Data Table Schema

airports

Important details (name, state, identifier, latitude, longitude, etc.) on various US airports.
322 rows & 6 columns. Size: ~0.1MB. Source: US Department of Transportation.

Field	Type	Description
airport_id	STRING	3-letter airport identifier
airport_name	STRING	Name of the airport
city	STRING	Airport's city
state	STRING	Airport's state
latitude	FLOAT	Airport's latitude
longitude	FLOAT	Airport's longitude

events_US

Public events from around the US throughout 2017.
1,151 rows & 4 columns. Size: ~0.1MB. Source: Shore Fire Media.

Field	Type	Description
date	STRING	Date. Format DD/MM/YYYY.
city	STRING	Name of city event takes place in
state	STRING	Name of state event takes place in
event_name	STRING	Name of event

fares

Airline fare distributions for each quarter-route-airline combination in 2017 with a bucket size of $10.
80,823 rows & 255 columns. Size: ~44MB. Source: US Department of Transportation.

Field	Type	Description
quarter	INTEGER	Financial quarter of the datum
airline_id	STRING	Airline identifier
origin_airport	STRING	Unique 3-character identifier of origin airport
destination_airport	STRING	Unique 3-character identifier of destination airport
distance	INTEGER	Distance between airports, in miles
fare_bucket_x (249 total)	INTEGER	Number of passengers whose fares were between x and x+10 dollars. The first column excludes values below $10. Last column includes all values above $2500

flight_traffic

Information about delays for US domestic flights in 2017. *Note: in order to keep the dataset size manageable, the provided data is a 10% unbiased sample of the raw data. If using flight count metrics, remember to multiply quantities by 10 to approximate the actual data.*
~600,000 rows & 24 columns. Size: ~40MB. Source: Bureau of Transportation Statistics.

Field	Type	Description
year	INTEGER	Year
month	INTEGER	Month
day	INTEGER	Day of month
airline_id	STRING	Airline identifier
origin_airport	STRING	Origin airport code
destination_airport	STRING	Destination airport code
scheduled_departure	INTEGER	Scheduled departure time. Format HHMM
actual_departure	INTEGER	Actual departure time. Format HHMM
taxi_out	INTEGER	Taxi-out time, in minutes
wheels_off	INTEGER	Wheels off time. Format HHMM
wheels_on	INTEGER	Wheels on time. Format HHMM
taxi_in	INTEGER	Taxi-in time, in minutes
scheduled_arrival	INTEGER	Scheduled arrival time. Format HHMM
actual_arrival	INTEGER	Actual arrival time. Format HHMM
cancelled	BINARY	1 if flight was cancelled, 0 otherwise
diverted	BINARY	1 if flight was diverted, 0 otherwise
scheduled_elapsed	INTEGER	Scheduled elapsed time of flight, in minutes
actual_elapsed	INTEGER	Actual elapsed time of flight, in minutes
distance	INTEGER	Distance between airports, in miles
airline_delay	INTEGER	Delay due to the airline, in minutes
weather_delay	INTEGER	Delay due to weather, in minutes
air_system_delay	INTEGER	Delay due to the National Air System, in minutes
security_delay	INTEGER	Delay due to security, in minutes
aircraft_delay	INTEGER	Delay due to late aircraft, in minutes

stock_prices

Daily closing stock prices of various US airlines from late-2016 to early-2018.
380 rows & 10 columns. Size: ~0.1MB. Source: Alpha Vantage.

Field	Type	Description
timestamp	STRING	Date. Format MM/DD/YY.
airline_stock_price (9 total)	FLOAT	Stock closing price

weather

Weather data (temperature, wind, precipitation, cloud cover, etc.) collected at various US airports every 6 hours through 2017.
~375,000 rows & 12 columns. Size: ~40MB. Source: National Centers for Environmental Information.

Field	Type	Description
airport_id	STRING	3-letter airport identifier
airport_name	STRING	Airport's name
latitude	FLOAT	Airport's latitude
longitude	FLOAT	Airport's longitude
elevation	FLOAT	Airport's elevation, in meters
datetime	STRING	Date and time of the measurement. Format YYYY-MM-DDTHH:MM:SS
temperature	FLOAT	Air temperature surrounding the airport, in degrees Celsius
visibility	INTEGER	Visibility, in meters
wind_direction	INTEGER	Horizontal wind direction, in degrees from North clockwise
wind_speed	FLOAT	Horizontal wind speed in meters per second
snow_depth	INTEGER	Depth of snow accumulation in millimeters
cloud_status	INTEGER	Code (0 – 9) representing the cloud cover status

***cloud_status** key:*
0: None
1: ACSL (Altocumulus Standing Lenticular)
2: ACCAS (Altocumulus Castelanus)
3: TCU (Towering Cumulus)
4: MDT CU (Moderate Cumulus)
5: CB/CB MAM DISTANT (Cumulonimbus or Cumulonimbus Mammatus visible)
6: CB/CBMAM (Cumulonimbus or Cumulonimbus Mammatus within 20 nautical miles)
7: Unknown
9: Missing