USA'S AIRPORT FLIGHTS ANALYSIS FOCUSED ON CHICAGO O'HARE INTERNATIONAL AIRPORT (ORD) AND DALLAS LOVE FIELD AIRPORT (DAL)

Andreadis Georgios

Gkaveras Dimitrios

Moscholios Filippos-Michael
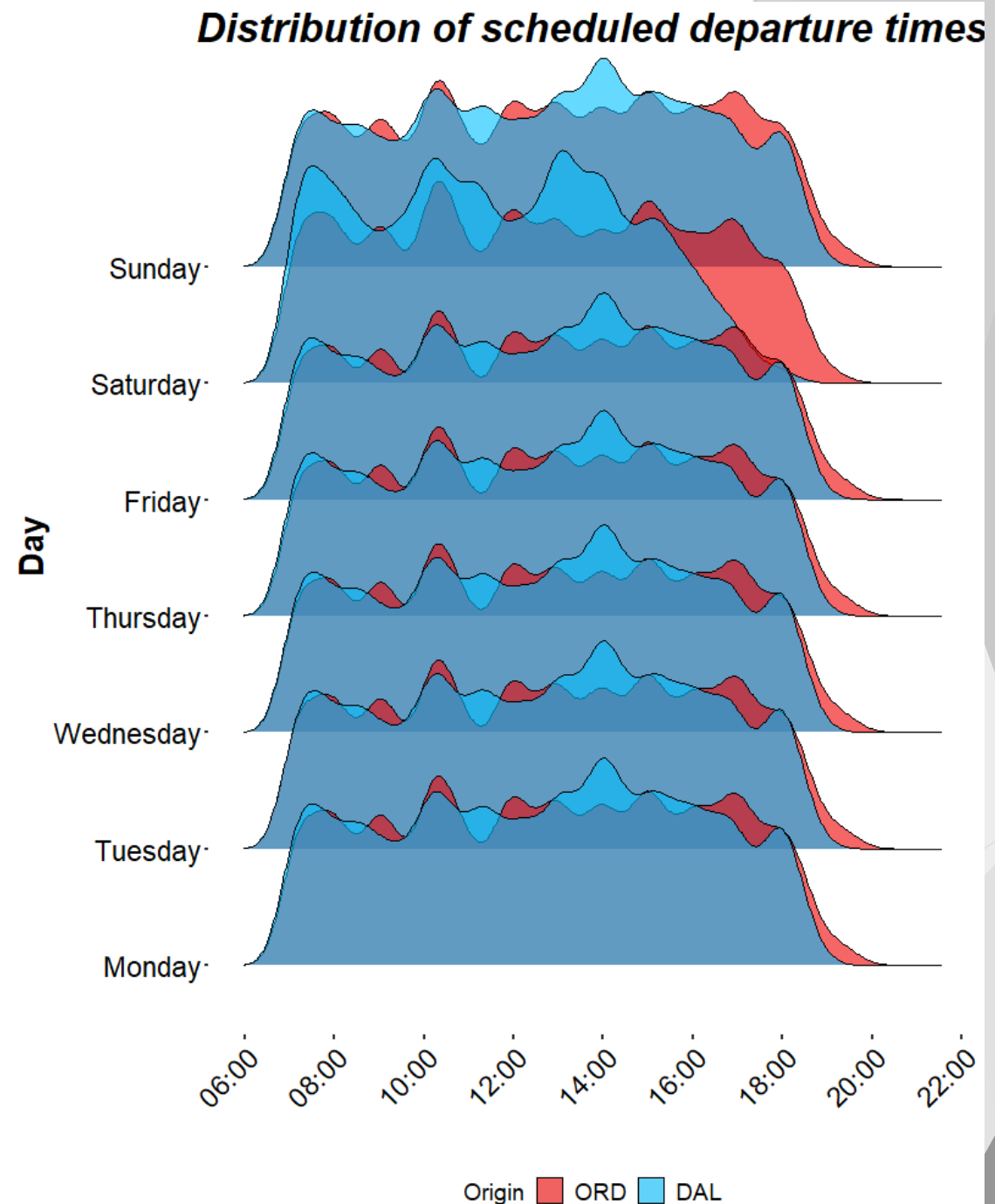
Tsakiri Dimitra

Introduction: A view of the data

The data consists of flight arrival and departure details for all commercial flights within the USA, from January 2004 to December 2007. Our target airports are Chicago O'Hare International Airport (ORD) and Dallas Love Field Airport (DAL), the former being a large Airport and the latter being a medium Airport according to Wikipedia's rank for busiest US Airports by total passenger boardings. The goal of this analysis is to compare the aforementioned airports and highlight their positive and negative elements.

Evolution of destinations from Chicago-ORD and Dallas-DAL and the arrival time behavior.
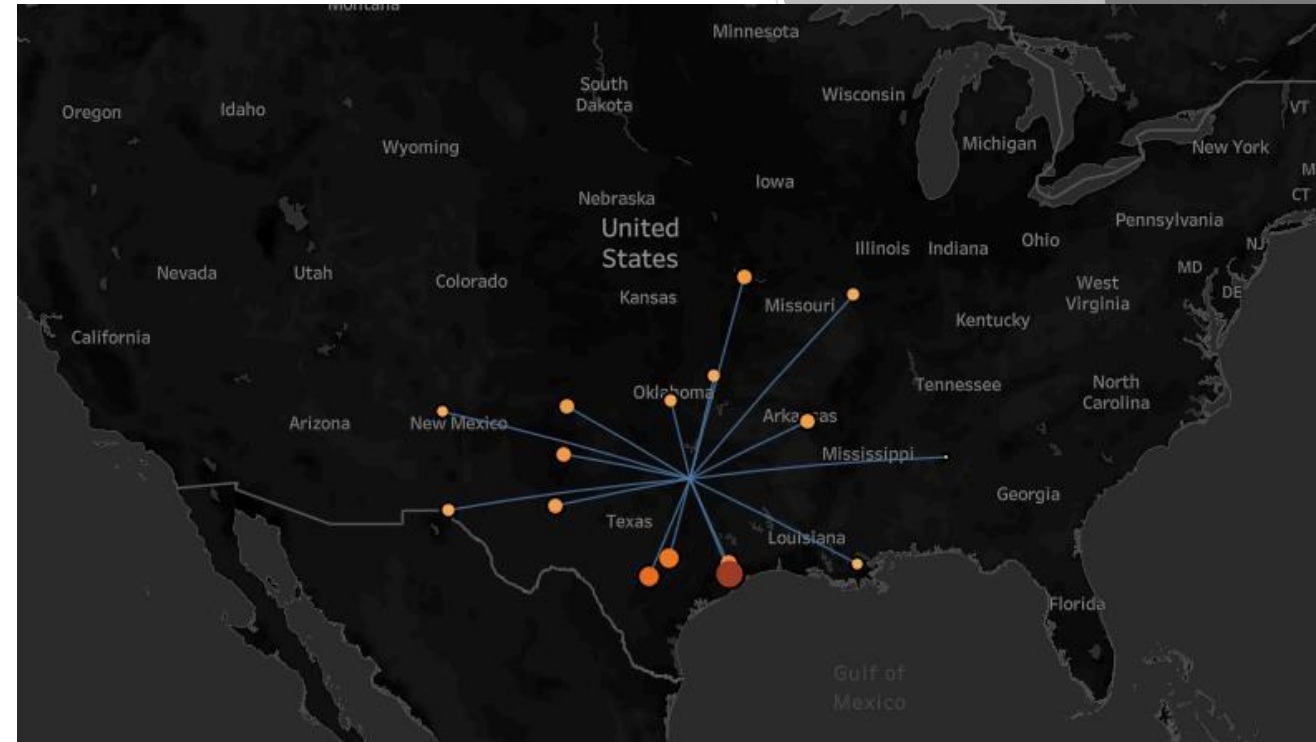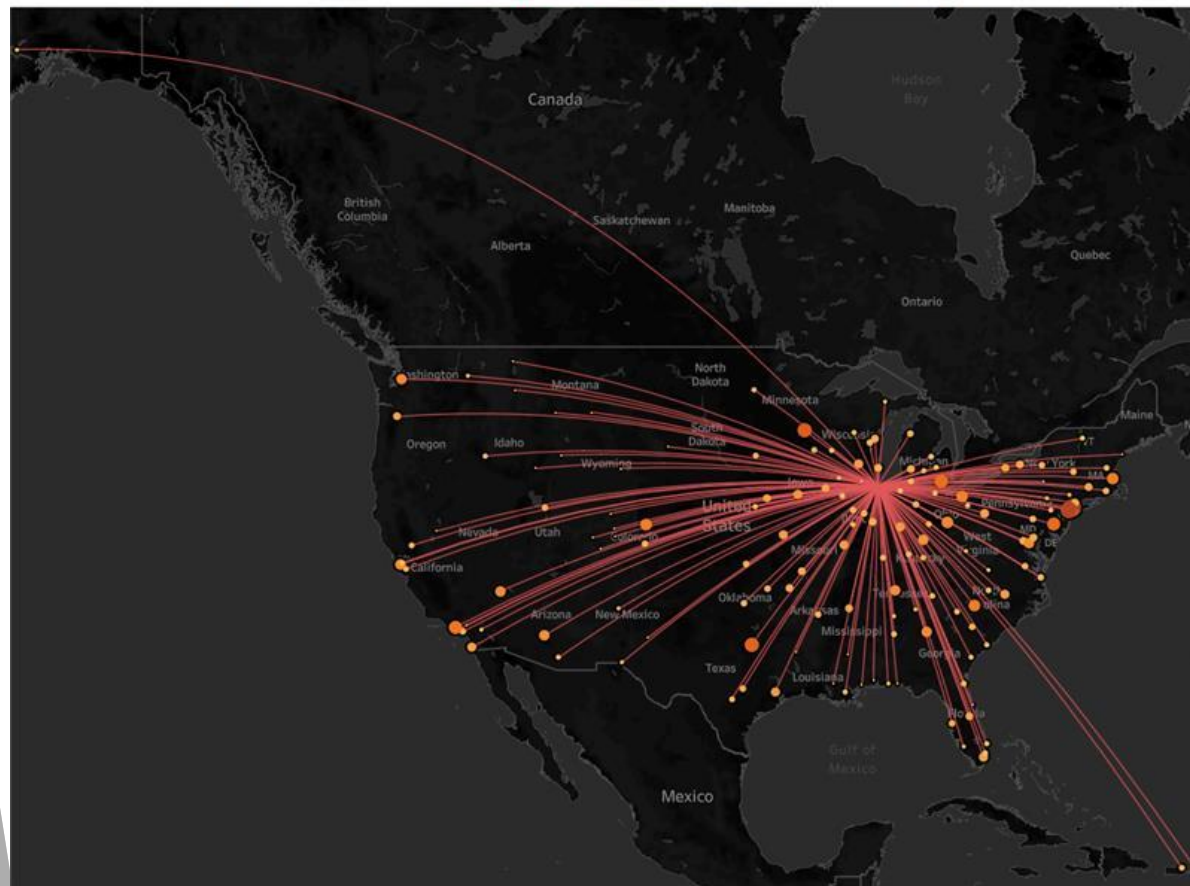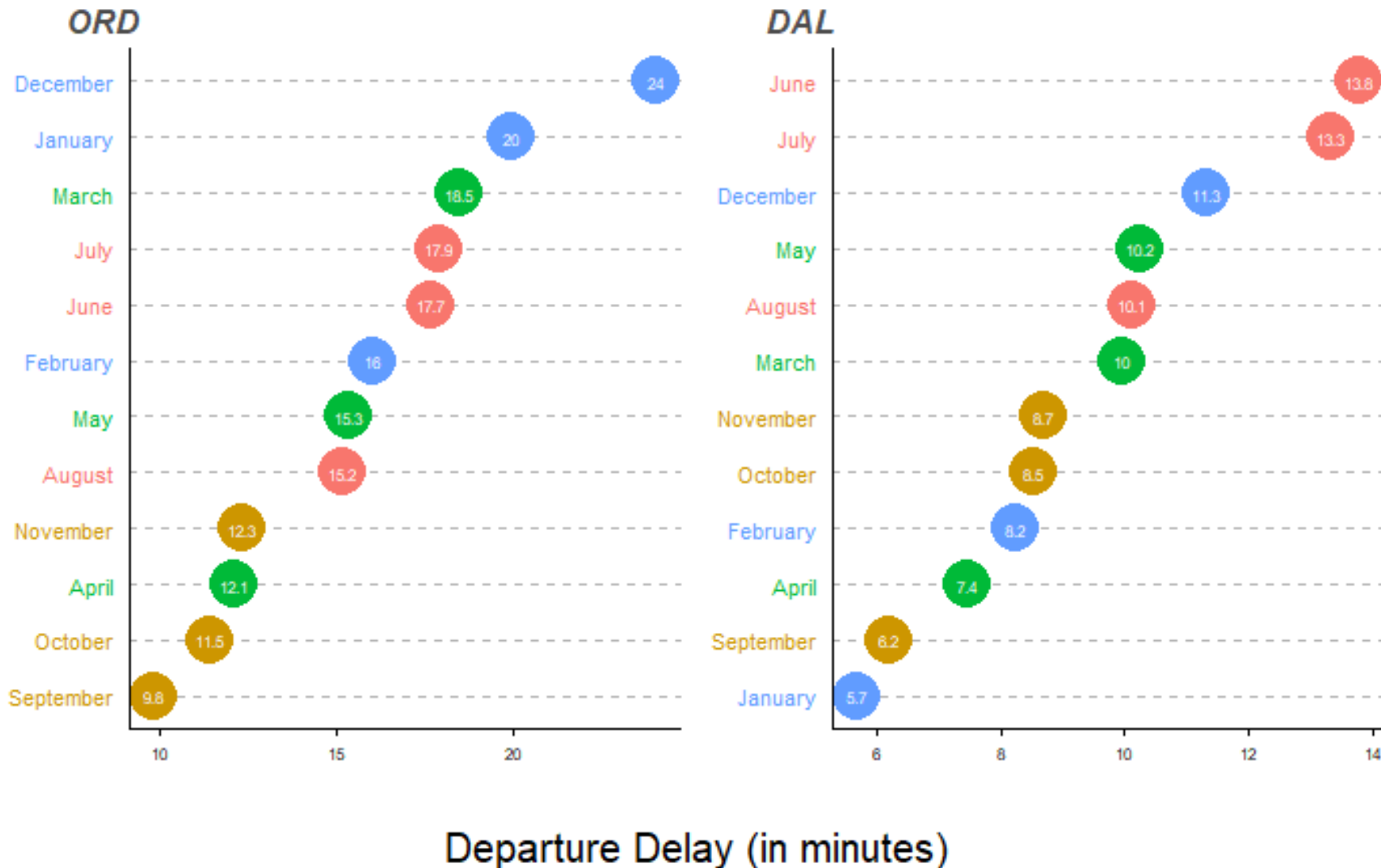
Here we can see that O'Hare Airport has a lot more flights than the Dallas one, hence the observed arrival delay is also larger regarding the colour and the size of the destination points. This plot is interactive and is made using Tableau. The user can choose the year and the origin Airport he wants to be displayed as well as filter the flights based on their distance from the origin Airport.

In this graph we can observe the distribution of scheduled departure times for each day of the week and each airport. We can say that the distribution for O'Hare Airport is similar throughout the week and relatively uniform throughout the day. This is not the case with Dallas Airport, which has a relatively stable schedule for the weekdays, but differs on the weekend where we can see peaks.
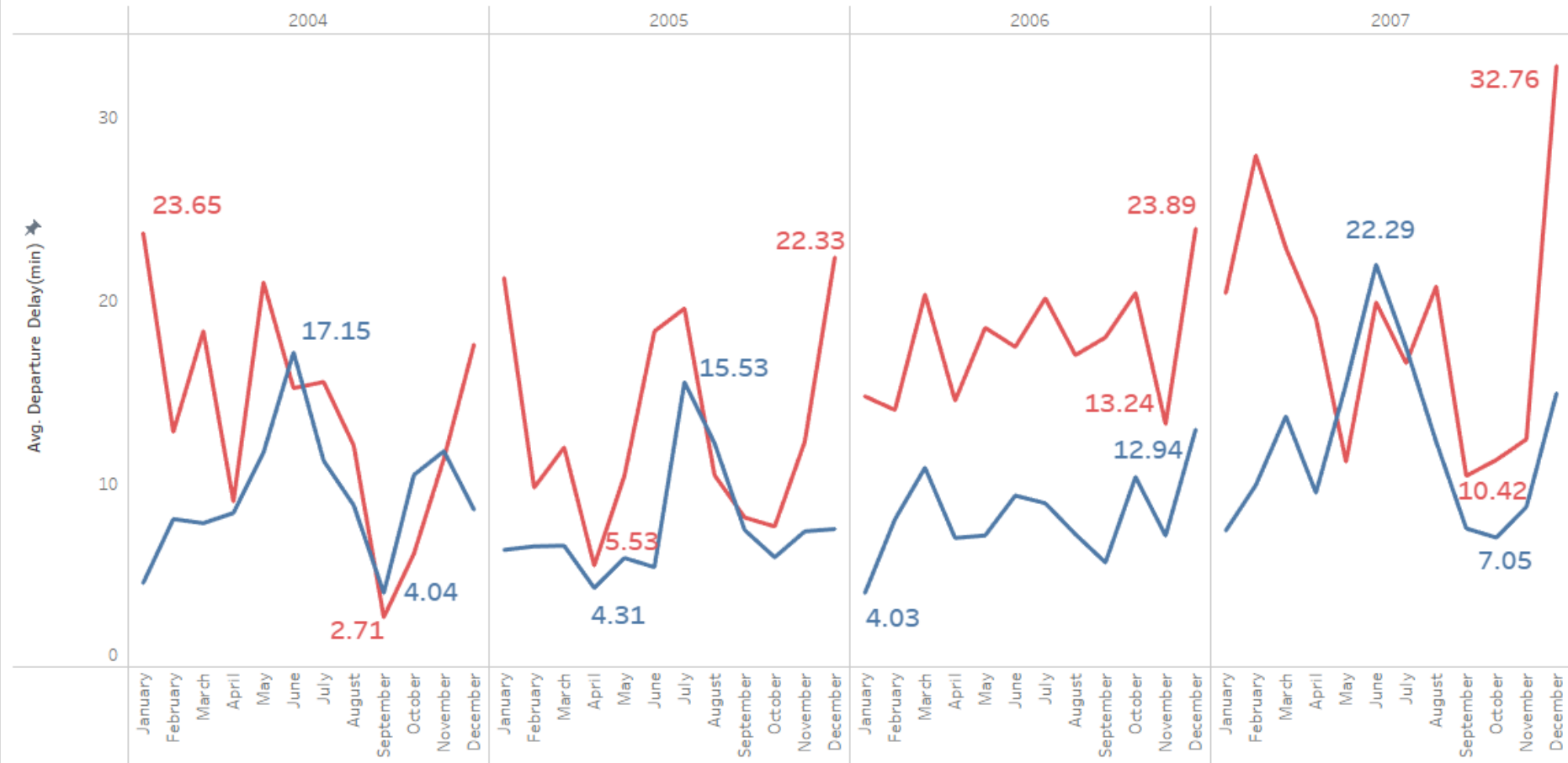

Distribution of scheduled departure times

Averaged Departure Delay

ORD

Month	Departure Delay
December	24
January	20
March	18.5
July	17.9
June	17.7
February	16
May	15.3
August	15.2
November	12.3
April	12.1
October	11.5
September	9.8

DAL

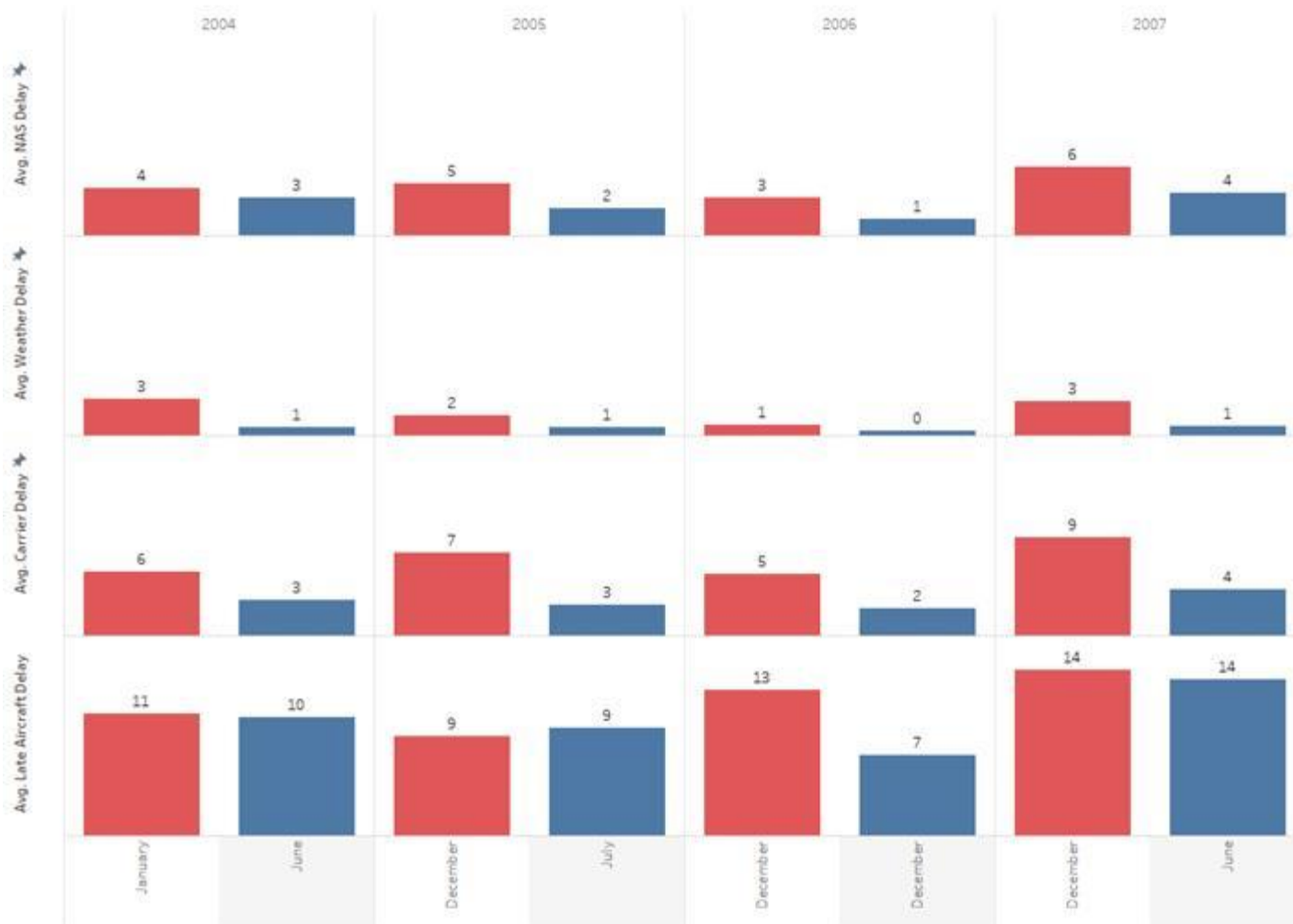| Month	Departure Delay
June	13.8
July	13.3
December	11.3
May	10.2
August	10.1
March	10
November	8.7
October	8.5
February	8.2
April	7.4
September	6.2
January	5.7

Departure Delay (in minutes)

In this plot, the average departure delay of the flights for the two target airports are easily comparable. Generally, the average delay in minutes for the years 2004-2007 is smaller per month for Dallas airport. This is expected, because, as we mentioned earlier, Dallas airport is smaller, thus it is easier to be more punctual than a larger airport as Chicago's, which has more passengers and long-distance flights. Also, for both airports the two most delayed months, are months that people usually travel for vacations, for Dallas being two summer months and for Chicago being two of the winter months.

In this graph we observe the departure delay per month and year and examine the overall behavior of each airport. We can see an increasing trend from October to January for O'Hare airport. On the other hand, Dallas airport departure delay tend to increase from April to July. Moreover, we can again confirm though the graph that Chicago's departure delays are larger than Dallas'. This plot is interactive and is made using Tableau. The user can choose the year and the month he wants to be displayed.

Which airport has the longest departure delays, the Chicago-ORD or the Dallas-DAL ?

Which factors affected the largest departure delay (in min) so far for the busiest months per year, for Chicago-ORD and Dallas-DAL?

Diving further into the results of the previous plot, it is interesting to examine which specific delay reasons seems to affect more the largest reported departure delays. Specifically, considering the max values highlighted from the previous graph, we analyze the months that these values occur per year, the rest reasons of delays. We conclude that the major factor for both airports is the aircraft delay and in 2007 the largest delay taking place.

Arrival Delay of Flights from some important airports
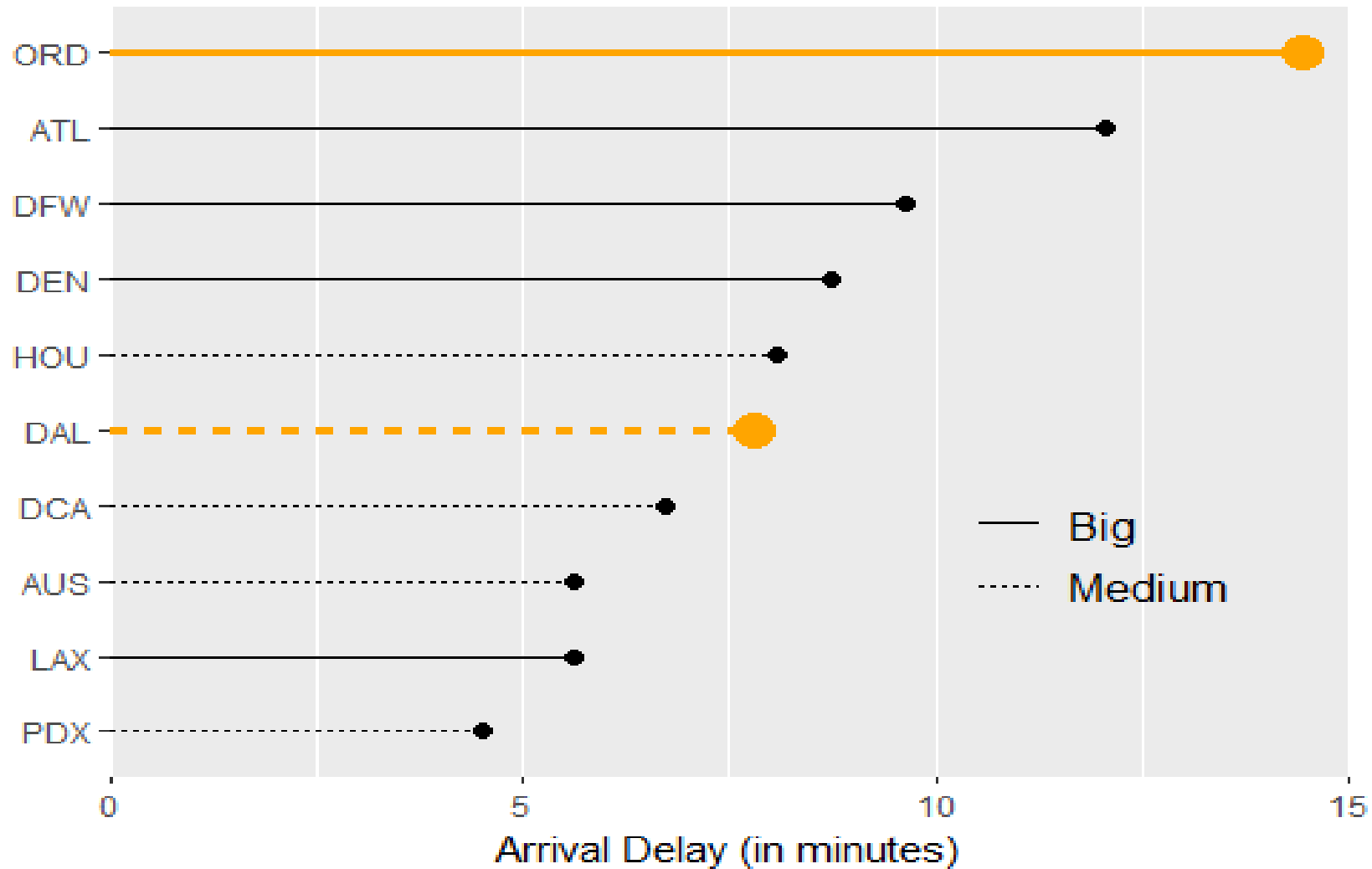
Flights from **O'Hare** airport have the longest average delay, from all other 4 biggest airports in US. The arrival delay is almost 15 minutes, which is approximately the maximum acceptable delay. This should cause some concern to the airport management. Furthermore, flights from Dallas arrive at their destination with an average of almost 8 minutes delay. We can clearly see that the size of the airport affects the delay of flights.
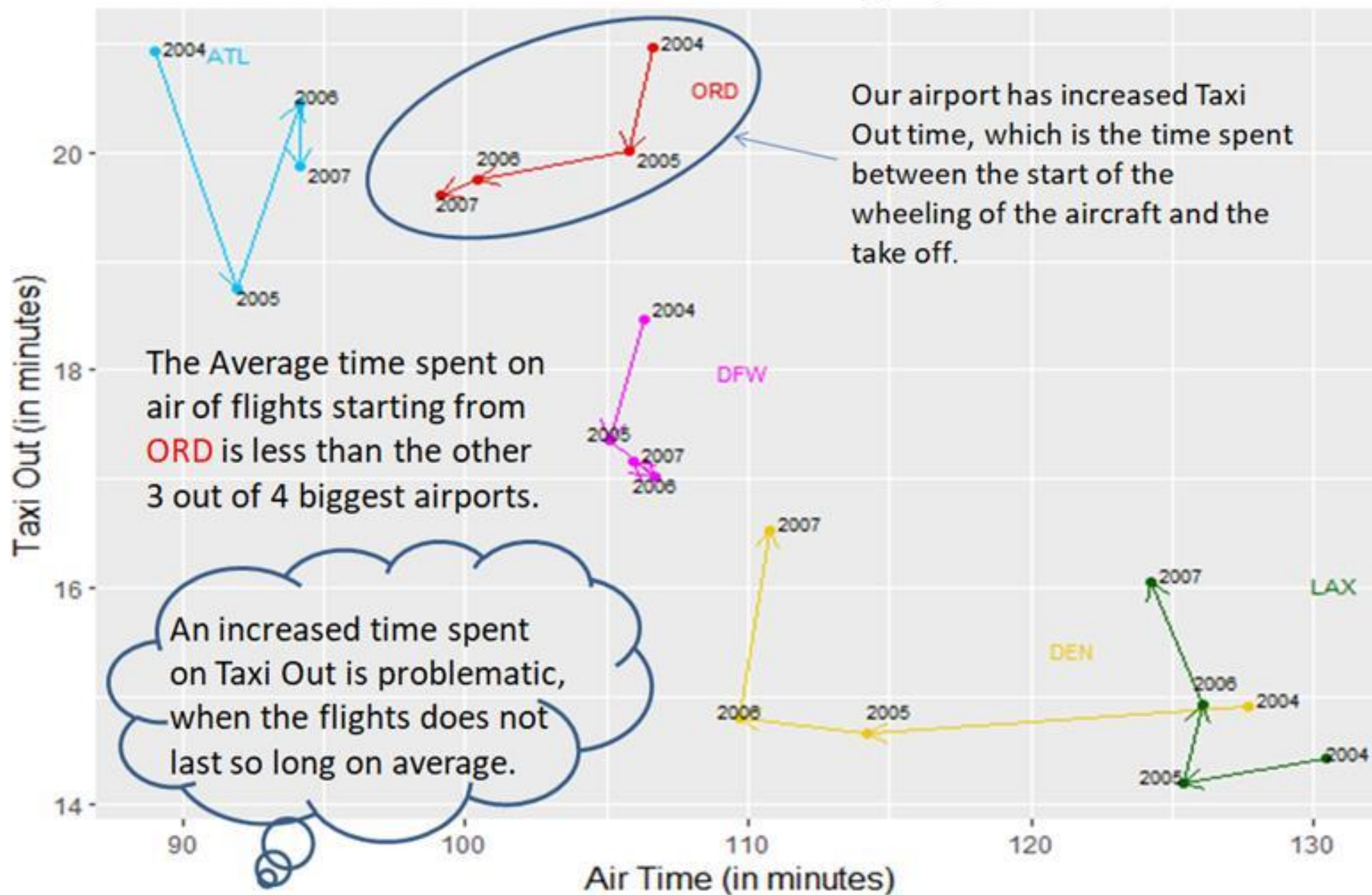
Difference of average arrival delay
Per Month and Year
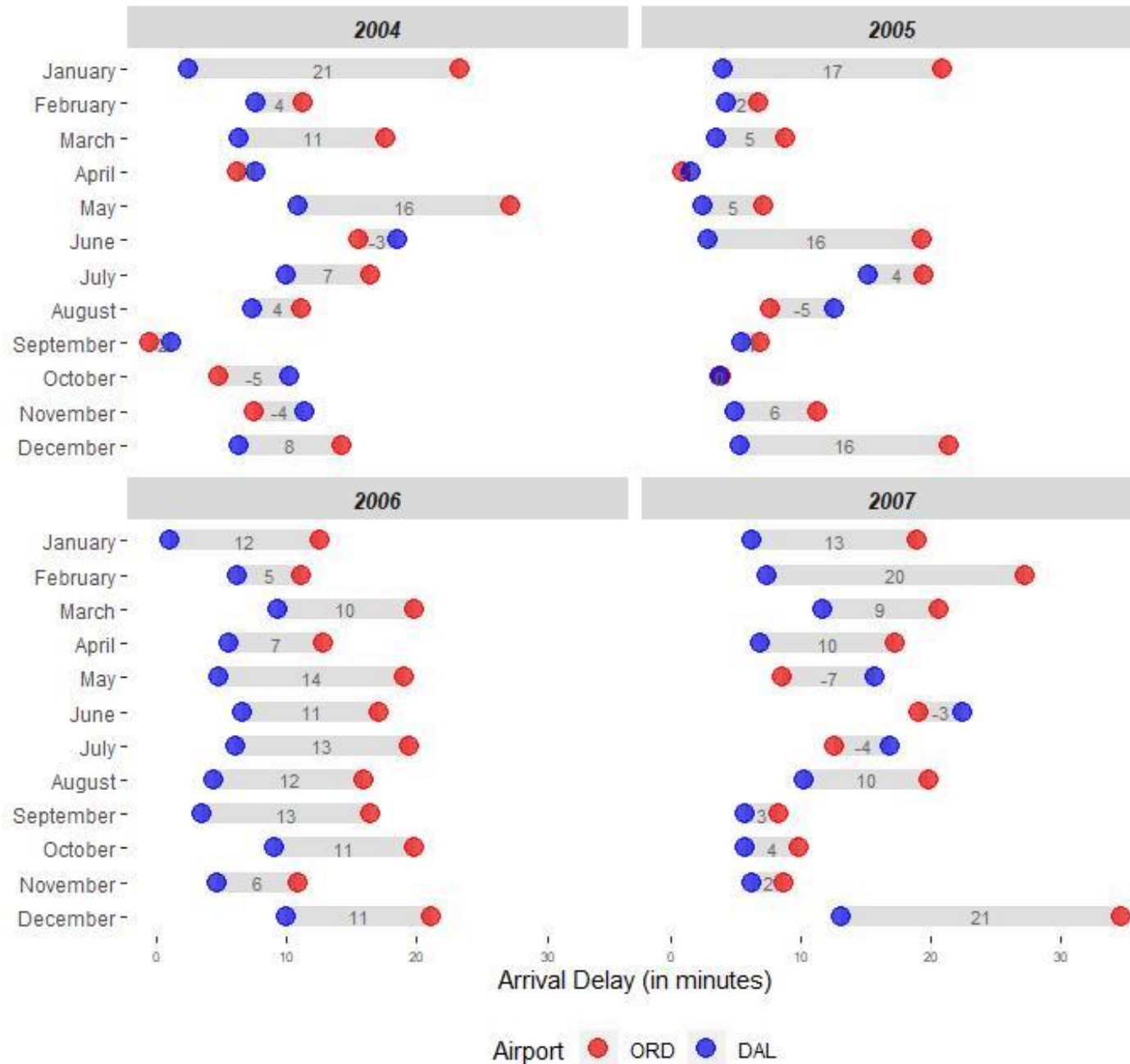
This dumbbell plot is a more analytical version of the previous plot. It shows the average actual and scheduled arrival time difference (average arrival delay) per month for every year between 2004 and 2007 for our two target airports. Additionally, it shows the distance of the two differences in minutes. It is worth mentioning that in most cases the difference for Dallas airport is smaller than Chicago's and also less than 10 minutes. What is more, there are some months, like January of 2004 or December of 2007 , that the average arrival delay is more than 15 minutes for O'Hare airport which is disappointing.

Correlation of Taxi Out and Air Time for the 5 big airports

Now we try to examine the number of diverted flights which are the flights that have been rerouted to a new location per month. We can see the percentage of the rerouted flights per month for each airport from 2004 to 2007. It is evident that only few flights were diverted from their original destination. But despite the rarity of diverted flights, we can spot some behaviour for each airport. In particular, for O'Hare airport, we can observe an increased number of diverted flights during the summer months, May, December and January. A similar pattern appear for Dallas where the most instances occur on May, June and December. In addition, if we compare the two airports, we can see that the rate of flights from Dallas that change their destination is consistently smaller.

Diverted Flights
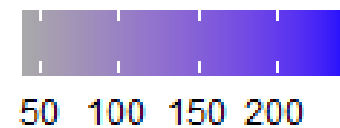
Cancellations

ORD
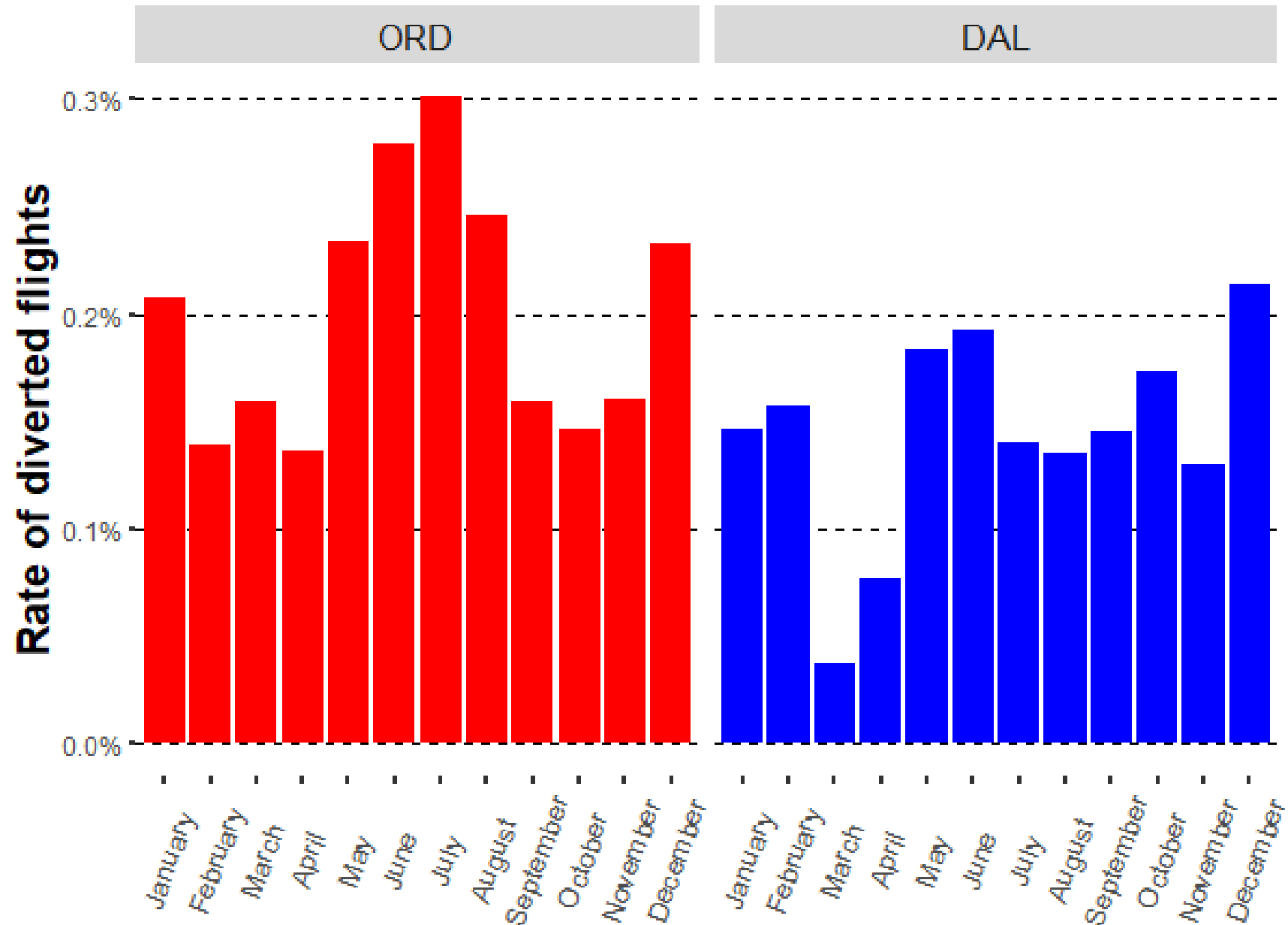
Month	2004	2005	2006	2007
January	2545	2139	988	1268
February	903	640	742	2929
March	998	491	1013	1885
April	380	330	722	1612
May	1928	601	1123	491
June	755	593	679	1432
July	973	868	1137	1008
August	844	448	883	1337
September	491	426	1131	458
October	386	363	1490	485
November	611	382	867	529
December	911	1185	2140	3181

DAL
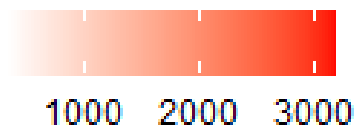
| Month	2004	2005	2006	2007
January	144	113	100	248
February	128	80	57	105
March	82	72	63	82
April	72	39	51	50
May	106	62	58	63
June	164	50	81	111
July	93	92	78	81
August	89	139	94	57
September	59	126	74	62
October	98	53	114	47
November	76	47	102	48
December	75	98	98	77

Another thing we wanted to explore was the cancellations for the two airports. For the O'Hare airport, we notice an increased number of cancelled flights during the winter months. Especially in December of 2007 there is a very high amount of cancellations which seems to be correlated with a high average departure delay shown on previous plots. Now about Dallas, the cancellations seems to be more uniformly distributed with a spike appearing in January 2007. Finally, regarding both airports it seems like there is not a certain pattern of cancellations among the months and the years observed.

Cancellation Reasons
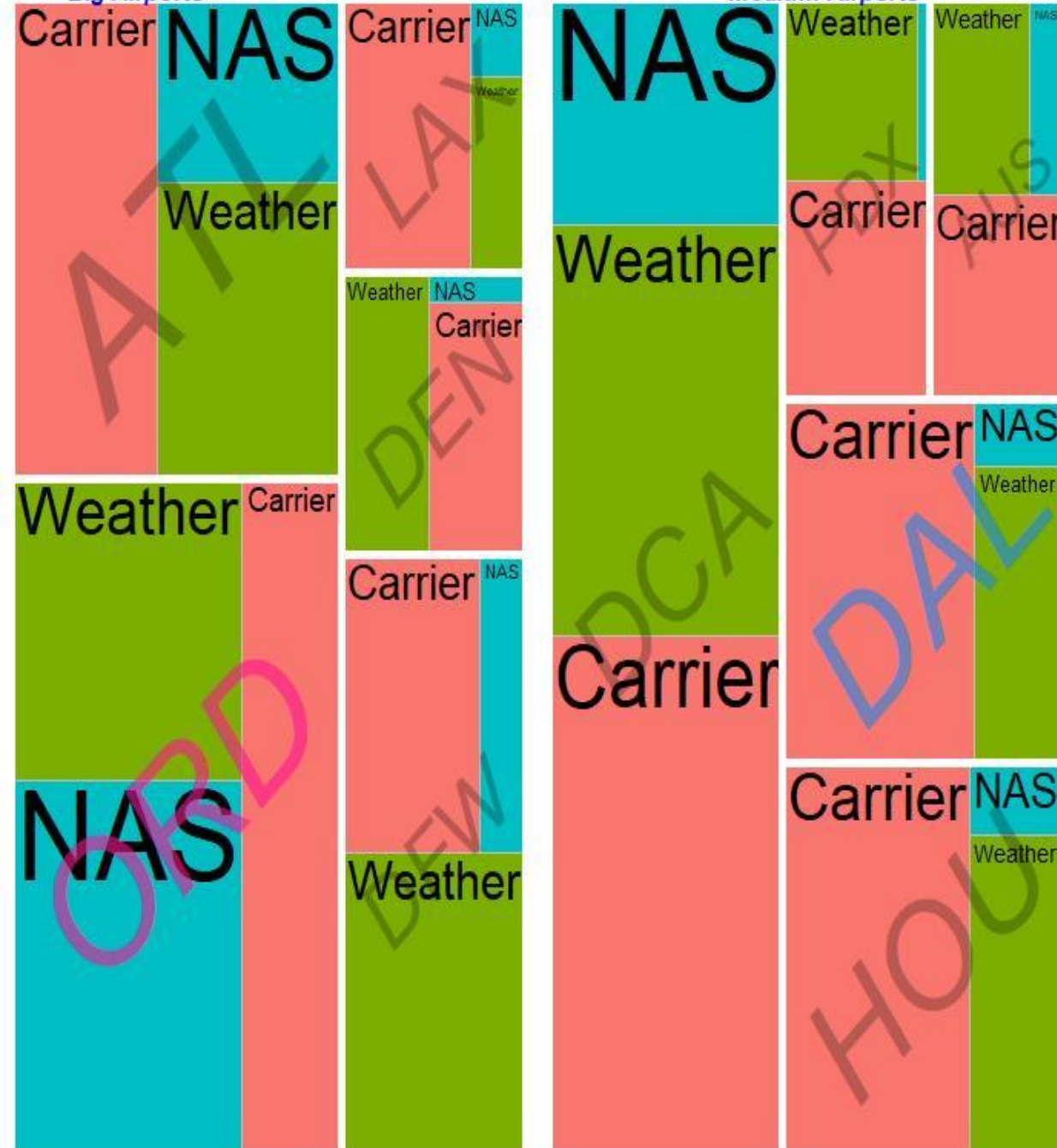
Taking a closer look to the cancellation reasons, it is clear that the increased number of cancelled flight in winter months for O'Hare airport is probably due to the bad weather of Chicago. In addition, we can notice a consistent amount of cancellations happening because of the carrier, in all airports and an even more intense problem in medium airports as it is the most common reason. Finally in big airports, it seems that there is not a standard distribution among the cancellation reasons.

Contribution of each member

Each member of the team worked on at least 2 graphs, with the work being distributed evenly. Also the entirety of the plots were discussed and altered based on the opinion of every member and finally the report was made by all the members of the group through a call.