



# US DOMESTIC FLIGHTS DELAY ANALYSIS

ILYA KVYATKOVSKIY

DATA SCIENCE IMMERSIVE CAPSTONE PROJECT

GENERAL ASSEMBLY NYC, DECEMBER 2019

# PROJECT STATEMENT

---

- ▶ 777mi domestic US passengers in 2018 (5% growth to 2017)
- ▶ 8.4mi domestic flights (2.6% growth to 2017)
- ▶ Delays approach 20%

Could we anticipate a delay?

What are the triggers?

Can we reduce delays?



Data - US Bureau of Transportation Statistics, photo - Caroline Selfors/Unsplash

# PROJECT WORK FLOW

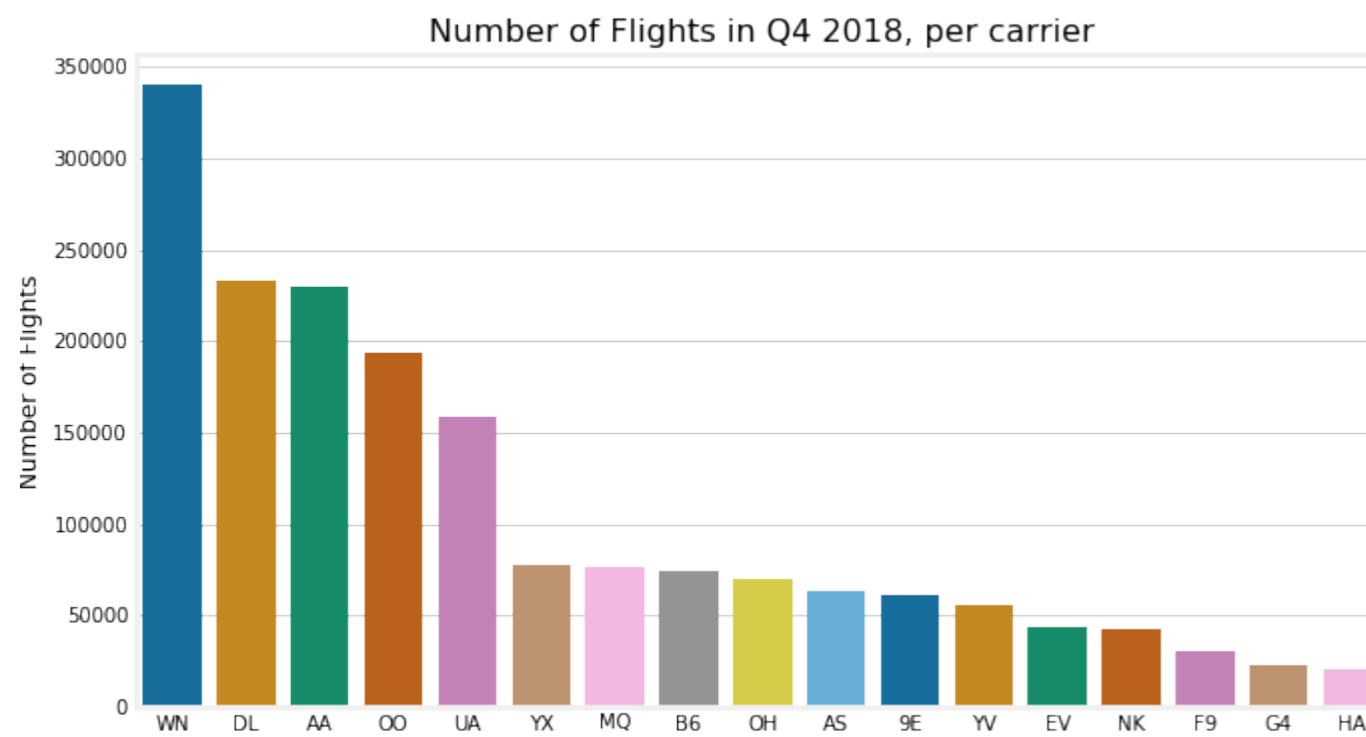
---

- ▶ Domain Analysis
- ▶ Data Exploration
- ▶ Modeling
- ▶ Results Evaluation
- ▶ Future Steps



Olivier Sjorström/Unsplash

# US DOMESTIC FLIGHTS Q4 2018



**Total US domestic flights 2018:**

Year 8.391.400

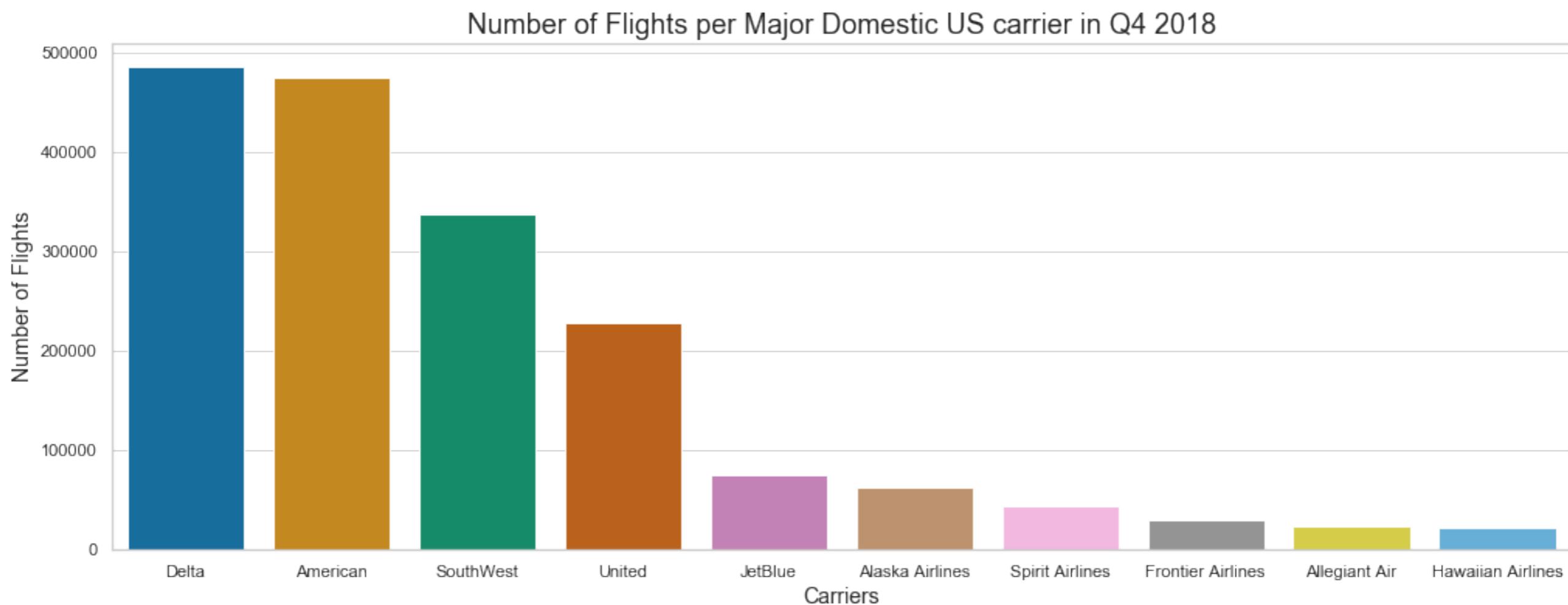
Q4 1.796.121

Monthly ~700.000

Daily ~ 23.000

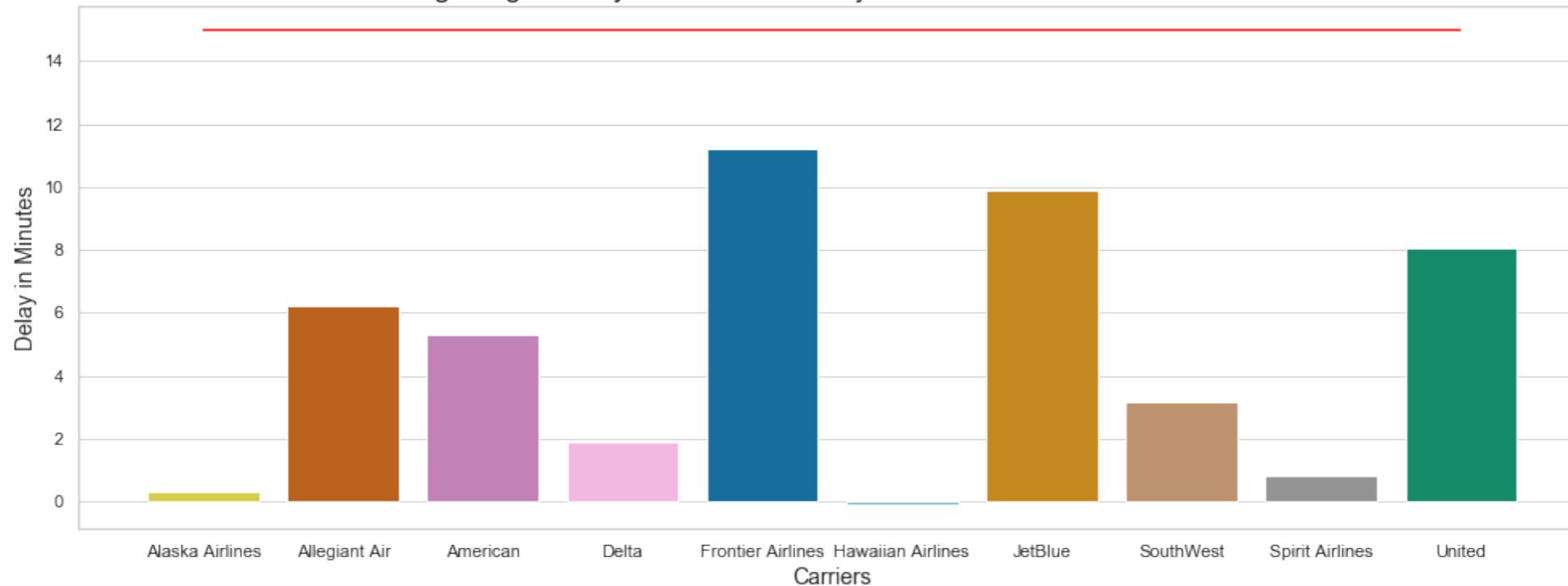
Hourly ~ 1000

Per minute ~ 16

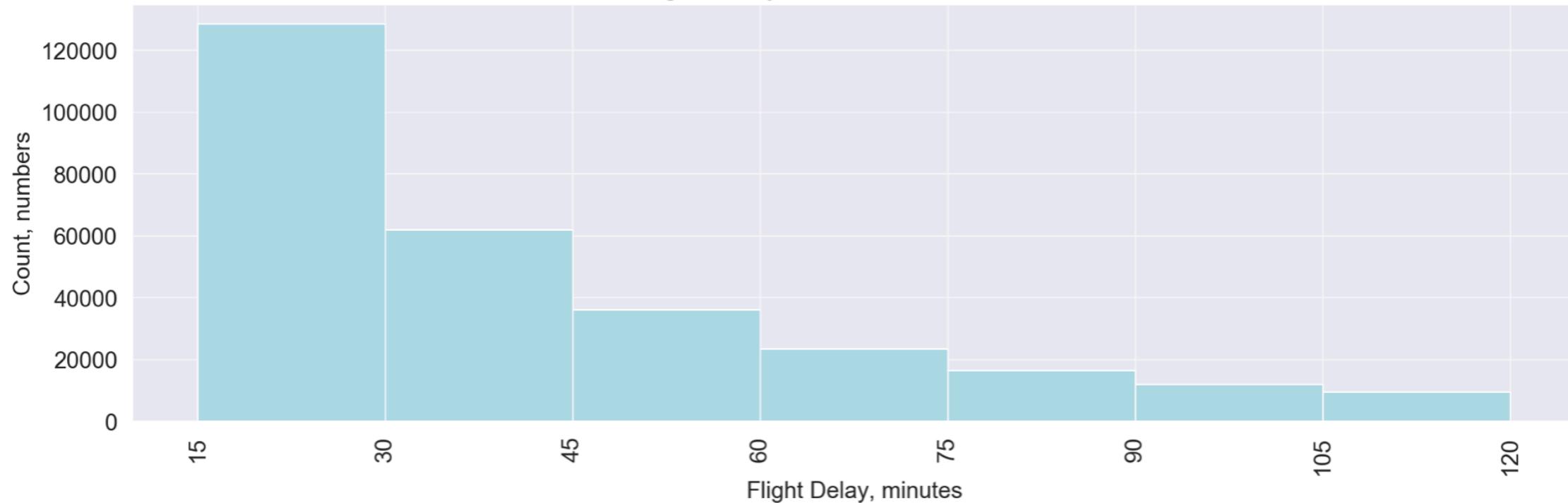


# DOMESTIC FLIGHTS DELAYS Q4 2018

Average Flight Delay in Minutes for Major Domestic US carrier in Q4 2018



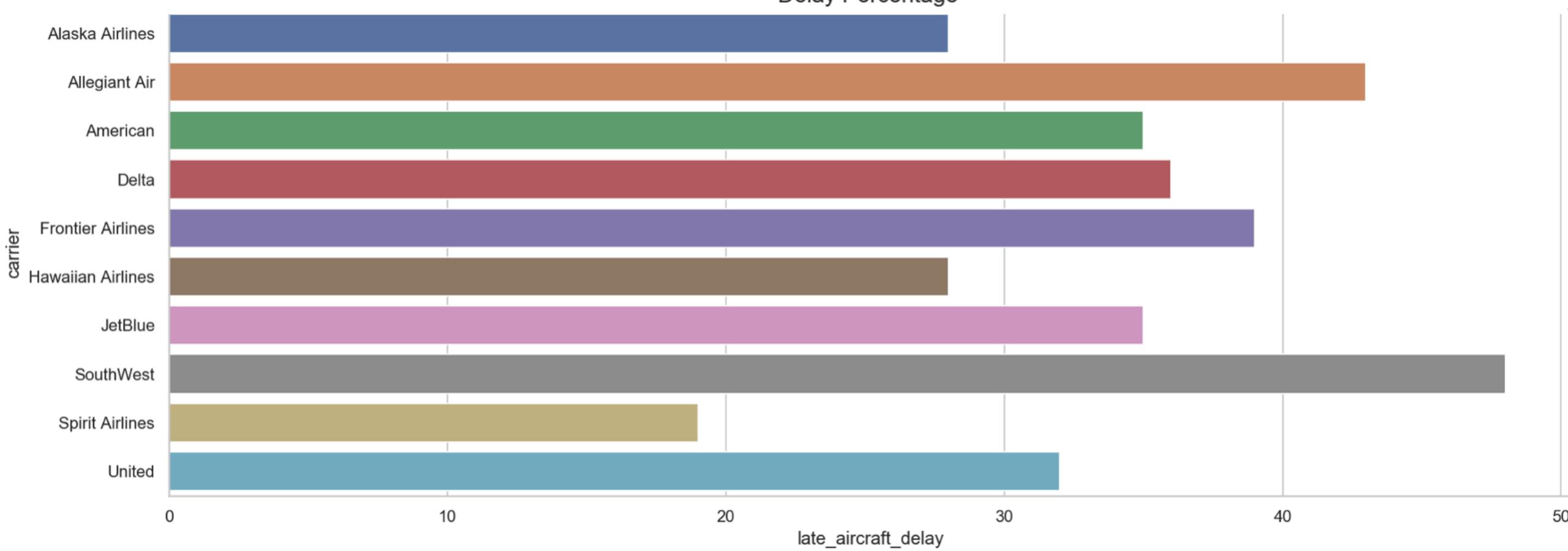
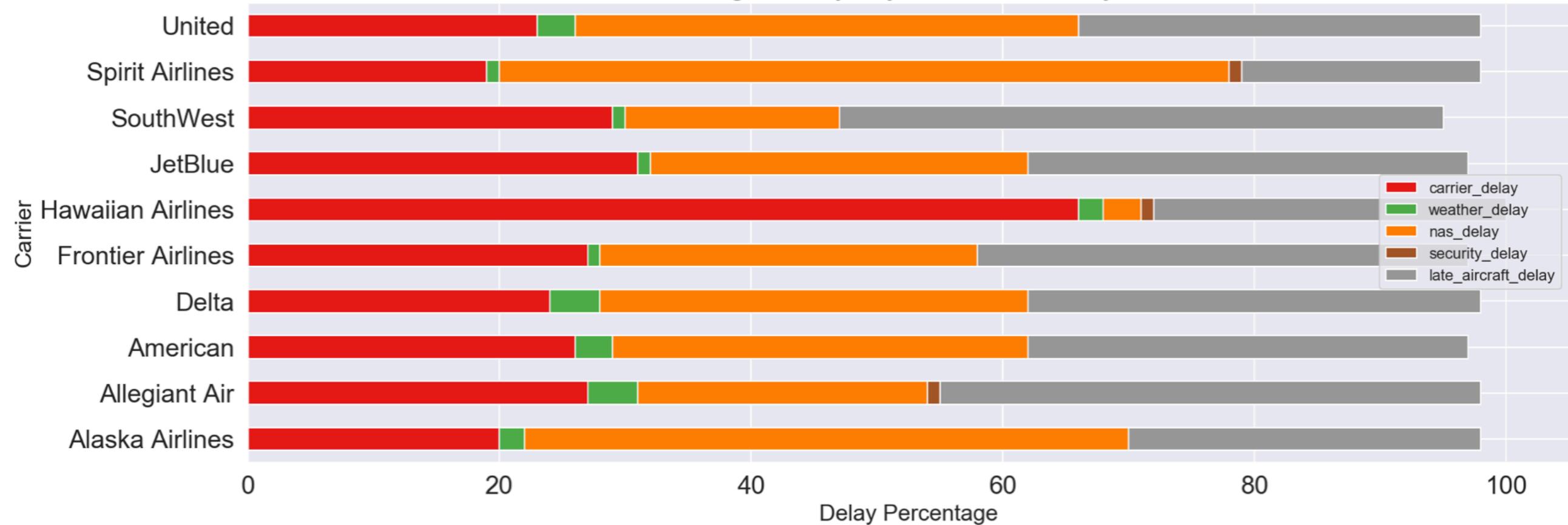
Total Flight Delay between 15 and 120 Minutes



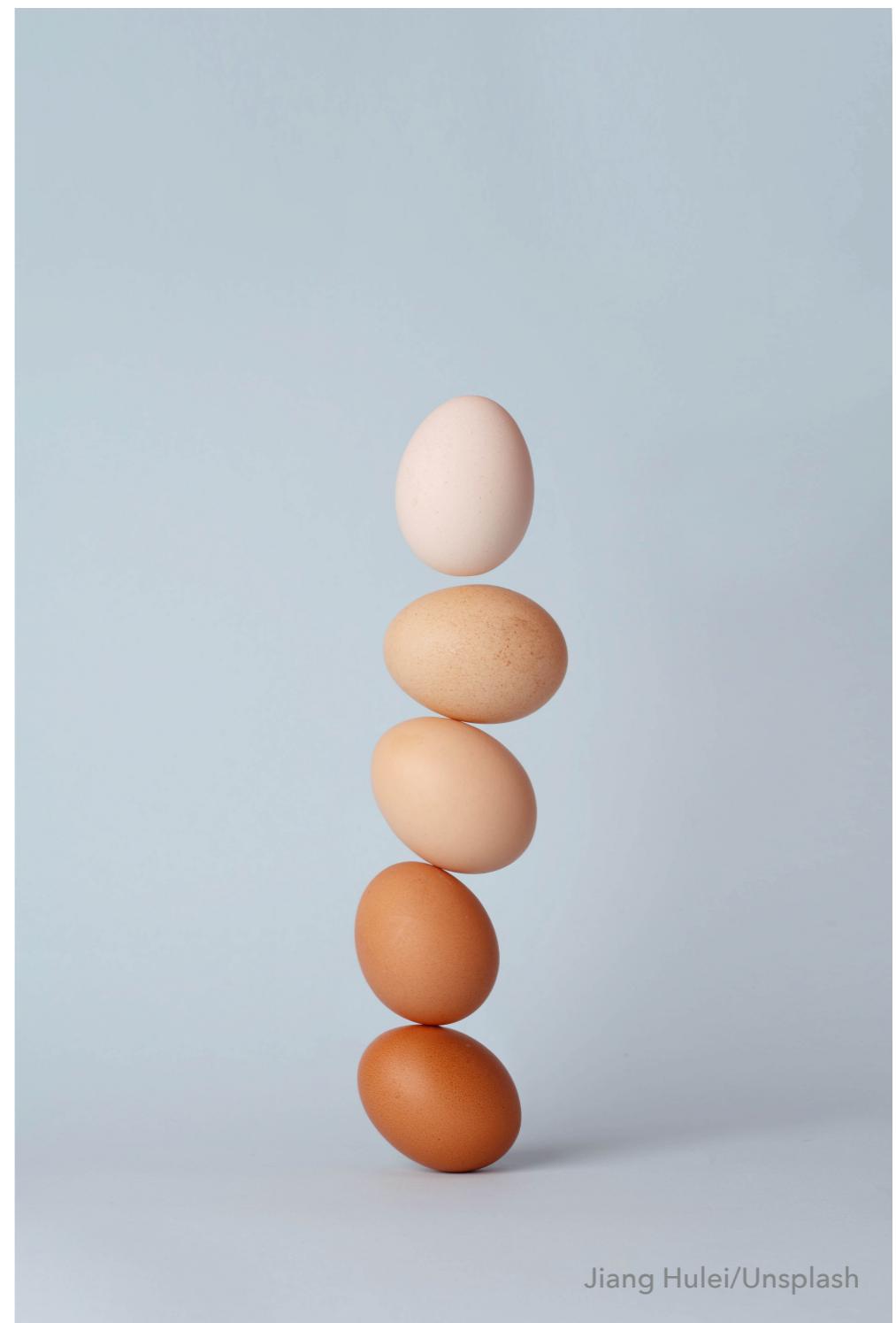
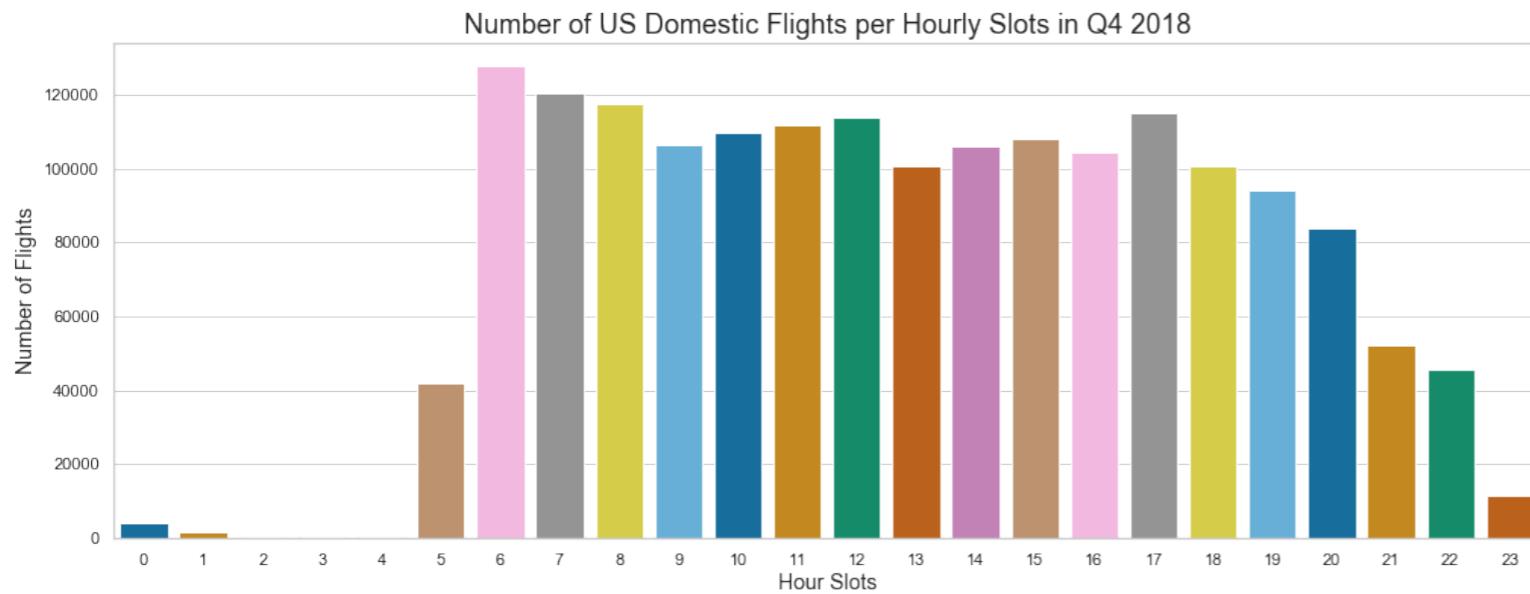
**Mean Flight Delay in Q4 - 4.25 minutes, Delayed Flights ~ 18%**

# DOMESTIC FLIGHTS DELAYS Q4 2018

Flight Delays by Reason of Delay



# FEATURE ENGINEERING AND PRE-PROCESSING



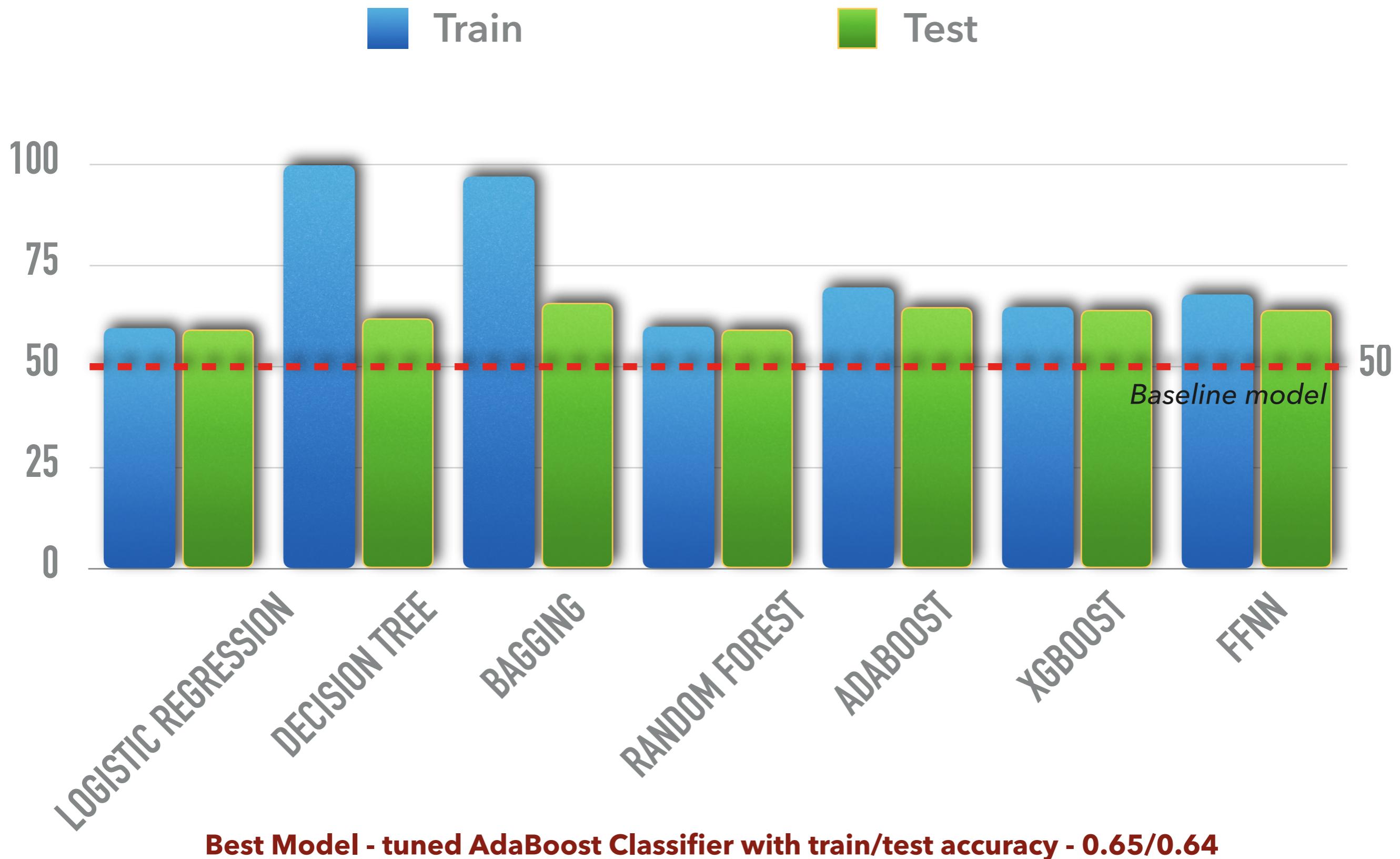
Original Dataset Size 1.796.121x111

Cleaned Dataset Size 1.777.932x27

(1% observations dropped)

Modeling Dataset **622.232x19**

# CLASSIFIERS



# REGRESSION MODEL

---



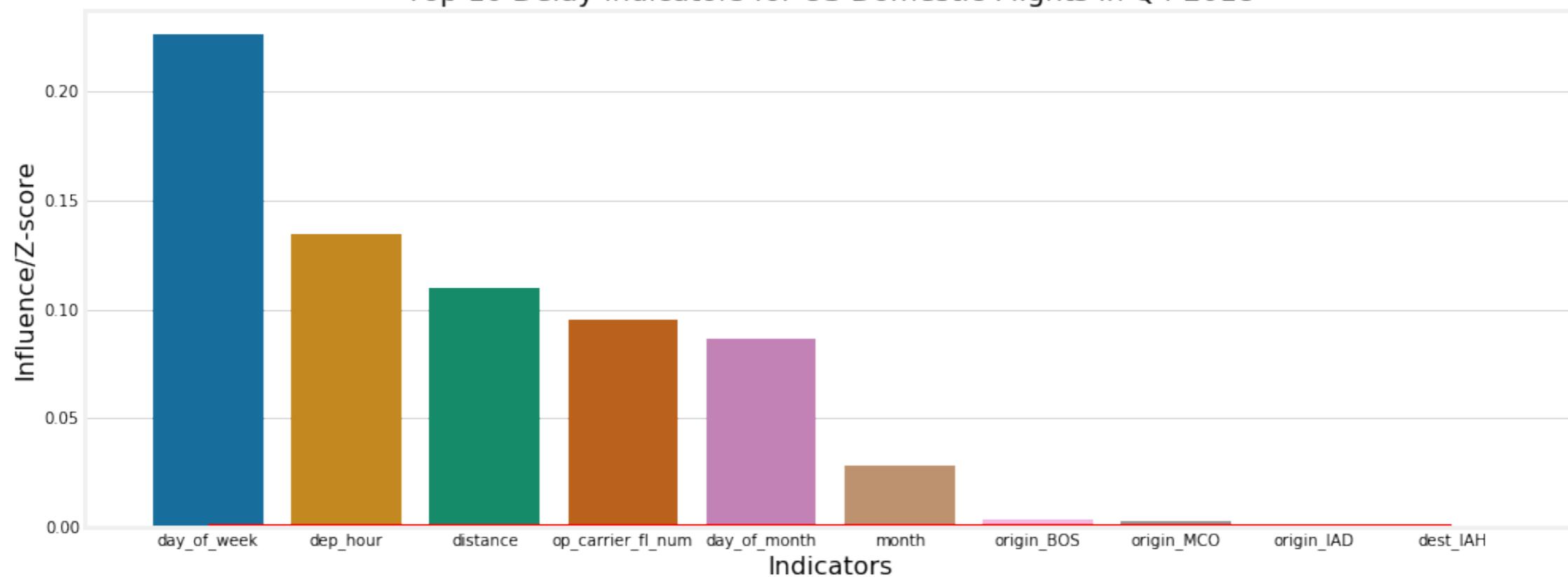
**FFNN Regressor performance predicting a delayed flight's delay duration:**

**RMSE ~ 25 minutes**

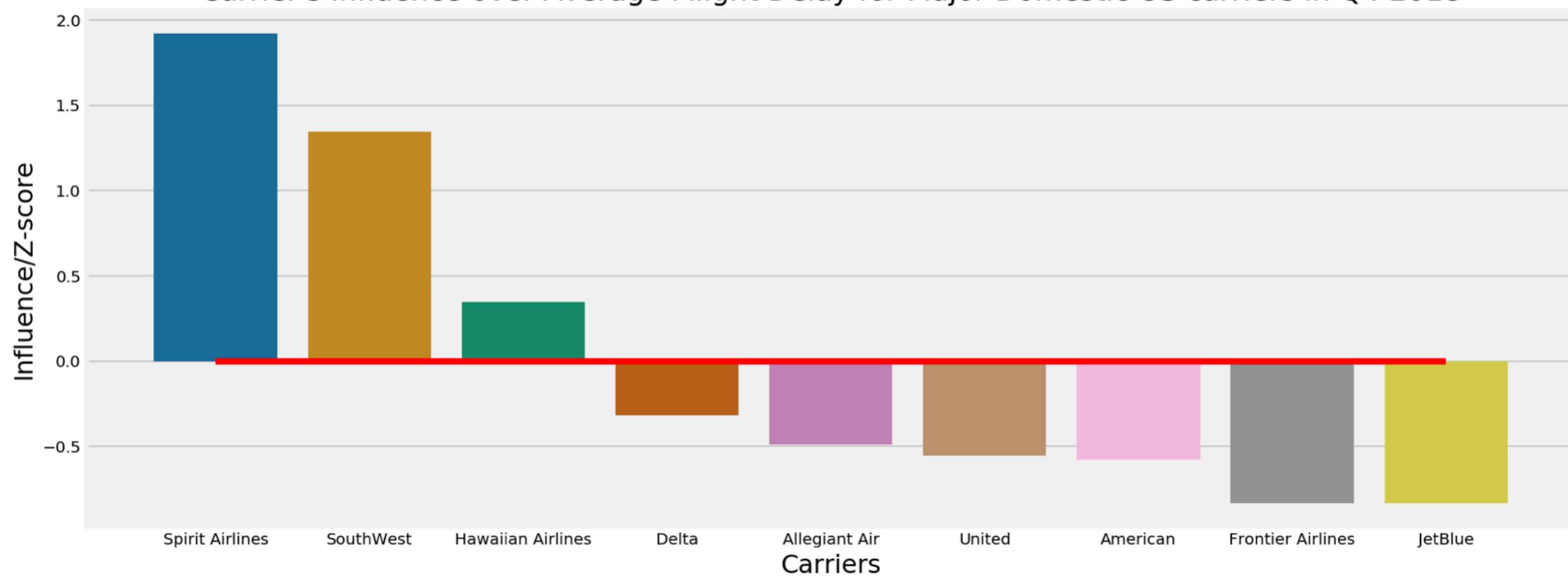
**Mean delay ~40 minutes**

# STRONGEST PREDICTORS AND AIRLINE INFLUENCE

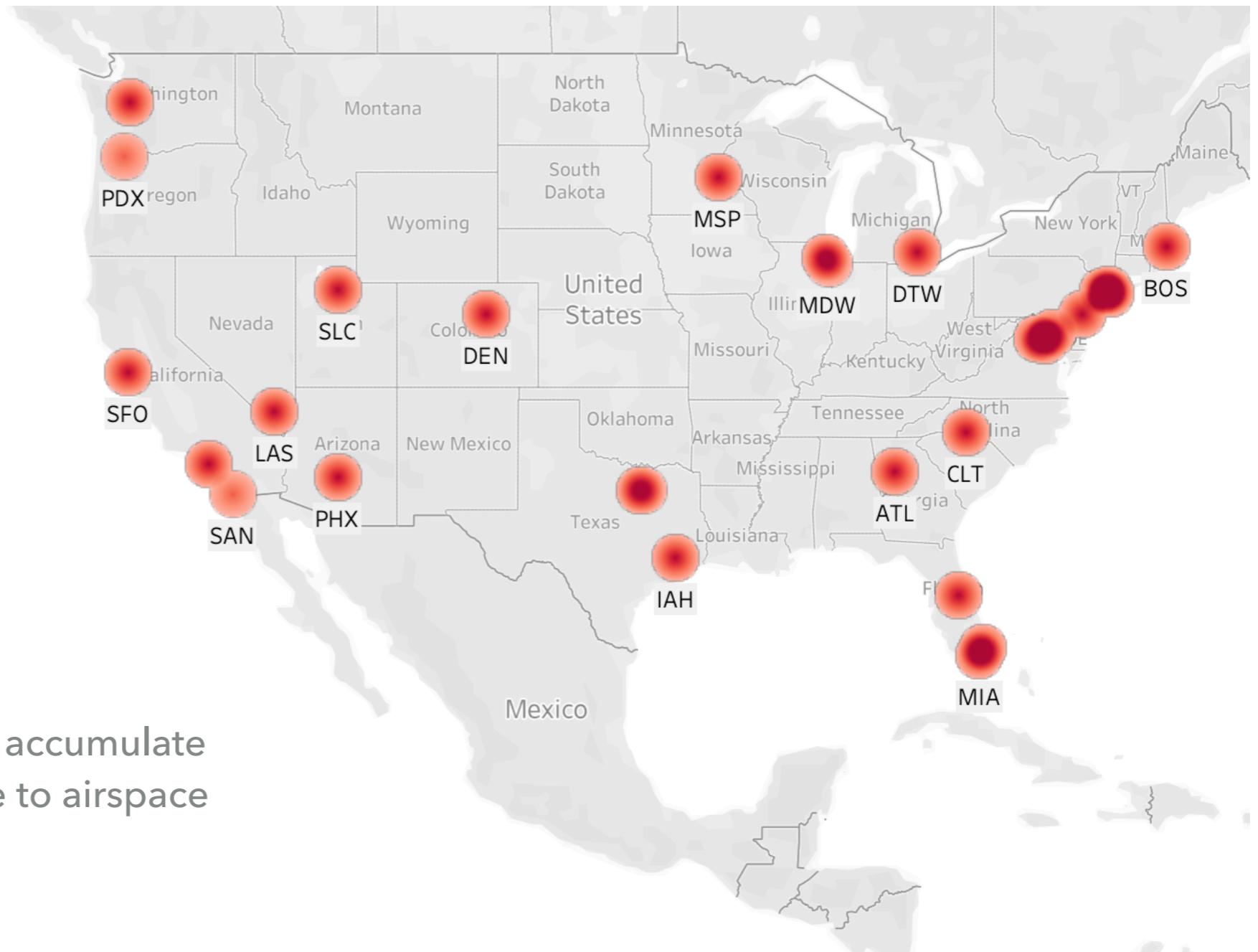
Top 10 Delay Indicators for US Domestic Flights in Q4 2018



Carrier's Influence over Average Flight Delay for Major Domestic US carriers in Q4 2018

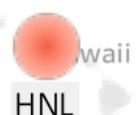


# AIRPORT INFLUENCE



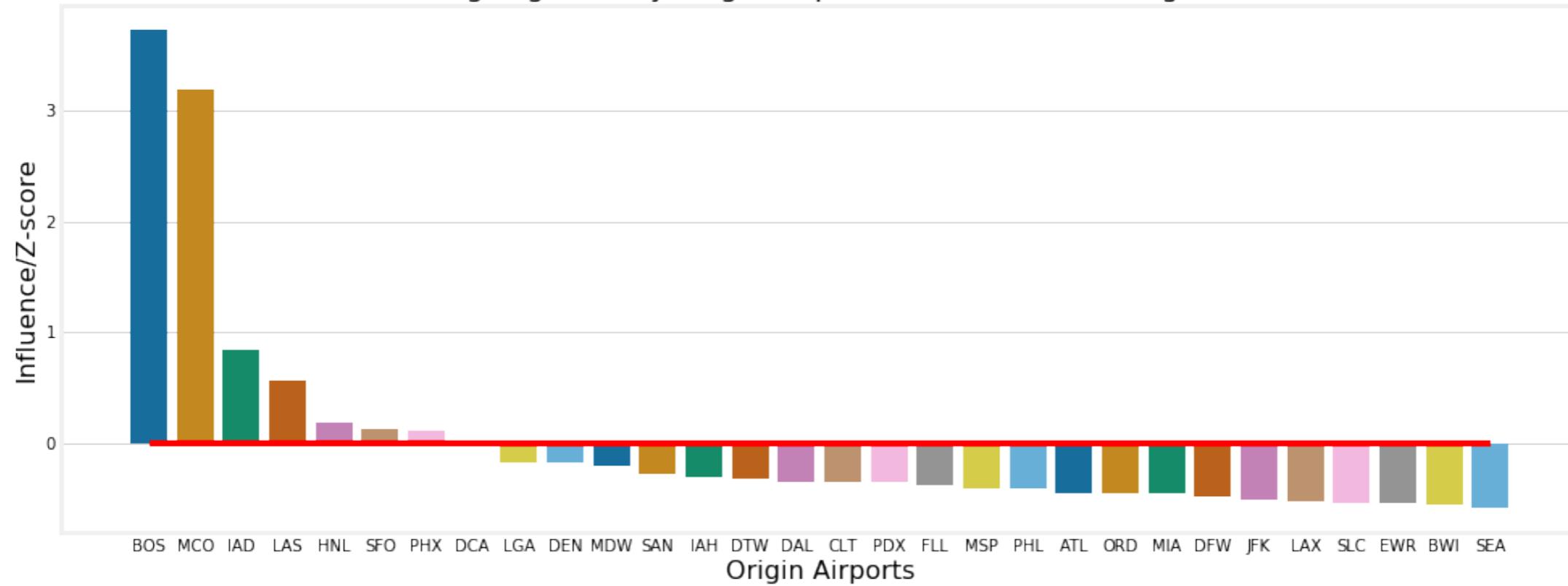
## Observations:

- ▶ Agglomerated airports accumulate each other's delays due to airspace constraints
- ▶ Weather impact
- ▶ Geographical position (sea, mountains)
- ▶ Infrastructure concerns

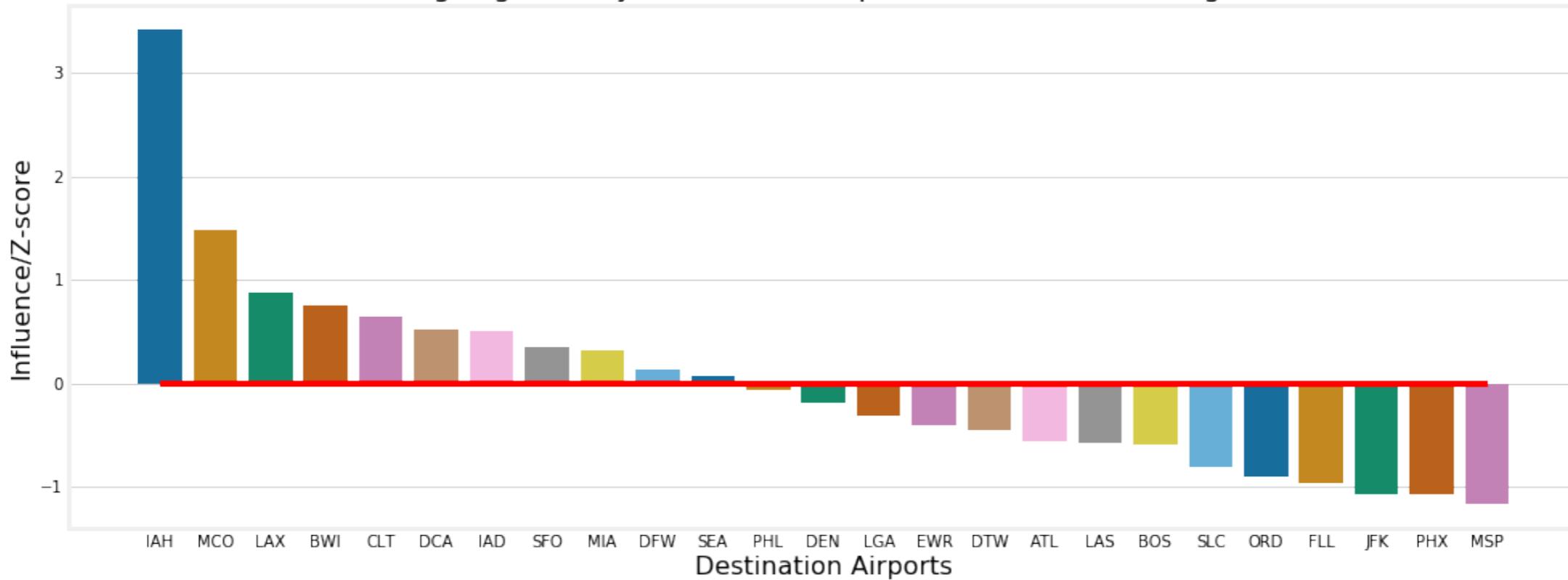


# AIRPORT INFLUENCE

Most Influencing Flight Delay Origin Airport for US Domestic Flights in Q4 2018



Most Influencing Flight Delay Destination Airport for US Domestic Flights in Q4 2018



# CONCLUSIONS AND FUTURE STEPS

---

- ▶ Further data exploration with feature engineering
- ▶ Classifiers tuning
- ▶ Regression models
- ▶ Weather factor
- ▶ International flight data incorporation
- ▶ GPU usage

