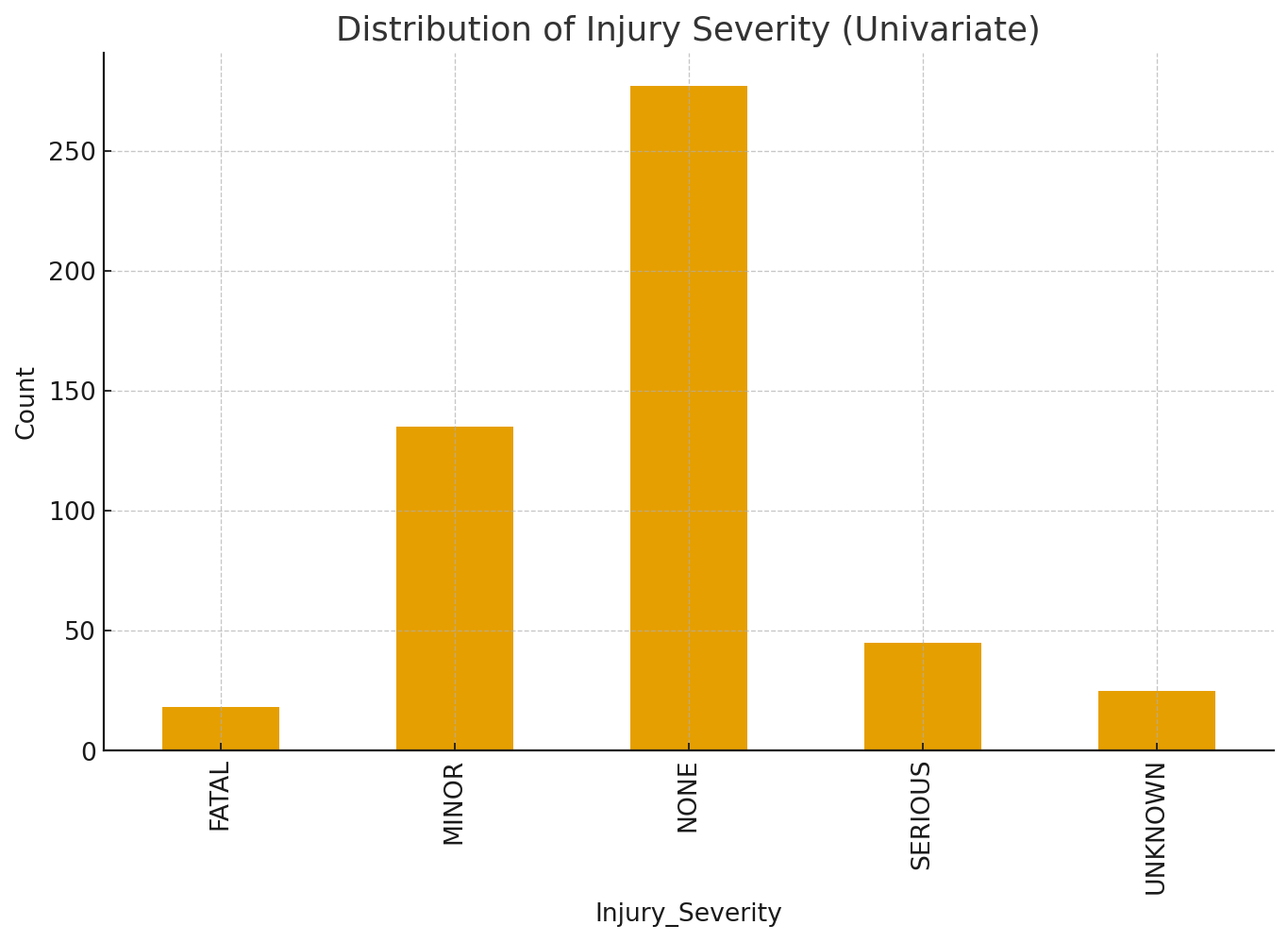
# TeamX\_LastNames\_DataEXP\_10Graphs

Dataset: FAA ASIAS Preliminary Accident & Incident Reports (10-business-day window).

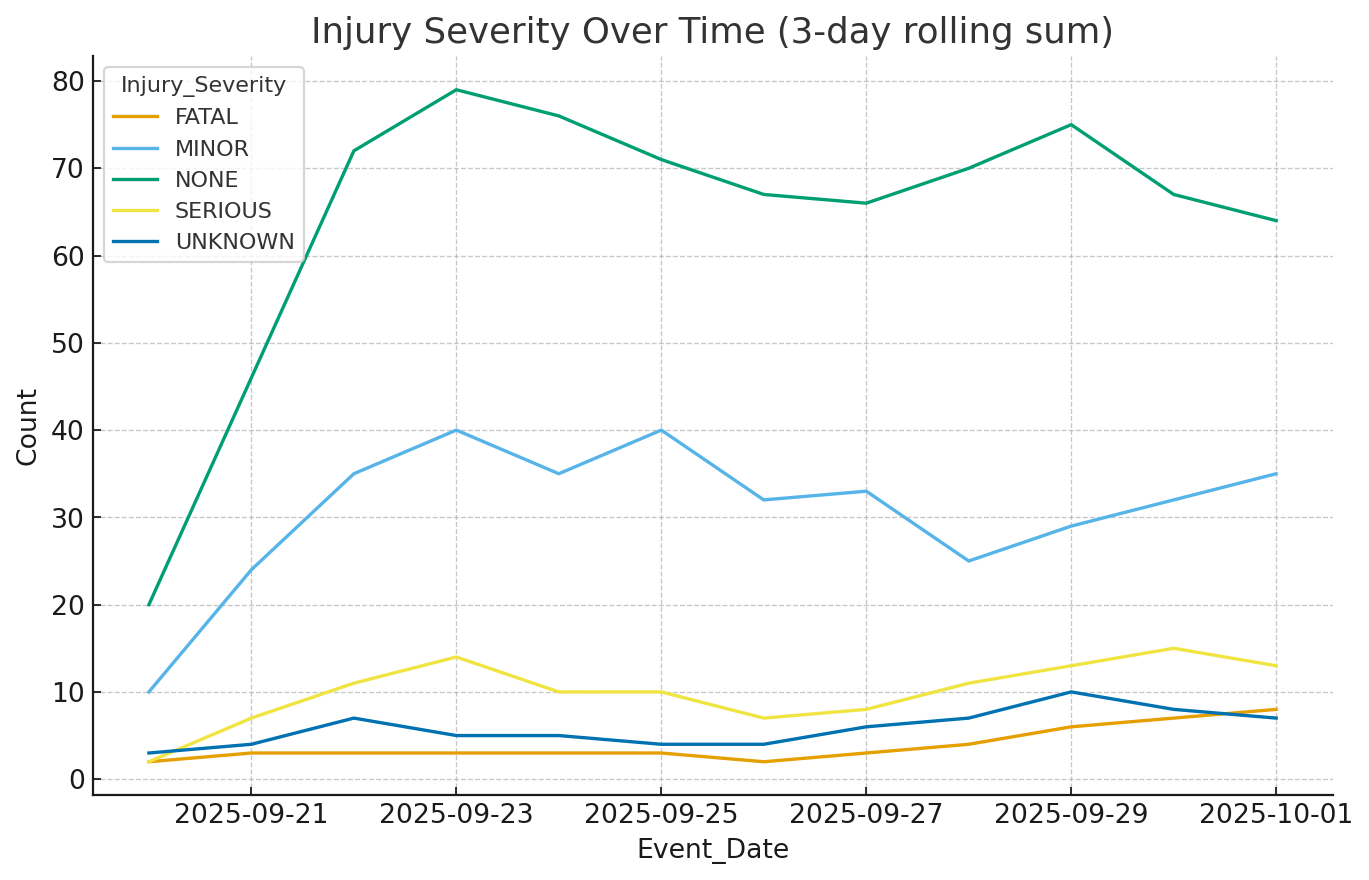
Response variable: Injury\_Severity. The following graphics explain the distribution and drivers of injury outcomes, plus data-quality diagnostics.

## Univariate 1 — Injury Severity Distribution



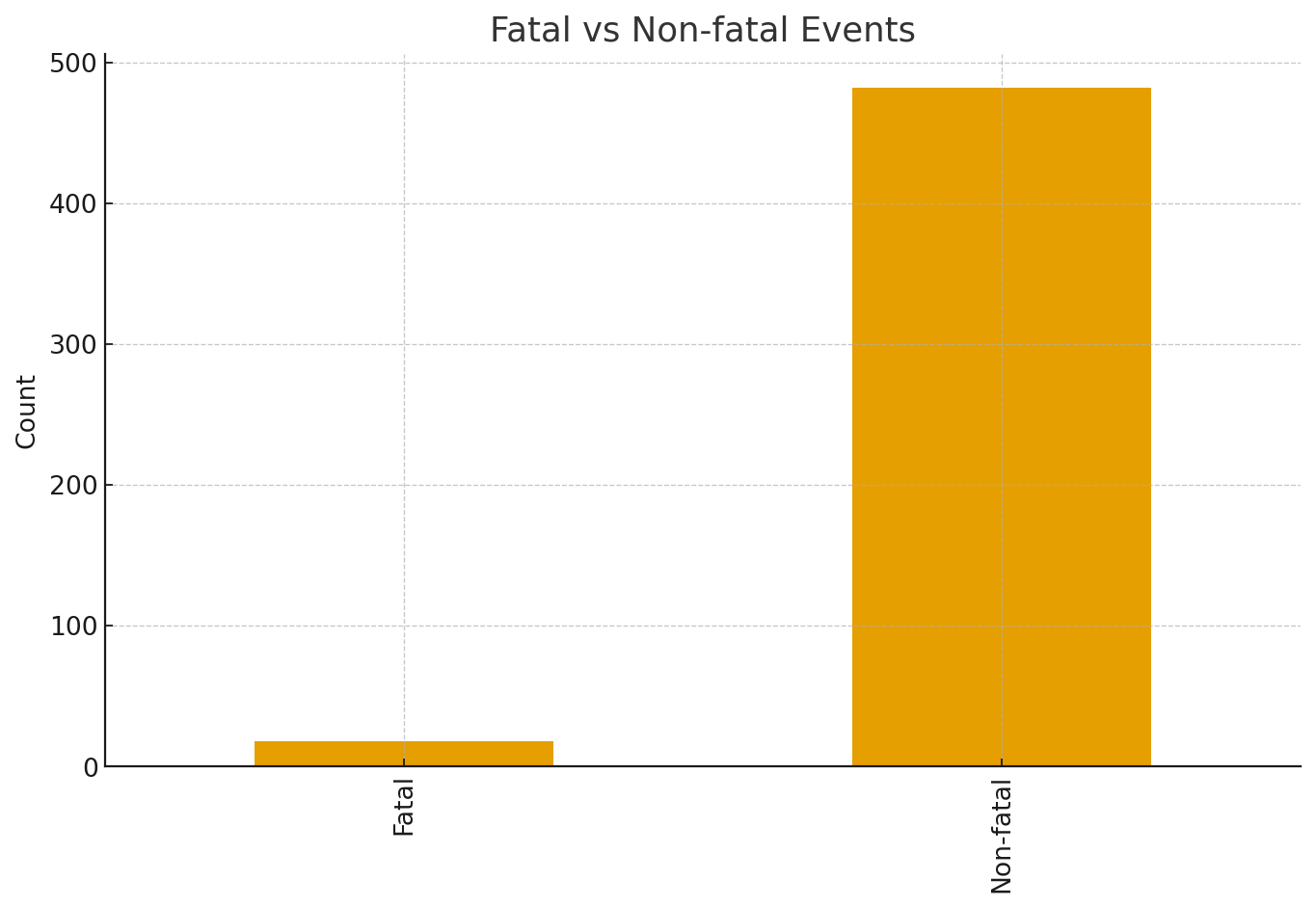
What: Count of events by injury severity. Why: Establishes baseline class balance; heavy skew toward NONE/MINOR suggests rare severe events.

## Univariate 2 — Injury Severity Over Time



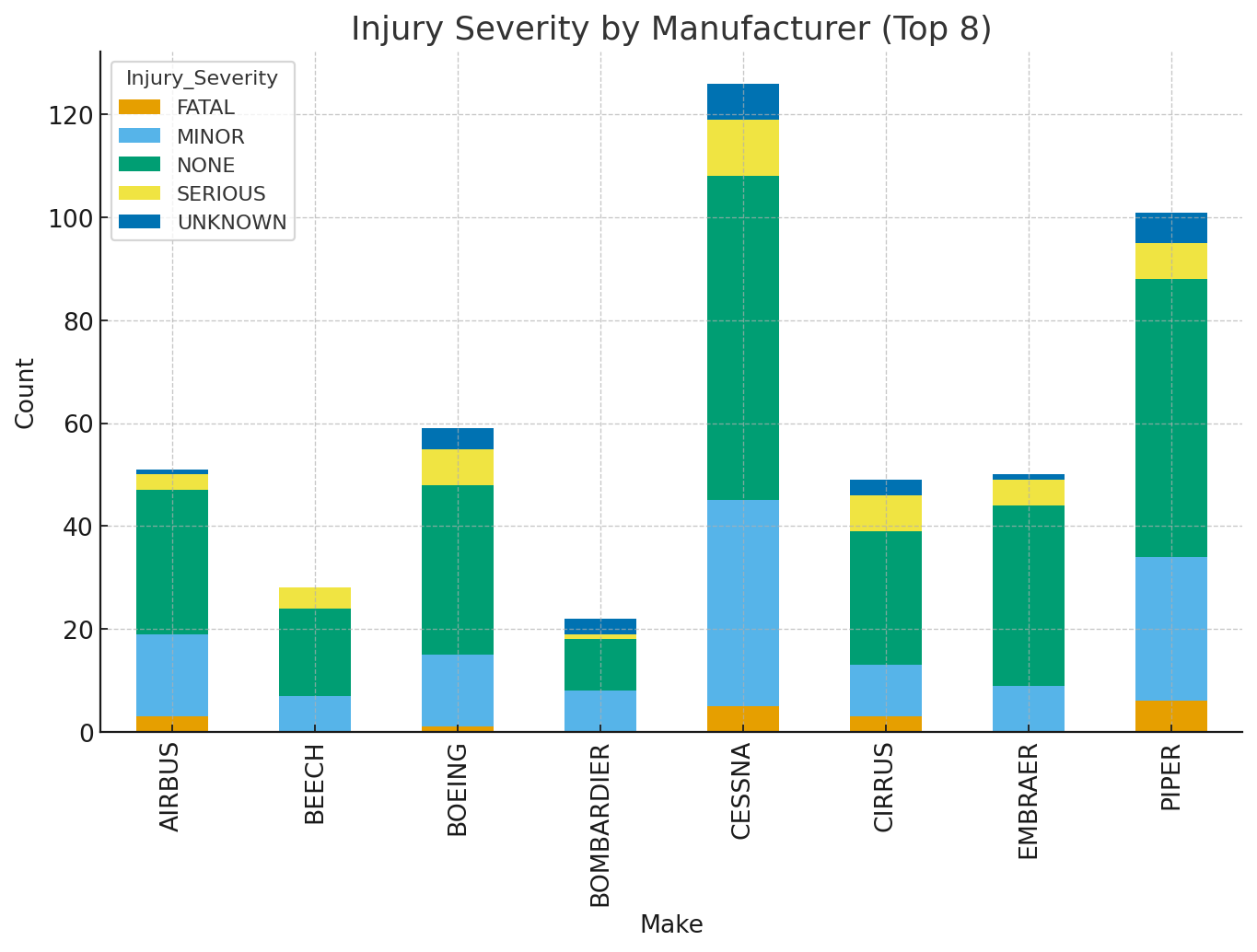
What: 3-day rolling sums by severity. Why: Reveals temporal spikes; helps plan rolling-window monitoring and alerts.

## Univariate 3 — Fatal vs. Non-fatal



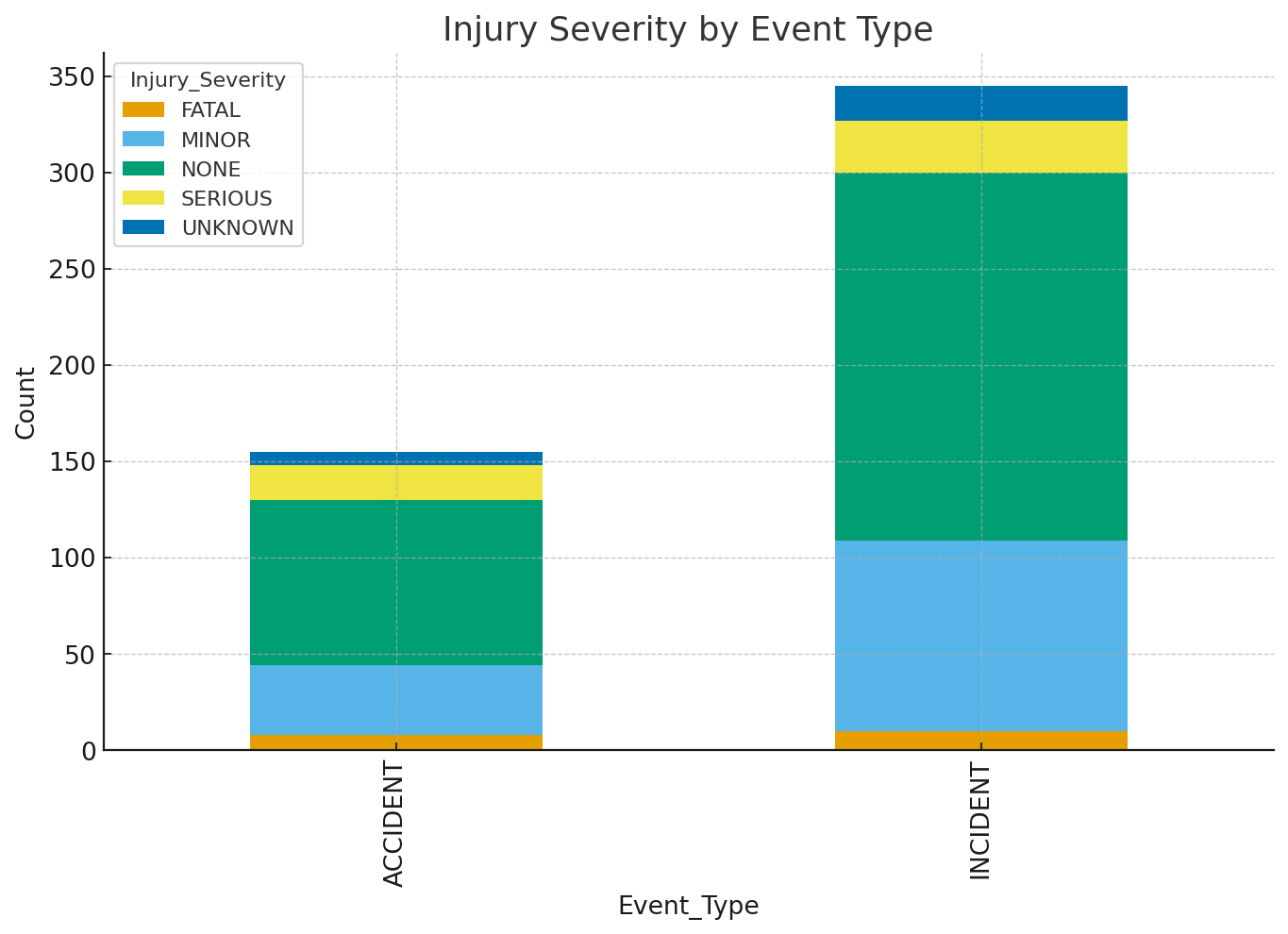
What: Binary view of fatal vs non-fatal outcomes. Why: Frames the primary safety outcome and its rarity.

## Bivariate 1 — Severity by Manufacturer (Top 8)



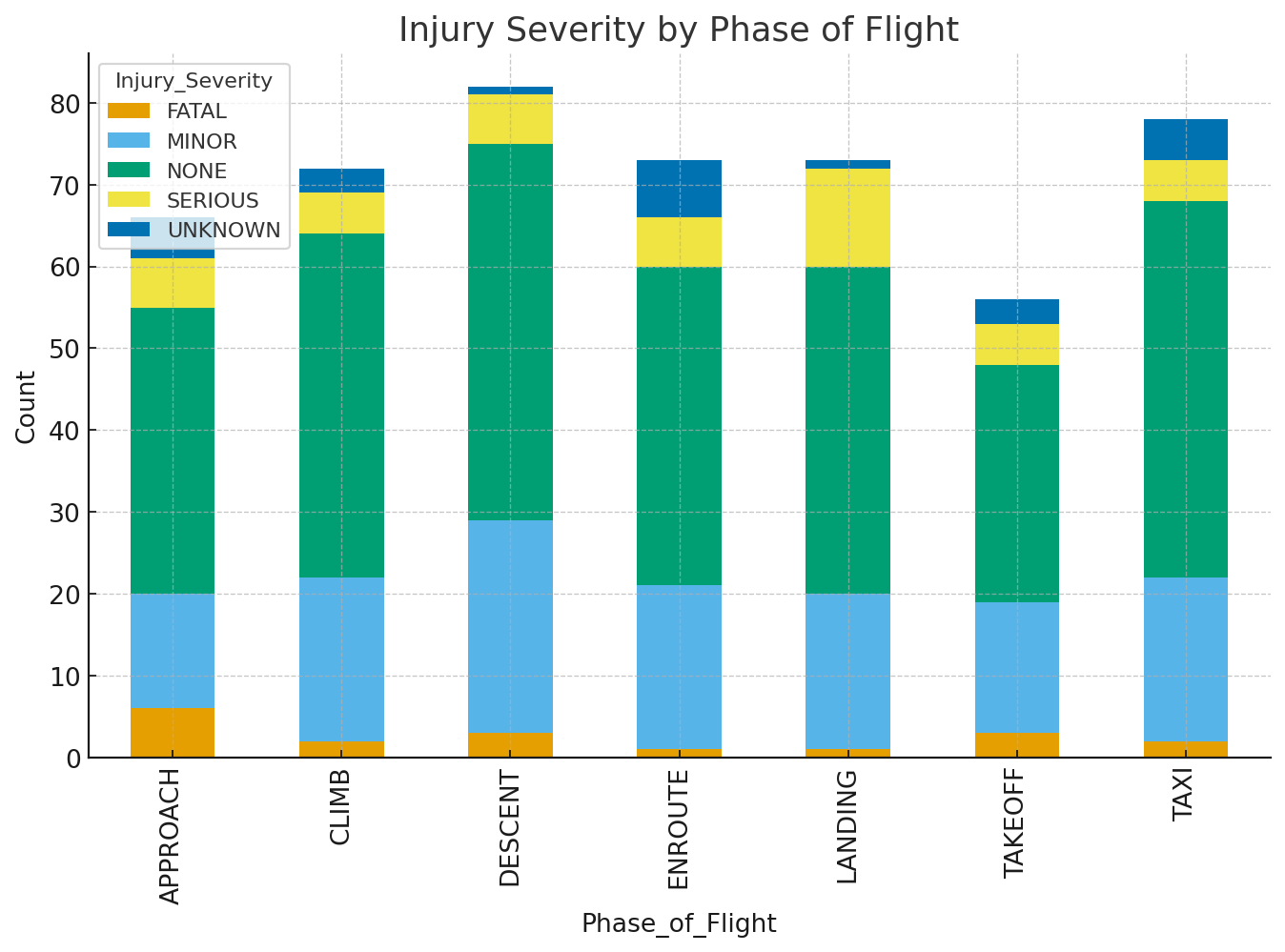
What: Stacked counts by make and severity. Why: Identifies manufacturers associated with relatively higher severe outcomes (case-mix awareness).

## Bivariate 2 — Severity by Event Type



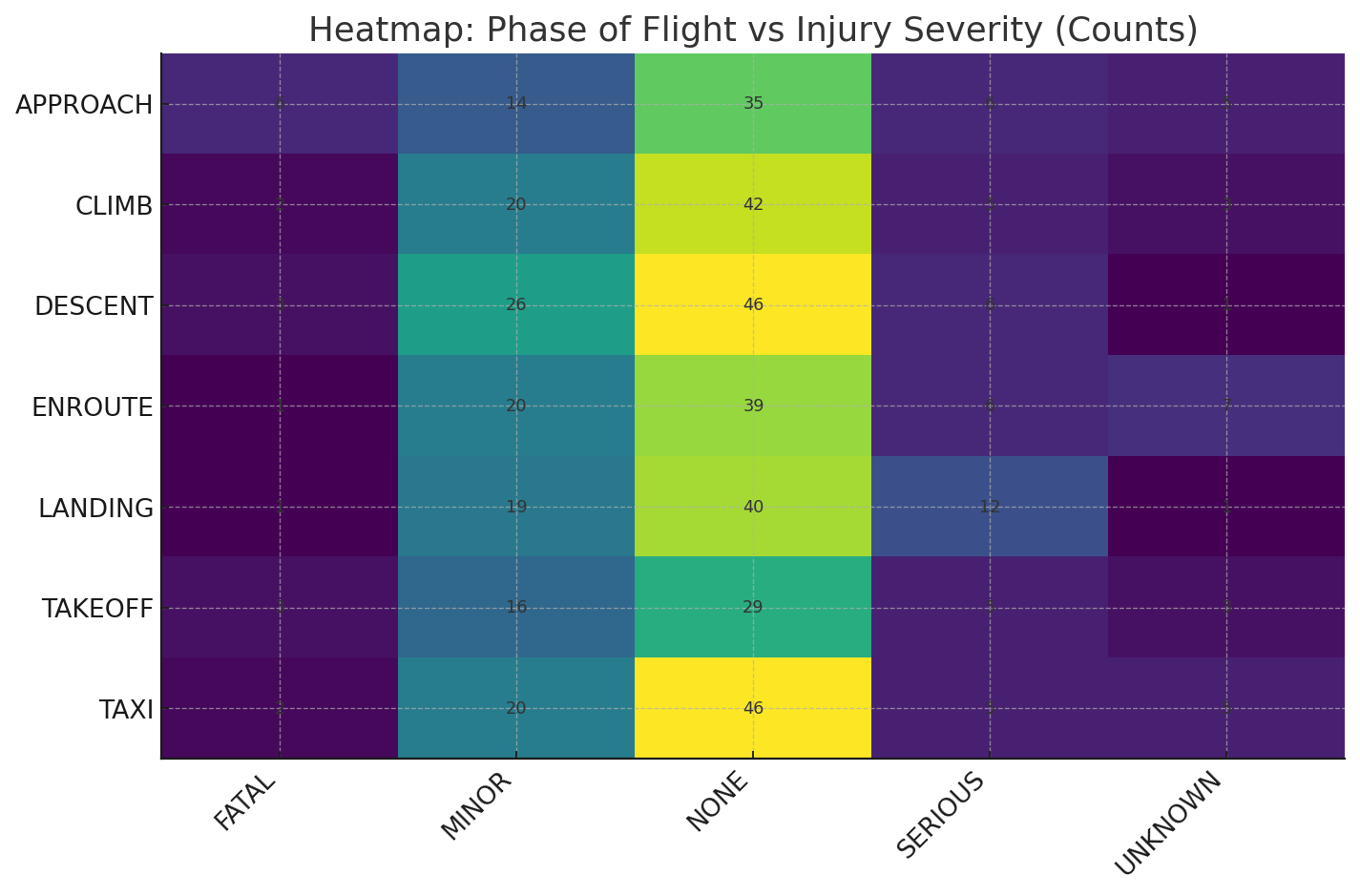
What: Severity distribution split by INCIDENT vs ACCIDENT. Why: Confirms expectation that ACCIDENT skews more severe.

## Bivariate 3 — Severity by Phase of Flight



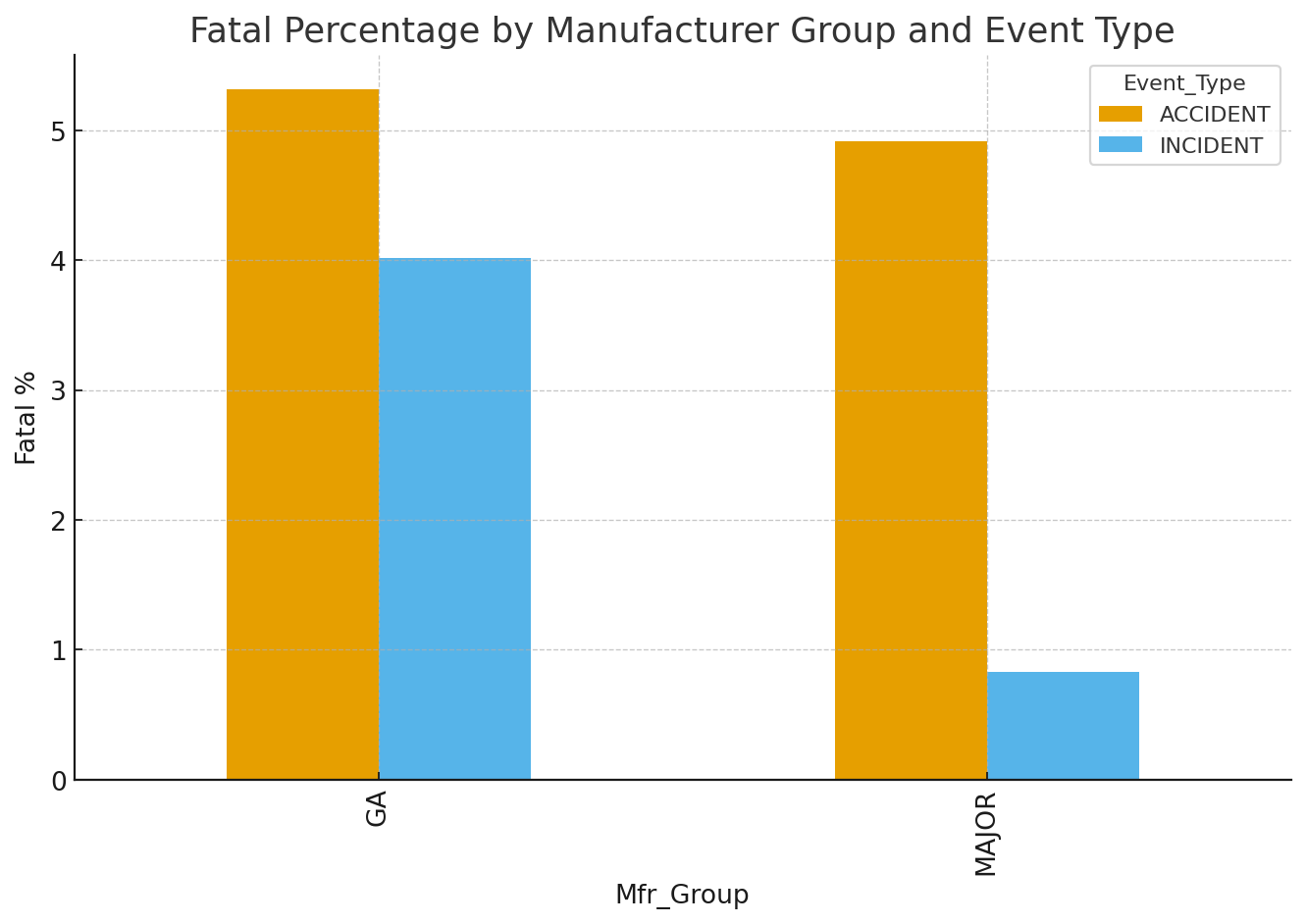
What: Severity across flight phases. Why: Highlights operational phases with disproportionate severe outcomes (e.g., approach/landing).

## High-Dimensional 1 — Heatmap: Phase × Severity



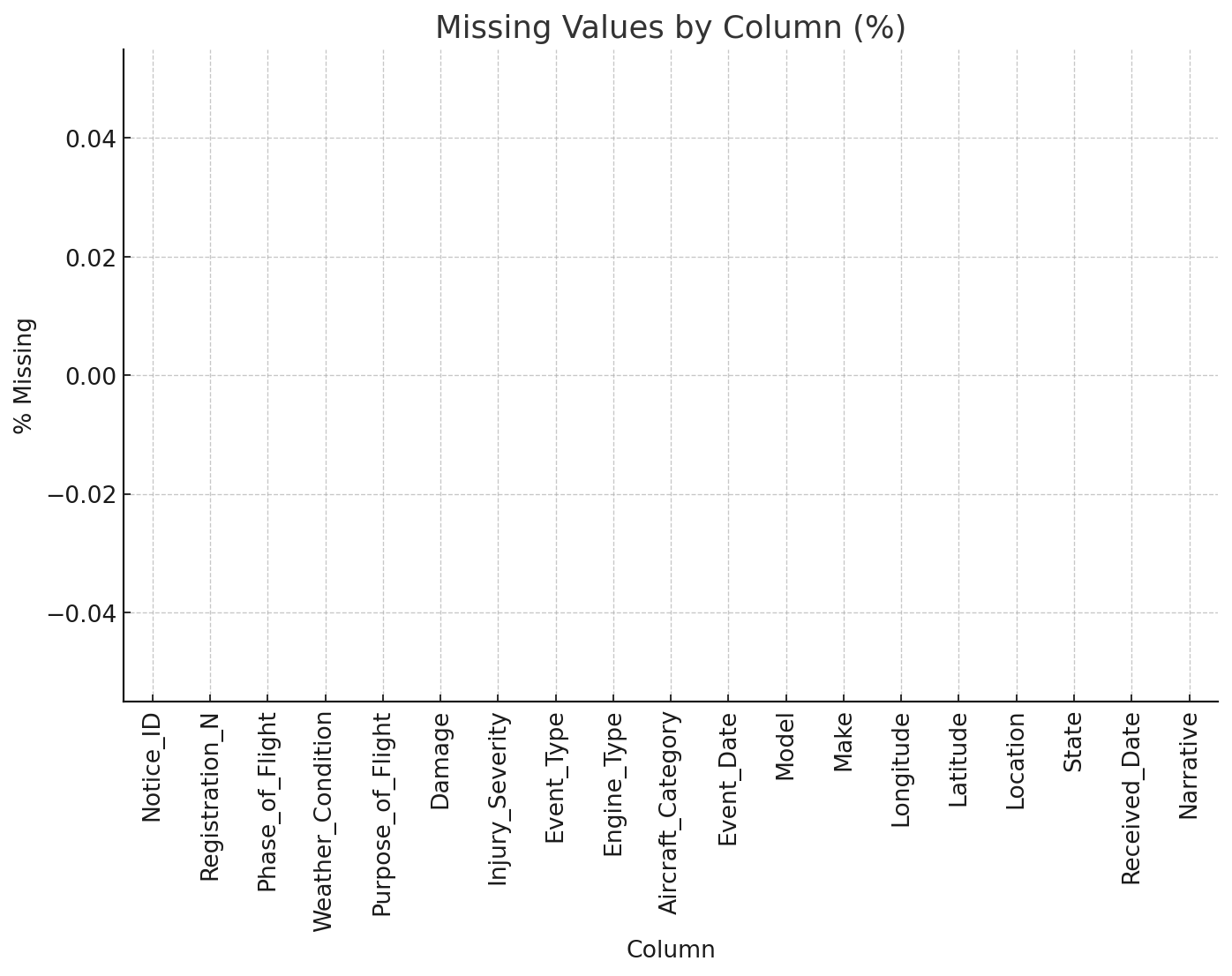
What: Matrix of counts for Phase\_of\_Flight by Injury\_Severity. Why: Combines two predictors to surface high-risk intersections.

## High-Dimensional 2 — Fatal % by Mfr Group × Event Type



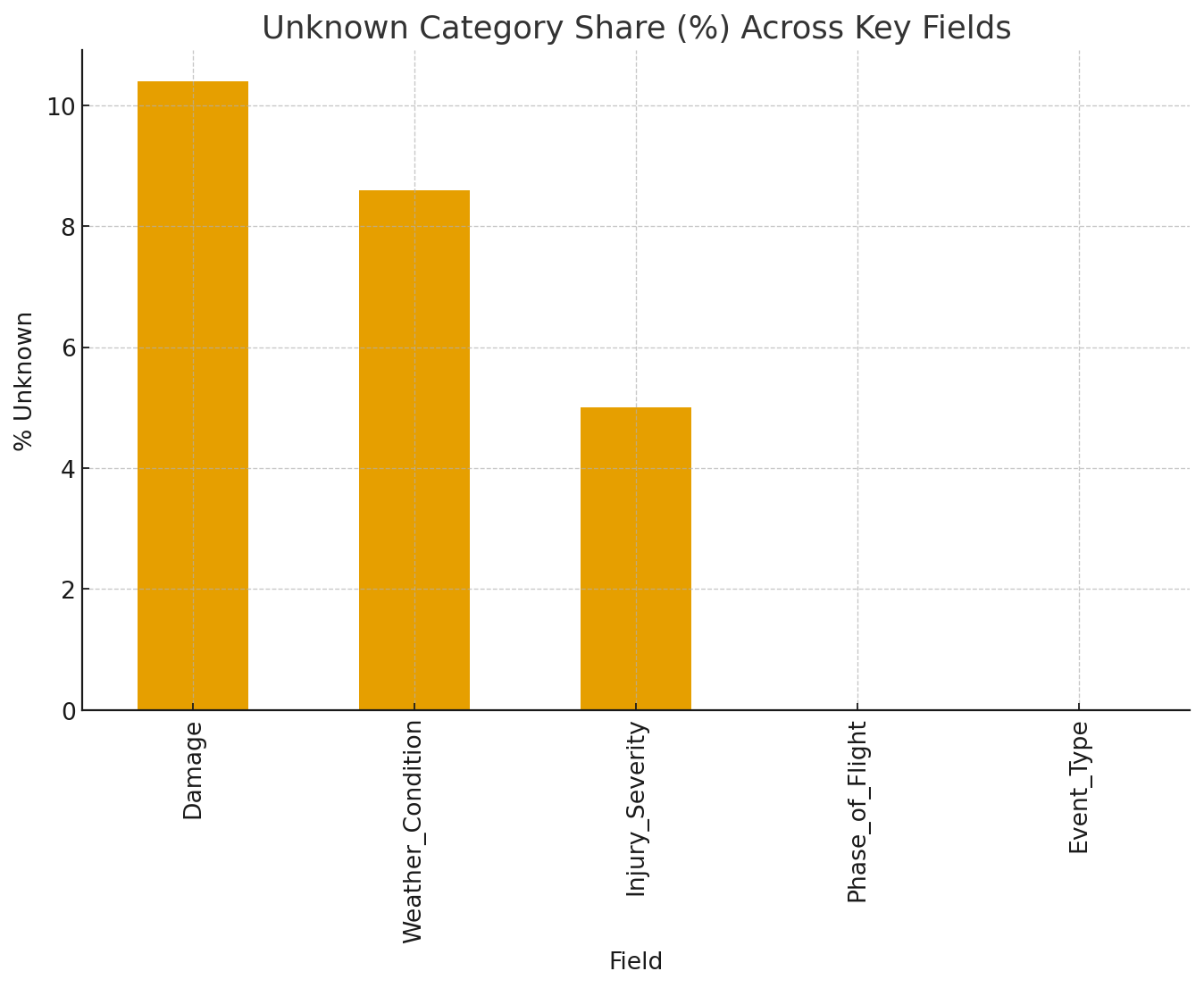
What: Fatal rate by manufacturer group and event type. Why: Multi-factor comparison to prioritize safety review segments.

## Data Issue 1 — Missingness by Column



What: % missing across fields. Why: Guides imputation/exclusion decisions; high missingness may bias results.

## Data Issue 2 — 'Unknown' Category Prevalence



What: Share of 'UNKNOWN' across key categoricals. Why: Flags low-quality categorical coding that weakens inference.

## Descriptive Statistics — Before Join

rows unique\_notice\_ids pct\_missing\_Injury\_Severity  
 500 500 0.0

## Descriptive Statistics — After Join

rows unique\_notice\_ids pct\_missing\_Injury\_Severity pct\_missing\_Mfr\_Group  
 500 500 0.0 0.0