

Variable	Description	Additional commentary
id	Unique identifier made up of the combination of dt_prediction_date, dt_target_date and cat_i_flight number	
dt_prediction_date	Date on which the prediction is being made on	E.g. On November 21st 2016, we want to run the model to predict 4 weeks in the future, Dec 19th
dt_target_date	Date being predicted	In that case, dt_prediction_date = 2016-11-21, dt_target_date = 2016-12-20,
s_model_type	How far in the future the prediction is for <i>i.e. dt_target_date - dt_prediction_date</i>	s_model_type = '04 weeks'
cat_case_type	Indicator for whether the case is a target case to be predicted or an explanatory case	
cat_i_flightno	Anonymised transformation of the flight number	E.g. EI125, 06:40 Aer Lingus, Dublin to London Heathrow could be transformed to 1567 Please note that the same flight number will be transformed differently for different model types to avoid any leakage between models
dt_flight_date	Scheduled flight date	E.g. 2016-11-18
num_flight_year	Year of scheduled flight date	
num_flight_month	Month of scheduled flight date	
num_flight_weekofyear	Calendar week of scheduled flight date	Derived from dt_flight_date
num_flight_dayofweek	Day of week of scheduled flight date	
cat_sdt_hour	Scheduled hour of departure	E.g. In the case of EI125 above, this would be 6
cat_i_airport	Anonymised destination airport name	E.g. London Heathrow could be represented by code 101
cat_i_city	Anonymised destination city	E.g. London could be represented by code 234
cat_destination_group_id	Anonymised region indicator	E.g. North America could be represented as code 9. Western Europe could be represented as code 4
cat_longhaul_ind	Indicator for whether the flight is longhaul or not	Value of 1 indicates that the flight is a longhaul flight
num_airport_other_ops	Number of other operators who flew to the same destination airport on this date <i>(based on dt_flight_date and cat_i_flight_number)</i>	
num_airport_other_flights	Number of other flight numbers flying to the same destination airport on this date <i>(based on dt_flight_date and cat_i_flight_number)</i>	Based on dt_flight_date and the destination for cat_i_flight_number
num_city_other_ops	Number of other operators who flew to the same destination city on this date <i>(based on dt_flight_date and cat_i_flight_number)</i>	This gives a proxy for competition / options on a route
num_city_other_flights	Number of other flight numbers flying to the same destination city on this date <i>(based on dt_flight_date and cat_i_flight_number)</i>	
cat_s_plane_capacity	Passenger capacity of plane	
cat_flight_class_type_id	Anonymised indicator for whether the flight is a scheduled or chartered flight	
ord_leisure	Ordinal variable where higher values indicate a larger percentage of passengers flying for leisure purposes	
ord_irish_residents	Ordinal variable where higher values indicate a larger percentage of Irish residents on the flight	
ord_trip_duration	Ordinal variable where higher values indicate a longer average trip duration for passengers on this flight	
ord_female	Ordinal variable where higher values indicate a larger percentage of female passengers on the flight	This is derived from research data related to the flight destination
ord_party_size	Ordinal variable where higher values indicate a larger average party size on the flight	
ord_bag_checkin	Ordinal variable where higher values indicate a larger percentage of passengers checking in a bag on the flight	
ord_arrive_by_car	Ordinal variable where higher values indicate a larger percentage of passengers who arrive to the airport by car for the flight	
num_pax_[]_mins_before_sdt	The number of passengers who present at security screening in the specified 15 minute window prior to scheduled departure time	E.g. if EI125 is departing at 06:40, the column num_pax_210_224_mins_before_sdt relates to the number of pax who present at security screening the window between 2 hrs 45 min and 2 hrs 30 min before the flight is scheduled to depart i.e. between 02:56 and 03:10