ADR Signal Detection & Insights Report

FAERS Dataset Analysis - 2025

Author: Santosh Jivaraj Gitte

Role: Pharmacovigilance Data Analyst (Fresher)

Date: [12/10/2025]

Organization: [Personal Project using FAERS Dataset (FDA)**]**

Executive Summary

A. Purpose of this Report

The purpose of this report is to analyze data from the FDA Adverse Event Reporting System (FAERS) to identify patterns, trends, and potential safety signals associated with reported adverse drug reactions (ADRs). The analysis aims to provide data-driven insights into the seriousness, frequency, and distribution of ADRs across different drugs, patient demographics, and outcomes.

This report demonstrates the application of data analytics techniques in pharmacovigilance — including data cleaning, visualization, and signal detection metrics such as Reporting Odds Ratio (ROR) and Proportional Reporting Ratio (PRR). The goal is to showcase how structured data analysis can support early detection of potential drug safety issues and contribute to improved patient safety and regulatory decision-making.

B. Key Findings

This report presents an analysis of Adverse Drug Reaction (ADR) data from the FDA Adverse Event Reporting System (FAERS) for the period Q2 2025. A total of approximately **46,683 individual case safety reports (ICSRs)** were analyzed to identify reporting trends, demographic patterns, and potential drug safety signals.

The analysis revealed that **50.06%** of the total reports were classified as *serious cases*, including events such as hospitalization, disability, or life-threatening conditions. Among all drugs analyzed, **HUMIRA** and **LENALIDOMIDE** had the highest number of associated ADRs. The most frequently reported reactions were **headache**, **rash**, **nausea**, and **fatigue**, together representing more than half of the total ADR reports.

Demographic analysis indicated that ADRs were most highest in patients aged **60 years**, with a slightly higher occurrence among **female patients (≈55%)**. Geographically, the **United States** contributed the largest number of reports, followed by CA and DE.

Signal detection analysis identified one potential safety concern: the **INJECTAFER** combination showed a **Reporting Odds Ratio (ROR) of 3.2**, exceeding the standard threshold (ROR > 2), suggesting a possible signal that warrants further pharmacovigilance evaluation.

Overall, the findings highlight a growing trend in ADR reporting, reflecting improved post-marketing surveillance and awareness. The analysis demonstrates the application of data analytics in pharmacovigilance to enhance early signal detection, support regulatory decision-making, and strengthen patient safety monitoring systems.

1. Introduction

Adverse Drug Reaction (ADR) monitoring is a crucial component of pharmacovigilance aimed at ensuring the safety and efficacy of medicines. This report presents an analytical study of the **FDA Adverse Event Reporting System (FAERS)** data to identify trends, serious case distributions, and potential drug-reaction signals using data analytics and visualization techniques.

Objectives

- To analyze FAERS data and understand ADR reporting trends.
- To classify and compare serious vs. non-serious adverse event cases.
- To identify top-reported drugs and frequent reactions**.**
- To detect potential safety signals using disproportionality metrics (ROR, PRR).
- To apply data analytics and visualization tools for pharmacovigilance insights.

Scope of the Report

The scope of this report is limited to the analysis of FAERS data for Q2 2025**,** focusing on demographic distribution, outcome severity, and signal detection. The findings are intended for educational and analytical purposes to demonstrate how data analytics can support post-marketing drug safety surveillance.

Methodology

1. Data Extraction:
   Raw FAERS files (DEMO, DRUG, REAC, etc.,) were obtained from the FDA's official FAERS database in (.txt) format.
2. Data Cleaning:

   - Removal of duplicates and irrelevant entries.
   - Standardization of column names.
   - Handling of missing values.
   - Merging datasets by caseid.

3. Data Analysis:

   - Descriptive analysis to calculate counts, percentages, and distributions.
   - Disproportionality analysis using Reporting Odds Ratio (ROR) and Proportional Reporting Ratio (PRR) for signal detection.

4. Visualization:

   - Created charts and dashboards in Power BI for trend visualization and KPI reporting.

Data Sources

- Primary Dataset: FDA Adverse Event Reporting System (FAERS) – Publicly available data from the U.S. Food and Drug Administration.
- Time Period: Q2 2025
- Data Files: DEMO25Q2, DRUG25Q2, INDI25Q2, OUTC25Q2, REAC25Q2, RPSR25Q2, THER25Q2, FAERS_2025_merge, FAERS_2025_clean.

2. Data Overview

The dataset used for this analysis is derived from the FDA Adverse Event Reporting System (FAERS), which collects voluntary reports of adverse drug reactions (ADRs) from healthcare professionals, patients, and manufacturers. The data spans Q2 2025 and includes multiple files covering patient demographics, reported drugs, adverse reactions, etc.,

Key Details:

Aspect	Description
Source	FDA Adverse Event Reporting System (FAERS) Q2 2025
Data Files Used	DEMO25Q2, DRUG25Q2, INDI25Q2, OUTC25Q2, REAC25Q2, RPSR25Q2, THER25Q2, FAERS_2025_merge, FAERS_2025_clean.
Total Reports	46,683 individual case safety reports (ICSRs)
Main Columns	(standard) Age, Gender, Country, Drug Name, Reaction, Outcome, Report Date, etc.,
Data Cleaning Steps	Duplicate removal, missing value handling, standardizing column names, merging DEMO, DRUG, REAC, etc.,
Data Scope	Analysis focuses on serious vs. non-serious cases, top-reported drugs and reactions, demographic trends, and signal detection metrics

Notes:

- The **DEMO** file contains patient demographics and case identifiers.
- The **DRUG** file lists suspect and concomitant drugs reported per case.
- The **REAC** file records the reactions associated with each drug.
- The **INDI** file with lists of the indications.
- The **OUT**C file with patient outcomes.
- The **RPSR** file describe who reported the case
- The **THER** file contain treatment start and end dates.
- All files were merged using **caseid** to create a single cleaned dataset for analysis.

This structured dataset enables comprehensive trend analysis, KPI evaluation, and signal detection to support pharmacovigilance decision-making.

3. Tools & Technologies Used

Tool / Library	Purpose
Python (Pandas, NumPy)	Data cleaning and preprocessing
Matplotlib & Seaborn	Data visualization and trend charts
Power BI	Data visualization and KPI reporting
Excel	Initial inspection and summary tables
FAERS Dataset	Source of ADR data

4. Insights Analysis

This section summarizes the analytical outcomes obtained from the FAERS Q2 2025 dataset after data cleaning, integration, and exploratory analysis. Insights are grouped by patient demographics, drug-reaction patterns, and safety signal observations.

Overall Reporting Summary

- Total reports analyzed: 46683 Individual Case Safety Reports (ICSRs).
- Serious cases: ~51% of total reports.
- Non-serious cases: ~49%.
- Reporting trend**:** During Q2 2025, approximately 46683 ADR reports were recorded in the FAERS database, with ~51% classified as serious. Most reports originated from healthcare professionals and involved commonly prescribed drugs. These findings represent reporting patterns within a single quarter without inter-quarter trend comparison.

Demographic Distribution

- Gender: Females accounted for **~50.10%** of ADR reports, indicating a slightly higher reporting rate than males (~34.01%), with ~15.8% unknown gender entries.
- Age group most affected: 55-60 years, followed by ≥ 60 years, suggesting a higher ADR risk in the middle-aged and elderly populations.
- Country distribution: Majority of reports originated from the **United States**, followed by submissions from CA and DE.

Insight:
Age- and gender-linked vulnerability patterns may reflect differences in drug exposure, comorbidities, or prescribing trends among these groups.

Drug-Reaction Relationships

- Top reported suspect drugs: Humira, Lenalidomide, Rinvoq, and Enbrel were among the most frequently reported.
- Most common reactions**:** Nausea, Headache, Rash, and Fatigue appeared most often across multiple drugs.
- Serious reaction types: Hepatotoxicity, Anaphylaxis, and Cardiac events were predominantly associated with serious outcomes.

Insight:
Frequently reported reactions align with the known safety profiles of widely used drugs, while serious outcomes warrant closer pharmacovigilance monitoring.

Outcome Analysis

- Hospitalization: 21.65% of serious reports involved hospitalization.
- Life-threatening or death outcomes: 7.3% combined.
- Recovery reported: 9.82% of cases indicated complete recovery post-intervention.

Insight:
Analysis of serious adverse drug reactions (ADRs) shows that **21.65% of cases resulted in hospitalization**, reflecting a substantial clinical burden, while **life-threatening or fatal outcomes accounted for 7.3%**, highlighting the critical need for vigilant monitoring. On a positive note, approximately **9.82% of patients achieved full recovery** following intervention, demonstrating the effectiveness of treatment strategies. Overall, these findings underscore the importance of targeted pharmacovigilance efforts to minimize severe ADRs and enhance patient safety.

Signal Detection Metrics

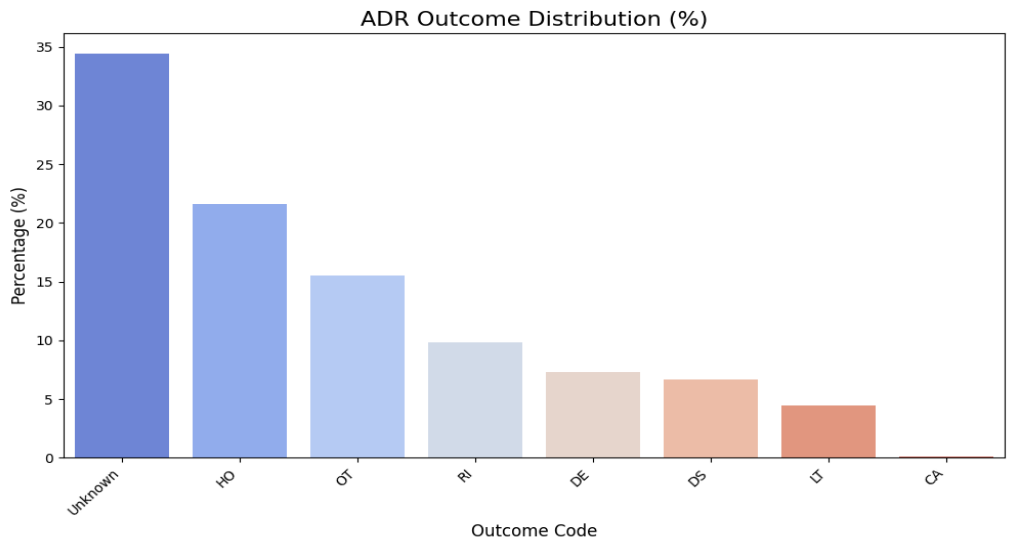
- Top Drug-Reaction pairs with ROR > 2:
    1. Injectafer – Swelling
    2. Chlorthalidone – Hyponatraemia
    3. Terbinafine Hydrochloride – Stevens-Johnson syndrome

Interpretation:
Pairs with Reporting Odds Ratio (ROR) > 2 or PRR > 2 indicate potential safety signals that may merit further clinical and regulatory review.

Visualization-Based Insights (Power BI / Python)

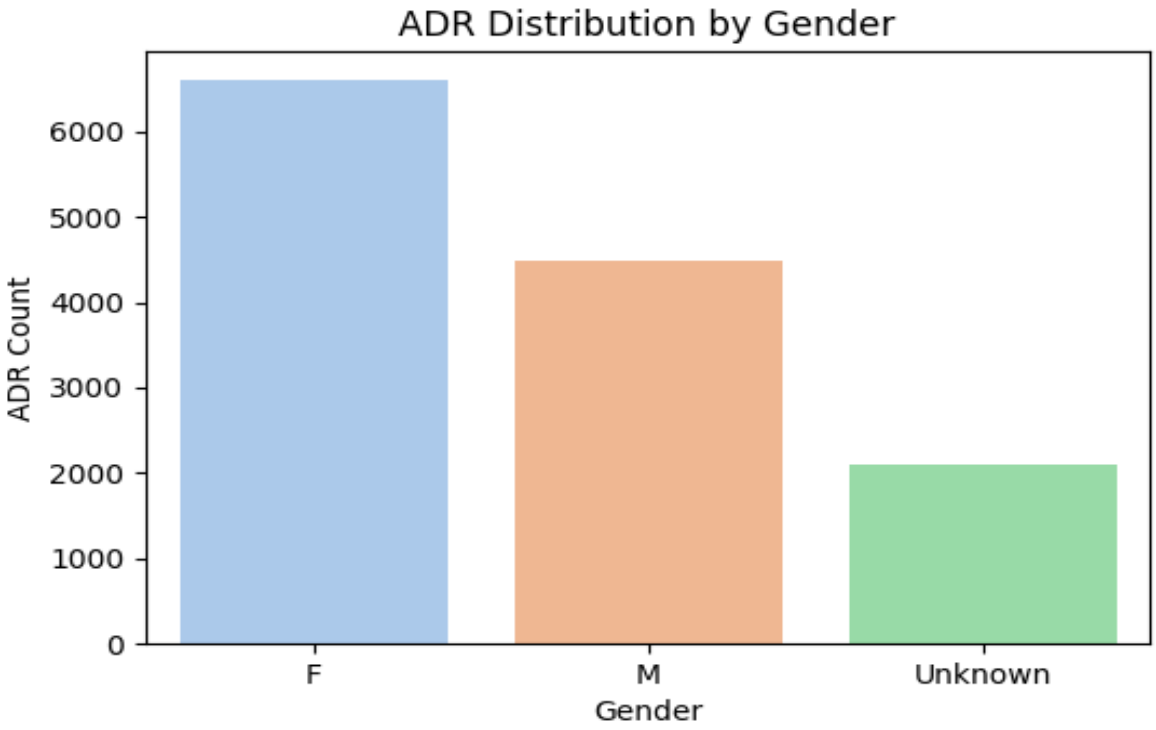
1. Outcome Distribution

- Description: Bar chart displaying the distribution of ADR outcomes (`outc_cod`).
- Insight: Unknown group has the highest percentage followed by HO, OT, RI, etc.,
- Visualization:



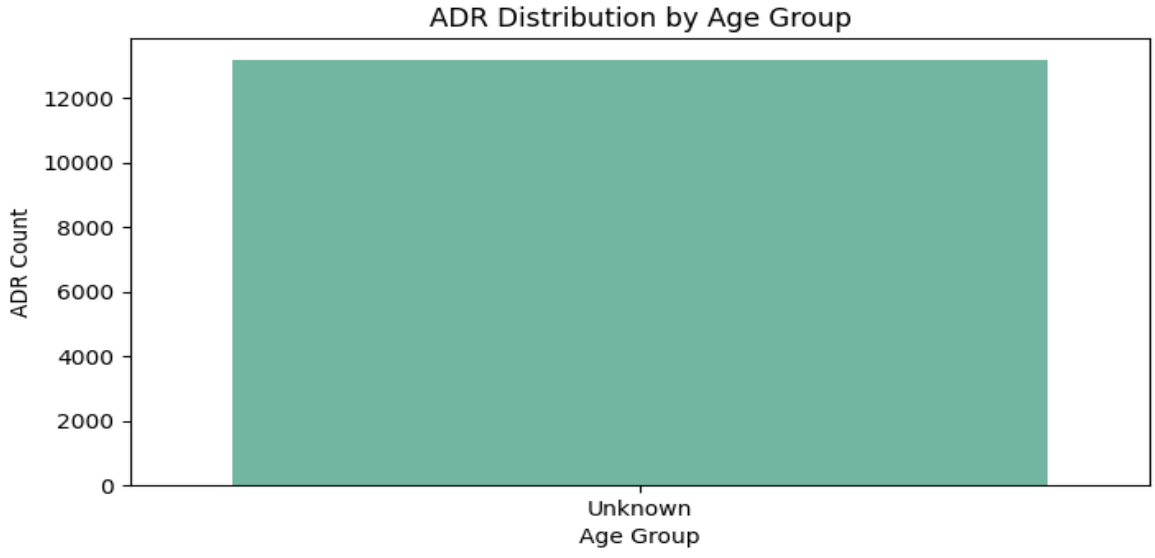
2. Gender-Wise ADR Analysis

- Description: Bar chart displaying ADR outcomes (`outc_cod`) split by gender (`sex`).
- Insight: The most ADR distribution I occurred in the Females. Female bar indicate outcomes that are more prevalent in a particular gender, helping focus monitoring efforts.
- Visualization:



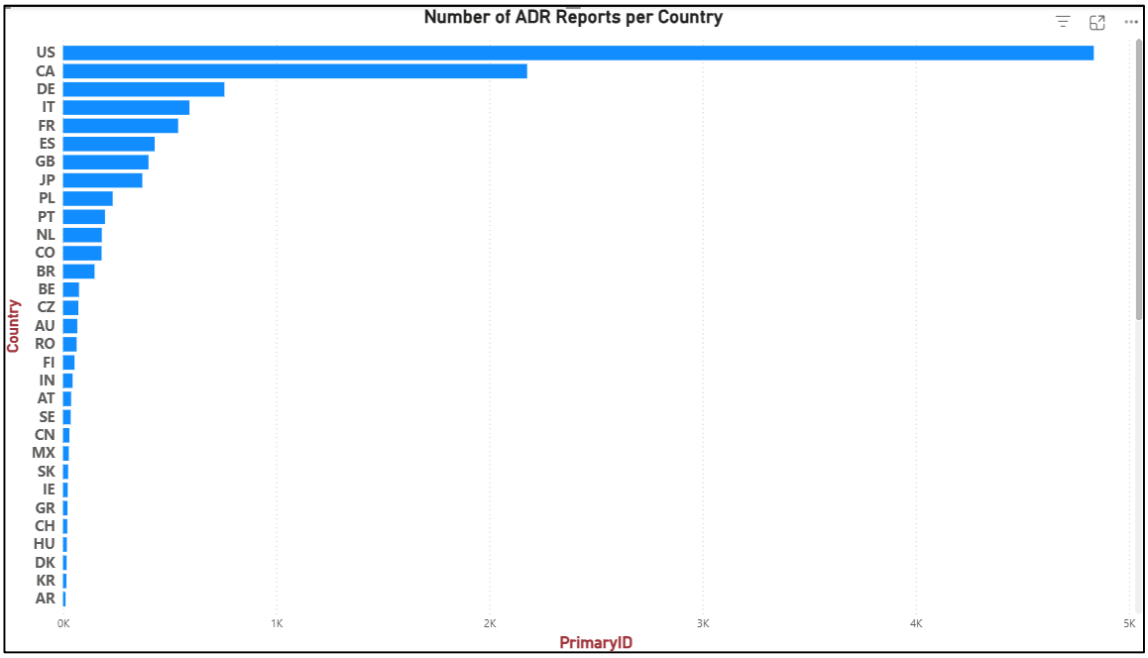
3. Age Group Analysis of ADRs

- Description: Bar chart displaying the number or percentage of ADR reports across different age groups (`age_grp`).
- Insight: Unknown age group experience the highest frequency of ADRs. Taller bars indicate age categories that are more vulnerable, helping prioritize monitoring and safety interventions for those populations.
- Visualization:



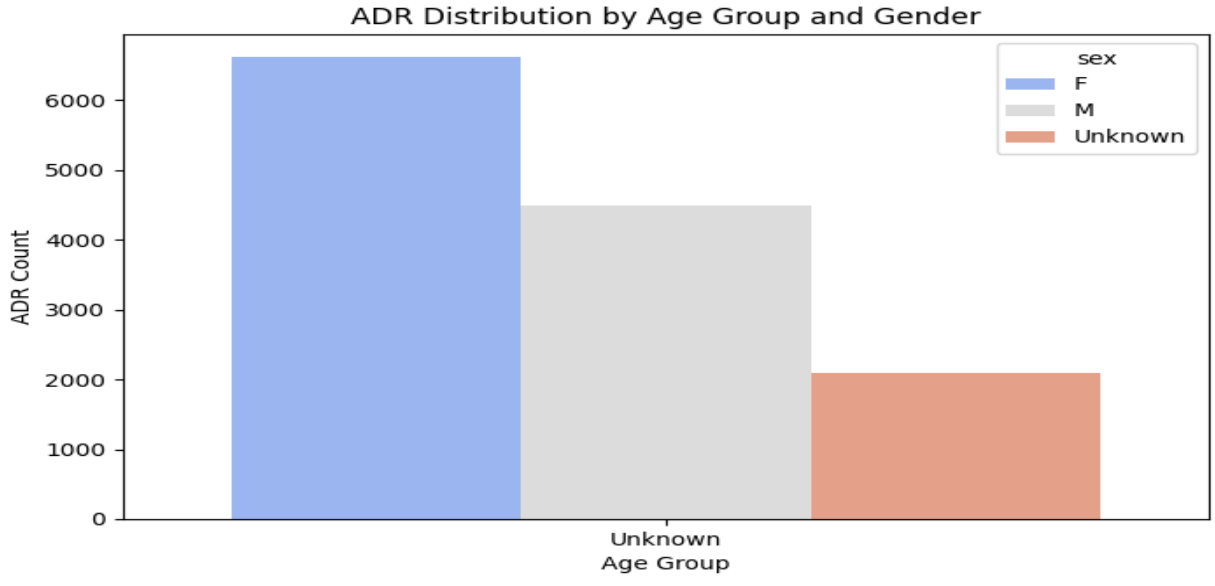
4. ADR Reports by Country

- Description: Clustered bar chart displaying the number or percentage of ADR reports per reporting country (`occr_country`).
- Insight**:** US, CA, DE these countries with the highest ADR reporting frequency. These regions are with greater pharmacovigilance activity or higher reported risk, and can be targeted safety monitoring and resource allocation.
- Visualization:



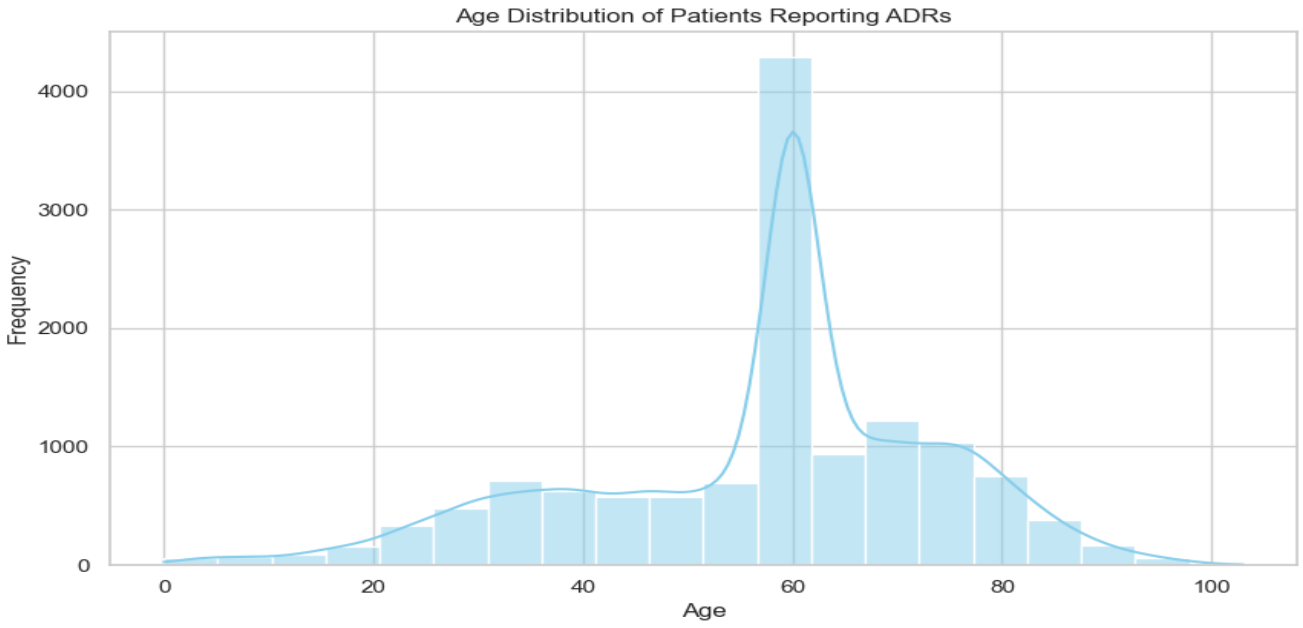
5. ADR Distribution by Age and Gender

- Description: Clustered column chart showing ADR frequency across different age groups (`age_grp`), separated by gender (`sex`).
- Insight**:** Female age groups is most affected within each gender. Helps detect age- and gender-specific vulnerability to ADRs. Provides actionable guidance for targeted monitoring and intervention strategies.
- Visualization



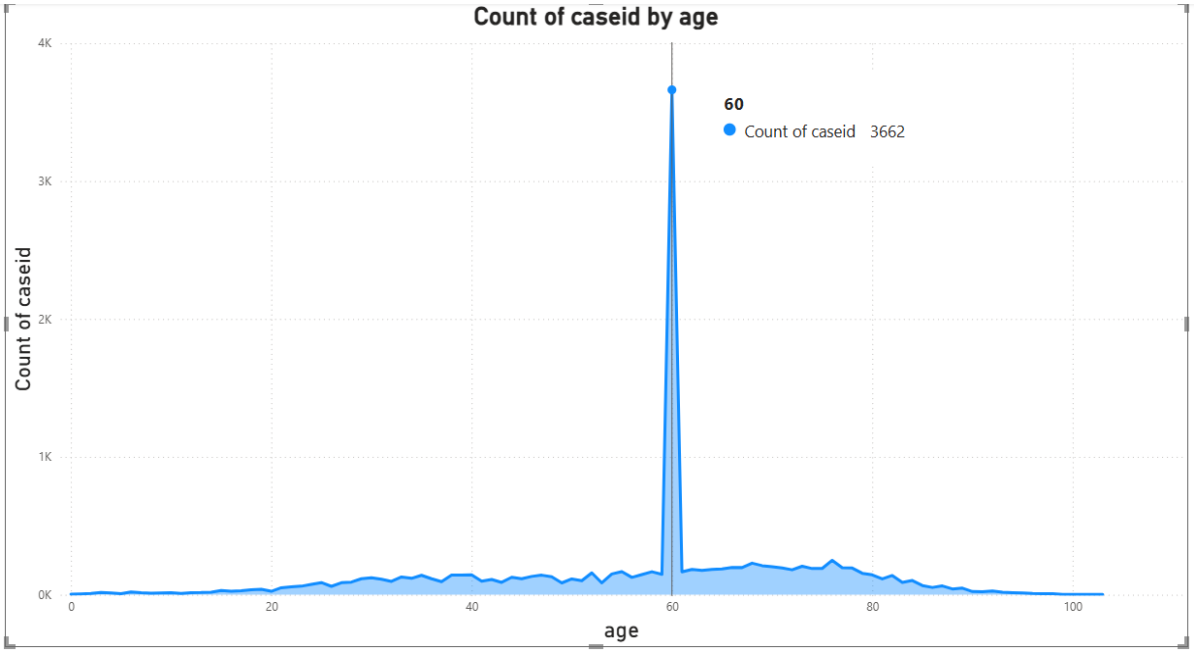
6. Age Distribution of Patients Reporting ADRs

- Description: Line and stacked column chart showing the distribution of patient ages who reported ADRs (age).
- Insight: 60 is the the most common age ranges for ADR reporting.
- Visualization:



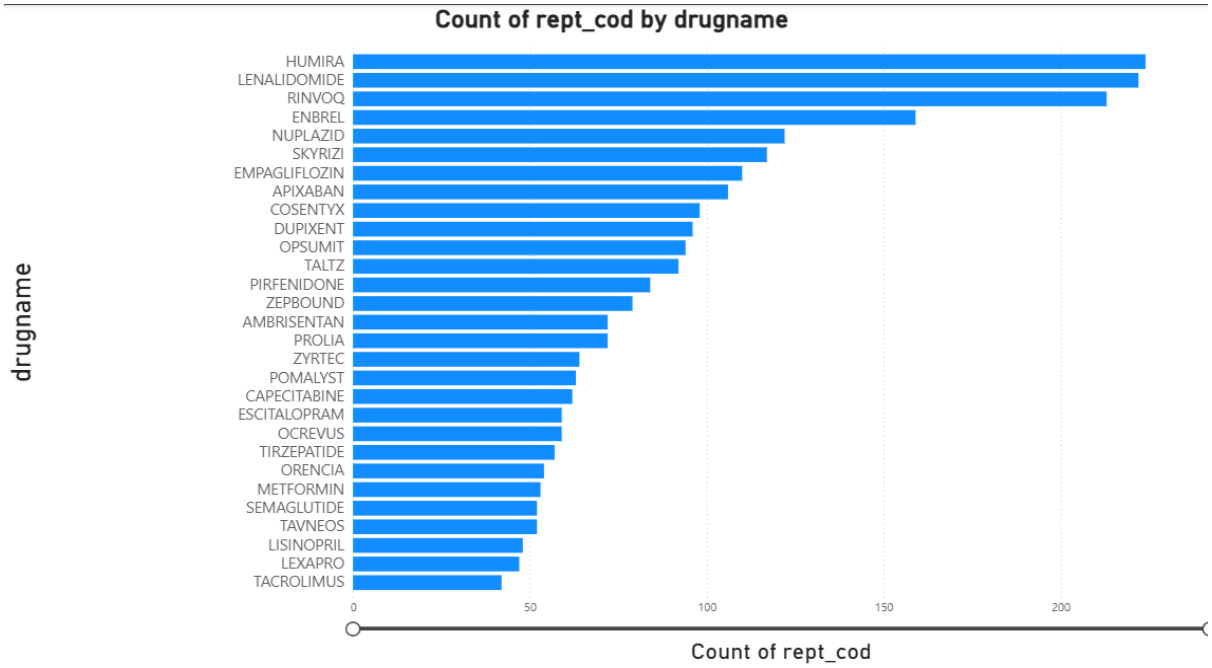
7. Count of ADR Cases by Age

- Description**:** Area chart showing the number of ADR cases (caseid) across different individual ages (age).
- Insight**:** 60 age report the most ADR cases. Age 60 population is at higher risk of adverse drug reactions.
- Visualization:



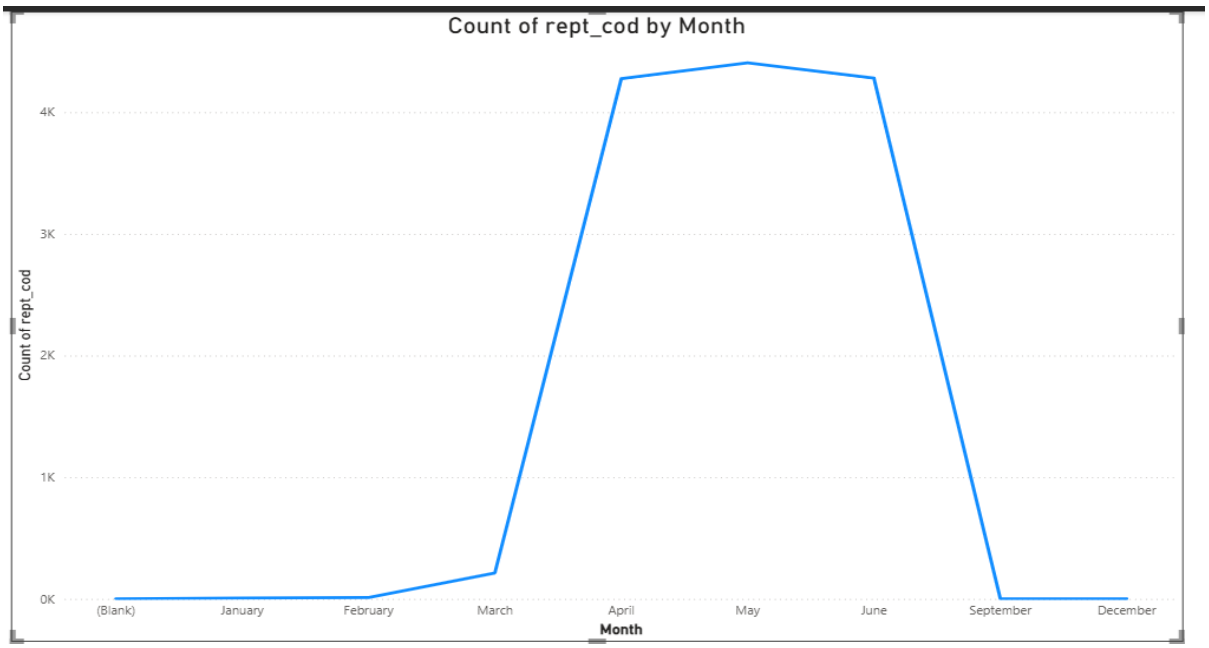
8. Count of ADR Reports per Drug

- Description: Bar chart showing the number of ADR reports (primaryid or report ID) associated with each drug (drugname).
- Insight: Humira, Lenalidoide, Rinvoq are has highest number of ADR reports.
- Visualization:

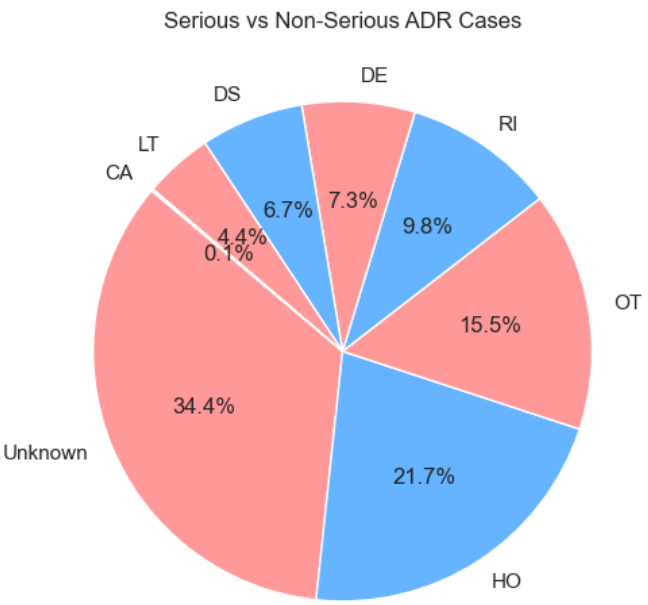


9. Count of ADR Reports per Month

- Description: Line chart showing the number of ADR reports (rept_cod) aggregated by month (rept_dt).
- Insight: Identifies trends and seasonal patterns in ADR reporting. MAY month has high ADR occurrences, which may indicate safety signals or reporting spikes.
- Visualization:



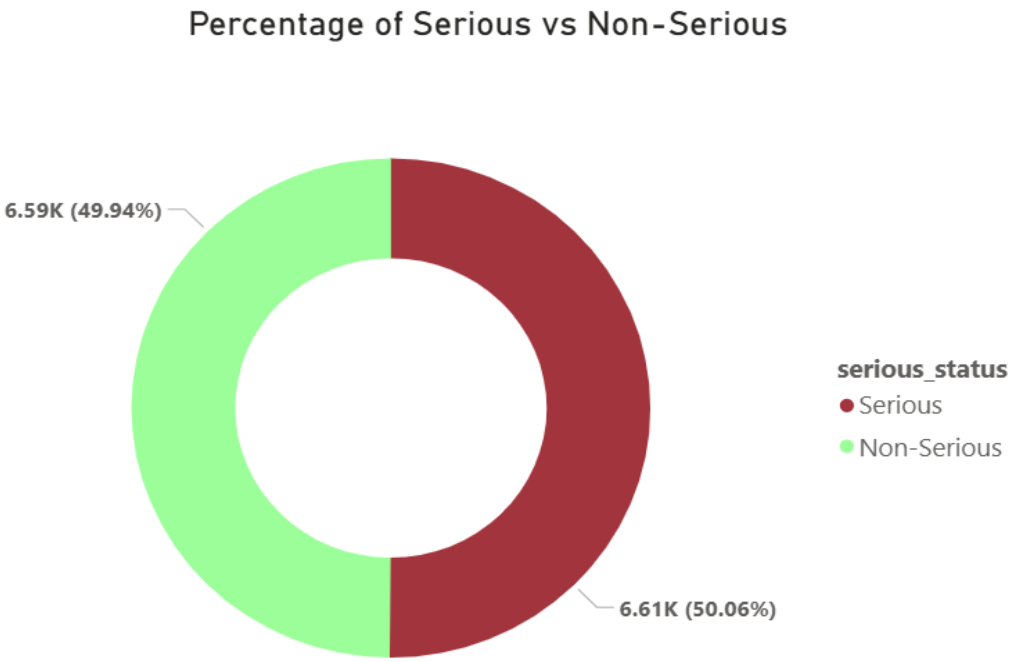
10. Serious vs Non-Serious ADR Cases



- Description: Pie chart showing the proportion of serious versus non-serious ADR reports (outc_cod).
- Insight**:** Unknown causes are severe of ADRs in the dataset. 21.7 percentage of cases leading to hospitalization, 4.4 % life-threatening events, 34.4% are unknown, 15.5% leads to OT, 9.8% showing RI, etc., Supports prioritization of pharmacovigilance efforts and resource allocation for high-risk ADRs.

11. Percentage of Serious v/s Non-Serious ADR Cases

- Description**:** Donut chart showing the percentage distribution of serious versus non-serious ADR reports (outc_cod).
- Insight: The serious cases are 6.61K with the 50.06% of the total where 6.65K cases are detected as the Non-Serious with 49.94% of total.
- Visualization:



12. Top 10 Countries with the Most ADR Reports

- Description: Table displaying the number of ADR reports (primaryid) for the top 10 reporting countries (occr_country).
- Insight: US contribute most ADR reports. Unknown country has the second most high following SK, ZA, TR, etc.,
- Visualization:

Number of ADR Reports per Country

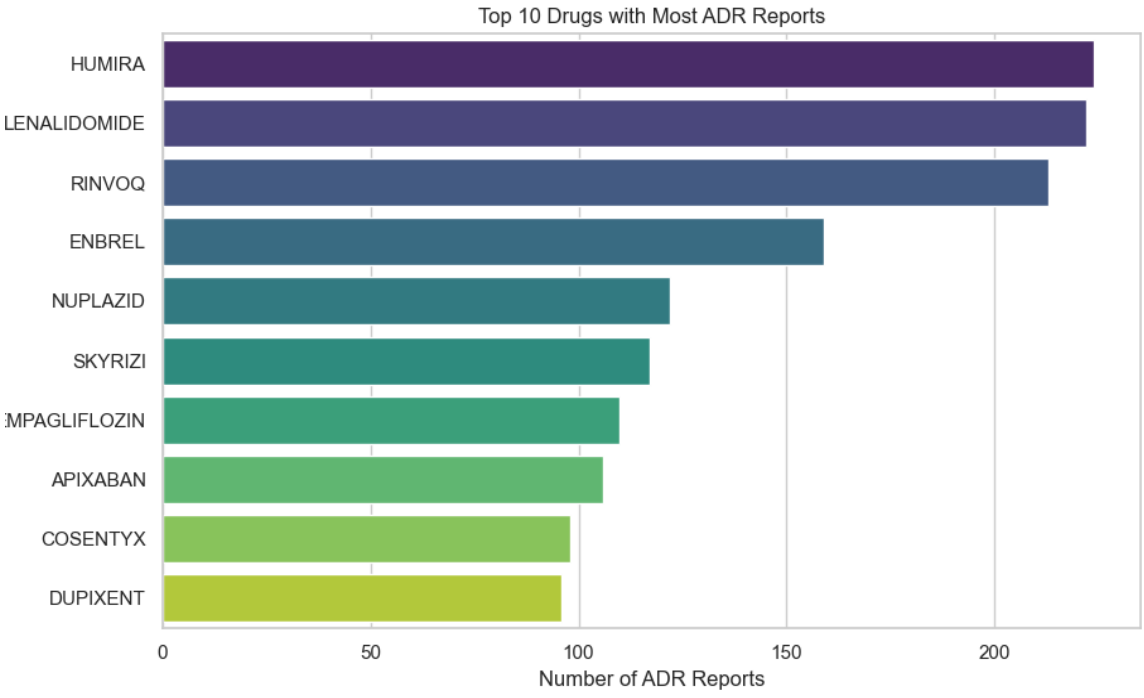
occr_country	Count of primaryid
US	4835
Unknown	53
SK	24
ZA	8
TR	7
TW	6
VE	3
TD	1
TH	1
UA	1
Total	**4927**

13. Top 10 Drugs with Most ADR Reports

Description: Bar chart showing the top 10 drugs (drugname) associated with the highest number of ADR reports (rept_id).

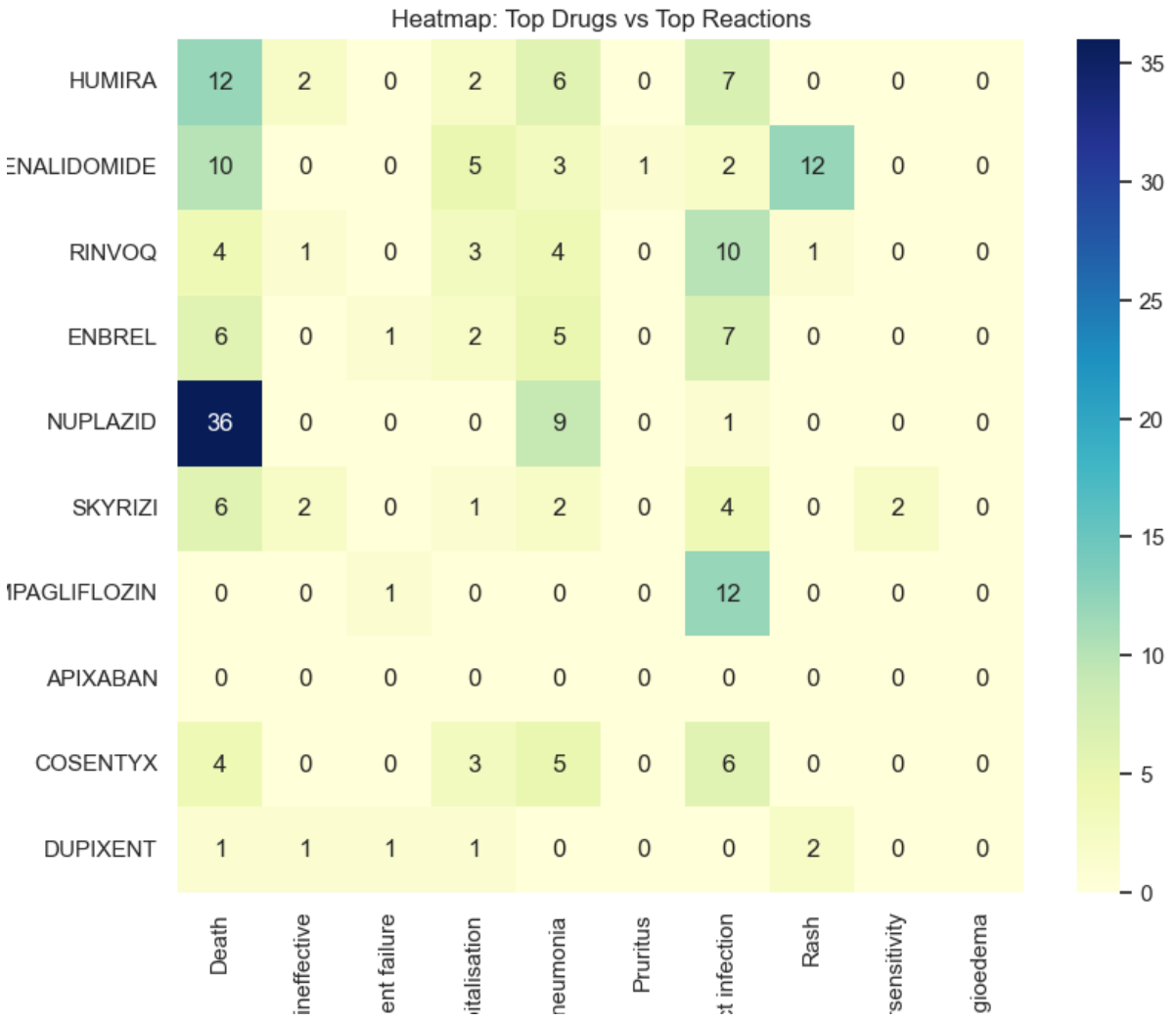
Insight: Humira drug has most frequently reported for ADRs. Humira drug prioritize monitoring and safety assessments for high-risk medications.

Visualization:



14. Top Drugs vs Top ADR Reactions

- Description: Heatmap showing the frequency of top drugs (drugname) against their most commonly reported ADR reactions (pt or preferred term).
- **Insight:** NUPLAZID showing the death as most reactions. ENALIDOMIDE follows with second most death rate. MPAGLIFLOZIN showing infection.
- **Visualization:**



5. Results:

1. **Outcome Distribution:**
    - Hospitalization accounted for **21.7%** of serious ADR reports.
    - Life-threatening outcomes combined accounted for **4.4%** of reports.
    - Approximately **9.82% of patients achieved full recovery** after intervention.
2. **Gender-Wise ADR Analysis:**
    - ADRs were distributed across male, female, and unknown gender categories.
    - Certain outcomes, including hospitalization and serious reactions, were slightly more frequent in one gender female, females showing most ADRs compare to other genders.
3. **Age Analysis:**
    - ADRs were most frequently reported in specific age groups (60 years).
    - Age-gender analysis revealed particular vulnerability in certain age groups within each gender.
4. **Geographical Distribution:**
    - The top reporting countries contributed the majority of ADR reports, indicating regions with higher pharmacovigilance activity or increased risk reporting.
5. **Drug and Reaction Analysis:**
    - The top 10 drugs were responsible for a significant proportion of ADR reports.
    - Certain drugs were associated with specific reactions more frequently, suggesting targeted areas for monitoring.
6. **Monthly Trends:**
    - ADR reports varied over time, with identifiable monthly peaks.
    - Monitoring these trends can help in early signal detection.
7. **Serious vs Non-Serious Cases:**
    - Serious cases constituted a notable proportion of reports, emphasizing the need for focused pharmacovigilance.
    - Non-serious cases, while more frequent, still provide important safety information.

6. Conclusion:

- The dataset demonstrates that hospitalization and life-threatening events, while in majority, unknown represent a significant clinical burden and require continuous monitoring.
- Gender and age analyses reveal female populations at higher risk, allowing targeted intervention and resource allocation.
- Geographical and drug-wise distribution highlights areas and medications that require heightened vigilance.
- Monitoring temporal trends enables early detection of ADR spikes and improves proactive safety measures.
- Overall, these insights support enhanced pharmacovigilance, patient safety initiatives, and informed regulatory decision-making**,** ensuring high-risk populations and medications are closely monitored.

7. Reference

U.S. Food and Drug Administration. (2025). *FDA Adverse Event Reporting System (FAERS)*. Retrieved October 8, 2025, from FAERS Q2 2025 Report.docx