NYC Recycling Rates

Evelyn Cates

08/07/2019

## **Introduction**

When it comes to single use plastics, we have a global problem that’s reaching a breaking point. Some countries have even stopped importing trash because they are unable to process the high quantities they have. Single use plastic is created with the iinstantention of lasting for years completely disregarding the impact it has on our environment. In the project I use data to understand how much plastic is being recycled in NYC.

In order to understand the problem I found the NYC Open Data source Recycling Diversion and Capture Rates. The Diversion rate is categorized as the percentage of total municipal solid waste collected by the Department of Sanitation that is disposed of by recycling. The Capture Rate is the percentage of total Paper or Metal/Glass/Plastic in the waste stream that is disposed of by recycling.

### Importing and Cleaning the Data

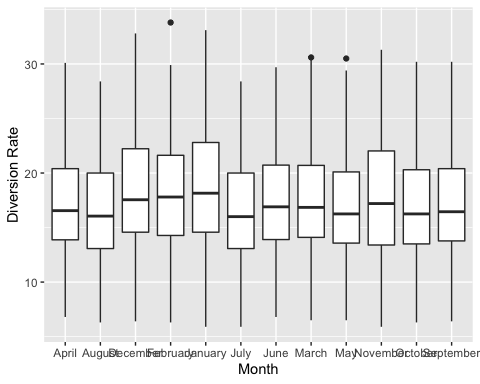
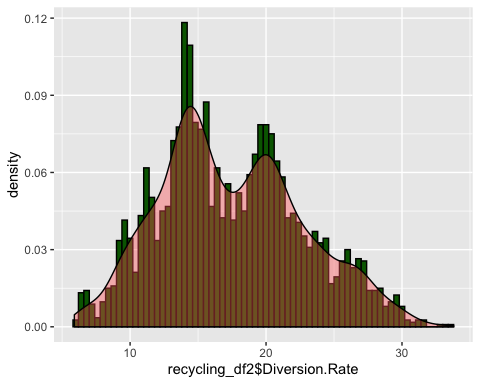
Once I imported the data, I used na.omit to remove any empty cells. Since we are currently finishing up Fiscal Year 2019 and the dataset is not updated with Fiscal Year 2020 I didn’t not remove any dates. I also modified the name of the column in order to find them easier.

## Zone District Fiscal.Month.Number Fiscal.Year   
## Length:2832 Length:2832 Min. : 1.00 Min. :2016   
## Class :character Class :character 1st Qu.: 3.75 1st Qu.:2017   
## Mode :character Mode :character Median : 6.50 Median :2018   
## Mean : 6.50 Mean :2018   
## 3rd Qu.: 9.25 3rd Qu.:2018   
## Max. :12.00 Max. :2019   
## Month.Name Diversion.Rate Capture.Rate Capture.Rate.1   
## Length:2832 Min. : 5.90 Min. :16.90 Min. : 20.40   
## Class :character 1st Qu.:13.80 1st Qu.:31.90 1st Qu.: 48.40   
## Mode :character Median :16.90 Median :37.75 Median : 61.60   
## Mean :17.48 Mean :38.43 Mean : 61.22   
## 3rd Qu.:21.00 3rd Qu.:45.20 3rd Qu.: 74.42   
## Max. :33.80 Max. :72.60 Max. :101.80   
## Capture.Rate.2   
## Min. :18.50   
## 1st Qu.:39.70   
## Median :46.30   
## Mean :46.16   
## 3rd Qu.:52.62   
## Max. :75.50

## [1] 0

## Observations: 2,832  
## Variables: 9  
## $ Zone <chr> "Brooklyn North", "Brooklyn North", "Brookly…  
## $ District <chr> "BKN01", "BKN02", "BKN03", "BKN04", "BKN05",…  
## $ Fiscal.Month.Number <int> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, …  
## $ Fiscal.Year <int> 2019, 2019, 2019, 2019, 2019, 2019, 2019, 20…  
## $ Month.Name <chr> "April", "April", "April", "April", "April",…  
## $ Diversion.Rate <dbl> 14.7, 20.0, 12.2, 15.5, 10.1, 16.4, 14.0, 9.…  
## $ Capture.Rate <dbl> 44.9, 34.2, 33.5, 35.2, 22.3, 44.5, 34.0, 20…  
## $ Capture.Rate.1 <dbl> 43.0, 57.9, 44.9, 68.5, 45.1, 59.4, 47.9, 35…  
## $ Capture.Rate.2 <dbl> 44.1, 41.2, 38.2, 48.8, 31.5, 50.6, 39.2, 26…

### Data Exploration and Analysis



## # A tibble: 4 x 2  
## Fiscal.Year diversion\_rate\_fy  
## <int> <dbl>  
## 1 2016 16.7  
## 2 2017 17.3  
## 3 2018 17.9  
## 4 2019 18.0

##   
## Call:  
## lm(formula = Diversion.Rate + Capture.Rate ~ Zone, data = recycling\_df2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.045 -8.163 -0.637 7.755 36.368   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 45.1627 0.4344 103.97 < 2e-16 \*\*\*  
## ZoneBrooklyn North 2.5017 0.6635 3.77 0.000166 \*\*\*  
## ZoneBrooklyn South 19.0820 0.6635 28.76 < 2e-16 \*\*\*  
## ZoneManhattan 16.2740 0.6143 26.49 < 2e-16 \*\*\*  
## ZoneQueens East 11.5698 0.7156 16.17 < 2e-16 \*\*\*  
## ZoneQueens West 14.9403 0.7156 20.88 < 2e-16 \*\*\*  
## ZoneStaten Island 19.7387 0.9713 20.32 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.42 on 2825 degrees of freedom  
## Multiple R-squared: 0.3444, Adjusted R-squared: 0.343   
## F-statistic: 247.3 on 6 and 2825 DF, p-value: < 2.2e-16

###Analysis Impact This dataset allowed me to start looking at how we can create models for environmentalism in a new way. By using an exploratory method in the beginning of the project I was able to look at the plastic recycling way by zone in NYC and start asking additional questions that lead me to discover additional discrepancies. This was a smaller dataset than I originally thought and had a challenging time building a regression model. I orginally thought the recycling amounts were going to be lower because of the consumer and while that isn’t completely wrong, this analysis takes into account another component not considered much when talking about recycling, but our waste management system. There is a process and until we understand what the impact is from each part of the process we can’t improve the outcome.