: 'E6F6F9';
   entities: { description: { urls: [] } },
   profile_sidebar_border_color: 'DBE9ED',
id_str: '39871078',
profile_background_color: 'DBE9ED',
   listed_count: 3,
is_translation_enabled: false,
   utc_offset: null,
statuses_count: 27643,
description: 'Piece of Sunshine, Inner peace, Entertainer.',
  friends_count: 436,
location: "In God's Hands",
profile_link_color: 'CC3366',
profile_image_url: 'http://pbs.twimg.com/profile_images/1176003516676891904/7WyNrPkk_normal.jpg',
   Tottowing: Taise,
geo_enabled: true,
profile_banner_url: 'https://pbs.twimg.com/profile_banners/39871078/1560231788',
profile_background_image_url: 'http://abs.twimg.com/images/themes/theme17/bg.gif',
   screen_name:
   profile_background_tile: true,
favourites_count: 5849,
name: 'Sweet Holy Honey '',
   notifications: false,
   url: null,
created_at: 'Wed May 13 23:40:46 +0000 2009',
contributors_enabled: false,
   time_zone: null,
protected: false,
   default_profile: false, is_translator: false
geo: null,
in_reply_to_user_id_str: null,
created_at: 'Mon Feb 03 19:55:10 +0000 2020'.
in_reply_to_status_id_str: null,
place: null,
metadata: { iso_language_code: 'en', result_type: 'recent' }
```

We can find the distinct values for a specific field by using the distinct() command. For example, let's find the distinct values for lang (language): db.weather.distinct("lang")

```
my_db> db.weather.distinct("lang")
[ 'en', 'pt' ]
```

or find what the text of the tweets say: db.weather.distinct("text")

```
_db> db.weather.distinct("text")
           The Australian wildfires sent one of the largest planes of smoke... higher than certain satellites have ever obser. https://t.co/c00aCpFapl
history in 1997, the record for the coldest temperature was recorded in Seag, Yuken Territories, hitting a low e. https://t.co/sSQllbm/ry
**Latest available observations (Nepal time)**\n" *
                                       /ms, le:Ms)(" +
secamp: Temp = -9°C\n" +
sp II: Temp = -19°C, Max gust_ https://t.co/hxM0FeC5%a',
test available observations (Mepal time)+*\n" +
                                     /el, IT:45\n' +
secang: Teng = -18°C\n' +
secang: Teng = -18°C\n' +
secang: Teng = -18°C, Max quet. https://t.co/MYZECUMYMO',
test available observations (Hepal time)+>\n' +
                        Case 11: Tamp = -10°C, Max gust = 9m/s\m' +

[No. 1]

[No. 1]

[No. 1]

[No. 1]

[No. 1]

[No. 1]

[No. 2005]

[No. 1]

[No. 2005]

[No. 2
                              In the://t.co/%pTUGAF58",

otallamberg What about all these areas that are having one of the coldest winters over? \n' +

ou didn't mention thes. https://t.co/irfrFFMChr',

onloceaun These who more surprised by the halftime show, haven't over watched Latin American programming and un... https://t.co/lumwAuGeVF

stazz68 Thanks Joe 1 With the meather here so crary 66°during the day and 4° below zero the next the Earometric. https://t.co/lumwAuGeVF
                              relability in: +

colding Jordani'n' +

colding Jordani'n' +

colding gorgoous as ever this morning doing the traffic reports. Just like you did. https://t.co/sSciyWanWC',

ctin)Tipe Bo you and 0c harichello ever get into meather arguments? https://t.co/sWbMc19MD',

chrushAEFO ever heard of the Ventusky meather app?? If not check it out, it does cest $2.90 the',

cutscapald' if only I'd have known you had a stern to meather' is one of the best lyries I've ever heard. I doet. https://t.co/ivpISIbMNS'

Trades12 raw you imagine ? sidest I've ever been in my life ?? on a cruise ship and it wasn't from bud meat. https://t.co/W6aE295IS'

torGheather @doeclandTrust I'll have a look, thanks. 'n' +

as ever-warmy of claims that occasional episodes of apport. https://t.co/MF2BMc7amu',

asmbob @armsonic @chinodwanz @ghreciella @lollypes21s You can't force cycling an Lagesians we don't need it. The https://t.co/191aYY18DQ'

playConic @promoteconics @toposbconics @Mebconicspdates @MebConicNetwark Larby you.\n' +
             Canada Disabania Delicoformuma Operacialla (Mollysessia voi can't force cycling on Lagacians we don't need it. Th. https://t.co/s9laYV18DQ1.
SeplayConic Operactocomics Operacobounics Observational Control of Core cycling on Lagacians we don't need it. Th. https://t.co/s9laYV18DQ1.
"In" -
"It's been raining pretty ms. https://t.co/s9laYDyV18Q0',
INSELEC Shit like this clearly coming from people she have never played anything but Sunday football. Hes's be ms. https://t.co/S8BC2FM50.
Towardint Operation only in your dreams be does not have a clum about econfur. He is the marse I have ever seen."
TackorRossIon (McCanadoportalum I think during the entire Deflatogate sepa, the NT. realized that svery game ever p. https://t.co/S8BC2FM50.
TackorRossIon (McCanadoportalum I think during the entire Deflatogate sepa, the NT. realized that svery game ever p. https://t.co/S8BCQUINTALIZED CONTROLOGY.
TackorRossIon (McCanadoportalum I think during the entire Deflatogate sepa, the NT. realized that svery game ever p. https://t.co/S8BCQUINTALIZED CONTROLOGY.
TackorRossIon (McCanadoportalum I think during the entire Deflatogate sepa, the NT. realized that svery game ever p. https://t.co/S8BCQUINTALIZED CONTROLOGY.
TackorRossIon (McCanadoportalum I think during the entire think sever book that the think of the second that the second the second the second that the second
                            In the solution of the side of road putted the food on the car and in cold seathur enjoy the fish, broad, tanatees and. https://t.co/mistbsvbsp is not sclinatechange, its sweather... https://t.co/mistbsvbsp is not sclinatechange, its sweather... https://t.co/mistbsvbsp is not sclinatechange, its sweather... https://t.co/mistbsvbsp is not sclinate in the side of road putted the food on the car and in cold seathur enjoy the fish, broad, tanatees and... https://t.co/mistbsvbsp is non fish 1, the Dyatley group started up Read 1879, called "Dead Rountain" by the indigeneur Ramei. With the m. https://t.co/mistbsvbsp sare packed. Ready for my longest flight ever 1 And most radical weather change ! This is gones be one to remem. https://t.co/smistbsvbsp it January bus duty weathur EVERS Bliennation Edsadpride https://t.co/clyq/lib326 https://t.co/sclwbspffAk.

"It was to be in a car accident because I have to go to school under this weather is! Please don't tell us ever. https://t.co/ibfe6IAEDw dost neather I've ever felt offcels may may colder than 25 https://t.co/mistbsvbsp.

"It was tell us sounder couldn't heat a Suppressal. Our February seather is usually protty mild. \n' +
```

Finding unique values within a nested document needs the use of the dot notation, for example, find a user's name: db.weather.distinct("user.name")

Some documents can have several layers of nesting, which can be accessed using the dot notation: db.weather.distinct("extended\_entities.media.type")

```
my_db> db.weather.distinct("extended_entities.media.type")
[ 'animated_gif', 'photo' ]
```

We can search for fields with a specific value using the find() command. For example, pick one of the names seen in the previous command and search for their tweets. For example: db.weather.find({"user.name":"Bri"})

```
db> db.weather.find({"user
      _id: ObjectId('65f261268c238a184dea5c8d'),
contributors: null,
truncated: false,
text: 'Wouldn't of imagined I'd ever be so
                ordinates: null,
urce: '<a href='http://twitter.com/download/iphono" rel="nofollow">Twitter for iPhone</a>',
reply_to_screen_name: null,
reply_to_user_id: null,
tweet_count: 0,
str: '1224998601369365584',
vorited: false,
ir: {
                                                                  ackground_image_url_https: 'https://abs.twing.com/images/themes/theme1/bg.png',
                                                       url: 'https://t.co/uBytdjNNhm',
indices: [ 6, 23 ],
expanded_url: 'http://youtu.be/Em4pWz8Am6e
display_url: 'youtu.be/Em4pWz8Am6e'
        description: { urts: t; },
},
followers_count: 416,
profile_sidebar_border_color: 'CRREED',
id_str: '192253750',
profile_background_color: 'CRREED',
listed_count: 1,
is translation_enabled: false,
utc_offset: null,
statuses_count: 2056,
description: 'just a young balla city livin',
friends_count: 514,
location: 'New York: NV',
profile_link_color: 'IDAIF2',
profile_link_color: 'IDAIF2',
profile_link_color: 'IDAIF2',
profile_link_color: 'IDAIF2',
profile_link_color: 'http://pbs.twimg.com/profile_banners/198253756/1577926383',
profile_banner_url: 'https://pbs.twimg.com/profile_banners/198253756/1577926383',
profile_banner_url: 'https://pbs.twimg.com/profile_banners/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/themes/the
               screen_name: 'BrianaClugston',
lang: null,
profile background_tile: false,
favourites_count: 5051,
            name: 'Bri',
notifications: false,
url: 'https://t.ce/uBytdjRNAme',
created_at: 'Sun Oct 03 19:53:29 +0000 2018',
contributors_enabled: false,
time_zone: null,
protected: false,
default_profile: true,
default_profile: true,
is_translator: false
               o: null,
_reply_to_user_id_str: null,
         ng: 'en', 'Mon Feb 83 18:26:81 +8888 2820',
reated_at: 'Mon Feb 83 18:26:81 +8888 2820',
n_repty_to_status_id_str: null,
lace: null,
lace: null,
etadata: [ iso_language_code: 'en', result_type: 'recent' ]
```

We can specify a second argument to the find() command to only show specific field(s) in the result. Let's repeat the previous search, but only show the tweet\_ID field: db.weather.find({"user.name":"Bri"}, {id: 1}).pretty()

The \_id field is objectID for every document and will be different for each user. We can remove it from the results with the following filter: db.weather.find({"user.name":"Bri"}, {id: 1, \_id:0}).pretty()

```
my_db> db.weather.find({"user.name":"Bri"},{id:1,_id:0}).pretty()
[ { id: Long('1224398651369365584') } ]
```

For example, if we search for the value sun in the tweet's text field, there are no results: db.weather.find({text: "sun"}) However, if we search using a regular expression, there are many results: db.weather.find({text: /sun/}, {text: 1})

```
my_db> db.weather.find({text:"sun"})
my_db> db.weather.find({text:"sun"}, {text:1})

my_db> db.weather.find({text:"sun"}, {text:1})

{
    id: ObjectId('65f261268c238a164dea5bdc'),
    text: 'I really might have seasonal depression because the sun is out?, the weather is great, and I feel more motivated than ever.'

}

id: ObjectId('65f261268c238a164dea5c84'),
    text: 'Ever notice how the weather can impact our moods so much? A sunny day in Pittsburgh. everyone is out & about! Tha_ https://t.co/rya6uN3oeI'

id: ObjectId('65f261268c238a164dea5cba'),
    text: 'RT @capitalmeather: Ever hear the saying that a butterfly flapping its wings in China could bring rain instead of sunshine to New York? Tur_'

id: ObjectId('65f261268c238a164dea5cdu'),
    text: 'RT @capitalmeather: Ever hear the saying that a butterfly flapping its wings in China could bring rain instead of sunshine to New York? Tur_'

id: ObjectId('65f261268c238a164dea5cdu'),
    text: 'RT @capitalmeather: Ever hear the saying that a butterfly flapping its wings in China could bring rain instead of sunshine to New York? Tur_'

}
```

Note, the quotes have gone, otherwise it will try to search for "/sun/" The difference between the queries is that the first searched for where the text field value was exactly equal to sun, and the second searched for where the field value contained sun. We can append .count() to the command to count the number of results: db.weather.find({text: /sun/}).count() How many included rain in the text instead: db.weather.find({text: /rain/}).count()

```
my_db> db.weather.find({text:/sun/}).count()
4
my_db> db.weather.find({text:/rnin/}).count()
13
```

For example, the previous search for sun in the text field can also be written using \$regex: db.weather.find({ text: { \$regex: /sun/ } }).count()

```
my_db> db.weather.find({text:{$regex:/sun/}}).count()
4
my_db> db.weather.find({text:{$regex:/sun/1}}).count()
9
my_db> db.weather.find({text:{$regex:/*RT/1}}).count()
86

my_db> db.weather.find({text:{$regex:/*RT/1}}).count()
8
```

To include a regular expression in an \$in query expression, you can only use JavaScript regular expression objects (i.e. /pattern/). \$in is similar to the IN clause in SQL, where you can search for a value that matches any value in the list.

For example, find any texts that starts with sun in any case or contains rain: db.weather.find({ text: { \$in: [/^sun/i, /rain/] } }, {text: 1} )

```
### Objection (Septimental State State of the State St
```

```
my_db> db.weather.find({text:{$regex:/cold.*night/,$options:"si"}},{_id:0, text:1}).count()
0
```

Let's create the index using createIndex():

```
my_db> db.weather.createIndex({text:"text"}
text_text
```

Next, we can use the \$text operator to search the collection. We can perform the previous query to find the documents containing snow:

```
my_db> db.weather.find({$text:{$search:"snow"}}).count()
6
my_db> db.weather.find({$text:{$search:"snow -ice"}}).count()
a
```

The function getIndexes() can be used to check what indexes exist on a collection:

Indexes can also be dropped using the dropIndex() function.

```
my_db> db.weather.dropIndex("text_text")
{ nIndexesWas: 2, ok: 1 }
```

## **Further Queries: Operators**

MongoDB can also search for field values matching a specific criteria. For example, we can find where the retweet\_count is greater than 20:

```
my_db> db.weather.find({"user.location":{$qq: null}}).count()

my_db> db.weather.find({"user.goo_enabled":{$qq:true}}).count()

my_db> db.weather.find({"user.goo_enabled":{$qq:true}}).count()

my_db> db.weather.find({"user.goo_enabled":{$qq:true}}).count()

my_db> db.weather.find({"user.followers_count":{$tte:10}}).count()

my_db> db.weather.find({"retweeted_status.user.followers_count":{$qt:2000}}).count()

my_db> db.weather.find({$where:"this.user.followers_count>this.user.friends_count"}).count()

my_db> db.weather.find({$and:{{text:/icw/},{"retweeted_status.retweet_count":{$qt:50}}})).count()

my_db> db.weather.aggregate({{$qroup:{id:null,totalAmount:{$sum:"$retweeted_status.rotweet_count"},count:{$sum:1}}}))

my_db> db.weather.find({$and:{{text:/icw/,{"retweeted_status.rotweet_count":{$qt:50}}}}).count()

my_db> db.weather.find({$and:{{text:/icw/,{"retweeted_count":{$qt:50}}}}}).count()

my_db> db.weather.find({$and:{{text:/icw/,{"retweeted_cou
```