



*Christopher Giler*

# Predicting Airline On-Time Performance

Using Flight Schedules and  
Weather Forecasts

# Project Objective

- ❖ Can we use weather and airport data together to reliably predict airline arrival delays?
- ❖ Delay = Actual arrival at least 15 minutes later than scheduled arrival time.



---

# Data Sources

---

## ❖ Airline Data

❖ Bureau of Transportation  
Statistics

❖ Open Flights

**15**

Years  
(2010 - 2017)

**87.2 million**

Flight Records

---

## ❖ Weather Data

❖ National Oceanic and  
Atmospheric Administration

**105,000**

Weather Stations  
(Worldwide)

**3.5 million**

Daily Weather Records

**48**

Weather Metrics  
(Precipitation, Snowfall, Wind Speed, etc.)

---

# Data Sources

---

## ❖ Airline Data

❖ Bureau of Transportation  
Statistics

❖ Open Flights

8

Years  
(2010 - 2017)

200,000

Flight Records

(Delta Flights leaving LaGuardia Airport)

---

## ❖ Weather Data

❖ National Oceanic and  
Atmospheric Administration

105,000

Weather Stations  
(Worldwide)

3.5 million

Daily Weather Records

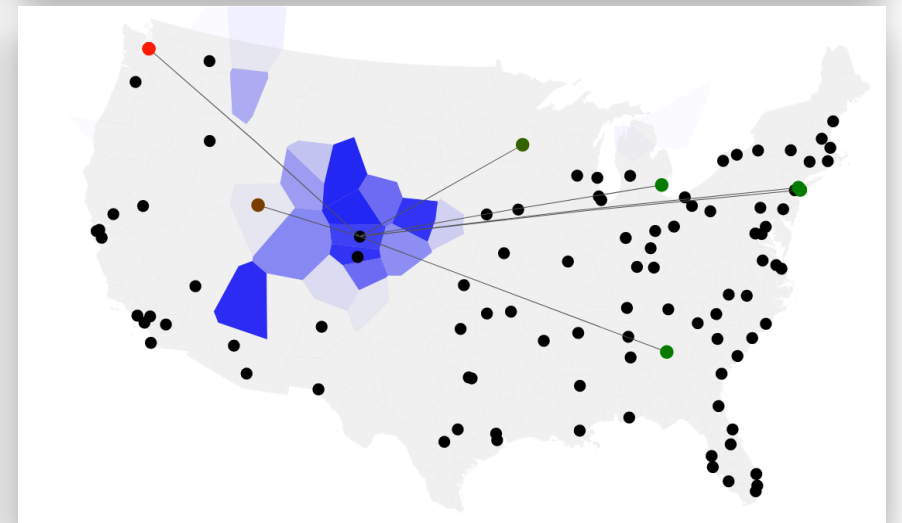
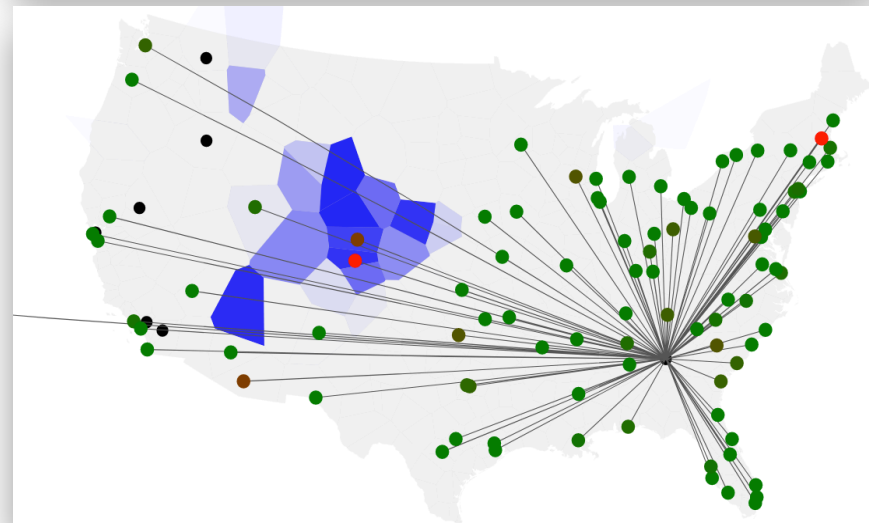
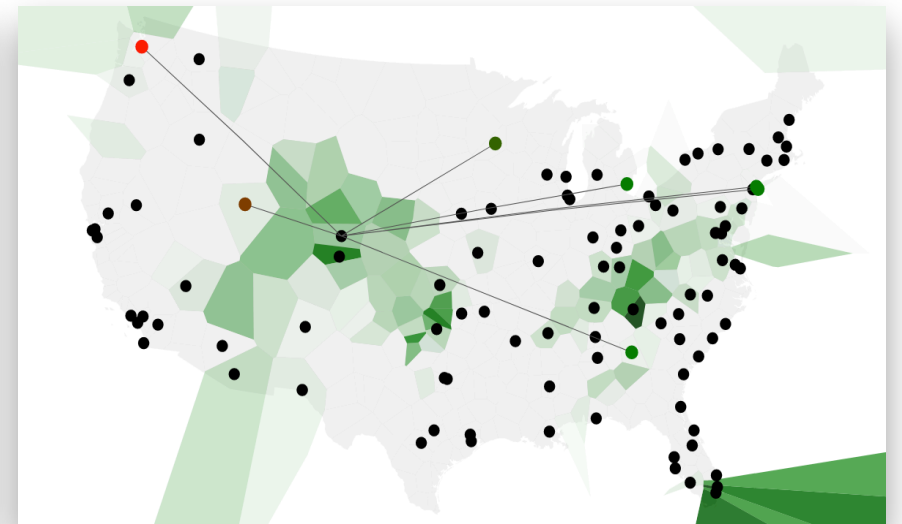
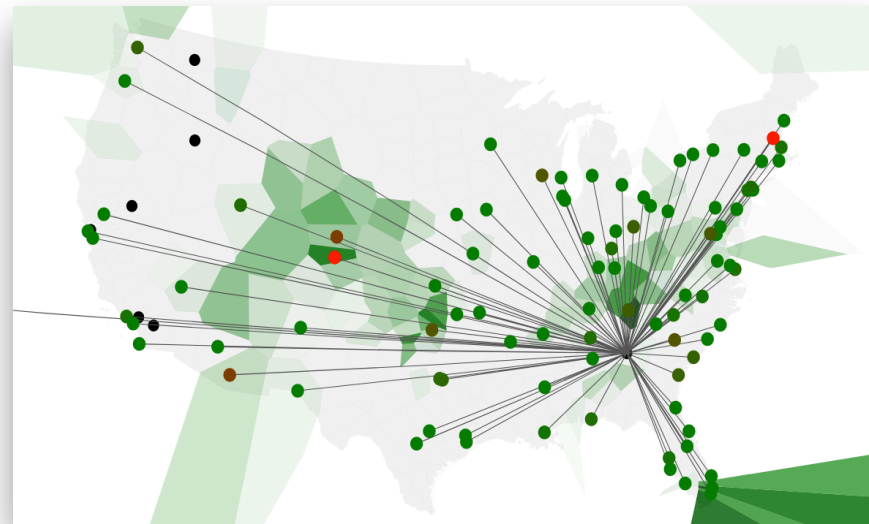
48

Weather Metrics  
(Precipitation, Snowfall, Wind Speed, etc.)



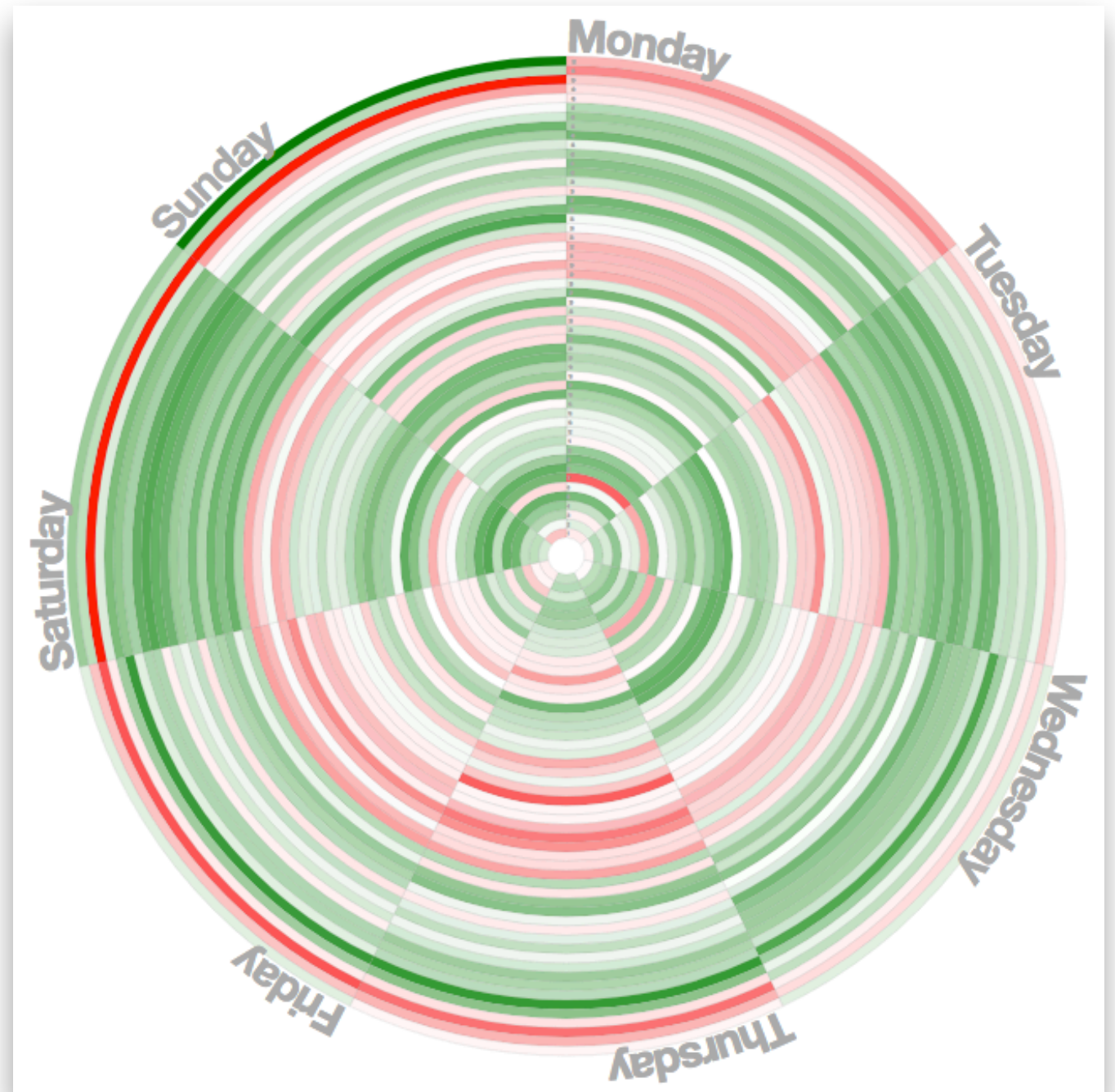
# Data Selection - Weather

- ❖ Data varies among airports and carriers
- ❖ Data shown: Delta Airlines flights on 2016-02-01.
- ❖ Included Metrics:
  - ❖ Precipitation
  - ❖ Snow Fall
  - ❖ Temperature
  - ❖ Wind Speed
  - ❖ Destination Airport



# Data Selection - Day of the Year

- ❖ Daily flight data for all LGA flights in 2016
- ❖ Visualization split by:
  - ❖ Week of the year
  - ❖ Day of the week
- ❖ Flight delays impacted by:
  - ❖ Seasonality
  - ❖ Day of the week



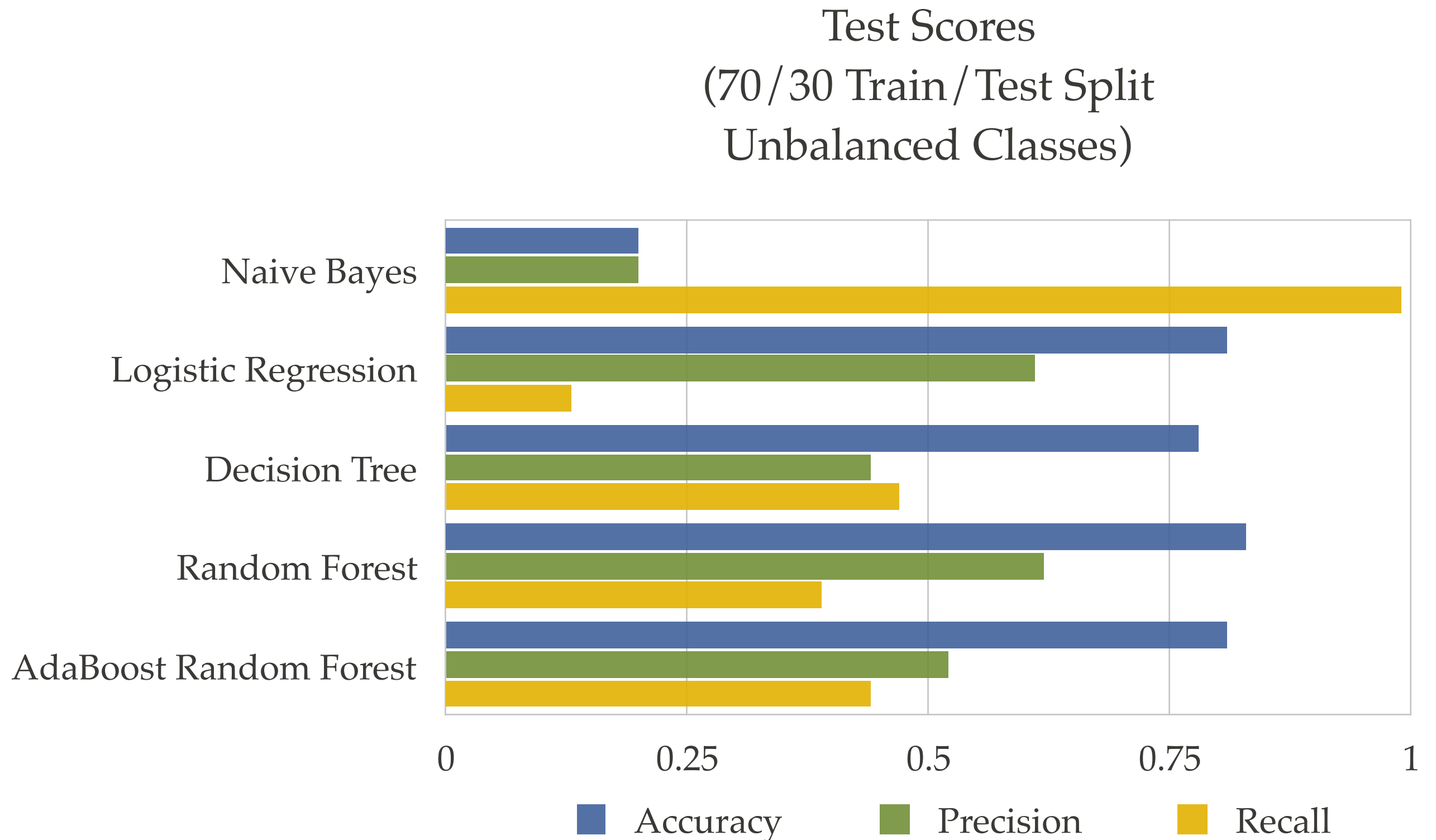
---

# Criteria for Success

---

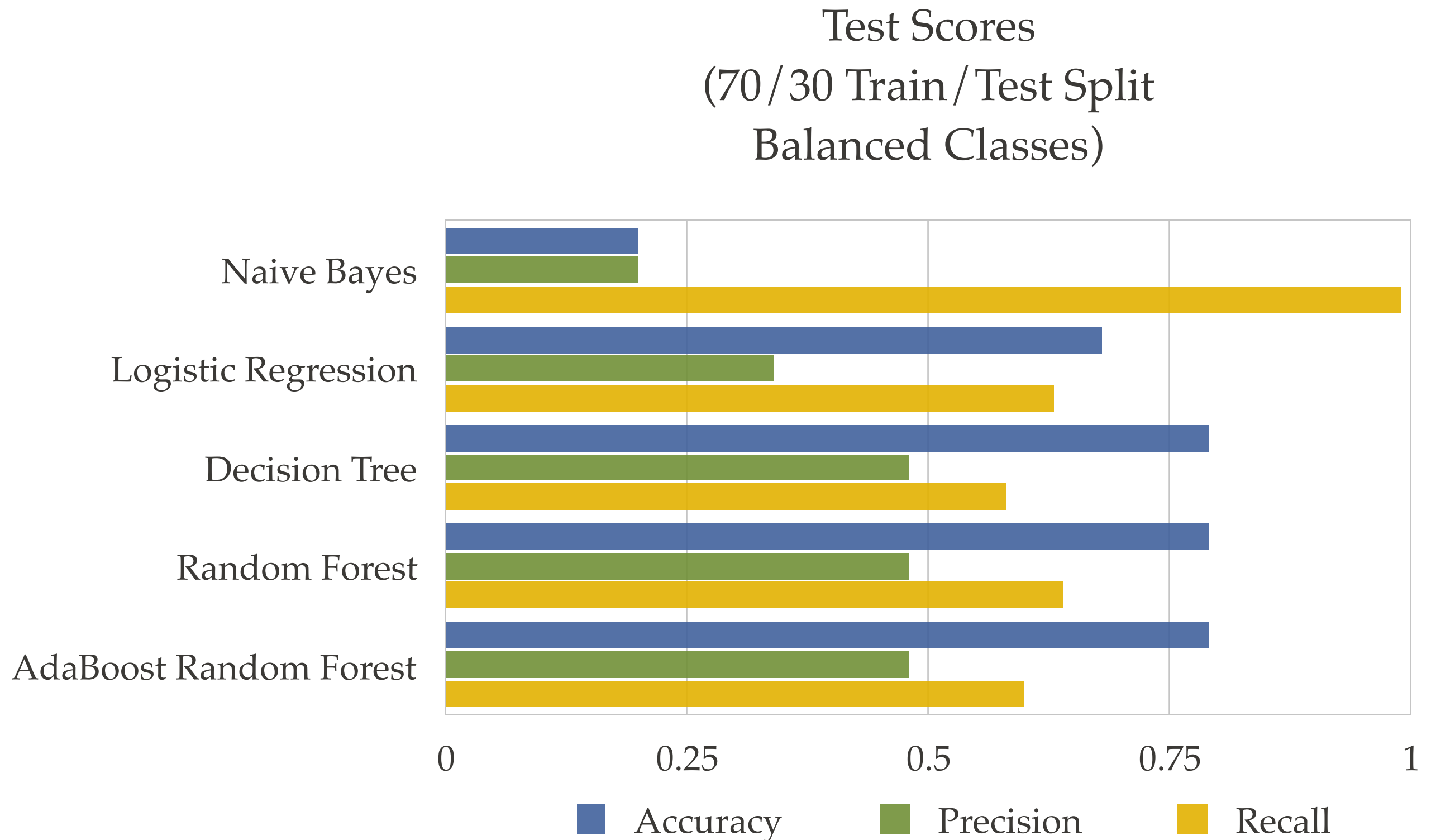
- ❖ Selecting Metric
  - ❖ **Accuracy:** Probability that our predictions are correct.
  - ❖ **Precision:** Probability that our predicted delays are actually delays.
  - ❖ **Recall:** Probability that all actual delays are predicted.
- ❖ Benefits of Higher Recall Weighting:
  - ❖ Higher awareness of potential delays for passengers
  - ❖ More warning to airline pilots / crew for potential delays

# Algorithm Selection





# Algorithm Selection



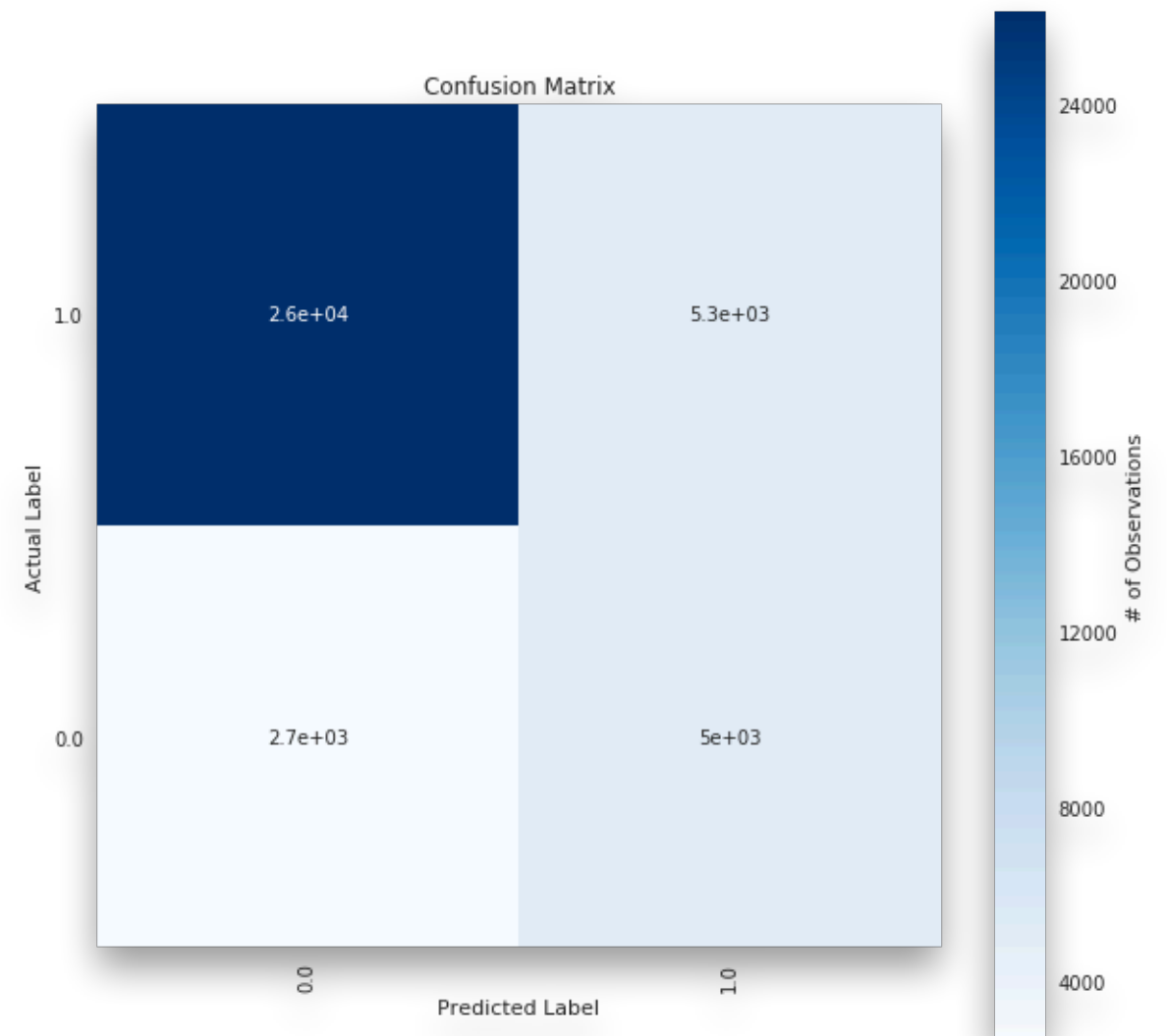
# Selected Features

- ❖ Airline Scheduling

- ❖ Departure / Arrival Times
- ❖ Flight Time
- ❖ Day of the Year / Week

- ❖ Weather Forecasting

- ❖ Precipitation (Origin and Destination)
- ❖ Average Wind Speed / Direction (Origin and Destination)
- ❖ Ambient Temperature



## Cross Validation Score (K = 7)

Accuracy	0.80
Precision	0.49
Recall	0.67

---

# Further Work

---

- ❖ Web Application
  - ❖ Connect to live current weather data for up-to-date predictions
  - ❖ Connect to weather forecasts for future predictions
- ❖ Big Data Processing / Distributed Computing
  - ❖ Manage training models over large datasets (All U.S. airports and airline carriers)
  - ❖ Provide hosting of web app

Thank You

*Questions?*