

MSc Business Analytics
**MIS41040 Business Decision Support
System**



**UCD Michael Smurfit Graduate Business
School**

Tableau Assignment

Team 27

Lecturer: Dr. Peter Keenan

Name of Student	Student ID	Email ID
Adithya Vivek	22200615	adithya.vivek@ucdconnect.ie
Anish M Nargund	21200076	anish.nargund@ucdconnect.ie
Sudeep Tandon	22200426	sudeep.tandon@ucdconnect.ie

Contents

Introduction.....	1
Data Analysis and Preparation.....	1
Dashboard Interpretation	2
Overview by State	4
Overview by Airports	5
Overview by Airlines	6

Introduction

Flight delays are a prevalent problem in the aviation industry, causing several issues for passengers, airlines, and airports. As a result, a decision support system (DSS) for flight delays may give airlines and other stakeholders information and analysis to assist them in making informed decisions. Real-time data integration, in which the DSS acquires and combines real-time data from diverse sources such as weather reports, air traffic control, and airline systems, is one of the crucial characteristics of a flight delay DSS. In addition, by examining historical data and present situations, predictive modelling and scenario analysis can assist airlines in assessing the possible impact of various scenarios, such as rerouting or cancelling flights, on their overall operations.

Data visualization is crucial in a DSS for flight delays. The visualization displays complex data in an understandable and straightforward style, allowing decision-makers to absorb essential insights easily. It can also assist in identifying patterns and trends that may not be obvious in raw data, such as the influence of weather on flight delays. Visualization assists in monitoring real-time performance, such as on-time performance, and identifying areas for improvement. Identifying outliers, such as planes having lengthier delays than typical, can assist decision-makers in responding rapidly to these events and minimizing their impact.

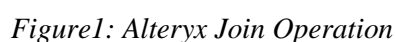
Dashboard and reporting are two powerful ways of visualizing data in DSS. Dashboard and reporting features enable airlines, passengers, and other stakeholders to monitor flight status and measure essential performance metrics such as on-time performance and customer satisfaction, to name a few, to make well-informed choices. To make intelligent choices, the DSS would need feature-rich data. For example, flight, weather, air traffic, maintenance, passenger, and other critical data are essential for the DSS to help with decision-making.

Data Analysis and Preparation

The data was collected from the TranStats data library's Marketing Carrier On-Time Performance for the Year 2018 data table of the "On-Time" database. Because the data is enormous, we analyze July month's data, a subset of the raw data. There were 119 columns in all. After studying it, we used the most relevant and beneficial columns for decision-making. For the date and time information, we use the columns Quarter, Month, DayofMonth, DayOfWeek, and FlightDate, followed by the columns Operating_Airline, Tail_Number, Flight_Number_Operating_Airline, OriginCityMarketID, Origin, OriginStateName, DestCityMarketID, Dest, and DestStateName for flight information. Next, columns offer information on probable causes of delays such as CarrierDelay, WeatherDelay, NASDelay, SecurityDelay, LateAircraftDelay, DivAirportLandings, DivReachedDest, and DivActualElapsedTime. Finally, we also considered additional fields like DepTime, DepDelay, TaxiOut, TaxiIn, ArrTime, ArrDelay, Cancelled, CancellationCode, Diverted, ActualElapsedTime, AirTime, Distance, and DistanceGroup to complete the analysis. With this, we could bring the number of columns to 35.

Since the raw data is given per month, we needed to perform some rudimentary data processing. We used Python programming language for preparing the data by using its pandas library. First, we imported all the CSV files for each month given in the raw data for the year 2018 and clubbed all the files into one master file. Next, we drop the unwanted columns from the file

We were also given additional CSV files, namely *Airlines.csv* and *Airports.csv*. These files contain some critical fields like the latitude and longitude of the airports and the airline names corresponding to their codes. There were a total of 52 additional airports that were found in the raw data but were missing in the *airports.csv* file. To ensure the completeness of the data we populated the all columns for the missing values of these airports. We use Alteryx software, widely used for data processing, to integrate these files. We form two joins, one for each file, to add the columns by entering the related field like *IATA_CODE* to *Dest* and *Origin* and airline *Code* to perform the join operating to bring the corresponding fields respectively. We then carry out some fact checks like number records and columns. We then extract the data into a Tableau Extract format called ".hyper", which is easier for the Tableau software to read since the size is small compared to the CSV file making it effortless to process.

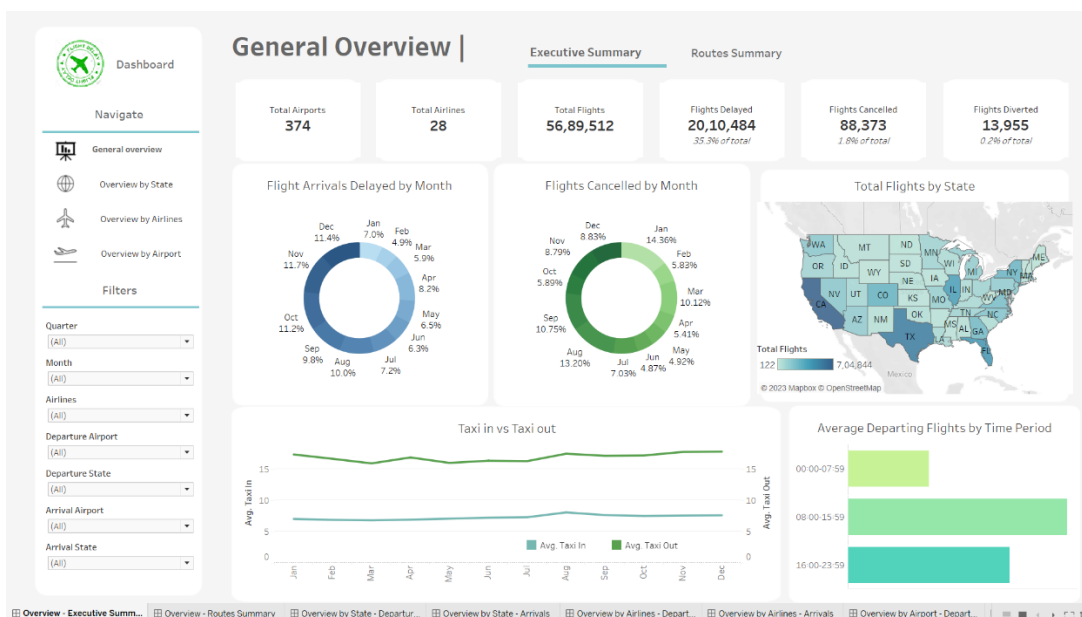
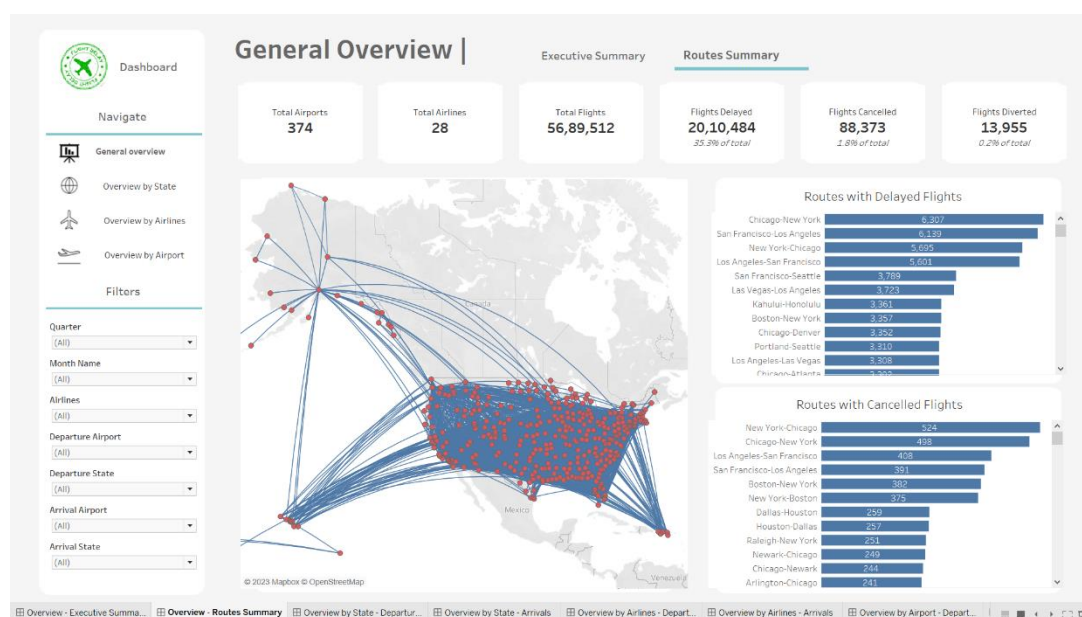


The home screen of the dashboard is a general overview of flight delays. It is a DSS tool that visually represents flight delay data. On the widgets at the top of the dashboard, essential metrics like the total number of flights, airlines, airports, delayed flights, cancelled flights, and diverted flights are present under the executive summary tab. We also see the proportion of total flights in percentage. We also see a density-based map of the United States, which shows the number of flights in each state. Finally, we also see two pie charts that show the percentage of flights that are delayed or cancelled each month.

2

of the day- [00:00-07:59], [08:00-15:59] and [16:00-23:59]. Under the “Routes Summary” tab, we can identify the most delayed flight routes based on the origin and destination state. The two adjacent bar charts show the number of delayed and cancelled flight routes.

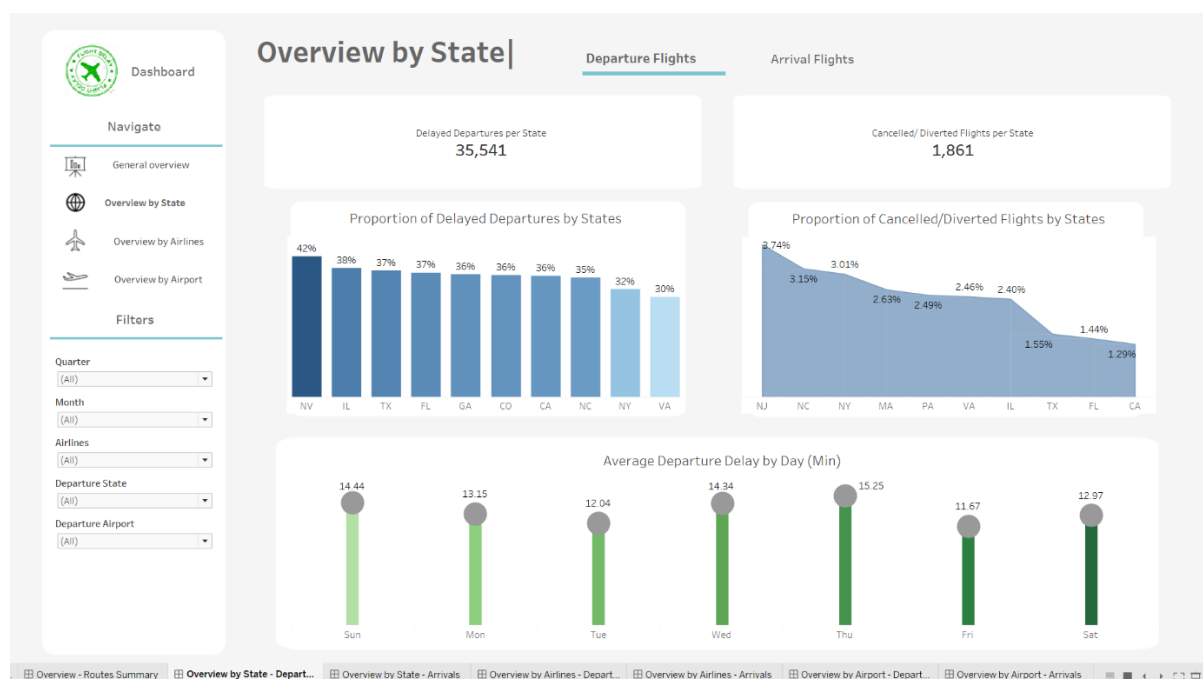
The general overview dashboard on flight delays is customizable to meet the needs of different users. For example, using the filter on the left below the navigation tab, we can view the flight delays based on quarters, months, airline, departure state, origin state, departure airport and destination airport. The overall goal of this dashboard is to assist users in immediately understanding the state of flight delays and identifying any patterns or trends that may be influencing flight delays. In addition, we can assess the flight delays based on states, airlines, and airports to get into more granularity. We can navigate to these from the left side of the dashboard under the navigate tab.

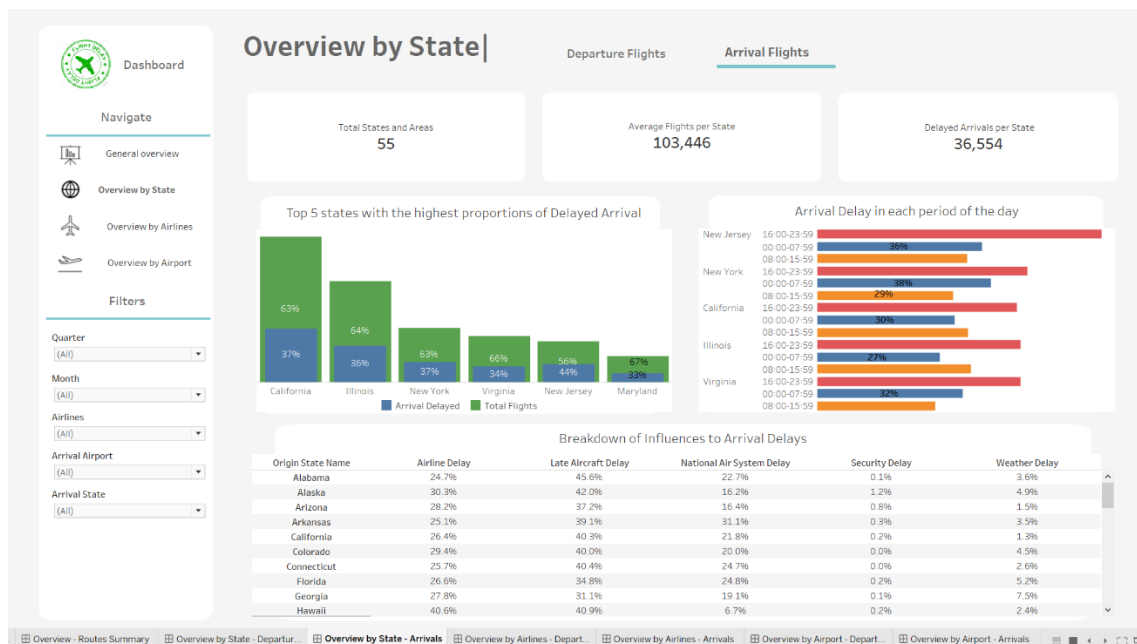


Overview by State

The Overview by State dashboard is a graphical depiction of data that shows aircraft delays in various states. The components of the dashboard include bar charts showing the "Top 5 states with the highest proportions of Delayed Arrival," "Arrival Delay in each period of the day," "Proportion of Delayed Departures by State," and "Average Departure Delay by Day (In Mins)." Furthermore, the table "Breakdown of Influences to Arrival Delay" provides insights into the frequency and severity of disruptions in air travel for each state, and the table "Breakdown of Influences to Arrival Delay" provides insights into the various factors that contribute to flight arrival delays. The dashboard based on states can assist decision makers such as the government, airlines, and other stakeholders in identifying areas for improvement in performance.

This data can assist them in prioritizing resource allocation to those specific regions to decrease delays. The dashboard can compare different states' or airports' on-time performance. Airlines and airports can look at trends in flight delays in other states by utilizing the dashboard over time. The information from the dashboard can be used to identify long-term issues and areas for improvement. Travellers can also use the dashboard to plan for probable flight delays. For example, if passengers know that flights from a specific state are routinely delayed, they may schedule a ticket from another state to prevent delays. Airlines may utilize the dashboard to give consumers real-time updates on flight delays and to manage customer expectations proactively. Airlines can boost customer satisfaction even in-flight delays by informing consumers and giving rebooking choices.

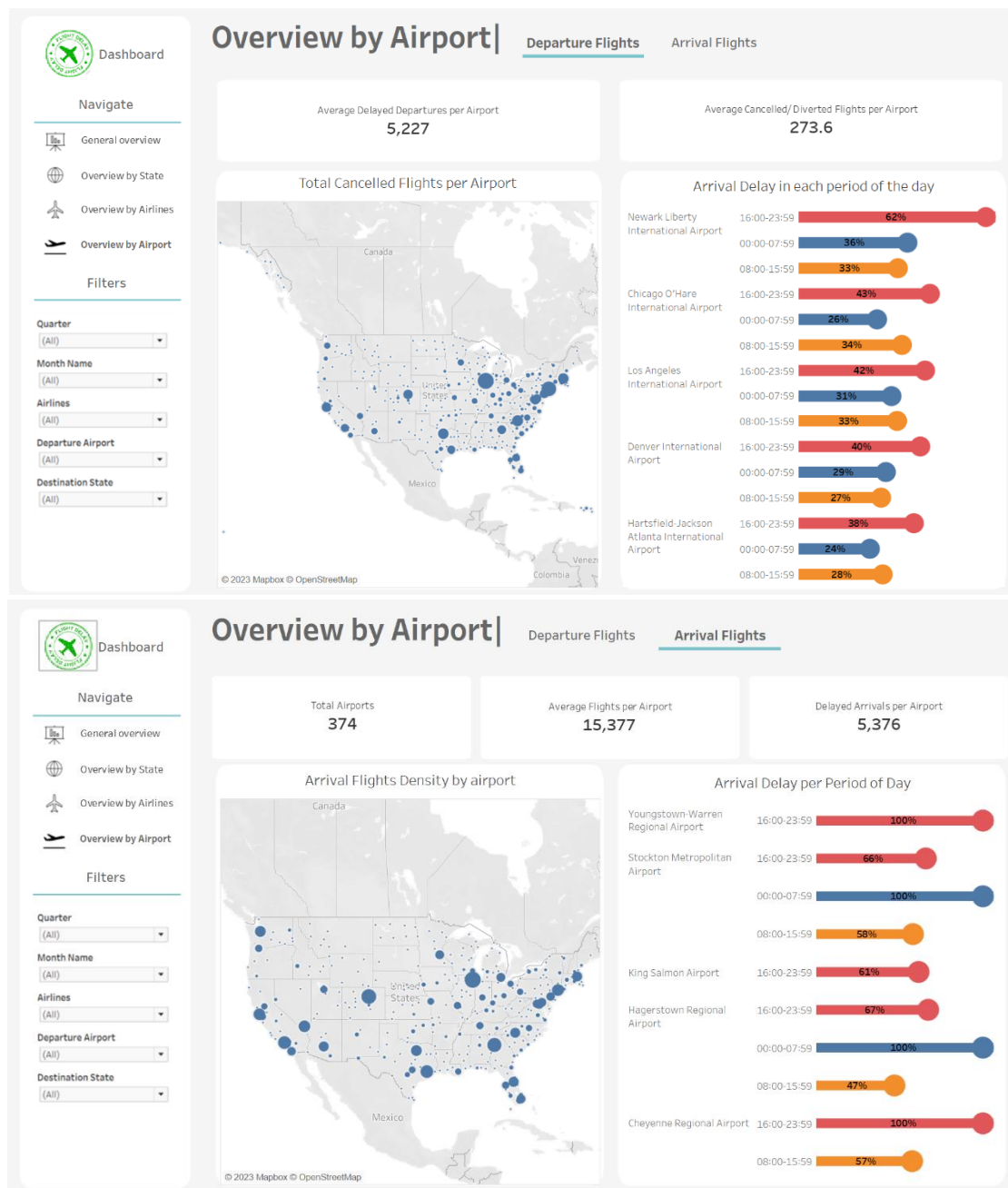




Overview by Airports

The Airport's dashboard is similar to the Overview of States. But this provides further in-depth details of the different airports in every state. This dashboard helps visualise the "Arrival Flights Density by airport", "Total Cancelled Flights per airport", "Proportion of Departure Delayed per Period of the Day", "Proportion Arrival Delays per Period of the Day", and "Proportion of Cancelled/Diverted Flights by Airport" with the help of maps and bar charts. This visualisation can provide insights into the busiest airports and those that receive a high volume of arrival and departure flights. It can also help identify airports with capacity issues due to high incoming and outgoing flight density, leading to potential delays or cancellations. It provides insights into the peak hours when most flights experience arrival and departure delays.

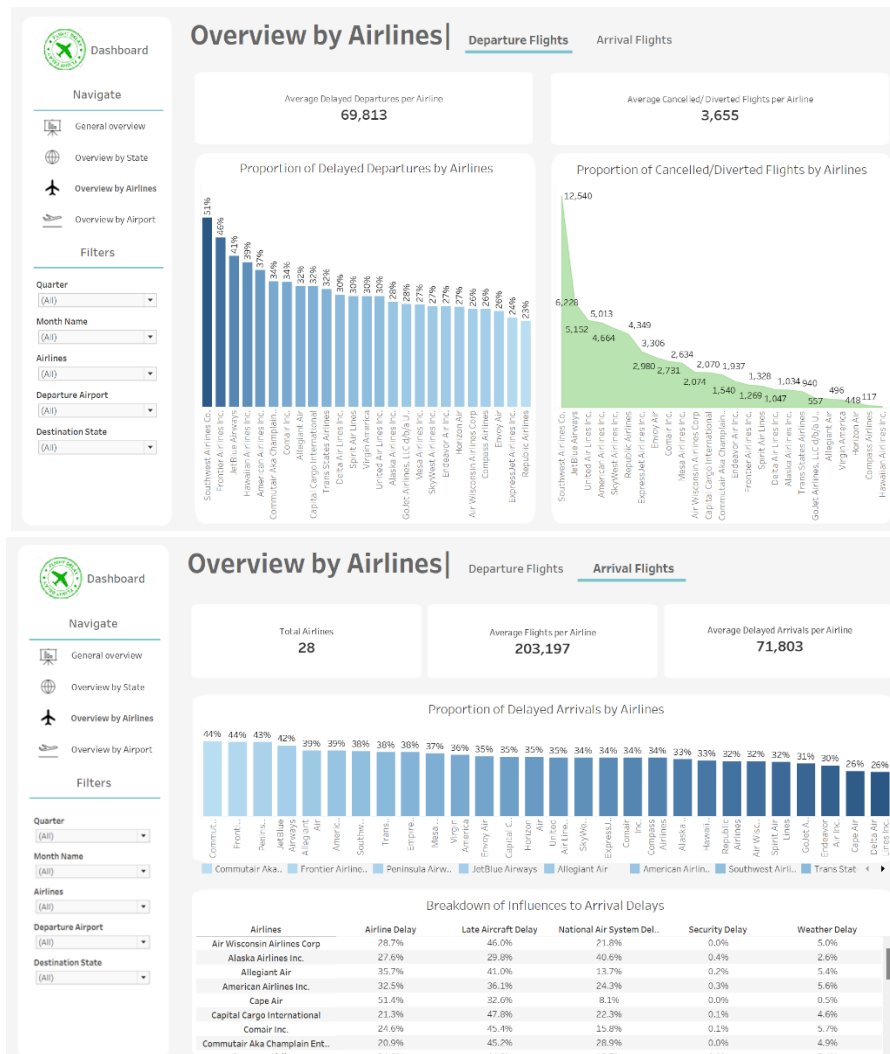
The airport-based flight delays dashboard can give vital information that can be utilised to improve flight planning and management. It can assist airlines and passengers in avoiding inefficient airports. The dashboard allows users to compare delay rates between airports, which is valuable information for airlines and customers when deciding where to fly from or to. It can also help airlines and airports analyse trends in flight delays over time, allowing them to spot patterns and alter their operations accordingly. The dashboard can reveal the causes of aircraft delays at various airports, such as weather, air traffic control, or airline-related difficulties. This information may be utilised to create strategies for reducing delays and increasing overall efficiency.



Overview by Airlines

The dashboard on airline flight delays gives an easy-to-read and intuitive way to assess and understand airline performance regarding flight delays. With bar charts, the dashboard visualizes the proportion of delayed arrivals and departures by airlines. It shows the proportion of flights for each airline that were delayed in arrival or departure. It also aids in visualizing the percentage of flights cancelled or diverted for each airline in the data. The table "Breakdown of Influences on Arrival Delays" displays the elements contributing to flight delays for each airline. It aids in determining which airlines are experiencing the most significant delays and what reasons lead to the delays.

To gain insights from the dashboard based on airlines, we analyze the data offered in the dashboard and find patterns, trends, and significant drivers of flight delays. By inspecting the dashboard, we can determine which airlines have the longest and shortest average delay durations. These airlines may be worth investigating to see why their performance is better or worse than others. It gives information about the seasonality of airline delays. The dashboard determines the reasons for aircraft delays, such as weather, mechanical difficulties, or air traffic control. Finally, we can detect which airlines are operating well and which need to improve their performance by using the dashboard's filters to compare the performance of different airlines.



Result and Conclusion

From the general overview of the flight delays dashboard, we can see that there are a total of 374 airports, 28 airlines, and 5689512 total flights. We see that the total number of flights delayed are 2010484 in 2018 constitutes 35.3 % and the flights cancelled were 88373 (1.8%) and the diverted flights are 13955 (0.2%). The month of November saw the maximum 11.7%

of arrival delay and the month of August had 12.92 % of departure delay. The maximum number of flights cancelled was 13.2% in the month of August. The average departing flights were maximum during the time [08:00-15:59] of the day. The state of California and Texas are the states with the maximum number of flights. Finally, we can compare the average Taxi In and Out time over the year where Taxi Out time was higher compared to Taxi In. We can also see that the flight route Chicago – New York faces the most delay while the route San Francisco – Los Angeles often gets cancelled.

In the States overview, delayed departure per state were observed at 35541. Cancelled or diverted flights per state is seen at 1861. The proportion of delayed departure by states is maximum for the state for Nevada (NV) which is at 42%. The proportion of cancelled or diverted flights is maximum for the state of New Jersey (NJ) at 3.74%. We see that the average departure delay by day (minutes) is maximum on Thursday which is 15.25 mins. Out of 103446 flights per state, delayed arrival per state is seen at 36554. The top three states with the highest proportions of delay are California, Illinois and New York. In the states of New Jersey, California, Illinois, New York and Virginia experience the maximum arrival delay is during [16:00-23:59]. From the break down of influences for arrival delays airline delay is maximum for Hawaii at 40.6%, Late aircraft delay is 54.3% in US virgin islands, National Air System delay is 37.8% at Main, Security delay is 1.2% at Alaska and Weather delay is 9.8% in Minnesota.

For Airline, we see that Southwest airline experiences the highest number of delays in departure at 51% and cancelled/ diverted flights at 12540 flights respectively. The proportion of Delayed Arrivals by Airlines is maximum for Commutair Aka Champlain Enterprises, Inc at 44%. The average delayed departures per airport is 5227 and average cancelled or diverted flights per airport is 273.6 minutes. Total cancelled flights is maximum at Chicago airport with 6314 flights. Average Flights per Airport is 15,377 and Delayed Arrivals per Airport is 5,376. Arrival Flights Density by airport is seen maximum at Chicago Airport with total Flights at 2,89,451.

Dashboard Link

Team 27

<https://public.tableau.com/app/profile/sudeep.tandon/viz/AirlinesDelayDescisionSupportSystemDashboard/Title?publish=yes>