Drug Reaction

Presented by:
Kanisha Shah, Alena Chan, Jada Dixon, Raiymbek Kokteubay

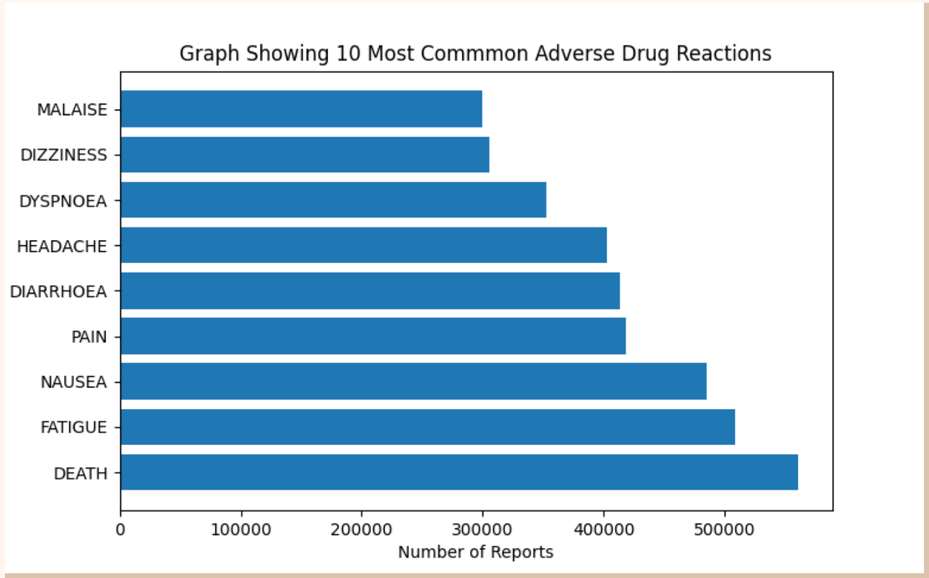
Patient Safety Track : Drug Adverse Events

Goal: Decrease rate of avoidable ADRs

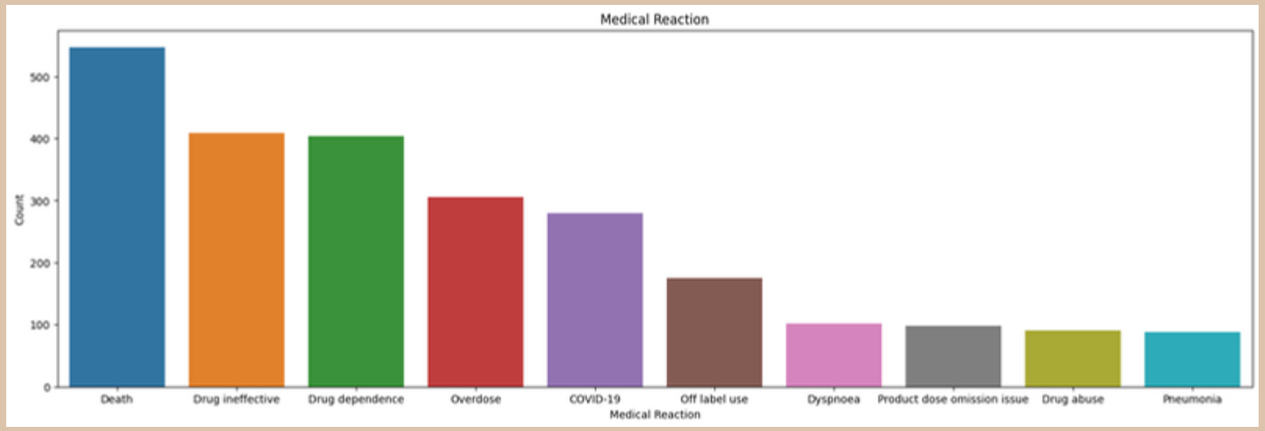
1. Analyze patterns in severe drug reaction and use this statistical evidence to inform medical decisions.

2. Train a classifier as a tool to predict possibility of a severe reaction based on drug and patient information.

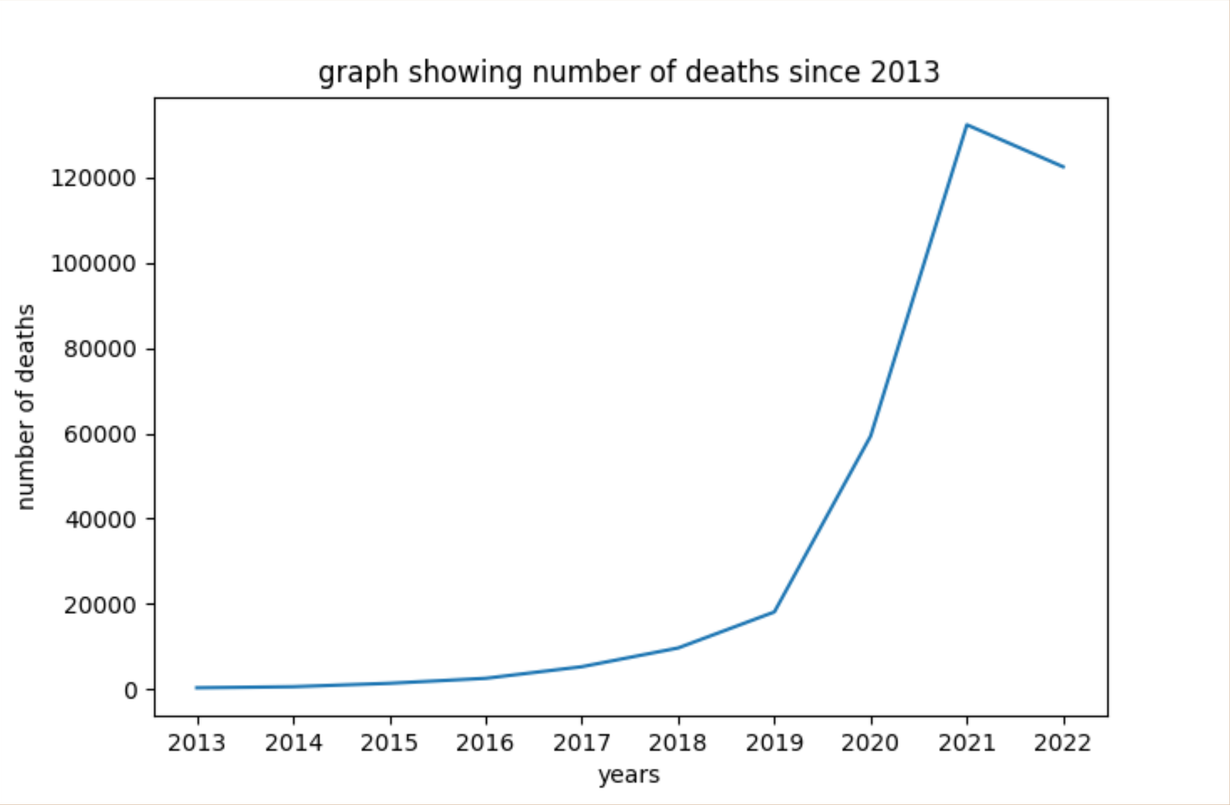
Possible Adverse Drug Reactions (ADRs)



Taking a drug always comes with a risk of possible unintended side effects. The graph shows some reports of adverse side effects dating from 2013 onwards.



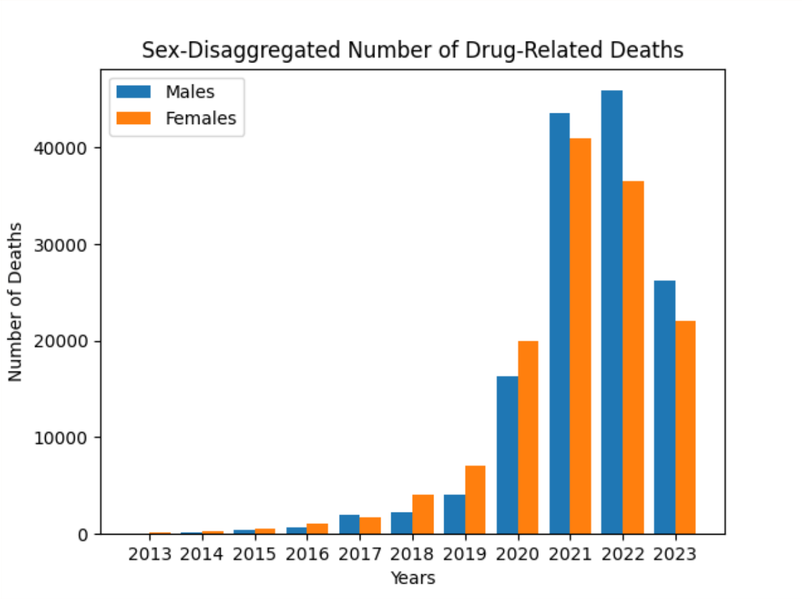
Trends in ADRs over time



The number of adverse reactions to drugs have been trending upwards since 2013, and so have the number of deaths.

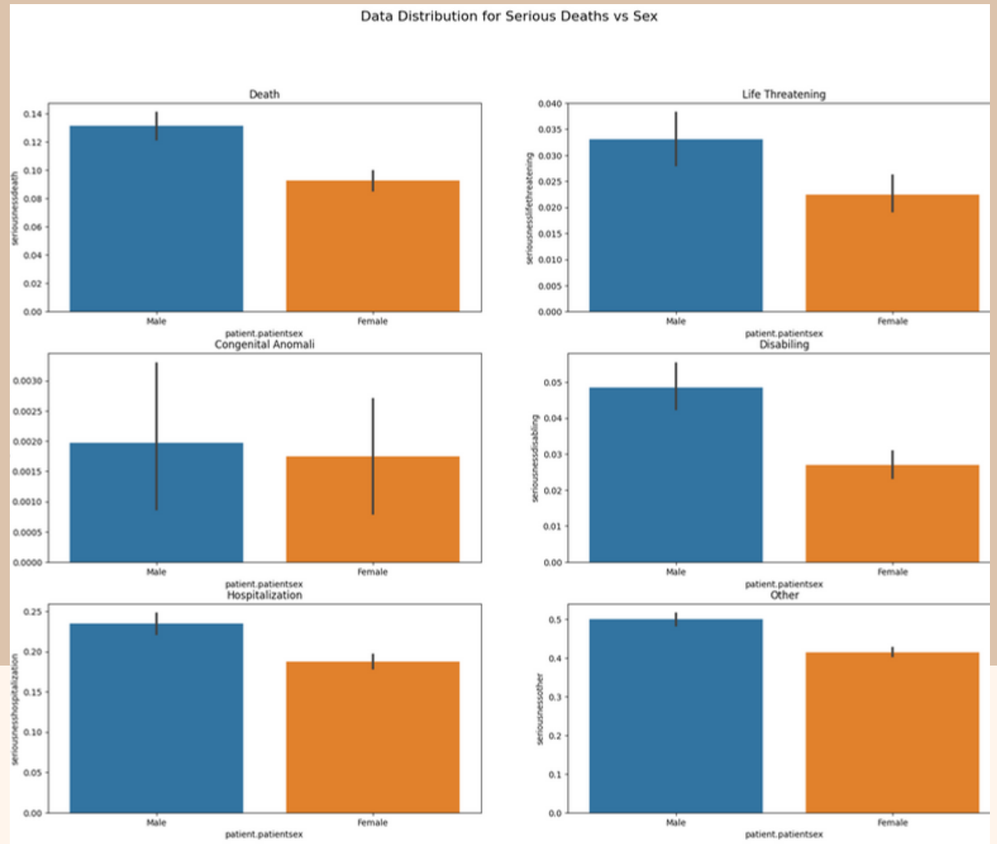
This led us to investigate if there are any factors correlating to likelihood of having a serious adverse reaction.

Trends in ARDs -Demographics

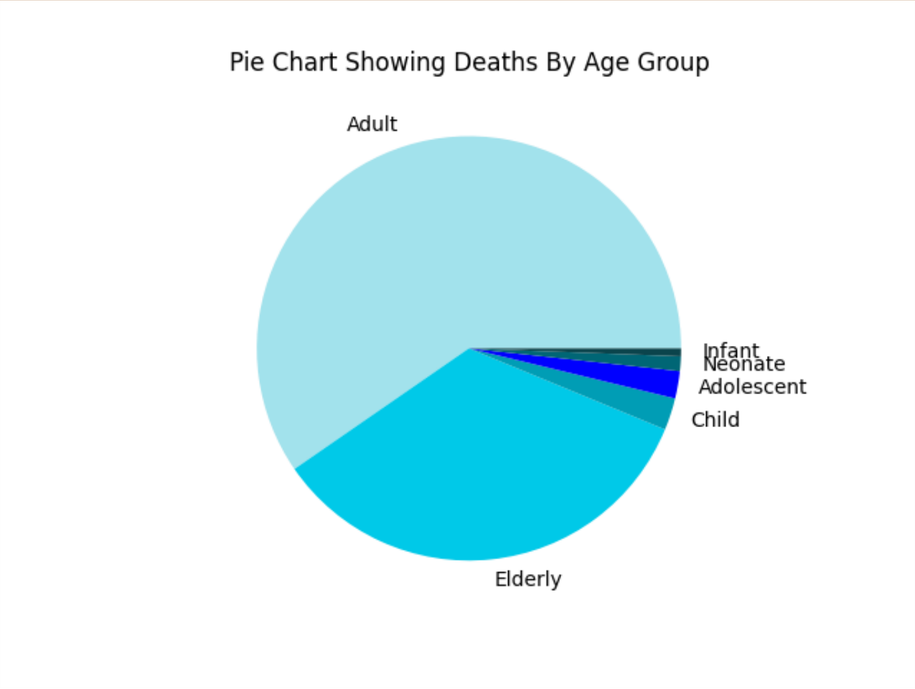


We found there to be no significant difference between drug-reported death rates of men and women.

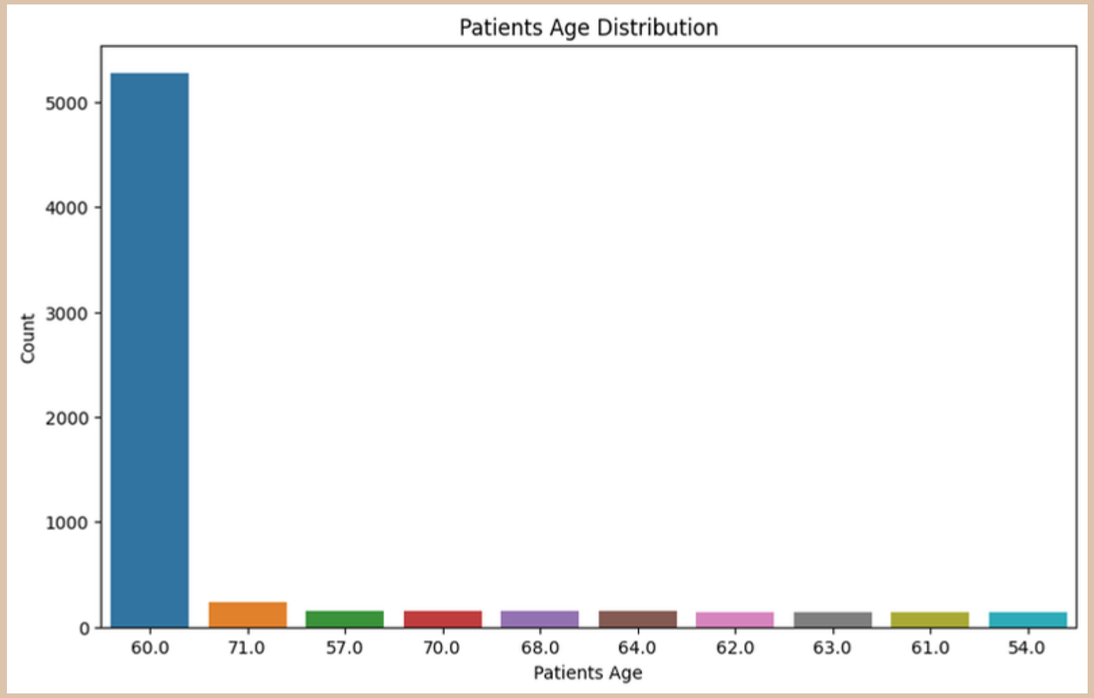
However, processed data demonstrates gender differences. Males had a slightly higher chance of severe side effects, including hospitalization and death.



Trends in ADRs - Demographics

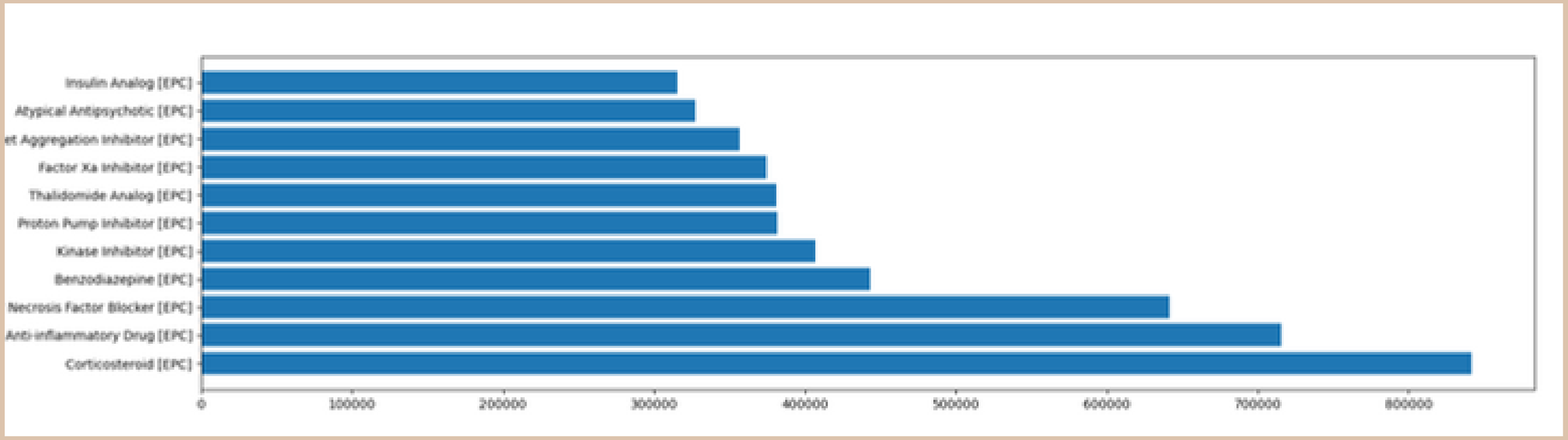


Adults had the highest number of deaths, followed by the elderly.



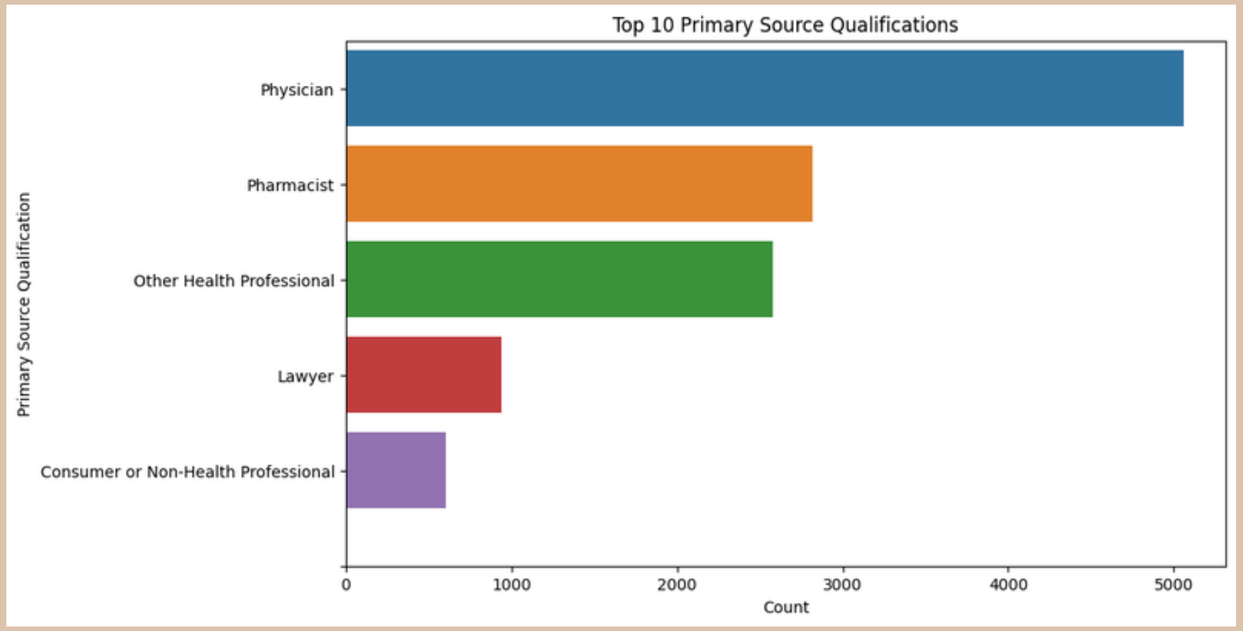
However, processed data shows more detail: most of the people affected were adults between 54 and 71, with a majority being 60 years old.

Trends - Drug Classes and Death Reports



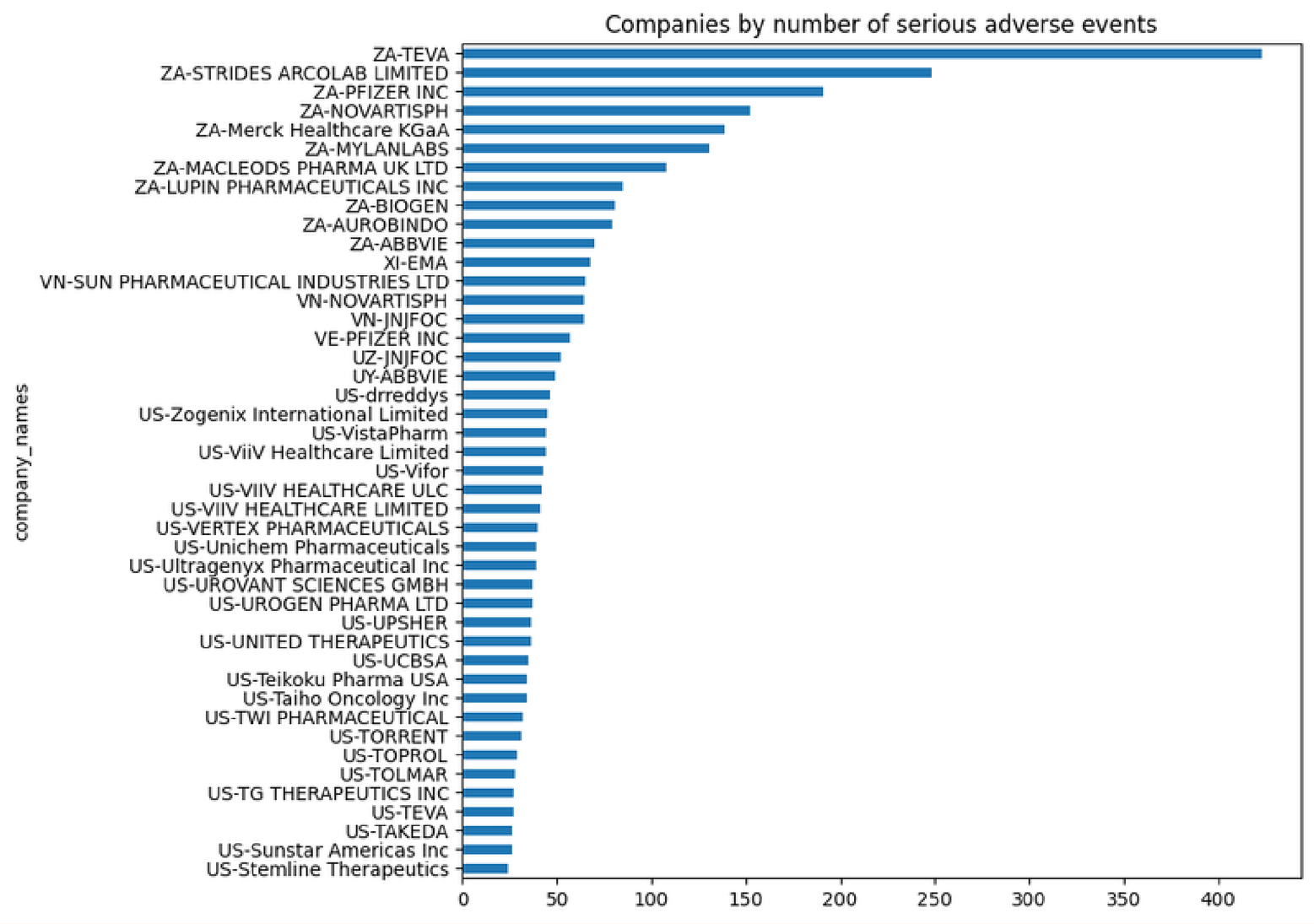
Corticosteroids were found to have the highest amount of associated deaths, followed by Anti-Inflammatory Drugs.

Trends in ADRs - Reports



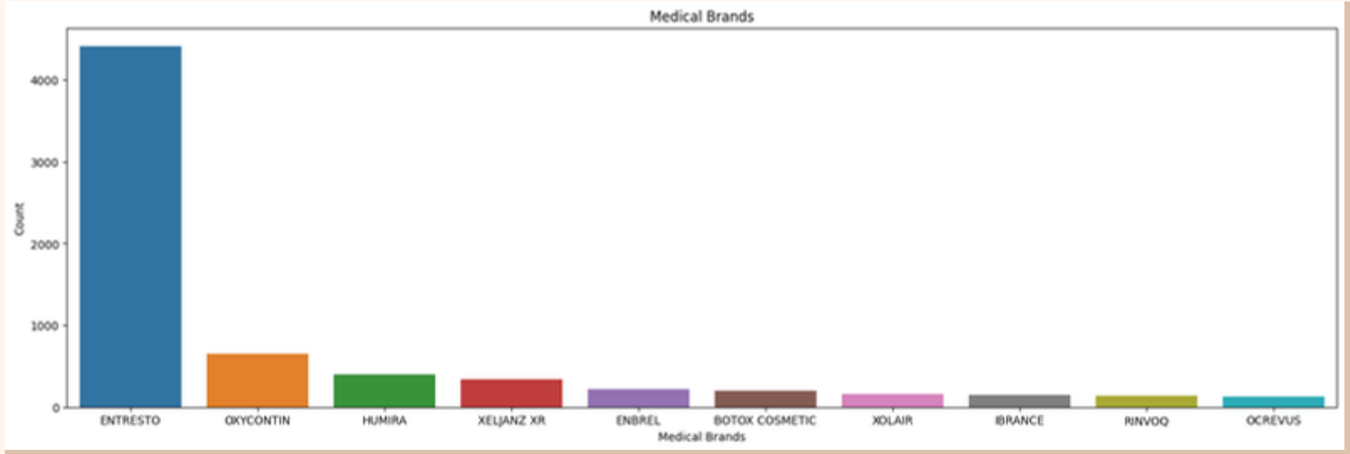
Physicians were major source of reports.

Trends in ADRs - Manufacturers



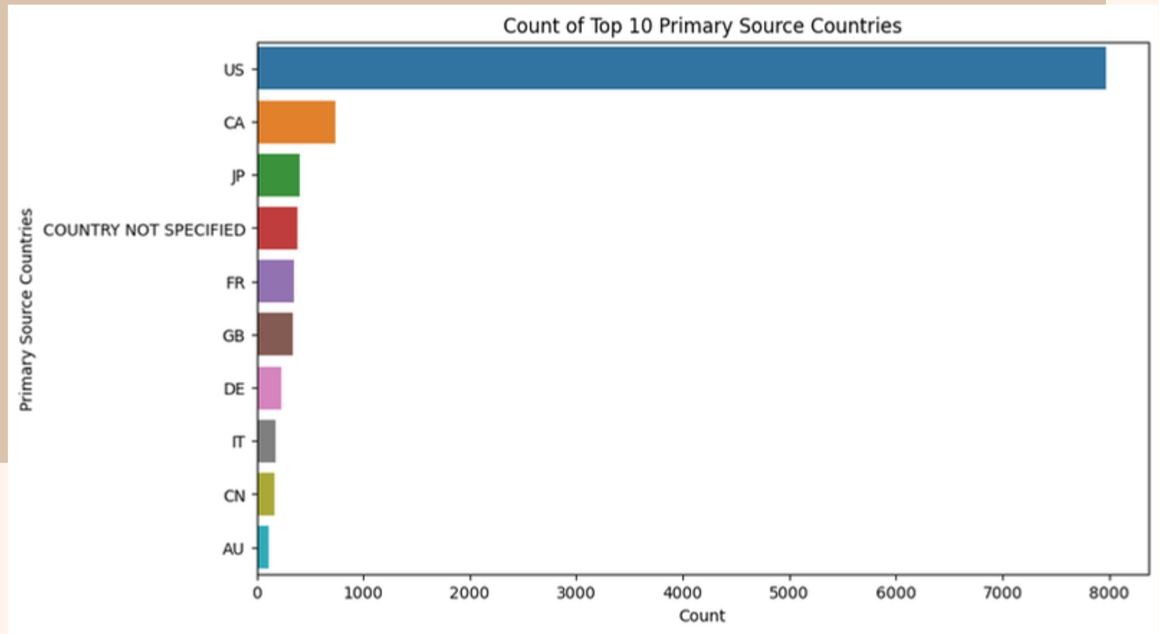
ZA-Teva's produce was responsible for most number of serious events.

Trends in ARDs - Outcomes



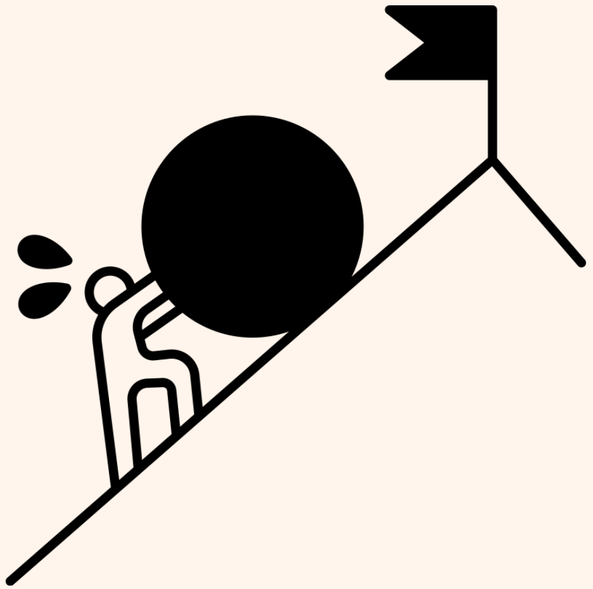
ENTRESTO was the brands who produced the most lethal outcomes

Most of reported ADR were associated with US-produced drugs.



Challenges

- Poorly structured dataset
- Duplicate labels
- Hard to use data format



Pre-Processing Data

- We have **dropped unwanted columns** after proper analysis
- Handled **Null values** by replacing with its mode value
- We noticed a trend in **occurcountry** and **companynumb** values. We combined information to remove redundancy
- We **filled missing values** in the 'patient.patientonsetage' column based on the mode of the corresponding age group in 'patient.patientagegroup'.
- **Mapped and visualized** data

Classifiers

CatBoost

XG Boost

Random Forest Classifier

Metrics

- Label Encoder
- Train Test Split (80-20)%
- Accuracy & F1 score as Metric

Variables

Target Variable: Serious

Input Variables: Occurring Country, Patient age, Patient weight, Patient sex, Brand Name, Medicinal product, Reaction Outcome, Reaction Med Arp

Results

Result 1

We identified common **trends** in ADR statistics using our **cleaned dataset**.

Result 2

We trained a classifier to **predict** possibility of serious health **hazard** associated with a drug.

Accuracy and MAE

Accuracy

CatBoost : 0.86

XGBoost : 0.86

Random Forest : 0.83

Mean Absolute Error

CatBoost : 0.14

XGBoost : 0.14

Random Forest : 0.17

Applications

- **Minimize medical error**: based on our understanding of causes of adverse drug effects, we can direct inspector authorities to questionable manufacturers, as well as inform patients.
- Reduce **distrust in medical care** that followed the pandemic.
- Reduce **bias** in medicine.

Thankyou