



# Data Understanding and Analysis

## Data Preparation- Exploratory data analysis

- Deleted one column
- Deleted duplicate data
- new index set
- renamed 4 columns
- changed crush\_date and fatalities column datatypes



## Data analysis

The data analysis used comprised the following:

- Column Chart of Top 5 Most Common Aircraft Types
- Column Chart of Fatalities by Operator
- Bar Chart of Damage Classification Distribution

## Data visualization

The following visualizations were done:

- Barchart of top 5 Most Common Aircraft Types
- Bar chart of fatalities by Operator
- line chart of accident frequency overtime
- damage classification distribution
- Fatal vs. non-fatal accidents
- Top accident locations

## Conclusion

- The top 5 most common aircraft types dominate the dataset, suggesting that incident reporting is concentrated among widely used models.
- A small number of operators account for disproportionately high fatalities
- The majority of incidents fall under Write-Off and Substantial damage categories, while Minor and No Damage are relatively rare. The presence of missing/unknown entries highlights gaps in reporting consistency and data quality.



## Recommendations

- There is need for deeper analysis on the top 5 aircraft types to identify whether their high incident counts are purely due to usage volume or linked to specific safety vulnerabilities.
- Strengthen reporting standards to minimize "Unknown" or "Missing" classifications.



## Releases

No releases published

[Create a new release](#)

## Packages

No packages published

[Publish your first package](#)