NoSQL
Business Meeting

Indu Abhilash, Aryana Far, Jun Park

Table of Contents

01

Introduction

Proposing NoSQL for AGM’s future

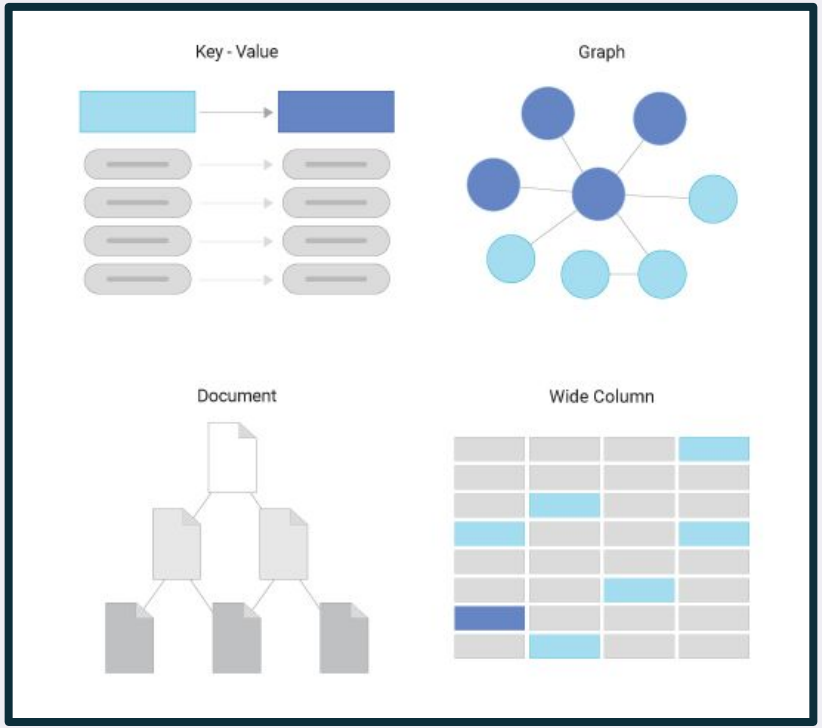
A Data-Driven Future

As seasoned data engineers at AGM, we are proposing some new and exciting business ideas that will leverage NoSQL databases to propel the future of our company.

Why NoSQL?

Here is what NoSQL permits:

- Schema-less database allowing for unstructured data
- Supports various data models
  - document-oriented (MongoDB)
  - key-value pairs (Redis)
  - graph databases (Neo4j).
- Good for complex and inconsistent data



02

Bart–Truck Hybrid Delivery

Leveraging Neo4j and Redis

Bart-Truck Hybrid Delivery

Neo4j

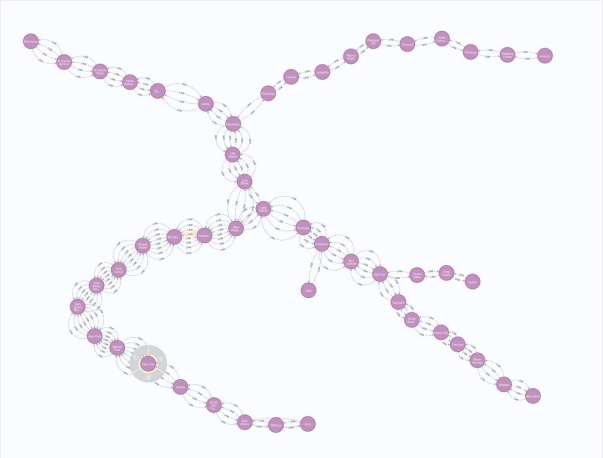
- Bart lines are best represented as connected graph (Neo4j)
- Each node represent a bart station
- Each edge represents a bart line connecting two stations
- Weight of an edge represents time to get from one station to another

Redis

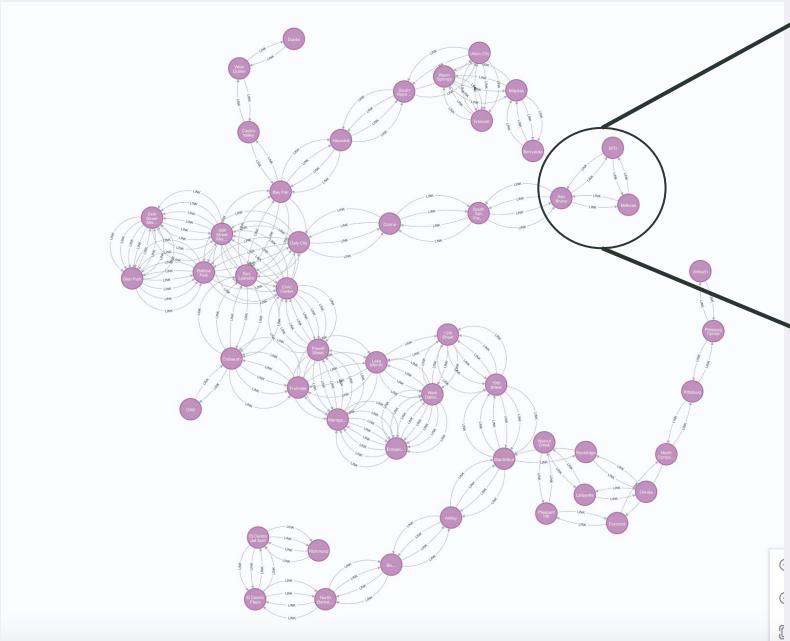
- Using Redis for more instantaneous data analytics for identifying ideal local delivery routes
- Traffic data updated every minute within local truck delivery routes
- Real time traffic data/analysis
- Quicker delivery times translate to more deliveries and more profits

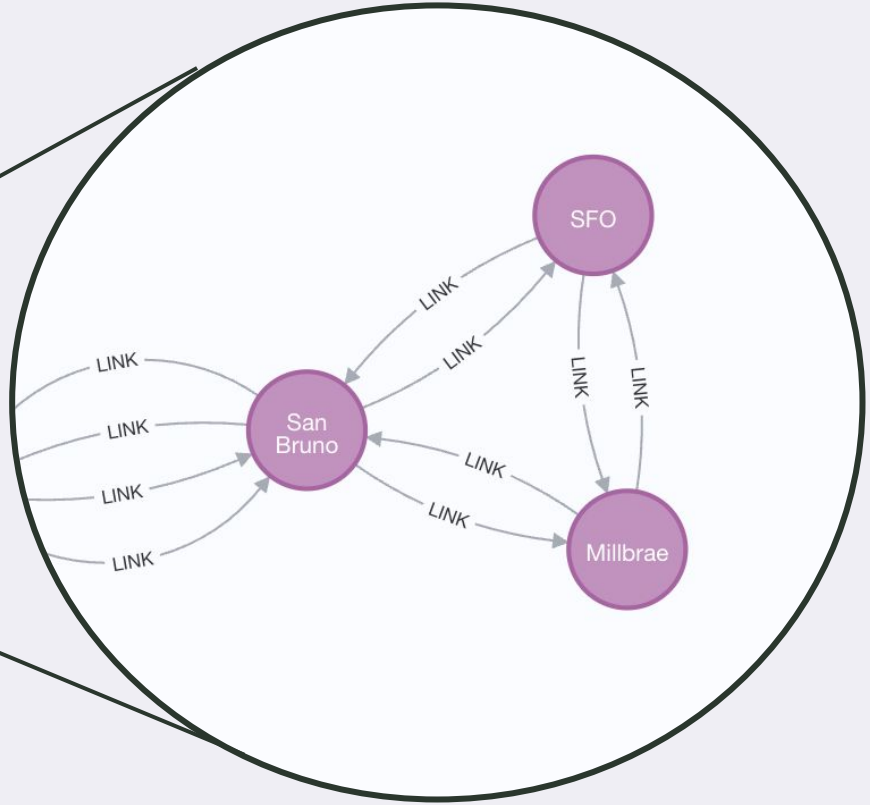
Why Neo4j?

- Graphs inherently follow the same structure as a map
- Neo4j is a graph database
- Neo4j can better handle connected relationships and efficient graph traversals
- Minimizes number of queries and I/Os

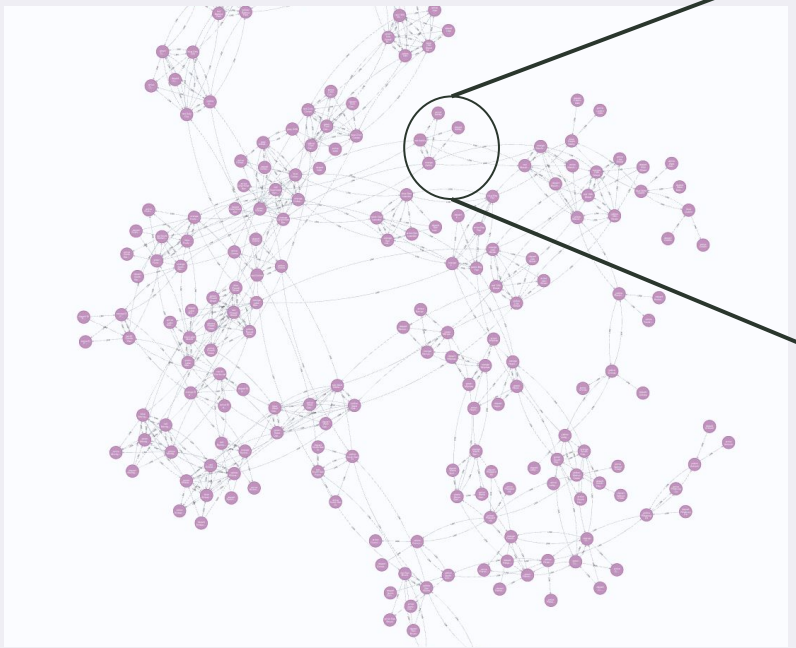


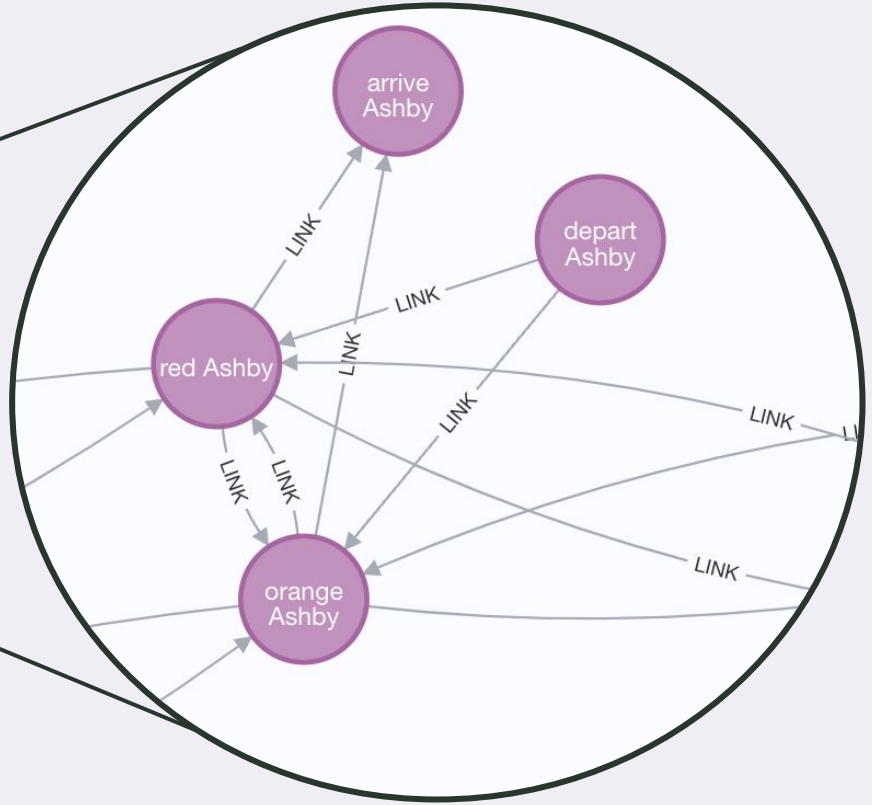
Graph 1




SFO
LINK
LINK
LINK
LINK
LINK
LINK
LINK
LINK
San Bruno
Millbrae

Graph 2




arrive Ashby
depart Ashby
red Ashby
orange Ashby
LINK
LINK
LINK
LINK
LINK
LINK
LINK
LINK
LINK

Our Algorithms

```
depart San Bruno, 0, 0
yellow San Bruno, 0, 0
yellow South San Francisco, 240, 240
yellow Colma, 180, 420
yellow Daly City, 240, 660
yellow Balboa Park, 240, 900
green Balboa Park, 48, 948
green Glen Park, 120, 1068
green 24th Street Mission, 180, 1248
green 16th Street Mission, 120, 1368
green Civic Center, 180, 1548
green Powell Street, 60, 1608
green Montgomery Street, 120, 1728
green Embarcadero, 60, 1788
green West Oakland, 420, 2208
green Lake Merritt, 360, 2568
green Fruitvale, 300, 2868
green Coliseum, 240, 3108
green San Leandro, 240, 3348
green Bay Fair, 240, 3588
green Hayward, 240, 3828
green South Hayward, 240, 4068
arrive South Hayward, 0, 4068
```

Weighted Shortest Paths: best route for inventory to take from a given BART stop closest to the AGM distribution center

Nodes: <line color> station name, <depart/arrive> station name

Weight: Travel time between stations

- e.g San Bruno to South Hayward

	name	betweenness
0	MacArthur	1176.0
1	12th Street	1116.0
2	19th Street	1088.0
3	Lake Merritt	1020.0
4	Fruitvale	980.0
5	West Oakland	980.0
6	Coliseum	960.0
7	Embarcadero	936.0
8	Montgomery Street	888.0
9	Powell Street	836.0
10	San Leandro	836.0

Betweenness Centrality: high betweenness nodes will have a high impact if they are shut down or have delays

Nodes: station name

Edge: Each connection between station

	name	degree
0	16th Street Mission	8.0
1	24th Street Mission	8.0
2	Balboa Park	8.0
3	Civic Center	8.0
4	Embarcadero	8.0
5	Glen Park	8.0
6	Montgomery Street	8.0
7	Powell Street	8.0
8	West Oakland	8.0
9	Coliseum	7.0
10	12th Street	6.0

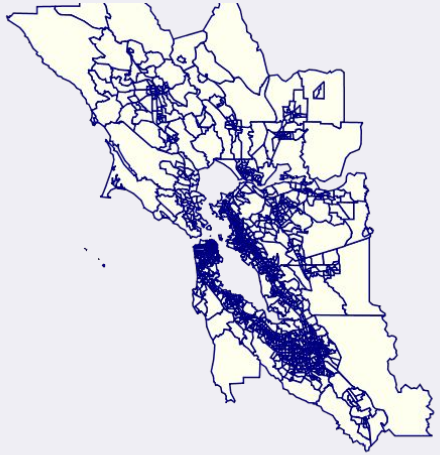
Degree Centrality: find BART stations that are overloaded with connections and likely to be busy and crowded

Nodes: station name

Edge: Each connection between station

Why Redis?

- Storing real time traffic data using Redis for active traffic analyses for rerouting delivery trucks
- Relational databases can’t query and process through real time, high volume data as efficiently as Redis, leading to longer routing times
- Identifying real time traffic doesn’t rely on complex queries
- Routes can be pre-determined using other tools, but Redis will help navigate through active traffic data



03

Recommendation Tool

Leveraging MongoDB

Quicker Consumer Analytics

Analyze customers’ purchasing habits

Present recommendations

Why MongoDB?

Complexity

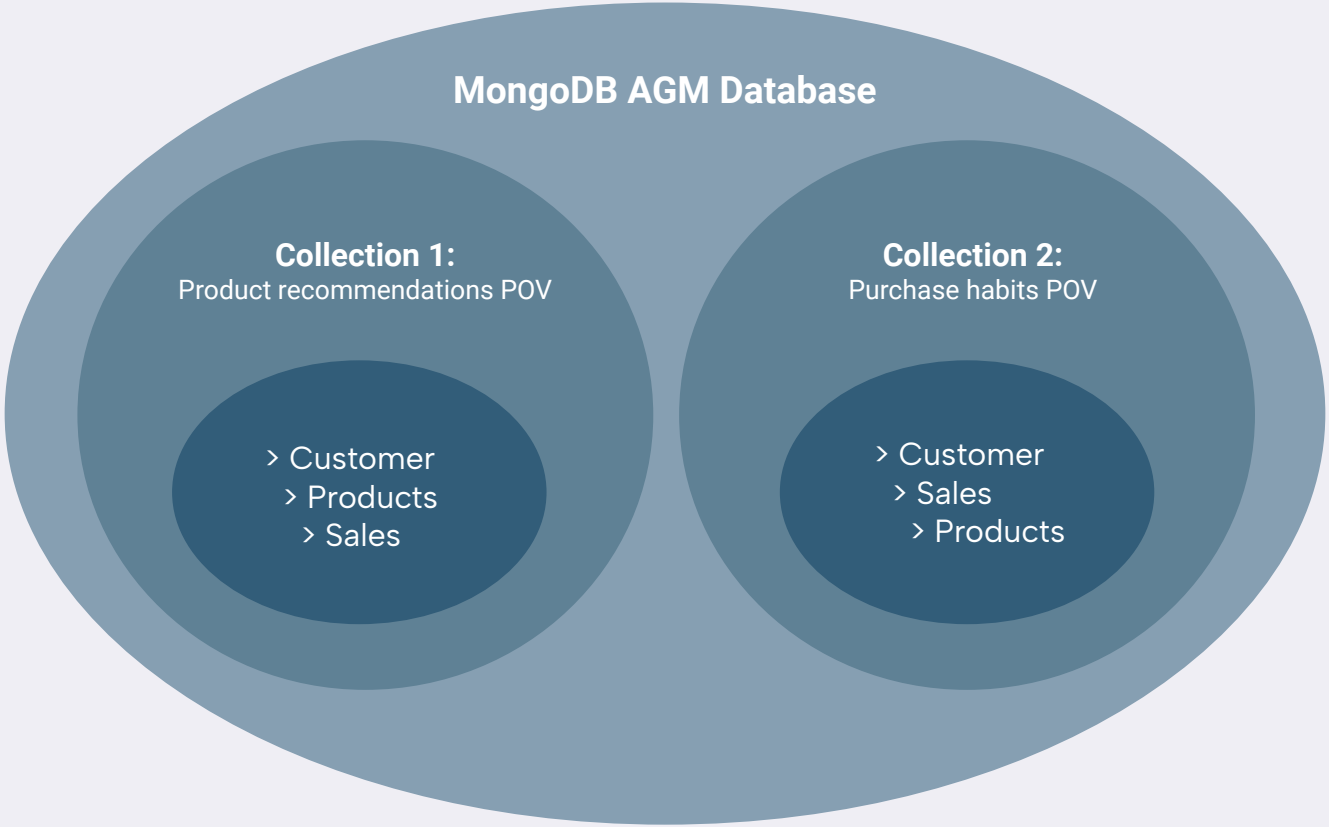
Customizable nested structuring of data allows for very specific, complex analyses

Scalability

Not as much computing power necessary

Ability to handle high volumes of data

MongoDB Overview



Yes SQL?
No, SQL!