Almond-Survey-2020-

Analysis of Almond Survey results. This repository contains a survey response dataset that was compiled from a survey distributed to almond producers and farm managers throughout California as well as an analysis of the response data.

Summary

The purpose of this repository is to provide a guideline for the Almond Survey Results and Analyses report. This repository contains the raw survey dataset along with processed datasets that were cleaned for specific analyses.

The analysis goal of the Almond Survey Results dataset is to examine the relationships between respondent demographics and cover crop adoption. Planting a cover crop is one of several bee-friendly practices that can be implemented on almond orchards to protect and support managed honey bee health. Evaluating factors that may affect cover crop adoption will provide greater insight into the barriers almond producers and farm managers are facing regarding the implementation of bee-friendly practices on their farms.

The following research questions will inform the analysis goals:

- Where are the respondents' almond orchards located?
- Which demographic factors affect whether or not respondents have planted cover crop in the last 5 years?
- How does region affect whether or not the respondents have planted cover crop in the last 5 years?
- How does respondent role in the almond operation affect whether or not the respondents have planted cover crop in the last 5 years?
- How does respondent age affect whether or not the respondents have planted cover crop in the last 5 years?
- How does the size of the almond operation affect whether or not the respondents have planted cover crop in the last 5 years?

Investigators

Emily McNamara, Duke University's Nicholas School of the Environment Email: emily.mcnamara@duke.edu

Keywords

*Cover crop: Temporary forage planted between tree rows

*Region: The location of the respondant's almond orchard(s)

*Role in Operation: Respondent could select 'Owner, not responsible for day-to-day management', 'Owner/Operator', or 'Farm Manger (not owner)' to describe the role he or she has in the almond operation.

Database Information

This dataset contains data of 301 completed responses from a survey that was distributed to almond producers and farm managers throughout California. The survey was launched on December 10th, 2019 and was closed on February 5th, 2020. Data were collected using Qualtrics.

csv files were saved as 'Almond_Survey_Results_raw.csv', 'Almond_Survey_Data_Project_Processed.csv' , and 'Almond_Survey_Numeric_Answers_Processed.csv'

Folder structure, file formats, and naming conventions

Data/Processed: This folder contains datasets that were processed from the raw dataset as csv files

Data/Raw: This folder contains the raw survey dataset as a csv file

Output: This folder contains output information from the analyses and visualizations produced from specific code files.

Code: This folder contains the code used for the analyses

File Formats

- All data files are in csv format
- Code files are in R Markdown or R Script documents

Naming Conventions

Files are named according to the following naming convention: `databasename_datatype_details_stage.format`, where:

databasename refers to the database from where the data originated

datatype is a description of data

details are additional descriptive details, particularly important for processed data

stagerefers to the stage in data management pipelines (e.g., raw, cleaned, or processed)

format is a non-proprietary file format (e.g., .csv, .txt)

Metadata

Almond Survey Results Dataset

Column	Description
End Date	Date the respondent completed submitted the survey
Role in Operation	Respondent's role in operation ('owner, not responsible for
County	County the almond orchard(s) was located (Counties in California)
Regions	Region in which the county was located (Sacramento Valley, Delta,
Total Yield Bearing Acreage	Total amount of acreage with almonds that are mature enough to
Pollinator Manager	The person in charge of pollination management decisions (Farm

Column	Description
Cover Crop Grown	Whether or not the respondent has grown cover crop in the last 5
Cover Crop Seeds	Description of how the respondent acquired cover crop seed (Private
Cover Crop Satisfaction	Respondent's level of satisfaction with cover crop (Not satisfied,
Cover Crop Interest	Respondent's level of interest in planting cover crop if he/she had
Cover Crop Concerns	Respondent's concerns with plant-ing/maintaining cover crop
Cover Crop Incentives	Possible incentives that may assist respondent in planting cover
Water Source	The water source used to irrigate the respondent's almond
PPH Grown	Whether or not the respondent has permanent pollinator habitat
PPH Satisfaction	Respondent's level of satisfaction with permanent pollinator
PPH Interest	Respondent's level of interest in planting permanent pollinator
PPH Concerns	Respondent's concerns with plant-ing/maintaining permanent
PPH Incentives	Possible incentives that may assist respondent in planting

Column	Description
Pollination	How the respondent pollinated his/her almond orchard in 2019 (Our
Beekeeper Location	Where the bee hives came from if the respondent rented honey bees
Rental Price	Highest rental fee/ per bee hive the respondent paid in 2019 ($)
Age	The age range of the respondent

Scripts and code

The file called Data Wrangling was used to process the data from raw form to processed form.

Quality assurance/quality control

For quality assurance and to ensure relevant responses from the survey, survey respondents had to be an almond producer or farm manager who farms one or more acres of almonds in California to qualify for the survey. Once the survey was closed, the dataset was filtered for 100% completed responses and responses 2.5 minutes and over for quality control purposes. The IP address column was then sorted to identify responses that had duplicate IP address. Responses with duplicate IP addresses were highlighted and then each of the responses were analyzed to determine whether or not the respondent had taken the survey more than once. If it was determined that a respondent had taken the survey more than once, all responses from that respondent were deleted. After the quality assurance/quality control process, there were 301 completed responses in the dataset.