

The background of the slide is a stylized landscape. It features a light yellow-green sky with two white, fluffy clouds. In the foreground, there are two large trees with brown trunks and green, rounded canopies. Behind the trees are dark green, rolling hills. The overall style is simple and illustrative.

SAVE NEW YORK CITY TREES

Tree Census ! 2015

TABLE OF CONTENTS



01

ABOUT PROJECT

Data source and cleaning

02

EDA

Explatory data analysis

03

MODEL

Machine learning models
developed

04

RECOMMENDATIONS

Results and next steps
recommendations

01

ABOUT PROJECT


Tree Census ! 2015



A stylized landscape illustration. On the left, a large tree with a brown trunk and a full, rounded green canopy stands on a dark green hill. In the background, there are rolling dark green hills and a white cloud on the right side of the sky. The sky is a light yellow-green color.

INTRODUCTION

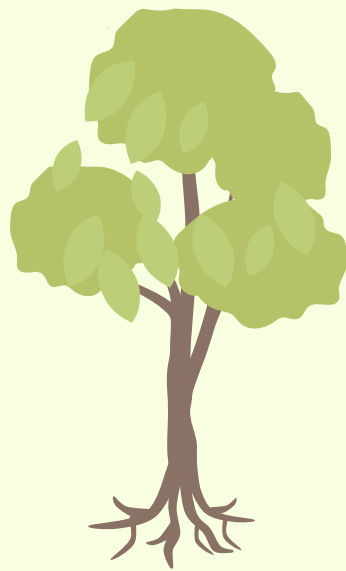
The aim of this project is to support New York City Parks Department efforts of Urban Forest to reduce climate and air pollution by predicting the health of the trees. The early detection will provide an opportunity to mitigate any conditions that are causing poor health in a tree.



“Each year NYC's street trees reduce summer electricity use from shading by 50,000 MWh.”

Every tree matters !

Problem Statement : Predict the health status of trees in New York City



Good



Fair



Poor



02

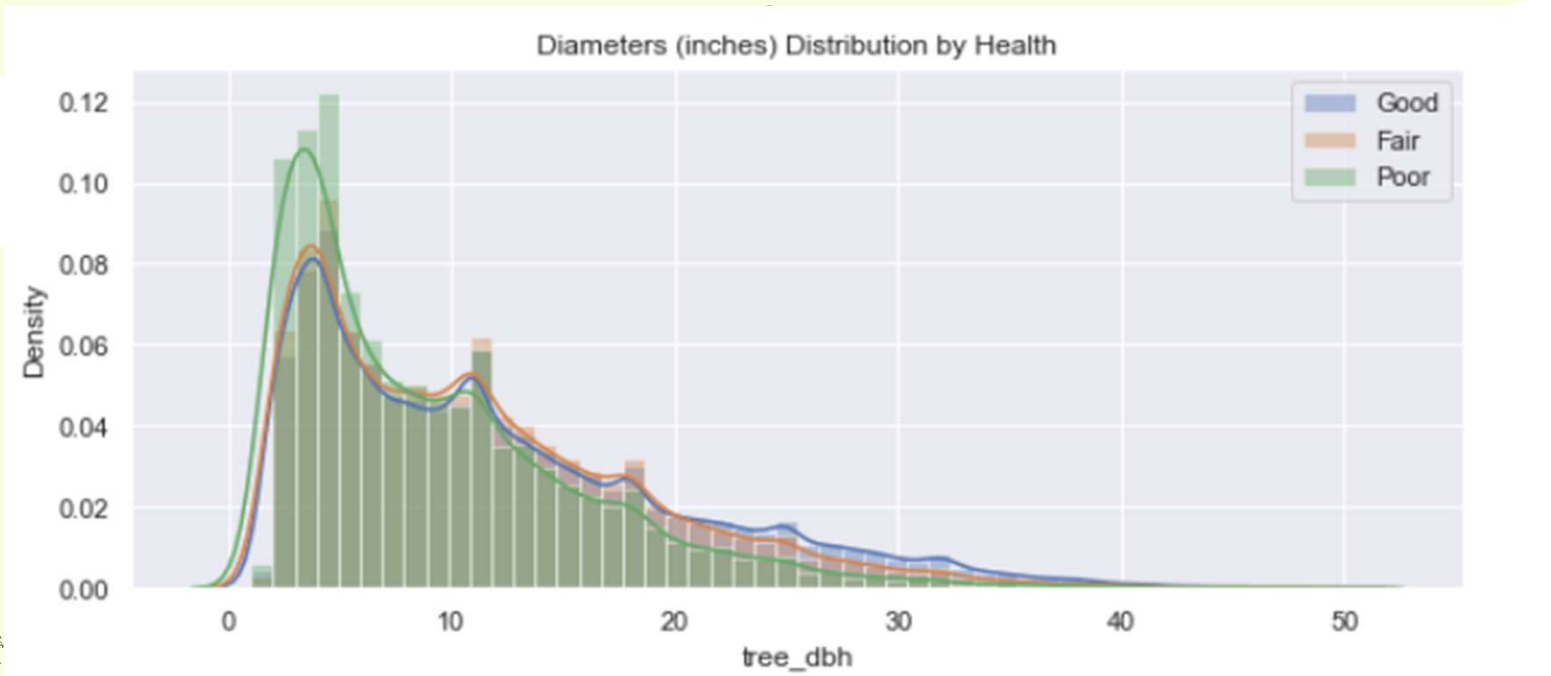
EDA

Exploratory Data Analysis

EDA

- One City | Five Boroughs | 666,134 Trees
- 81% of trees are in good health
- Most common species – London Planetree
- Chinese Chestnut, Atlas Cedar and Red Pine have least poor health trees.

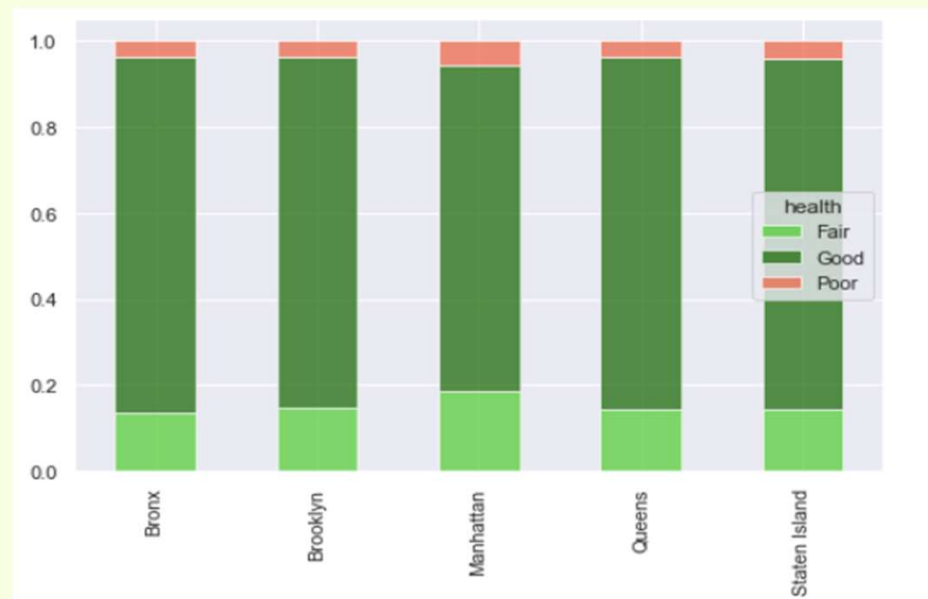
Most trees are less than 50 inches in diameter



EDA

- Most of the trees are planted on the curb
- Most common problem in the tree is related to stones
- Manhattan has highest number of trees in poor health

Tree Health by Borough



Problems in tree

01

BRANCH

Problems caused by lights,
wire, shoes, or other
branches problems

02

TRUNK

Problems caused by wire,
rope, lights, or other trunk
problems

03

ROOT

Problems caused by paving
stones, metal grates, other
root problems

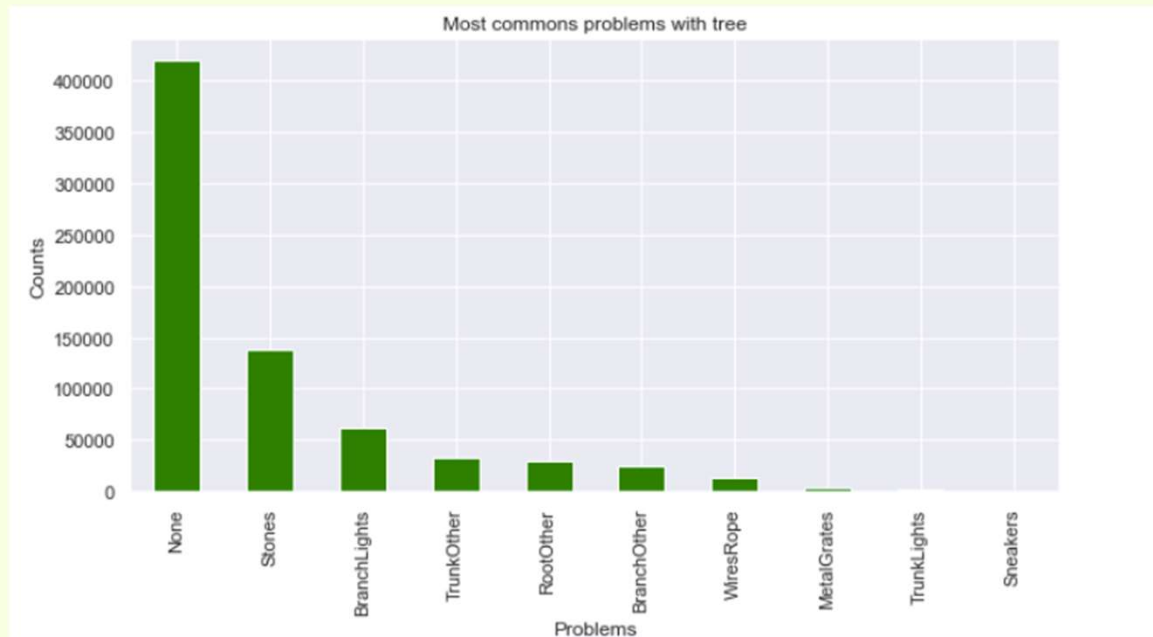
04

OTHERS

Gurads around trees,
sidewalk damage adjacent
to tree



Most common problems in trees





03

MODEL

Predicting the health of trees

Intial Models – Comparsions

Models	Logistic Regression		K-Nearest Neighbor		Decision Tree		Random Forest		Gradient Boosting	
	Train Accuracy - 0.81		Train Accuracy - 0.85		Train Accuracy - 0.99		Train Accuracy - 0.99		Train Accuracy – 0.81	
	Test Accuracy - 0.81		Test Accuracy - 0.80		Test Accuracy - 0.74		Test Accuracy - 0.81		Test Accuracy – 0.81	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
Good	1.00	0.00	0.94	0.85	0.85	0.86	0.96	0.85	1.00	0.81
Fair	0.00	0.00	0.25	0.40	0.31	0.30	0.20	0.47	0.01	0.51
Poor	0.00	0.00	0.06	0.38	0.18	0.17	0.12	0.39	0.03	0.48

Dealing with data imbalance

Random Forest Models	Random Oversampling		SMOTE		Borderline SMOTE		AdaSYN	
	Train Accuracy - 0.81		Train Accuracy - 0.99		Train Accuracy - 0.99		Train Accuracy - 0.99	
	Test Accuracy - 0.81		Test Accuracy - 0.87		Test Accuracy - 0.89		Test Accuracy - 0.86	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
Good	0.99	0.89	0.93	0.85	0.93	0.86	0.93	0.85
Fair	0.91	0.99	0.79	0.87	0.83	0.90	0.77	0.87
Poor	1.00	0.98	0.91	0.90	0.93	0.94	0.91	0.89

Random Forest Model

Post Random Sampling and Grid SearchCV

Random Forest		
Train Accuracy -		
Test Accuracy -		
	Recall	Precision
Good		
Fair		
Poor		

04

Next Steps





Next Steps

- Incorporate data from previous census years - 2005, 1995
- Explore deep learning models

THANKS!

Do you have any questions?

parijat.bhardwaj@gmail.com



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**

Please keep this slide for attribution