

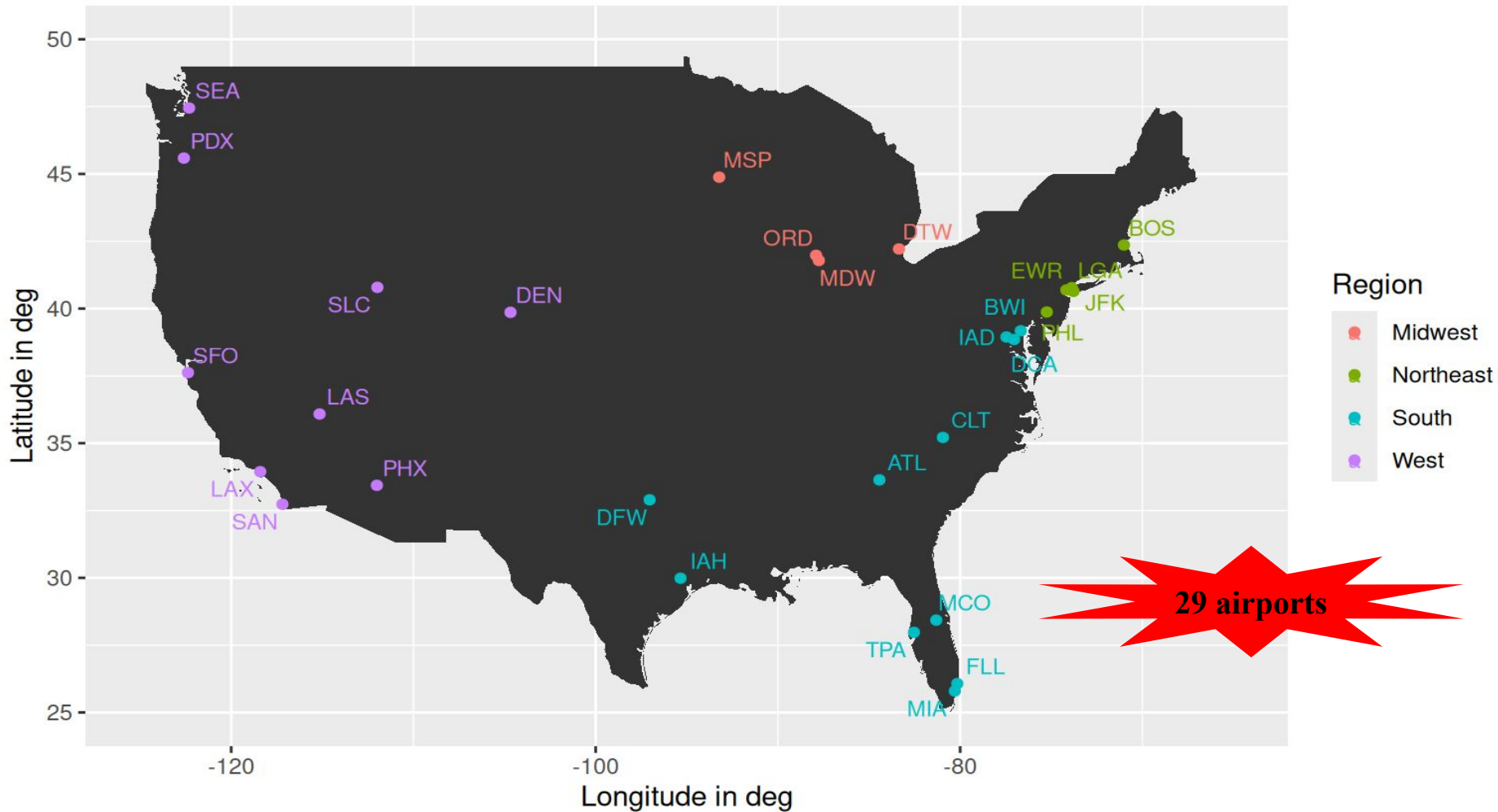
Costco Airlines

John, Eva, Pallavi, and Joseph

Introduction

- Our data set `airlines` (CORGIS Dataset Project)
 - flight, airline, and delay data
 - major US airports
 - 2003–2016, monthly
- Supplemental data sets
 - geographic coordinate data for airports (OurAirports)
 - geographic region data (US Census Bureau)
- We will analyse how 1) **geography** and 2) **offered airlines** at different US airports have an effect on flight delays, 1) **# of flights delayed** and 2) **total minutes delayed**.

Airports in the CORGIS Dataset





**summary
stats**

mean delays/airport/month = 2402 flights

mean delay minutes/airport/month = 135977 minutes

mean proportion of flights delayed = 0.20

RQ1A: "Are certain regions of the U.S. more or less likely to experience delays?"

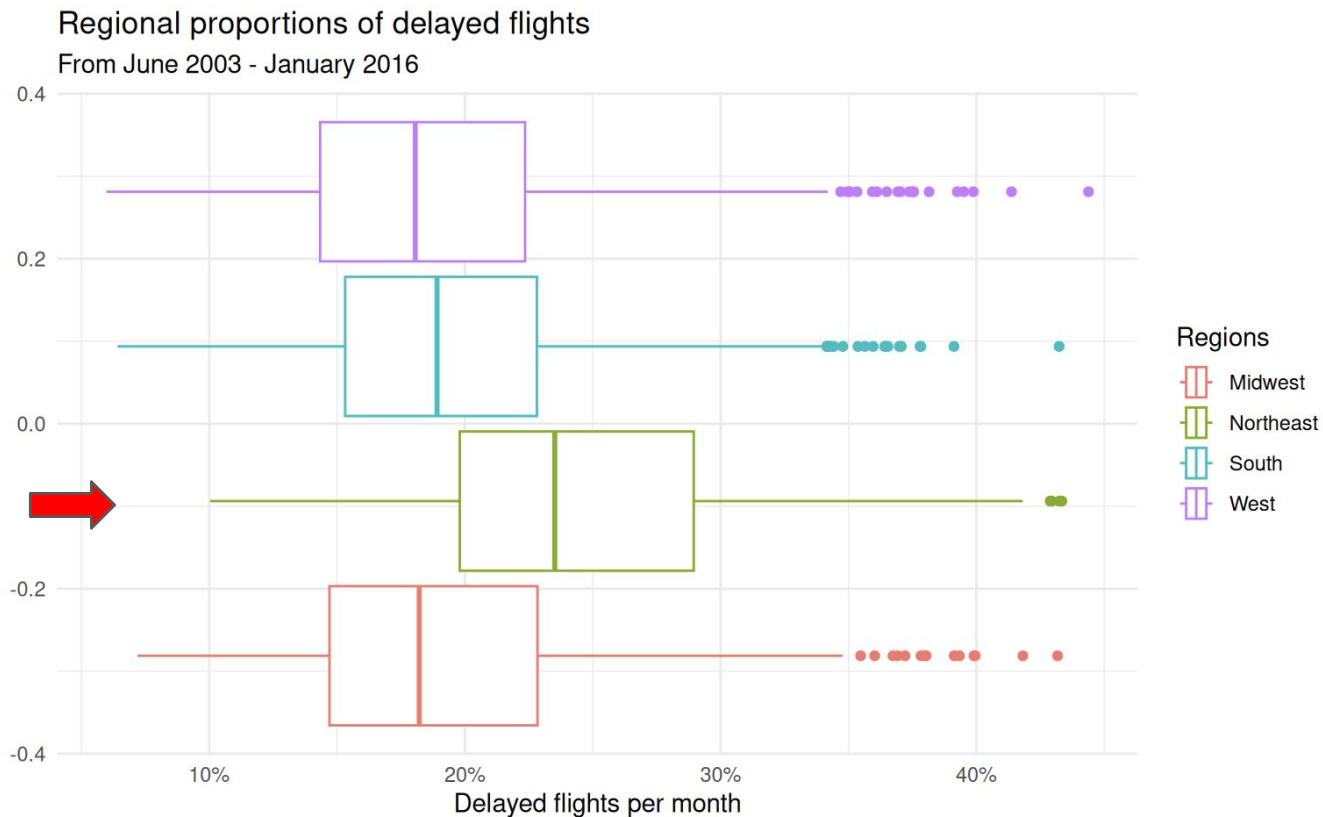


Figure 1

RQ1B: "Patterns in type of delays?"

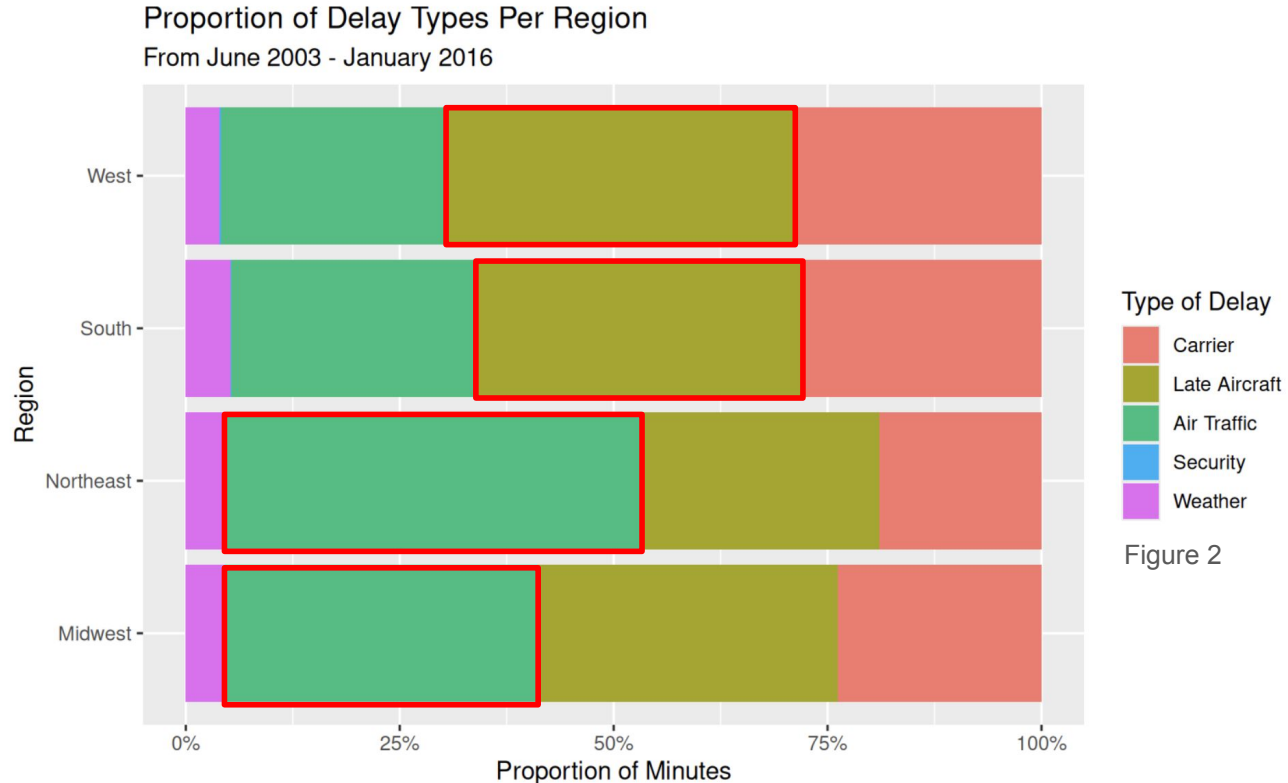
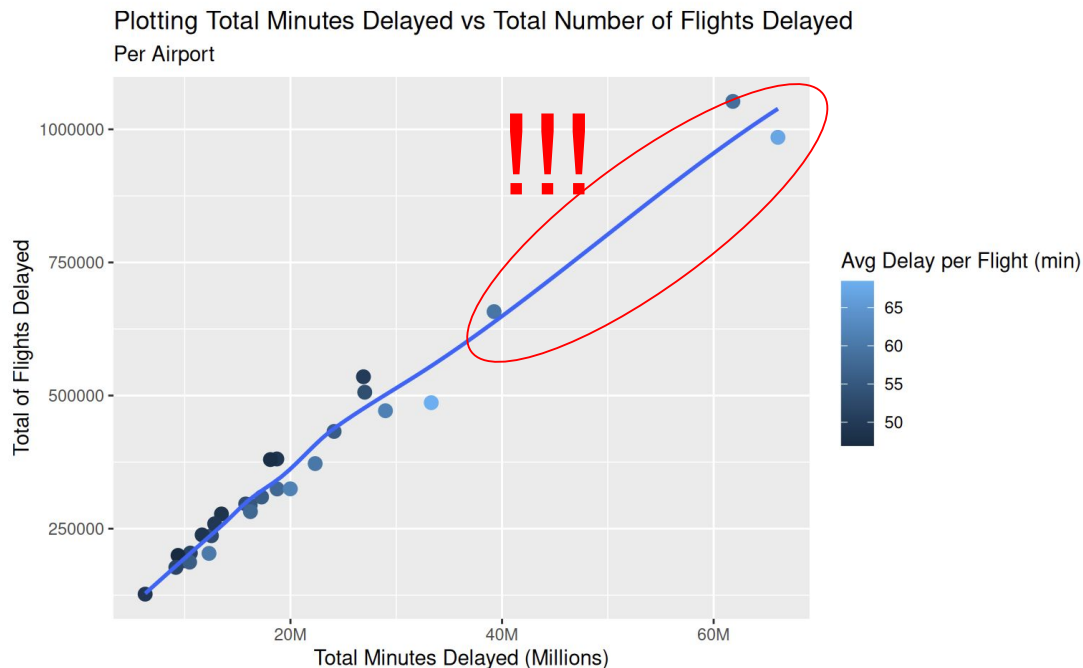


Figure 2

Note: Air traffic delays heavily influenced by weather

RQ2: "How closely do two metrics of delay (minutes delayed & flights delayed) for airports correlate to each other for in the US? If there are trends in correlation, what variables might explain them?" (Pallavi)



- Strong positive correlation between total minutes delayed and total flights delayed
- Contests **domino effect** hypothesis
- Linear growth shows that length of delay grows at relatively consistent rate regardless of if previous flight

Figure 3

RQ3: Does the proportion of delayed flights at different airports in the US vary by the airlines at each airport?

Range: 4.9%

Top 5:

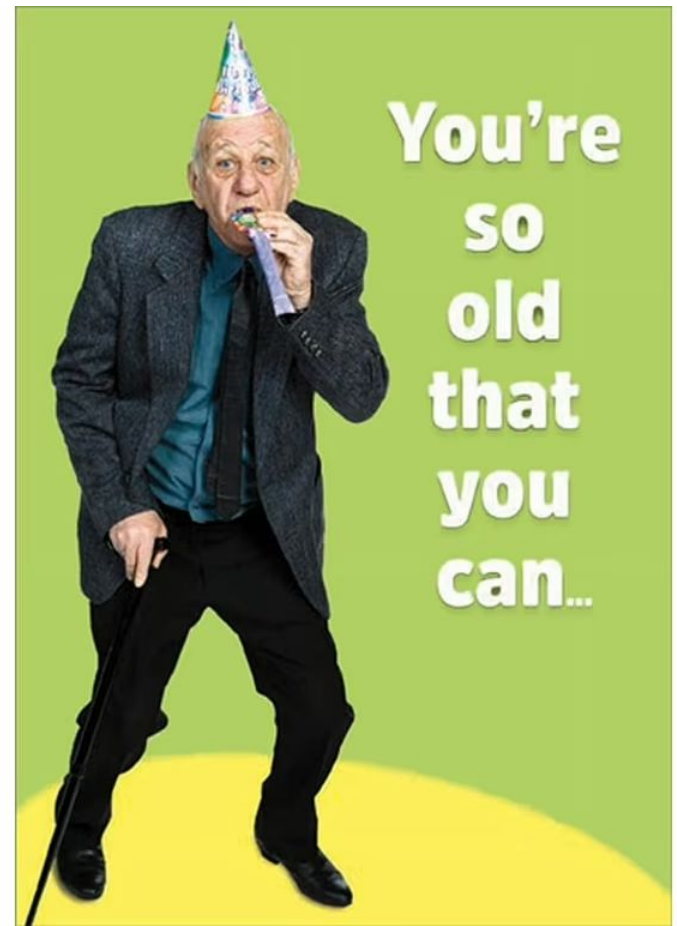
1. Mesa Airlines (24.42%)
2. Northwest Airlines (24.30%)
3. Pinnacle Airlines (23.82%)
4. Comair (23.27%)
5. Independence Air (23.06%)

Note: Metric based on airports with carriers, not carrier-specific data

carriers	Average Proportion of Flights Not On time
Endeavor Air Inc.	19.53
Spirit Air Lines	20.13
Southwest Airlines Co.	20.78
SkyWest Airlines Inc.	20.88
Hawaiian Airlines Inc.	21.01
Frontier Airlines Inc.	21.22
Envoy Air	21.39
America West Airlines Inc.	21.58
Alaska Airlines Inc.	21.81
Delta Air Lines Inc.	21.96
Virgin America	21.97
American Airlines Inc.	22.01
United Air Lines Inc.	22.04
ATA Airlines d/b/a ATA	22.07
JetBlue Airways	22.18
US Airways Inc.	22.25
AirTran Airways Corporation	22.36
ExpressJet Airlines Inc.	22.38
Atlantic Coast Airlines	22.46
Continental Air Lines Inc.	22.67
Atlantic Southeast Airlines	22.83
Aloha Airlines Inc.	22.84
American Eagle Airlines Inc.	22.97
Mesa Airlines Inc.	23.06
Northwest Airlines Inc.	23.27
Pinnacle Airlines Inc.	23.82
Comair Inc.	24.30
Independence Air	24.42

Dataset Limitations

- Data only spans 2003–2016 — likely outdated given COVID, funding shifts to DHS/TSA, and changing airline operations
- Only large airports were included, which may bias findings toward busier, higher-delay airports



Interpretation Limitations

Correlation \neq Causation

- Our scatterplot shows a strong positive relationship between minutes delayed and flights delayed, but we **cannot conclude this is causal**



Data Reliability Concerns

- Much of the data is self-reported by airlines to the FAA
- Airlines may be incentivized to attribute delays to factors out of their control
 - Weather rather than operational failures
- Questionable reporting accuracy

Future Work

- Use a data set that goes up to 2026
- Group by airport size and traffic volume when comparing carriers — current analysis treats all airports equally regardless of busyness

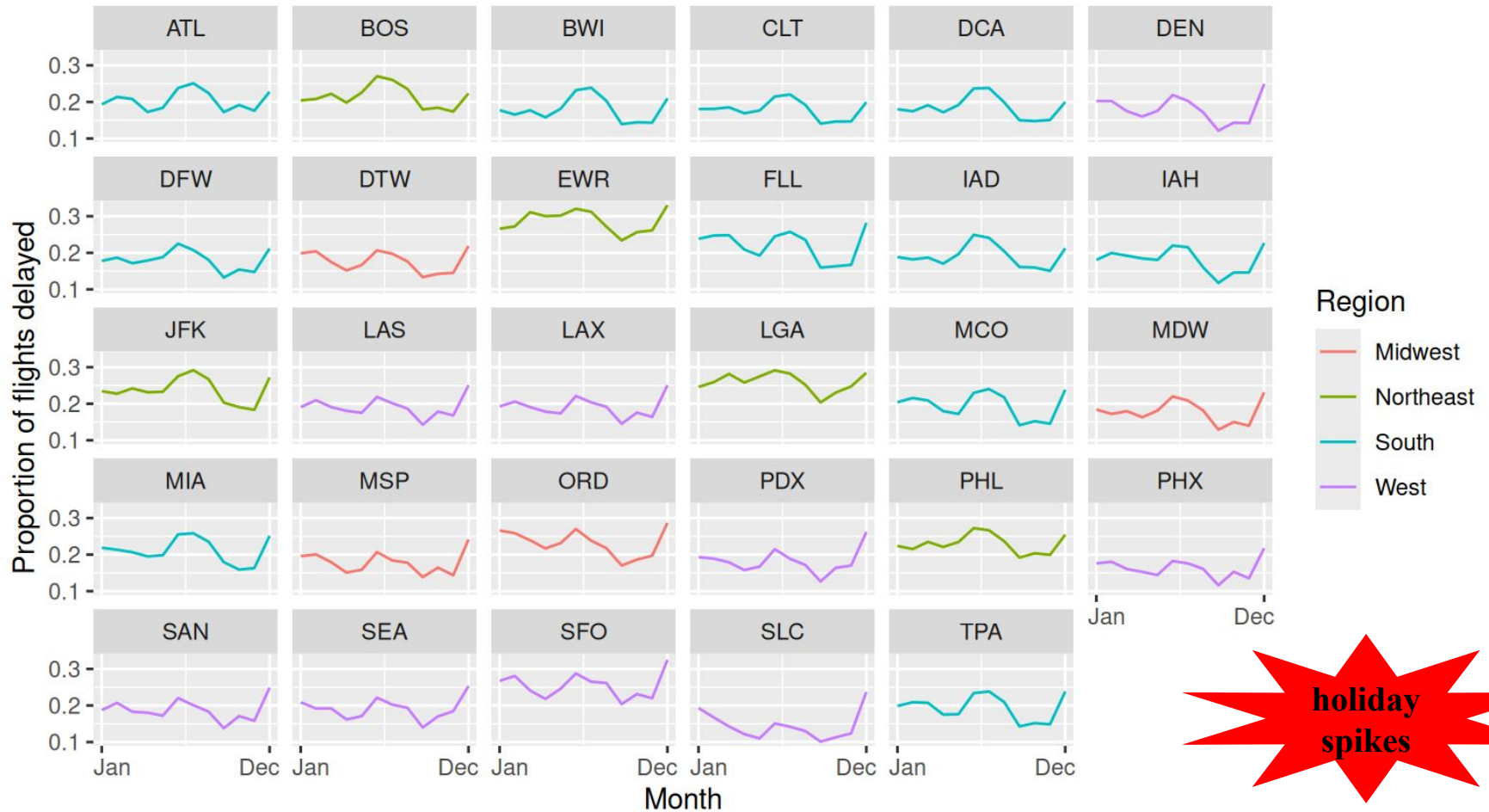
Why? To account for mass social shifts such as COVID, and to identify how much the airline itself drives delays, versus the airport it operates in

Conclusion

- Weather and late aircraft leading factor of flight delays
 - Correlates with Northeast having higher chance of flight delays
- **Linear** relationship between total amount of flights delayed and total minutes delayed.
 - Challenges the domino effect theory that minutes delayed would increase exponentially.

Flight delay proportion monthly trends at US airports

2003-2016 monthly means



Flight delay proportion monthly trends at US airports

Averaged across 2003-2016, 29 airports (95% confidence band)

