

Postgres DB

You are given a JSON dataset of medicines. Your task is to design and implement a **search system** using PostgreSQL that supports:

- **Prefix search** (e.g., typing Ava should return Avastin)
- **Substring search** (e.g., Injection should match all injection medicines)
- **Full-text search** (e.g., antibiotic should match relevant medicines)
- **Fuzzy Search** (typo-tolerant search) Example: typing "Avastn" (typo) should still return "Avastin".

The challenge tests your ability to design the schema, indexes, and queries efficiently while also building a benchmark to evaluate performance.

Dataset

Download from this link.

https://drive.google.com/file/d/11aYU10aRSrf1PWR8wOfrE1EoScSz_d54/view?usp=sharing

You will be provided with a JSON dump of medicines (sample included).

Each record contains fields like:

- id, sku_id, name
- manufacturer_name, marketer_name
- type, price, pack_size_label
- short_composition
- is_discontinued, available

Deliverables

Each participant/team must submit a **GitHub repository** containing:

1. **Schema & Setup**
 - SQL schema (schema.sql) with indexes
 - Data import script (import_data.py or SQL COPY command)
2. **Search API**
 - Implement a simple REST API (FastAPI) with endpoints:
 - /search/prefix?q=...

- /search/substring?q=...
- /search/fulltext?q=...
- /search/fussy?q=...

3. Benchmark Report (benchmark.md)

- Document query performance (latency, throughput, indexes used).
- Use the **fixed query set** below for benchmarking.

4. Run Instructions

- README.md with clear steps.
- Detailed description on what approach was used in order to achieve the performance.

Evaluation Plan

Each submission must upload a file named `submission.json` by running the script to evaluate `benchmark_queries.json` in the following format

```
submission.json

{
  "results":
  {
    "1": ["Avastin", "Avastin Injection"],
    "2": ["Paracetamol"],
    "3": ["Clexane Injection", "Avastin Injection"],
    "4": ["Amoxicillin", "Azithromycin", "Ciprofloxacin"],
    "5": ["Ibuprofen", "Diclofenac", "Paracetamol"]
  }
}
```