

# 311 Service Requests

Monica Puerto, MS Data Science AU University  
GASP 9/23/2019

311 Service Requests Analysis

Github: @monipip3

# What is a 311 Service Request?

A solicitation to the local government to remediate a non emergency issue that occurred within that local jurisdiction.

Requests range and are not exclusive to the following :  
noise complaints, abandoned vehicles, dead animal removal, potholes / sinkholes, snow removal, street light repairs, missed garbage complaint, sidewalk repair, homeless services, etc.

# Where are 311 Services Available?

- The service launched first in Baltimore, Maryland in 1996 after concluding that 60% of the prior years 1.8 million 911 emergency phone calls were not an emergency.
  - Now 80 cities in United States have 311
  - Canada has ~20 participating cities.
  - Latin American countries like Panama and Costa Rica have also launched this 311 service.

# 311 Open Data

“Open 311” is about providing an API standard for future web applications that is accessible and decentralized. With the explosion of mobile data, most cities have stored their millionth 311 request in their 311 database.

## Why is this important?

- Transparency
- Accountability
- Accessibility
- Infrastructure for Operations

# Hypothesis

Wealthier neighborhoods utilize the 311 requests more than neighborhoods where poverty is present. Income has a positive affect on the rate of 311 requests per 1000 people and poverty has a negative affect.

# Datasets

**American Community Survey(ACS)- 5 year estimates.** Using an R package called 'acs'.The package queries from the Census Bureau API.

- total population, male counts
- median household income
- median age
- white count, and black count

**311 Data Sets 2016-2018** Using the Socrata API and an R package called 'RSocrata':

- Baltimore
- New York City
- [DC](#) downloaded via [opendata.dc.gov](https://opendata.dc.gov)

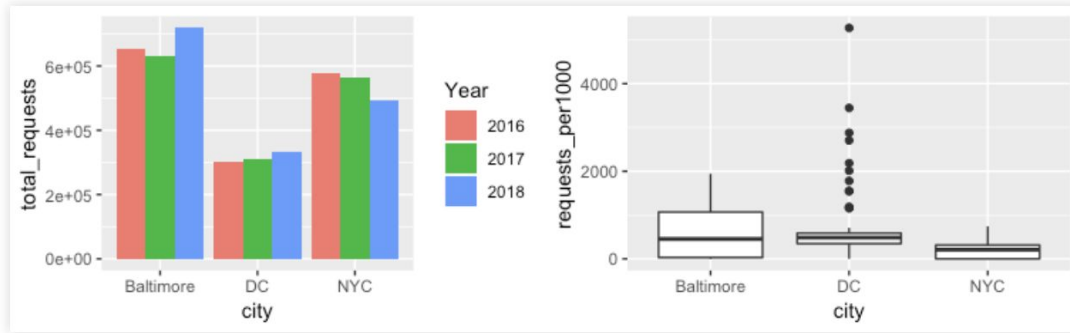
# Data Challenges

- Each 311 data set had to be cleaned
- New York City is way too big (20M rows) ; had to aggregate at GET request.
- There is not a nationwide reporting standard of 311 data labeling
- Learning the ACS package functions and the structure of American Community Survey (table names, variable names, variable codes)

city	unique_service_request_categories	unique_agencies
Baltimore	286	19
NYC	238	30
DC	159	13

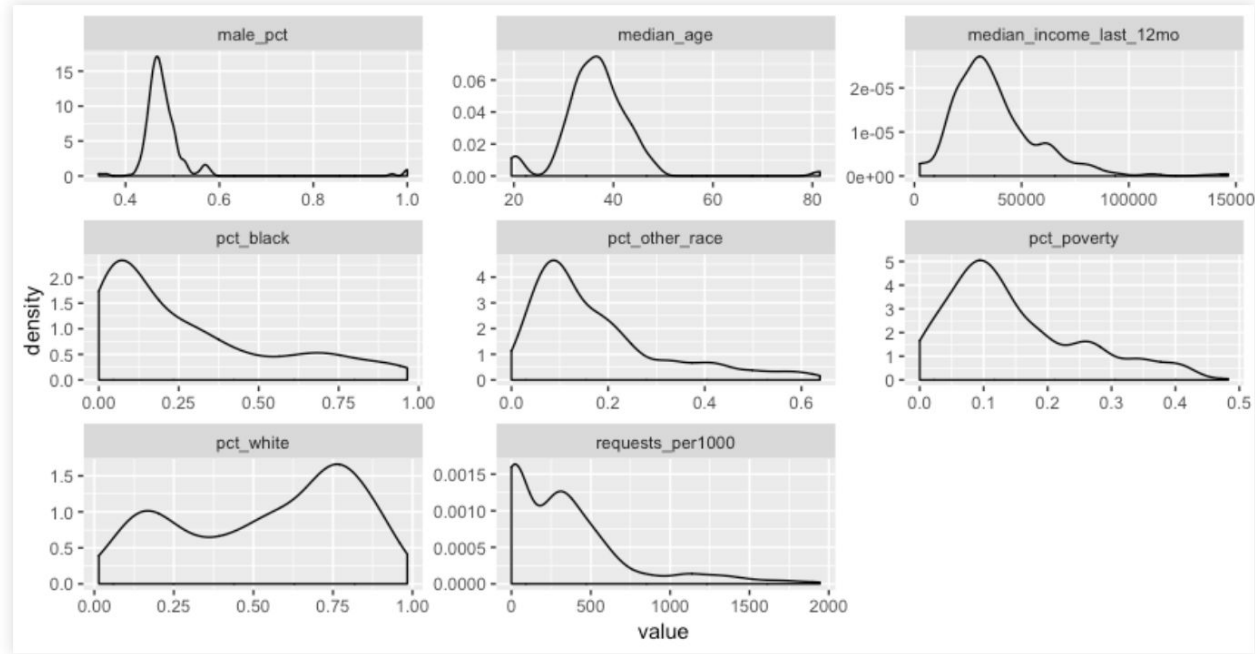
# 311 Request Size

- Baltimore had the most logged requests despite smallest population.
- DC, had the highest average rate of 311 requests per 1000 people out of the 3 cities.
  - DC outliers due to museums/universities/downtown area (removed them later on)





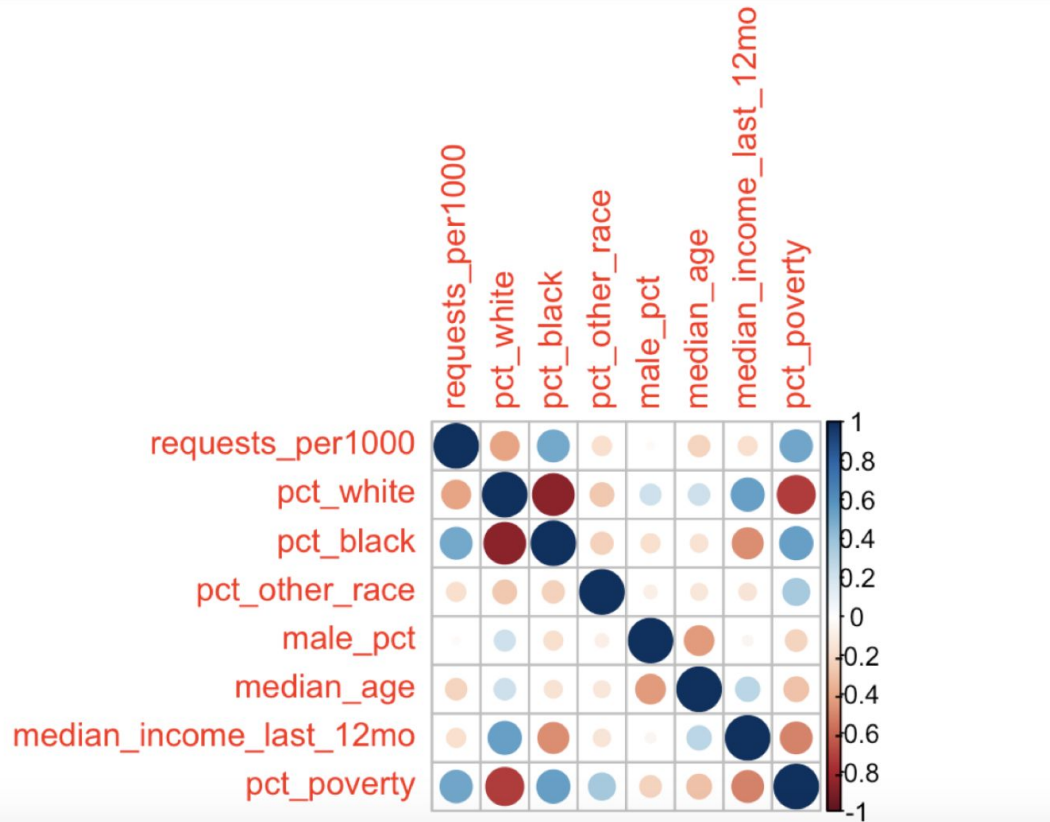
# Distribution of all Variables



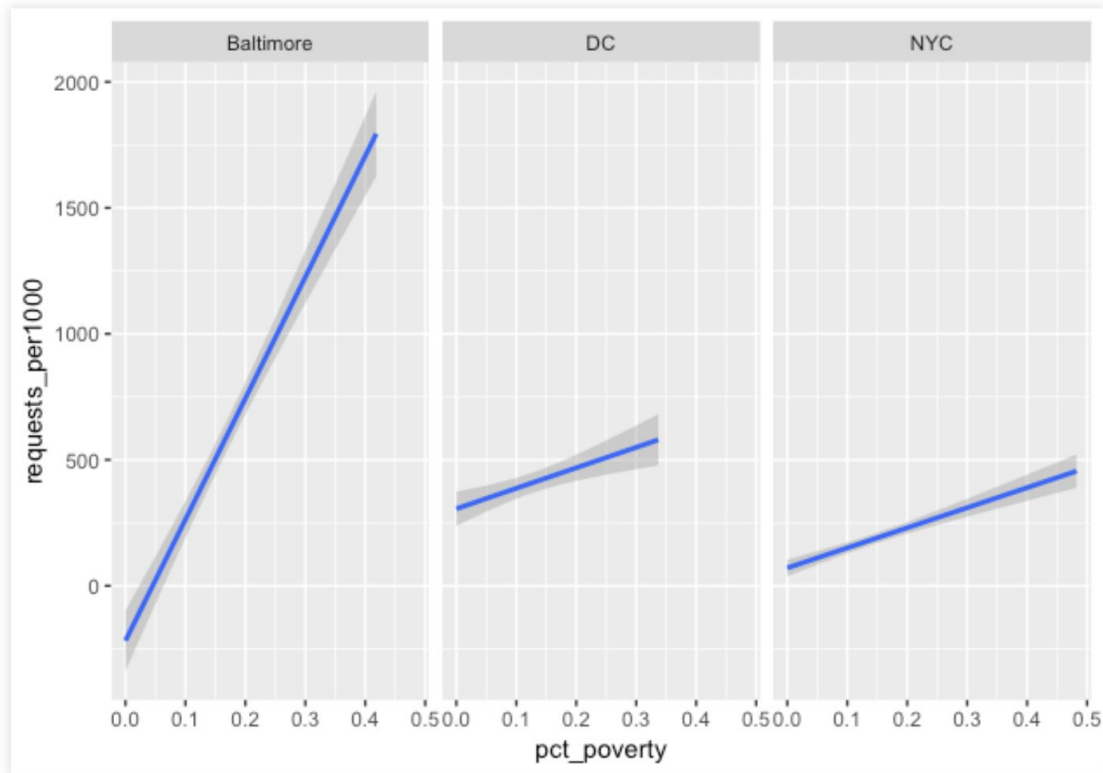
# Transforming the Data

- Create table of requests, by year, zipcode, and city.
- Convert dependent variable; a count of 311 requests by zipcode to requests per 1,000 people by utilizing the population estimate from ACS(2013-2017).
- Create other race count by :  
$$\text{\( Total Population \)} - \text{\( White Count \)} - \text{\( Black Count \)}$$
- Transform the acs count data into percentages from acs of Race, divided by population. Eg: 
$$\text{\( \frac{White}{Total Population} \)}$$
- Convert the city & years into factors

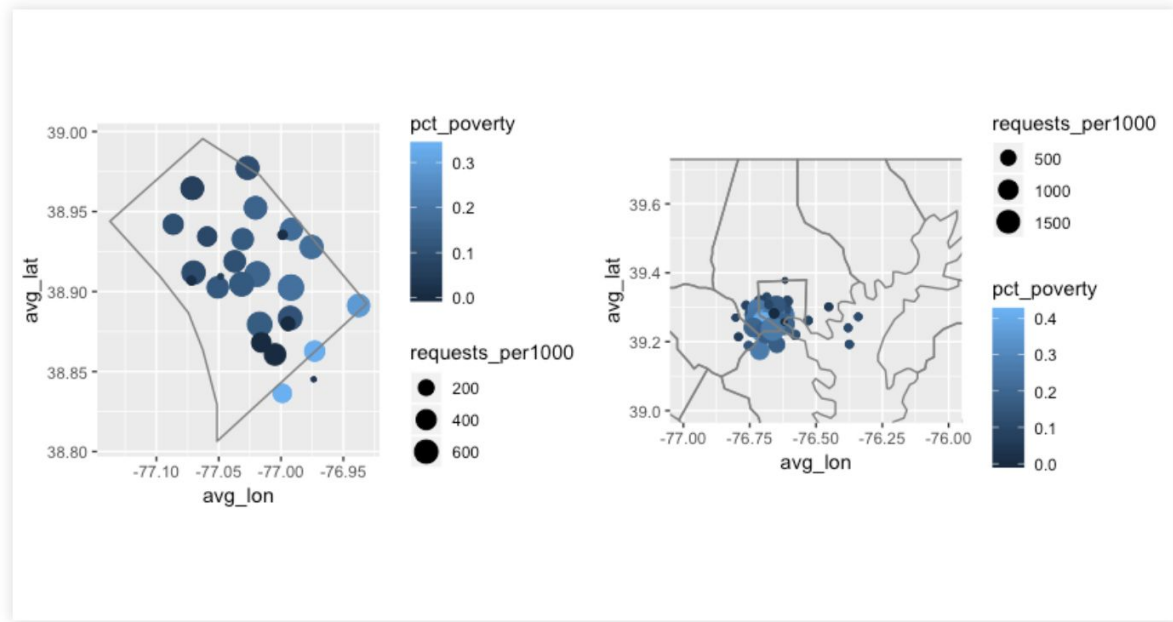
# Correlation



# Poverty vs 311 by city



# Maps



# Using Stepwise to find the best model

```
Call:
lm(formula = requests_per1000 ~ pct_white + pct_black + male_pct +
    median_income_last_12mo + city + pct_poverty, data = data_model_final)
```

Residuals:

Min	1Q	Median	3Q	Max
-646.45	-146.12	0.66	140.51	1032.25

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-6.749e+02	1.716e+02	-3.933	9.90e-05	***
pct_white	3.275e+02	1.451e+02	2.258	0.024510	*
pct_black	5.219e+02	1.339e+02	3.899	0.000114	***
male_pct	6.887e+02	1.845e+02	3.734	0.000217	***
median_income_last_12mo	5.666e-03	8.350e-04	6.785	4.26e-11	***
cityDC	-1.627e+02	4.223e+01	-3.852	0.000137	***
cityNYC	-3.149e+02	4.112e+01	-7.658	1.47e-13	***
pct_poverty	2.313e+03	1.928e+02	11.998	< 2e-16	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 262.5 on 394 degrees of freedom

Multiple R-squared: 0.5304, Adjusted R-squared: 0.5221

F-statistic: 63.58 on 7 and 394 DF, p-value: < 2.2e-16

# Conclusion

- Poverty had the strongest relationship/affect on the rate of 311 requests.
- 1% of poverty results in an increase of 23 service requests per 1000 people
  - According to Pew Research “In 2016, one-fifth of adults living in households earning less than \$30,000 a year were “smartphone-only” internet users” and “The vast majority of Americans (95%) now own a cellphone of some kind”.

# Next Steps

## Include other cities in analysis

- Have separate models for types of cities. Use KNN to figure out similar cities with ACS data
- Inspire a standardization of 311 data and more research on 311 data that can help urban planning.

