Evaluating NoSQL Models on Nuclear Energy Dataset

Group 45

Asia Montico 1966494
Beatrice Mazzocchi 1982122

SAPIENZA
UNIVERSITÀ DI ROMA

The dataset includes power plant data such as fuel type, capacity, coordinates, generation statistics, and national energy outputs. Our research questions span a range of query types:

- Aggregations by fuel type or country
- Temporal trends in generation
- Comparisons across fuel types (e.g., nuclear vs solar)
- Mortality estimates per energy source
- These questions informed our evaluation of each model’s analytical capabilities.

Mongo DB vs Neo4j

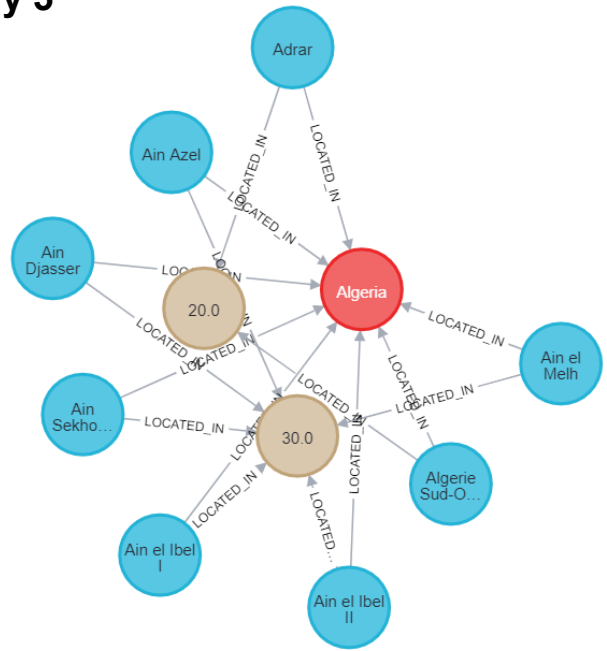
Query	Title	Techinical reasons
1	Capacity by fuel type	Requires GROUP BY and sum on attributes; no relationships between nodes
2	Top 10 countries by nuclear capacity	Aggregation and sorting; simple tabular analysis, not relational
3	Percentage of nuclear over total capacity	Needs two aggregations and a global comparison; lacks relationship logic
8	Relative error estimated vs actual	Simple calculation between two attributes; no node-to-node relation

When Neo4j Fits → Query 5: Geographic Banding

While most analytical queries are not well-suited for Neo4j, **Query 5** aligns with its strengths.

Why?
It explores the **geographic distribution** of nuclear power plants.
Relies on **explicit relationships**:

•(:PowerPlant)-[:LOCATED_IN]->(:Country)



Output:

lat_band	lon_band	plant_count	capacity_mw	countries_in_band
45–50°	0–5°	9	27890.0	{France}
35–40°	125–130°	6	23076.0	{South Korea}
35–40°	135–140°	6	16386.0	{Japan}

Aspect	Neo4j	MongoDB
Coordinate transformations	More readable (uses floor() and string concatenation)	More nested and technical ($floor, $concat, $toString, $add)
Aggregation	via WITH and RETURN	Requires multiple stages ($group, $project, $sort)
Country set handling	collect(DISTINCT ...)	$addToSet + $reduce for sorting and concatenation
Final sorting	ORDER BY	$sort
Conceptual effort	Simpler if relationships are already modeled	More verbose, but structured and flexible

From SQL Tables to MongoDB Collections

PPG	Power Plant Global
WNEG	World Nuclear Energy Gen.
DEATH	Death Rates per Energy Source
STAT	US Nuclear Statistics
UPRICE	Uranium Purchase Price
UP	Uranium Production