Meet the Team

Alekya Ghanta

Aishwarya Jain

Saimah Khan

Vikash Gujju

Srikar Alluri

Agenda

01	Executive Summary
02	Data Overview & Manipulation
03	Market Analysis – Findings & Insights
04	Modeling
05	Alert Mechanism
06	Challenges & Workarounds
07	Recommendations & Opportunities

01

Executive Summary

Executive Summary

Client Overview

QL2 provides the world's best solution for competitive pricing data, helping airline clients to reduce the time and money it takes to make informed and profitable **pricing decisions**

Experience
20+ years defining the data harvesting industry for airlines, expanded into 26 additional verticals

Scale
600 million+ records collected daily from 1000's of sites globally

200B+ data records

2+ years of historical pricing records from 400 companies across the world (150 public)

36 Million+ Products collected

5,000+ Sites monitored - easily available and always adding more

Business Problem

- Unclear impact of competitors and their pricing strategies on the markets

- Unknown price volatility trends across competitors

- Are pricing changes a resultant of inflection points in the data?

Objectives

- Identify distinctive patterns in international markets and assess how the carriers are influencing each others pricing strategies across various routes

- Implement a real-time monitoring system to identify shifts in pricing trends, capturing key inflection points

ROBERT H. SMITH
SCHOOL OF BUSINESS

Analysis Roadmap

Market Analysis

Analyze historical data to uncover pricing patterns for short, medium, and long-haul flights, aiming to understand the factors influencing airfare dynamics across different travel distances

Comparative Analysis

Comprehend the pricing behavior of competitors to discern patterns observed among airlines catering to the same or comparable market

Volatility Analysis

Assess pricing & volatility patterns of different carriers across different geographical locations

Predictive Modelling

Develop predictive models to forecast price difference among various airlines and understand dynamic shifts in the market

Alert Mechanism

Build a data model which can detect inflection points leading to shift in the pricing data

02

Data Overview & Manipulation

Data Source

We used data from QL2's database on Snowflake :

	COLLECT_DATE	FROM_AIRPORT	TO_AIRPORT	CXR	DDATE	DTIME	DROUTE	FARE
1	2021-01-06 08:41:00.000	PHX	HNL	HA	2021-02-19	08:00	08:00 - PHX :: 11:55 - HNL	245.00
2	2021-01-12 08:08:00.000	LAX	OGG	HA	2021-04-11	08:00	08:00 - LAX :: 10:50 - OGG	169.00
3	2021-01-06 04:11:00.000	HNL	SFO	HA	2021-01-13	14:15	14:15 - HNL :: 21:15 - SFO	248.00
4	2021-01-06 05:16:00.000	PHX	HNL	HA	2021-01-15	08:00	08:00 - PHX :: 11:50 - HNL	377.00
5	2021-01-06 02:13:00.000	LAX	OGG	HA	2021-01-09	08:00	08:00 - LAX :: 11:55 - OGG	248.00
6	2021-01-08 01:20:00.000	SEA	HNL	HA	2021-01-08	11:30	11:30 - SEA :: 15:45 - HNL	322.00
7	2020-09-24 22:15:00.000	FLL	ORD	NK	2020-10-01	20:20	20:20 - FLL :: 22:34 - ORD	86.09
8	2021-03-28 05:49:00.000	SDQ	FLL	NK	2021-04-21	13:42	13:42 - SDQ :: 16:04 - FLL	109.35
9	2021-03-28 05:49:00.000	SDQ	FLL	NK	2021-04-21	13:42	13:42 - SDQ :: 16:04 - FLL	109.35
10	2020-09-24 23:46:00.000	FLL	ORD	NK	2020-10-02	20:20	20:20 - FLL :: 22:34 - ORD	51.39
11	2021-01-06 01:49:00.000	SAN	HNL	HA	2021-01-08	07:40	07:40 - SAN :: 12:00 - HNL	278.00

Additional Data Sources :

Market Segmentation
IATA codes to Country and Region mappings

Passenger Count
Passengers travel count by time period from TSA & APIS Air Traffic Monitor

Manual Data Collection
Collected data manually on direct flights to Tokyo and London from Orbitz website

ROBERT H. SMITH
SCHOOL OF BUSINESS

Data Overview & Extraction

Before Data Cleaning :

392 Mn +	**142**	**1785**
AirFare data collected from 2018 to present	Airports tracked across 7 continents of the world	Unique Routes tracked

- The size of our primary table is above **2.5 GB** (.hyper tableau extract)
- Connected QL2's cmu_airfare schema to Python & Tableau for efficient data extraction
- Standardized externally sourced IATA codes & collated with airfare data to understand the arrival & departure locations
- Used the arrival & departure airport columns to remove domestic routes
- Narrowed down the data by filtering the records collected from the year 2023 to date

Data Cleansing

After Data Cleaning :

5.6 Mn + AirFare data after filtering for data collected after 2023	**12** Airport codes shortlisted based on data availability	**23** Unique Routes chosen to study pricing trends

✈ Removed data points with blank Arrival/ Departure Airports & routes with same Origin & Destination Airport Codes

✈ Deduplicated data using collection_date, departure_date, flight_time, carrier, origin, destination & fare columns

✈ Reduced the dataset's dimensionality by removing unnecessary columns like Website source, Flight number etc, resulting in a dataset with 8 columns

Feature Engineering

New Variables Created :

Geographic
- Departure Airport Country
- Arrival Airport Country

Temporal
- Days to Departure - No of days between the booking date and the flight date
- Day of the week
- Time of the day

Travel
- Duration of Travel - Used to segment the routes into short, medium & long
- Air miles - Distance between origin and destination airports

Economic
- Price per mile - Calculated as the total fare divided by air miles

Market Segmentation

4 destinations have been selected for this airfare project based on a thoughtful combination of popularity, data availability & distance

SHORT   MEDIUM   LONG

Los Cabos, Mexico

Cancun, Mexico

London, UK

Tokyo, Japan

03.a
Market Analysis - Japan

# LAX To NRT: United Airlines Pricing Pattern
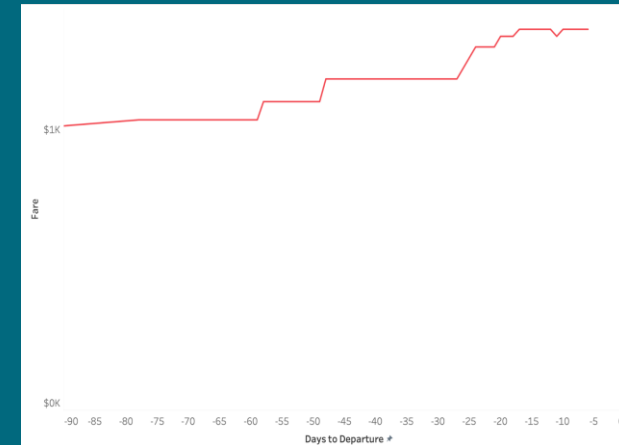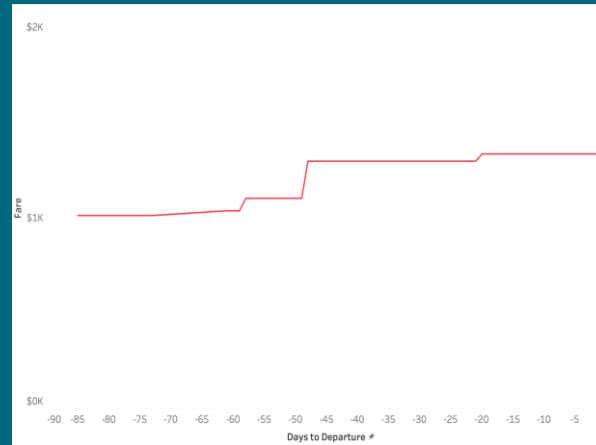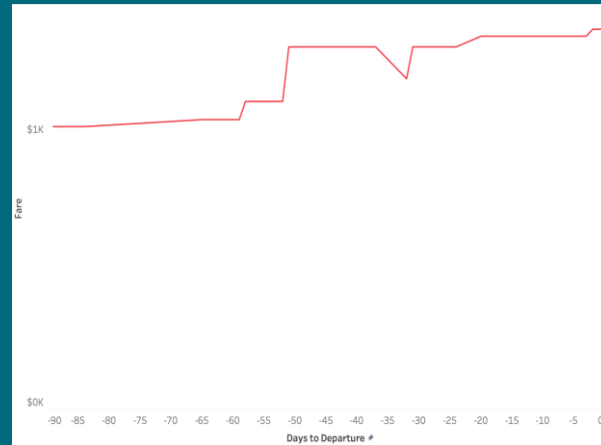
🌐 United Airlines consistently maintains stable pricing patterns on the LAX to NRT (Los Angeles International Airport to Tokyo Narita International Airport) route

🌐 Given the absence of competitors like American and Delta on the LAX to NRT route for United Airlines, there is no imperative for the airline to engage in aggressive price fluctuations, allowing for a more consistent pricing approach
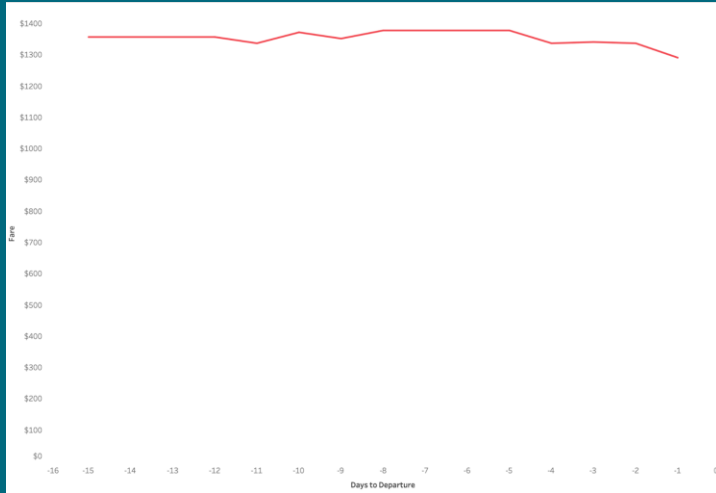
Delta & United pricing patterns - LAX to HND
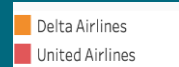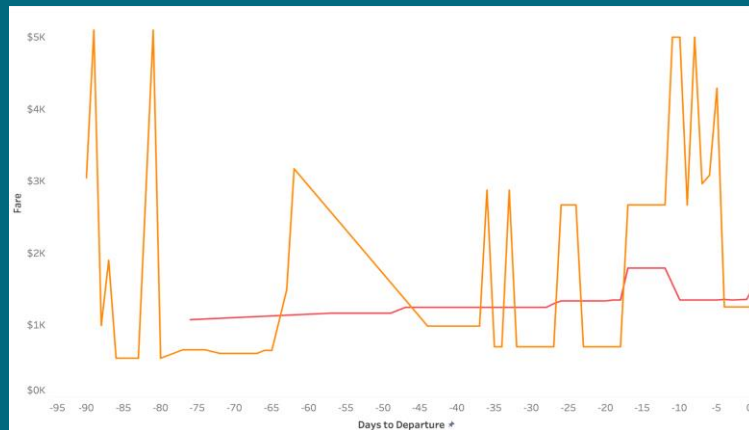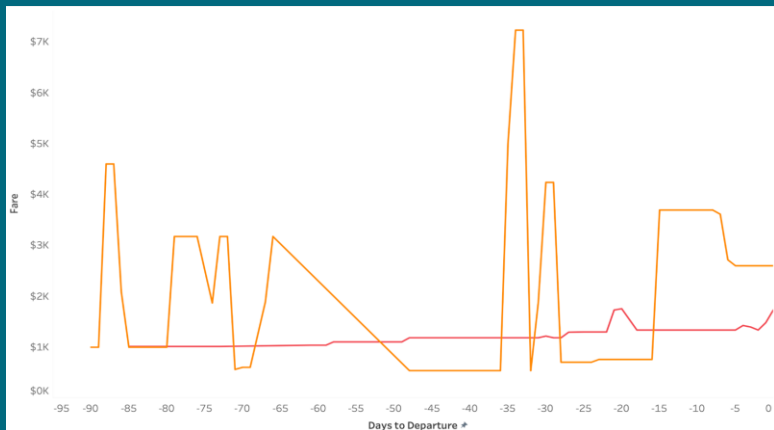
Delta Airlines:
- Exhibits significant fluctuations in ticket prices for the LAX to HND (Los Angeles International Airport to Tokyo Haneda Airport) route
- The airline may aggressively adjust prices to remain competitive in the market owing to the fact that the market is also served by American and United airlines as well
- Employs a more dynamic pricing strategy where day to day changes in demand and other market factors are considered

United Airlines:
- United Airlines, in contrast, maintains a more stable pricing structure for the LAX to HND route
- United's fares on this route are less prone to frequent changes, providing passengers with a more predictable and consistent pricing experience
- The airline may prioritize maintaining a steady market position and customer loyalty over engaging in frequent price fluctuations

Manually collected data substantiates our previous findings

- The manually collected data substantiates the observation of Delta's aggressive pricing behavior, showcasing the airline's propensity for frequent adjustments in ticket prices
- The consistency observed across multiple data points emphasizes United's steady fare structure for the specified route
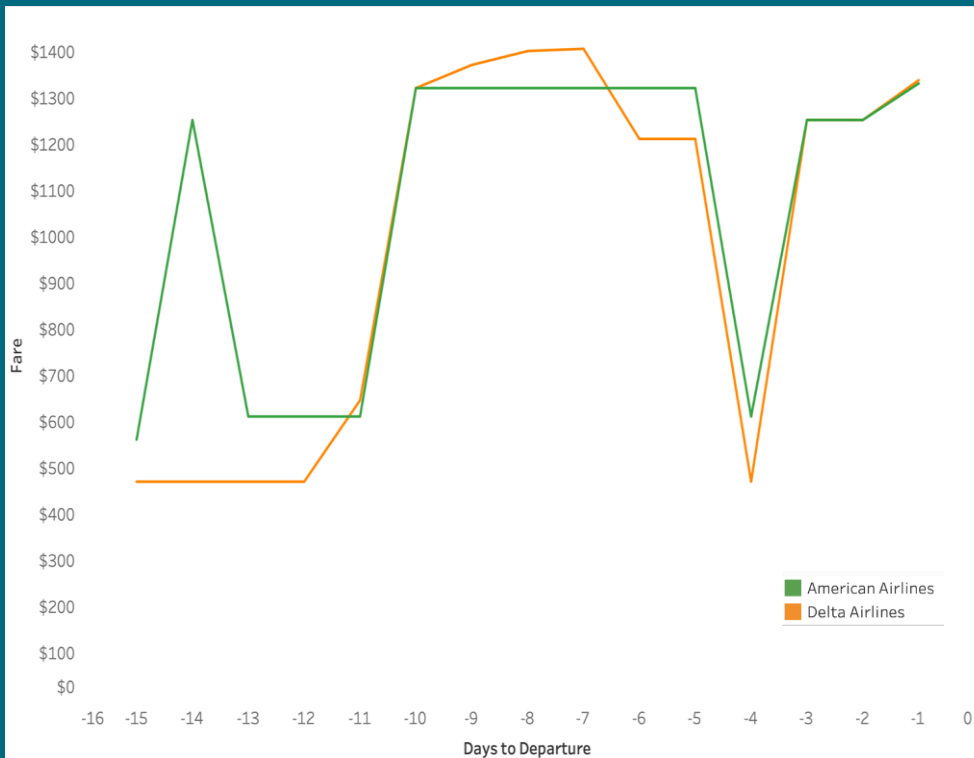- Analysis of the collected data may reveal patterns in Delta's dynamic pricing, reflecting fluctuations influenced by factors such as demand peaks, promotional periods, or market competition

American & Delta exhibit synchronous price movement (LAX – HND)

The simultaneous adjustments suggest a mutual awareness or responsiveness between the two airlines, implying that changes in one airline's pricing may trigger a swift response from the other

The concurrent price adjustments may be a deliberate strategy to uphold price parity, preventing either airline from presenting a significant advantage to prospective passengers
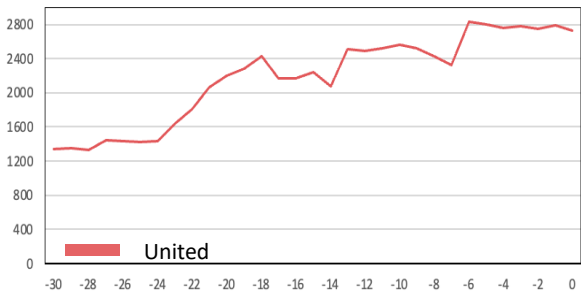
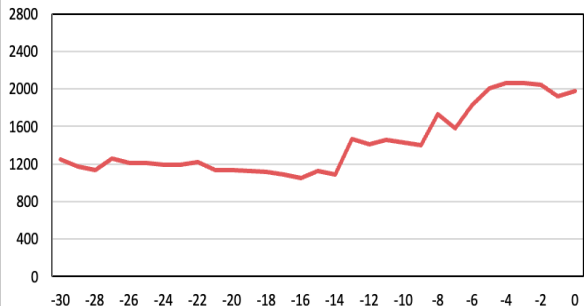# 03.b
Market Analysis – UK (London)

United remains stable harmonizing with the demand on the ORD-LHR route

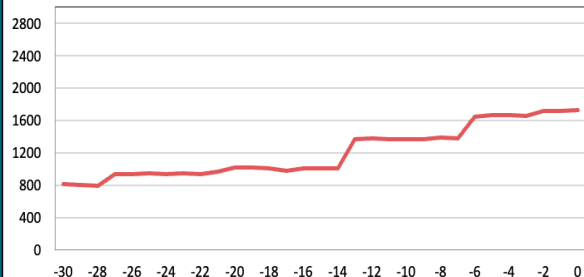## Average Fare trends in the ORD-LHR Route :
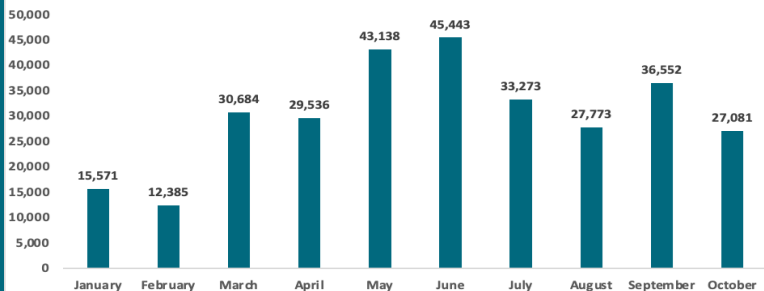

Average Fare - June 2023

Average Fare - July 2023

Average Fare - August 2023

Demand trends in the ORD-LHR Route :

Passengers travelled in ORD to LHR Route (Year 2023)

▷ The absence of direct competition from Delta in this segment enables United to focus on matching with airlines like American Airlines & British Airways(One world alliance), while responding to demand dynamics in the market

▷ The pricing trends remains relatively stable but tends to decrease with a drop in demand, aligning with typical airline pricing strategies in competitive but not oversaturated markets

ROBERT H. SMITH
SCHOOL OF BUSINESS

Price Volatility

Boston to London:
Delta's presence may intensify competition, leading to more frequent pricing adjustments by United to maintain competitiveness
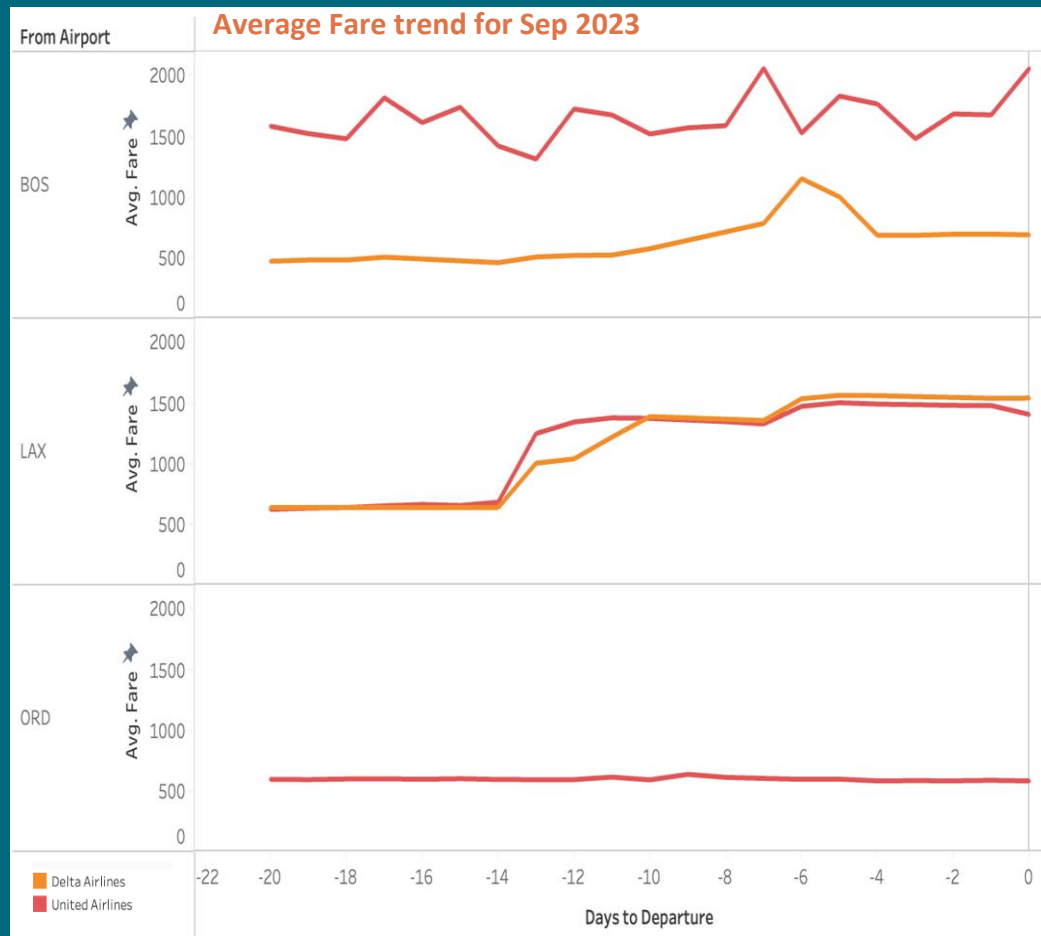
Los Angeles to London:
Similar price trends may indicate that both United and Delta are responding to common market factors affecting the Los Angeles to London route
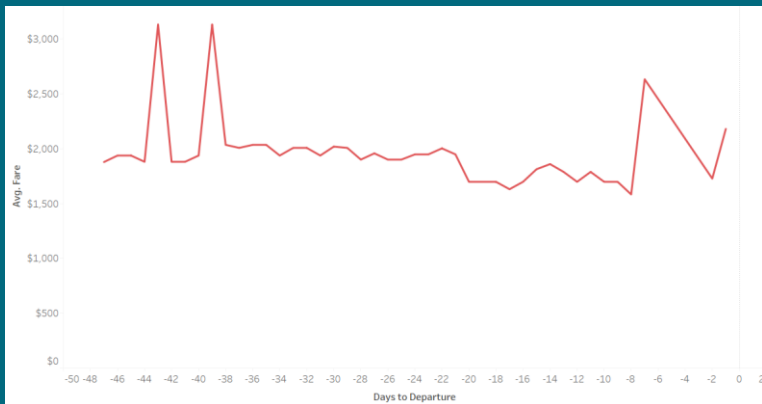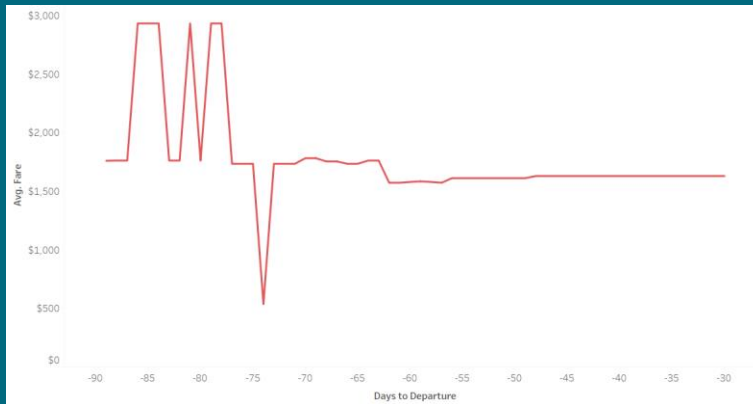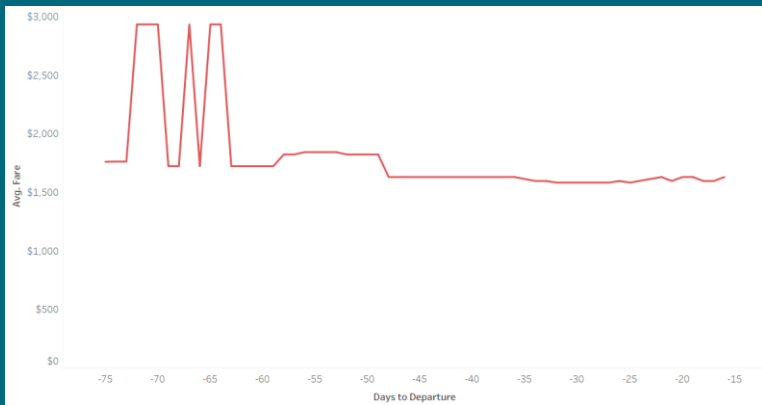
Chicago to London:
The absence of Delta may result in lower competitive pressure, allowing United to maintain stable prices

United Airlines exhibits significant volatility well before departure

▷ Early volatility peaks suggest that United adjusts prices dynamically, possibly responding to anticipated demand fluctuations or strategic pricing adjustments

▷ United strategically varies fares in response to different stages of the booking timeline. Understanding these dynamics allows for better anticipation and adaptation to market conditions

▷ Conduct a deeper dive into the factors influencing volatility during early, stable, and late stages

ROBERT H. SMITH
SCHOOL OF BUSINESS

Alliance Influence: A Comparative Analysis of American & British Airlines (NYC to LHR)

One World Alliance Connection:
The parallel pricing patterns suggest a potential correlation between the two carriers, prompting an exploration into the influence of their alliance

Collaborative Pricing Strategies:
Alliance membership often involves collaboration in various aspects, including pricing strategies, route coordination, and shared customer benefits

Joint Marketing Initiatives:
Alliance members may engage in joint marketing efforts, influencing fare structures and promotional activities

Shared Market intelligence:
Shared information may contribute to aligning pricing strategies based on a mutual understanding of market conditions

03.c

Market Analysis – Mexico
(Event Data Analysis)

Leisure travel - US to Mexico

- Tourists flock to Mexico for its vibrant culture, stunning beaches, and ancient ruins. From the bustling metropolis of Mexico City to the idyllic Mayan ruins of Chichen Itza, Mexico has something to offer everyone

- From the dataset there are two major routes Chicago to Cancun, and Los Angeles to Los Cabos that we have analysed

- Most popular festivals:
  - Festival Internacional Cervantino (100000)
  - Guelaguetza (80000)
  - Fiesta del Sol (60000)
  - Cumbre Tajín (40000)
  - Festival de Cine de Morelia (20000)

- October draws the highest number of tourists to Mexico, thanks to a combination of factors, including pleasant weather (sunny days with comfortable temperatures), and a hurricane-free season

United Airlines- Peak Season Pricing
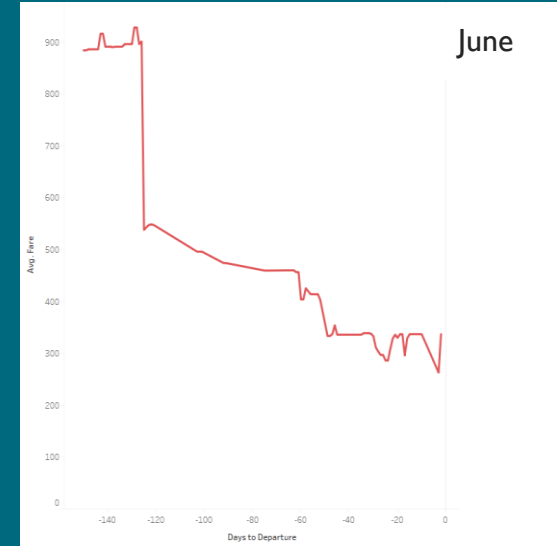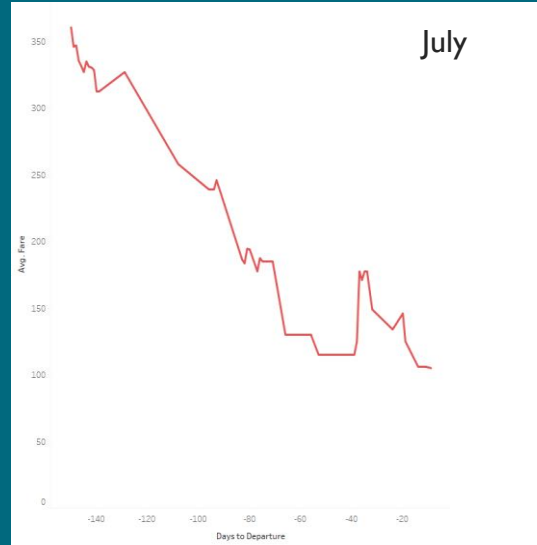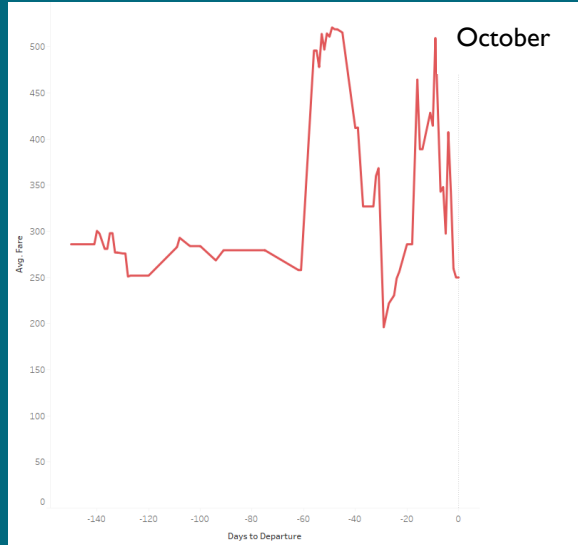

October hosts major cultural and pop events, driving travel. This variable demand may cause flight price fluctuations, as attendees may book tickets at different times, some in advance and others last-minute

Summer travelers often plan and book well in advance. Airlines, recognizing this pattern, may initially set higher fares, contributing to the subsequent downward trend

United Airlines in Multiple Markets

ORD to CUN

LAX to SJD
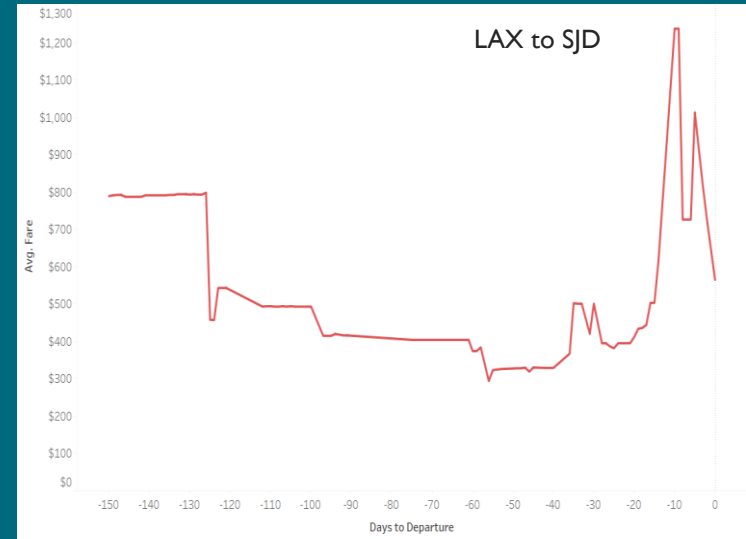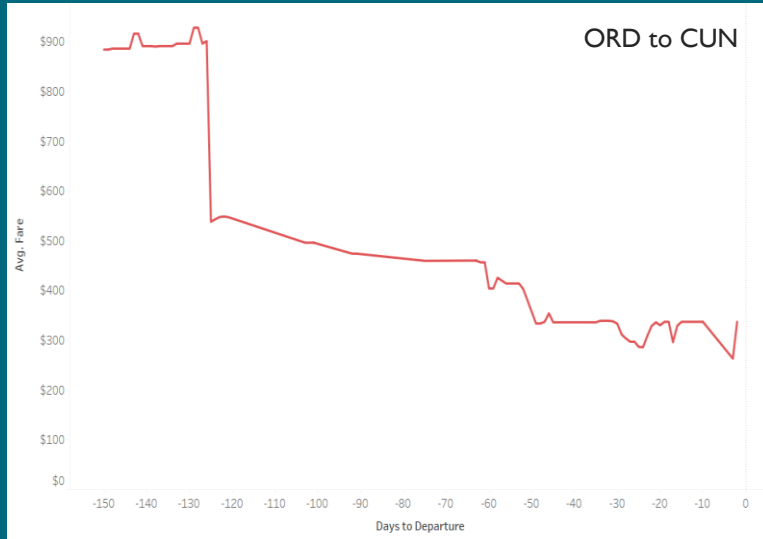
▷ The continuous downward trend in flight prices from Chicago to Cancun could be attributed to the prevalent behavior among travelers on this route to plan and book their tickets well in advance

▷ Los Cabos is considered a more upscale holiday destination in comparison to Cancun. Consequently, travelers heading to Los Cabos may consist of a mixed group, including individuals who prefer to book their tickets closer to the departure date

03.c

Market Analysis – Mexico
(Peak Season Analysis)

TSA reports that 7 out of 10 highest volume days fall in June & July

National Press Release
Monday, November 13, 2023

Top 10 Highest Volume Days

Rank	Date	Total Passenger Volume	Rank	Date	Total Passenger Volume
1	6/30/2023	2,884,783	6	7/7/2019	2,795,369
2	12/1/2019	2,882,915	7	5/24/2019	2,792,670
3	7/30/2023	2,815,068	8	7/28/2023	2,787,676
4	7/23/2023	2,800,368	9	6/16/2023	2,785,332
5	10/15/2023	2,797,326	10	7/21/2023	2,782,146

This table represents the highest passenger screening volumes recorded by TSA since its establishment in 2001.

Reasons for High Volume in June and July:

- **Summer Travel Peak**: June and July are peak months for summer travel, contributing to increased passenger numbers
- **School Breaks & Vacation Season**: Many schools have summer breaks during this period, leading to family vacations and higher air travel demand

Significance of Focusing on June and July for Analyses:

- **Peak Market Activity**: Concentrating analyses during peak months provides insights into carriers performance during high-demand periods
- **Understanding Price Trends**: Summer travel is associated with dynamic pricing Analyzing this period helps understand price trends and factors influencing fares
- **Volatility Exploration**: Increased travel can lead to greater volatility in pricing. Analyzing volatility during peak months aids in understanding market dynamics

Robert H. Smith School of Business

Contrasting Pricing Strategies between Delta & Alaska (US - MEX)
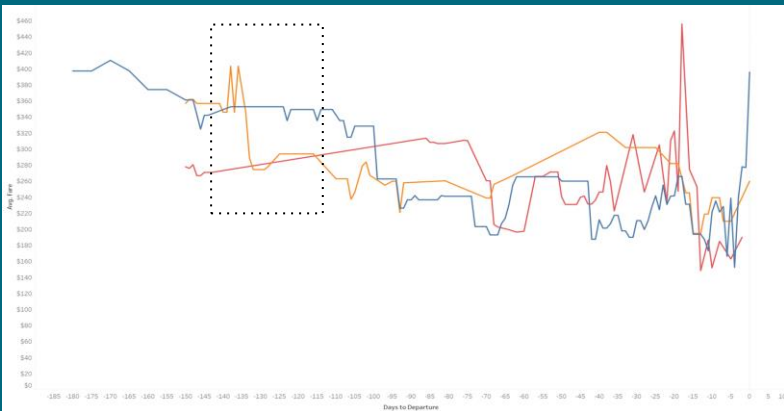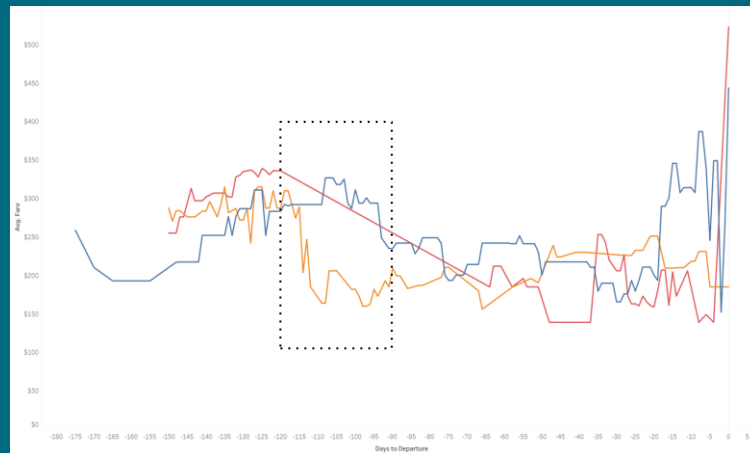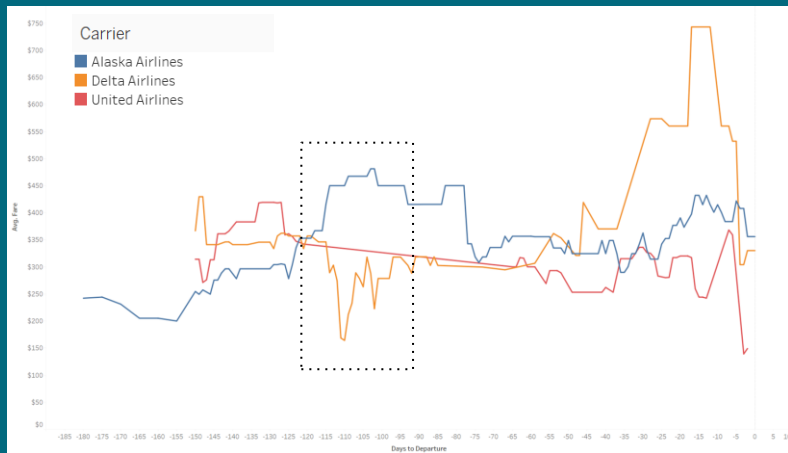
▷ **Within the 130 to 90 days before departure**:
- Delta was consistently pricing lower than Delta Airlines

▷ **Possible Causes**:
- **Competitive Dynamics**: Differences may stem from competitive interactions between Alaska and Delta
- **Market Positioning**: Varied market positioning and target customer segments could lead to divergent pricing approaches
- **Demand Sensitivity**: Fluctuations in demand may influence airline fares

Previous Day-to-Day Price Volatility
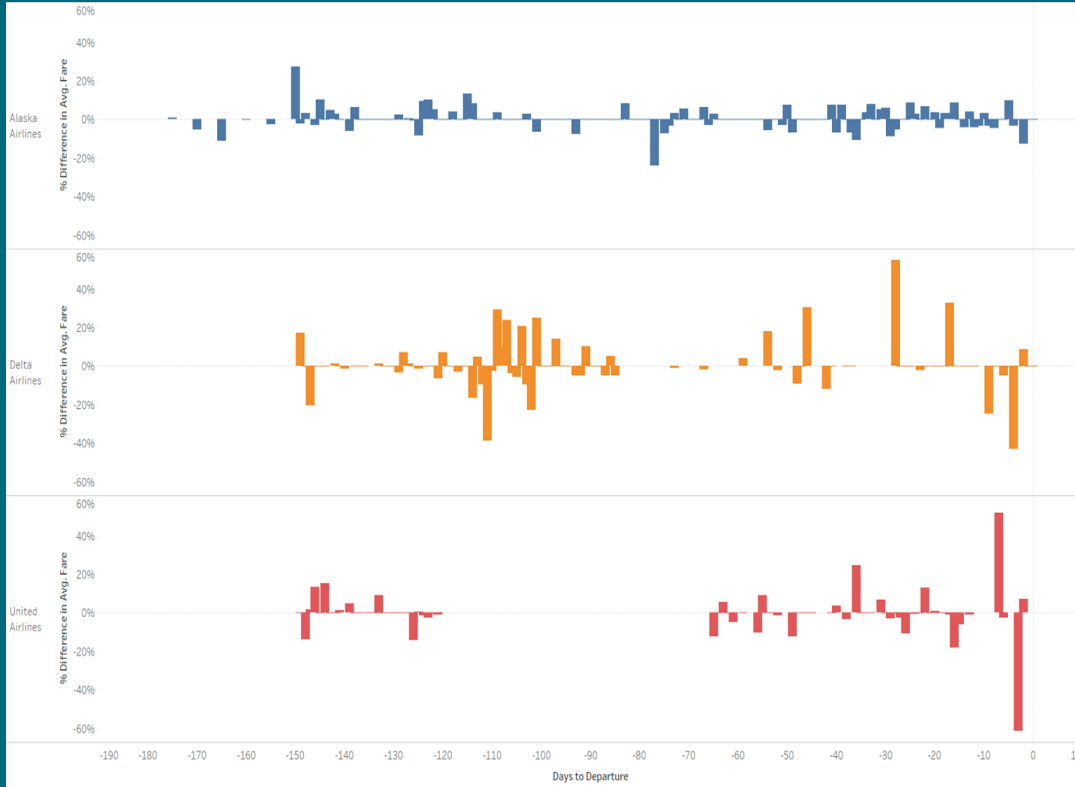
▶ **Delta Airlines:**
Demonstrates higher day-to-day price fluctuations, particularly between 130 and 90 days before departure. Suggesting that Delta's prices are more susceptible to short-term changes in demand and competitor offerings

▶ **United Airlines:**
Exhibits increased volatility closer to departure dates, indicating that United's pricing strategy becomes more dynamic and responsive to immediate booking trends
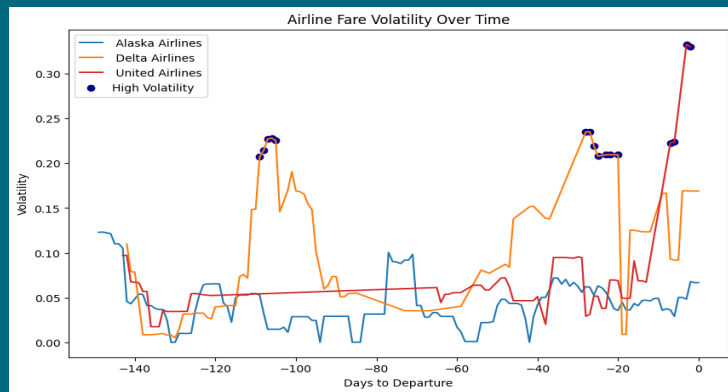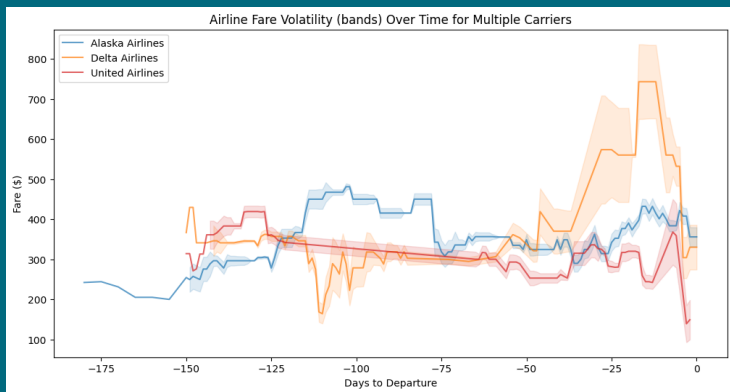
▶ **Alaska Airlines:**
Maintains relatively stable daily price fluctuations throughout the analyzed timeframe. This suggests a more consistent and less reactive approach to pricing compared to Delta and United

Volatility – Rolling Time Period

Airline Fare Volatility (bands) Over Time for Multiple Carriers

Airline Fare Volatility Over Time

▷ **Delta Airlines:** Maintains the highest average volatility across the analyzed time frame, especially within the 130-90 day window and closer to departure.This confirms the observation of Delta's price sensitivity to short-term market dynamics

▷ **United Airlines:** Suggests a similar shift towards dynamic pricing closer to travel dates

▷ **Alaska Airlines:** Continues to exhibit the lowest average volatility, showcasing a consistent and less responsive pricing strategy compared to Delta & United

▷ **Additional Observations:**
- Both Delta and United appear to prioritize dynamic pricing strategies, adjusting their fares frequently based on market condition
- Alaska Airlines, in contrast, seems to favor a more stable pricing strategy, potentially focusing on attracting budget-conscious travelers seeking predictable fares
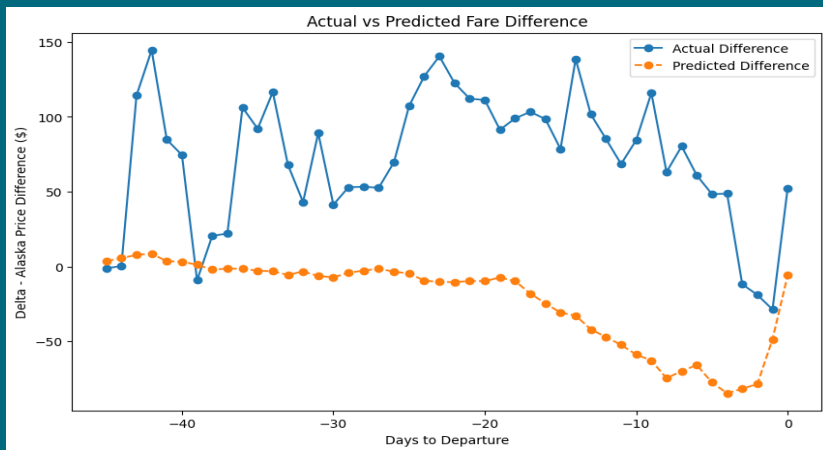
04

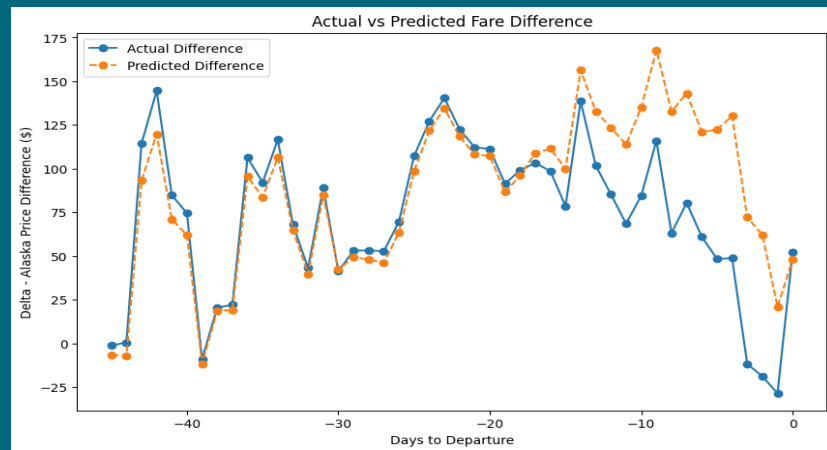Modeling

Modeling and Hypothesis Testing

Hypothesis:
- **Initial Observation:** Delta consistently prices itself $100 higher than Alaska 130-150 days before departure.
- **Initial Hypothesis:** Expect the price difference to persist closer to departure

Model 1: Linear Regression using Days to Departure & Alaska price

Model 2: Linear Regression using Days to Departure & Delta price

Hypothesis testing using Regression Models:
- **Finding**: Model 2 (Delta) outperformed Model 1 in predicting the price difference.
- **Possible Insights**: Delta's price Is a strong predictor & its strategy appears to be responsive to market conditions, leading to fluctuating price differences
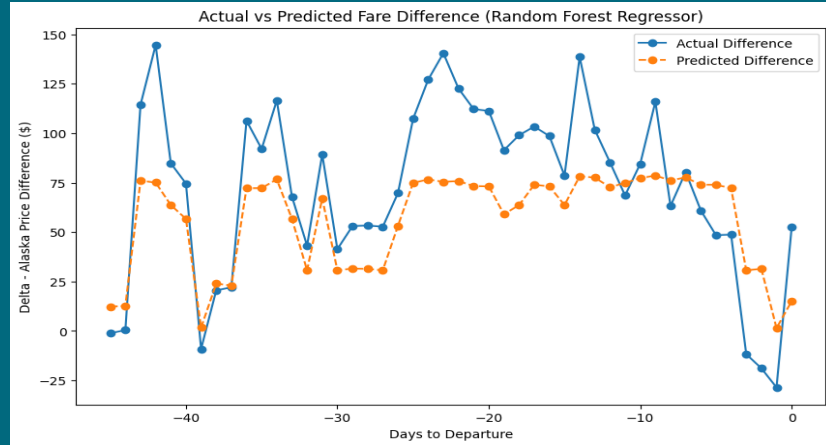
Modeling and Hypothesis Testing (Contd.)

Model Improvement:
- Random Forest Regressor to improve prediction accuracy
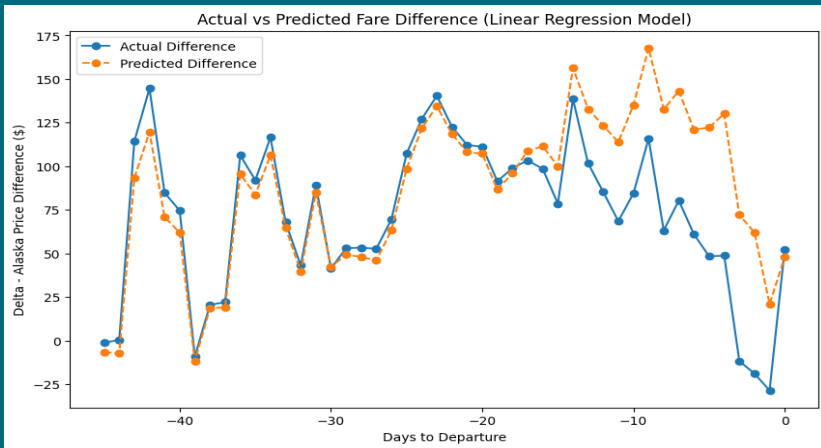
Linear Regression model with MSE: 1162.51   Random Forest Regressor with MSE: 963.93

Hypothesis Validation:
- **Finding**: Actual price differences were predicted to be much lower than that and sometimes even negative
- **Conclusion**: The hypothesis was proven wrong through model predictions

Next Steps:
- Add more features and perform feature importance analysis
- Continue to build advanced statistical models and continuously refine models based on emerging data and market changes

05

Price Inflection
Alerting Mechanism

Understanding Inflection Points in Flight Pricing

Characteristics:
Observation that disrupts an established prevailing pattern noticed in the preceding period, indicating an impending shift in the trend's direction
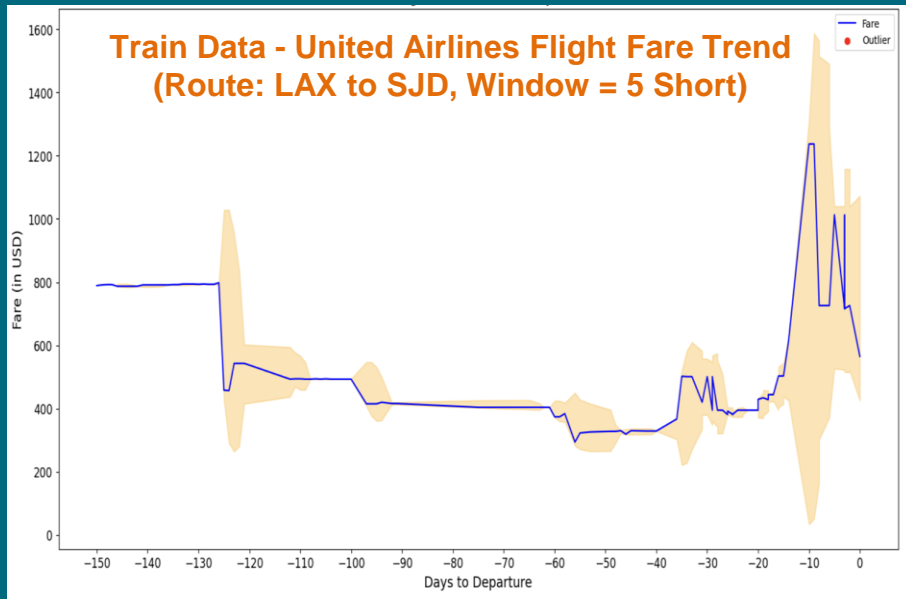
Complexity:
Data can be noisy with many fluctuations & does not follow any distribution ruling out the conventional methods of detecting Inflection points
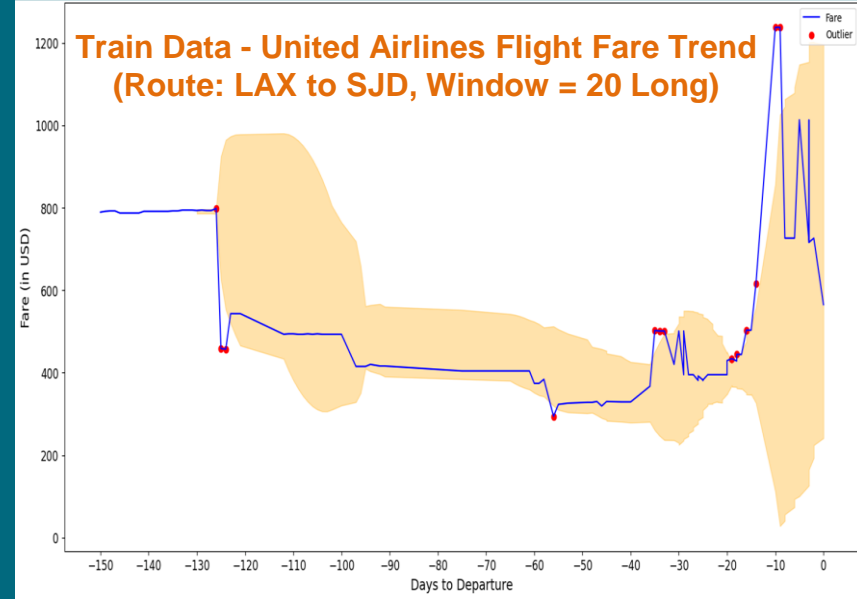
Real Time Monitoring:
Data Alert can be established to track fluctuations around Running Moving Average to identify Unusual Events

Both short & long window models do not address the gap in price responsiveness & precision in inflection detection

Train Data - United Airlines Flight Fare Trend
(Route: LAX to SJD, Window = 5 Short)

VS

Train Data - United Airlines Flight Fare Trend
(Route: LAX to SJD, Window = 20 Long)

▷ The upper and lower bands are calculated as two rolling standard deviations above and below the moving average with the area between representing the volatility of fares

▷ Shorter window size is more responsive to immediate changes in prices, resulting in overfitting while a longer window results in over smoothing making it less reactive to short-term fare fluctuations and more reflective of the long-term trend

Train Data - United Airlines Flight Fare Trend
(Route: LAX to SJD)
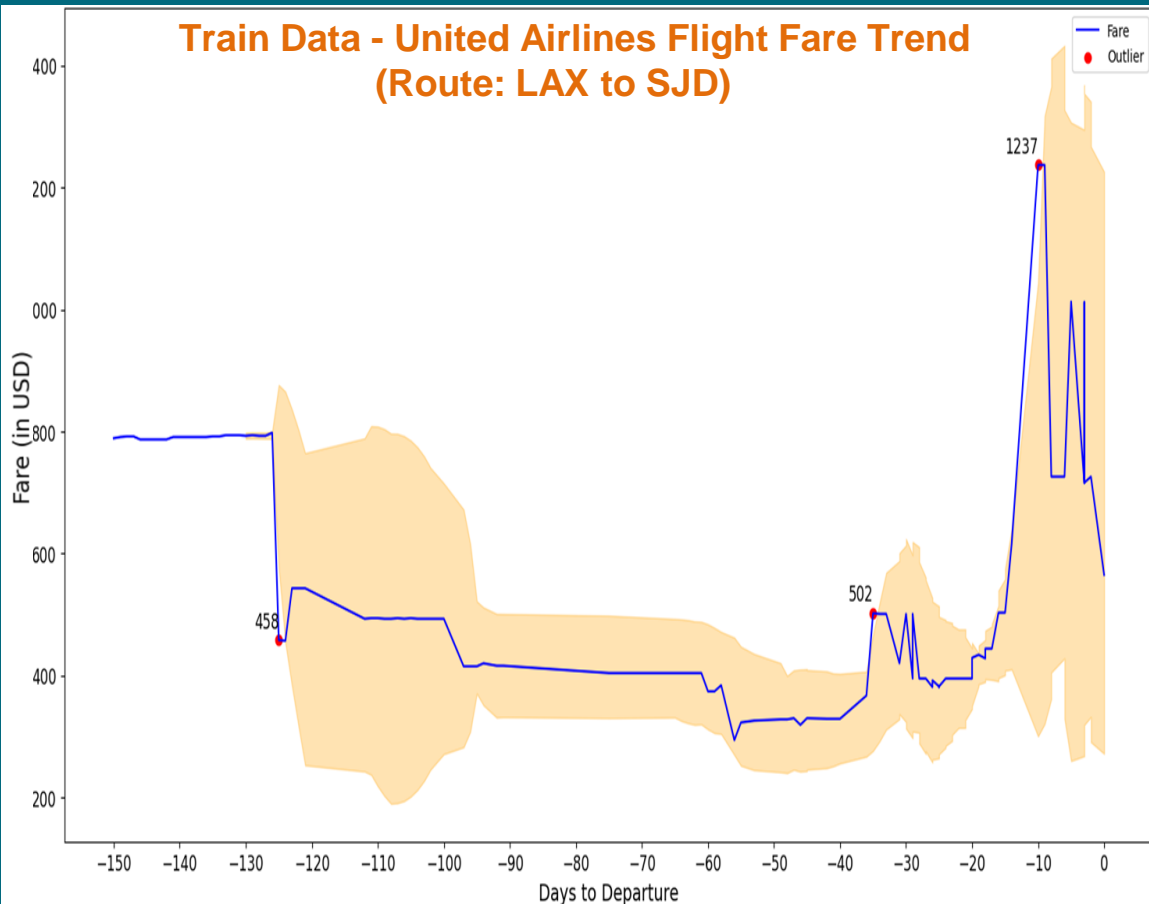
Using different windows for MA & volatility yield significantly better results

▷ Tested two different windows combinations to estimate the running moving average & running standard deviation

▷ A smaller window for moving average will make it more responsive to price changes, while a larger window for standard deviation can provide a more stable measure of volatility

▷ Stored procedure can be implemented with Snowflake's alerting system & task scheduling in place to get real-time alerts when an inflection is detected

Test Data - United Airlines Flight Fare Trend
(Route: LAX to SJD)
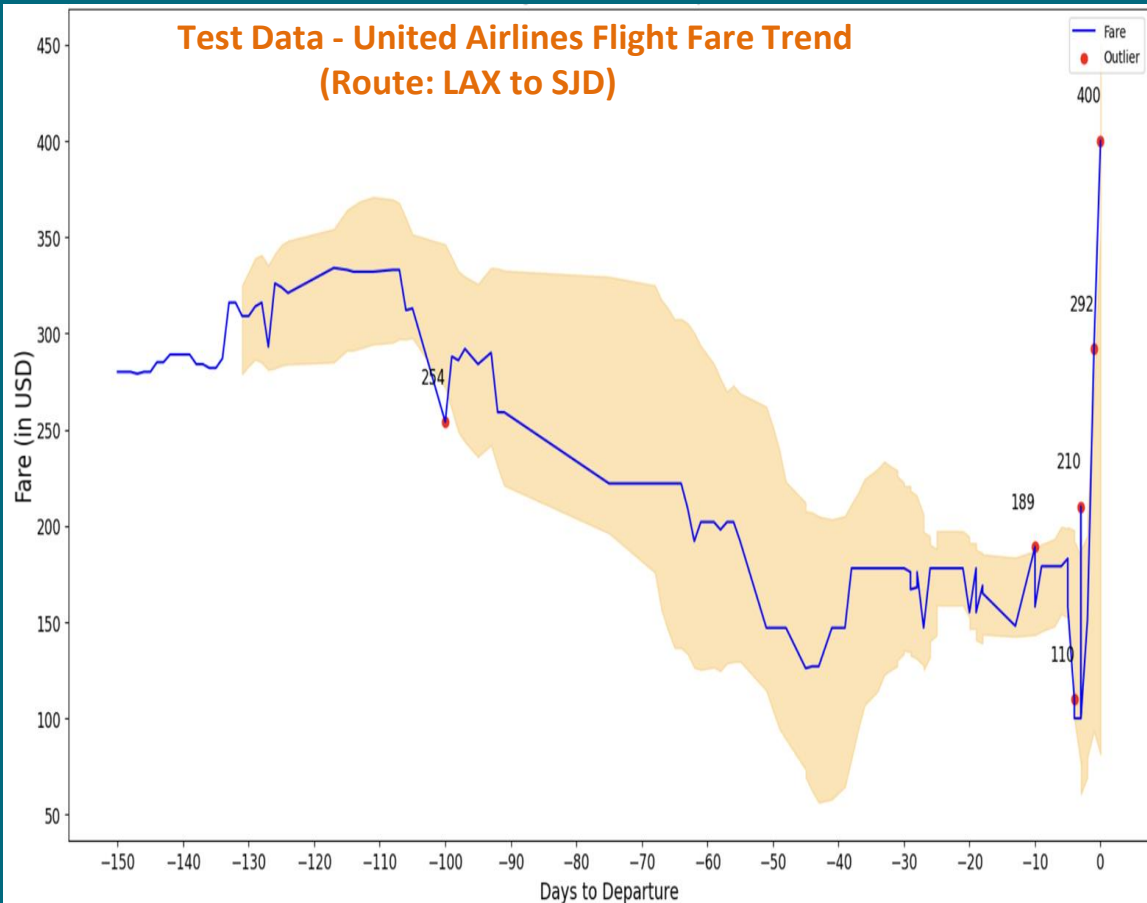
Over-smoothing effects in areas of intense volatility

▷ Somewhat distinguishes between genuine anomalies and usual fare fluctuations

▷ However, the model fails in highly volatile fare zones leading to inaccurate flagging of regular fluctuations

Scope for Further Analysis

Integrate seasonal effects on the model, owing to its significant impact on demand which heavily influences the pricing decisions

Adjust the window values for different 'Days to Departure' categories, aiding in efficiently capturing the high price volatility periods

Expand the historical data used in conjunction with these adjustments for a better accuracy of data model

06

Challenges
&
Workarounds

Challenges & Workarounds

Huge Data set:
- Workaround:
  Filtering by latest time period to focus analysis on the most recent data to reduce the dataset size

Missing Data for Many International Markets:
- Workaround:
  Selected three specific markets based on criteria such as flight duration, travel type (leisure or business), and data completeness

Challenges & Workarounds (Contd.)

Missing Data for Many Carriers in Multiple Routes:
- Workaround:
  - Analysis on single carrier in multiple markets to mitigate the impact of missing data
  - Manual price collection for specific routes for carriers to gain insights into their pricing trends

Assessing the Influence of Airfare on Car Rental Prices:
- Workaround:
  Leveraged business domain expertise to conclude that in international markets, there might not be a significant correlation between airfare and car rental prices

07

Recommendations
&
Opportunities

Recommendations & Opportunities

Integrate Demand Estimates
QL2 must collect historic daily demand estimates for US to international routes from APIS Air Traffic Monitor along with Airfare data to get a better understanding of price fluctuations

Cross-Market Learning
Continuously learning about airlines strategies in different markets to stay stable vs volatile will help clients in identifying unique opportunities

Inter-Airline Price Influence Modeling
Develop models to forecast price difference among various airlines where significant price influence on each other is observed to understand competitors behavior pattern

Price inflection alerting system
Enhance the price inflection alert system in line with the client's response strategy to competition, and market this as a solution for reducing manual price monitoring

01 02 03 04

Thank you