

Who Calls for Help?

Statistical evidence of disparities in citizen-government interactions using geo-spatial survey and 311 data from Kansas City

Constantine E. Kontokosta, PhD

New York University
Tandon School of Engineering
Center for Urban Science and Progress (CUSP)
Brooklyn, NY
ckontokosta@nyu.edu

Boyeong Hong

New York University
Tandon School of Engineering
Center for Urban Science and Progress (CUSP)
Brooklyn, NY
boyeong.hong@nyu.edu

ABSTRACT

Local governments are increasingly turning to '311' citizen complaints and service reports to provide a real-time condition assessment of the city. When combined with machine learning and predictive analytics, these data can be mined to gain new insight into city service needs and potential problems, including quality-of-life issues like rodent infestations, illegally converted buildings, and heat and hot water outages. However, we know that people do not complain at the same rate, and the willingness to interact with local government varies with socioeconomic, demographic, and cultural traits, as well as individual exposures to negative conditions. The goal of this preliminary work with the Kansas City Office of Performance Management (OPM) is to understand the disparities in the use of 311 and to identify the individual characteristics that impact citizen-government engagement and interaction. To do so, we use results from a citizen satisfaction survey administered by OPM that consists of 21,046 individual responses from 2014 to 2017. We combine these data with more than 500,000 311 reports and approximately 200,000 code enforcement violations to develop a statistical model to predict the likelihood of 311 reporting about neighborhood and property issues (code enforcement) based on household demographic and socioeconomic characteristics, as well as measures of community engagement and political participation. Furthermore, we explore the relationship between 311 usage and citizen satisfaction with city government, testing the hypothesis that there is a positive relationship between citizen participation in government processes (co-production) and overall citizen satisfaction. The insights of our preliminary work form the basis of acknowledging and accounting for the data bias problem of the citizen self-reported data, and contribute to the more equitable allocation and delivery of city services.

Keywords

311 complaint; citizen engagement; bias; machine learning; urban analytics

1. INTRODUCTION

Many North American cities, including New York City, Chicago, Toronto, Washington, DC, and Kansas City, have implemented 311 systems to manage citizen complaints and service requests [9, 7, 6, 5]. These platforms provide 24/7 customer service to collect citizen reports and to disseminate them to relevant city departments for improved city services [1, 5]. As such, 311 can be a crucial link between citizens and government and represent an example of co-production through digital technology [9]. Because these citizen reports provide a real-time condition assessment of the city, local governments are analyzing these data to understand and forecast problems, service demands, and quality-of-life issues, such as rodent infestations, illegally converted buildings, and heat and hot water outages [5, 10].

However, we know that people do not complain at the same rate. Some of this may simply be a function of the problems they are exposed to - a neighborhood with better conditions should illicit fewer complaints per person. At the same time, two individuals facing similar conditions may have different responses based on their expectations for what the conditions should be. For instance, an individual accustomed to seeing rodents in their building may be less likely to complain than someone seeing a rodent in their apartment for the first time. In addition, individuals may have different levels of trust in government, differing expectations that the government will actually respond, and socio-cultural traits that make them more or less likely to report a problem. While we explored this topic previously using NYC311 data, we build substantially on that work by exploring individual resident survey data, and linking household characteristics, preferences, and opinions and the use of the 311 system in Kansas City (KC).

The goal of this preliminary work with the KC Office of Performance Management (OPM) is to understand the disparities in the use of 311 and to identify the individual characteristics that impact citizen-government engagement and interaction. To do so, we use results from a citizen satisfaction survey that consists of 21,046 individual responses from

2014 to 2017. The survey is a stratified sample across city council districts based on a balanced population against census demographics. We combine these data with more than 500,000 311 reports and approximately 200,000 code enforcement violations to develop a statistical model to predict the likelihood of 311 reporting about neighborhood and property issues based on household demographic and socioeconomic characteristics, as well as measures of community engagement and political participation. Furthermore, we explore the relationship between 311 usage and citizen satisfaction with city government, testing the hypothesis that there is a positive relationship between citizen participation in government processes and overall citizen satisfaction [3, 11]. We expect that the insights of this preliminary work will form the basis of acknowledging and accounting for bias in citizen self-reported data, and contribute to the efficient and fair delivery of city services in KC. This paper is organized as follows: section two describes our data and methodology, section three presents our preliminary findings, and implications and future work are discussed in section four.

2. DATA AND METHODS

In order to identify disparities in citizen-government interactions, we acquire and integrate a wide range of KC datasets, described in Table 1. OPM provided the results from a citizen satisfaction survey conducted between 2014 to 2017, consisting of 21,046 individual responses. The City administers a citizen survey about the level of satisfaction with city services, use of city services, and household characteristics each fiscal year. The survey samples are stratified across the city council districts based on a statistically significant random sample of the balanced population against census demographics, and responses are collected via mail, phone, and webpage [2]. The format of the questionnaire is a five-level Likert scale with 5 being "very satisfied" and 1 being "very dissatisfied". We integrate these survey responses with KC 311 reports and property maintenance code violations data from 2014 to 2017 extracted from the KC Open Data platform. If a resident reports a code enforcement problem related to a specific property, a 311 case is created and, as needed, an inspection is conducted by the Neighborhoods and Housing Services Department (NHS). If inspectors do not find any violations, the 311 case is closed; otherwise, if a violation is found, each (there can be a one-to-many rela-

tionship between 311 case and property violations) is entered into the property violation database. Figure 2 represents the Sankey diagram of the complaint-violation process flow.

Figure 1 visualizes the spatial patterns of 311 reports and property violations (showing a heat map, total volume, and population-normalized volume, respectively). According to the citizen satisfaction survey data, 11,041 (52.4%) respondents have used 311 at least once. In 2016, a total of 135,798 311 complaints were reported, and 53.2% (72,200) of these were related to neighborhood and property issues (code enforcement) handled by NHS. Fully 82.8% (59,807) of these code enforcement-related reports were closed before an inspection, or not identified as a violating condition following an inspection. The remaining 12,394 (17.2%) reports resulted in 62,342 property violations, as any single 311 report can be associated with one or multiple violations (up to 80 based on the analyzed data). These represent 99.9% of all property violations issued in 2016, indicating that effectively all property violations result from citizen reports through the KC 311 system. The ratio of the number of 311 reports resulting in property violations to the total number of 311 reports varies considerably across the city (see Figure 3), suggesting that there are non-trivial socio-spatial variations in reporting behavior.

We use the citizen satisfaction survey data to identify statistical relationships between the demographics of 311 users and non-users. Since the survey data is a statistically significant random sample across neighborhoods, there should not be *a priori* differences in household characteristics between 311 users and non-users if there is no disparity in 311 reporting due to individual factors. We split the data into two groups - 311 users and non-users - based on the survey question "Have you used KC 311" and extract relevant variables from the survey questionnaires relating to demographic and socioeconomic characteristics, as well as measures of community engagement and political participation described in Table 2. We also use the survey results to explore the relationship between co-production (311 usage) and overall citizen satisfaction. We use t-tests, Chi-square tests, and Mann-Whitney U-tests depending on variable types to determine whether there is statistically significant difference between two groups.

Table 1: Description of Dataset used

Dataset	Records	Time range	Temporal resolution	Spatial resolution	Source and description
Citizen satisfaction survey	21,046	2014 - 2017	Quarterly	Individual (household)	Citizen survey on level of satisfaction with city services, use of city services, and household characteristics each fiscal year provided by Kansas City (Missouri)
311 service requests	1.19M	2013 - present	Daily	GPS (X,Y)	All citizen service requests and non-emergency complaints, provided by KC 311 through OpenData KC
Property violations	610k	2015 - present	Daily	GPS (X,Y)	Neighborhood code enforcement violations issued by the KC Neighborhood and Housing Service (NHS), provided by KC NHS through OpenData KC
Census population	246	2010	Annual	Neighborhood	KC Census population in 2010, U.S. Decennial Census and American Community Survey

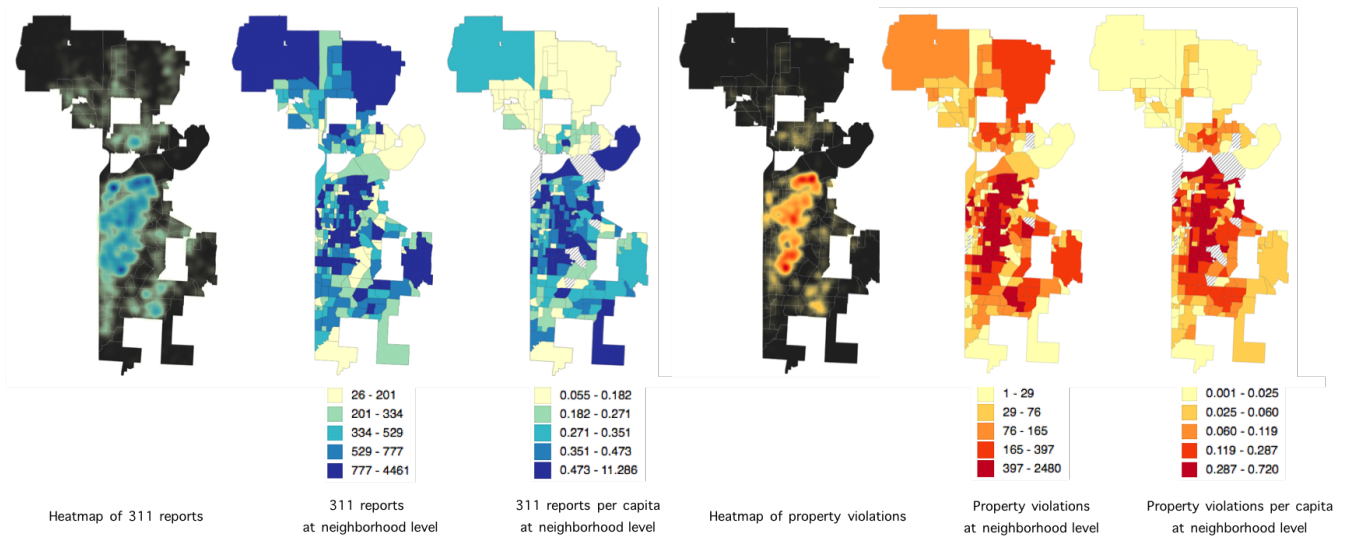


Figure 1: Spatial patterns of 311 reports and property violations in 2016 (Total 311 service requests in 2016: 135,789; total property violations in 2016: 62,391)

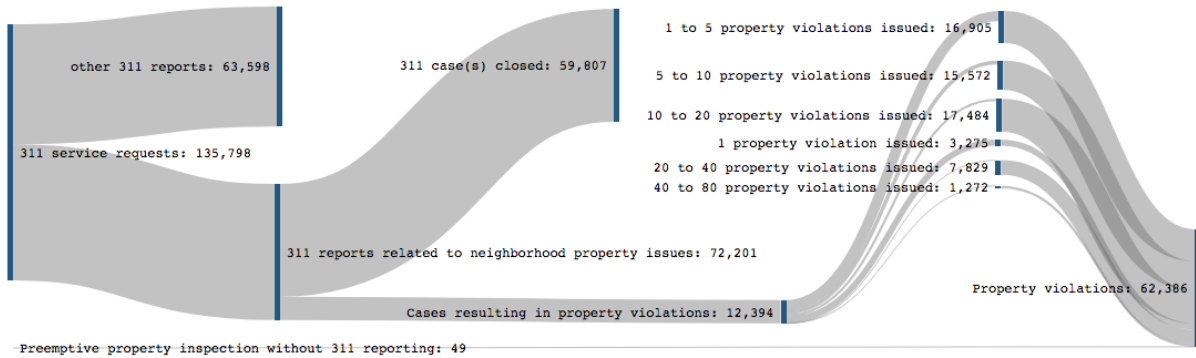


Figure 2: Sankey diagram of 311 service requests and property violations workflow

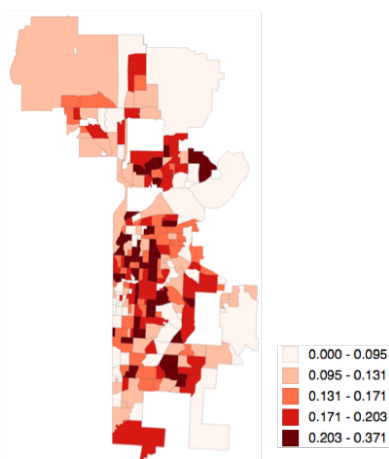


Figure 3: Ratio of the number of 311 reports resulting in property violations to the total number of 311 reports, by neighborhood

Turning to our primary research question, we develop a preliminary logistic regression model to predict the individual likelihood of 311 reporting. The dependent variable is a binary representation of 311 use based on survey responses.

- 1 (311 user) : Respondent who has used 311
- 0 (311 non-user) : Respondent who has not used 311

Statistically significant features from the previous analysis (t-test, Chi-square test, and Mann-Whitney U-test) are used as the predictors (independent variables) for our logistic regression model, specified as:

$$\text{Logit}(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n \quad (1)$$

where y is a binary for individual 311 usage and x_i is the set of individual demographic and socioeconomic characteristics, level of satisfaction with the city services, and measures of civic engagement and political participation. The coefficients β_i are used to estimate the odds of 311 use for an individual respondent.

3. PRELIMINARY RESULTS

Table 2: Input variables

Category	Variable	Questionnaire
Demographic/socioeconomic characteristics	Gender	Gender (Categorical)
Demographic/socioeconomic characteristics	Age	What is your age (Categorical)
Demographic/socioeconomic characteristics	Income	Total annual household income (Categorical)
Demographic/socioeconomic characteristics	Race/Ethnicity 1	Describe your race ethnicity (Categorical)
Demographic/socioeconomic characteristics	Race/Ethnicity 2	Are you Hispanic Latino Spanish (Categorical)
Demographic/socioeconomic characteristics	Dwelling type	Type of dwelling live in (Categorical)
Demographic/socioeconomic characteristics	Home ownership	Own or rent current residence (Categorical)
Demographic/socioeconomic characteristics	Household with children 1	Have children live in KCMO (Categorical)
Demographic/socioeconomic characteristics	Household with children 2	Type of K12 school attend (Categorical)
Demographic/socioeconomic characteristics	Long-term residents in KC	How long have lived in KCMO (Continuous)
Demographic/socioeconomic characteristics	Pet	Own at least one cat or dog (Categorical)
Demographic/socioeconomic characteristics	Internet access	Have access to Internet at home (Categorical)
Demographic/socioeconomic characteristics	Recycling	How often use curbside recycling (Categorical)
Civic engagement	Channel 2	Have watched Chn 2 last year (Categorical)
Civic engagement	City website	Visited city’s website (Categorical)
Civic engagement	Community center	Visited KCMO community center (Categorical)
Civic engagement	Park	Visited any parks in KCMO (Categorical)
Civic engagement	Public meeting	Attended watched KCMO public mtgs (Categorical)
Civic engagement	Public transportation	Used public transportation (Categorical)
Political participation	Voting	Vote in election (Categorical)
Citizen satisfaction	Neighborhood service	Quality of neighborhood service (Ordinal)
Citizen satisfaction	Quality of life	Quality of life in the city (Ordinal)
Citizen satisfaction	Safety	Feeling of safety in the city (Ordinal)
Citizen satisfaction	Physical appearance	Physical appearance of neighborhood (Ordinal)

Table 3: Results of the two-sample tests for 311 use (*) p-value = 0.00, ** p-value < 0.01, * p-value < 0.05)**

Feature	Test method	Test value
Gender	Chi-square	35.21***
Age	Chi-square	382.88***
Income	Chi-square	122.17***
Race/Ethnicity1	Chi-square	367.47***
Race/Ethnicity2	Chi-square	134.89***
Dwelling type	Chi-square	351.28***
Homeownership	Chi-square	994.98***
Household with children1	Chi-square	47.14***
Household with children2	Chi-square	41.28***
Long-term resident	T-square	19.58***
Pet	Chi-square	115.36***
Internet access	Chi-square	9253.53***
Recycling	Chi-square	1143.65***
Channel 2	Chi-square	191.56***
City webiste	Chi-square	10262.8***
Community center	Chi-square	7734.6***
Park	Chi-square	8207.31***
Public meeting	Chi-square	578.34***
Public transportation	Chi-square	533.82***
Voting	Chi-square	492.10***
Neighborhood service	Mann-Withney	40801257.0***
Quality of life	Mann-Withney	47628351.0***
Safety	Mann-Withney	47628351.0***
Physical appearance	Mann-Withney	7370220.0***

Table 3 is the results of statistical tests. From the two-sample hypothesis tests, we observe statistically significant differences between 311 users and non-users across several

features. We find that the 311 user group are comprised of a larger proportion of females, middle income households, single family housing residents, homeowners, households with kids, parents whose children attend private or charter schools, long-term residents, households with a pet, households with Internet access, and households that participate in the curbside recycling program often. Interestingly, people who have watched KC Channel 2 (KCMO government channel providing coverage of city council meetings, city events, and other programs), who use city infrastructure (parks and public transit), who have attended city public meetings, and have participated in elections tend to contact 311 more regularly. With respect to civic engagement and political participation, our results support previous evidence that there is a positive relationship between e-government participation and real-world civic involvement [4]. We further find an inverse relationship between 311 usage and citizen satisfaction with city government. The 311 user group tends to be more dissatisfied with their local government and public services compared to the non-user group. Although intuitive, this finding is contrary to previous literature.

Figure 4 and Table 4 show the results of the preliminary logistic regression model. The model performs reasonably well (AUC is 0.69), and we observe statistically significant relationships between individual factors and the likelihood of 311 use. We find that citizens who regularly use city infrastructure (public transit, city parks) and who actively participate in political events are up to approximately 25% more likely to contact 311 to report neighborhood issues, controlling for other factors. With respect to individual/household demographic and socioeconomic characteristics, we observe that gender, age, income, race, dwelling type, pet ownership, and curbside recycling participation are statistically signifi-

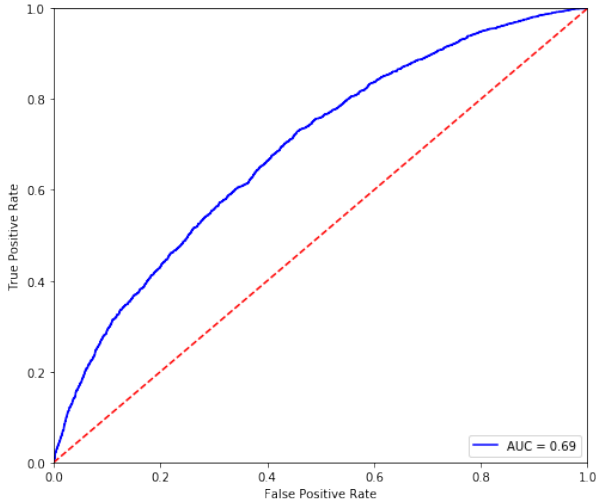


Figure 4: ROC (Receiver Operating Characteristics) Curve of the logistic regression model

Table 4: Results of the logistic regression model for 311 use (*) p-value = 0.00, ** p-value < 0.01, * p-value < 0.05)**

Feature	Coef.	Odds ratio
Female	0.2140***	1.2386
Age group under 25	-0.5159***	0.5969
Middle income household	-0.0177	0.9824
Black	0.4582***	1.5812
Asian	-0.4699***	0.6250
Multifamily housing	-0.4433***	0.6419
Renter	-0.2875***	0.7501
Weekly curbside recycling	0.6471***	1.9100
Pet	0.1627***	1.1767
Vote in election	0.1665***	1.1812
Attend public meeting	0.2204***	1.2466
Use public transportation	0.2352***	1.2651
DO NOT use city park	-0.1137**	0.8924
Visit community center	0.1916***	1.2112
DO NOT use city website	-0.5057***	0.6031
Watched Chn2	0.2086***	1.3615
Level of satisfaction with city services	-0.1870***	0.8294

cant explanatory variables. The odds of a female using 311 are 1.24 times that of a male, and Black residents have a 58% higher odds of using 311. Households with at least one pet are generally more likely to contact 311, as are those who often participate in a curbside recycling program. Those that live in multifamily housing or rental units are less likely to report property issues through 311, and Asian residents have 38% lower odds of complaining through 311. Overall, these results indicate that individual demographic and socioeconomic characteristics, civic engagement and political participation have a non-trivial effect on reporting behavior.

4. DISCUSSION AND CONCLUSION

Our preliminary findings indicate statistical evidence of disparities in citizen-government interactions via the KC 311

system based on individual satisfaction survey data. In previous work, explaining and quantifying the unequal use of the 311 system has been a challenge without individual user data. Our results presented here clearly demonstrate that citizen-government interaction through the 311 system in KC does vary depending on demographic and socioeconomic characteristics, civic engagement and political participation, and citizen satisfaction with their government. Based on this evidence, we demonstrate that there is bias in complaints received through 311. As such, training predictive city service delivery models on these data would lead to an inequitable distribution of service provision, leading to over-allocation of resources to households and neighborhoods that are more likely to report problems.

Further improvements in our work include accounting for the interactions between individual factors, as well as analyzing geo-spatial relationships to understand disparities in the use of 311 across report/complaint types and neighborhoods. Integrating citizen satisfaction survey data and additional datasets (property tax data, 311, property violations, etc.) will provide deeper insights into our research question. Finally, we will explore social network effects on reporting behavior, the likelihood of using city services, and the overall level of satisfaction with city government.

Our study is critical to ensuring equitable and fair provision of city services when using predictive analytics for city management. Since 311 data are shown not to be representative of the entire population, nor necessarily reflective of actual problems and local conditions, city agencies need to recognize the potential for, and magnitude of, sample bias in self-reported data. Our work forms the basis for acknowledging and accounting for this data bias, and contributes to the KC 311 strategy of resource allocation and proactive inspections for efficient and equitable government services. Additionally, city departments will be able to encourage more widespread use of 311 through targeted outreach and community education for under-served residents and neighborhoods. Consequently, this work is a significant starting point for more equitable, fair, and effective city service delivery.

5. ACKNOWLEDGMENTS

The authors thank the the Kansas City Office of Performance Management for providing the data used in this study. This work was supported, in part, by a grant from the MacArthur Foundation. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of any supporting institution. All errors remain the authors.

6. REFERENCES

- [1] About 311, our mission. <http://www1.nyc.gov/311/our-mission.page>. Accessed: 2017-07-08.
- [2] Citizen and business satisfaction survey results. <http://kcmo.gov/data/citizen-satisfaction-survey-results/>. Accessed: 2018-07-03.
- [3] B. Y. Clark et al. Do 311 systems shape citizen satisfaction with local government? 2014.
- [4] S. Kang and S. Gearhart. E-government and civic

- engagement: How is citizens' use of city web sites related with civic involvement and political behaviors? *Journal of Broadcasting & Electronic Media*, 54(3):443–462, 2010.
- [5] C. Kontokosta, B. Hong, and K. Korsberg. Equity in 311 reporting: Understanding socio-spatial differentials in the propensity to complain. *arXiv preprint arXiv:1710.02452*, 2017.
- [6] K. Layne and J. Lee. Developing fully functional e-government: A four stage model. *Government information quarterly*, 18(2):122–136, 2001.
- [7] D. L. McClure. Statement of david l. mcclure, us general accounting office, before the subcommittee on government management, information and technology, committee on government reform, house of representatives. *Committee on Government Reform*, 2000.
- [8] S. L. Minkoff. Nyc 311: A tract-level analysis of citizen's government contacting in new york city. *Urban Affairs Review*, 52(2):211–246, 2016.
- [9] D. T. O'Brien. 311 hotlines, territoriality, and the collaborative maintenance of the urban commons: Examining the intersection of a coproduction policy and evolved human behavior. *Evolutionary Behavioral Sciences*, 10(2):123, 2016.
- [10] C. Q. C. K. Wang, Lingjing and S. Sobolevsky. Structure of 311 service requests as a signature of urban location. *arXiv preprint arXiv:1611.06660*, 2016.
- [11] E. W. Welch, C. C. Hinnant, and M. J. Moon. Linking citizen satisfaction with e-government and trust in government. *Journal of public administration research and theory*, 15(3):371–391, 2004.