Getting Closer to Consumer...

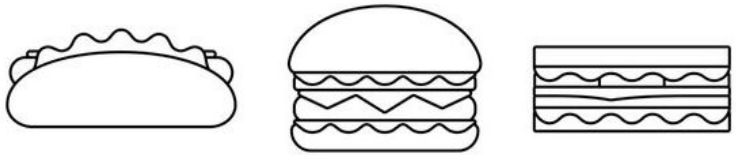
Technology Proposal For AGM To Improve The Delivery Service

Data Science -New Tech Team: Peeti Sriwongsanguan , Ahmad Umair, Nishika Abeytunge

https://www.fitnesstipblog.com/2500-calories-meals-preparation/

Current Status

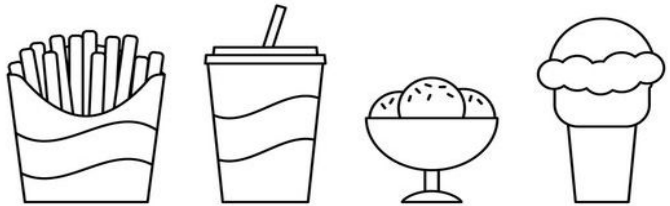
- Acme Gourmet Meals restaurant is located in downtown Berkeley
- Peek is the designated third party delivery service provider
- Leadership see a high growth potential through improving delivery service
- Tech base: Relational database containing information linke sales, customers and Product fed internally, Peek data uploaded separately



Objective Of The Analysis

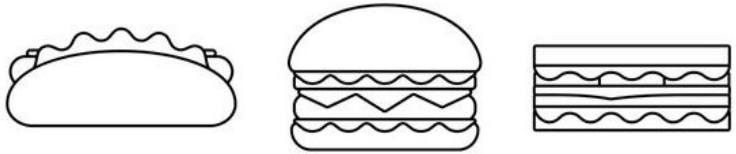
Propose a NoSql database solution which will facilitate maintaining better delivery options so that AGM can

- Increase the customer base
- Improve the customer satisfaction
- Improve the profitability



Why NoSql?

- NoSql Databases are scalable- will help accommodating changes efficiently
- Graph database suited for delivery routing options
- JSON objects reduce the no of tables and easy to use in other applications
- Characteristics are more desired for the cloud databases
- Can handle unstructured, semi-structured or structured data
- Easy handle of big data

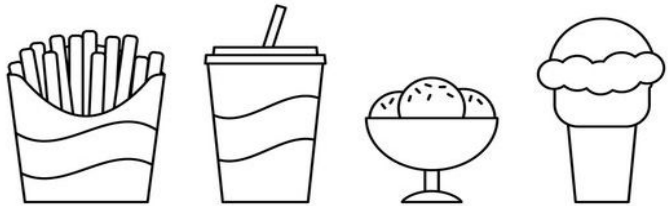


NoSql Database - Primary: Neo4j

Is an open-source, NoSQL, native graph database uses Cypher as the query language

We can use a Neo4j graphs to store and analyse data about distribution , while using relational database to create business data tables from json and csv. These graphs will help in

- Providing graphical views of best routing options
- Identifying the heavy traffic routes
- Deciding the shortest path
- Adapt new relationships easily
- Use SQL-Like query language



Neo4j: Business Case Scenario #1

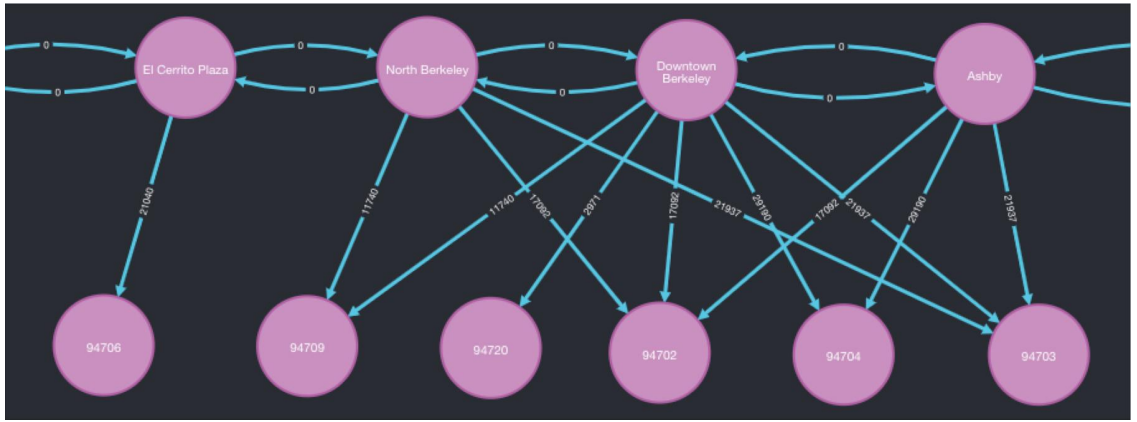
Add pickup locations at the BART stations, based on:

- Zip code population within *X miles* from the station
- Travel time at least *Y minutes* between the stations

Find all zip codes within ***X miles*** from each station → Find population of the respective zip codes → Neo4j Graph Creation → ...

Neo4j graph: X = 1 mile in this example

```
Nodes: 91
Relationships: 169
```



Neo4j: Business Case Scenario #1 (Continued...)

Add pickup locations at the BART stations, based on the population within *X miles* from the station.

...

Degree or Betweenness Centrality **(Weighted)**

	station	degree
0	Powell Street	140730.0
1	Glen Park	115068.0
2	24th Street Mission	108915.0
3	Balboa Park	106589.0
4	Pittsburg Center	96081.0
5	Montgomery Street	85465.0
6	Downtown Berkeley	82930.0
7	South Hayward	79235.0
8	Civic Center	74898.0
9	Union City	74601.0
10	Fremont	73855.0

...
...
...

All stations within ***Y minutes*** of a specific station **(Separate Neo4j graph: time weighted)**

	station	within_7_minutes
0	12th Street	[19th Street, Lake Merritt, MacArthur, West Oa...
1	16th Street Mission	[24th Street Mission, Civic Center, Glen Park,...
2	19th Street	[12th Street, Lake Merritt, MacArthur]
3	24th Street Mission	[16th Street Mission, Balboa Park, Civic Cente...
4	Antioch	[]
5	Ashby	[Downtown Berkeley, MacArthur, North Berkeley]
6	Balboa Park	[24th Street Mission, Daly City, Glen Park]
7	Bay Fair	[Castro Valley, Hayward, San Leandro]
8	Berryessa	[Milpitas]
9	Castro Valley	[Bay Fair]
10	Civic Center	[16th Street Mission, 24th Street Mission, Emb...

...
...
...

Note: X = 1 mile in this example

Remove stations within ***Y minutes*** of the higher degree station

	station	degree	within_7_minutes
0	Powell Street	140730.0	[16th Street Mission, 24th Street Mission, Civ...
1	Glen Park	115068.0	[16th Street Mission, 24th Street Mission, Bal...
2	Pittsburg Center	96081.0	[]
3	Downtown Berkeley	82930.0	[Ashby, El Cerrito Plaza, North Berkeley]
4	South Hayward	79235.0	[Hayward, Union City]
5	Fremont	73855.0	[Union City, Warm Springs]
6	Colma	64887.0	[Daly City, South San Francisco]
7	Dublin	54420.0	[West Dublin]
8	Fruitvale	52299.0	[Coliseum, Lake Merritt]
9	San Leandro	48088.0	[Bay Fair, Coliseum]
10	San Bruno	43124.0	[SFO, South San Francisco]
11	Rockridge	39852.0	[MacArthur, Orinda]
12	North Concord	38568.0	[Concord, Pittsburg]
13	Lafayette	29639.0	[Orinda, Walnut Creek]
14	Berryessa	28726.0	[Milpitas]
15	West Oakland	26254.0	[12th Street, Lake Merritt]
16	Pleasant Hill	22734.0	[Concord, Walnut Creek]
17	Millbrae	22710.0	[SFO]
18	19th Street	16062.0	[12th Street, Lake Merritt, MacArthur]

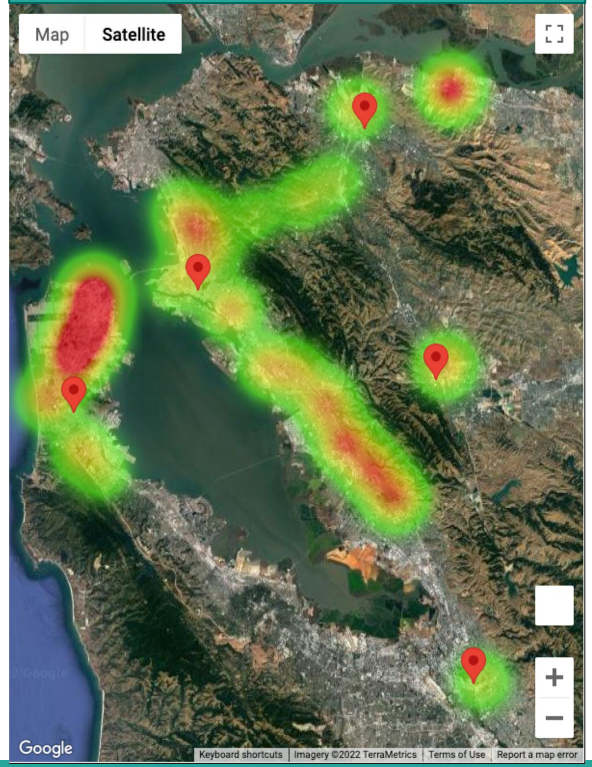
Google maps API snapshot with locations and population heat map

Neo4j: Business Case Scenario #1 (Continued...)

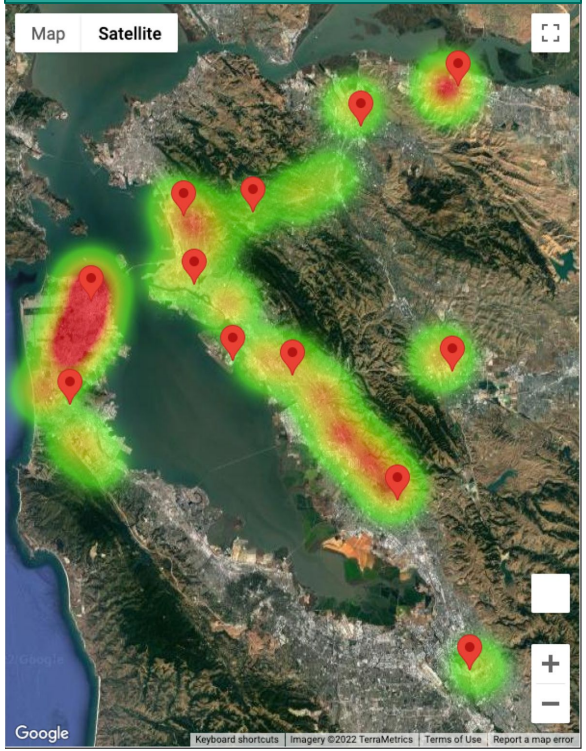
Adding pickup locations at the BART stations, based on the population within *X miles* from the station and travel time at least *Y minutes*.

Degree Centrality with X = 10 miles, Y = 10-30 minutes

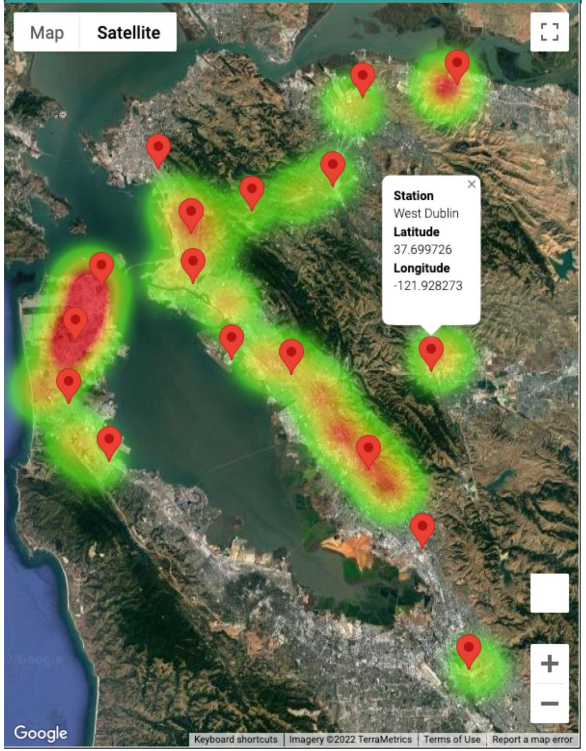
X = 10 miles, Y = 30 minutes



X = 10 miles, Y = 15 minutes



X = 10 miles, Y = 10 minutes

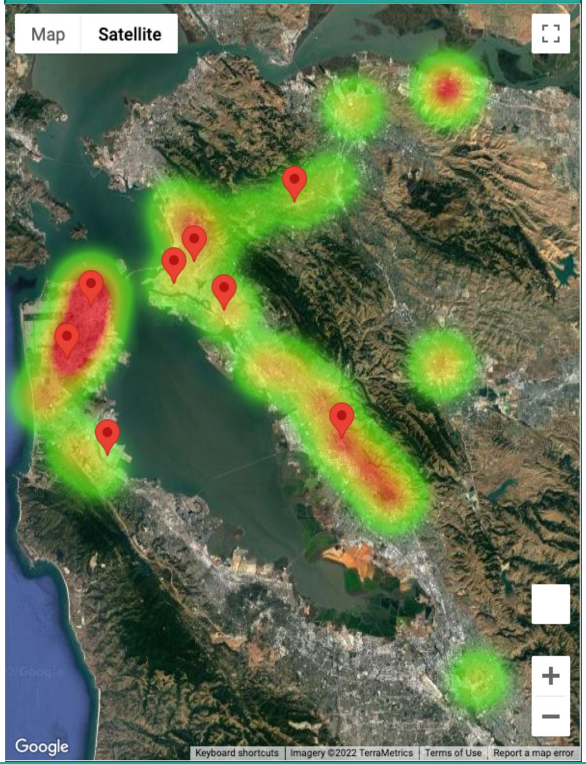


Neo4j: Business Case Scenario #1 (Continued...)

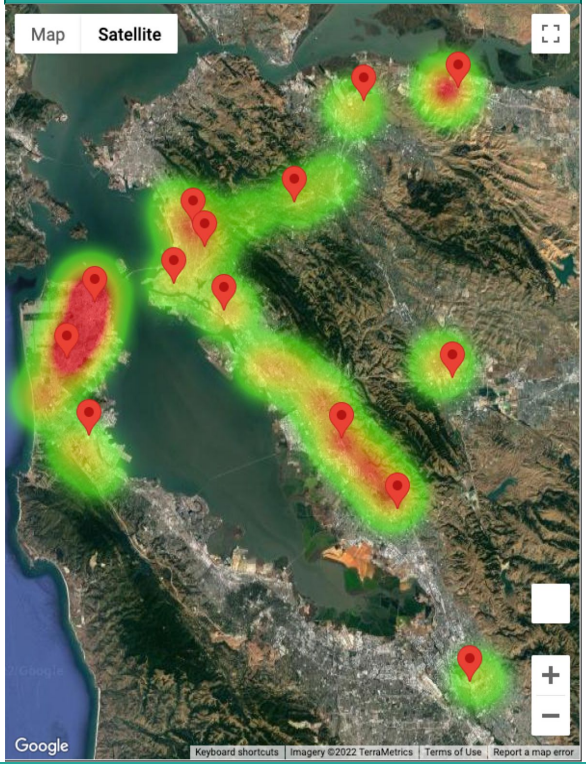
Adding pickup locations at the BART stations, based on the population within *X miles* from the station and travel time at least *Y minutes*.

Degree Centrality with X = 0.5 to 10 miles, Y = 10 minutes

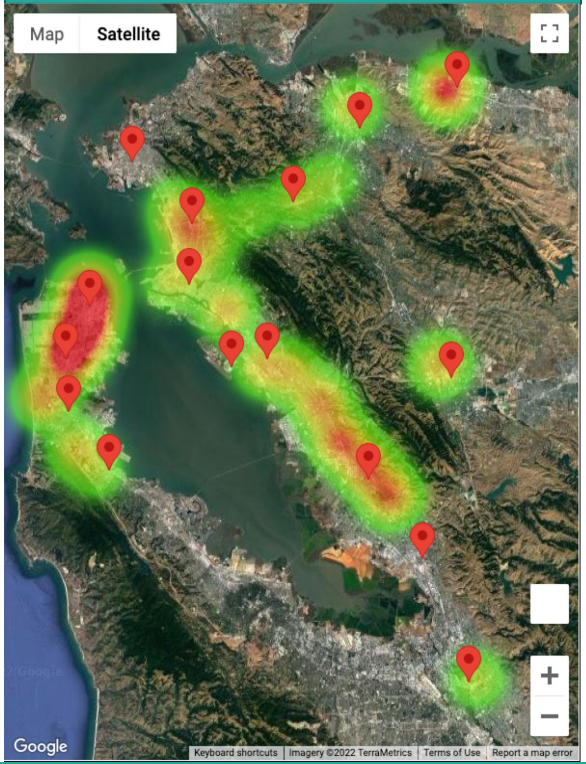
X = 0.5 mile, Y = 10 minutes



X = 1 mile, Y = 10 minutes

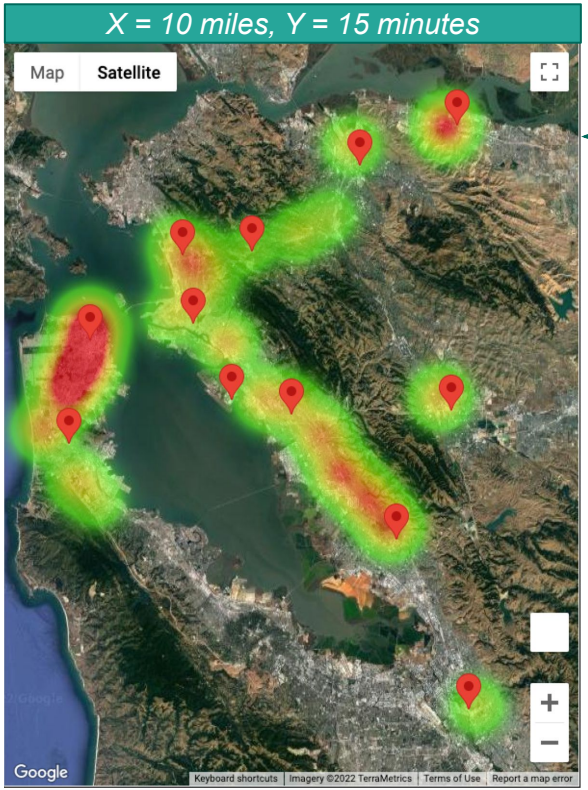


X = 10 miles, Y = 10 minutes



Neo4j: Business Case Scenario #2

Adding pickup locations at the BART stations, based on the *current customers population* within *X miles* from the station and travel at least *Y minutes*.

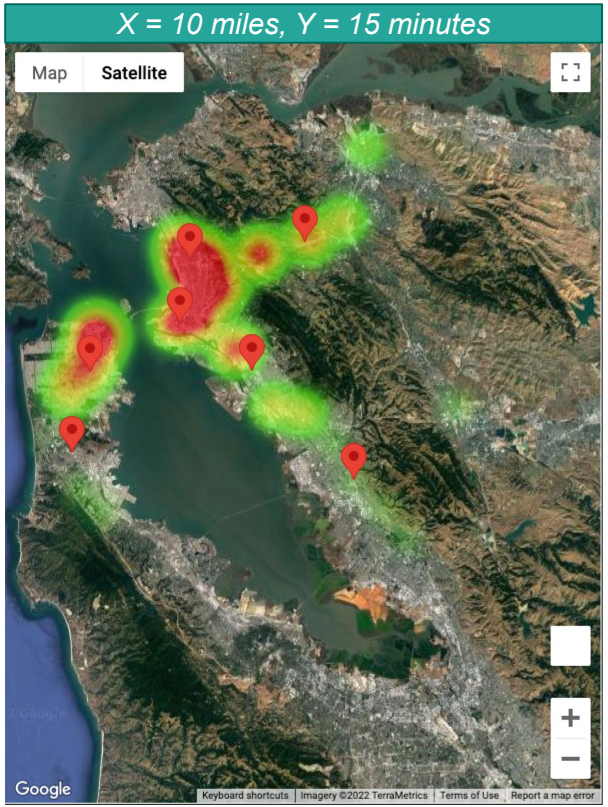


Previous case

Weights = total population in the zip code

Current case

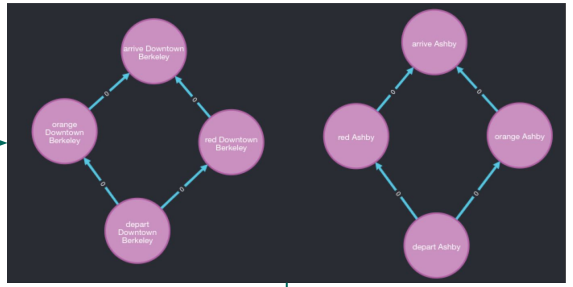
Weights = total customer population in the zip code

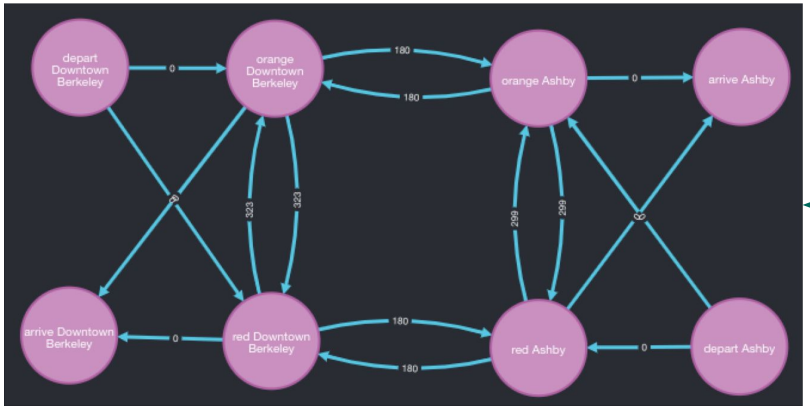


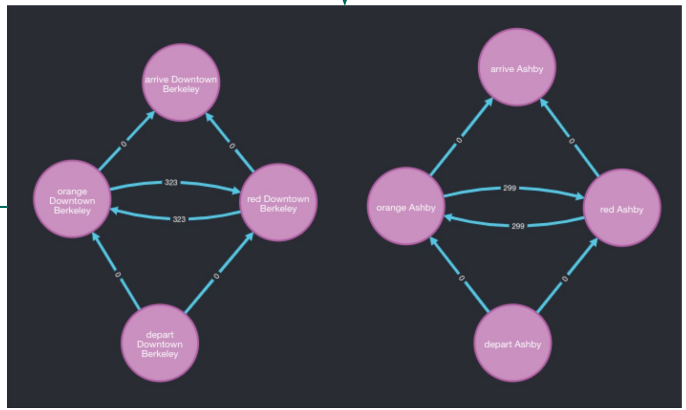
Neo4j: Business Case Scenario #3

Efficiently delivering items to specific BART stations, from Downtown Berkeley Store.

Neo4j Graph Structure Into







All stations/lines

```
----------------------
Nodes: 214
Relationships: 652
----------------------
```

Neo4j: Business Case Scenario #3 (Continued)

Efficiently delivering items to specific BART stations, from Downtown Berkeley Station

...

Station list based on some algorithm

Depart = Downtown Berkeley
Arrive = Iterate through the station list
—
Yen's algorithm Shortest path

Randomly select a path from the lowest total cost paths

	station	degree	within_7_minutes
0	Powell Street	140730.0	[16th Street Mission, 24th Street Mission, Civ...
1	Glen Park	115068.0	[16th Street Mission, 24th Street Mission, Bal...
2	Pittsburg Center	96081.0	[]
3	Downtown Berkeley	82930.0	[Ashby, El Cerrito Plaza, North Berkeley]
4	South Hayward	79235.0	[Hayward, Union City]
5	Fremont	73855.0	[Union City, Warm Springs]
6	Colma	64887.0	[Daly City, South San Francisco]
7	Dublin	54420.0	[West Dublin]
8	Fruitvale	52299.0	[Coliseum, Lake Merritt]
9	San Leandro	48088.0	[Bay Fair, Coliseum]
10	San Bruno	43124.0	[SFO, South San Francisco]
10	Civic Center		[16th Street Mission, 24th Street Mission, Emb...

...
...
...

Note: Here we are taking the stations list we got from business cases scenario #1.

	from	to	totalCost	nodes
0	depart Downtown Berkeley	arrive Powell Street	1620.0	[depart Downtown Berkeley, red Downtown Berkel...
1	depart Downtown Berkeley	arrive Powell Street	1679.0	[depart Downtown Berkeley, orange Downtown Ber...
2	depart Downtown Berkeley	arrive Powell Street	1679.0	[depart Downtown Berkeley, red Downtown Berkel...
3	depart Downtown Berkeley	arrive Powell Street	1679.0	[depart Downtown Berkeley, orange Downtown Ber...
4	depart Downtown Berkeley	arrive Powell Street	1687.0	[depart Downtown Berkeley, red Downtown Berkel...

Note: Here we have taken the fixed travel/transfer times, based on the BART csv files.

If we have access to real-time travel and line-transfer times, we will be able to get the real-time shortest paths according based on this proof of concept.

```
[('depart Downtown Berkeley',
  'arrive Powell Street',
  1620.0,
  ['depart Downtown Berkeley',
   'red Downtown Berkeley',
   'red Ashby',
   'red MacArthur',
   'red 19th Street',
   'red 12th Street',
   'red West Oakland',
   'red Embarcadero',
   'red Montgomery Street',
   'red Powell Street',
   'arrive Powell Street'],
  [0.0,
   0.0,
   180.0,
   420.0,
   600.0,
   720.0,
   1020.0,
   1440.0,
   1500.0,
   1620.0,
   1620.0])]
```

Shortest Path

Time (s)

Summary of Neo4j Recommendations

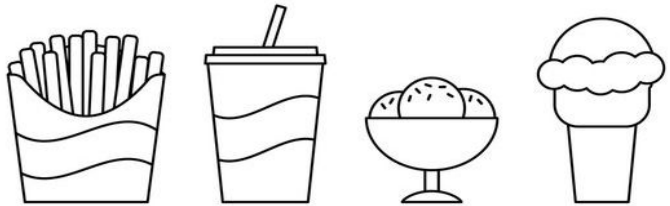
We presented various business cases aligned with the AGM executives vision :

- Add pickup locations based on the population surrounding the BART
  - Our recommendation is to have a pickup stations at (30+ minutes apart):
    - South San Francisco
    - Concord
    - West Dublin
    - Berryessa
- Efficiently delivering to customers using shortest possible paths
  - Once we have real-time BART travel time, our algorithm will provide all shortest paths - and one of the paths will be randomly selected

MongoDB: Business Use Case

MongoDB can help AGM:

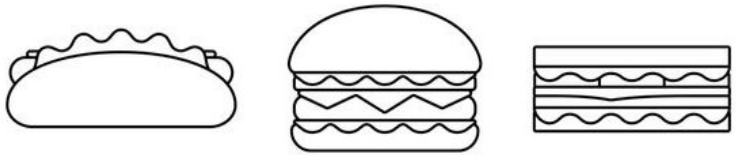
- Product Data management
- Provide a search suggestion feature
- Storing and managing huge amounts of customer data
- Run queries faster compared to relational database
- Real-time analytics



Redis: Business Use Case

Consumers expectation in tracking their deliveries, password resets etc require real time interactions with the database.

- Higher frequency database interactions for real time delivery tracking
- Ability to deal with multiple data format
- Cons : Will be expensive for higher memory requirements



MongoDB vs Redis

- Speed
- RAM
- Scalability
- Storage

Secondary DB Recommendation: MongoDB