Airline

# AIRLINE DELAYS

**TEAM:**

Ashley, Gopi, Kiena, Tracy, Zainab

**BOARDING PASS**

- **FLIGHT**
  B345
- **GATE**
  D8
- **SEAT**
  29E

Airline

# Project Overview

**Airlines Dataset to predict Delays & Airline Delays with Weather and Airport Detail**

In the present world the major components of a transportation system include a passenger airline. With time, we have evolved and improved the airline transportation system and operations. However, even in today's day and time, flight delays cause a lot of inconvenience to our modern passengers. Every year approximately 25-30% of flights are delayed, costing passengers and the operations approximately more than $28 billion in money and their time.


TEAM:
Ashley, Gopi, Kiena, Tracy, Zainab


Flights:
FREEPIK | FLATICON | STORYSET | WEPIK | VIDFY

Airline

# Purpose/Problem Statement

The purpose of this project is to predict whether a given flight will be delayed, given the information of the scheduled departure.

**01** Which airline has the most delayed flights?

**02** Which routes have the most delayed occurrences?

**03** Will flying time impact delay occurrences?

**04** Which date of the week will have the most delay? Weekend vs Weekday

**05** Which airport (departure/arrival) is the worst?

Airline

# Purpose/Problem Statement

## NOTE:

By conducting an analysis of flight times and metrics regarding airlines, day of travel, and airport arrivals/ departures we can predict what flights are more likely to experience delays. Once patterns are identified, solutions can be developed to address the issues.

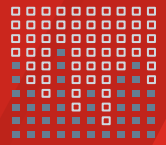
BOARDING PASS

- FLIGHT
  B345
- GATE
  D8
- SEAT
  29E

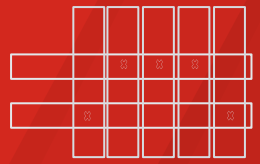
Airline

# The Roles

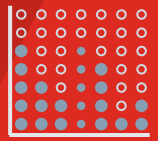
Ashley & Tracy



Gopi



Kiena



Zainab

**NOTE:**

Ashley, Gopi, Kiena, Tracy, Zainab

**BOARDING PASS**

- **FLIGHT**
  B345
- **GATE**
  D8
- **SEAT**
  29E

Airline

BOARDING PASS

# Proposal of Machine Learning Model

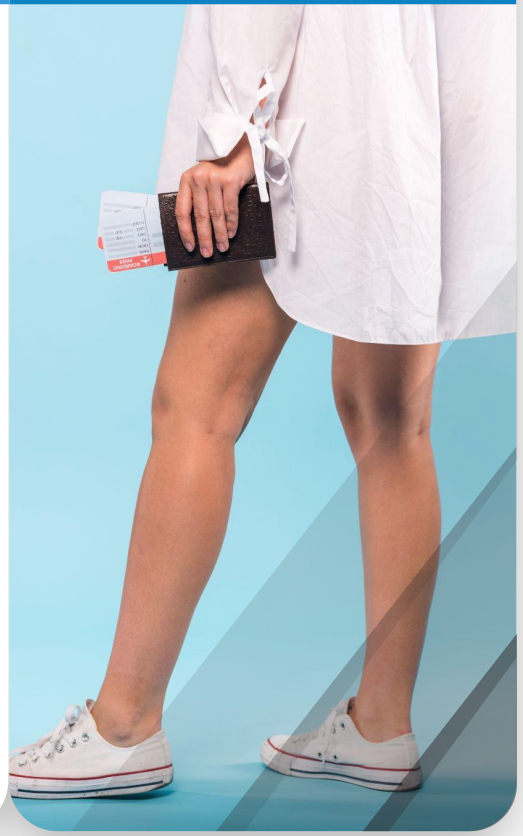
## NOTE:

Supervised Learning- Logistic Regression- binary result Delayed/Not delayed

- **What variables will be used for the machine learning model?**

  y= "Dep_Del15"

  X= other columns (Carrier and Airport columns need to be converted to numerical data)



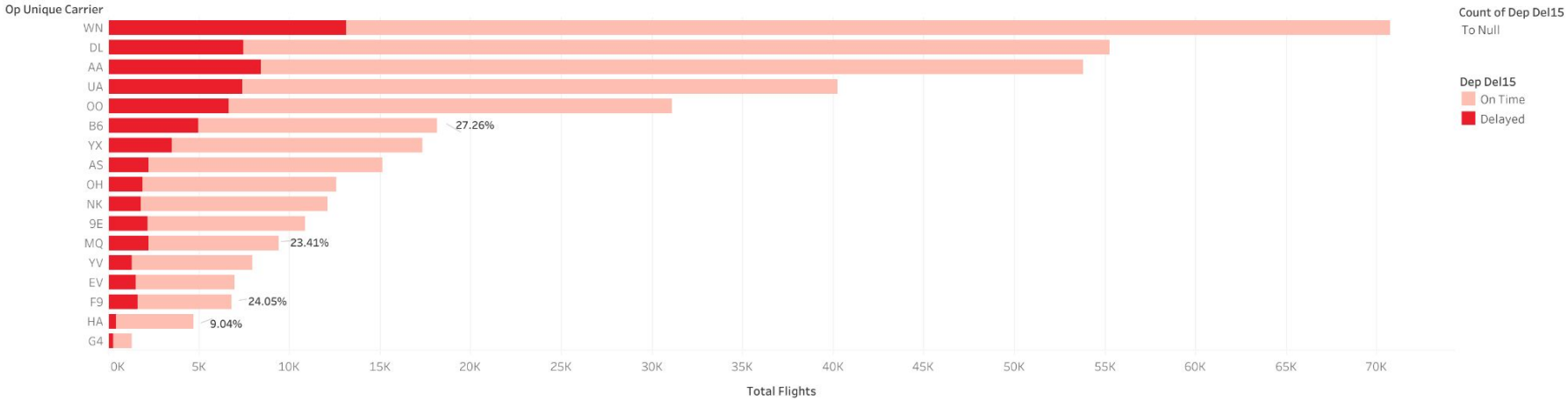
# Proposal of Machine Learning Model

## NOTE:

1) Create a model with LogisticRegression().
2) Train the model with model.fit().
3) Make predictions with model.predict().
4) Validate the model with accuracy_score().

Airline

# Percentage of Delayed Flights by Airline




## Summary Results of Delayed Flights by Airline

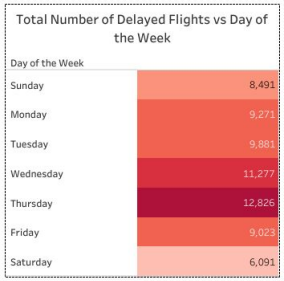
The top 3 airlines with the highest percentage of delayed flights to total flights are: JetBlue (B6), Frontier Airlines (F9) and Envoy Air (MQ).

The airline with the lowest percentage of delayed flights to total flights is Hawaiian Airlines (HA).

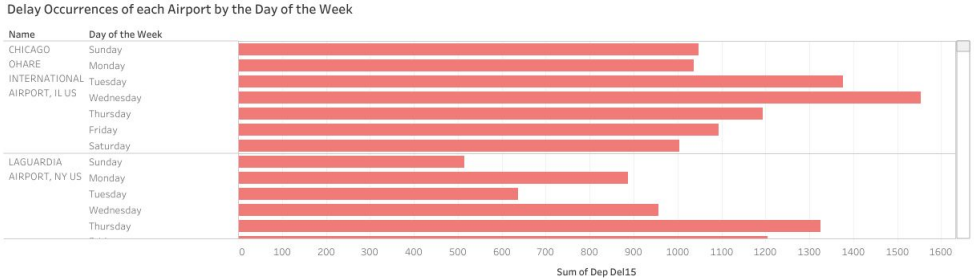
The majority of the listed airlines has a percentage of delayed flights that falls in between 14% and 21%.

# Airline

# Flight Delays by Day of the Week and Airport







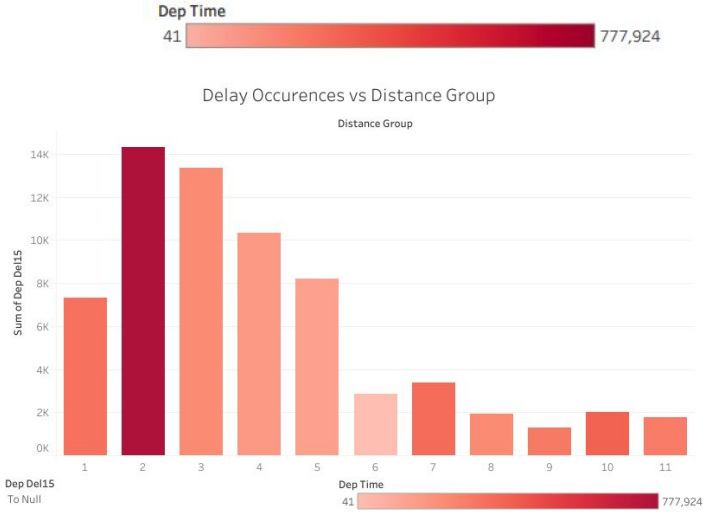


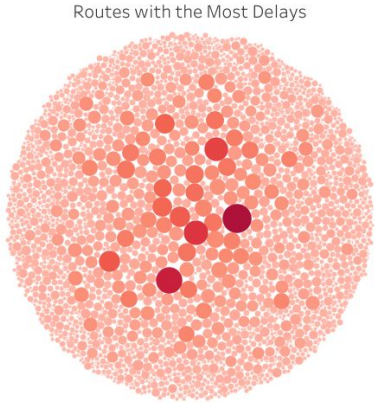
The Chicago O'Hare airport had the most overall flight delays and had the second most flights in total. Atlanta Hartfield Jackson had the most flights of all the listed airports and had the third most flight delays. Minneapolis St. Paul airport had the least amount of flights and flight delays of all the listed airports.

Overall, the busiest airports and busiest weekdays have the most amount of delayed flights. Thursday has the most amount of flights of the week as well as the most amount of delayed flights. Saturday has the least amount of total flights as well as the least amount of delayed flights. The majority of delayed flights occurred during the weekdays.

# Airline

# Flight Delays and Route Distances




## The Top 5 Routes with the Most Delays

**01** Chicago, IL to New York, NY

**02** New York, NY to Chicago, IL

**03** Boston, MA to New York, NY

**04** Los Angeles, CA to San Francisco, CA

**05** San Francisco, CA to Los Angeles, CA

The 11 distance groups in the graph below represent different lengths of flight distance, with distance group 1 covering a distance of 0 - 250 km. Each consecutive group has an increased distance of 250 km

Distance group 2 (250 - 499 km) has the most occurrences of delayed flights followed by distance group 3. The longer distanced flights, beginning from distance group 6 (1250 km) to distance group 11 (4999 km), have the least amount delayed flights

Routes going in and out of New York, NY appear to face the most delays.