# Zvartnots Project

2023-12-10

## Flight Pathways and Beyond: Unveiling Zvartnots Airport Data

### Introduction:

This project explores the intricate dynamics of aviation punctuality using a comprehensive dataset from Zvartnots International Airlines, encompassing both departure and arrival details. The objective is to analyze patterns and derive meaningful insights, particularly focusing on the potential correlation between delays and weather conditions.

### Objectives:

We aim to explore the distribution of Zvartnots International Airlines flights across destinations, assess the airline's performance, and analyze temporal patterns based on days of the week and departure times. The hypotheses to be tested include investigating the impact of geographical locations on delays, differences in Zvartnots International Airlines' punctuality, the temporal influence on delays, and the contribution of weather conditions to flight delays.

### Methodology:

The analysis involves exploratory data analysis, statistical testing, and visualization techniques to uncover trends and relationships within the dataset. This study is valuable for various stakeholders within the aviation industry, including airline operators, airport authorities, and policymakers. These insights can inform strategies for improving overall flight efficiency, enhancing airline performance, and ensuring passenger satisfaction. Additionally, researchers and analysts interested in aviation operations and punctuality trends can find this analysis beneficial for academic and industry-related studies.

### Who Can Use and Benefit from This Project:

Airline Operators: Airlines can leverage the insights gained from this project to optimize their flight schedules, improve punctuality, and enhance overall operational efficiency. By understanding the patterns revealed in the analysis, operators can make informed decisions to minimize delays and improve the overall passenger experience.

Airport Authorities: Airport management teams can use the findings to better allocate resources, streamline processes, and address specific challenges related to punctuality. This knowledge can contribute to the development of more effective operational strategies, ultimately benefiting both airlines and passengers.

Policymakers: Policymakers in the aviation industry can utilize the results of this analysis to inform regulations and policies aimed at improving overall system efficiency. By understanding the factors influencing punctuality, policymakers can contribute to the creation of an environment that fosters better on-time performance.

Passengers: Travelers stand to benefit from this project as it provides insights into potential delays and their causes. Armed with this information, passengers can make more informed decisions about their travel plans, including choosing optimal departure times and considering alternative routes during periods of expected delays.

Academic Researchers: Scholars and researchers specializing in aviation operations, transportation, and data analysis can use the results of this study as a valuable resource for academic purposes. The dataset and methodology employed in this project provide a foundation for further research and analysis within the field.

Industry Analysts: Professionals and analysts focused on the aviation industry can gain valuable market intelligence from the trends identified in this project. Understanding the correlation between delays and various factors, including weather conditions, can assist analysts in making predictions, identifying potential risks, and recommending strategies for sustained operational excellence.

## Datasets:

The main dataset was systematically gathered over a two-month period by daily web scraping of Zvartnots International Airlines' website. This ongoing data collection approach ensured a comprehensive and up-to-date representation of the airline's departure operations over an extended timeframe. By scraping the website every day for two months, we accumulated a rich dataset that captures the dynamics of daily departures, enabling a thorough analysis of patterns and trends in Zvartnots International Airlines' flight punctuality.

```
##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##     last_plot

## The following object is masked from 'package:stats':
##
##     filter

## The following object is masked from 'package:graphics':
##
##     layout

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```
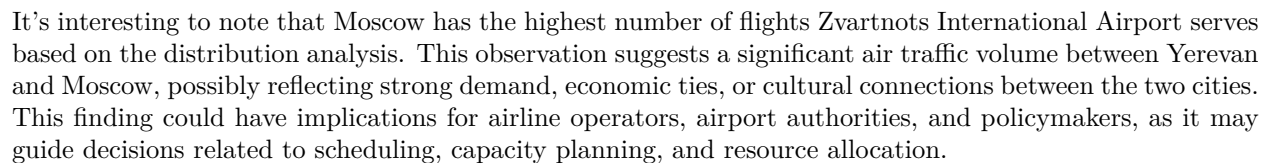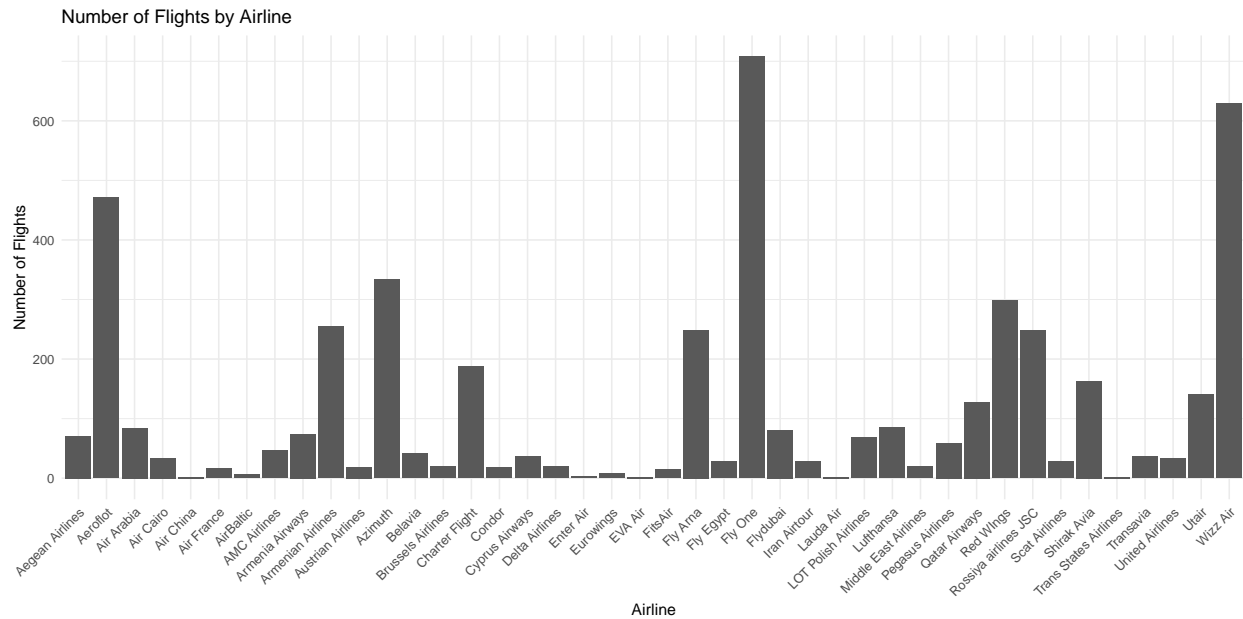
Data Cleaning:

In the data cleaning process, several steps were taken to enhance the quality and utility of the dataset obtained through web scraping of Zvartnots International Airlines' website over a two-month period. Initially, the dataset was merged with a country dataset. A city-country mapping was created to update destination

information accurately. The Scheduled column was converted to datetime, and date and time components were extracted into separate columns. An airline mapping was established to add the 'Airline' column. A binary 'Armenian' column was introduced based on Armenian airlines. The 'Date' column was converted to datetime, and a 'DayOfWeek' column was added. The 'Status' column was updated to categorize flights as 'Ontime' or according to specific delay statuses, and duplicate and irrelevant rows were removed. Additionally, specific dates were excluded, and the final dataset was saved as 'merged_data.csv.' These cleaning steps aim to ensure data accuracy, completeness, and relevance for subsequent analysis.

Now, lets analyze the distribution of flight destinations.



It's interesting to note that Moscow has the highest number of flights Zvartnots International Airport serves based on the distribution analysis. This observation suggests a significant air traffic volume between Yerevan and Moscow, possibly reflecting strong demand, economic ties, or cultural connections between the two cities. This finding could have implications for airline operators, airport authorities, and policymakers, as it may guide decisions related to scheduling, capacity planning, and resource allocation.

Now, let's dive into the interactive map to gain insights into the regions that have the most flight connections with Armenia.

From the interactive map, a noticeable trend emerges with Russia, having the highest number of flights, which is logical as from the previous plot we see that top countries are Russian countries such as Moscow, Sochi, Mineralniye Vody, St. Peterburg. This saturates the color scale and makes it challenging to discern variations among other countries. To gain a more nuanced understanding of flight connections, we have to temporarily exclude Russia from the visualization.

Upon re-examining the map without Russia, we observe a clearer distribution of flights to various countries. It becomes evident that the United Arab Emirates (UAE) is one of the prominent destinations with a substantial number of flights from Zvartnots International Airport. This shift in focus allows us to analyze and highlight regions beyond Russia, providing valuable insights into the diverse flight connections with Armenia.

Number of Flights by Airline



Upon analyzing the distribution of flights among different airlines, it is evident that "Fly One," an Armenian airline, stands out with the highest number of flights.

Upon closer examination of the map, it's evident that Fly One, our Armenian airline, is strategically emphasizing flights to nearby cities. The destinations served by Fly One are marked in royal blue, and you'll notice Yerevan, our capital, standing out in red. This visual representation underscores Fly One's commitment to efficiently serve closer cities in the region.

Now let´s go deeply and understand the performance of other Armenian airlines:

Distribution of Armenian and International Flights
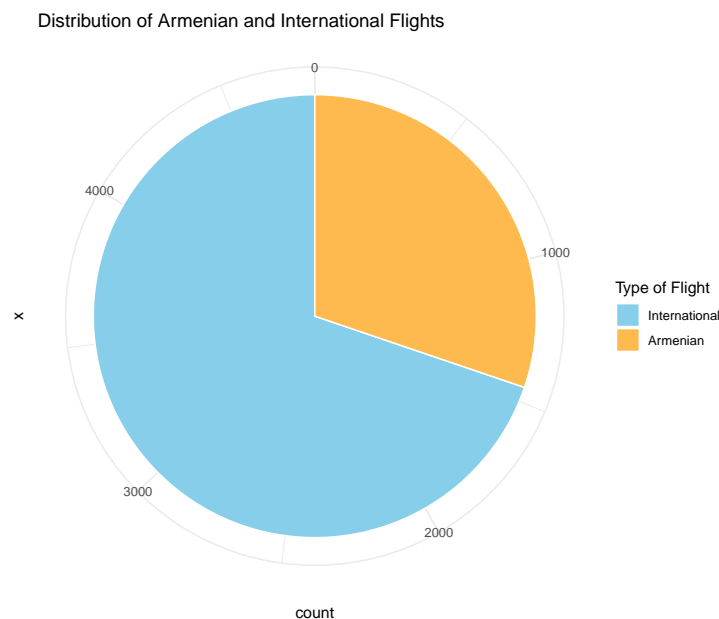


Examining the distribution of flights in our dataset, it's logical to observe that the majority fall into two categories: Armenian and international. Given that we currently have representation from only five Armenian airlines, this distinction becomes even more apparent. The accompanying pie chart visualizes this breakdown, with the slices representing the proportion of international and Armenian flights. It's a concise yet insightful glimpse into the overall composition of flights in our data, reflecting the influence of the limited number of Armenian airlines.

```
## Warning: 'stat(count)' was deprecated in ggplot2 3.4.0.
## i Please use 'after_stat(count)' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Distribution of Flights Among Armenian Airlines
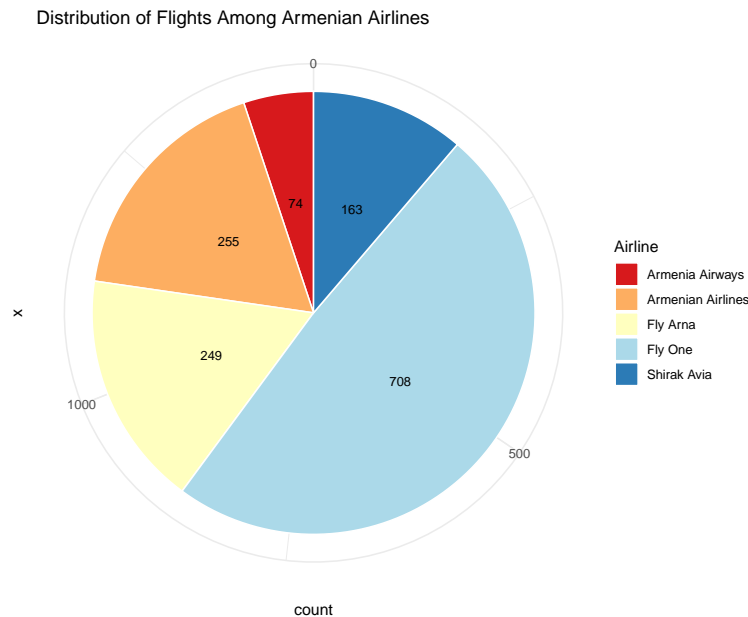


In the breakdown of flights among Armenian airlines, Fly One takes the lead with the majority of flights, followed by Armenian Airlines. Armenia Airways, on the other hand, represents a smaller share. This insight, gleaned from the accompanying pie chart, provides a clear picture of the distribution of flights among these Armenian carriers.

Now, let's delve into a
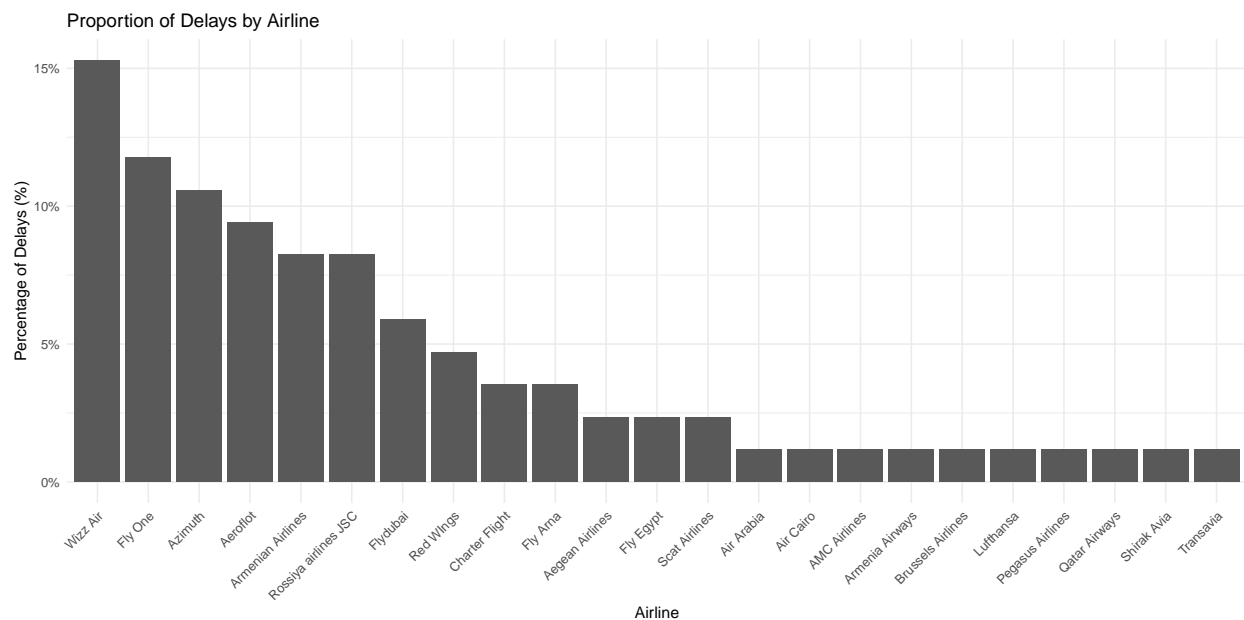


After analyzing the proportion of delays for each airline, it is apparent that Wizz Air, FlyOne, and Azimuth are among the airlines experiencing the highest delays. This insight sheds light on the specific carriers facing

challenges in maintaining on-time performance and prompts further investigation into the factors contributing to delays within these airlines. Addressing these issues can lead to improved operational efficiency and enhanced passenger satisfaction.

Proportion of Delayed Flights by Destination
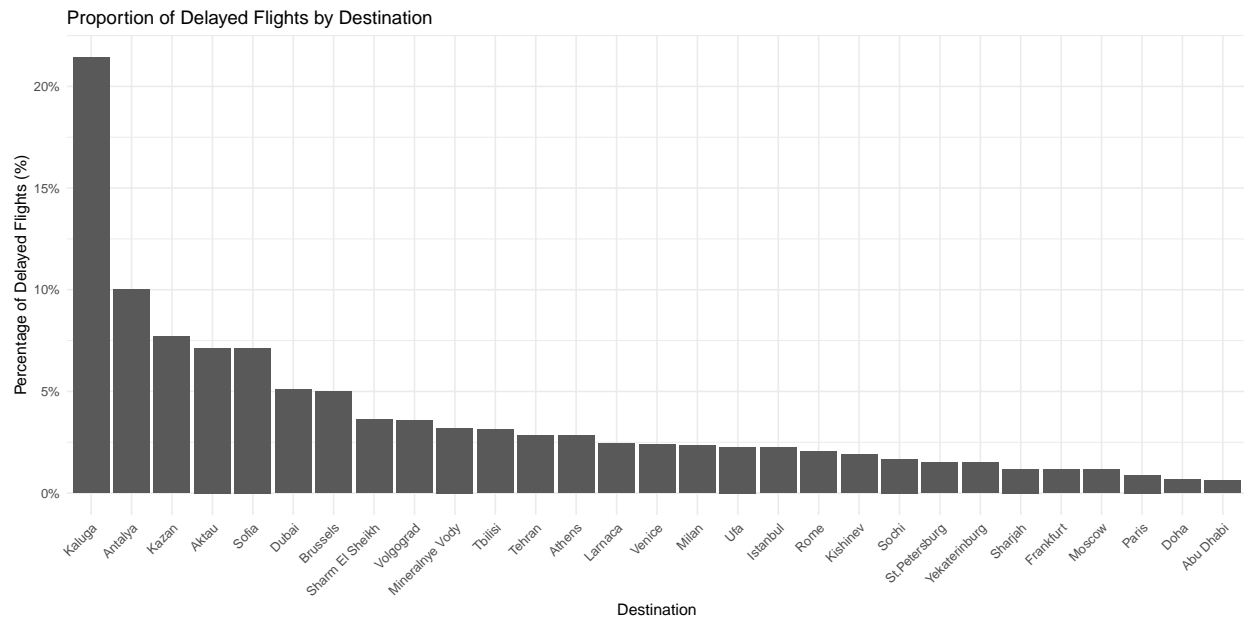


After analyzing the proportion of delayed flights for each destination, it is evident that Kaluga, Antalya, and Kazan are among the destinations experiencing the highest delays.

Now, lets analyze peak hours of flight activity:

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Flights Time Series by Hour
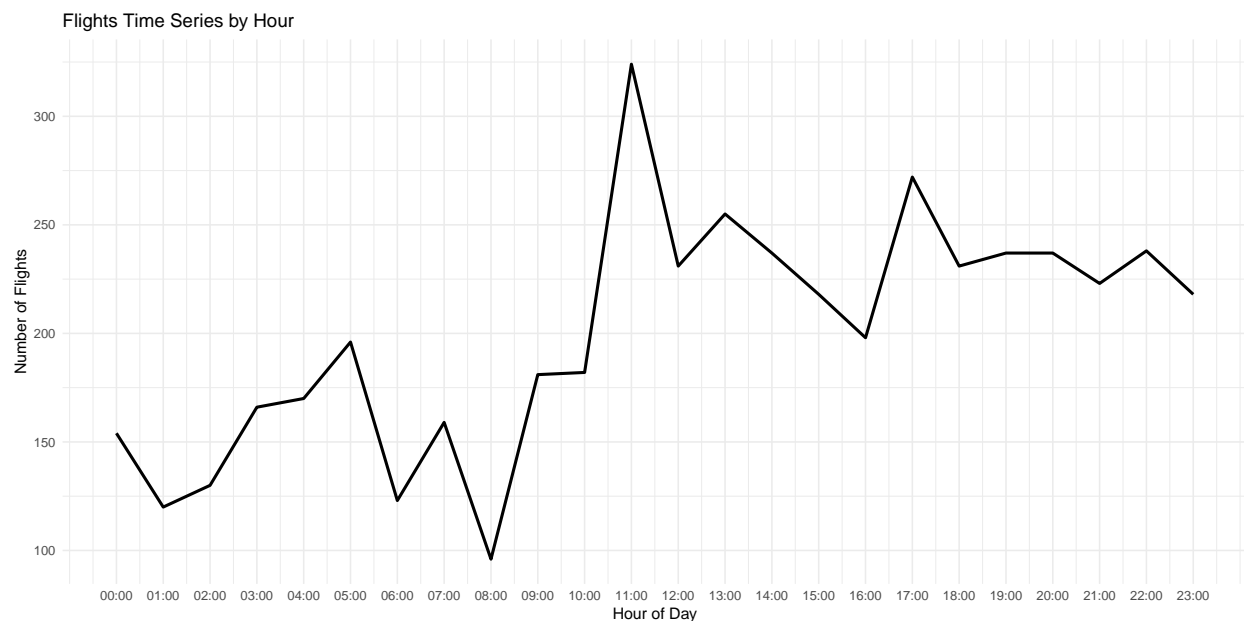
From the observed plot, it is evident that the peak hour for flights at Zvartnots International Airport occurs at 11:00, showcasing the highest concentration of scheduled departures or arrivals during this specific time frame. Conversely, the hour with the least flight activity is identified as 08:00, indicating a comparatively lower number of flights during this period. This information provides valuable insights into the temporal distribution of flight operations, aiding in resource management and logistical planning for both the airport and airline operators.

Let´s go deeper and compare this with the delayed flights.
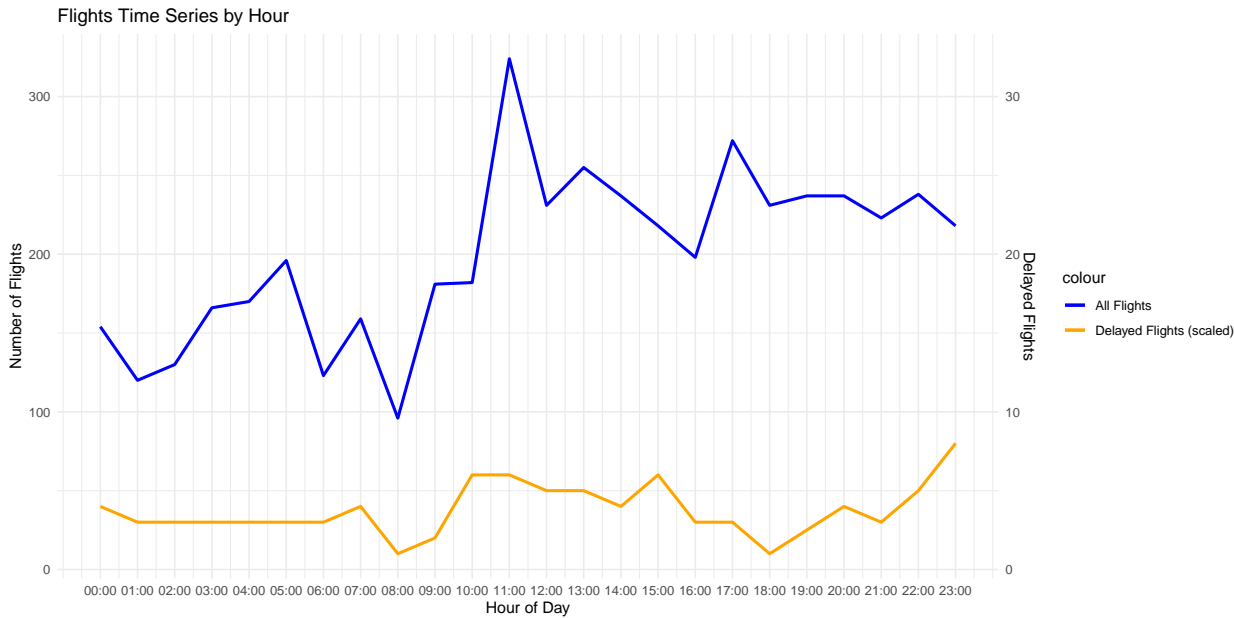


The correlation between flight volume and delays is a well-documented phenomenon. During peak hours, when there is a higher concentration of flights, various factors such as air traffic congestion, limited runway availability, and increased passenger activity can contribute to delays. These challenges are often more pronounced during busy periods, leading to a higher likelihood of delays.

Conversely, during off-peak hours when the number of flights is reduced, the overall demand on airport infrastructure decreases, resulting in a more streamlined and efficient flow of air traffic. This can lead to fewer delays as the system has more capacity to handle the existing demand.

Understanding the relationship between flight volume and delays is crucial for aviation stakeholders to implement effective strategies to manage and mitigate delays during busy periods. It underscores the importance of optimizing airport and airspace capacity to accommodate the varying demands of different times of the day.

Now, let´s analyze how are the flights distributed among weekdays.

Flight Counts by Day of the Week
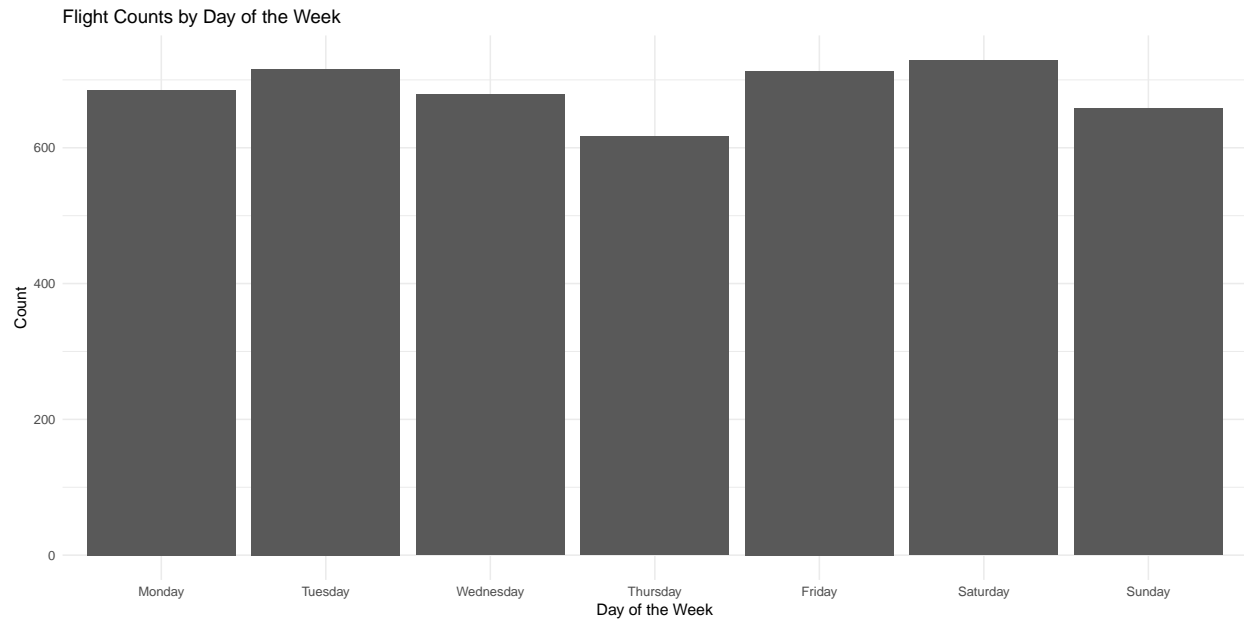


From the analysis of flight distribution across weekdays, it is evident that Saturday experiences the highest volume of flights compared to other days of the week. This observation provides valuable insights into the temporal patterns of flight activity, allowing for a better understanding of the busiest and comparatively quieter days at Zvartnots International Airport.

Let´s see the week performance for the airlines

```
## 'summarise()' has grouped output by 'DayOfWeek'. You can override using the
## '.groups' argument.
## Selecting by TotalFlights
```

Number of Flights by DayOfWeek and Airline
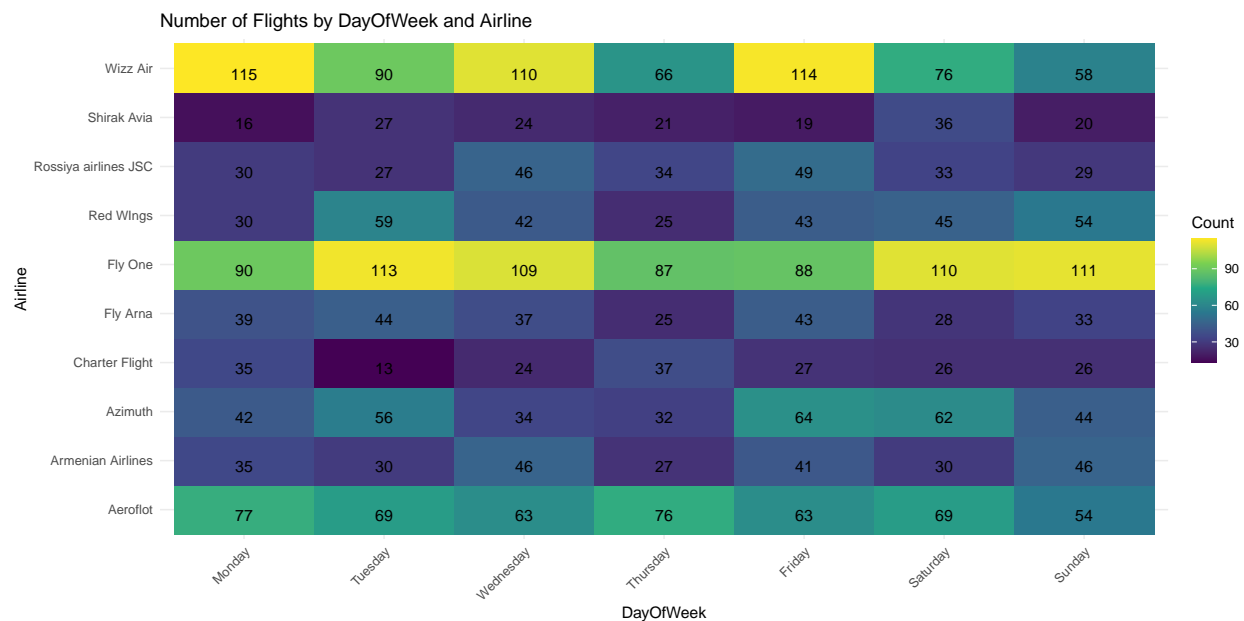


Based on the provided heatmap, we observe distinct patterns in the distribution of flights for different airlines across weekdays. For instance, Fly One tends to complete the majority of its flights on Tuesday, while Wizz Air shows a more evenly distributed pattern with peaks on Monday and Friday. This visualization

effectively highlights the weekday preferences of each airline, providing insights into their scheduling practices and operational trends. Now, let´s look at the delays


Distribution of Delays by Day of the Week
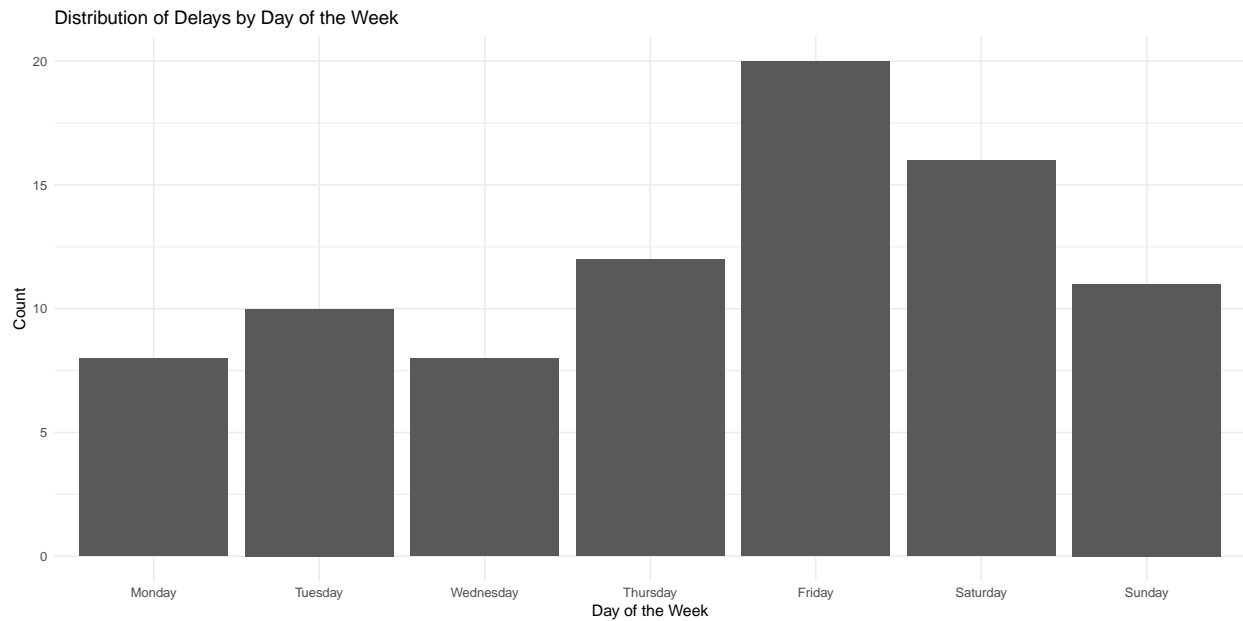
We see that the most of delays are being on fridays, bur previously we saw that the most flights were on saturdays so we can conclude that the delayed flights of fridays are being scheduled for saturdays. The analysis of the heatmap reveals an interesting observation: while the majority of flights occur on Saturdays, a higher concentration of delays is observed on Fridays. This pattern suggests a potential scheduling strategy where flights with delays on Fridays may be rescheduled for Saturdays. This correlation between high flight volumes on Saturdays and increased delays on Fridays could indicate operational decisions made by airlines to manage and accommodate delayed flights by rescheduling them to the following day.

Now, let´s analyze the distribution of flights during this time period and do time series analyses.


Time Series of Number of Flights
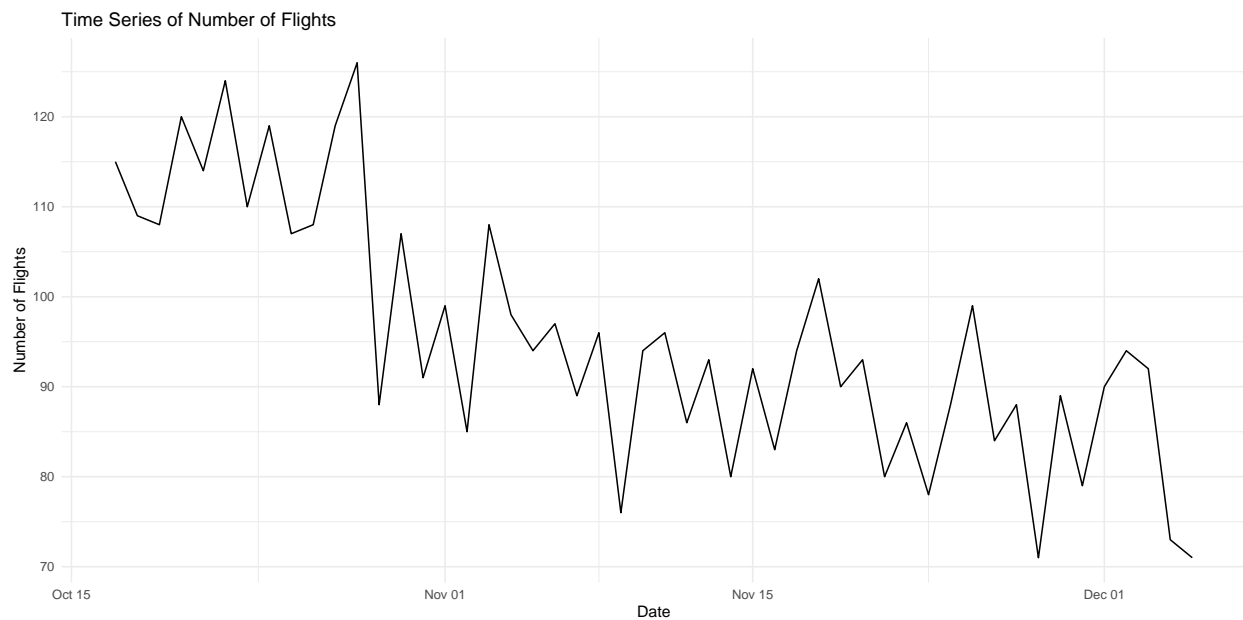
Based on the time series plot, it's evident that there is an overall decreasing pattern, suggesting a potential decline in the number of flights over time. Additionally, the observed seasonality in the data aligns with the logical scheduling of flights on specific weekdays with fixed days and times for every week. This seasonality is

a common characteristic of airline schedules, reflecting the structured and recurring nature of flight planning. The insights derived from the time series plot contribute to a better understanding of the temporal trends in flight activity, aiding in strategic decision-making for airline operations and resource allocation.

If you want to see more analyses concerning this, you can check the appendix part and see our future and deeper analyses and predictions.

At the inception of our project, we embarked on a comprehensive exploration of the potential correlation between weather conditions and flight delays. Collecting data diligently for a span of two months, we aimed to analyze any discernible patterns. However, upon meticulous examination, we found that the weather conditions during this period exhibited no substantial variations that could significantly impact flight operations. Given the absence of noteworthy changes, we made an informed decision to exclude the weather-related findings from our current project. It is essential to note that this decision was made to maintain the focus on the primary objectives and insights derived from the extensive analysis of aviation punctuality. Despite this, we wish to highlight that our team remains committed to continuing this project beyond the course duration. Future iterations will include a thorough investigation into the correlations between weather conditions and flight delays, enhancing the depth and breadth of our analysis. While the current project primarily focuses on aviation punctuality, we acknowledge the substantial effort invested in collecting and analyzing weather data. As a testament to our commitment, we present key findings from our weather dataset, providing a glimpse into the extensive work undertaken in this aspect of the project.

The three plots provide a visual representation of weather conditions over five weeks in late 2023. The first graph tracks cloud cover, with a scale from 0 for clear skies to 1 for complete coverage, showing frequent changes. The second displays wind gusts, with peaks indicating strong wind events. The third shows wind speed, highlighting more consistent wind conditions with occasional increases. These plots illustrate variations in cloudiness, wind bursts, and general wind flow, important for weather forecasting and related activities. Together, they detail the changing atmospheric conditions observed during the period these plots offer valuable insights into weather dynamics, critical for sectors such as agriculture, aviation, and outdoor event planning.

This wind rose represents the distribution of wind speed and direction. Each segment's length indicates how often the wind blows from a particular direction, with longer segments representing a more frequent occurrence. The colors likely represent different wind speed ranges, with darker shades such as red perhaps denoting higher speeds and lighter shades like orange and gray indicating lower speeds. Directions are marked around the circle (N for North, E for East, S for South, W for West, and their intermediates). A slider at the top suggests that the user can select a range of wind speeds to display, enhancing the plot's interactivity.

We use this plot to understand in general the distribution of weather and weather conditions.

Reference Links:

https://www.zvartnots.aero    https://plotly.com/r/    https://r-graph-gallery.com    https://shiny.posit.com https://www.rstudio.com/products/shiny/