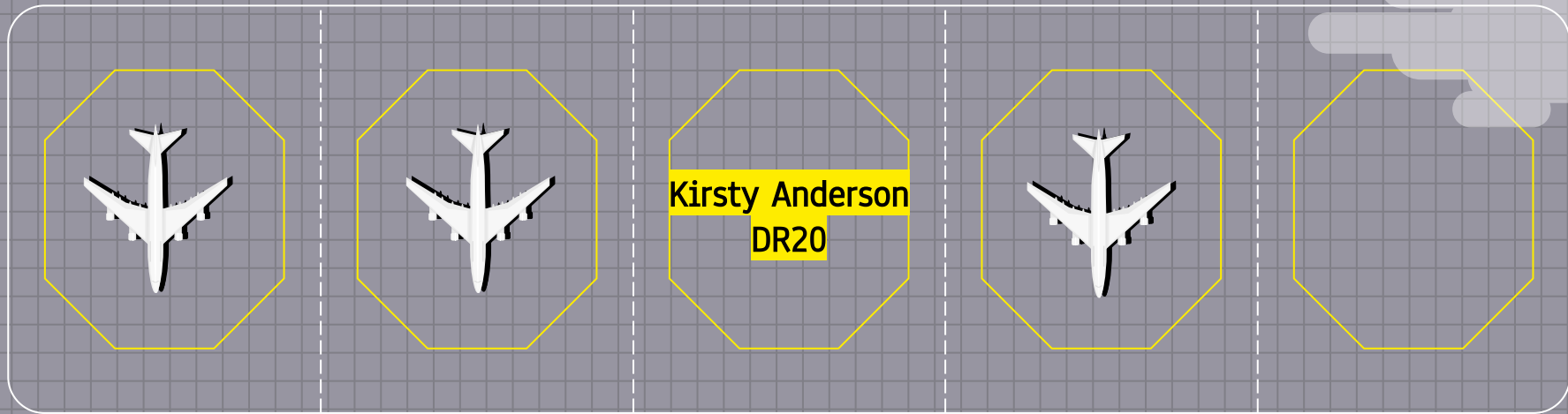


# Investigating Flight Delays at Newark Airport



# Business Questions



What type of weather causes delays?



How serious are weather delays?



What other factors cause delays?



# What data was used?

**01**

## Flight Information

Date of flight, route, delay times

**02**

## Weather Observations

Various weather metrics at a given time point across 3 airports

**03**

## Plane Information

Used to understand plane manufacturer, and age

## Airport Data

Airport codes and geospatial data for 3 Airports Investigated

**04**

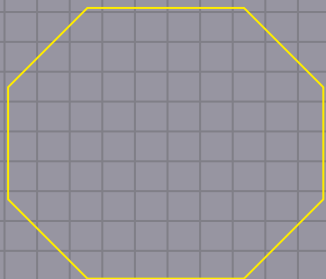
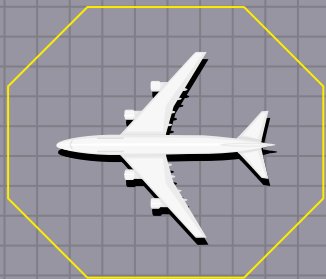
## Airline Codes

Airline codes relating to flight information

**05**



# Key Performance Indicators



- Length of delay

- The difference between the scheduled departure time and the actual departure time

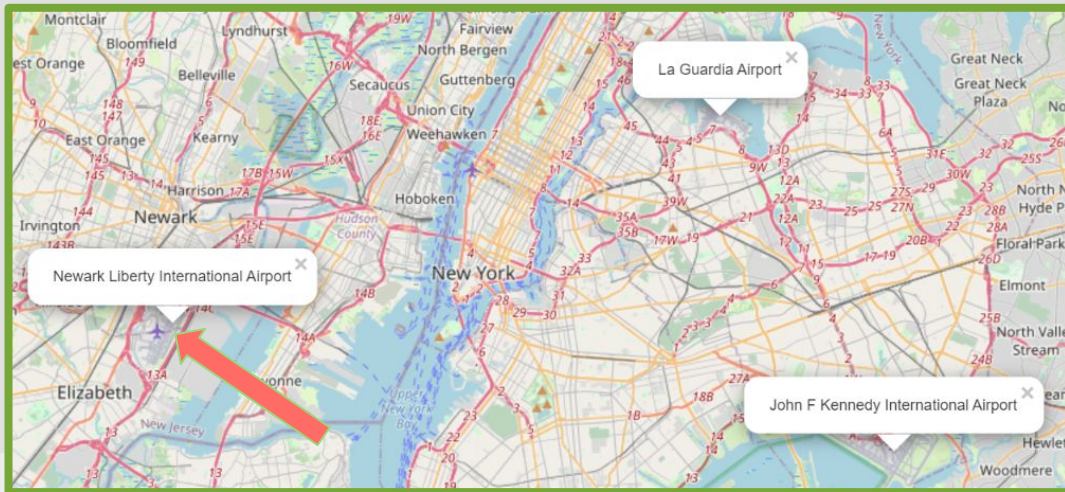
- Delayed Binary

- If the length of delay is  $>0$ , the flight is marked as delayed, otherwise marked as not delayed.

- Percentage of Delayed Flights

- The percentage of flights delayed within a specific group.

# Newark Airport



## Average Weather



61°F  
(16°C)



0.5mm  
(Slight)

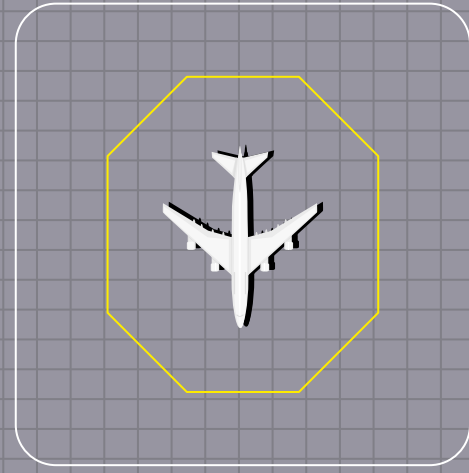


9 mph  
(Breeze)

**116,000** Flights in 2017

**01**

**What weather  
types cause  
delays?**



# Weather Types

01

## Humidity

% of moisture in the air

02

## Dewpoint

The temperature where the air would be saturated with water

03

## Temperature

°F

## Wind

Direction, Speed (mph)

04

## Pressure

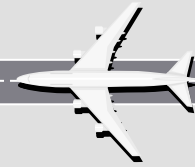
mBar

05

## Visibility

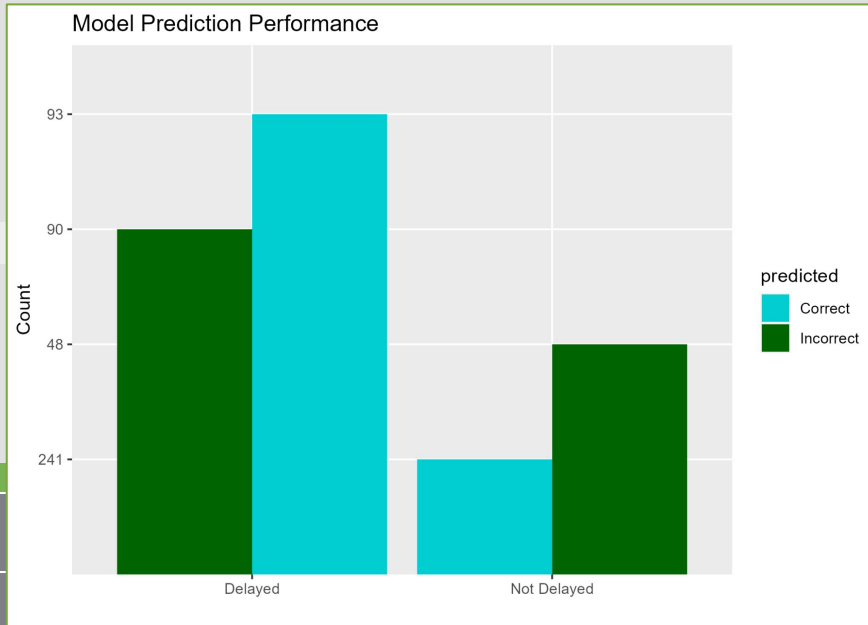
1 - 10 miles

06



# Predicting a Flight Delay from Weather

- Created a model using a technique called random forest to predict if flights will be delayed based on weather observations
- This helps to understand how these factors affect whether a flight is likely to be delayed or not
- The model reveals each weather variables' importance to the prediction
- Unfortunately the predictive ability of the model was poor, suggesting that there are other factors that impact the possibility of a flight delay

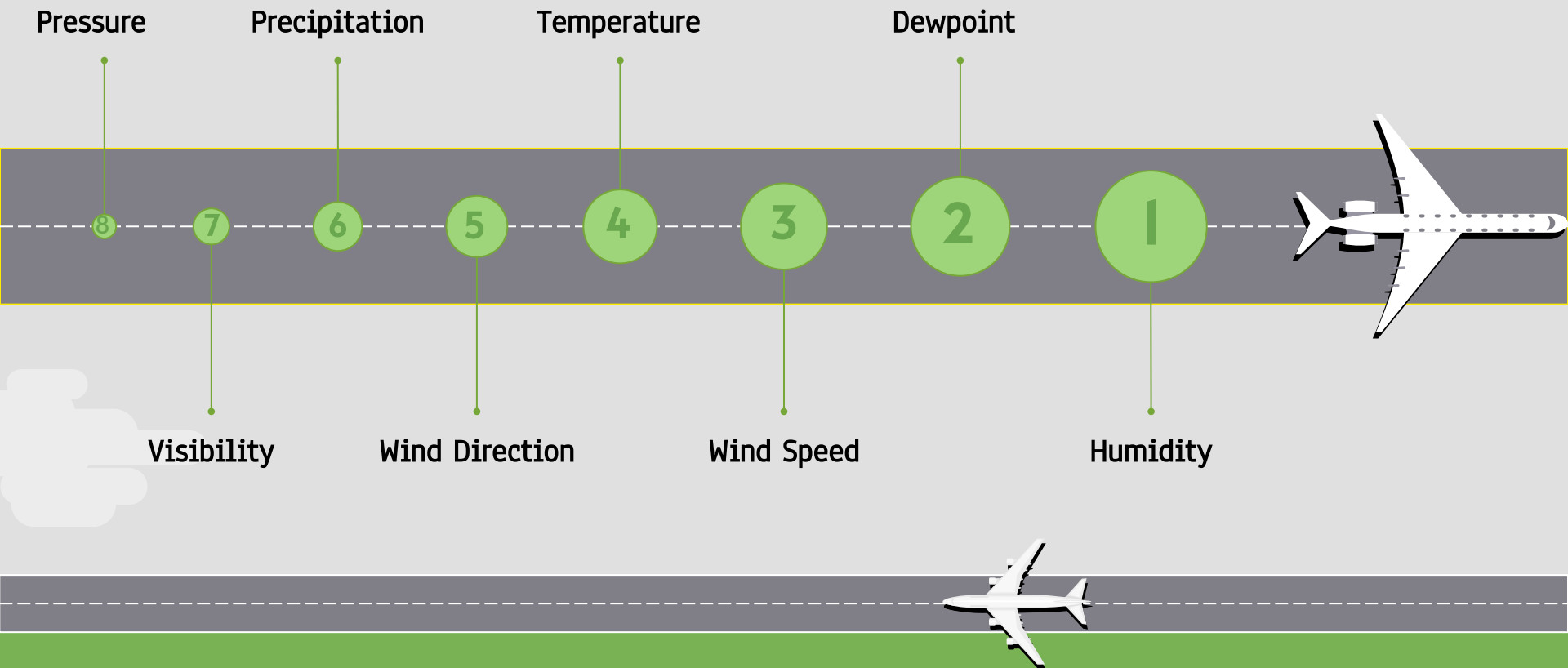


**AUC = 0.67**

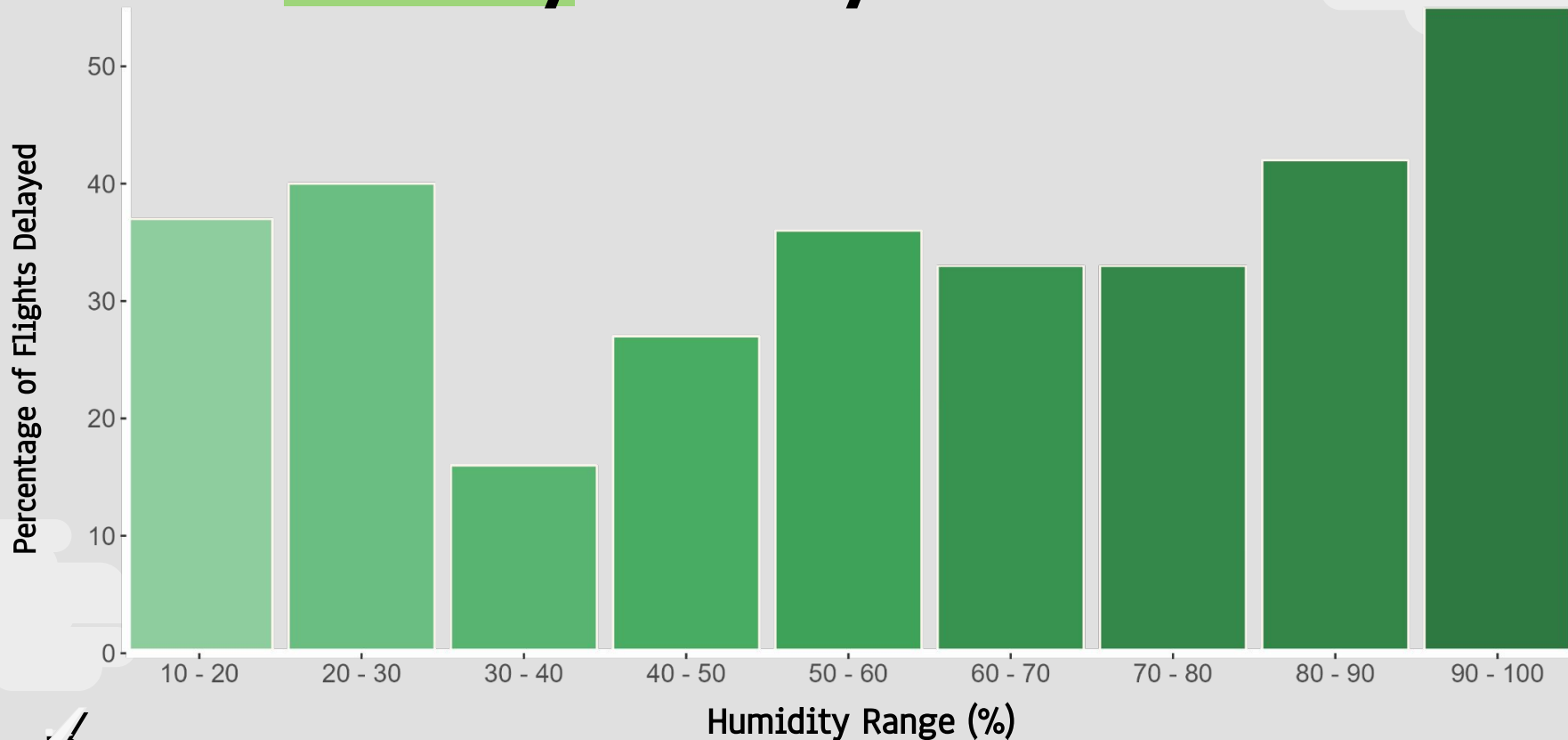
AUC is a measure of accuracy, the closer to 1 the better



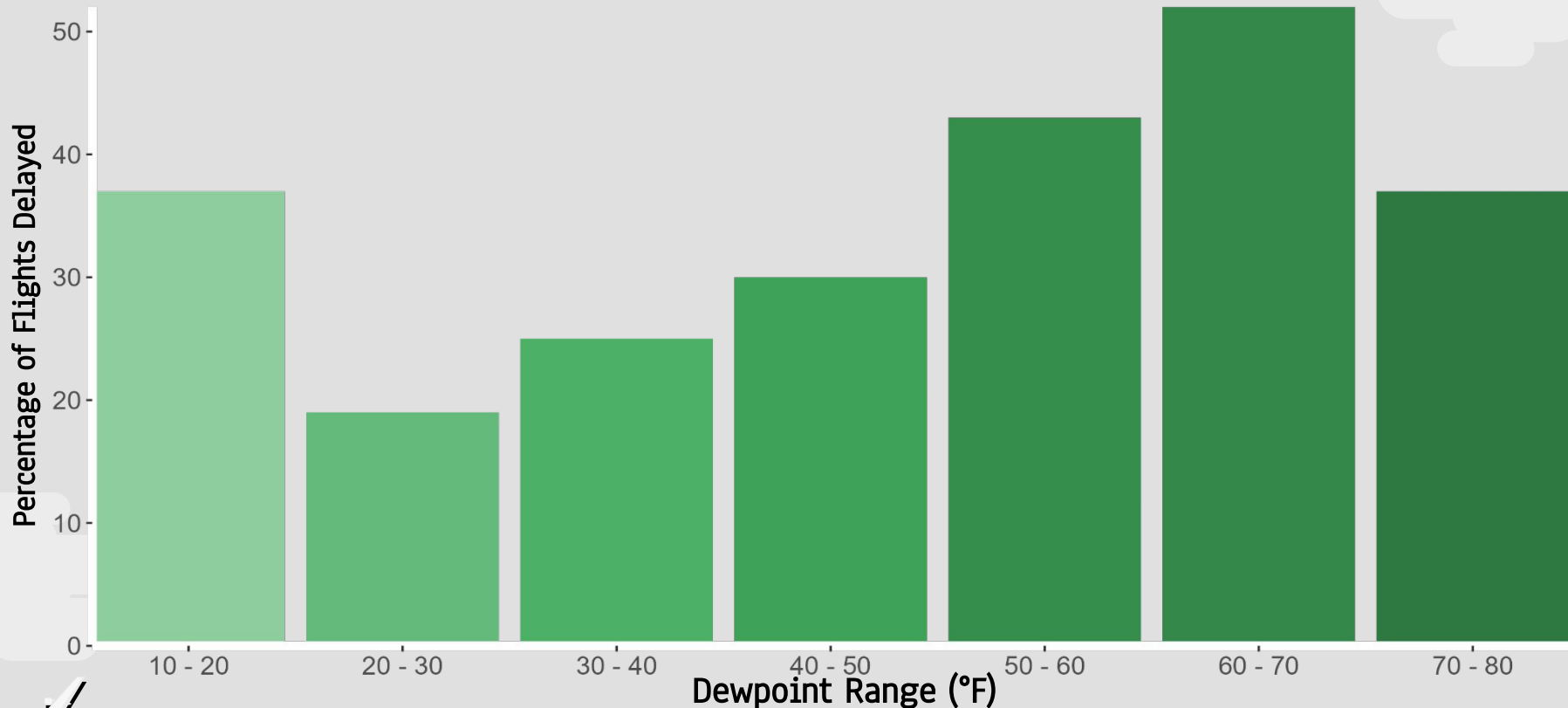
# Weather Impact on the Likelihood of Delay



# Effect of Humidity on Delay

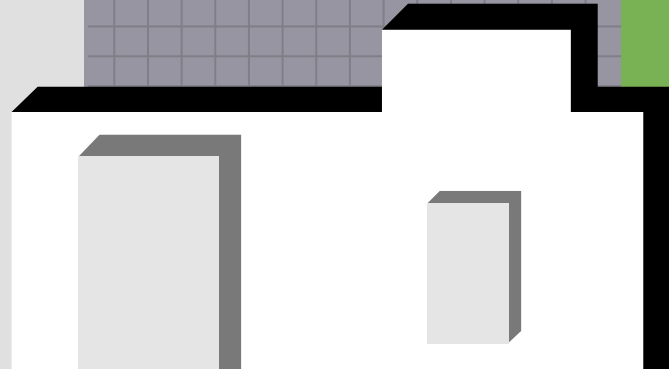
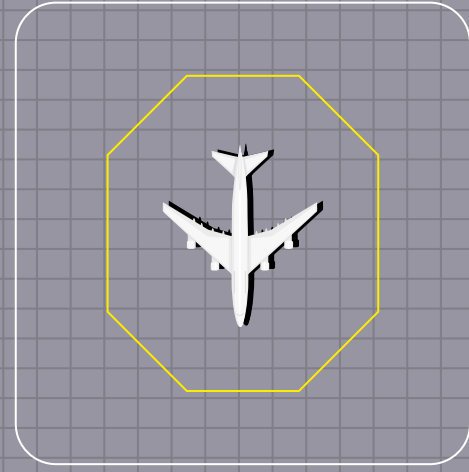


# Effect of Dewpoint on Delay



**02**

**How serious are  
weather related  
delays?**



# Predicting Length of Delay from Weather

- Created a **linear** regression predictive model
- The model selects weather variables to use as predictors,
- Variables are selected based on their connection with delay length and their connection with the other variables
- A higher connection = higher impact on length of delay
- Findings:



**Dewpoint**



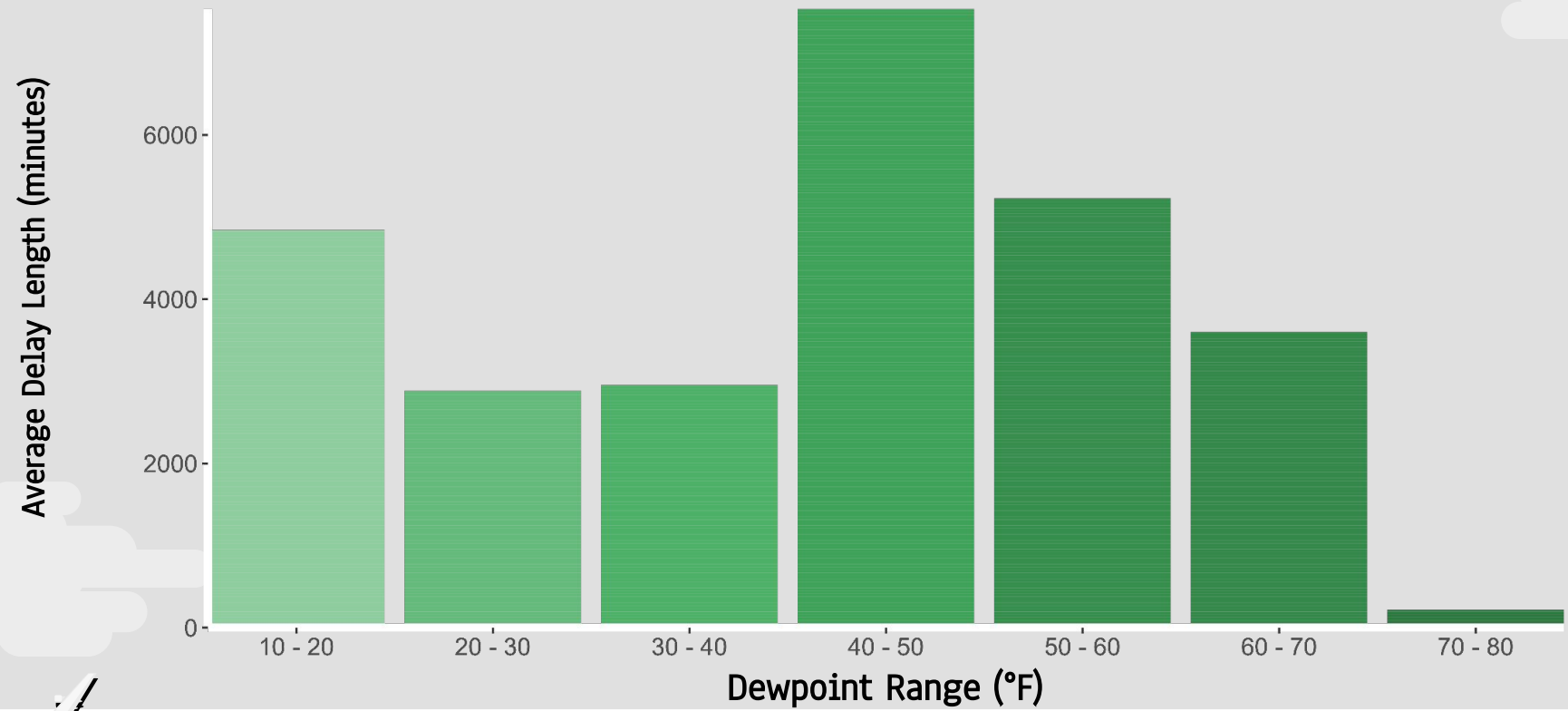
**Wind Speed**



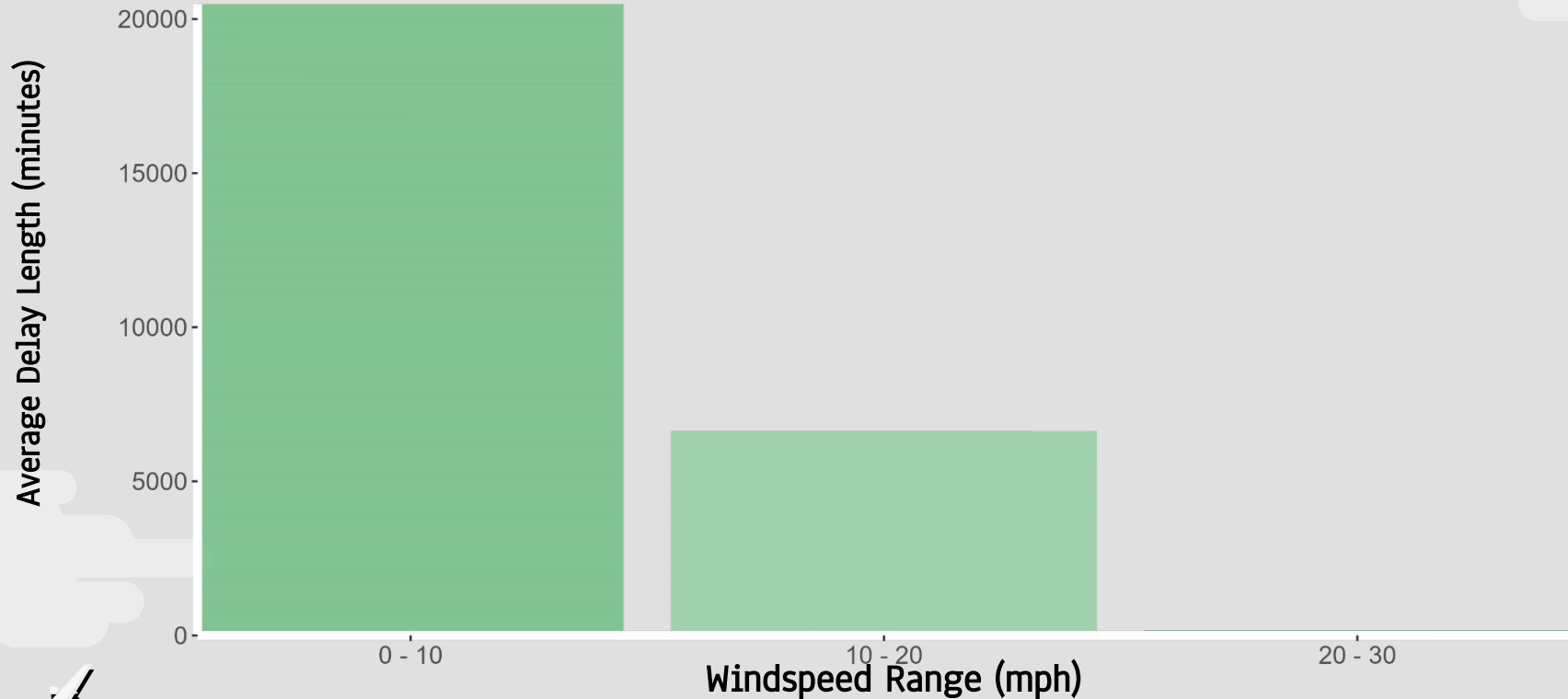
**Precipitation**



# Effect of Dewpoint on Delay Length

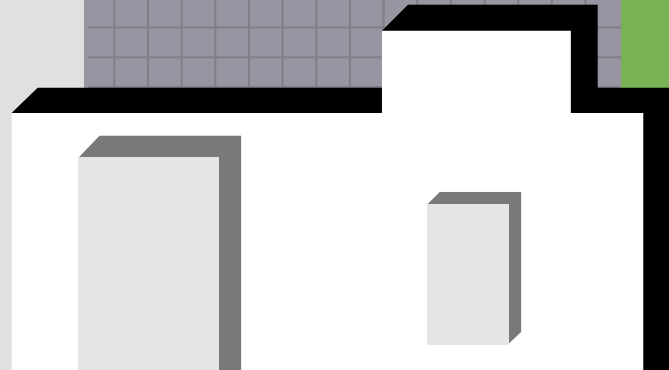
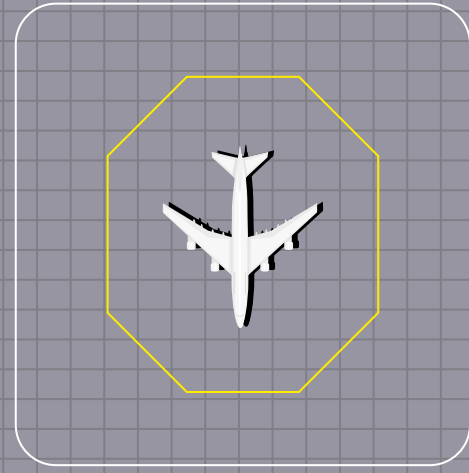


# Effect of Wind Speed on Delay Length



**03**

**What other  
factors cause  
delays?**





# Other Details

01

## Plane Details

Plane age and manufacturer

02

## Time Details

Month, day, hour of flight

## Airline

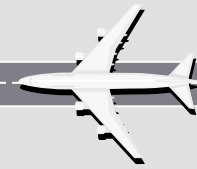
Airline “carrier” of given flight

03

## Route

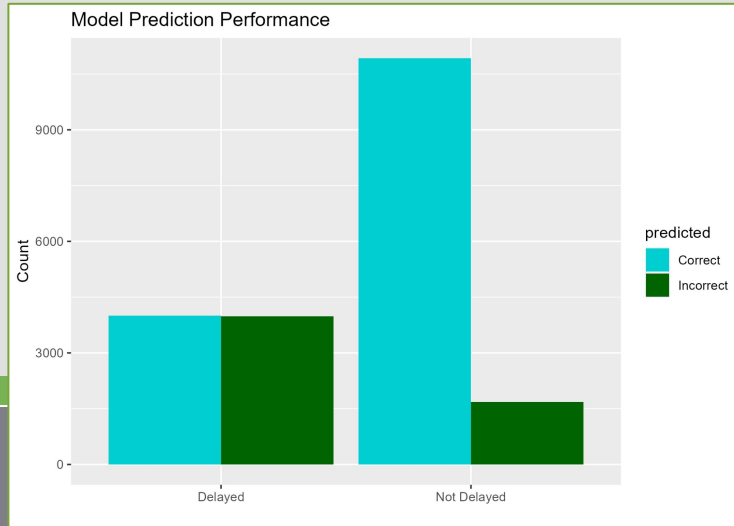
Origin and destination, and calculated distance

04



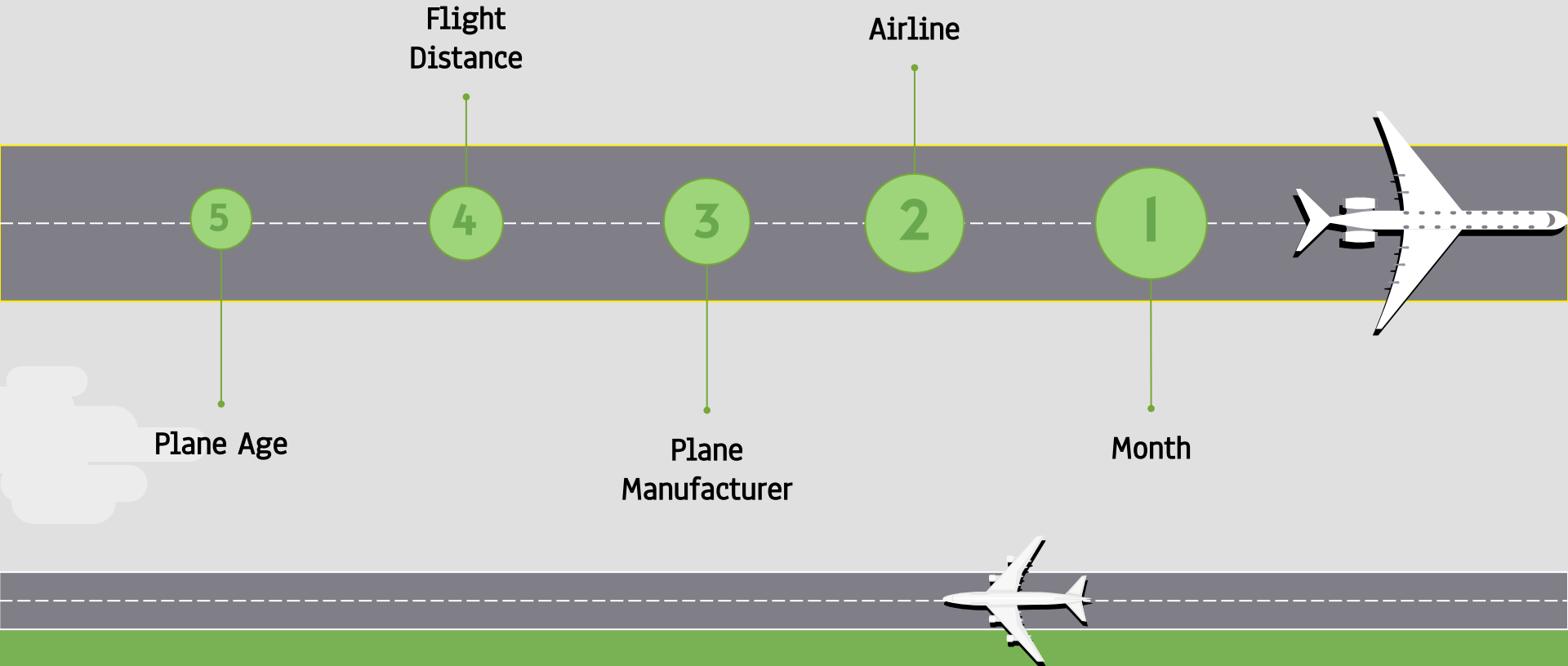
# Predicting a Flight Delay from Other Details

- Similar to weather, using a logical regression predictive model using a random forest technique to analyse the relationship between delayed flights and general observations
- The model reveals each variables' importance to the prediction
- Unfortunately the predictive ability of the model was poor, suggesting that there are other factors that impact the possibility of a flight delay

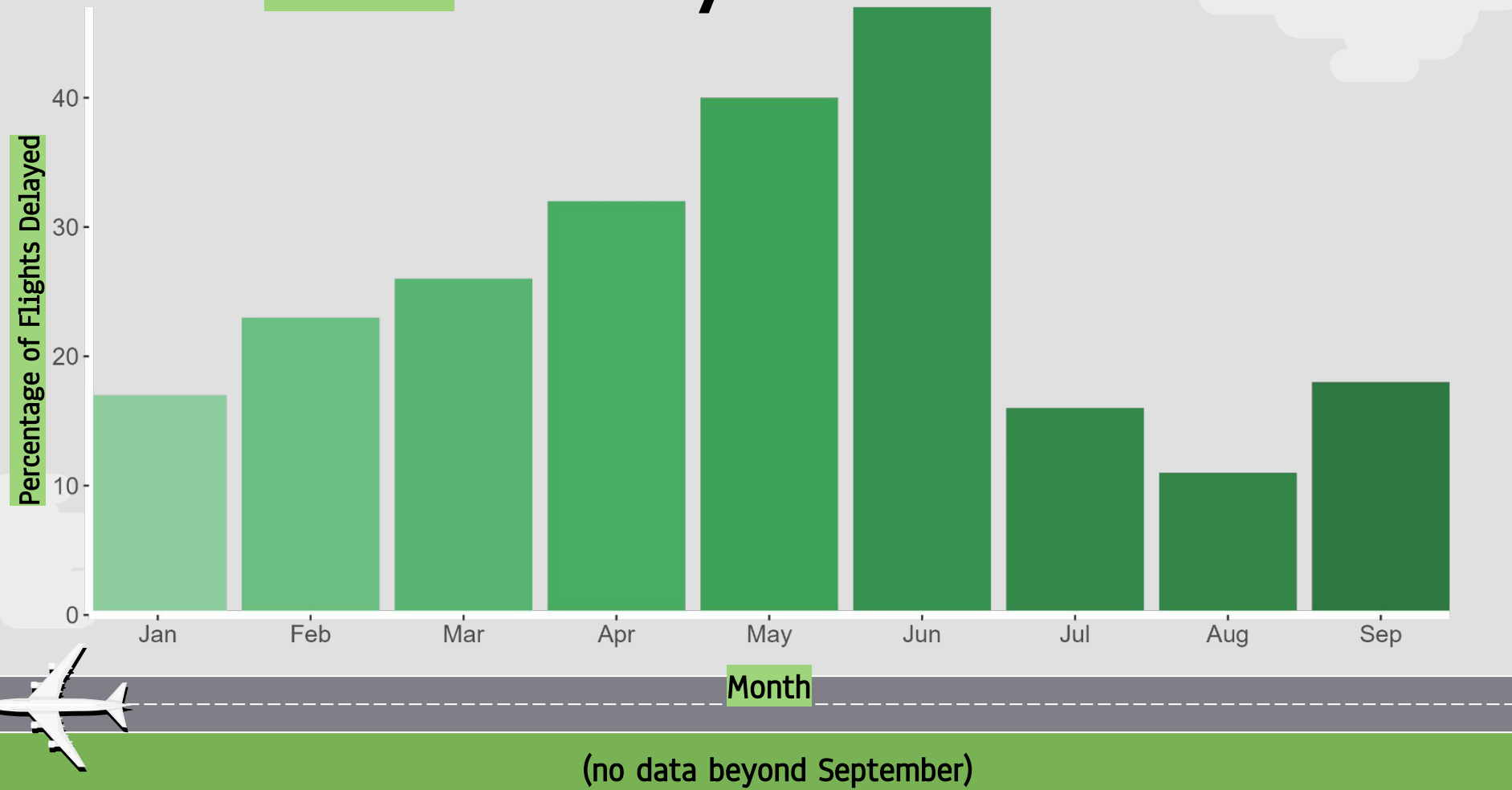


**AUC = 0.68**

# Other Impacts on the Likelihood of Delay

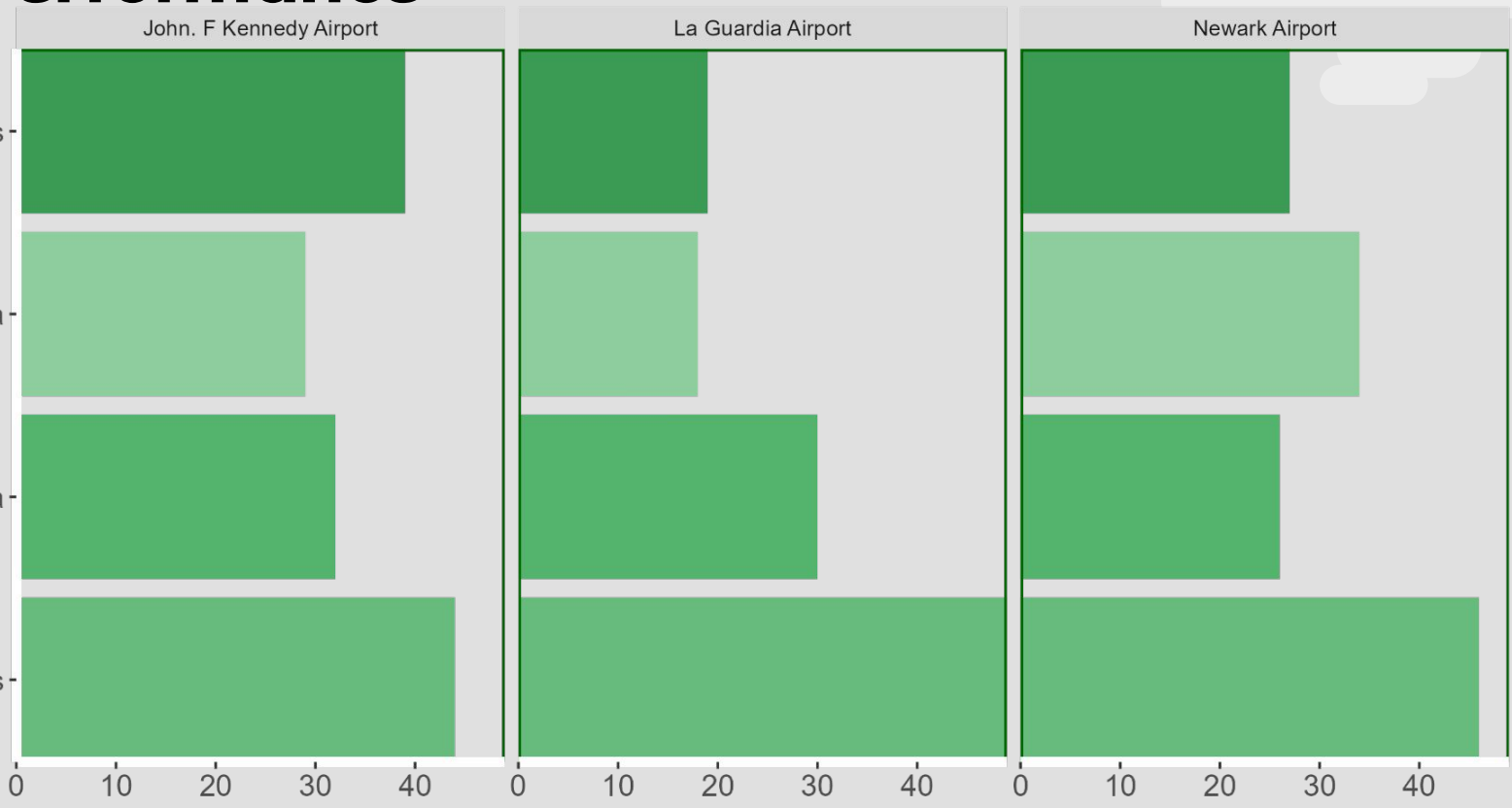


# Effect of Month on Delay



# Airline Performance

Airline

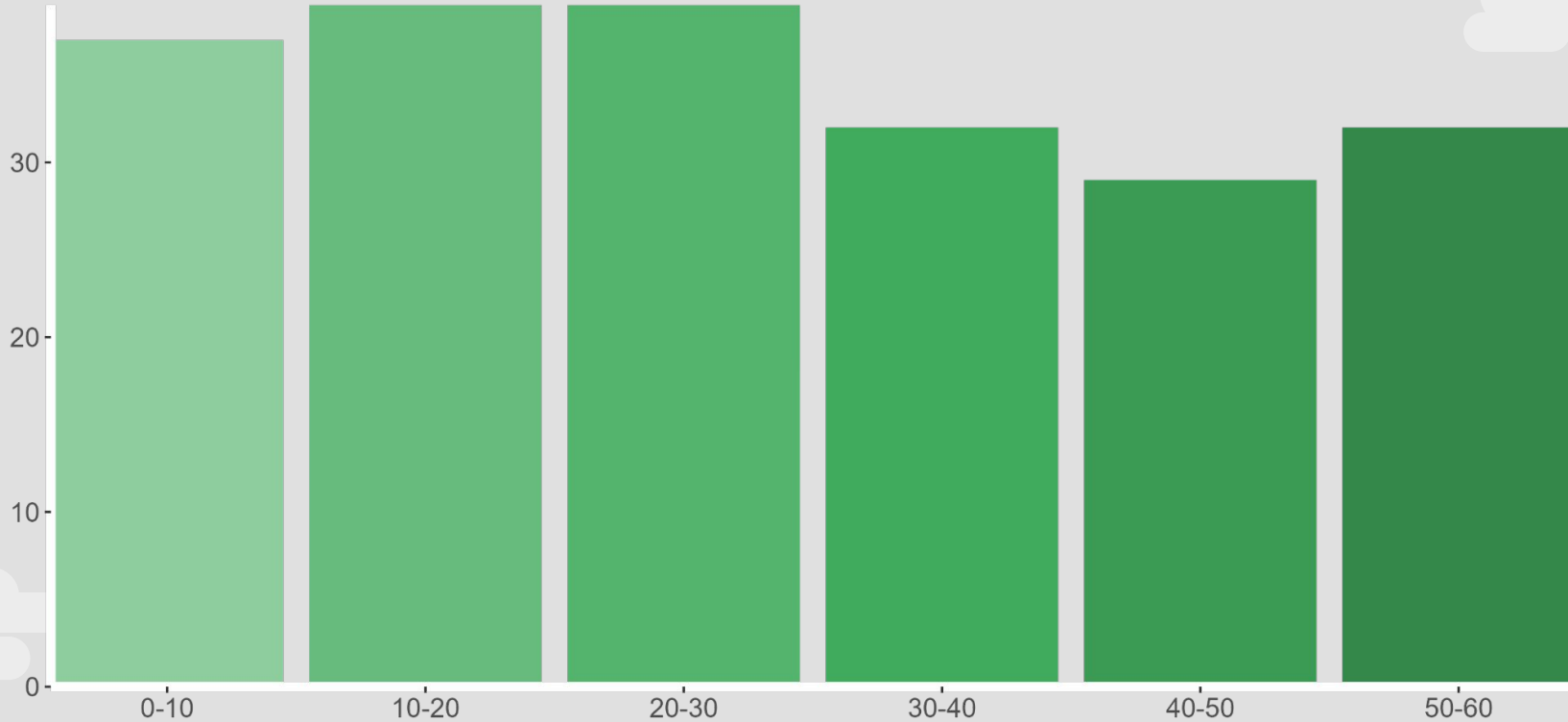


Percentage of Flights Delayed



# Plane Age

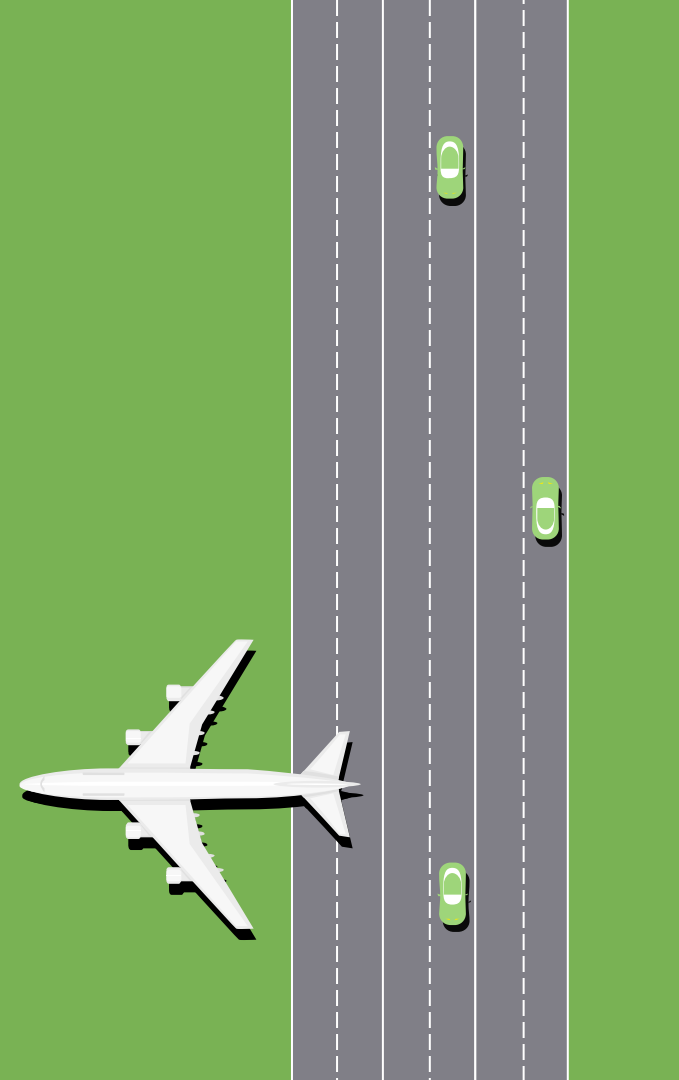
Percentage of Flights Delays



Plane Age (in 2017)

# Conclusions

- We need more data!
  - The predictive models show that we cannot accurately predict a flight delay from weather alone
- There is a relationship between weather and delay
  - Humidity and dewpoint are the key players
- Other factors are also impacting delay
  - Particularly month of departure, airline and plane manufacturer.



# Topics to Return To



**Humidity &  
Dewpoint**



**Collect More  
Data**



**Overall  
Model**



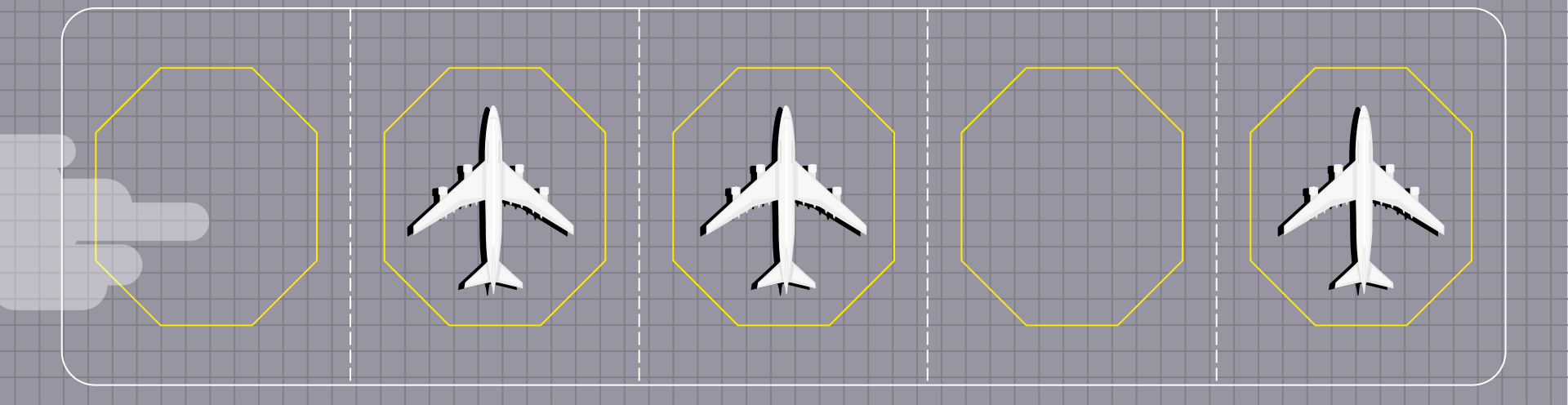
**Deeper Dive**



20







# Thankyou!

Any questions?



