

Argon AI Code Screen
Exercise Time: <2 hours

Problem Overview

Sarah, an executive at Pharma Co., is in charge of a drug called *Luminarex*, an immunotherapy which has not yet been approved by the FDA. Sarah's drug is currently in phase 3 of Clinical Trials, specifically targeting a disease called *Non Small Cell Lung Cancer (NSCLC)*, and Sarah hopes that her drug will be approved soon.

Sarah wants to develop a better understanding of her competitive landscape. Specifically, she wants to know about other pharma companies with clinical trials in the NSCLC space.

She currently uses [ClinicalTrials.gov](https://clinicaltrials.gov) to do her research, but is frustrated by how difficult it is to find the clinical trials she is looking for.

Note: NSCLC has many different representations in the dataset. For example, it could be "non small cell lung cancer", "non small cell lung carcinoma", "NSCLC", "carcinoma of the lungs, non small cell", etc. How do we capture all the relevant clinical trials for searches on any disease?

Task

Your task is to build an application that can help Sarah do research on her competitive space.

Using this dataset, build a web application that allows Sarah to search through the dataset and list out the titles of relevant clinical trials.

You are allowed to use any framework you want. Feel free to load the dataset directly into memory so that you don't have to set up a DB. You do not need to deploy your code.

Here's an example of an interface that could work, but you are welcome to come up with any product experience that you think makes sense.

Once completed, email over your code in a .zip format or share your github repo with recruiting@argon-ai.com.

The image shows a simplified user interface for searching clinical trials. At the top, there is a light gray rectangular box containing the text "non small cell lung cancer". To the right of this box is a red button with the word "Search" in white. Below the search box, there are five stacked red rectangular buttons, each containing the text "Clinical Trial Title #1" through "Clinical Trial Title #5" respectively. The entire interface is set against a light gray background.

Example UI

Bonus points:

- *How do we allow her to search for NSCLC trials -AND- immunotherapy related drugs?*
- *How would you deploy your software?*
- *What are the alternatives to loading the dataset into memory, and why would you want to use those alternatives?*
- *How do we evaluate completeness of results?*

Dataset

1. Go to <https://clinicaltrials.gov/search>
2. Click on the download icon (in between "None Selected" and "Bookmark")

Search Results

Viewing 1-10 out of 504,145 studies

Showing results for: **All**

Sort studies by  Relevance 

None Selected





|

<input type="checkbox"/>	Study Title
--------------------------	-------------

- a.
3. Select either CSV or JSON, whichever your choice
 4. Select all data fields
 5. Download the **top 10,000** trials by scrolling to the bottom and clicking “Download”