Secrets of Avoiding Flight Delays at NYC

Group 18

October 25, 2019

DATA SCIENCE 4 ALL

POWERED BY correlation one

Can we help passengers in NYC area to avoid flight delays with data analytics?

1. Overview of NYC Airports - EWR, JFK and LGA
2. Delay Patterns in Time
3. Weather Condition Analysis
4. Airline Analysis

Team Members

Sinem Uysal

PhD student in Operations Research and Financial Engineering at Princeton University

Mengya (Mia) Hu

PhD student in Mechanical & Aerospace Engineering at Princeton University

Sonal Savaliya

MS Computer Science, Pace University

Jianing Lu

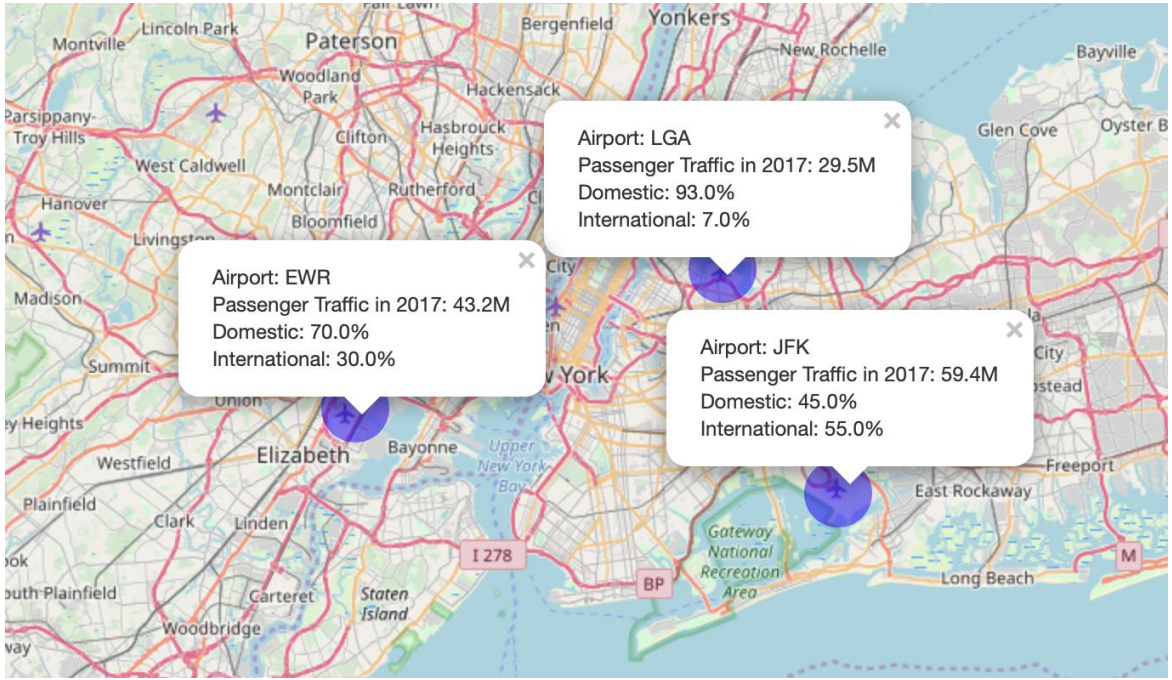
PhD student in Computational Chemistry at New York University

Overview of NYC Airports

Sinem

1. Passenger traffic and flight records in 2017
2. Frequency of delayed flights in departures and arrivals
3. Breakdown of delay reasons

3 Major Airports in NYC Area - EWR, JFK and LGA



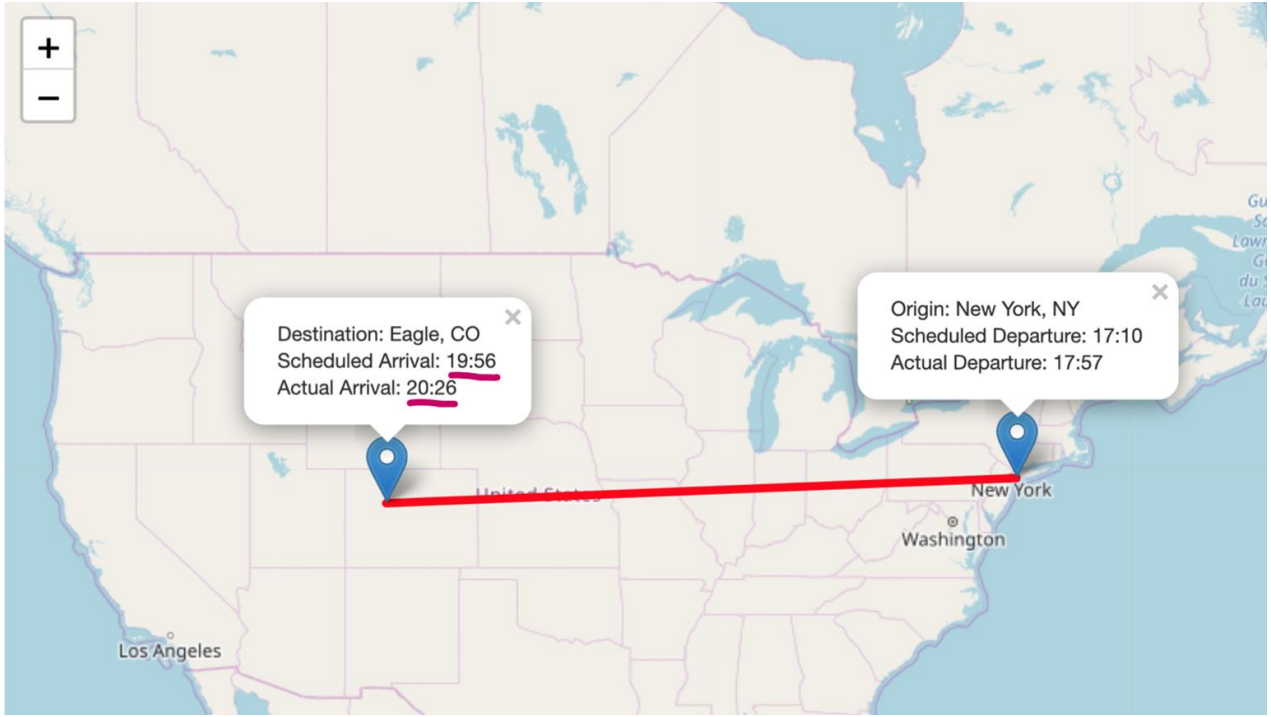
Dataset: Delays for US domestic flights in 2017 from *Bureau of Transportation Statistics*

Domestic flights in 2017

1. EWR ~ 23.4K
2. JFK ~ 18.9K
3. LGA ~ 18.6K

Passenger traffic at each airport is from Airport Traffic Report of The Port Authority of NY & NJ

What Does a Delayed Flight Mean?



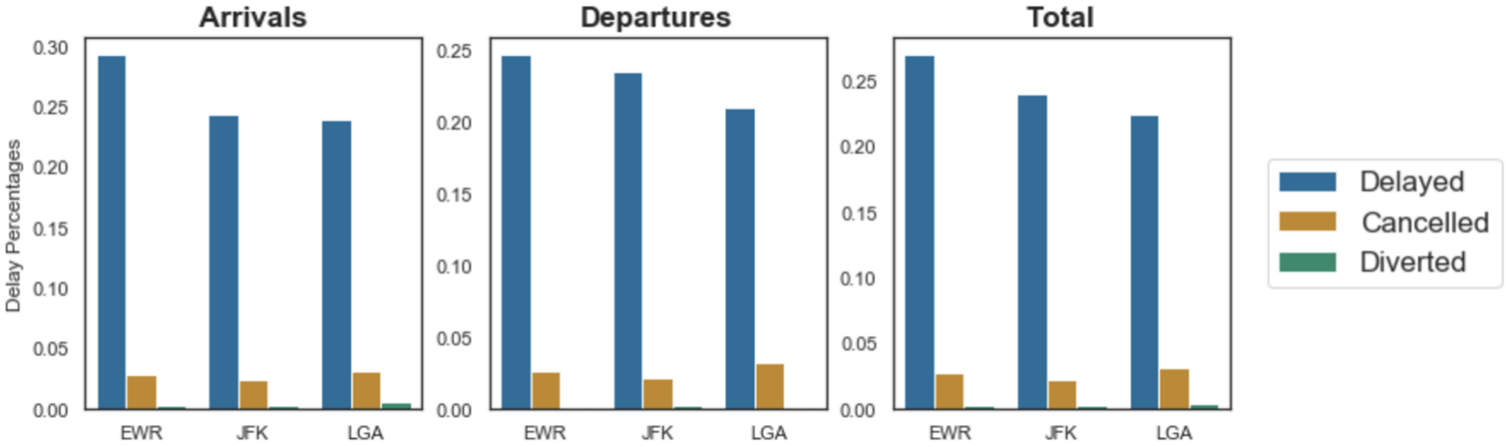
The *Federal Aviation Administration* considers a flight to be delayed when it is **15 minutes later than its scheduled time***

Delay is measured based on difference in arrival time.

This flight recorded as delayed for **30 min** in the dataset.

** Flight cancellation and delay Wikipedia entry*

Delay Percentages in 3 Airports

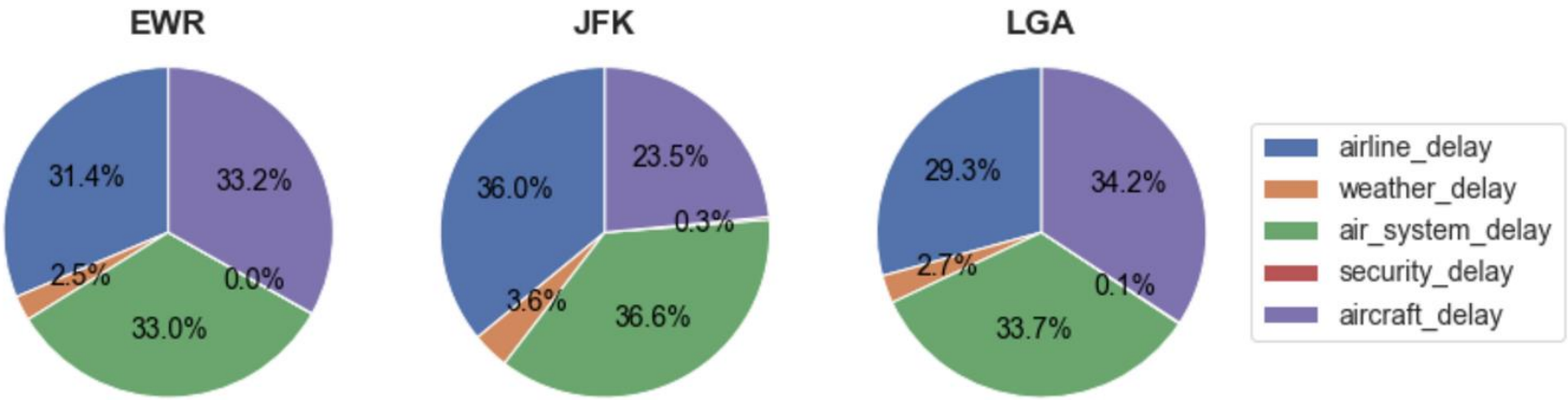


Average delay time
in total flights (min)

EWR: 73
JFK: 77.6
LGA: 75.6

- Cancelled and diverted flights are ~ 3%
- EWR has the highest delay percentage among 3 airports - **26.9%** (total flights)
- From now on we will look at **delays at departure flights**

Delay Reasons



airline_delay: maintenance or crew problems, aircraft cleaning, baggage loading, fueling, etc.
weather_delay: extreme weather conditions such as tornado, blizzard or hurricane
air_system_delay: non-extreme weather conditions, airport operations, heavy traffic volume, and air traffic control
aircraft_delay: late-arriving craft
security_delay: evacuation of a terminal or concourse, re-boarding of aircraft because of security breach etc.

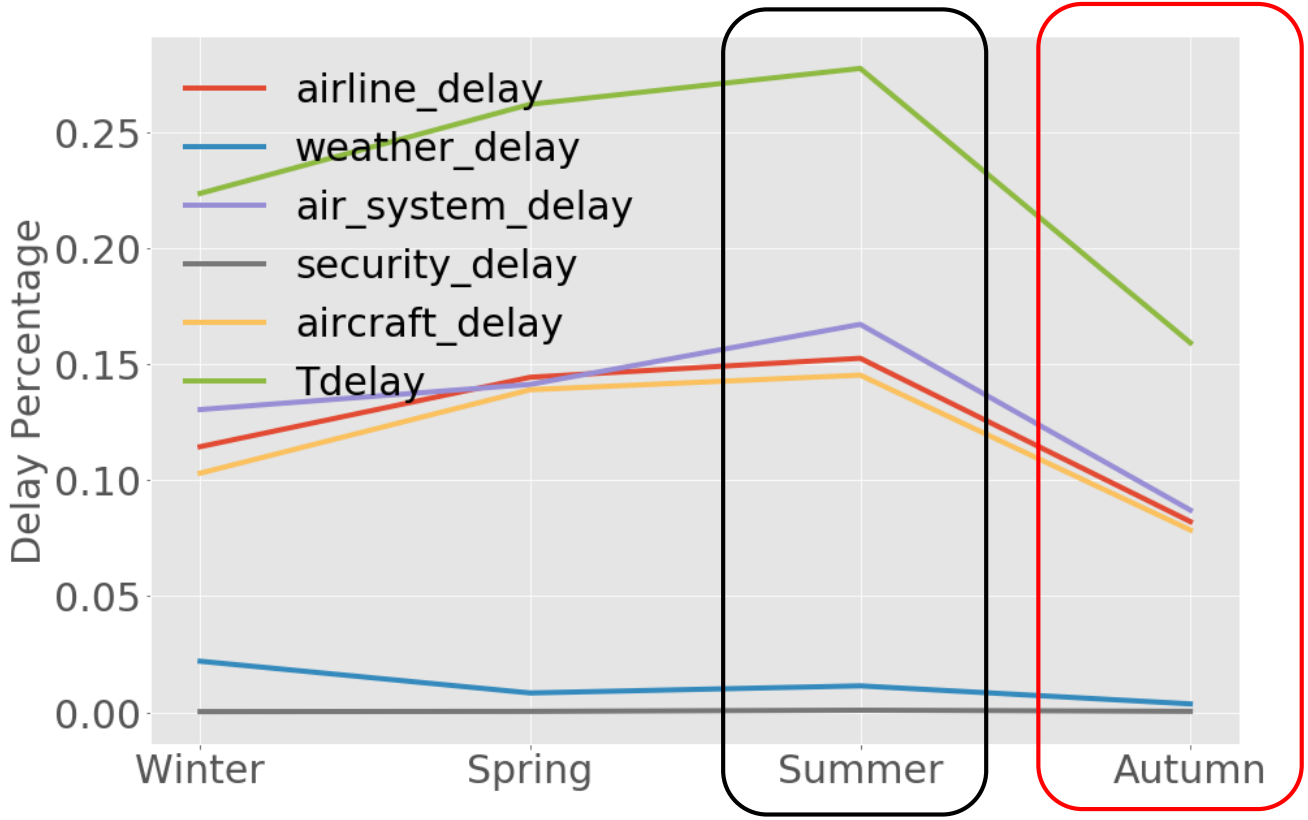
Delay Patterns in Time

Mengya (Mia)

1. Delay Percentage
    a. Seasons
    b. Months
    c. Days in a Week
    d. Time in a Day
2. Average Delay Time (if Delayed)
    a. Seasons
    b. Months
    c. Days in a Week
    d. Time in a Day

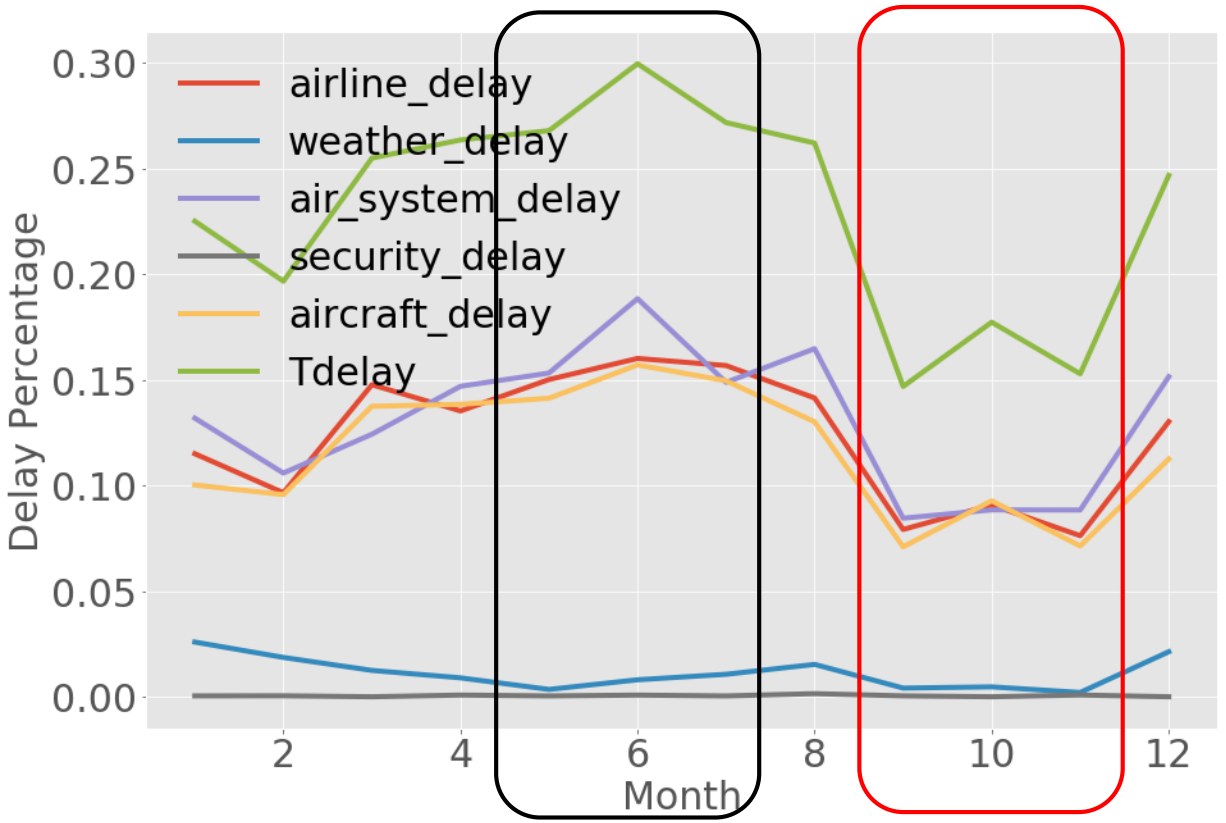
Delay Percentage: Seasons

Autumn is the best;
Summer is the worst.



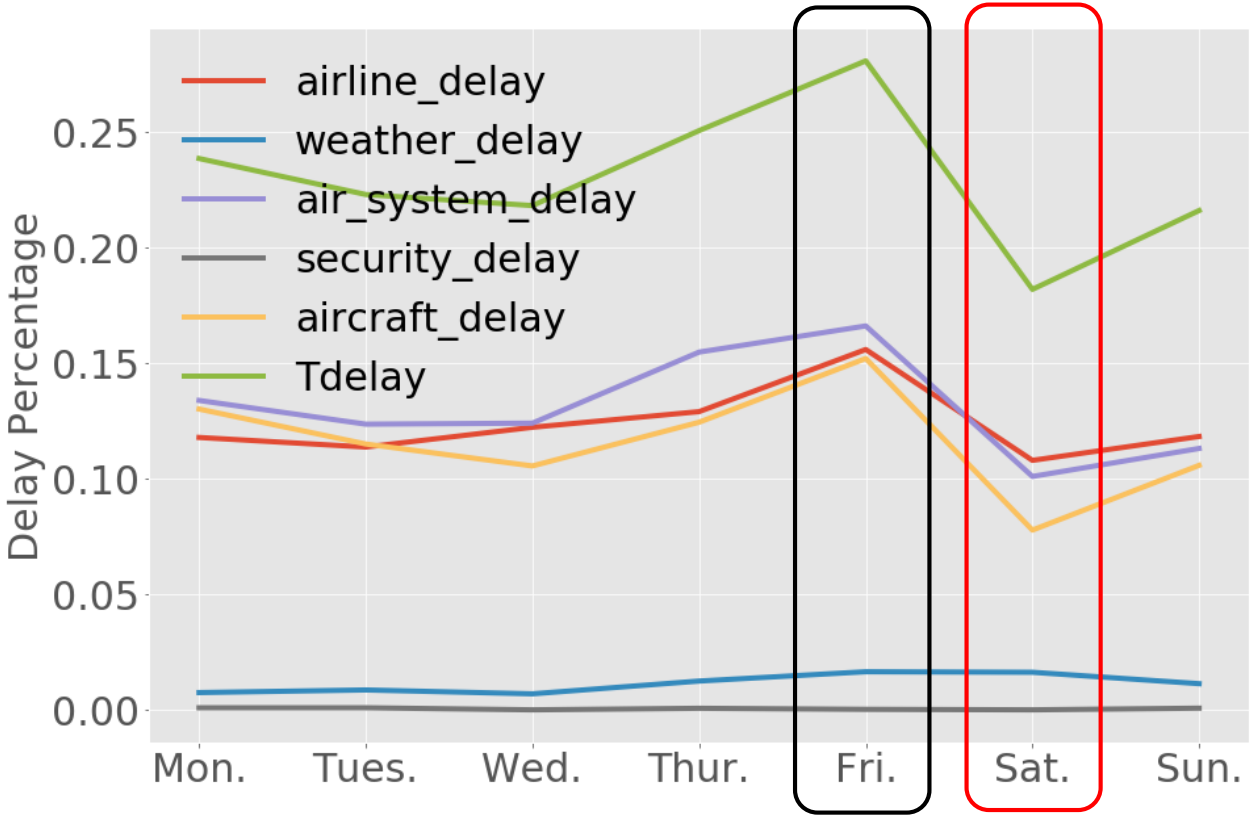
Delay Percentage: Months

Sept. to Nov. is the best;
June is the worst.



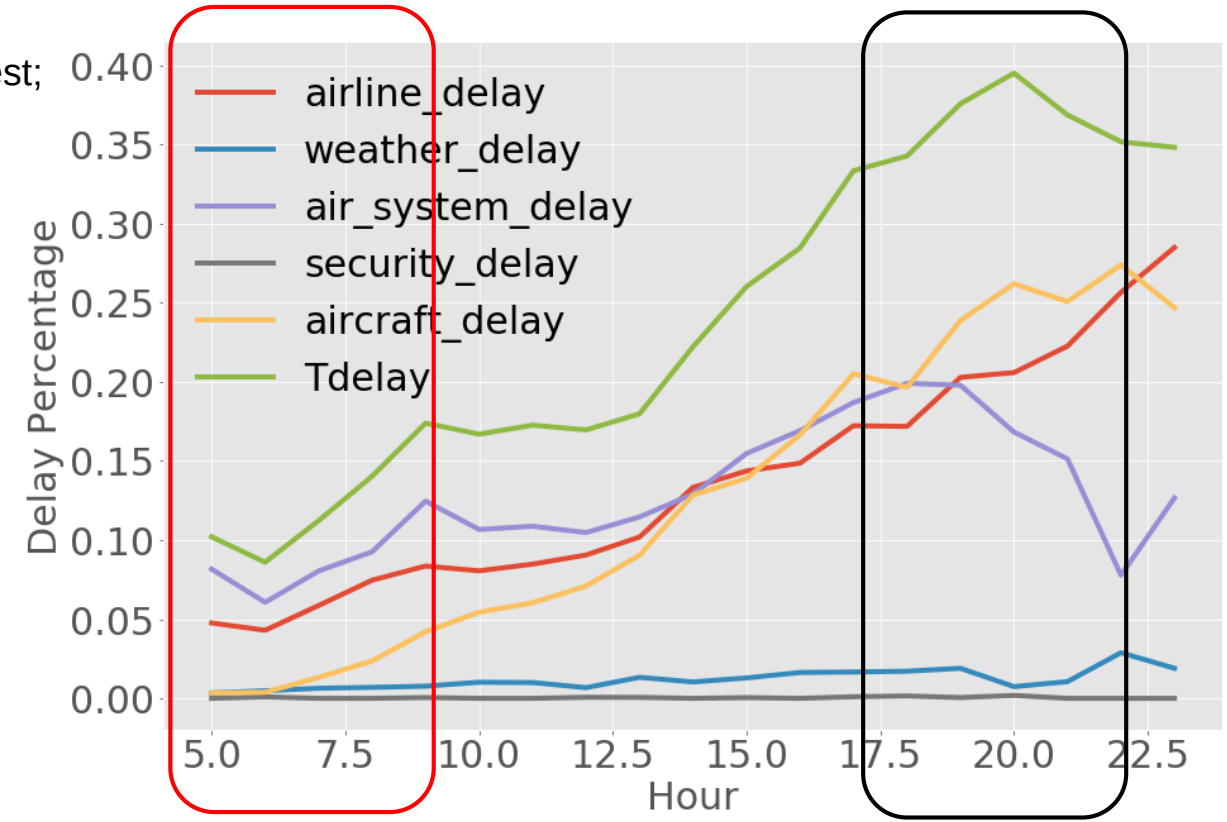
Delay Percentage: Days in a Week

Sat. is the best;
Fri. is the worst.

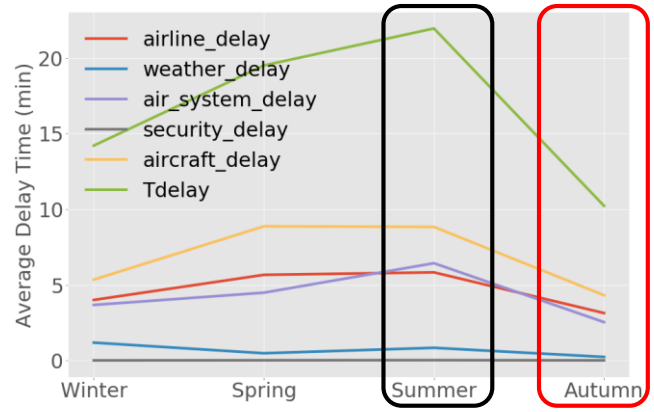


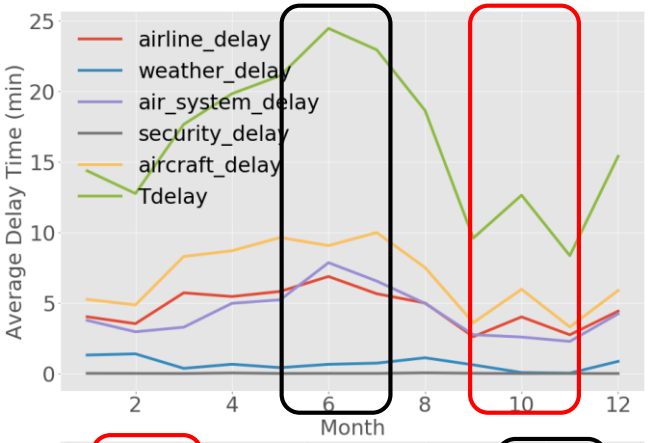
Delay Percentage: Time in a Day

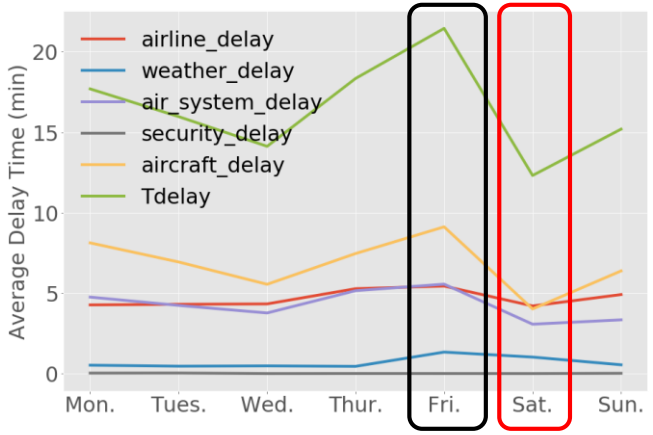
Early in the morning is the best;
Around 8pm is the worst.

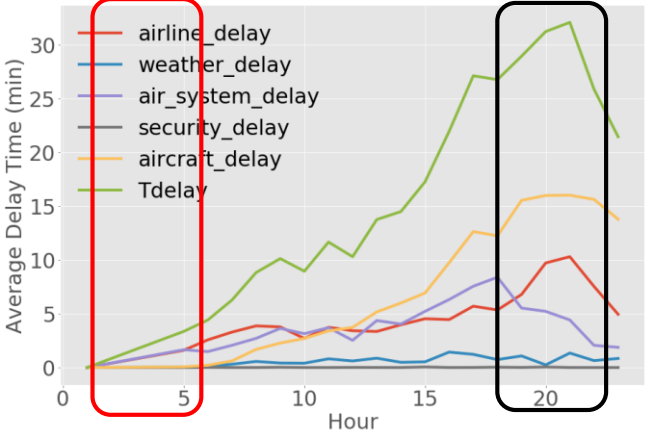


Average Delay Time (if Delayed)







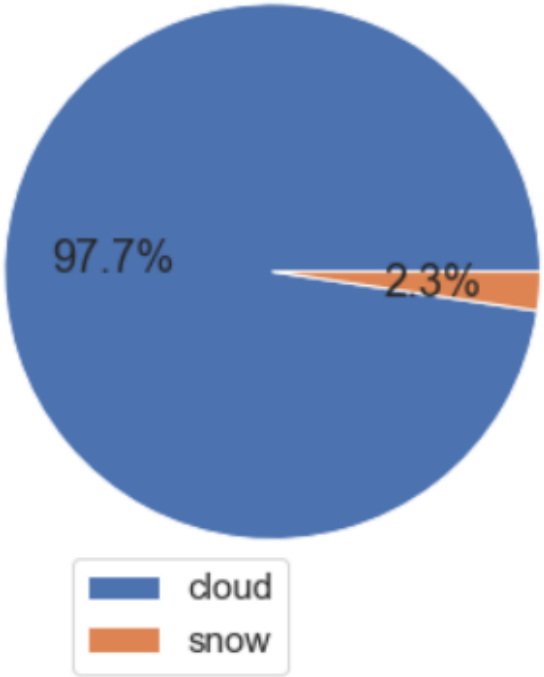


Weather Condition Analysis

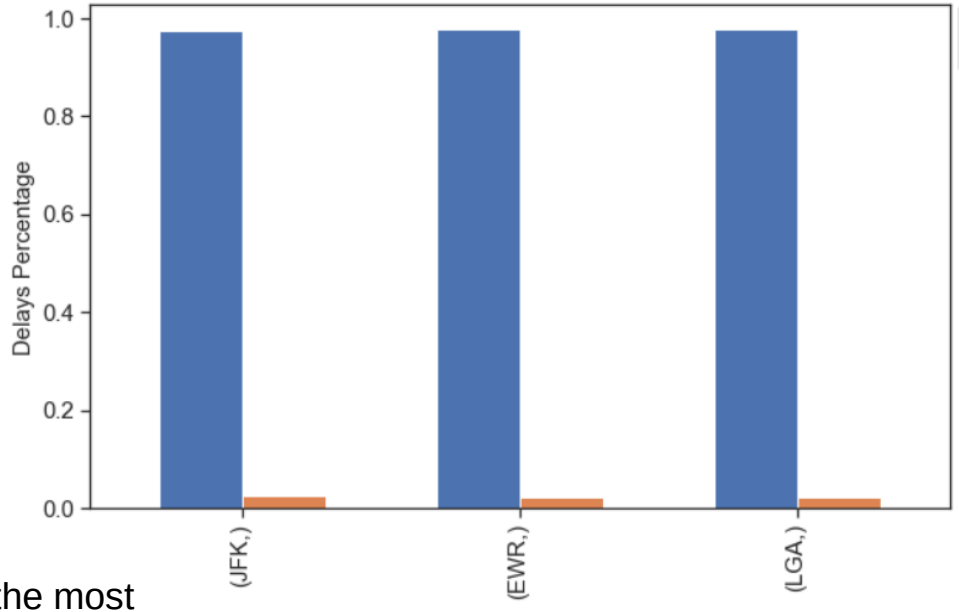
Sonal

1. Which weather condition is the worst
2. Weather delays frequency per month
3. weather conditions per airport

Weather Conditions Effect on Major NYC Airports

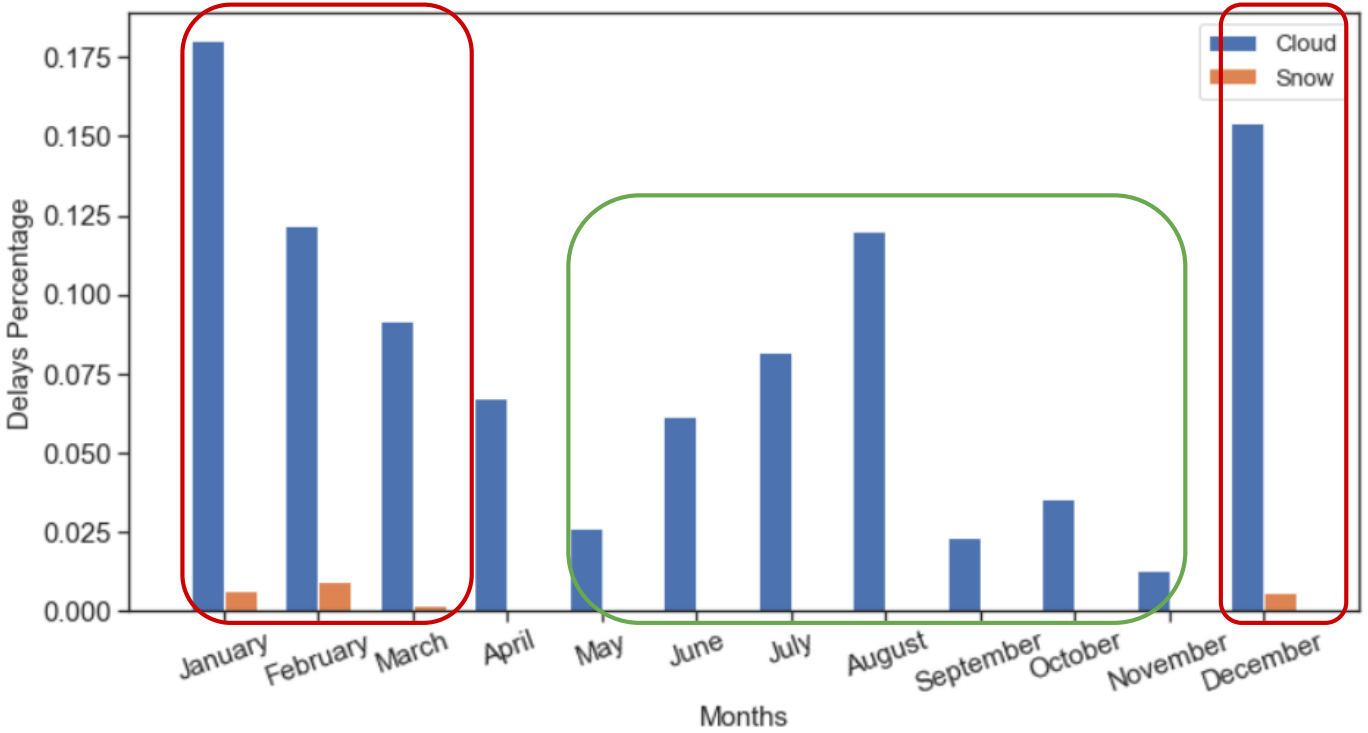


Dataset: Weather data collected at various US airports every 6 hours through 2017.
National Centers for Environmental Information



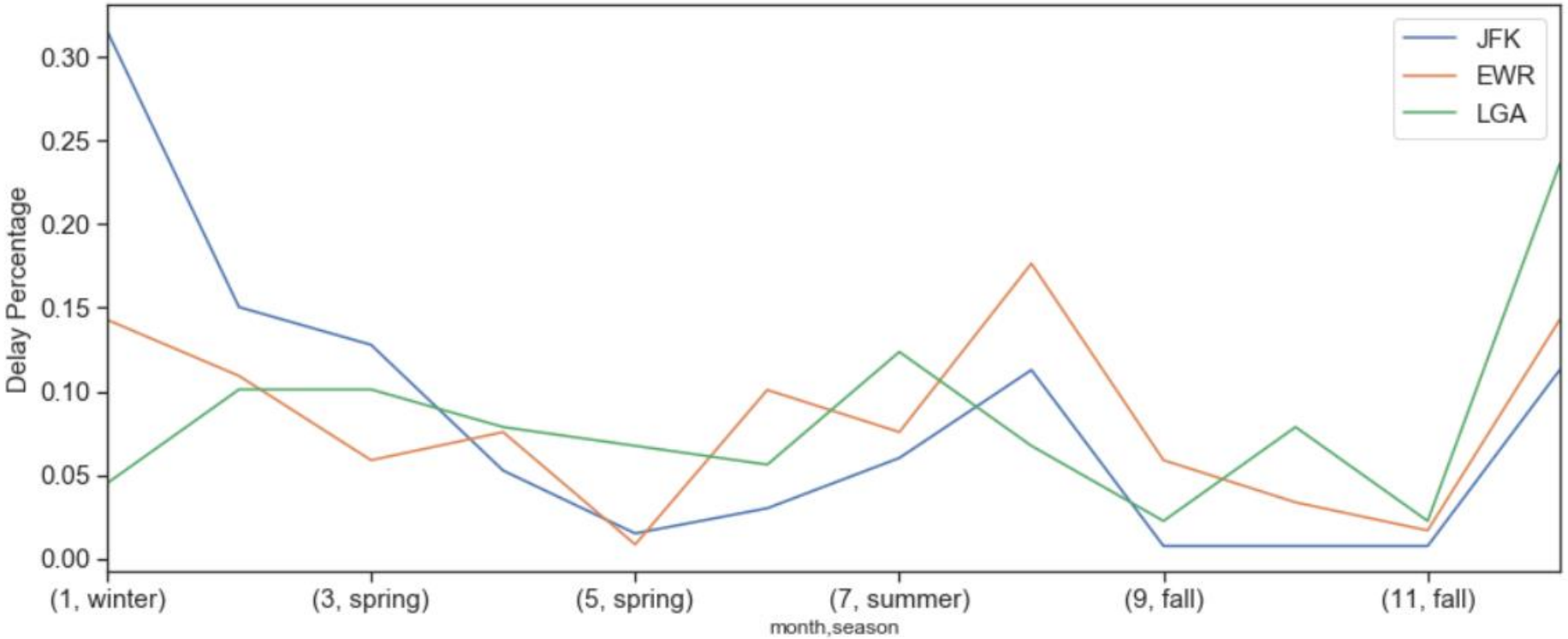
Cloudy weather condition affects the most
All three airports have **the similar effect for cloudy and snow** as they have same geographical region

Weather Condition: Months



Mostly **cloudy weather in winter and summer**
Snow affects during winter only

Weather Condition: Seasons



JFK is bad in winter. However, **LGA has less delays percentage** during extreme weather like winter

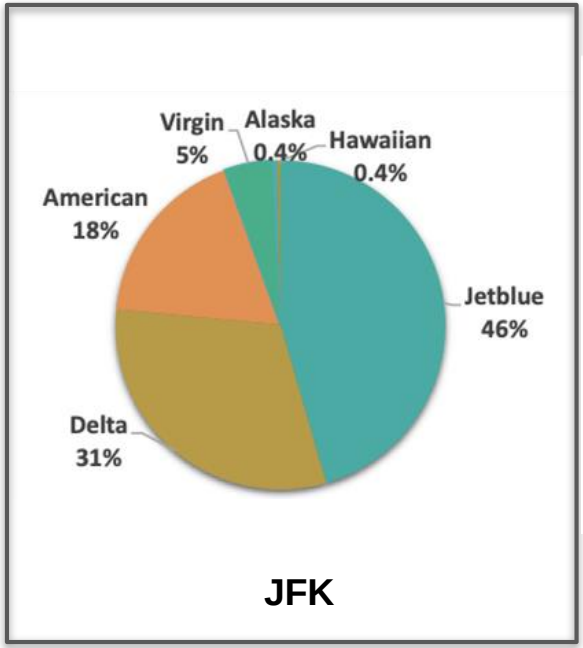
Airline Analysis

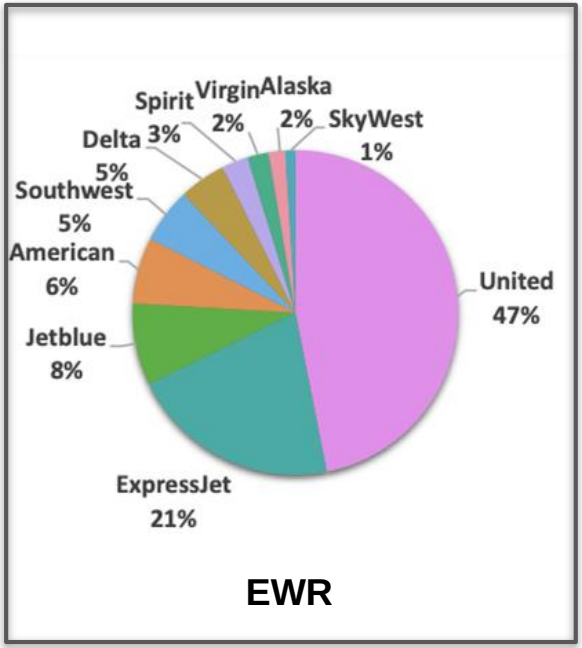
Jianing

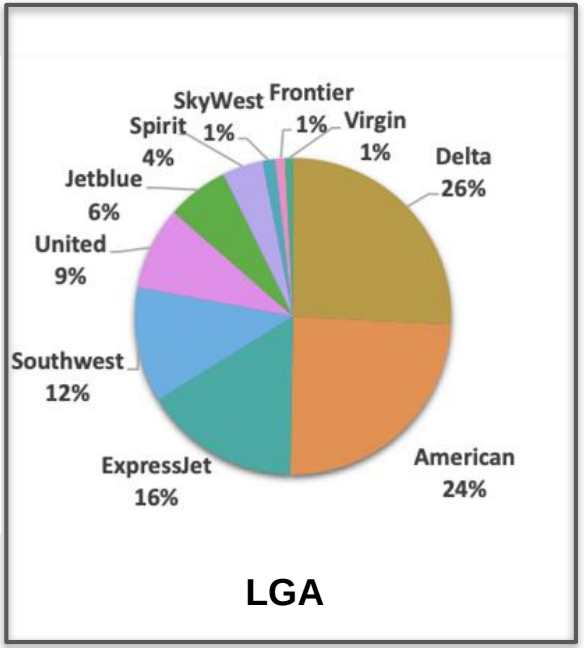
1. Airline Rank
2. Airline Delay Analysis

Airlines in NYC Airports

- Airline Percentage for NYC Airports





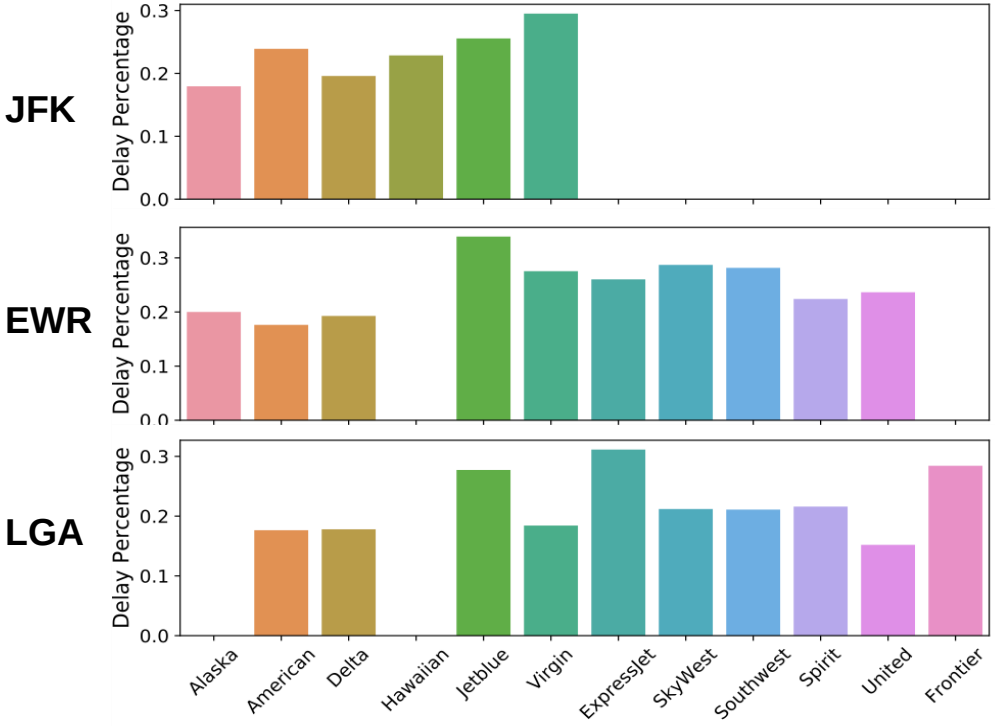


Airline Rank

- Airline Rank in 2017

Rank of airline is from [The Best and Worst US Airlines in 2018]

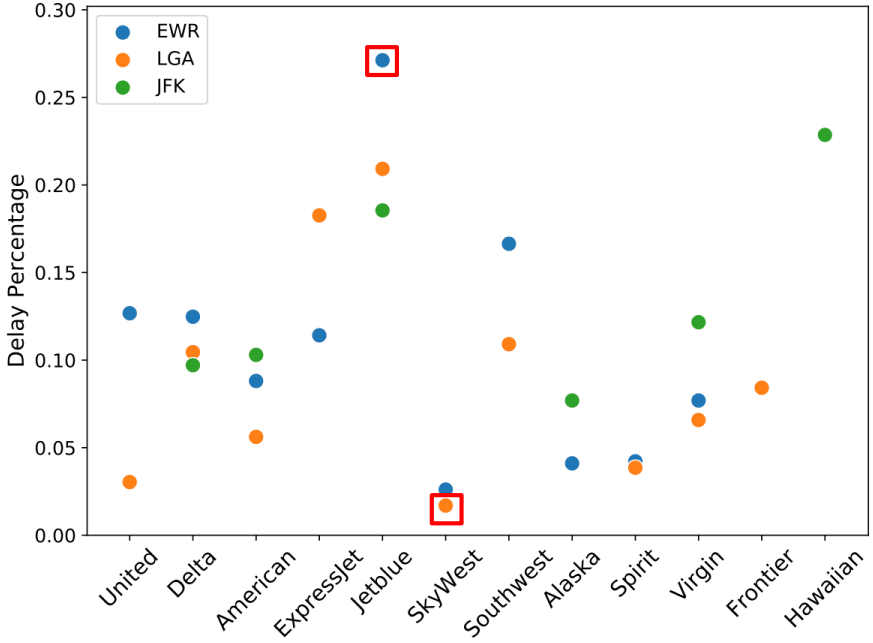
- Airline Rank for NYC Airports



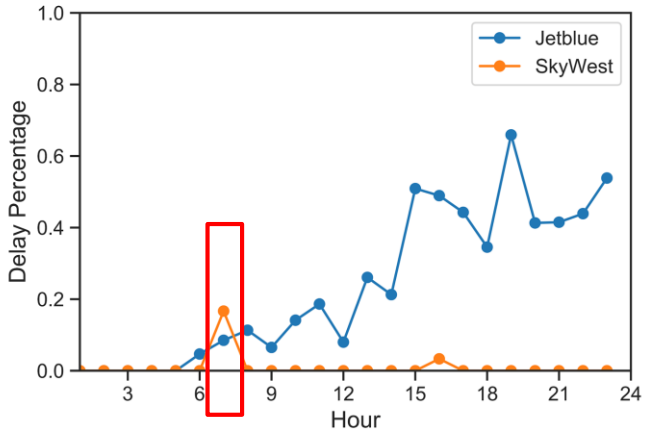
Delta, American, United are the best airlines for **JFK, EWR, LGA**

Airline Delay Analysis

- Airline_Delay Percentage

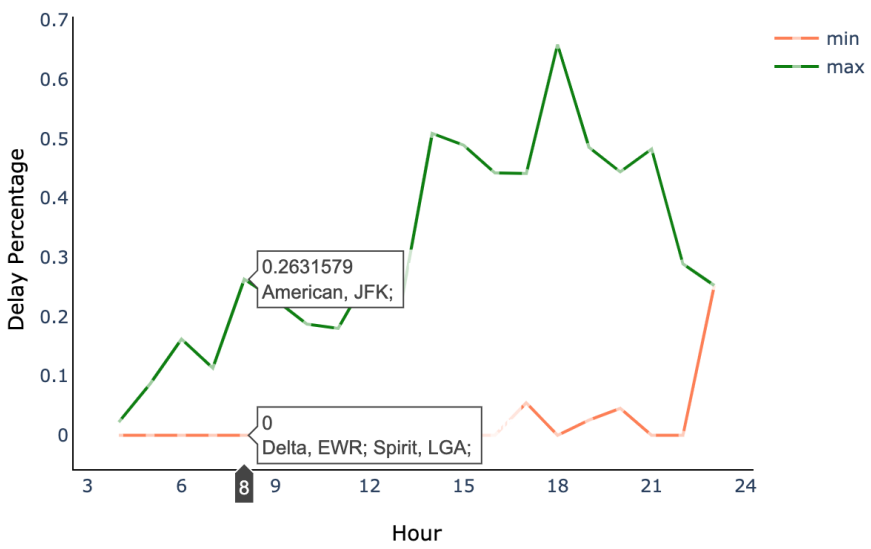


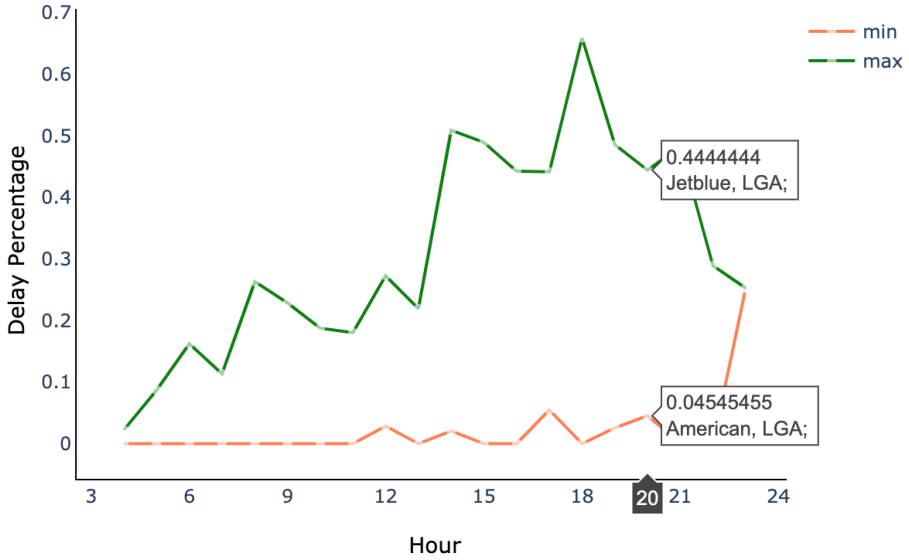
Time?



Airline_delay: maintenance or crew problems, aircraft cleaning, baggage loading, fueling, etc.

Best and Worst on Airline Delay : Time in a Day





Best / Worst airline ≠ always **Best / Worst** performance. **Time** is important.

Summary of Findings

1. **Early in the morning, saturday and autumn** are the best time to avoid delays. Late in the evening, Friday and summer are the worst..
2. **LGA and EWR** are the best airports to avoid delays due to extreme weather. JFK performs worst in extreme weather condition.
3. Generally, **Delta, American, United** are the best airlines for **JFK, EWR, LGA**, respectively. With consideration of time, you can further optimize your airline choice.

Thank You!

DATA SCIENCE 4 ALL

POWERED BY correlation one

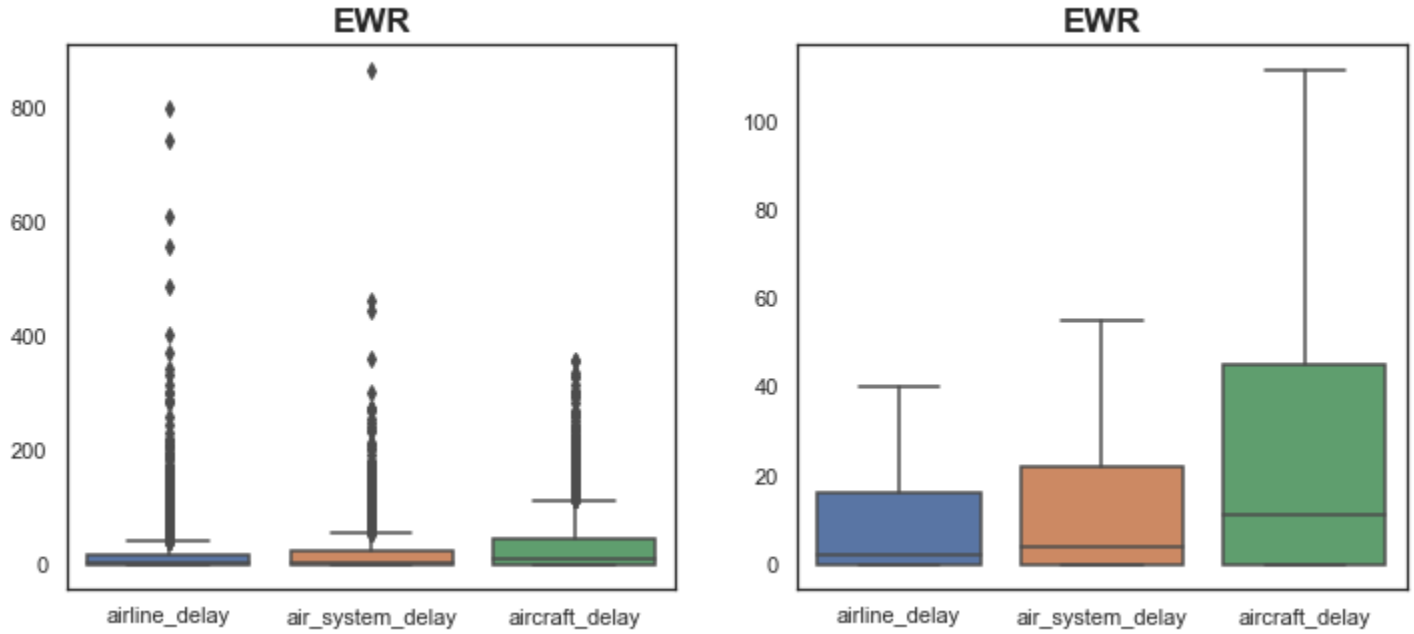
APPENDIX

Delayed flight data in page 6

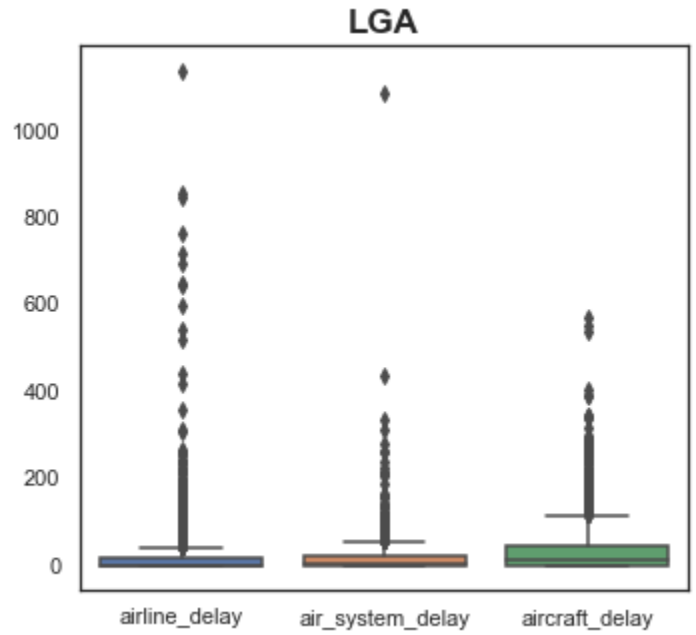
In [375]:
```
display(first_delay)
```

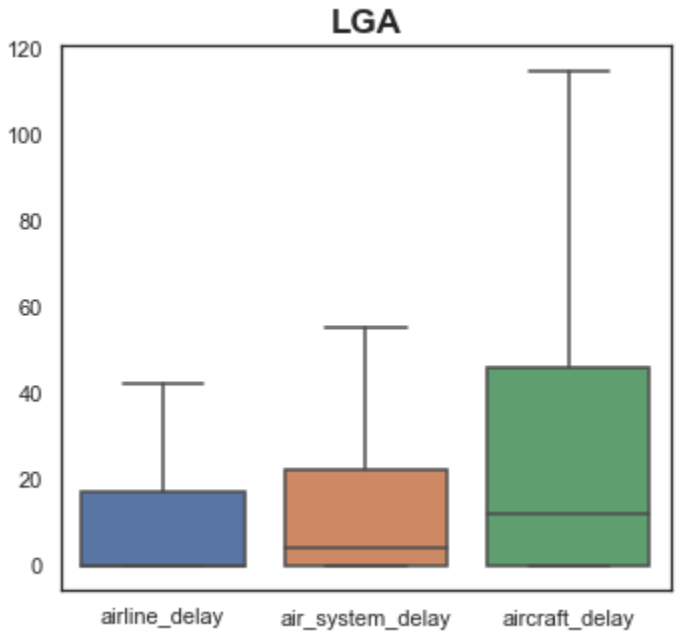
```
year                                    2017
month                                      1
day                                       22
airline_id                                AA
origin_airport                           JFK
destination_airport                      EGE
scheduled_departure                     1710
actual_departure                        1757
taxi_out                                  22
wheels_off                              1819
wheels_on                               2021
taxi_in                                    5
scheduled_arrival                       1956
actual_arrival                          2026
cancelled                                  0
diverted                                   0
scheduled_elapsed                        286
actual_elapsed                           269
distance                                1746
airline_delay                              0
weather_delay                              0
air_system_delay                           0
security_delay                             0
aircraft_delay                            30
airline_name           American Airlines Inc.
Name: 656, dtype: object
```

Delay time distribution

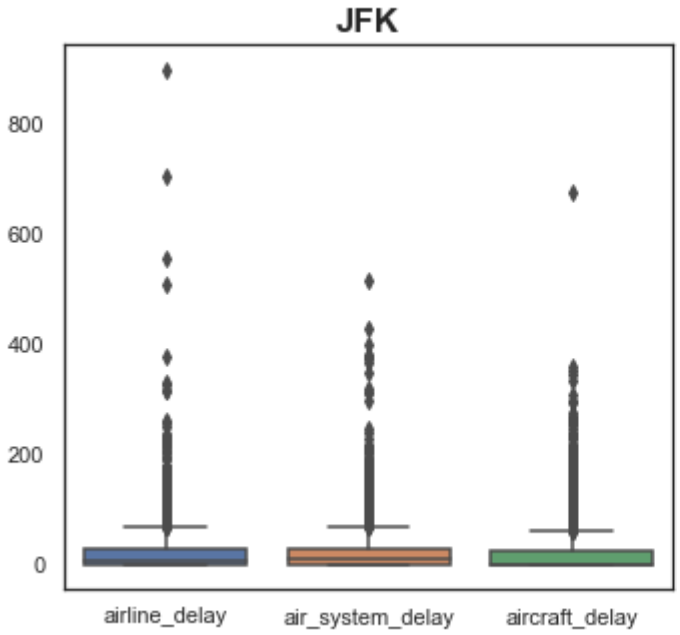


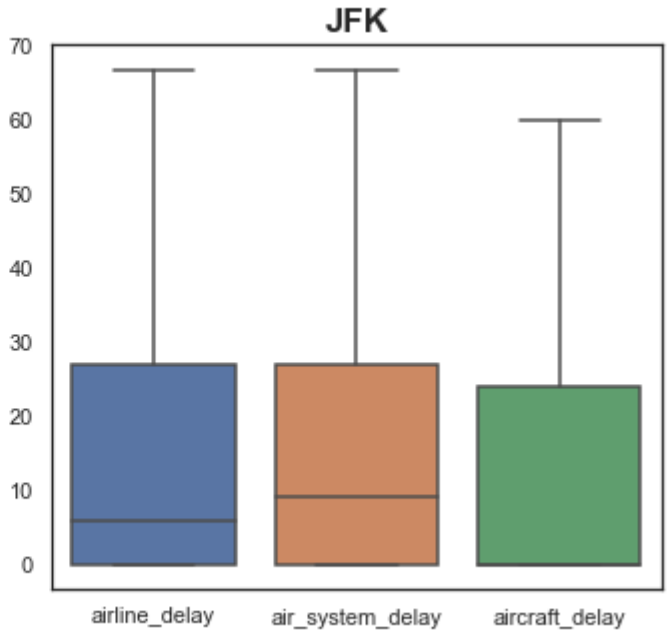
Delay time distribution





Delay time distribution





Hypothesis Testing for Airlines in NYC Airports

JFK:Delta

	Contrast	A	B	Paired	Parametric	T	dof	Tail	p-unc	p-corr	p-adjust	BF10	hedges
1	airline_name	American	Delta	False	True	3.394	3324.45	two-sided	6.960593e-04	1.044089e-02	bonf	10.645	0.106
5	airline_name	Jetblue	Delta	False	True	6.012	6626.26	two-sided	1.934644e-09	2.901966e-08	bonf	1.753e+06	0.142
9	airline_name	Delta	Alaska	False	True	0.262	39.07	two-sided	7.950366e-01	1.000000e+00	bonf	0.179	0.041
10	airline_name	Delta	Virgin	False	True	-4.334	549.77	two-sided	1.743733e-05	2.615600e-04	bonf	610.863	-0.245
11	airline_name	Delta	Hawaiian	False	True	-0.452	34.71	two-sided	6.543927e-01	1.000000e+00	bonf	0.2	-0.082

EWR: American

	Contrast	A	B	Paired	Parametric	T	dof	Tail	p-unc	p-corr	p-adjust	BF10	hedges
0	airline_name	American	Jetblue	False	True	-7.928	1724.97	two-sided	3.956474e-15	1.780413e-13	bonf	1.095e+12	-0.375
1	airline_name	American	Delta	False	True	-0.757	1145.70	two-sided	4.489063e-01	1.000000e+00	bonf	0.083	-0.043
2	airline_name	American	ExpressJet	False	True	-5.144	1470.31	two-sided	3.048430e-07	1.371793e-05	bonf	2.262e+04	-0.197
3	airline_name	American	Spirit	False	True	-1.742	522.07	two-sided	8.202689e-02	1.000000e+00	bonf	0.335	-0.122
4	airline_name	American	SkyWest	False	True	-2.488	138.94	two-sided	1.403242e-02	6.314589e-01	bonf	2.178	-0.283
5	airline_name	American	United	False	True	-4.052	1058.90	two-sided	5.458871e-05	2.456492e-03	bonf	150.009	-0.143
6	airline_name	American	Southwest	False	True	-4.695	1262.95	two-sided	2.956784e-06	1.330553e-04	bonf	3022.298	-0.254
7	airline_name	American	Alaska	False	True	-0.749	288.88	two-sided	4.545515e-01	1.000000e+00	bonf	0.117	-0.062
8	airline_name	American	Virgin	False	True	-3.136	367.13	two-sided	1.849750e-03	8.323875e-02	bonf	9.954	-0.249

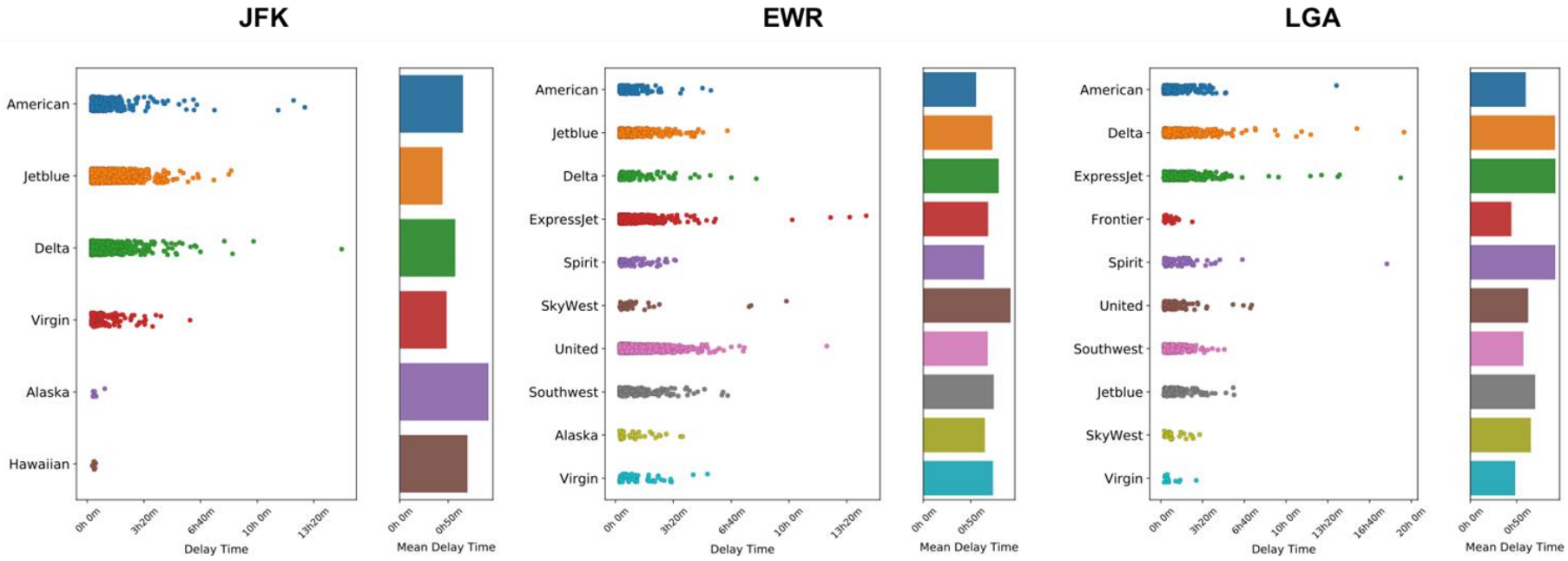
LGA:United

	Contrast	A	B	Paired	Parametric	T	dof	Tail	p-unc	p-corr	p-adjust	BF10	hedges
4	airline_name	American	United	False	True	1.626	1449.15	two-sided	1.042592e-01	1.000000e+00	bonf	0.172	0.065
12	airline_name	Delta	United	False	True	1.740	1420.36	two-sided	8.212896e-02	1.000000e+00	bonf	0.207	0.069
19	airline_name	ExpressJet	United	False	True	9.086	1979.24	two-sided	2.420162e-19	1.089073e-17	bonf	1.558e+16	0.371
25	airline_name	Frontier	United	False	True	2.743	108.64	two-sided	7.131778e-03	3.209300e-01	bonf	4.384	0.357
30	airline_name	Spirit	United	False	True	2.615	685.49	two-sided	9.113944e-03	4.101275e-01	bonf	2.004	0.170
35	airline_name	United	Southwest	False	True	-3.326	1812.14	two-sided	8.973980e-04	4.038291e-02	bonf	12.482	-0.152
36	airline_name	United	Jetblue	False	True	-5.578	1096.27	two-sided	3.056004e-08	1.375202e-06	bonf	2.519e+05	-0.314
37	airline_name	United	SkyWest	False	True	-1.504	145.02	two-sided	1.348483e-01	1.000000e+00	bonf	0.325	-0.164
38	airline_name	United	Virgin	False	True	-0.694	87.67	two-sided	4.894423e-01	1.000000e+00	bonf	0.166	-0.089

Best-worst for Airline_delay

	Contrast	A	B	Paired	Parametric	T	dof	Tail	p-unc	BF10	hedges
0	airline_name	Jetblue	SkyWest	False	True	13.598	559.77	two-sided	1.355807e-36	2.928e+35	0.602

Delay Time for Different Airlines



Summary for Airline

1. Generally, Delta, American, United are the best airlines for JFK, EWR, LGA, respectively.
2. Considering the pure performance of airline companies, Jetblue is the worst and SkyWest is the best.
3. You can further choose specific airline for different time of a day using our analysis result to avoid high probability of airline_delay.

In air system delay, there is more than 50% are related to weather

Causes of National Aviation System Delays
National (January - December, 2017)

Most Recent Month Year to Date View Individual Months View Tabular Version

More Topics:

- On-Time Arrival Performance
- Flight Delays by Cause
- Weather's Share of Delayed Flights
- Weather's Share of National Aviation System (NAS) Delays