

# Airline Data Warehouse Documentation



<b>Ahmed NasrEldin</b>	<b>an2071497@gmail.com</b>
<b>Ahmed Otifi</b>	<b>Ahmedotifi11112000@gmail.com</b>
<b>Bassant Hassan</b>	<b>Bassanti355@gmail.com</b>
<b>Mohamed sakr</b>	<b>sakrm276@gmail.com</b>
<b>Mariam Eid</b>	<b>mariiameidd@gmail.com</b>
<b>Rodina Ashraf</b>	<b>rodinaashraf515@gmail.com</b>

# 1. Overview

This document provides a comprehensive view of the Airline Data Warehouse, including business processes, dimensional modeling, fact and dimension tables, and reporting metrics.

## Business Goals

- Monitor and improve reservation management.
- Analyze frequent flyer behavior and loyalty programs.
- Track customer interactions and complaint resolutions.
- Optimize financial and operational performance.

## 2. Data Flow Architecture

This data warehouse integrates AI-generated insights, cloud storage, a data warehouse, and a business intelligence tool, as illustrated in the diagram below:

1. **ChatGPT (AI Model)** – The process begins with AI-generated or external data being collected.
2. **AWS S3 (Storage Layer)** – The data is stored in an Amazon S3 bucket, which acts as a scalable data lake.
3. **Amazon Redshift (Data Warehouse)** – Data from S3 is loaded into Redshift for structured querying, analysis, and transformation.
4. **Power BI (Visualization and Reporting)** – Processed data in Redshift is then used in Power BI for reporting, dashboards, and business intelligence insights.

This setup enables AI-driven insights to be stored, processed, and visualized efficiently.



## 3. Dimensional Modeling Process

### 3.1 Business Processes

- **Flight Reservation:** Tracks ticket bookings, revenue, and costs.
- **Frequent Flyer Loyalty Program:** Captures passenger rewards and promotions.
- **Flight Activity:** Monitors trip performance, fuel usage, and seat occupancy.
- **Customer Care Interactions:** Tracks customer service interactions and complaint resolutions.

## 3.2 Grain Level

- **Reservation Fact:** One row per reservation.
- **Flight Activity Fact:** Per flight trip.
- **Frequent Flyer Fact:** Per loyalty program transaction.
- **Customer Care Fact:** Per customer interaction or complaint resolution.

## 4. Dimension Tables

- **Passenger Dimension:** Stores passenger demographics (age, gender, occupation, loyalty tier, etc.).
- **Airport Dimension:** Includes airport code, name, city, and country.
- **Aircraft Dimension:** Captures aircraft model, capacity, seating arrangement, and manufacturer.
- **Booking Channel Dimension:** Identifies reservation channels (website, travel agency, etc.).
- **Fare Basis Dimension:** Stores fare class and promotional offers.
- **Complaint Category Dimension:** Tracks complaint types and resolution timeframes.
- **Customer Experience Dimension:** Includes service aspects like priority check-in and boarding.
- **Date & Time Dimensions:** Used for analyzing trends across time.

## 5. Fact Tables

### 5.1 Reservation Fact Table

- **Grain:** One row per reservation.
- **Key Metrics:** Revenue, profit, fare basis, payment method, airport fees, taxes, fuel cost, and total cost.

### 5.2 Frequent Flyer Fact Table

- **Grain:** Tracks loyalty transactions.
- **Key Metrics:** Points earned/redeemed, operation type (flight, upgrade, cashback reward).

### 5.3 Flight Activity Fact Table

- **Grain:** One row per flight trip.
- **Key Metrics:** Flight duration, booked seats, fuel consumption, baggage weight.

### 5.4 Customer Care Fact Table

- **Grain:** One row per customer interaction.
- **Key Metrics:** Satisfaction score, response time, resolution status.

## 6. Reporting & Business Questions

- **Flight Reservation Analysis**
  - What are the preferred booking channels?
  - What is the total net revenue by quarter?
- **Frequent Flyer Insights**
  - How often do frequent flyers upgrade their seats?
  - What flights do frequent flyers mostly take?
- **Customer Service Performance**
  - What is the resolution time for complaints?
  - What airports have the highest complaint rates?

## 7. Physical Model & Optimization

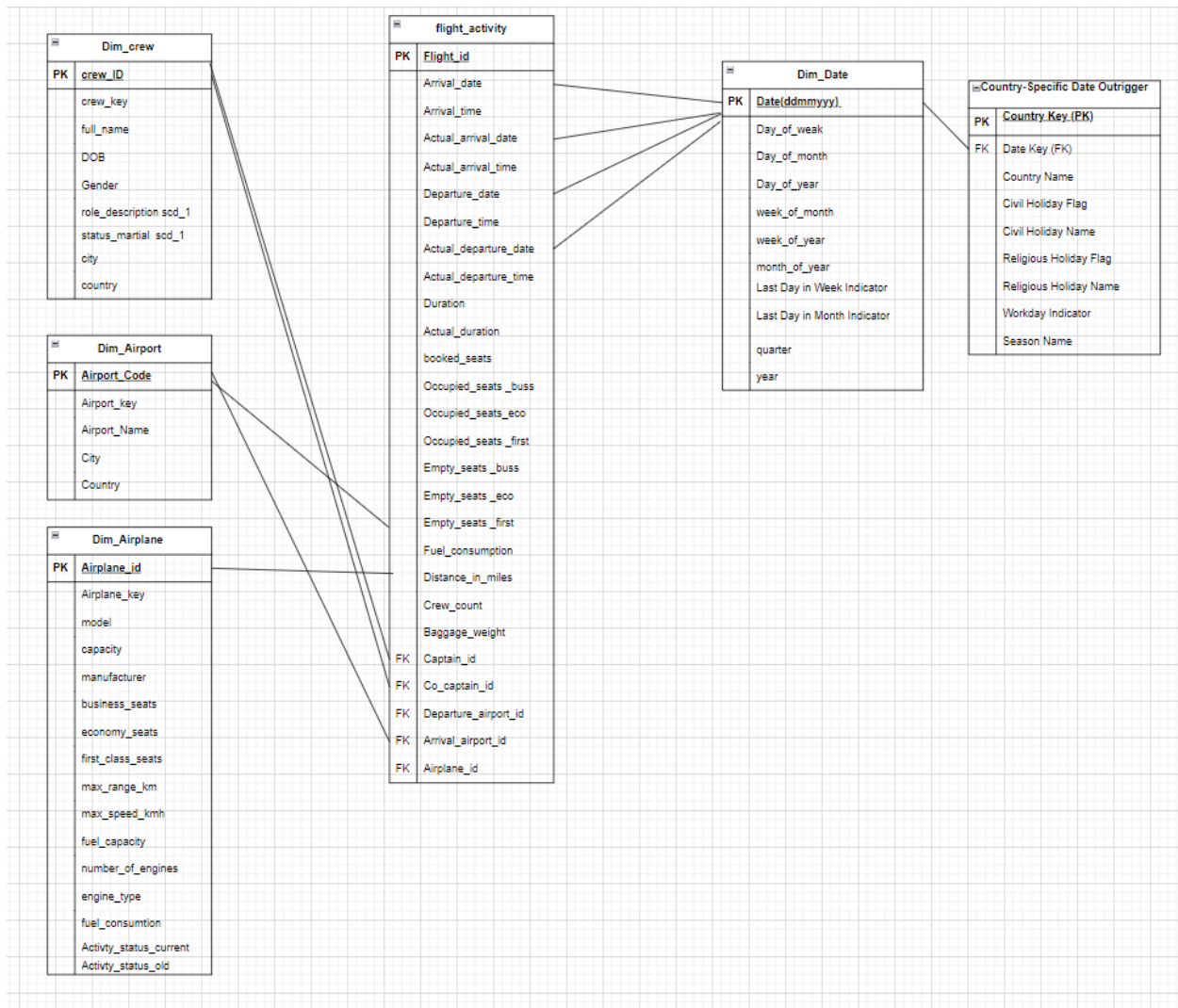
- **Amazon Redshift Distribution Strategies**
  - **ALL Distribution** for small dimension tables (Passenger, Date, Airport, etc.).
  - **KEY Distribution** for fact tables (using Passenger\_ID, Employee\_ID as distribution keys).

## 8. Conclusion

The data warehouse model ensures structured and scalable analytics for airline operations, allowing for efficient tracking of reservations, customer care interactions, flight activities, and loyalty programs.

# Logical Model

Flight Activity:



Loyalty Program and overNightStay

