Final Report

Isaac Slagel and Jack Welsh

Introduction

According to an ASPCA estimate, over 6.5 millions animals enter a shelter system every year. (Cite 1). After entering the shelter system animals are either rehabilitated and sent to homes or euthanized. Animal shelters, while running mostly on revenue from donations and adoption fees, routinely run into perminant defieciets and must be supported with money from local governments (Cite 2). Knowing more about the outcomes of animals from these shelters could help us minimize the percent of animals that need to be euthanized. Further, by identifying which animals are less likely to be adopted, we can help shelters allocate resources to minimize the time these animals spend in the shelter system.

Our research looks into animal shelters in the Dallas area using a dataset provided by the city of Dallas on dallasopendata.com. This data includes information about every animal that has been brought into an animal shelter in the Dallas area. There are 61634 individual animals in this dataset with with information about when and where the animal was found, the health of the animal on arrival, and whether the animal was adopted, euthinized or transfered.

Recent research on animal shelters have consistantly shown that outcomes for pitbulls are consistantly worse than other dog breeds (Patronek & Crowe 2018, Lepper et al 2002). We are interested in seeing if these this trends exist in the Dallas animal shelter system.

Materials and Methods

We accessed our data from the Dallas Open Data website. This website contains many public datasets, including animal shelter records. Anyone can request an access key to this dataset and then download it. We downloaded the data, selected the variables of interest to us, and created a .csv file in the CreateDataset.Rmd file. Our dataset is saved as adoptions.csv. The initial variables from this dataset are described below in Table 1.

Table 1: Description of Variables

Variable Name	Variable Role	Variable Type	Range of Values	Units
animal_breed	explanatory	categorical	296 unique breeds	NA
animal_origin	explanatory	categorical	4 sources of shelter animals	NA
animal_type	explanatory	categorical	5 species of animal	NA
chip_status	explanatory	bianary	(0,1)	NA
intake_type	potential confounder	categorical	how animal came to be at the shelter	NA
outcome_type	reponse	categorical	how animals was removed from shelter	NA
intake_condition	potential confounder	categorical	keyword description of animal status at intake	NA
outcome_condition	potential confounder	categorical	keyword description of animal status at outcome	NA
intake_date	response	date	(2017-10-01, 2019-04-03)	y-m-d
outcome_date	response	date	(2017-10-01, 2019-04-03)	y-m-d

We used the imported data described in Table 1 to create a fleet of indicator variables: `adopted`, `chip_status`, `summer`, `treatable_intake`, `adopted`, `dead`. These indicator variables were created to streamline our modleing process. They allow us to increase the interpretability of variables with multiple categories (like outcome_type).

In order to fit a binomial regression, we had to summarise our dataset with respect to certain variables we controlled for in our model. One interesting issue that we encountered was the significance of the dataset we fit on SE estimates. If we summarized our data for each specific model, parameter estimates had increased SEs and more insignificant p-values. However, if we fit all of our models on the same, more extensive, summary table, SEs were small and parameters were more likely to be significant. We decided to proceed with the latter option, as it allows us to carry out nested F tests and would be more like a modeling process for true grouped data. This could mean that our SEs are artificaly low.

We used binomial regression to compare how the outcomes of pitbulls differ from other breeds of dog. In fitting this model we had to control for a plethora of cofounding variables including season of outcome, chip status, and intake condition of the animal. We also considered a multilevel model to account for difference in district correlation. Next we used a binomial regression to compare how the outcomes of dogs and cats differ, after controlling for season of outcome, chip status, and intake condition of the animal. We used findings from our exploratory data analysis to guide decisions on variable selection. Additionally we refered to models fit in the literature to guide decisions about model inclusion.

We also fit multilevel logistic regression models to compare ourcomes after accounting for correlation within city council districts. Our modeling process took on a similar approach to the binomial regression, we examined differences in odds of dying for pitbulls and non-pitbulls after accounting for differences in chip status, and season.

To investigate the ammount of time that dogs are spending in the animal shelter before dying we used cox proportional hazards. Cox proportional hazards is a semi parametric model used in survival to analyze how variables of interest affect the time till an event occuring. As the name suggests this is commonly used to investigate what factors affect time till death. This is done by comparing hazard rates using a hazard ratio. The hazard ratio is the ratio of hazard rates between groups, a hazard rate is the rate at which the outcome is happening. So, this type of model investigates how variables affect the rate at which outcomes happen.

The cox proportional hazards model is a semiparametric model, this means that this model does not make assumptions about the underlying distribution of data, but it does assume that the hazard ratios for any two individuals is constant over time. In other words, the risk of the outcome of interest happening for one individual at every time point must be proportionally constant to all other individuals.

This model follows the general form

$$h(t) = h_0(t) * e^{b_1 * x_1 + b_2 * x_2 + \dots + b_p * x_p}$$

.

where $h_0(t)$ is called the baseline hazard rate and is an estimation of the hazard rate of an individual at time t with x_1 to x_p all equal to 0 having the outcome of interest already happened. b_1 to b_p are the hazard ratio estimates. These are estimations of how much the hazard rate will change at time t compared to the baseline hazard rate at time t.
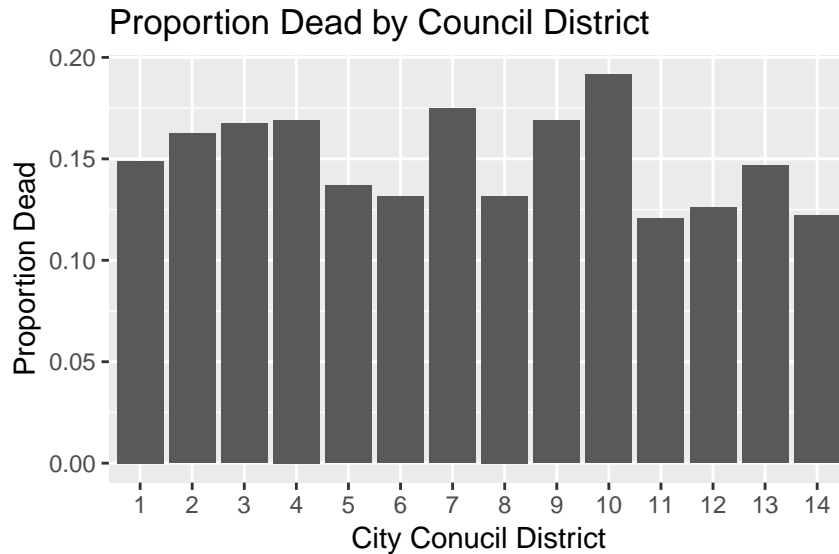
Results

EDA Results: Pitbull

To look into how pit bulls are handled within animal shelters, we decided to explore how the outcomes of pit bulls may differ from non-pitbulls. Table 1 describes the differences in outcome rates for non-pitbulls compared to pitbulls.

Table 2: Pitbull Outcomes

Outcome	Pitbull (%)	Non-Pitbull (%)
Adoption	32.03	35.73
Euthanized	28.42	10.25
Returned to owner	22.25	30.96
Transfer	10.92	17.56
Foster	2.25	1.64
Other	1.88	1.59
Dead on arrival	0.93	0.82
Treatment	0.77	1.01
Died	0.51	0.39
Missing	0.04	0.06

We see that pit bulls, while only adopted at a slightly lower rate (~5%), are euthanized at well over double the rate of other dogs. Further, we see that other dogs have a much higher chance of being transferred to another facility or returned to their owner. To get a better understanding of this relationship, but we likely need to control for things like chip status, intake condition, and season of outcome in a binomial regression.

EDA Results: Effect of Council District

EDA Results: Dog and Cat Differences

Binomial Models for Dog Outcomes

Table 3: Modeling Dog Outcomes in Dallas Animal Shelters

	Dependent variable:		
	Proportion of dogs who died		
	(1)	(2)	(3)
Intercept	0.116***	0.111***	0.552***
	(0.069,0.186)	(0.069,0.170)	(0.454,0.669)
Pitbull	3.440***	3.424***	3.489***
	(1.795,6.557)	(1.905,6.130)	(3.022,4.027)
Scannable Chip	0.789	0.799	0.781***
	(0.377,1.566)	(0.412,1.483)	(0.667,0.911)
Summer Outcome	1.461	1.447	1.478***
	(0.725,2.852)	(0.771,2.649)	(1.271,1.718)
Contagious		7.286**	3.975***
		(1.324,44.137)	(2.568,6.168)
Treatable At Intake			0.161***
			(0.133,0.196)
Overdisperson Parameter	139.72	111.46	6.27
Nested F Test		$F: 5.1142^*$	$F: 313.62^{***}$
Note:			*p<0.1; **p<0.05; ***p<0.01

We see in Table 3 that all of our models have variable estimates that are significant including Model 3 which accounted for whether or not an animal was deemed "treatable" at the time of intake. Further, by two F tests we see each of our models is an improvement on the one before it (Model 1 → Model 2: F = 5.1141, p-value = 0.031) (Model 2 → Model 3: F = 313.62, p-value <<< .05).

Procceding with model 3 as our best model of dog outcomes from the animal shelter, we glean several insights about the relationships between certain characteristics of dogs and those animal's outcomes from the shelter. For instance, if a dog is a pitbull, we expect it's odds of leaving the animal shelter dead increase by 348%, after controlling for chip status, season, and intake condition. Further, if a dog is a pitbull we are 95% confident that the true increase in odds of death at outcome is between 302% and 402%. Yikes! On the other hand, if a dog comes into the shelter with a scannable chip, we expect the animals odds of dying within the shelter system to drop by about 22%, after controlling for pitbull, season, and intake condition. Interestingly enough, if a dog has its outcome in the summer, the odds of that outcome being death are about 47% higher, controlling for breed, chip status, and intake condition.

Multilevel Model for Dog Outcomes

Our model has two levels,

-Level 1
$$log(\frac{p_{ij}}{1 - p_{ij}}) = a_i + b_i * summer_{ij} + c_i * pitbull_{ij} + d_i * chipstatus_{ij}$$

where

-Level 2 $a_i = \alpha_0 + u_i$ $b_i = \beta_0 + v_i$ $c_i = \gamma_0 + w_i$ $d_i = \delta_0 + x_i$

4

where

$$\begin{bmatrix} u_i \\ v_i \\ w_i \\ x_i \end{bmatrix} = N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & & & \\ \sigma_{uv} & \sigma_v^2 & & \\ \sigma_{uvw} & \sigma_{vw} & \sigma_w^2 & \\ \sigma_{uvwx} & \sigma{vwx} & \sigma_{wx} & \sigma_x^2 \end{bmatrix} \right)$$

This model was chosen based on information from our exploratory data analysis. . It is important to note that the odds of a pitbull dying in the animal shelter are 3.42 times higher than for non pitbulls accounting for season, chip status, and intake condition, see table 4. As expected, the odds of a contagious dog dying at the animal shelter are 3.96 time higher than a non contagious dog accounting for whether the dog is treatable, the season, chip status and pitbull. Also, unsuprisingly if a dog enters the animal shelter with a treatable intake the odds of dying in the animal shelter decrease by 84% controlling for whether the dog is contagious on intake, the season, and chip status.

Table 4: Predicting Dead Outcomes using Multilevel Logistic Regression

	Dependent variable:
	out_dead
Constant	0.506***
	(0.089)
pitbull	3.362***
	(0.051)
chip_status	0.780***
	(0.040)
summer	1.418***
	(0.053)
contagious	3.994***
	(0.089)
treatable	0.182*
	(0.108)
Observations	45,098
Log Likelihood	-16,709.350
Akaike Inf. Crit.	33,460.710
Bayesian Inf. Crit.	33,643.750
Note:	*p<0.1; **p<0.05; ***p<0.01

Modeling Dog and Cat Differences

In our modeling of Dog and Cat differences we used used binomial regression. To model this we considered the following model:

$$log(\frac{p_i}{1 - p_i}) = \beta_0 + \beta_1 DogYN + \beta_i(Confounder_i)$$

We considered several models with different confounders controled for based on findings in the literature and our EDA. We were not able to find any significant difference in the odds of death between dogs and cat. We conclude that there is not evidence of a significant difference between dogs and cats in odds of dying in an animal shelter after accounting for chip_status, season of outcome, and intake condition. These models can be found in the Dog Cat portion of our appendix.

Modeling Time Till Death for Dogs

Discussion

- Begin with an accurate summary statement; describe how the results help answer your research questions and what was most interesting from your analysis. In fact, the first paragraph of the Discussion is very important – in professional journals, it is often the first and sometimes the only paragraph that is read in a paper. After the first sentence highlights primary results, the remainder of the first paragraph might compare your results to others in the literature or include interesting secondary results.

- Discuss possible implications of the results in the context of the research question.

- Make a statement regarding potential confounding variables in your study

- Make a statement about the generalizability of your results. Don't give generic statements of possible causation and generalizability, but thoughtfully discuss relevant issues – confounding variables, representativeness of the sample, etc.

Conclusions about animal outcomes of animals in shelters from this report can only be extended to the city of Dallas. While our sample size is totally representative for this population, we cannot account for differences that would be encountered in other municipalities, say higher adoption rates or less stigma against certain breeds of dog. Further, we cannot be sure that the trends observed in this paper will continue into the future. Anti-pitbull advocacy groups like National Pitbull Victims Awareness and pro-pitbull advocacy groups like Love-a-Bull are at odds trying to sway public opinion and policy reguarding the handling of pitbulls in animal shelters. If changes in public percecption of pitbulls or the policy that surrounds them changes, our model will lose it's generalizability to Dallas.

- Identify any limitations of your study. Discuss the potential impact of such limitations on the conclusions.

One limitation of our study is that we do not know what happens to animals after they have been transferred out of the animal shelter. In our analysis we have viewed this outcome as a live outcome, but for all we know these animals could go on to another shelter where they are euthanized. Depending on the validity of our assumption, we could be adding some bias to our conclusions. Additionally, there is not consistant infomration on the criteria used by shelters to evaluate animals intake conditions. We look for key words that indicated the health of the animal. However we are not sure how consistant these evaluations are across all Dallas animal shelters.

- Identify strengths and weaknesses of your analysis.

- Make suggestions for future research. Identify important next steps that a researcher could take to build on your work.

Future work could be done in Dallas to help minimize the number of animals which need to be euthanized. By identifying specific at-risk groups of animals, more resources could be allocated to advocating for these animals or transfering them to other No-Kill shelters. We would recomend any additional research to work directly with the Dallas Animal Shelter group. By having a direct contact with the group doing data collection, a statistician will be able to better identify pressing issues of the group and verify certain assumptions that we have not been able to verify.

Works Cited

1. https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender
2. https://rucore.libraries.rutgers.edu/rutgers-lib/38418/PDF/1/play/

Patronek, G. J., & Crowe, A. (2018). Factors Associated with High Live Release for Dogs at a Large, Open-Admission, Municipal Shelter. Animals : an open access journal from MDPI, 8(4), 45. doi:10.3390/ani8040045

Lepper, M., Kass, P. H., & Hart, L. A. (2002). Prediction of adoption versus euthanasia among dogs and cats in a California animal shelter. Journal of Applied Animal Welfare Science, 5(1), 29-42.

Posage, J. M., Bartlett, P. C., & Thomas, D. K. (1998). Determining factors for successful adoption of dogs from an animal shelter. Journal of the American Veterinary Medical Association, 213(4), 478-482.

Lampe, R., & Witte, T. H. (2015). Speed of dog adoption: Impact of online photo traits. Journal of applied animal welfare science, 18(4), 343-354.

5. Annotated Appendix

- Tables and figures that are informative but were not referenced specifically in the main report. Include a short annotation – one or two sentences on what they show.
- R scripts and output (annotated) so that I can trace how you constructed your final data set, what models you ran to produce the results quoted in your report, and what intermediate models you also considered.
- Description of statistical modeling steps that were not included in the main body of your report. Possible entries here include: How you handled missing data. Evaluation of assumptions. Outlier analysis and how you decided to deal with any outliers along with rationale for your decision. Describe hypotheses testing you performed during model building and how you decided on the explanatory variables you ultimately included in your final model. Assessment of the final model.
- How you went from the model output in R to interpretations in your report (e.g. exponentiate coefficients, then take inverse)
- Anticipate questions someone might have after reading your report, and make sure those questions can be answered with information in the appendix.
- A citation for each reference article (in APA format or something similar) you included in your proposal. Also include a link, if appropriate. Remember that you must have the entire paper and not just an abstract, and at least two must be from peer-reviewed journals.

Dogs vs Cats

```
cat_dog_binom <- adoptions %>%
  filter(!str_detect(intake_subtype, "(DEAD)|(DIED)")) %>%
  filter(dog == 1 | cat == 1) %>%
  mutate(out_dead = outcome_type %in% c("DEAD ON ARRIVAL", "EUTHANIZED", "DIED"),
         summer = ifelse(month %in% c(5, 6, 7, 8, 9), 1, 0))%>%
  mutate(chip_status = ifelse(chip_status=="SCAN CHIP", 1, 0)) %>%
  group_by(dog, chip_status, summer, stray) %>%
  summarize(prop_dead = sum(out_dead)/n(), count = n())

####### MODEL 1: dog #########

dogcat_model1_binom <- glm(prop_dead ~ dog, weights = count, family = binomial, data = cat_dog_binom)
```

```
dogcat_model1_quasi <- glm(prop_dead ~ dog, weights = count, family = quasibinomial, data = cat_dog_bin
```

```
summary(dogcat_model1_binom)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog, family = binomial, data = cat_dog_binom,
##     weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -19.837   -7.385   -2.250    4.985   15.639
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.43741    0.02199  -65.38   <2e-16 ***
## dog         -0.30509    0.02563  -11.90   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1745.6  on 16  degrees of freedom
## AIC: 1866
##
## Number of Fisher Scoring iterations: 4
```

```
exp(confint(dogcat_model1_binom))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.2274800 0.2479563
## dog         0.7010303 0.7751213
```

```
summary(dogcat_model1_quasi)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog, family = quasibinomial, data = cat_dog_binom,
##     weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -19.837   -7.385   -2.250    4.985   15.639
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.4374     0.2249  -6.393 8.91e-06 ***
## dog          -0.3051     0.2621  -1.164    0.261
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 104.5899)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1745.6  on 16  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model1_quasi))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %     97.5 %
## (Intercept) 0.1495373 0.3624932
## dog         0.4457027 1.2501118
```

```
####### MODEL 2: dog + summer #########
```

```r
dogcat_model2_binom <- glm(prop_dead ~ dog + summer, weights = count, family = binomial, data = cat_dog_

dogcat_model2_quasi <- glm(prop_dead ~ dog + summer, weights = count, family = quasibinomial, data = ca

summary(dogcat_model2_binom)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + summer, family = binomial, data = cat_dog_binom,
##     weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -17.003   -6.915   -5.412    5.530   19.155
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.61541    0.02452 -65.870   <2e-16 ***
## dog         -0.24951    0.02590  -9.632   <2e-16 ***
## summer       0.42766    0.02392  17.879   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1434.4  on 15  degrees of freedom
## AIC: 1556.8
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model2_binom))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.1894425 0.2085590
## dog         0.7407010 0.8198698
## summer      1.4633398 1.6071906
```

```r
summary(dogcat_model2_quasi)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + summer, family = quasibinomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q     Max
## -17.003  -6.915   -5.412    5.530   19.155
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.6154     0.2344  -6.892 5.13e-06 ***
## dog          -0.2495     0.2476  -1.008    0.330
## summer        0.4277     0.2286   1.871    0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 91.33309)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1434.4  on 15  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model2_quasi))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.1232172 0.3096656
## dog         0.4842912 1.2820599
## summer      0.9734565 2.3902024
```

```r
####### MODEL 3: dog + chip ######

dogcat_model3_binom <- glm(prop_dead ~ dog + chip_status, weights = count, family = binomial, data = ca

dogcat_model3_quasi <- glm(prop_dead ~ dog + chip_status, weights = count, family = quasibinomial, data

summary(dogcat_model3_binom)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + chip_status, family = binomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -15.6677   -4.1391    0.1846   10.8318   14.4792
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.40636    0.02211 -63.599   <2e-16 ***
## dog         -0.24565    0.02620  -9.377   <2e-16 ***
## chip_status -0.29305    0.02828 -10.363   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1750.8  on 15  degrees of freedom
## Residual deviance: 1503.0  on 13  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: 1625.4
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model3_binom))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.2345951 0.2558385
## dog         0.7431337 0.8235017
## chip_status 0.7056151 0.7883420
```

```r
summary(dogcat_model3_quasi)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + chip_status, family = quasibinomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -15.6677   -4.1391    0.1846   10.8318   14.4792
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.4064     0.2391  -5.883 5.38e-05 ***
## dog          -0.2456     0.2832  -0.867    0.401
## chip_status  -0.2931     0.3057  -0.959    0.355
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 116.867)
##
##     Null deviance: 1750.8  on 15  degrees of freedom
## Residual deviance: 1503.0  on 13  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```
exp(confint(dogcat_model3_quasi))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %     97.5 %
## (Intercept) 0.1496550 0.3838132
## dog         0.4540004 1.3839138
## chip_status 0.3987708 1.3312323
```

```
####### MODEL 4: summer + dog + summer:dog ######
```

```
dogcat_model4_binom <- glm(prop_dead ~ summer+dog+ summer:dog, weights = count, family = binomial, data
```

```
dogcat_model4_quasi <- glm(prop_dead ~ summer+dog+ summer:dog, weights = count, family = quasibinomial,
```

```
summary(dogcat_model4_binom)
```

```
##
## Call:
## glm(formula = prop_dead ~ summer + dog + summer:dog, family = binomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -17.392   -7.749   -4.789    6.527   18.670
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.67600    0.03030 -55.313  < 2e-16 ***
## summer       0.56018    0.04434  12.634  < 2e-16 ***
## dog         -0.17185    0.03420  -5.024 5.05e-07 ***
## summer:dog  -0.18718    0.05275  -3.549 0.000387 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1421.8  on 14  degrees of freedom
## AIC: 1546.2
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model4_binom))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %     97.5 %
## (Intercept) 0.1762591 0.1984903
## summer      1.6052408 1.9099840
## dog         0.7877336 0.9007624
## summer:dog  0.7478053 0.9195843
```

```r
summary(dogcat_model4_quasi)
```

```
##
## Call:
## glm(formula = prop_dead ~ summer + dog + summer:dog, family = quasibinomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -17.392   -7.749   -4.789    6.527   18.670
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.6760     0.2978  -5.628 6.23e-05 ***
## summer        0.5602     0.4358   1.285    0.220
## dog          -0.1718     0.3362  -0.511    0.617
## summer:dog   -0.1872     0.5185  -0.361    0.723
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 96.60753)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1421.8  on 14  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model4_quasi))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.09987833 0.324086
## summer      0.74071253 4.144038
## dog         0.44621757 1.682512
## summer:dog  0.29772816 2.292399
```

```r
####### MODEL 5: stray ######
```

```
dogcat_model5_binom <- glm(prop_dead ~ dog + stray, weights = count, family = binomial, data = cat_dog_

dogcat_model5_quasi <- glm(prop_dead ~ dog + stray, weights = count, family = quasibinomial, data = cat_

summary(dogcat_model5_binom)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + stray, family = binomial, data = cat_dog_binom,
##     weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -16.892   -6.265   -2.337    5.256   18.556
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.23402    0.02551  -48.38   <2e-16 ***
## dog         -0.28633    0.02571  -11.13   <2e-16 ***
## stray       -0.34943    0.02308  -15.14   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1519.7  on 15  degrees of freedom
## AIC: 1642.1
##
## Number of Fisher Scoring iterations: 4
```

```
exp(confint(dogcat_model5_binom))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.2768768 0.3059909
## dog         0.7141890 0.7899317
## stray       0.6739322 0.7377448
```

```
summary(dogcat_model5_quasi)
```

```
##
## Call:
## glm(formula = prop_dead ~ dog + stray, family = quasibinomial,
##     data = cat_dog_binom, weights = count)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -16.892   -6.265   -2.337    5.256   18.556
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.2340     0.2487  -4.961 0.000171 ***
## dog          -0.2863     0.2508  -1.142 0.271435
## stray        -0.3494     0.2251  -1.553 0.141379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 95.10904)
##
##     Null deviance: 1883.0  on 17  degrees of freedom
## Residual deviance: 1519.7  on 15  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

```r
exp(confint(dogcat_model5_quasi))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.1755655 0.4670297
## dog         0.4639650 1.2440641
## stray       0.4546178 1.1008372
```