

Analysis of International Airline Ticket Prices

FROM **norwegian** 

PREDICTING INT'L AIRFARE – SUMMARY

FLIGHT DATA

90

days starting

June 2, 2019

4,400+
flights

PRODUCTION

Predicting Airfare
via
Linear Reg. Model

Producing:



Flask

NORWEGIAN AIR FLIGHT ITINERARIES

I ✕ OAK

JFK



LAX

Los Angeles World Airports



Gatwick
LONDON AIRPORT



Aerpuerto de Barcelona-El Prat

OVERVIEW – DATA COLLECTION

STANDARD PAGE FORMAT

Flights to Copenhagen from New York-JFK

Select departure and start reservation. All fares per person in USD. Restrictions, baggage fee and credit card surcharge

Fly from New York-JFK (JFK) ▾ ☐ Return Outbound 05 ▾ Jun 2019 ▾ ☐ Adults 1 ▾ Children 1 ▾
Fly to Copenhagen (CPH) ▾ ☐ Show direct only Return 05 ▾ Jun 2019 ▾ ☐ Infants

Outbound: New York-JFK - Copenhagen

Wednesday 5. Jun 2019

Compare fare types

Departure	Arrival	Economy cabin All fares per person in USD			Premium cabin All fares per person in USD	
		LowFare	LowFare+	Flex	Premium	PremiumFlex
23:00 New York-JFK Direct Duration: 7h 35m ① The 787 Dreamliner	13:06 +1 Copenhagen	469.90	559.90	764.90	-	1,094.90
23:00 New York-JFK 1 stop Duration: 11h 25m ① The 787 Dreamliner ② 1 stop (2h 35m) in London	18:25 +1 Copenhagen	516.00	606.00	1,066.00	1,096.00 Premium on parts of route	1,286.00 Premium on parts of route
22:25 New York-JFK 1 stop Duration: 12h 0m ① The 787 Dreamliner ② 1 stop (3h 10m) in London	18:25 +1 Copenhagen	566.00	656.00	1,066.00	1,096.00 Premium on parts of route	1,286.00 Premium on parts of route
23:00 New York-JFK 1 stop Duration: 13h 15m ① The 787 Dreamliner ② 1 stop (4h 25m) in London	18:16 +1 Copenhagen	523.60 Only 2 seats with this price	613.60 Only 2 seats with this price	1,066.00	1,103.60 Premium on parts of route	1,286.00 Premium on parts of route
23:55 New York-JFK 1 stop Duration: 13h 45m ① The 787 Dreamliner ② 1 stop (3h 25m) in Madrid	18:40 +1 Copenhagen	600.00	690.00	1,178.90	760.00 Premium on parts of route	1,208.90 Premium on parts of route

WARNING:

Site structure sometimes changes....

[Notice only 3 price columns here]

Flights to Copenhagen from Los Angeles

Select departure and start reservation. All fares per person in USD. Restrictions, baggage fee and credit card surcharge

Fly from Los Angeles (LAX) ▾ ☐ Return Outbound 09 ▾ Jun 2019 ▾ ☐ Adults 1 ▾ Children 1 ▾
Fly to Copenhagen (CPH) ▾ ☐ Show direct only Return 09 ▾ Jun 2019 ▾ ☐ Infants

Outbound: Los Angeles - Copenhagen

Sunday 9. Jun 2019

Compare fare types

Departure	Arrival	Details	All fares per person in USD		
			LowFare	LowFare+	Flex
20:30 Los Angeles ① The 787 Dreamliner ② 1 stop (1h 40m) in Oslo	18:35 +1 Copenhagen	1 stop Duration: 13h 5m	Sold out	Sold out	Sold out
19:30 Los Angeles ① The 787 Dreamliner ② 1 stop (1h 35m) in London	18:15 +1 Copenhagen	1 stop Duration: 13h 45m	950.00 Only 2 seats with this price	1,040.00 Only 2 seats with this price	1,186.00
20:30 Los Angeles ① The 787 Dreamliner ② 1 stop (2h 55m) in Oslo	19:50 +1 Copenhagen	1 stop Duration: 14h 20m	Sold out	Sold out	Sold out
19:30 Los Angeles ① The 787 Dreamliner ② 1 stop (6h 55m) in London	23:35 +1 Copenhagen	1 stop Duration: 19h 5m	836.00 Only 1 seat with this price	926.00 Only 1 seat with this price	1,186.00

DATA EXPLORATORY ANALYSIS

DATA SUMMARY

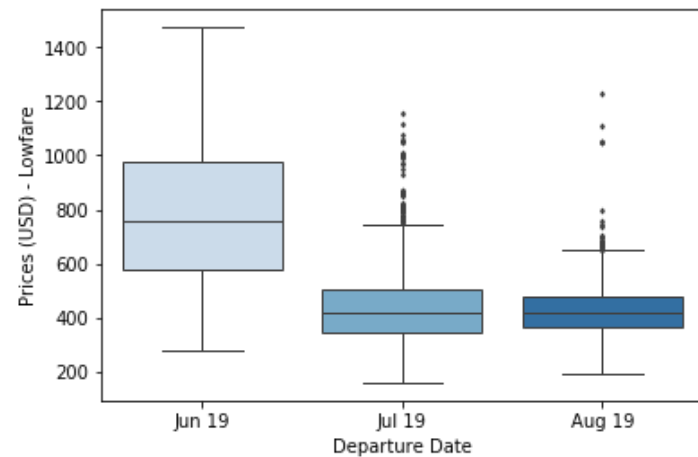
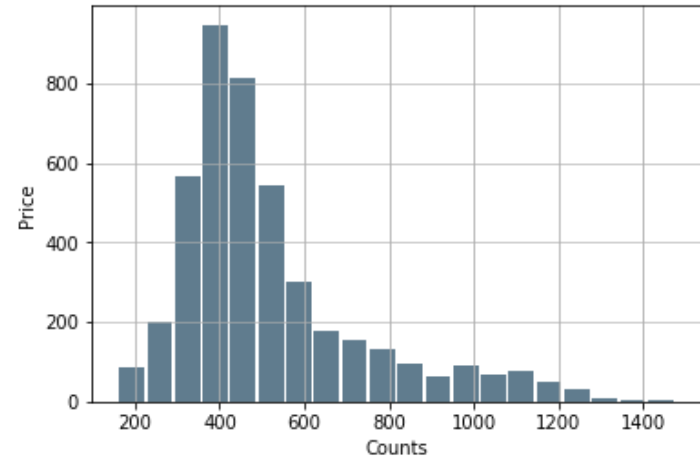
UNFILTERED
4,847 x 34

CLEANED
4,413 x 9

3,600
Connecting
813
Direct Flights

orig_port_coded	dest_port_coded	days_to_flight	depart_yr	depart_mo	depart_day	duration_total_min	stops	prices_lowfare_USD
0	0	1.0	2019	6	3	690	1	1172.9
0	0	1.0	2019	6	3	750	1	772.9
0	0	1.0	2019	6	3	760	1	910.9
0	0	1.0	2019	6	3	920	1	1172.9
0	0	1.0	2019	6	3	980	1	772.9

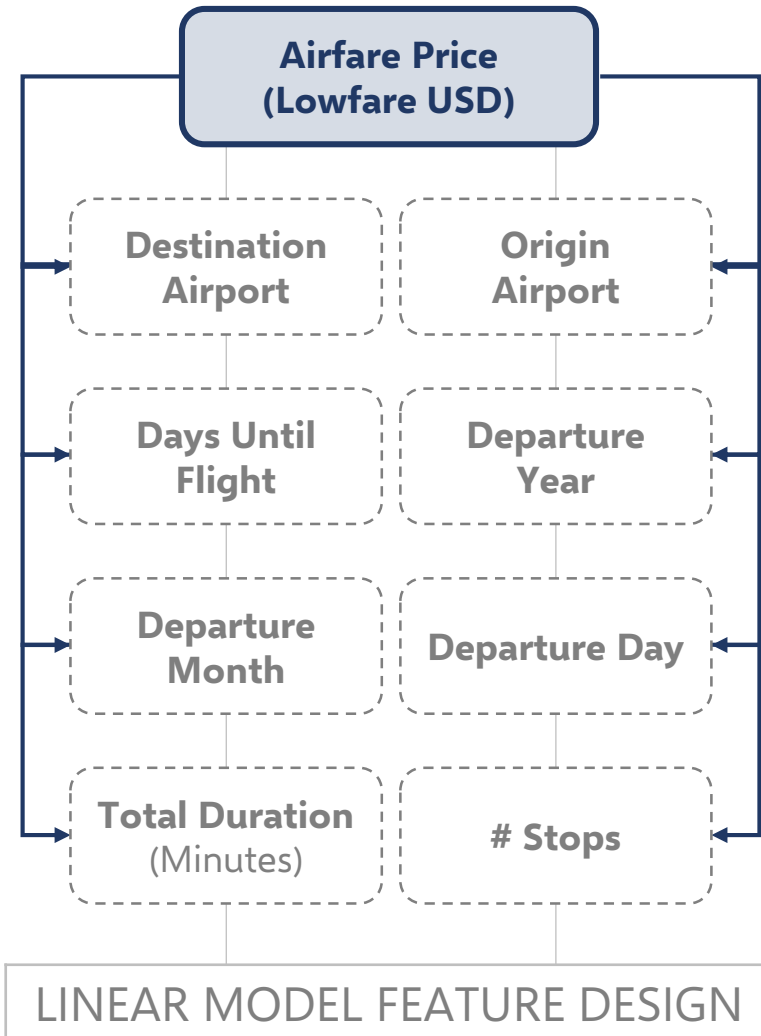
PRICES



STOP LOCATION

London	1,109	Madrid	225
Oslo	840	Amsterdam	129
Stockholm	532	Paris	101
Copenhagen	361	Rome	64
Barcelona	239		

PREDICTIVE LINEAR MODEL & FLASK API



Flask +



Amazon EC2

REQUEST USING `curl`

```
curl -H "Content-Type: application/json" -X POST -d '{"orig_port_encoded": "0", "dest_port_encoded": "0", "days_to_flight": "10", "depart_yr": "2019", "depart_mo": "6", "depart_day": "13", "duration_total_min": "750", "stops": "1"}' "http://ec2-54-202-68-148.us-west-2.compute.amazonaws.com:5000/predict_airfare"
```

LAUNCHING FLASK API

```
[ec2-user@ip-172-31-27-175 docker]$ docker-compose up
Recreating docker_flask_1 ... done
Attaching to docker_flask_1
Flask_1   * Serving Flask app "server" (lazy loading)
Flask_1   * Environment: production
Flask_1   WARNING: Do not use the development server in a production environment.
Flask_1   Use a production WSGI server instead.
Flask_1   * Debug mode: on
Flask_1   * Running on http://0.0.0.0:5000/ (Press CTRL+C to quit)
Flask_1   * Restarting with stat
Flask_1   * Debugger is active!
Flask_1   * Debugger PIN: 690-542-898
Flask_1   72.197.183.122 - - [03/Jun/2019 07:51:28] "GET / HTTP/1.1" 200 -
```

EXAMPLE: REQUEST PREDICTION

```
Amazon Linux AMI
https://aws.amazon.com/amazon-linux-ami/2018.03-release-notes/
[ec2-user@ip-172-31-27-175 ~]$ curl -H "Content-Type: application/json" -X POST -d '{"orig_port_encoded": "0", "dest_port_encoded": "0", "days_to_flight": "10", "depart_yr": "2019", "depart_mo": "6", "depart_day": "13", "duration_total_min": "750", "stops": "1"}' "http://ec2-54-202-68-148.us-west-2.compute.amazonaws.com:5000/predict_airfare"
{"airfare_prediction": 432.46229448441675}
```

CHALLENGES & FUTURE IMPROVEMENTS

SITE STRUCTURE CHANGES



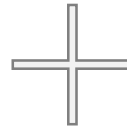
EXTREMELY SLOW SCRAPER



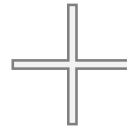
SEMI-CLUNKY CODE



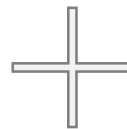
IMPROVEMENTS



Add FX comparisons /
access FX API



Run scraper daily at
consistent times to
capture changes in ticket
prices



Get better at  python™

DISCLAIMER

All mentions of Norwegian (Norwegian Air Shuttle) and uses of logos, colors and font schemes are registered trademarks of Norwegian Air Shuttle and subject to all trademark and copyright laws.

All data presented in this analysis is collected in adherence to Norwegian Air Shuttle data policies and used for personal academic and educational purposes only.