



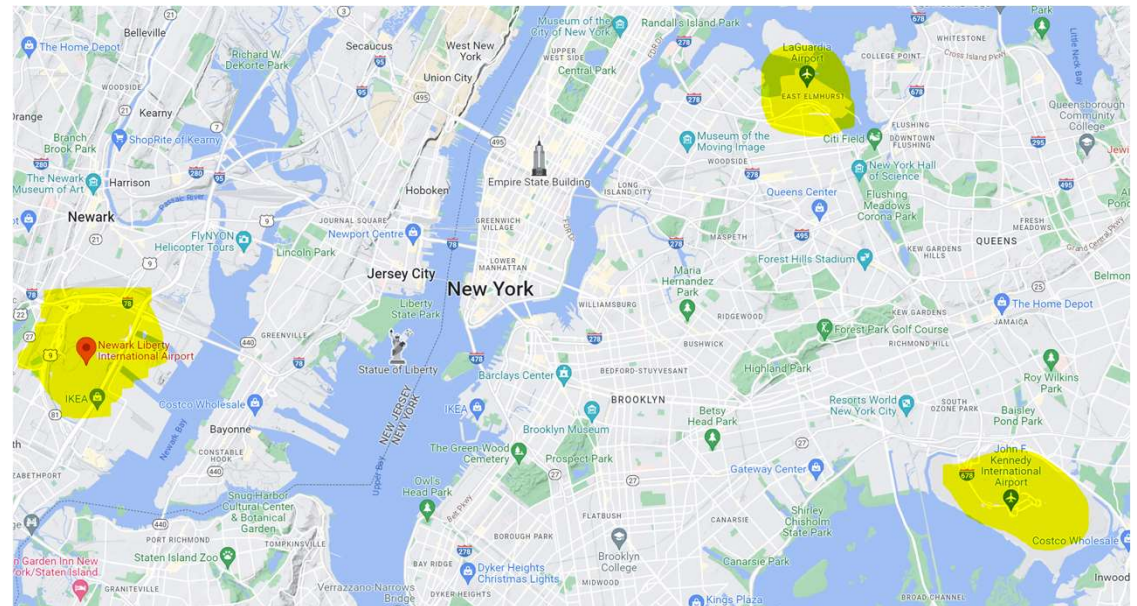
# NEWARK AIRPORT (EWR) FLIGHT DELAYS

MANDIP FARMAHAN (DR20)

21 JUNE 2023

# NEW YORK AND NEW JERSEY AIRPORTS

- 3 main airports
  - Newark Liberty International Airport (EWR)
  - John F. Kennedy International Airport (JFK)
  - LaGuardia Airport (LGA)
- Sources of delay
  - Weather
  - Runway slots
  - Flight paths



# PROJECT REMIT

“You have been hired by Newark airport to investigate the effect of weather on aeroplane departure delays\*. They want to invest in improving facilities so that aircraft can take off in more types of weather. As part of investigating the effect of weather you should investigate other factors to understand how important weather is in comparison to them.”

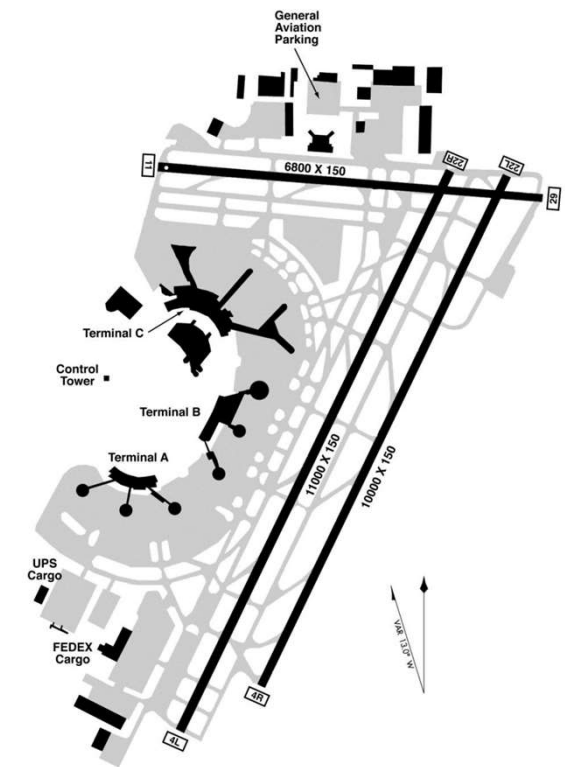
- The scope of this project will therefore focus on the following:
  - How many flights are affected and is this different from the neighbouring airports (JFK and LGA)?
  - Are these delays weather related and if so, what elements are causing these delays?
  - If not weather related, what other factors are causing these delays?
  - Binary classification of delayed vs not delayed

\* U.S. Department of Transportation Federal Aviation Administration (FAA) Order JO 7210.55F states the following:

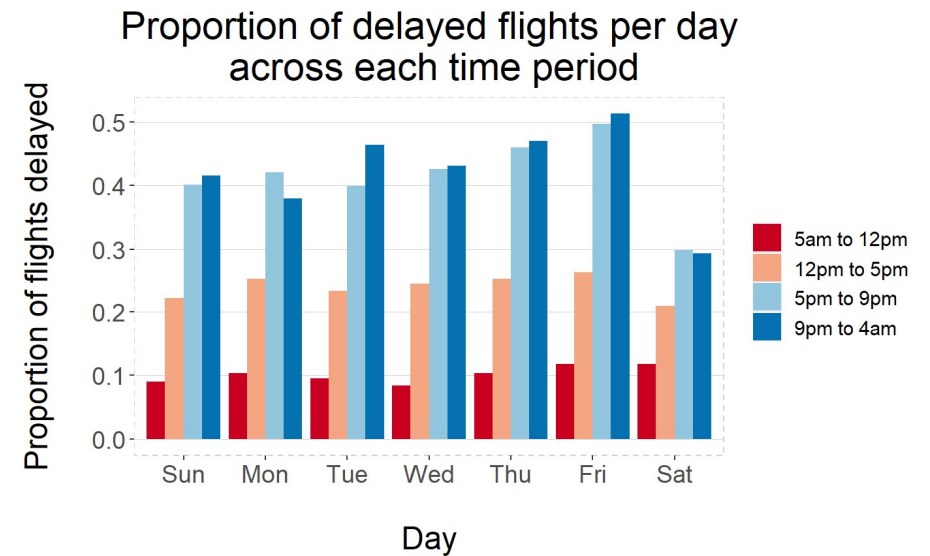
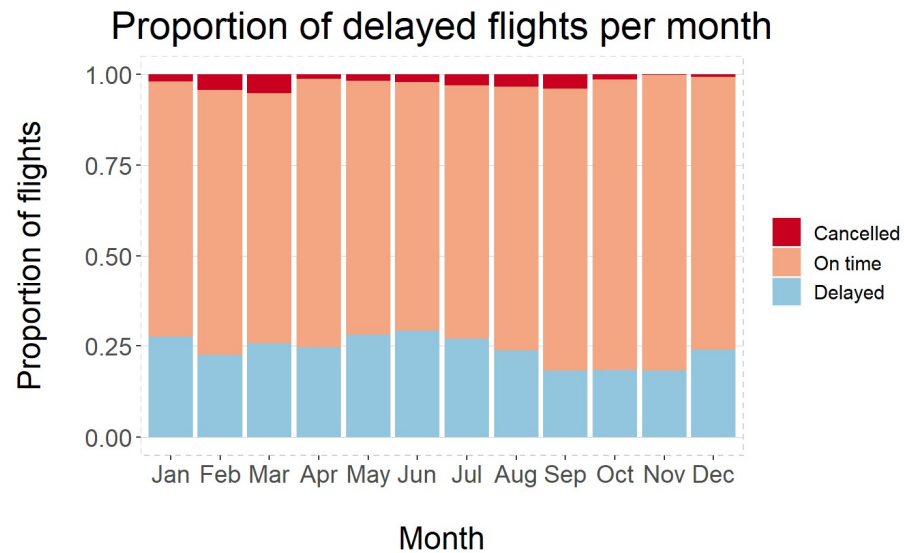
"Delays to instrument flight rules (IFR) traffic of 15 minutes or more, which result from the ATC system detaining an aircraft at the gate, short of the runway, on the runway, on a taxiway, or in a holding configuration anywhere enroute must be reported."

# EWR AIRPORT BACKGROUND

- Opened in 1928
- 3 terminals and 3 runways
  - 11/29 is used by smaller aircraft or when there are strong crosswinds on the two main runways
  - 4R/22L (opened 1952) is used primarily for landings
  - 4L/22R (opened 1970) is used primarily for departures
  - Port Authority completed a \$97 million dollar rehabilitation project of 4L/22R in 2014
- EWR was previously a FAA-designated Level 3 (slot-coordinated) airport
- In 2016, FAA re-designated EWR a Level 2 (schedule facilitated) airport
- 2017 annual departure metrics (proportions):
  - On time – 0.734
  - Delayed – 0.241
  - Cancelled – 0.025



# EWR AIRPORT DELAYS



- Is there a holiday effect?

$$H_0 : \pi_{\text{holiday\_delays}} - \pi_{\text{not\_holiday\_delays}} \leq 0$$

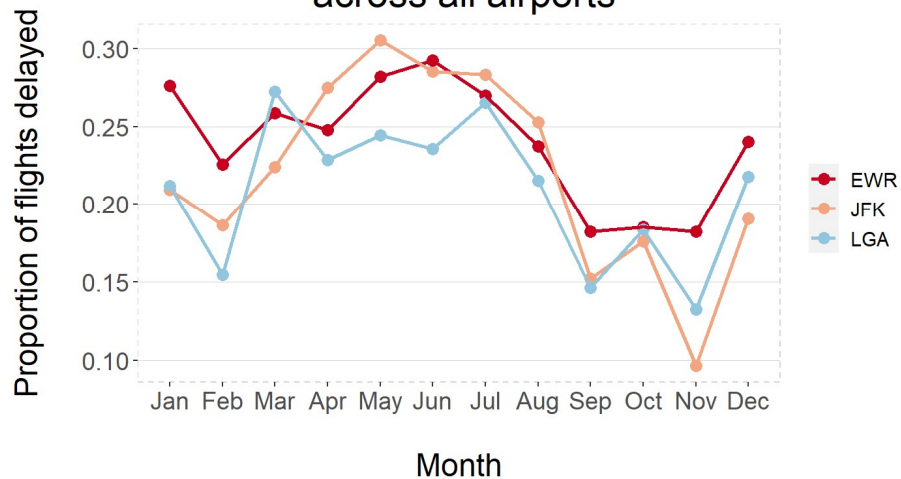
$$H_a : \pi_{\text{holiday\_delays}} - \pi_{\text{not\_holiday\_delays}} > 0$$

- There are significantly more delays during the holiday season

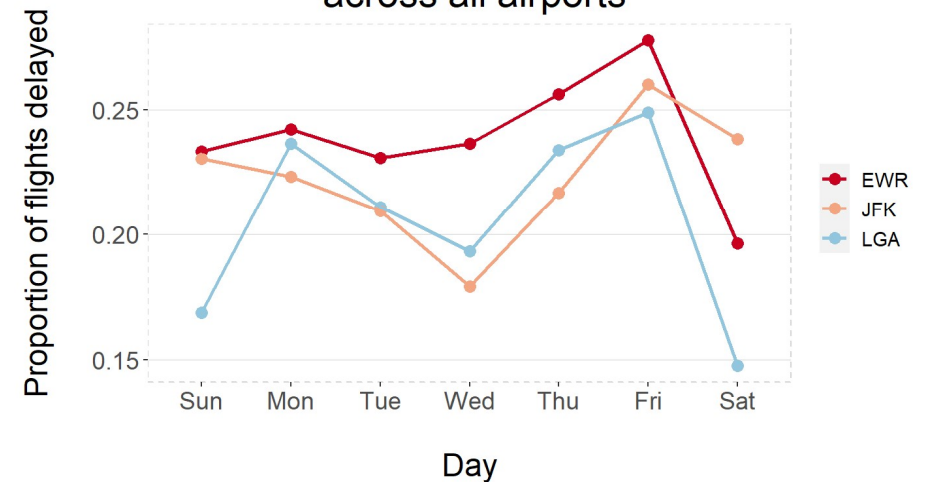


# COMPARISON OF AIRPORT DELAYS

Proportion of delayed flights per month across all airports



Proportion of delayed flights per day across all airports



- Is there an airport effect?

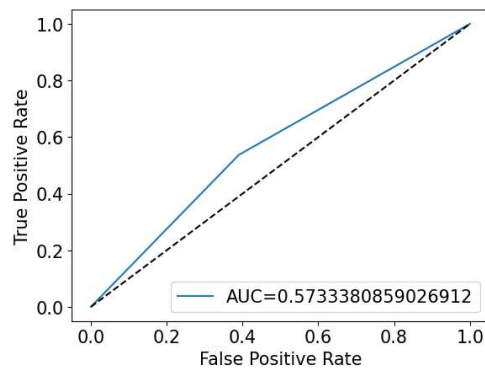
$$H_0 : \pi_{\text{newark\_delays}} - \pi_{\text{other\_delays}} \leq 0$$

$$H_a : \pi_{\text{newark\_delays}} - \pi_{\text{other\_delays}} > 0$$

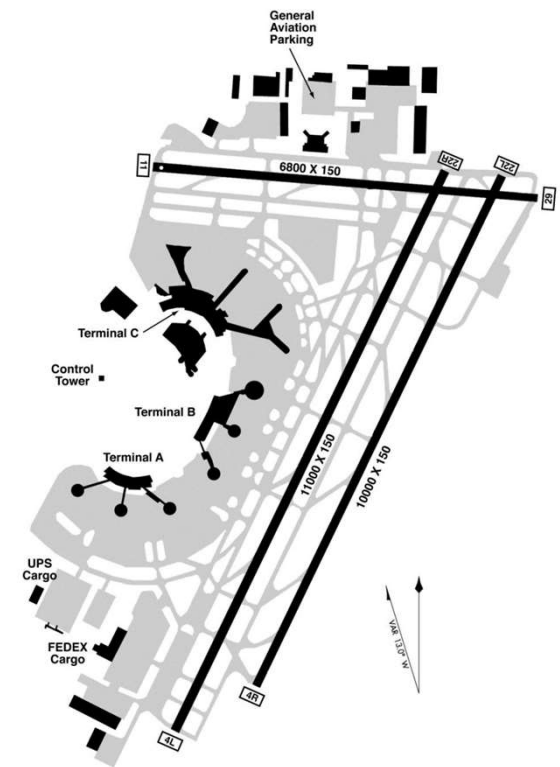
- Proportion of delays at EWR are significantly higher than neighbouring airports

# WEATHER RELATED DELAY FACTORS

- Original data:
  - Wind direction (wind\_dir), wind speed (wind\_speed) and visibility (visib)
- Alternative data sourced (NOAA and Meteostat):
  - Temperature (tavg), precipitation (prcp), snow fall (snow) and snow depth (snwd)
  - Air pressure (pres)
- Random Forest Classifier Model built in Python after model and parameter tuning

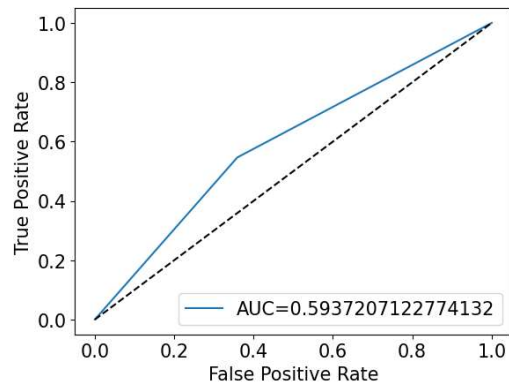


Variable	Importance
prcp	0.36
pres	0.22
visib	0.18
tavg	0.11
snwd	0.07
wind_dir	0.04
wind_speed	0.02
snow	0



# OTHER DELAY FACTORS

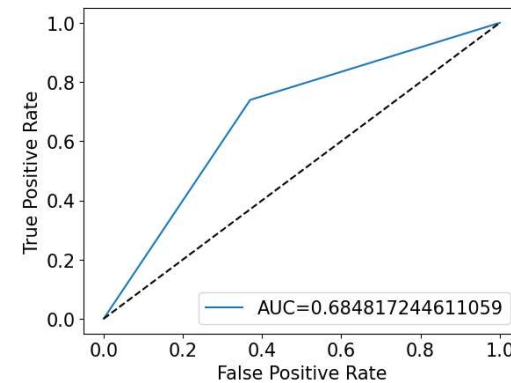
- Weather
- Plane and flight related factors
  - Destination, flight distance, flight direction, carrier, type of plane, seats



Variable	Importance
prcp	0.2
pres	0.11
visib	0.1
tavg	0.07
carrier_JetBlue	0.07
snwd	0.06
snow	0.04
seats	0.04
wind_dir	0.03
distance	0.03

Only top 10 shown

- Weather
- Plane and flight related factors
  - Destination, flight distance, flight direction, carrier, type of plane, seats
- Airport traffic
  - Month, day of week, hour, cycle, flights per day, flights per hour



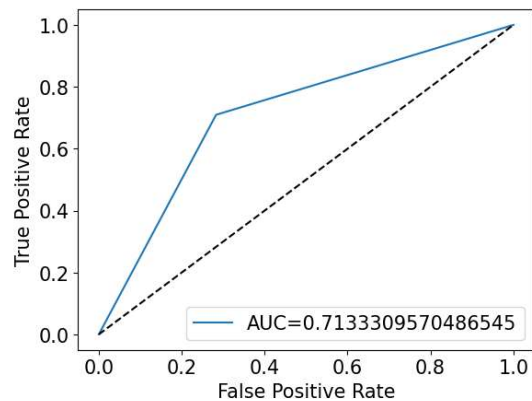
Variable	Importance
hour	0.24
cycle_Morning	0.22
cycle_Evening	0.12
prcp	0.06
pres	0.05
cycle_Night	0.05
month	0.03
flights_per_day	0.03
flights_per_hour	0.03
visib	0.02

Only top 10 shown



# FINAL MODEL

- Final predictors:
  - Weather (4): precipitation (prcp), air pressure (pres), temperature (tavg), visibility (visib)
  - Airport traffic (4): time of day (hour), flights per day, flights per hour, month



Variable	Importance
hour	0.59
flights_per_day	0.11
prcp	0.09
pres	0.07
tavg	0.06
month	0.04
flights_per_hour	0.04
visib	0.01

EWR model with JFK data

- AUC = 0.640

EWR model with LGA data

- AUC = 0.658

# CONCLUSIONS AND RECOMMENDATIONS

- 24% of flights from EWR in 2017 had departure delays
- Weather alone is not responsible for departure delays
  - Weather related airport improvements unlikely to cause a decrease in departure delays
- Time of flight (59%) and number flights per day are most important
  - 40% of flights are in the morning and only 5% at night
  - Flights should be more evenly spread across the day or reduce total number of flights
- Delays in general are worse at EWR than LGA and JFK
  - More flights leave from EWR
  - Final model indicates flights from all NY and NJ airports are affected by similar factors
  - To fully investigate the influence of wind speed and direction, terminal-specific data is required

THANK YOU

