Classifying Pet Product Reviews

Project Repository: https://github.com/halemitch/CS3-Project/tree/main
A DS-4002 Project by Haley Mitchell

Have you ever worried that the pet supplies you are giving your pets are causing them harm? Pet owners often have to put faith in trusted companies to make these pet supplies, believing that they will keep the safety of their pets in mind. This is unfortunately not the reality of all pet product companies. According to Thixton, "in the last five years…the FDA issued 45 different [pet food] recall announcements [2]" which equates to "69,364,194 [2]" lbs of pet food. Although the causes for sickness and the effects differed, the fact of the matter is that many pets were harmed as a direct result of these companies' products [2].

To portray the gravity of poorly made/contaminated pet supplies, an article from 2019 states that the FDA warned that Midwestern Pet Foods Inc.'s dry pet food was "linked to the deaths of 130 dogs and…may have sickened an additional 220 pets [1]". This is the extensive damage that one company's contaminated pet products can have.

As a data scientist, your task is to analyze pet supply reviews to identify which words distinguish good from bad reviews. You will start by performing a basic EDA, which will help you understand the data you will be using. Subsequently, you will be creating a logistic regression model to classify Amazon pet supplies reviews into two groups: "bad" (star ratings 1-3) and "good" (star ratings 4-5). You will also do text vectorization using TF-IDF (Term Frequency-Inverse Document Frequency), where the TF component will highlight the frequency of a word and the IDF component will highlight distinctive words [2]. The results you provide could be pivotal towards helping pet owners distinguish how good a pet product is directly from the language of the reviews.

[1]   Grullón-Paz, "Recalled Pet Food May Be Linked to 130 Dog Deaths, F.D.A. Says"
        New York Times, para. 1-2, August 19, 2021. [Online], Available:
        https://www.nytimes.com/2021/08/19/business/midwestern-pet-food-fda.html [Accessed Sept. 11, 2025].

[2]   S. Thixton, "45 Pet Food Recalls over the Past Five Years," *Truth About Pet Food,* July 17, 2025. [Online],
        Available: https://truthaboutpetfood.com/45-pet-food-recalls-over-the-past-five-years/. [Accessed Sept. 11,
        2025].