



NoSQL Databases

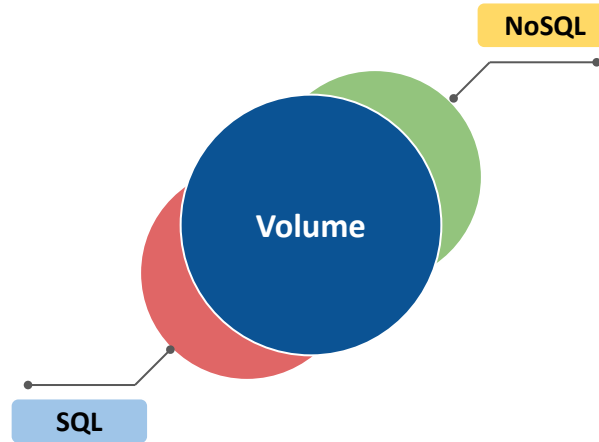
This file is meant for personal use by maneel.chauhan@coforge.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Agenda

- Why SQL is not enough?
- Problem Definition
- Learning Journey
- Setting up MongoDB Compass
- Data Understanding
- Exploratory Data Analysis
- Problem Solution
- Summary

Why SQL databases are not enough?

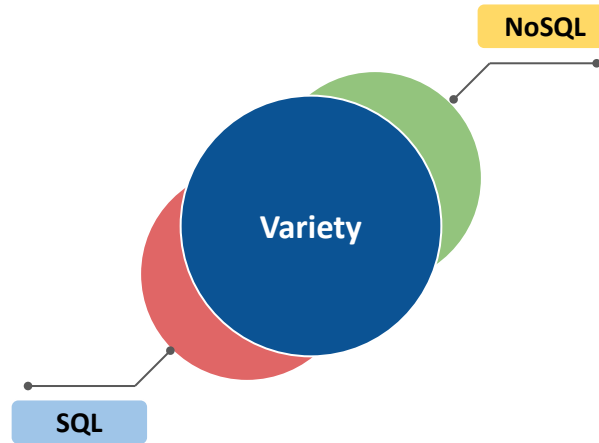
Traditional SQL databases like MySQL are not based out of the cloud. So scaling when the volume of data increases is not possible.



MongoDB or other NoSQL databases are horizontally scalable when the load increases - ie. you can add more resources to increase the storage and processing capacity when you need it.

Why SQL databases are not enough?

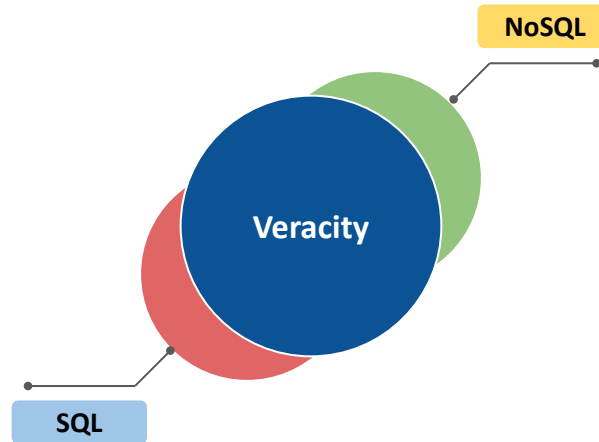
SQL databases are not efficient with handling semi-structured and unstructured data as the schema or the structure of data can keep changing over time & it is not an easy task for us make sure the defined schemas keep up with it!



NoSQL databases offer a more flexible path to handling semi-structured and unstructured data. They are highly programmatic in nature so you can dynamically evolve the schema of your tables as the data changes.

Why SQL databases are not enough?

Today's data can be inconsistent. SQL databases cannot deal with these uncertainties. Once you define a DDL, you have to input data in the same format and cannot keep changing it



NoSQL does not need a schema to be defined in the first place to store data. You can dynamically change things on the go to suit your analytical needs

Problem Definition

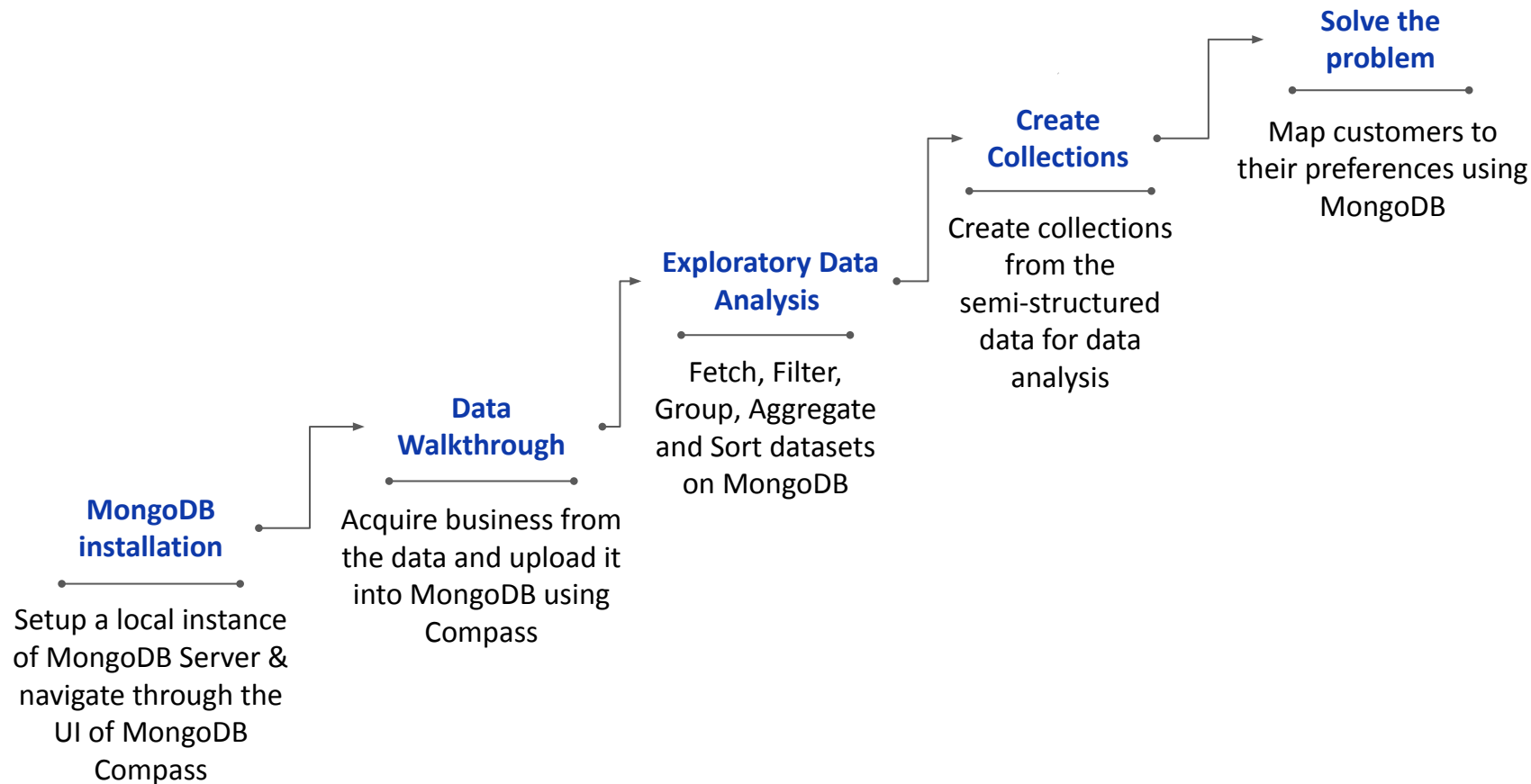
Business Context:

G-Flix is an OTT platform which has millions of customers across ~190 countries. In the past 2 months they have acquired 1 million new customers for their platform. As part of analytics initiative G-Flix wants to recommend movies & TV shows from the pool of ~8800 records from their library.

Objective:

As a Data Scientist your objective is to understand the customer preferences and map them with the Movies & TV Shows in our library, based on data acquired on the customer related to genre and artist they prefer and finally provide the right recommendations.

Learning Journey

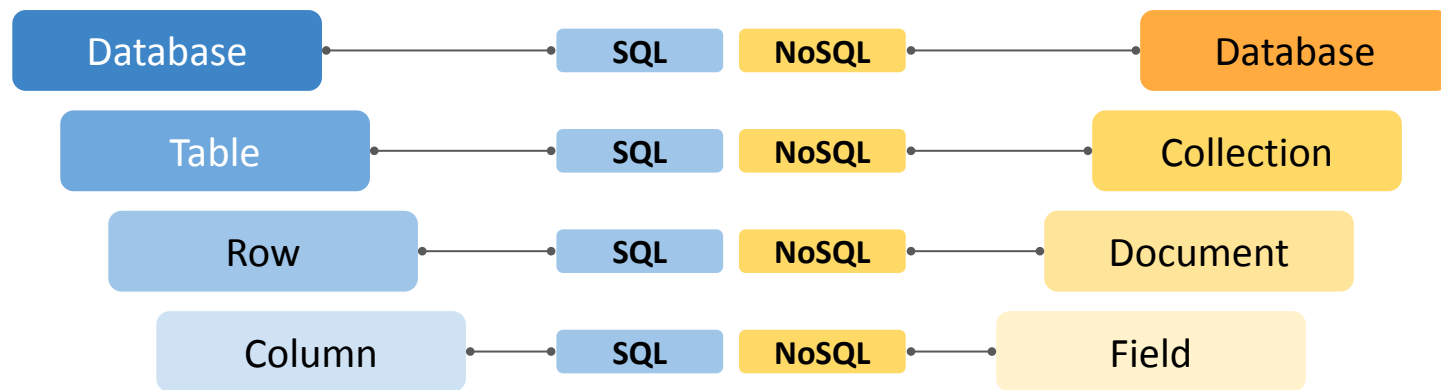




Databases, Collections, and Documents

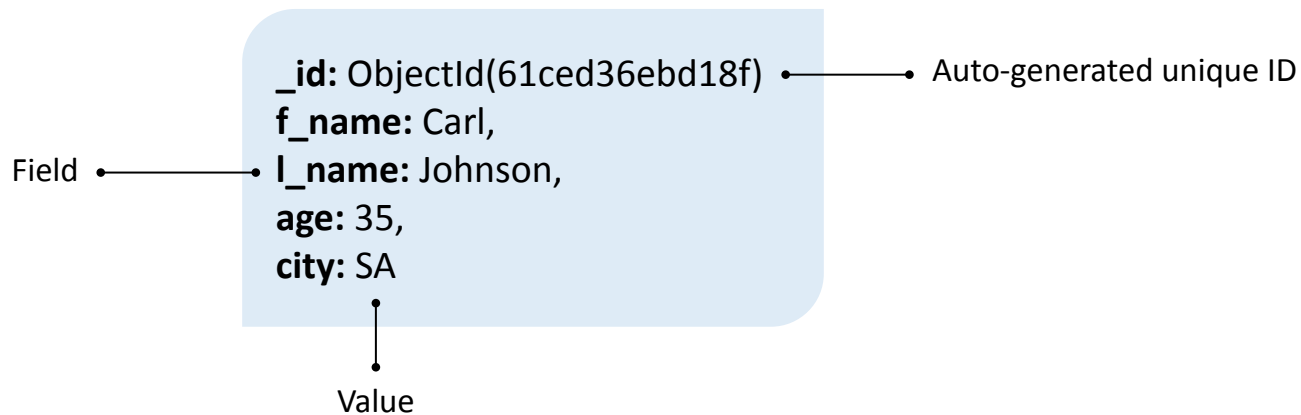
This file is meant for personal use by maneel.chauhan@coforge.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Mapping terms from SQL-DB to MongoDB



Documents in MongoDB

A sample **Document**



Collections and Databases in MongoDB

A sample **Collection**

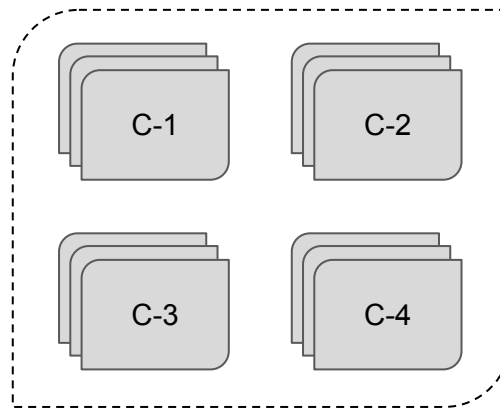
_id: ObjectId(61ced36ebd18f)
f_name: Carl,
l_name: Johnson,
age: 35,
city: SA

_id: ObjectId(61ced36ebd18f)
f_name: Tommy,
l_name: Vercetti,
age: 40,
city: VC

_id: ObjectId(61ced36ebd18f)
f_name: Niko,
l_name: Bellic,
age: 30,
city: LC



Multiple collections make a **Database**

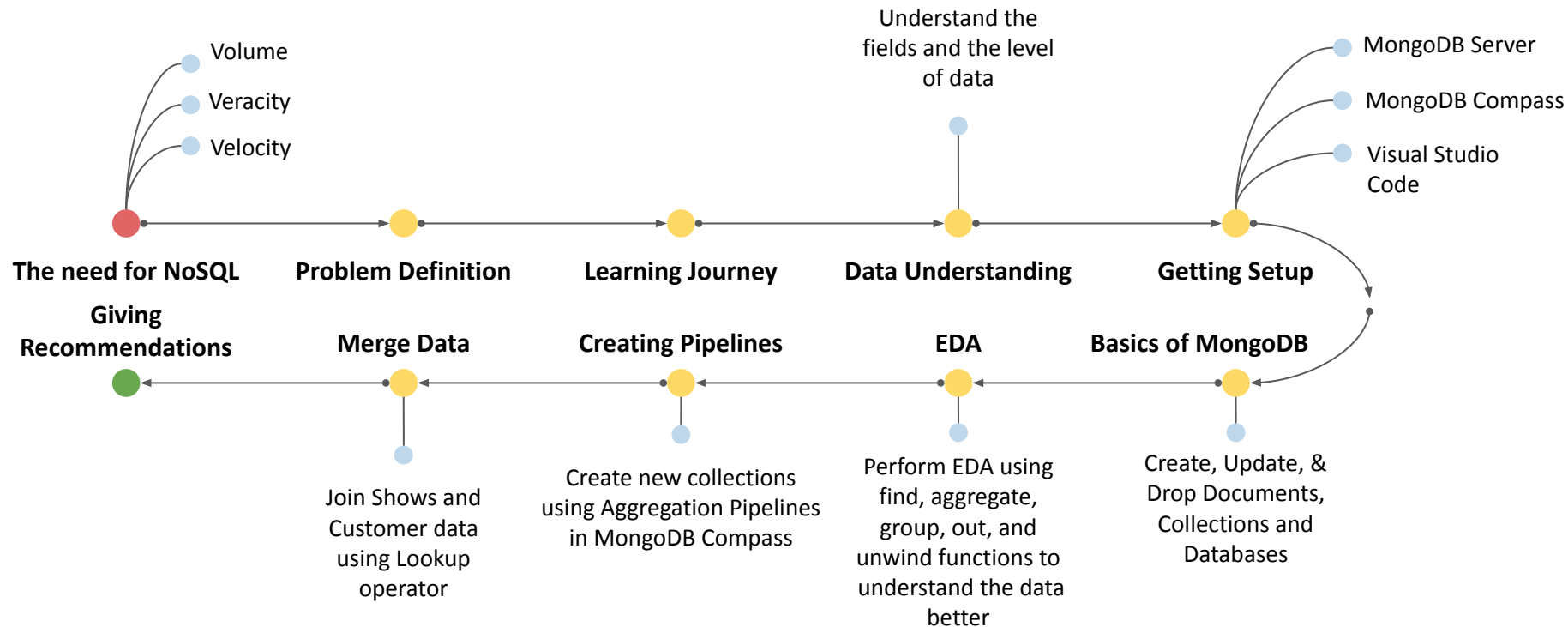


...

...This file is meant for personal use by maneel.chauhan@coforge.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Summary





Happy Learning !

