OneSky: Dynamic Pricing

Presenters: Payton Boliek, Paige Rasmussen, Mariah Marin, Allison Acosta
Project Advisors: Michael Campobasso, Dr. Mihhail Berezovski

OneSky

EMBRY-RIDDLE
Aeronautical University™

Introduction

Dynamic pricing tool algorithms have been widely used for increasing a company's revenue and customer satisfaction while maintaining competitiveness in many industries. In the private aviation industry, the objective of the dynamic pricing tool algorithm is to calculate flight prices for a customer based on a multitude of factors including: the type of private jet, the arrival and departure dates and locations, and the time proximity to a special event or holiday. For our project, a dynamic pricing tool is being created in a collaboration with OneSky, particularly OneSky Network, in order to maximize the possible revenue for every day of the year. OneSky is a combined resource company that manages several private jet charters including PrivateFly, SentientJet, FlexJet, and so on. The main goal of the dynamic pricing tool for OneSky is to maximize the profit made by these private charters.
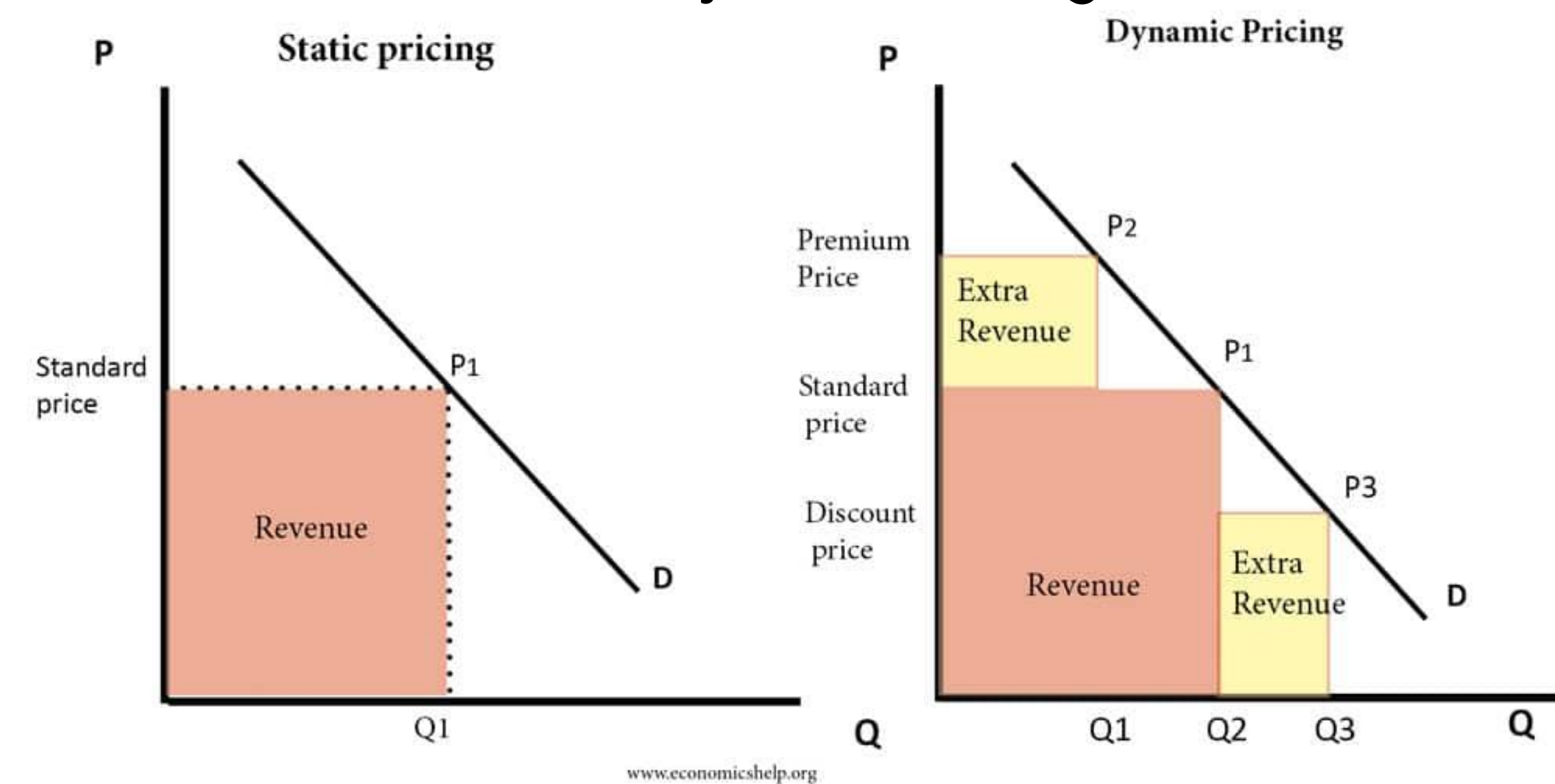
Static vs. Dynamic Pricing

Figure 1: The two graphs shown above represent the difference between static and dynamic pricing. Static pricing will not yield as much extra revenue as dynamic pricing. This is due to the fact that in dynamic pricing, the price of goods can be varied. Therefore, using dynamic pricing allows for a greater area underneath the demand curve to be accessed, which allows for extra revenue to be gained.

Description of Data

The original data provided was 333MB in size and contained about 4.8 million unique flight data points, so there were reoccurring complications in trying to open all of the data at once. In order to address the problem, the data was divided into 4 quarters for each year for the three consecutive years: 2017, 2018 and 2019. Thus, a separate file was created for each quarter in order for it to be manageable by the processing systems. Afterwards, any outliers from the data were removed since they provide no insight into trends (see Table 1).

Data Statistics

File Size	333MB
Number of Unique Flight Data Points	≈4.8 million
Number of Flights in 2017	1,684,207
Number of Flights in 2018	1,694,462
Number of Flights in 2019	1,654,914
Total Number of Flights	5,033,583
Number of Unique Airports	6176

Table 1: The information above describes the flight data provided by OneSky quantitatively. The data statistics provided do not include any outliers, which are airports that contained 2 or less flights throughout the 3-year time period, or any flights with null departure or arrival locations. Only the data that provided useful trend information was accounted for.

Number of Departures per Day From 2017 - 2019

$$y = -1E\text{-}13x^6 + 2E\text{-}08x^5 - 0.0027x^4 + 156.16x^3 - 5E\text{+}06x^2 + 9E\text{+}10x - 6E\text{+}14$$
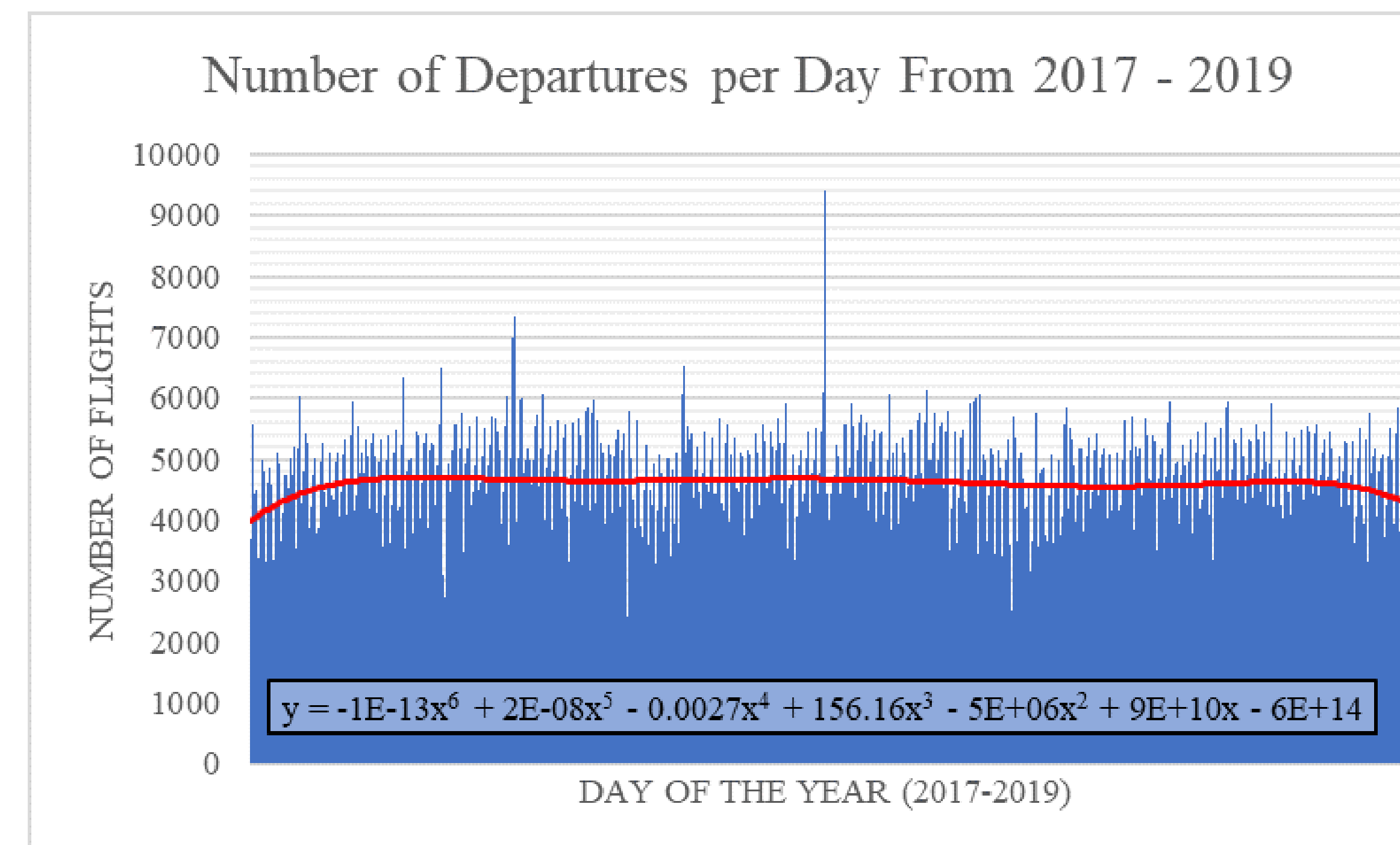
Figure 2: The graph above contains the number of flights for each day of the year from 2017 to 2019. Within the data shown, outliers (airports that contained 2 or less flights throughout the 3-year time period) have been removed, so the dynamic pricing tool will be based upon this newly shown data. The red dotted line shows the best fit curve, which is used to obtain a baseline price. The many peaks and valleys seen above are to be normalized and flattened using our algorithm.

Main Strategy

In order to create a useful dynamic pricing tool, we will need to calculate and assign coefficients or multipliers to every possible aspect affecting flight demand, including day of the week, month, year, and events. These coefficients will then be kept in a database, called the event calendar. The steps of our strategy to achieve this are as follows:

- Normalize the flight data in order to obtain a flat curve, from which we can then calculate the values of multipliers (for more detail of this process, see section Algorithm Description)
- Cross examine the leftover peaks in flight data after normalization with events to account for any deviations in demand from any baseline trends
- List the found coefficients in the event calendar, beginning with global and regional events and narrowing it down to local events

Currently, days and weeks have been normalized, so next the coefficients of the months will be calculated (see Fig. 3). Also, the first rendition of a heat map has been generated, so now it must be applied to the entire data set (see Fig. 4).
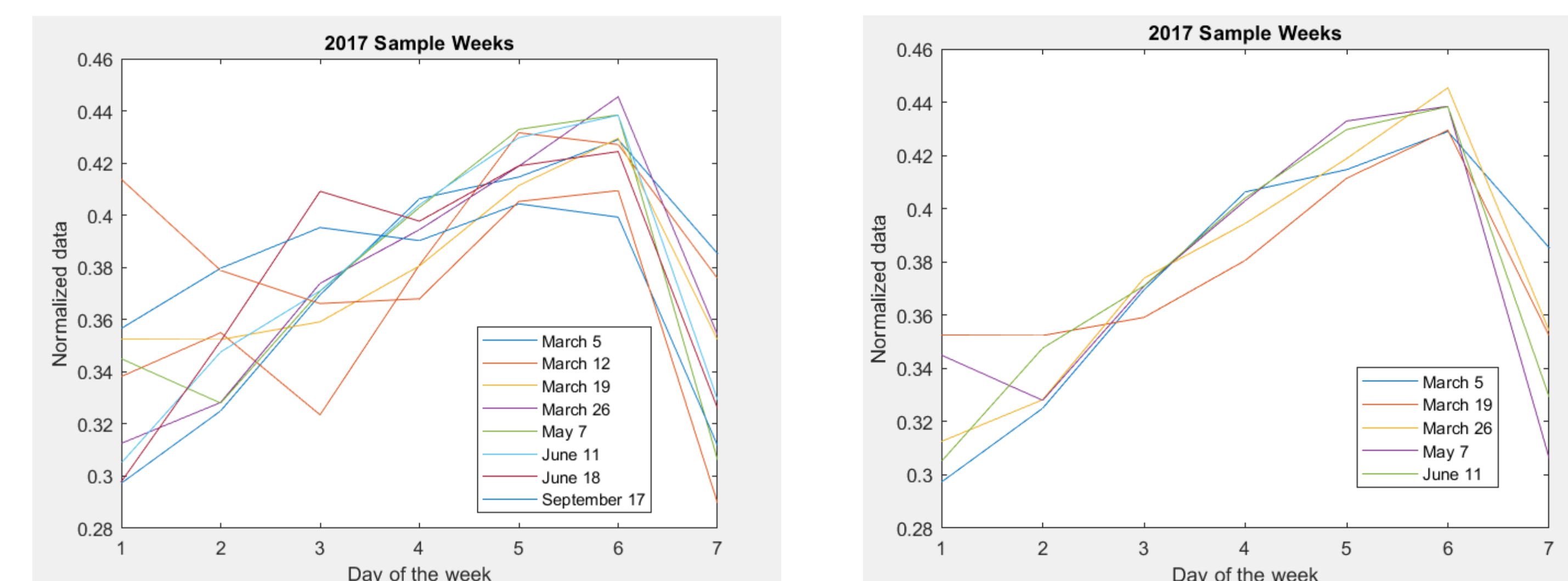

Figure 3: (Left) The graph above shows multiple weeks of 2017, which were chosen at random. Stacked on top of one another, is it clear that there are variations within days of the week that will affect the average value for a multiplier that will be used for normalization. Some of the weeks shown are also being affected by event traffic. (Right) This graph shows more carefully chosen weeks, which are weeks that are not as influenced by event traffic (in other words, these are more "normal" weeks). It is from these normal weeks that will allow for more static numbers for multipliers since they all follow approximately the same trend.

Algorithm Description

The purpose of our algorithm is to normalize the data given into a flat curve. By doing this, it will be possible to pick out variations in flight demand caused by the corresponding day, week, month, or event and assign multipliers. The steps in order to accomplish this are as follows:

- For each "average" week, identify an average coefficient of demand for each day,, then divide each week by their respective coefficients to normalize each week (for example, see Fig. 3)
- Since the number of flights per month will vary, coefficients of demand will then be found for each month, which can then be divided out to normalize the months
- Yearly coefficients will then be found based off of the three years of data provided; then the coefficients can be divided out to normalize the entire data set
- After normalization, the only demand peaks left will be those corresponding to specific events, so the events can be identified and their respective coefficients can be found
- Dividing the event coefficients out will produce a completely flat data curve, from which the multipliers can be reapplied in order to produce the original data (see Fig. 2)

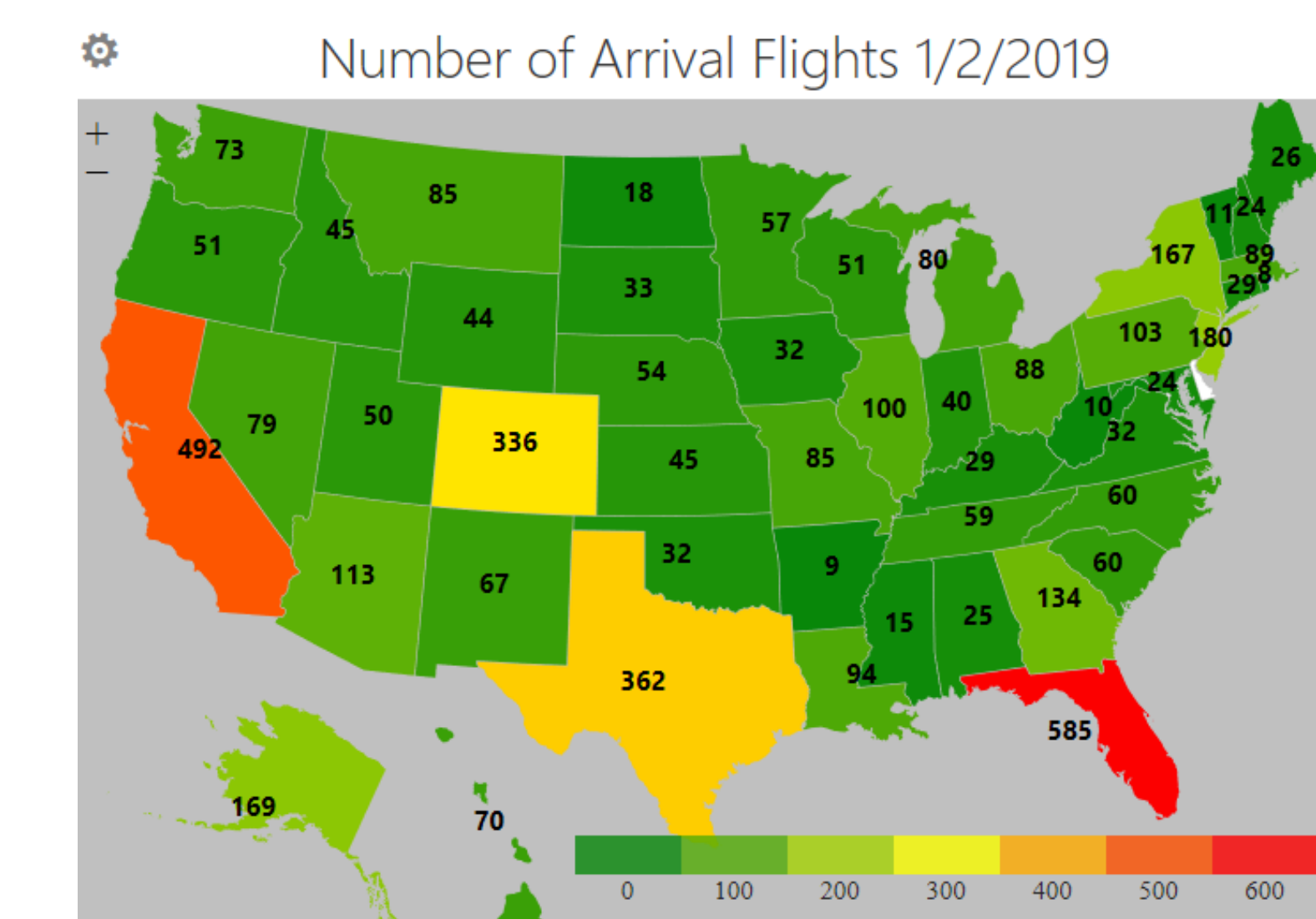## Number of Arrival Flights 1/2/2019

Figure 4: The image above contains the number of arrivals on January 2, 2019, which is normally associated with one particular event: New Years Day. The flight data shows trends towards a few particular regions,, which are California, Colorado, Texas, and Florida, with some smaller trending towards the Northeastern states. These trends are likely appearing due to the fact that these states have the largest populations., so many people are travelling home or to their work locations after the holiday break. These states will likely remain hotspots throughout the year due to their high population levels.

Findings

From normalizing the weeks, the multipliers for each day of the week can be calculated. So far, only the multiplier for Fridays has been found, which is approximately 0.42±0.02. The coefficients for the rest of the days are still being determined. Also, from the heat map (Fig. 4), it is possible to see flight trends towards particular locations in association with certain events. In the figure shown, it is clear that the states with the highest populations also have the highest amount of flight traffic during a major holiday (New Years Day), which is a trend that will likely reoccur with country-wide events throughout the year. Following the same processes as we have thus far, we will soon have the rest of the coefficients for days of the week, months, years, and events and will be able to complete the dynamic pricing tool.

Resources

"Helping to Simplify Economics." Economics Help. 2020. Web. https://www.economicshelp.org/?attachment_id/page/7/&attachment_id
Source of data: Campobasso, M. OneSky Network. 2020.
"About OneSky." OneSky Network. 2020. Web. http://www.oneskynetwork.com/