Data Modeling

Elements of Data Integration (CS 333E)

Feb 27, 2026

What is an entity?

In a DB, every table is called an entity.

Most entities represent real-world things. Examples:

Places

Movies

Products

They represent distinct, identifiable concepts in the world.

Each entity is represented with a schema.

Each entity has a unique identifier called a Primary Key (abbreviated to PK).

Entities do not contain other entities.

Entities link to other entities. These links are called relationships.

There are three relationship types: one-to-one, one-many, and many-to-many.
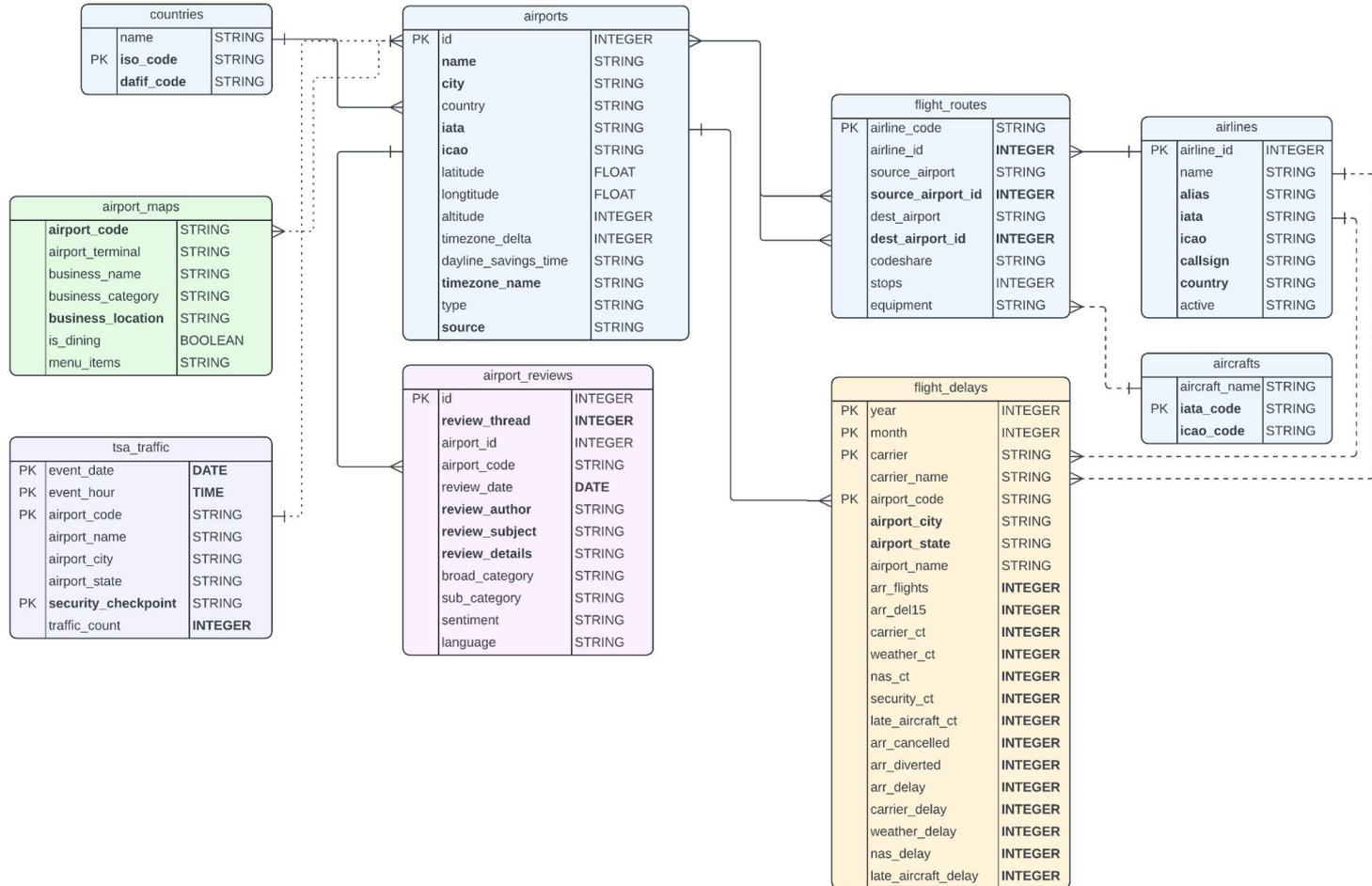
One-to-one and one-to-many relationships are represented with Foreign Keys (abbreviated to FK).

Many-to-many relationships are represented as their own tables called Junction Tables.

What is a core entity?

- Important entities in the database

- They have valuable features

- They are referenced by other entities in the databases

- They may be represented as a table, multiple tables, or even a set of columns within one or more tables in the staging layer

- To identify the core entities in your database, you need some domain knowledge and a sense of how users will want to query your database.

Air Travel Data Model - Staging Layer as of Project 5

countries

	name	STRING
PK	iso_code	STRING
	dafif_code	STRING

airports

PK	id	INTEGER
	name	STRING
	city	STRING
	country	STRING
	iata	STRING
	icao	STRING
	latitude	FLOAT
	longtitude	FLOAT
	altitude	INTEGER
	timezone_delta	INTEGER
	dayline_savings_time	STRING
	timezone_name	STRING
	type	STRING
	source	STRING

flight_routes

PK	airline_code	STRING
	airline_id	INTEGER
	source_airport	STRING
	source_airport_id	INTEGER
	dest_airport	STRING
	dest_airport_id	INTEGER
	codeshare	STRING
	stops	INTEGER
	equipment	STRING

airlines

PK	airline_id	INTEGER
	name	STRING
	alias	STRING
	iata	STRING
	icao	STRING
	callsign	STRING
	country	STRING
	active	STRING

airport_maps

	airport_code	STRING
	airport_terminal	STRING
	business_name	STRING
	business_category	STRING
	business_location	STRING
	is_dining	BOOLEAN
	menu_items	STRING

airport_reviews

PK	id	INTEGER
	review_thread	INTEGER
	airport_id	INTEGER
	airport_code	STRING
	review_date	DATE
	review_author	STRING
	review_subject	STRING
	review_details	STRING
	broad_category	STRING
	sub_category	STRING
	sentiment	STRING
	language	STRING

aircrafts

	aircraft_name	STRING
PK	iata_code	STRING
	icao_code	STRING

tsa_traffic

PK	event_date	DATE
PK	event_hour	TIME
PK	airport_code	STRING
	airport_name	STRING
	airport_city	STRING
	airport_state	STRING
PK	security_checkpoint	STRING
	traffic_count	INTEGER

flight_delays

PK	year	INTEGER
PK	month	INTEGER
PK	carrier	STRING
	carrier_name	STRING
PK	airport_code	STRING
	airport_city	STRING
	airport_state	STRING
	airport_name	STRING
	arr_flights	INTEGER
	arr_del15	INTEGER
	carrier_ct	INTEGER
	weather_ct	INTEGER
	nas_ct	INTEGER
	security_ct	INTEGER
	late_aircraft_ct	INTEGER
	arr_cancelled	INTEGER
	arr_diverted	INTEGER
	arr_delay	INTEGER
	carrier_delay	INTEGER
	weather_delay	INTEGER
	nas_delay	INTEGER
	late_aircraft_delay	INTEGER

Transformation process: Start with current entities

Original Entities (Staging)		
- Country - Airport - Flight Route - Airline - Aircraft - Airport Maps - TSA Traffic - Airport Reviews - Flight Delays		

Transformation process: Promote core entities

Original entities (staging)	Core entities	
● Country ● Airport ● Flight Route ● Airline ● Aircraft ● Airport Maps ● TSA Traffic ● Airport Reviews ● Flight Delays	● Country ● Airport ● Airline ● Aircraft	

Transformation process: Identify new entities

Original entities (staging)	Core entities	
● Country ● Airport ● Flight Route ● Airline ● Aircraft ● Airport Maps ● TSA Traffic ● Airport Reviews ● Flight Delays	● Country ● Airport ● Airline ● Aircraft New entities ● Airport Business ● Airport Terminal ?	

Transformation process: Add remaining entities

Original entities (staging)	Core entities	Final entity list
● Country ● Airport ● Flight Route ● Airline ● Aircraft ● Airport Maps ● TSA Traffic ● Airport Reviews ● Flight Delays	● Country ● Airport ● Airline ● Aircraft New entities ● Airport Business ● Airport Terminal ?	● **Country** ● **Airport** ● Flight Route ● **Airline** ● **Aircraft** ● Airport Maps ● **Airport Business** ● TSA Traffic ● Airport Reviews ● Flight Delays

Air Travel Data Model - Normalized Layer as of Project 6