Streamlining AGM Business Needs (NoSQL Databases)

Jasmine Teo & Sally Fang
Data 205 Section 099

Table of contents

01 Neo4j

02 MongoDB

03 Redis

04 Moving Forward

01

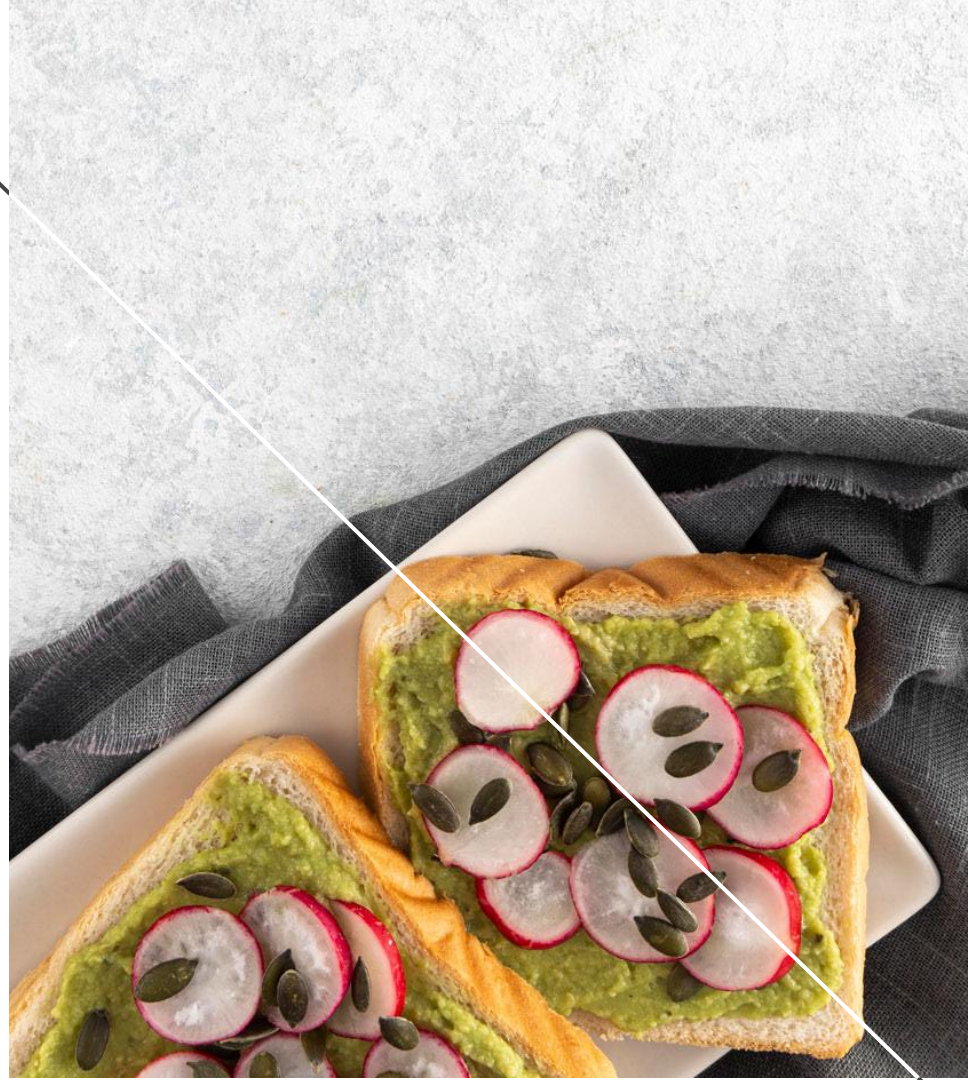
Neo4j

Business Example #1

Neo4j Business Example:

AGM is interested in adding pickup locations at the BART stations and also using BART as transportation for delivery.

We recommend using Neo4j to find the most ideal pickup locations in terms of efficiency and importance, as well as mapping the best route for delivery.



Utilizing Neo4j for AGM Services

Shortest Path Algorithm

Finding the optimal path to transport meals using Bart

Weighted Betweenness Centrality

Finding the most critical BART stations

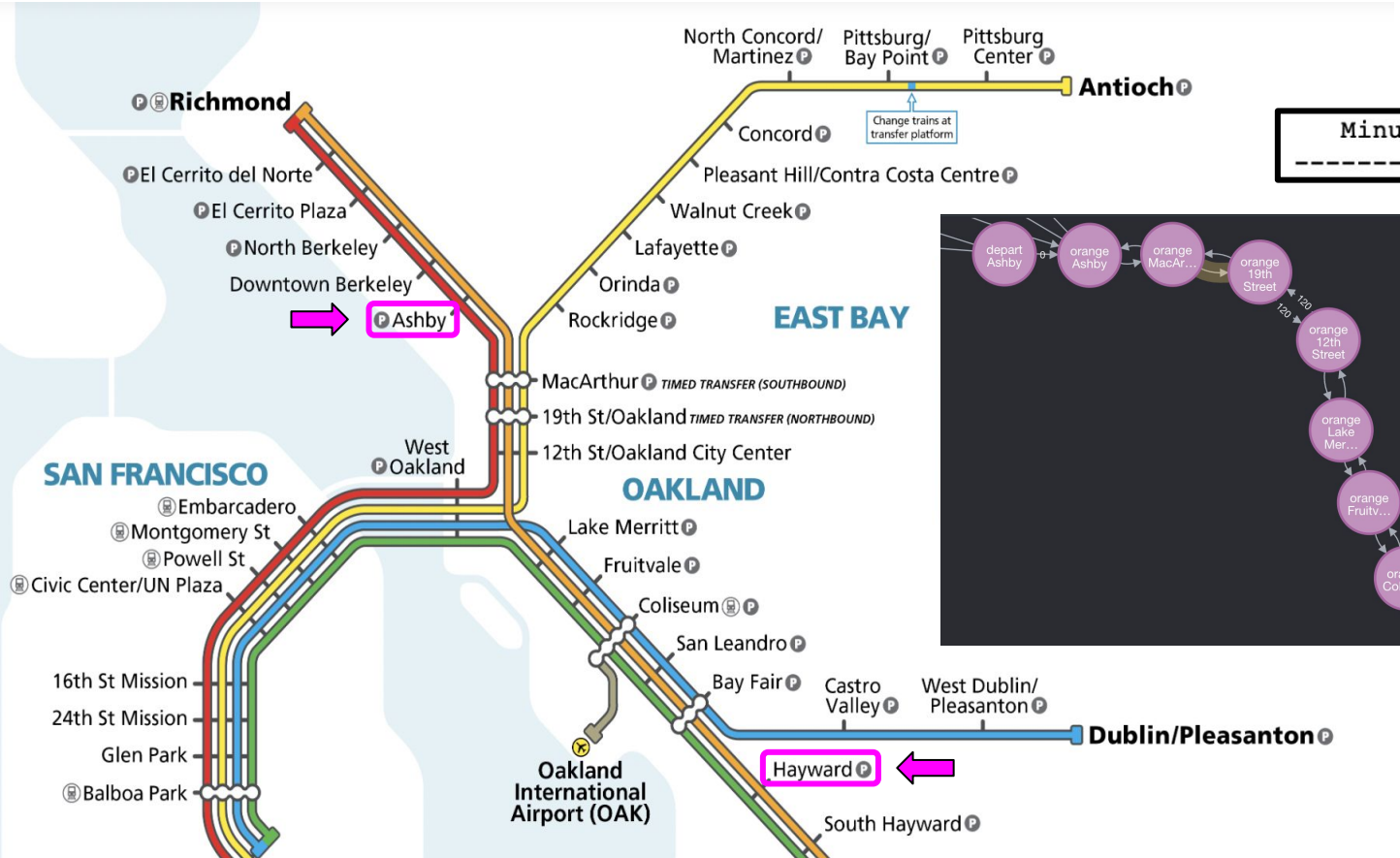
Louvain Modularity

Finding the clusters of BART stations

Shortest Path Algorithm

- Use *Dijkstra Source-Target Shortest Path* algorithm:
  - Shortest path between a source and a target
  - Return an ordered list of bart stations
- Use this to find the **optimal path** to transport meals, assessed by:
  - **Time:**: compare the estimated duration of travel to other BART routes and trucks delivery routes
  - **Cost**: compare to the cost of delivering by truck or other means of transportation

Shortest Path: Ashby ->Hayward

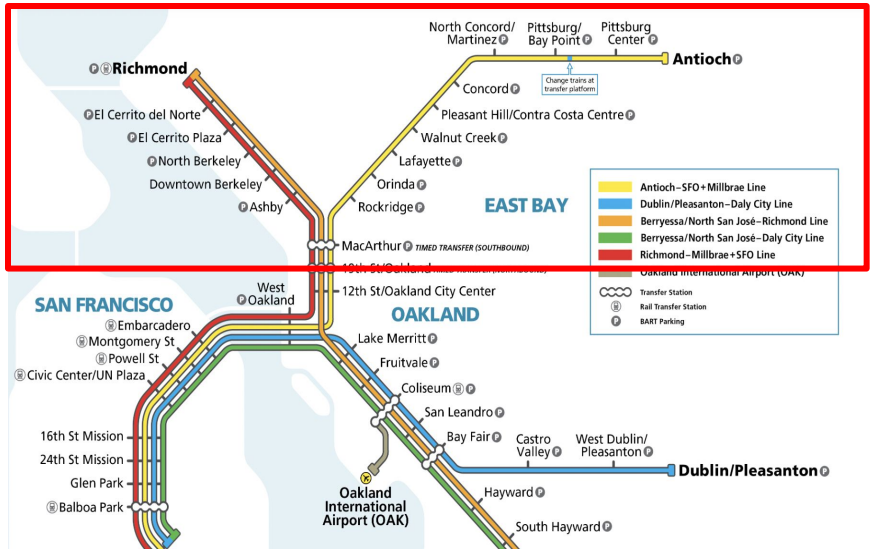


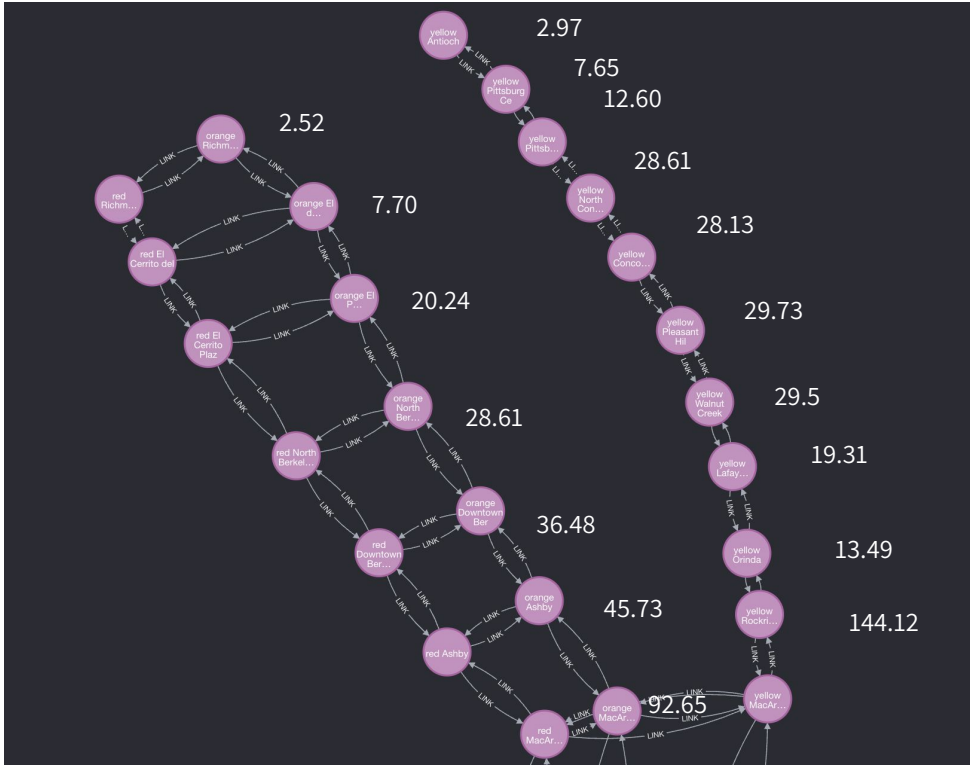

```
Minutes:  33.0
--------------------------------
```

Weighted Betweenness Centrality

- Use *Weighted Betweenness Centrality,* taking population density into account
  - What is the node's **ability to act as a bridge**
  - Assign a weight based on the population density
  - Calculate the number of **shortest paths** between station pairs that past through specific station
- Use cases:
  - Identify **key stations** with heavy foot traffic
  - Assess **network resilience**
  - **Visualize** BART lines and stations

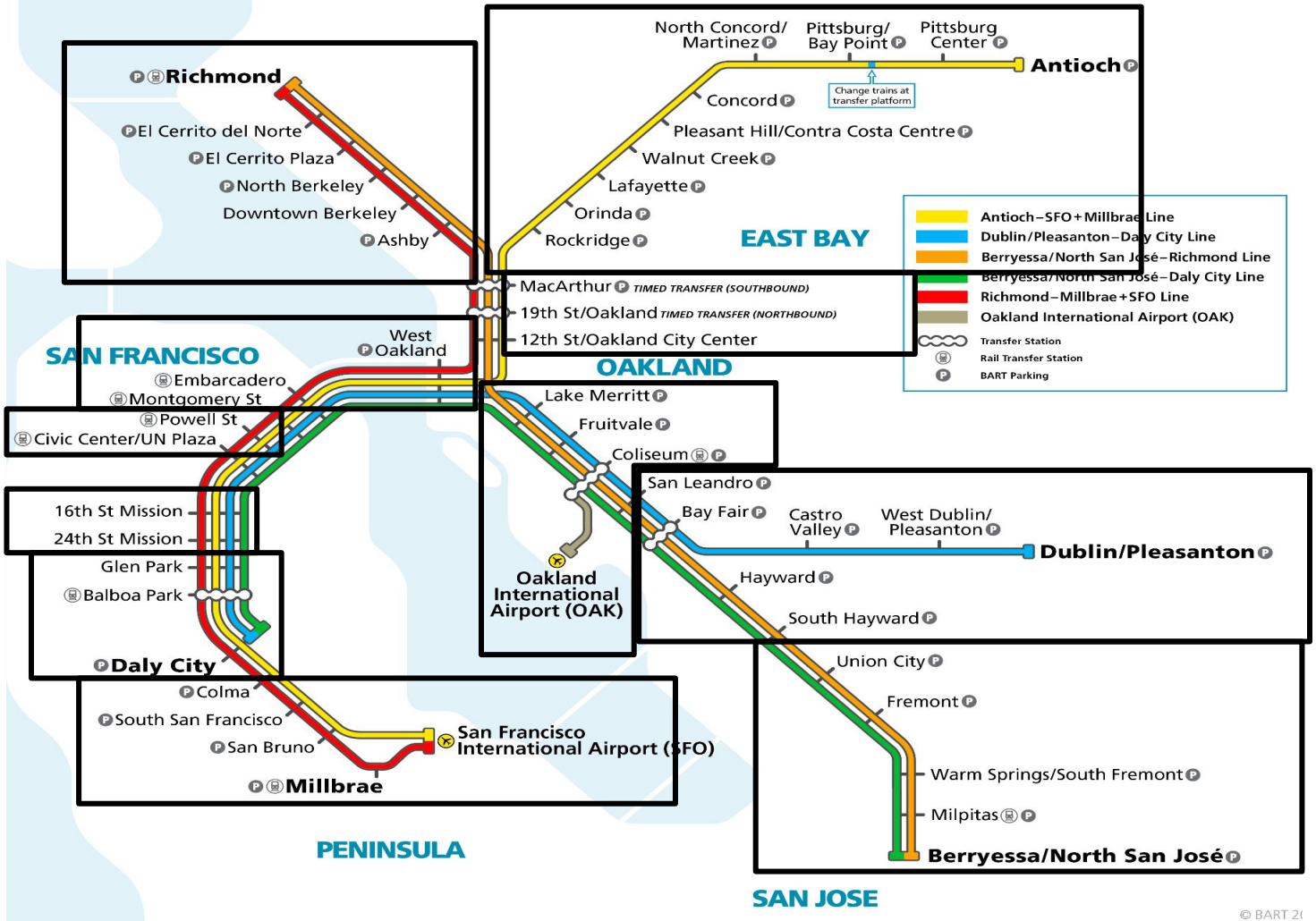
MacArthur -> Richmond; Antioch





Community Detection: Louvain Modularity

- Use *Louvain Modularity* to **detect clusters** and optimize modularity in large networks quickly
  - **Label Propagation algorithm (LPA)**
  - **Maximize a modularity score** for each community
- Identify clusters of pickup locations that are grouped together
  - **Optimize** delivery routes
  - **Reduce** transportation costs


Richmond
El Cerrito del Norte
El Cerrito Plaza
North Berkeley
Downtown Berkeley
Ashby
North Concord/ Martinez
Pittsburg/ Bay Point
Pittsburg Center
Antioch
Change trains at transfer platform
Concord
Pleasant Hill/Contra Costa Centre
Walnut Creek
Lafayette
Orinda
Rockridge
EAST BAY
MacArthur TIMED TRANSFER (SOUTHBOUND)
19th St/Oakland TIMED TRANSFER (NORTHBOUND)
12th St/Oakland City Center
OAKLAND
Antioch–SFO+Millbrae Line
Dublin/Pleasanton–Daly City Line
Berryessa/North San José–Richmond Line
Berryessa/North San José–Daly City Line
Richmond–Millbrae+SFO Line
Oakland International Airport (OAK)
Transfer Station
Rail Transfer Station
BART Parking
SAN FRANCISCO
West Oakland
Embarcadero
Montgomery St
Powell St
Civic Center/UN Plaza
16th St Mission
24th St Mission
Glen Park
Balboa Park
Daly City
Lake Merritt
Fruitvale
Coliseum
Oakland International Airport (OAK)
San Leandro
Bay Fair
Castro Valley
West Dublin/ Pleasanton
Dublin/Pleasanton
Hayward
South Hayward
Union City
Fremont
Warm Springs/South Fremont
Milpitas
Berryessa/North San José
SAN JOSE
Colma
South San Francisco
San Bruno
San Francisco International Airport (SFO)
Millbrae
PENINSULA

Why Not Relational Database?

- **Modeling Flexibility:**
  - Want to model BART network system as nodes
  - Relationships between nodes
- **Fast Queries:**
  - Neo4j is graph-based and more efficient,
  - Traverse node relationships in one operation
- **Scalability:**
  - Relational databases may struggle as the size of the database grows,
- **Visualization:**
  - Graphical representation of BART network system
  - Easier to interpret and understand

02

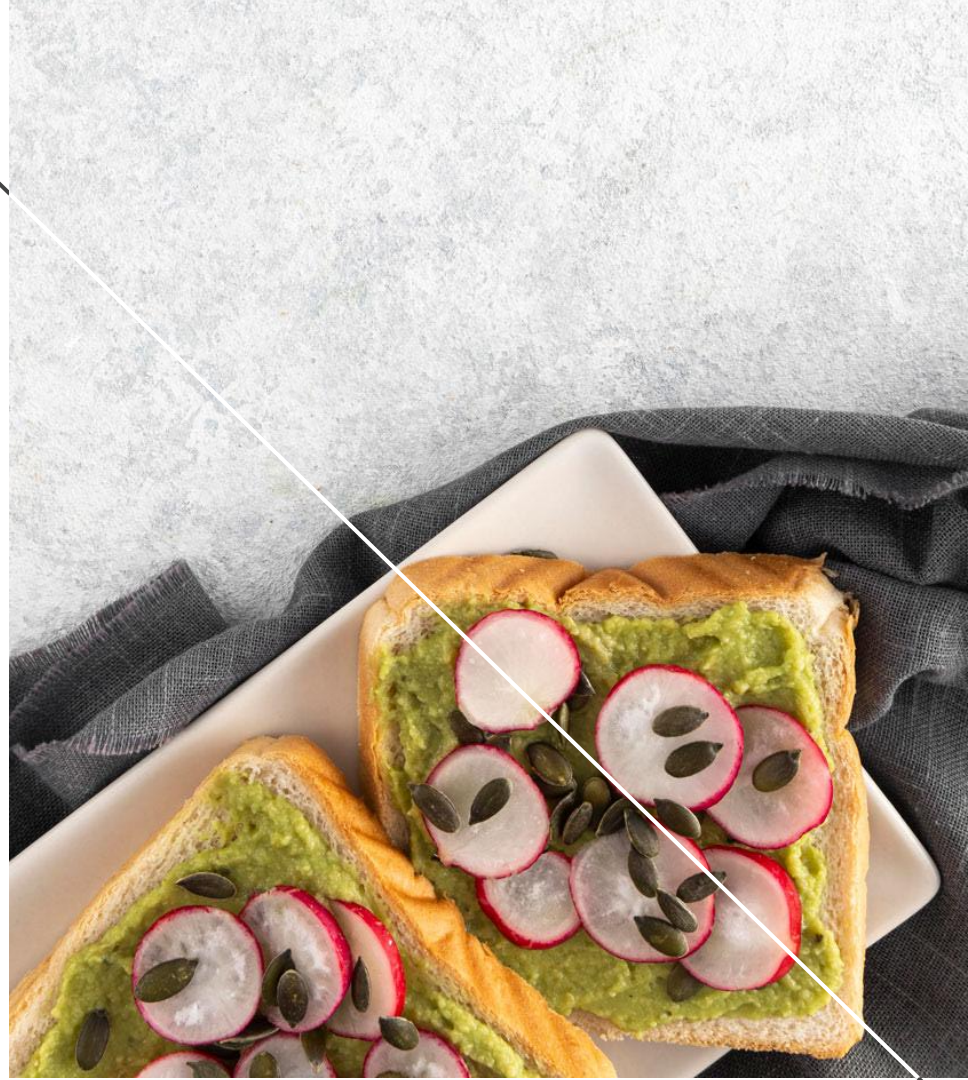
MongoDB

Business Example #2

MongoDB Business Example:

AGM is interested in adding pickup locations at the BART stations and also using BART as transportation for delivery.

We recommend using MongoDB to store the shortest path computed for transportation using BART as well as the heuristics for computing a shortest path for trucks, delivery drones, and robots.

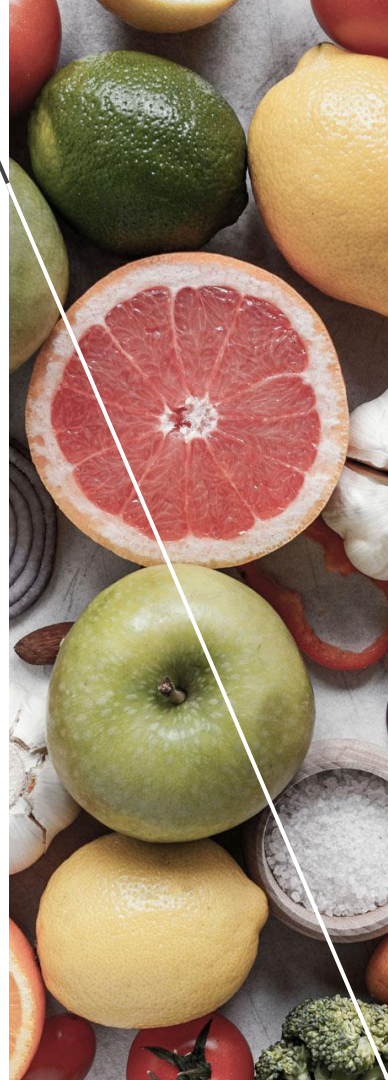


Why MongoDB?

- **Store shortest path for BART**
- **Store the heuristics for trucks, delivery drones, and robots**
    - For trucks: weight, size, and road restrictions specific to trucks
    - For delivery drones: flight range
    - For robots: speed, mobility, road restrictions specific to delivery robots
- **Storing Multiple POVs**
    - AGM POV: Store administrative data
    - Customer POV: Store data related to customer orders
    - Truck Driver POV: Store driver data

Why Not Relational Database?

- **Data Structure & Flexibility:**
  - MongoDB index and query Geospatial data
  - Frequent changes in data structures
- **Scalability:**
  - MongoDB scale up to handle large volumes of traffic data
  - Optimize delivery routes
- **Cost:**
  - MongoDB is cheaper to main large amounts of data
  - Require less hardware



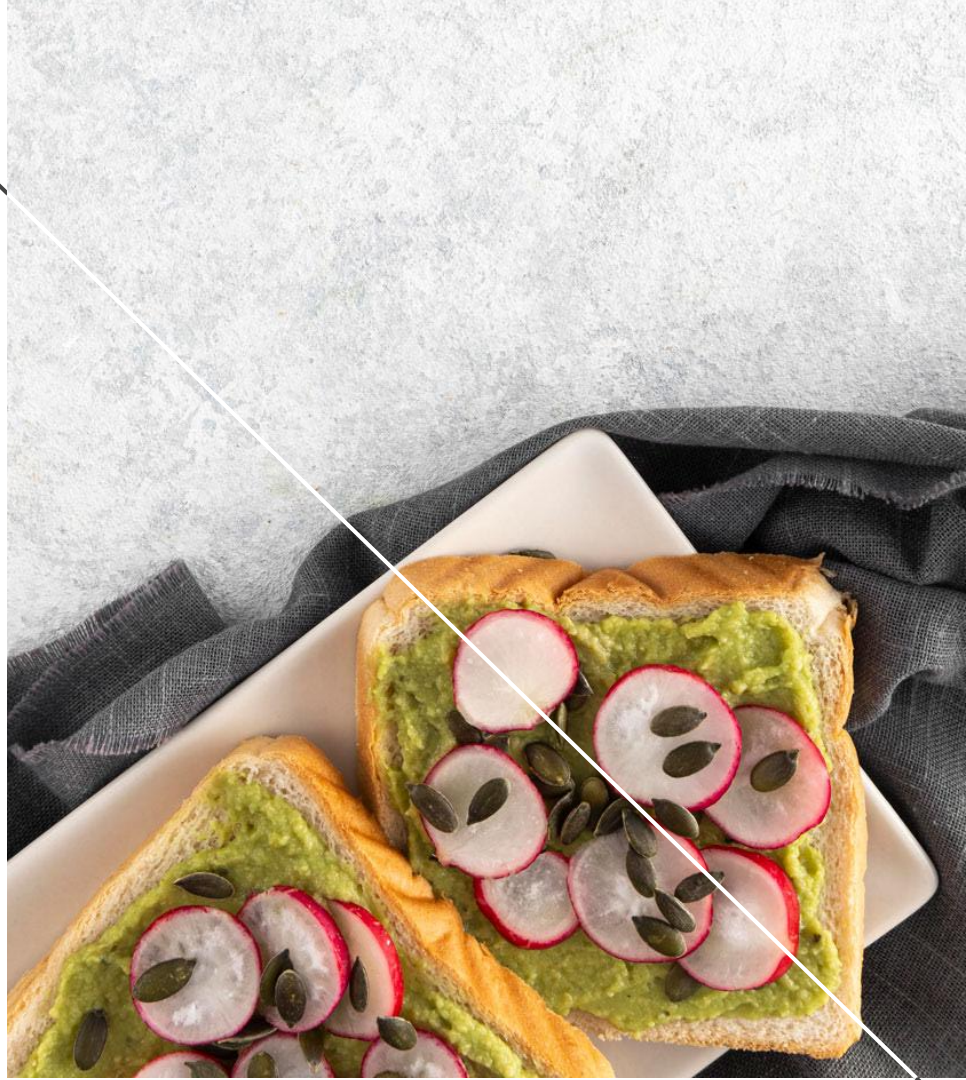
03 Redis

Business Example #3

Redis Business Example:

AGM is interested in adding pickup locations at the BART stations and also using BART as transportation for delivery.

We recommend using Redis to store live updates and real time traffic conditions for delivery, provide location based recommendations to customers, and using those information to optimize routes for delivery, as well as detecting fraudulent orders and transactions



Why Redis?

- **Store live updates of orders**
  - Track live status of food delivery
  - Update the inventory status
- **Real time traffic conditions**
  - Store real time traffic conditions
  - Provide live updates and instructions
  - Accurate delivery estimates and adjust delivery route
- **Fraud Detection**
  - Monitor order / delivery patterns and detect unusual activity

Why Not Relational Database?

- Real-time data processing and retrieval
- Concurrent updates
- Scalability

Moving Forward

We highly recommend to the data science team at AGM to utilize **NoSQL databases** in implementing the executives' vision for the future.

NoSQL databases provide a **fast, efficient, scalable** approach to handle real-time traffic updates, geospatial data, and large volumes of unstructured data.

Thanks!