* **Database: airlines**

CREATE DATABASE airlines

WITH

OWNER = postgres

ENCODING = 'UTF8'

LC\_COLLATE = 'English\_United States.1252'

LC\_CTYPE = 'English\_United States.1252'

TABLESPACE = pg\_default

CONNECTION LIMIT = -1;

* **Table: public.flights**

CREATE TABLE public.flights

(

fl\_date date,

op\_carrier text COLLATE pg\_catalog."default",

op\_carrier\_fl\_number text COLLATE pg\_catalog."default",

origin text COLLATE pg\_catalog."default",

dest text COLLATE pg\_catalog."default",

crs\_dep\_time numeric(4,0),

dep\_time numeric(4,0),

dep\_delay numeric(4,0),

taxi\_out numeric(4,0),

wheels\_off numeric(4,0),

wheels\_on numeric(4,0),

taxi\_in numeric(4,0),

crs\_arr\_time numeric(4,0),

arr\_time numeric(4,0),

arr\_delay numeric(4,0),

canceled numeric(1,0),

cancelation\_code text COLLATE pg\_catalog."default",

diverted numeric(1,0),

crs\_elapsed\_time numeric(4,0),

actual\_elapsed\_time numeric(4,0),

air\_time numeric(4,0),

distance numeric(4,0),

carrier\_delay numeric(4,0),

weather\_delay numeric(4,0),

nas\_delay numeric(4,0),

security\_delay numeric(4,0),

late\_aircraft\_delay numeric(4,0)

)

TABLESPACE pg\_default;

ALTER TABLE public.flights

OWNER to postgres;

* **Import data**
* **Add new columns**
* Alter table flights add column dep\_hour integer, add column dep\_min integer;
* Update dep\_hour and dep\_min columns

with fl\_time as

(

select distinct dep\_time as dt,

case WHEN length(dep\_time::text) <=2 THEN 0::text

ELSE substr(dep\_time::text, 1, (length(dep\_time::text)-2))::text

END as hr,

right(dep\_time::text, 2)::integer as mn

from flights

)

Update flights

SET dep\_hour = fl\_time.hr::integer,

dep\_min = fl\_time.mn::integer

FROM fl\_time

where flights.dep\_time = fl\_time.dt;

* Alter table flights add column season text;
* Update flights

set season = CASE

WHEN date\_part('month',fl\_date) in (12,1,2) THEN 'winter'

WHEN date\_part('month',fl\_date) in (3,4,5) THEN 'spring'

WHEN date\_part('month',fl\_date) in (6,7,8) THEN 'summer'

WHEN date\_part('month',fl\_date) in (9,10,11) THEN 'fall'

END

* Alter table flights add column day\_part text;
* Update flights

set day\_part = CASE

WHEN dep\_hour >= 5 and dep\_hour < 12 THEN 'morning'

WHEN dep\_hour >= 12 and dep\_hour < 17 THEN 'afternoon'

WHEN dep\_hour >= 17 and dep\_hour < 22 THEN 'evening'

ELSE 'night'

END

* Alter table flights add column flight\_length text;
* Update flights

set flight\_length = CASE

WHEN distance < 1500 THEN 'short haul'

WHEN dep\_hour > 5000 THEN 'long haul'

ELSE 'medium haul'

END

* Alter table flights add column op\_code integer;
* Update flights

SET op\_code = (select op\_code from carriers where carriers.op\_carrier = flights.op\_carrier)

* Alter table flights add column origin\_id integer, add column dest\_id integer;
* Update flights

SET origin\_id = (select loc\_id from locations where locations.loc\_name = flights.origin),

dest\_id = (select loc\_id from locations where locations.loc\_name = flights.dest)

* Alter table flights add column delayed integer;
* Update flights

SET delayed = case WHEN carrier\_delay is not null or weather\_delay is not null or nas\_delay is not null or security\_delay is not null or late\_aircraft\_delay is not null

THEN 1

ELSE 0

END

* Alter table flights add column fl\_month integer, add column fl\_day integer;
* Update flights

set fl\_month = date\_part('month',fl\_date),

fl\_day = date\_part('day',fl\_date)

* **Delete columns**
* Delete from flights where actual\_elapsed\_time is null