Kevin Pennell
3736175
CS2704 Final Project Proposal
March 30th 2025

Data set:
https://opendata.vancouver.ca/explore/dataset/animal-control-inventory-lost-and-found/table/?disjunctive.color&disjunctive.breed&sort=date

GitHub Repo:
https://github.com/KevinPennell/CS2704-final-project

Hypothesis:

I would like to observe the lost and found data and find out what breeds of dogs and cats and other pets are more likely to be lost and then found, opposed to never found.
My hypothesis is that dogs will show higher rates of being found due to the nature of loyalty in canines. However, i expect the rates of being found to be like cats, and should show a percentage within 5% of each other, with dogs having a slight edge. I may observe other animals if the data set is large enough, but my exploration will mainly consist of analyzing dogs and cats

Testing plans:

My plan is to test the data on a breed level, i would like to create a bar graph that will represent the percentages of found pets. It should be simple to calculate. However, in examining the data i may need to spend a lot of work sorting the data. Cats are listed as 'Cat - breed' while dogs just list the breed, and there's a few other animals in the data set, so I may need to manually code in cases for data if it does not belong to cats or dogs