

AeroUS

Group 4-

Aishwarya Chaubal

Bhupinder Jagwani

Nivedita Thapa

Shruti Garg





CONTENTS

- 01** Business Case Scenario
- 02** Swim Lane Diagram
- 03** Our Approach
- 04** Technology Overview
- 05** Analytics
- 06** Airline Modeling System
- 07** ETL Process
- 08** Tableau Demo

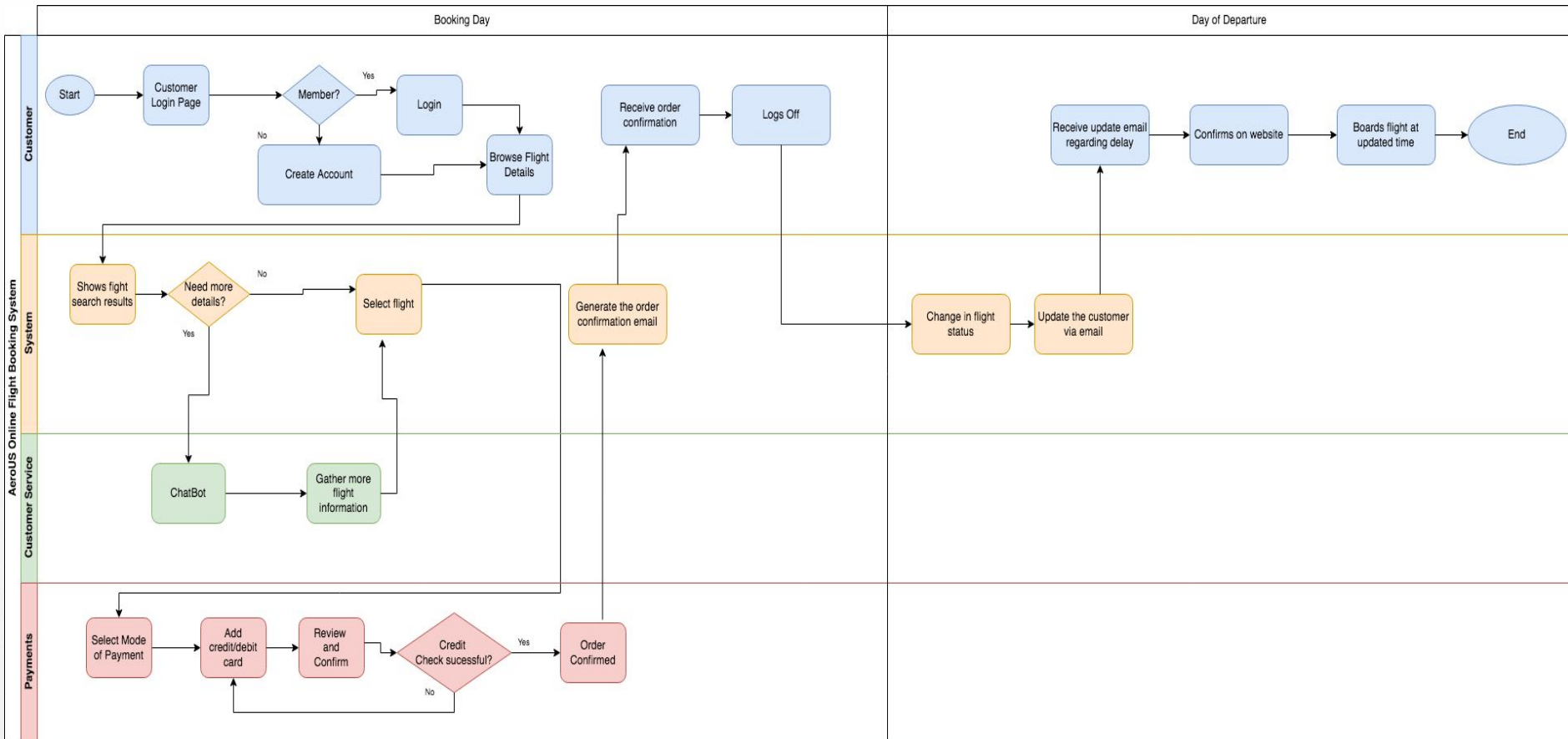
Business Case Scenario



“John wanted to book a flight from San Francisco to New York to visit his parents for Christmas. He knew that AeroUS Airlines was having a sale for Christmas and decided to book the flight 4 months in advance. He logged on to the site, signed in with his email ID and password, and started searching for the flights. He entered the source and destination, the required travel date, the number of passengers, and the class he wanted to travel in. He found a few flights departing at night(which he preferred), was within his budget and seemed like a good option for him. However, he had a few doubts regarding baggage allowance and decided to use the AeroUS chatbot to resolve his doubts. He asked the agent about the check-in baggage allowance and the charges for an extra bag. After getting all his queries answered from the chatbot, John selected the preferred flight. He navigated to check-out, entered his personal information, billing and mailing address, credit card details, and finally purchased the ticket. His ticket details were emailed to him. Satisfied with the services provided and the ease of the entire process, he logged off.

On the day of his flight departure, he received an email update from AeroUS stating that the flight would be delayed by 3 hours due to bad weather. John went on AeroUS’s website to confirm the delay, and decided to leave for the airport at a later time than previously planned.”

Swimlane Diagram



Our Approach



Data Collection & Profiling



Data Cleansing



Data Modeling



Data Integration

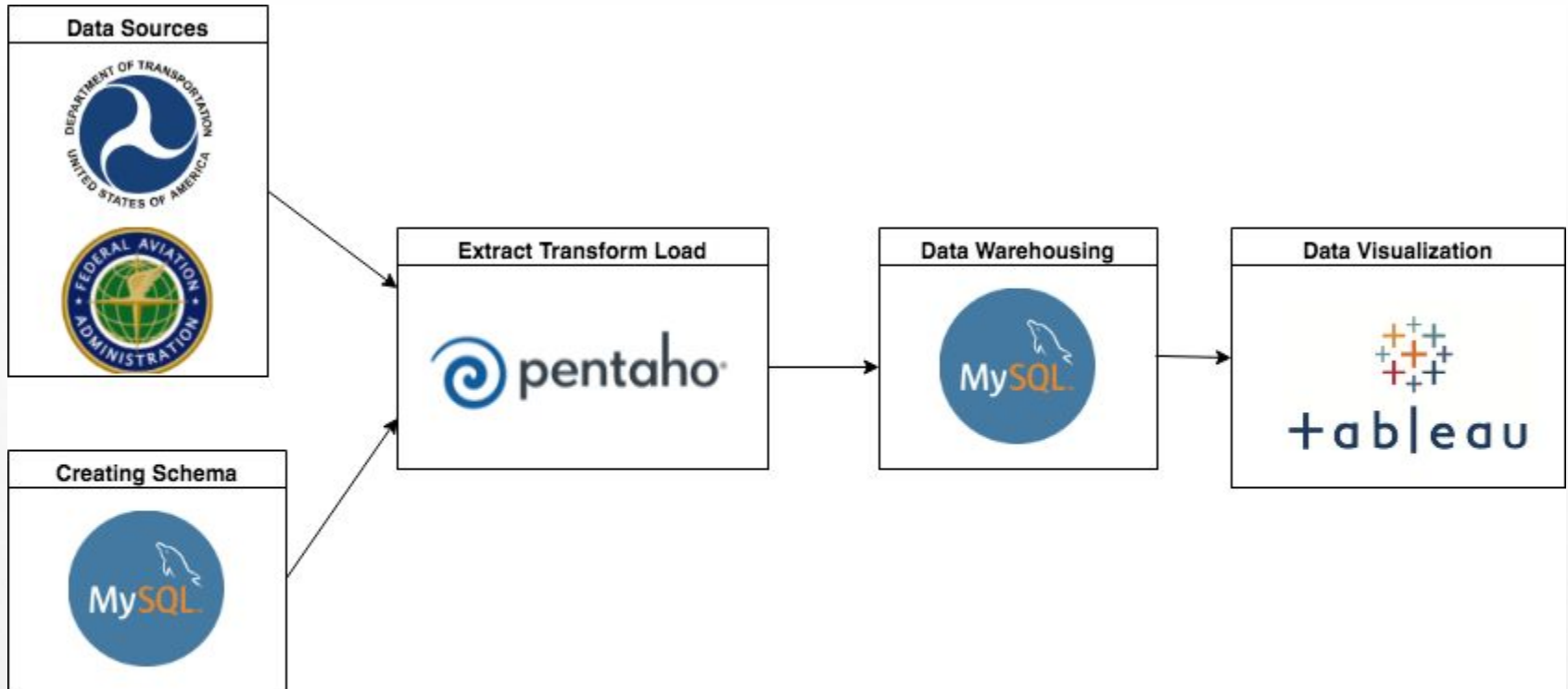


Data Analytics



Data Insights

Technology Overview



Analytics - Lead and Lag



Lead Measures:

1. How many left on time in a day?
2. What are the causes of delay per day?
3. What are the airports where delays were caused?
4. Maximum and Minimum Total Sales per Week



VS



Lag Measures:

1. Increase in the number of flights in the last 1 year?
2. How many left on time in a month?
3. Cause of delay per year?
4. Top 10 Airport names where delays were in a month?
5. Number of bookings via third-party website in the past month
6. Total yearly sales
7. Total sales per quarter

Analytics - Descriptive and Predictive



Descriptive Analytics:

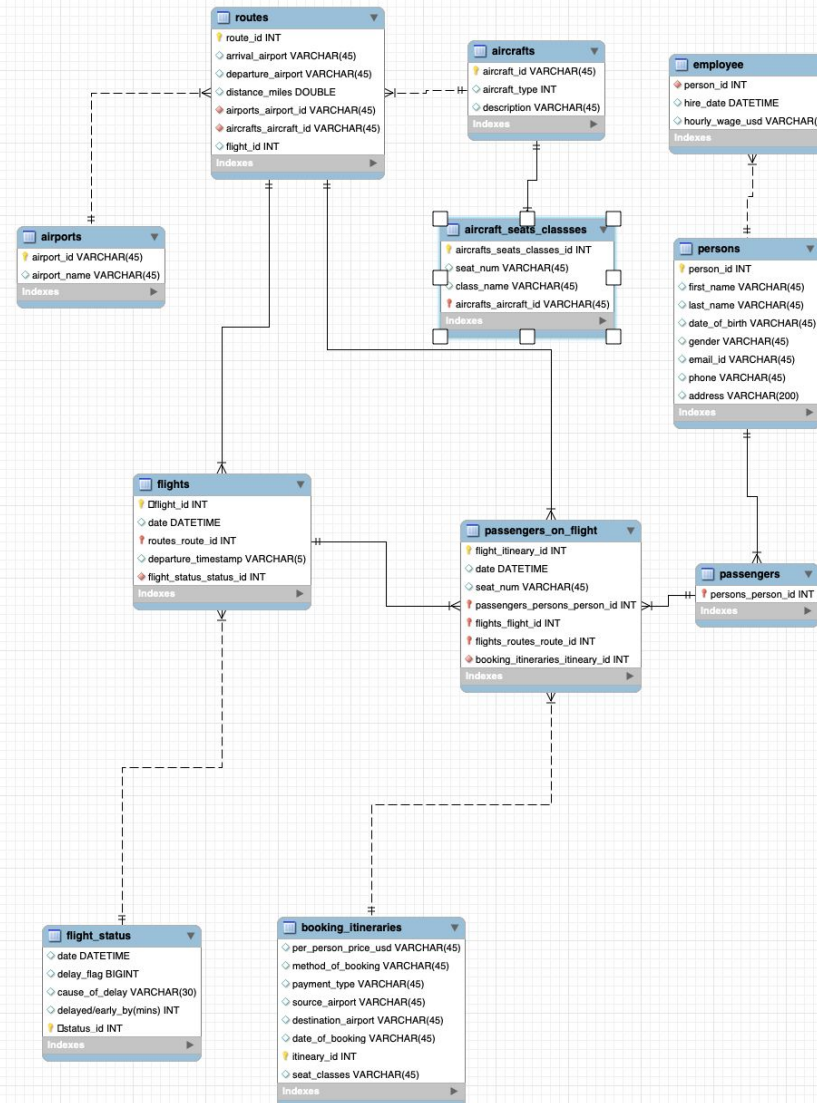
1. Revenue increase/decrease over the period of 1 year
2. Increase in the number of flights in the last 1 year
3. Increase in the number of new routes
4. Top 5 fully-booked flight routes in the last year
5. Revenue generated by each airline service category (cargo/passenger) in the past 1 year.



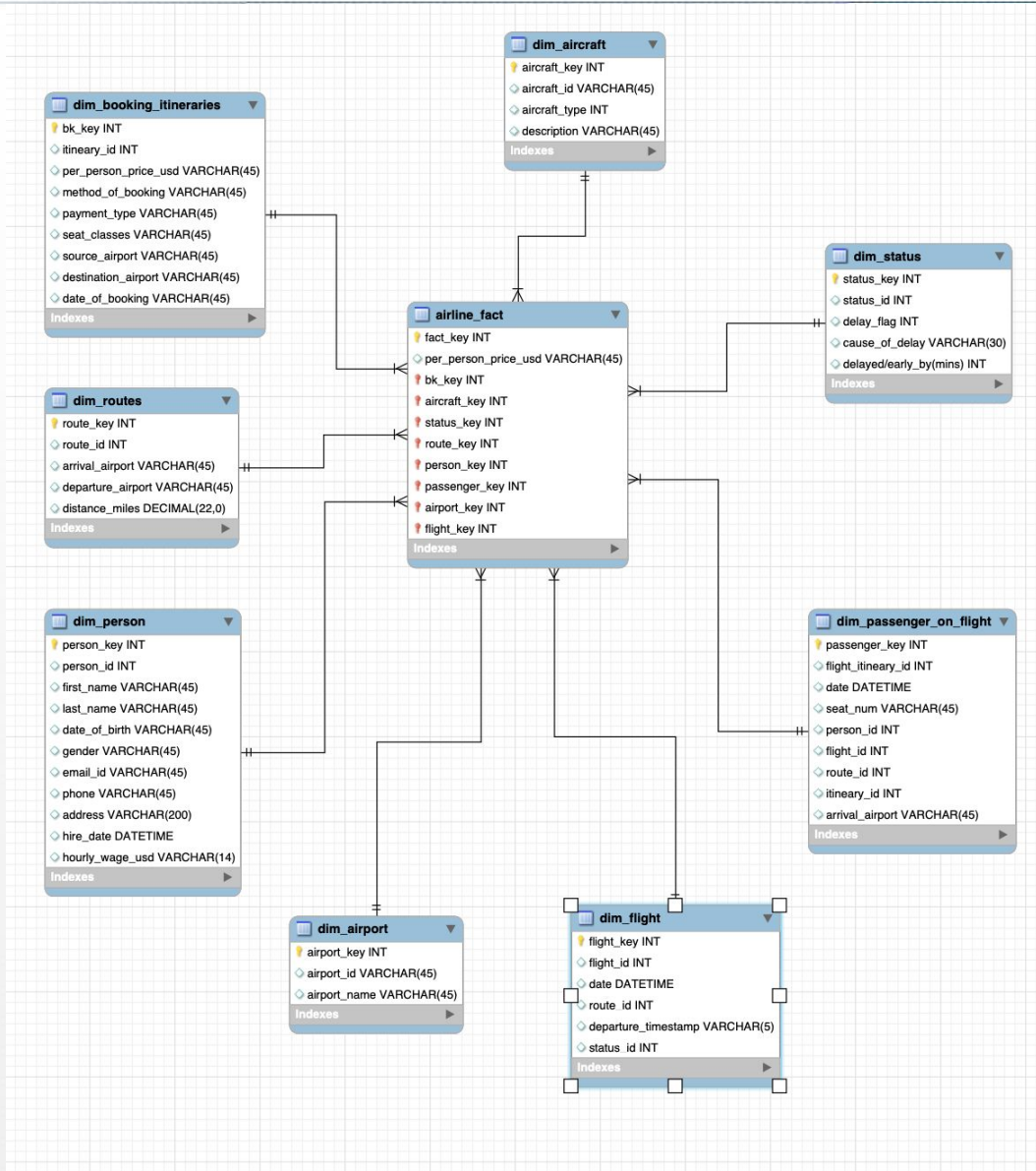
Predictive Analytics:

1. Cost prediction based on average aircraft delay projections
2. Offers or upgrades based on customer history
3. Increase in the flight frequency based on analysing overbooked flights for certain routes.
4. Forecast sales based on past promotions (Eg: sales during Christmas, Thanksgiving promotion)
5. Reduction in expenses by analysing past marketing costs (Eg: targeted marketing based on demographics/ economic status)
6. Analyze flight data based on aircraft type & weight, weather and the route, etc. to predict the optimal amount of fuel needed for the flight

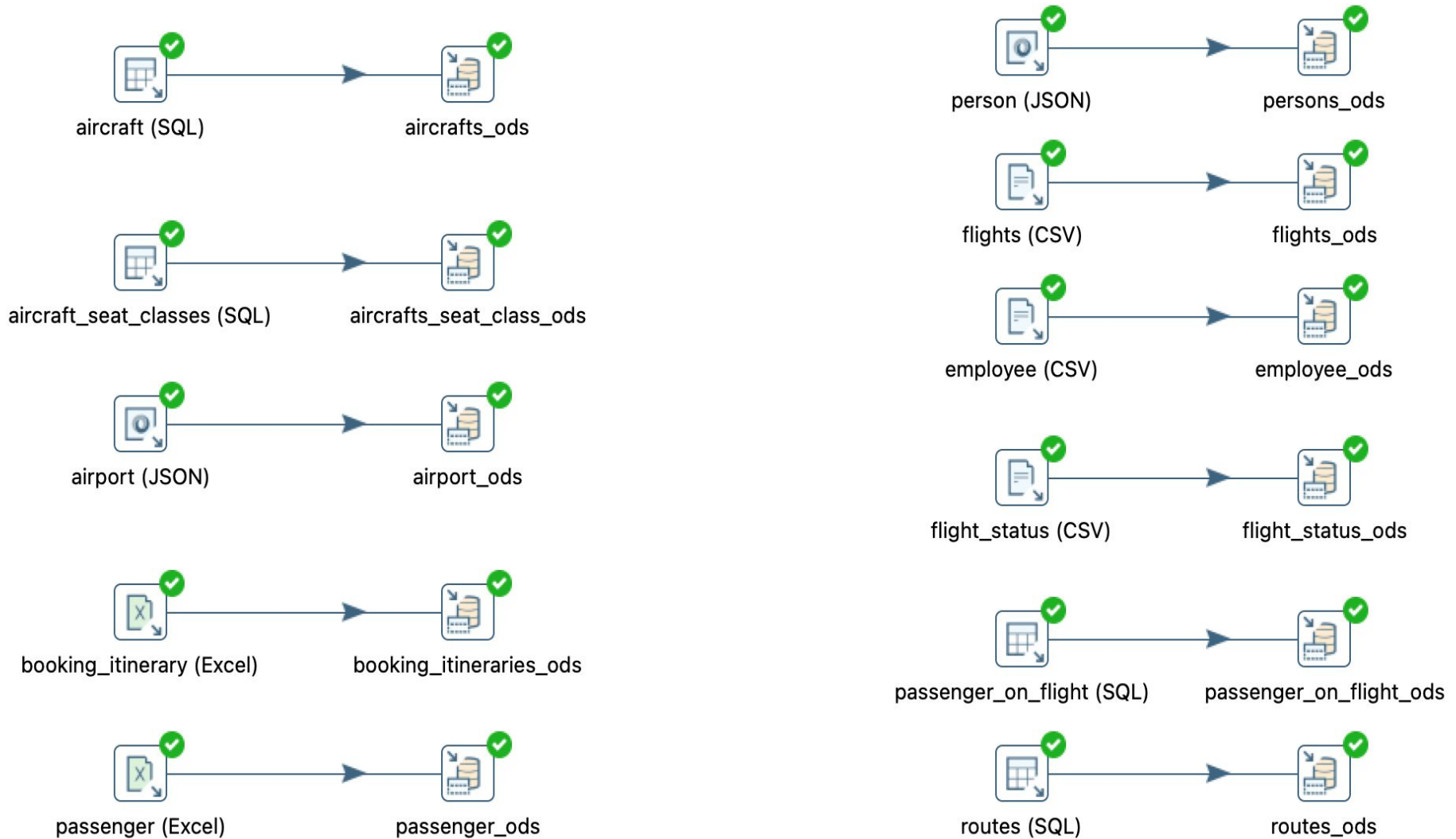
Airline Modeling System - OLTP Data Model



Airline Modeling System - Dimensional Data Model



ETL : Extract & Transform



ETL : Load



Combination lookup/update

Step name: `load_dim_airport`

Connection: `aerous wh` Edit... New... Wizard...

Target schema: `aerous_wh` Browse...

Target table: `dim_airport` Browse...

Commit size: `100` Cache size: `9999`

Pre-load the cache? ☐

Key fields (to look up row in table):

#	Dimension field	Field in stream
1	<code>airport_id</code>	<code>airport_id</code>
2	<code>airport_name</code>	<code>airport_name</code>

Technical key field: `airport_key`

Creation of technical key

☐ Use table maximum + 1

☐ Use sequence

☒ Use auto increment field

Remove lookup fields? ☐

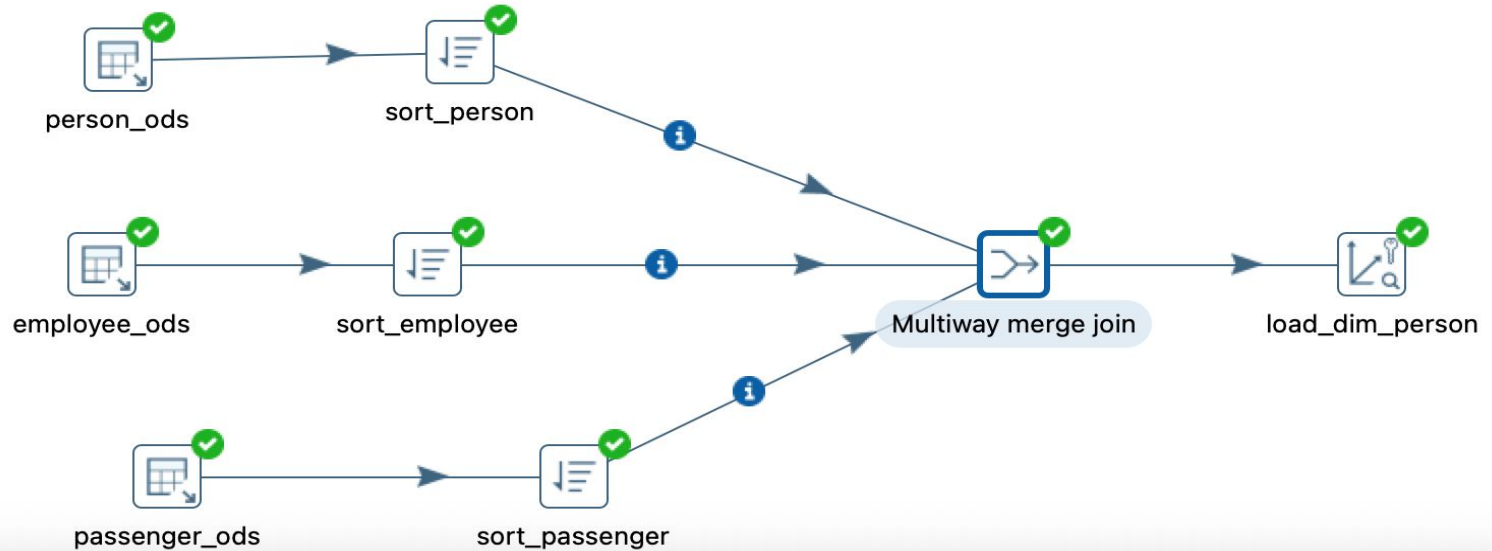
Use hashcode? ☐

Hashcode field in table

Date of last update field (optional)

Help OK Cancel Get Fields SQL

ETL : Transform & Load



Multiway merge join

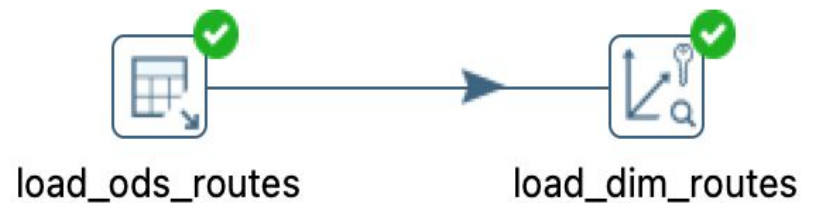
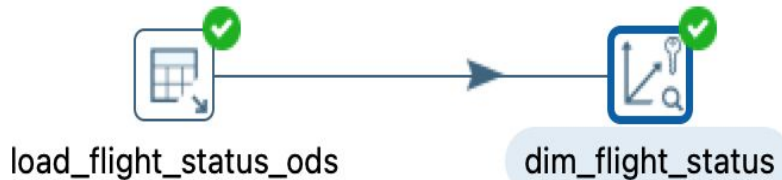
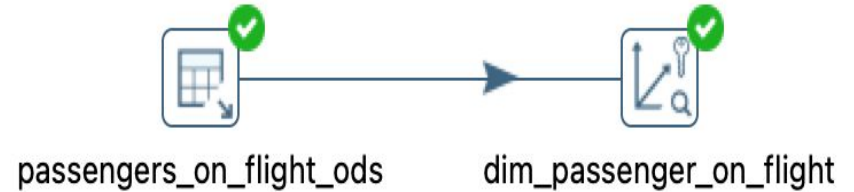
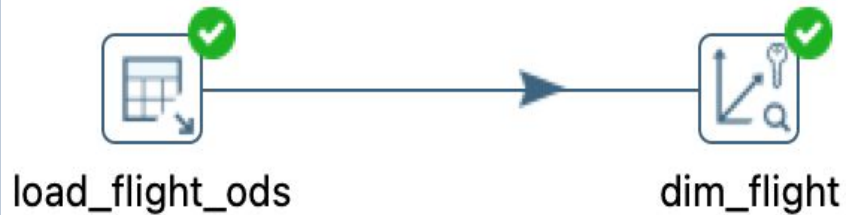
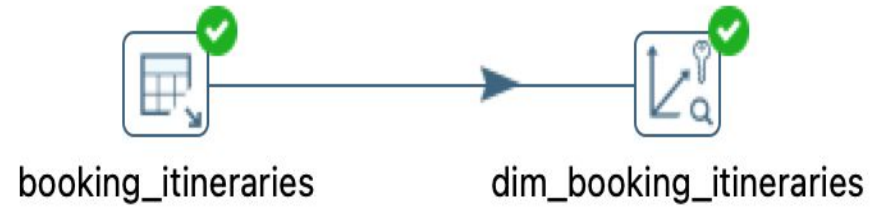
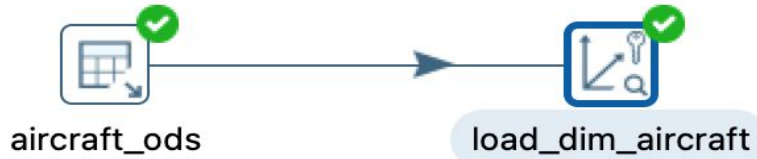
Step name	Input Step1	Input Step2	Input Step3	Join Type:
Multiway merge join	sort_person	sort_employee	sort_passenger	FULL OUTER

Join Keys

Join Keys	Select Keys
person_id	Select Keys
persons_person_id	Select Keys
persons_person_id	Select Keys

Help OK Cancel

ETL : Transform & Load



ETL : Transform & Load



Table input

Step name Table input

Connection aerous ods Edit... New... Wizard...

SQL

Get SQL select statement...

```
SELECT
flights_ods.flight_id
, flights_ods.date
, routes_ods.route_id
, flights_ods.flight_status_status_id
, routes_ods.airports_airport_id
, routes_ods.aircrafts_aircraft_id
, passenger_on_flight_ods.booking_itineraries_itineary_id
, passenger_on_flight_ods.flight_itineary_id
, passenger_on_flight_ods.passengers_persons_person_id
FROM flights_ods
INNER JOIN routes_ods
ON routes_ods.route_id = flights_ods.routes_route_id
LEFT OUTER JOIN passenger_on_flight_ods
ON passenger_on_flight_ods.flights_routes_route_id = flights_ods.routes_route_id
WHERE aerous_ods.flights_ods.date > '2017-01-01';
```

Line 1 Column 0

☐ Store column info in step meta data

☐ Enable lazy conversion

☐ Replace variables in script?

Insert data from step ▼

☐ Execute for each row?

Limit size 0

Help OK Preview Cancel

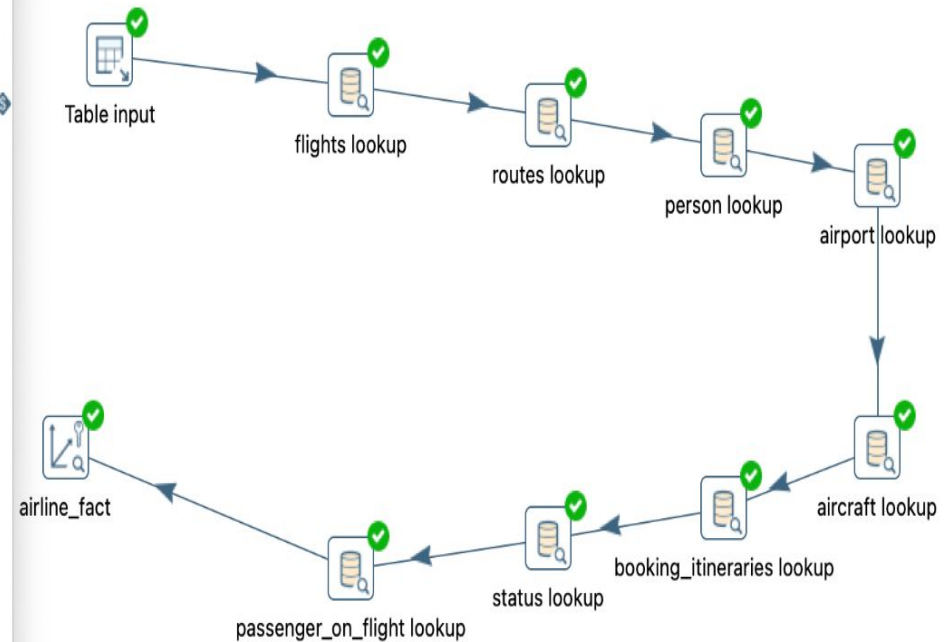
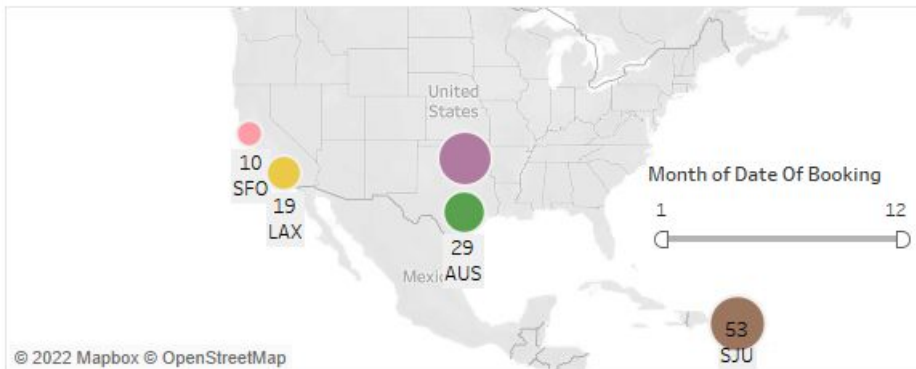


Tableau Demo - Descriptive Analysis

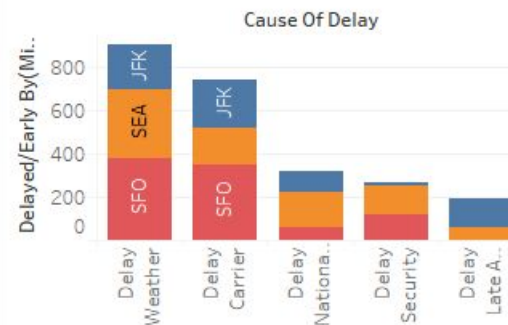


AeroUS Descriptive Analysis

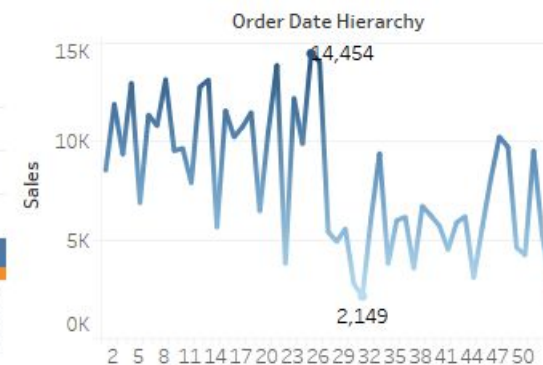
Top 5 Destination Airports



Top 3 airport having max delays in past year



Sales per Week



Increase in the number of flights in the last 1 year



Year of Date (Dim Pa..

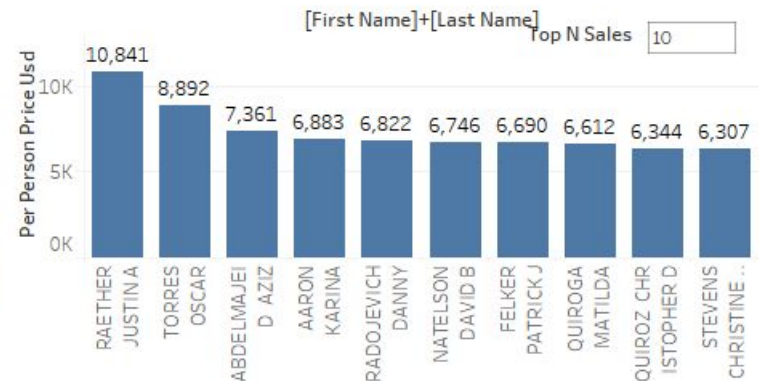
☐ 2018
☒ 2019
☒ 2020

Year
2019 2019

Month/Qua.. W...

Week
(All)

Top N customers generating max sales



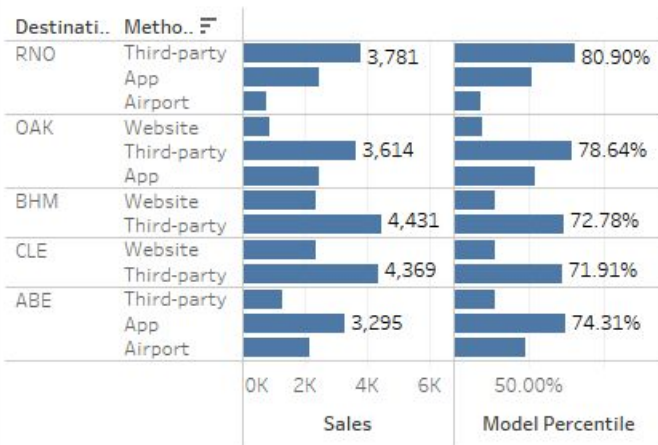
Top N Sales 10

Per Person Price Usd

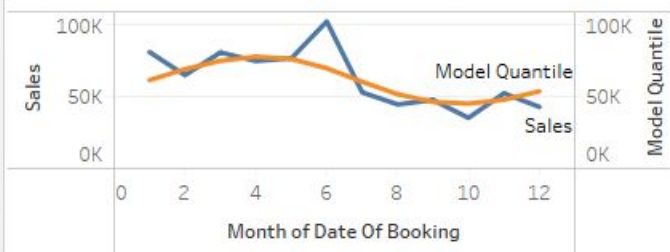
Tableau Demo - Predictive Analysis



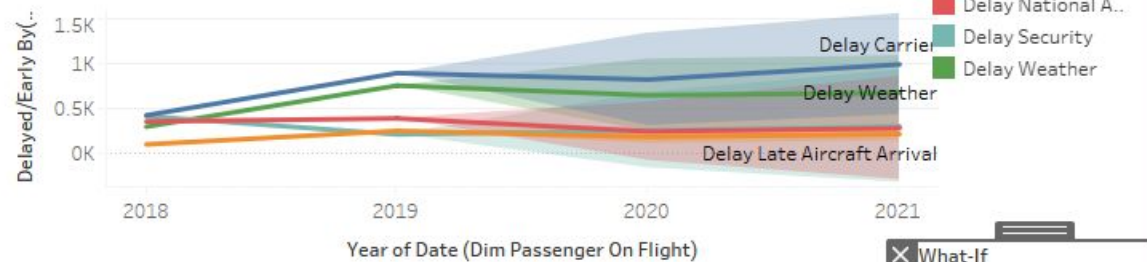
Model Percentile for Top 5 Destination Airports and Method of Booking



Model Quantile for Sum of Sales



AeroUS Predictive Analysis
Delay Forecast per Cause of Delay



What-If Analysis for Monthly Sales



Sales Forecast for Booking Methods



Key Learnings & Future Scope



Key Learnings:



Importance of good (clean, consistent) data



Ease that Data Transformation tools like Pentaho and Tableau have brought



Data Visualization helps in comprehending data at a glance and also brings focus to the quality of data

Future Scope:

More data would help us answer the following business questions:



Number of booking cancellations 24 hrs before departure (Lead Measure)



Customer/Employee Satisfaction (Lag Measure)



Offers or upgrades based on customer history (Predictive)



Top 5 fully-booked flight routes in the last year (Descriptive)



THANK YOU

