The Data Open 2022

When Will Your 311 Request Be Resolved?
Reducing Delays in 311 Services

Team 15
Weiji Li, David Lang, Felix Yang, Bo Xia
March 27, 2022

1 Topic question

1.1 Background of 311 Requests

In recent years, many cities including New York, Washington DC and Austin, TX have implemented municipal service requests systems through which citizens can request government assistance for a wide range of issues. These are more commonly known as 311 requests and they include Illegal Parking, Unsanitary conditions and even Refunds for misfiled parking tickets. These requests are systematically filed and made public to citizens as a way to promote engagement with the government and act as a reflection of improvements to the quality of life of citizens.

1.2 Importance of 311 Requests to Quality of Life

In fact, the release of such data to the public has led to various recent research and exploration into how such data reflect the socio-economic and political background of the community. There has been extensive research using 311 request data, from predicting local real estate prices to rationalizing neighborhood distress and its relationship with the condition of neighborhood These indicate that 311 requests are in fact capable of reflecting the needs of citizens and represent a crucial indicator of their quality of life through rising cost of living and even further negative externalities like opioid abuse.

1.3 Significance of our findings

We recognised that the timeliness and efficiency of the government in responding to 311 requests is therefore another important cog in maintaining the quality of life for citizens. More specifically, we wish to investigate the level of compliance with Service License Agreements (SLAs), which indicate the maximum time within which the agency responsible for handling the request will take action on and close a request of a particular category. Timely service is crucial to all stakeholders involved and is key to high customer satisfaction and improving quality of life, while also reducing disruptions to the operational efficiency of businesses. In addition, failing to meet the expected time for delivery of service (DoS) elicits a negative emotion in individuals which exponentially worsens their quality of life on top of tangible inconveniences that result in these service requests. Given the nature of 311 requests are so diverse and common, we seek to provide possible solutions to effectively ensure better service quality and thus quality of living.

1.4 Key Questions

1. **How accurately do the resolution time of 311 services in (NYC and Washington DC) follow their Service License Agreements?**
2. **What are the similarities and differences between the factors of resolution delay time in 311 services in NYC and DC?**
3. **How can we accurately predict the resolution time of 311 service requests?**

2 Executive Summary

2.1 Key Finding

Based on our analysis and comparison between the DC and NYC 311 data, we found that NYC is more consistent than DC in terms of completing service requests by the expected date set by the Service License Agreements with smaller variance in delays for complaints of a single type. However, each 311 request of that type in NYC can be expected to be more delayed in most cases (higher median delay) than for a similar request in Washington DC. However, looking at the data from both datasets, both cities still experienced significant delays in service time for several complaint types.

We also suspected that geography and seasonality are important factors affecting the time delays in both cities and saw that the projected changes in delays motivated actionable insights towards improving the service delay times in both cities.

Finally, we found that in Washington DC, many slow services are having service improvements under initiatives like PaveDC, while most slow services in New York City are not being improved.

2.2 Solutions

Based on the key findings we had from the data, we came up with several applicable solutions to increase the efficiency of 311 requests and predict the resolution time for different 311 requests.

1. Reallocate resources from area of less average delays to area of larger average delays
2. Reduce the SLA time for requests that are resolved much earlier than expected
3. Use the ML model to precisely predict the resolution time for each 311 request to forecast future resource allocation
4. Implement initiatives similar to PaveDC in NYC to reduce delays for highly delayed services like Unsanitary Conditions.

3 Technical Exposition

General Approach:

We did our analysis based on the 311 request data for both NYC and DC. For DC, we also included the 311 request data from 2017 - 2020 that we found online to support our analysis and modeling.

In our analysis, we first identified whether or not there were complaint types in each city for which service was much faster or much slower than the rate detailed in the Service License Agreement (SLA). From that analysis, we dove deeper into characterizing some of the requests that consistently saw delay times.

Next, we investigated what factors could influence the expected delay for each type of complaint. This was done by investigating whether delays differed significantly by zip code, investigating the relationship between the number of such requests and the delay to be expected and finally investigating temporal trends in delays.

3.1 Data Cleaning & Exploration

3.1.1 Data processing

As our analysis is heavily based on the amount of time expected for each request to be completed compared to SLAs, only data for completed requests was used. Requests lacking zip code data were also excluded to allow for analysis of the influence of geography on delays.

To obtain the time allotted for each request by SLAs, we used the Service Due Date column for Washington DC requests. For New York, outside data of SLA expectations from https://data.cityofnewyork.us were used as the expected completion time is not listed for most complaints within the dataset. However, this outside data did not include all types of complaints found in the 311 dataset, so requests of these types were also excluded. Datasets from Austin, Texas were not analyzed as data on SLA agreements could not be found.

3.1.2 Identifying Complaint Types of Interest

To identify service types that were normally delayed, service types that had a median delay of over 1 day were identified. This is because shorter delays were considered far less significant from a practical standpoint as the request would still generally finish on the same day as expected. The median was used as having a median delay over 1 would suggest that at least 50% of such requests experienced significant delays.

In contrast, to identify service types that were completed significantly more quickly, service types that had 95% of requests completed at least a day early were identified. These were hypothesized to be services for which more stringent expectations could feasibly be achieved to elevate standards of service. A higher cutoff was used compared to delayed requests as we prioritized avoiding delays significantly more than raising expectations and thus believed that expectations should only be raised if they can be met at least 95% of the time.

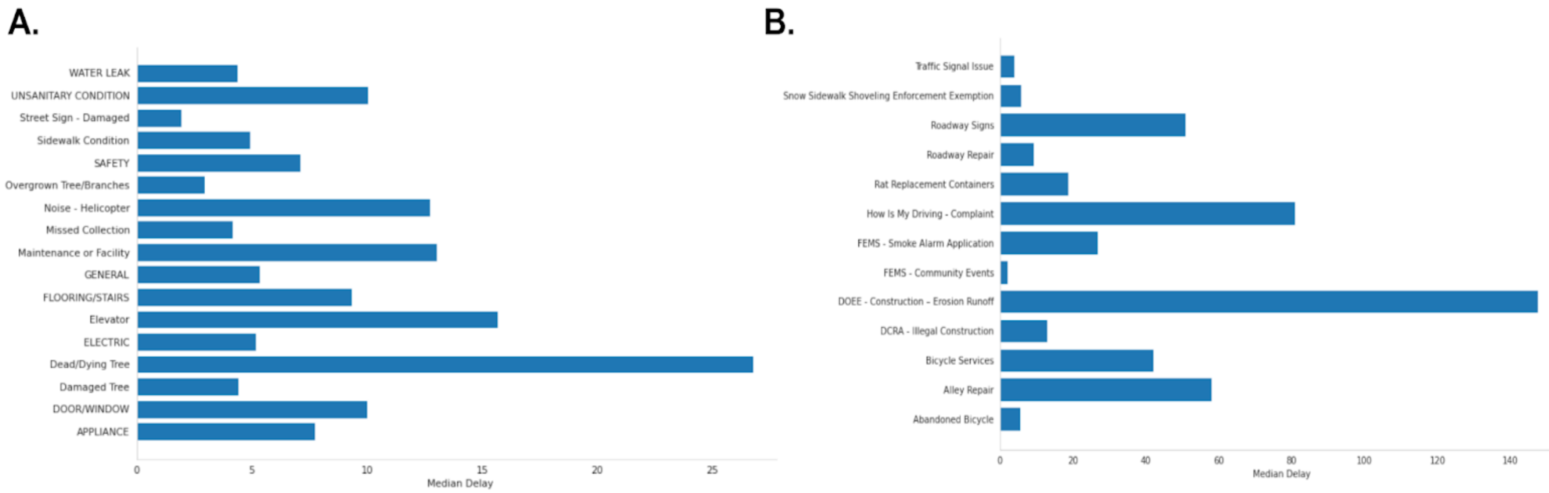

Figure 1. Commonly delayed 311 request types in A) NYC and B) Washington DC and median delay time in days. NYC has more request types that are commonly delayed but median delays for delayed services are significantly smaller than in Washington DC.

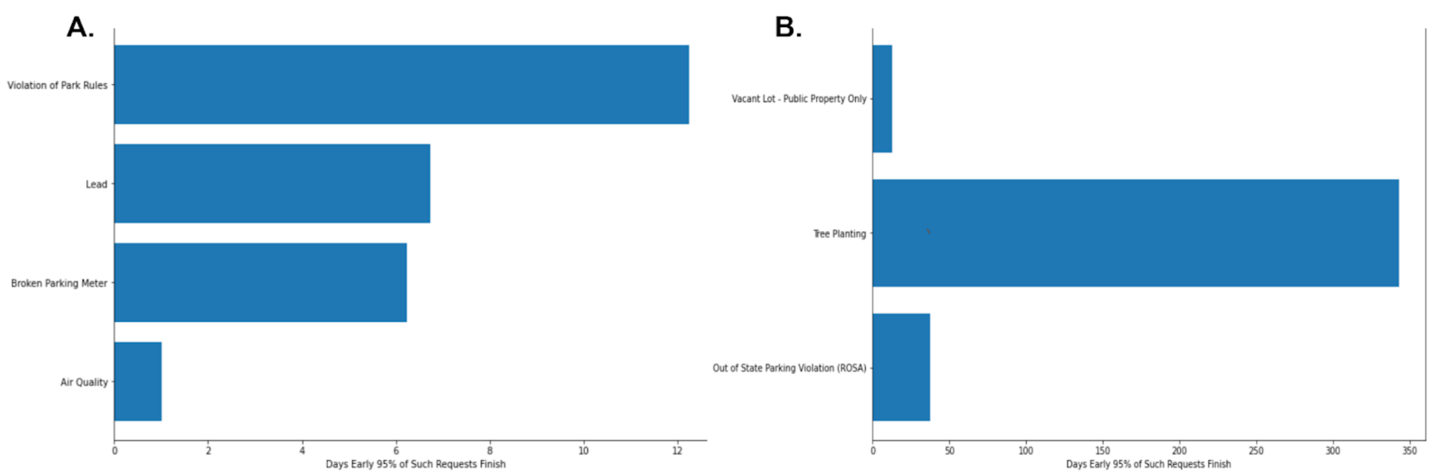

Figure 2. Common 311 request types that are completed early in A) NYC and B) Washington DC and the number of days early which 95% of such requests finish.

The results of this analysis are in Figures 1 and 2. We found that overall, the magnitude of “time deltas”, or the difference between the time for a request to be processed and the time allotted for the request, was much larger in Washington DC. This can be seen as the highest median delay in NYC is around 30 days while the highest in Washington DC is well over 100 days. Similarly, NYC complaints that were completed early were generally at most 1-2 weeks early while in Washington DC, certain types of requests such as Tree Planting and ROSA could be expected to be completed months in advance. However, we found that more types of requests commonly experience delays in NYC, with 17 request types commonly being delayed in comparison to 13 in Washington DC. This suggests that the time it takes for a specific request type in NYC varies less but is more delayed in most cases than in Washington DC. Analysis of plots of the distribution of delays for the same type of request in each city supported this hypothesis, as shown in Figure 3.

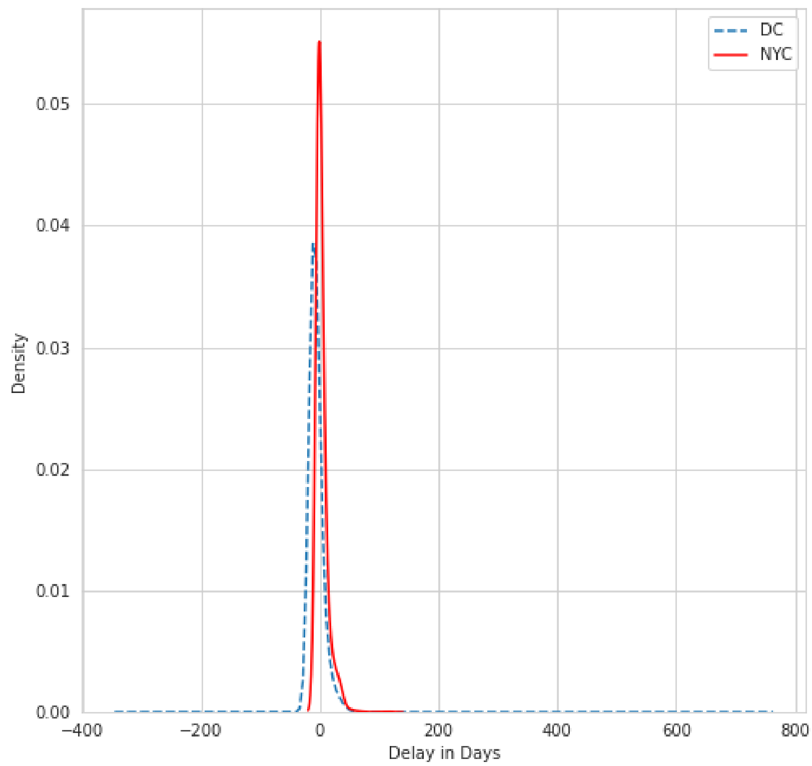

Figure 3: Density plot of delays for "Illegal Dumping" requests in Washington DC and New York City. The range of delays for Illegal Dumping requests is much larger in Washington DC, but the peak of the density plot occurs earlier, suggesting that the request may be less delayed in most cases.

3.2 Factors influencing Delays

We hypothesized that other than the type of complaint, the geographic location of the request, the date and time it was submitted could influence the time needed for a request to be processed. This is because certain weather related complaints like snow removal would logically take longer in winter months, while the distance of the complaint from offices for departments handling such complaints might affect the amount of time needed to process the request.

To analyze the effect of geography on delays, we created a heatmap of the mean delay for each service type for each zip code. The results of this analysis are in Figure 4. It can be seen that for both cities, delays differ by service type, but that the difference is more significant in Washington DC.

Interestingly, for both cities, there are certain complaint types where service does not depend much on zip code while other complaint types have service can depend significantly on zip code. An example of this is shown in Figure 5.

Overall, these results support our hypothesis that geography can influence service times. However, it is unclear why certain request types have delays independent of geography, which could be a topic of future research.

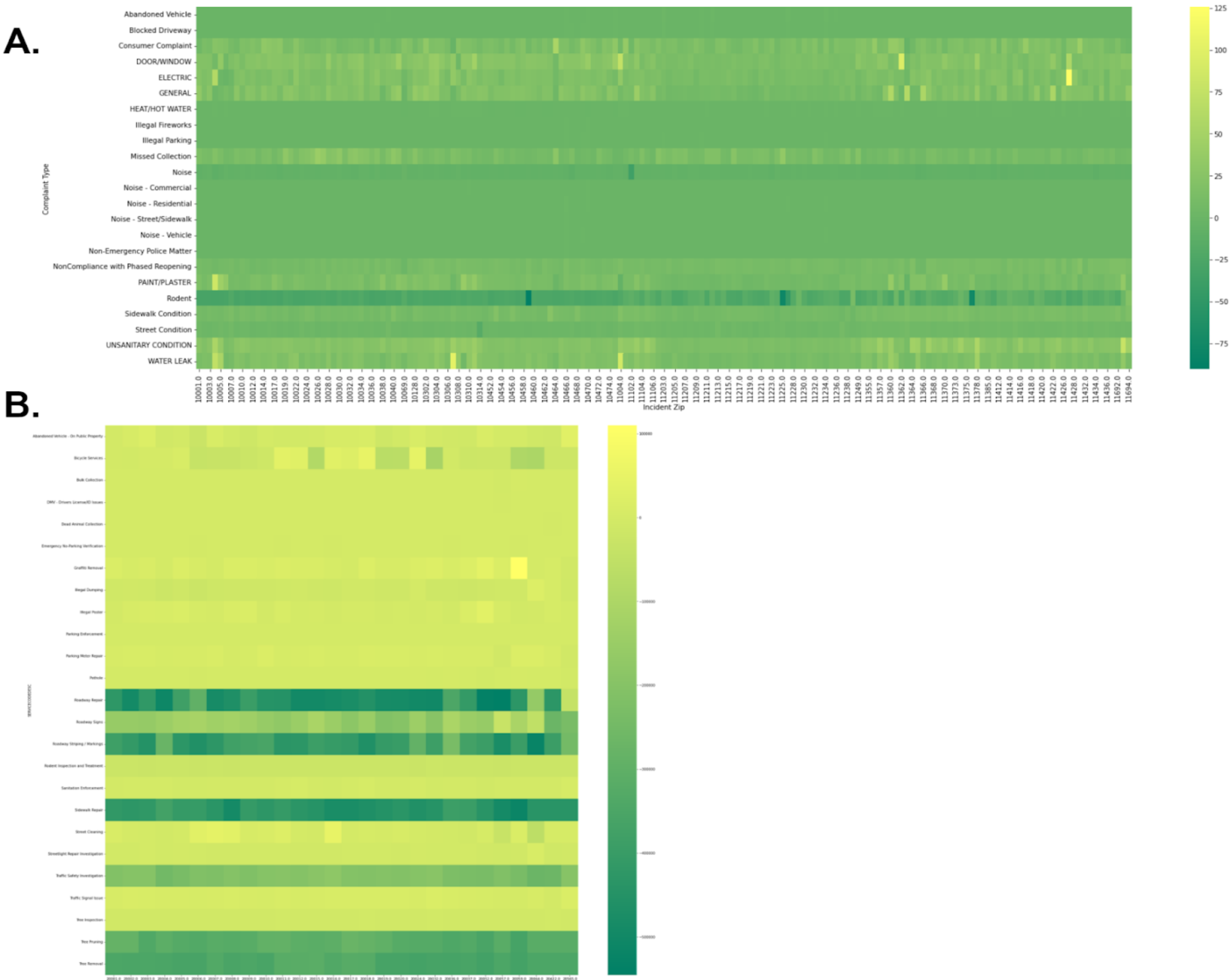

Figure 4: Heatmaps of mean time delay by service type and zip code for the 25 most common service types for A) NYC and B) Washington DC.

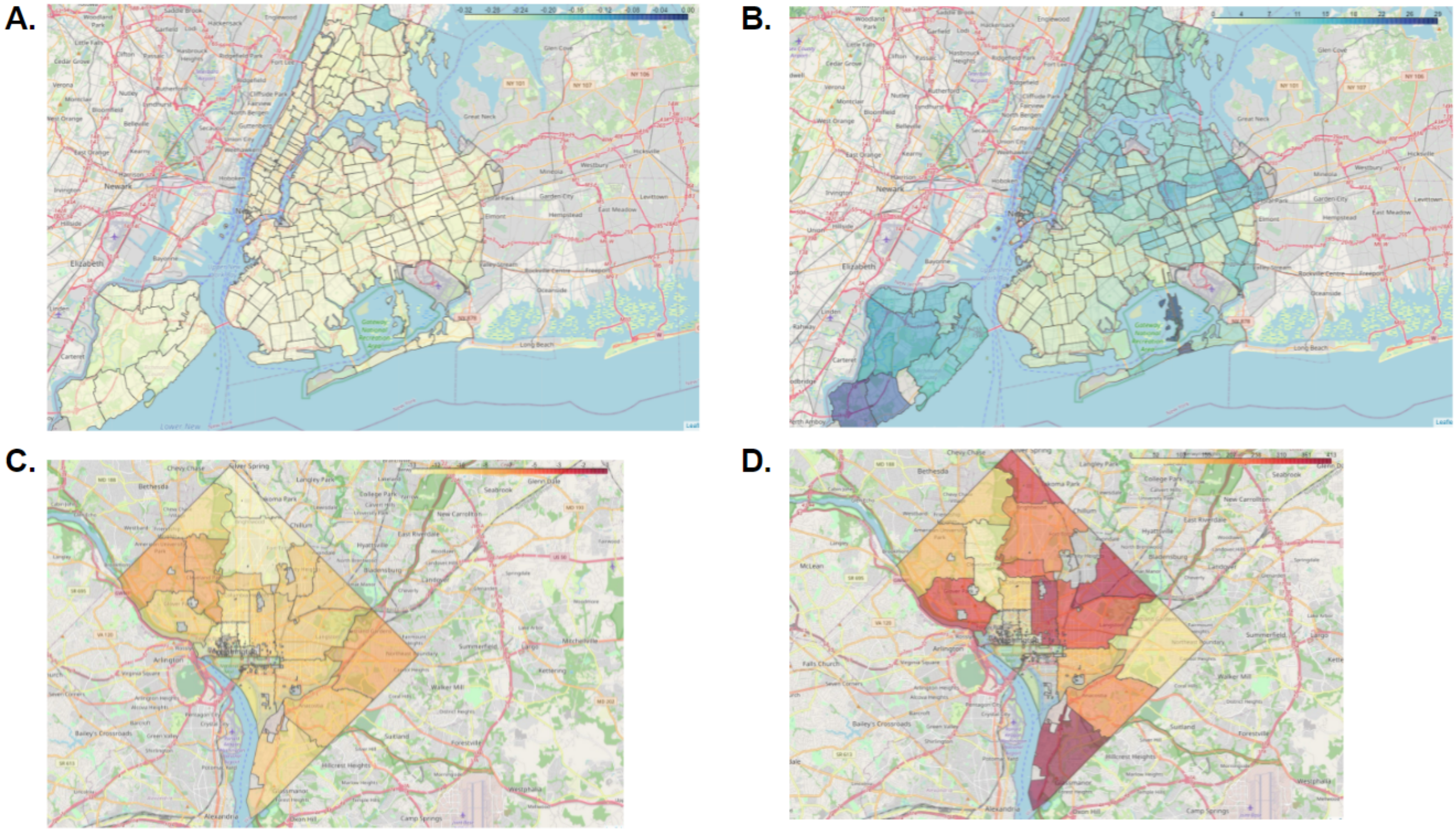

Figure 5: Certain service types experience geography dependent delays. A) Vehicle noise complaints in New York and C) Bulk collection complaints in Washington DC, have delays that do not vary significantly based on zip code. B) Unsanitary condition complaints in New York and D) Roadway repair complaints in Washington DC have delays that vary based on zip code.

We also wanted to investigate the effects of the number of similar requests in the same zip code on the delay time. To do this, the Spearman correlation coefficients and their respective p-values between the proportion of requests of a certain type in each zip code and the median time delay for that request type in that zip code. We hypothesized that if the correlation was positive, areas with many requests of that type (for example, bulk collection) are causing requests to be backed up and thus there may need to be reallocation of workers to reduce these delays. In contrast, if the correlation is negative, resources may be allocated too heavily into regions where such requests are common. From the coefficients and p-values calculated, Roadway Repair requests saw a 0.25 correlation coefficient with p-value 0.0027. In addition to this, 10 out of the top 25 most common request types saw a statistically significant correlation between the average delay time of the request and the amount of requests that were being made on a zip code basis which illustrates the impact of the number of similar requests based on the request type.

Finally, we investigated the effects of the month and year at which the complaint was filed on the delay time to investigate the seasonality of delays as well as whether there are trends in the delay time. For each request type found to be commonly delayed in Figure 1, the median time delay for each month of data available was plotted (from 1/2021-3/2022 for Washington DC and 1/2020-3/2022 for NYC). Then, the trend in the delay was classified as either decreasing, stable or increasing based on the spearman correlation coefficients between the date and median delay. The results of this analysis are in Table 1 and examples of the plots generated can be found in Figure 6.

Washington DC		
Decreasing Delay	**Stable Delay**	**Increasing Delay**
Alley Repair	Abandoned Bicycle	
Bicycle Services	DOEE - Erosion Runoff	
How is my driving - Complaint		
Roadway Repair		
Roadway signs		
Traffic Signal Issue		

New York City		
Decreasing Delay	**Stable Delay**	**Increasing Delay**
Electric	Unsanitary Condition	Sidewalk Condition
Street Sign - Damaged	Water Leak	
	Safety	
	Appliance	
	Door/window	
	Damaged Tree	
	Elevator	
	Flooring/stairs	
	General	
	Maintenance or facility	
	Overgrown Branches	

Table 1: Trends in delays for commonly delayed services. Spearman correlation coefficients between date and median delay in days for each service type were calculated and the sign and p value used to categorize each service type as having increasingly long delays (+ sign, p<0.05), having decreasing delays (- sign, p<0.05) or having stable delays (p<0.05).

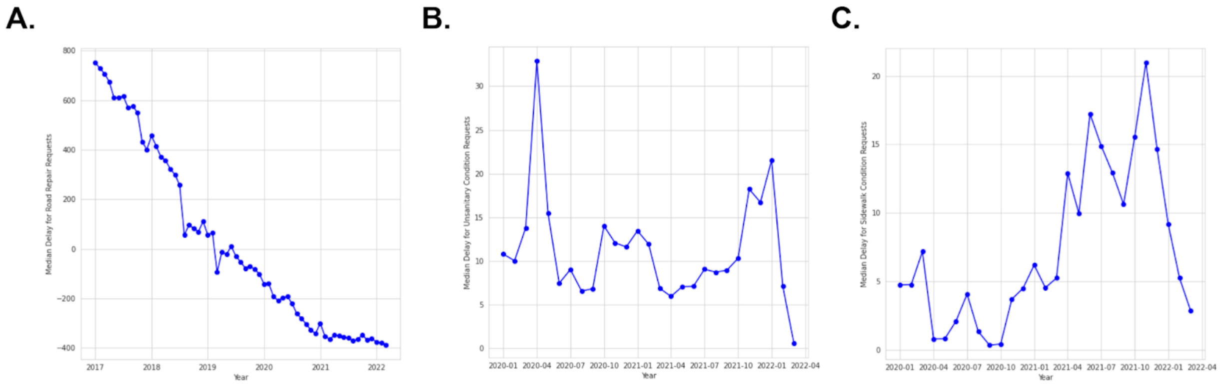

Figure 6: Examples of A) Decreasing delay (Roadway Repair in DC), r = -0.99, p = 5.4e-55, B) Stable delay or no trend (Unsanitary Conditions in NYC), r = -0.15, p = 0.47, C) Increasing delay (Sidewalk repair in NYC), r = -0.593, p = 0.0011.

Interestingly, Washington DC has most of its currently slow services being improved at a rapid rate. As seen in Figure 7A), the median delay for roadway repair complaints has decreased from over 600 to -400 days between 2017 and the present. This specific finding may be linked to the PaveDC initiative, which was launched in 2015 and uses 311 data to systematically prioritize roadway repair requests and has led to record numbers for miles of roadway paved in 2019[1].

3.3 Key Insights from Data

- In both cities, there are services that can be expected to be delayed and requests that can be expected to be completed early
- 311 delays vary with season and zip code for certain times of complaints.
- NYC seems to have a more mature 311 system with more consistent service times. However, this also means that commonly delayed services are not having their delays reduced year after year.
- DC has much more inconsistent service, with service occasionally being extremely fast/slow, but is seeing the median delay for consistently delayed services decrease
- Initiatives like PaveDC can significantly reduce delays users experience

Recommendations:

- Requests that can be expected to finish significantly earlier such as lead in NYC and Tree Planting in Washington DC may want to implement more stringent SLAs to increase quality of life for citizens
- Governments may want to use data of delay distributions across seasons and zip codes to reallocate workers where they are needed most to provide timely service
- DC should implement more consistent filing/quality control to make sure there is more consistency in request completion times.
- NYC may want to consider launching initiatives similar to PaveDC to reduce wait times for their commonly delayed services, focusing on the most detrimental delays such as for complaints involving safety, unsanitary conditions.

3.4 Prediction Model

Seeing the issues about the misalignment between expected complete time versus the actual complete time, we decide to build an AI model to predict the expected complete more accurately. As a result, the customers can have a better user experience when knowing a more accurate estimated complete date.

3.4.1 Task Statement

First of all, we need to clearly define the task. We want to **predict the timeframe for the resolution time**. Timeframe means "within 12 hours", "within 1 day", "within in 3 days", etc. We did the classification task here, instead of regression. The reason is that customers usually just need the timeframe information and the exact time is impossible to predict given the entropy of the dataset. The label to predict is: "6 hours", "12 hours", "1 day", "3 days", "7 days", "30 days" and "more than 30 days".

3.4.2 Dataset

We used our complete collection of DC 311 requests data from 2017 to 2021. It has 1662176 non-empty data points with both expected resolution date and actual resolution date. We use an 80/20 train-test split to train and test our model.

3.4.3 Feature Selection

Next, we need to select features to train our AI models. From our previous analysis, we found that some factors have some relationship with the actual resolution time. For example, the **category of the requests** has been seen to have a relationship with the actual resolution time. For instance, the time for power outages is usually less than that for roadway service. Moreover, the **time in the day of the request** will influence how quickly the request could be handled. After an iterative testing approach, we decided on the features listed in Table 2 to train our model.

Parameter	Type
Service Code Description	Categorical
Service Type Code Description	Categorical
Organization	Categorical
Zip Code	Categorical
Created Hour	Categorical
Created Month	Categorical
Service License Agreements	Continuous

Table 2: Features used to train the model

3.4.4 Model

Our next step is to find an appropriate model for our task. We initially attempted a linear classifier and an SVM classifier. Lastly, we trained a neural network classifier (**MLPClassifier from sklearn**). After the experiment, we concluded that the neural network performs the best. We believe the reason for that is the result and the features have a really complex and non-linear relationship. Also, some features might have a relationship with each other. It can also infer unseen relationships on unseen data as well, thus making the model generalize and predict on unseen data.

Loss Function:

$$- \sum_{c=1}^{M} y_{(o,c)} log(p_{o,c})$$

Symbol	Definition
M	Number of classes
o	Observation
c	Class label from the model
log	Natural logarithmic
y	Binary indicator (0 or 1) if class label **c** is the correct classification for observation **o**
p	Predicted probability observation **o** is of class **c**

Table 3: Loss function symbol

3.4.5 Model Results

The accuracy that we achieved from the model is **65.64%** on the test dataset. We also calculated the previous expected resolution time, which is at only 37.34%. It demonstrated that our model achieves a nearly double accuracy on prediction the resolution time.

Figure 7 shows the distribution of the number of 311 requests of different time frames, which includes the actual time frame, predicted timeframe and the expected timeframe from the dataset. We can see that there is much closer alignment between the actual and predicted than the actual and expected. Besides, we also added an additional category denoting "within 6 hours", which could give the customers a more precise estimation of resolution time.

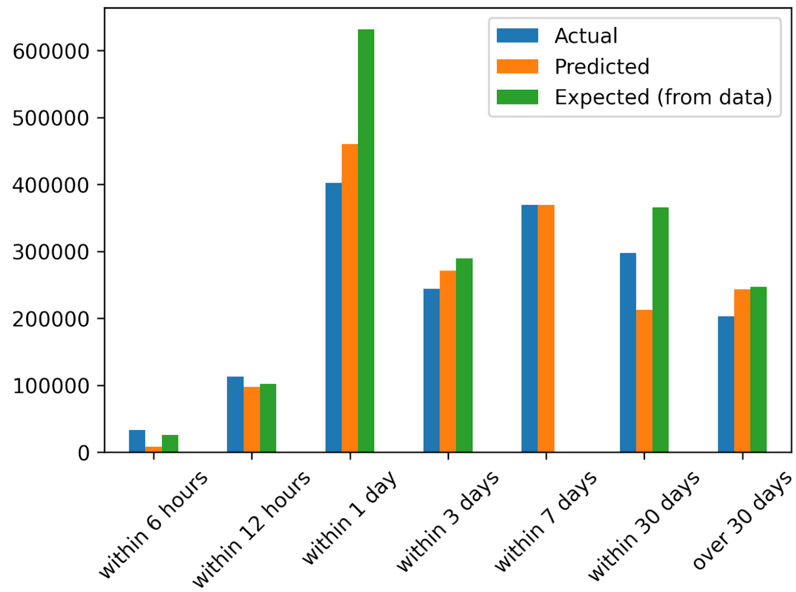

Figure 7: Number of 311 Requests of Different Timeframe

3.4.6 Limitations

Despite the higher accuracy we achieved from the neural network, there are some tradeoffs in return which is loss of interpretability. We do not know which of the factors in the NN contributed most to the prediction of resolution time due to the nature of the multi-layered NN. Another limitation in our methodology is also the fact that we trained it mainly using DC data to train the model which might not be generalizable to other contexts like NYC.

3.5 Key Insights from Model

- There is inefficiency with regards to the expected time provided by agency
- Our model for the 311 request system will largely increase the alignment between actual and predicting resolution time.
- Better allocation of resources into achieving an expected resolution time or providing more realistic resolution time expectations for the citizens.

4 Areas for Further Analysis

4.1 Different prediction models for interpretation

We could also incorporate decision trees for more interpretable models to non-technical audiences. On top of that decision trees could also act as a feature selection tool in earlier stages to create a better NN model.

4.2 Further Time Series Modeling

We could also conduct further times series analysis using ARIMA models and time series regression to predict counts of different requests in the future and compare it with our currentmodel. Instead of a normal training test split, we could also use forward-chaining as a form of splitting our data.

4.3 Demographic and population

Another possible direction could be to include data on demographics of each region. We can add extra dimensions to our models in terms of predicting what kinds of issues are encountered by different demographics. This could add more meaning to the service problems faced in different regions. On top of that population density could be useful for different cities since DC and NYC are all urbanized cities while there might be differences in types of issues and even time delta for rural regions.

4.4 Geospatial Research

There could be more efforts done with respect to finding the relationship between service requests and geography. Since zip code is only one proxy of geography and remains at a primitive level, other more detailed forms of geographical features could be incorporated into the analysis and modeling in the future.

5 References

Additional Data:

Open data DC. Open Data DC. (n.d.). Retrieved March 24, 2022, from https://opendata.dc.gov/datasets/

City of New York, N. Y. C. O. D. (n.d.). *NYC open data*. NYC Open Data WP Engine. Retrieved March 27, 2022, from https://data.cityofnewyork.us/

External Research and Findings:

1. *GIS helps prioritize repairs for Washington, DC's heavily trafficked roadways*. Esri. (2020, June 25). Retrieved March 26, 2022, from https://www.esri.com/about/newsroom/arcnews/gis-helps-prioritize-repairs-for-washington-dcs-heavily-trafficked-roadways/
2. Li, Y., Hyder, A., Southerland, L. T., Hammond, G., Porr, A., & Miller, H. J. (2020, November 11). *311 service requests as indicators of neighborhood distress and opioid use disorder*. Nature News. Retrieved March 27, 2022, from https://www.nature.com/articles/s41598-020-76685-z
3. Minkoff, S. L. (2015). NYC 311. *Urban Affairs Review*, *52*(2), 211–246. https://doi.org/10.1177/1078087415577796
4. Wang, L., Qian, C., Kats, P., Kontokosta, C., & Sobolevsky, S. (n.d.). *Structure of 311 service requests as a signature of urban location*. PLOS ONE. Retrieved March 27, 2022, from https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0186314
5. White, A., & Trump, K.-S. (2016). The promises and pitfalls of 311 Data. *Urban Affairs Review*, *54*(4), 794–823. https://doi.org/10.1177/1078087416673202
6. *23-305 service level agreements.* American Legal Publishing Corporation. (n.d.). Retrieved March 27, 2022, from https://codelibrary.amlegal.com/codes/newyorkcity/latest/NYCadmin/0-0-0-114437