Project Proposal: Team 1

**Group:**

* Ranjani, Saiana, Lauren

**Background:**

Our team was interested in exploring healthcare data. We extracted pharmaceutical spending data across the globe to determine how much each country is spending on off the shelf pharmaceutical medicines. Data was extracted from CSV and JSON sources, transformed using Python, Pandas and SQL. Data was then loaded into PostgreSQL. SQLAlchemy was used with Flask to deploy results to a pharmaceutical web application.

**Extraction:**

* In **Pharmaceutical Spending**, we will be pulling % of pharmaceutical spending for countries across the globe.
* In **Population**, we will be retrieving population data for countries across the globe.
* Datasets will be merged using **country code.**

|  |  |  |  |
| --- | --- | --- | --- |
| **DATA** | **LINK** | **TYPE** | **SIZE** |
| **PHARMACEUTICAL\_SPENDING** | [**https://data.oecd.org/healthres/pharmaceutical-spending.htm#indicator-chart**](https://data.oecd.org/healthres/pharmaceutical-spending.htm#indicator-chart) | **CSV** | **240 KB** |
| **POPULATION** | [**https://datahub.io/core/population#pandas**](https://datahub.io/core/population#pandas) | **JSON** | **1 MB** |

**Transformation:**

* Python
* Pandas
* SQL

**Loading:**

* Database: PostgreSQL

**Presentation:**

* HTML file using Flask