information

[https://github.com/Haroldinho/nasa_airport_throughput2024] Harold Nikoue, Second Author Atlanta, GA – December 3, 2024

DATA EXPLORATION

December 3, 2024

EXPLORATORY DATA ANALYSIS: DATA CONTENT

- FUSER
    - Configs Data Set (D-ATIS Data)
    - Runways Data Set (Arrival Departure Detection) TARGET
    - First Position Data Set
    - TBFM Data Set (Time-Based Flow Management Data)
    - TFM Track Data Set (Traffic Flow Management Data)
    - ETD Data Set (Estimated Time of Departure)
    - LAMP Data Set (Local Aviation MOS Program Data)
    - MFS Data Set (FAA SWIM Feeds)
- METAR
- Terminal Aerodrome Forecast (TAF)
- Convective Weather Avoidance Model (CWAM)

FUSER

- FUSER dataset consists of 8 types of CSV files, each providing various flight and airport-related data
- Files are named using the format {airport}_date_range.{file_type}_{data_set}.csv
- Data is broken down between test and training set

FUSER TRAINING DATA

Data is broken down by airport

- KATL
- KCLT
- KDEN
- KDFW
- KJFK
- KMEM
- KORD
- KPHX
- KSEA
- location: FUSER_train_{airport}/{airport}/{airport}_{year}-{month}-{day}_{year}-{month}-{day}.{data_set}
- For each day, we have one of the 8 datasets listed above (configs_data_set, runways_data_set, first_postion_data_set, TBFM_data_set, TFM_track_data_set, ETD_data_set, LAMP_data_set, MFS_data_set)

FUSER: CONFIGS_DATA_SET

Contains airport configuration and weather details extracted from D-ATIS messages, used for understanding airport configurations.

Columns

- 'airport_id'
- 'data_header'
- 'src_addr'
- 'datis_time'
- 'start_time'
- 'weather_report'
- 'departure_runways'
- 'arrival_runways'

- 'timestamp_source_received'
- 'timestamp_source_processed'
- 'invalid_departure_runways'
- 'invalid_arrival_runways'
- 'departure_runway_string'
- 'arrival_runway_string'
- 'airport_configuration_name'

FUSER: RUNWAYS_DATA_SET

Columns

- 'gufi'
- 'arrival_runway_actual_time'
- 'arrival_runway_actual'
- 'departure_runway_actual_time'
- 'departure_runway_actual'

FUSER: FIRST_POSITION_DATA_SET

Columns

- 'gufi'
- 'time_first_tracked'

FUSER: TBFM_DATA_SET

Columns

- 'gufi'
- 'timestamp'
- 'arrival_runway_sta'

FUSER: TFM_TRACK_DATA_SET

Columns

- 'gufi'
- 'timestamp'
- arrival_runway_estimated_time'

FUSER: ETD_DATA_SET

Columns

- 'gufi'
- 'timestamp'
- 'departure_runway_estimated_time