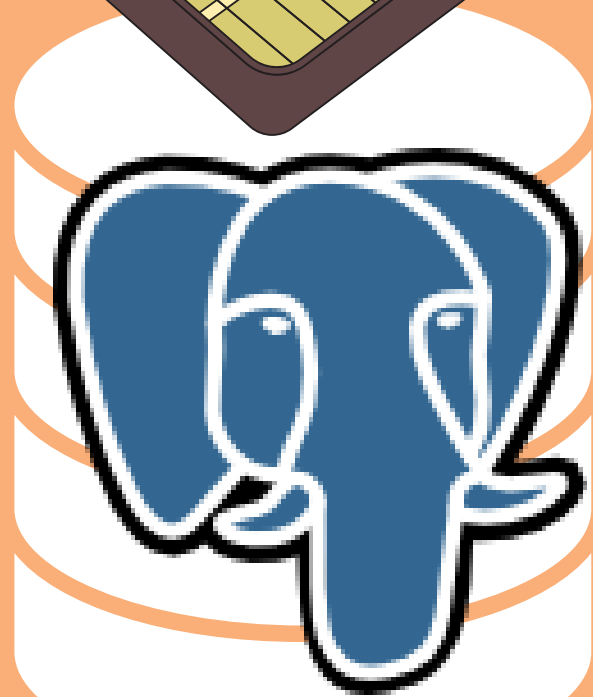


# MASTERING DATA MODELING WITH AIRBNB DATASET ON POSTGRES DB

## A PRACTICAL GUIDE TO DIM AND FACT TABLES



123	id
ABC	name
123	host_id
ABC	host_name
123	neighbourhood_group
ABC	neighbourhood
123	latitude
123	longitude
ABC	room_type
123	price
123	minimum_nights
123	number_of_reviews
ABC	last_review
123	reviews_per_month
123	calculated_host_listings_count

123	listing_id
ABC	date
ABC	available
ABC	price

123	listing_id
123	id
ABC	date
123	reviewer_id
ABC	reviewer_name
ABC	comments

# **DATAMODELING OBJECTIVE TO ANSWER QUESTIONS THAT MATTER**

**TO FIND WHICH HOTEL  
IS HAVING THE  
HIGHEST REVENUE**

**ON WHICH DATE  
THERE WERE MORE  
BOOKINGS**

1. CONNECT THESE DATA TABLES WITH DIVERSE DIMENSIONS IN A WAY TO EFFICIENTLY QUERY THEM.
2. REDUCE THE AMOUNT OF DATA THAT IS TOUCHED DURING THE QUERY
3. SUBSEQUENTLY INCREASE THE SPEED OF QUERY RESPONSE.

**HOW THEIR PRICING IS  
CORRELATED WITH  
REVIEW RATING**

**REVIEW QTY OF EACH  
PROPERTY**

# TYPES OF SCHEMA : IN DATA MODELING

## 6 TYPES OF SCHEMAS

1. THE **FLAT MODEL** IS FOR SMALL, SIMPLE APPLICATIONS.
2. THE **HIERARCHICAL MODEL** IS FOR NESTED DATA, LIKE XML OR JSON.
3. THE **NETWORK MODEL** IS USEFUL IN MAPPING AND SPATIAL DATA, ALSO FOR DEPICTING WORKFLOWS.
4. THE **RELATIONAL MODEL** BEST REFLECTS OBJECT-ORIENTED PROGRAMMING APPLICATIONS.
5. THE **STAR SCHEMA** AND **SNOWFLAKE SCHEMA** ARE FOR ANALYZING LARGE DATASETS.

## STAR SCHEMAS

THE STAR SCHEMA IS A DESIGN APPROACH FOR STORING AND ANALYZING MASSIVE AMOUNTS OF DATA, AND IT RELIES ON THE USAGE OF "FACTS" AND "DIMENSIONS."

A "**FACT**" IS A NUMERICAL DATA POINT THAT DRIVES BUSINESS PROCESSES,

A "**DIMENSION**" IS A DESCRIPTION OF THAT FACT. USING CAR SALES NUMBERS, FOR EXAMPLE,

THE "**FACT**" **TABLE** WOULD CONTAIN INFORMATION ABOUT THE NUMBER OF UNITS SOLD, AND A CORRESPONDING "**DIMENSIONAL**" **TABLE** WOULD HAVE THE COLORS OF THOSE CARS.

# STEPS FOR DATA MODELING AIRBNB DATASET

- 1.GET THE SCHEMA FROM THE AIRBNB DATABASE THAT IS ALREADY IN PG
- 2.THE DIMTABLES AND FACTTABLE ARE ALREADY IN THE SCHEMA
- 3.UNDERSTAND THE SCHEMA AND DECIDE HOW TO CONNECT FACTTABLE AND DIMENSION TABLE
- 4.THINK AND EXPLAIN HOW THE FINAL SCHEMA WILL HELP IN IMPROVING THE OLAP PROCESS

**WHAT ARE THE MAIN ENTITIES/CONCEPTS IN MY APPLICATION?**

- LISTING (LOCATION, HOST, PROPERTY DETAILS)
- REVIEW (COMMENT AND STARS)
- DATE / DAY DETAILS (DATE, DAY, MONTH, YEAR)

**HOW DO THEY RELATE TO EACH OTHER?**

- 1.REVIEW, REVIEW STARS, PRICE, SALES ARE THE DRIVERS
- 2.REVIEWS ARE GIVEN BY CUSTOMERS
- 3.CUSTOMERS ARE RELATED TO THE LISTINGS THEY STAY
- 4.PROPERTY HAVE HOSTS, LOCATION
- 5.HOSTS CAN HAVE MULTIPLE PROPERTIES
- 6.LISTING WILL HAVE ONE PROPERTY LINKED TO IT

# PRACTICAL @ DBEAVER

## CREATING THE SCHEMA & ER DIAGRAM

listings	
123	id
ABC	name
123	host_id
ABC	host_name
123	neighbourhood_group
ABC	neighbourhood
123	latitude
123	longitude
ABC	room_type
123	price
123	minimum_nights
123	number_of_reviews
ABC	last_review
123	reviews_per_month
123	calculated_host_listings_count
123	availability_365

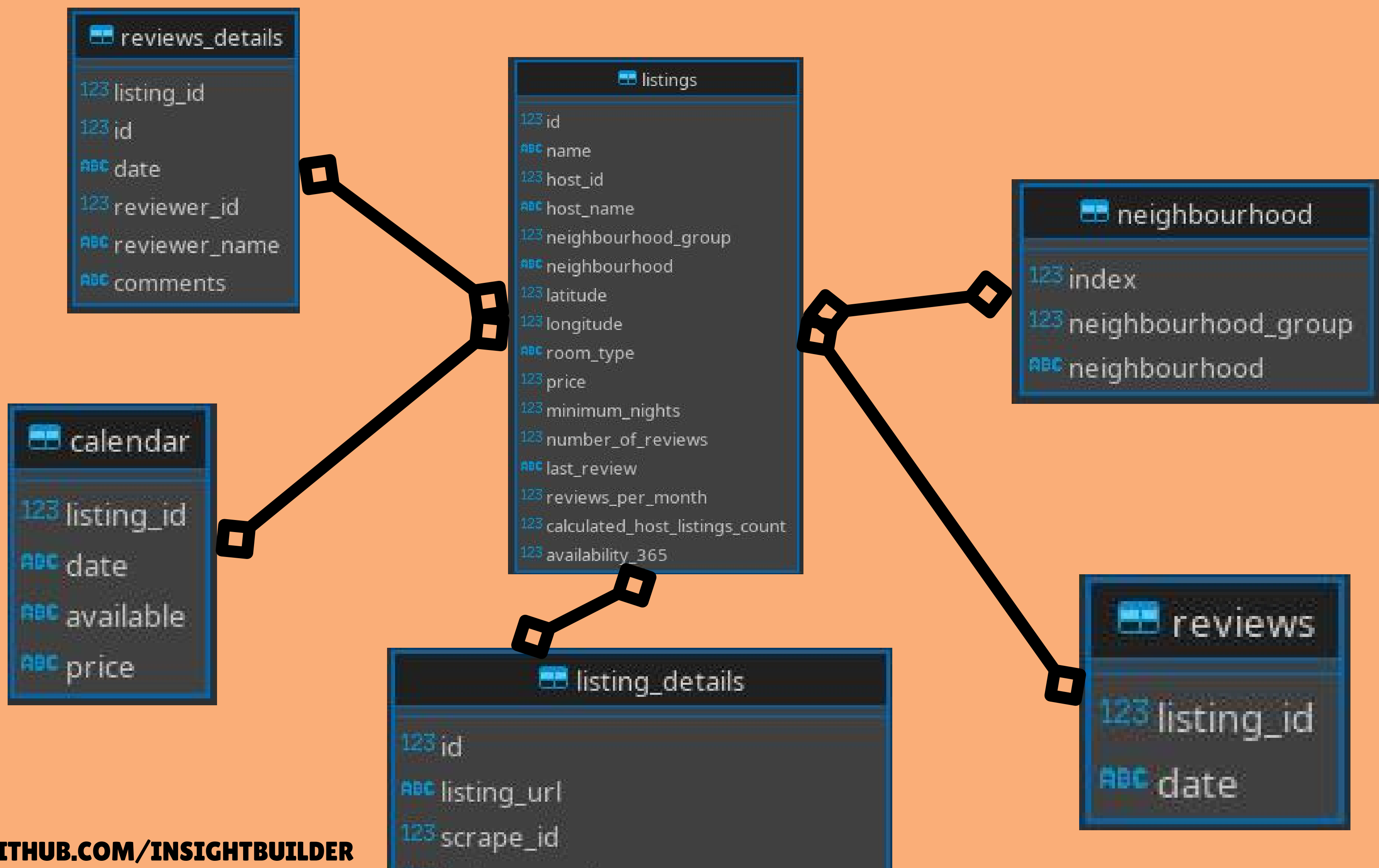
calendar	
123	listing_id
ABC	date
ABC	available
ABC	price

neighbourhood	
123	index
123	neighbourhood_group
ABC	neighbourhood

reviews	
123	listing_id
ABC	date

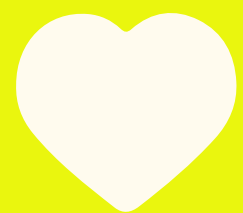
reviews_details	
123	listing_id
123	id
ABC	date
123	reviewer_id
ABC	reviewer_name
ABC	comments

listing_details	
123	id
ABC	listing_url
123	scrape_id
ABC	last_scraped
ABC	name

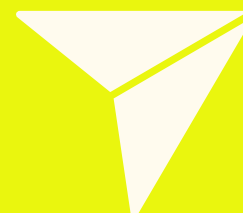


**PRACTICE**

**PRACTICE**



**LIKE**



**SHARE**



**SUBSCRIBE**

**PRACTICE**

**PRACTICE**