WHAT DATABASE SHOULD I USE

By Boyang Bai

INTRODUCTION

Relational Database

- Long history and most widely used, many choices
- Using relational model to store data, e.g. entities and relations
- Data need to be structured
- Adopting SQL like query language

Non-relational Database

- NoSql, not only sql, eliminating the use of the standardized SQL
- More scalable and providing superior performance
- Schema free, using various models to store data, good for unstructured data
- Models: Key/value, Column, Document, Graph

RDBMS CHOICES

SQLite
https://sqlite.org/

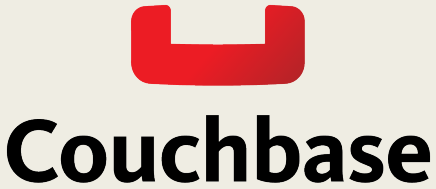
- Advantage:
  - *Self-contained, file-based database relational database*
  - *Standard SQL syntax*
- Disadvantage:
  - *No user management*
  - *Low scalability, concurrency*
- When to use:
  - *Single user light weight application (400K to 500K HTTP requests per day)*
  - *Local test*

Cassandra
http://cassandra.apache.org/

- Advantage:
  - *Powerful for big data, high scalability/availability*
- Disadvantage:
  - *Quite complex that need expert knowledge to maintain*
  - *Consistency not guaranteed*
- When to use:
  - *Handling extremely huge throughput (First for FB inbox)*
  - *Low consistency requirement*

Couchbase

http://www.couchbase.com/



- Advantage:
  - *N1QL, similar to SQL*
  - *Built-in in-memory cache functionality*
- Disadvantage:
  - *Consistency not guaranteed*
- When to use:
  - *Familiar with SQL but want high performance*
  - *in-memory cache together with data storage (Mango+Redis)*

Couchbase

http://www.couchbase.com/

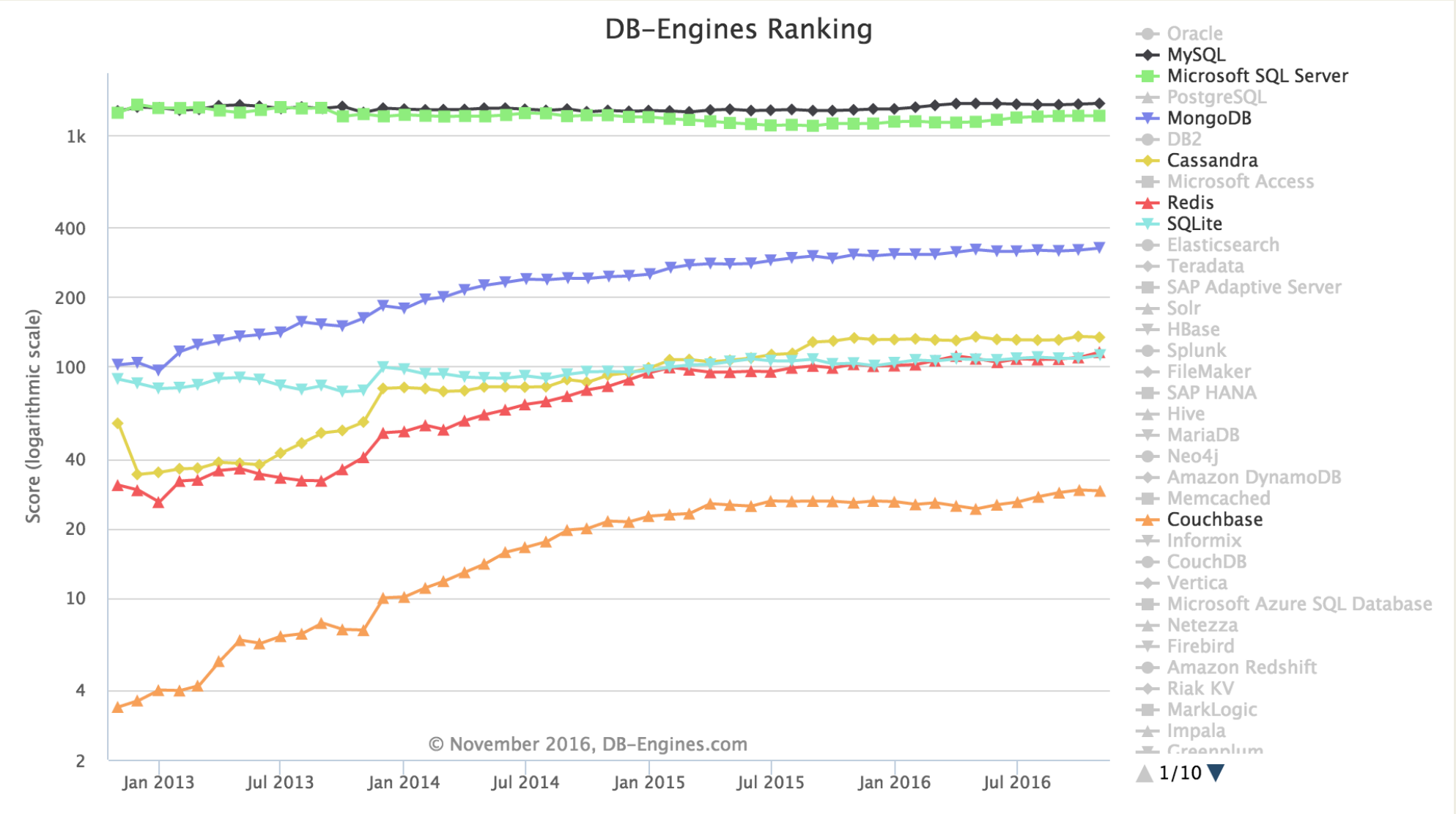
SQL

```
SELECT breweries.name AS brewery,
   count(*) AS cnt
FROM beers
INNER JOIN breweries
ON beer.brewery_id = breweries.id
WHERE beers.type = "beer" AND
   breweries.type = "brewery" AND
   beers.style = "American-Style Imperial Stout"
GROUP BY breweries.name
HAVING count(*) > 2
ORDER BY cnt DESC;
```

N1QL

```
SELECT breweries.name AS brewery,
   count(*) AS cnt
FROM `beer-sample` beers
INNER JOIN `beer-sample` breweries
ON KEYS beers.brewery_id
WHERE beers.type = "beer" AND
   breweries.type = "brewery" AND
   beers.style = "American-Style Imperial Stout"
GROUP BY breweries.name
HAVING count(*) > 2
ORDER BY cnt DESC;
```

Which Database is more popular?



THANKS!