

Visualizing Air Travel Delays^{*}

Aishwarya Krishnakumar^{1,2}, Weidi Lou^{1,3}, and Zixin Zhang^{1,4}

¹ University of Southern California, Los Angeles CA 90007, USA

² kris184@usc.edu

³ wlou@usc.edu

⁴ zzhang09@usc.edu

Abstract. This paper is written in correspondence with our demo, video, and presentation for the DSCI-554 Fall 2021 class. The project is about visualizing air travel delays.

Keywords: Airline · Airport · Departure Delay · Carriers · Air Traffic

1 Introduction

1.1 Motivation

The United States of America has the largest airline system worldwide [4]. In order to be optimal, airports and airlines aim to complete the maximum number of flights possible while simultaneously adhering to safety measures. In order to do this, flights at times may have to be delayed.

We chose to do our final project on air travel since it's widely used by people around the country, especially us college students. We want to create a visualization to present data about flight delays, primarily domestic. By creating a dashboard, users will be able to understand the reasons why flights are delayed, which airports around the country deal with the most delays, which travel dates see the most delays, and which airlines have the most delays. These pieces of information can be understood better through graphs and maps, which is why we think having a strong visualization is important.

1.2 Scope

The end user for our project includes current and future passengers, airport employees, and airline employees who are interested in learning about flight delays. Our project was inspired by many of our own travel journeys and our own experiences with delayed flights.

This dashboard is useful because it not only informs users of the reasons behind flight delays, but also allows them to interact with the information. This can help them with planning their own flight trip.

^{*} Student project for USC's DSCI-554 Fall 2021.

1.3 Our Project Contributions Compared to Existing Work

Our project was inspired by works done by the Federal Aviation Administration and by the flight tracking application, FlightAware. Although our contributions to this project are inspired by existing work such as FlightAware, we have made it unique as we focus our information on delayed flights. Our contributions are discussed more in Section 2.

1.4 Introduction of Sections

Throughout this paper, we will discuss the many parts to this project. In Section 2, we present related visualization work. In Section 3, we present the dataset used. In Section 4, we explain how we designed and the process we used to build the system and in Section 5, we showcase the main aspects of the website we have built. Finally, we conclude the paper with Section 6.

2 Related Work

2.1 Similar Work in Same Topic

Similar works of the same topic have been done by the Federal Aviation Administration [6] and by the flight tracking application, FlightAware [1]. The Federal Aviation Administration has a website which contains some visualization about general airport conditions. For those specific flight requests, it provides the link to airline.

FlightAware has a real-time worldwide flight traffic visualization on the map. It also includes some other visualizations about data trend such as airplane velocity and elapsed flight time. FlightAware’s map, however, only shows the flight route and does not provide any information about delayed flight routes.

2.2 Similar Work in Other Areas

Similar works have been done in other areas by the Metropolitan Transportation Authority [2] and the California Highway Patrol [5]. The Metropolitan Transportation Authority has an online dashboard that tracks subway performance in the state of New York. This is similar to our project because it provides a bar chart of performance over time.

The California Highway Patrol maintains a dashboard to track where they have conducted stops throughout the state of California. This work is similar to our project in the sense that both include interactive maps with tooltips.

2.3 Our Project in Relation to the Similar Work

Similar works have been done by the Federal Aviation Administration and by the flight tracking application, FlightAware. We decided to model our map based on the visualizations in FlightAware. However, the map in FlightAware only shows

the flight route. It does not provide any information about delayed flight routes. Therefore, in our project, we decided to create a map that shows delayed flight routes.

The Federal Aviation Administration has a website which contains some visualization about general airport conditions. For those specific flight requests, it provides the link to airline. In our project, we provide additional information about a certain delayed flight when the route is hovered over. Although our contributions to this project are inspired by existing work such as FlightAware, we have made it unique as we focus our information on delayed flights.

3 Data

3.1 Explanation of Dataset

Our project's dataset is from the United States Bureau of Transportation [7]. This dataset includes flight date, airline identification codes, individual flight numbers, origin airport, destination airport, departure time, arrival time, and delay times. The dataset categorizes the causes of delays in five expansive categories: Air Carrier, Weather, National Aviation System (NAS), Security, and Late-Arriving Aircraft. These categories encompass all possible reasons for a flight to be delayed.

Another dataset we used was from the Humanitarian Data Exchange [3]. This data had the local code, longitude, and latitude of the airports. This dataset had 27,917 rows of data. This dataset was necessary since our initial dataset did not have the physical locations of the airports.

3.2 Explanation of Pre-Processing

The data on the United States Bureau of Transportation's website allowed us to filter based on carrier, airport, and date. We ended up choosing all carriers, all major airports, and the dates of August 2021. We then downloaded the data as a csv file. In total, we had 81,779 rows of data.

After doing that, we realized that the dataset did not include the physical locations of the airports. This led us to finding the second dataset from the Humanitarian Data Exchange. We connected both datasets by the common variable of airport code.

4 Approach

4.1 Design Process

In order to create a thoughtful design that adequately portrayed our desired story, we decided to meet through Zoom to first decide the goal and overall story of our project. Then, we had to decide what graphs would best fit each part of the story. Afterwards, we first sketched out a rough layout to decide the topics. More detail about the background behind each graph is in the next subsection.

4.2 Design Considerations

We knew that we wanted to focus our story on air travel, specifically domestic delays. However, we also wanted the users to be able to learn about the locations with most delays, dates in August 2021 with most delays, and carriers that had most delays. To best represent the information, we decided to create an interactive line graph to show carrier delays, a donut chart to show the reasons why flights are delayed, USA maps to show the delayed flight routes, and a treemap to show location by state of most delays. In the beginning of the project, we thought about doing an animated treemap. However, we did not find true meaning to that and instead decided to create a static treemap with small multiples so that comparisons can be made by the user.

5 System

5.1 Technologies and Methods

For the demo, we utilized vue.js as the framework of the web application. BootstrapVue was also used for making the application responsive and mobile friendly. For plotting the beautiful curve of the flight path, we applied turf.js.

In terms of the paper, we used overleaf.com, and for the presentation, we used Sozi and Inkscape. The video was added to YouTube and created through screen recording tools.

For project management, our team decided to use Zoom for weekly meetings. We also were in communication through Slack.

5.2 Highlights of Demo and Website

Like we briefly stated earlier in this paper, our demo consists of maps, a donut chart, a line chart, and a tree map. To make the website more clear and user-friendly, we added instruction and explanation buttons for each chart or map.

The donut chart represents the percentage of each type of delay a flight can experience. This chart is interactive since the user can hover over each section to see the information. The line chart is where we compare the airline carriers. Each color represents a different carrier. This is also interactive with animated line drawing. The tree map's goal is to show the users how each state differs in the number of departure delays at its airports. Data from both weekend and weekday are used to generate tree maps, so that the users can compare the difference.

We have two maps to show the details of each flight. One is the D3 map. After selecting the date, the user can click on the state to view the delayed flights departing from that state. When the user hovers their mouse over the path, the delay information about the flight is shown. The other map is the responsive Mapbox map. This map asks users to input the date and the airport code which they want to learn about. Then, the map plots the paths of those matching flights. By clicking on one single path, the delay information is shown on the

page. We have two maps with similar function because we want the map to be friendly to both PC users and mobile users.

We also optimized user interaction on the maps. On the D3 map, the delayed flights are shown in various colors. Different colors represent different causes of delay. In the Mapbox map, the paths of delayed flights are shown in a beautiful curve. Users can see detailed information of each flight by clicking on the path.

6 Conclusion

6.1 Contributions

We decided to split the work by section while also allowing each other to put a little time in each other's tasks. In order to create a design for the dashboard, we all met through Zoom and discussed the story that we wanted to show. Then, we decided what graphs would best explain the story. We created a rough sketch of how we wanted the dashboard to look like.

In terms of individual tasks, Aishwarya focused on creating the powerpoints and writing the paper. She also created an interactive donut chart. Zixin focused her efforts on the line chart and the tree map. She also gave input for the powerpoints. Weidi created the d3 map and the mapbox map and contributed in writing about the maps. The video and in-class presentation was a group effort.

6.2 Potential Improvements and Changes

To expand on this project, we would add more dates instead of just one month. In addition, we would expand to international flights as well. In order to accomplish this, we would need to expand our both our flight dataset and airport location dataset tremendously.

In terms of the functionality, we would like to allow users to save the information of a certain path and be able to compare the saved flight to another flight. Another potential improvement is allowing the user to be able to look at nearby airports to compare the flight delay times to the same destination.

References

1. Flightaware, last accessed Nov 2021
2. Authority, M.T.: Subway dashboard performance (2021), last accessed Nov 2021
3. Exchange, H.D.: Airport data - airports in the united states of america (2021), last accessed Nov 2021
4. Mazareanu, E.: Passenger airlines in the u.s. - statistics facts (2021)
5. Patrol, C.H.: 2019 racial and identity profiling act (ripa) (2019), last accessed Nov 2021
6. of Transportation, U.S.D.: Federal aviation administration - national airspace system status (2021), last accessed Nov 2021
7. of Transportation, U.S.N.B.: Flight data - on-time : Marketing carrier on-time performance (beginning january 2018) (2021), last accessed Nov 2021