

# Finding Flight Delay Trends

DAT500 Project Group 17

---

Bhakti Prabhu & Stephan F. W. Brandasu

Faculty of Science and Technology

# Why?

## the use case

As a traveller:

1. Not miss any important meetings/events/functions.
2. Pre-plan journey
3. Have idea of buffer time while flight booking

As an airline:

1. know when to increase workforce.
2. Opportunity to improve over competitors.

# What?

## the dataset

Sourced from the United States Bureau of Transportation statistics. The dataset is a total of 5GB for 3 years of data

```
1 FL_DATE 1/1/2022 12:00:00 AM
2 OP_UNIQUE_CARRIER 9E
3 OP_CARRIER_AIRLINE_ID 20363
4 OP_CARRIER 9E
5 TAIL_NUM N138EV
6 OP_CARRIER_FL_NUM 4732
7 ORIGIN_AIRPORT_ID 10135
8 ORIGIN_AIRPORT_SEQ_ID 1013506
9 ORIGIN_CITY_MARKET_ID 30135
10 ORIGIN ABE
11 DEST_AIRPORT_ID 11433
12 DEST_AIRPORT_SEQ_ID 1143302
13 DEST_CITY_MARKET_ID 31295
14 DEST DTW
15 ARR_DELAY -15.00
16 ARR_DELAY_NEW 0.00
17 CANCELLED 0.00
18 CARRIER_DELAY
19 WEATHER_DELAY
20 NAS_DELAY
21 SECURITY_DELAY
22 LATE_AIRCRAFT_DELAY
```

Note: delay is measured in minutes and cancelled is 0 for no and 1 for yes

Note: some fields don't always contain data, this was just the first entry from the dataset

# How?

## high level plan

1. Finding High probability of flight delays in a period
  - Which days in a month
  - Which month in a year
  - Which quarter in a year
2. Find Trend in Flight Delays from year to year