Hippocrates Analytics: Clinical Trial Matching Engine

Live Dashboard: [Link to your published Power BI dashboard]

Project Overview

Hippocrates Analytics is a data analysis prototype designed to solve a critical bottleneck in healthcare: matching oncology patients with relevant clinical trials. Manually searching for trials is a time-consuming process for clinicians. This project automates that workflow by using a Python-based matching engine and a dynamic Power BI dashboard.

The tool processes real-world clinical trial data from ClinicalTrials.gov and matches it against a dataset of mock patient profiles based on specific criteria like diagnosis, biomarkers, and age. The result is a powerful, interactive dashboard that allows users to instantly view potential trial matches for each patient, potentially reducing search time by over 95%.

This project demonstrates a practical, end-to-end data analysis workflow, from data sourcing and cleaning to logical matching and professional-grade visualization.

***Tech Stack***

**Data Analysis:Python (Pandas)**

**Data Sourcing: ClinicalTrials.gov (CSV Export)**

**Data Visualization & Dashboard:\*\* Microsoft Power BI**

**Features**

Automated Matching Engine: A Python script that filters thousands of clinical trials based on complex rules (patient condition, biomarkers, age).

Interactive Dashboard:A multi-page Power BI report that allows for deep-dive analysis and interactive filtering.

Executive Summary Page: High-level KPIs for administrators, including total patients, total trials, and successful match rates.

Patient Analysis Page: A detailed view of patient demographics and clinical data.

Clinical Trial Scout Tool:The core feature of the dashboard—click on a patient to instantly see a list of their matched clinical trials.

-‘

**How to Run This Project**

To run this project on your local machine, follow these steps:

Prerequisites

Python 3.x installed

Pandas library installed (`pip install pandas`)

Microsoft Power BI Desktop installed

**Setup**

1. Clone the repository:

```bash

git clone [https://github.com/your-username/your-repo-name.git](https://github.com/your-username/your-repo-name.git)

```

2. Navigate to the project directory:

```bash

cd your-repo-name

```

3. Download the Clinical Trial Data:

Go to [ClinicalTrials.gov](https://classic.clinicaltrials.gov/ct2/search) and search for recruiting trials (e.g., for "Glioblastoma" in the "United States").

Download the results as a CSV and save it in the project folder.

4. Run the Python Scripts:

First, clean the downloaded trial data:

```bash

python clean\_trials\_data.py

```

Next, generate the mock patient data:

```bash

python create\_patients.py

```

Finally, run the matching engine:

```bash

python trial\_matcher.py

```

5. View the Dashboard:

\* Open the `Hippocrates\_Dashboard.pbix` file in Power BI Desktop.

\* Click "Refresh" on the Home ribbon to load all the newly generated CSV data.

**Project Structure**

├── Hippocrates\_Dashboard.pbix # The Power BI dashboard file

├── create\_patients.py # Script to generate mock patient data

├── clean\_trials\_data.py # Script to clean the raw clinical trial data

├── trial\_matcher.py # The core matching engine script

├── patient\_diagnostics\_mock.csv # (Generated) Mock patient data

├── cleaned\_clinical\_trials.csv # (Generated) Cleaned trial data

├── successful\_matches.csv # (Generated) Final patient-trial matches

└── README.md # This README file