

# SQL for Data Science

---

## Module 1: Introduction to Databases

*In this module you will learn about some fundamentals of databases and why we need them*

- **Why do we need databases?**
- **What is a database?**
- **Some properties of a Good Database**
- **Types of databases – Relational Vs NOSQL databases**
- **Companies using MySQL**

# Why do we need databases?

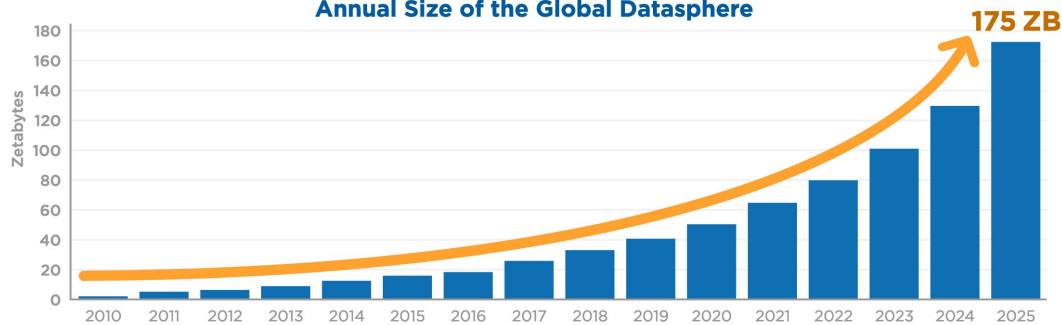
*The rapid explosion of data creates a necessity to have a place to store and retrieve data efficiently*

## The age of Data



## The Global Data Explosion

Annual Size of the Global Datasphere



Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018

## Why we need databases?

**Store**

**Update**

**Delete**

**Search**

*The rapid explosion of data creates a necessity to have a place to store and retrieve data efficiently*

## Quick Wikipedia Search

### Database

From Wikipedia, the free encyclopedia

*For a topical guide to this subject, see [Outline of databases](#).*

A **database** is an organized collection of [data](#), generally stored and accessed electronically from a computer system. Where databases are more complex they are often developed using formal [design and modeling](#) techniques.

### Important Concepts

- **It's a collection of data**
- **Its Organized**
- **Stored and Accessed Electronically**

## Scalability



## Elasticity



## Security & Governance



## Data Integrity



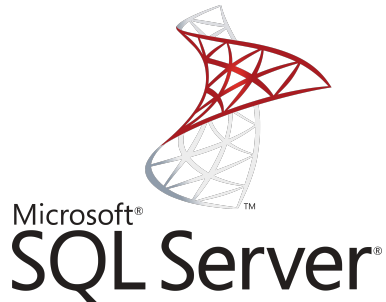
# Types of Databases – Relational vs NoSQL

*Examples of Relational and NoSQL Databases used in the industry*

## Relational Databases



ORACLE



## NoSQL Databases



*The data in Relational databases is stored in rows and columns*

## Database Structure

**Table 1: UserDetails**

User id	Name	Skill Id
1	Anand	10
2	Dimple	20



**UserData**

**Table 2: SkillMapping**

Skill Id	Skills
10	Machine Learning, Python
20	Marketing, Google Analytics

## How data is stored in NoSQL databases

*The NoSQL way of defining tables is much more flexible and uses JSON*

```
{  
  "id":1,  
  "name":"Anand",  
  "skills":["Machine Learning", "Python"]  
},  
  
{  
  "id":2,  
  "name":"Dimple",  
  "skills":["Marketing", "Google Analytics"],  
  "education":{  
    "Graduation": { "School": "IIT Delhi",  
                    "CGPA": 9.0  
                  }  
  }  
}
```



## Companies using MySQL (Continued)

**80% of the companies worldwide are using MySQL.** A list of selected companies using MySQL

Source: <https://www.mysql.com/customers/industry/>

### Aerospace



### Financial Services



### Government



### Manufacturing



## Companies using MySQL (Continued)

*80% of the companies worldwide are using MySQL. A list of selected companies using MySQL*

Source: <https://www.mysql.com/customers/industry/>

### Travel

**Booking.com**

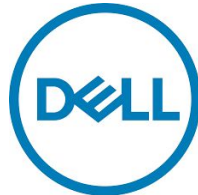


**Lufthansa**

### Media & Entertainment

**BBC** **NETFLIX**

### Technology: Hardware & Software



**IBM**



**Microsoft**

### Internet companies



# Thank you slide

---