Assessing Predictors of Citizen Reports of Dead Animals in Cincinnati

Nate Wessel

December 1, 2016

Abstract

This exploratory analysis looks at a geolocated dataset of calls to the City of Cincinnati regarding dead animals over a period of about four years. Regression analysis is employed to discover variables associated with higher or lower local reporting volumes. The results indicate that the spatial patterns observable in the data may have at least as much to do with human dispositions toward calling the city or disposing of a pet as they do with the likelihood of an animal dying on a public right of way.

1 Introduction

This paper seeks to explore an interesting dataset from the City of Cincinnati describing citizen-reported animal deaths between the years 2012 and 2016. The data consists of reports of calls to the City requesting that a body be removed from some location which the caller identifies. The reports are made by citizens, but logged and geolocated by City employees as part of the process of dispatching a range of city services like replacing road signs and filling in potholes.

My interest in this topic comes from my husband's practice as a taxidermist in Cincinnati and my own experience working with him and collecting bodies from the streets of this city. One thing I learned in the process of becoming comfortable working with dead things myself is just how uncomfortable many people are with bodies and with the idea of their own mortality which the presence of dead bodies seems to raise. The reports in this dataset often include vivid comments by callers describing an often invisible side of the city - a side that some are quick to enlist the City in hiding away.

Many of the reports clearly describe what is often termed "roadkill": the violent end of an animal's life as it is crushed beneath the wheels of, or broken by the impact with a vehicle moving at unnatural speed and presumably out of the control of its human(e?) occupants. But there seems to be another category, beside the flattened bodies of the former; these records are typified by domestic animals like cats and dogs, and an apparent concern for their dignity. Bodies here have evidently been handled, placed in boxes or bags and set neatly at the curb. My own suburban, childhood experience with pets does not allow me to understand why anyone would call city services rather than performing a backyard burial or passing the body to a veterinarian for cremation or an otherwise more dignified disposal. Yet I can't deny that these seem to be records of domestic animals.

The spatial patterns evident in the dataset (see Figure 1) raise several interesting questions for me:

1. What explanatory factors are associated with the violent roadside death of animals?

2. What factors explain the willingness of people to report these events?

3. What factors might explain the surprisingly civic disposal of what appear to have been pets?

I will attempt to answer these questions through spatially aware regression models that take the location and frequency of reports as the variable to be explained.

2 Literature Review

2.1 Roadkill

The effects of roadways on wildlife have been well documented; Coffin [2] and Forman [5] offer good overviews of the literature, and there is now an entire textbook available on the topic of "road ecology" [4]. Beside directly destroying habitat by paving over it, roads form major barriers to movement for many species and fragment breeding populations. They encourage the

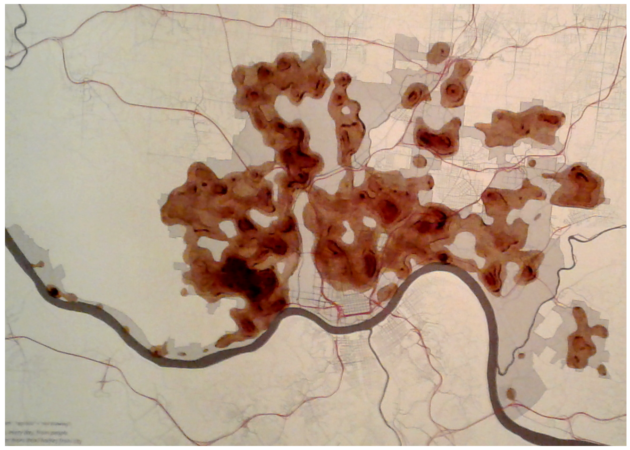


Figure 1: Map of kernel density surface of roadkill data, with successive contour bands overprinted in blood. The city boundary is visible in grey. No records outside the city are included.

spread of invasive generalist species, which may thrive in cleared roadside scrub, and they diminish the numbers of species that require larger undisturbed spaces for survival. Beside these indirect effects, cars and trucks likely kill more animals directly than does human hunting, with an estimated one million vertebrate roadkills each day in the United States alone [5]. Some species, like the Florida Panther or the Key Deer have been observed to lose ten percent or more of their entire population each year to collisions with cars [5].

The phenomenon of ”roadkill”, or roadway mortality, has been studied in a in a variety of environments, for many types of animals [3, 8, 7, 6]. Several variables have been found to be reliable predictors of roadkill events, including road speed limits [3, 5], road width [5], warmer weather [3], habitat diversity and proximity to wetlands [3, 5]. Many studies are based on surveys conducted by periodically surveying roads. Several studies have pointed out that such technique may be deeply flawed. For example, Antworth et al. [1] experimentally placed dead baby chickens and snakes on several roadways in Florida and observed them over 36 hours. In that time between 60% and 97% of the specimens were collected by scavengers; the position of the body in the roadway also had an effect on scavenging rates. Another study, by Santos et al. [14], conducted daily surveys of road segments in Portugal, identifying and tracking roadkill to see how long bodies persisted in the roadway. They found that on average, the bodies disappeared within a day of being sighted, and that this depended to some degree on the taxon, body size, position on the roadway, and many other factors.

2.2 Civic Reporting

The second big question this paper needs to address is the matter of who is likely to call a city's services line to report a dead animal. The literature on this topic does not seem quite as well developed as that for roadkill, but we do have some guides. A study of New York City 311 reports [10] developed and tested a theoretical framework for propensity to make a report. First, there must be a perceived problem. Second, there must be an awareness of the program and a propensity to call. The author suggests several measures that he expects would be positively related to propensity such as homeownership, presence of children in the household, and higher incomes. Other factors were then taken as measures of neighbourhood condition. In regards to requests for government goods, homeownership and families with children did tend to make slightly more requests, though the effect of increasing income was somewhat less clear [10]. A study using data from Boston's 311 system found that homeowners were three times as likely to report an issue to the city than renters [12]. That study also found that callers had a strong tendency to make reports in a very small area around their own home - no more than a few blocks away [12]. One final study, by O'brien et al. combines the same Boston 311 data with a survey of the actual users. The authors find that the biggest motivation for making reports comes from a desire to benefit the local community, and that considerations of property values were less important. Callers expressed a desire to build community, and also noted that calling 311 did little to help connect them with others. There was also little perception of the service as a mechanism for resisting the encroachments of outsiders or enforcing norms [13].

2.3 Pet Disposal

The third potential cause of reports in this dataset takes us into waters far less explored, at least by quantitative geographers. This is the issue of potential pet disposal. My excursion into the academic literature here turned up a few interesting but irrelevant pieces on the concepts of pets and death, and their particular salience in the Victorian era. However, the more interesting results were online answers to questions such as “My cat has died. What do I do?”. Many websites recommend veterinarians or professional pet disposal services, but many also note that home burial is illegal in some areas. Notably, some city 311 sites (eg in Toronto [9] or New York [11]) offer options for disposal, either for free or for a fee. From the New York City website:

> You can also put a dead animal out for pickup by the Department of Sanitation on your garbage day. The remains must be placed in a heavy-duty black plastic bag or double plastic bag and a note should be taped to the bag stating its contents (for example, “dead dog” or “dead cat”). Animals that may have been rabid should not be put in the trash.[11]

So, apparently there is plenty of precedent for a civic pet disposal, however I am left with little theoretical guidance on predicting the spatial occurrence of such events.

3 Methods

The dataset on reported animal deaths will be explored using regression analysis. The literature review has suggested several potentially related variables which will be used as independent predictors. I will start with simple OLS regression, but since this is a spatial dataset I expect that autocorrelation will exist among the model residuals. A spatial model will be employed to account for this. I will not bother with techniques for estimating spatially varying coefficients because I don’t expect that the City of Cincinnati is large enough to contain meaningful differences in effect sizes. I would expect such variation to exist on a more regional scale and Cincinnati is a very small city (boundary) relative to the metropolitan region of which it is the center.

3.1 Data

3.1.1 Dependent Variable

The dependent variable for this study comes from a dataset published by the City of Cincinnati. It consists of reports of any request for city services between January 2012 and June 2016. From this dataset, I am interested in the records reporting a dead animal: a total of 6,422 reports. Records are timestamped and geolocated, both by address and latitude/longitude. A simple kernel density surface from the point data can be seen in figure 1. There is clear clustering on the west side of the city, and notably, there are many large areas of the city where there are few or no records.

Since the dataset consists of point data, it needed to be aggregated into areas and converted into a rate of reporting so that regression could be employed. I chose not to take the simplest option, binning the points into preexisting census boundaries, because such boundaries are typically broken along major roads; I would expect roadkill to happen along major roads, and the side of the road on which the report is made should not have any effect on the measured rate of animal death in a zone. To avoid this effect, I used a centroidal Voronoi tessellation of the city boundary, dividing the area into 547 zones of roughly equal size (Figure 2). Such a tessellation disregards internal boundary features and only considers the boundary defining the city itself. Reports were aggregated to these new boundaries and divided by the area in square kilometers to yield a *report_rate* variable (Figure 2). This variable had a very strong positive skew with 62 zero observations and a maximum of 179 reports per square kilometer. To deal with this and other large outliers, the square root of this variable ($report\_rate^{1/2}$) will be used as the dependent variable throughout the analysis that follows.

3.1.2 Roadkill Variables

Several variables were created as predictors of roadkill events. Tree cover is taken as a proxy for natural habitat and derived from a USGS raster dataset on tree canopy cover. The presence and nature of roadways is taken from OpenStreetMap data, and divided into three categories: "streets" for residential and minor, slower streets, "roads" for primary, secondary and tertiary classified ways, and "motorway" for highways and trunk roads. These roadway variables represent the length in kilometers of all segments intersecting with the area, divided by the area in square kilometers.

Variable	Description
tree_cover	% covered by tree canopy
$tree\_cover^2$	square of the above
street_length	km/km^2 of minor streets
road_length	km/km^2 of major roads
motorway_length	km/km^2 of motorways

In a scatterplot, the *tree_cover* variable appeared to have a non-linear

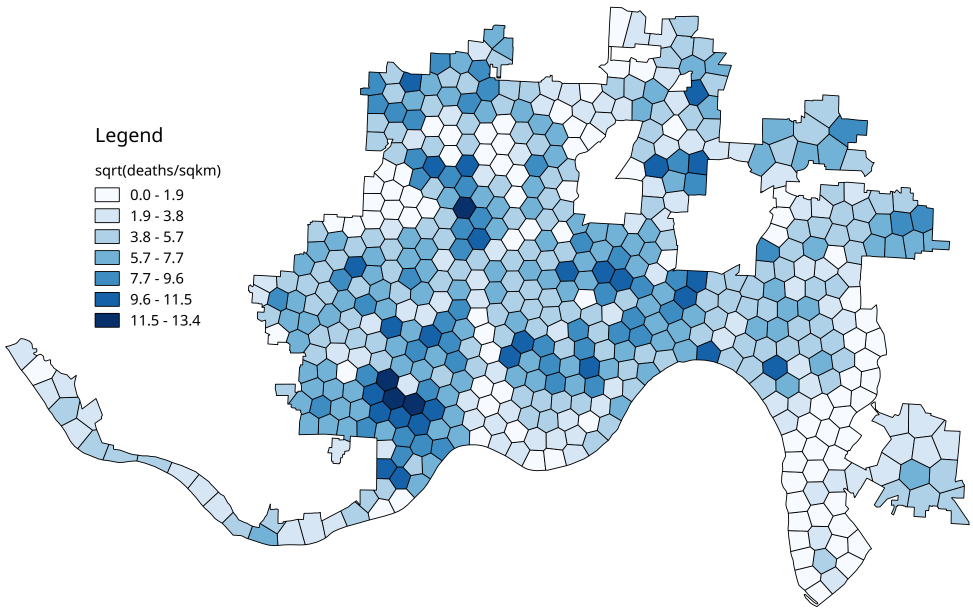


Figure 2: Map of $report\_rate^{1/2}$ variable showing centroidal voronoi tessellation bins. Average bin size is 0.38 square kilometers.

relationship to the reporting rate, so a squared term was added. A water or wetland variable would have been included, but natural water features are not a significant presence in Cincinnati and no data was available on other potential sources of water.

3.1.3 Reporting Variables

Other variables were created as predictors of the inclination of people in the area to call city services for any reason.

Variable	Description
per_capita_income	mean per-capita income, thousands
mean_improvement	average appraised SF home value, thousands. Null set to mean
stdev_improvement	st.dev of appraised SF home value, thousands. Null set to mean
pct_rented	% of units rented vs owner occupied
pct_rented^2	square of the above
pop_density	residential population density (square root transformed)
job_density	workplace density (cube root transformed)
institutional	% institutionally managed land

The first two variables, *per_capita_income* and *mean_improvement* are essentially measures of income or wealth and are expected to be positively related to report frequency. *stdev_improvement* measures differences in home values within a zone and is expected to be positively related. Both improved value variables had some missing values where no single family homes exist and these were recoded to the mean of the remaining values.

percent_rented gives the percentage of housing units that are renter occupied and is expected to be negatively related to report frequency. In a simple scatterplot, the *pct_rented* variable appeared to have a non-linear relationship to the reporting rate, so a squared term was added. *job_density* and *pop_density* are simple density measures and are both expected to relate positively with report frequency, though residential density should probably be more strongly related since people are generally more protective of their homes.

The final variable here, *institutional*, gives the percentage of the area covered by institutionally owned land that might reasonably be expected to have its own service personnel to handle dead animals on or near the property: railyards, cemeteries, airports, universities, colleges, and primary and secondary schools. The impetus for this variable comes from my informal observation of the mapped pattern; reports seem very infrequent over certain large railyards and airports.

per_capita_income, *pct_rented*, *pop_density*, and *job_density* were derived from ACS Census data at the finest available spatial resolution and apportioned into the new aggregation units. *mean_improvement* was taken from a 2010 dataset of assessed values from the Hamilton County auditor's office. *institutional* was derived from OpenStreetMap data.

4 Results

The first model, a simple OLS regression, uses all of the above variables to predict the rate of reporting of dead animals per square kilometer. Results are shown in the table below.

	β	p
α	-1.6166	.0386
tree_cover	0.0785	.0...
tree_cover^2	-0.0009	.0...
street_length	0.3268	.0...
road_length	0.3346	.0...
motorway_length	0.0484	.1920
per_capita_income	-0.0029	.6861
mean_improvement	-0.0072	.0...
stdev_improvement	0.0034	.0136
pct_rented	0.1076	.0...
pct_rented^2	-0.0008	.0...
pop_density$^{1/2}$	0.0525	.0...
job_density$^{1/3}$	-0.1997	.0...
institutional	0.0151	.0052
adjusted r^2	0.5835	.0...
residual Moran's I	0.1940	.0...

The model does a surprisingly good job of explaining variation in report frequency, with an adjusted r^2 of 0.58. Variables related to both reporting propensity and roadkill factors are are highly significant for the most part.

The road factor variables show essentially the expected signs, with the exception of *motorway_length*. Other studies have clearly linked traffic volume and speed to roadkill, however I might speculate that through the City of Cincinnati at least, highways are quite thoroughly separated from the rest of the world by walls, fences, or elevation. Or perhaps they are simply administered by a different agency. *street_length* and *road_length* are both positive and highly significant, and perhaps surprisingly not actually inter-correlated ($r = 0.023$). Both variables have essentially the same estimated coefficient here, though I would have expected higher speed roads to demonstrate a stronger effect than slower residential streets. The other roadkill variable in the model, *tree_cover*, is a bit harder to interpret from the coefficients, but also hows the expected relationship. Reports are more frequent at intermediate values of *tree_cover* and less frequent at either extreme (totally wooded versus no natural habitat at all).

Considering now the variables related to human propensities to phone the city, we see a slightly less intuitive interpretation. Only one of the variables describing wealth or income was really very significant, and the sign is not as expected. *per_capita_income* may show little significance because it comes from tract level data, while *mean_improvement* comes from parcel level data, which probably better describes the local level of investment in a neighborhood. It is interesting though that *mean_improvement* has a negative relationship with reporting rate. As discussed by Minkoff,[10], we might expect that income should be positively related to ability (time, leisure) or propensity (investment) to call the city, but it may also be negatively related to the actual neighborhood conditions that precipitate reporting. One other possible theory is that the likelihood of disposing of a pet via 311 goes up as incomes go down, as lower-income people may be less likely to have a regular veterinarian in their employ to whom they might turn when Death swings his dreaded scythe. *stdev_improvement* shows a relatively weak though significant relation to reporting frequency and should probably just be ignored for now.

The variable describing rental tenancy has a somewhat surprising relationship to reporting frequency, similar to that of *tree_cover*, where intermediate values are associated with more reports of dead animals than either extreme. My expectation was that all else being equal, homeowners would be more protective of their neighborhood and more likely to call the city. As noted by Minkoff [10] however, there is a potential for circularity here, in that neighborhoods with a higher propensity to report issues may actually end up making fewer reports once they successfully fix most of their issues by dint of their persistence.

The variables on density at home and work indicate that reports are much more likely to be made from residential locations than work locations. I had expected *pop_density* to be a strong positive predictor, however I was a little surprised by the highly significant negative effect of *job_density* on reporting frequency. The *institutional* variable was significant, but did not have the expected sign for reasons that I don't yet understand.

4.1 Accounting for spatial effects

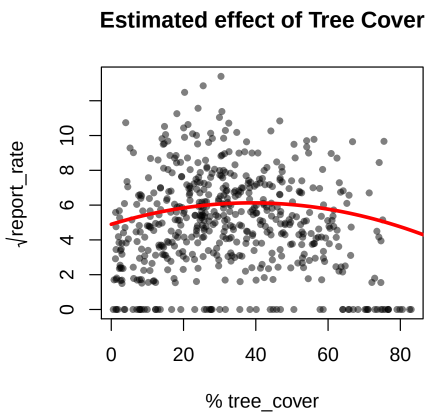
Again, the model fit with the simple OLS is quite good ($r^2 = 0.58$), however significant autocorrelation among the residuals ($I = 0.194$) indicates a problem that needs to be addressed by a more advanced spatial model. To address this issue, I make use of both a spatial error model an a spatial lag model. I have reason to expect that the autocorrelation among the residuals could be due to either a missing variable, such as percent of households with pets, or to a spatial contagion effect such as the phenomenon of scavengers getting hit by cars themselves or of local social spillovers such as neighbors talking about their interactions with city services. Quite possibly both effects are present, and an initial try with both modelling techniques yielded highly significant results in ether case. To attempt to locate the most appropriate model, I made use of the robust model tests available through R's *spdep* package.

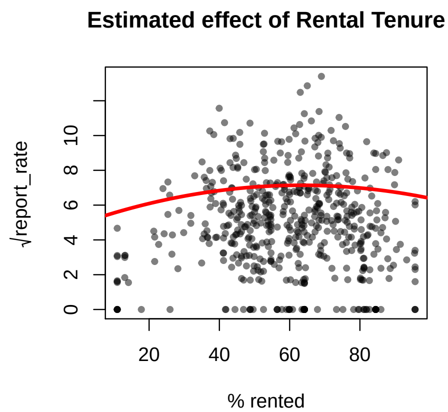
	p
Robust Error Model	0.052
Robust Lag Model	0.002

The results indicate that a spatial lag model is best suited to this data. To simplify things, I removed the variables which were not significant in the original OLS model and which do not become significant in the spatial model: *motorway_length* and *per_capita_income*. The results of the spatial lag model follow and plots visualizing selected effects from the model can be found in figure 3.

	estimate	p
α	-1.3267	.0512
ρ	0.3426	.0...
tree_cover	0.0632	.0...
$tree\_cover^2$	-0.0008	.0...
street_length	0.2784	.0...
road_length	0.3054	.0...
mean_improvement	-0.0060	.0...
stdev_improvement	0.0028	.0278
pct_rented	0.0708	.0002
pct_rented^2	-0.0005	.0010
$pop\_density^{1/2}$	0.0339	.0...
$job\_density^{1/3}$	-0.1807	.0...
institutional	0.0074	.1338
residual Moran's I	0.0279	.1388

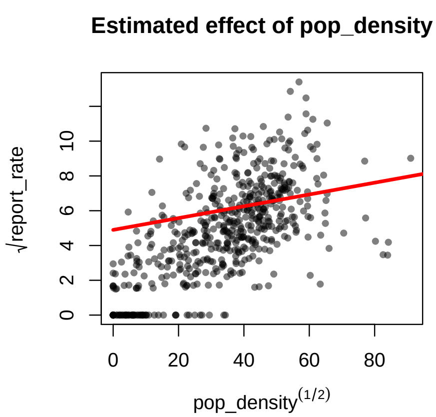
The spatial lag ρ is highly significant and it is worth noting that estimated coefficients have not changed substantially between the lag model and the OLS model. Autocorrelation among the residuals from the lag model have

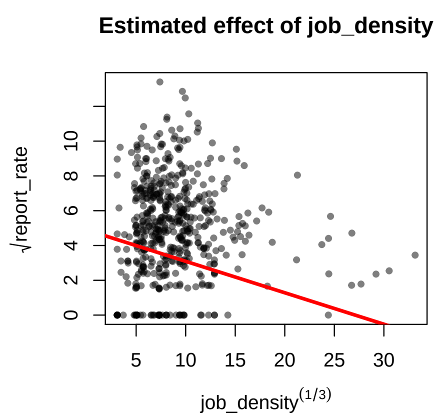




(a) $y = \mu + 0.0632 * tree\_cover + -0.0008 * tree\_cover^2$.

(b) $y = \mu + 0.0708 * pct\_rented + -0.0005 * pct\_rented^2$.





(c) $y = \mu + 0.0339 * pop\_density^{1/2}$.

(d) $y = \mu + -0.1807 * job\_density^{1/3}$.

Figure 3: Plots visually showing estimated variable effects estimated by the lag model. Noet the curvilinear effects produced by the squared terms for *tree_cover* and *pct_rented*.

been reduced to insignificance.

5 Discussion

The results of the two regression models suggest that all three potential explanatory factors are likely at play in producing the observed frequency of reports about dead animals in Cincinnati. Variables introduced by expectations about roadkill events are highly significant and moving in the expected directions. The *tree_cover* variable indicates that reports are most likely where natural and human habitats are intermixed (Figure 3a). Reports are also more likely where there are streets, and especially in the final model, where there are higher-speed and higher-traffic roads.

On the other hand, the variables meant to explain general reporting propensity are also very significant. Reporting rates are highest where there are the most people (Figure 3c), but lower where more people are working rather than living (Figure 3d). This matches with an expectation that people are more likely to report things within a few blocks of where they reside. However the strong negative coefficient for job density demands further examination. It's possible that places with more jobs may be places where people on staff would just pick up a dead animal as part of their duties.

However, to speak to the third potential explanation for report frequency, the places where people work are also unlikely places for pets to die. Conversely, they should be more likely to die where people live. The unexpected negative coefficient for *mean_improvement* may also point in the direction of pet deaths, as it's conceivable that civic disposal (as opposed to burial or cremation) is more likely in poorer neighborhoods. Other studies had suggested a positive relationship between incomes and reporting frequency, so it's possible that we are observing a stronger effect going in the other direction and cancelling that out.

The effect of rental tenancy is an interesting standout. Rather than reporting frequency being higher in areas with more homeowners, it appears to be higher where there is a mix of renters and owners (Figure 3b), though the effect is not as strong as others in the model. Possible explanations for the effect of rental tenancy mix are elusive.

The significance of the lag model over the error model indicates that some of the residual errors in a non-spatial model may be due to contagion-type effects where nearby reporting frequency may actually directly influence the likelihood of reporting nearby.

In any event, I seem to have been rather deeply mistaken when I earlier interpretted the dataset to represent primarily the probability of an animal dying after being struck by a car (Figure 1). I'm quite convinced now that the effect observed is at least as much the result of people's home locations, attitudes toward reporting problems to the city, and the distribution of the pet population.

References

[1] Rebecca L Antworth, David A Pike, and Ernest E Stevens. "Hit and run: effects of scavenging on estimates of roadkilled vertebrates". In: *Southeastern Naturalist* 4.4 (2005), pp. 647–656.

[2] Alisa Coffin. "Form Roadkill to Road Ecology: A Review of the Ecological Effect of Roads". In: *Journal of Transport Geography* (2007), pp. 396–406.

[3] Robert G. Farmer and Ron J. Brooks. "Integrated Risk Factors for Vertebrate Roadkill in Southern Ontario". In: *Journal of Wildlife Management* (2012).

[4] Richard TT Forman. *Road ecology: science and solutions*. Island Press, 2003.

[5] Richard TT Forman and Lauren E Alexander. "Roads and their major ecological effects". In: *Annual review of ecology and systematics* (1998), pp. 207–C2.

[6] Luís Gomes et al. "Identification methods and deterministic factors of owl roadkill hotspot locations in Mediterranean landscapes". In: *Ecological Research* 24.2 (2009), pp. 355–370.

[7] Kari E Gunson, Giorgos Mountrakis, and Lindi J Quackenbush. "Spatial wildlife-vehicle collision models: A review of current work and its application to transportation mitigation projects". In: *Journal of environmental management* 92.4 (2011), pp. 1074–1082.

[8] Alistair J Hobday and Melinda L Minstrell. "Distribution and abundance of roadkill on Tasmanian highways: human management options". In: *Wildlife Research* 35.7 (2008), pp. 712–726.

[9] *Living or deceased pets - animal surrender*. City of Toronto. 2016. URL: `http://www.toronto.ca/311/knowledgebase/19/101000048019.html`.

[10] Scott L Minkoff. “NYC 311 A Tract-Level Analysis of Citizen–Government Contacting in New York City”. In: *Urban Affairs Review* 52.2 (2016), pp. 211–246.

[11] City of New York. *Dead animal*. 2016. URL: `http://www1.nyc.gov/nyc-resources/service/1487/dead-animal`.

[12] Daniel Tumminelli O’Brien. “Custodians and custodianship in urban neighborhoods: a methodology using reports of public issues received by a citys 311 hotline”. In: *Environment and Behavior* (2013).

[13] Daniel Tumminelli O’Brien, Eric Gordon, and Jessica Baldwin. “Caring about the community, counteracting disorder: 311 reports of public issues as expressions of territoriality”. In: *Journal of Environmental Psychology* 40 (2014), pp. 320–330.

[14] Sara M Santos, Filipe Carvalho, and António Mira. “How long do the dead survive on the road? Carcass persistence probability and implications for road-kill monitoring surveys”. In: *PLoS One* 6.9 (2011), e25383.