# Coding Exercise: Healthcare Cost Navigator (MVP)

Build a basic, functional web service that enables patients to search for hospitals offering MS-DRG procedures, view estimated prices & quality ratings, and interact with an AI assistant for natural language queries.

* The interface should be minimal (raw HTML or plain JSON responses, no styling).
* Please use **Python 3.11, FastAPI, async SQLAlchemy, PostgreSQL, and OpenAI API**
* **Deadline:** Please submit your deliverables within 72 hours of replying to the Next Steps email. The task should take < 4 hours with the use of LLMs / AI copilots like Claude3.5 Sonnet, or Cursor with gpt-4o.

**Deliverables**

1. **Loom Videos (2)**
   * Record a video of your *webcam and screen* as you work on the task
   * Include a 2 minute or less demo of the working application.
2. **Source Code**
   * Provide a Git repository with granular commit history
   * Include ETL script, API implementation, and database migrations
3. **README**
   * Docker Compose setup instructions
   * Database seeding instructions
   * Sample cURL commands for all endpoints
   * 5+ example prompts that the AI assistant can answer
   * Architecture decisions and trade-offs
   * Short recorded clip/GIF showing /providers and /ask endpoints working

**Features for the app to support:**

* Design and implement a database schema for:
  + Search for hospitals offering a given **MS-DRG** procedure within a radius of a ZIP code and view estimated prices & quality signals.
  + Star ratings (provider\_id as FK, rating 1-10) - create mock ratings data and join by the provider Rndrng\_Prvdr\_CCN column
* ETL script (etl.py) that:
  + Reads the provided CSV file given to you
  + Clean up the data if needed.
  + Loads data into PostgreSQL tables
* REST API endpoints:
  + GET /providers - Search hospitals by DRG, ZIP code, and radius\_km
    - Returns hospitals sorted by average\_covered\_charges
    - Implements DRG description matching using ILIKE or fuzzy search
  + POST /ask - Natural language interface
    - Accepts questions like "Who is cheapest for DRG 470 within 25 miles of 10001?"
    - Uses OpenAI to convert NL to SQL queries
    - Returns grounded answers based on database results
* AI Assistant capabilities:
  + Answer cost-related queries ("What's the cheapest hospital for knee replacement near me?")
  + Answer quality-related queries ("Which hospitals have the best ratings for heart surgery?")
  + Handle out-of-scope questions appropriately
  + Support at least 5 example prompts (document in README)

**Sample Data:**

* Use provided CMS file: **sample\_prices\_ny.csv** (15k-row NY-only sample)
* Key columns: provider\_id, provider\_name, provider\_city/state/zip\_code, ms\_drg\_definition, total\_discharges, average\_covered\_charges, average\_total\_payments, average\_medicare\_payments
* Generate mock star ratings (1-10) for each provider. Bonus points: Find and use actual Medicare star ratings.

**Download the CMS data you will be using from** [**this link**](https://catalog.data.gov/dataset/medicare-inpatient-hospitals-by-provider-and-service-9af02/resource/e51cf14c-615a-4efe-ba6b-3a3ef15dcfb0)**.**

Example Schema:

| **Column** | **Example** | **Description** |
| --- | --- | --- |
| provider\_id | 330123 | CMS ID for the hospital |
| provider\_name | CLEVELAND CLINIC | Hospital name |
| provider\_city | NEW YORK | Hospital city |
| provider\_state | NY | Hospital state |
| provider\_zip\_code | 10032 | Hospital ZIP code for radius queries |
| ms\_drg\_definition | 470 – Major Joint Replacement w/o MCC | Inpatient procedure group |
| total\_discharges | 1539 | Volume indicator |
| average\_covered\_charges | 84621 | Avg. hospital bill for the DRG |
| average\_total\_payments | 21515 | Total paid (hospital + patient + insurer) |
| average\_medicare\_payments | 19024 | Portion paid by Medicare Part A |

**Example Interactions:**

GET /providers?drg=470&zip=10001&radius\_km=40

Response: List of hospitals with knee replacement procedures, sorted by cost

POST /ask

Body: {"question": "Who has the best ratings for heart surgery near 10032?"}

Response: Based on data, Mount Sinai Hospital (rating: 9/10) and NYU Langone (rating: 8.5/10) have the highest ratings for cardiac procedures near 10032.

POST /ask

Body: {"question": "What's the weather today?"}

Response: I can only help with hospital pricing and quality information. Please ask about medical procedures, costs, or hospital ratings.

**Evaluation Criteria**

* **Database Design**: Efficient schema with proper indexes for radius queries and text search
* **ETL Implementation**: Clean data processing and loading strategy
* **Technical Skills**: Proficiency in Python, FastAPI, async SQLAlchemy, and PostgreSQL
* **AI Integration**: Effective use of OpenAI for NL→SQL conversion with proper grounding
* **Code Quality**: Clean, maintainable async Python code
* **Testing**: Working end-to-end functionality demonstration