Project – 3

Emily-Karan-Peng-Ravi

Scaling AGM's Business with NoSQL

UC Berkeley School of Information

Agenda

1. Overview
2. NoSQL Tools
3. Business Applications
4. Recommendations

Overview:

- Introduce NoSQL Databases like Neo4j, Redis and MongoDB
- Provide examples of how these tools can be applied to help AGM's grow profitably

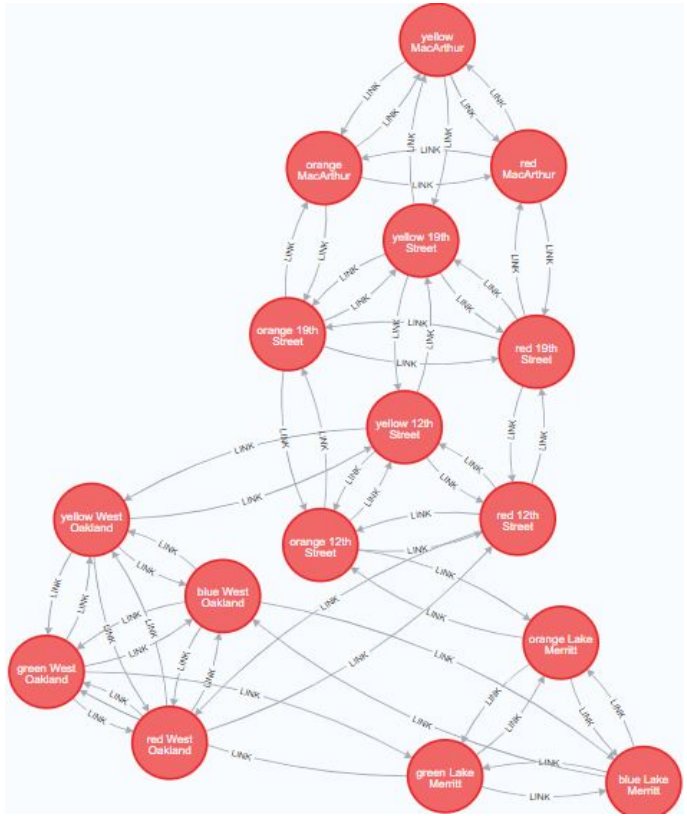
A graph database that stores data as nodes and relationships; leverages mathematical graph theory to classify, connect, and analyze data effectively

Redis is an in-memory, NoSQL key/value store that is used primarily as an application cache or quick-response database.

MongoDB is a non-relational document database that provides support for JSON-like storage

NoSQL Databases are gaining popularity with their unique advantages

Business Application: Neo4J – Visualization of BART



Neo4j visualizes data with interconnected nodes and quantified relationships

Business Application: Neo4J – Closeness Centrality

	station_name	closeness
0	West Oakland	0.159924
1	Embarcadero	0.154470
2	Lake Merritt	0.154067
3	12th Street	0.153819
4	Montgomery Street	0.150407
5	Fruitvale	0.147292
6	Powell Street	0.146942
7	19th Street	0.145794
8	Civic Center	0.143685
9	Coliseum	0.142317

Closeness Centrality can help identify strategically located stations for efficient goods distribution and regional coverage.

Business Application: Neo4J – Betweenness Centrality

	station_name	betweenness
0	Rockridge	5509.000000
1	MacArthur	5239.000000
2	Orinda	4997.000000
3	Lafayette	4469.000000
4	Walnut Creek	3925.000000
5	12th Street	3752.333333
6	Lake Merritt	3723.000000
7	West Oakland	3647.750000
8	Coliseum	3603.750000
9	19th Street	3535.000000

Betweenness centrality can help identify critical transfer hubs for connecting regions, pickup/drop-off points and ensuring efficient deliveries.

Business Application: Neo4J – Clustering Coefficient Analysis

	station_name	clustering_coefficient
0	Daly City	0.514286
1	Coliseum	0.471429
2	West Oakland	0.428571
3	Embarcadero	0.428571
4	24th Street Mission	0.428571
5	Balboa Park	0.428571
6	Powell Street	0.428571
7	Montgomery Street	0.428571
8	Civic Center	0.428571
9	16th Street Mission	0.428571
10	Glen Park	0.428571

Clustering coefficient analysis can help identify tightly connected local clusters ideal for efficient deliveries and future expansion.

Business Application: Neo4J – Consolidated Results

Rank	Station Name	Exits
1	Embarcadero	15,420
2	Daly City	4,042
3	MacArthur	3,356
4	West Oakland	2,970
5	Lake Merritt	2,721
6	Coliseum	2,323
7	Rockridge	2,116
8	Orinda	1,061

Final Decision:

1. Top 3 Stations from 3 algorithms
2. Ordered by Average Exits volume
3. Choose the top 3 volume stations

Neo4J can help reach most customers, fast delivery and closeness to customers

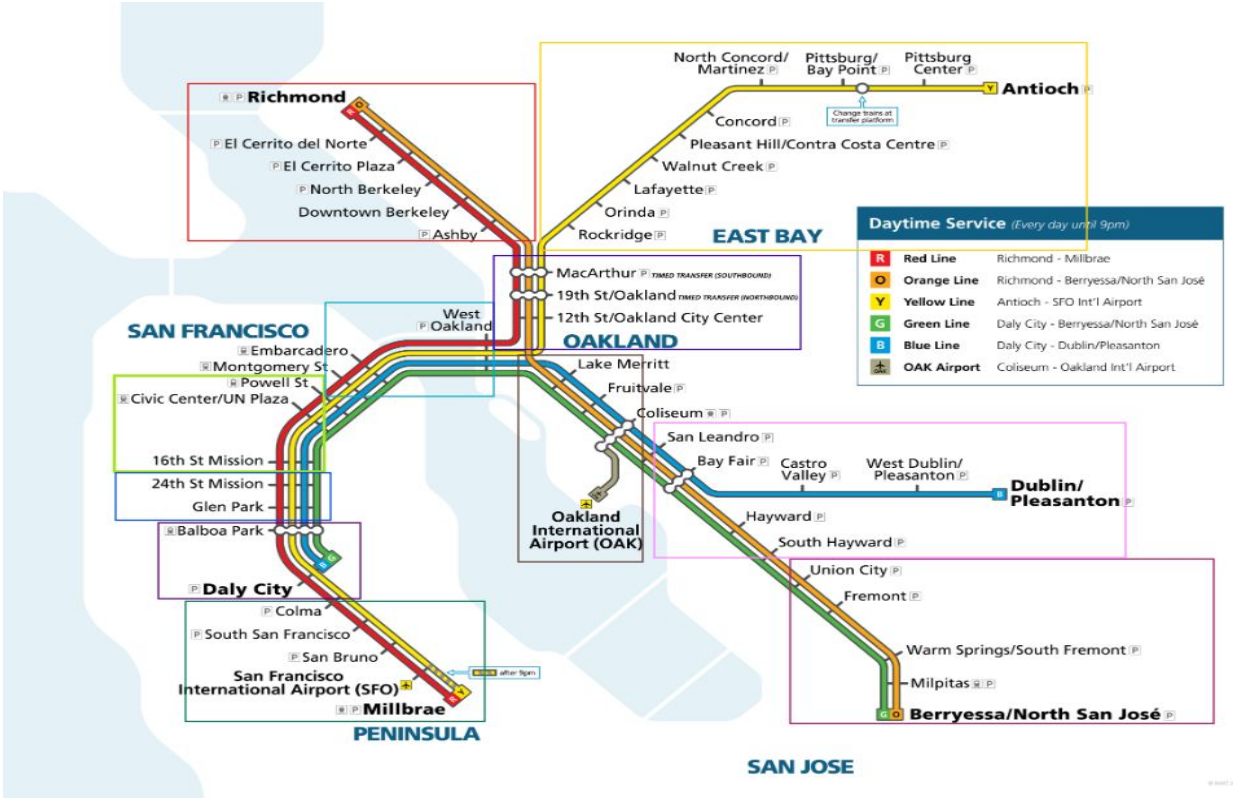
Business Application: Neo4J – Page Rank

	station_name	page_rank
0	Pittsburg Center	1.040071
1	West Dublin	1.039954
2	Pittsburg	1.032033
3	North Concord	1.031899
4	Concord	1.031897
5	Pleasant Hill	1.031897
6	Walnut Creek	1.031897
7	Lafayette	1.031895
8	Orinda	1.031779
9	Millbrae	1.026996

Page rank can help identify popular and well connected store locations or distribution points

Business Application: Neo4J – Louvain Modularity

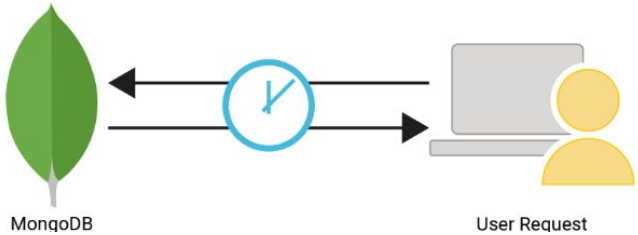
	station_name	community	intermediate_community
0	Walnut Creek	184	[184, 184, 184]
1	Rockridge	184	[170, 158, 184]
2	Pleasant Hill	184	[164, 184, 184]
3	Pittsburg Center	184	[162, 162, 184]
4	Pittsburg	184	[160, 162, 184]
5	Orinda	184	[158, 158, 184]
6	North Concord	184	[154, 154, 184]
7	Lafayette	184	[140, 158, 184]
8	Concord	184	[118, 154, 184]
9	Antioch	184	[100, 162, 184]
10	Warm Springs	182	[186, 148, 182]
11	Union City	182	[182, 182, 182]
12	Milpitas	182	[148, 148, 182]
13	Fremont	182	[132, 182, 182]
14	Berryessa	182	[108, 148, 182]

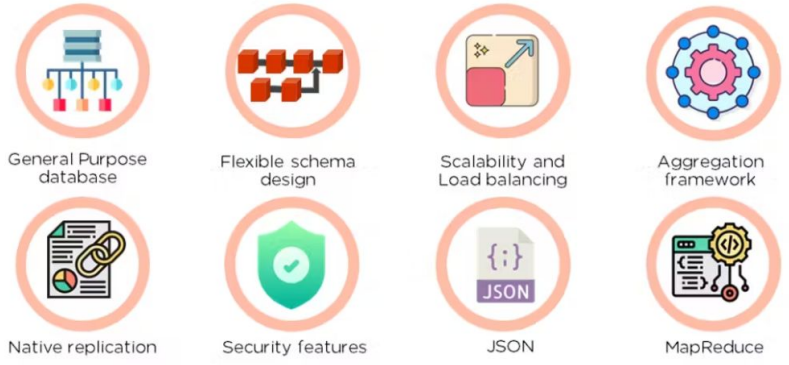


Louvain Modularity helps identify high density communities helping to pick new store locations and distribution points

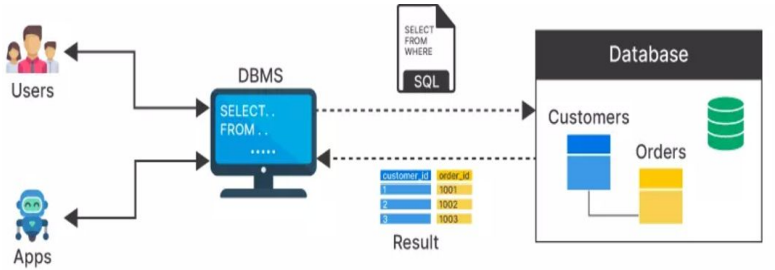
Business Application: MongoDB – High Speed Data Access

- MongoDB is a non-relational document database that provides support for JSON-like storage (document style instead of a traditional row and column format).
- Business Case: BART delivery route pre-computation system.





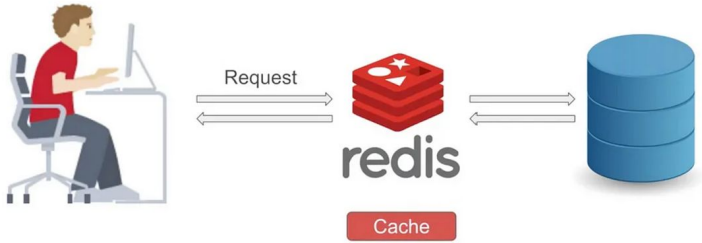
https://sarthaksavvy.hashnode.dev/everything-you-should-know-about-mongodb

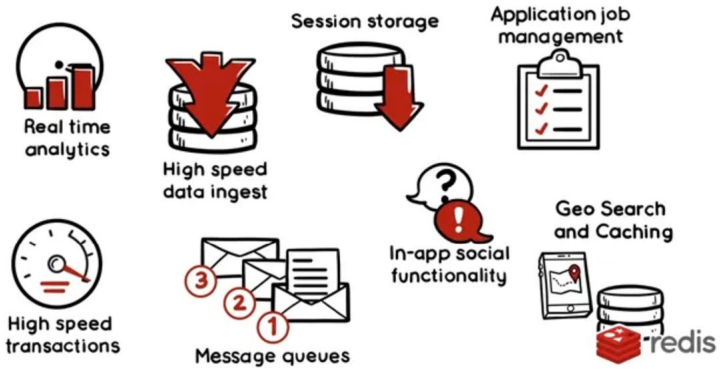


Point of view tables provide quick access to data

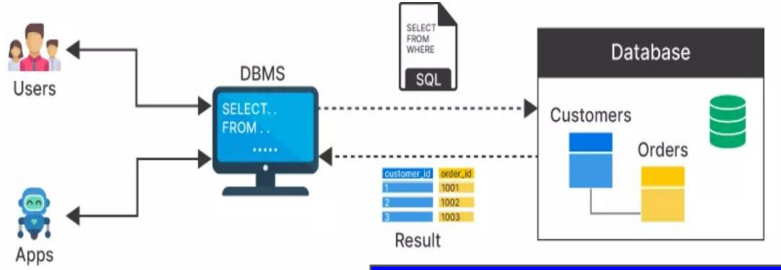
Business Application: Redis – Speedy Transactions

- Redis (Remote Dictionary Server) is an open source, in-memory, NoSQL key/value store that is used primarily as an application cache or quick-response database.
- Business case: Real Time - order taking and marketing activities





https://levelup.gitconnected.com/redis-vs-other-databases-an-in-depth-comparison-of-sql-and-nosql-solutions-7c4a9ca9183



Real-time! No.
Real-time! No.
Real-time! No.

Data stored in memory making transactions real time

Recommendations & Next Steps:

- AGM should take full advantage of NoSQL tools to modernize the data management and enable profitable growth
- Neo4j's ability to visualize the data can help build & optimize the network, which can also be used to show/tell its usefulness
- Redis manages real time sales data, which when combined with the power of MongoDB can be used to reach customers real time with marketing campaign

- **Next Steps:**

  - Initiate a six month evaluation of Neo4J in combination with MongoDB and Redis
  - Set a SMART* business goals and evaluate success against current business baseline performance
  - If evaluation is a success, implement NoSQL tools in combination with AI algorithms to create a competitive advantage over peers in the service area

NoSQL can help build the company of the future

* Specific, Measurable, Achievable, Relevant and Time-Bound