

An Exploration into Expedia

By: Amanda, Jose, Maddie, & Nathan



What is Expedia?

Online Travel Agency

- Owned by Expedia Group

Used to book:

- airline tickets
- hotel reservations
- car rentals
- cruise ships
- vacation packages

The screenshot displays the Expedia website interface. At the top, the Expedia logo is visible. Below it, there are three images: a hotel room, a hotel lobby, and a hotel restaurant. The main content area shows the details for 'Forest Lodge Retreat', including a 4.3/5 rating, a 'Member Price available' badge, and a 'Reserve' button. The price is listed as USD\$624. Below the hotel details, there is a calendar view showing rates for different room types: Queen Room, Harbour View Twin, and Corporate rate. The calendar shows rates for various dates, with some dates highlighted in red.

Forest Lodge Retreat
4.3/5 Excellent
Guests rated this property 4.3/5 for cleanliness.
See all 1,682 reviews >

Popular amenities
Pool, Free WiFi, Air conditioning, Gym, Parking included

Room, 2 Queen Beds
Fully refundable ⓘ
Before Mon., 18 Apr.
USD\$624
includes taxes & fees
Price details > Reserve

Go to date < > >>

	TODAY	TUE	WED	THU	FRI	SAT	SUN
Queen Room	10	10	10	10	10	10	10
Early bird 2 CHANNELS ⚡	RATES ⓘ 🗑️ ✖️	100	100	100	100	100	100
Corporate rate 3 CHANNELS ⚡	RATES ⓘ 🗑️ ✖️	100	100	100	100	100	100
Holiday rate 3 CHANNELS ⚡	RATES ⓘ 🗑️ ✖️	10	10	10	10	10	10
Harbour View Twin	10	10	10	10	10	10	10
Early bird 2 CHANNELS ⚡	RATES ⓘ 🗑️ ✖️	100	100	100	100	100	100
Corporate rate 3 CHANNELS ⚡	RATES ⓘ 🗑️ ✖️	100	100	100	100	100	100



How does Expedia make money?



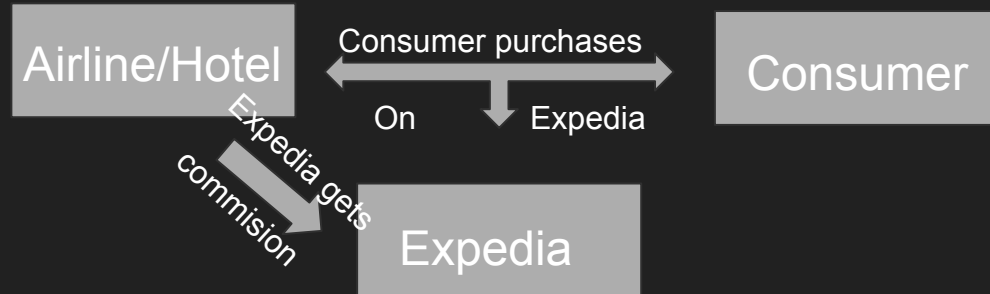
Expedia Group includes:



- Expedia, Hotels.com, Vrbo, Travelocity, Hotwire.com, Orbitz, Ebookers, CheapTickets, CarRentals.com, Expedia Cruises, Wotif, and Trivago

Expedia uses:

- Merchant Model
- Agency Model
- Advertising Model



How are flights scheduled?

Airline Hubs

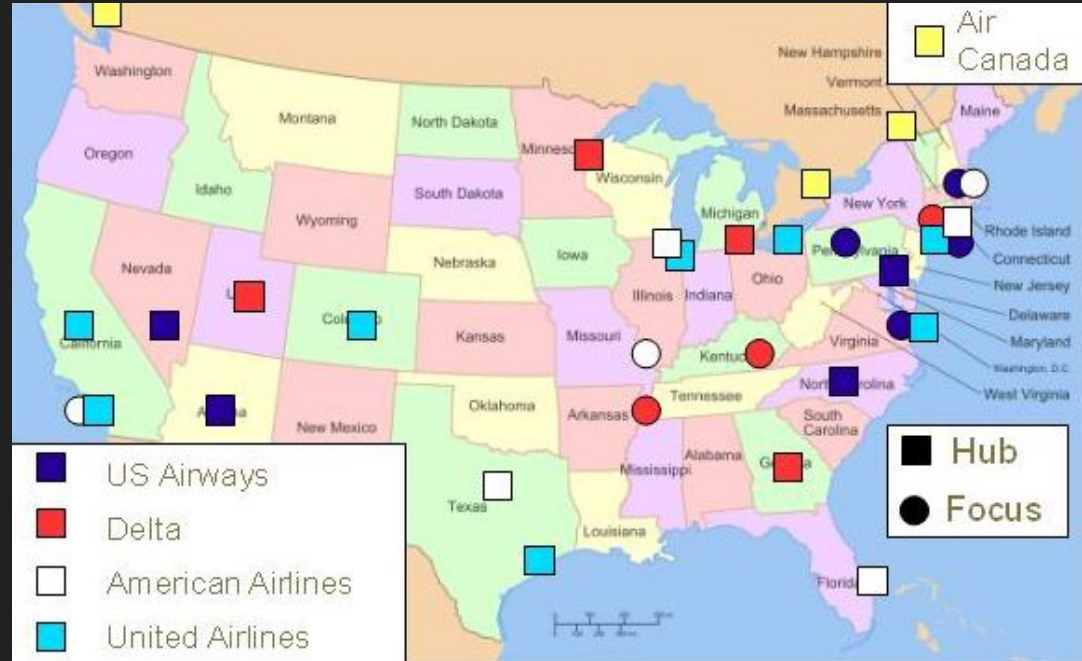
- Connections can be cheaper

Hub Banking

- Average day has ~10 'banked' times

Gate-to-Gate

- More "early" arrivals



How does Expedia get its flights?

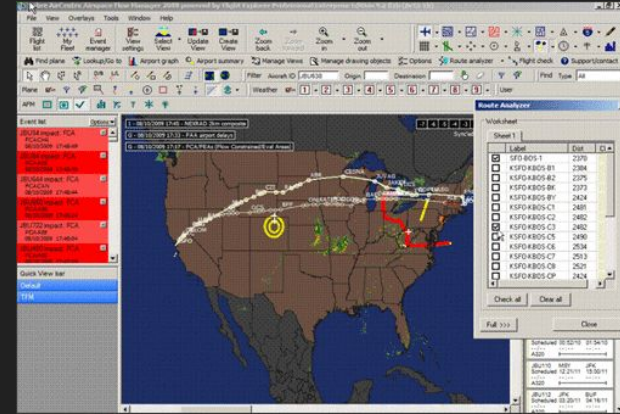
Global Distribution System (GDS)

- Sabre, Amadeus and Travelport

Wholesale flight resale

Agency commission on sold tickets

Organized by Best Fare Search (BFS)

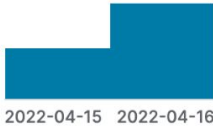
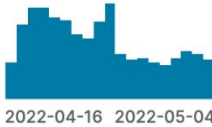


<https://youtu.be/AlRnoym6QeE>

A screenshot of the Expedia website's flight search interface. The top navigation bar includes the Expedia logo, a 'More travel' dropdown, and links for English, List your property, Support, Trips, and Katie. Below this, there's a horizontal menu with tabs for Stays, Flights (selected), Cars, Packages, Things to do, and Cruises. The main search area has a 'Roundtrip' button selected, along with 'One-way' and 'Multi-city' options. A 'traveler' dropdown is set to '1 traveler' and the 'Economy' class is selected. The search criteria include 'Leaving from' and 'Going to' fields, both with location pins, and date pickers for 'Departing May 24' and 'Returning May 25'. A large blue 'Search' button is at the bottom.

Our Data Set

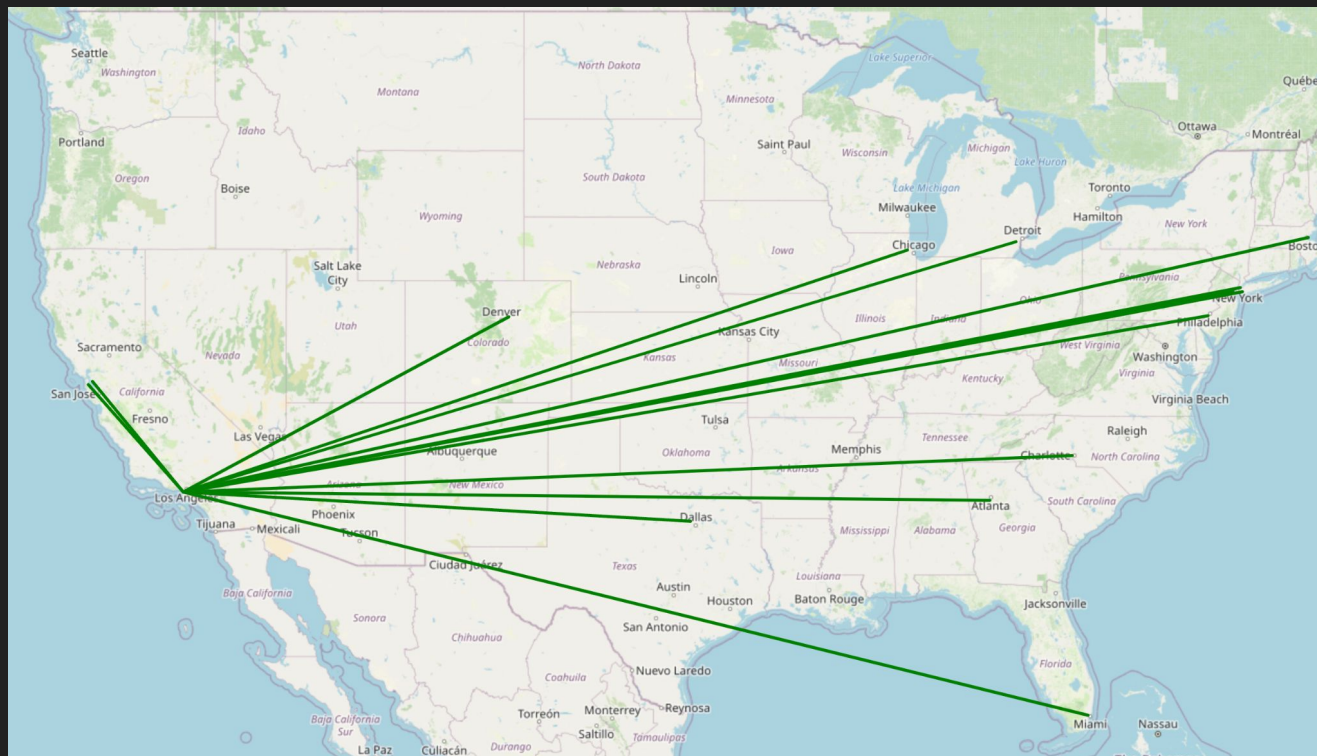
- The full dataset has 6 million rows and 27 columns
- We utilized a subset of this dataset that is 173,483 rows
- Features of the dataset include info such as starting airport, destination airport, price of flight, distances, type of airlines, time of travel, etc

itineraries-min-250k.csv (94.94 MB)			
Detail	Compact	Column	
🔍 legId	📅 searchDate	📅 flightDate	📍 startingAirport
173483 unique values			LAX 10% LGA 8% Other (206195) 82%
9ca0e81111c683bec1012473feefd28f	2022-04-16	2022-04-17	ATL
98685953630e772a098941b71906592b	2022-04-16	2022-04-17	ATL
98d90cbc32bfb05c2fc32897c7c1087	2022-04-16	2022-04-17	ATL
969a269d38eae583f455486fa90877b4	2022-04-16	2022-04-17	ATL
980370cf27c89b40d2833a1d5afc9751	2022-04-16	2022-04-17	ATL

All Possible Flight Paths from LAX

15 Airports Represented:

- LAX
- LGA
- BOS
- ORD
- CLT
- DFW
- SFO
- ATL
- MIA
- PHL
- DTW
- DEN
- EWR
- JFK
- OAK



Top 5 Most Popular Flights (~2000 searches)

LGA→LAX

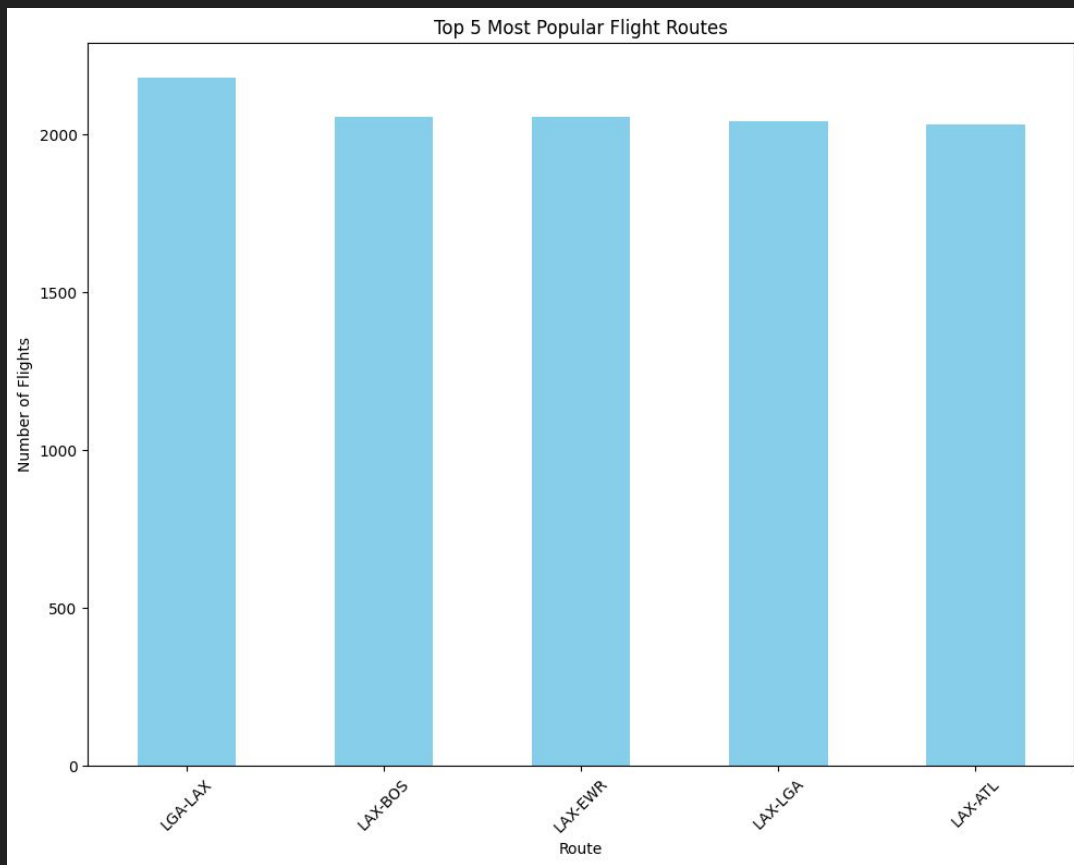
The other 4 flights start at
LAX

LAX→BOS

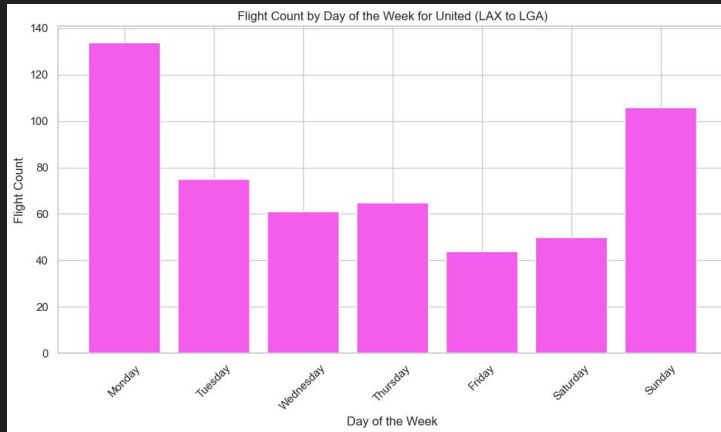
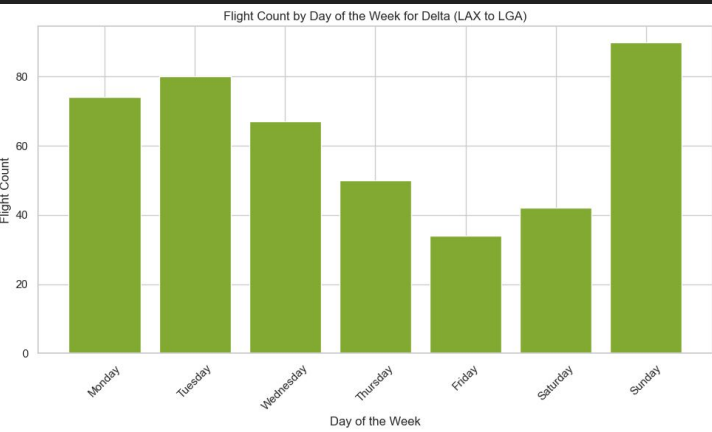
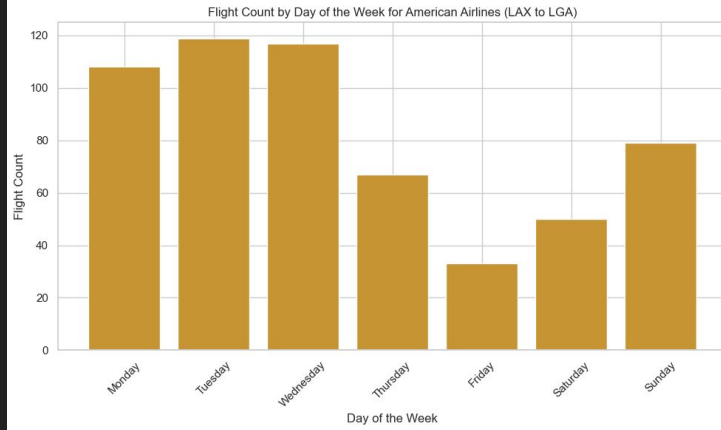
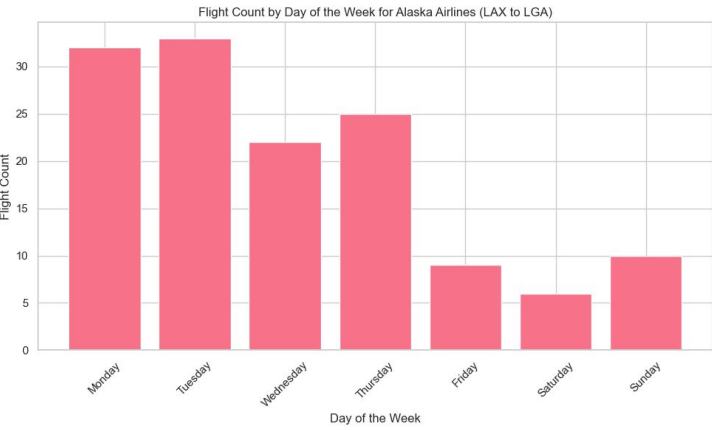
LAX→EWR

LAX→LGA

LAX→ATL

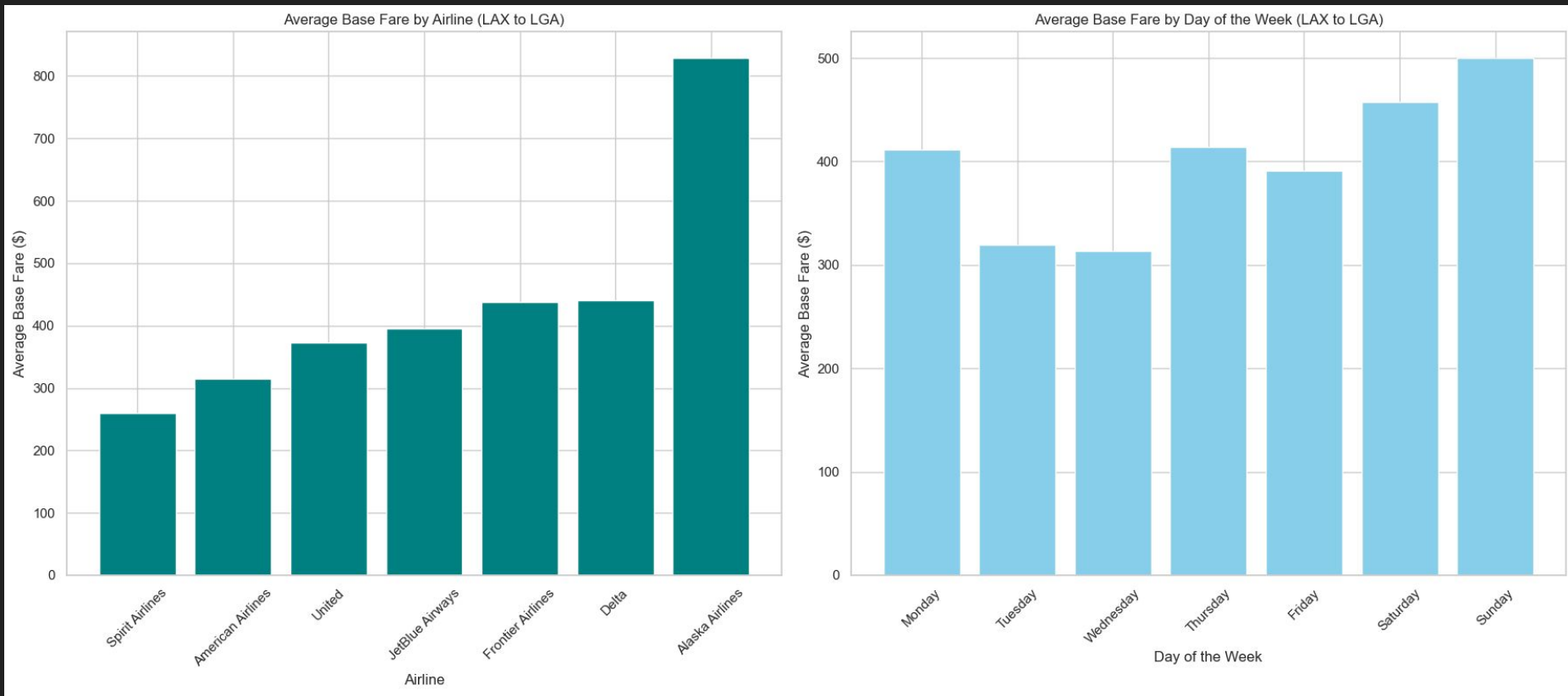


Most flights per day of the week:



- Monday: United
- Tuesday: AA
- Wednesday: AA
- Thursday: Delta/AA/United
- Friday: AA/Delta
- Saturday: AA
- Sunday: Delta

- Why is American Airlines cheaper than Alaska Airlines from LAX to LGA, when they have greater amount of flights, and operates more than any other airlines on Tuesday, Wednesday, Thursday, Friday, and Saturday?

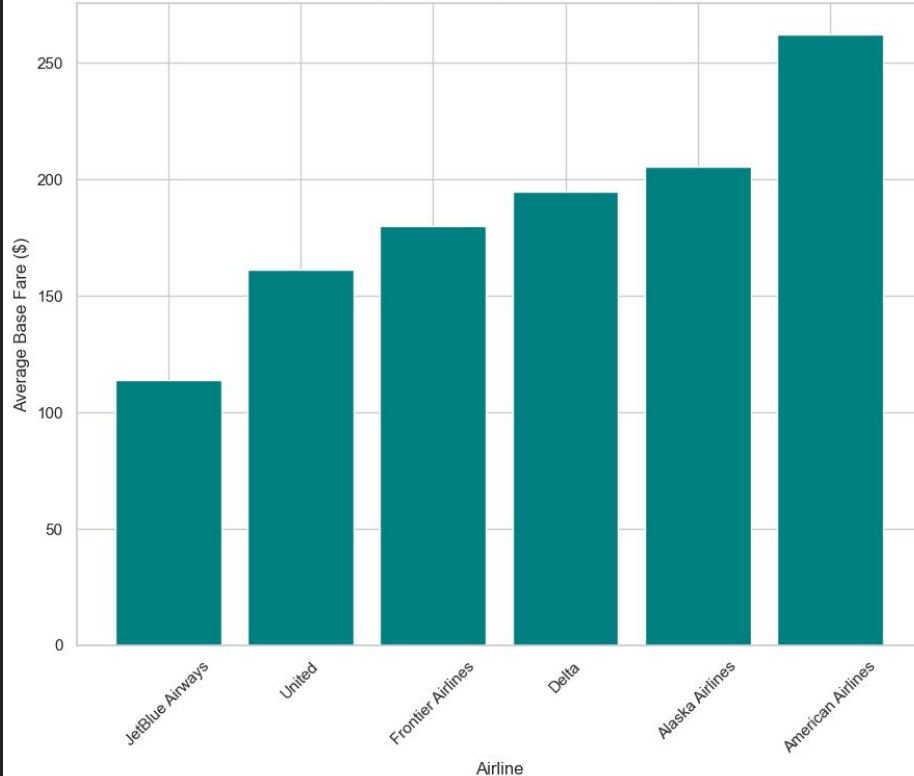


AA vs. Alaska Airlines from LAX to LGA

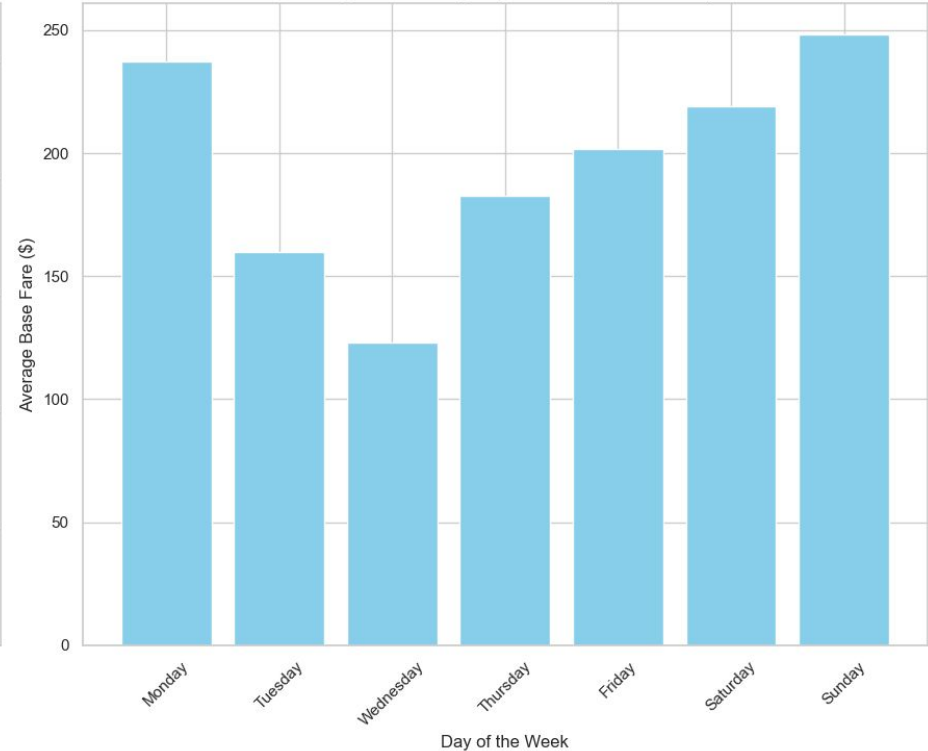
- Basic Economy
 - American Airlines' basic option is slightly **cheaper**, earns flyers more redeemable miles than Alaska and includes upgrade privileges that Alaska doesn't offer
- Reliability:
 - Alaska beats out American Airlines when it comes to reliability, with a higher percentage of **on-time flights** and a lower percentage of flight cancellations and mishandled luggage
- Fees:
 - **Extra fees** can really put a damper on your travel experience — and Alaska passengers can generally expect to pay **less fees** than American Airlines flyers.
- Loyalty program:
 - Alaska has the best **elite status and reward program** in the U.S., making Alaska's loyalty program the clear winner.
- In the end:
 - Alaska Airlines comes out on top of American Airlines in a head-to-head comparison, in large part due to Alaska's generous elite program, where **miles and status are earned based on miles flown** instead of dollars spent.

- Why is American Airlines more expensive when the destination is SFO?
 - its competition between airlines. At airports that don't have much competition, you can safely assume you'll see higher prices.

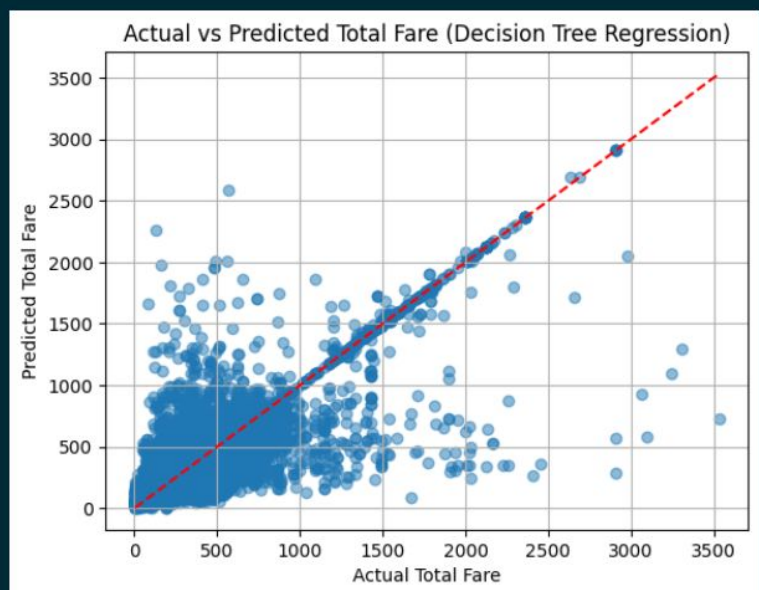
Average Base Fare by Airline (LAX to SFO)



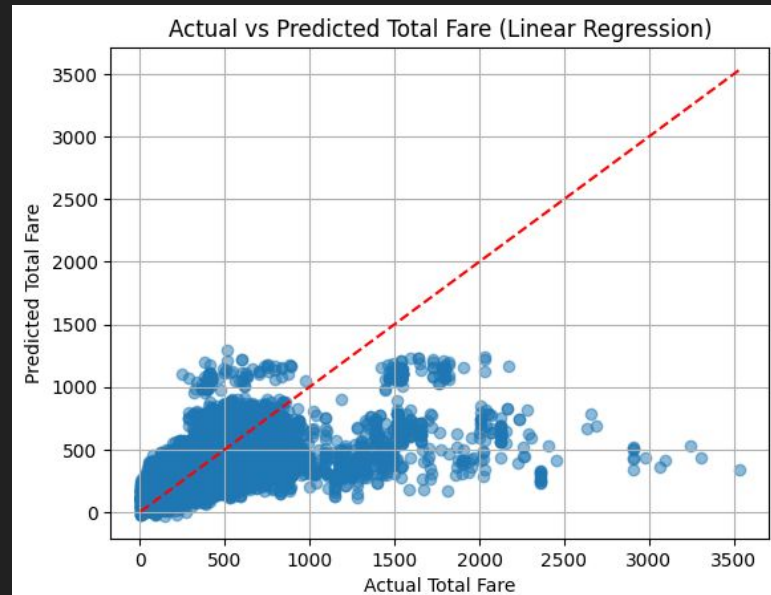
Average Base Fare by Day of the Week (LAX to SFO)



ML Supervised Learning (Predicting Total Fare)



Decision Tree Regression Mean Squared Error: 12976.523228641252
Decision Tree Regression R-squared: 0.7076209010185737
Decision Tree Regression Mean Absolute Error: 54.56830393943562

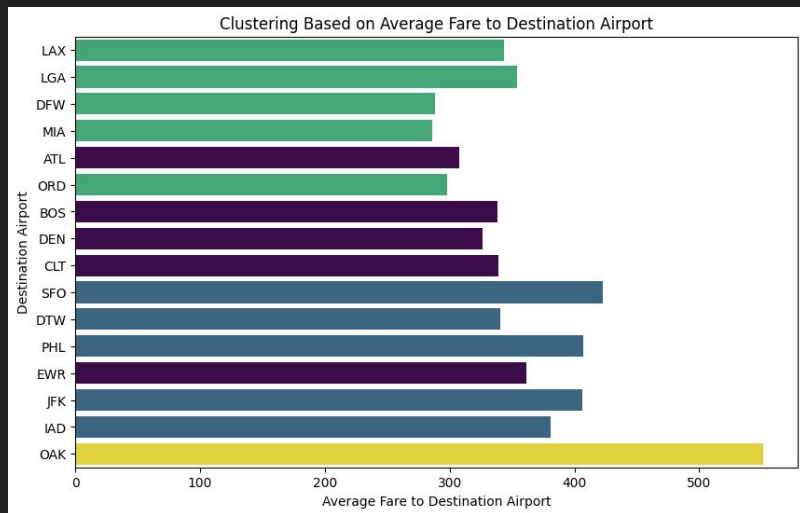


Linear Regression Mean Squared Error: 25732.913007623054
Linear Regression R-squared: 0.42020171452935073
Linear Regression Mean Absolute Error: 95.71731860747254

- Data isn't linear, so hard to predict with linear regression, had to use other models (XGBoost, Decision Tree, Gradient Boosting)
- But overall doesn't do a terrible job

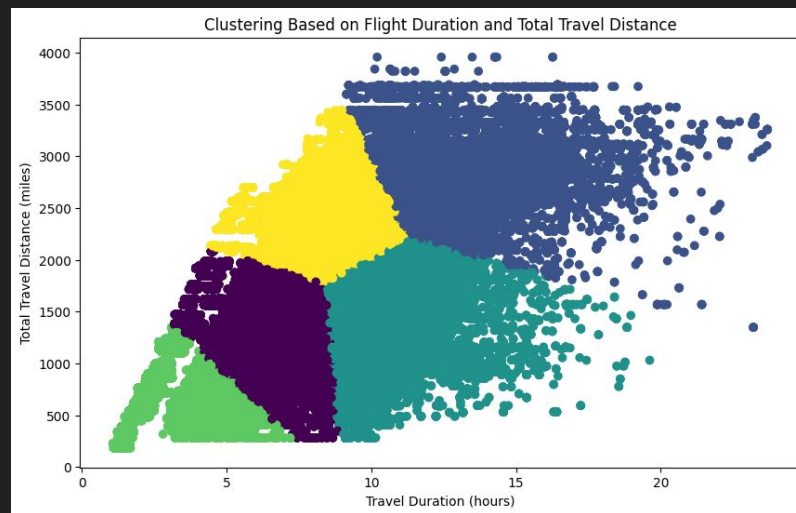
ML Unsupervised Learning (Cluster)

- Green Cluster: LAX, LGA, DFW, ORD
 - Major hubs for domestic/international travel
 - Tourist areas/large metropolitan
- Purple Cluster: ATL, BOS, DEN, CLT, EWR
 - Very similar to green, main difference is potentially less options in terms of airlines and less flight activity compared to the green cluster
- Blue Cluster: SFO, DTW, PHL, JFK, IAD
 - Overall the highest average total fare
 - Most airports here are relatively close to the coast
- Yellow Cluster: OAK
 - The highest total fare



Inertia: 18.007651187000057

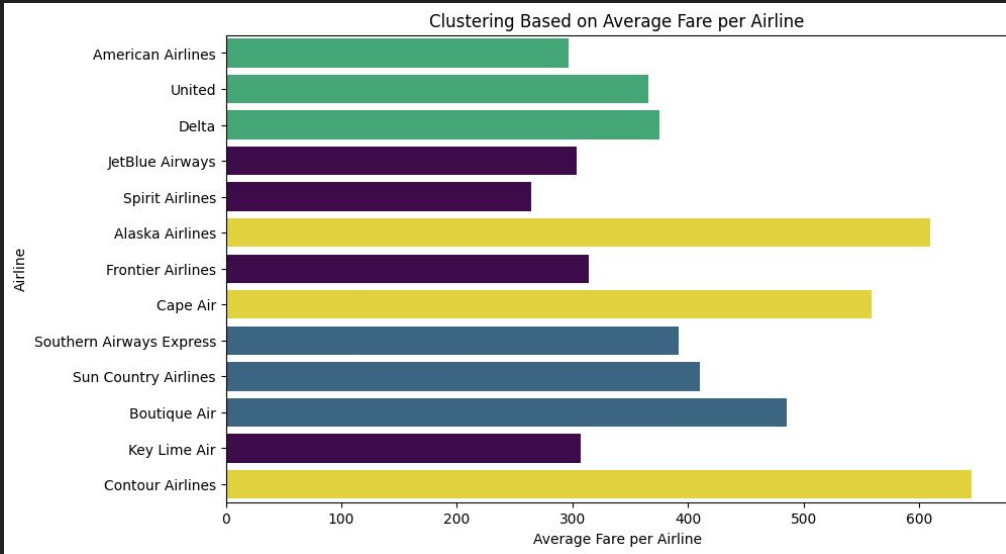
Silhouette Score: 0.05516756876964028



Inertia: 81997.85455522545

Silhouette Score: 0.23990601175699003

ML Unsupervised Learning (Cluster) Continue...



Inertia: 6.536563252066205

Silhouette Score: 0.23990601175699003

- Green (American, United, Delta)
 - “Legacy” carriers, large route network
 - High flight frequencies
 - Travel durations more often than not are longer than the airlines below
 - Hub and spoke system (centralized area for flights for lots of different destinations)
- Purple (Jetblue, Spirit, Frontier, Key Lime)
 - Lower average ticket prices compared to the legacy carriers
 - Operate point to point route (direct flights between cities without hub) instead of hub and spoke system
- Yellow (Alaska, Cape, Contour)
 - Prob the weirdest cluster cause alaska is here but reasons could be price
 - These flights are more regional potentially, so shorter routes
 - Speciality airlines (specifically cape and contour) focus on cargo and leisure travel
- Blue (Southern, Sun Country, Boutique)
 - Even Smaller carries historically, limited routes
 - Focus on travel packages

Summary/Conclusion

- Popular searches
 - More route options
 - Wholesale purchased flights trying to be offloaded
- Pricing
 - Airline Hubs
 - LGA is American Airlines hub
 - OAK is not a hub
 - Day of the week
 - Wednesdays are cheapest
 - Sundays are consistently expensive
 - Airlines
 - Larger airlines seem cheaper

Sources:

- <https://www.nerdwallet.com/article/travel/alaska-airlines-vs-american-airlines>
- <https://thriftytraveler.com/guides/why-are-flights-from-my-airport-so-expensive/#:~:text=If%20there's%20one%20thing%20that,you'll%20see%20higher%20prices>
- <https://www.linkedin.com/pulse/understanding-hub-and-spoke-vs-point-to-point/>
- <https://www.expedia.com/stories/science-travel-bill-nye/>
- <https://www.siteminder.com/r/global-distribution-system/>
- <https://medium.com/expedia-group-tech/how-expedia-finds-your-flights-an-overview-802cd5e1fb6c#:~:text=Many%20flight%20searches%20on%20Expedia,technology%20providers%20for%20flight%20search.>
- [https://businessmodelanalyst.com/expedia-business-model/#:~:text=agency%2C%20and%20advertising.-,Merchant%20Model%20\(Reselling%20of%20travel%20packages%20and%20commission%20fees\).at%20a%20slightly%20higher%20price.](https://businessmodelanalyst.com/expedia-business-model/#:~:text=agency%2C%20and%20advertising.-,Merchant%20Model%20(Reselling%20of%20travel%20packages%20and%20commission%20fees).at%20a%20slightly%20higher%20price.)
- <https://welcome.expediagroup.com/en/about-us/about-expedia-group-travel-brands>