

PROJECT DESCRIPTION

Data description:

Link: <https://github.com/the-pudding/data/tree/master/dog-shelters>

This is data about adoptable dogs in each of the US states and the District of Columbia. The main information on the two datasets in adoptable dogs are shown below:

`allDogDescriptions.csv`:

Data collected from the PetFinder API for all adoptable dogs in each state on September 20, 2019. There are 58,180 rows in this dataset. Each row represents an individual adoptable dog in the US on September 20, 2019. Each dog has a unique ID number. And there are 36 columns in this dataset, including id and some detail information on dogs.

`dogTravel.csv`:

This file aims to show where those dogs are available and where they came from. There are 2,460 rows in this dataset. Each row represents a single dog that was available for adoption somewhere in the US. Each of these dogs is described as having been moved from another location to their current location. And there are 8 columns in this dataset.

Plan:

Firstly, I am going to do some data tidying. Since there are some code in Github to clean the data, I will learn from it but do something different. And then, as the data shown, the two datasets can be joined via id column. From the data I want to show some information like people in which state are more likely to adopt dogs, or what kind of foster home is suitable for the adoptable dogs.