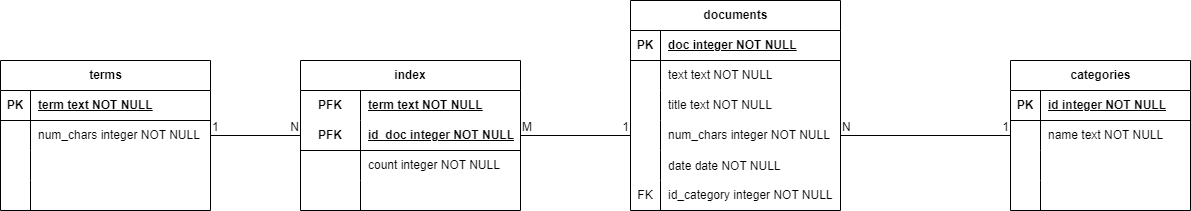
Bronco ID: 014644097

Last Name: Pham

First Name: Nathan



* 1. SELECT COUNT(\*) FROM documents
  2. SELECT COUNT(\*) FROM documents INNER JOIN categories ON categories.name = 'Sports' WHERE documents.id\_category = categories.id
  3. SELECT COUNT(\*) FROM terms
  4. SELECT SUM(index.count) FROM index INNER JOIN documents ON documents.title = 'Arizona' WHERE index.id\_doc = documents.doc
  5. SELECT SUM(index.count) FROM index INNER JOIN documents ON index.id\_doc = documents.doc INNER JOIN categories ON documents.id\_category = categories.id WHERE categories.name = 'Seasons'
  6. SELECT SUM(index.count) FROM index WHERE index.term = 'months'
  7. SELECT documents.title, SUM(index.count) FROM index INNER JOIN documents ON documents.doc = index.id\_doc GROUP BY documents.title ORDER BY SUM(index.count) DESC LIMIT 1
  8. SELECT index.term, SUM(index.count) FROM index GROUP BY index.term ORDER By SUM(index.count) DESC LIMIT 1

1. <https://github.com/nathanvpham/CS4250Assignment2>
2. Full document is on GitHub

Example:

{

"\_id": {

"$oid": "6537226b6ae93fe487b27e30"

},

"id": 1,

"title": "Exercise",

"text": "Baseball is played during summer months",

"num\_chars": 34,

"date": "2023-10-03",

"categories": {

"category": "Sports"

},

"index": {

"baseball": 1,

"is": 1,

"played": 1,

"during": 1,

"summer": 1,

"months": 1

},

"terms": [

{

"term": "baseball",

"num\_chars": 8

},

{

"term": "is",

"num\_chars": 2

},

{

"term": "played",

"num\_chars": 6

},

{

"term": "during",

"num\_chars": 6

},

{

"term": "summer",

"num\_chars": 6

},

{

"term": "months",

"num\_chars": 6

}

]

}

* 1. db.documents.find()
  2. db.documents.find({categories:{category:"Sports"}})
  3. db.documents.distinct("terms.term")
  4. db.getCollection('documents').aggregate(

[

{

$match: { 'categories.category': 'Seasons' }

},

{ $unwind: { path: '$terms' } },

{ $project: { 'terms.term': 1 } }

],

{ maxTimeMS: 60000, allowDiskUse: true }

);

* 1. db.getCollection('documents').aggregate(

[

{ $match: { title: 'Arizona' } },

{ $unwind: { path: '$terms' } },

{ $project: { 'terms.term': 1 } }

],

{ maxTimeMS: 60000, allowDiskUse: true }

);

* 1. db.getCollection('documents').aggregate(

[

{ $unwind: { path: '$terms' } },

{ $match: { 'terms.term': 'months' } },

{ $count: 'count' }

],

{ maxTimeMS: 60000, allowDiskUse: true }

);

db.getCollection('documents').aggregate(

[

{ $unwind: { path: '$terms' } },

{

$group: {

\_id: '$title',

count: { $sum: 1 }

}

},

{ $sort: { count: -1 } }

],

{ maxTimeMS: 60000, allowDiskUse: true }

);

db.getCollection('documents').aggregate(

[

{ $unwind: { path: '$terms' } },

{ $project: { terms: 1 } },

{

$group: {

\_id: '$terms',

count: { $sum: 1 }

}

},

{ $sort: { count: -1 } }

],

{ maxTimeMS: 60000, allowDiskUse: true }

);

1. <https://github.com/nathanvpham/CS4250Assignment2>