



## Methodology

- 1 Data
  - Flight data from Bureau of Transportation
    Statistics
  - Weather data from DarkSky API

#### Modeling

- Logistic Regression
- Random Forest Classifier

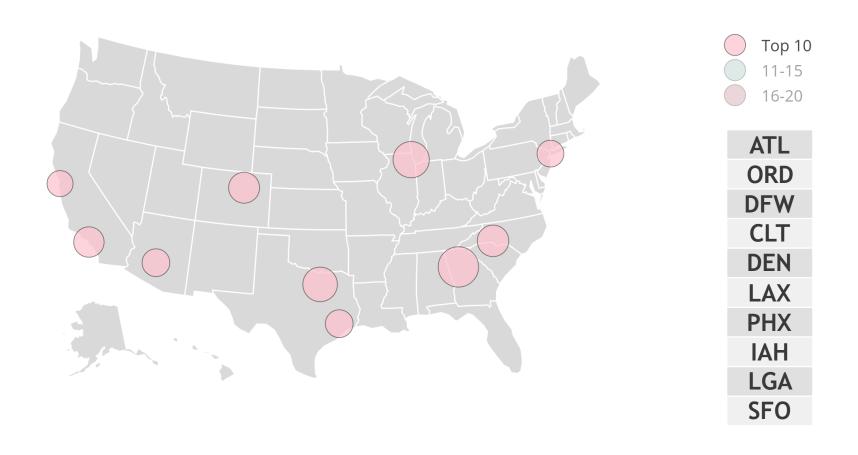
Results and

**Conclusions** 

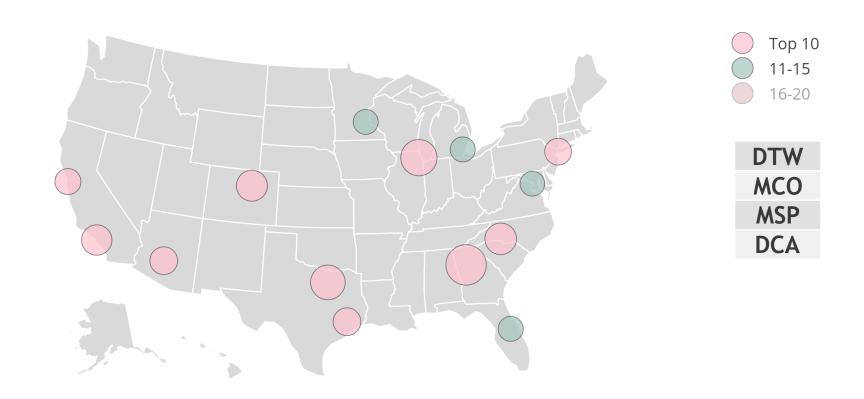
XGBoost

EDA and Feature Engineering and Selection

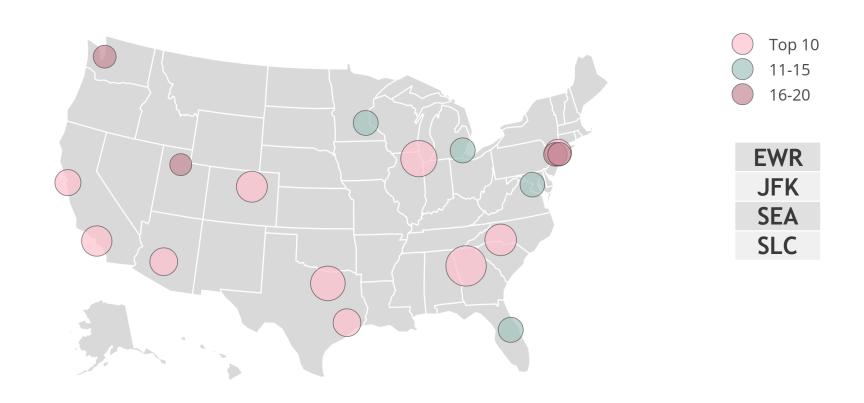
Top 20 Busiest Airports by Traffic



Top 20 Busiest Airports by Traffic



Top 20 Busiest Airports by Traffic



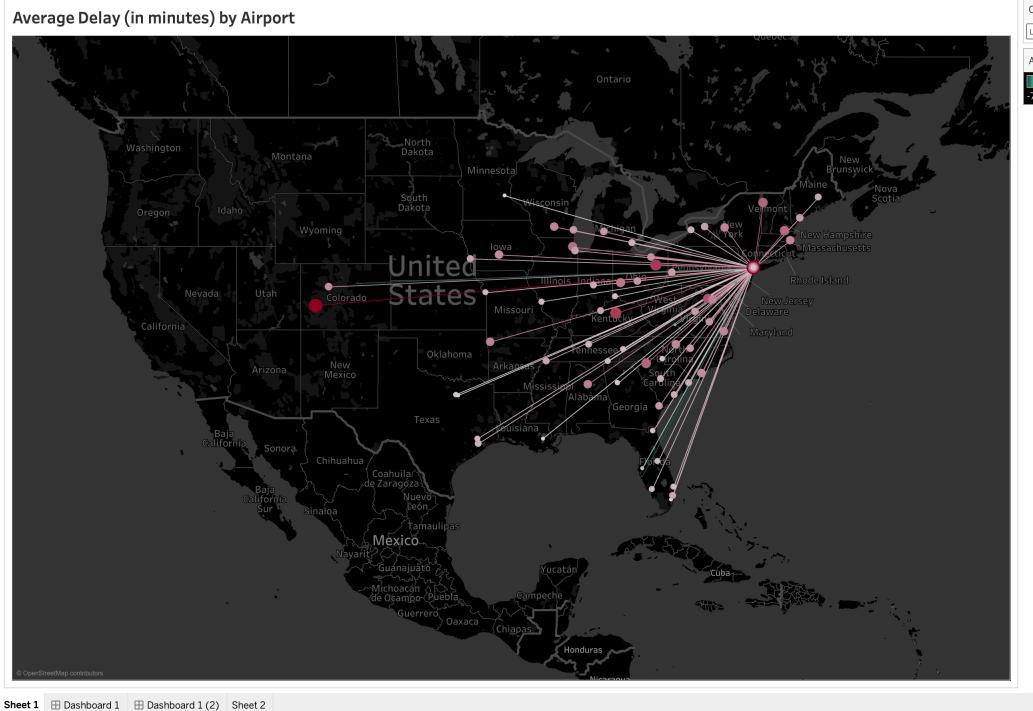
### Results



F1 Score: 0.71 Accuracy: 0.90







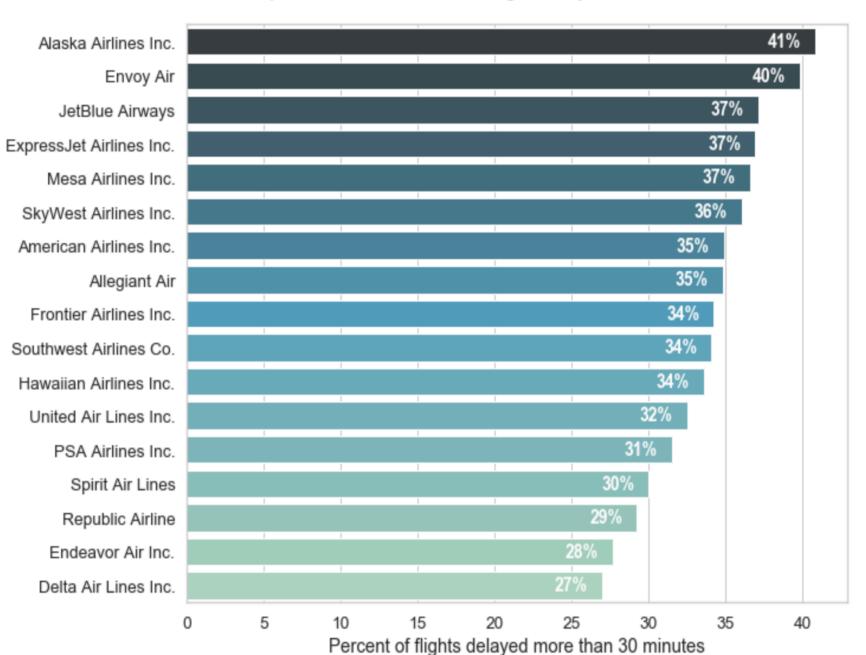
# Thank You

....

# Appendix

.....

Top 20 Airlines and Percent of flights delayed more than 30 minutes





Top 20 Airports and Average Delay in Minutes 31.16 SFO 28.05 ORD 25.56 MSP 21.63 **EWR** 20.93 BOS 19.42 SEA 19.08 DFW LAS 19.02 18.86 LGA 16.96 DTW 16.95 LAX 16.75 DEN 16.49 IAH 16.38 JFK 15.99 SLC 15.53 DCA 15.30 MCO .....I....... 15.28 PHX 13.38 ATL 1111 12.84 CLT 15 5 10 20 25 30 0 Average Delay in Minutes