IDENTIFYING LOW-RISK AIRCRAFT USING AVIATION ACCIDENT DATA

PHASE 1 DATA SCIENCE PROJECT

BY: FATUMA TARI

DATE: JULY 2025

WHY DOES THIS ANALYSIS MATTERS?

The company is expanding into the aviation industry and needs to decide which aircraft types are safest to purchase. The team lacks internal data expertise and is relying on this analysis to:

- Understand patterns in aviation accidents
- Identify aircraft with the **lowest safety risk**
- Make data-driven purchasing decisions

ANALYSIS APROACH

1. Dataset: National Transportation Safety Board (NTSB) accident reports from 1962–2023
2. Cleaned and prepared the dataset using Python
3. Conducted exploratory data analysis
4. Built an interactive Tableau dashboard
5. Delivered three actionable business recommendations

DATA OVERVIEW

What the Data Looks Like

1. Over 90,000 aviation incidents from 1962 to 2023
2. **Key fields:** event id, investigation type, accident number, event date, location, country, latitude, longitude, airport code, airport name, injury severity, aircraft damage, aircraft category, registration number, make, model, amateur built, number of engines, engine type, far description, schedule, purpose of flight, air carrier, total fatal injuries, total serious injuries, total minor injuries, total uninjured, weather condition, broad phase of flight, report status, publication date.
3. Data had many missing values and inconsistencies

DATA CLEANING STRATEGY

1. Dropped irrelevant columns (e.g., airport code, registration number)
2. Filled missing values with 'Unknown' or mode where relevant
3. Standardized inconsistent categories (e.g., “Fatal(1)” → “Fatal”)
4. Converted date formats and extracted year for analysis

BEFORE DATA CLEANING

injury_severity	count
Non-Fatal	67357
Fatal(1)	6167
Fatal	5262
Fatal(2)	3711
Incident	2219
...	...
Fatal(107)	1
Fatal(57)	1
Fatal(89)	1
Fatal(199)	1
Fatal(114)	1

110 rows × 1 columns

dtype: int64

AFTER DATA CLEANING

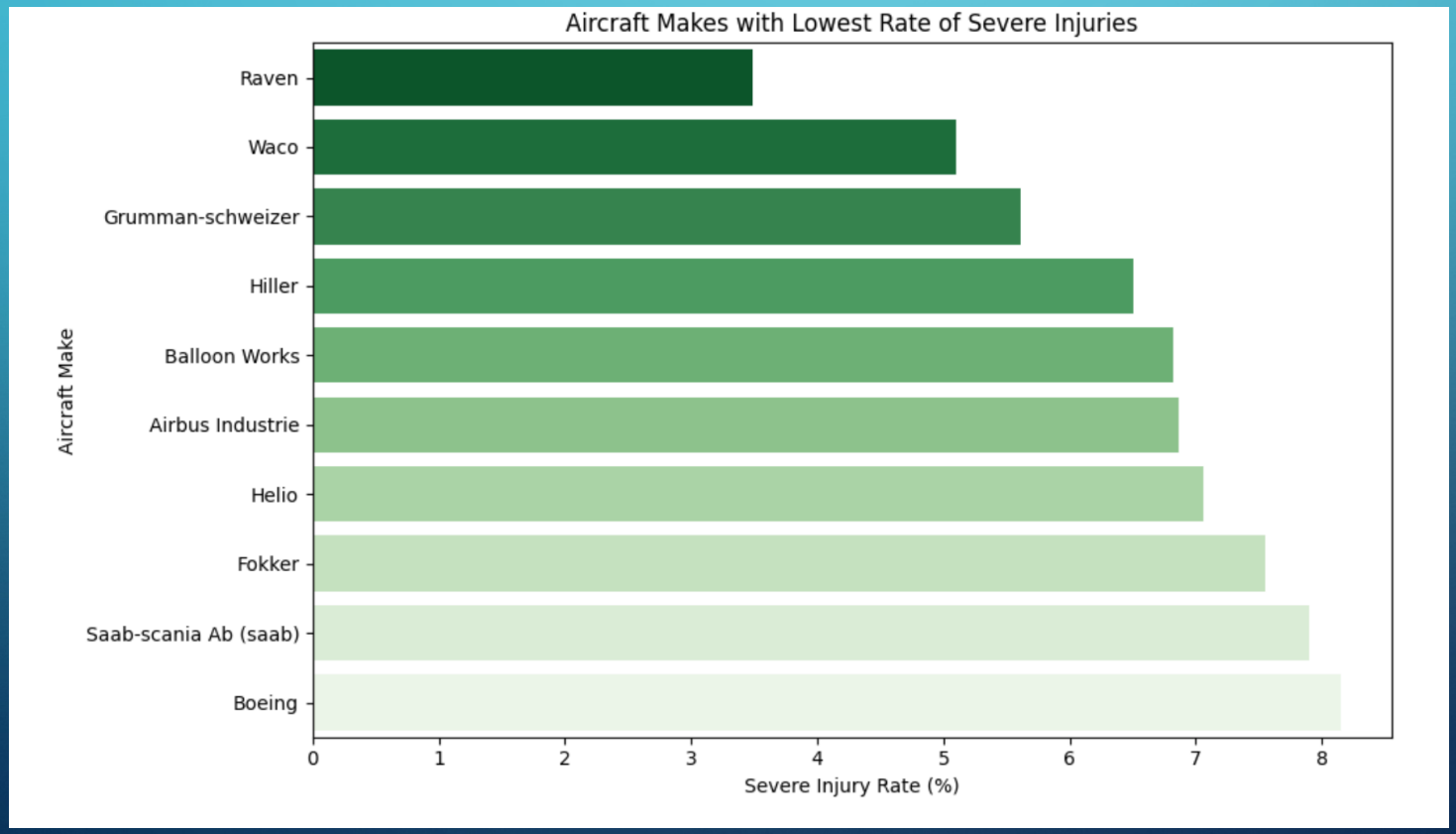
injury_severity	count
Non-Fatal	67357
Fatal	17826
Incident	2219
Unknown	1000
Minor	218
Serious	173
Unavailable	96

dtype: int64

AIRCRAFT MAKE VS. INJURY SEVERITY

Which Aircraft Makes Are Safest?

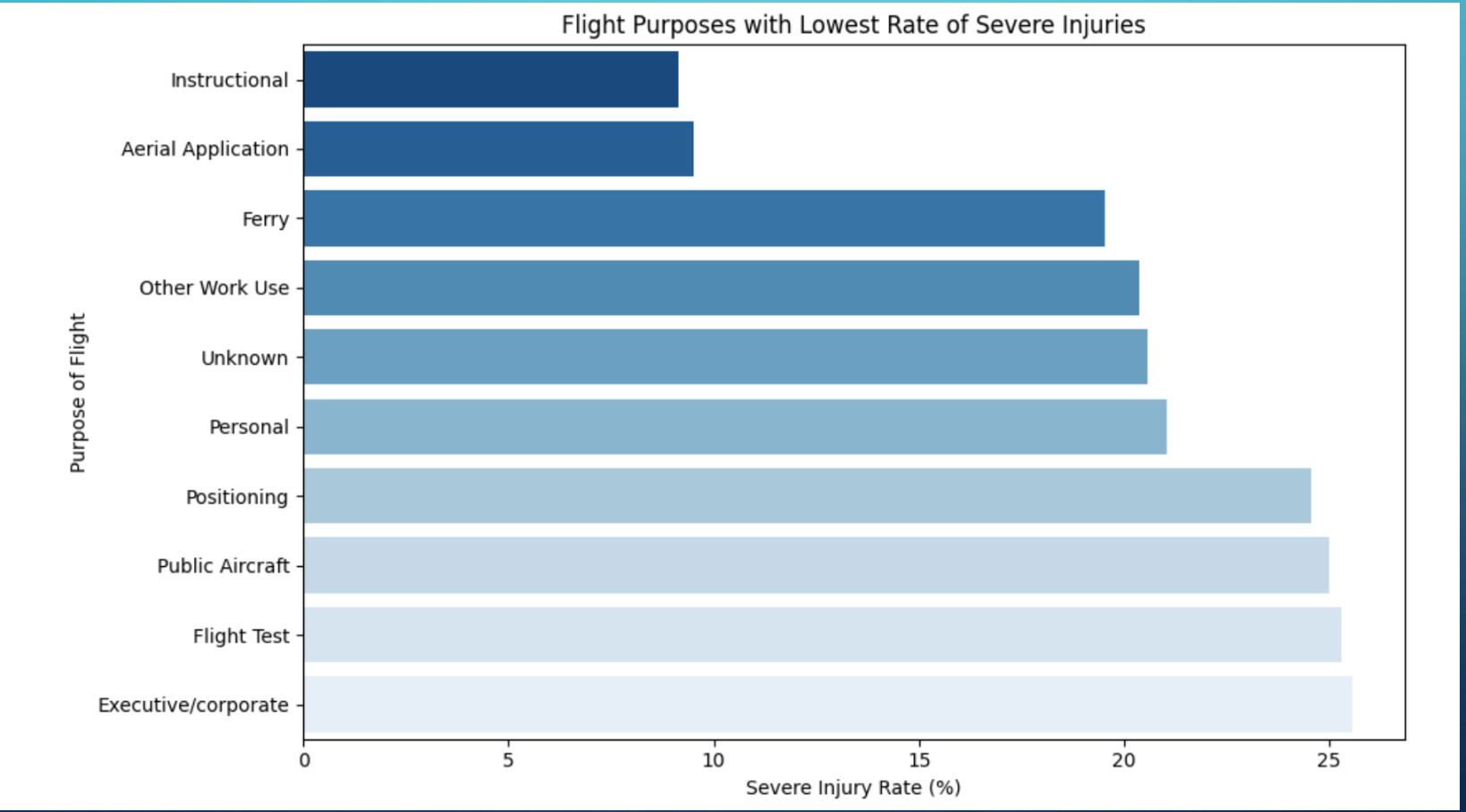
- We analysed the **top 10 aircraft makes** with the most accidents
- Compared their distribution across injury severities
- Some popular makes **Boeing** have **higher fatality rates** than others



PURPOSE OF FLIGHT VS. INJURY SEVERITY

Does Flight Purpose Affect Risk?

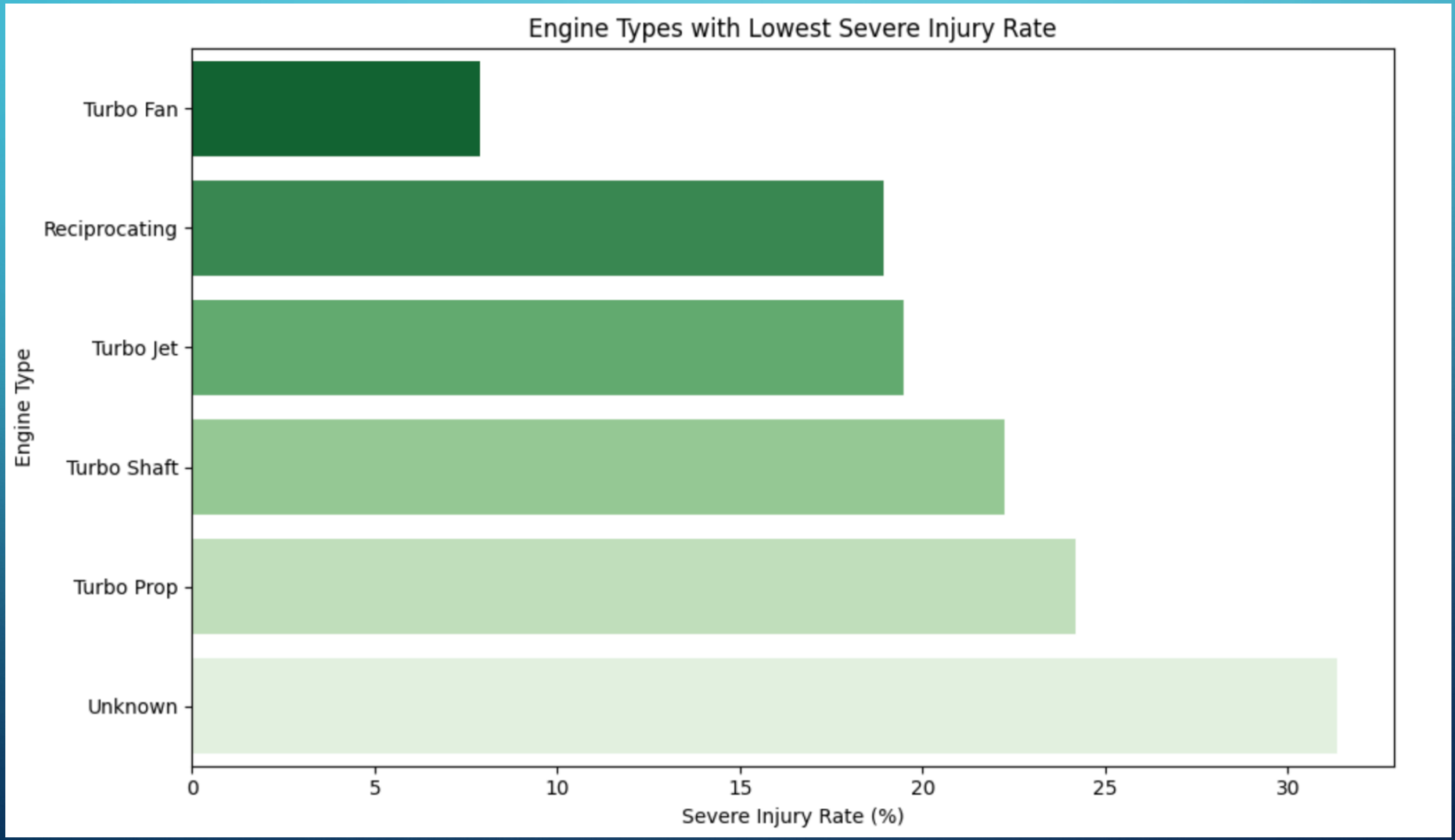
- **Corporate** and **Flight test flights** had the highest rate of severe injuries
- **Instructional and Aerial application flights** showed lower injury severity overall



ENGINE TYPE VS. INJURY SEVERITY

Are Some Engines Safer Than Others?

- Aircraft with **turbo prop** showed higher fatality rates
- **Turbofan** had better safety outcomes



KEY BUSINESS RECOMMENDATIONS

- **Favor Makes with Low Fatality Rates:** Aircraft such as Grumman ACFT COR-SCHWEIZER, Airbus, STINSON, and Raven show lower fatal/serious outcomes.
- **Avoid High-Risk Engine Types:** Certain engine types like turboprop and turboshaft, are linked to higher severity outcomes while turbofan is the safest.
- **Focus on Commercial/Business Flights:** Flight purposes like instructional, aerial application and public aircraft(local) had significantly higher severe injury rates

NEXT STEPS

Moving Forward

- Conduct **cost-benefit analysis** on recommended aircraft types
- Continue to **monitor safety data** post-purchase
- Explore **pilot experience and maintenance** data for further analysis

THANK YOU!

Open For Questions?

Contact: Fatuma Tari

Email: fatumatari@gmail.com

GitHub: https://github.com/Fatumatari/dsc-phase-1-project-v3/tree/master