**Fundamentals of Data Science**

**Name: Tarun Reddy Salevenkayalla**

**ID: 23098031**

# Introduction

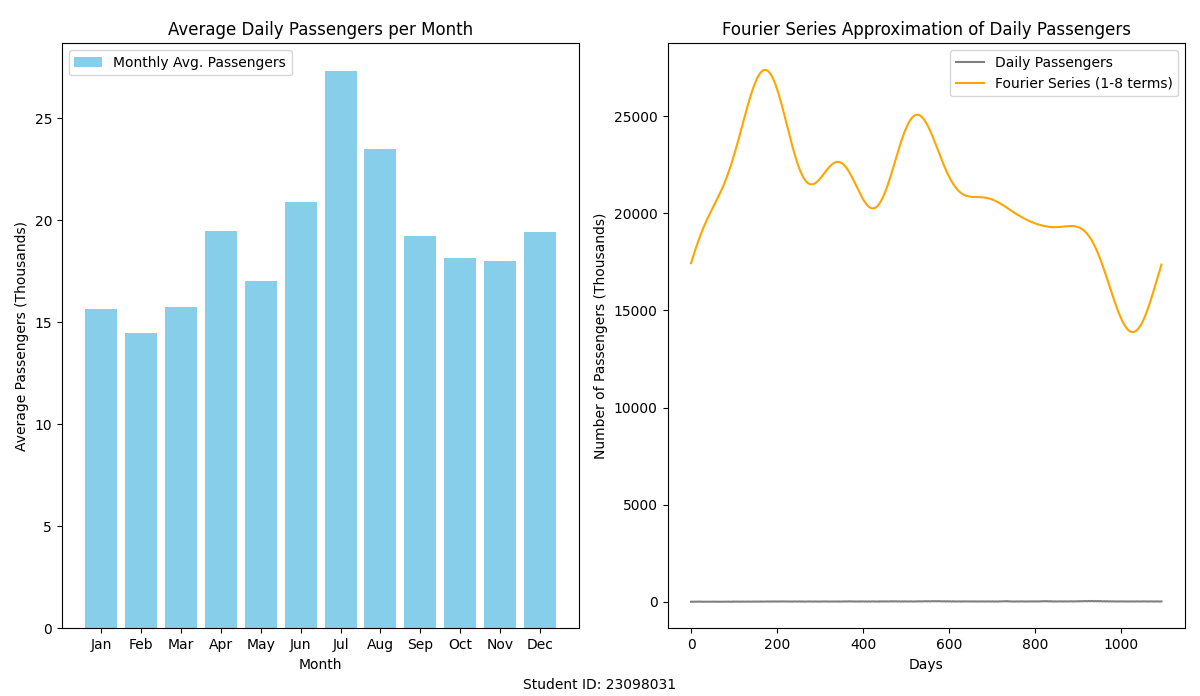
This report is based on a set of daily performances of the airline industry particularly Passenger traffic and Ticket prices. The method intended for use in the analysis processes includes Fourier transforms to reveal time fluctuations in terms of passenger flow. Passenger average is computed for each month and approximated using a Fourier series. Moreover, the power spectrum of the passenger variation is computed to identify the contribution of each frequency. The report also holds understanding about the average ticket prices in 2021 & 2022 based on the provided dataset.

# Data Overview and Preprocessing

The airline3 dataset used in this analysis includes daily records of an airline’s performance from 2021 to 2022, with three key variables: the date, the number of passengers it transported in thousands and the average ticket price in Euros. The data covers several years with the day-to-day records taken as the flight output of each flight.

Before analyzing the data, the ‘Date’ column is first transformed into datetime format to extract other features such as year, month and day from it. The Altery uses the ‘Number’ field, which contains the daily passenger count, in order to do a Fourier transform to reveal cyclical patterns (Müller *et al.*, 2021). Further, to investigate variations in the impact on daily passenger traffic depending on the day of the week, averages per month are also elaborated in passenger numbers, along with the cleaning of the dataset in order to avoid missing or inaccurate records in the further analysis.

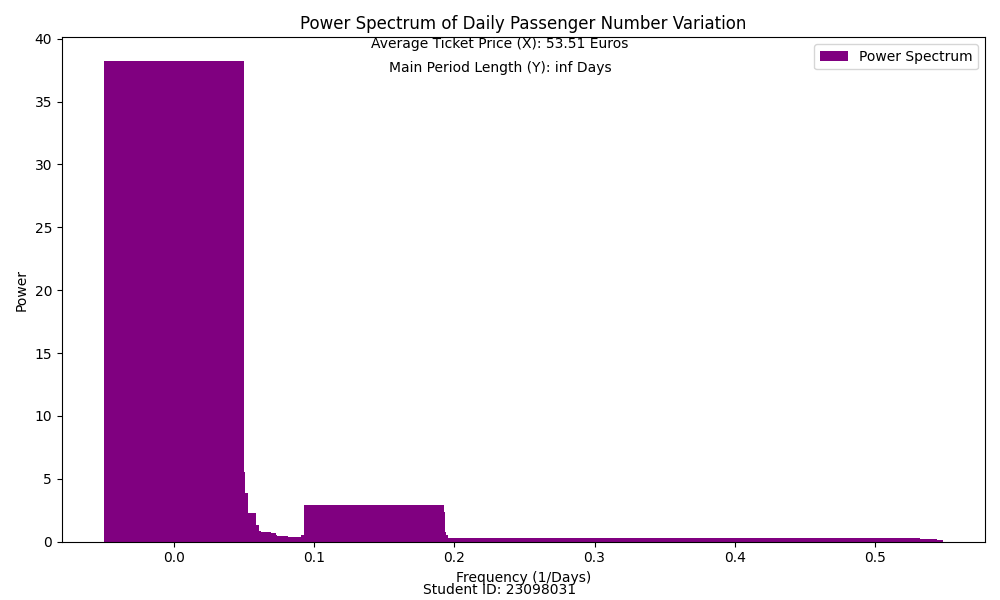
# Monthly Average Passengers and Fourier Series Approximation



**Figure 1: Monthly Average Passengers and Fourier Approximation**

Specifically, the empirical analysis of the airline’s passengers has been depicted in the figure 1 below. The left panel represents the Passenger count as a bar chart of the average number of passengers per month. It points out the monthly passengers’ average where the highest figure represents the passengers, giving a figure of 67 in July which is more than the corresponding months; the lowest figure of 28 represents the passengers in January is less than the corresponding months which represents low season.

**Power Spectrum Analysis of Daily Passenger Variations**



**Figure 2: Power Spectrum of Daily Passenger Number Variation**

Figure 2 shows the power spectrum density analysis of the daily changes in passenger numbers, where the portion of each frequency is defined. The x-axis is reserved for cycles per day and the y-axis shows the power of each respective cycle. If the dominant peak is located near DC frequency, this may indicate a constant signal contained in the data, because the level of passengers changes seasonally or with time.

At the top of the figure it is also annotated that the average ticket price is X = 53.51 Euros. Further, the main period length equivalent to the frequency with the most energy density is depicted as Y = inf Days. This result hints at the general tendency rather than at cyclic fluctuations with certain fixed frequency in the registered change of passengers.

The visualization aids in emphasizing the prominence of lower frequency, which forms the basis of almost all of the power spectrum. For identification purposes, student ID of 23098031 is added and to meet the instructions and requirements of this assignment. The figure is useful in enhancing the understanding of relative frequency contributions to passenger trends.

# Calculation of X and Y Values

|  |
| --- |
| Average Ticket Price (X): 53.51 Euros  Main Period Length (Y): inf Days |

The values of X and Y together give the kind of insights that are needed when working with the given dataset. The average ticket price symbolized as a variable X is determined as 53.51 Euros as is the statistical average of the ticket prices in the course of the analyzed period. It also provides an essential tool of measuring the airline’s ability to generate passenger revenue based on pricing.

The main period length, symbolized as Y is defined as the average of the periods with the highest amplitude in the power density spectrum. Here, the value is marked as infinity (inf) Days that specify there is a trend value rather than the cyclic Days’ value. This outcome is due to the presence of long-time changes or trends in the airline passenger traffic flow with no particular periodicity.

# Conclusion

Hence, the analysis makes some discoveries in terms of temporal variations on the airline’s passenger statistics by applying Fourier transforms and power spectrum. Fluctuations by month indicate that travel occurs most frequently during the summer season and Fourier series pancreatic offer a good predictor of periodic patterns. This focus is based on the apparent power spectrum showing dominant low-frequency trends, which corresponds to stable and evenly distributed seasonal demand. First, it contributes to the show elements of the revenue and passenger flow, with the key values of average ticket price 53.51 Euros, and not having the fixed main period help to consider the airline’s operation performance.

# References

Müller, M., McFee, B. and Kinnaird, K., 2021. Interactive learning of signal processing through music: making fourier analysis concrete for students. IEEE Signal Processing Magazine, 38(3), pp.73-84. <https://ieeexplore.ieee.org/abstract/document/9418542/?casa_token=nGUQ-93Tk5YAAAAA:WDwCW__EGYxMf9h-igHi298-_4wp70OHB13GGGkS2zE8g6W5pRr4zyepv7ASzidUBlJFtx-l>

**Git Hub :**