

Bay Area

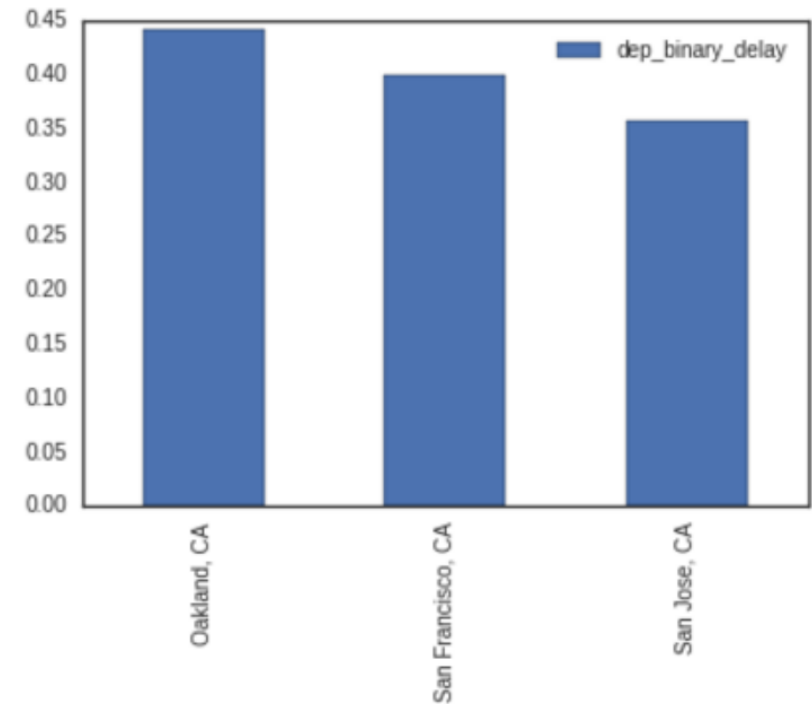
Flight Delays



HAL WRIGHT
GA – DSI #1
JUNE 216

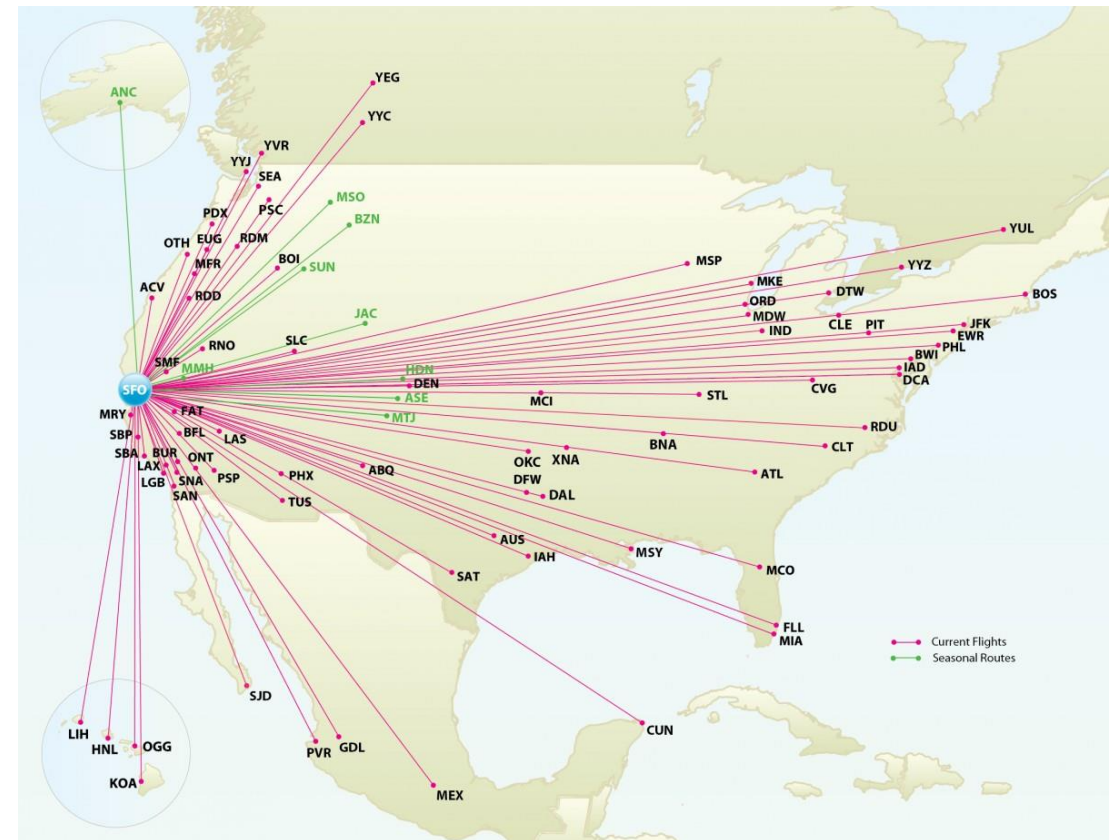
Project Description and Goals

- 39% of all US Domestic Flights were delayed in 2015
- My Goal in the project was to model flight delays to understand what factors influence departure delays the most
- Then ultimately to create a tool that would help consumers avoid flight delays



Describing the Data / Dataset

- The Research and Innovative Technology Administration (RITA) has published on time performance data for all US domestic flights from January 1987 to January 2016.
- Dataset has over 168 million flights, 30+ airports with 160+ columns
 - Paired down to 21 relevant columns for departure delay
 - 185k Flights
- My focus was on comparing 3 major airports in the bay area
 - 21 common destinations from those airports



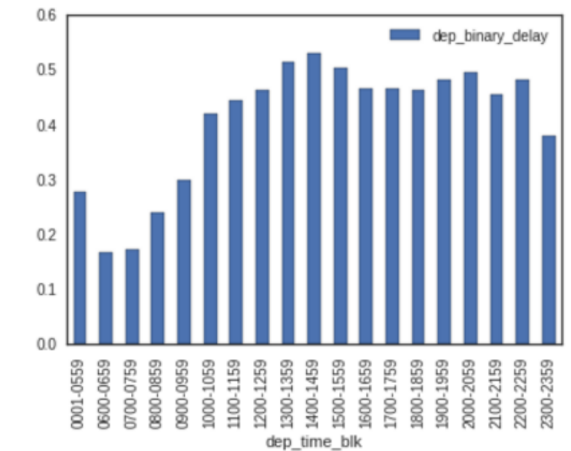
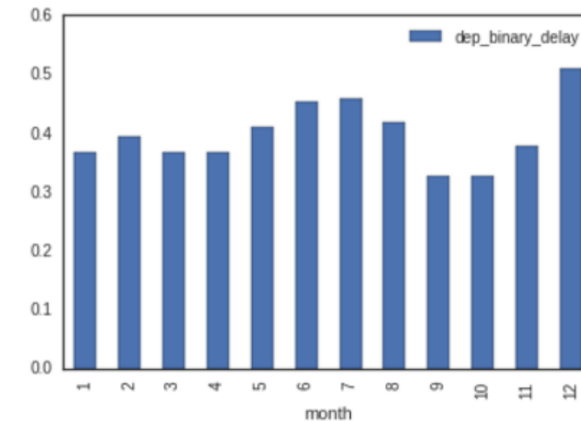
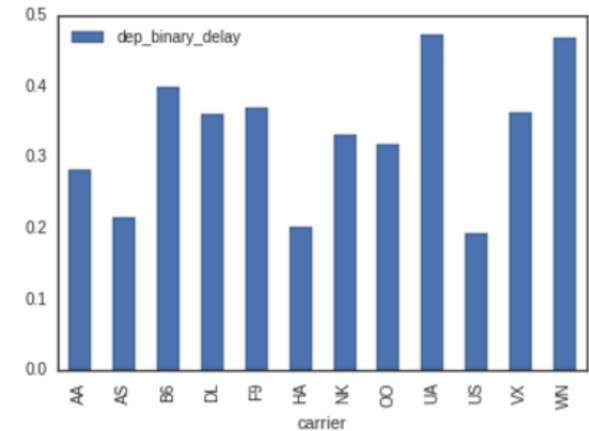
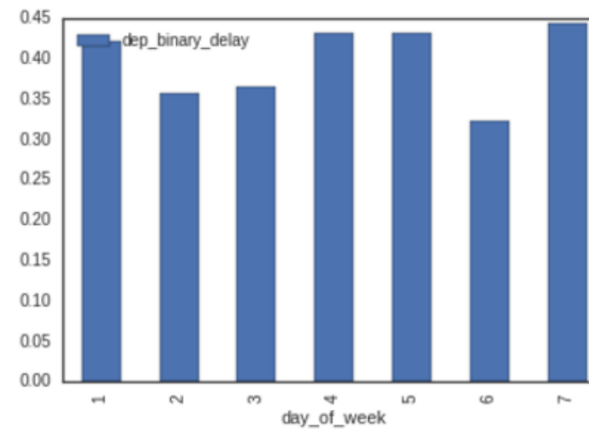
Selecting Predictors (EDA)

- Categorical Variables

- Origin Airport
- Month
- Day of Week
- Time of Day
- Carrier

- Continuous Variable

- Distance

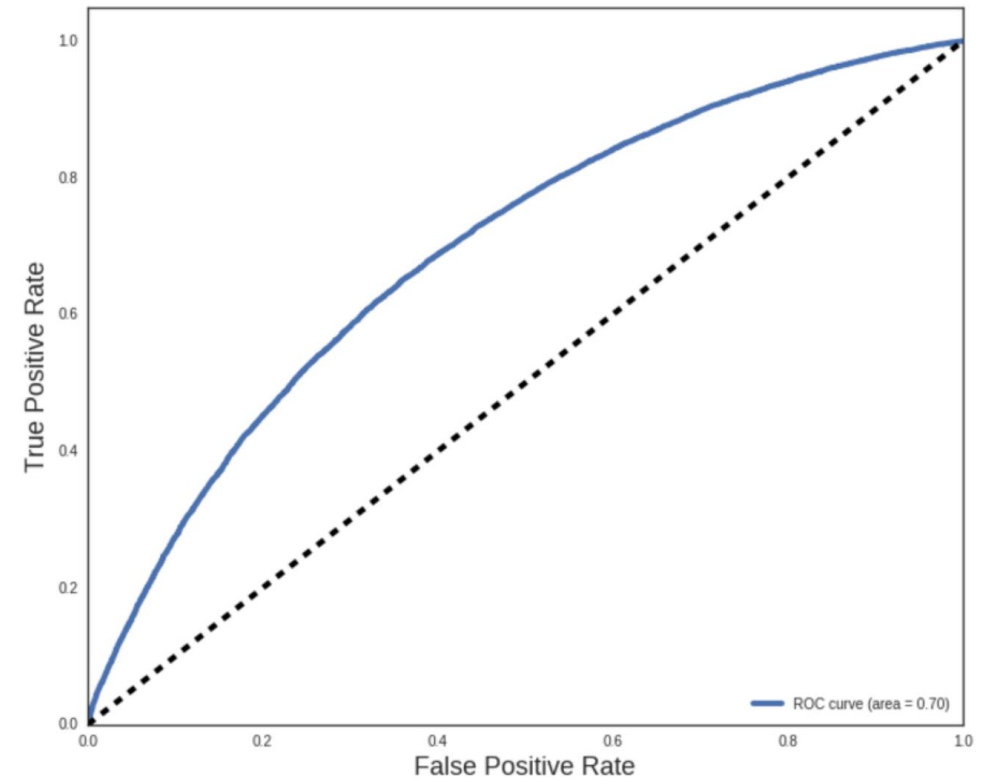


Best Model & Performance

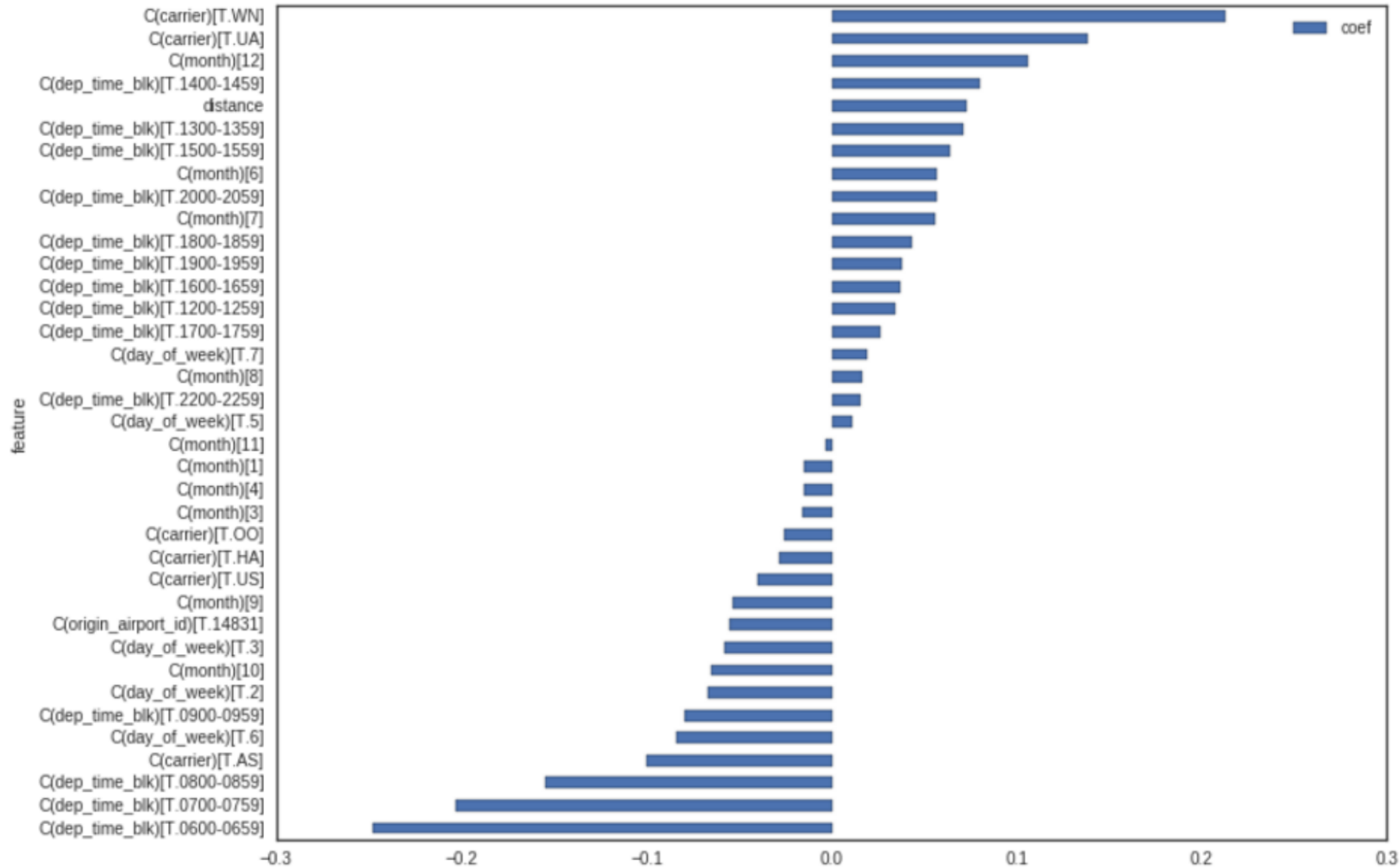
- Tried Three Models:
 - SGD Classifier, Decision Tree Classifier, Random Forest Classifier
- Stochastic Gradient Descent Classifier was best performer

	Predicted On Time	Predicted Delayed
Actually On Time	28659	4740
Actually Delayed	13974	7760

	precision	recall	f1-score	support
On Time	0.67	0.86	0.75	33399
Delayed	0.62	0.36	0.45	21734
avg / total	0.65	0.66	0.64	55133



Coefficient Interpretation



User Input

× Dest. City: Atlanta, GA ▼

Airline: DL ▼

Month: 1 ▼

Day of the Week: 4 ▼

Hour of Flight: 2300-2359 ▼

Submit Info

Based on your selection there is a 35.06% chance of delay at the San Jose Airport.
Based on your selection there is a 38.66% chance of delay at the San Francisco Airport.
Based on your selection there is a 38.2% chance of delay at the Oakland Airport.

I suggest you fly out of San Jose to avoid delays

Thank You
