# TEAM BAMF (Bootcamp Analytics Mystery Files)

# Project 1 Proposal: Airport Delays During Government Shutdown

*LeAnne Sevier, Luciana Roberts, Lillian Bolfer, Mike Machata, Rob Phillips*

# Hypothesis

“As the longest government shutdown in US history continues, Americans are restructuring their travel plans. National parks are either closed, trashed, or both, museum doors remain shut, and some people have even had to reschedule their weddings. But the place where perhaps the greatest number of people are feeling the effects of the partial shutdown is at the airport, which is stressful even when functioning at full capacity.” This comes from an article published on January 17th on Vox.com by Aditi Shrikant.

The most recent government shutdown spanned 35 days, from December 22, 2018 to January 25, 2018 and is the longest in history. Every sector of the US economy was impacted by the furloughing of government employees and resources or the essential services employees being expected to work with no pay. The financial and human impact will be the subject of research and study for years to come.

According to the Committee for a Responsible Federal Budget, “air travel was strained as a result of air traffic controllers and Transportation Security Administration (TSA) agents working without pay. Travelers faced longer lines as some TSA agents did not report to work and security checkpoints were closed, while the absence of 10 air traffic controllers temporarily stopped travel at LaGuardia Airport and caused delays at several major airports.”

Our group’s goal was to help shed insight on the impact of the government shutdown on air travel and subsequent delays. To accomplish this, we sampled flight delay data within the shutdown period for two major international airports, Hartsfield-Jackson Atlanta and Orlando International Airport, for one of the largest domestic airlines, Delta Air. Our expectation is that reduced staffing of TSA agents in security and air traffic controllers effect the volume and duration of flight departure delays.

Q: What data is available that might indicate how the government shutdown impacted the airports?

Although there are government data resources pertinent to air travel, and more specifically flight delays, this data is only current through November 2018. Our group turned to a commercial flight data service, Cirium Flightstats APIs, to obtain the data we needed.

Q: How can we source and visualize flight delay information across the USA, during the period of a government shutdown?

1) increase in number of delayed flights

2) length of flight delay

# A Guide To Air Travel During The Government Shutdown

As the longest [**government shutdown in US history continues**](https://www.vox.com/2019/1/3/18165728/new-congress-2019-democrats-government-shutdown), Americans are restructuring their travel plans.

**How is the shutdown affecting agents?**

TSA agents make up 51,000 of the 420,000 federal employees who are deemed “essential,” and while they will be paid for their work eventually, they have no clue when exactly that day will come.

**Are airport lines longer?**

Short answer: maybe. Although it depends on the airport, many major hubs have reported longer lines. “While national average wait times are within normal TSA times of 30 minutes for standard lanes ... some airports experienced longer than usual wait times,” [**TSA said in a statement**](https://www.tsa.gov/news/releases/2019/01/15/tsa-statement-checkpoint-operations-january-15).

At Hartsfield-Jackson Atlanta on January 15, travelers waited more than an hour in security lines. And at Dallas Love Field Airport, travelers waited 44 minutes.

**Are flights getting delayed or canceled?**

Not yet. Flight delays may be caused by winter storms across the Midwest, but so far, the government shutdown has not led to widespread delays or cancellations. There is, however, potential for future delays because of [**air traffic controllers**](https://qz.com/1522138/us-government-shutdowns-effect-on-air-travel/). Air traffic controllers are also essential employees, and therefore have also been working without paychecks. If they start calling in sick, the government may have to limit the amount of air traffic, but it hasn’t come to that yet

[United Airlines](https://www.cnbc.com/quotes/?symbol=UAL) CEO Oscar Munoz on Wednesday said the airline is getting worried about the partial U.S. government shutdown as it drags on but added that the company isn't yet seeing a "significant" impact on bookings.

Munoz's comments come a day after the CEO of rival [Delta Air Lines](https://www.cnbc.com/quotes/?symbol=DAL), Ed Bastian, said the shutdown that began Dec. 22 will cost the airline $25 million this month.

"There is some impact there," Munoz told CNBC's [Phil LeBeau](https://www.cnbc.com/phil-lebeau/) in an interview Wednesday. "It's not discernible and it's not significant. Clearly the longer this goes, of course there's going to be impact, and we do worry about that.“

On Tuesday, Delta's Bastian [said the airline is losing $25 million in revenue this month alone](https://www.cnbc.com/2019/01/15/delta-ceo-government-shutdown-is-costing-the-airline-25-million-this-month.html) because fewer government employees and contractors are traveling.

[Transportation Security Administration](https://www.cnbc.com/tsa/) officers, air traffic controllers and federal safety inspectors are among the some 420,000 government employees who are deemed essential and have been ordered to work without a regular paycheck.

Major airports in Atlanta, Miami, Houston and Washington, D.C., this week said they closed passenger screening lanes or checkpoints as more TSA officers than usual were absent.

* Dec 26 - Holiday travel should not be impacted by the partial federal government shutdown - https://www.usatoday.com/story/news/politics/2018/12/26/government-shutdown-impact-your-holiday-travel-plans/2414075002/
* Jan 8 - Air travelers start to feel effects of government shutdown - https://www.chicagotribune.com/news/nationworld/ct-tsa-airports-government-shutdown-20190108-story.html
* Jan 15 - TSA experienced a national rate of 6.8 percent of unscheduled absences compared to a 2.5 percent rate one year ago on the same day - https://www.tsa.gov/news/releases/2019/01/15/tsa-statement-checkpoint-operations-january-15
* Jan 16 - Several airports in the US have been forced to closed terminals due to a shortage of security workers following the partial shutdown of the US Federal Government. Airports affected include Miami International Airport in Florida and George Bush International Airport in Houston, Texas - https://www.airport-technology.com/news/us-government-shutdown-airport/
* Jan 25 - 10 air traffic controllers called in sick, six in northern Virginia and four in Florida, temporarily shutting down travel at New York's La Guardia airport and causing delays at other major hubs, including in New Jersey, Philadelphia, Orlando and Atlanta - https://www.cnn.com/2019/01/25/us/air-traffic-controller-shortage-faa/index.html

# Asking The Right Questions

**What data can we analyze to determine impacts on flights during the period of the government shutdown?**

* Cirium Flightstats API

<https://helpdesk.cirium.com/hc/en-us/categories/202631018-FlightStats-by-Cirium>

**What are we looking for in the data specifically?**

* An increase in the number of flights delayed
* An increase in the delayed departure times

**How do we want to frame or scope the investigation?**

* Limit the analysis to one Airline (Delta Airlines)
* Limit the analysis to two Airports (Major Hub Atlanta and Orlando)
* Afternoon flights where delays might accumulate (2pm-8pm)
* First and final weeks of the government shutdown period

**What are we comparing this data to?**

* Data for the first and final weeks of the previous year

Date Range A: Dec 22 2017 – Jan 25 2018

Date Range B: Dec 22 2018 – Jan 25 2019

**Potential variables not considered:**

* Air Line issues: The number one cause of flight delays are factors within the air line’s control. Issues such as maintenance, crew and hiring, cleaning, fueling, and baggage loading.
* Inclement weather. Major weather events such as tornados, snow storms, heavy rain, and fast winds make it unsafe for an airplane to take off.
* Airport security issues. Acts of vandalism, terrorism, threats, and extended security screenings will cause major delays in flights
* Heavy airport volume. Heavier than average volume in travelers can cause long lines for ticketing, TSA screening, and general operations. Months with major holidays such as December will have heavier airport traveler volume.
* Late arriving aircraft. When a previous flight arrives late it has a ripple effect causing other departing flights to be delayed.
* Airport staffing issues. Malperformance or absence of TSA, ticketing staff, and air traffic controllers may have contributed to delays.
* The government shutdown started in December 2018. Due to the Christmas and New Year holidays, December has higher than average air travel volume which could contribute to the delays in this time period.
* https://www.bts.gov/topics/airlines-and-airports/understanding-reporting-causes-flight-delays-and-cancellations

# Finding The Right Data

**Bureau of Transportation Statistics** – as part of the Department of Transportation, collects and provides statistics and analysis for commercial aviation, multimodal freight activity, and transportation economics

* + Has very detailed data that can be manipulated online for different data sets, as well as the ability to download detailed data
  + One noted consequence of the shutdown is delays to statistical analysis and reporting that is vital to most business sectors
  + We were unable to use this site as the last updates were only through November 2018

**Cirium Flightstats Developer Center** – a commercial provider of global flight data through analysis and API services

* + Access to real-time and historical information used by mobile developers, airports, travel agencies, airlines and more
  + BUYER BEWARE! Although easy to obtain and digest, the API services are not all free! Historical data is an example!
  + BUT, we were able to connect to the API to collect for the specific timeframes, airports, and airline we needed, extracting the data in JSON format

**FLIGHT DELAY DATA – WHERE ART THOU?**

Cirium’s API supports retrieval of current Flight Delay data at a specified airport

* **SOURCE:** <https://developer.flightstats.com/api-docs/delayindex/v1>
* **JUPYTER:** [AirportDelays.ipynb](https://github.com/leannesevier/BAMF-Project-7/blob/master/Xtra%20Files/AirportDelays.ipynb) OUTPUT: [Delay\_Request\_MCO.txt](https://github.com/leannesevier/BAMF-Project-7/blob/master/Xtra%20Files/Delay_Request_MCO.txt)

The flight delay response json provided some interesting metrics on flight delays (including bins)

however free access is limited to 100 current flights a month, so we had to use the Historical Flight Status API

Cirium’s Historical Flight Status API supports retrieval of Historical airport status (departures) by date, airport, airline, and time period (up to 6 hour block) – but there is a usage fee so we had to limit our data requests.

* **SOURCE:** <https://developer.flightstats.com/api-docs/historical-flight-status/v3>
* **API DOC:** [https://developer.flightstats.com/api-docs/historical-flight-status/v3/historicalFlightStatusResponse#extendedOptions](https://developer.flightstats.com/api-docs/historical-flight-status/v3/historicalFlightStatusResponse)
* **JUPYTER:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/Xtra%20Files/MCO9Read.ipynb>

**OUTPUT:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/Xtra%20Files/Historic_Request_MCO_20190122_9.json>

# Data Exploration And Clean Up

## INCOMPLETE DATA!!!

Our first attempt at extracting “departureGateDelayMinutes” encountered some problems. The returned data did not always include a value for departureGateDelayMinutes, so we had to rely on ScheduledDeparture and ActualDeparture date and time to calculate the delay minutes using python.

**JUPYTER:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/FlightCounter.ipynb>

The next challenge we encountered was that for very large data returns (specifically for Atlanta Airport), the json response was incomplete. Saving the json response to a text file revealed that the aircraft equipment strings of data were omitted from the end of the file. Fortunately no flight data was missing or corrupted. But when saving the text files with the extension ‘.json’, Visual Code also red underlined file content that was not json compliant. As such, our team then went about repairing the files, cleaning out non json content, and saving each text file as a **json file:**

<https://github.com/leannesevier/BAMF-Project-7/blob/master/Historic_Request_ATL_20171222_14.json>

## ARE WE THERE YET?

The next challenge was not immediately apparent. While our team loaded each json file and quickly scanned through the hundreds of flights per day, some peculiarities arose. Whilst we included some python code to quickly calculate the delay time in minutes, a small number of flights appeared to have extremely early departures. This was actually due to evening flights delaying through to the next day. So rather than extracting and comparing the hour and minutes for the scheduled and actual flights, we commented this code out, and utilized the time and datetime modules to correctly compare the scheduled versus actual.

**JUPYTER:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/FlightCounter_v2.ipynb>

## IT’S ABOUT TIME !

With the date and time comparison corrected, we then went ahead and created a for loop for each airport, and a for loop for each date in the date range, so we now have a pandas dataframe for each Date Range (A and B).

**JUPYTER:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/FlightCounter_DateRangeA.ipynb>

**JUPYTER:** <https://github.com/leannesevier/BAMF-Project-7/blob/master/FlightCounter_DateRangeB.ipynb>

# Analysis

## FIRST FACTS

There is a subtle increase in the number of flights comparing Date Range B to Date Range A. According to publications on Atlanta and Orlando airports, there has been a gradual increase of passengers over this period in the order of about 3%, and about 1.6% of flights. In the same reports, Delta Airlines is identified as constituting approximately 70% of all flights.

**SOURCE:** <https://www.atl.com/wp-content/uploads/2018/12/ATL-Traffic-Report-Nov-2018.pdf>

**SOURCE:** <https://www.orlandoairports.net/press/2018/09/18/passenger-traffic-numbers-continued-record-pace-in-july-at-orlando-international-airport/>

There is approximately a 7% increase in the number of flights delayed between the two date ranges, but insufficient information to suggest that this is due to the government shutdown. This is complex because there are several variables which may or may not contribute to this increase. Publications list several factors that may be contributors.

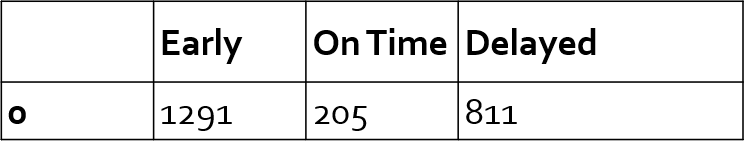
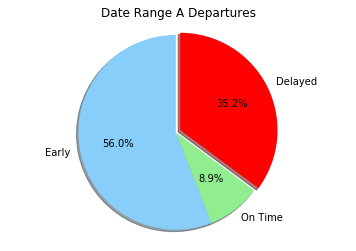
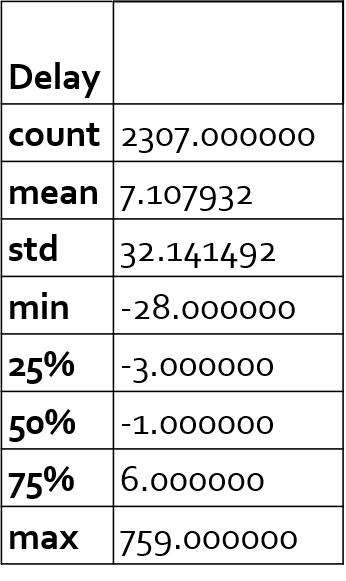
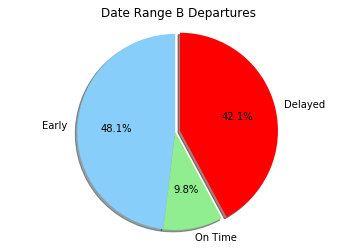
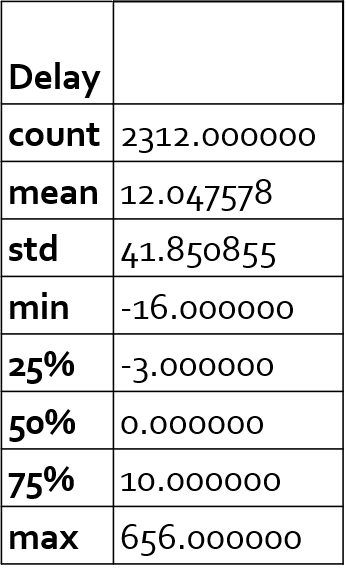


Figure 1: Date Range A Departures



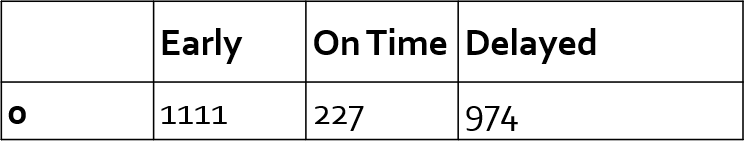


Figure 2: Date Range B Departures

There is however, a significant increase in the number of flights delayed when comparing the final weeks of both date ranges. In fact, the final week of the government shutdown period exhibits double the number of delayed flights compared with the same week of the previous year. This definitely requires deeper investigation.

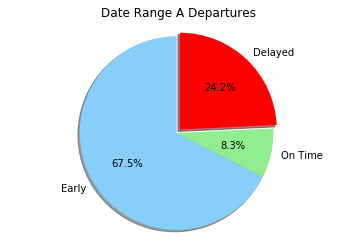
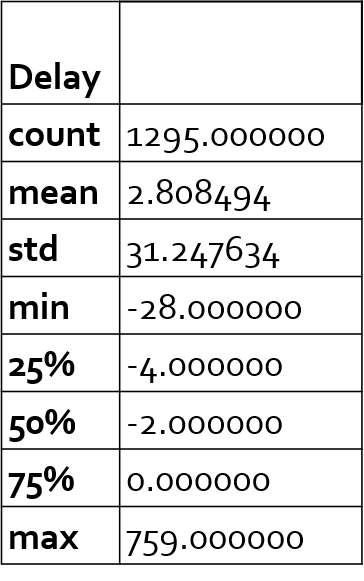




Figure 3: Date Range A Final Week Departures

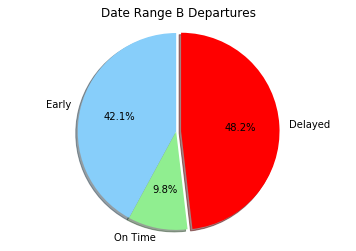
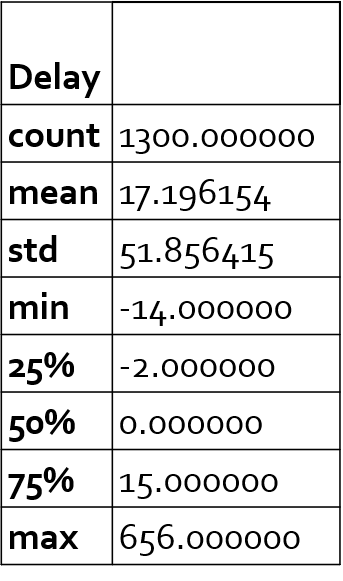




Figure 4: Date Range B Final Week Departures

# Findings And Conclusions

## OBSERVATIONS

In the date range A, MCO airport has handled less than 8.5 % of the Delta flights comparing with the numbers of Delta flights in ATL airport. Even that MCO airport shows a higher proportional number of delayed departure flights.

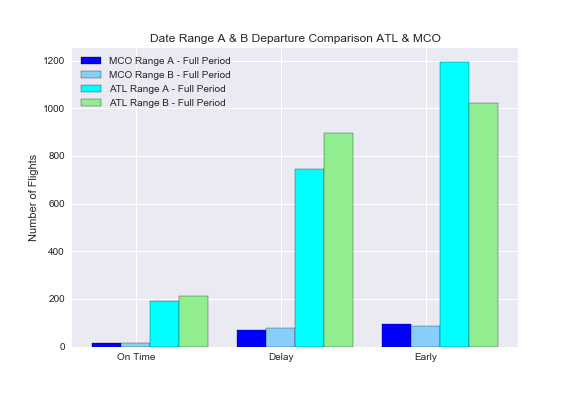


Figure 5: Date Range A & B Departure Comparison by Airport

In the shutdown period (range B) this difference has increased, showing that even operating a much higher number of flights, Atlanta airport has had a smaller impact with the shutdown than Orlando airport, when it comes to flight departures delays.

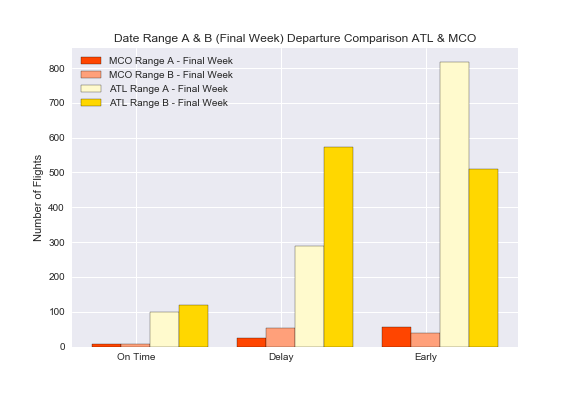


Figure 6: Date Range A & B (Final Week) Departure Comparison by Airport

In the last week of the shutdown the number of delayed departure flights has increased, representing over half of total delays for full shutdown period. (MCO full period 79, 52 in the last week – ATL full period 895, 574 in the last week) The increase might be related with TSA sickouts reaching 10% on Jan 20, 2019.

https://www.forbes.com/sites/michaelgoldstein/2019/01/22/us-airlines-take-stock-market-hit-as-government-shutdown-reaches-32-days/#68ff706a79de

Comparing the numbers of flights in each date range per airport, we can conclude that MCO airport has being more affected by the shutdown in number of flights departure delays, than Atlanta airport, however, in the subject “delays” historically MCO airport has been behind ATL airport.

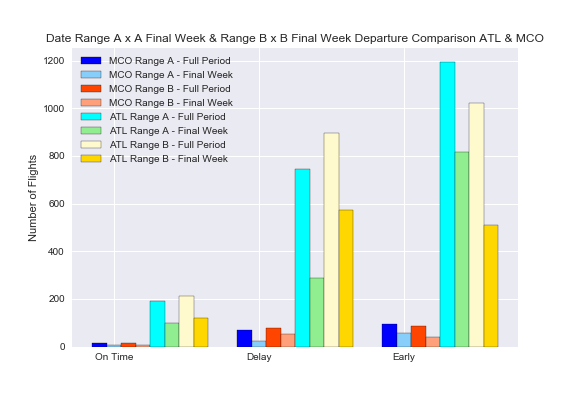


Figure 7: Final Week Departure Status Comparison by Airport

The MCO average delays for customs/security is 24,45 min. against 12,27 in ATL airport. The average departures delays in MCO is 68 min. against 58 in ATL.

Atlanta Airport has a Delay Index of B – (401,78) and MCO Delay Index is D (517,80).

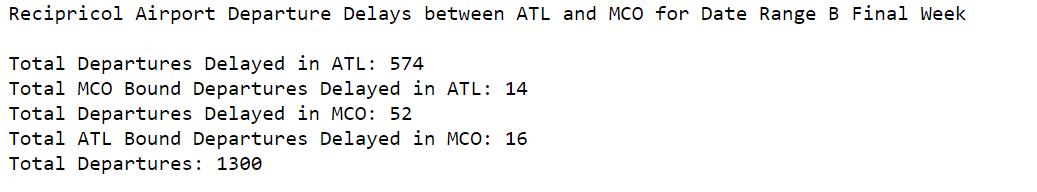
Based on these numbers we can assume that ATL airport was better prepared to deal with the shutdown than MCO airport.

https://www.orlandoweekly.com/Blogs/archives/2018/11/21/orlando-international-airport-ranked-among-most-likely-to-delay-your-flight

## AIRPORT TO AIRPORT IMPACT ANALYSIS

**What evidence is there that delays at one airport cause delays at another airport?**

Based on the data we collected, we decided to examine any possible causality of delay from MCO to ATL and vice versa.



According to our data queries, only 4.8% (14+16) /( 574+52) of delayed flights out of ATL and MCO were bound for the other airport. This implies that on average, delays incurred at one airport do not incur significant delays at the other airports. But each airport has a different volume of airline flights originating from other airports (both domestic and international).

While Atlanta only received a comparative 2.8% of its delayed flights from MCO, Orlando received a comparative 27% of its delayed flights from ATL.

**So what does the data reveal for each airport?**

## ORLANDO AIRPORT

## 

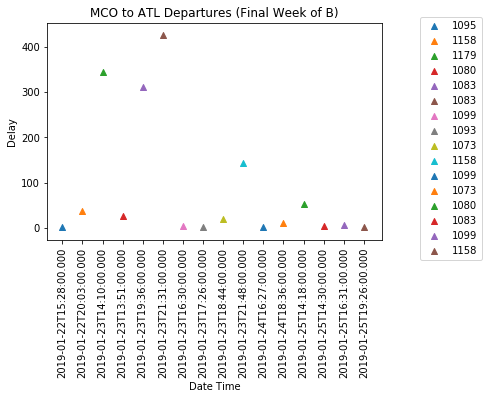


Figure 8: Flight Number by Delay (minutes) Scatter Plot for

Delayed MCO Flights to ATL by Final Week Day

The most significant ATL bound flight delays occurred on January 23rd where all but two of the eight delayed flights greatly exceeded mean delay durations of 12 (Range B) and 17 minutes (Range B Final Week).

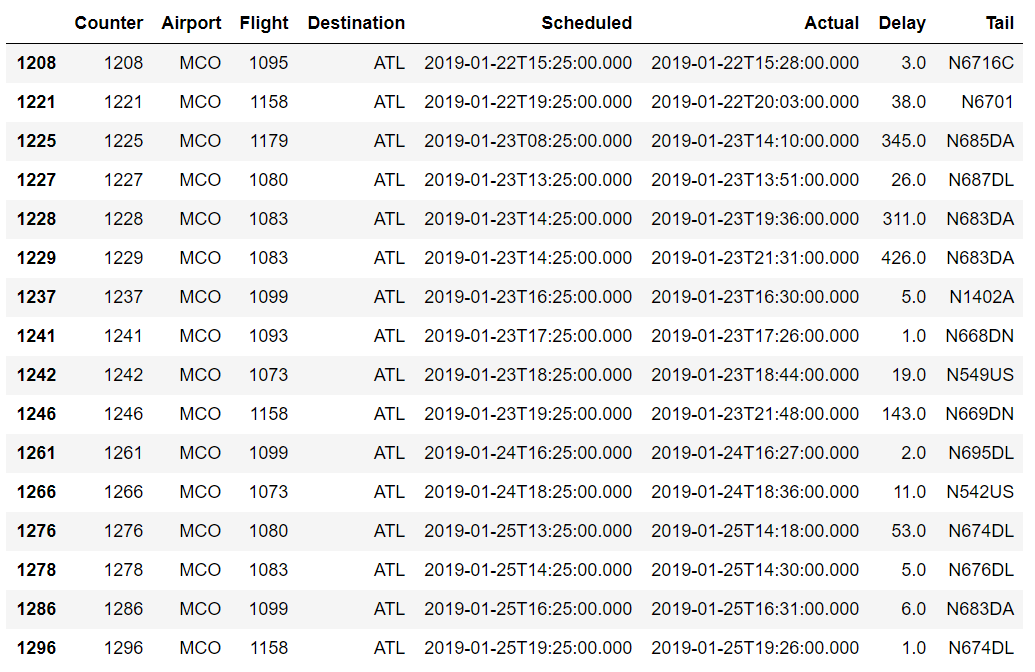


Figure 9: Delayed MCO Flights to ATL by Final Week Day Data Table

Flight 1083 / N683DA appeared to have particular difficulties with accumulative delays.

## ATLANTA AIRPORT

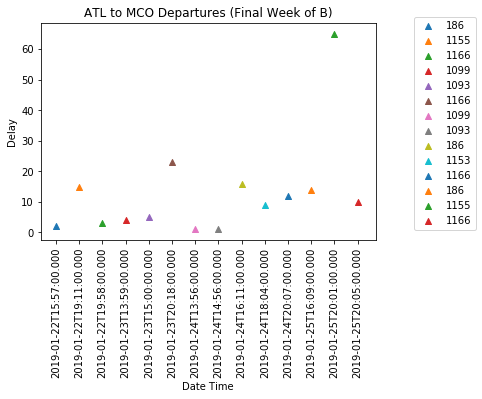


Figure 10: Flight Number by Delay (minutes) Scatter Plot for

Delayed ATL Flights to MCO by Final Week Day

Considering ATL is a major hub (particularly for Delta Airlines), the mean delay for MCO bound aircraft is less than 15 minutes, which almost all fall within typical delay durations. While the 65-minute delay is outside our 8pm limit.

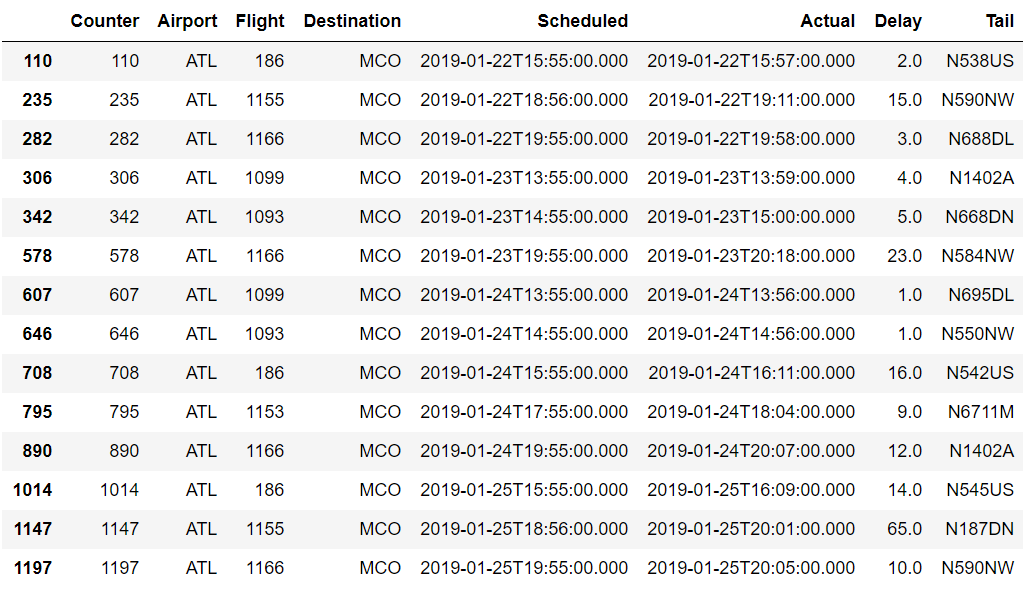


Figure 10: Delayed ATL Flights to MCO by Final Week Day Data Table

Based on the unique tail numbers, it appears that Delta utilizes a reserve of aircraft at ATL to operationally optimize for earliest possible departures.

# Summary

1. Watch Out! Data Analysts Beware! – Some API vendors charge by transaction (10 cents per flight)
2. Government shutdown effects on airports and airlines is gradual. There was no significant disruption to the number of flights or passengers, and delay durations in the first week of shutdown was comparable to the non-government shutdown period.
3. Delta Airlines appears to be able to operationally optimize aircraft at the Major Hub of Atlanta to reduce departures delay times. They are less dependent on specific aircraft making timely connections. But this does not imply that all passengers are always able to make their connections through Atlanta Airport.
4. Delayed flights bound for Atlanta from Orlando experience significantly larger delays compared with connections in Atlanta. This might be due to lower availability of air traffic controllers at MCO compared with ATL. It may also be due to the volume of passengers originating in Orlando, unable to board timely, due to extensive security checkpoint delays.
5. If the government elects to privatize air traffic controllers and airport security and customs officers, this might possibly alleviate some of the airport-based flight delays. But there is insufficient data available to directly correlate all of these factors.
6. If I do need to travel during a government shutdown in the future, I would plan to book a morning flight from Orlando, with enough time to make an afternoon connection in Atlanta. I would be less concerned with my connection being delayed in Atlanta.

# Works Cited

“Q&A: Everything You Should Know About Government Shutdowns.” Committee for a Responsible Federal Budget, Committee for a Responsible Federal Budget, 27 Feb. 2019, www.crfb.org/papers/qa-everything-you-should-know-about-government-shutdowns.

Shrikant, Aditi. “Everything You Need to Know about Airports during the Government Shutdown.” *Vox.com*, Vox Media, 17 Jan. 2019, [www.vox.com/the-goods/2019/1/16/18183634/tsa-government-shutdown-travel-flying](http://www.vox.com/the-goods/2019/1/16/18183634/tsa-government-shutdown-travel-flying).

<https://www.ajc.com/news/world/will-airports-affected-possible-partial-government-shutdown/wQRRXS5s5By3daQFnz12bN/>

<https://www.cnbc.com/2019/01/16/united-ceo-worried-about-shutdown-but-impact-not-yet-significant.html>