Understanding the Dataset

The dataset provided consists of **three interconnected tables** — *Events, Devices, and Manufacturers*. These tables together describe regulatory actions, device details, and manufacturer information. Below is a focused explanation of the most important fields chosen for analysis.

1. Events Dataset

The **Events table** contains records of regulatory actions (such as recalls, safety alerts, and field safety notices). It is the **primary source of failure information**.

Key Fields

- **id*** → Unique identifier for each event.
- **action** → Type of regulatory action (recall, safety alert, FSN, correction, removal, etc.).
- **action_classification*** → Severity of action (Class I: most serious, Class II: moderate, Class III: least serious).
- **action_level** → Level at which the action applies (device, batch, model, or company-wide).
- **action_summary*** → Short description of the issue or defect.
- **country*** → Country where the event was reported or initiated.
- **date_initiated_by_firm** → Date the manufacturer/distributor started the action.
- **date_posted** → Date when the authority made the event public.
- **date_terminated** → End date when the recall or action was closed.
- **determined_cause*** → Cause of failure (e.g., design flaw, manufacturing defect, labeling issue).
- **icij_notes** → Notes and additional information curated by ICIJ.
- **reason*** → Regulatory reason behind the action.
- **status** → Current state of the action (ongoing, completed, terminated).

- **type*** → Type/category of event (recall, correction, safety notice, etc.).
- **device_id*** → Foreign key linking to the Devices table.

2. Devices Dataset

The **Devices table** contains metadata describing medical devices that appear in the events.

Key Fields

- **id*** → Unique identifier for each device (linked to Events).
- **classification*** → Regulatory classification (general category of the device).
- **code*** → Internal model or device code.
- **description** → Free-text description of the device.
- **implanted*** → Boolean indicating if the device is implantable (yes/no).
- **name** → Device's commercial name or label.
- **risk_class*** → Regulatory risk class of the device (I, IIa, IIb, III).
- **country*** → Country where the device was distributed/marketed.
- **manufacturer_id** → Foreign key linking to the Manufacturers table.

3. Manufacturers Dataset

The **Manufacturers table** provides company-level details for device producers, suppliers, or distributors.

Key Fields

- **id*** → Unique identifier for each manufacturer (linked to Devices).
- **address** → Location/address of the manufacturer.
- **name*** → Name of the company or manufacturer.
- **parent_company*** → Parent corporate group (if applicable).

4. Relationships Between Tables

- **Events ↔ Devices:**
  Connected through `device_id`. Each event is linked to the specific device involved.

- **Devices ↔ Manufacturers:**
  Connected through `manufacturer_id`. Each device is linked to its producer.

- **Flow of information:**
  `Event (failure/recall) → Device (characteristics) → Manufacturer (responsible company)`

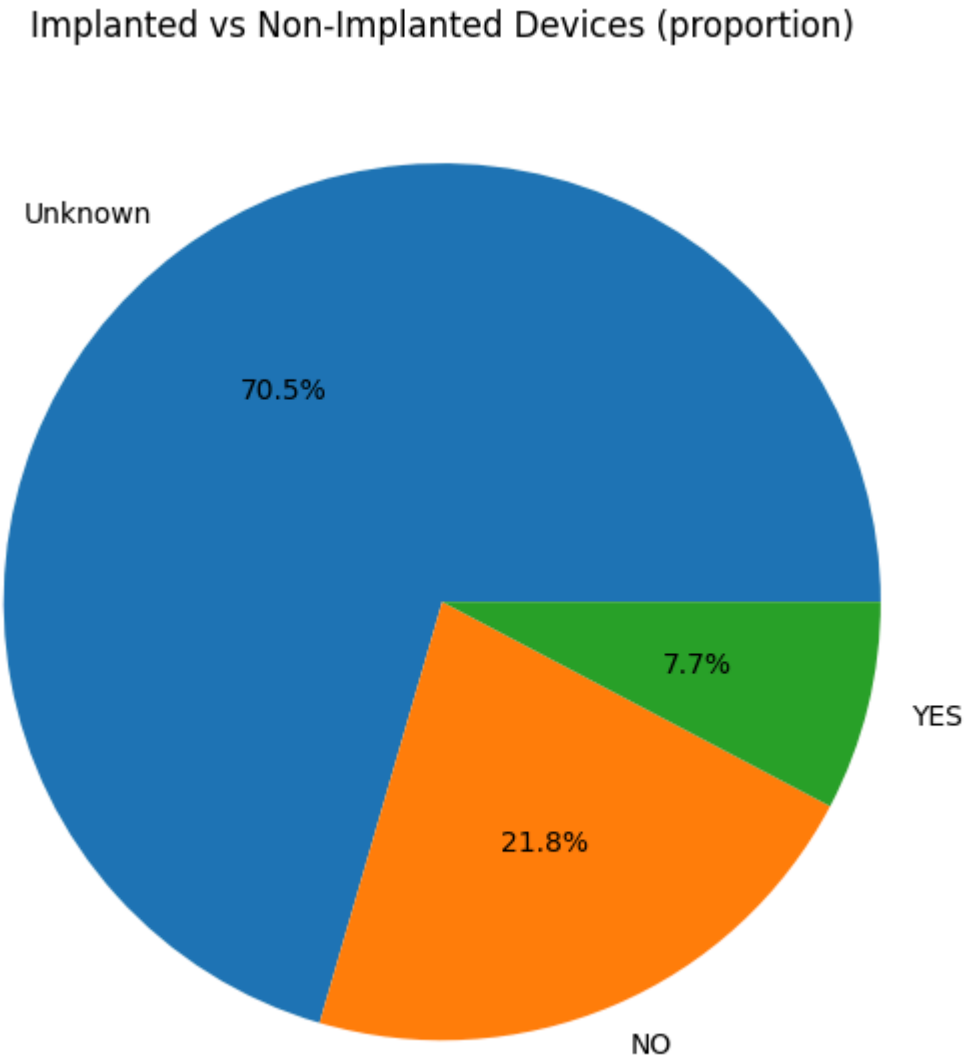
5. Why These Features Matter

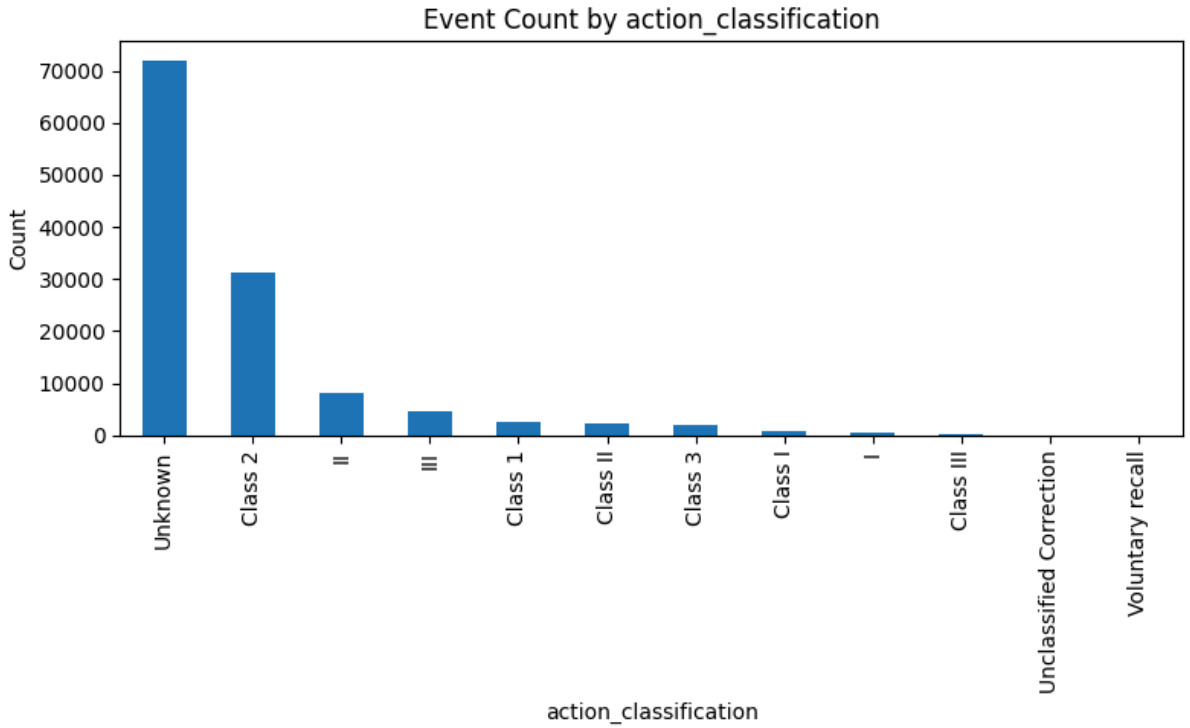
- **Events:** Tell *what went wrong*, how severe it was, and why.

- **Devices:** Tell *what kind of product failed* and its inherent risk factors.

- **Manufacturers:** Tell *who was responsible* and whether history shows reliability or negligence.

VISUALIZATION




Event Count by action_classification
Count
70000
60000
50000
40000
30000
20000
10000
0
Unknown
Class 2
II
III
Class 1
Class II
Class 3
Class I
I
Class III
Unclassified Correction
Voluntary recall
action_classification