**Exploratory analysis of Airline's past dataset**

**Introduction:**

This dataset consists of data which is related to the timings of the flight's departure and arrival and the information related to source and destination. Additionally, it also consists of customer basic details like the name, passenger id, gender, and age.

More details like the name of the pilot, and flight status (On time, canceled, delayed).

**Data Inspection:**

Dataset Size: 98620 records

Columns: 15

Data Types:

Numeric: Age

Categorical: Passenger ID, First Name, Last Name, Gender, Age, Nationality, Airport Name, Airport Country Code, Country Name, Airport Continent, Continents, Departure Date, Arrival Airport, Pilot Name, Flight Status

Date: Departure Date

**Incorrect values:**

Arrival Airport has incorrect values “0” and “–“ which need to be handled with the help of other columns or can be filtered out as a part of data cleaning.

**Descriptive Statistics:**

Summary statistics for 'Age':

count 98619.000000

mean 45.504021

std 25.929849

min 1.000000

25% 23.000000

50% 46.000000

75% 68.000000

max 90.000000

Name: Age, datatype: float64

A graph with a line and a line

Description automatically generated

**Data visualization:**

This graph represents the age distribution calculated from the data available in the dataset.

A graph of age distribution

Description automatically generated

**Gender distribution:**

Gender

Male 49598

Female 49021

**A graph of a number of blue rectangular objects

Description automatically generated with medium confidence**

**Temporal Analysis:**

Yearly flight counts:

Year

2022.0 59658

The dataset contains data for the year 2022 only.

**Geographical Analysis:**

Top 10 countries by flight counts:

Country Name

United States 22104

Australia 6370

Canada 5424

Brazil 4504

Papua New Guinea 4081

China 2779

Indonesia 2358

Russian Federation 2247

Colombia 1643

India 1486

**A graph of blue squares

Description automatically generated**

**Top 10 arrival airports by airport code:**

**A graph of blue bars

Description automatically generated**

**Segmentation:** This data shows the conclusion of the flight's journey (Delayed, On-time, Cancelled). A graph of flight status

Description automatically generated

**Conclusion:**

To sum up, the dataset's exploratory analysis has given rise to a fundamental comprehension of the main characteristics, distributions, and connections within the data. For further research, predictive modeling, or decision-making about flight operations, passenger experiences, or airline performance, these insights can be used as a springboard. The analysis's conclusions can direct future research and help the aviation sector make data-driven decisions.