[12/12/20]

[Enrique Otanez (fictional) head of the NYC parks and recreation department.]

**Executive:**

While finding the best trees to fill the stumps and dead trees in Hunt’s Point, the sophora, ginkgo, eastern redbud, and callery pear are the best choices to use for replacement according to this analysis.

**Background:**

New York parks and recreation department has updated its operations with data analytics. After reading an article called "The 10 Worst Neighborhoods in New York City for 2020" by Sam Sparkes of RoadSnacks back in January(Sparks, Sam), the governor wanted something done about Hunt’s Point.

**Purposes and Questions:**

This project's purpose and overarching goal are to better what the worst neighborhood in NYC is as of 2020 (Sparks, Sam). The most pertinent problem looks to be stumps and dead trees. What immediately stood out was that although there could be damage shown on tree trunks, branches, and roots, choosing a familiar and liked tree by the citizens of Hunt's Point was pertinent. Although further analysis of damage to the tree itself is crucial, the KPI's selected were chosen because they would consider the overall conditions of the trees and the consequences of putting the trees into the ground on city sidewalks.

What trees show the most consistent stewardship and the highest amount of stewardship?

What trees are consistently on or off the curb, and what trees show improvement from on to off the curb?

What trees show any harm or help from guards placed on them and improvements from none to helpful and harmful to helpful guards?

What trees show good, fair, and poor health, and what trees show health improvement?

What trees show sidewalk damage, and what trees show improvement from sidewalk damage to no sidewalk damage?

**Methods and Data Source:**

The dataset is about all the trees in NYC and it is a very accurate representation of the city trees of as there are three ways NYC parks and recreation were able to receive the data, community volunteer, NYC parks and recreation staff, and partner organizations but the data was Kaggle.com. General information about the data recorded can help show the block id and find out accuracy based on what user (volunteer, NYC parks staff, or partner organization) recorded the record. Finding out the most consistently damaged trees could be insightful and the general location data can help narrow down inside boroughs and neighborhoods where they can be broken down further. We used these KPI's for this first iteration of analysis: steward, curb\_loc, guards, health, and sidewalk. Steward is the amount of recorded stewardship that was shown to a tree. The specific method to be used is comparing each variable KPI to the types of trees and of each of the sub-variables see what trees shown improvement. For instance, looking at how health is broken up between good, fair and poor health, running an analysis of each and then comparing to see what trees did better by percentage and moving up in the list. There are other methods that were tried like comparing curb\_loc analysis directly to the trees picked in the stewardship analysis, but it was decided to look at the rest of the KPI’s objectively in a non-comparison way. The other was to compare KPI’s but needing something more on the base level, that approach was ignored for the time being.

**Results and Recommendations:**

The results are based on what trees consistently showed improvements across all KPI's used for the overall business problem analysis so overall improvements between KPI's, the trees with the most improvements will be selected. These trees are the sophora, ginkgo, eastern redbud, and callery pear in order. The sophora was chosen for the number one spot because one of its improvements were in stewardship, our main KPI. There should be no actions taken based on this analysis alone. More analysis must be done with the remaining possible KPI's not used in this analysis. The next steps would then be to enact the above. Consider the consequences, good or bad, of implementing the findings now rather than what should be done before implementation, which is testing the current KPIs against each other for more insight and testing the remaining KPIs. The innovative solution currently is to do surveys about this current tree and what tree they like independently. This is to find out the tree the citizens like the most so far.

**Limitations:**

Firstly, the results are inconclusive, mostly because of the dataset's accuracy issue and how the analysis was conducted. There may be some decisions that were made based on the assumptions that were not sound. Why compare the trees that improved to stewardship? The safe assumption is that it will be more liked, but that is hardly the case for the real world. The percentages often changed between visualizations to represent either the total percentage of trees selected or the entirety of trees in Hunt’s Point, interpreting improvement via percentage was a challenge and a better measure could be determined. The overabundance of certain trees we also a problem. Possibly normalizing the data or found a different decision to select trees.

References:

1. Sparkes, Sam. “These Are The 10 Worst Neighborhoods In New York For 2019.” *RoadSnacks*, 12 Jan. 2020, [www.roadsnacks.net/worst-neighborhoods-in-new-york-city/](http://www.roadsnacks.net/worst-neighborhoods-in-new-york-city/) (Sparks, Sam)