Reducing delays to improve customer satisfaction

Analyses to help reduce delays for an airline client

Context

Client is currently #3 by volume of flights flown. They are keen to understand on how best to improve their business performance across metrics such as customer satisfaction etc.

1. Data – there are 12 files corresponding to 1 year of data that captures flights by origin:destination ids and airline code
2. You are more than welcome to add any 3rd party data sources

Questions

1. Which airline has the 5th highest delay?
2. Which day of the week is the 2nd most popular?
3. Which is the 4th most flown sector?
4. Add any other analyses you believe would shed additional light on our client and their competition
5. What do you believe are the root causes of their issues?
6. What do you recommend the airline addresses?
7. How would you enhance the study given additional time

1. Specific questions
2. Additional analyses
3. Recommendation
4. Next steps

1. **Specific questions**
2. **Additional analyses**
3. **Recommendation**
4. **Next steps**

As there are multiple airlines serving the same airlines, through the analysis each airline will be evaluated as a standalone

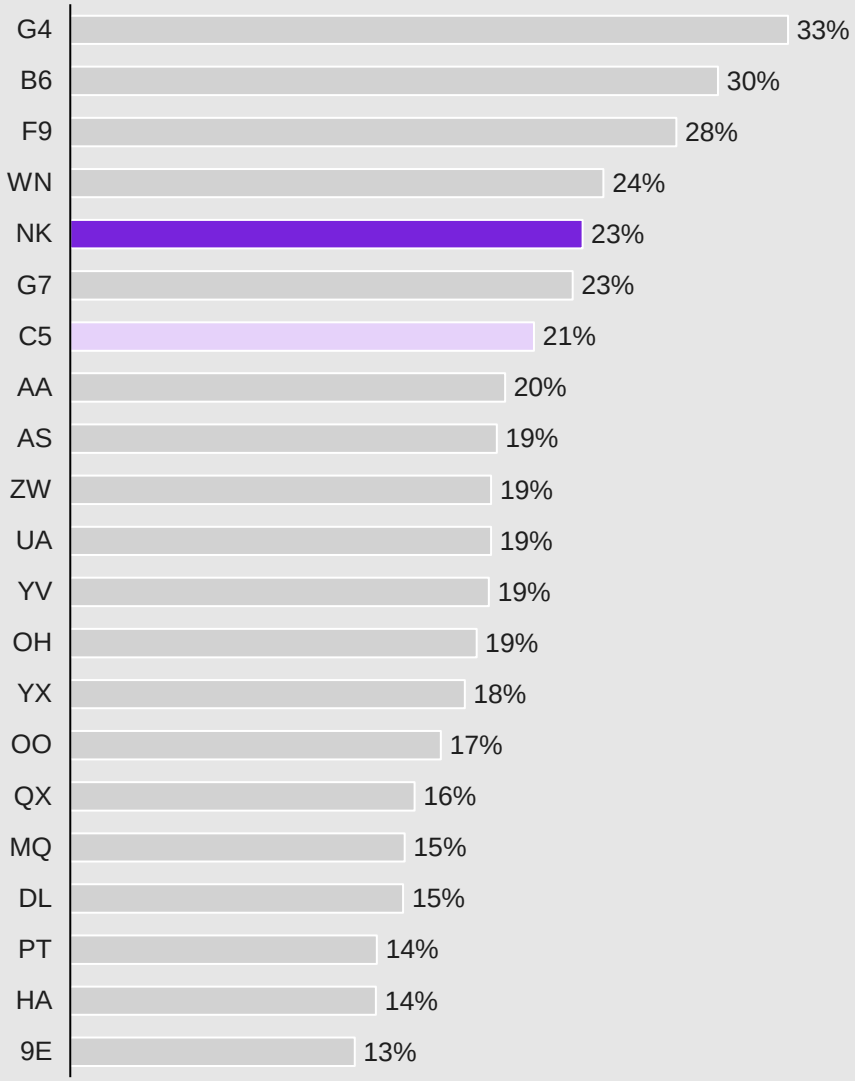
	American Airlines	Alaska Airlines	Delta Airlines	United Airlines
Skywest	✓	✓	✓	✓
Republic Airways	✓		✓	✓
Mesa Air	✓			✓
Endeavor			✓	
CommuteAir				✓
GoJet				✓
Envoy	✓			
PSA	✓			
Pierdmont	✓			
Horizon		✓		
Air Wisconsin				✓

1. Airline with 5th highest delay

Despite NK being the 5th airline with highest percentage of delays, C5 is the 5th airline with highest expected delay time

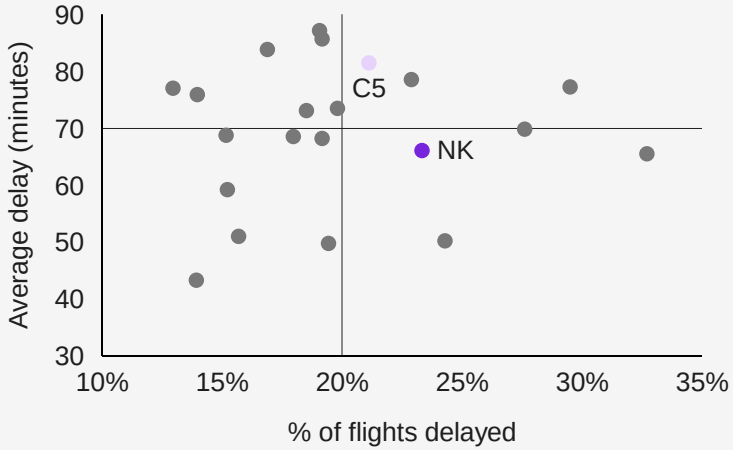
NK is the 5th airline with the highest percentage of delayed flights

% of flights delayed by airline



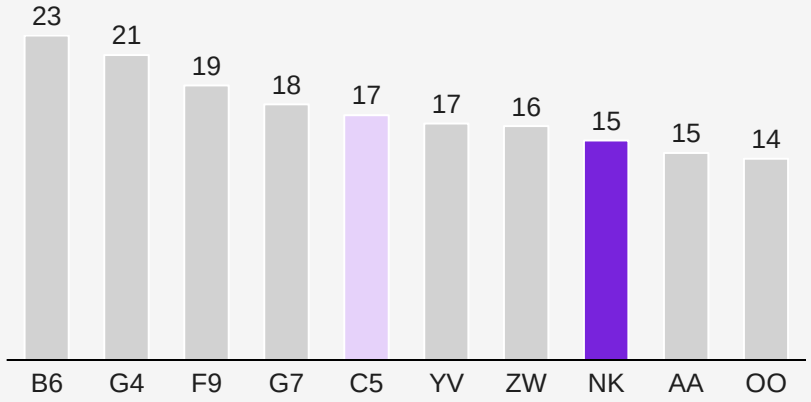
Nonetheless, its average delay is lower than industry average

Average delay (min) vs % of flights delayed by airline



Allowing them to have a better performance in expected delay time than others

Expected delay (min): % delay x average delay

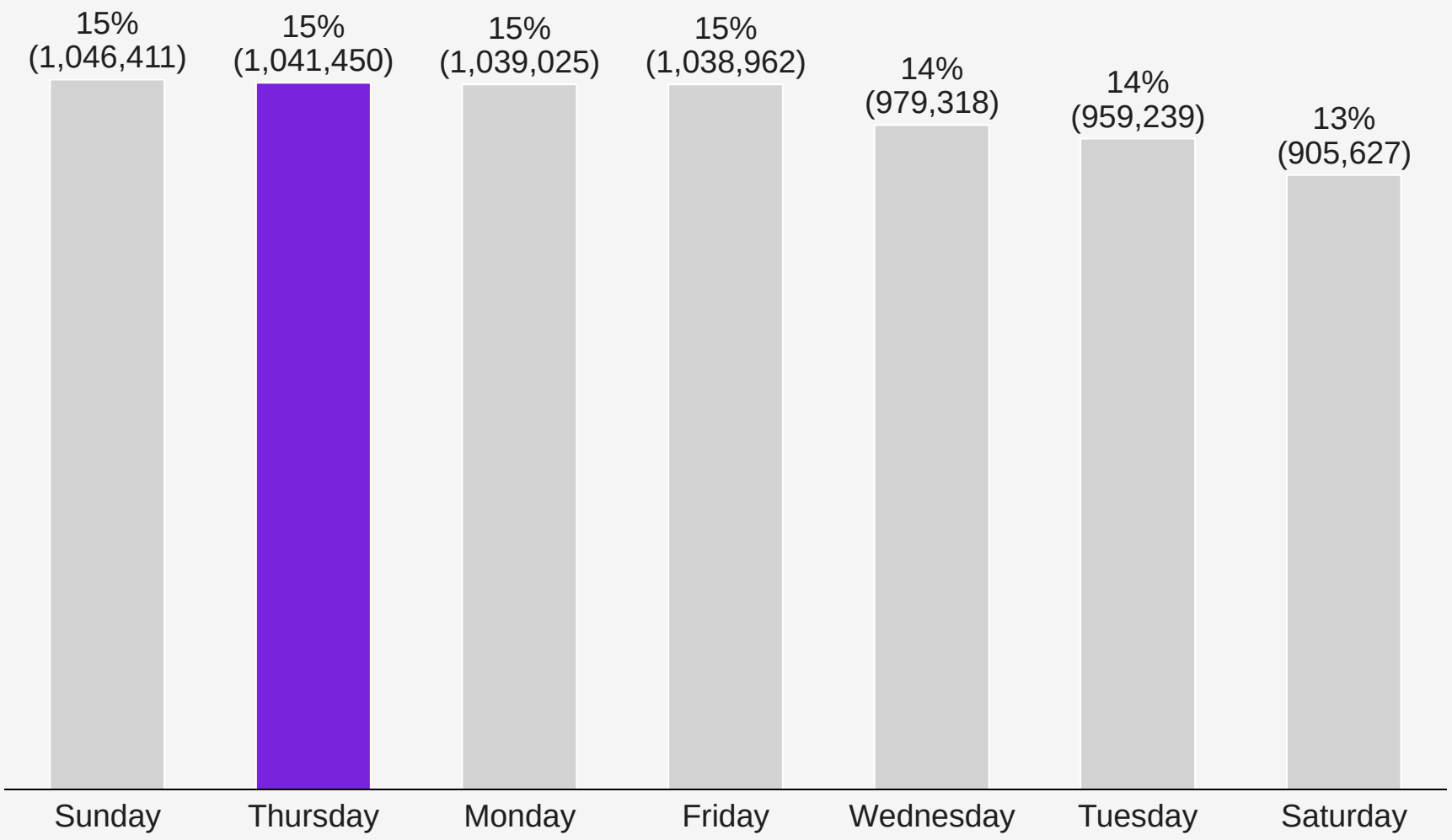


2. 2nd most popular day

While Thursday is the second most popular day, we are only looking at volume of flights and not number of people flying per day

Number of flights per day of week is relatively stable

Distribution of the number of flights per day of week

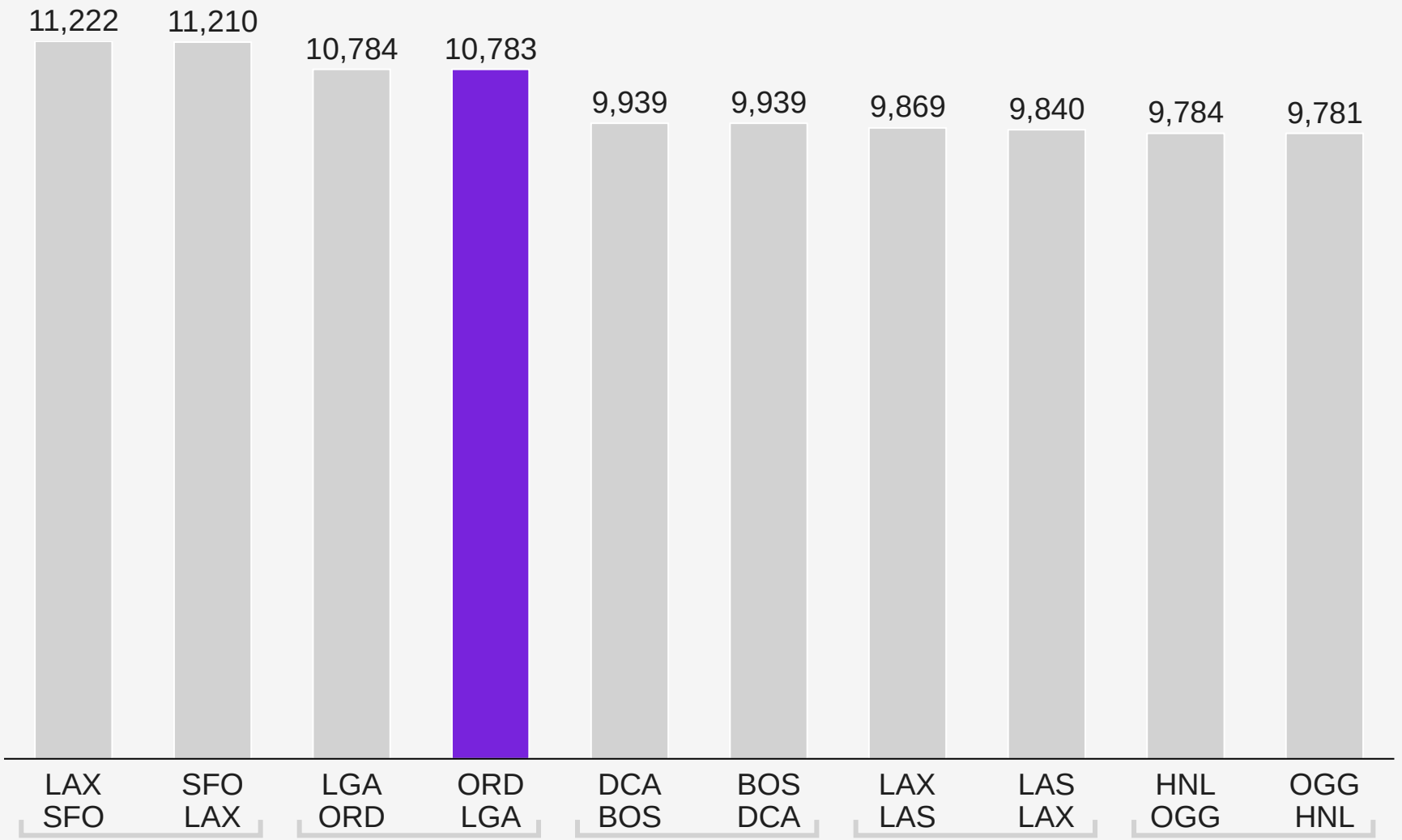


3. 4th most flown sector

Chicago O’Hare to La Guardia is the fourth most flown route

Outbound and inbound routes tend to be aligned on the number of flights they have

Number of flights per route



From research it was found that sectors are usually interpreted as a directed graph. This means that LAX-SFO and SFO-LAX are two different sectors, even if they are between the same two points.

1. Specific questions
2. **Additional analyses**
3. Recommendation
4. Next steps

Executive summary

1. AA has an opportunity to improve cancellations rate by airport by identifying best practices from top performing airports

- AA has similar performance across reasons for cancelations than the market but in terms of cancellations rate it is underperforming compared to similarly sized airlines
- LAX, PHL, and PHX are the airports with best cancellation rate given their size and are potential centers of excellence to learn best practices
- Still, AA is underperforming in its most important airports in cancellation rate compared to the competition and has an opportunity to improve

2. There are opportunities to further improve cancellations rate by looking at cancellations by sector, which could yield better results than optimizing airports

- It seems as if there are differences in how sectors are managed even if they depart from the same airport, which yields differences in cancellation rates for flights departing from the same location
- This implies that AA can look for best practices not only at the airport level but also at the sector level

3. While big airport delays are robust against delayed flights, they are having issues with turnaround time, on the other hand, smaller AA's airports struggle with both issues

- Despite AA having higher delay rates, its average delay time is on par with the competition; nonetheless, this seems to be due to a reduced delay time of late aircraft
- Nonetheless, while big airports are performing well on late aircraft delays, the rest are not, and they are facing double pressure to improve delays be reducing carrier as well as late aircraft

4. From the big airlines, AA is the one that requires the highest number of stops to connect all the airports in its network, suggesting that its network connectivity is lagging

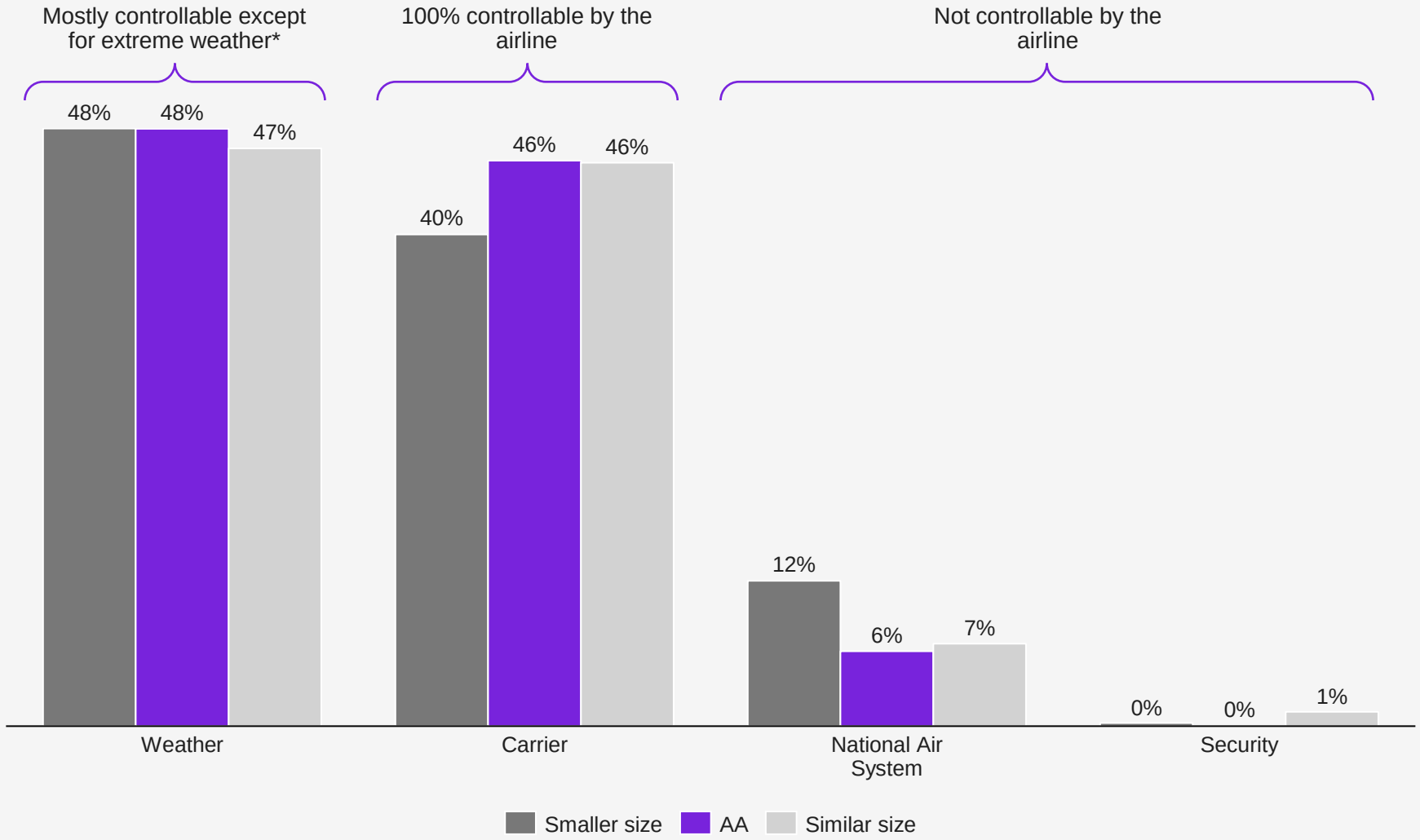
- Given the increase in number of stops required, the probability of being exposed to delays or cancellations in a trip also increases, playing an important role in reducing customer's exposure to delays and cancellations

1. Cancellations by airport

AA seems to follow market averages in both controllable and uncontrollable factors for canceling flights

There are four reasons for which a flight can be canceled, two of these are controllable by the airline and they represent ~ 90% of total cancellations

% of flights canceled by type of reason



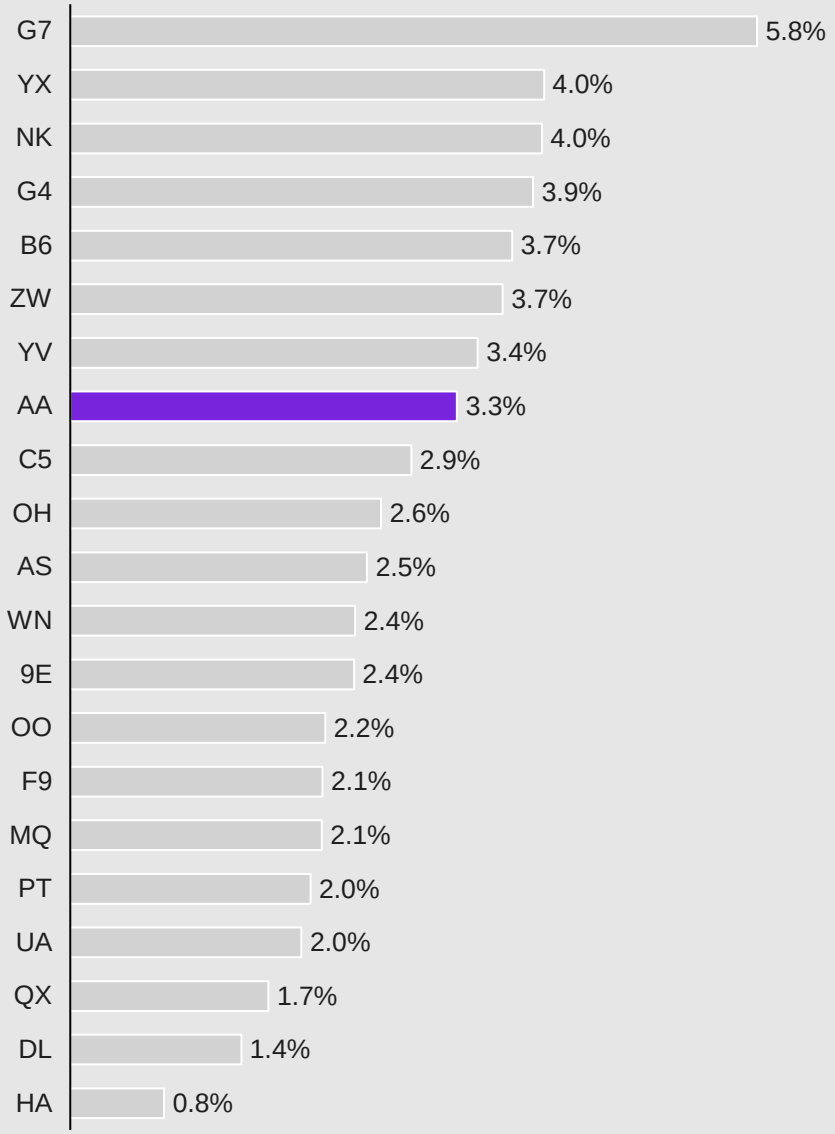
* It is being assumed that weather cancellations can be controlled as the airline decides when to cancel a flight due to weather conditions except in the cases of extreme weather. Additionally, there is no federal requirement on what constitutes weather.

1. Cancellations by airport

Nonetheless, AA has mid-table performance on cancellations, and it is not performing as well as similarly sized airlines

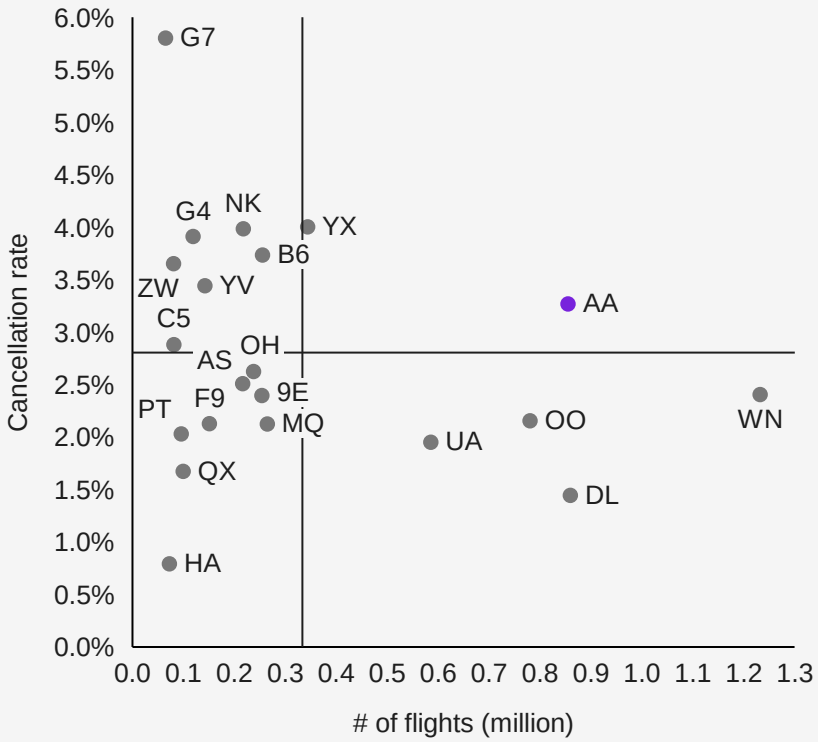
AA has a mid-table performance in the airline industry in the cancellations ratio

% of cancellations rate by airline



Nonetheless, similar sized airlines have significantly lower cancellation rates

$ of cancellation rate vs number of flights by airline



Average cancellation rate

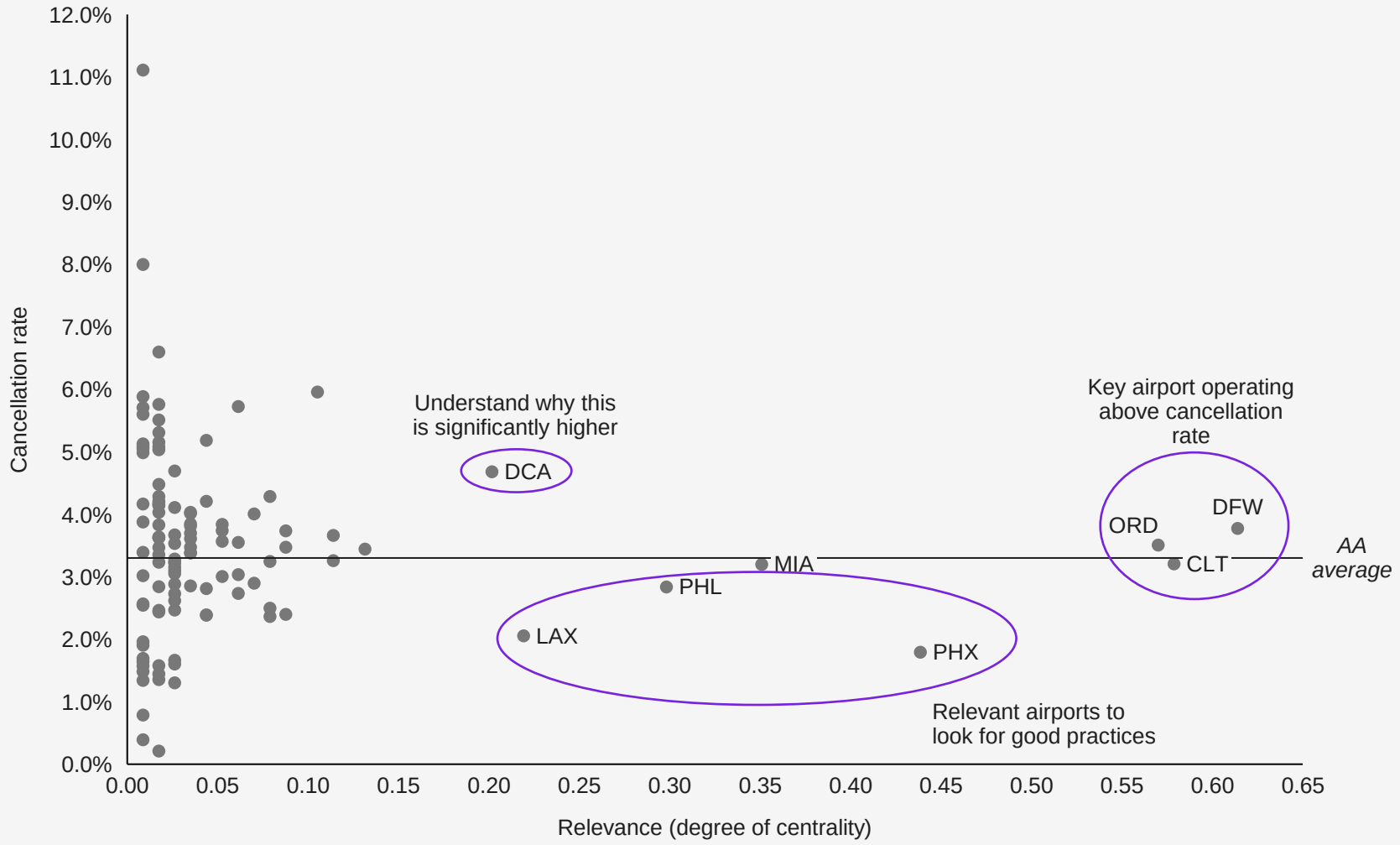
$$\frac{\textit{AA: 3.3\%}}{\textit{Peers: 2.0\%}} = \textit{1.6}$$

1. Cancellations by airport

At airport level for AA there are significant differences in rates, indicating opportunity to implement best practices

AA's more relevant airports at close or above average in cancellations rate, meaning that a change in these could have a big effect on the number of flights and passengers

Cancellation rate vs relevance of airport (measured using degree of centrality)

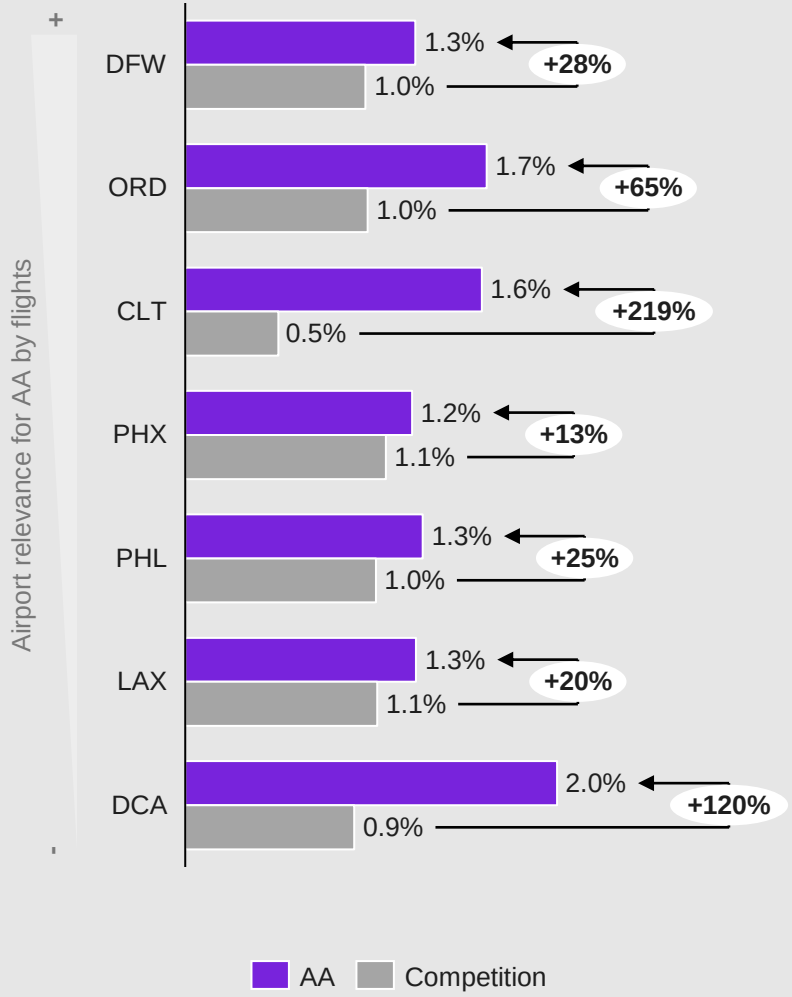


1. Cancellations by airport

And while best performing airports are not better than the competition, the gap is smaller than in the other ones

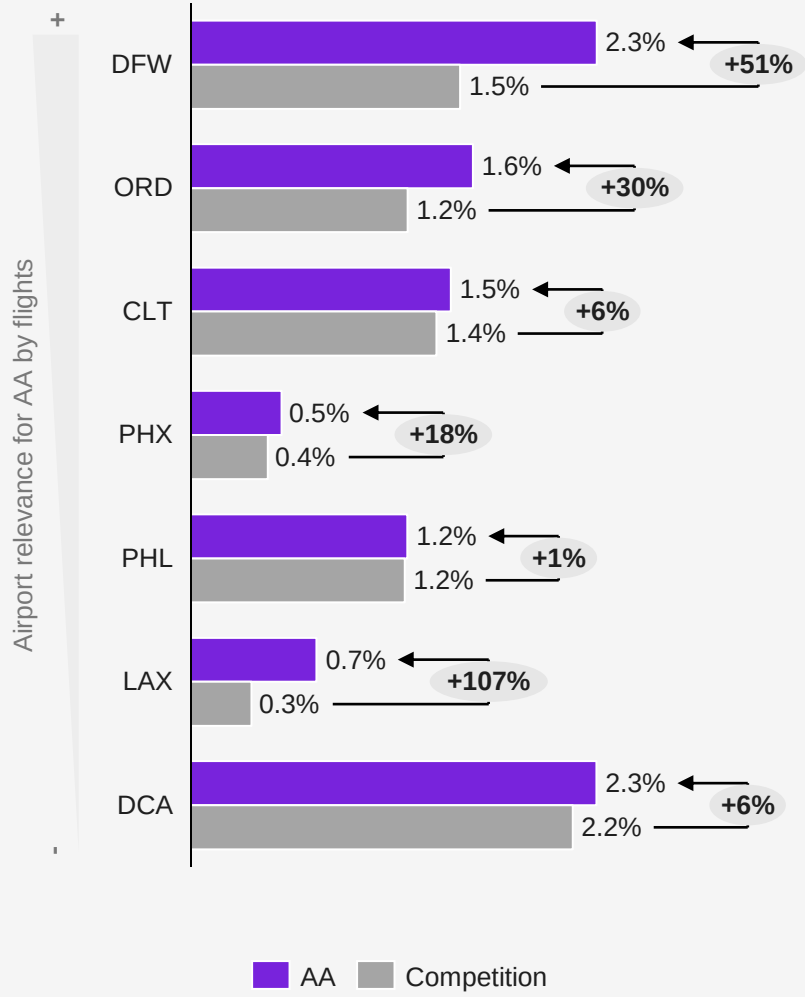
In carrier cancellations DFW, ORD, CLT, and DCA are lagging significantly

% of carrier cancellations



In weather cancellations we see big gaps in DFW and LAX

% of weather cancellations

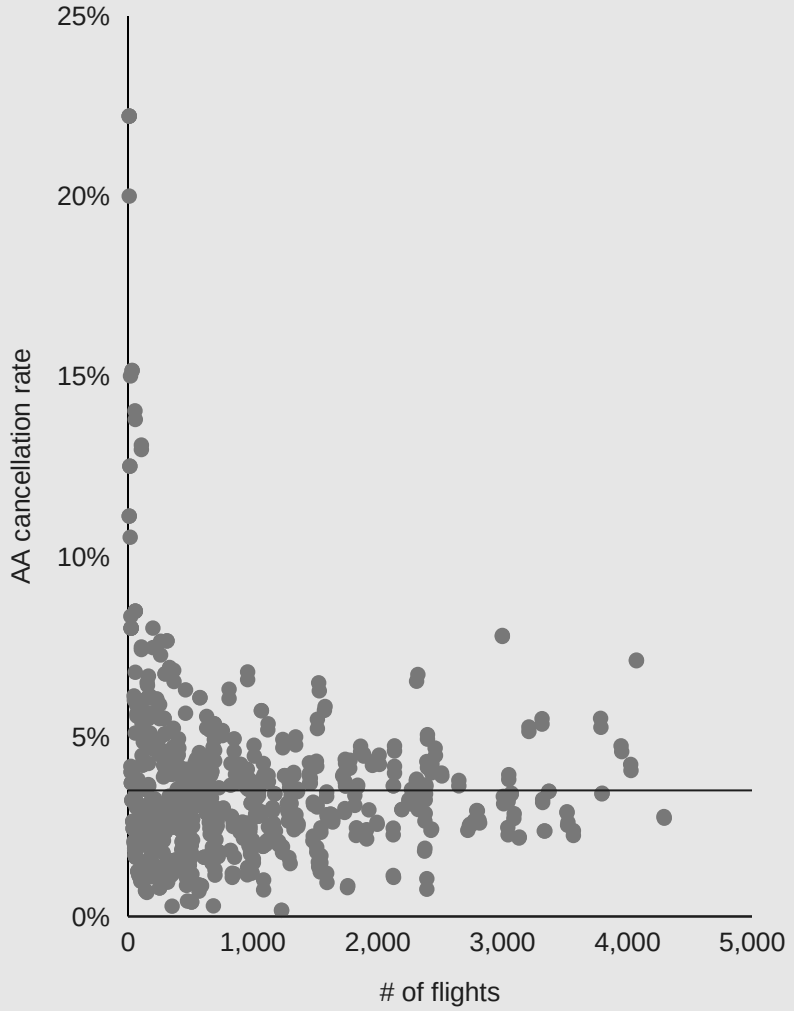


2. Cancellations by sector

In sectors the story is similar, where AA tends to have lower performance than its competitors on the same routes

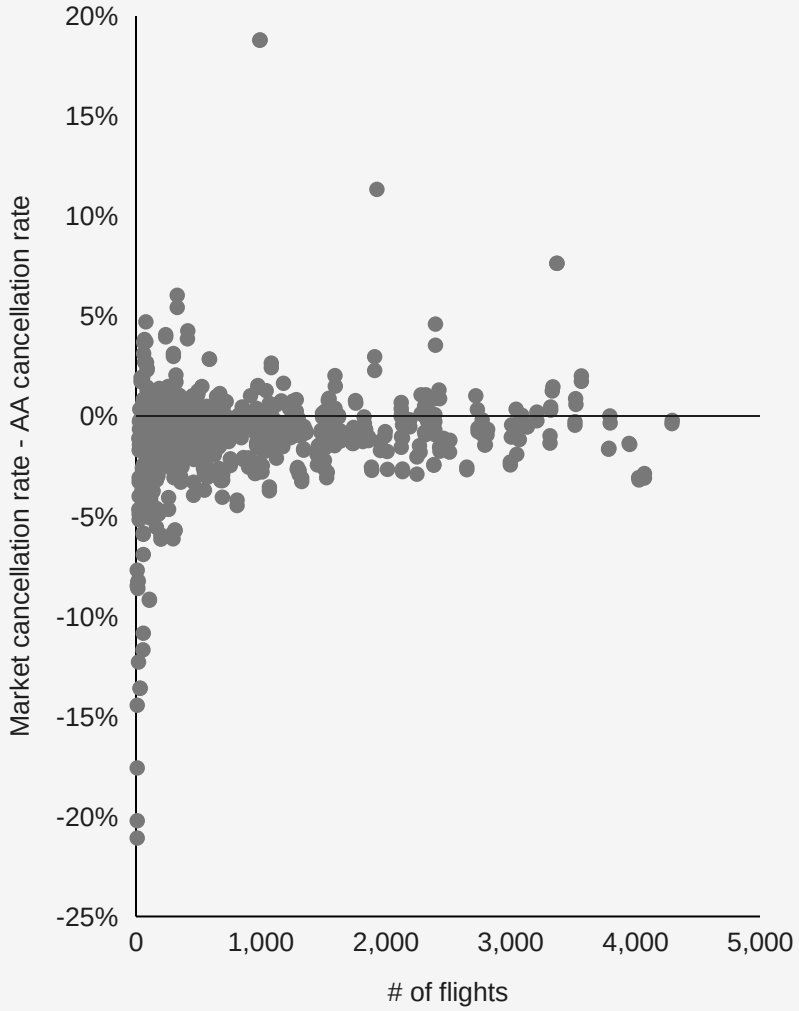
More relevant sectors tend to have higher than average cancellation rates

% of cancellations vs # of flights for AA



Comparing versus the competition we see that AA tends to underperform

Difference between cancellations from the market and AA vs flights

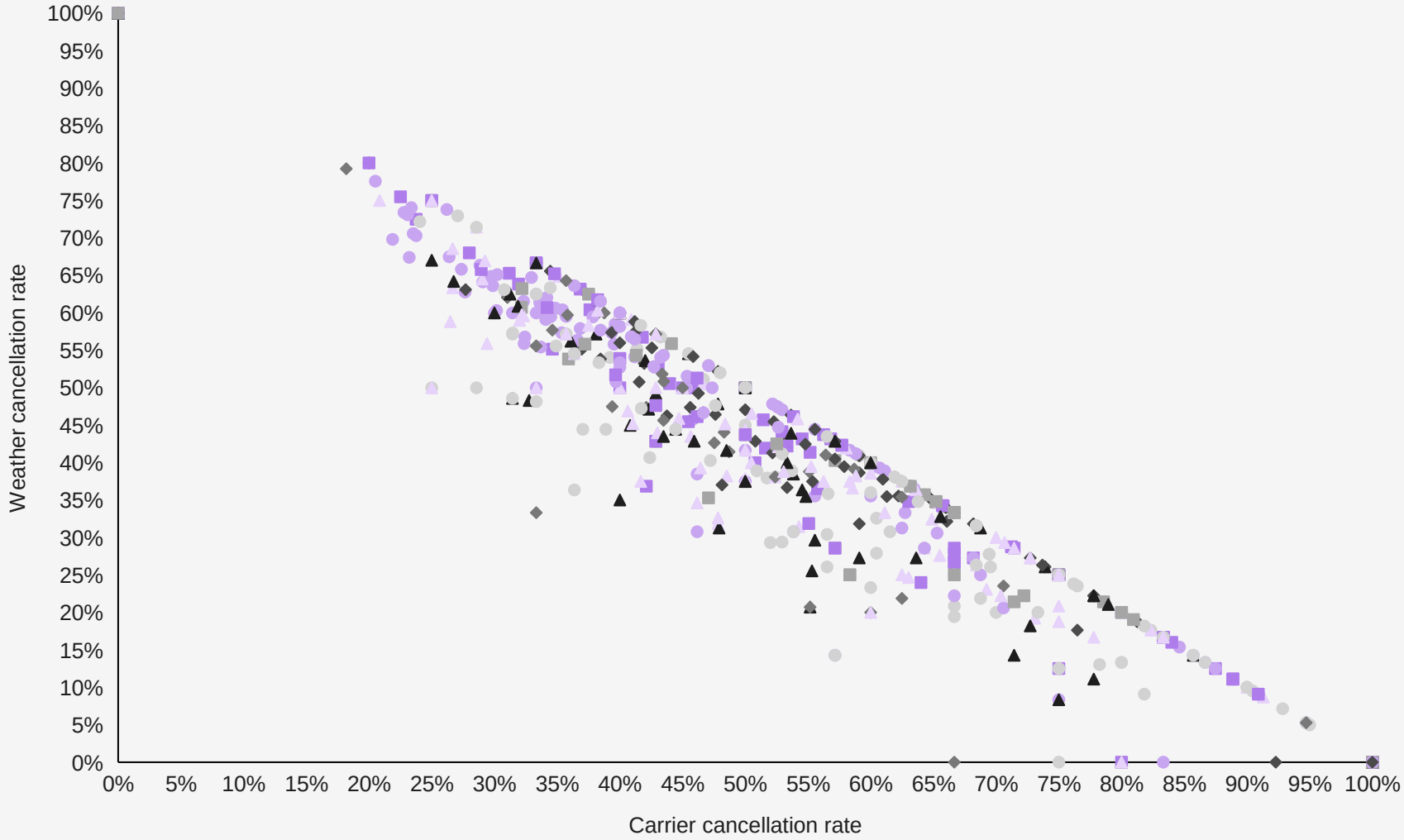


2. Cancellations by sector

Additionally, the lack of segments suggest that what drives cancellation rates is the sector and not the airport

If airport was a confounding factor, we would expect to see clusters of routes changing together, which is not the case, suggesting that airport is not a major driver of the rates

% of weather cancellations by % of carrier cancellations by airpor



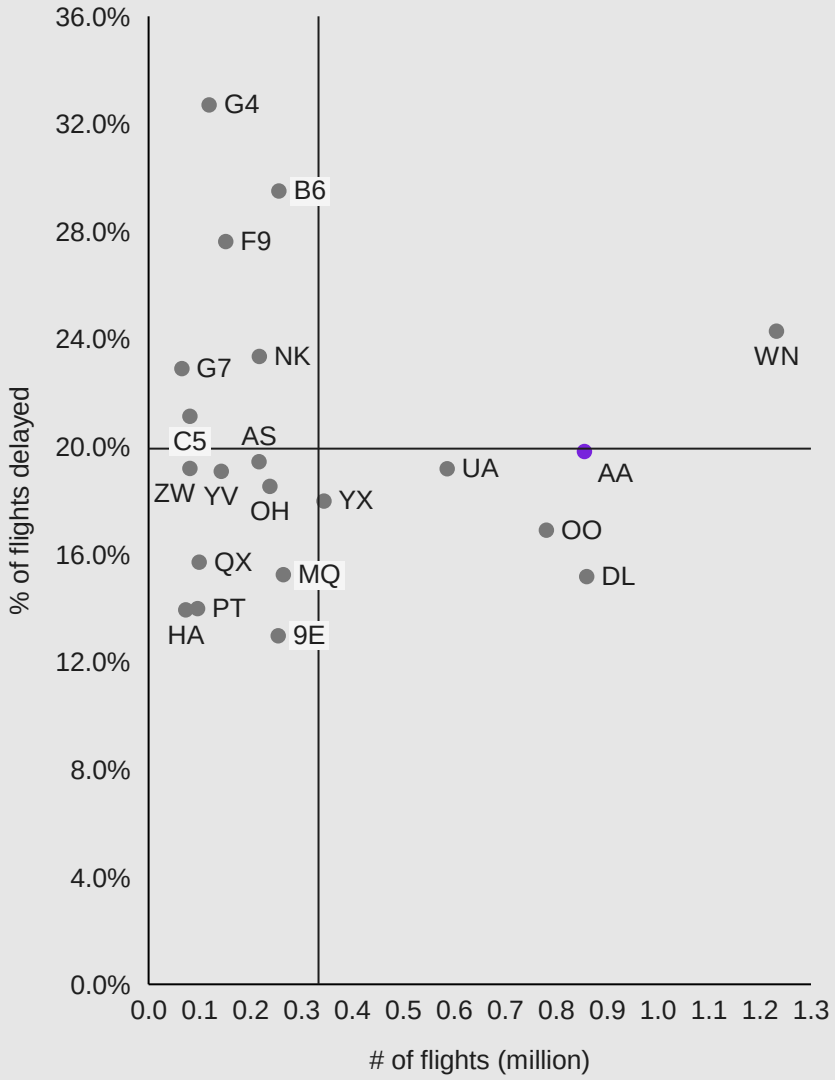
Color and shape represents airport
The correlation between carrier and weather should not be taken into account as both variables usually make close to 90% of cancellations and that behavior could be expected

3. Delays by airport

AA has a higher delay rate than its peers but also a higher average delay due to a significantly higher late aircraft average delay

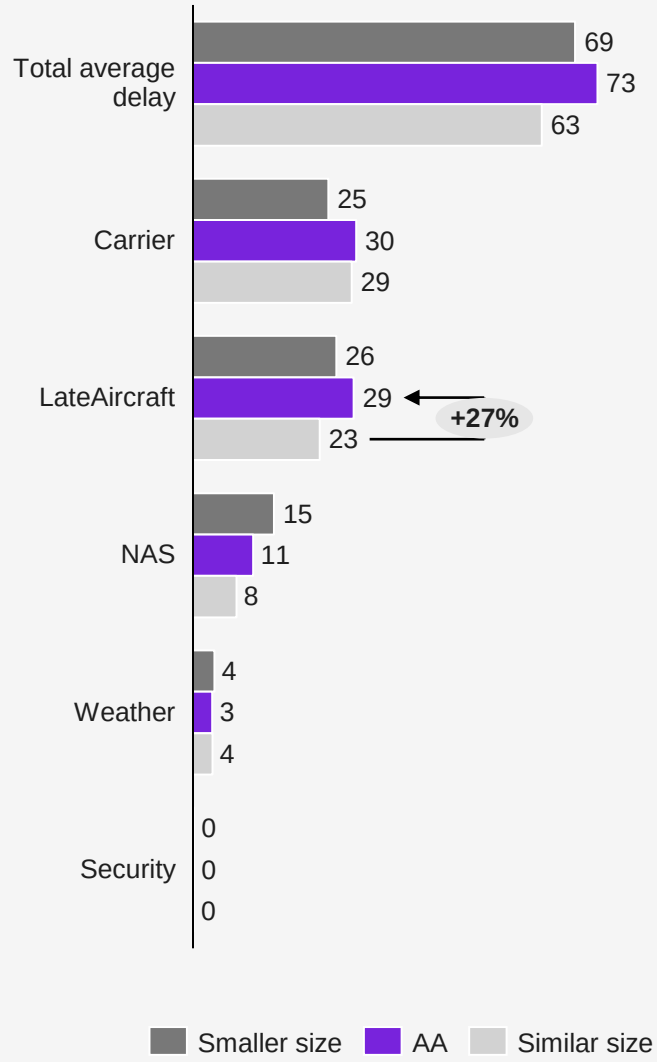
AA's delay ratio is at market average, but similar sized airlines are performing better

% of flights delayed vs number of flights per airline



AA is mostly lagging in delays due to late aircraft compared to the industry

Average delay time by reason (min)

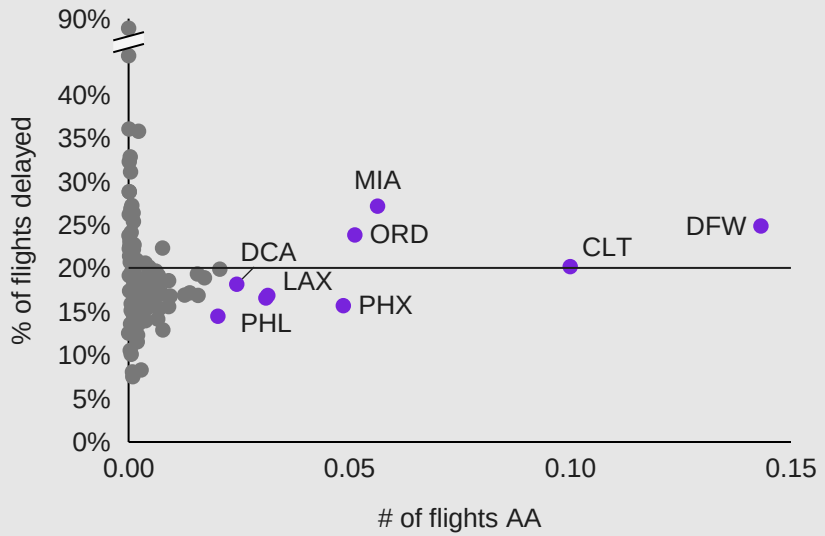


3. Delays by airport

Despite AA having higher delay rates in its main airports compared to themselves and the market, in time they tend to perform well

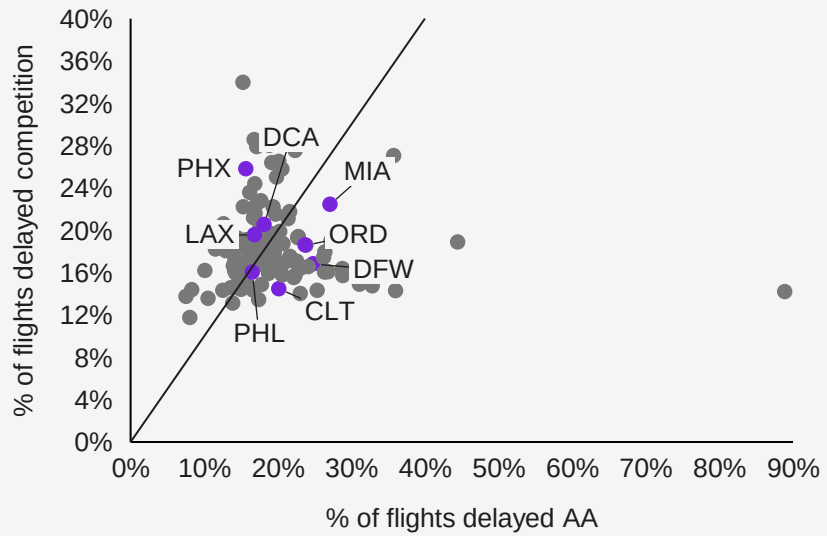
AA's most important airports tend to be over average for delays ratio

% of delayed flights vs volume of flights per airport for AA



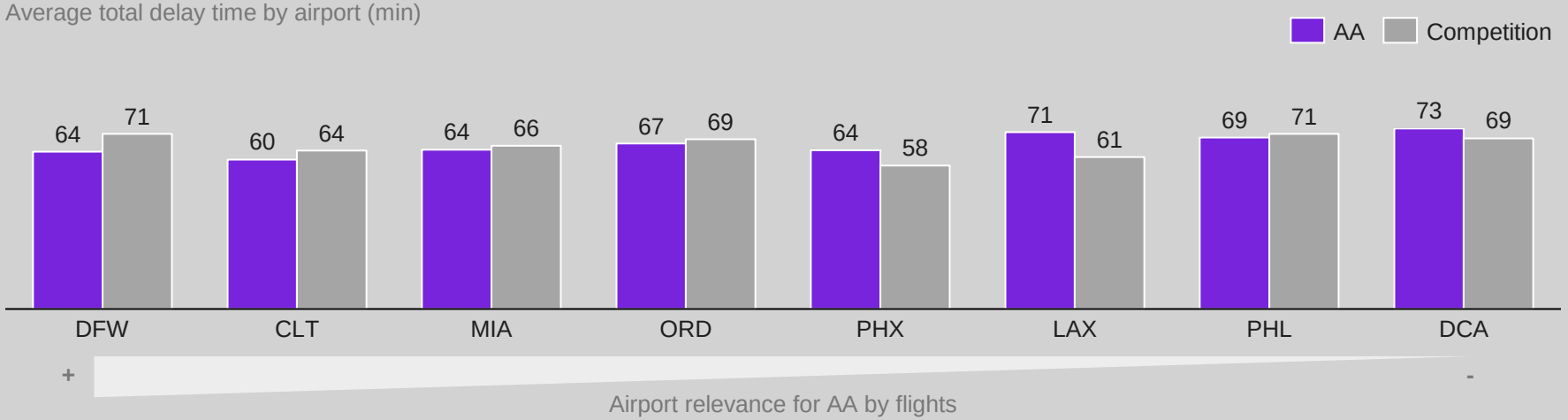
AA has higher delay rates in 5 out of 8 of its main airports compared to the competition

% delay flights competition vs % delay flights AA by airport



AA is mostly performing slightly better than the competition in its main airports and even when it lags, the differences are not big enough to be considered practically significant

Average total delay time by airport (min)

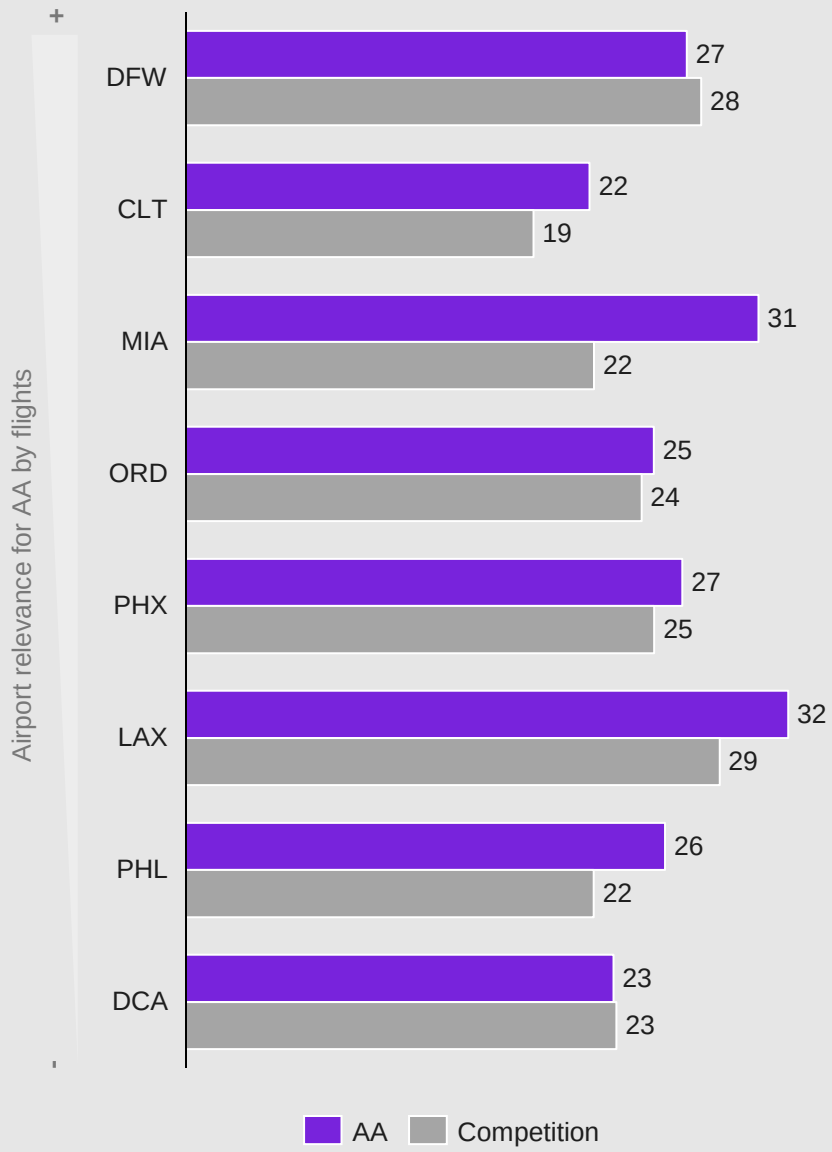


3. Delays by airport

Given that AA is having higher delay times due to carrier, it seems that they are facing an issue in turnaround time of the planes

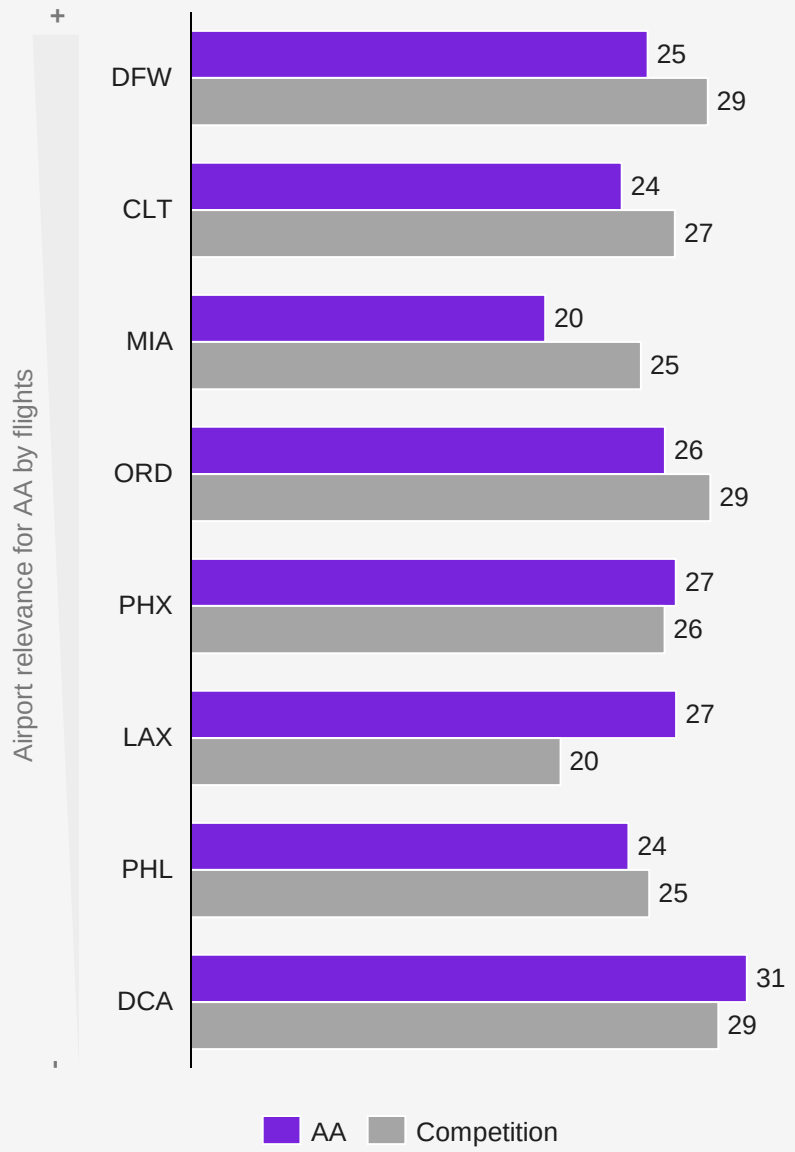
AA tends to have higher delay times due to the carrier than its competition

Average delay time caused by carrier (min)



AA has better performance in late aircraft delay times, helping reduce total delay time

Average delay time caused by late aircraft (min)

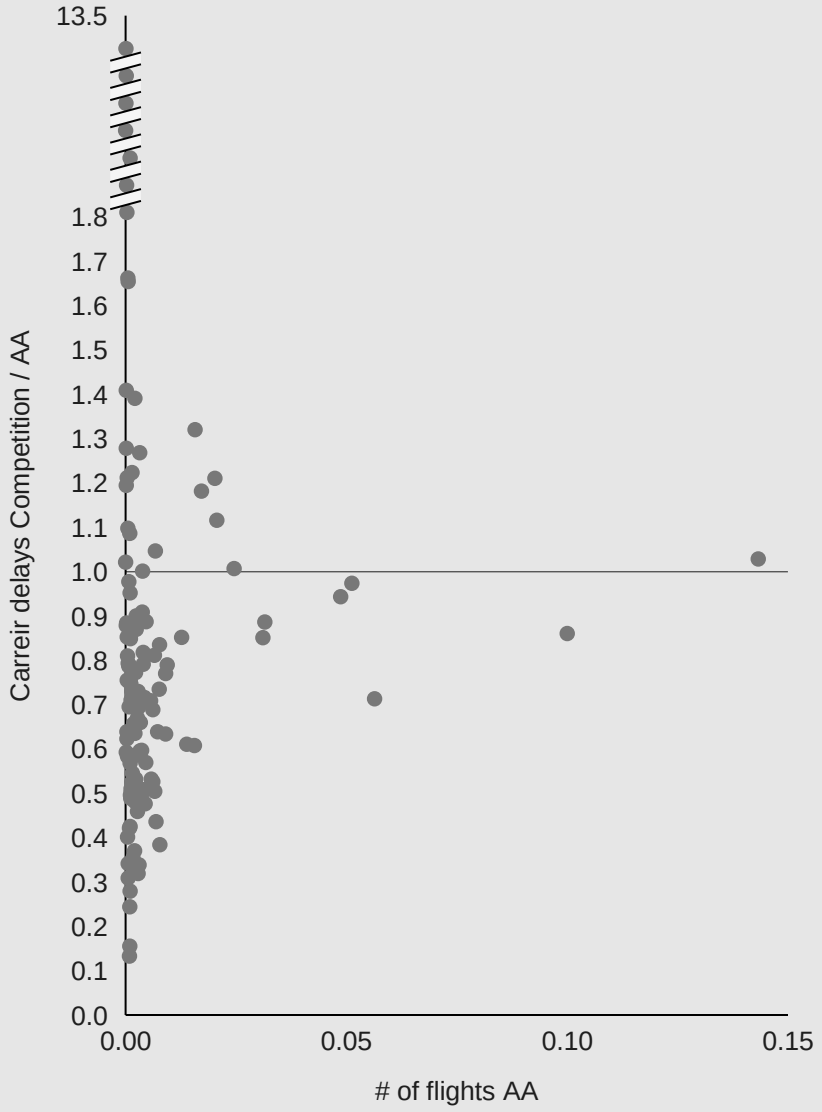


3. Delays by airport

Finally, the only airports performing well on aircraft delays are the large ones, the rest, need to close the gap on carrier and late aircraft

Most of AA's airports, independent of size need to improve on carrier delays

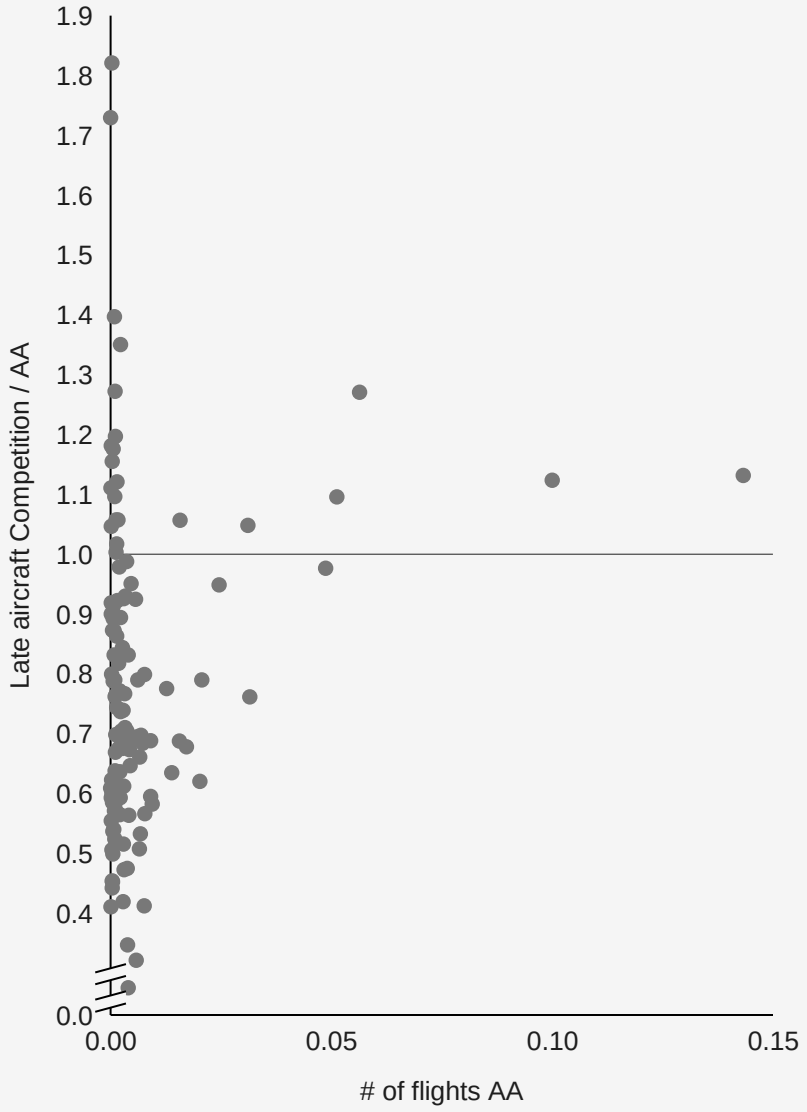
Ratio of competition to AA average carrier delay by flights



Y axis has been cut in both graph to ease visualization due to extreme values

Only AA's bigger airports are performing well on late aircraft delays

Ratio of competition to AA average late aircraft delay by flights



4. Network connectivity

AA's hub lags connectivity which generates more layovers, increasing exposure to delays and cancellations

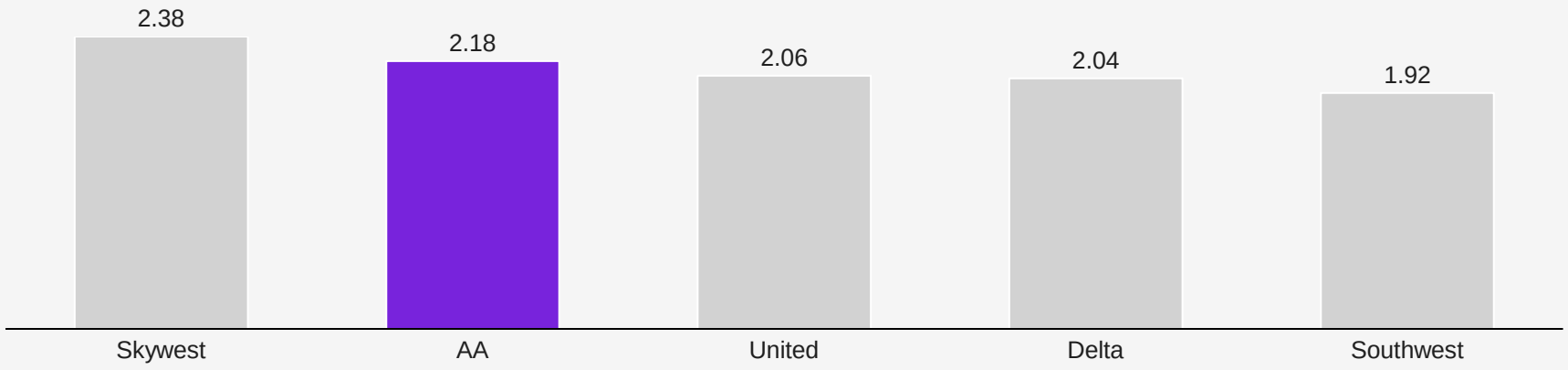
AA is the only big airline for which its hub it's not the best-connected airport

Hub (measure by network centrality[1]) and Best-connected airport (measured by eigenvector centrality[2]) by airline

Airline	Hub	Best connected airport
Southwest	Denver International Airport	Denver International Airport
Delta	Hartsfield-Jackson Atlanta International Airport	Hartsfield-Jackson Atlanta International Airport
American Airlines	Dallas/Fort Worth International Airport	Chicago O'Hare International Airport
SkyWest Airlines	Denver International Airport	Denver International Airport
United	Chicago O'Hare International Airport	Chicago O'Hare International Airport

Excluding Skywest, which is regional, AA is the airline with the highest require number of stops

Average number of stops required to connect all the possible routes in the network



1/ Network centrality is the node in a network that has the most connections, this would be the hub of the airline
2/ Eigenvector centrality accounts for the number of connections of each node but also if the nodes to which it connects are well connected

1. Specific questions
2. Additional analyses
3. **Recommendation**
4. Next steps

AA should focus on three initiatives to improve on delays and cancellations and evaluate the potential expansion of its network selectively

Reduce cancellations rates across airports and sectors

- Improve cancellations rate in DFW, ORD, and DCA by:
  - Understanding differences in operations and policies between these airports and LAX, PHL, and PHX as the latter have similar volumes of operations but significantly better cancellation rates
- Improve cancellations rate in roundtrip routes of LGA-ORD, DFW-LGA, and MIA-LGA by:
  - Understanding differences in operations and policies between these routes and LAX-DFW which has a low cancellation rate and is on par with the competition
- Assess the meteorology team from AA to understand the accuracy of its predictions and benchmark against competition as well as weather-related policies and procedures to understand drivers of higher weather cancellations

Reduce carrier delays in main airports

- Reduce carrier delay times in MIA, LAX, and DFW by:
  - Understanding differences in operations and policies between these airports and DCA
- Once MIA, LAX, DFW has been diagnosed and solved this effort can be escalated to the rest of the main airports
- From the lessons learned from these airports, escalate to the next group of relevant airports

Reduce late aircraft delays in airport network

- Benchmark good practices of main airports to ensure that late aircraft delays are below the competition
- Understand what are the drivers of the late aircraft delays and identify levers that can be pulled to reduced the delays

1. Specific questions
2. Additional analyses
3. Recommendation
4. **Next steps**

Next steps

Additional data and sources of information

Operational data

- Carrier, weather, and late aircraft cancellations and delays detail to understand the underlying cause of the delay/cancellation (for example crew missing, etc.)
- Plane (and history) used in each flight to see if there is a pattern of cancellations or delays based on plane model/year/maintenance history/etc.
- Relationship of flights to understand which routes are flown together and identify potential drivers for cancellations or delays

Customer data

- Customer database to understand behavior and how delays and cancellations affects retention and lifetime value
- Past studies they might have on customer satisfaction, segments, drivers, and satisfaction

Interviews

- Interviews with operations, scheduling, maintenance, pilots, and flight crew to understand from their perspective what drives cancellations and delays as well as what explains the differences in performance in airports and routes

Potentially required market research study

- Survey to understand customer segments and quantify the importance of delays and cancellations on customer satisfaction

Additional analyses

Operational data

- The operational data could help us understand what is driving cancellations and delays
- A predictive model could be built to get a probability of a flight being delayed or cancelled. The key in this model would be to include variables that can be activated in time to prevent this from happening. Just knowing that a flight has a high probability of being cancelled without being able to do something would not be adding value

Customer data

- Using the customer database, we could estimate the impact on CLV of the delays and cancellations, this would provide the uplift/retention of revenue we could expect by increasing the performance
- This would ultimately help in the assessment of the feasibility of increasing performance (cost-benefit analysis) and would allow to develop a prioritization of initiatives based on returns

Interviews

- Interviews would allow to pinpoint the root cause of the issues and generate hypotheses that we could later validate using data

Potentially required market research study

- Using the survey, we could run a segmentation analysis to understand how important being on time is for the market and for the segment that AA is targeting. This could be done with factor analysis, regression, or conjoint (multinomial logit)