

Lost and Found Pets in King County 2016

CSE 160 - Final Project
Josephine Le and Allyson Kline

Research Questions

1. *How many pets have been lost and how many have been found?*

Since the dataset covers information about lost and found animals, it is important to provide an overview of this information. We can keep a count to calculate the amount of pets lost and the amount found.

2. *Where are most pets found?*

Perhaps some cities are more monitored than others for lost or stray animals. If someone loses a pet, they may have a better chance of him or her being found in a certain city. We can have a counter keep track of how many animals are found per city and compare them amongst each other to get the top three cities in King County where pets are found.

3. *Is there a correlation between the day of the week and the amount of pets lost?*

Maybe pets are more susceptible to being lost on certain days of the week. This information can help keep pet owners more alert. We can compute the amount of lost pets per day then compare each day to see if there are any patterns.

4. *How many adoptable animals are there and where are they located?*

To end our report, we would like to calculate how many adoptable pets there currently are in King County to bring awareness to people how many animals need a loving forever home.

5. *How are adoptable pets described?*

Since our dataset contains descriptions of most of the animals, we would like to compile this information into a word cloud to further share information about the currently available pets for adoption.

Motivation & Background

Pets are an important part of many modern homes. Most pets even feel more like family members than simple animals. Losing a beloved pet can be a troublesome situation causing a great deal of stress and worrying. Answering the questions above can help pet owners be more alert and know where they have a better chance of a lost pet being found. In addition, there are also many animals that need to be adopted into caring homes. Adopting a pet can help one learn and practice responsibility and time

management along with creating an emotional bond and sense of friendship. By sharing the currently available adoptable pets, we can provide awareness of the pets who are ready to love and be loved.

Dataset

<https://data.kingcounty.gov/Pets/Lost-found-adoptable-pets/yaai-7frk>

The dataset we found contains data concerning lost, found, and adoptable pets in King County, Washington from 2014-2016. It is in CSV (comma-separated values) format and is publicly available online at data.kingcounty.gov under the category, "Pets". It is updated daily, so pets lost and/or found that have been returned or already adopted no longer appear in the list.

Methodology

Our methodology will be based on answering each one of our research questions. First, the data will need to be read in and cleaned up for easier data wrangling and computation. Blank entries will be ignored. For each of our research questions, we will:

1. Count of the total amounts of lost pets as well as found pets to be calculated and showcased on a visual chart
2. For each given city, the amount of found pets will be calculated and compared against each other to determine the city with the most found pets. Each city will be graphed along with their respective found pet counts.
3. For each day, we will compute how many pets were lost, group dates by day of the week, and compare the data to see if there are any correlating patterns. This information will be plotted along with a best fit line.
4. Calculate the total amount of adoptable pets there currently are in King County and collect their current locations. This information will be graphed on a map.
5. Descriptions of each adoptable pet will be collected and put into a word cloud visual representation.

Work Plan

1. Clean the data (2 hours)
2. Gather the necessary data to answer the first 3 research questions & test code (5 hours)
3. Gather the necessary data to answer the last 2 research questions & test code (3 hours)
4. Create visual representations for each question (4 hours)
5. Compile entire report and practice presentation (2 hours)

We will hopefully spend about half the time coding and collaborating in person together. When this is not the case, we will share files through Google Drive or a shared Github repository, keeping a log of what has been done or changed and what still needs to be accomplished. We will keep in contact through means of online communication to clarify any questions or concerns.