



Ali Syed Chief Data Scientist NYC's 3 Airports Flights' Delay Data Analysis Findings

Introduction

Project

• This dataset contains information about all flights that departed from and arrived at NYC's 3 airports e.g. Newark (EWR), JFK and LaGuardia (LGA)

Objective

 Analyze Flights' delays (both departure and arrival) from New York City's 3 Airports

Scope

• Dataset contain a total of 336,776 flights data operated through NYC's 3 airports

Summary of Findings

Flight Delays findings

- Newark Airport (EWR) has highest number of flights (120,835)
 ; followed by JFK (111,279) and LGA (104,662);
- * LaGuardia (LGA) is the most efficient airport among three airports in NYC.
- EWR has highest average delay both in departure (15.1 minutes);
- * Median delay (both depart and arrived) for all three airports are negative indicates i.e. more than half of the flights depart/arrived early.
- * No significant difference in efficiency levels of all three NY's airports has been noted as delay Variation (SD) for both arrival and departure are almost equal for all three airports;

Summary of Findings (cont...)

- * Most of the flights (over 50%) for all three airports were delayed for only few minutes;
- * Highest number of flights delayed (both departure and arrival) in summer months (i.e. June and July) followed by December;
- * Most of the flights delayed in PM hours while Flights arriving in the AM hours are, on average, couple of minutes early;
- Airplanes' average speed is higher in summer months

Popular Destinations

- * Most frequent destination from NY's airports (all 3 combined) was Orlando (ORD) followed by Atlanta (ALT);
- *❖ JFK to LAX is the most popular route followed by LGA to ALT;*

Summary of Findings (cont...)

- Most frequent destination from EWR was ORD followed by BOS;
- Most popular destination from LGA was ALT followed by ORD;
- No significant difference in efficiency levels of all three NY's airports has been noted as delay Variation (SD) for both arrival and departure are almost equal for all three airports;

Flights Operated by Months:

- Additional flights operated in the Schools' holidays months i.e. July, August, December, March;
- * EWR and LGA airports has more flights in Spring and Summer months (i.e. March to August) compare fall/winter months;
- JFK airport is more busier in the school holidays months (i.e. March, July and August);

Dataset Info()

# of Records	336,776
# of Variables	19

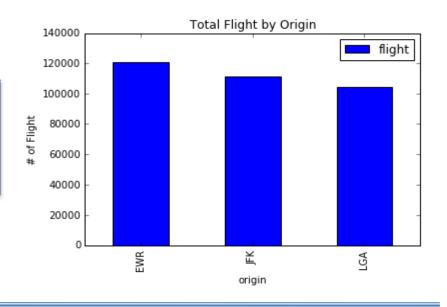
Variable	Description
year	2015
month	1-12
day	Day of the month (1-31)
dep_time	Departure times, local timezone
sched_dep_time	Scheduled departure time
dep_delay	Departure delay, in minutes, Negative times represent early departures
arr_time	Arrival times, local timezone
sched_arr-time	Scheduled departure time
arr_delay	Arrival delay, in minutes, Negative times represent early arrivals
carrier	Two letter carrier abbreviation
flight	Flight number
tailnum	Plane tail number
origin, dest	Airport codes for origin and destination
air_time	Amount of time spent in the air, in minutes.
distance	Distance flown, in miles.
hour, minute	Time of departure broken in to hour and mins.
time_hour	Timestamp

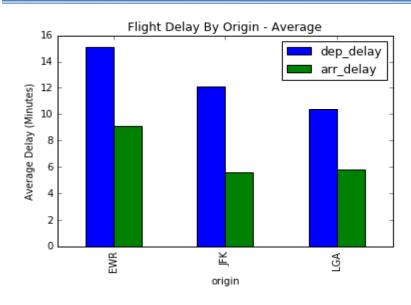
Dataset Describe()

Variables	count	mean	std	min	25%	50%	75%	max
year	336,776	-	-	-	-	-	-	-
month	336,776	-	-	•	-	-	1	-
day	336,776	-	-	-	-	-	1	-
dep_time	328,521	-	-	-	-	-	1	-
sched_dep_time	336,776	1,344.3	467.3	106.0	906.0	1,359.0	1,729.0	2,359.0
dep_delay	328,521	12.6	40.2	-43.0	-5.0	-2.0	11.0	1,301.0
arr_time	328,063	1,502.1	533.3	1.0	1,104.0	1,535.0	1,940.0	2,400.0
sched_arr_time	336,776	1,536.4	497.5	1.0	1,124.0	1,556.0	1,945.0	2,359.0
arr_delay	327,346	6.9	44.6	-86.0	-17.0	-5.0	14.0	1,272.0
flight	336,776	1,971.9	1,632.5	1.0	553.0	1,496.0	3,465.0	8,500.0
air_time	327,346	150.7	93.7	20.0	82.0	129.0	192.0	695.0
distance	336,776	1,039.9	733.2	17.0	502.0	872.0	1,389.0	4,983.0
hour	336,776	-	-	-	-	-	-	-
minute	336,776	-	-	-	-	-	-	-

of Flights and Flights Delay by Origin

• Newark Airport (EWR) has highest number of flights (120,835); followed by JFK (111,279) and LGA (104,662)

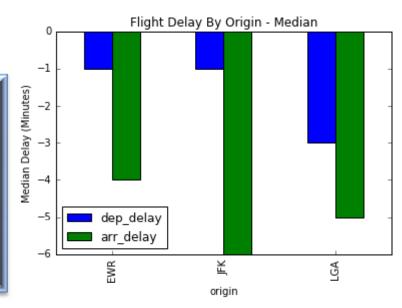


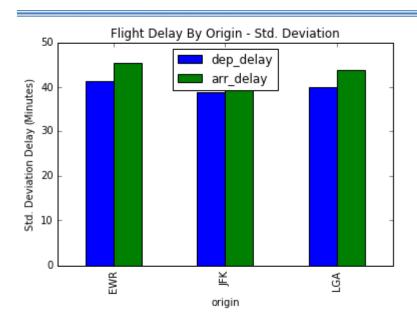


- EWR has highest average delay both in departure (15.1 minutes) and arrival (9.1 minutes); while LGA has lowest average delay (depart=1.3 min; arrival=5.8 min)
- LGA is the most efficient airport among three airports in NYC.

Flights Delay – Median and Standard Deviation

- Median delay for all three airports are negative indicates that more than half of the flights depart/arrived early. It highlights the efficiency of all three airports
- Note it is only logical that departure delay is lower than arrival and in range of 1 to 3 minutes early



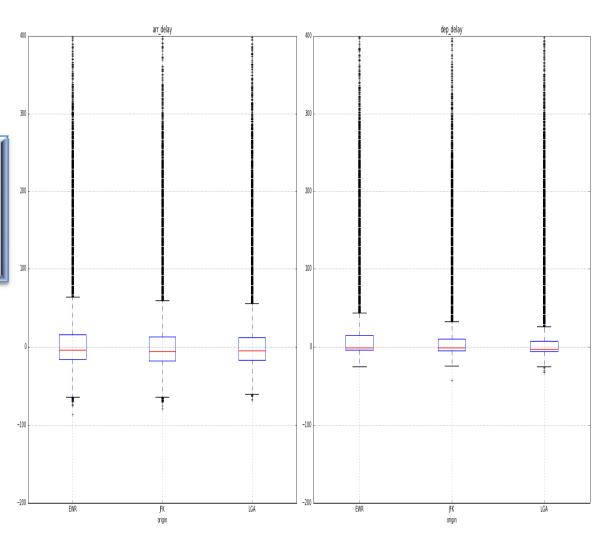


 Standard Deviation for delay for both arrival and departure are almost equal, indicates that all three airports have been working on pretty much on same standards;

Departure and Arrival Delay - Boxplots

 Box-plots of departure and arrival delays validated our findings that most of the flights (almost 50%) were delayed for only few

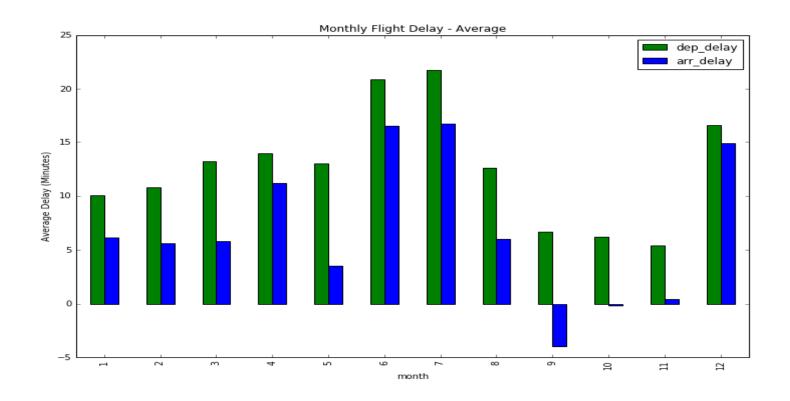
minutes.



Boxplot grouped by origin

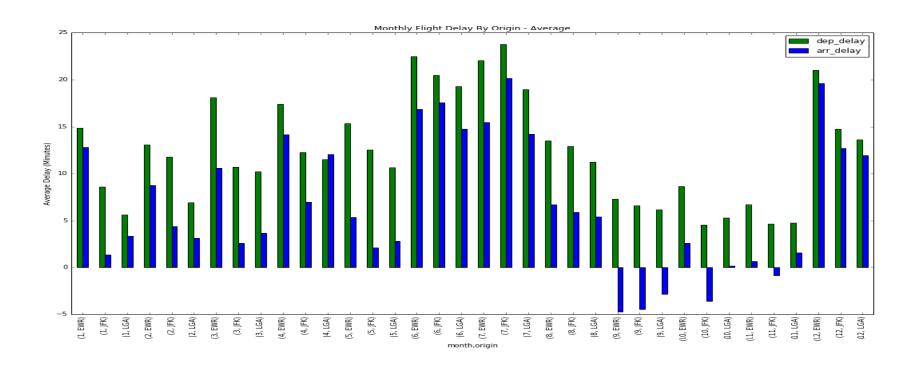
Average Flights Delay by months

Highest number of flights delayed (both departure and arrival) in summer months (i.e. June and July) followed by December



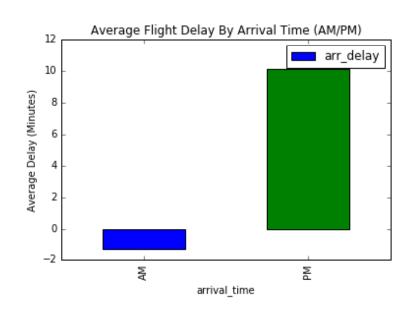
Average Flights Delay by Origin

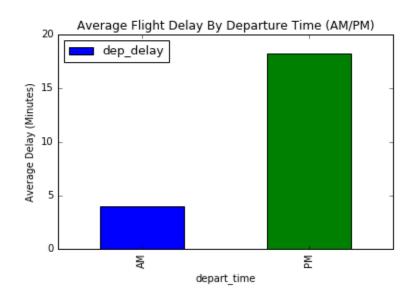
 Summer months have seen largest delays for all three airports both for arrival and departure



Flights Delayed Times (AM / PM)

 Most of the flights delayed in PM hours

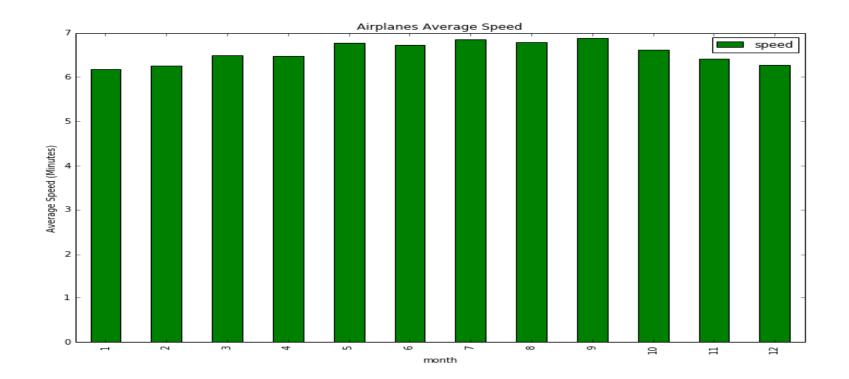




 Flights arriving on average couple of minutes early in the AM hours.

Airplanes Average Speed by months

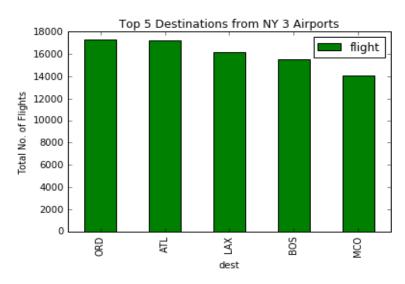
Airplanes' average speed is higher in summer months



Top 5 Destinations

 Most frequent destination from NYC airports (for all three of them) are ORD closely followed by ALT.

JFK to LAX is the most popular route followed by LGA to ALT



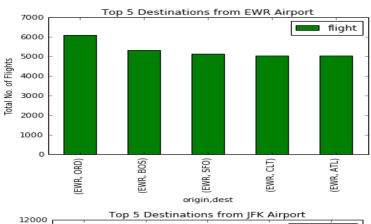


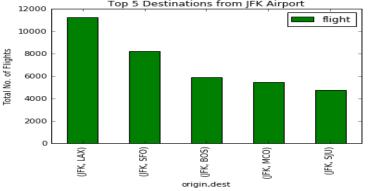
Top 5 Destinations by Origin

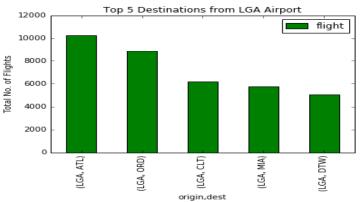
 Most frequent destination from EWR was ORD followed by BOS.

JFK to LAX is the most popular route

 Most popular destination from LGA was ALT followed by ORD.

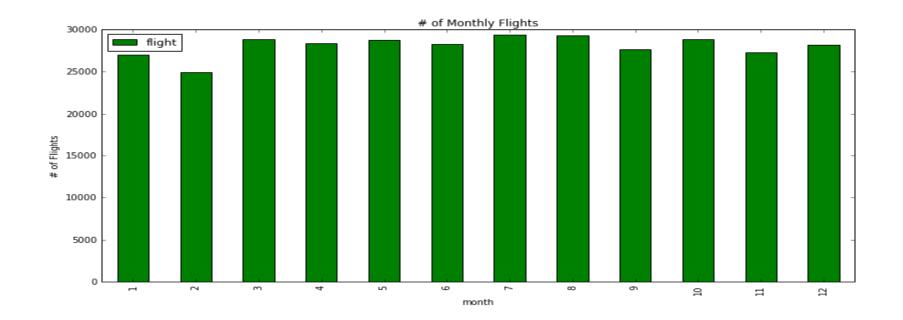






of Flights by months from all 3 NYC's Airports

 Additional number of flight were operated in the months that normally have schools closed i.e. July, August, December, March

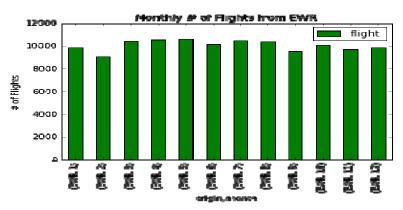


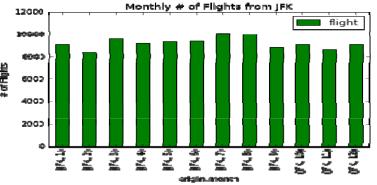
of Flights by Origin

 EWR airport has more flights in Spring and Summer months (i.e. March to August) compare to other months.

 JFK airport is more busier in the school holidays months (i.e. March, July and August)

 LGA airport is generally busier in spring and summer months with peaks in month of March and in October.







Thank You

Data is the new oil &

Analytics is the combustion engine

