

GETTING STARTED *with* mongoDB

Ramesh S



The background is a solid olive green. At the top, there are three white, stylized clouds of varying shapes. In the center, the text "Let's Learn From Home!" is written in a bold, dark blue font. Below the text, there is a large, white, stylized mountain range. At the base of the mountains, there are two dark blue, stylized plants with multiple leaves, one on the left and one on the right.

Let's Learn From Home!



01

INTRODUCTION

Get to know mongoDB
as NoSQL DB from your
understanding

02

mongoDB CLIENTS

Get to know the big players
whose use mongoDB

03

RDMS vs mongoDB Terminology

Know what is what in mongoDB
lingo

04

mongoDB BINARIES

See what mongoDB brings
out-of-box

05

mongoDB DATA TYPES

Data Types used in mongoDB

06

STORAGE ENGINES


Know the types of Storage Engines mongoDB uses

07

JavaScript in mongoDB

Know how to run






“I’m not a fan of MongoDB, you see. But to the way it was implemented and designed, I’m more than a fan.”

— **Whimsical Anonymous**



01. INTRODUCTION

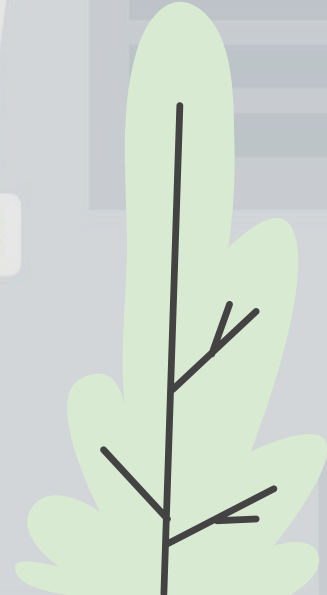
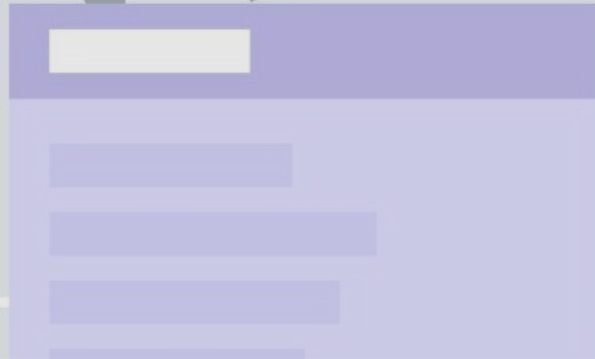
Get to know mongoDB
as NoSQL DB from your understanding.



What is a Database?



WHAT IS AN RDBMS?



WHEN DO RDBMSs FAIL?



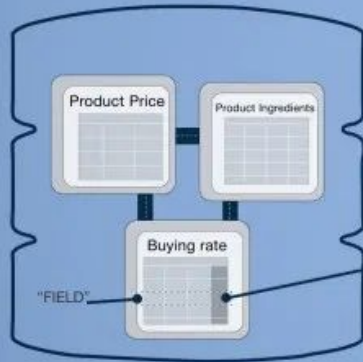
WHAT IS A NoSQL DATABASE?



SQL

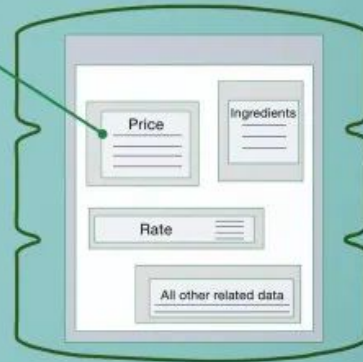
VS

NoSQL

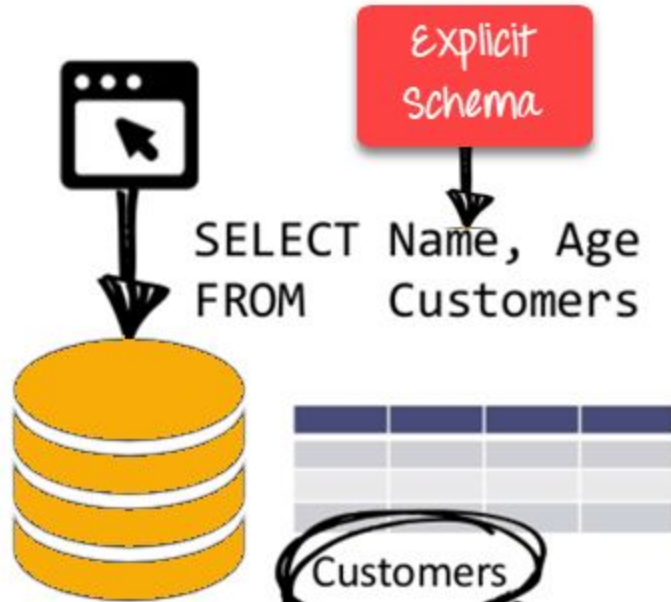


A RELATIONAL DATABASE TABLE
ORGANIZES STRUCTURED DATA
FIELDS INTO DEFINED COLUMNS.

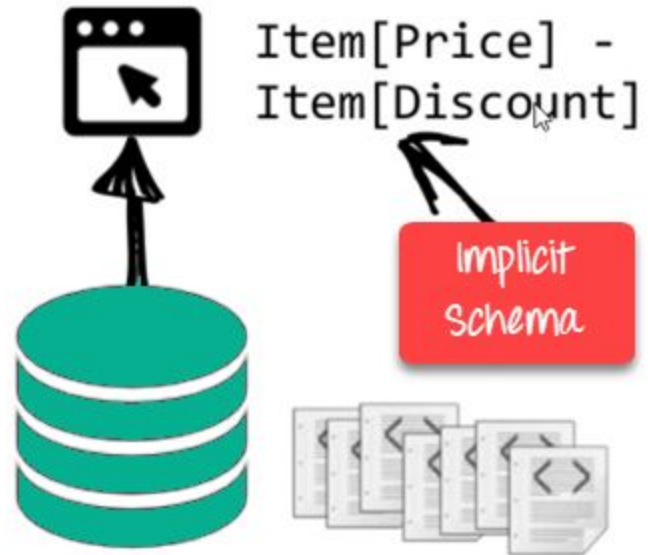
A NON-RELATIONAL DATABASE
DOES NOT INCORPORATE THE
TABLE MODEL. INSTEAD, DATA
CAN BE STORED IN A SINGLE
DOCUMENT FILE.



RDBMS:



NoSQL DB:



WHAT PROBLEMS DO NoSQL DBs SOLVE?



INTRO TO HUMONGOUS DATABASE

What is MongoDB?

Why use MongoDB?

Who is behind MongoDB?

Where is it suitable to use?

Who are using it?



NoSQL LANDSCAPE

Key-value



Graph database



Document-oriented



Column family



WHAT TO USE WHEN?

Use Key-Value Database if your application,

- is more performance driven, since the key-value stores are capable of providing much higher performances
- data model is not hierarchical

Use Graph Database

- If you want to store relationship information as a first-class entity

WHAT TO USE WHEN?

Use a Document-Store NoSQL Database if

- Your data schema is subject to frequent changes
- Your system will frequently need data in a computed or aggregate form
- You want to implement pagination of query results efficiently

WHAT TO USE WHEN?

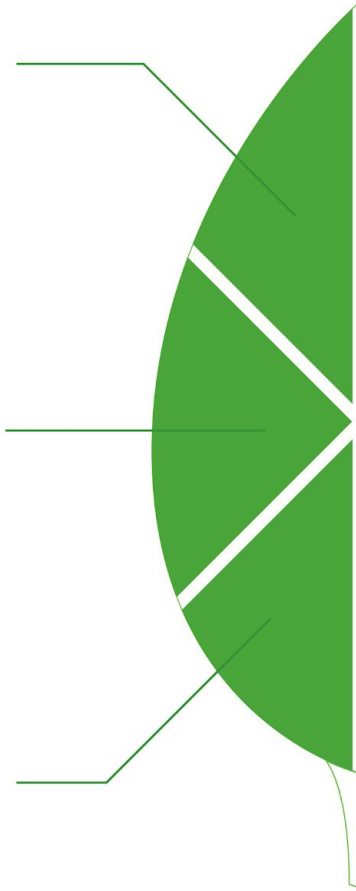
Use a Columnar NoSQL Database if your application has

- Queries that involve only a few columns
- Aggregation queries against vast amounts of data
- Column-wise compression

Expressive
Query Language

Strong
Consistency

Secondary
Indexes

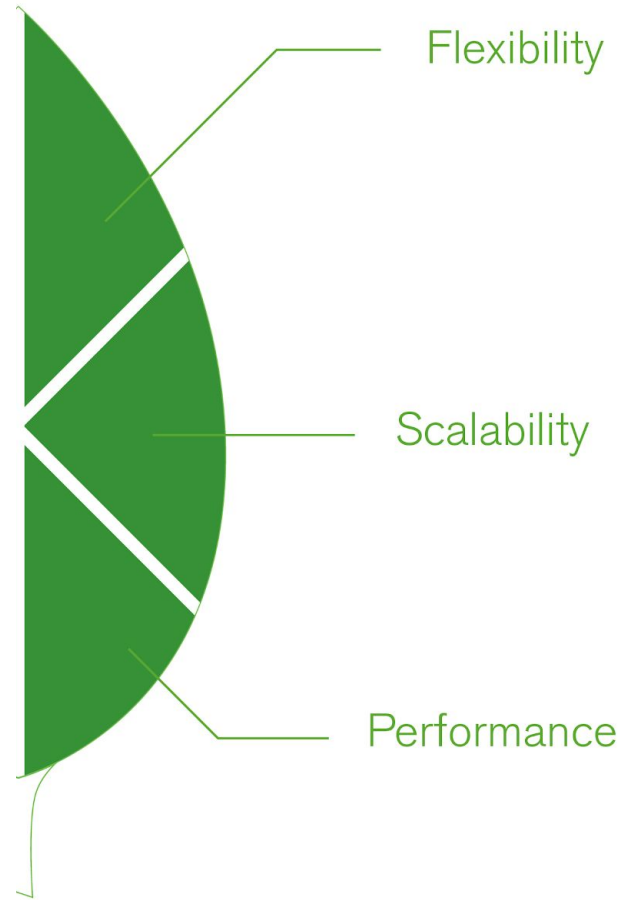


STRENGTHS OF **RELATIONAL** DATABASES

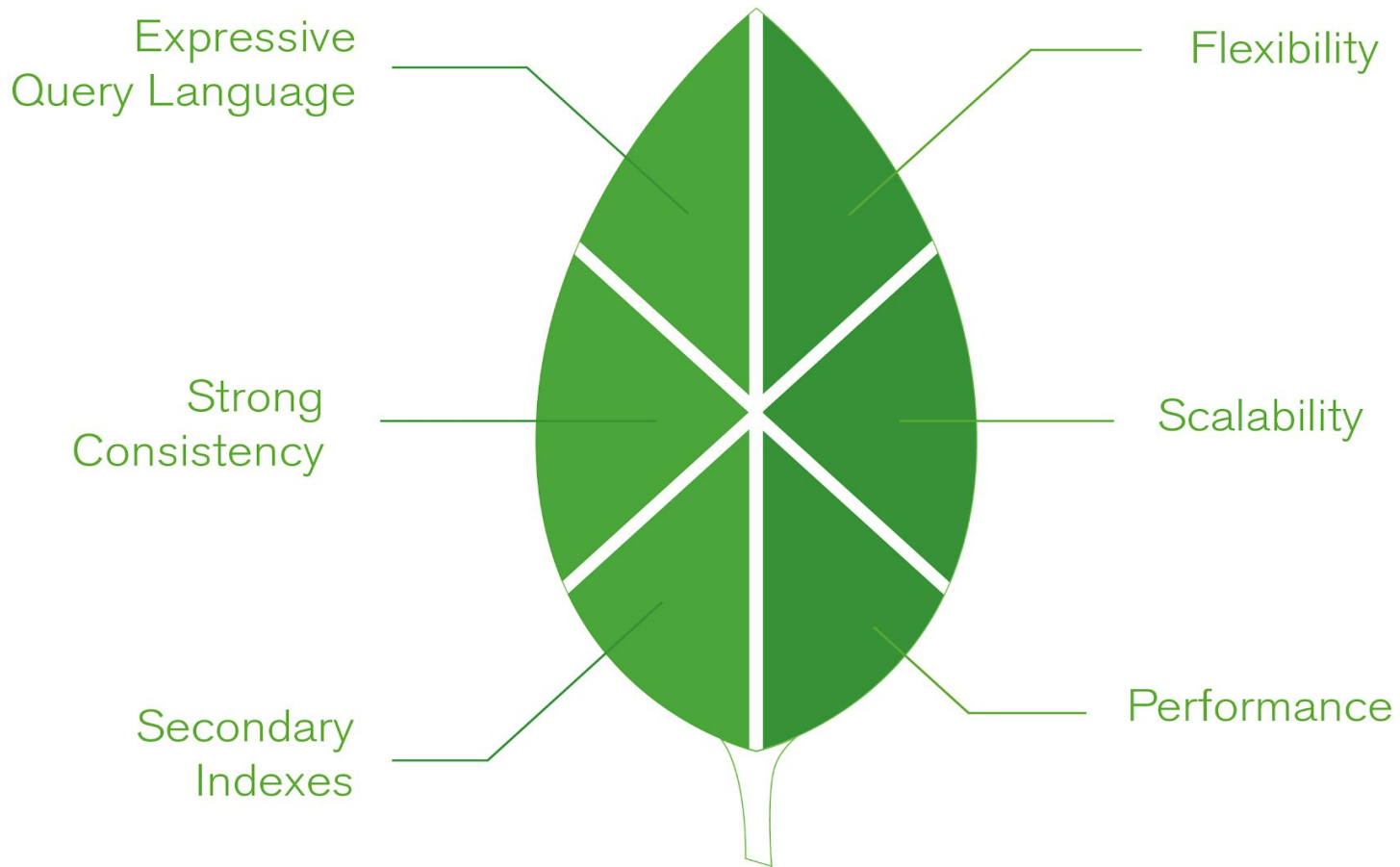




STRENGTHS OF NoSQL DATABASES



mongoDB NEXUS ARCHITECTURE



WHAT IS mongoDB?



WHAT IS mongoDB?

mongoDB is a NoSQL Database 






WHAT IS mongoDB?

mongoDB is a NoSQL Database 
that stores data 



WHAT IS mongoDB?

mongoDB is a NoSQL Database 
that stores data 
in the form of documents 

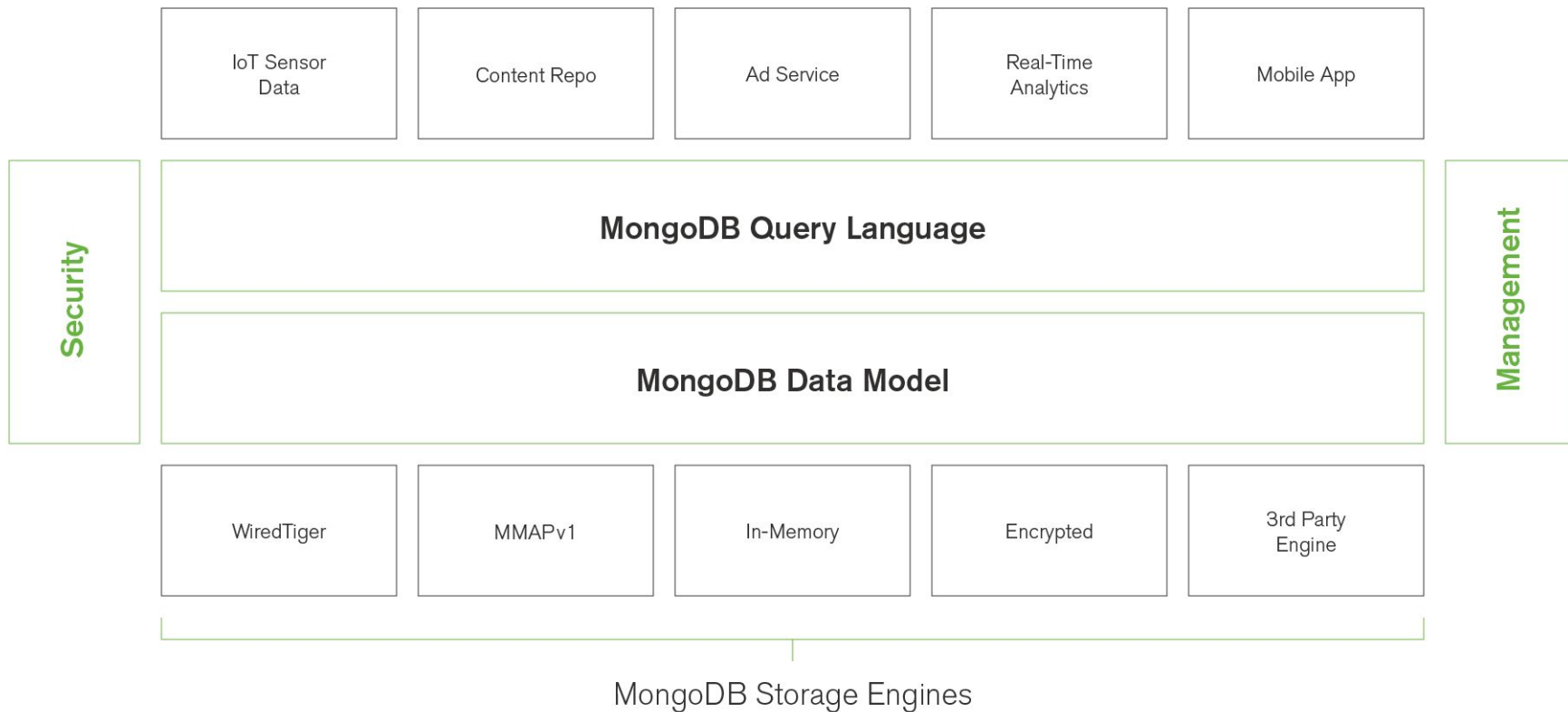


WHAT IS mongoDB?

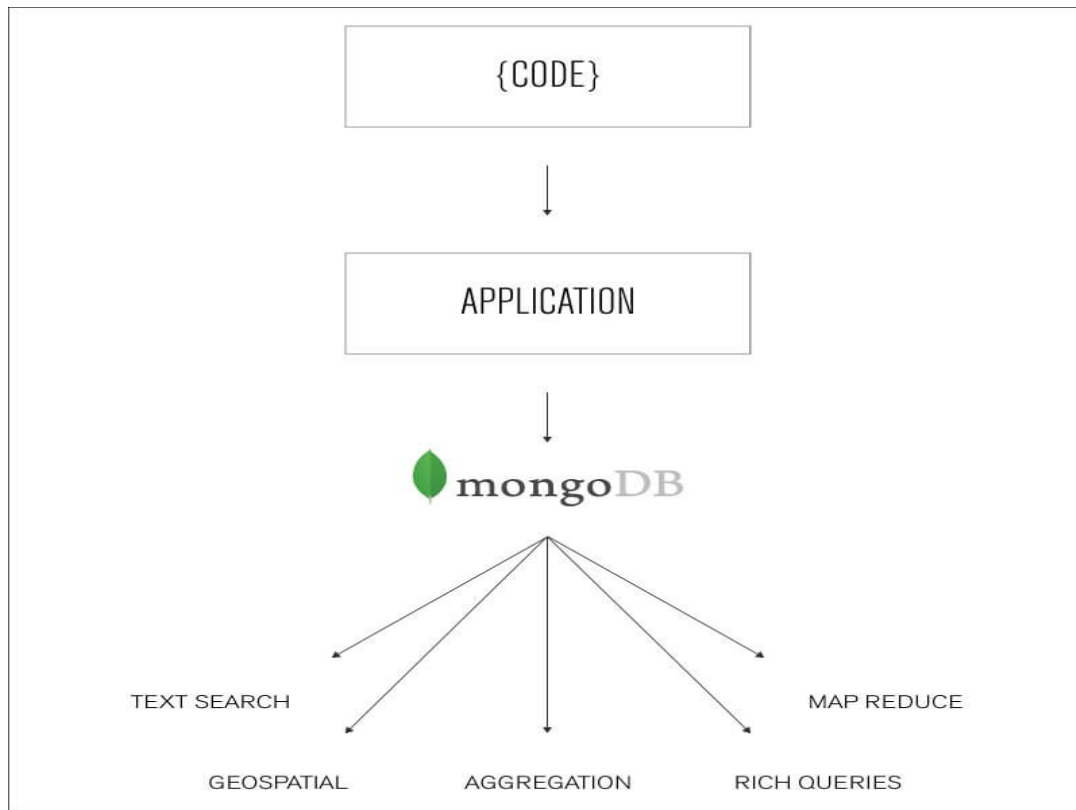
mongoDB is a NoSQL Database 
that stores data 
in the form of JSON documents 



PLUGGABLE STORAGE ARCHITECTURE

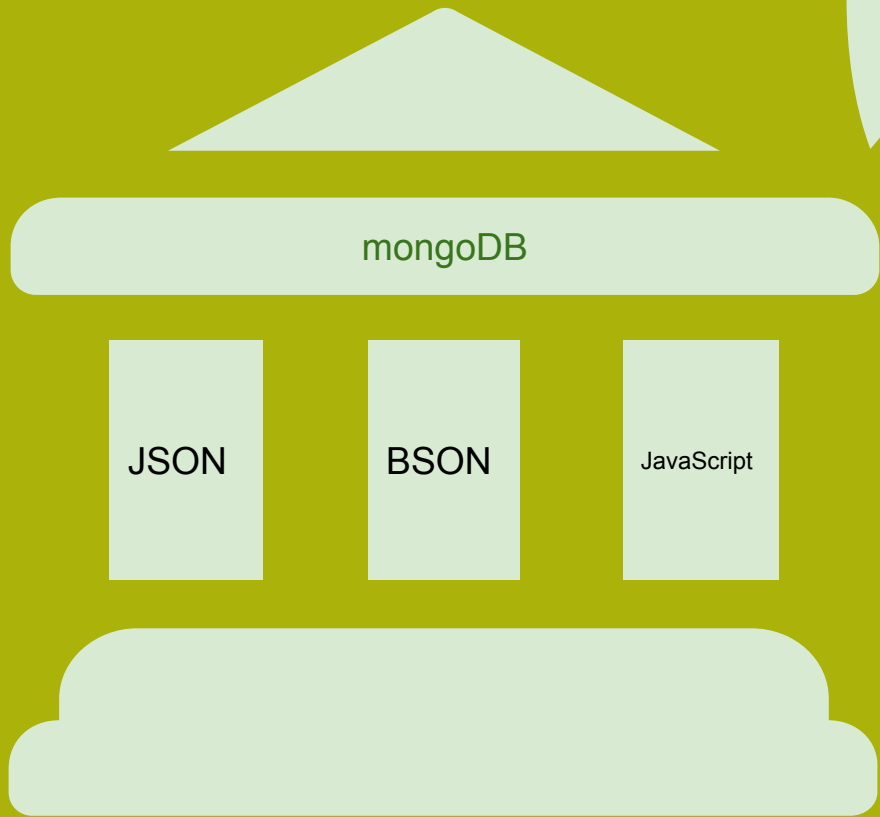


SUPPORT FOR **AD HOC** QUERIES



	MongoDB	Relational	Key-value
Key-value Queries	Yes	Yes	Yes
Secondary Indexes	Yes	Yes	No
Index Intersection	Yes	Yes	No
Range Queries	Yes	Yes	No
Geospatial	Yes	Expensive Add-on	No
Text Search	Yes	Expensive Add-on	No
Aggregation	Yes	Yes	No
MapReduce	Yes	No	Yes
Idiomatic Drivers	Yes	No	No
Left Outer JOINS (\$Lookup)	Yes	Yes	No

PILLARS OF mongoDB





BASIC JSON

```
{“Name”:”Ramesh”}  
{“Name”:”Ramesh”, “Location”:”Hyderabad”}  
{“Name”:”Ramesh”, “Location”:”Hyderabad”, “Years of experience”:20}  
{“Name”:”Ramesh”, “Skills”: [“MongoDB”, “MySQL”, “Neo4j”]}
```

I/O PERFORMANCE

L1 cache reference	0.5 ns!
Branch mispredict	5 ns!
L2 cache reference	7 ns!
Mutex lock/unlock	25 ns!
Main memory reference	100 ns!
Compress 1K bytes with Zippy	3,000 ns!
Send 2K bytes over 1 Gbps network	20,000 ns!
Read 1 MB sequentially from memory	250,000 ns!
Round trip within same datacenter	500,000 ns!
Disk seek	10,000,000 ns!
Read 1 MB sequentially from disk	20,000,000 ns!
Send packet CA→Netherlands→CA	150,000,000 ns

UNDERSTANDING DISK LOCALITY

RDBMS table JOIN

VS

MongoDB JSON Document access



02. **mongoDB** **CLIENTS**

Get to know the big players whose use mongoDB



CLIENTS e-commerce



Multinational eCommerce store delivering all media metadata at 99.999% availability.



A leading clothing and accessories retailer was able to develop a new MongoDB-based purchase order system in just 75 days, a record for the company.

CLIENTS e-commerce



Online speciality retailer powers its ecommerce platform with MongoDB

The Chico's logo consists of the word "chico's" in a bold, lowercase, sans-serif font. The letters are closely spaced, and the apostrophe in "chico's" is positioned between the 'i' and 'c'.

Women's retail apparel relaunched its software application on MongoDB to scale for holiday demand and provide a highly engaging shopping experience.

CLIENTS GOVERNMENT



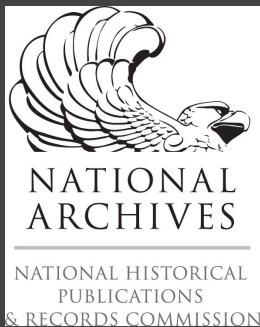
Government agency moves from wireframe to production app with MongoDB in weeks, not months or years



Driver & Vehicle
Licensing
Agency

UK's Driver & Vehicle Licensing Agency uses MongoDB's scalability to digitise vital driving license data and make it available online for the first time.

CLIENTS GOVERNMENT



The National Archives use MongoDB to digitize First World War diaries and other documents from British history.



MongoDB's flexibility powers solar weather forecasting for the UK Met Office.

CLIENTS TECHNOLOGY



The world's best-selling sports video game franchise relies on MongoDB to scale to millions of players



Adobe

World's #1 Content Management solution relies on MongoDB for petabyte scale data management in the cloud.

CLIENTS TECHNOLOGY



Worldwide leader in networking powers its collaborative workspace on MongoDB



MongoDB is the core database underpinning SAP's Platform-as-a-Service content management system.

CLIENTS FINANCIAL



Leading consumer, corporate and investment bank replaces three decades of relational databases with MongoDB, for increased agility, scalability, and cost-efficiencies.



MongoDB supports the global bank's enterprise data service which is underpinning several core trading systems

CLIENTS FINANCIAL



MongoDB is the database powering the Loop accounting suite, used by KPMG's 4,800 CPAs.



HM Revenue
& Customs

UK government delivers Tax Platform that cuts wait times in half, serves 9.2m people and saves £8m in operating costs

mongoDB IN NUMBERS



Since its first MongoDB project in 2012, Baidu has grown its cluster to 600 nodes storing 200 billion documents and 1PB of data, powering over 100 apps



Amadeus is the world's largest technology company dedicated to the travel industry, handling 630 million bookings and 16 billion passengers every year.

The company's travel systems rely on over 100 MongoDB clusters, totaling 570+ nodes, with the largest cluster managing over 100TB of data.

mongoDB IN NUMBERS

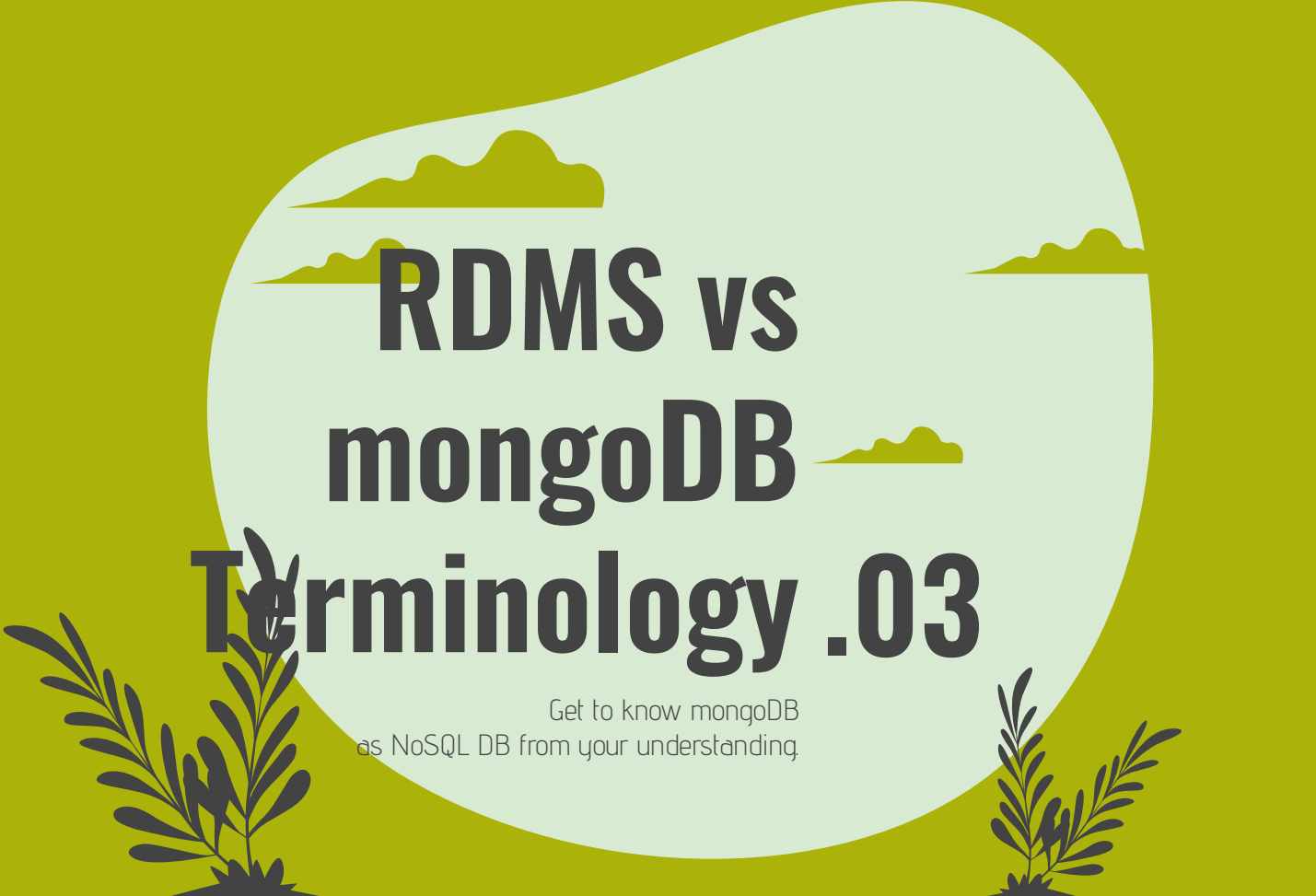


To serve millions of players, EA's Spearhead development studio selected MongoDB to store user data and game state. Auto-sharding makes it simple to scale MongoDB across EA's 250+ servers with no limits to growth as EA FIFA wins more fans.



Foursquare is used by over 50 million people worldwide, who have checked in over 6 billion times, with millions more added every day.

MongoDB is Foursquare's main database, supporting hundreds of thousands of operations per second and storing all check-ins and history, user, and venue data along with reviews.

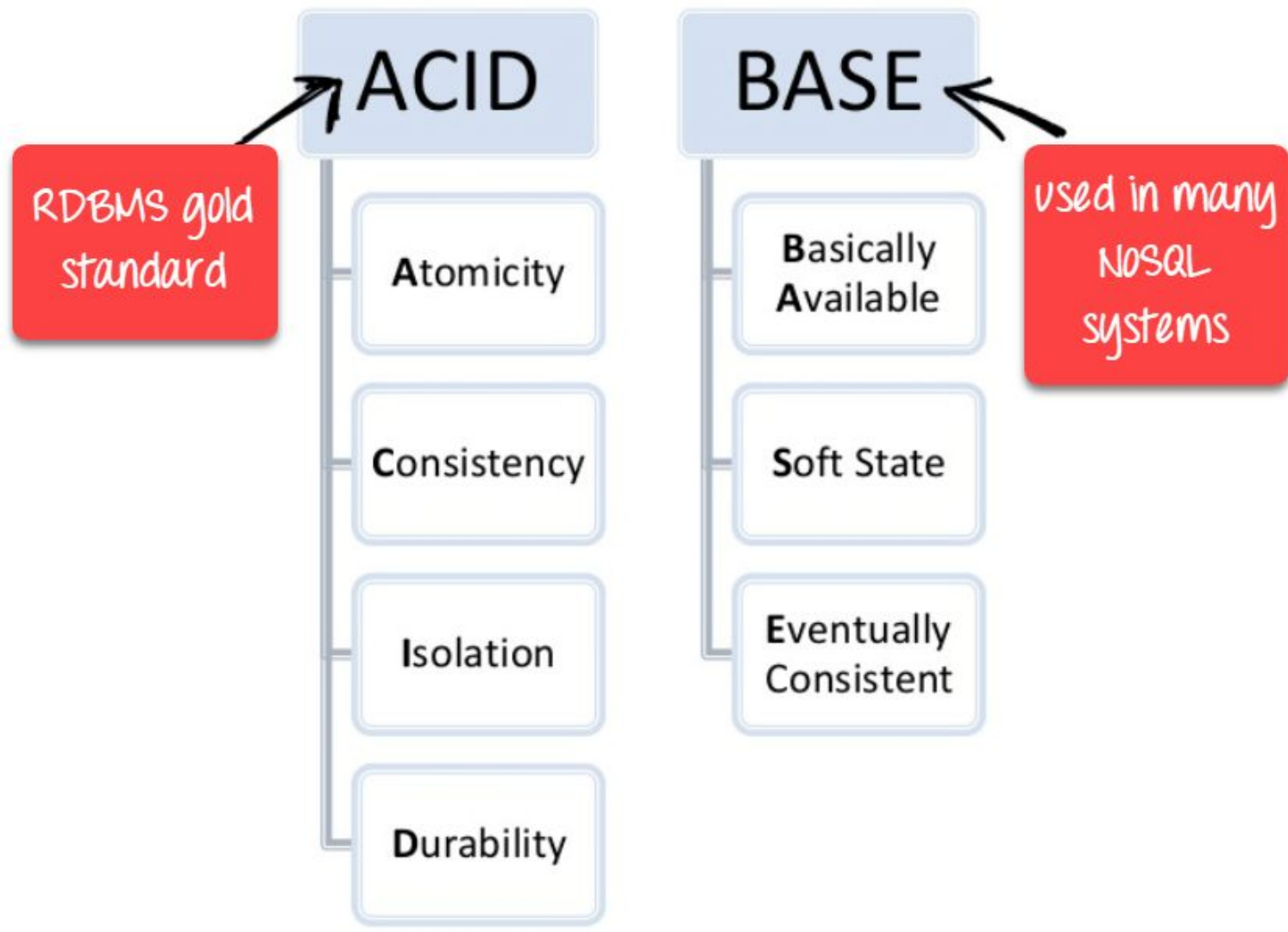
The background is a solid olive green. In the center is a large, light green, irregularly shaped circle. Inside this circle are stylized, dark green cloud-like shapes. At the bottom left and right corners of the circle, there are dark green, stylized leafy branches. The text is centered within the circle.

RDMS vs mongoDB Terminology .03

Get to know mongoDB
as NoSQL DB from your understanding.

RDBMS vs MongoDB Terminology

RDBMS		MongoDB
Database	→	Database
Table, View	→	Collection
Row	→	Documents (JSON, BSON)
Column	→	Field
Index	→	Index
Join	→	Embedded Document
Foreign Key	→	Reference
Partition	→	Shard





04. **mongoDB** **BINARIES**

See what mongoDB brings out-of-box



mongod

mongod is the primary daemon process for the MongoDB system.

It handles data requests, manages data access, and performs background management operations.

mongo

MongoDB offers an interactive shell called **mongo** which lets developers view, insert, remove, and update data in their databases, as well as get replication information, set up sharding, shut down servers, execute JavaScript, and more.

Administrative information can also be accessed through a web interface, a simple webpage that serves information about the current server status.

By default, this interface is 1000 ports above the database port (28017).

mongostat

mongostat is a command-line tool that displays a summary list of status statistics for a currently running MongoDB instance: how many inserts, updates, removes, queries, and commands were performed, as well as what percentage of the time the database was locked and how much memory it is using.

This tool is similar to the UNIX/Linux **vmstat** utility.

mongotop

mongotop MongoDB instance spends reading and writing data. mongotop provides statistics on the per-collection level.

By default, mongotop returns values every second.

This tool is similar to the UNIX/Linux **top** utility.

mongostat

mongostat is a command-line tool that displays a summary list of status statistics for a currently running MongoDB instance: how many inserts, updates, removes, queries, and commands were performed, as well as what percentage of the time the database was locked and how much memory it is using.

This tool is similar to the UNIX/Linux **vmstat** utility.

mongoimport, mongoexport

`mongoimport` is a command-line utility to import content from a JSON, CSV, or TSV export created by `mongoexport` or potentially other third-party data exports.

mongodump, **mongorestore**

mongodump is a command-line utility for creating a binary export of the contents of a Mongo database;

mongorestore can be used to reload a database dump



mongoDB

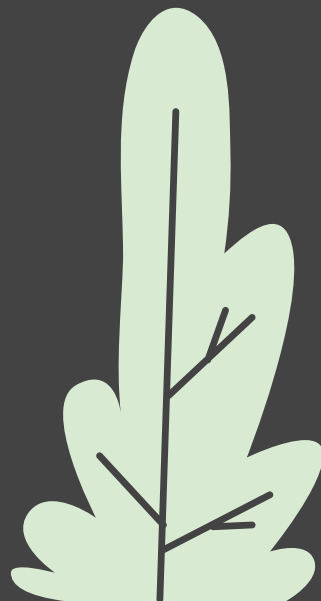
05. DATA TYPES

Data Types used in mongoDB

mongodb is

Case Sensitive

Type Sensitive



“true”

**WHAT TYPE
IS THIS?**

“1”

**WHAT TYPE IS
THIS?**

false

**WHAT TYPE
IS THIS?**

```
{“married”:true,  
  “happy”:true}
```

**WHAT TYPE IS
THIS?**

BEWARE!

```
1  != "1"  
2.05 != "2.05"  
true != "true"
```





DATA TYPES - FOR 4.2



`null`

Null can be used to represent both a null value and a nonexistent field:

```
{"x" : null}
```

DATA TYPES - FOR 4.2

boolean

There is a boolean type, which will be used for the values
'true' and 'false':

```
{"x" : true}
```





DATA TYPES - FOR 4.2



numberInt

```
{"x" : 3}
```

numberLong (64-bit integer)

```
{"x" : 246011064}
```

DATA TYPES - FOR 4.2

numberDecimal (64-bit floating point number)

```
{"x" : 3.1456}
```





DATA TYPES - FOR 4.2



numberInt

```
{"x" : 3}
```

numberLong (64-bit integer)

```
{"x" : 246011064}
```

DATA TYPES - FOR 4.2

object id

An object id is a unique 12-byte 1D for documents.

```
{"x" : ObjectId() }
```





DATA TYPES - FOR 4.2

date

Dates are stored as milliseconds since the epoch. The time zone is not stored:

```
{"x" : new Date() }
```

DATA TYPES - FOR 4.2

regular expression

Documents can contain regular expressions, using JavaScript's regular expression

```
{"x" : /fivestar/i}
```





DATA TYPES - FOR 4.2



code

Documents can also contain JavaScript code:

```
{"x" : function() { /* ... */ }}
```

DATA TYPES - FOR 4.2

binary data

Binary data is a string of arbitrary bytes



DATA TYPES - FOR 4.2

array

Sets or lists of values can be represented as arrays:

```
{"x" : ["a", "b", "c"]}
```





DATA TYPES - FOR 4.2

embedded document

Documents can contain entire documents, embedded as values in a parent document:

```
{"x" : {"five" : "star"}}
```

AWESOME: ObjectId

BSON (binary-encoded JSON)

Unique: combination of

- ❖ Time
- ❖ Machine
- ❖ Process id
- ❖ Incrementer

Mongo will automatically add this `_id` field

Language-specific drivers automatically add this, too (BEFORE saving!)

Your documents get a timestamp for free



MongoDB Supported **BSON** Data Types

Type	Number	Alias	Notes
Double	1	“double”	
String	2	“string”	
Object	3	“object”	
Array	4	“array”	
Binary data	5	“binData”	
Undefined	6	“undefined”	Deprecated.
ObjectId	7	“objectId”	

MongoDB Supported **BSON** Data Types

Type	Number	Alias	Notes
Boolean	8	“bool”	
Date	9	“date”	
Null	10	“null”	
Regular Expression	11	“regex”	
DBPointer	12	“dbPointer”	Deprecated
JavaScript	13	“javascript”	
Symbol	14	“symbol”	Deprecated

MongoDB Supported **BSON** Data Types

Type	Number	Alias	Notes
JavaScript (with scope)	15	“javascriptWithScope”	
32-bit integer	16	“int”	
Timestamp	17	“timestamp”	
64-bit integer	18	“long”	
Decimal128	19	“decimal”	New in 3.4
Min key	-1	“minKey”	
Max key	127	“maxKey”	

Executing a JavaScript file:

1. You can specify a `.js` file to the mongo shell, and mongo will execute the JavaScript directly

```
mongo localhost:27017/test myjsfile.js
```

2. Execute a `.js` file from within the mongo shell, using the `load()` function

```
load("myjstest.js")
```

How to Run JavaScript in MongoDB

Shell Helpers	JavaScript Equivalents
show dbs, show databases	db.adminCommand('listDatabases')
use <db>	db = db.getSiblingDB('<db>')
show collections	db.getCollectionNames()
show users	db.getUsers()
show log <logname>	db.adminCommand({ 'getLog' : '<logname>' })
show logs	db.adminCommand({ 'getLog' : '*' })
it	<pre>cursor = db.collection.find() if (cursor.hasNext()){ cursor.next(); }</pre>


How to Run JavaScript in MongoDB

Scripts for the mongo shell manipulate data in MongoDB or perform administrative operations

Eg. To instantiate a new connection:

```
conn = new Mongo();  
db = connect("localhost:27020/myDatabase");
```

How to Run JavaScript in MongoDB



mongoDB STORAGE 06. ENGINES

Know the types of Storage Engines mongoDB
uses



WiredTiger Storage Engine

Starting in MongoDB 3.2, the WiredTiger storage engine is the default storage engine

Features:

- Document Level Concurrency
- Snapshots and Checkpoints
- Journaling
- Compression
- Memory Usage

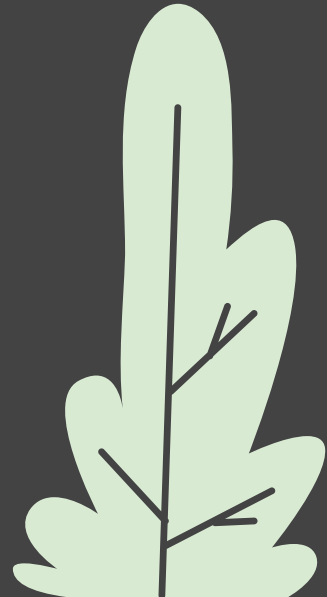
MMAP Storage Engine

Memory Mapped Storage Engine uses memory mapped files as its storage engine

Memory mapped files allow MongoDB to delegate the handling of Virtual Memory to the operating system instead of explicitly managing memory itself

Why MMAP is Deprecated?

Doesn't support Data Compression
Encryption is not possible



Storage engine comparison

	MongoDB WiredTiger	MongoDB MMAPv1
Write Performance	Excellent Document-Level Concurrency Control	Good Collection-Level Concurrency Control
Read Performance	Excellent	Excellent
Compression Support	Yes	No
MongoDB Query Language Support	Yes	Yes
Secondary Index Support	Yes	Yes
Replication Support	Yes	Yes
Sharding Support	Yes	Yes
Ops Manager & MMS Support	Yes All features including deployment, upgrade backup, restore, and monitoring	Yes All features including deployment, upgrade backup, restore, and monitoring
Security Controls	Yes	Yes
Platform Availability	Linux, Windows, Mac OS X	Linux, Windows, Mac OS X, Solaris (x86)



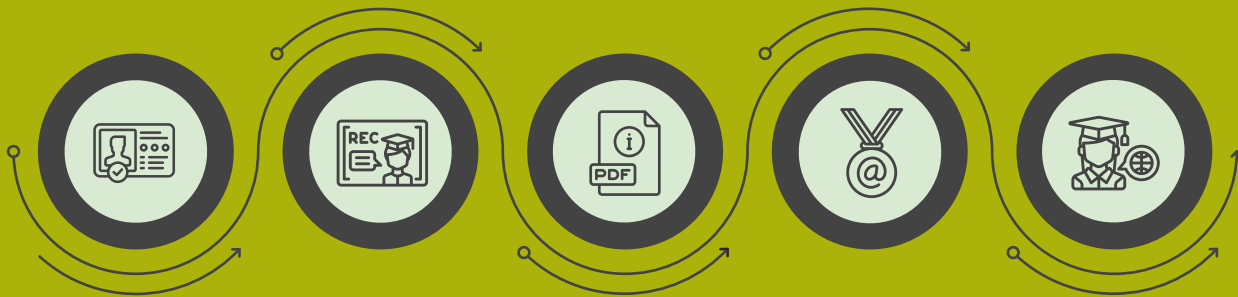
A TIMELINE ALWAYS WORKS FINE

MARS

Despite being red, Mars is a cold place, not hot. It's full of iron oxide dust

VENUS

Venus has a beautiful name and is the second planet from the Sun



Mercury is the closest planet to the Sun and is only a bit larger than our Moon

MERCURY

Neptune is the farthest planet from the Sun and the fourth-largest by diameter

NEPTUNE

Saturn is the ringed one. It's composed mostly of hydrogen and helium

SATURN



THANKS!

Does anyone have any questions?

rsannareddy@gmail.com

+91 98859 70033