# Data Description

We can find the \*\*all\*\* information about specific attributes in this file.

### Table \*\*flights\*\*

Variables:

- \*\*fl\_date\*\*: Flight Date (yyyy-mm-dd) Do we need to change to yyyy, day of week, holiday

- \*\*mkt\_unique\_carrier\*\*: Unique Marketing Carrier Code. When the same code has been used by multiple carriers, a numeric suffix is used for earlier users, for example, PA, PA(1), PA(2). Use this field for analysis across a range of years.

- \*\*branded\_code\_share\*\*: Reporting Carrier Operated or Branded Code Share Partners

- \*\*mkt\_carrier\*\*: Code assigned by IATA and commonly used to identify a carrier. As the same code may have been assigned to different carriers over time, the code is not always unique. For analysis, use the Unique Carrier Code.

- \*\*mkt\_carrier\_fl\_num\*\*: Flight Number

- \*\*op\_unique\_carrier\*\*: Unique Scheduled Operating Carrier Code. When the same code has been used by multiple carriers, a numeric suffix is used for earlier users,for example, PA, PA(1), PA(2). Use this field for analysis across a range of years.

- \*\*tail\_num\*\*: Tail Number

- \*\*op\_carrier\_fl\_num\*\*: Flight Number

- \*\*origin\_airport\_id\*\*: Origin Airport, Airport ID. An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.

- \*\*origin\*\*: Origin Airport

- \*\*origin\_city\_name\*\*: Origin Airport, City Name

- \*\*dest\_airport\_id\*\*: Destination Airport, Airport ID. An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.

- \*\*dest\*\*: Destination Airport

- \*\*dest\_city\_name\*\*: Destination Airport, City Name

- \*\*crs\_dep\_time\*\*: CRS Departure Time (local time: hhmm)

- \*\*dep\_time\*\*: Actual Departure Time (local time: hhmm)

- \*\*dep\_delay\*\*: Difference in minutes between scheduled and actual departure time. Early departures show negative numbers.

- \*\*taxi\_out\*\*: Taxi Out Time, in Minutes

- \*\*wheels\_off\*\*: Wheels Off Time (local time: hhmm)

- \*\*wheels\_on\*\*: Wheels On Time (local time: hhmm)

- \*\*taxi\_in\*\*: Taxi In Time, in Minutes

- \*\*crs\_arr\_time\*\*: CRS Arrival Time (local time: hhmm)

- \*\*arr\_time\*\*: Actual Arrival Time (local time: hhmm)

- \*\*arr\_delay\*\*: Difference in minutes between scheduled and actual arrival time. Early arrivals show negative numbers.

- \*\*cancelled\*\*: Cancelled Flight Indicator (1=Yes)

- \*\*cancellation\_code\*\*: Specifies The Reason For Cancellation

- \*\*diverted\*\*: Diverted Flight Indicator (1=Yes)

- \*\*dup\*\*: Duplicate flag marked Y if the flight is swapped based on Form-3A data

- \*\*crs\_elapsed\_time\*\*: CRS Elapsed Time of Flight, in Minutes

- \*\*actual\_elapsed\_time\*\*: Elapsed Time of Flight, in Minutes

- \*\*air\_time\*\*: Flight Time, in Minutes

- \*\*flights\*\*: Number of Flights

- \*\*distance\*\*: Distance between airports (miles)

- \*\*carrier\_delay\*\*: Carrier Delay, in Minutes

- \*\*weather\_delay\*\*: Weather Delay, in Minutes

- \*\*nas\_delay\*\*: National Air System Delay, in Minutes

- \*\*security\_delay\*\*: Security Delay, in Minutes

- \*\*late\_aircraft\_delay\*\*: Late Aircraft Delay, in Minutes

- \*\*first\_dep\_time\*\*: First Gate Departure Time at Origin Airport

- \*\*total\_add\_gtime\*\*: Total Ground Time Away from Gate for Gate Return or Cancelled Flight

- \*\*longest\_add\_gtime\*\*: Longest Time Away from Gate for Gate Return or Cancelled Flight

### Table \*\*flights\_test\*\*

This table consists of subset of columns from table flights. It represents flights from January 2020 which will be used for evaluation. Therefore, we are missing some features that we are not suppossed to know before the flight lands.

Variables:

- \*\*fl\_date\*\*: Flight Date (yyyy-mm-dd) as only one month, consider day of week, holiday, month

- \*\*mkt\_unique\_carrier\*\*: Unique Marketing Carrier Code. When the same code has been used by multiple carriers, a numeric suffix is used for earlier users, for example, PA, PA(1), PA(2). Use this field for analysis across a range of years.

- \*\*branded\_code\_share\*\*: Reporting Carrier Operated or Branded Code Share Partners

- \*\*mkt\_carrier\*\*: Code assigned by IATA and commonly used to identify a carrier. As the same code may have been assigned to different carriers over time, the code is not always unique. For analysis, use the Unique Carrier Code.

- \*\*mkt\_carrier\_fl\_num\*\*: Flight Number

- \*\*op\_unique\_carrier\*\*: Unique Scheduled Operating Carrier Code. When the same code has been used by multiple carriers, a numeric suffix is used for earlier users,for example, PA, PA(1), PA(2). Use this field for analysis across a range of years.

- \*\*tail\_num\*\*: Tail Number

- \*\*op\_carrier\_fl\_num\*\*: Flight Number

- \*\*origin\_airport\_id\*\*: Origin Airport, Airport ID. An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.

- \*\*origin\*\*: Origin Airport

- \*\*origin\_city\_name\*\*: Origin Airport, City Name

- \*\*dest\_airport\_id\*\*: Destination Airport, Airport ID. An identification number assigned by US DOT to identify a unique airport. Use this field for airport analysis across a range of years because an airport can change its airport code and airport codes can be reused.

- \*\*dest\*\*: Destination Airport

- \*\*dest\_city\_name\*\*: Destination Airport, City Name

- \*\*crs\_dep\_time\*\*: CRS Departure Time (local time: hhmm)

- \*\*crs\_arr\_time\*\*: CRS Arrival Time (local time: hhmm)

- \*\*dup\*\*: Duplicate flag marked Y if the flight is swapped based on Form-3A data

- \*\*crs\_elapsed\_time\*\*: CRS Elapsed Time of Flight, in Minutes. Computerized Reservations Systems (CRS)

- \*\*flights\*\*: Number of Flights

- \*\*distance\*\*: Distance between airports (miles)