

Project Proposal

Jefferson Ratliff, Sydney Stitt, Evi Jonas

2024-02-23

Data Set: Week 18 Dallas Animals

```
readr::read_csv("week18_dallas_animals.csv")%>%  
  head(n = 5)
```

```
## # A tibble: 5 x 13  
##   animal_id animal_type animal_breed   month   year census_tract council_district  
##   <chr>      <chr>      <chr>      <chr> <dbl> <chr>      <chr>  
## 1 A0979593 DOG        RHOD RIDGEBACK FEB     2017 1502        2  
## 2 A0743013 DOG        YORKSHIRE TERR NOV     2016 13609       11  
## 3 A1004433 BIRD        CHICKEN      AUG     2017 10803        3  
## 4 A0969724 DOG        GERM SHEPHERD DEC     2016 7102        2  
## 5 A0981479 DOG        GERM SHEPHERD FEB     2017 6001        4  
## # i 6 more variables: intake_type <chr>, intake_date <chr>, outcome_type <chr>,  
## #   outcome_date <chr>, chip_status <chr>, animal_origin <chr>
```

This data set uses information from a variety of different animal shelters across the City of Dallas, upon their intake to the shelter. Each observation/row is an individual animal entered by staff into the shelter system. The data is across all animal types including dogs, cats and other wildlife. The data collections the animals health, chip status, breed/species, as well as the time of their intake and the form in which the animal was received. We chose this data set because we are dog lovers and one of our group members is from Dallas and volunteered at a local shelter. It also seemed to have well collected data with a lot of diversity within key variables like dog breed and intake form.

Variables:

- 1) Animal_id (ID number for each animals)
- 2) Animal_type (dog, cat, wildlife)
- 3) Animal_Breed
- 4) Month (intake month)
- 5) Month (intake Year)
- 6) Census_tract (Geographical Location)
- 7) Council_district (Geographical Boundries)
- 8) Intake_type
- 9) Outcome_type
- 10) Outcome_date(When animals leave the shelter)
- 11) Chip_status (is the animal chipped?)
- 12) animal_origin (How they got the animal)

##2 Research Questions:

1. Which city council districts have the highest rate of adoption? (council_district and outcome_type)
Filter outcome_type for observations only including adoptions using dplyr. Quantify how many observations of adoptions there are in each district. Graph using bar graph with ggplot2.
2. What types of animal intake tend to stay longer in shelters? Mutate new variable for duration of stay with dplyr, bar graph with error bars for average stay (y axis) and type of intake (x axis) with ggplot2