**Team Members: Project Group 4**

* Emily Bowers
* Kelly Carter
* Joe Cotton
* Cory Lingerfelt

**Project Description/Outline:**

The safety and efficacy of pharmaceutical products are critical concerns for regulatory agencies like the U.S. Food & Drug Administration (FDA). When issues arise that jeopardize the health of consumers, the FDA intervenes through recall actions, either voluntary or mandated, to mitigate risks. By examining drug recall enforcement reports since XXXX, we can gain insights into the trends, patterns, and impacts of these recalls on public health and regulatory oversight.

**Questions Addressed with Dataset:**

1. **Temporal Trends:**
   * Analysis of drug recall enforcement reports over a certain period of time. Determine trend of frequency and severity of recalls.
   * Question: Certain years or timeframes may exhibit spikes in recall activity.
   * Question: Heat map of U.S. sates with highest frequency of recall activity.
2. **Product Characteristics:**
   * Volume & frequency of product types recalled
   * Volume & frequency of prescription vs over-the-counter drugs recalled
   * Distribution of drug recall by drug route, classification
3. **Recall Characteristics:**
   * Determine if there are common reasons for drug recalls, i.e. contamination, labeling errors, manufacturing defects, etc.
   * Are specific companies reporting and/or experiencing recalls more than others
   * Determine distribution of voluntary vs mandated recalls
   * Volume & frequency of Reasons for recall,
   * Distribution of Severity of recall classification

**Dataset(s) to be Utilized:**

OpenFDA Drug Recall Enforcement Reports (/drug/enforcement)

(<https://open.fda.gov/apis/drug/enforcement/download/>)

**Rough Breakdown of Tasks:**

* Clean and organize the data, check for null data and duplicate data
  + Use Python to clean and format dataset
  + Inspect dataset in Pandas data frame
* Create a database in SQL – Postgres SQL
* Data Analysis
  + Create code & run code
  + Creating and saving csv or JSON file
* Visualization
  + Create & run code using Python
  + Flask backend with Interactive API routes that serve back Python or JavaScript created plots
* Delivery
  + Include a method for reading the data from the database and displaying it for future use with either Pandas DataFrame or Flask API with JSON output
* Create a write-up summarizing major findings and implications.
  + An overview of the project and its purpose
  + Instructions on how to use and interact with the project
  + At least one paragraph summarizing efforts for ethical considerations made in the project
  + References for the data source(s)
  + References for any code used that is not your own
* Presentation
  + Slide presentation