Examining Public Comments on the Labeling of Dairy Products

Final Project
94-775 Unstructured Data Analytics
Spring 2019

By: David Contreras,
Honghua Li, Batshur Gootiiz, Alisa Hartle,
Nora Murray

Table of Contents

Project Overview – Background

DAIRY INDUSTRY IN CRISIS

1. **DECREASED CONSUMPTION OF MILK:**

Americans' consumption of milk has been steadily decreasing since the 1970s. More recently, milk sales fell from 54.5 million pounds in 2009 to 45.5 million pounds in 2017[1].

2. **CHANGES IN PRODUCTION TRENDS:**

Changes in industrial agriculture has made producing milk less profitable for individual dairy farmers. Since 1980, the number of small family-owned dairies has dropped by between 5% to 9% each year[2].

1. Agricultural Marketing Service; 2005 to 2017
2. Sananes, Rebeka (2017) "As Big Milk Moves In, Family-Owned U.S. Dairy Farms rapidly fold". Retrieved on 3/2/2019 from: https://www.npr.org/sections/thesalt/2017/01/11/509135189/as-big-milk-moves-in-family-owned-u-s-dairy-farms-rapidly-fold

Project Overview – Background

FOOD & DRUG ADMINISTRATION MISSION

*The Food and Drug Administration is responsible for protecting the public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices; and by **ensuring the safety of our nation's food supply**, cosmetics, and products that emit radiation.*

Project Overview – Background Cont.

DAIRY INDUSTRY TRIES TO MAINTAIN ITS MARKET SHARE BY:

1. **LOBBYING THE FOOD & DRUG ADMINISTRATION**

2. **MAINTAINING AN EXCLUSIVE CLAIM TO USE THE 'DAIRY' TERMS SUCH AS:**

   MILK
   CHEESE
   YOGURT
   BUTTER

Project Overview – Policy Question

In determining the final labeling status of milk, in September 2018 the FDA has asked for public comment on the following topics:

- *How do you use plant-based products?*
- *What is your understanding of dairy terms like milk, yogurt, and cheese when they are used to label plant-based products?*
- *Do you understand the nutritional characteristics of plant-based products? Do you know how they're different from each other?*
- *Do you know how their nutritional qualities compare with dairy products?*

The FDA will use the public comments to help inform their final decision on whether the terms dairy and milk should be reserved for animal-derived products.

Project Overview – Policy Question

Our research aims to add insight to the questions posed by the FDA:

- **Do people support or oppose increased regulation for dairy product labeling?**

- **Can we uncover the affiliations of those submitting comments and understand how their support differs?**

Methodology

- **Data retrieved using the API from regulations.gov**

- **Total comments:** 14,010

- **Data type:** plain-text comments and attachments

- **Methods used:**

  - Topic Modelling

  - Entity Analysis

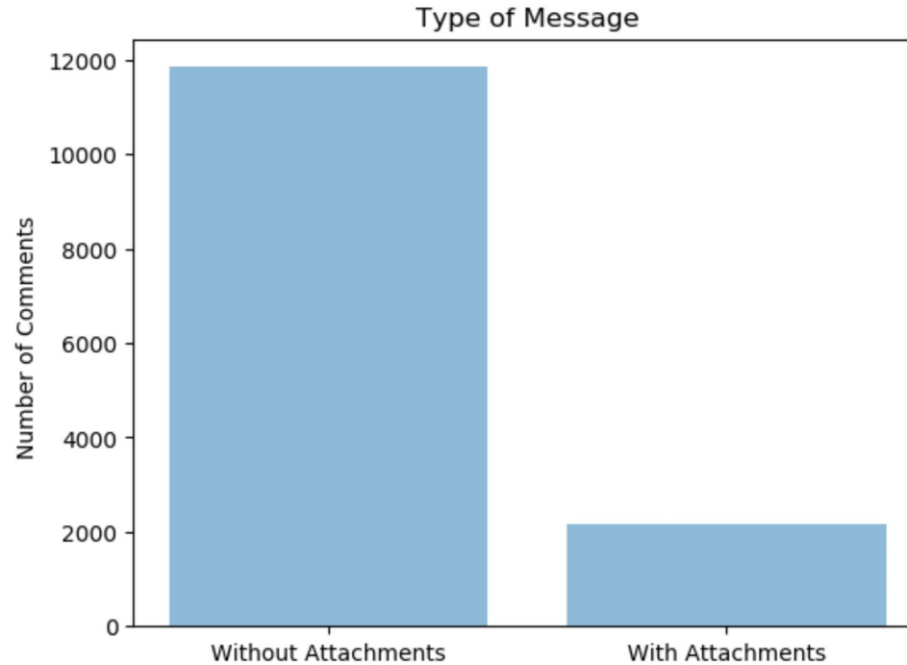 - Sentiment Analysis

  - Predictive Modeling

Sample Data

{"allowLateComment":false,"commentDueDate":null,"commentOnDoc":{"documentId":"FDA-2018-N-3522-0001","documentType":"NOTICES","title":"Use of the Names of Dairy Foods in the Labeling of Plant-Based Products"},"openForComment":false,"postedDate":"2018-10-01T00:00:00-04:00","receivedDate":"2018-09-28T00:00:00-04:00","status":"Posted","docketTitle":{"label":"Docket Title","value":"Use of Dairy Terms in the Labeling of Plant-Based Products"},"docketType":{"label":"Docket Type","value":"Nonrulemaking"},"documentType":{"label":"Document Type","value":"Public Submission"},"docSubType":{"label":"Document SubType","value":"Electronic Regulation from Form"},"attachmentCount":{"label":"Attachment Count","value":"0"},"agencyName":{"label":"Agency Name","value":"Food and Drug Administration"},"title":{"label":"Document Title","value":"Comment from Shannon Vaughn"},"trackingNumber":{"label":"Comment tracking Number","value":"1k2-95oc-c2vd"},"commentCategory":{"label":"comment_category","value":"Health Professional - A0007"},"docketId":{"label":"Docket ID","value":"FDA-2018-N-3522"},"comment":{"label":"Comment","value":"Yes I do know the difference when milk or yogurt is included in the names of plant-based products. As well as understanding differences between the basic nature, characteristics, ingredients, and nutritional content of plant-based products and their dairy counterpart.\n\nI do not purchase animal dairy products for nutritional reasons, same for plant-based dairy products. "},"documentId":{"label":"Document ID","value":"FDA-2018-N-3522-0004"},"numItemsRecieved":{"label":"Number of Comments Received","value":"1"},"agencyAcronym":{"label":"Agency","value":"FDA"}}
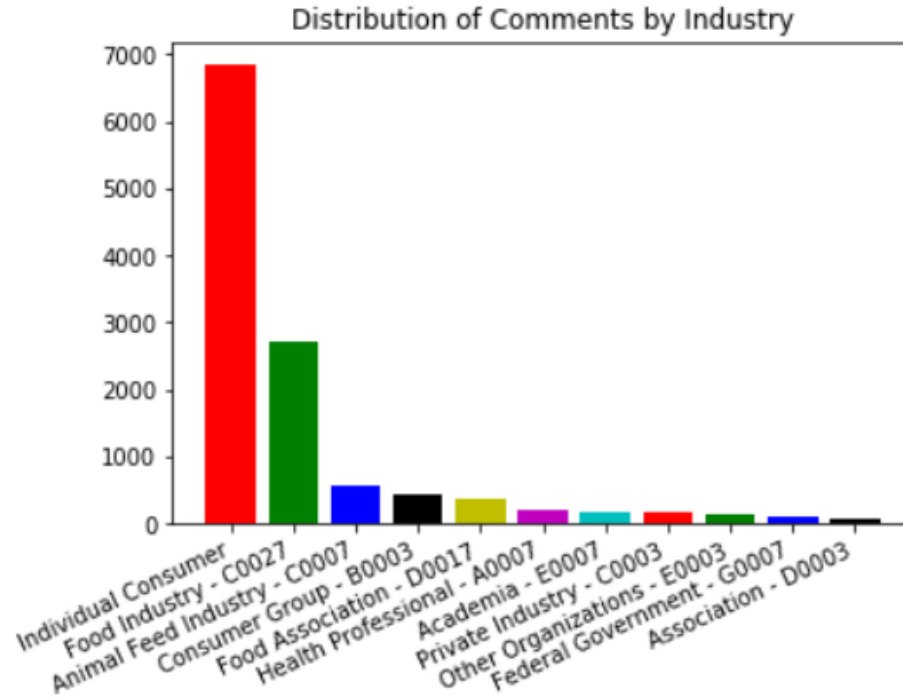
Executive Summary

- Entity analysis shows dominance of interest groups and information sources in comments.

- Topic modeling reveals breakdown along the themes of labelling regulations, diet, and animal abuse.

- Prediction modeling provides insight on dairy labeling opinions in the larger context.

- After filtering by the Food Industry, entity analysis and topic modeling shows that the Food Industry isn't just representative of the dairy industry, but includes viewpoints from the health food/plant-based alternative sector

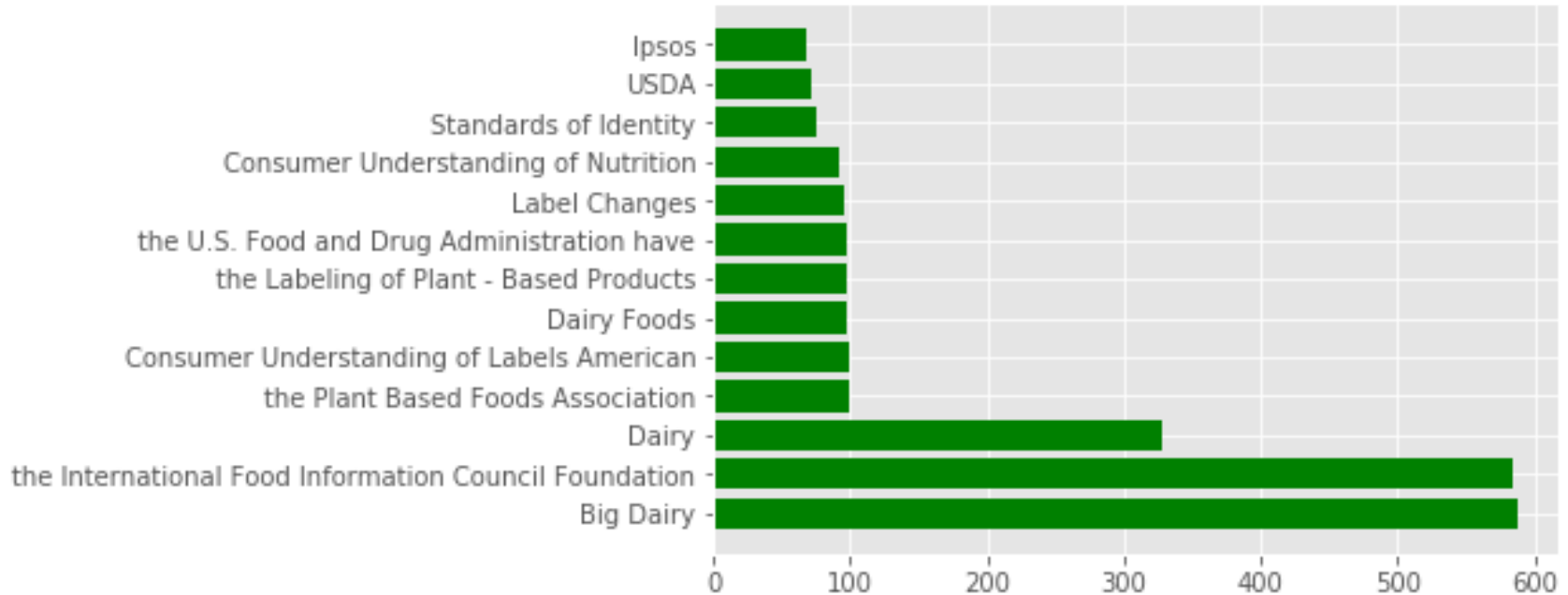# MAJORITY OF PUBLIC COMMENTS SUBMITTED ELECTRONICALLY

Comments submitted with attachments are a good proxy to represent organizational interests in the form of form letters or mass mailer campaigns.

MAJORITY OF PUBLIC COMMENTS SUBMITTED BY INDIVIDUAL CONSUMERS
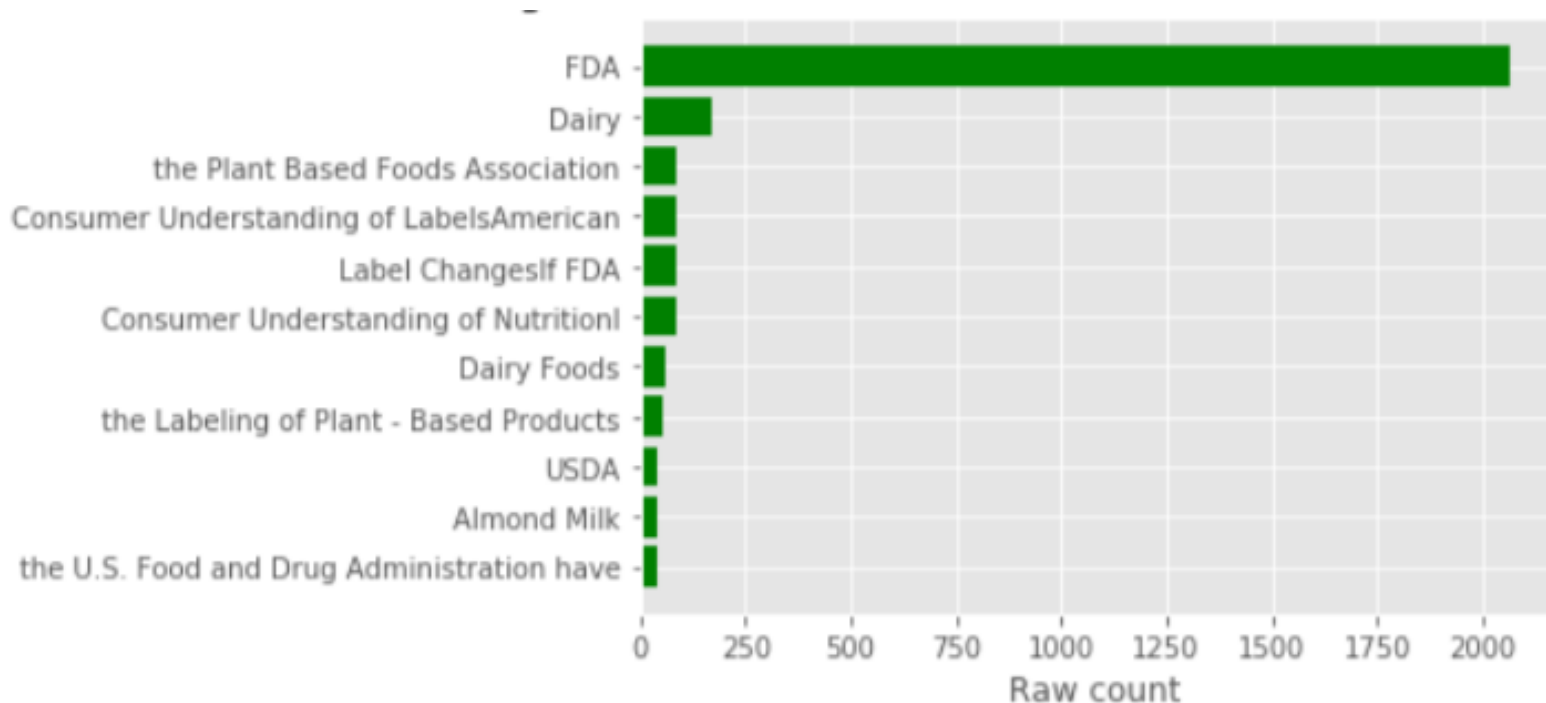

Distribution of Comments by Industry

While not all comments contained information regarding who submitted the comment, more than half of the comments were submitted by individual consumers. Approximately 1 in 4 comments with submitter information were submitted from food industry representatives.
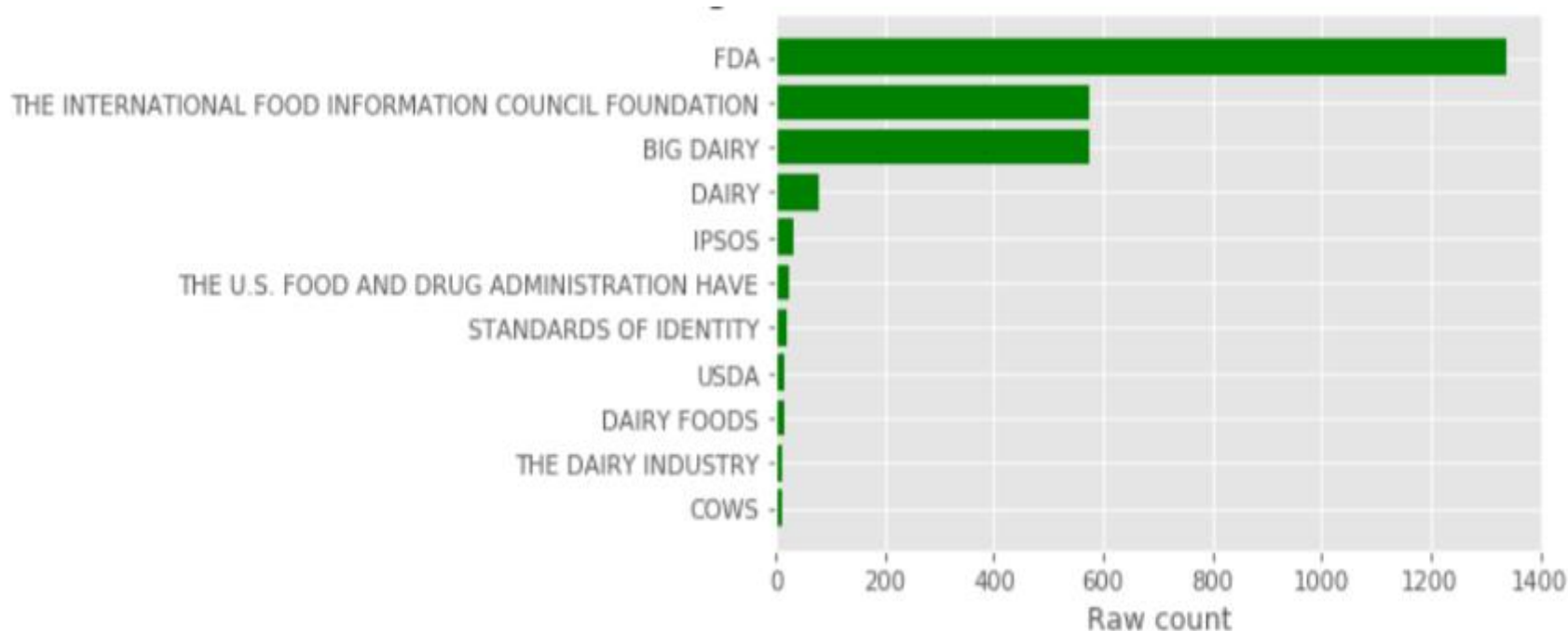
TOP ORGANIZATION ENTITIES IN PUBLIC COMMENTS

"Big Dairy" is mentioned in roughly 600 comments. Comparatively, The Plant Based Foods Association is mentioned in just 100 comments. The International Food Information Council is mentioned in approximated 5% of comments, suggesting that the organizational mouthpiece for food labeling may be having an effect on consumers.

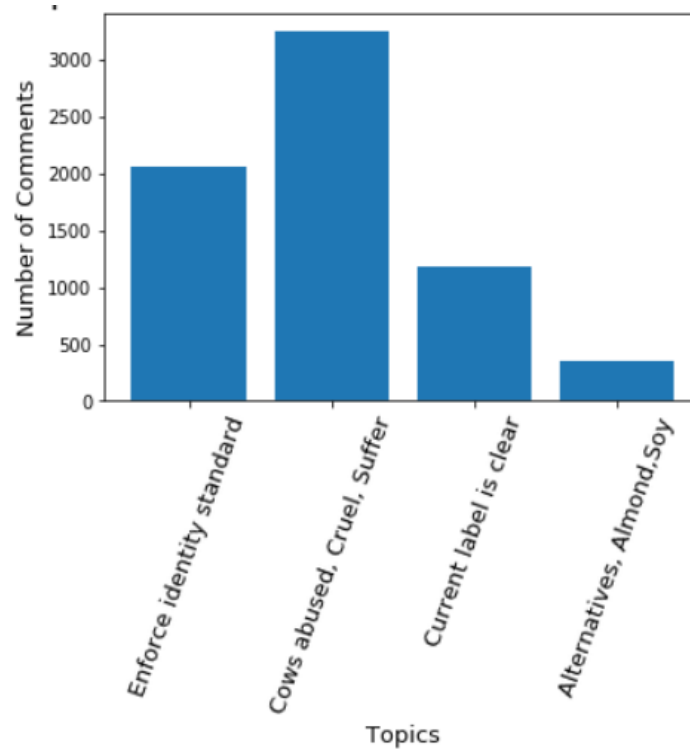# TOP ORGANIZATION ENTITIES IN COMMENTS FROM CONSUMERS

Organization entity analysis shows the predominance of entities like the FDA and references to labelling. However, entity analysis falls short for individual consumers in that packages like SPACY can't differentiate between abbreviations (FDA) and the entity organization's full name.

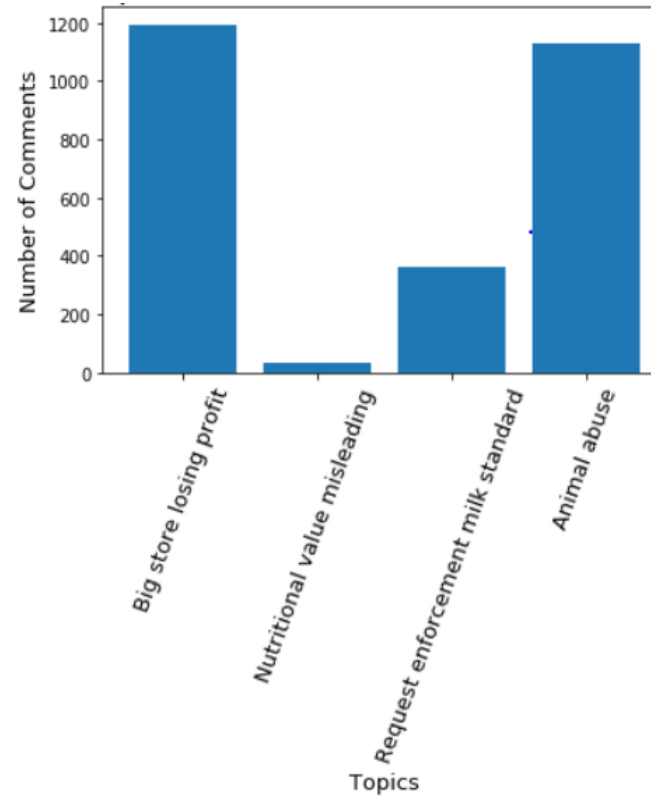# TOP ORGANIZATION ENTITIES IN COMMENTS FROM FOOD INDUSTRY

Again, we see the pattern of the Dairy Industry's market research appearing as top entities (IPSOS and International Food Information Council). The third most common entity in the comments was "Big Dairy"—a term we associate with critics of the dairy industry. This suggests that comments submitted by representatives of the food industry may also include an industry sector that is more aligned with health foods or plant-based alternatives.
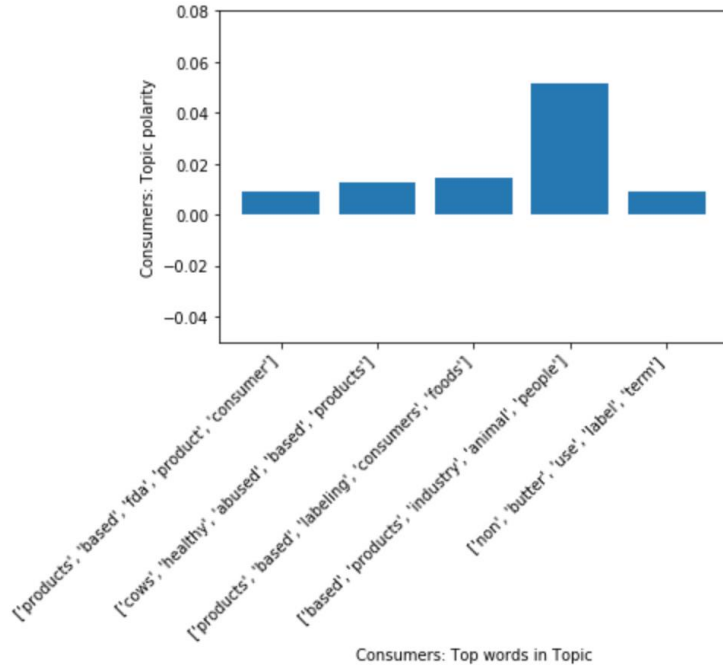
TOPICS FROM INDIVIDUAL CONSUMERS

LDA topic modeling filtered for individual consumers , shows that the top topics for individual consumers center around enforcing the identity standard of milk, animal abuse, support for the current standard of regulations, as well as comments about non-dairy alternatives such as almond milk and soy milk.

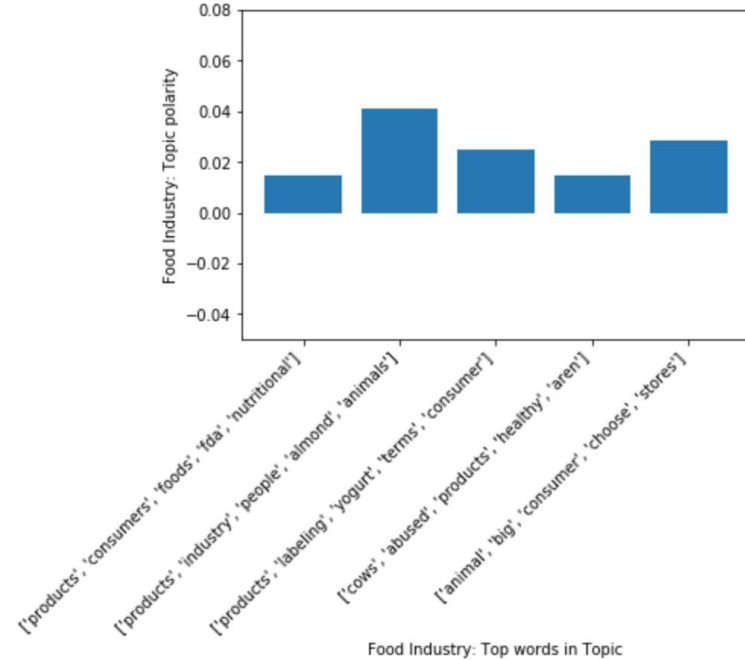# TOPICS FROM THE FOOD INDUSTRY

LDA topic modeling filtered for food industry members shows a bifurcation of support for regulatory changes. We hypothesized that comments from the food industry would be representative of the dairy food industry and their support for an enforcement of identity standards for milk. However, it appears that comments from the food industry might also include the health food industry, given the presence of animal abuse as a topic.

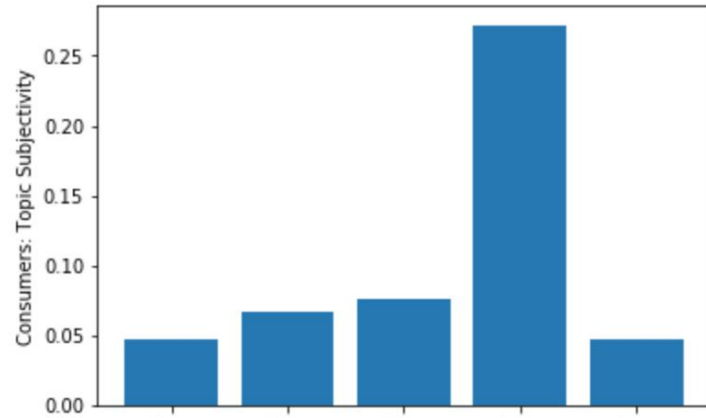# LDA TOPIC –SENTIMENT FINDS MOST COMMENTS AS NEUTRAL
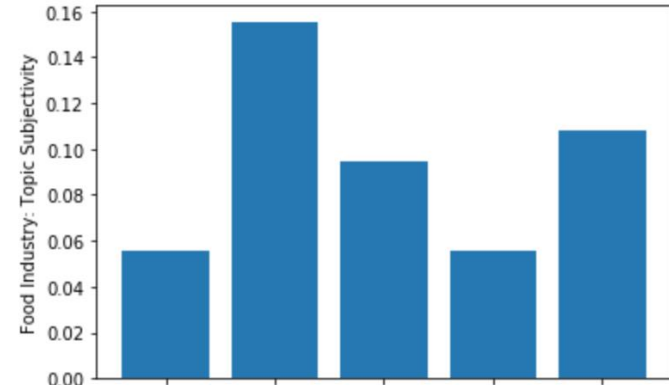

Individual Consumers

Food Industry

The Topic Polarity Score and Topic Subjectivity scores ranges between –1 and +1, with –1 as negative and +1 as positive. While both show neutral comments overall, the highest rated Topic Subjectivity Topic relates to support for plant based milk as a benefit for the environment.

TOPIC SUBJECTIVITY REMAINS NEUTRAL FOR EACH AUDIENCE

Individual Consumers

Food Industry

Position Types for Predictive Modeling

Description of Position Types:

Against: against the labeling changes (or "let plant-based milks continue to be milk")

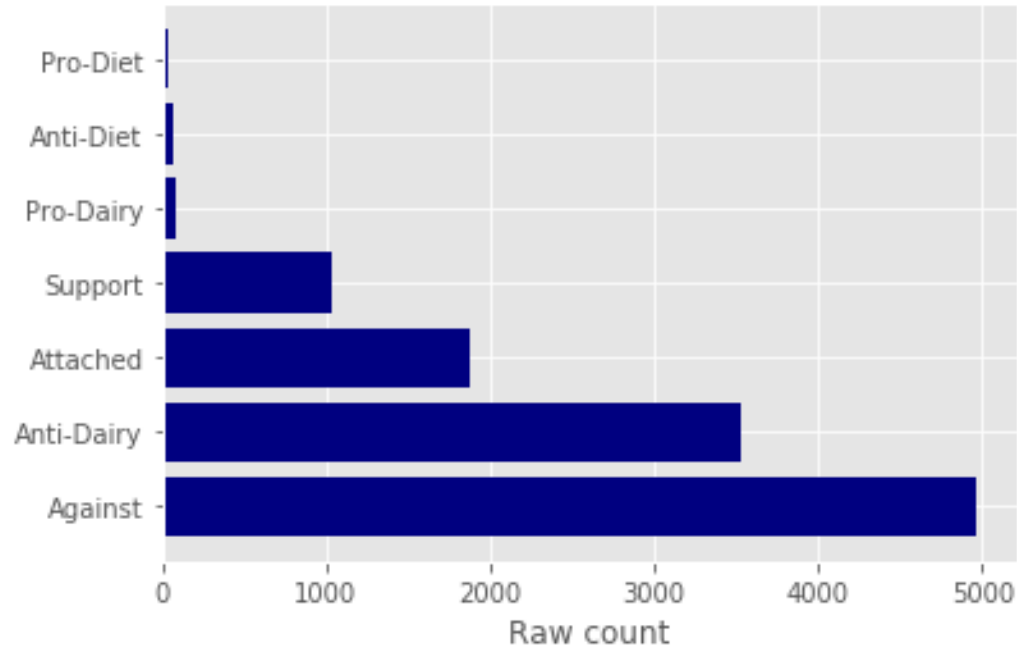**Support**: supporting the labeling distinctions (or "plant milk should be juice")

Anti-Dairy: comments that are primarily about hating the dairy industry ("I want my milk from plants, not abused cows")

Attached: only attachments included, so no text to analyze

Anti-Diet: comments that primarily refer to non-dairy dietary habits (such as "I drink soy milk because I'm allergic")

Pro-Diet: comments that only talk about the benefits of dairy in the diet

PREDICTIVE MODELING OF TEXT COMMENTS

Against : 4969 **Anti-Dairy** : 3535 **Attached** : 1883 **Support** : 1036
Pro-Dairy : 82 **Anti-Diet** : 63 **Pro-Diet** : 37

Conclusion

- Initial findings show opposition to imposing stricter identity standard to plant-based milks

- Individual consumers and some food industry members expressed concern about animal abuse in dairy production

- Submitter Types are not monolithic—our analysis suggests that comments from the Food Industry include health-conscious/plant-based sectors, not just representative from the Dairy Industry as we hypothesized

Policy Recommendations

- Our analysis should be interpreted as informative rather than prescriptive.

- The onus is on the FDA to determine which voices and opinions matter when changing identity standards of food for regulatory purposes.

- The nature of public comments means that they don't map back perfectly to the questions posed by the FDA. Our analysis found a large number of comments critical of the dairy industry as a whole, rather than addressing regulations and nutritional content of plant-based milk.
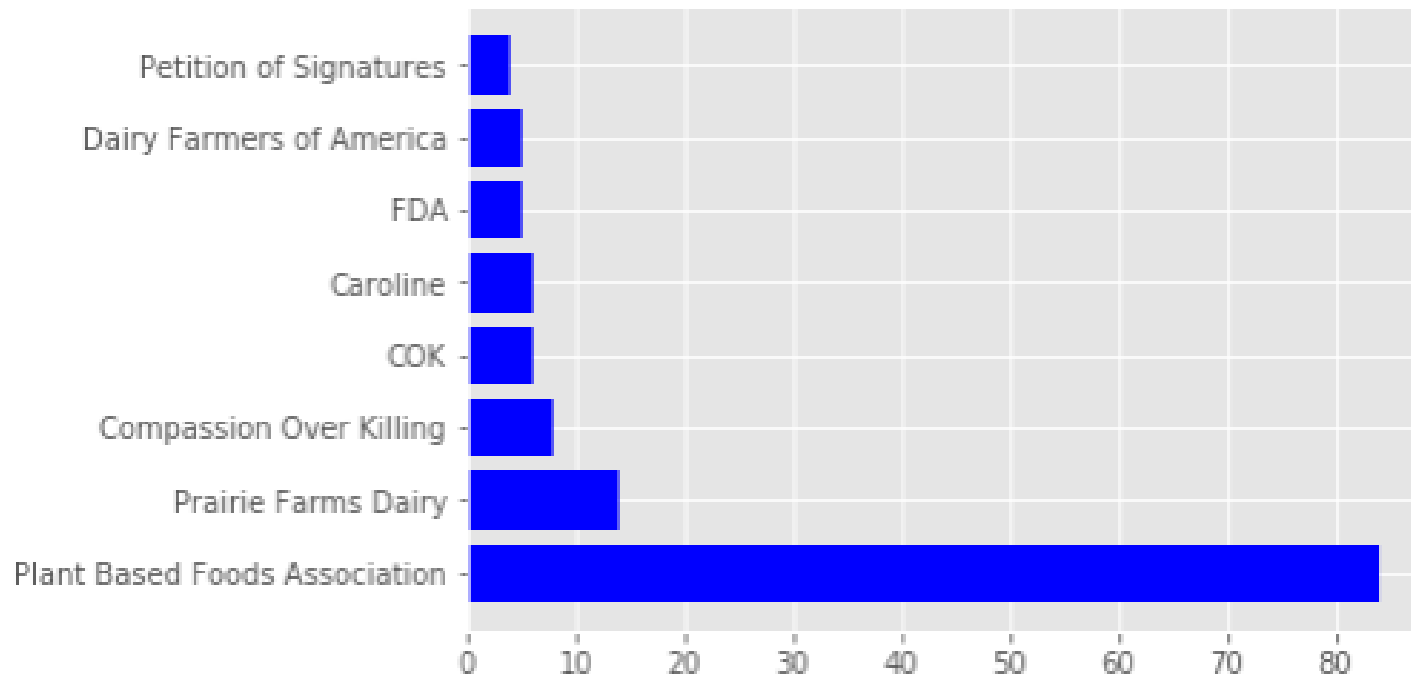
got questions?

Appendix

Federal Register under docket number: FDA-2018-N-3522. The period of public commenting was from 9/28/2018 through 11/27/2018. The comment period was extended through 1/28/2019 resulting in a total of 11,903 plain text comments.

Public comments were extracted for project use on 2/18/2019 using an API from regulations.gov. Comment meta data was extracted using scraper tool on 3/2/2019.
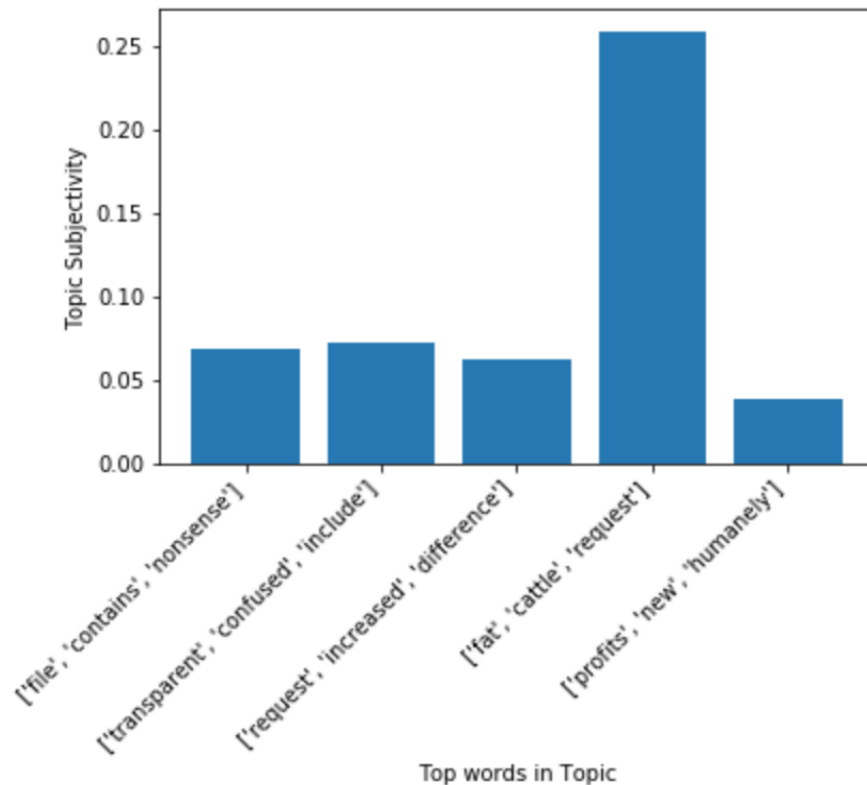
Although comments may be submitted in hard copy and appear in the data set as an attached, our research is limited to the plain text comments submitted electronically. Public comments submitted as mass mailing comments have been curated by the FDA to reflect 1 instance of a text document.
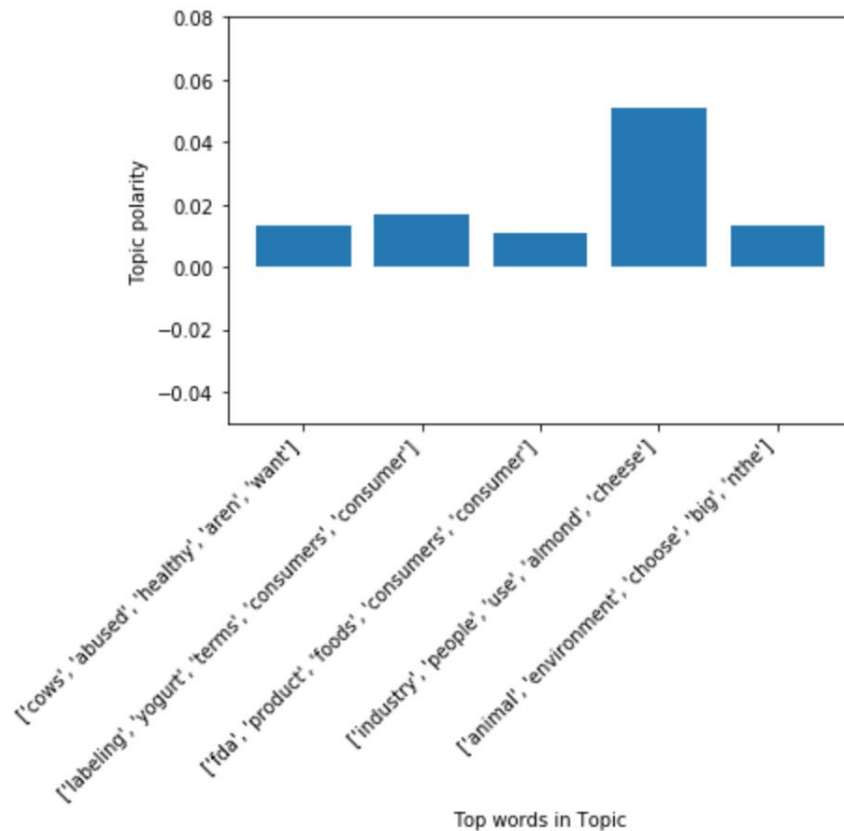
ORGANIZATIONS PRESENT IN COMMENT TITLES

TOPIC SUBJECTIVITY– ALL COMMENTS

Text(0, 0.5, 'Topic Subjectivity')

TOPIC POLARITY– ALL COMMENTS

COMPLETE COMMENT SUBMISSION TYPE

: Text(0.5, 1.0, 'Distribution of Comments by Industry')

Distribution of Comments by Industry