Global Aviation Data Analysis Report

1. Introduction

This report presents an exploratory data analysis (EDA) of global aviation datasets, focusing on airports, airlines, and flight routes. The aim is to uncover insights into airline connectivity, airport traffic, and geographic distributions to support strategic decision-making.

2. Dataset Overview

Airports Dataset: Contains detailed information on airport locations, codes, and characteristics.

Airlines Dataset: Includes airline identifiers, country of operation, and activity status.

Routes Dataset: Details the flight routes between airports operated by different airlines.

3. Data Cleaning Summary

Checked and removed duplicate rows.

Validated geographic coordinates (latitude/longitude).

Ensured data type consistency for each column.

Verified route airport codes exist in the airports dataset.

4. Summary Statistics (Numerical Variables)

Airports

| Variable | Mean | Std | Min | Max |
| --- | --- | --- | --- | --- |
| Latitude | XX.XX | X.XX | -90.0 | 90.0 |
| Longitude | XX.XX | X.XX | -180 | 180 |
| Altitude | XXX.X | XX.X | -XXX | XXXX |

Routes

| Variable | Mean | Std | Min | Max |
| --- | --- | --- | --- | --- |
| Stops | 0.1 | 0.3 | 0 | 2 |
| Distance\_km | XXXX | XXX | XX | XXXX |

5. Frequency Distributions (Categorical Variables)

Airlines

Activity Status:

Active: XXX airlines

Inactive: XXX airlines

Top 5 Countries by Airline Count:

Country A: XX airlines

Country B: XX airlines

Routes

Top 5 Airlines by Route Count:

Airline A: XXX routes

Airline B: XXX routes

6. Key Findings

Top Airlines:

Airline A operates the most routes (XXX).

Busiest Airports:

Airport X has the highest number of connections (XXXX).

Airline Country Distribution:

Most airlines are based in Country X.

7. Geographic Distribution Analysis

Map of Global Airport Locations: Scatter plot with country coloring.

Route Density Map: Visualizing high-density corridors.

8. Airline Activity by Country

Bar Chart: Airlines per country (active vs. inactive).

Pie Chart: Overall airline activity status breakdown.

9. Visualizations Gallery

Global Airport Map (colored by country)

Bar chart: Top Airlines by Route Count

Bar chart: Top Airports by Connections

Histogram: Route Distance Distribution

10. Conclusion

The analysis revealed significant hubs, dominant airlines, and key geographic trends. These insights can inform route planning, infrastructure investment, and policy-making in global aviation.

*Note: Visualizations and tables can be embedded as images or interactive plots where applicable.*

A graph with a red and green rectangle

AI-generated content may be incorrect.A graph with green bars

AI-generated content may be incorrect.A graph of the number of airlines

AI-generated content may be incorrect.Airlines Numerical Dataset

|  |  |
| --- | --- |
|  | airline\_id |
| count | 6160.00 |
| mean | 4154.07 |
| std | 4507.42 |
| min | 2.00 |
| 25% | 1542.75 |
| 50% | 3083.50 |
| 75% | 4629.25 |
| max | 21317.00 |

Airports Numerical Dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | airport\_id | latitude | longitude | altitude | timezone |
| count | 7543.00 | 7543.00 | 7543.00 | 7543.00 | 7213.00 |
| mean | 5043.54 | 25.84 | -2.42 | 1004.71 | 0.22 |
| std | 3684.47 | 28.43 | 86.42 | 1611.78 | 5.71 |
| min | 1.00 | -90.00 | -179.88 | -1266.00 | -12.00 |
| 25% | 1951.50 | 6.95 | -79.39 | 62.50 | -5.00 |
| 50% | 3982.00 | 34.13 | 5.71 | 352.00 | 1.00 |
| 75% | 7576.50 | 47.37 | 53.27 | 1180.50 | 3.50 |
| max | 13726.00 | 82.52 | 179.95 | 14472.00 | 13.00 |

Routes Numerical Dataset

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | airline\_id | source\_airport\_id | dest\_airport\_id | stops |
| count | 67184.00 | 67443.00 | 67442.00 | 67663.00 |
| mean | 3520.74 | 2693.69 | 2695.74 | 0.00 |
| std | 3527.47 | 1672.39 | 1673.31 | 0.01 |
| min | 10.00 | 1.00 | 1.00 | 0.00 |
| 25% | 1355.00 | 1382.00 | 1382.00 | 0.00 |
| 50% | 2987.00 | 2964.00 | 2964.50 | 0.00 |
| 75% | 4609.00 | 3670.00 | 3670.00 | 0.00 |
| max | 21012.00 | 11922.00 | 11922.00 | 1.00 |

Airline and Routes Count

|  |  |  |
| --- | --- | --- |
|  | airline | route\_count |
| 199 | FR | 2484 |
| 96 | AA | 2354 |
| 479 | UA | 2180 |
| 160 | DL | 1981 |
| 490 | US | 1960 |
| 150 | CZ | 1454 |
| 332 | MU | 1263 |
| 140 | CA | 1260 |
| 522 | WN | 1146 |
| 476 | U2 | 1130 |

Top Ten Busiest Airports by Connections

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | IATA Code | Airport Name | Country | Connections |
| 165 | ATL | Hartsfield Jackson Atlanta International Airport | United States | 1826 |
| 2128 | ORD | Chicago O'Hare International Airport | United States | 1108 |
| 2200 | PEK | Beijing Capital International Airport | China | 1069 |
| 1626 | LHR | London Heathrow Airport | United Kingdom | 1051 |
| 463 | CDG | Charles de Gaulle International Airport | France | 1041 |
| 890 | FRA | Frankfurt am Main Airport | Germany | 990 |
| 1573 | LAX | Los Angeles International Airport | United States | 990 |
| 667 | DFW | Dallas Fort Worth International Airport | United States | 936 |
| 1293 | JFK | John F Kennedy International Airport | United States | 911 |
| 110 | AMS | Amsterdam Airport Schiphol | Netherlands | 903 |

Top 10 Countries by Airport Count

|  |  |  |
| --- | --- | --- |
|  | Country | Number of Airports |
| 0 | United States | 1503 |
| 1 | Canada | 428 |
| 2 | Australia | 328 |
| 3 | Brazil | 260 |
| 4 | Russia | 251 |
| 5 | Germany | 248 |
| 6 | France | 217 |
| 7 | China | 207 |
| 8 | United Kingdom | 165 |
| 9 | India | 145 |

