How to find min sup for 311 requests?

### Analyze Data:

(some requests are not taken into account because they do not have a neighborhood assigned)

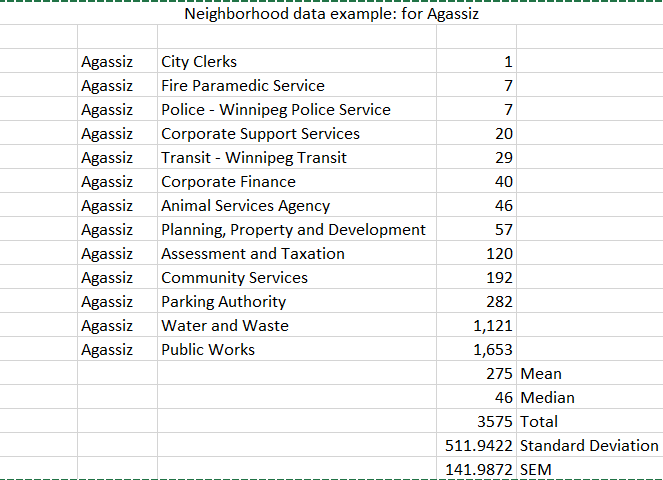
There are 243 different neighborhoods in Winnipeg, the neighborhoods receiving the most requests is:

| William Whyte | 73,431 |
| --- | --- |

And the neighborhood receiving the least request is:

| West Perimeter South | 13 |
| --- | --- |

In each neighborhood we can calculate the median and mean of each type of requests, using Agassiz as an example:



Since the mean is really high(rule out every other request type except for the top 3, or do we want that?)

We could use Minsup = median/total requests per neighborhood for the minimum support for this neighborhood.

* In this case minsup = 46/3575 = 0.0128

We could also use Minsup = mean/total requests per neighborhood for the minimum support for this neighborhood

* In this case minsup = 275/3575 = 0.076

For more data refer to: <https://docs.google.com/spreadsheets/d/1hkuQD_pZqRjt8_2ZrWkNUmNvHiFVfaw5/edit?gid=149167380#gid=149167380>

1. How to auto adjust minsup for different sizes of data?

We can use a formula to set minsup higher when the data size is lower.

Refer to <https://data-mining.philippe-fournier-viger.com/how-to-auto-adjust-the-minimum-support-threshold-according-to-the-data-size/>