

# Influential factors of days spent at the shelter

Hanfan Chen 2703334C, Zhaohao Li 2611602L, Zhenhao Qiao 2612824Q, Chao Wang 2518608W, Rachan

## 1 Introduction

## 2 Exploratory Data Analysis

	animal_type	month	year	intake_type	outcome_type	chip_status
1	CAT	9	2017	STRAY	ADOPTION	UNABLE TO SCAN
2	DOG	6	2017	STRAY	EUTHANIZED	SCAN NO CHIP
3	DOG	12	2016	STRAY	ADOPTION	SCAN NO CHIP
4	DOG	9	2017	STRAY	ADOPTION	SCAN NO CHIP
5	CAT	11	2016	OWNER SURRENDER	ADOPTION	SCAN CHIP
6	CAT	12	2016	OWNER SURRENDER	ADOPTION	SCAN NO CHIP

	time_at_shelter
1	9
2	4
3	21
4	4
5	7
6	4

Column 1 :

[1] "BIRD" "CAT" "DOG" "WILDLIFE"

Column 2 :

[1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" "11" "12"

Column 3 :

[1] "2016" "2017"

Column 4 :

[1] "CONFISCATED" "OWNER SURRENDER" "STRAY"

Column 5 :

[1] "ADOPTION" "DIED" "EUTHANIZED"  
[4] "FOSTER" "RETURNED TO OWNER"

Column 6 :

[1] "SCAN CHIP" "SCAN NO CHIP" "UNABLE TO SCAN"

[1] 317

Excess zeros in raw data may cause overdispersion in Poisson regression. The hurdle model is suggested to fit.

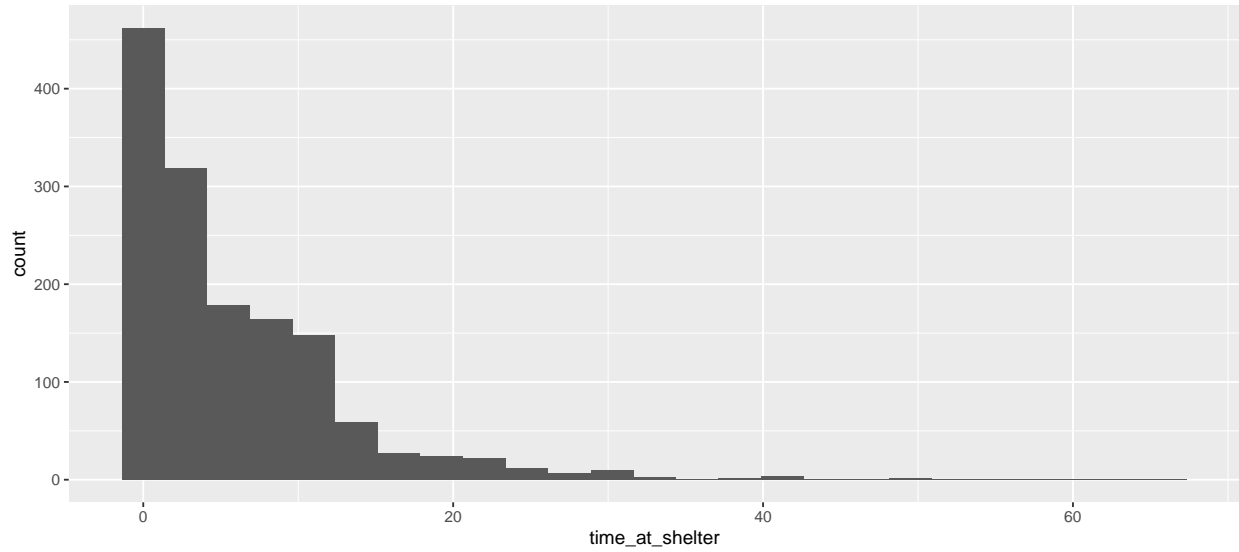


Figure 1: The histogram of day time at shelter

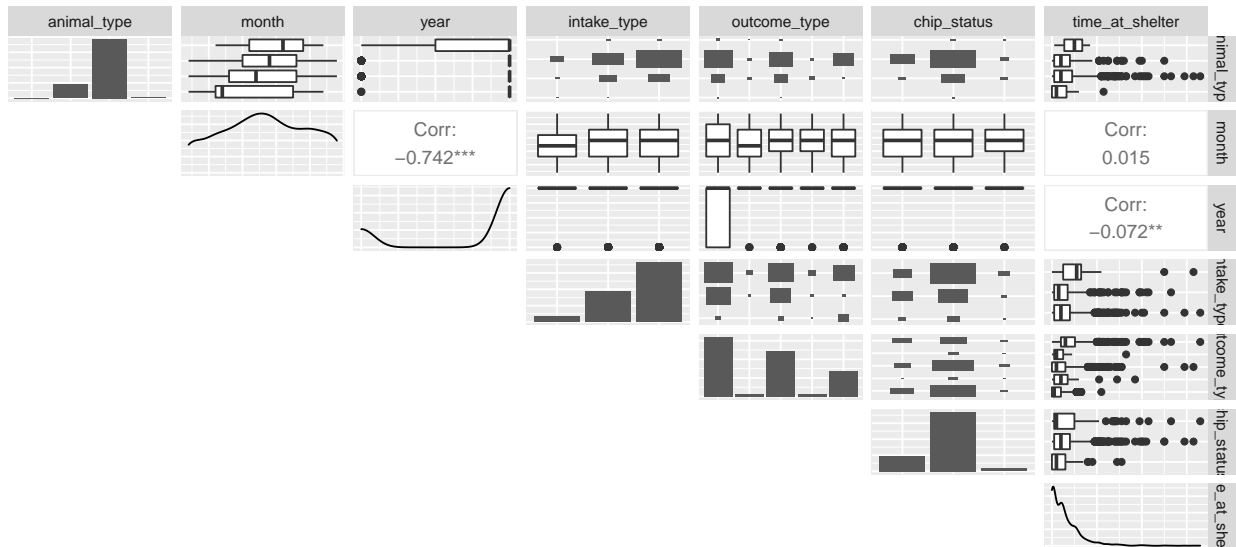


Figure 2: Pair plots of the variables

Note that year and month are strongly correlated, suggesting multicollinearity. As the data were collected over the period of a year, month and year represent the same variable, namely when the animal was admitted to the shelter. Therefore, year shall be omitted from the model.

The other explanatory variables are all categorical and their box plots are shown. The median time at shelter appears to be low for all the explanatory variables, which is due to the median time at shelter being 4.

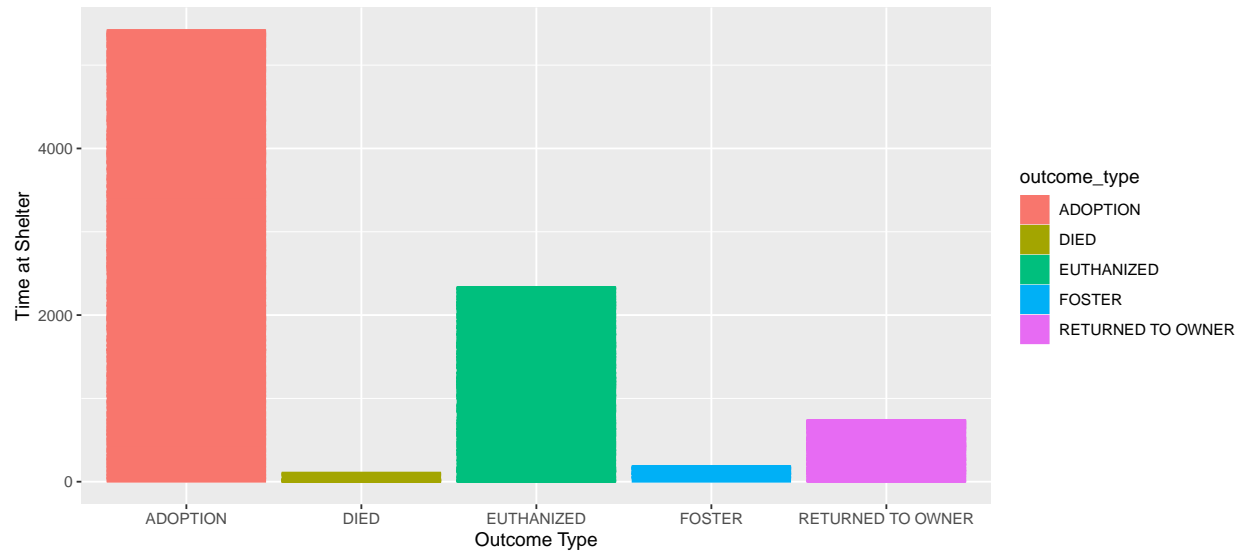


Figure 3: Option 1: Bar plot of outcome type vs time at shelter

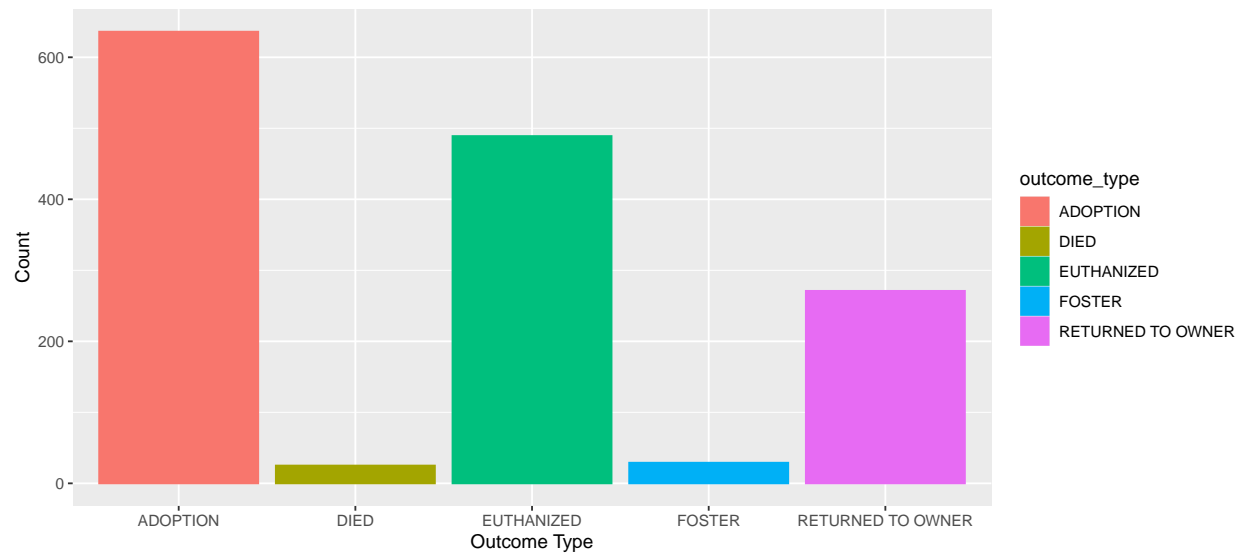


Figure 4: Option 2: Bar plot of outcome type vs count

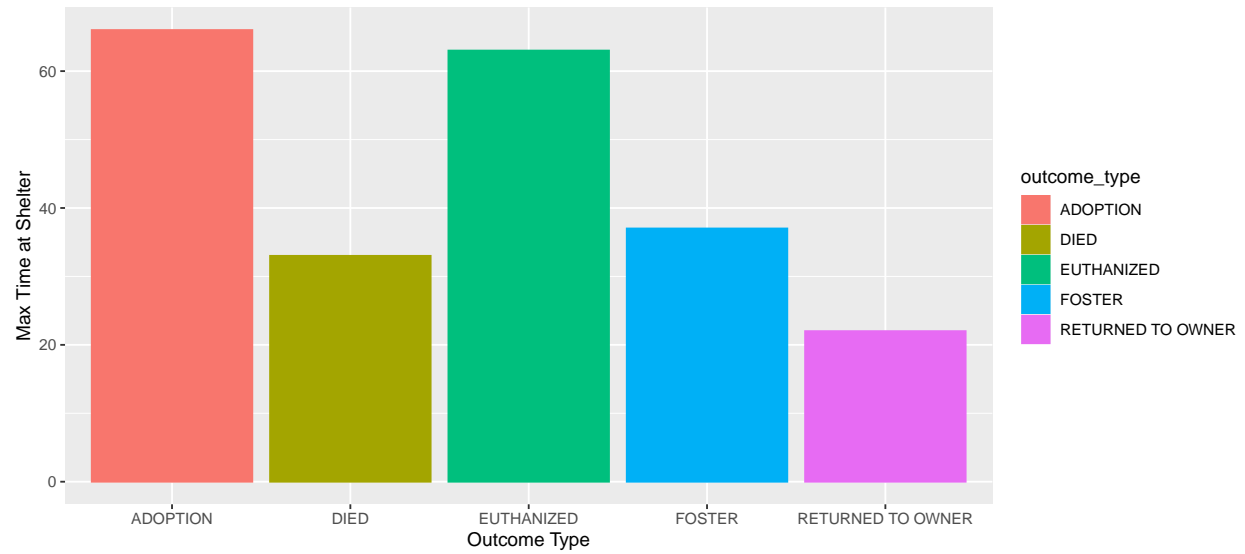


Figure 5: Option 3: Bar plot of outcome type vs max time at shelter

This is just so you can see that there's no overlap between the months and years.

	month
1	12
2	11
7	10

	month
1	9
2	6
5	1
6	2
7	4
11	3
13	5
22	8
31	7

### 3 Formal Data Analysis

### 4 Conclusions