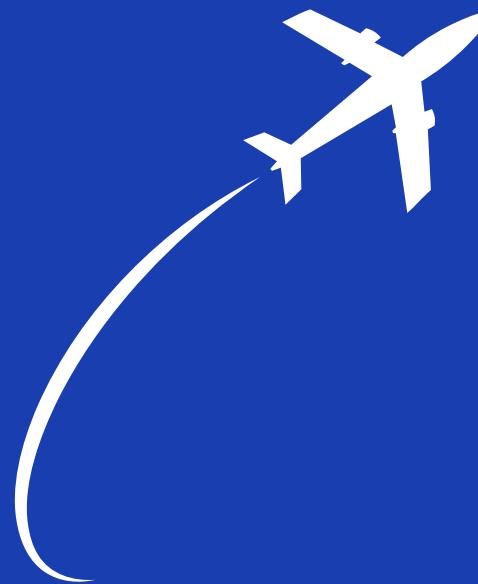


# US Flight Delay

SIC 6 Team 9



# About Us:



Ahmed Ashraf



Ahmed Ibrahim

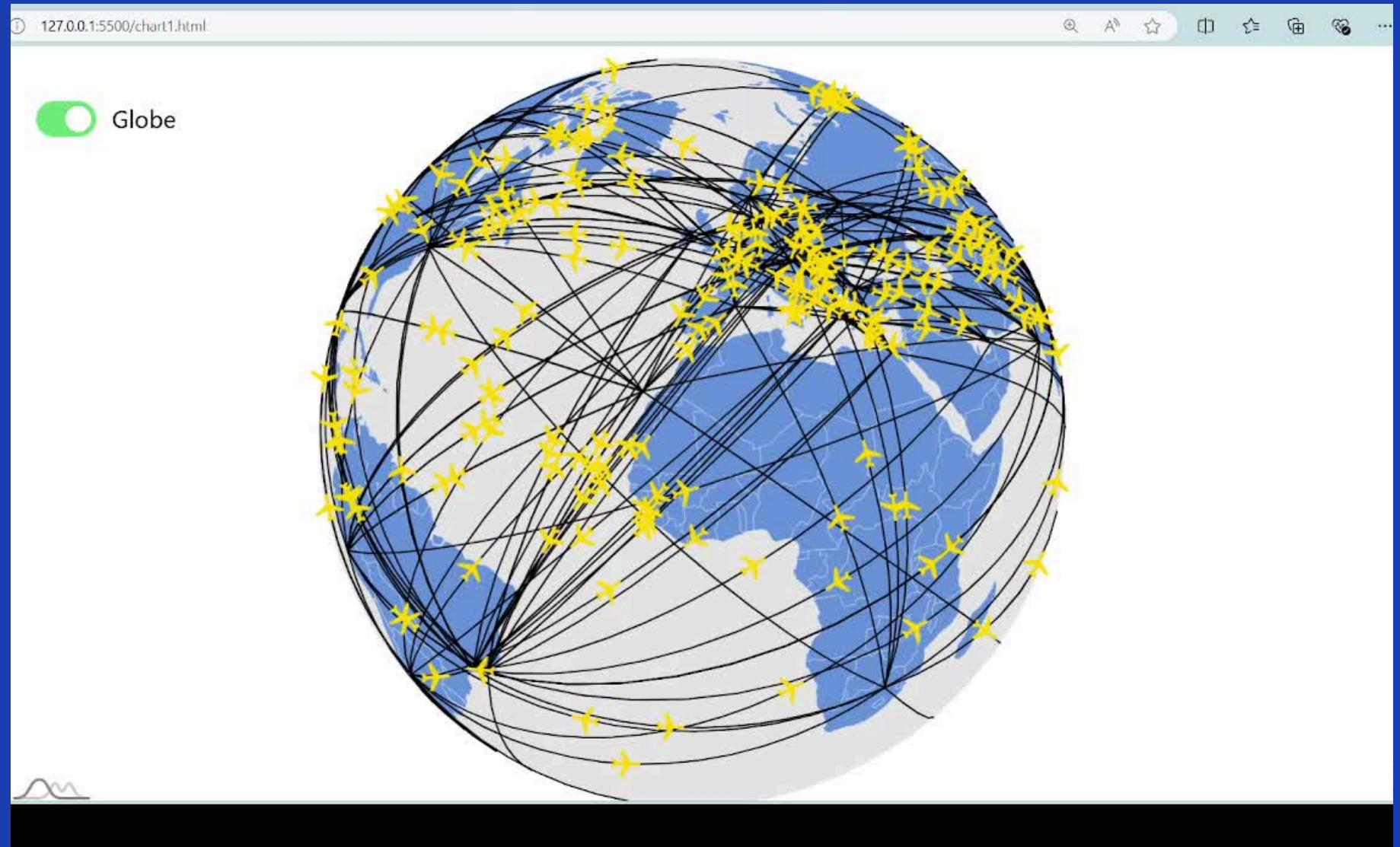


Farah Mostafa



## Under Supervision:

Eng: Mohamed Elshaghouny



# Agenda

- | Problem Statement
- | Business Case
- | Health Case
- | Data Description
- | Process
- | Dashboard
- | EDA & Business Analysis
- | Modeling
- | Deployment

## Problem Statement

Flight delays cause significant challenges for airlines, leading to higher operational costs, reduced customer satisfaction, and potential safety risks.



Addressing key causes such as weather, airport congestion, and outdated systems can improve on-time performance, enhancing efficiency and profitability.

## Business Case

---

- In 2023, the economic loss due to flight delays in the U.S. costs averaged **\$100.80 per minute**. Delays in flights result in additional costs for airlines, including additional gates and ground staff, lost productivity,
- The overall cost to the U.S. economy due to flight delays was estimated to be around **\$33 billion**.

## **Impact on Passenger Health**

- **In 2023, Frontier Airlines** experienced over **30%** of flights delayed. These delays not only disrupt travel plans but also cause increased passenger stress.
- **Health Risks:** Delays can lead to **exhaustion, dehydration, and heightened anxiety** due to prolonged waiting times and travel disruptions.
- **Safety Concerns:** In one case, **oxygen** masks were deployed due to pressurization issues

# Project 1

## Machine Learning Project

# Data Description

## US\_flights.csv

Field	Description	Example
FlightDate	The date of the flight.	2024-09-09
Day_Of_Week	The day of the week when the flight occurred.	Monday
Airline	The name or code of the airline operating the flight.	AA for American Airlines
Tail_Number	The unique identifier (registration number) of the aircraft.	N123AA
Dep_Airport	The IATA code of the departure airport.	JFK
Dep_CityName	The name of the city where the departure airport is located.	New York
DepTime_Label	The categorized or labeled departure time.	Morning
Dep_Delay	The number of minutes the flight was delayed at departure.	15
Dep_Delay_Tag	Categorical representation of the departure delay.	On-time
Dep_Delay_Type	The type or reason for the departure delay.	Operational
Arr_Airport	The IATA code of the arrival airport.	LAX
Arr_CityName	The name of the city where the arrival airport is located.	Los Angeles
Arr_Delay	The number of minutes the flight was delayed at arrival.	10
Arr_Delay_Type	The type or reason for the arrival delay.	Weather
Flight_Duration	The duration of the flight.	300 minutes
Distance_type	The classification of flight distance.	Long-haul
Delay_Carrier	The number of minutes the delay is attributed to the airline/carrier.	5
Delay_Weather	The number of minutes the delay is attributed to weather conditions.	10
Delay_NAS	The number of minutes the delay is attributed to National Airspace System issues.	0
Delay_Security	The number of minutes the delay is attributed to security-related issues.	0
Delay_LastAircraft	The number of minutes the delay is attributed to delays in the aircraft's previous flight.	20
Manufacturer	The company that manufactured the aircraft.	Boeing
Model	The model of the aircraft.	Boeing 737
Aircraft_age	The age of the aircraft in years.	12 years

## weather\_meteo\_by\_airport.csv

time	The specific time or timestamp associated with the flight record.	12:30 PM
tavg	The average temperature on the day of the flight.	22°C
tmin	The minimum temperature on the day of the flight.	18°C
tmax	The maximum temperature on the day of the flight.	28°C
prcp	The amount of precipitation on the day of the flight.	2 mm
snow	The amount of snowfall on the day of the flight.	0 mm
wdir	The wind direction.	270°
wspd	The wind speed.	15 km/h
pres	The atmospheric pressure.	1015 hPa
airport_id	A unique identifier for the airport.	12345

## airports\_geolocation.csv

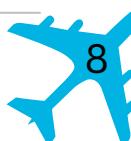
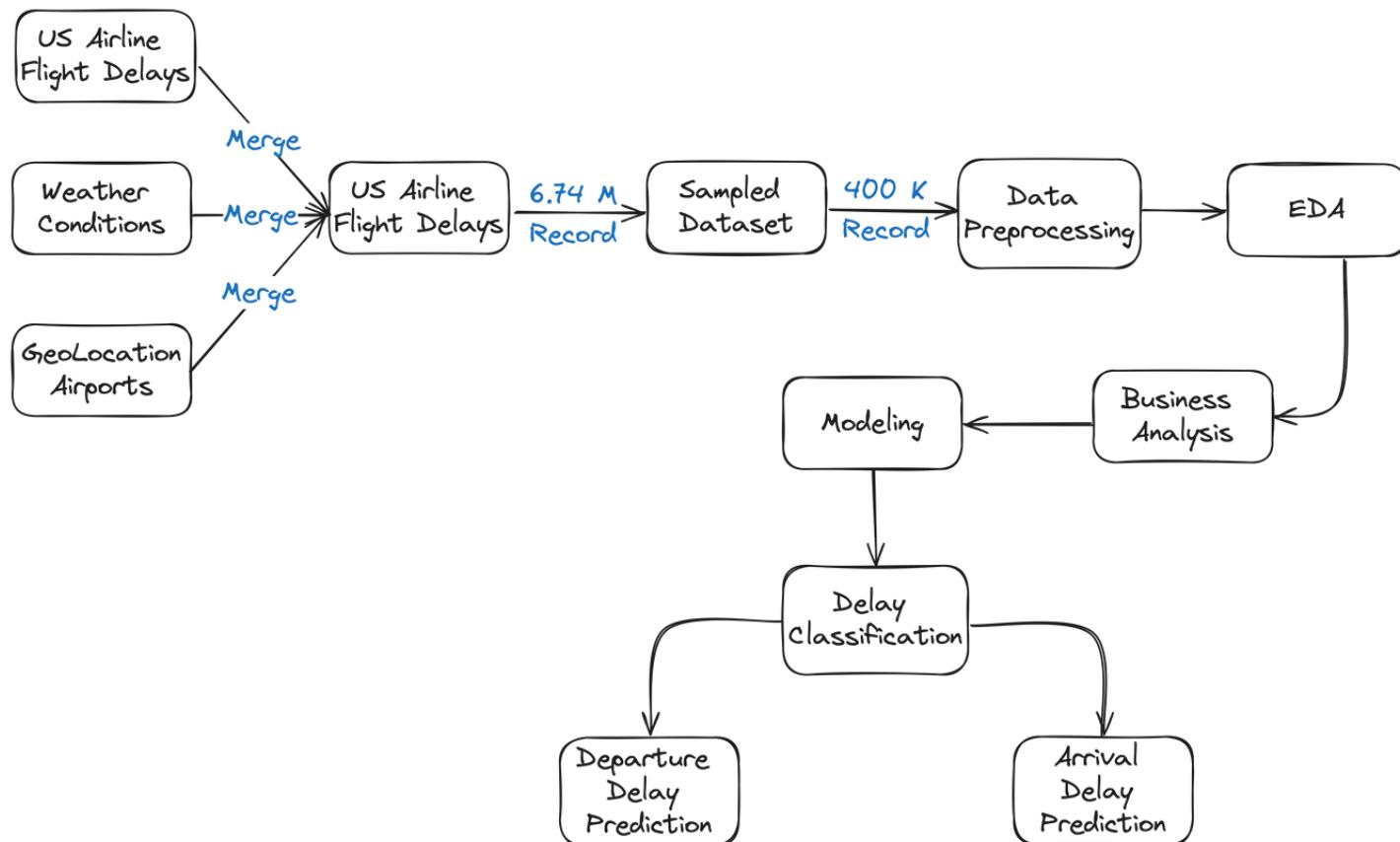
IATA_CODE	The International Air Transport Association code for the airport.	JFK
AIRPORT	The name of the airport.	John F. Kennedy International Airport
CITY	The city where the airport is located.	New York
STATE	The state or region where the airport is located.	NY
COUNTRY	The country where the airport is located.	USA
LATITUDE	The latitude coordinate of the airport.	40.6413° N
LONGITUDE	The longitude coordinate of the airport.	73.7781° W



# Project 1

## Machine Learning Project

## Process



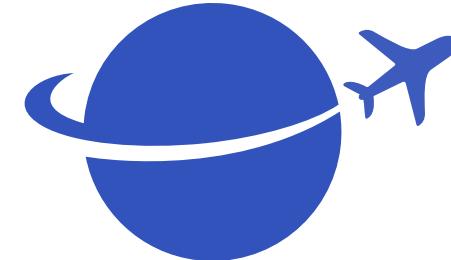
# Project 1

## Machine Learning Project

# Dashboard

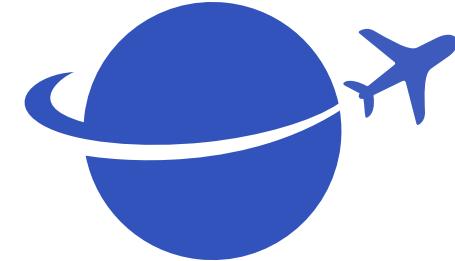


## **EDA & Business Analysis**



- 1. Which Airlines Have the Highest Delay Rates**
- 2. Which Days of the Week See the Most Delays?**
- 3. How Does the Aircraft Manufacturer Impact Flight Delays?**
- 4. Which States Experience the Most Flight Delays?**
- 5. How Do Wind Speed Impact Delays Across Regions?**
- 6. Does Temperature Variation Throughout the Day Affect Flight Delays?**

## **EDA & Business Analysis**



- 7.Which U.S. Cities Have the Highest Carrier-Related Flight with Departure and Arrival Delays?**
- 8.Which Types of Flight Delays Are Most Common?**
- 9.How Does Weather (Temp) Impact Departure Delays?**

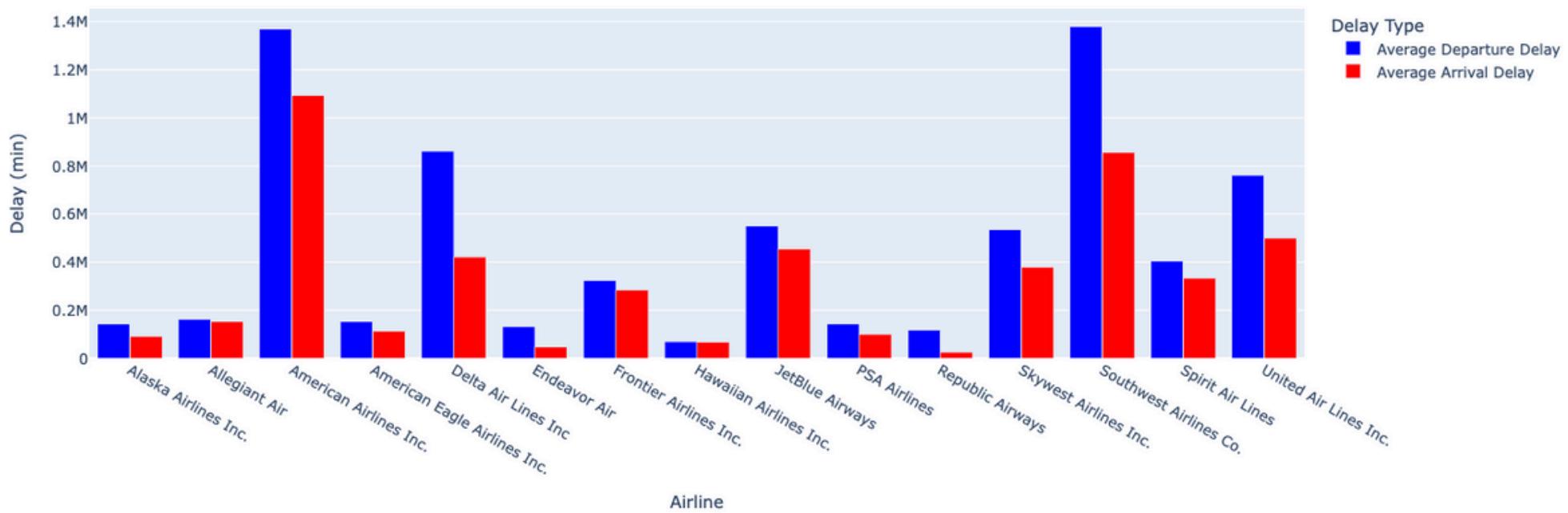
# Project 1

## Machine Learning Project

1

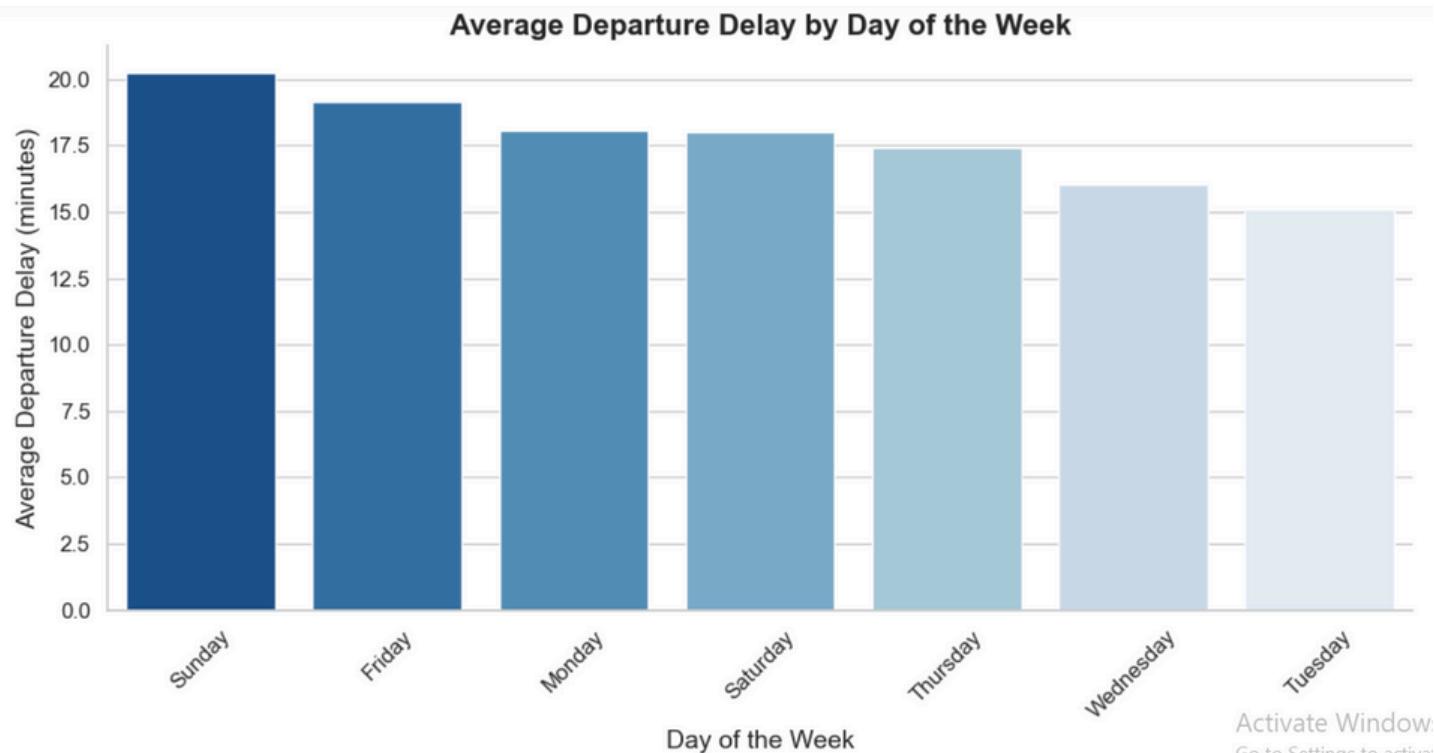
### Which Airlines Have the Highest Delay Rates

Average Departure and Arrival Delays by Airline



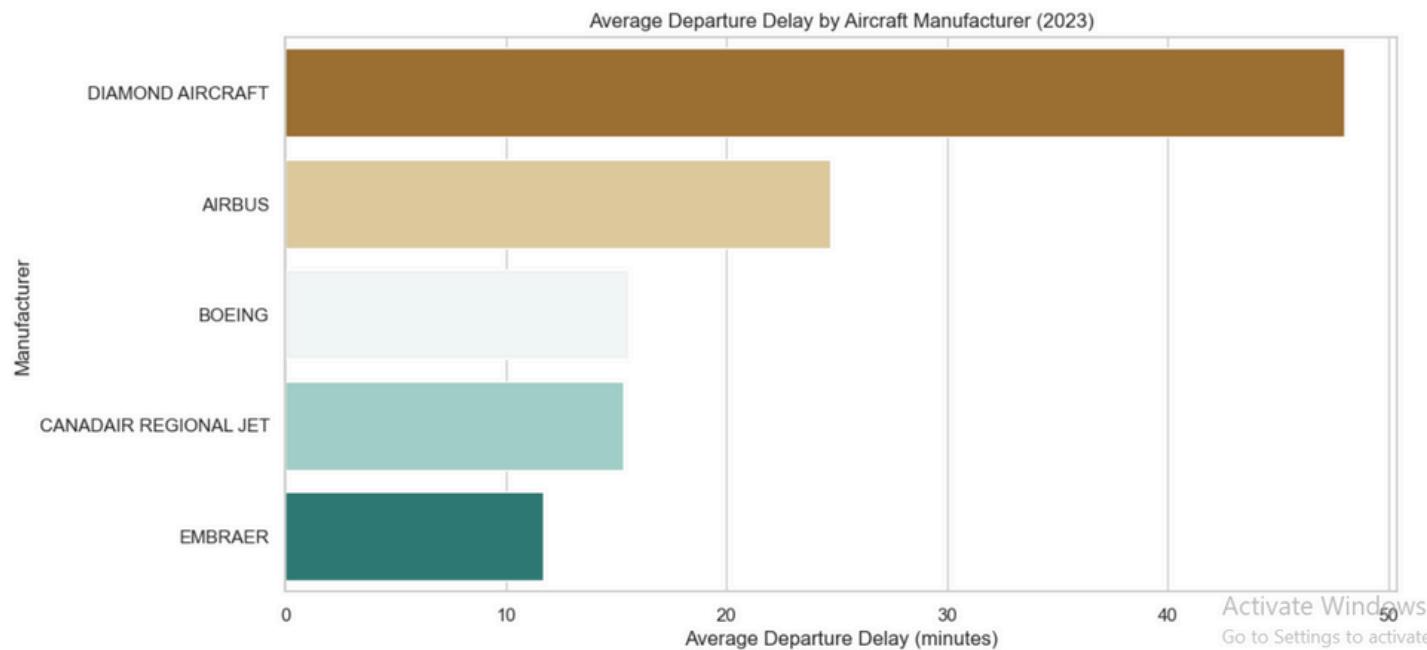
2

## Which Days of the Week See the Most Delays?



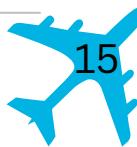
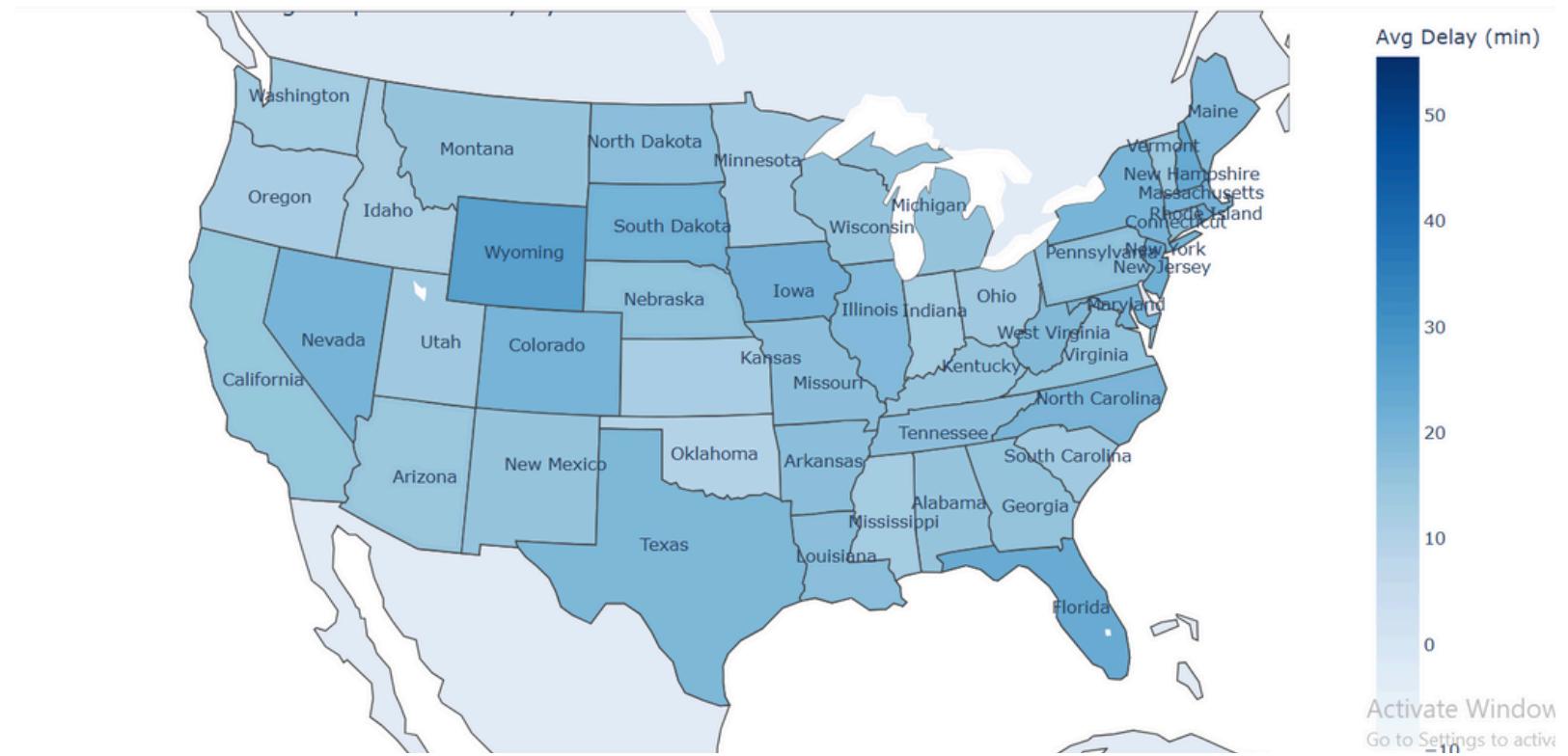
3

## How Does the Aircraft Manufacturer Impact Flight Delays?



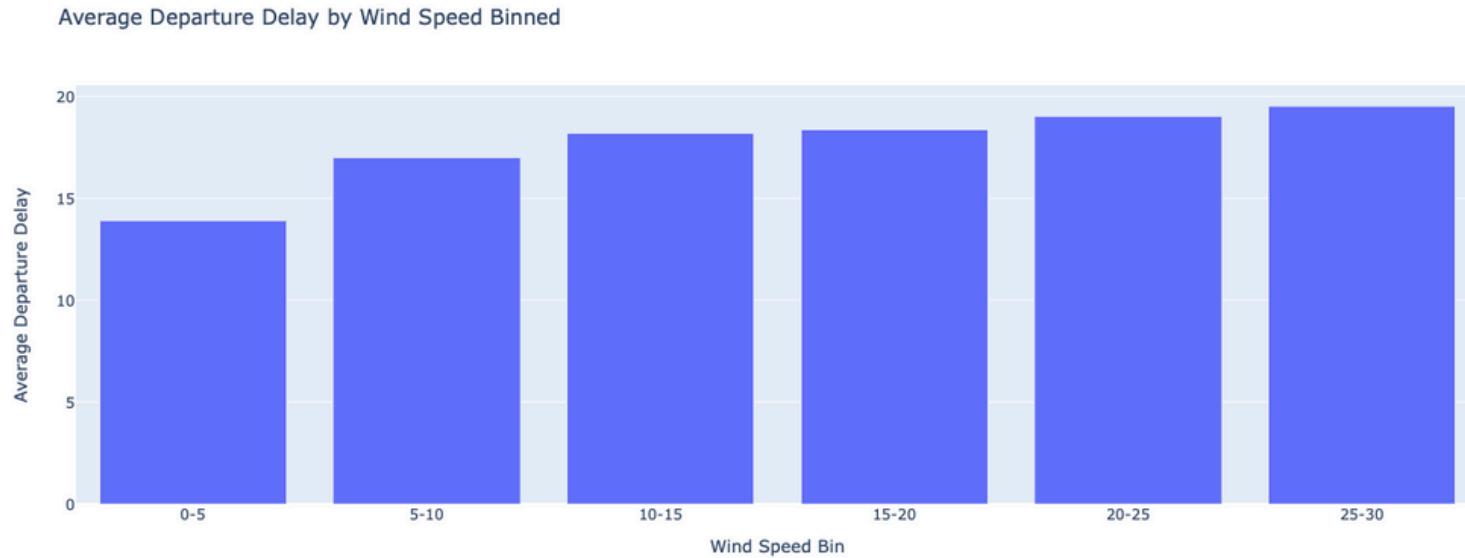
4

# Which States Experience the Most Flight Delays?



5

## How Do Wind Speed Impact Delays Across Regions?

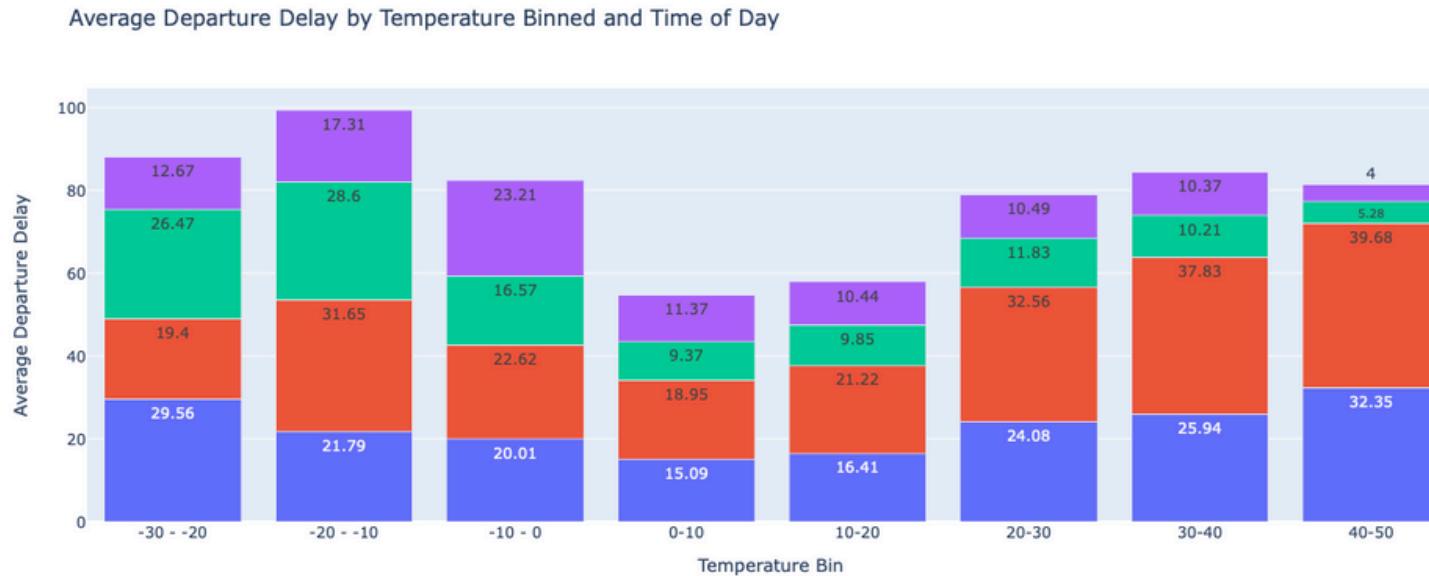


# Project 1

## Machine Learning Project

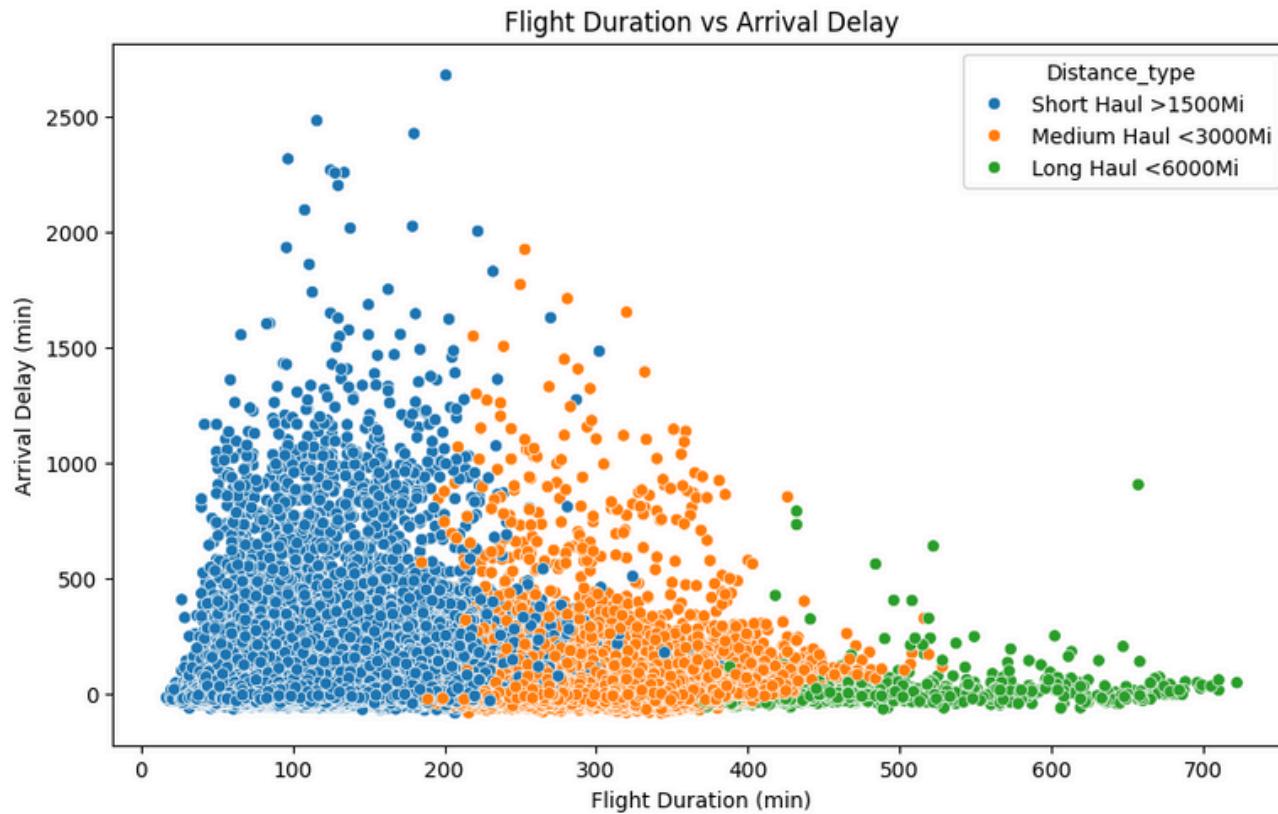
6

## Does Temperature Variation Throughout the Day Affect Flight Delays?



7

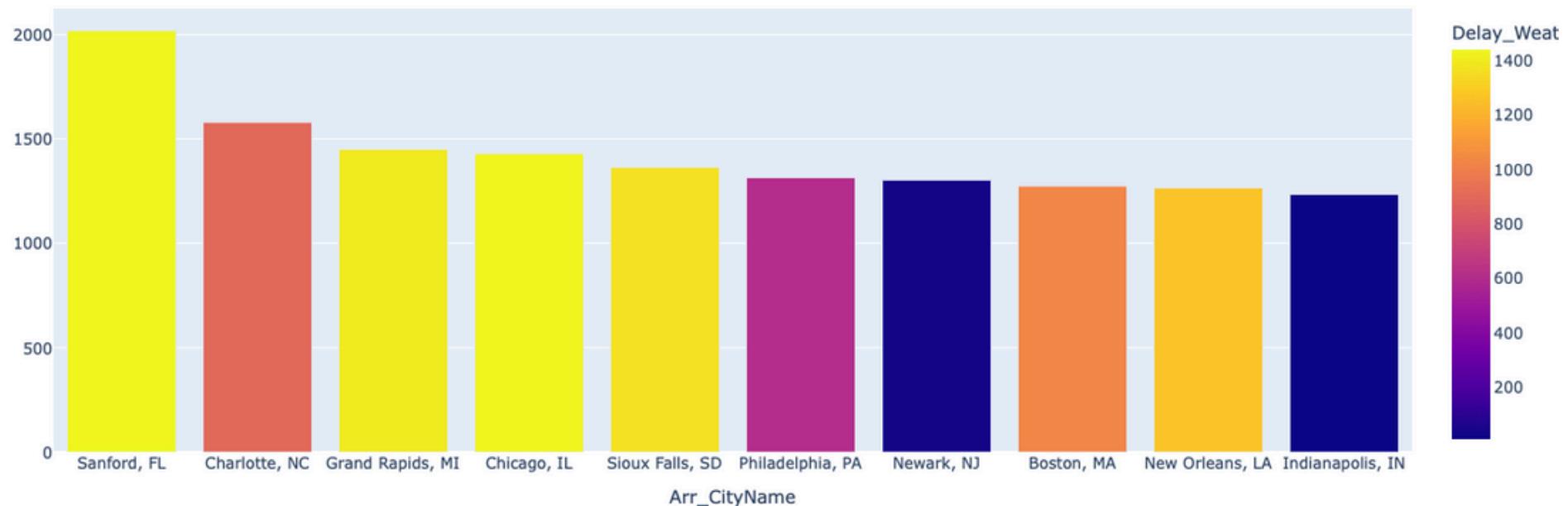
## Is there correlation between Flight duration and arrival delay?



8

## Which U.S. Cities Have the Highest Carrier-Related Flight with Departure and Arrival Delays?

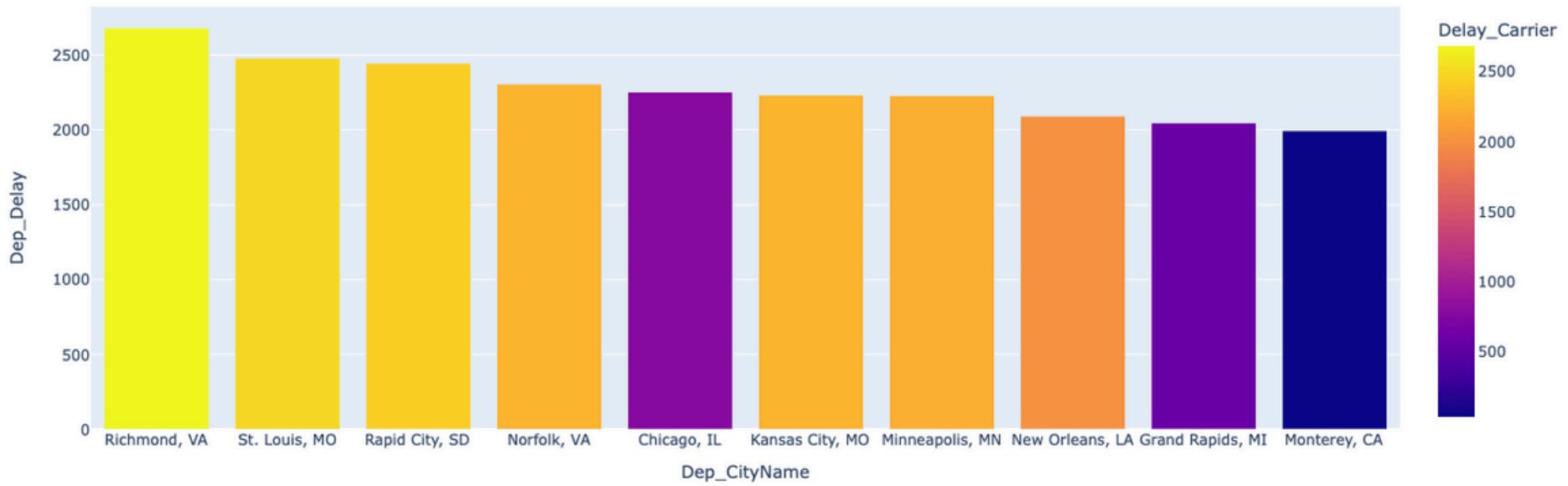
Top 10 Cities with Highest Rates of Arrival Delays



# Project 1

## Machine Learning Project

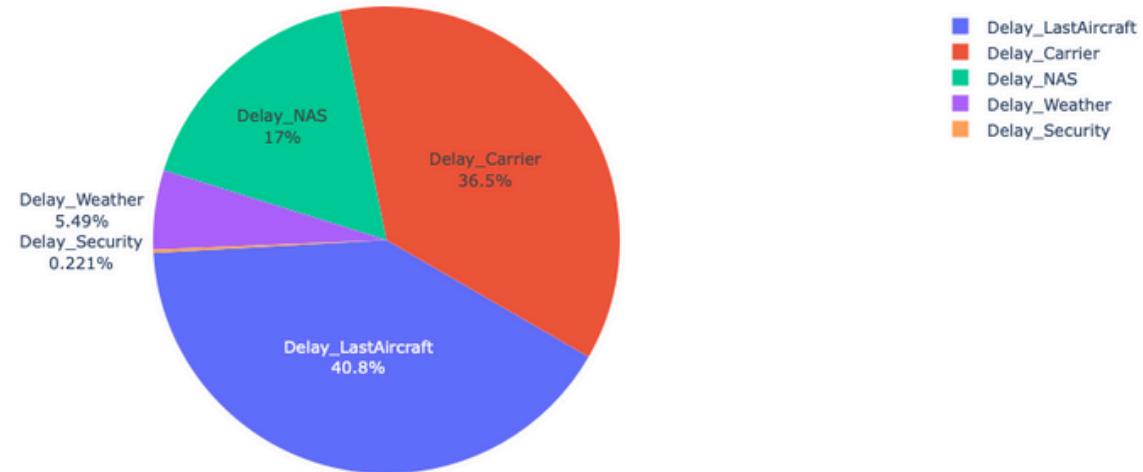
Top 10 Cities with Highest Rates of Departure Delays



9

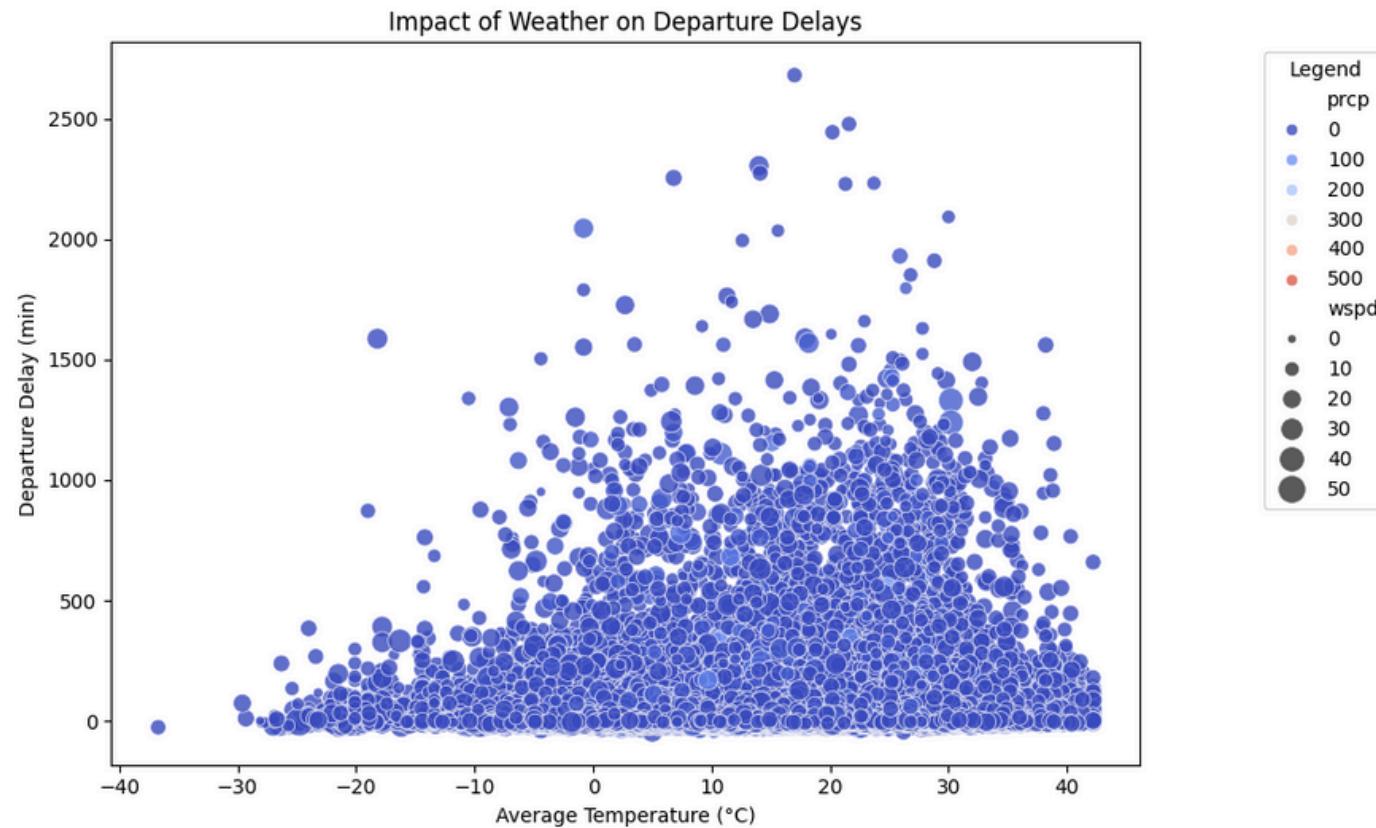
## Which Types of Flight Delays Are Most Common?

Distribution of Delay Types



10

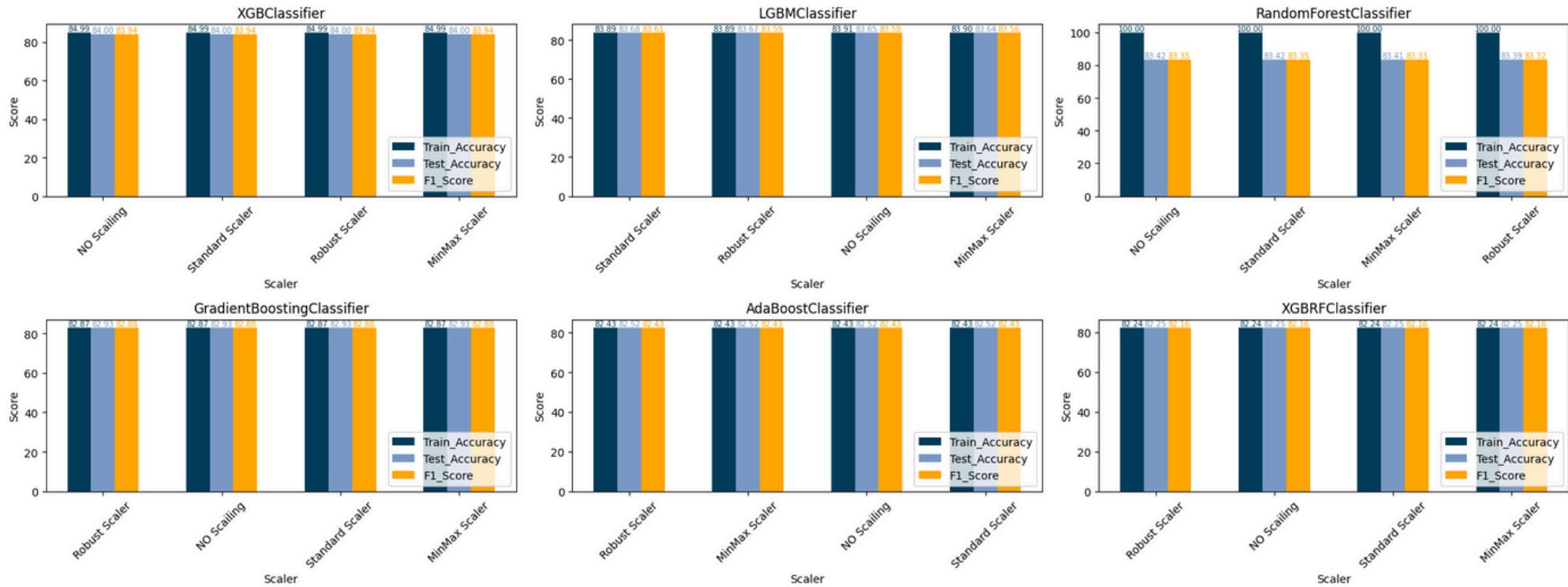
## How Does Weather (Temp) Impact Departure Delays?



# Project 1

## Machine Learning Project

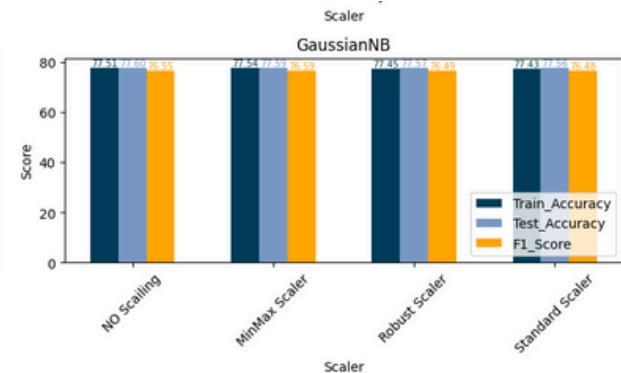
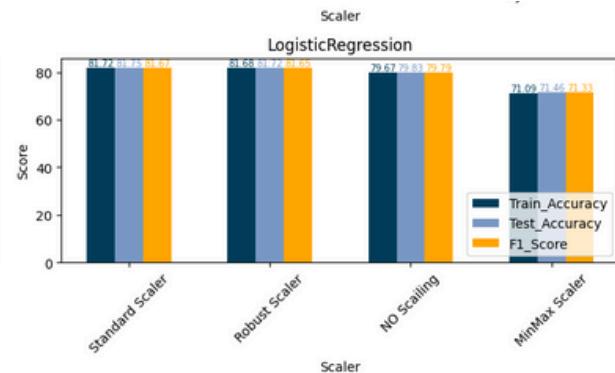
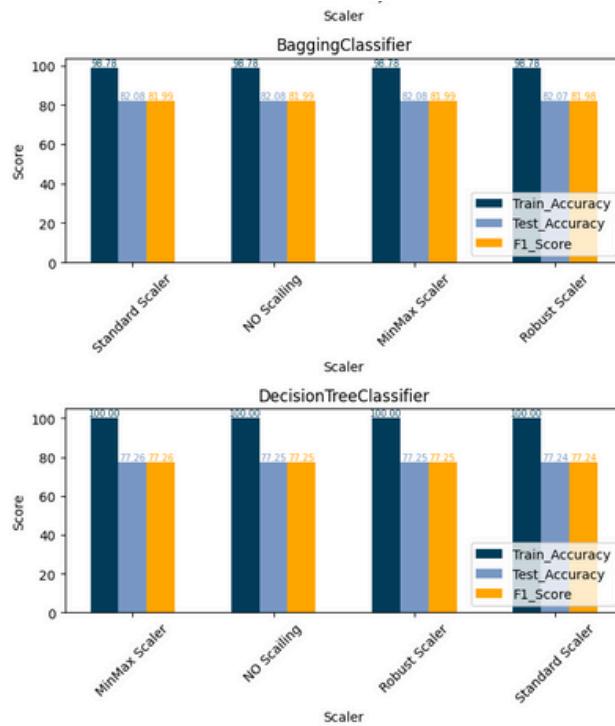
# Modeling



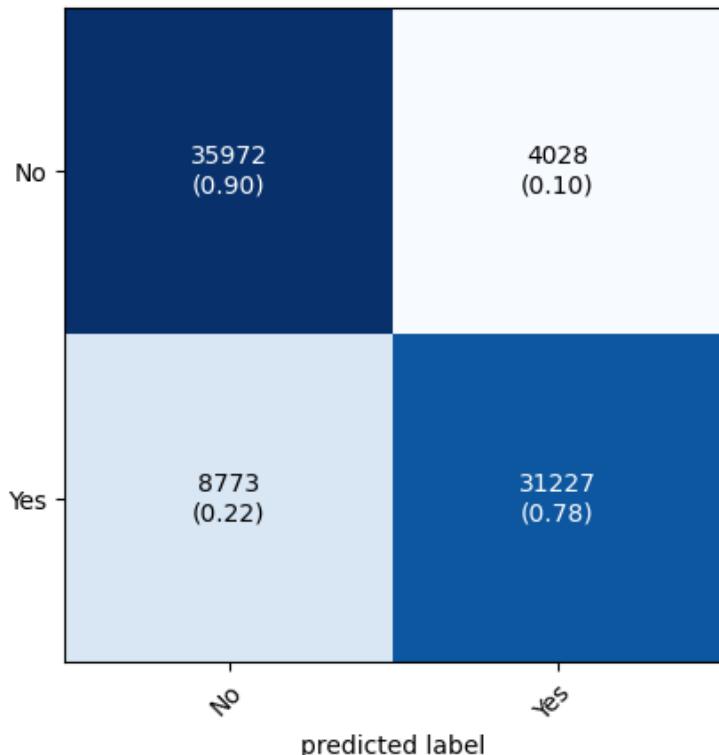
# Project 1

## Machine Learning Project

# Modeling



# Classification



		precision	recall	f1-score	support
		No	0.80	0.90	0.85
accuracy	Yes	0.89	0.78	0.83	40000
	macro avg	0.84	0.84	0.84	80000
	weighted avg	0.84	0.84	0.84	80000

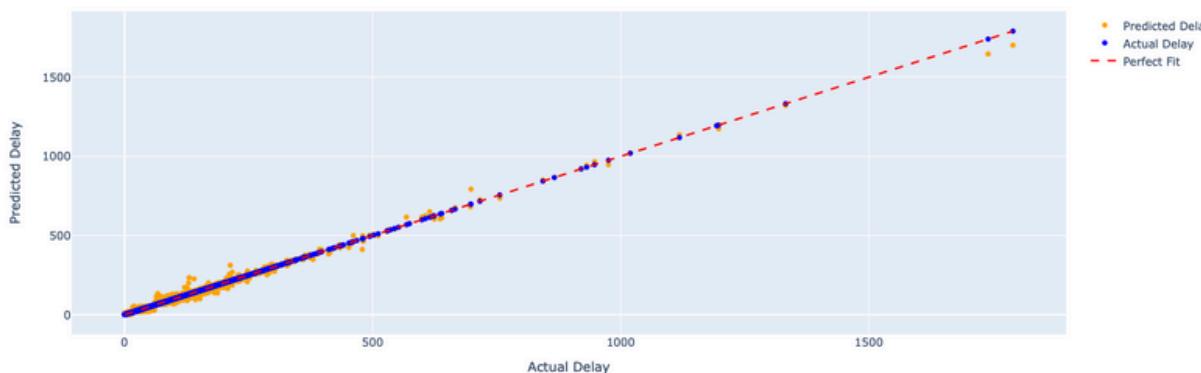


# Project 1

## Machine Learning Project

# Departure Regression

Actual vs Predicted Departure Delay



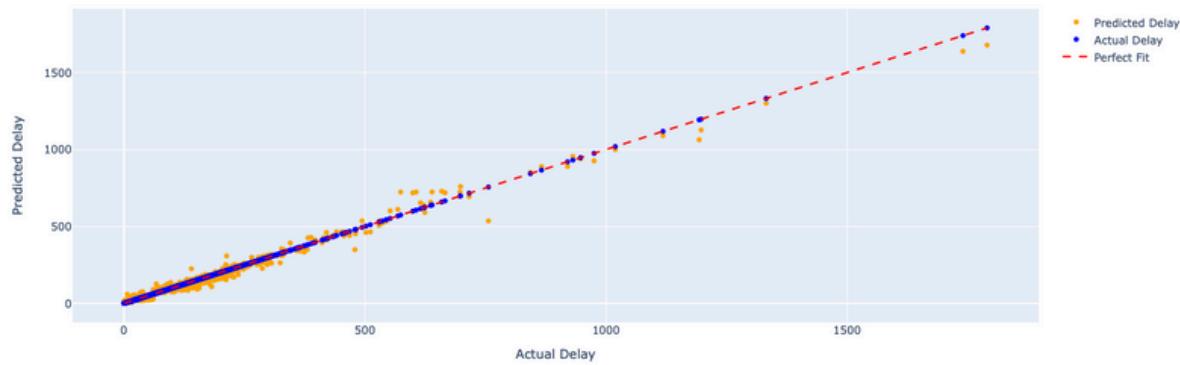
Model-Name	RMSE	R2_Score	MAE	Inference Time (ms)
RandomForestRegressor	7.775396	99.163893	4.886793	198.779345
GradientBoostingRegressor	8.011948	99.112245	5.679474	12.837887
BaggingRegressor	8.208510	99.068151	5.112168	26.554585
DecisionTreeRegressor	10.992108	98.328990	6.691643	5.024195
LGBMRegressor	11.203162	98.264206	5.221853	41.761875
Ridge	13.552559	97.459849	9.402284	3.659248
LinearRegression	13.552592	97.459836	9.402531	3.486633
Lasso	13.980949	97.296725	9.565284	3.253937
XGBRegressor	15.998452	96.460248	5.630310	21.680593
AdaBoostRegressor	18.567204	95.232288	14.838796	20.307541

# Project 1

## Machine Learning Project

# Arrival Regression

Actual vs Predicted Arrival Delay



Model-Name	RMSE	R2_Score	MAE	Inference Time (ms)
GradientBoostingRegressor	11.024389	98.319161	7.104620	10.396004
RandomForestRegressor	12.216214	97.936093	5.715394	194.059372
BaggingRegressor	13.435282	97.503621	6.060375	26.226997
Ridge	13.880576	97.335400	9.682106	9.928942
LinearRegression	13.880613	97.335386	9.682349	4.047632
Lasso	14.303144	97.170694	9.876387	3.023624
XGBRegressor	18.529701	95.251529	6.131650	21.267414
LGBMRegressor	19.117907	94.945273	6.228120	40.725470
DecisionTreeRegressor	19.800555	94.577848	7.800672	4.851580
AdaBoostRegressor	267.817654	-891.962777	264.979443	30.853271

# Project 1

## Machine Learning Project

# Deployment



This app aims to assist in three things:

- 1- Predict if there is delay in the departure of the flight.
- 2- How many minutes there will be a delay in the departure?
- 3- How much time there will be as a delay in the arrival of flight?

Day of Week (1 for Monday and 7 for Sunday)  
1 (highlighted)

Arrival Delay Type  
Low <5min

Airline Company  
Southwest Airlines Co.

Flight Distance  
Short Haul >1500Mi

Delay Carrier

Departure Time Period  
Evening

Delay Weather

Departure Delay Type  
Low <5min

Departure NAS

Arrival Delay  
0

Share WhatsApp Image... ⓘ ⌂ ⌂ ⌂ ⌂

FlightDate	2023-06-23 00:00:00
Day_Of_Week	5
Airline	Southwest Airlines Co.
Tail_Number	N936WN
Dep_Airport	DEN
Dep_CityName	Denver, CO
DepTime_label	Night
Dep_Delay	-2
Dep_Delay_Tag	0
Dep_Delay_Type	Low <5min
Arr_Airport	DAL
Arr_CityName	Dallas, TX
Arr_Delay	-5
Arr_Delay_Type	Low <5min
Flight_Duration	107
Distance_type	Short Haul >1500Mi
Delay_Carrier	0
Delay_Weather	0
Delay_NAS	0
Delay_Security	0
Delay_LastAircraft	0
Manufacturer	BOEING
Model	737 NG
Aircraft_ago	15
LONGITUDE	-104.667
Month	6
Day	23
Quarter	2
Name:	230867, dtype: object



# Project 1

## Machine Learning Project

Link to kaggle Notebook



**Hit that upvote button like you're snoozing through a flight delay—let's keep this data journey on time!** 😎✈️

SAMSUNG



Thank You