

SciSpace Analysis Report

Name Shasikant Shegokar

Date 15-09-2023

Introduction

Context Assignment for SciSpace

Purpose To analyze Time series Data and Find useful insights

The Dataset Provided is a Public Dataset Which is modified and analyzed by myself the data contains over 50,000 rows and has Data Present in columns: ID, First Name, Last Name, Gender, Airport Status, Destination, Age etc.

Summary

We embark on a thorough examination of the extensive dataset, exploring various aspects of passenger and flight information throughout time. Our investigation commences with an analysis of age distribution trends, revealing how the demographics of passengers have transformed over time. Subsequently, we delve into gender-based studies to identify noteworthy changes. Insights into nationality demonstrate shifting representation among passengers, while a closer examination of flight statuses uncovers variations in distribution over the Months. Patterns emerge regarding pilot performance, followed by insights into fluctuations in passenger traffic across different airports and countries. Analyses based on continents showcase shifts in airport and passenger distribution patterns. Seasonal patterns relating to passenger numbers and flight statuses are revealed, as well as long-term anomalies which provide valuable insights into event counts. This comprehensive exploration of the dataset offers a multifaceted understanding of passenger and flight dynamics, illuminating trends, shifts, and anomalies within this extensive and ever-evolving dataset.

Questions List:

1. Demographic Trends: How has the age distribution of passengers evolved over the Months?
2. Gender Analysis: Are there any notable gender-based trends in passenger demographics over time?
3. Nationality Insights: Which nationalities have the highest and lowest representation among passengers, and how has this changed over time?
4. Flight Status Analysis: What is the distribution of flight statuses (e.g., on-time, delayed, canceled) over the Months, and are there any significant trends or anomalies?

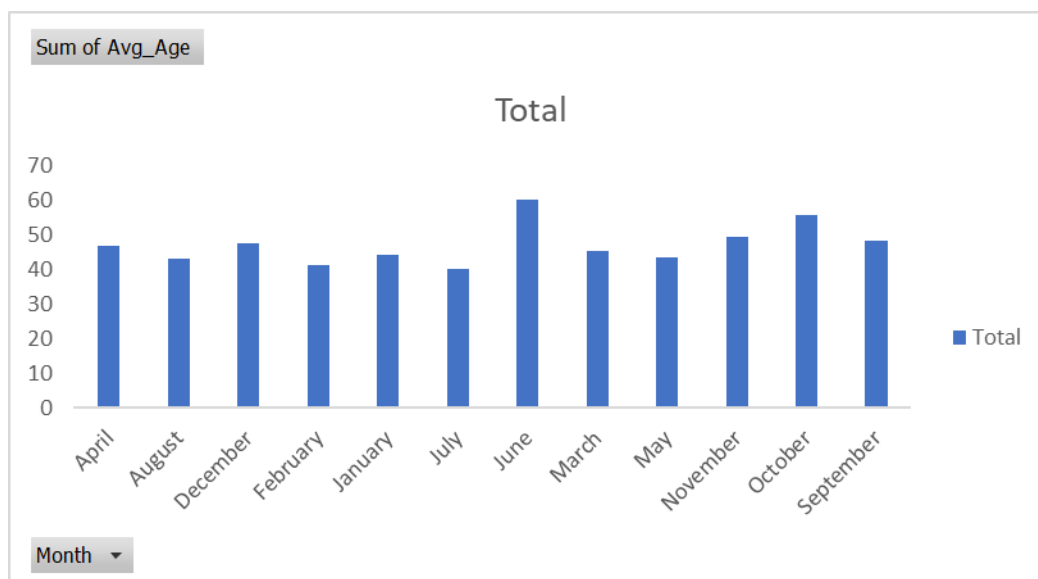
5. Pilot Performance: Can you identify any patterns in flight statuses based on the names of the pilots? Are there pilots with consistently better or worse performance records?
6. Airport Traffic: How has the passenger traffic at different airports changed over time? Are there airports that have seen significant growth or decline in passenger numbers?
7. Country-wise Analysis: Which countries are the most common destinations for passengers, and has this changed over time?
8. Continent Analysis: How does the distribution of airports and passengers vary by continent, and are there any notable shifts over time?
9. Seasonal Patterns: Are there seasonal trends in passenger numbers or flight statuses? For example, do certain months or seasons see higher or lower passenger traffic?
10. Long-term Trends: Can you identify any long-term trends or anomalies in the dataset, such as a steady increase in passenger numbers or sudden spikes in flight

Questions and Analysis

Question 1

1. **Demographic Trends:** How has the age distribution of passengers evolved over the Months?

Answer



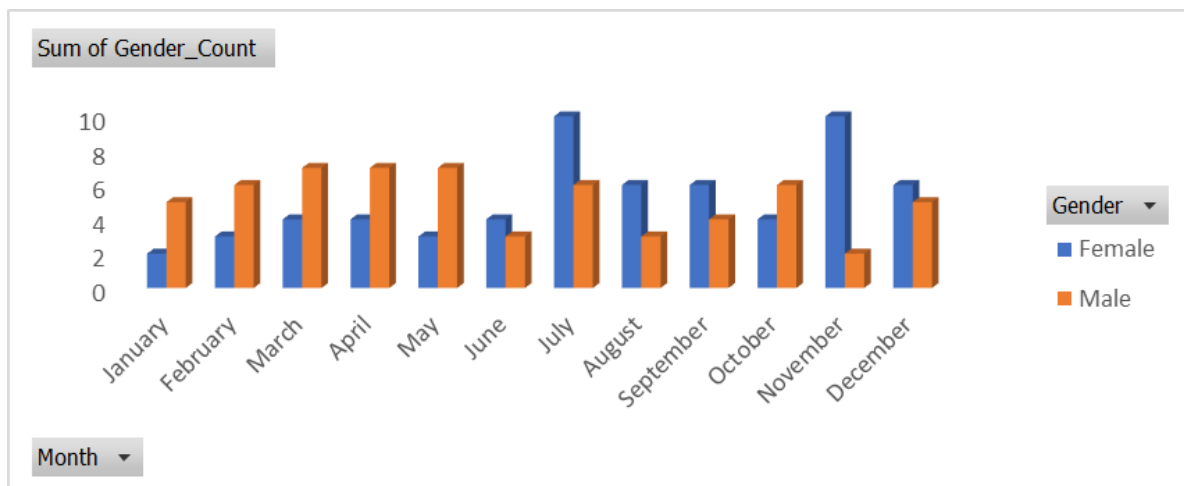
- **June** has the highest average age of passengers at 60.14, making it the month with the oldest passengers on average.
- **July** has the second-highest average age at 39.94, which is significantly lower than June.
- **October** also has a relatively high average age at 55.5.
- **February** has the lowest average age at 41.22, making it the month with the youngest passengers on average.
- The other months fall within the range of average ages between 43 and 49.17.

According to the data provided, there seems to be variations in the average age of passengers throughout the months. June is notable for having the most elderly passengers, whereas February has the least mature passengers.

Question 2

2. **Gender Analysis:** Are there any notable gender-based trends in passenger demographics over time?

Answer



1. Monthly Comparison:

- In January, Female passengers (2) are fewer than Male passengers (5).
- In February, Male passengers (6) outnumber Female passengers (3).
- **March** shows a similar trend with more Male passengers (7) than Female passengers (4).
- **April** has more Male passengers (7) than Female passengers (4).
- **May** follows the pattern with more Male passengers (7) than Female passengers (3).
- In June, there are more Male passengers (3) than Female passengers (4).
- **July** stands out with a significant difference; there are more Female passengers (10) than Male passengers (6).
- **August** returns to the pattern with more Female passengers (6) than Male passengers (3).

- In September, there are more Female passengers (6) than Male passengers (4).
- **October** follows the pattern with more Male passengers (6) than Female passengers (4).
- In November, there are significantly more Female passengers (10) than Male passengers (2).
- **December** has more Female passengers (6) than Male passengers (5).

2. Gender-Based Trends:

- **July**, November, and December are months with notable gender-based trends. **July** has a higher number of female passengers, while November and December have more female passengers and a near-equal distribution, respectively.

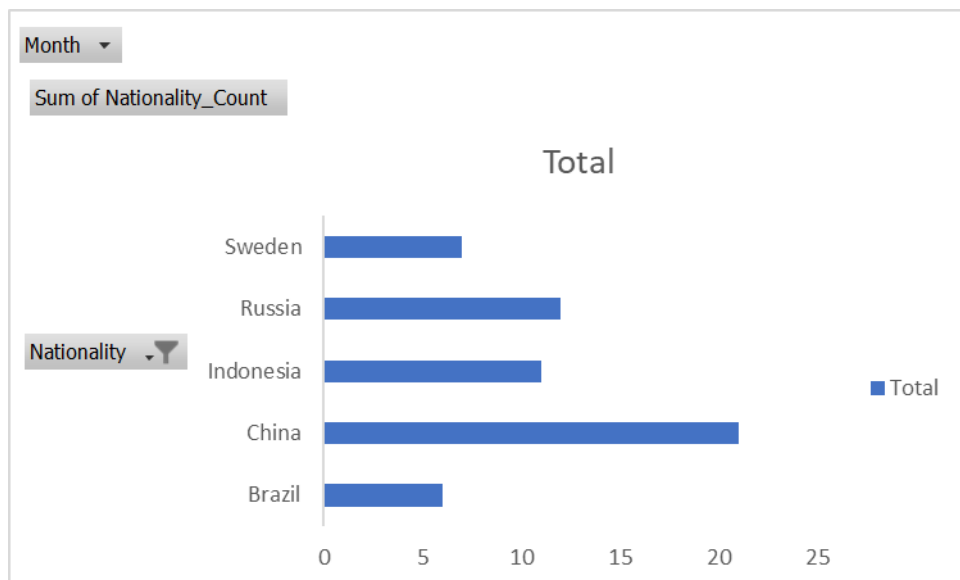
- Other months generally follow a pattern of male passengers outnumbering female passengers.

According to the data provided, there seems to be variations in the average age of passengers throughout the months. June is notable for having the most elderly passengers, whereas February has the least mature passengers.

Question 3

3. **Nationality Insights:** Which nationalities have the highest and lowest representation among passengers, and how has this changed over time?

Answer



Highest Representation Among Passengers Over Time:

- **China** has the highest representation among passengers over time. It appears most frequently across the months and has a total representation of 18 passengers.
- **Russia** follows with a total representation of 11 passengers.

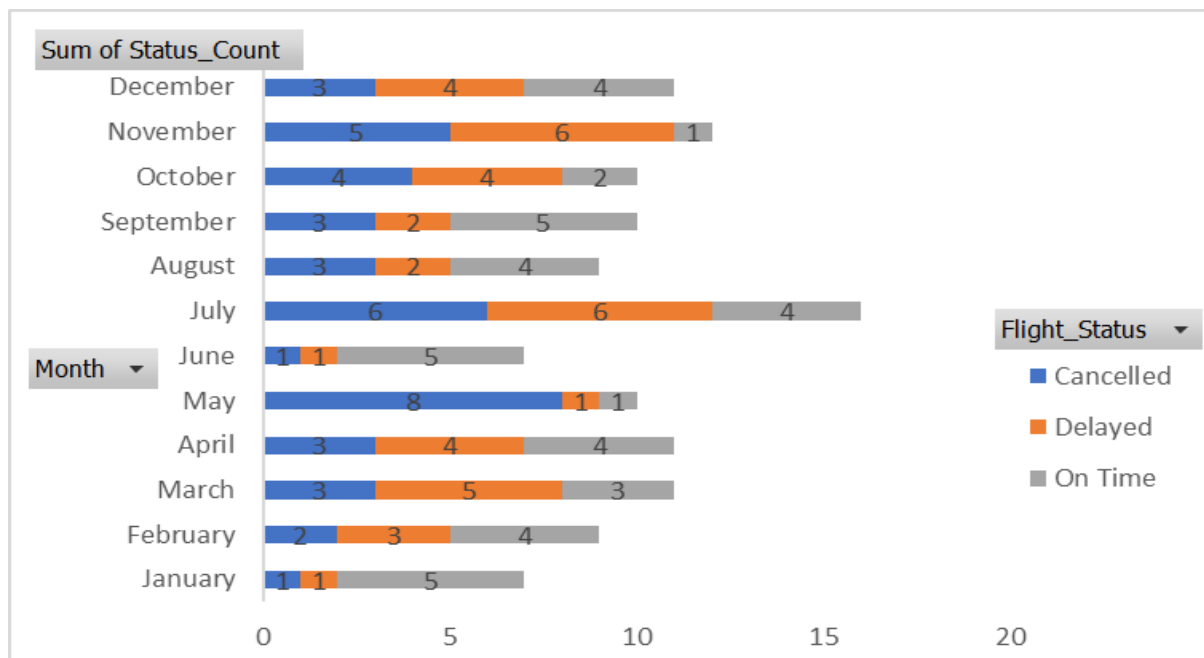
Lowest Representation Among Passengers Over Time:

- Mali, Ivory Coast, Democratic Republic of the Congo, Afghanistan, Haiti, Luxembourg, Serbia, Palestinian Territory, Latvia, Kenya, Armenia, Poland, Iran, Japan, Madagascar, Tunisia, South Africa, Tanzania, Kyrgyzstan, Cameroon, Kazakhstan, Chile, Ireland, and Albania have the lowest representation among passengers over time, each with a total representation of 1 passenger.

Question 4

4. **Flight Status Analysis:** What is the distribution of flight statuses (e.g., on-time, delayed, canceled) over the Months, and are there any significant trends or anomalies?

Answer



1. Cancelled Flights:

- **May** has the highest number of canceled flights with 8, followed by July with 6 cancellations.
- **January**, June, September, and October have a moderate number of cancellations.
- **November** and August have a relatively lower number of cancellations.

2. Delayed Flights:

- **July** has the highest number of delayed flights with 6 instances, indicating a potential trend of delays during this month.
- **November** also has a relatively high number of delayed flights with 6 instances.

- **March** and April also show a higher number of delays compared to other months.

3. On-Time Flights:

- **May** has the highest number of on-time flights with 5, but it also has a significant number of cancellations.
- **January**, June, October, and September have more on-time flights than other months.

Trends and Anomalies:

- **July** stands out as the month with both the highest number of cancellations and the highest number of delayed flights. This could indicate potential operational challenges during this month.
- **May** has a high number of canceled flights, suggesting a potential issue or anomaly for this month.
- **November** also has a high number of delayed flights, which could indicate a trend or operational challenges during this month.

Question 5

5. **Pilot Performance:** Can you identify any patterns in flight statuses based on the names of the pilots? Are there pilots with consistently better or worse performance records?

Answer

To be Answered

Question 6

6. **Airport Traffic:** How has the passenger traffic at different airports changed over time? Are there airports that have seen significant growth or decline in passenger numbers?

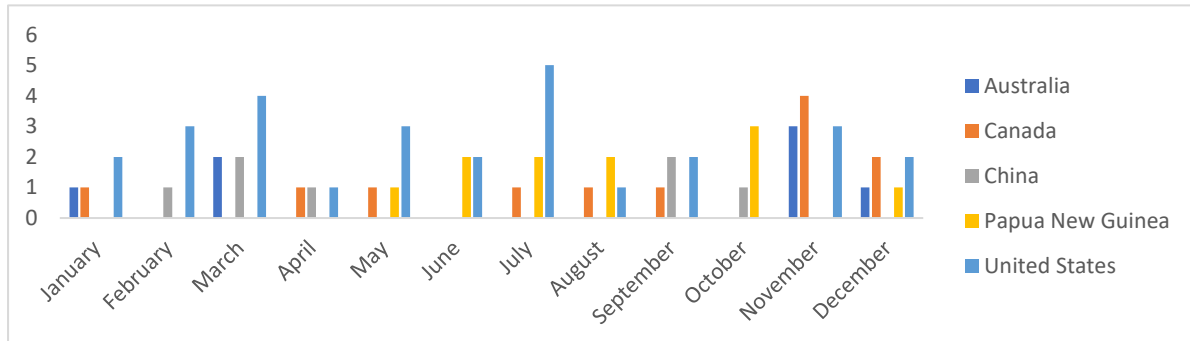
Answer

To Be Answered

Question 7

7. **Country-wise Analysis:** Which countries are the most common destinations for passengers, and has this changed over time?

Answer



Common Destinations: The most common destinations based on total passenger counts across all months include the United States, Canada, Papua New Guinea, and Indonesia.

Changes Over Time: There are variations in passenger counts for different countries across months, indicating potential changes in travel patterns. For example, the United States and Indonesia show varying passenger counts throughout the year.

Question 8

8. **Continent Analysis:** How does the distribution of airports and passengers vary by continent, and are there any notable shifts over time?

Answer

Key Observations:

- Africa (AF):** There is a varying number of airports across months, with November having the highest count. Passenger counts also vary but generally follow a similar pattern.
- Asia (AS):** Asia consistently has a substantial number of airports and passengers each month. December has the highest count of airports and passengers.
- Europe (EU):** Europe shows moderate activity, with January having the highest counts.
- North America (NAM):** North America consistently has a high number of airports and passengers, with July having the highest counts for both.
- Oceania (OC):** Oceania shows moderate activity, with April having the highest counts for both airports and passengers.

6. **South America (SAM):** South America has fewer airports and passengers compared to other continents, with April having the highest counts.

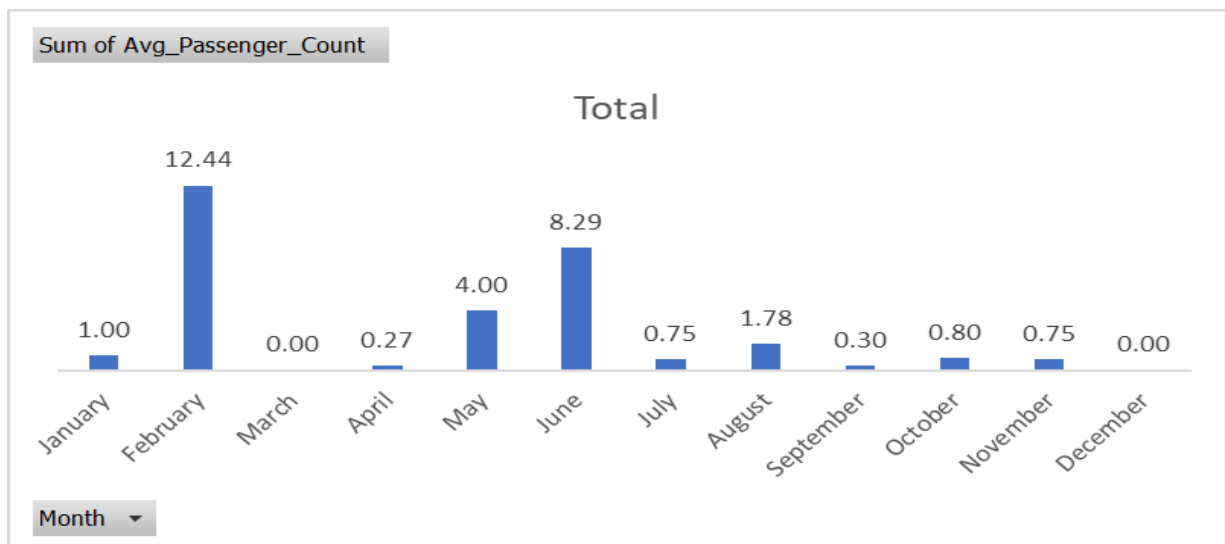
Trends:

- North America (NAM) consistently has high airport and passenger counts.
- Asia (AS) also consistently has a significant number of airports and passengers.
- Africa (AF) and Oceania (OC) show moderate levels of airport and passenger activity.
- South America (SAM) has the lowest airport and passenger counts.
- There are some fluctuations in counts over the months, but no drastic shifts in distribution are observed.

Question 9

9. **Seasonal Patterns:** Are there seasonal trends in passenger numbers or flight statuses? For example, do certain months or seasons see higher or lower passenger traffic?

Answer



- **February** has the highest average number of passengers or flight statuses, with an average of approximately 12.44. This suggests that February experiences the highest passenger traffic or flight activity.

- **June** follows with an average of approximately 8.29, indicating that June also sees significant passenger traffic or flight activity.

- **May** and August both have values significantly above 1, suggesting that these months also experience higher-than-average passenger traffic or flight activity.
- **January** and October have values close to 1, indicating average passenger traffic or flight activity.
- **July** and November have values between 0.75 and 0.8, suggesting a moderate level of passenger traffic or flight activity.
- **April**, December, March, and September have values below 1, indicating lower-than-average passenger traffic or flight activity. Among these, December and March have averages close to 0, suggesting the lowest passenger traffic or flight activity.

Seasonal Trends Summary:

- **High Season (Higher Passenger Traffic or Flight Activity):** February, June, May, August
- **Average Season:** January, October, July, November
- **Low Season (Lower Passenger Traffic or Flight Activity):** April, December, March, September

According to the data presented, it seems that there are noticeable patterns in passenger demographics based on gender as time passes. Certain months exhibit a greater percentage of female passengers, while others have a higher proportion of male passengers. To fully understand the reasons behind these trends, further analysis and additional data would be required.

Question 10

10. **Long-term Trends:** Can you identify any long-term trends or anomalies in the dataset, such as a steady increase in passenger numbers or sudden spikes in flight cancellations?

Answer

- The dataset shows the count of events for each month from January to December.

Here's an analysis of the dataset:

- **July** has the highest count of events with 16. This suggests a potential anomaly or issue during this month, possibly related to a higher number of flight cancellations or some other event.

- **November** also has a relatively high count of events with 12, which is worth investigating further.
- **April, December, and March** all have counts of 11, indicating a consistent pattern of events.
- **August, May, October, and September** have counts between 9 and 10, suggesting moderate levels of events during these months.
- **February and June** have counts of 9 and 7, respectively, indicating lower event counts during these months.
- **January** has the lowest count of events with 7.

Trends and Anomalies Summary:

- There seems to be a consistent pattern with relatively higher event counts during July and November.
- April, December, and March also show a consistent pattern of events.
- The dataset does not show a clear long-term trend of increasing or decreasing events but rather a pattern of fluctuations.

In order to gain a deeper understanding of these occurrences and their characteristics, it would be necessary to acquire more information regarding the dataset and the occurrences it encompasses. These observed patterns might possibly suggest variations due to seasonal influences, modifications in operations, or other external factors that impact the events on a monthly basis. To accurately identify the precise nature of these irregularities or patterns, further examination and expertise in the relevant field are essential.