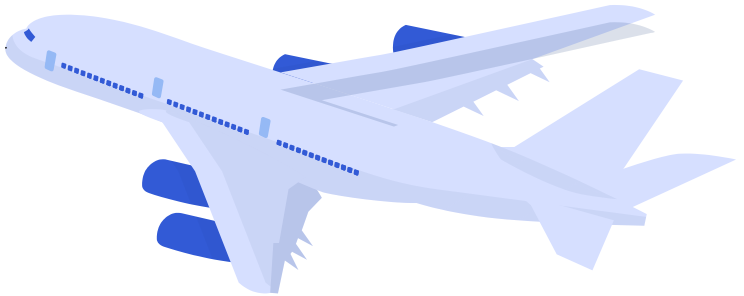
Super Travel



Group 3: Yanlin Zhou, Ziting Guo, Xueqi Zhou, Candice Huang

Super Travel

SuperTravel is a third-party platform that provides customers with a one-stop-shop for airline tickets, rental cars, and hotel bookings.

SuperTravel has approached us **to create a comprehensive and efficient database** that can quickly **update travel and user data, record reservation information, incorporate user reviews and feedback, and store company information**. By providing SuperTravel with a centralized platform for their data, our improved **database design will help them to access and analyze data quickly and efficiently**, leading to better business decisions and improved customer service.

Data Overview

Real Data

	travelCode	userCode	from	to	flightType	price	time	distance	agency	date
850	425	4	Aracaju (SE)	Campo Grande (MS)	firstClass	1389.37	1.69	650.10	CloudFy	09/26/2019
851	425	4	Campo Grande (MS)	Aracaju (SE)	firstClass	1581.80	1.69	650.10	CloudFy	09/29/2019
8994	4497	45	Aracaju (SE)	Brasilia (DF)	premium	755.62	1.11	425.98	CloudFy	10/08/2020
8995	4497	45	Brasilia (DF)	Aracaju (SE)	premium	987.07	1.11	425.98	CloudFy	10/10/2020
9458	4729	47	Aracaju (SE)	Rio de Janeiro (RJ)	economic	717.04	1.55	597.61	CloudFy	04/09/2020

Simulated Data

user_id	first_name	last_name	gender	email	phone_number	credit_card	address_id
1	Renae	Wong	Female	rwong0@wikia.com	391-305-7828	3.54938E+15	1
2	Garwood	Langmuir	Male	glangmuir1@seesaa.net	106-414-5723	6.70995E+15	2
3	Nate	Gulberg	Male	ngulberg2@amazon.com	923-154-5798	3.37942E+14	3
4	Galen	Shillitto	Male	gshillitto3@cbslocal.com	289-413-9218	5.55811E+15	4

Normalization Plan

1 Goal

Design a centralized database that addresses SuperTravel's needs and improves their decision-making process

2 Research

Researched vacation planning tools such as Booking.com and TripAdvisor to understand the standard data requirements for a travel booking database.

3 Design

Design a database that centralizes SuperTravel's data and enables efficient access to customer data, booking information, user reviews, and feedback

4 ETL

Extract, transform, and load data from multiple sources into a unified format. Use real data as much as possible and generate virtual data to fill in the gaps.

Database Design

1

Goal

SuperTravel database involves creating 15 tables that adhere to the third normal form (3NF) to ensure data consistency and accuracy.

2

Data Preparation

Acquire actual travel data from kaggle.
Try to use real data as much as possible and generate virtual data to fill in the gaps.

3

Design

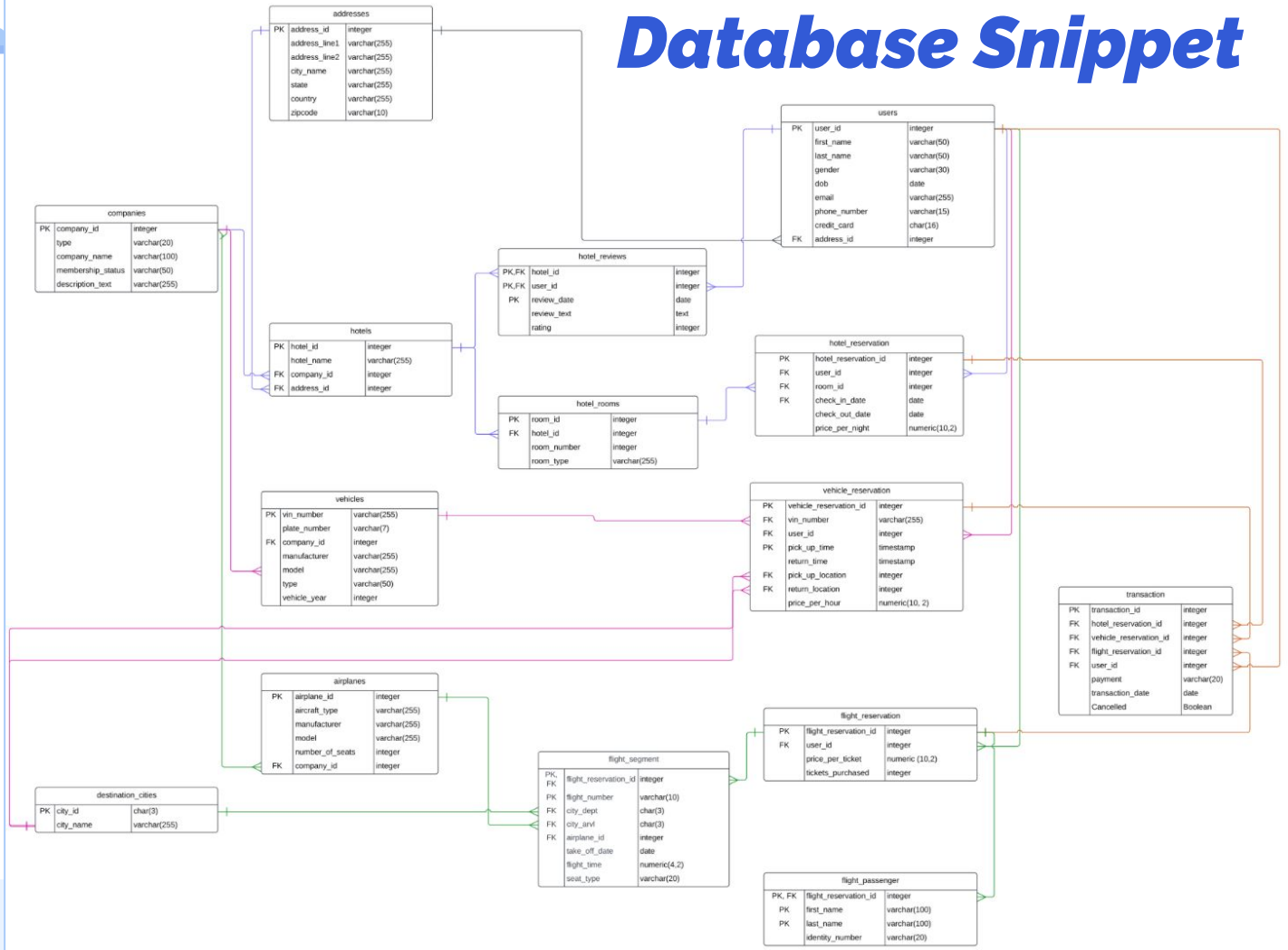
Includes tables for users, companies, addresses, destination cities, flights, flight segments, vehicles, car rental companies, hotels, hotel room types, hotel rooms, hotel reservations, vehicle reservations, flight reservations, and transactions.

4

Benefit

Allow users to book flights, rental cars, and hotels, make reviews, manage reservations, and make transactions. Help SuperTravel tailor their offerings to user preferences.

Database Snippet



ETL Process

Extract

Extract data from CSV files and simulated data through Python.

Real data (flights, hotels and car rental)

Simulated data (users, Flight_passager, reviews, etc)

Transform

Utilizing Python to clean the data, select relevant columns, and create new dataframes with the database schema designed to fulfill SuperTravel's needs.

Load

Load transformed data into SQL.

After data quality check, load aggregated data into 15 designed 3NF tables.

Analytical Process

Rational

The purpose of the analytical procedures is to enable efficient access to customer data, booking information, user feedback, and generally satisfy the data storage and retrieval need of our client

Customer Need

By using analytical procedures, SuperTravel can receive regular reports through Metabase that will help them provide better services to their customers, improve customer satisfaction, and increase their market share.

Insights

SuperTravel would enjoy a better understanding of customer preferences, and enhanced data consistency and accuracy, and make informed strategic decisions and tailor their offerings to companies and individual travelers.

Redundancy & Performance Concerns

Queries

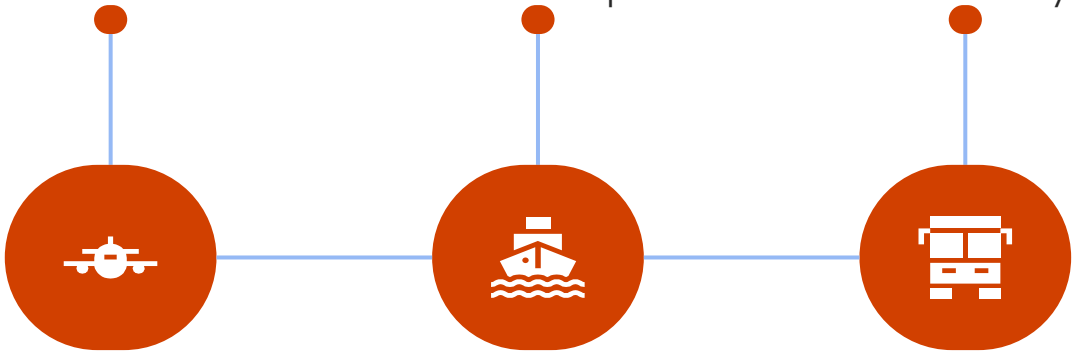
Ensure queries are as efficient as possible

Servers

Store data in distributed servers for performance and backups

NFs

Use 3NF tables in the database to maintain data consistency



Database Interaction

Q1

Q2

hotel_name	company_name	lo
Hotel AU	Mcneil-Bonilla	Reci

Rows 1-1 of 7

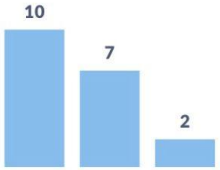
Q3

hotel_name	rating	review_text
Hotel BD	10	Exceptional

Rows 1-1 of 165

Q4

Q5



Q6

Q7

user_id	first_name	last_name
10	Angel	Fernier

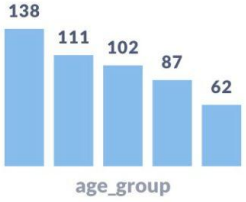
Rows 1-1 of 11

Q8

transaction_id	hotel_reservation_id
3	179

Rows 1-1 of 161

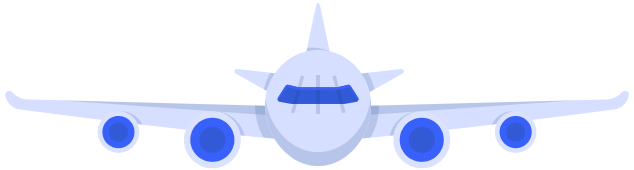
Q9



Q10

user_id	first_name	last_name
34	Federica	Hallums

Rows 1-1 of 6



Thank You !!!