

# PREDICTING AIRLINE DELAYS AT DIA

WHICH AIRLINE SHOULD I FLY?

# PROBLEM: CAN FLIGHT DELAYS BE PREDICTED?

THE PURPOSE OF THIS STUDY IS TO EXPLORE AND ANALYZE FLIGHT INFORMATION AND DETERMINE WHICH FACTORS INFLUENCE FLIGHT DELAYS AND USE STATISTICAL ANALYSIS TO DETERMINE ANY POTENTIAL CONSUMER RECOMMENDATIONS OR BUSINESS STRATEGIES

**DATA: BUREAU OF TRANSPORTATION STATISTICS**

**DATA DOMAIN: FLIGHT DELAYS**

**DATASET: ON-TIME PERFORMANCE COUNTED BY  
MINUTES, DELAY CLASSIFIED BY CAUSE, BY AIRLINE**

# FEATURES

```
['YEAR', 'MONTH', 'CARRIER', 'CARRIER_NAME', 'AIRPORT', 'AIRPORT_NAME',
'ARR_FLIGHTS', 'ARR_DEL15', 'CARRIER_CT', 'WEATHER_CT', 'NAS_CT',
'SECURITY_CT', 'LATE_AIRCRAFT_CT', 'ARR_CANCELLED', 'ARR_DIVERTED',
'ARR_DELAY', 'CARRIER_DELAY', 'WEATHER_DELAY', 'NAS_DELAY',
'SECURITY_DELAY', 'LATE_AIRCRAFT_DELAY']
```

# WHAT DATA DO WE NEED TO PREDICT FLIGHT DELAYS?

- HISTORICAL DATA
- ORGANIZED BY AIRLINE
- LIMITED BY AIRPORT
- REASONS FOR DELAY
- COUNT OF DELAY (>15 MINUTES)

# CUSTOMER CONCERNS

WHERE ARE YOU GOING?

- WHAT DESTINATIONS HAVE FREQUENT DELAYS, AND THE LONGEST DELAYS?

WHAT TIME OF DAY ARE YOU FLYING?

- WHAT FLIGHT TIMES ARE DELAYED MOST OFTEN, AND FOR THE LONGEST TIME?

WHAT AIRLINE ARE YOU FLYING?

- WHAT AIRLINES HAVE THE MOST FREQUENT AND LONGEST DELAYS?

# DATA SCIENCE CONCERNS

IS THERE A RELATIONSHIP BETWEEN DELAYS AND FLIGHT TIME?

IS THERE A RELATIONSHIP BETWEEN DELAYS AND AIRLINE?

IS THERE A RELATIONSHIP BETWEEN DELAYS AND DESTINATION?

**CAN WE CLASSIFY THE DELAYED FEATURE INTO DELAYED AND NOT DELAYED?**

DEFINE SUCCESS:  
INFORMATION HELPS THE CUSTOMER CHOOSE  
CARRIER, BASED ON DESTINATION AND TIME OF FLIGHT,  
WITH A REASONABLE IDEA OF EXPECTED DELAYS

I AM FLYING DELTA FROM DIA TO JFK AT NOON, WILL MY FLIGHT BE DELAYED? YES | NO

- IF YES, HOW LONG?

WHICH AIRLINE SHOULD I FLY ON CHRISTMAS EVE FROM DENVER TO NEW YORK

- SHOULD I FLY IN THE MORNING OR AT NIGHT?

- IF IT SNOWS, WILL I BE DELAYED MORE THAN AN HOUR?