SuperTravel-Travel Management Project

Group7 :
Boyu Wang, Caisen Mo, Qinyun Zhang, Zhenyu Wang

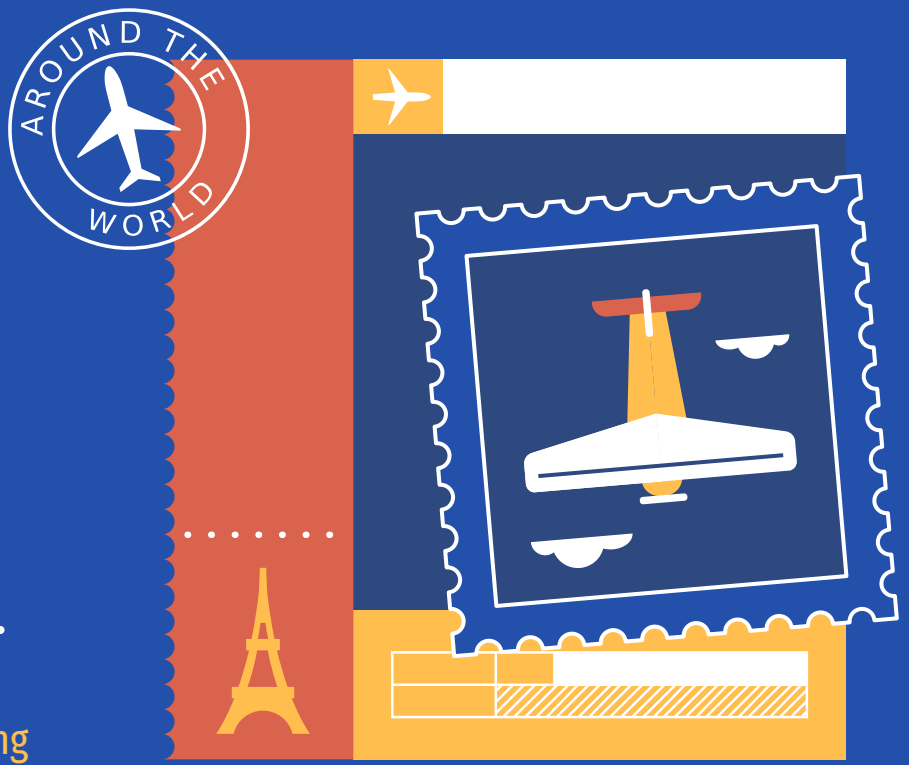


TABLE OF CONTENTS

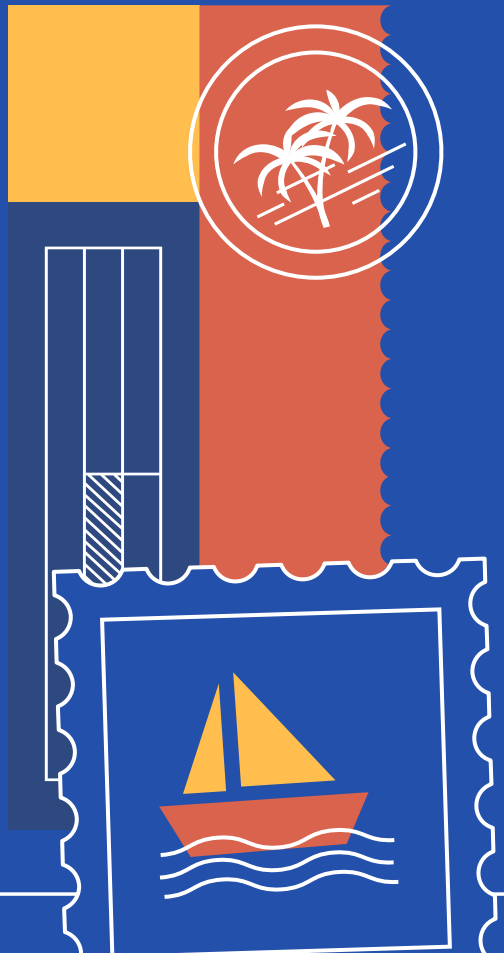
01 Client Scenario

02 Sample of Original Data

03 Normalization Plan

04 ETL Process

05 Database Interaction



Client Scenario

- Three main priorities that most people tend to focus on:
  - Flight
  - Hotel
  - Car rental
- An abundance of information is scattered across various websites and platforms.
- Our team members have gained firsthand experience with the intricacies and challenges of this process in real-life scenarios.
- Ctrip, a prominent Chinese travel service software that integrates three essential services: airfare purchases, hotel bookings, and vehicle rentals.

Sample Data

Flight

MONTH	DAY_OF_MONTH	DAY_OF_WEEK	OP_UNIQUE_CARRI	TAIL_NUM	DEST	DEP_DELAY	CRS_ELAPSED_TIN	DISTANCE	CRS_DEP_M	DEP_TIME_M	CRS_ARR_M
11	1	5	B6	N828JB	CHS	-1	124	636	324	323	448
11	1	5	B6	N992JB	LAX	-7	371	2475	340	333	531
11	1	5	B6	N959JB	FLL	40	181	1069	301	341	482
11	1	5	B6	N999JQ	MCO	-2	168	944	345	343	513
11	1	5	DL	N880DN	ATL	-4	139	760	360	356	499

Temperature	Dew Point	Humidity	Wind	Wind Speed	Wind Gust	Pressure	Condition	sch_dep	sch_arr	TAXI_OUT
48	34	58	W	25	38	29.86	Fair / Windy	9	17	14
48	34	58	W	25	38	29.86	Fair / Windy	9	17	15
48	34	58	W	25	38	29.86	Fair / Windy	9	17	22
48	34	58	W	25	38	29.86	Fair / Windy	9	17	12
46	32	58	W	24	35	29.91	Fair / Windy	9	17	13

Hotel

hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_n	arrival_date_day_of	stays_in_weekend_n	stays_in_week_night
Resort Hotel	0	342	2015	July	27	1	0	0
Resort Hotel	0	737	2015	July	27	1	0	0
Resort Hotel	0	7	2015	July	27	1	0	1
Resort Hotel	0	13	2015	July	27	1	0	1
Resort Hotel	0	14	2015	July	27	1	0	2

adults	children	babies	meal	country	market_segment	distribution_channel	is_repeated_guest	previous_cancellatic	previous_bookings_n	reserved_room_type
2	0	0	BB	PRT	Direct	Direct	0	0	0	C
2	0	0	BB	PRT	Direct	Direct	0	0	0	C
1	0	0	BB	GBR	Direct	Direct	0	0	0	A
1	0	0	BB	GBR	Corporate	Corporate	0	0	0	A
2	0	0	BB	GBR	Online TA	TA/TO	0	0	0	A

Car rental

fuelType	rating	renterTripsTaken	reviewCount	location.city	location.country	location.latitude	location.longitude
ELECTRIC	5	13	12	Seattle	US	47.449107	-122.308841
ELECTRIC	5	2	1	Tijeras	US	35.11106	-106.276551
HYBRID	4.92	28	24	Albuquerque	US	35.127163	-106.566681
GASOLINE	5	21	20	Albuquerque	US	35.149726	-106.711425
GASOLINE	5	3	1	Albuquerque	US	35.208659	-106.601008

location.state	owner.id	rate.daily	vehicle.make	vehicle.model	vehicle.type	vehicle.year
WA	12847615	135	Tesla	Model X	suv	2019
NM	15621242	190	Tesla	Model X	suv	2018
NM	10199256	35	Toyota	Prius	car	2012
NM	9365496	75	Ford	Mustang	car	2018
NM	3553565	47	Chrysler	Sebring	car	2010

Normalization & Database Schema

1

Users related tables

Users Table, Payment Info Table:

2

Flights related tables

Airline Table , Flight Table, Flight-User Relationship Table

3

Hotels related tables

Hotel Table, Room Type Table, Hotel Booking Table, Customer Type Table, Guest Table

4

Car rentals related tables

Renter Table, Vehicle Table, Rental Table, Car Rental User Table, Car Rental Reservation Table

ETL Process

1

Data Extraction

Users Table, Payment Info Table:

2

Data Cleaning

Airline Table , Flight Table, Flight-User Relationship Table

3

Data Transformation

Hotel Table, Room Type Table, Hotel Booking Table, Customer Type Table, Guest Table

4

Data Loading

Renter Table, Vehicle Table, Rental Table, Car Rental User Table, Car Rental Reservation Table

5

Data Loading

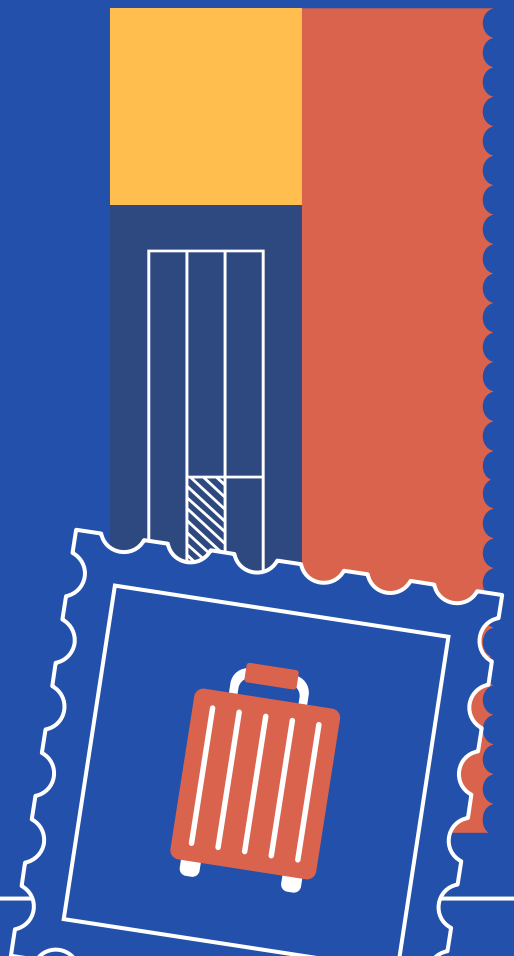
Renter Table, Vehicle Table, Rental Table, Car Rental User Table, Car Rental Reservation Table

10 Questions

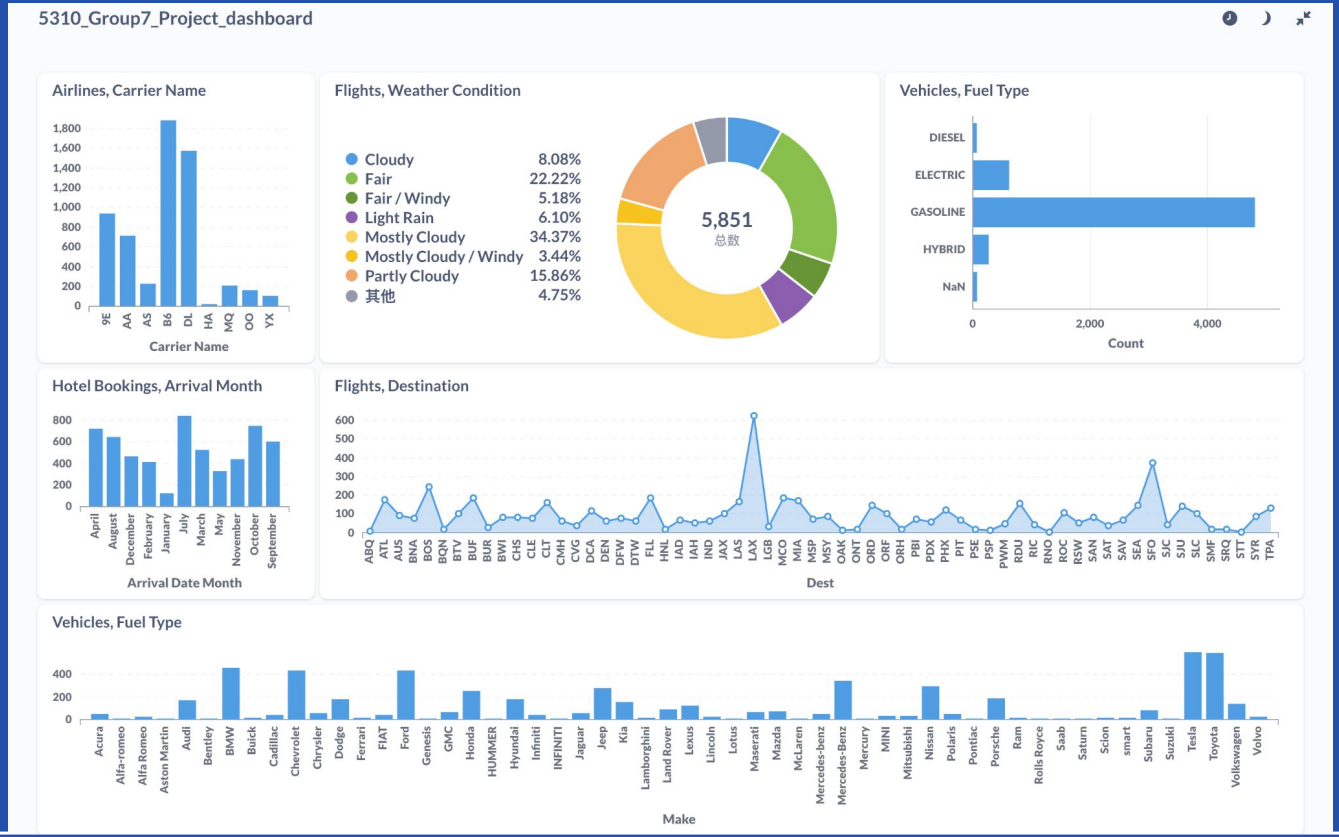
1. Which airline carrier is mostly used
2. What is average elapsed time for different destinations.
3. What is the most common departure time for different destinations.

4. Which room type (room_type_name) is most commonly booked for stays that include weekend nights (stays_in_weekend_nights)?
5. Which month (arrival_date_month) has the highest number of bookings among hotel bookings?
6. Are bookings made by guests requiring car parking spaces (required_car_parking_spaces) more likely to be canceled (is_canceled) compared to bookings that don't require parking?
7. What is the average lead time for hotel bookings, and does this vary based on the type of hotel (hotel_type)?

8. What is the average daily rate for each vehicle type (type), and how does it vary based on the fuel type (fuel_type)?
9. What is the most commonly rented car make and model, and what is its average daily rate?
10. What is the average number of trips taken by renters, and how does it correlate with their review counts?



Metabase Interaction



THANKS!