Weather-Related Delay Analysis at Harstfield Jackson International Airport

Samet Ayhan, Computer Science Department, UMD – sayhan@umd.edu
Scott Abromowitz, College of Information Studies, UMD – abromo@umd.edu
October 5, 2012

Introduction

The National Airspace System (NAS) is extremely efficient, yet the NAS is heavily loaded and often running close to capacity, as even small events can cause delays in the NAS that cascade throughout the system. Large events, mainly weather-related, can cause significant delays and inefficiencies. With these facts in mind, we decided to analyze weather-related delays at the Hartsfield Jackson International Airport (KATL) in Atlanta, Georgia because KATL is the world's busiest airport [1] and is one of the top five worsts weather affected airports in the United States. [2]

Data Sets

Our analysis uses two historical data sets:

- A severe-weather day (low cloud, poor visibility) – November 21, 2011
- A good-weather day –August 21, 2012

Our data sources are

- Aircraft Situation Display to Industry (ASDI) [3] with flight data messages and surveillance reports of 20 seconds frequency and
- Meteorological Aerodrome Reports (METAR) [4], an hourly surface aviation observation in a simple code that supplies the weather state at an airport.

The Bureau of Transportation Statistics (BTS) collects a wide range of monthly data from US-certified air carriers that report on-time arrival and departure data for non-stop domestic flights. Since the source of data comes from the airlines themselves, we decided to use a more objective data source, ASDI, because airlines typically use flexible padding times to refrain from reporting long delays. Nevertheless, one of the difficulties of sensor data is sensor data's massive size because the Federal Aviation Administration (FAA) distributes over seven million ASDI messages per day. In order to utilize only necessary pieces in this data, an extensive amount of data cleaning and transformation was performed. Some of the attributes included in this subset of ASDI data are flight date, flight time, aircraft type, equipment, departure airport, arrival airport, estimated arrival time, estimated departure time, actual arrival time,

actual departure time, aircraft id, air carrier, arrival state, departure state, delay, and ICAO code.

Equally, weather data included in this analysis contain meteorological attributes such as temperature, wind direction, wind speed, visibility, cloud type, cloud altitude, and ceiling.

Headlines

If You Can't See, You Can't Fly!

Key features, Insights and Significant Findings

We used Spotfire's scatter plot option to analyze the correlation between arrival delay and adverse weather condition (visibility) throughout the day (November 21, 2011). With trellis option, we were able to stack up two separate days to identify the impact of severe weather shown in Figure 1. The lighter circles represent poor visibility (0.25 mile) where the delays spike at around 3:00PM on an adverse-weather day. The bottom scatter plot shows no indication of weather related-delays because visibility is good (10+miles) throughout the day (August 21, 2012). Another insight indicated by these scatter plots is that less delays occur at early morning hours likely since no delays have cascaded yet.

Visualizations

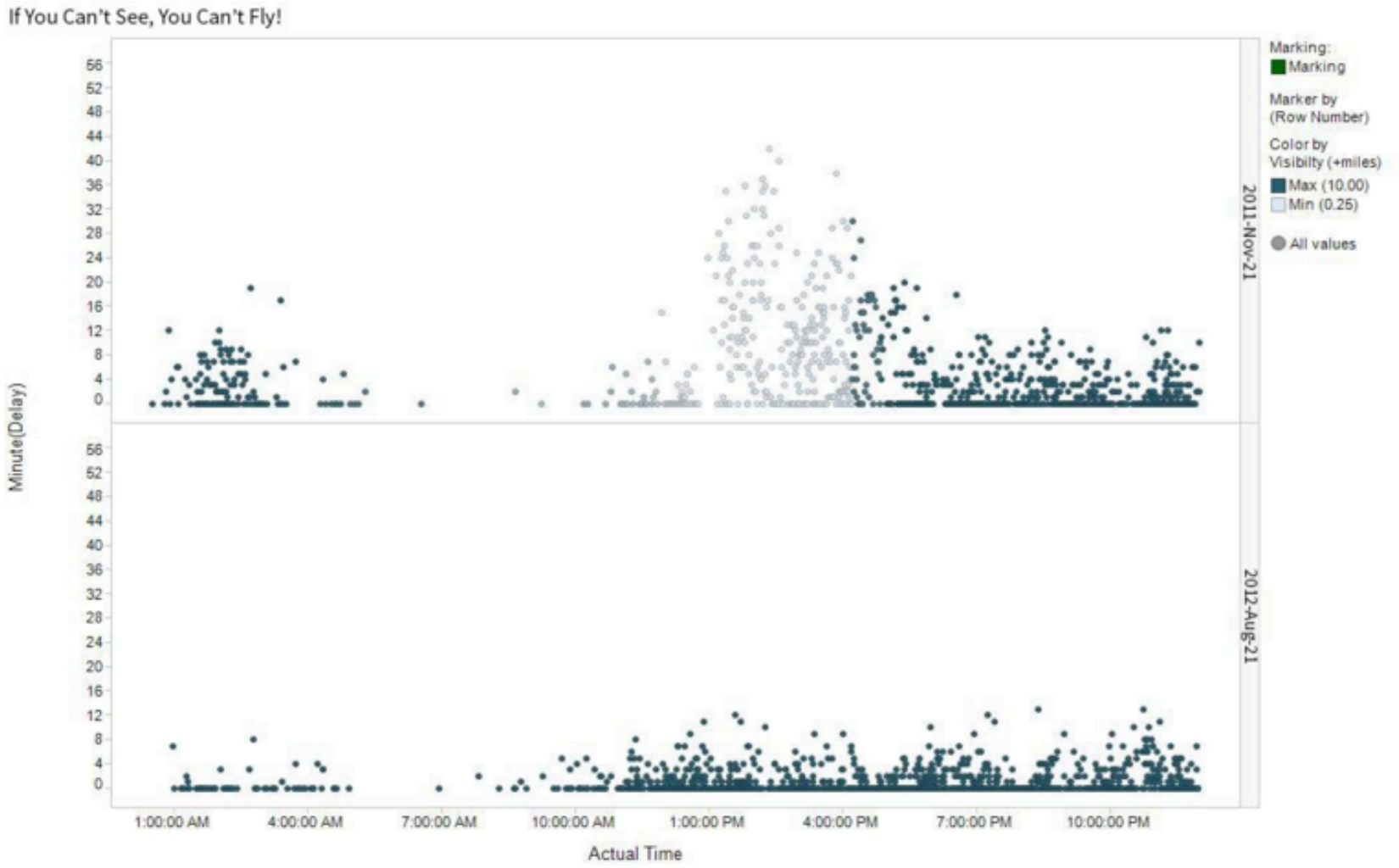

Figure -1: Weather Impact on Delays throughout the Day

The More Equipped the Aircraft the Less Likely It Will Delay!

Key features, Insights and Significant Findings

In the Figure 2 treemap, color represents the number of delays and rectangle size depicts the number of aircrafts in that specific type. We use treemap for this visualization because there is a hierarchy with aircraft type grouped based on the type of equipment on board. This chart basically reveals that most of the aircrafts departing from or arriving to KATL at these days have equipment capability Q, which means that such aircraft meet Required Navigational Performance (RNP) along with Reduced Vertical Separation Minimum (RVSM). This is considered the most equipped capability compared to other capabilities of L, W, and G; however, most delays occur with aircrafts that are equipped with type L, which indicates that they have Global Navigational Satellite System (GNSS) along with RVSM. As it is apparent, the least amount of delays occur with aircrafts that are equipped with capability G, which is the least equipped one among others. However, there are only total of 8 operations with this equipment type that experienced delays. Per these facts, we can infer that the more equipped an aircraft is the more likely it will experience fewer delays.

The full list of equipment types is available [5]. The full list of aircraft types is available [6].

Visualizations

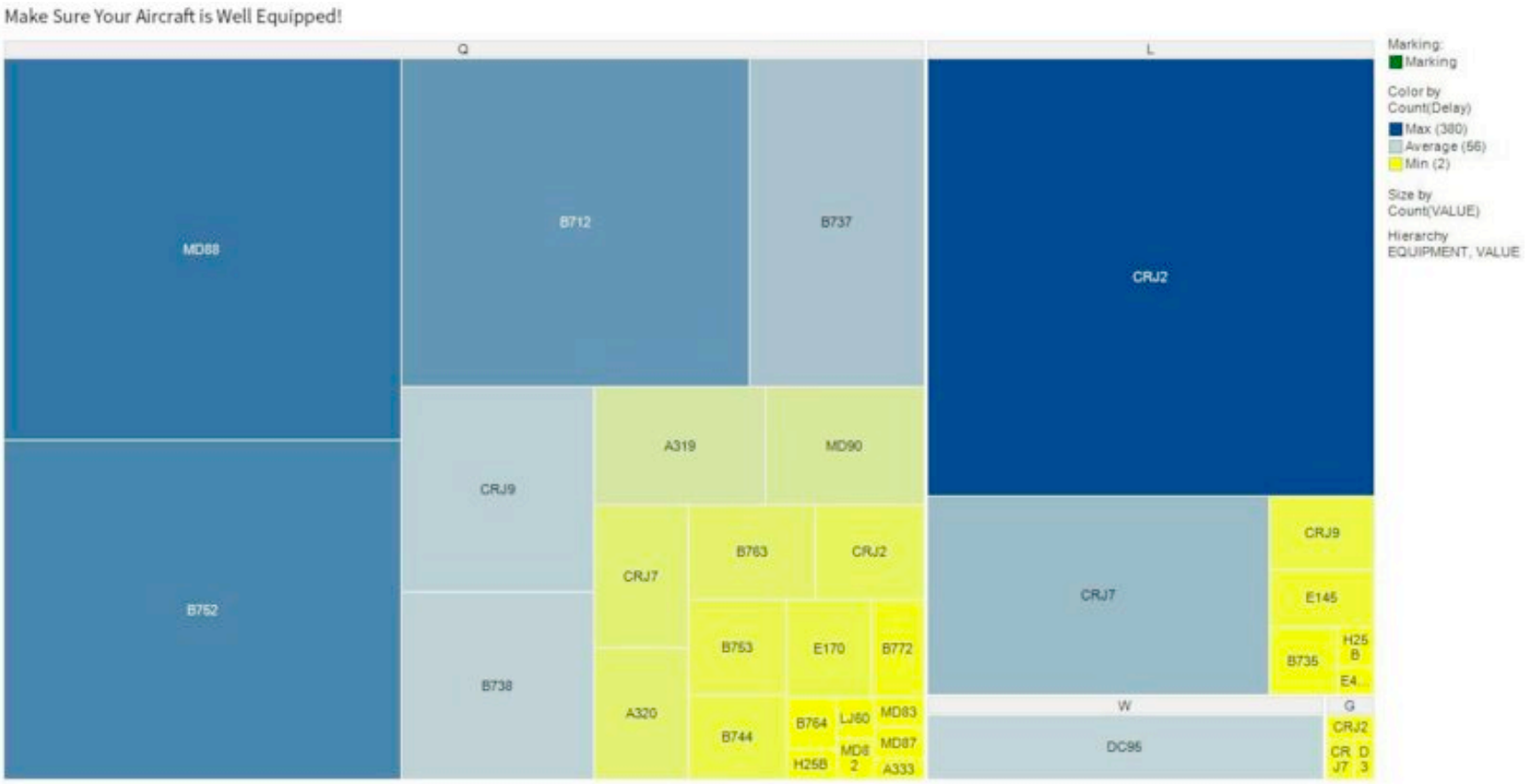

Figure -2: Correlation between Equipment on Board and Delays

They All Delay When the Weather Goes Bad!

Key features, Insights and Significant Findings

In Figure 3, the blue bars represent delay percentages during adverse weather conditions (November 21, 2011) and yellow bars represent delay percentages during good weather day (August 21, 2012). This bar chart is an illustration of the top six airlines with their highest percentage of delays in terms of departures and arrivals at KATL. This chart is a clear indicator of increased delay percentages for each airline because delay percentages during the good weather day (August 21, 2012) are always below delay percentages during the severe weather day (November 21, 2011). This chart is generated using Spotfire and Figure 4 is generated using Tableau utilizing the same dataset. We intended to use the same color-coding so that both tools' visualizations can be compared one to one showing the same visualization.

Visualizations

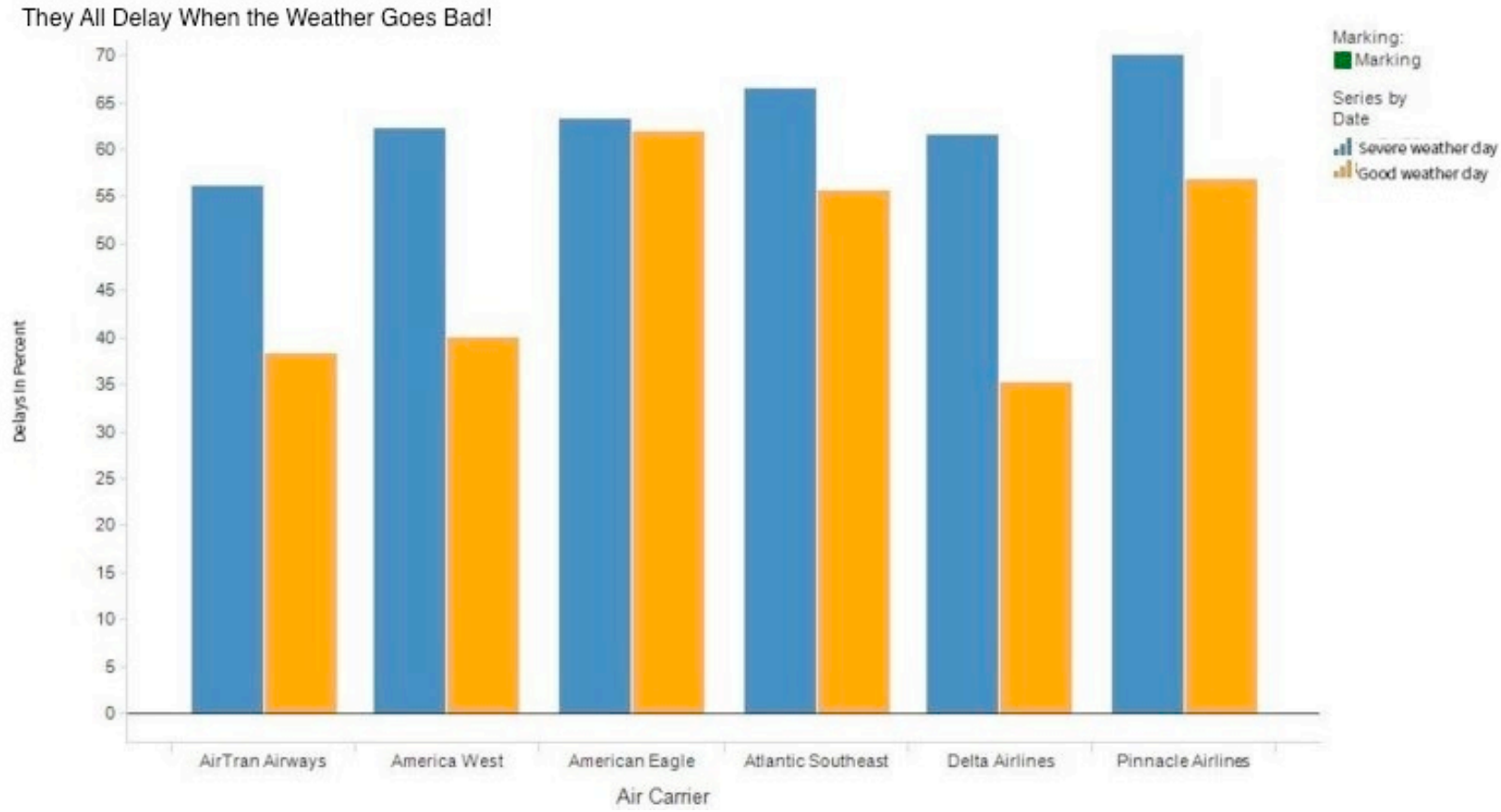

Figure 3 - Delay percentages per Airline (Spotfire)

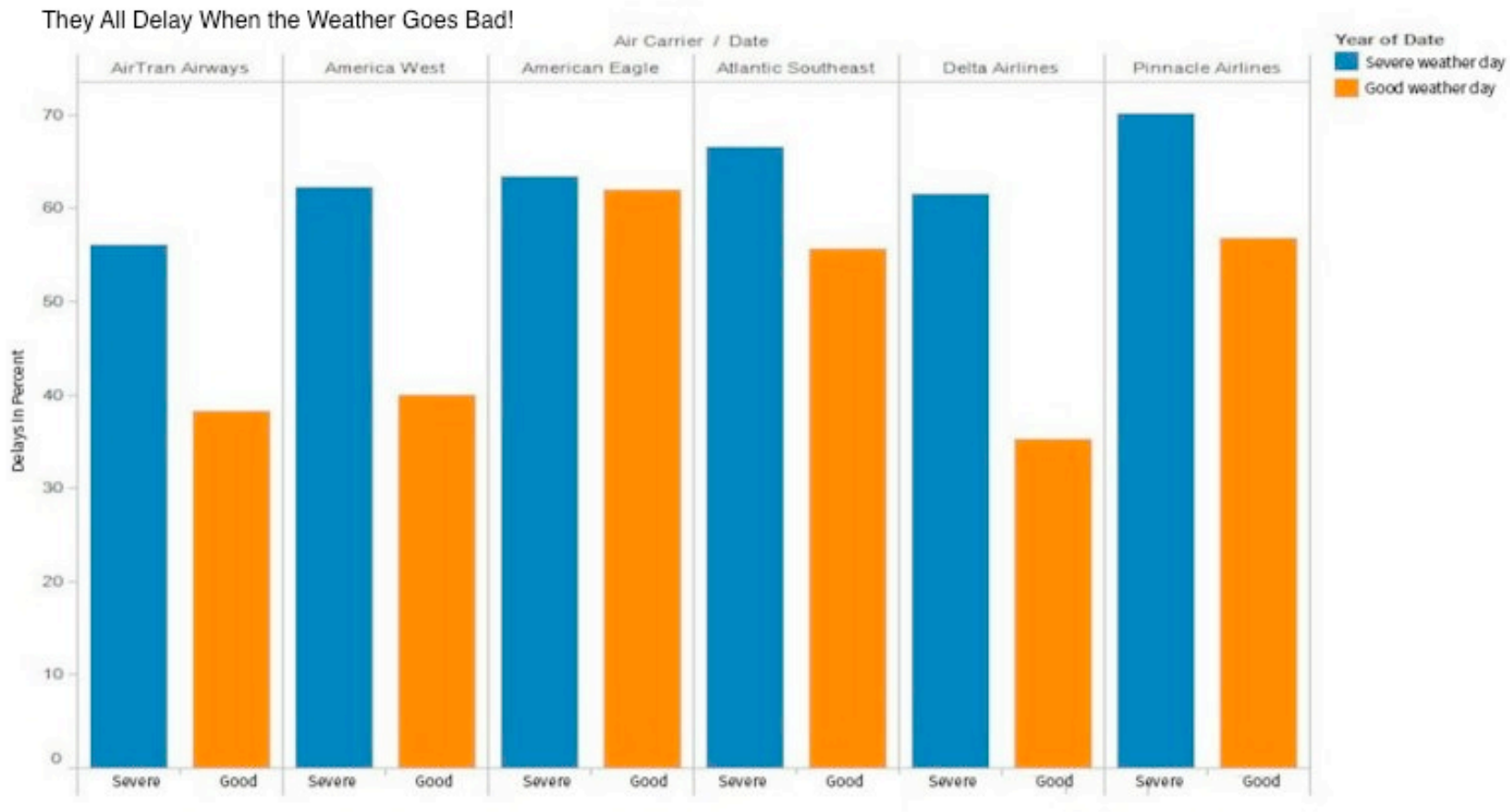

Figure 4 -Delay percentages per Airline (Tableau)

Additionally, we wanted to explore delays per arrival and departure state connecting to KATL. In this chart, Florida stands out with the maximum number of delays for both arrival and departure flights during good (August 21, 2012) and adverse weather conditions (November 21, 2011) partially due to the fact that Florida has the maximum number of connecting flights with KATL. In this trellis view, the top bar chart in Figure 5 represents the number of delays per departure state. Red bars illustrate delays in severe weather

conditions. Blue bars represent delays in good weather conditions per departure state. The bottom bar chart is illustrates departures. Given these charts, it is apparent that the number of delays per state remains relatively the same whether during either severe or good weather conditions. These visualizations also demonstrate correlation between arrival and departure delays per state.

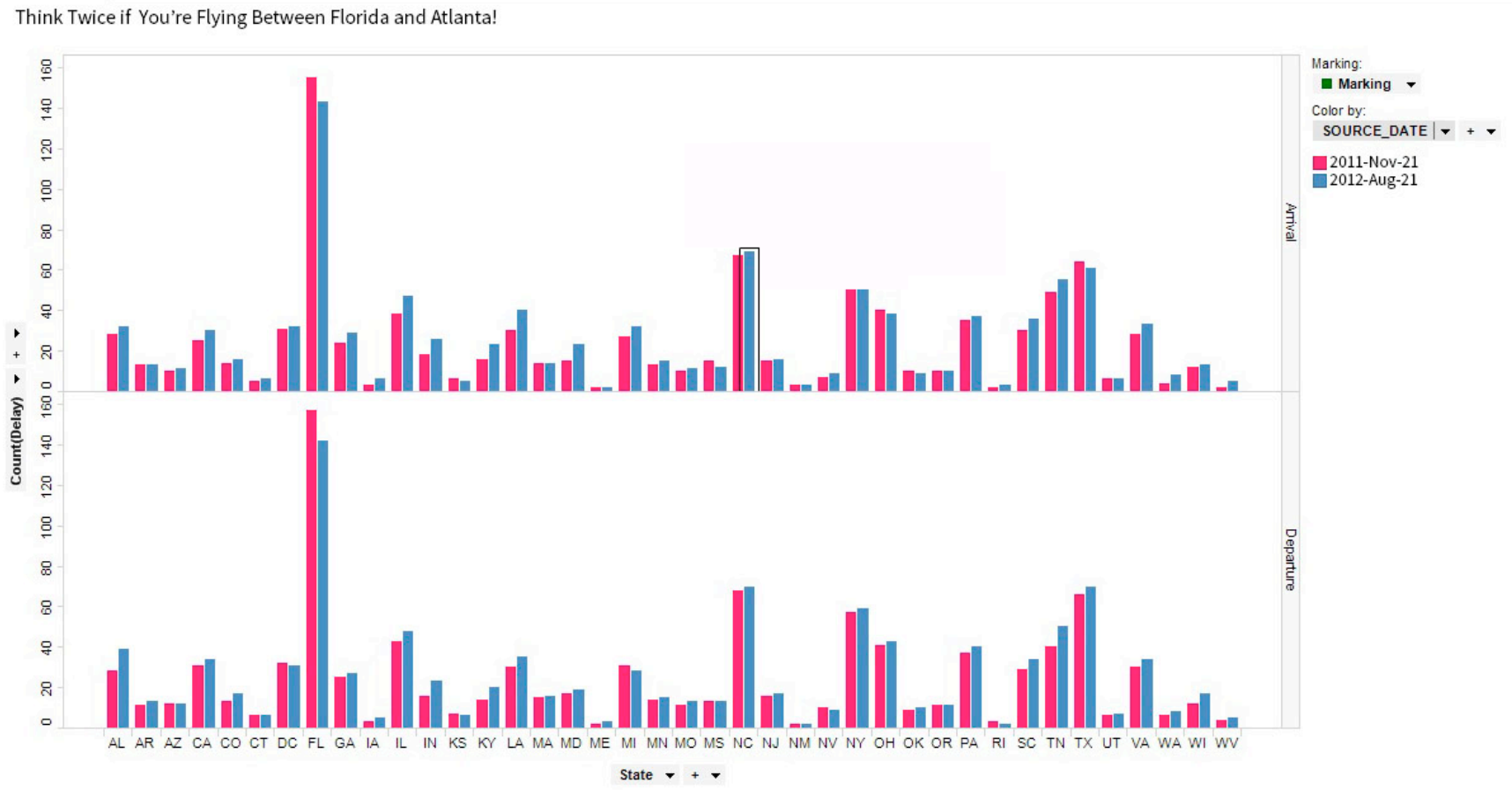

Figure 5 - Number of delays per State

Critique

Positives

While exploring each application's abilities, Spotfire was consistently more intuitive at performing simple to routine visualization tasks. Selecting a data set range (i.e., delay time range) is also easier in Spotfire because Spotfire's filter fields allow for faster and more direct data manipulation versus Tableau requiring a user to right click on the data and then create a parameter. Spotfire too suggests better insight than Tableau by automatically detecting an insight when importing data and selecting a visualization type.

Problems

Tableau's presentation of data in the form of dimensions and measures did not make sense until reading Tableau's help documents while Spotfire simply presented the data as one long list. While Spotfire often requires fewer actions to perform tasks, one feature that Spotfire lacks, but Tableau can perform is the ability to rename titles. In Spotfire users must export the visualizations and then import the image into a photo-editing program to rename the titles.

Suggestions

If we want to incorporate different visualization types, both tools, consequently, cause random changes to the axes when selecting different visualization types. It would be more consistent if the axes stay the same, whenever possible, when switching between visualization types. For color-deficient users, Tableau includes the ability to select a predefined palate, but in Spotfire users must manually select such colors. As such, Spotfire should add this function.

Conclusion

Given the above visualizations, we can conclude that severe weather has negative affect on flight delays regardless of the airline. As Figure 2 verifies, one way of mitigating this impact is equipping the aircraft with high-level capability such as Q.

References

1. http://en.wikipedia.org/wiki/Hartsfield%E2%80%93Jackson_Atlanta_International_Airport
2. http://www.11alive.com/news/article/214925/3/Atlanta-among-5-worst-weather-airports-in-US
3. http://www.fly.faa.gov/ASDI/asdi.html
4. http://aviationweather.gov/adds/metars/
5. http://www.gofir.com/general/rvsm/aircraft_equipment_suffix.htm
6. http://www.fly.faa.gov/ASDI/asdidocs/aircraft_types.txt