

flights dataset

Data for all flights that departed NYC (i.e. JFK, LGA or EWR) in 2013.

Variables

- **year** - year of departure
- **month** - month of departure
- **day** - day of departure
- **dep_time** - actual departure time (format HHMM or HMM), local tz
- **arr_time** - actual arrival time (format HHMM or HMM), local tz
- **sched_dep_time** - scheduled departure time (format HHMM or HMM), local tz
- **sched_arr_time** - scheduled arrival time (format HHMM or HMM), local tz
- **dep_delay** - departure delay in minutes (negative times represent early departures)
- **arr_delay** - arrival delay in minutes (negative times represent early arrivals)
- **carrier** - Two letter carrier abbreviation; See *airlines* data to get name
- **flight** - Flight number
- **tailnum** - Plane tail number; See *planes* data for additional metadata
- **origin** - origin airport; See *airports* data for additional metadata
- **dest** - destination airport; See *airports* data for additional metadata
- **air_time** - Amount of time spent in the air, in minutes
- **distance** - Distance between airports, in miles
- **hour** - Hour time of scheduled departure (use with *minute*)
- **minute** - Minute time of scheduled departure (use with *hour*)
- **time_hour** - Scheduled date and hour of the flight as a POSIXct date. Along with origin, can be used to join flights data to weather data.

planes dataset

Plane metadata for all plane tailnumbers found in the FAA aircraft registry. American Airways (AA) and Envoy Air (MQ) report fleet numbers rather than tail numbers so can't be matched.

Variables

- **tailnum** - Tail number
 - **year** - Year manufactured
 - **type** - Type of plane
 - **manufacturer** - Manufacturer
 - **model** - Model
 - **engines** - Number of engines
 - **seats** - Number of seats
 - **speed** - Average cruising speed in mph
 - **engine** - Type of engine
-

weather dataset

Hourly meterological data for LGA, JFK and EWR.

Variables

- **origin** - Weather station; Named origin to facilitate merging with *flights* data
 - **year, month, day, hour** Time of recording
 - **temp, dewp** - Temperature and dewpoint in F
 - **humid** - Relative humidity
 - **wind_dir** - Wind direction (in degrees)
 - **wind_speed, wind_gust** - Wind speed and gust speed (in mph)
 - **precip** - Precipitation, in inches.
 - **pressure** - Sea level pressure in millibars
 - **visib** - Visibility in miles
 - **time_hour** - Date and hour of the recording as a POSIXct date.
-

airports dataset

Useful metadata about airports

Variables

- **faa** - FAA airport code
 - **name** - Usual name of the airport
 - **lat, lon** - Location of airport
 - **alt** - Altitude, in feet
 - **tz** - Timezone offset from GMT
 - **dst** - Daylight savings time zone: A = Standard US DST: starts on the second Sunday of March, ends on the first Sunday of November; U = unknown; N = no dst
 - **tzzone** - IANA time zone, as determined by GeoNames webservice
-

airlines data

Look up airline names from their carrier codes.

Variables

- **carrier** - Two letter abbreviation
- **name** - Full name of carrier