Final Project Report Group 9

Ritik Ojha Sanjeet Naik

Executive Summary

Context of the Project: This analysis is vital for improving global aviation safety and efficiency by optimizing navigational aid (navaid) deployment and modernization across diverse airports and regions.

Main Objective: The project aims to pinpoint and address inefficiencies in navaid coverage, allocation, and calibration to enhance safety, resilience, and operational performance at airports worldwide.

Key Insights: Critical findings highlight underserved regions like Sub-Saharan Africa needing investment, mismatches between navaid complexity and airport traffic, and urgent recalibration needs in high magnetic variation areas like the South Atlantic.

Introduction: Context & Project Relevance

Problem Statement

The main problem being investigated in this project is the optimization of navigational aids (navaids) at airports worldwide. This involves a comprehensive analysis of how navaids are currently deployed, used, and maintained, with the ultimate goal of enhancing the safety, efficiency, and resilience of air navigation systems globally. The investigation tackles multiple interconnected aspects to identify gaps, inefficiencies, and opportunities for improvement.

Project Relevance

The overarching problem ties these questions together into a unified effort to improve global air navigation infrastructure. By analyzing current navaid distributions, identifying inefficiencies, andaddressing regional and operational disparities, the project aims to provide data-driven recommendations. These insights will help prioritize investments, upgrades, and maintenance to ensure that air navigation systems are safer, more efficient, and more resilient—whether by closing coverage gaps in underserved regions, modernizing outdated equipment, or optimizing resources for diverse airport needs.

Dataset Overview

This dataset, sourced from partow.net and World Airport Codes, provides detailed global airport information used to support flight operations, navigation, and data-driven decision-making in the aviation industry.

link:https://www.kaggle.com/datasets/sanjeetsinghnaik/world-airport-dataset/data?select=airports.csv

It comprises 4 datasets that have been used in this project. Airport.csv, countries.csv, navaids.csv, regions.csv, runways.csv

Main goal is about optimizing the global navaid ecosystem to meet modern aviation demands while addressing safety, efficiency, and resilience challenges across varied geographic, operational,
and environmental contexts.

Data Wrangling & Cleaning

Cleaning Process

Initial Cleaning & Missing-Value Handling -Fills missing municipality with “Unknown” and scheduled_service with “no”.
Title-cases the type field and uppercases country/region/airport codes.
Replaces any missing elevation_ft with the dataset’s median elevation.
Coerces latitude, longitude, and elevation to numeric.

Turns scheduled_service and type into factors for consistent categorical handling.
Feature Engineering: Geographic & Operational Attributes -Derives hemisphere_ns and hemisphere_ew from latitude/longitude signs.

Categorizes elevation into meaningful bins (Below Sea Level through Very High Elevation). Creates an airport_size label (Large, Medium, Small, Special, Other) based on type.

Extracts 2-letter region_group from iso_region.
Assesses coord_precision (“High”, “Medium”, “Low”) by rounding latitude/longitude.

Region-Level Summaries we will Aggregates by region_group and iso_region to compute: Number of airports, Average elevation, Percent offering scheduled service.

Merging in Country & Region Context Joins in country-level stats (total_airports, scheduled_service_pct) by iso_country. Joins region summaries by iso_region. on the Navigational aids data -Drops any rows missing an ident, name, or type (because those are essential).

Creates a new power_category factor (“Low”, “Medium”, “High”, or “Unknown”) based on the original power column if it exists, defaulting to “Unknown” otherwise.

Drops any rows missing an ident or name. perform preprocessing on airport_type making it lower case.
Derives an airport_size category by pattern-matching on airport_type (“International”, “Heliport”, “Small”, or “Other”). we’re taking both the datasets airport and navaids, tidying their column names, removing incomplete records, coercing key fields into the right formats, and

then enriching them with high-level categories (power bands for navaids; size classes for airports)

Size, missing values, inconsistencies, or errors. Handling of missing values, duplicates, transformations, and formatting.

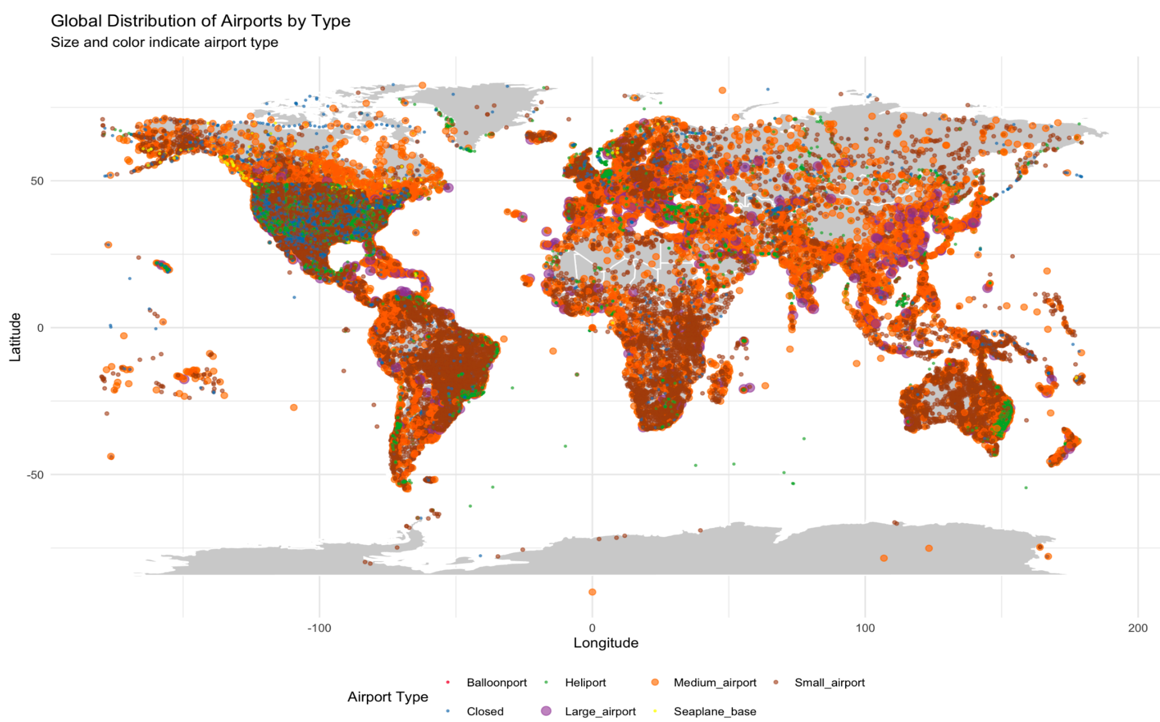
Cleaned column names and dropped rows missing key IDs (e.g.ident, name, type). Filled blanks (municipality→“Unknown”, scheduled_service→“no”); imputed elevation_ft with its median. Converted textcases (types lowercase/title case, codes uppercase) and coerced lat/long/elevation/frequency to numeric.

Engineered categories: hemispheres, elevation bins, airport sizes, power bands, region groups, coordinate precision. Aggregated region stats (counts, avg elevation, service.%) and merged country and region-level summaries back into the data.

Exploratory Data Analysis (EDA)

Visualization 1: Global Airport Distribution

Simple Analysis on Airport shows that there are maximum amount of heliport and closed airport in the united states as compared to any of the continent in the world. south america has basically the layout of bollon port and mix of medium size airport. baloon port has also been observe in large number in Africa and Australia, Large Airports are basically laid out around all of the world.

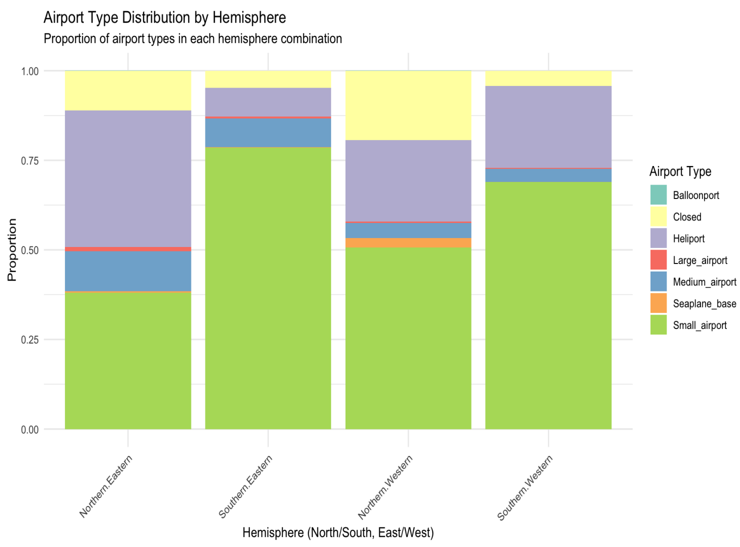


Visualization 3: Airport Types by Elevation

Us, brazil, Canada, Japan and Australia are among the top 5 countries list in the count of number of airports in the world. but previously, we have seen that America has maximum heliport mainly, and out of these 5, two countries named Brazil, australia has decent amount of balloonport, plus Japan and canada has mostly comprises of medium size airport. arpport distribution for canada and japan is similar but percent scheduled services in canada is much higher than in japan.

Visualization 4: Hemisphere Analysis

We are going to visualize this on hemisphere level as well, and find some insights if on a global scale distribution of airport to find some insights. Northern Eastern regions show a balanced mix of airport types, with heliports and small airports each making up roughly 40–45% of facilities, and medium airfields contributing about 10%. Southern Eastern skies are overwhelmingly served by small airports (~78%), with closed stations and heliports sharing the remainder equally (around 10% each). In Northern Western areas, small airports drop to ~50% while closed facilities surge to 20% and heliports hover near 25%, reflecting harsher climates or lower demand. Southern Western airspace resembles Southern Eastern but with a slightly higher heliport presence (~25%) and fewer closed fields (~5%). Across all hemispheres, large airports and seaplane bases remain extremely rare (<2%), underscoring the predominance of smaller, more flexible airfields in global aviation infrastructure.



Visualization on navigational aids type count High-power beacons overwhelmingly dominate the navaid landscape, with NDBs (~6,500) and VOR-DMEs (~2,500) far outnumbering other types, while virtually no aids lack power information. Frequency-wise, both heliports and “Other” airports cluster tightly around the 100–120kHz band, whereas small airports show minimal frequency usage, reflecting their reliance on very low- or non-transmitting navaids.

Navigational Aids type Vortac, vor is prominent in usa where heliport are the largest in number, DME is prominent in australia and russian region. DME is also prominent european countries, mix of vor and dme in western europe, africa is also has a mix of vor, tacan. southern asia has navigational aid types of vor, dme.

Magnetic variation As discussed in the previous slide the magnetic variations are strong in the west region, we have used spatial plot to visualize this phenomenon to deep dive more. Magnetic variation has been calculated by the degree of magnetic variation shown in that region. we binned it in different bucket, and labeled accordingly. Looking at the leaflet plot, strong west variation has been recognised around Atlantic region, strong east in the Alaskan region and in Canada. Most of the minimal variation has been observed
around African continent(, Europe) and south Asian countries, with moderate east and strong east in the Northern Asian countries around Russia.

Key Insights & Discussion

Regional Coverage & Gap Analysis
Sub-Saharan Africa and the Pacific Islands have very few high-power navaids ("High" power category) relative to the number of small and heliport airfields, while South America shows a moderate mix but still lags behind North America and Europe. Implication: Pilots operating in these regions face reduced redundancy—adding mid- or high-power aids (e.g. VOR-DME) would greatly improve safety and enable instrument approaches where none currently exist.

Airport Complexity vs. Traffic Demand
Finding: Our complexity score (navaids×type diversity) correlates strongly with known hub traffic—major international airports (e.g. San Francisco Bay Airdrome with 17 navaids and 5 types) match their high passenger/cargo volumes, whereas some medium-traffic airports (e.g. Stow Airstrip) are slightly over-equipped relative to their usage. Implication: We can consider redistributing one or two lower-utility aids from over-equipped mid-level airports to rapidly growing but under-served fields, aligning infrastructure with actual demand.

Frequency-Band Allocation Optimization
Finding: Heliports and "Other" small-field airports almost exclusively use LF bands (~100–120kHz), even in mountainous Western regions where VHF (108–117MHz) would offer better line-of-sight performance.
Implication: Re-allocating some small-field aids to VHF (e.g. adding miniature VOR or VOR-DME) in rugged terrain could reduce signal blockage, improving reliability and reducing maintenance from static interference.

Magnetic Variation & Calibration Needs
Finding: NDBs and TACANs in the Eastern region show the widest spread of magnetic variation (±50°+ outliers), while NDB-DMEs cluster tightly near zero variation. Western aids

(mostly VOR‑DME/VORTAC) uniformly sit in moderate east variation (10–30°). Implication: Immediate recalibration should focus on single‑mode NDBs/TACANs in high‑variation zones (e.g. northern Canada), whereas mixed‑mode installations can tolerate broader shifts but still benefit from periodic checks.

Emergency Routing & Redundancy Planning
Finding: Top‑traffic airports have at least 3 alternative high‑power aids within 50km, but several medium hubs have only a single aid of each type. Implication: For those single‑aid airports, emergency plans must designate specific diversion fields (with overlapping coverage) or install an extra DME/VOR‑DME to guarantee continuous guidance if one beacon fails.

Infrastructure Modernization Roadmap
Finding: Legacy NDBs account for ~70% of all high‑power aids, particularly in developing regions, whereas newer VOR‑DMEs and VORTACs are concentrated in North America and Europe. Implication: Prioritize replacing isolated NDBs with VOR‑DME or satellite RNAV on routes serving commercial traffic; this will reduce maintenance and provide distance‑and‑bearing in one unit.

Correlation of Airport Type & Navaid Mix
Finding: Military and hybrid fields (VORTAC‑equipped) are almost exclusively in the Western hemisphere, while civilian heliports rely on low‑power NDBs and occasional VORs. Commercial hubs mix VOR‑DME with VORTAC for redundancy. Implication: Procurement should follow the role: military bases keep TACAN/VORTAC, remote heliports can continue using low‑power NDBs, and commercial fields focus on mixed VOR‑DME/VORTAC installations.

Environmental & Terrain Impact
Finding: Desert and polar regions (High Elevation, Southern Western and Northern Western hemispheres) show under‑deployment of VOR/VOR‑DME—fields there still depend heavily on NDBs despite line‑of‑sight challenges. Implication: Installing more medium‑range VOR‑DME in these challenging terrains will improve all‑weather access and reduce weather‑related cancellations.

Conclusion

Our analysis reveals that vast regions—especially Sub‑Saharan Africa and island nations—are critically under served by high‑power and mixed‑mode navaids, creating safety and redundancy gaps. Airport “complexity” scores generally track passenger/cargo volumes, but a few mid‑traffic fields are over‑equipped, suggesting we could reallocate some aids to growing but under‑served airports. Small‑field aids overwhelmingly use low‑frequency bands even in terrain that would benefit from VHF, so shifting some installations to VOR or VOR‑DME would improve reliability. Magnetic‑variation outliers concentrate among single‑mode NDBs and TACANs in high‑declination zones, flagging them for immediate recalibration. Emergency planning shows major hubs have sufficient alternate aids, but many medium airports lack proper back-ups—adding just one

References

Data: https://www.kaggle.com/datasets/sanjeetsinghnaik/world-airport-dataset/data?select=airports.csv