



УРОК 17.1. MONGODB ATLAS. ИНДЕКСЫ. НАПОЛНЕНИЕ КОЛЛЕКЦИЙ

MONGODB ATLAS	2
ЗАДАНИЕ ДЛЯ ЗАКРЕПЛЕНИЯ	3
ИНДЕКСЫ В MONGODB	5



MONGODB ATLAS

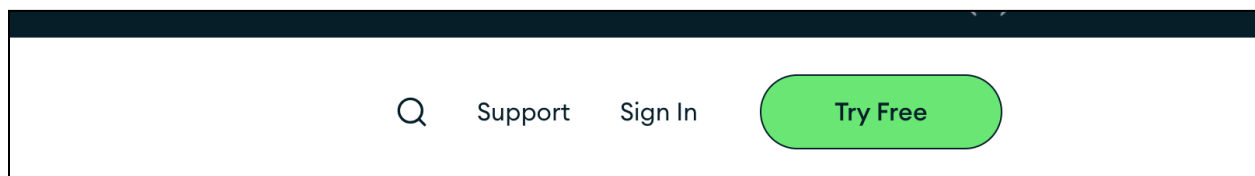
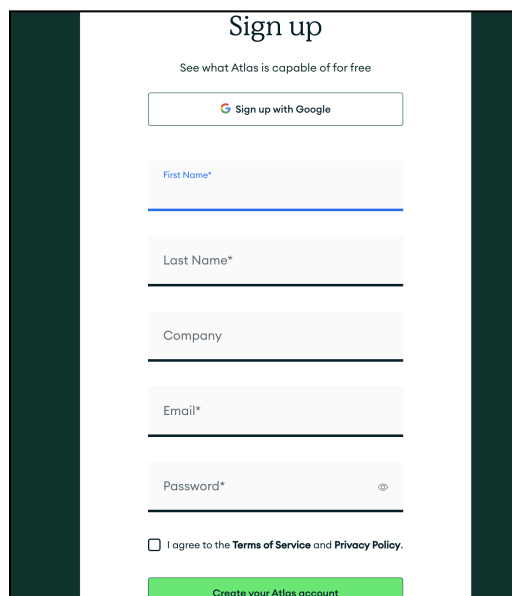
MongoDB Atlas — это облачный сервис от MongoDB, предназначенный для размещения, управления и масштабирования баз данных MongoDB в облачной среде. Сервис автоматизирует многие задачи, связанные с управлением базами данных, и позволяет интегрироваться с крупными облачными провайдерами, такими как AWS, Google Cloud и Microsoft Azure. Это обеспечивает гибкость в хранении и управлении данными.

MongoDB Atlas выделяется благодаря многоуровневой безопасности, включая шифрование и сетевую изоляцию, а также предлагает автоматическое резервное копирование и восстановление данных. Сервис также предоставляет инструменты для мониторинга и оптимизации производительности баз данных, что делает его надежным решением для управления данными в облаке.






ЗАДАНИЕ ДЛЯ ЗАКРЕПЛЕНИЯ


1. Перейти по ссылке <https://www.mongodb.com/> для регистрации
2. Выбрать Free atlas account

3. Выбрать название для проекта, посмотреть на пароль и доступ по ip адресу.
4. Нажать на Add Sample dataset для добавления данных для экспериментов.
5. После загрузки посмотреть на мониторинг и на новые данные.
6. Экспортировать для будущих экспериментов 2 коллекции - shipwrecks и sample_airbnb .



Andrew's Or... Access Manager  Billing

AtlasICH  Data Services App Services Charts

Overview

DEPLOYMENT

Database

Data Lake

SERVICES

Device Sync

Triggers

Data API

Data Federation

Atlas Search

Stream Processing

SECURITY

Backup

Database Access

Network Access

Advanced

New On Atlas 3

Goto

ANDREW'S ORG - 2024-01-09 > ATLASICH

Database Deployments

Find a database deployment...

Sample dataset successfully loaded. Access it in Collections or by connecting with the MongoDB Shell.

AtlasClusterICH

Connect View Monitoring Browse Collections ...

Visualize Your Data

Build dashboards and charts, and embed them in your apps with MongoDB Charts.

Dismiss


Explore Charts

R 0.05

W 0

Last 37 minutes


764.0/s



Connections 14.0

Last 42 minutes

14.0




In 119.1 B/s

Out 125.5 B/s

Last 37 minutes

1.0 MB/s




VERSION	REGION	CLUSTER TIER	TYPE	BACKUPS	LINKED APP SERVICES	ATLAS SQL	ATLAS SEARCH
6.0.12	AWS / Frankfurt (eu-central-1)	M0 Sandbox (General)	Replica Set - 3 nodes	Inactive	None Linked	Connect	Create Index


+ Add Tag

EDIT CONFIGURATION


FREE SHARED



Add Data



Load Sample Data



Data Modeling Templates

Featured Resources

GENERAL

[Get Started with Atlas](#)

[MongoDB Query API](#)

[Resources to Build AI-powered Apps](#)

[Tailored Solutions Library for Kickstarting Projects](#)

[Developer Center](#)



ИНДЕКСЫ В MONGODB

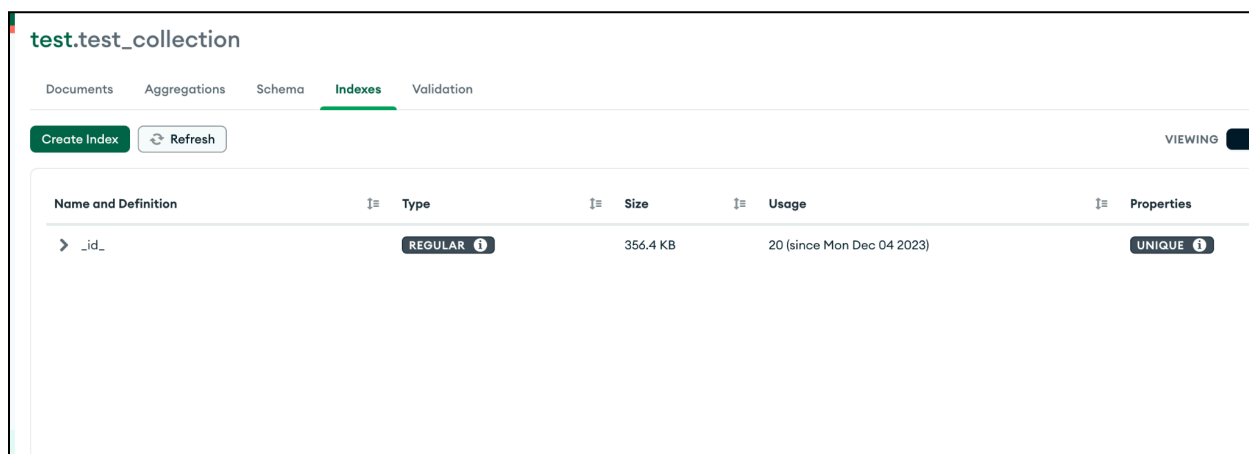


Индексы — это структуры данных, которые хранят небольшую часть набора данных коллекции в удобной для поиска форме.

Индексы значительно ускоряют поиск документов, так как индекс создается по определенному полю или полям и хранят упорядоченные значения определенного поля или набора полей.

Без индексов MongoDB должна выполнить сканирование коллекции, т. е. сканировать каждый документ в коллекции, чтобы выбрать те документы, которые соответствуют оператору запроса.

Автоматически создается индекс `_id_` на них можно посмотреть во вкладке `indexes`.



The screenshot shows the MongoDB Atlas interface for a collection named `test.test_collection`. The 'Indexes' tab is selected, showing a table of indexes. The table has columns for 'Name and Definition', 'Type', 'Size', 'Usage', and 'Properties'. A single index is listed: `_id_` with a 'REGULAR' type, a size of '356.4 KB', and usage of '20 (since Mon Dec 04 2023)'. The 'Properties' column indicates it is a 'UNIQUE' index.

Name and Definition	Type	Size	Usage	Properties
> <code>_id_</code>	REGULAR ⓘ	356.4 KB	20 (since Mon Dec 04 2023)	UNIQUE ⓘ

Один и тот же запрос `{price: {$gte: 60, $lte: 90}}` коллекция `sampleairbnb.listingsAndReviews` занимает совсем разное время.

Explain Plan

Explain provides key execution metrics that help diagnose slow queries and optimize index usage. [Learn more](#)

Visual Tree

Raw Output

> COLLSCAN

Returned **974**

Execution Time

141 ms

Documents Examined: **5555**

Explain Plan

Explain provides key execution metrics that help diagnose slow queries and optimize index usage. [Learn more](#)

Visual Tree

Raw Output

> FETCH

Returned **974**

Execution Time

0 ms

> IXSCAN

Returned **974**

Execution Time

2 ms

Index Name: **price_1**
Multi Key Index: **no**

Посмотрим на другую коллекцию и посмотрим на влияние индекса при выборке координат.

```
{coordinates: {$geoWithin: { $centerSphere: [ [ -96.48616143525156, 30.714748058919618 ], 0.07401030386454095 ]}}}
```



Результаты до :

> COLLSCAN

Returned 1023 Execution Time 3 ms

Documents Examined: 11095

Query Performance Summary

- 1023 documents returned
- 11095 documents examined
- 18 ms execution time
- Is not sorted in memory
- 0 index keys examined
- No index available for this query.

Создаем индекс :

```
db.shipwrecks.createIndex({coordinates: '2dsphere'})
```

Результаты после:

> FETCH

Returned 1023 Execution Time 0 ms

> IXSCAN

Returned 1529 Execution Time 0 ms

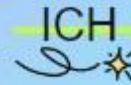
Index Name: coordinates_2dsphere

Multi Key Index: no

Query Performance Summary

- 1023 documents returned
- 1529 documents examined
- 3 ms execution time
- Is not sorted in memory
- 1535 index keys examined
- Query used the following index: coordinates (2dsphere)

Выполняем в mongosh :



```
db.shipwrecks.explain().find({coordinates: {$geoWithin: { $centerSphere: [ [ -96.48616143525156, 30.714748058919618 ], 0.07401030386454095 ]}}})
{
  explainVersion: '1',
  queryPlanner: {
    namespace: 'test.shipwrecks',
    indexFilterSet: false,
    parsedQuery: {
      coordinates: {
        '$geoWithin': {
          '$centerSphere': [
            [
              -96.48616143525156,
              30.714748058919618
            ],
            0.07401030386454095
          ]
        }
      }
    },
    queryHash: '9296542F',
    planCacheKey: '949A46C5',
    maxIndexedOrSolutionsReached: false,
    maxIndexedAndSolutionsReached: false,
    maxScansToExplodeReached: false,
    winningPlan: {
      stage: 'FETCH',
```

```
db.shipwrecks.explain().find({coordinates: {$geoWithin: { $centerSphere: [ [ -96.48616143525156, 30.714748058919618 ], 0.07401030386454095 ]}}})
```

Обращаем внимание на полученный ответ и на winningPlan:

- The IXSCAN
Этап указывает, что запрос использует индекс и какой индекс выбирается.
- The COLLSCAN
Этап указывает на то, что выполняется сканирование коллекции без использования каких-либо индексов.
- The FETCH
Этап указывает, что документы читаются из коллекции.
- The SORT
Этап указывает, что документы сортируются в памяти.

Через созданные индексы реализованы механизма защиты от дублированных документов.

Пример. У нас есть один из документов коллекции :

```
Unset
{
  "_id": { "$oid": "5ca4bbcea2dd94ee58162a6a" },
```




```
"username": "hillrachel",
"name": "Katherine David",
"address": "55711 Janet Plaza Apt. 865\nChristinachester, CT 62716",
"birthdate": { "$date": { "$numberLong": "582848134000" } },
"email": "timothy78@hotmail.com",
"Accounts": [
  { "$numberInt": "462501" },
  { "$numberInt": "228290" },
  { "$numberInt": "968786" },
  { "$numberInt": "515844" },
  { "$numberInt": "377292" }
],
"tier_and_details": {}
}
```

Мы создаем индекс и добавляем, что поле уникально :

```
db.customers.createIndex({email:1}, {unique:true})
```

Если мы попробуем добавить документ с таким же email - MongoDB вернет ошибку duplicate key error, и не добавит документ.

Пример ошибки из другой базы:

The screenshot shows the MongoDB Compass interface. On the left, a list of documents is visible. The main window displays the 'Insert Document' dialog for the 'test.test_collection'. The document being inserted is shown in a JSON format, with the following fields: `_id`, `age`, `workclass`, `total`, `education`, `education_num`, `marital_status`, `occupation`, `relationship`, `race`, `gender`, `capital_gain`, `capital_loss`, `hours_per_week`, `native_country`, and `income_bracket`. The error message at the bottom of the dialog indicates a duplicate key error: `E11000 duplicate key error collection: test.test_collection index: _id_ dup key: { _id: ObjectId('656dfc4d5afc911a8a7a4a98') }`. The 'Insert' button is highlighted in green.