

# Restaurant Roulette

## Damien Ramunno-Johnson



# Goal: Eat in Safe Restaurants

“Widespread violations found at Boston’s food spots” – Boston Globe (2015)

“Health and sanitary code violations found at nearly half of Boston’s eateries” – Fox25 (2015)

“Restaurants in Boston get infrequent inspections” – CommonWealth (2012)



# Restaurant Roulette

Will the restaurant pass its health inspection?

Check restaurant



# Serious violations found in Boston Restaurants



- Data was pulled from Boston, and Yelp.
- Less than half of health inspections pass.



# Pipeline

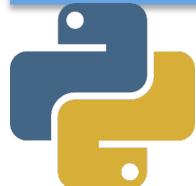
Clean and  
combine Yelp  
Data

Feature  
Engineering

Train classifier  
(Random Forest)

Predict  
current  
inspection  
status

Deliver  
results on  
web server



pandas

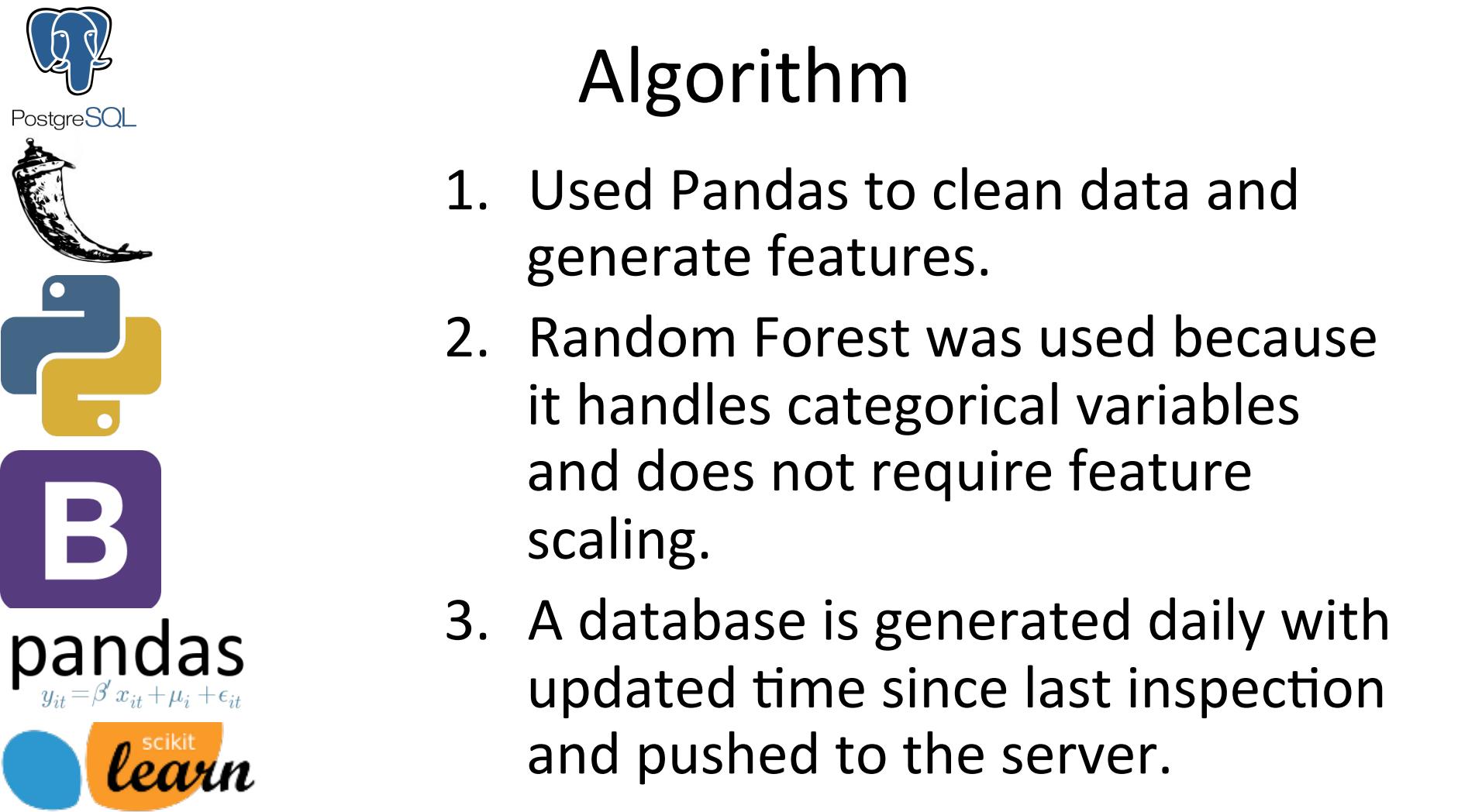
$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$

The scikit-learn logo, featuring a blue circle and an orange rectangle with the word "learn" in white.

The PostgreSQL logo, which is a stylized blue double-headed arrow.

The BeeWare logo, featuring a large white letter "B" on a purple square background.

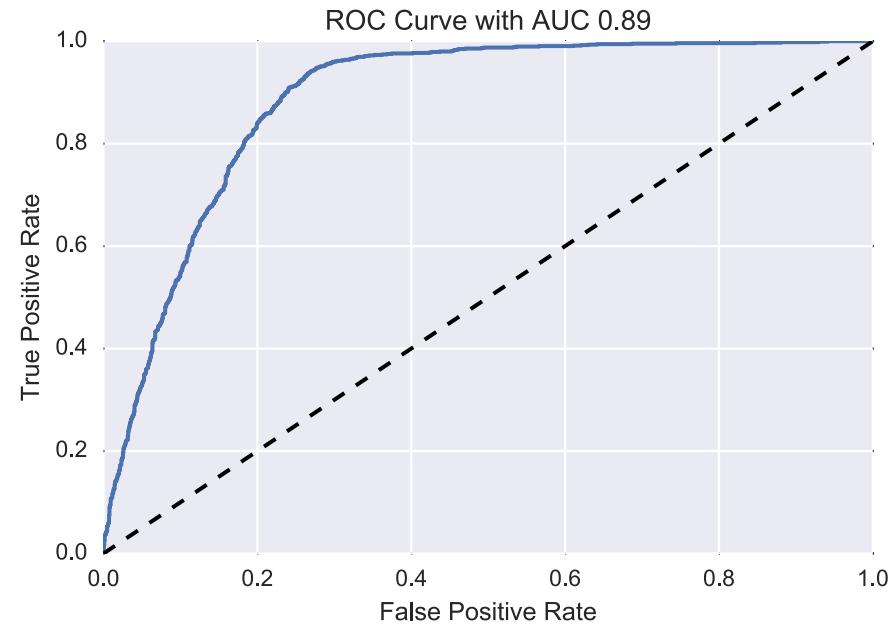
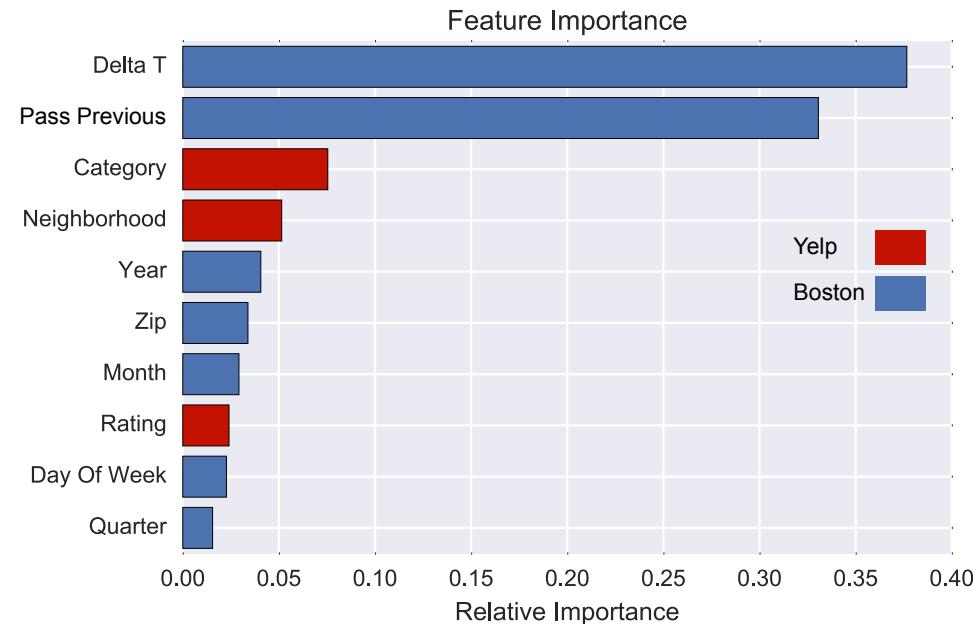




1. Used Pandas to clean data and generate features.
2. Random Forest was used because it handles categorical variables and does not require feature scaling.
3. A database is generated daily with updated time since last inspection and pushed to the server.

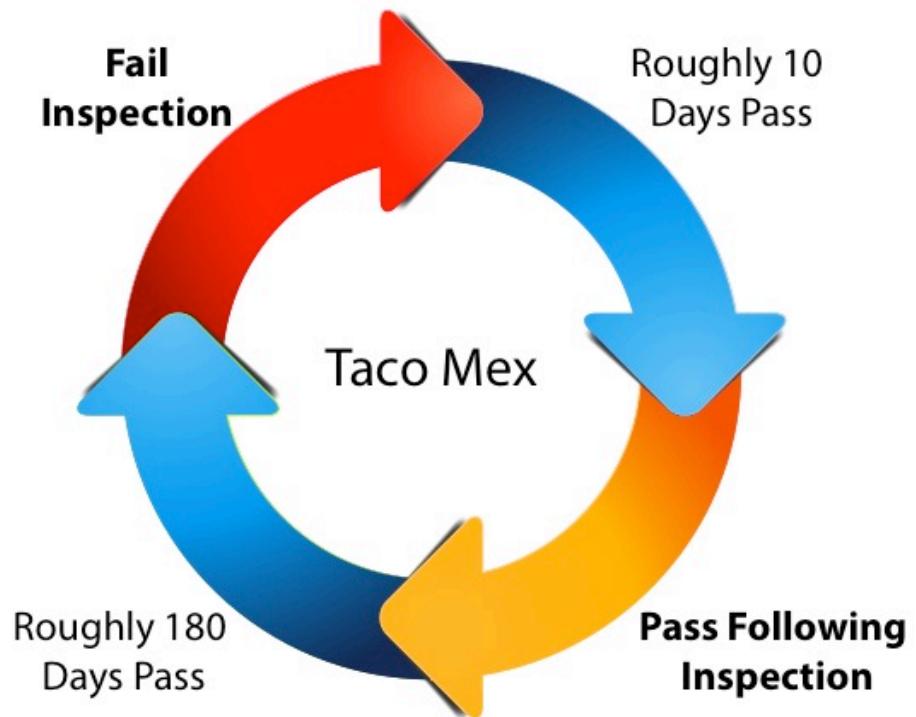
# Model Performance

- Validation was done with a 20% holdout.
- Accuracy of 81%, and AUC of .89



# A Pattern Emerges

- Taco Mex has a 75% chance of failing
- A pattern of failing and passing becomes clear



# About Me

