

# Predicting Resolution Time of 311 Calls in Minneapolis

By Langan Dee

## *What is the problem you want to solve?*

How can the City of Minneapolis improve the way they allocate resources and handle non-emergency service requests? Minneapolis, like many cities, has a 311 phone number that people can call to find information about the city's services, make complaints, or report problems. Questions can be related to almost any of the city's departments or services. For example, residents can ask about trash collection, voting locations, or zoning permits. They can report graffiti, potholes, stray animals, water pollution, or abandoned cars. Each call (or "case") is tracked by type, location, and date the issue was resolved. I will identify the most common cases and predict how long it will take to resolve each.

## *Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn't have done otherwise?*

My client is the City of Minneapolis. By identifying the most common 311 information requesting calls, the City can determine key information to make more prominent on their website. The City can use case resolution time predictions to decide where to efficiently allocate department resources. This has the potential to reduce staffing budget in multiple departments and improve response time to residents' concerns.

## *What data are you using? How will you acquire the data?*

The data used for this project comes from the City of Minneapolis Open Data policy, which provides records of 311 calls made each year. The data is available on [Open Minneapolis](#) and includes Case Type, Subject Name/Title, Date and Time, Geolocation, Resolution ("Open"/"Closed"), Date and Time closed. I will use information from the last 3-5 years.

Rows:

- 2015-2017: The datasets include 49,897-55,164 rows of data per year.
- 2018: The data is updated daily and could be used as a test dataset.

Columns:

- **CaseID**
- **OpenDateTime, ClosedDateTime, CaseStatus** - This indicates if and when a service request was resolved and closed
- **SubjectName** - This is closely associated with the department designation (ex. Vehicles and Commuting, Public Safety, Animal Related, etc.)

- **SubjectReason, SubjectType** - 2nd and 3rd level classification and description of the service request
- **XCoord, YCoord** - Coordinates of where the service request occurred

Historical weather data will also be incorporated as an extra dimension. Daily weather summary datasets will be collected from [Climate Data Online](#). I will explore the correlation of daily weather conditions (max and minimum temperatures, total precipitation, and snowfall) with incoming call types and resolution time.

*Briefly outline how you'll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.*

Part of the data wrangling process will be to decide how to handle calls that are missing data (such as location), and calls that are duplicated. For example, multiple neighbors might report a street light is out on their block. I might discard the duplicate calls, or find that it's a factor in determining how quickly a case will be closed.

To explore the data, I'll use a combination of inferential statistics and data visualization to examine each factor, such as type of call, time of day, month, weather conditions and location of service request. This could include mapping locations of the calls, and using time series.

After identifying trends and significant variables, I'll use ML to build a model to predict whether a case will be resolved, and how quickly.

*What are your deliverables? Typically, this includes code, a paper, or a slide deck.*

My deliverables will be a jupyter notebook and slide deck published to my github account. The note will include a report of my findings and related python code.