



Travel Delay Across Airports & Carriers



Team 2: Kairan Zhong, Polaris Chen, Sunny Sun, Xiran Li, Yue Zhang

Table of Contents



01

Business Problem & Dataset

02

Data Models Showcase

03

Data Analysis

04

Business Insights & Solution

Business Problem & Dataset

Initial Statement & Current Statement & Data Overview



Business Problems

Initial Statement

For **carriers and airports** in the U.S:

- Scalable, self-service business intelligence solutions
- decrease airline delay rates, cut operating costs, and increase customer satisfaction



Data Processing

After Data modeling, ETL and Data visualization:

- no positive relationship between delay rates and carriers' operating costs



Current Statement

Shift the perspective to **the passenger side**:

- build a recommendation engine for passengers
- provide them with travel advice that can effectively avoid the risk of delays

Dataset Overview



On-time performance for over **100K** Domestic Flights across **23** airports and **20** carriers in United States & **1100** Associated Financial Reports from **2017 -2022**

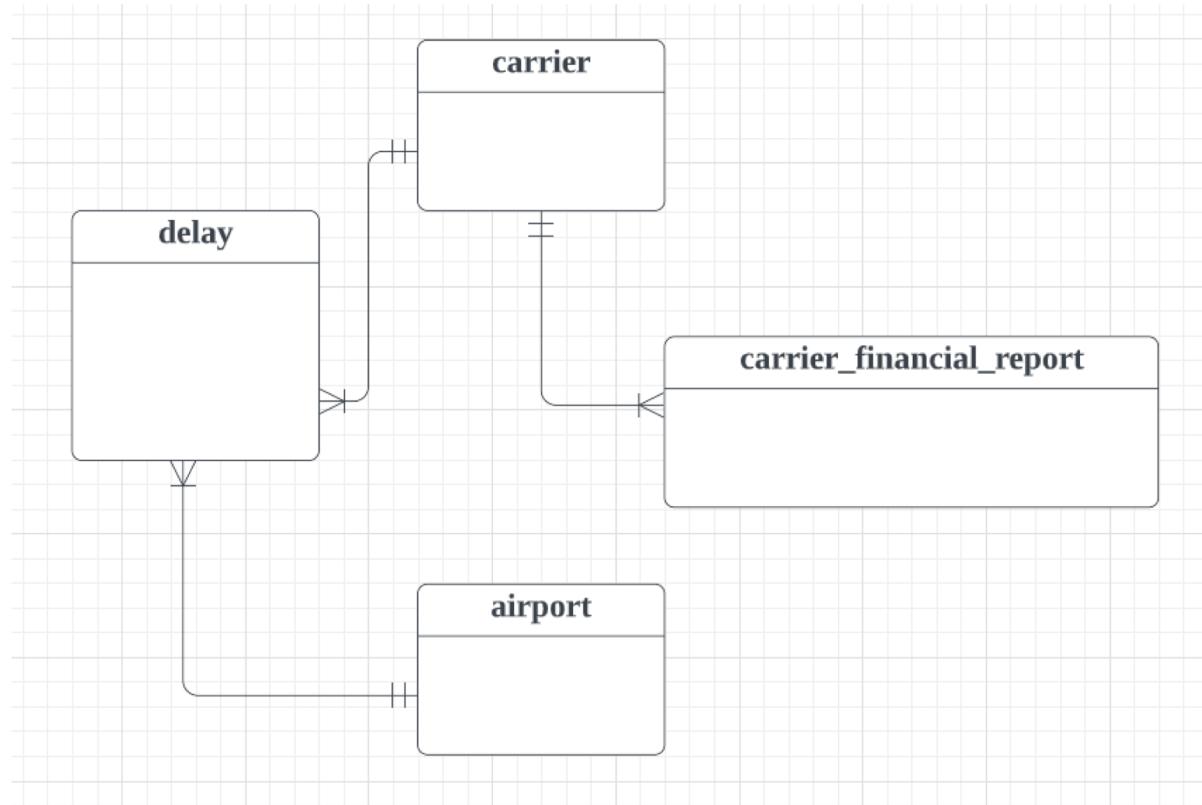
Dataset	Description	Data Source	Data Period	Anomalies
Airline_Delay_Cause, T_T100_MARKET_ALL_CARRIER, T_SCHEDULE_T3	On-time performance of domestic flights operated by large air carriers	The U.S. Department of Transportation's Bureau of Transportation Statistics	2017 -2022.05	High in analytical maturity, N.A. values account for less than 0.1%
T_F41SCHEDULE_B1	Quarterly operating balance statements for U.S. Air carriers			

Data Models

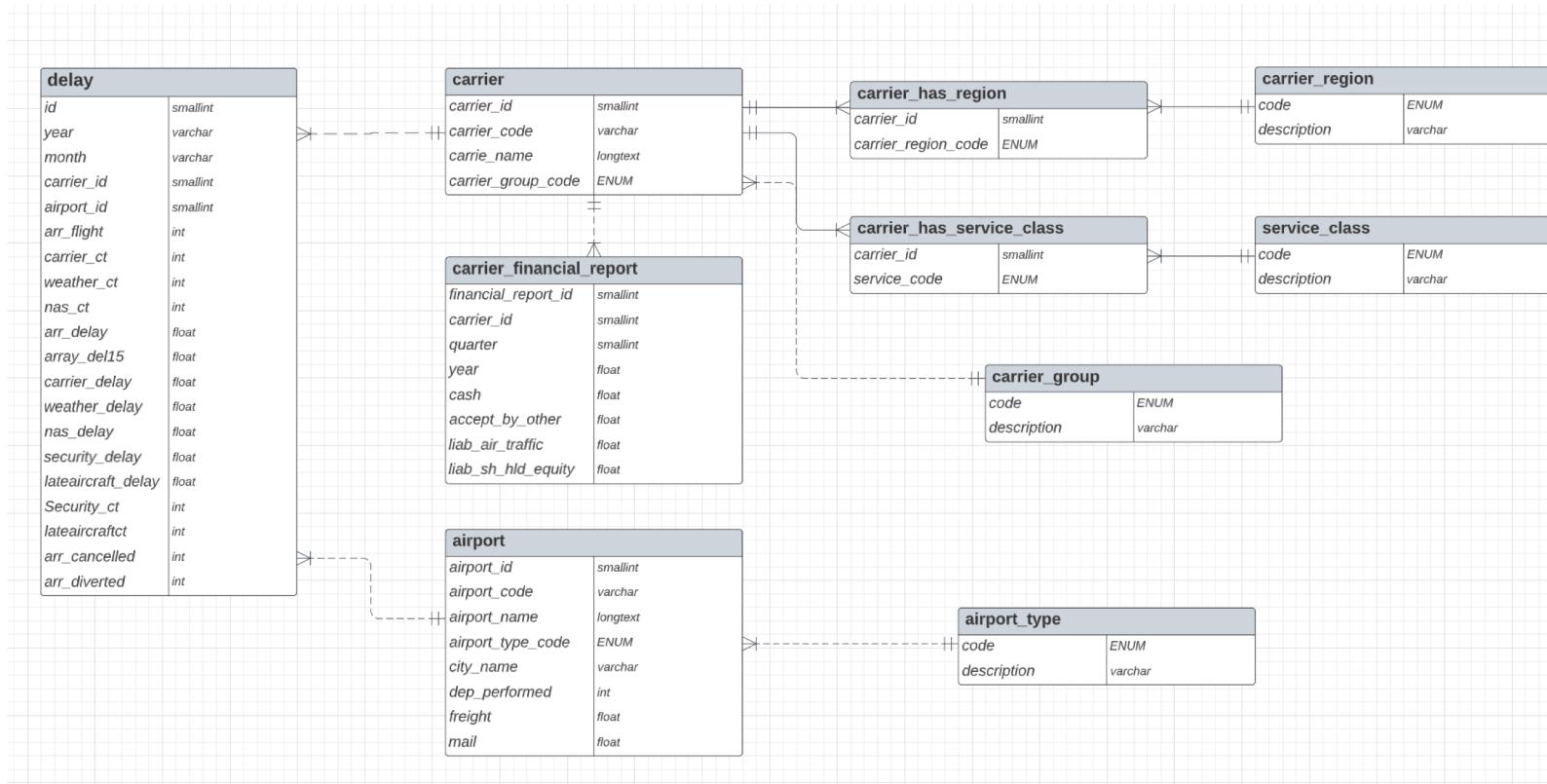
Conceptual Model & Logical Model & Physical Model



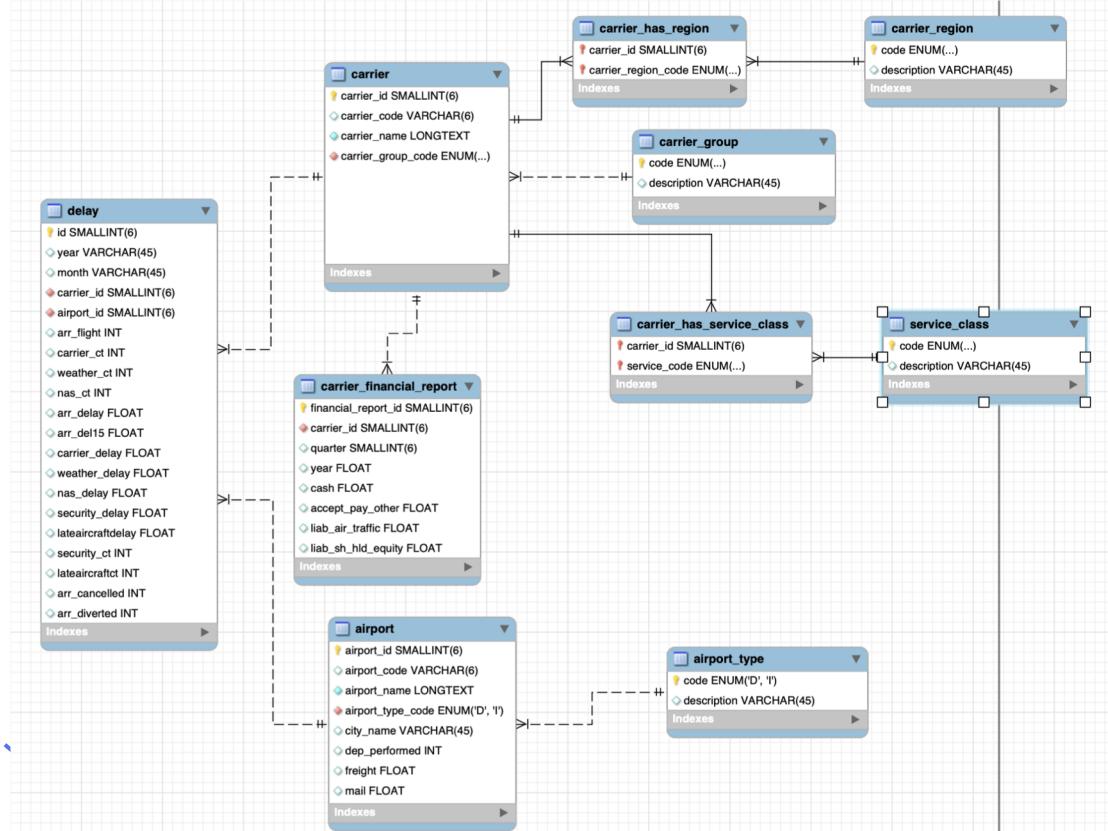
Conceptual Model



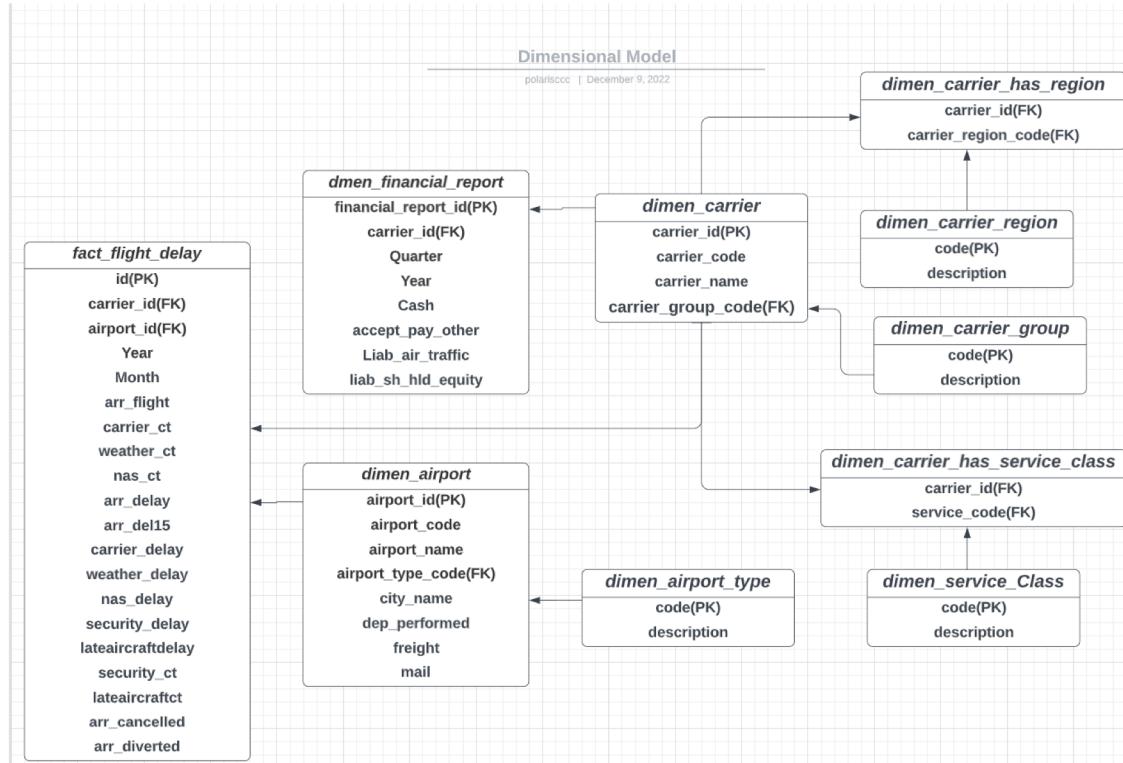
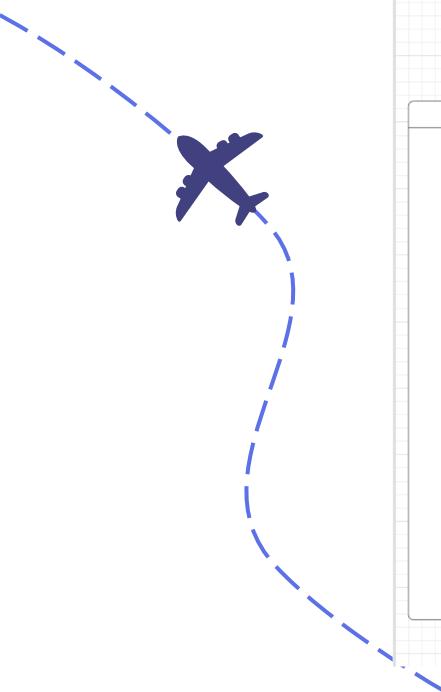
Logical Model



Relational Model



Dimensional Model



Implementation of ETL

Ingestion & Transformation & Loading & Cleaning



ETL: Data Issues

Problem

Carrier has M-M relationship
with region and service class



Original Dataset contains more
information than needed



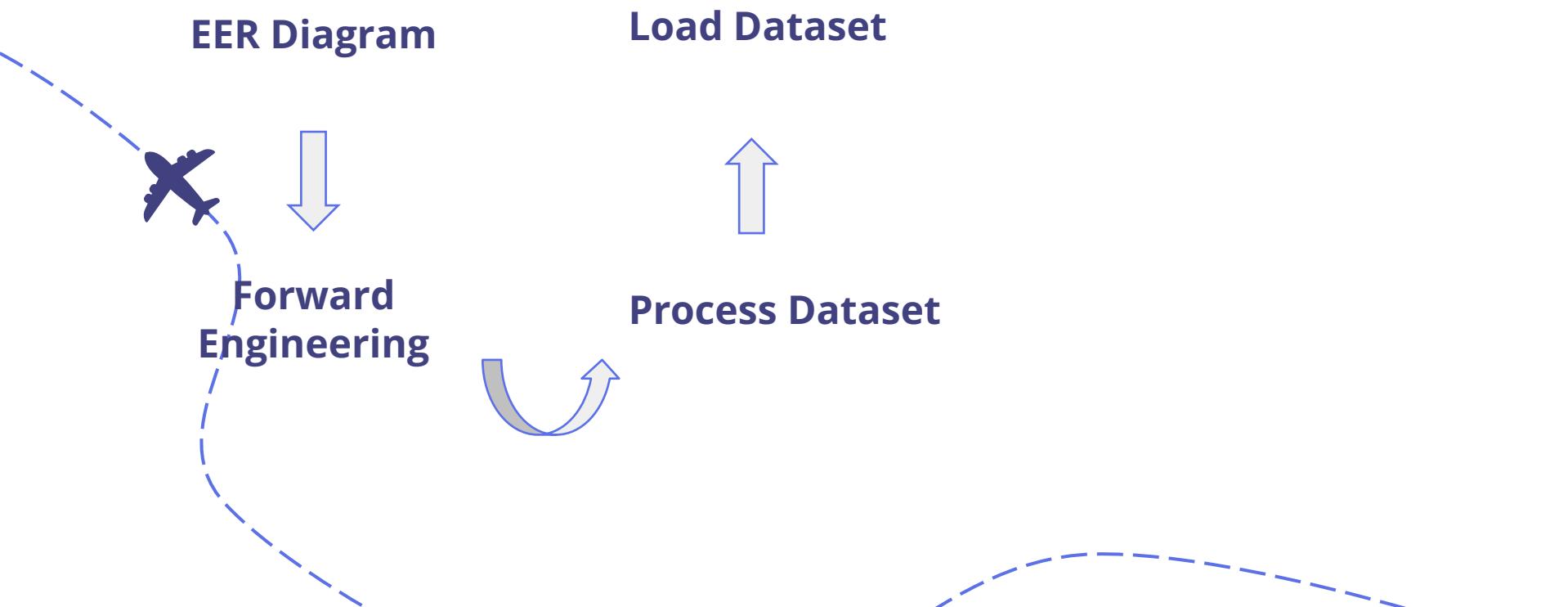
Solution

Add third table for
connection



Filter data when
loading into dataset

ETL: Process



ETL: Lesson Learned



- Add surrogate key for easier import
 - Normalize dataset into 3NF before loading data
 - Use data type LONGTEXT to store long strings





A dark blue airplane icon is positioned at the top of the dashed line, flying upwards.

Data Analysis

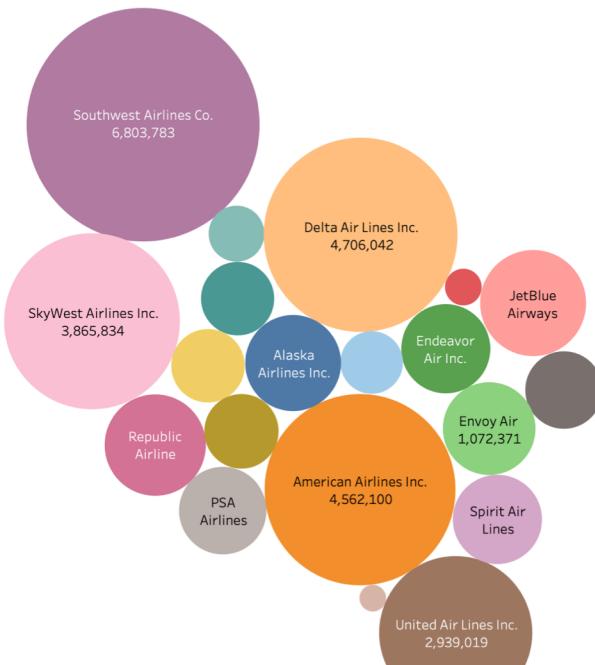
Airline Delay Records & Airline Financial Statements

Main Reasons of Delay

- Air Carriers (no crew)
 - ~ **35.8%** of Delay Time
- Weather
 - ~ **5.7%** of Delay Time
- Security Breach
 - ~**21.6%** of Delay Time
- Previous Late Flight
 - ~**36.7%** of Delay Time

Main Air Carriers

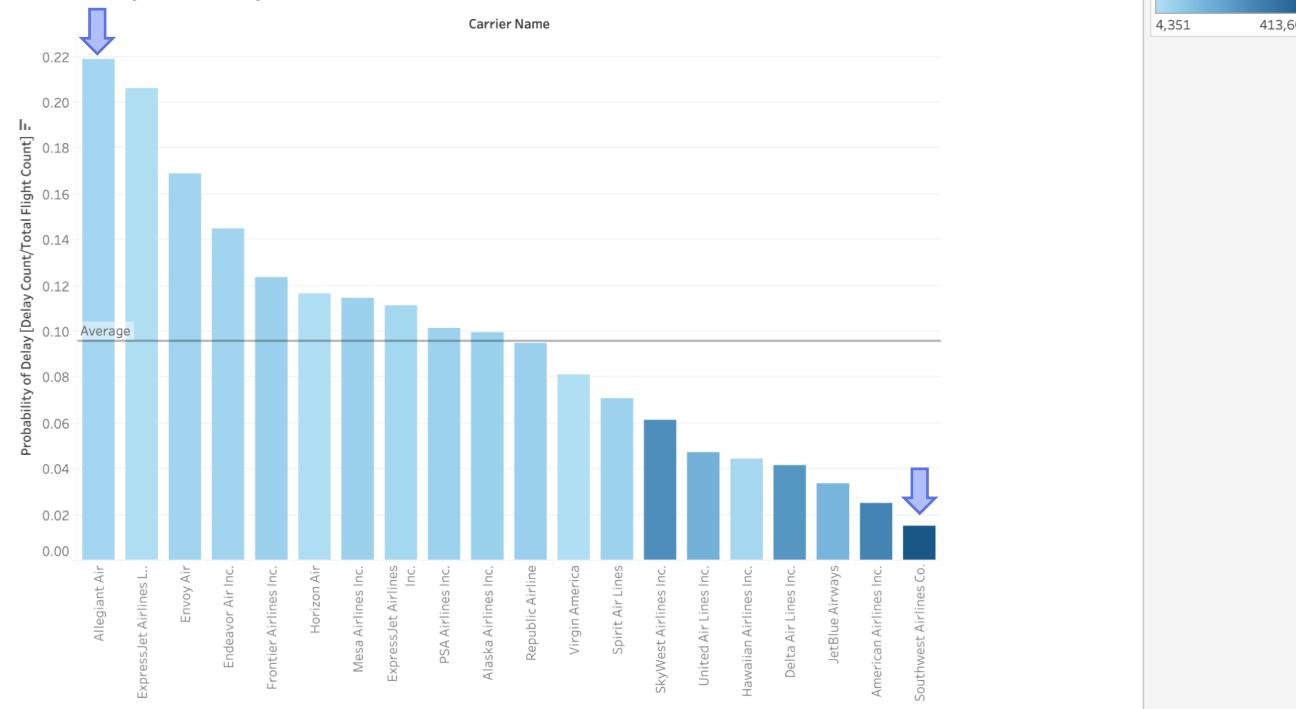
Total Flight Counts from 2017 - 2022



Carrier Delay Data Analysis



Carrier Delay Probability Distribution



Allegiant Air:

- Highest delay rate
- Low flight counts

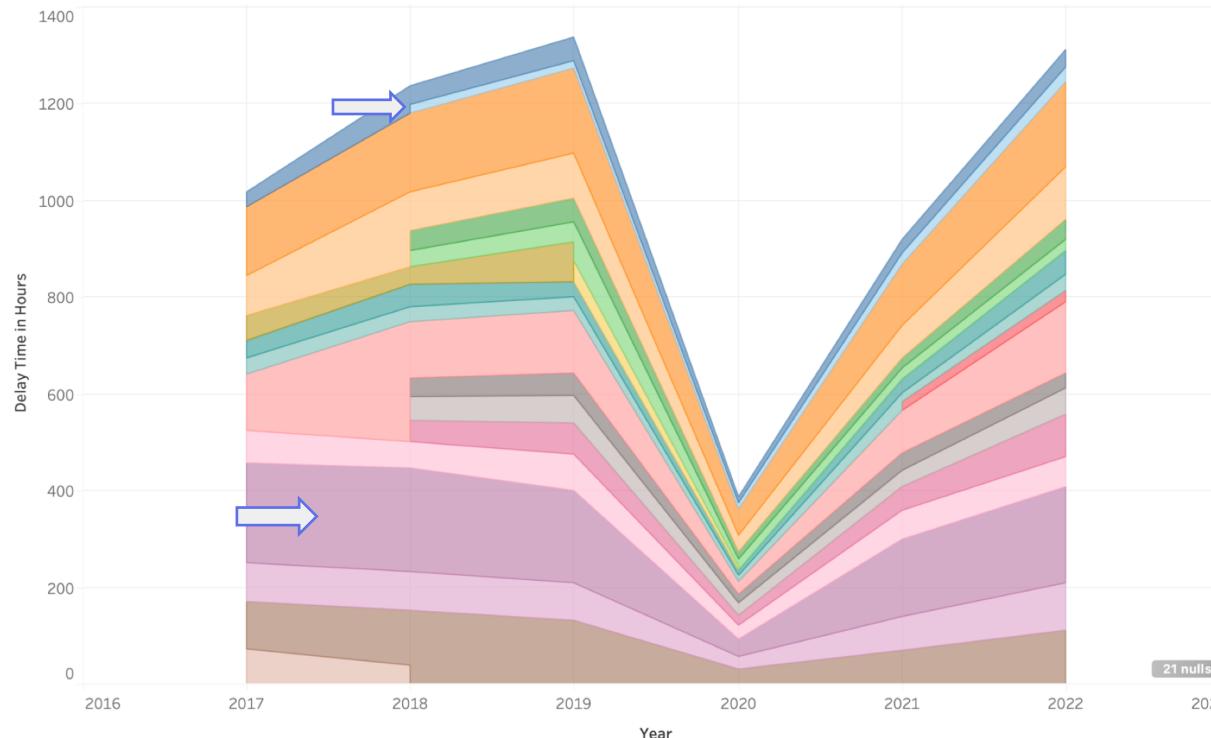
Southwest Airline:

- Highest flight counts
- Lowest delay rate

Carrier Delay Time



Average Delay Time per Month across Carriers



Carrier Name
Alaska Airlines Inc.
Allegiant Air
American Airlines Inc.
Delta Air Lines Inc.
Endeavor Air Inc.
Envoy Air
ExpressJet Airlines I.
ExpressJet Airlines L.
Frontier Airlines Inc.
Hawaiian Airlines Inc
Horizon Air
JetBlue Airways
Mesa Airlines Inc.
PSA Airlines Inc.
Republic Airline
SkyWest Airlines Inc.
Southwest Airlines C.
Spirit Air Lines
United Air Lines Inc.
Virgin America

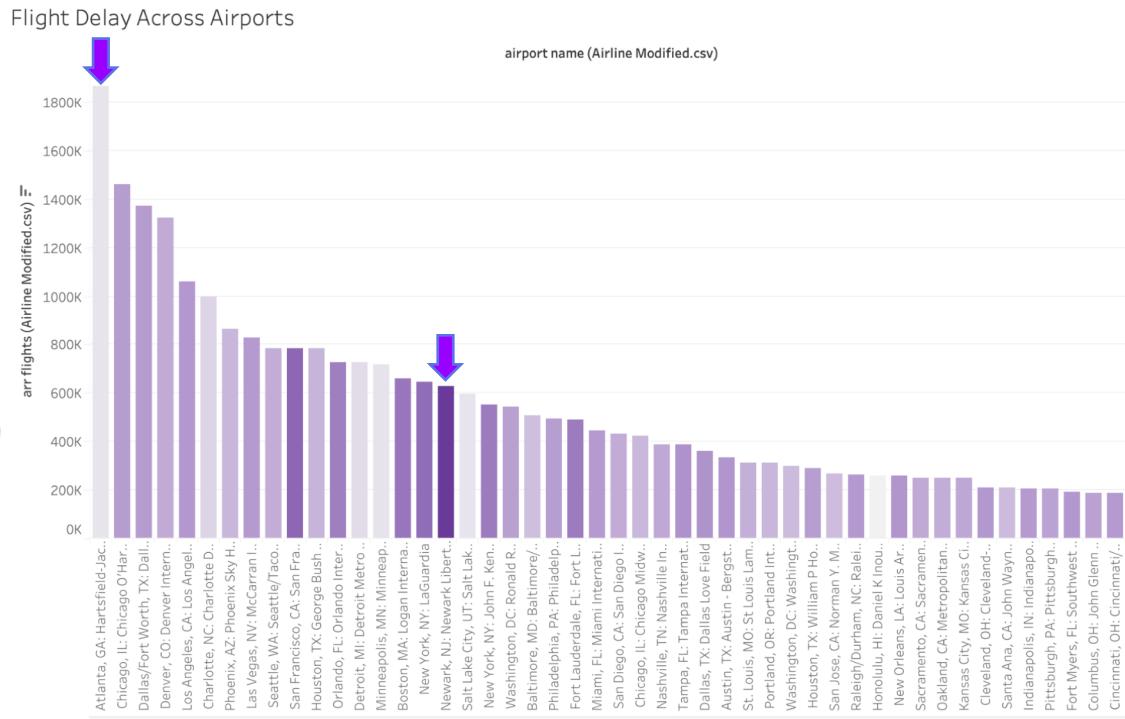
Allegiant Airline has 2nd highest average delay time.

Southwest has close to the least average delay time.

Airport Delay Analysis



Flight Delay Across Airports



Airport to fly from

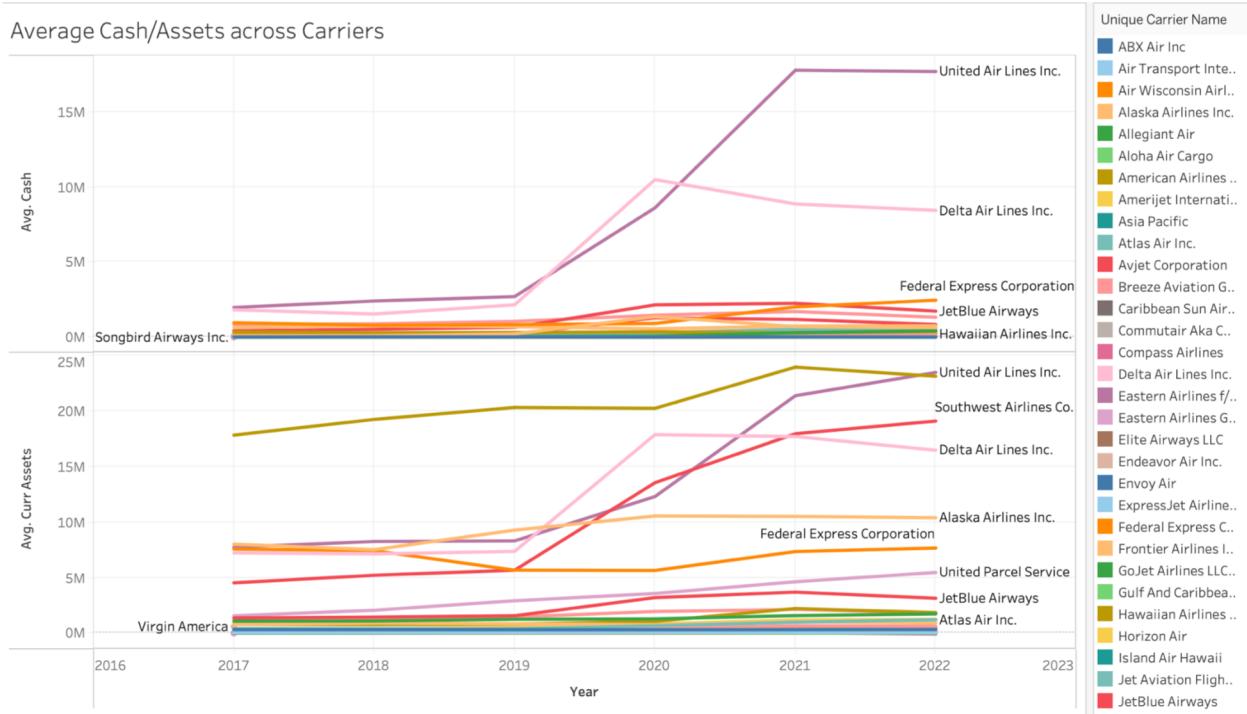
Worst:
Newark Liberty
International Airport

Best:
Hartsfield-Jackson
Atlanta International
Airport

Carrier's Cash Trend



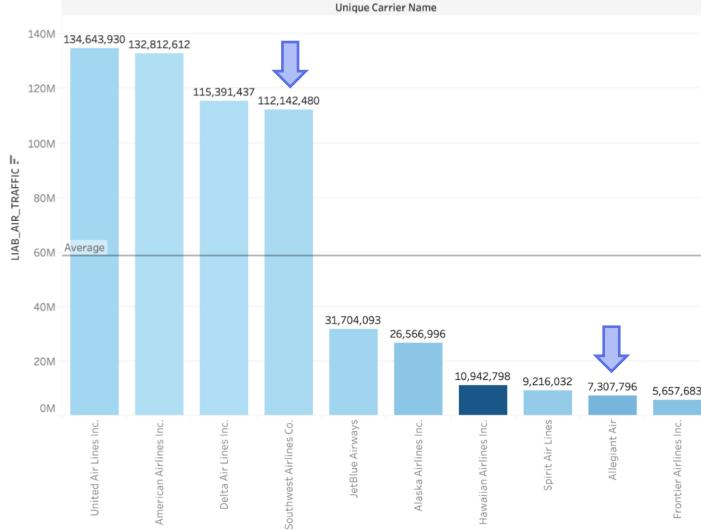
Average Cash/Assets across Carriers



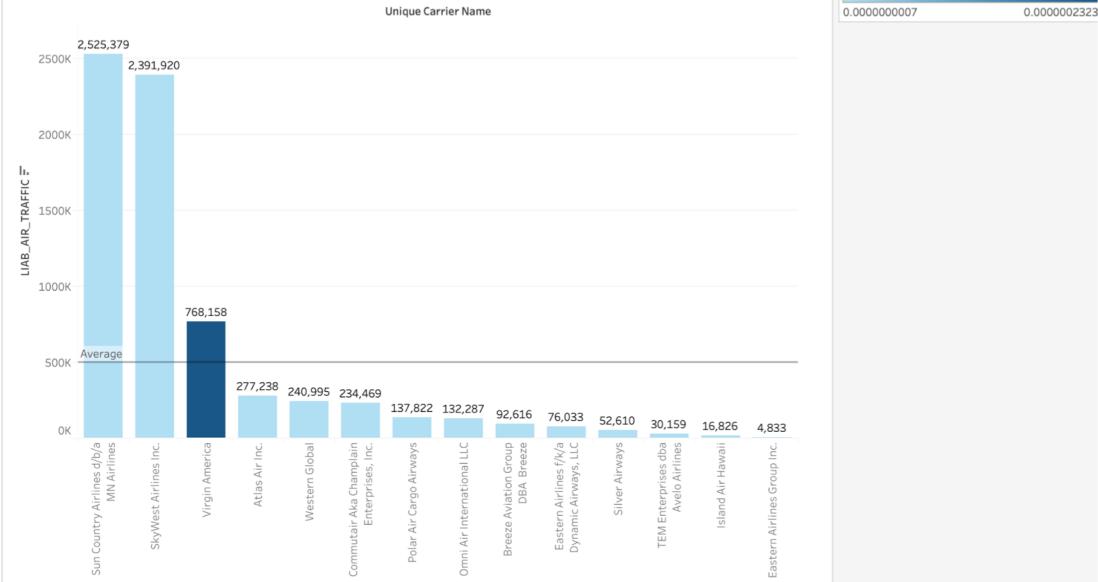
United Air,
Delta Air &
Southwest
all have
low delay rate
& high volume

Air Traffic Operational Costs

Air Traffic Operational Costs across Airlines (Above 3M)



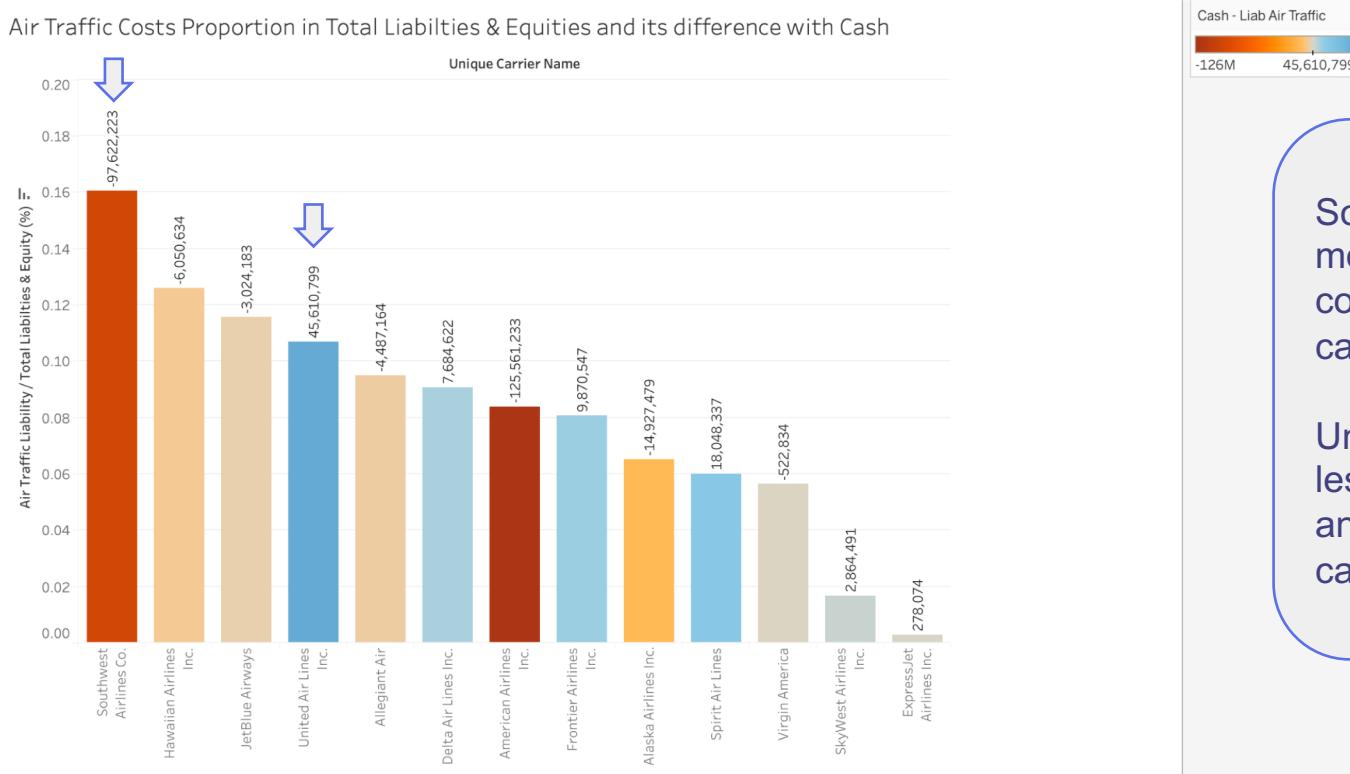
Air Traffic Operational Costs across Airlines (Below 3M)



Despite that **Allegiant Air has the highest delay rate**, it has **similar cost per flight** as other major carriers.

Air Traffic Costs Comparison

Air Traffic Costs Proportion in Total Liabilities & Equities and its difference with Cash



Southwest spends the most on air traffic costs and has little cash.

United Airlines spends less on air traffic costs and has much more cash.



Business Insights & Solution

Airport & Carrier Recommendation
Engine for Customers



BI & Solution



Insights

- Carrier delay time and delay probability has risen to a new high in 2022.
- Carrier with high delay rate may not have high air traffic cost (no compensation for customers)
- Southwest is the top carrier.



Solution

- Carrier Recommendation Engine
 - Input: Departure Airport, Available Carriers
 - Export: Best Recommended Carrier with low delay rate & short delay time.

Percentage of flights delayed by year (January through May)

Year	% delayed
2022	20.1%
2021	11.4%
2020	10.2%
2019	18.9%
2018	17.6%

Thanks for Listening

Questions?

