

Exploring Seattle “Find It, Fix It” requests with Python and NLP

Alternative title #1: how a Facebook argument accelerated my plan to learn NLP methods

Alternative title #2: you know what they say about assumptions....

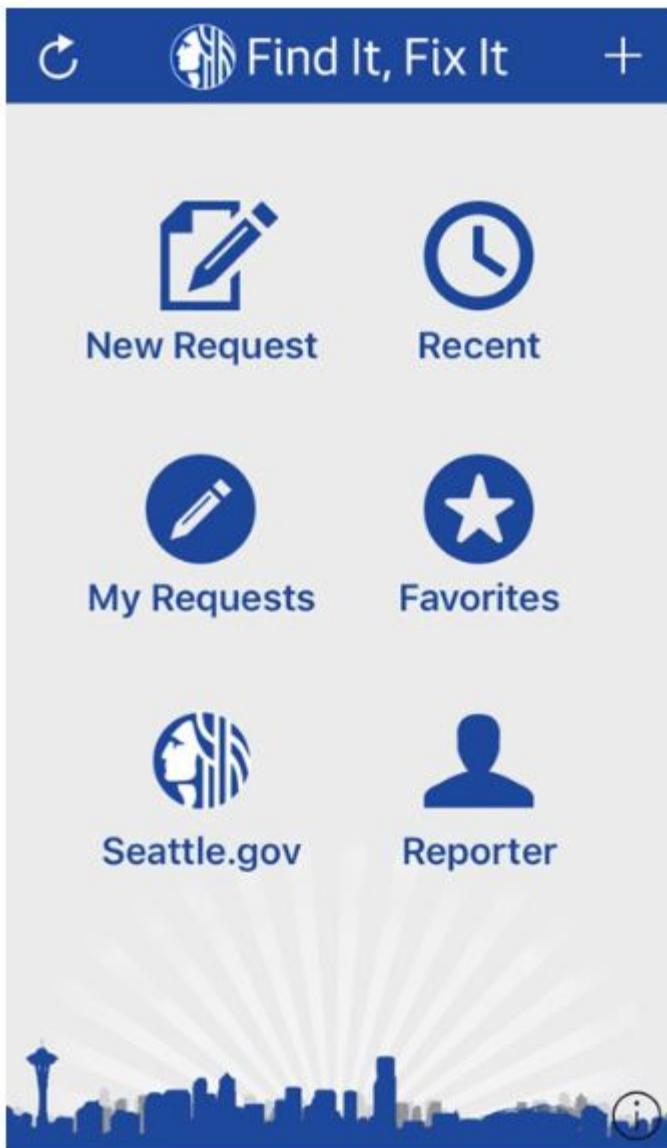


Abigail Schindler
pronouns: she, her, hers



This presentation was prepared using personal time/capacity. The opinions expressed in this article are my own and do not reflect the view of the National Institutes of Health, the Department of Health and Human Services, the Department of Veterans Affairs, or the United States government.

Find It, Fix It Service Request Mobile App



Find It, Fix It Service Request Mobile App



Why we can't have nice things....



Richmond Police
@TheQueerCrimer

Yesterdays Homeless Sweep

Search Menu



Seattle.gov
Mayor Jenny A. Durkan

Reporting - The City learns of camping/unauthorized encampments in several ways, including:

- Through calls from the public to the Customer Service Bureau, reports filed using the City's online Service Request Form and through the Find It, Fix It mobile app.
- From the Seattle Police Department and other City staff out in the field who encounter people camping on City-owned property.



3:36 PM · 7/20/19 · Twitter for iPhone

"This poster is not from the city, some asshole just made it up and said to use the app, which does not even have an option to report a homeless camp. They aren't doing sweeps based on reports and overloading the app will only get in the way of the people trying to respond to legitimate issues."

FIFI Seattle Public Disclosure request

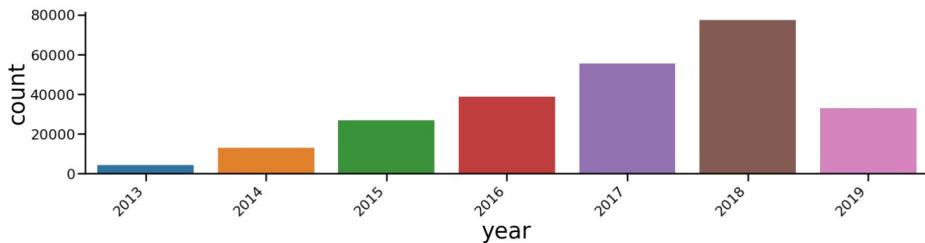
(248885, 30)

Out[6]:

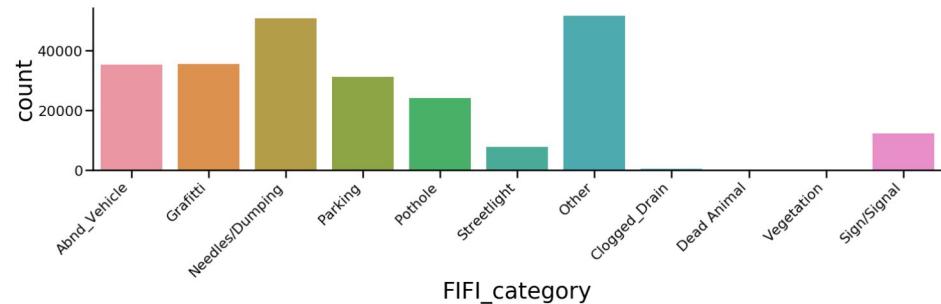
	Service_Request_Number	Created_Date	Location	Location_Details	Description
0	13-00036779	2013-05-09 20:32:33	5208 BEACON AVE S, SEATTLE, WA	5200 Beacon Ave S Seattle, WA 98108; XY: 12775...	Ford light blue
1	13-00064098	2013-08-13 18:05:17	2629 SW ANDOVER ST, SEATTLE, WA	26012799 SW Andover St, Seattle; XY: 1261975.9...	NaN
2	13-00064109	2013-08-13 18:57:44	402 BELMONT AVE E, SEATTLE, WA	327 Belmont Ave E, Seattle; XY: 1272796.132329...	I don't think this car has been moved in quite...
3	13-00064128	2013-08-13 20:24:22	5283 39TH AVE S, SEATTLE, WA	5239 39th Ave S Seattle, WA 98118; XY: 1282330...	Older Honda with flat tire

FIFI EDA

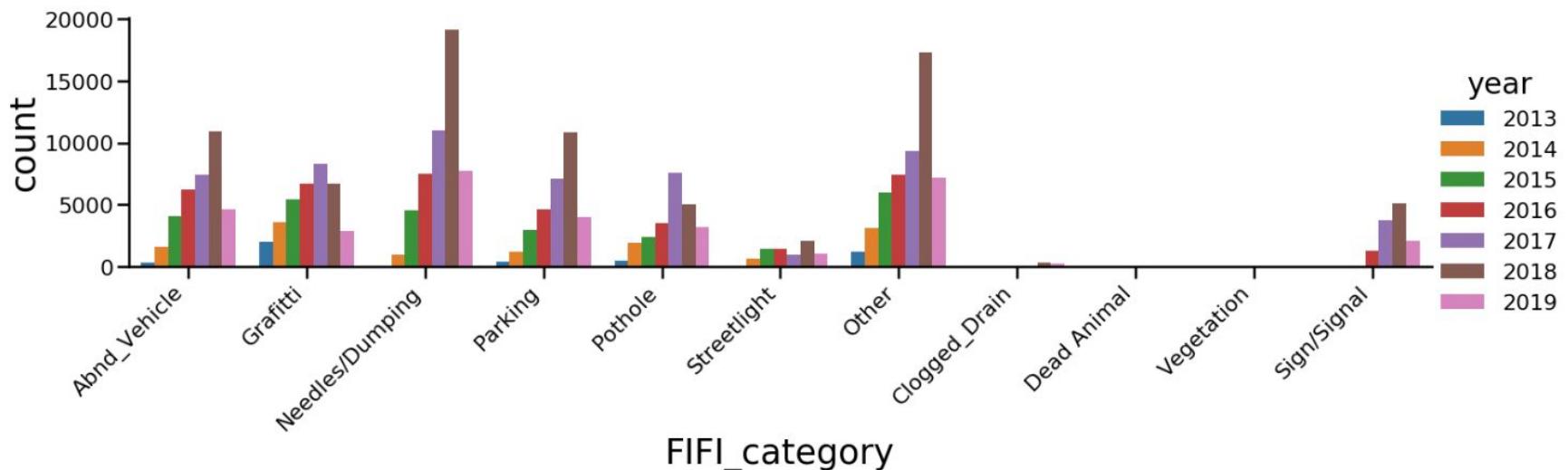
FIFI requests by year



FIFI requests by category

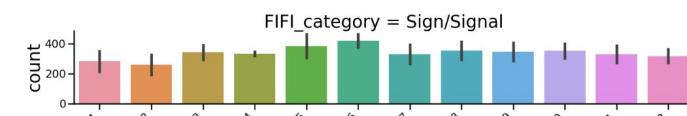
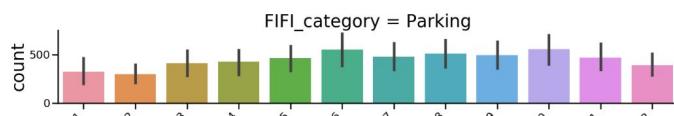
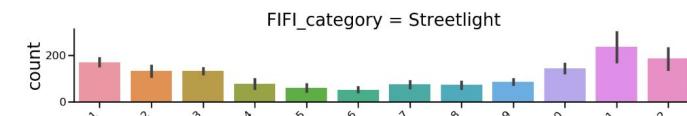
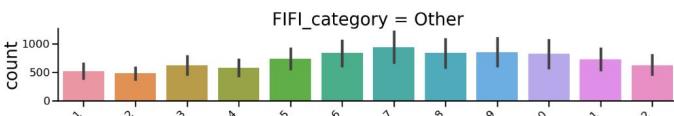
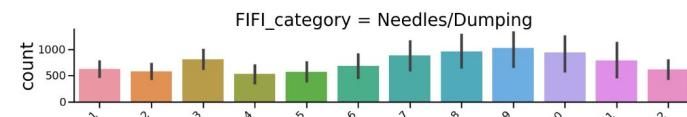
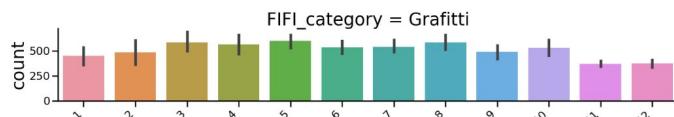
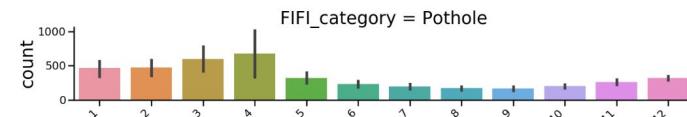
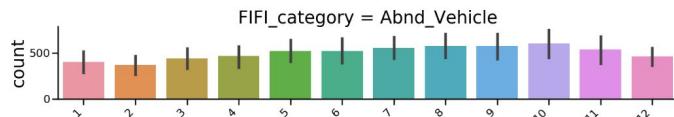


FIFI requests by category and year

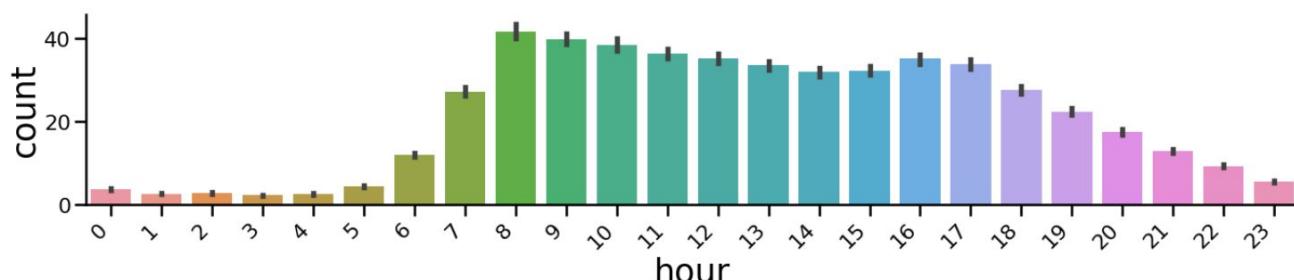


FIFI EDA

FIFI requests by month and category



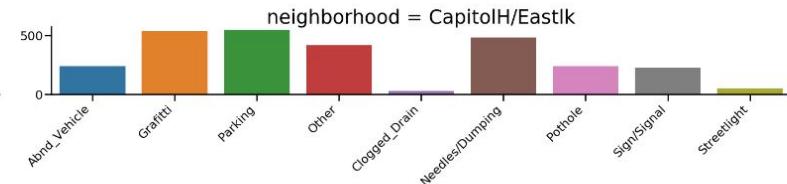
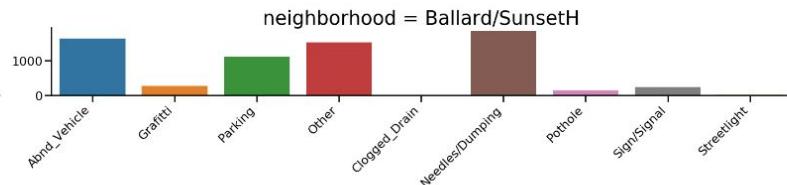
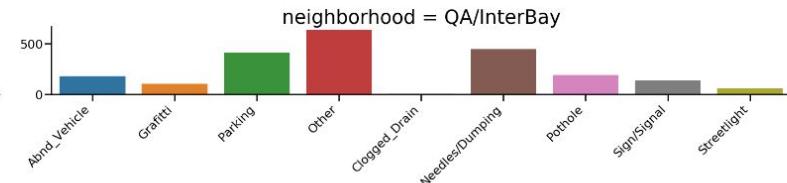
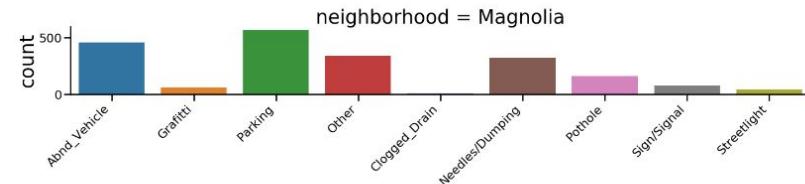
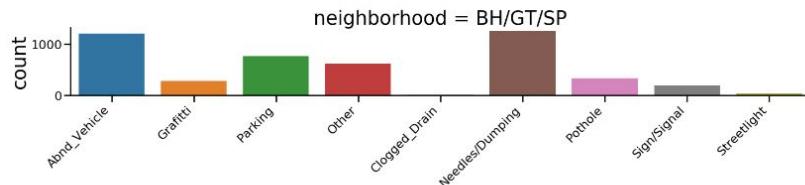
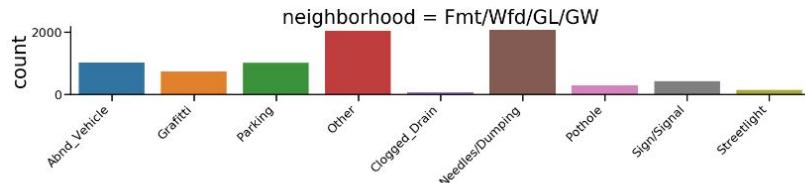
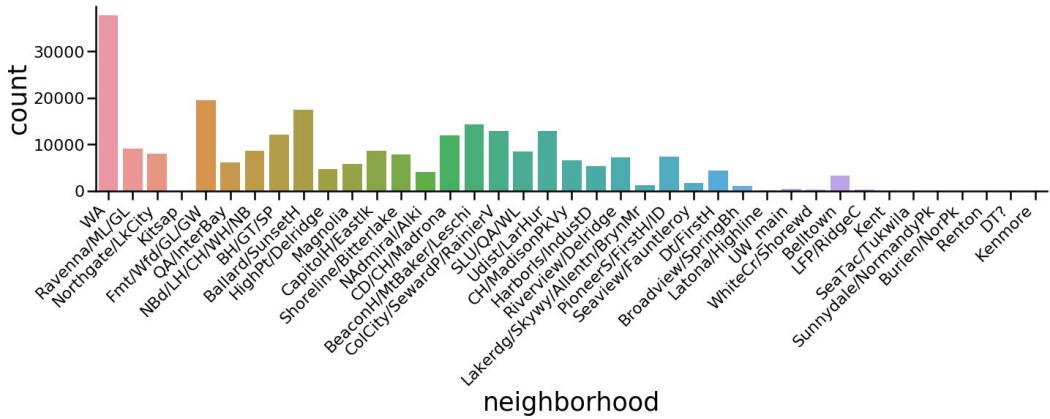
FIFI requests by time of day



FIFI EDA

	Location	count
22076	1906 9TH AVE, SEATTLE, WA 98101	322
30700	2480 BIRCH AVE N, SEATTLE, WA 98109	309
84355	8456 7TH AVE SW, SEATTLE, WA 98106	205
38774	311 DENNY WAY, SEATTLE, WA 98109	181
58606	4993 7TH AVE NE, SEATTLE, WA 98105	180
77169	731 S FIDALGO ST, SEATTLE, WA 98108	179
57711	4825 S MYRTLE ST, SEATTLE, WA 98118	177
53602	4455 SHILSHOLE AVE NW, SEATTLE, WA 98107	174
51760	4301 8TH AVE NW, SEATTLE, WA 98107	168
56108	4700 9TH AVE NW, SEATTLE, WA 98107	168
50304	4200 36TH AVE S, SEATTLE, WA 98118	163

FIFI requests by neighborhood



FIFI NLP

- Sentiment polarity score
- Processing pipeline
 - Remove punctuation
 - Make lowercase
 - Tokenize
 - Remove stop words
 - Lemmatize
 - Assign POS
- Word freq viz
 - FreqDist
 - WordCloud
- Topic modeling
 - Latent Dirichlet allocation (LDA)
 - gensim

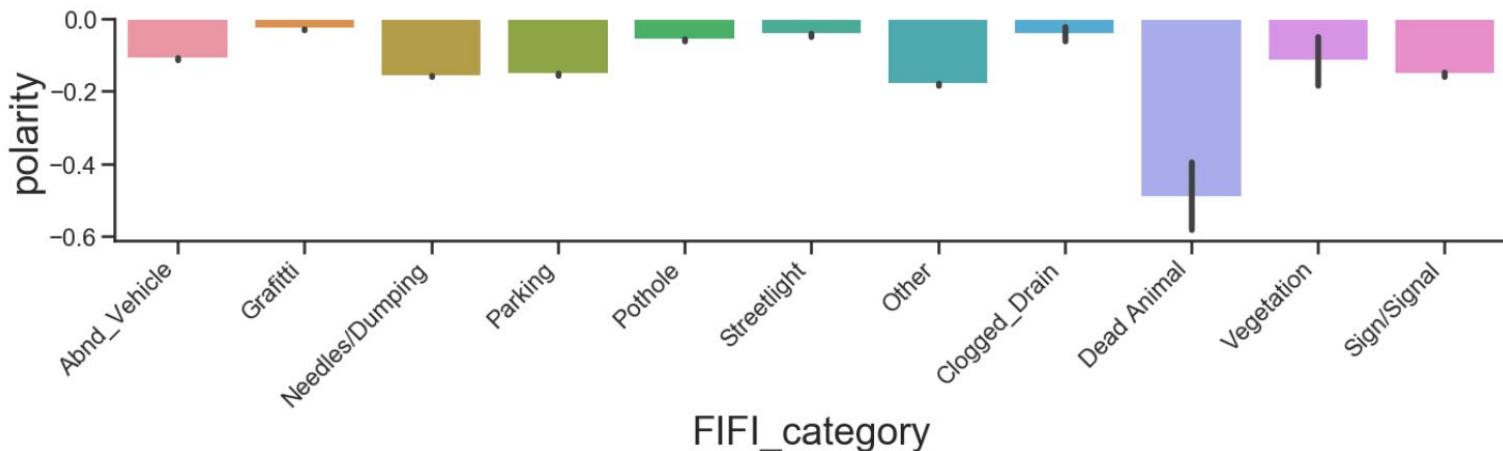
```
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.stem import WordNetLemmatizer
from nltk.sentiment.vader import SentimentIntensityAnalyzer
from nltk.probability import FreqDist

from wordcloud import WordCloud

import gensim
from gensim import corpora
```

FIFI Sentiment Analysis

Sentiment of FIFI requests by category



```
In [10]: data[data['polarity'] == data['polarity'].max()]['Description']
```

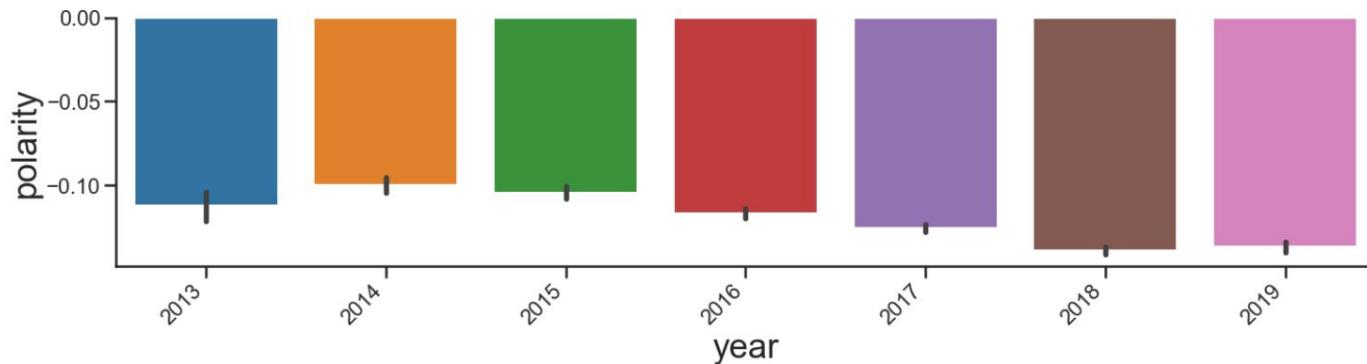
```
Out[10]: 215687    THANK YOU!!! for fixing this light!  It is terrific!! And really does the job.  I hope you can refer to this feedback that we, at Horizon House are VERY happy and appreciative of you following through and getting this one done. Thanks also for fixing the light across the street which seemed like a more straight forward repair or bulb replacement. But this light apparently need a whole new, different style fixture with which I am very pleased! Many thanks to everyone one involved to get this done!!!\nJohn Daniels, 206-679-0382.  Feel free to use my name about the success of "Find It, Fix It.  Thank You.  
Name: Description, dtype: object
```

```
In [11]: data[data['polarity'] == data['polarity'].min()]['Description']
```

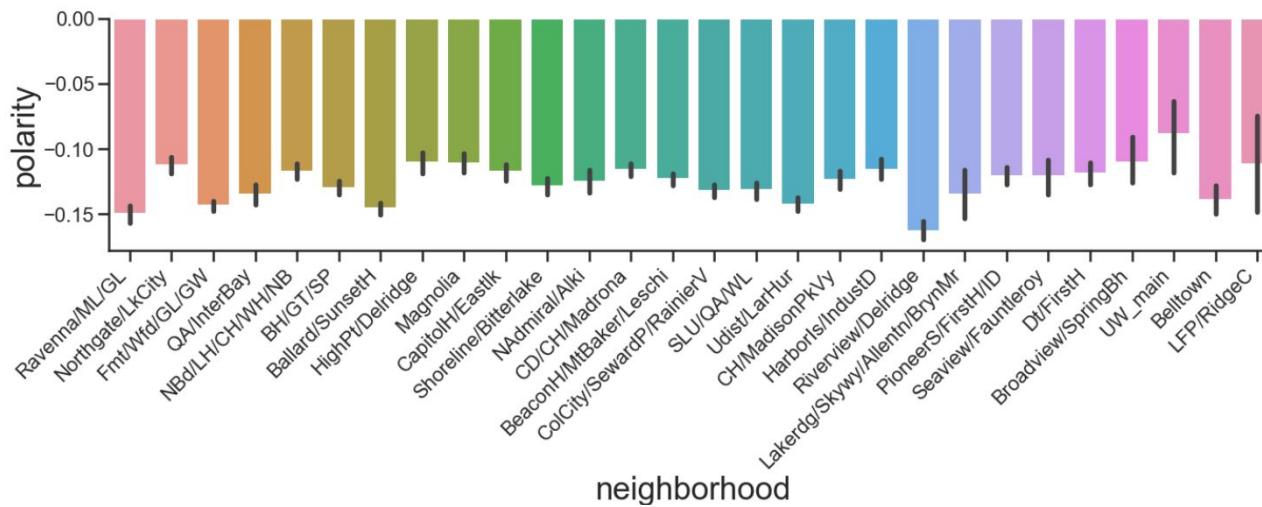
```
Out[11]: 164220    The referenced refers to curbside parking permissible between 3-6 pm M-F. On Rainier Ave S & S Kenyon to S Rose Street. I commute Rainier Ave S every day for the past ten years. There is no such other dangerous traffic flow obstacle as this. Rainier ave S has become horrendous for traffic volume and especially road rage. This small strip of permissible parking chokes off south bound commute traffic at a critical time and location. When people realize there are a couple cars ahead blocking the already congested lanes the rage to swerve left into the other lane is terrifying and very dangerous. Then once past the two block obstacle the rage is to get back into the right lane and get ahead of everyone. And of course in both instances commuters are resisting the lane change even though the car on the right is head straight for a collision with a parked car. Those who park on this narrow strip routinely disregard the time restrictions. And when outside the time restriction routinely will stop  
Name: Description, dtype: object
```

FIFI Sentiment Analysis

Sentiment of FIFI requests by category

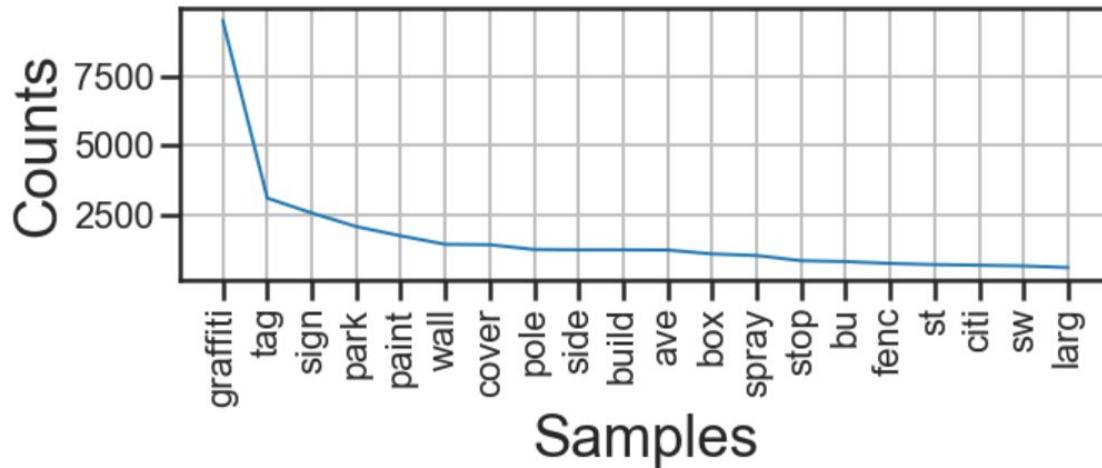


Sentiment of FIFI requests by neighborhood

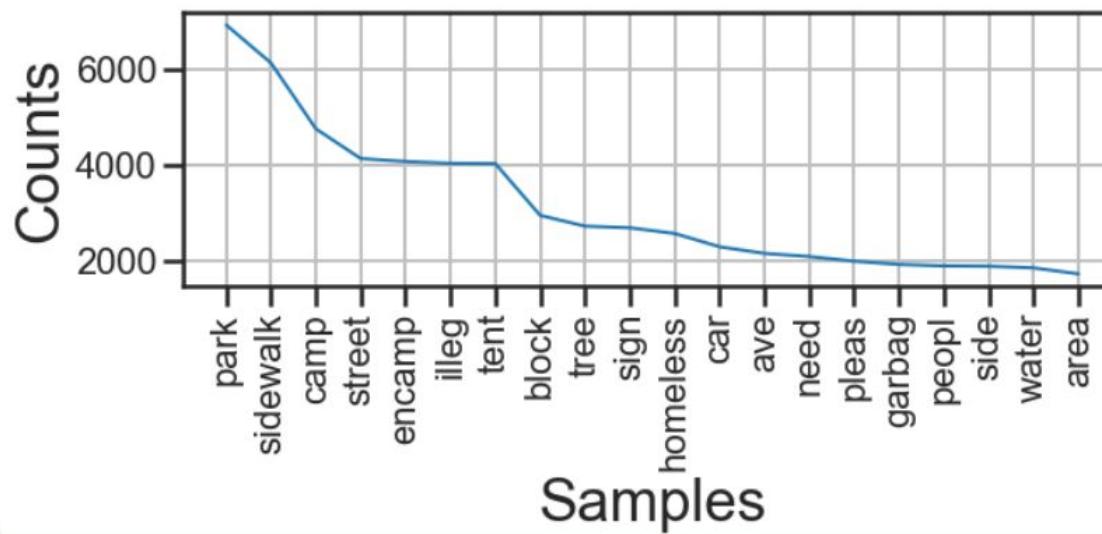


FIFI Word Frequency

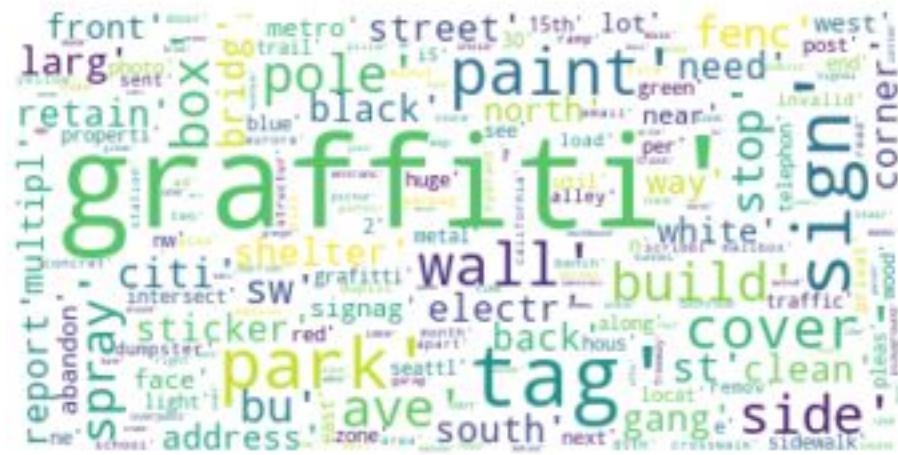
Graffiti



Other



FIFI Word Clouds



FIFI LDA Topic Modeling – all categories

```
Abnd_Vehicle
Words: 0.041*"park" + 0.022*"car" + 0.021*"week" + 0.021*"abandon" + 0.017*"rv" + 0.017*"vehicl" + 0.016*"white" + 0.014*"move"
+ 0.014*"van" + 0.013*"truck"

Graffiti
Words: 0.100*"graffiti" + 0.032*"tag" + 0.026*"sign" + 0.021*"park" + 0.017*"paint" + 0.016*"wall" + 0.014*"cover" + 0.013*"sid
e" + 0.013*"build" + 0.012*"ave"

Needles/Dumping
Words: 0.032*"garbag" + 0.030*"trash" + 0.027*"dump" + 0.018*"needl" + 0.013*"sidewalk" + 0.012*"park" + 0.012*"illeg" + 0.011
*"mattress" + 0.009*"street" + 0.009*"homeless"

Parking
Words: 0.069*"park" + 0.030*"sidewalk" + 0.022*"car" + 0.021*"block" + 0.017*"street" + 0.016*"pedestrian" + 0.013*"vehicl" +
0.011*"construct" + 0.010*"zone" + 0.010*"use"

Pothole
Words: 0.091*"pothol" + 0.021*"hole" + 0.018*"lane" + 0.018*"larg" + 0.013*"street" + 0.013*"road" + 0.013*"pot" + 0.012*"ave"
+ 0.012*"deep" + 0.009*"intersect"

Streetlight
Words: 0.095*"light" + 0.043*"pole" + 0.039*"street" + 0.026*"number" + 0.022*"flood" + 0.017*"ave" + 0.011*"streetlight" + 0.0
09*"lamp" + 0.009*"side" + 0.008*"corner"

Other
Words: 0.017*"park" + 0.016*"sidewalk" + 0.013*"camp" + 0.011*"encamp" + 0.011*"tent" + 0.011*"illeg" + 0.011*"street" + 0.008
*"block" + 0.007*"tree" + 0.007*"homeless"

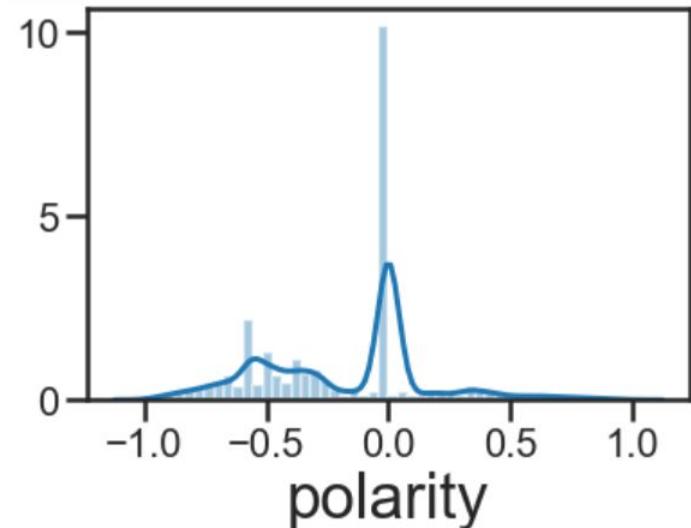
Clogged_Drain
Words: 0.075*"drain" + 0.063*"clog" + 0.038*"storm" + 0.021*"water" + 0.014*"grate" + 0.010*"sidewalk" + 0.009*"street" + 0.009
*"corner" + 0.008*"flood" + 0.007*"block"

Dead Animal
Words: 0.062*"dead" + 0.028*"raccoon" + 0.023*"side" + 0.021*"sidewalk" + 0.018*"park" + 0.018*"north" + 0.018*"street" + 0.018
*"ave" + 0.016*"rat" + 0.013*"east"

Vegetation
Words: 0.033*"sidewalk" + 0.031*"tree" + 0.031*"overgrown" + 0.025*"veget" + 0.023*"block" + 0.018*"grass" + 0.016*"need" + 0.0
15*"ave" + 0.013*"plant" + 0.013*"strip"

Sign/Signal
Words: 0.068*"sign" + 0.023*"park" + 0.015*"light" + 0.014*"signal" + 0.013*"stop" + 0.013*"zone" + 0.012*"need" + 0.011*"stree
t" + 0.011*"curb" + 0.010*"traffic"
```

FIFI LDA Topic Modeling – ‘other’ category



Other category requests with polarity > 0.5

```
Dictionary(5407 unique tokens: ['dog', 'love', 'market', 'park', 'see']...)
```

Topic: θ

Words: 0.015*"pleas" + 0.012*"park" + 0.010*"thank" + 0.008*"street" + 0.006*"sidewalk" + 0.006*"ave" + 0.006*"citi" + 0.005*"area" + 0.005*"car" + 0.005*"need"

Other category requests with polarity < -0.5

Dictionary(8644 unique tokens: ['alleyway', 'dump', 'illeg', 'area', 'creat']...)

Topic: 0

Words: 0.029*"illeg" + 0.017*"camp" + 0.017*"park" + 0.011*"encamp" + 0.011*"sidewalk" + 0.010*"block" + 0.009*"street" + 0.008*"sign" + 0.008*"tent" + 0.007*"car"

Thank you! Questions?

Abigail G. Schindler, Ph.D.
Assistant Professor
VA Puget Sound; University of Washington

grace311@gmail.com
aschind@uw.edu



Sentiment VADER

 [cjhutto / vaderSentiment](#)

Used by 846 Watch 118 Star 1,909 Fork 531

[Code](#) [Issues 14](#) [Pull requests 3](#) [Projects 0](#) [Wiki](#) [Security](#) [Insights](#)

VADER Sentiment Analysis. VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains.

108 commits 1 branch 1 release 8 contributors MIT

Branch: master ▾ New pull request Create new file Upload files Find File Clone or download ▾

Author	Commit Message	Date
	cjhutto Handling no as negation ...	Latest commit 5a6ccd9 12 days ago
	additional_resources Added support for emoji recognition (UTF-8 encoded)	last year
	vaderSentiment Handling no as negation	12 days ago
	.gitattributes update	3 years ago

LDA

