

# The use of Craigslist online community to measure an area's pet overpopulation problem and evaluate the effectiveness of spay/neuter programs

By

Lisa Wahl, Data Specialist, B. A. Computer Science

Robert A. Olsen, Financial Economist, Ph. D.

WAG (2009)

## Abstract

Evaluating the effectiveness of spay/neuter and other programs, as well as the overall pet overpopulation problem in an area and comparisons of different areas is an ongoing problem. Many factors can affect shelter intake and euthanasia. Shelter data is rarely available in a timely manner and sometimes not available at all. It rarely includes a breakdown by month or age, so is not useful for evaluating "Kitten Season," the clearest effect of a spay/neuter program. While the measurement of deaths per 1000 of the human population gives some method of evaluating the relative problem of communities and an indication of change over time. Its reliance on shelter data creates many problems. Monitoring Craigslist, however, gives a reading on the current situation, can be done with no reliance on shelters, and can be done as easily for distant or local communities. Initial investigation shows it to have great potential usefulness.

## Methodology

Craigslist is a centralized network of online communities, featuring free online classified advertisements – with sections devoted to jobs, housing, personals, for sale, services, community, gigs, résumés, and discussion forums.<sup>1</sup> Every week, at approximately the same time (10am Sunday morning) 3 counts are taken for each area being monitored of CL ads in the Community Pets section that come up in searches for the words "kitten" (K) "kittens" (S) and "dog" (D).

Craigslist Kitten Index (CLKI) =  $(K + 2*S)/D$

Craigslist areas have no clear geographic definition. People post on whatever CL they think most appropriate. In Oregon, for example, there's a CL for "OR Coast" but people on the Coast closest to Eugene or Portland might post in those cities instead. Some CL areas have tabs for sub-areas.

All of the areas monitored so far are entire CL areas, with two exceptions: 1) Tacoma is a sub-area of Seattle which is included because of the recently opened NSNRT spay/neuter clinic there. 2) Springfield does not have its own sub-area in the Eugene CL, but because of our interest in that area, we take its

---

<sup>1</sup> <http://en.wikipedia.org/wiki/Craigslist>

values by including the search term "Springfield" with the primary search term. Thus all the ads included in the Springfield numbers are also counted in Eugene.

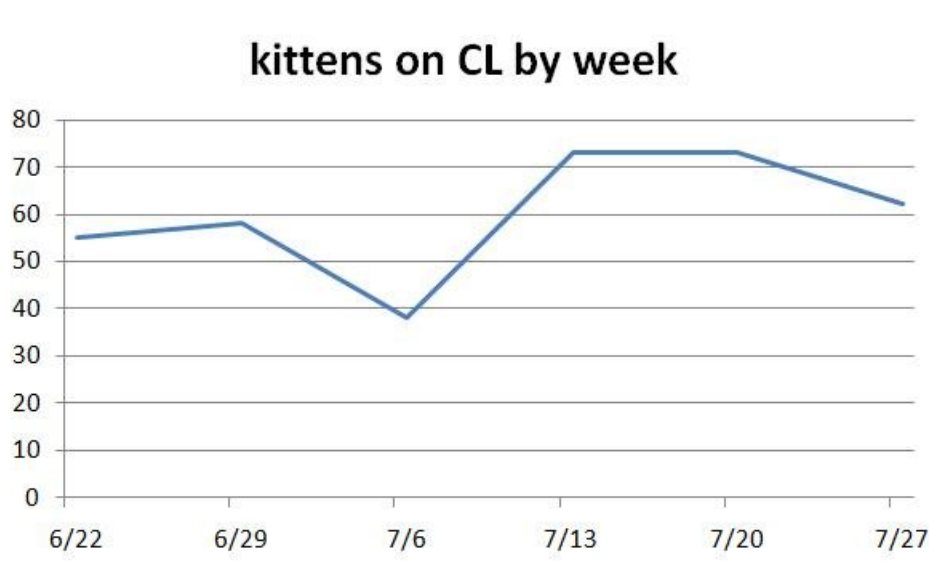
By using the Seamonkey browser and the ability to set up these searches as a set of bookmarked tabs, it requires only about 15 minutes a week to monitor 30 different areas.

Initially, we were monitoring only a few areas in Oregon, then we expanded it to all CL Oregon areas plus several selected pairs in other states where one area had an NSNRT s/n clinic and the other did not. Recently we added in many California areas because of the availability of CA shelter data, plus a second pair (the first being in California) including an area with a complete Maddie's Grant project. Therefore, we have data from Feb 13, 2009 – Oct 11, 2009 for a few areas, and from Mar 22, 2009 – Oct 11, 2009 for most of the others.

## Theory

The most likely first effect to be seen from a spay/neuter program is a decrease in "Kitten Season" that period of summer and fall when kittens born in the spring are looking for homes.

In 2008 we attempted to count the number of kittens looking for homes on the Eugene CL, eliminating duplicate ads. This proved to be very time consuming, but informative, with 39-72 new kittens advertised each week. At the same time, the local county shelter was taking in less than 100 cats per month.<sup>2</sup>



<sup>2</sup> <http://www.lanecounty.org/Animals/> Quarterly Statistical Report

CL areas have no definable geographic area, and their activity level is affected by many things outside of the size of the area served. An effort to estimate the population along county lines and measure posts/1000 shows a wide range in Oregon.

<u>city</u>	<u>Average dog posts</u>	<u>dog posts/1000 pop</u>	<u>population<sup>3</sup></u>	<u>area</u>
Medford	245.29	1.24	197,071	Jackson
Eugene*	360.57	1.07	337,870	lane county
Bend Prineville (SNIP)	175.71	0.88	199,321	central oregon
Corvallis	110.29	0.58	190,086	Benton/Linn
Salem	160.86	0.52	311,304	Marion
Portland	503.57	0.32	1,562,150	Portland (Multnomah, Washington and Clackamas Counties)
Roseburg	26.43	0.25	105,117	Douglas

Therefore, we consider the count for “dog” searches to be a good measure of that CL’s pet section activity level. Only about half of the ads that come up on such a search are dogs available for adoption (others being lost and found, grooming and pet sitting services offered) while 75-94% of ads that come up in the “kittens” search are multiple kittens needing homes. Thus, even though our search on “kitten” and “kittens” keywords doesn’t represent kittens for adoption only, the percentage is high enough, and higher than that for the “dogs” search that they provide a good measurement of kitten, and therefore cat overpopulation, not masked by variations in the dog overpopulation situation.

Oct 11 sample of proportion of ads of animals needing homes

	<u>kitten</u>	<u>kittens</u>	<u>dog</u>	<u>cat</u>	<u>ave KI</u>
Bend/Prineville	59%	75%	44%	41%	47%
Bakersfield		81%	49%		67%
Fresno	80%	80%	57%		87%
Eugene	68%	76%	52%	55%	92%
Stockton		91%	49%		93%
Portland	82%	86%	39%		97%
Salem	75%	92%	27%		117%
Modesto	73%	94%	59%		127%
Springfield	72%	82%	44%	55%	137%

Craigslist’s search feature works by whole word, so that the “kitten” search does not include ads in the “kittens” search unless both the singular and plural are used in the ad.

<sup>3</sup> from US Census estimates of population by county each year <http://factfinder.census.gov/>

The CLKI is positively correlated with % of the "kittens" ad count (K) at the 8 percent level. In other words, the higher the CLKI, the fewer of the ads included in the "count" are something other than kittens in need of homes.

The dog and kitten ad percentages are unrelated statistically. The kitten percentages are statistically larger than those of the dog. Dog percentages are statistically unrelated to the CLKI.

There is no periodicity in the dog data. The month to month values are not seasonally related. They follow a normal distribution over the period.

The cat data has significant periodicity. Clearly the cat numbers are statistically significantly higher during June to October. There is only one chance in a hundred that June to October cat numbers are NOT higher than the rest of the time period. That is they are significant at the 1% level.

All of this supports the usefulness of putting the D count in the denominator of the CLKI formula to give a number allowing comparison of different areas and that the CLKI gives a measurement of "Kitten Season."

## Validation

Several of the areas monitored from the beginning had very low CL activity and showed a great deal of random fluctuation, these have been dropped from analysis. Areas with very high levels of CL activity (areas of high population, such as Los Angeles and Atlanta) were not included because CL searches return a maximum of 1000 results.

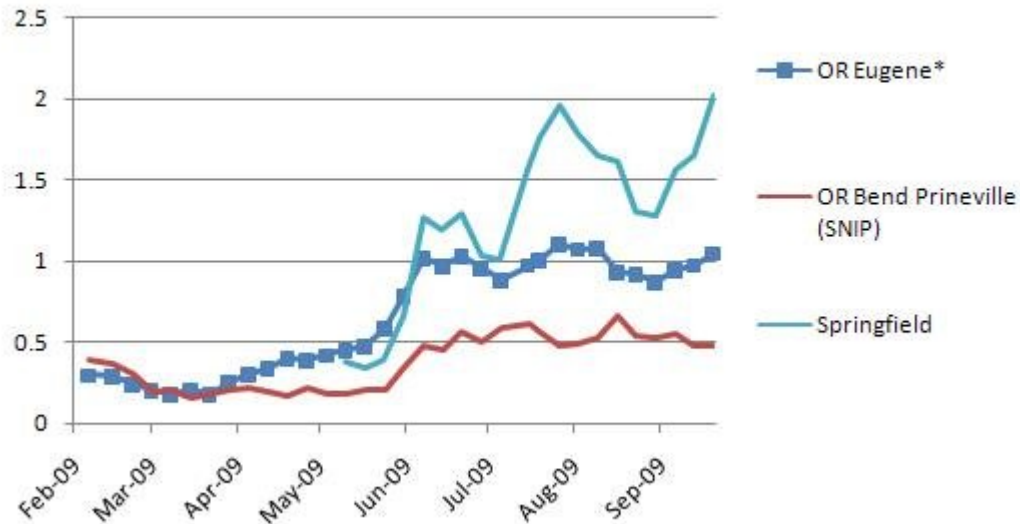
Most of the rest of the Oregon areas show remarkably similar CLKI values, illustrating "Kitten Season" effectively. (On many of these charts, cities are marked with an asterisk if they have NSNRT s/n clinics)

### CLKI similar in most of OR



However, several areas stand apart – the Prineville/Bend area where two low-cost s/n clinics have been operating starting in 2003 and 2005 respectively and where shelter intake has notably declined since <sup>4</sup> and the Springfield area, where the experience of local animal rescuers suggests the pet overpopulation problem is especially bad.

### Craigslist Kitten Index



Looking at data from California shelters<sup>5</sup> just one measurement of the CLKI (Sept 27, 2009) correlates significantly with 2008 deaths/1000, with the exception of Fresno.

<u>CL area</u>	<u>County</u>	<u>county pop</u>	<u>county deaths/1000</u>	<u>KI</u>
San Luis				
Obispo	San Luis Obispo	265,297	2.39	64%
Ventura	Ventura	797,740	4.95	45%
Santa Barbara	Santa Barbara	405,396	5.33	61%
Orange Co	Orange	3,010,759	5.71	82%
Sacramento*	Sacramento	1,394,154	13.39	119%
Chico	Butte	220,337	13.82	117%
Stockton	San Joaquin	672,388	20.17	83%
Inland Empire	Riverside/San Bernardino	4,115,871	23.67	67%
Bakersfield	Kern	800,458	29.68	69%
Redding	Shasta	180,214	30.38	140%
Modesto	Stanislaus	510,694	35.52	116%
Fresno*	Fresno	909,153	43.59	75%

<sup>4</sup> Unpublished papers by Lisa Wahl, coming soon to a theater near you

<sup>5</sup> California Department of Public Health, Veterinary Public Health

<http://www.cdph.ca.gov/HealthInfo/discond/Pages/LocalRabiesControlActivities.aspx>

For this California data, the deaths per 1000 and CLKI are positively correlated. The correlation coefficient is .51 and the result significant at the 8% level. That is, there is only about an 8% chance that a positive correlation does not exist. Each 1% increase in the CLKI is associated with a .2 increase in the death per 1000 number. Thus on average a 50 CLKI would be associated with death number of about 15. Likewise a CLKI of 0.0 would be associated with a death number of 5. This is particularly interesting since Merritt Clifton, editor of Animal People says the "natural" rate of shelter deaths in any given community for humane reasons is about five dogs and cats per 1,000 human residents per year.<sup>6</sup>

Deaths/1000 are not as closely correlated for Oregon shelters, even though these CLKI values are an average and, therefore, more likely to be an accurate picture of the area and they graphs look so similar. The correlation is .60 but it is only significant at the 20% level. Thus there is a 20% chance that there is no substantial correlation between the two sets of data based on this small sample. The shelter data for these figures,<sup>7</sup> however, is older than the California shelter data, being from 2005 or 2006. However, given our personal knowledge of the situations in many of these areas, we suspect the real reason for the differences relate to shelter policy, and the CLKI is a more accurate measurement than deaths/1000.

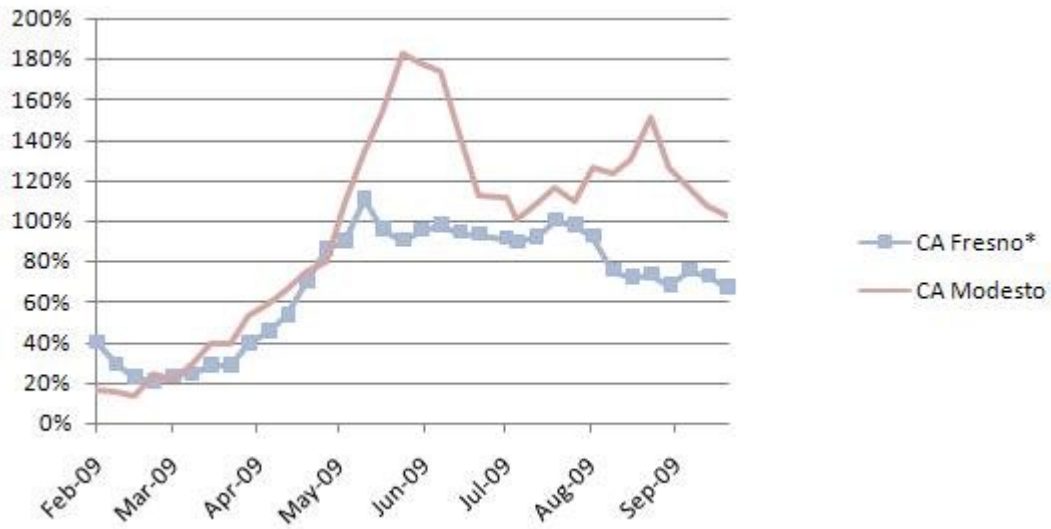
	deaths/1000	ave KI
Bend Prineville (SNIP)	4.67	46.92%
Portland	4.82	97.31%
Corvallis	7.51	82.73%
Eugene*	8.93	92.35%
Medford	18.79	87.53%
Salem	19.71	117.46%

The reason the current 2009 CLKI data for Fresno is at odds for the 2008 shelter data is probably the result of the HOPE NSNRT clinic there. Their CLKI data shows a dramatic difference from its "paired" city, Modesto.

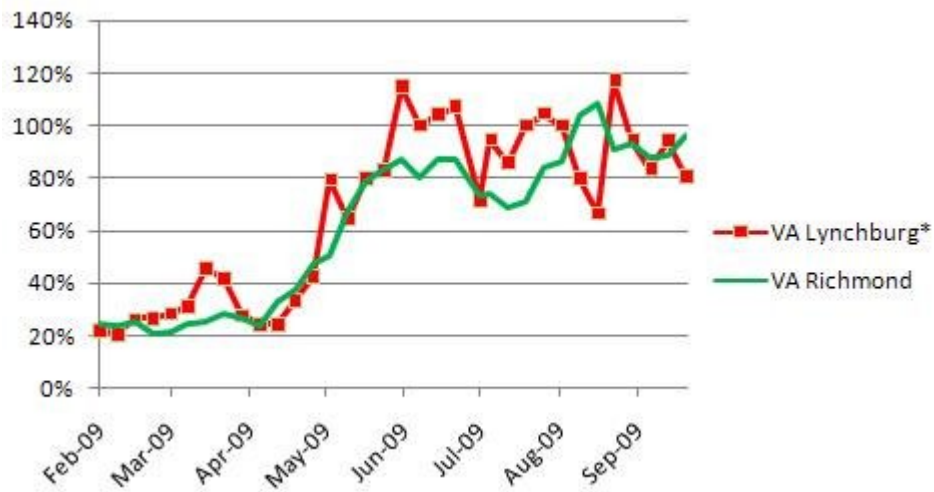
---

<sup>6</sup><http://www.sarasavesanimals.org/Resources/No%20kill.pdf> 'No kill' doesn't mean no killings The Register - Guard - Eugene, OR by Diane Dietz Apr 30, 2006

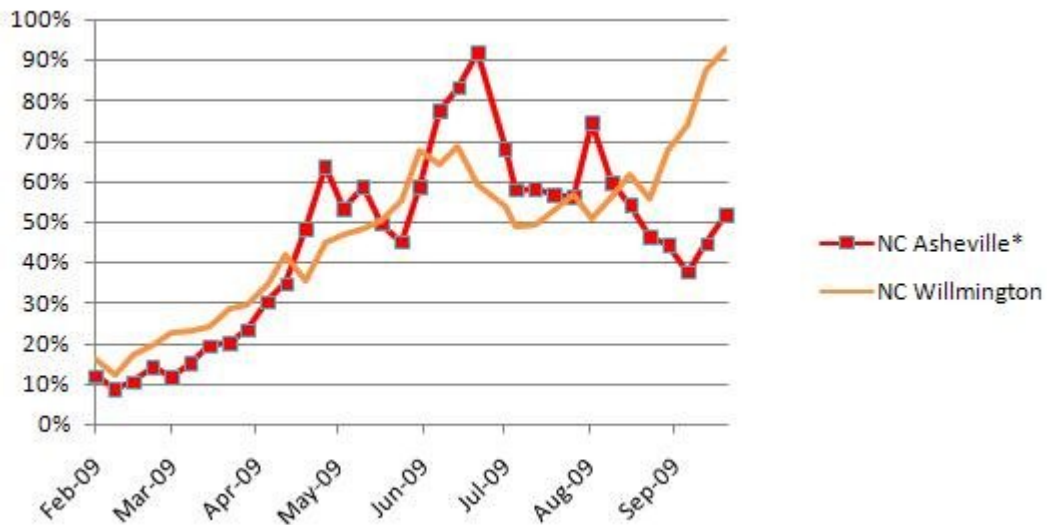
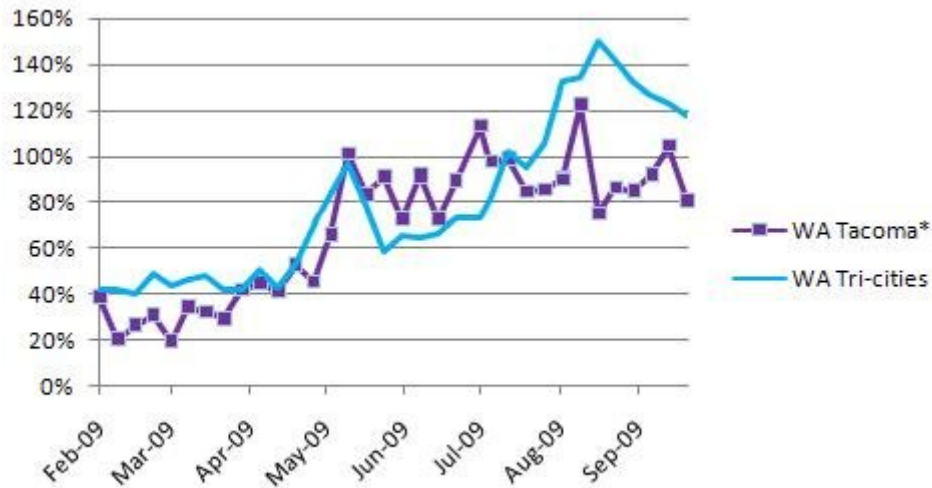
<sup>7</sup> Oregon Humane Society data



The other “paired” cities being monitored do not show as obvious a difference as do Fresno and Modesto. In one case (Virginia) we were unable to find a pair with enough CL activity to produce a smooth chart.



Of the other two pairs, we have less data on the area and the operating clinics and need to do further research. However, a look at the CLKI charts for them have some features of interest.



Note that Asheville is the home of the first NSNRT clinic, Humane Alliance. Their web page states "The euthanasia rate at our local shelter is down by 70% since our clinic opened in 1994." <sup>8</sup>

Here are the pairs with their average CLKI values.

59%	NC	Asheville*
60%	NC	Willmington
87%	CA	Fresno*
127%	CA	Modesto
89%	VA	Lynchburg*
82%	VA	Richmond
88%	WA	Tacoma*
99%	WA	Tri-cities

<sup>8</sup> [http://humanealliance.org/index.php?option=com\\_content&view=article&id=15&Itemid=43](http://humanealliance.org/index.php?option=com_content&view=article&id=15&Itemid=43)

Although not as convincing of the difference a s/n clinic makes as the Fresno chart suggests, there are a couple of interesting points in this data. One is how the pairs in NC and WA diverge toward the end of Kitten Season, just as the curve for Salem, while similar to other OR areas for most of the year, skyrockets during the late summer. Another is how dissimilar these pairs are, compared to the 5 urban Oregon shelters graphed earlier which are closer geographically. The pairs were intentionally chosen to be sufficiently geographically distant so that the activities of the spay/neuter clinic were unlikely to effect both, as well as being distant from any other NSNRT clinic. However, no other research was done on other spay/neuter programs in these areas. In fact, the choice to monitor the Tri-Cities (Kennewick on this map) area of Washington may have been a poor one because of projects now in NE Oregon and Walla Walla, WA.



## Future plans

Although the results from collecting data for less than a year are very promising, the use of the CLKI to monitor changes in an area has yet to be established. It will be especially interesting to compare next year's data to this year's in Fresno, in Eugene and Springfield where the work of the WAG NSNRT clinic is likely to make a change, and in Salem where a large NSNRT clinic is scheduled to open in Jan 2010.

The impressive correlation with Deaths per 1000 with a single CLKI value in California brings up the question of how often the CLKI needs to be measured to be useful. Would once a month provide as accurate a measurement as weekly? Or would that depend on the activity level of the CL of a given area?

## Conclusion

Monitoring Craigslist shows great promise as a way of evaluating the relative needs of different geographic areas as well as to monitor changes in a specific area.