

Analysis of Factors Affecting Cover Crop Adoption on Almond Orchards in California

<https://github.com/emac2020/Almond-Survey-2020->

Emily McNamara

Contents

1	Rationale and Research Questions	5
2	Dataset Information	6
2.1	Almond Survey Results Dataset	6
2.1.1	Data Content Information	6
2.2	Almond Survey Numeric Results Dataset	9
2.2.1	Data Content Information	9
2.3	Naming Conventions and File Formats	12
3	Exploratory Analysis and Wrangling	14
3.1	Data Wrangling: Almond Survey Response Dataset	14
3.2	Data Wrangling: Almond Survey Numeric Response Dataset	14
3.3	Data Exploration	15
4	Analysis	20
4.1	Question 1: THis part of analysis pertains to question 1, etc. (helps organize relevant info for final deliverable)	20
4.2	Question 2:	20
5	Summary and Conclusions	21
6	References	22

List of Tables

List of Figures

1 Rationale and Research Questions

California is the epicenter of global almond production, producing 80% of the world's almond supply. As the forecasted growth of consumer demand shows few signs of subsiding, farmers in California are converting farmland to almond orchards at a considerable rate - from 428,000 acres in 1996 to 1,170,000 acres in 2019. However, the industry is in a precarious position as almonds require 100% pollination from managed honey bees - a species that has witnessed significant decline in population since 2006. The decline in managed honey bees is attributed to poor nutrition, lack of diverse forage, stress from transpiration, and pesticide toxicity.

To mitigate the decline in managed honey bees and protect honey bee health on almond orchards, experts determined several bee-friendly practices that growers can adopt on their farms. One of these practices is planting temporary forage called 'cover crop' between tree rows to increase the diversity and abundance of nutrient and pollen sources for the honey bees.

This study uses data from a survey administered to almond producers and farm managers throughout California to identify potential barriers in adopting cover crop and understand why the practice of planting cover crops is not widely adopted by almond producers.

To meet the research objectives, the primary research questions are:

- 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?

2 Dataset Information

For the analysis, the following dataset was used:

2.1 Almond Survey Results Dataset

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.

The downloaded file was saved in the project folder path `./Data/Raw/Almond_Survey_Results_raw.csv` on 2020-04-02

2.1.1 Data Content Information

The dataset contains 24 columns, which are shown in Table 1.

2.1.1.1 Table 1: Almond Survey Response Dataset Content Information

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

Column	Description
Pollinator Manager	The person in charge of pollination management decisions (Farm
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

Column	Description
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
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

2.2 Almond Survey Numeric Results Dataset

This dataset contains the same responses as the Almond Survey Results Dataset in section 2.1., but the dataset was downloaded from Qualtrics in numerical answer form with split-answer columns.

The downloaded file was saved in the project folder path `./Data/Raw/Almond_Survey_Numeric_Answers_F` on 2020-04-02

2.2.1 Data Content Information

2.2.1.1 Table 1: Almond Survey Numeric Response Dataset Content Information

The dataset contains 48 columns, which are shown in Table 2.

Column	Description
End Date	Date the respondent completed submitted the survey
Role in Operation	Respondent's role in operation ('owner, not
Regions	Region in which the county was located (Sacramento
Tehama	County in California
Butte	County in California
Glenn	County in California
Colusa	County in California
Yuba	County in California
Sutter	County in California
Yolo	County in California
Solano	County in California

Column	Description
San Joaquin	County in California
Stanislaus	County in California
Madera	County in California
Merced	County in California
Fresno	County in California
Kings	County in California
Tulare	County in California
Kern	County in California
Sacramento	County in California
Total Yield Bearing Acreage	Total amount of acreage with almonds that are mature
Cover Crop Grown	Whether or not the respondent has grown cover crop in
Cover Crop Satisfaction	Respondent's level of satisfaction with cover crop
Cover Crop Interest	Respondent's level of interest in planting cover crop
ConcernCC_WaterAvailability	Answer choice for respondent concern for growing cover
ConcernCC_WaterExpense	Answer choice for respondent concern for growing cover

Column	Description
ConcernCC_IrrigationSystem	Answer choice for respondent concern for growing cover
ConcernCC_EffortTime	Answer choice for respondent concern for growing cover
ConcernCC_Labor	Answer choice for respondent concern for growing cover
ConcernCC_EquipmentCost	Answer choice for respondent concern for growing cover
ConcernCC_EquipmentAvailability	Answer choice for respondent concern for growing cover
ConcernCC_SeedCost	Answer choice for respondent concern for growing cover
ConcernCC_SoilType	Answer choice for respondent concern for growing cover
ConcernCC_FrostDamage	Answer choice for respondent concern for growing cover
ConcernCC_SupportPest	Answer choice for respondent concern for growing cover
ConcernCC_CompetingOperations	Answer choice for respondent concern for growing cover

Column	Description
ConcernCC_PhysicalInterference	Answer choice for respondent concern for growing cover
ConcernCC_NoConcern	Answer choice for respondent concern for growing cover
ConcernCC_PreferNotAnswer	Answer choice for respondent concern for growing cover
FutureIncentivesCC_AssociatedNonPollination	Answer choice for incentive to grow cover crop
FutureIncentivesCC_DecreasedRentalFee	Answer choice for incentive to grow cover crop
FutureIncentivesCC_FedCostShare	Answer choice for incentive to grow cover crop
FutureIncentivesCC_PrivateCostShare	Answer choice for incentive to grow cover crop
FutureIncentivesCC_Equipment	Answer choice for incentive to grow cover crop
FutureIncentivesCC_BeeStrength	Answer choice for incentive to grow cover crop
FutureIncentivesCC_None	Answer choice for incentive to grow cover crop
FutureIncentivesCC_PreferNotAnswer	Answer choice for incentive to grow cover crop
Age	Respondent age

2.3 Naming Conventions and File Formats

The files are named according to the following convention: Files are named according to the following naming convention: `dbname_datatype_details_stage.format`, where:

dbname 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

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

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

3 Exploratory Analysis and Wrangling

3.1 Data Wrangling: Almond Survey Response Dataset

The raw ‘Almond Survey Response Dataset’ and the ‘Almond Survey Numeric Response Dataset’ both contained unnecessary information for the overarching goals of this project. Thus, data regarding permanent pollinator habitat, the pollination manager, non-yield bearing acreage, and water sources was removed. The dates the respondents completed the survey were removed as well because the analyses do not involve ‘time’ as a parameter.

```
# Read in raw almond survey response data
almonds.project.raw <- read.csv("./Data/Raw/Almond_Survey_Results_Raw.csv")

# Look at column names
#colnames(almonds.project.raw)

# Select column names that only apply to cover crop analysis

almonds.project.CC.processed <- almonds.project.raw %>%
  dplyr::select(Role.in.Operation:Regions, Total.Yield.Bearing.Acreage, Acre.Ranges, Cov

# Fill all empty cells in almonds.project.CC.processed with 'NA' and name 'almonds.CC'
almonds.CC <- almonds.project.CC.processed %>% mutate_all(na_if, "")

# Save new processed dataset
#write.csv(almonds.CC, row.names = FALSE, file = #"./Data/Processed/Almond_Project_Cov
```

3.2 Data Wrangling: Almond Survey Numeric Response Dataset

```
# Read in almond survey numeric response data
almonds.numeric.raw <- read.csv("./Data/Raw/Almond_Survey_Numeric_Answers_Raw.csv")

# Look at column names
#colnames(almonds.numeric.raw)

# Choose columns for cover crop analyses that require numeric data

almonds.numeric <- almonds.numeric.raw %>%
  dplyr::select(Role.in.Operation, Regions, Total.Yield.Bearing, Cover.Crop.Grown, Age)

# Save new processed dataset
#write.csv(almonds.numeric, row.names = FALSE, file = #"./Data/Processed/Almond_Survey
```

The raw datasets now only contain data regarding respondents responses to cover crop adoption and interest as well as the respondents' answers to the demographic and pollination questions in the survey.

3.3 Data Exploration

After wrangling the data into a format that would allow me to answer my research questions more effectively, I was able to explore the processed datasets.

```
# Column names of both datasets
colnames(almonds.CC)

colnames(almonds.numeric)

# Structure of datasets

str(almonds.CC)

str(almonds.numeric)

# Class

class(almonds.CC$Total.Yield.Bearing.Acreage)
class(almonds.CC$Regions)

class(almonds.numeric$Total.Yield.Bearing)
class(almonds.numeric$Role.in.Operation)

# Dimensions of datasets

dim(almonds.CC)

dim(almonds.numeric)

# Head

head(almonds.CC)

head(almonds.numeric)

# Summary

summary(almonds.CC)
summary(almonds.CC$Regions)
summary(almonds.CC$Role.in.Operation)
```

```
summary(almonds.CC$Age)
summary(almonds.CC$Total.Yield.Bearing.Acreage)
summary(almonds.CC$Acre.Ranges)
summary(almonds.CC$Rental.Price)

summary(almonds.numeric)
summary(almonds.numeric$Regions)
summary(almonds.numeric$Total.Yield.Bearing)
```

Figure 1 plots the locations of the respondent’s almonds orchards by county in California. Because respondents were allowed to “check all that apply,” when selecting the counties where their almond orchards are located, some of the location responses in the dataset had several counties listed in one cell. To remedy this, I first created a new column in the dataset Excel spreadsheet called “County.Multiple,” then I copied the data from the “County” column into the new column and entered in “multiple” for location responses that contained more than one county. Thus, in **Figure 1**, “multiple” is used to describe the respondents who had almond orchards in several counties. This figure shows that a majority of the respondents farm almonds in Stanislaus, Madera, and Fresno. Since a majority of almond producers are from Southern California, I anticipated a higher response rate from these counties.

I then went through the spreadsheet again and noted that in the location responses that had multiple counties listed, the counties were close to one another and therefore, the counties could be categorized by Central Valley watershed basins. I chose to categorize the counties by watersheds provided the fact that the production of tree nuts requires a considerable amount of water, and due to the state’s water scarcity, water is a critical factor in determining managerial practices for farmers. Thus, I hypothesized that watershed basins are most representative of respondent behavior toward bee-friendly practices (i.e. cover crop).

The regional categories include Sacramento Valley, Delta, San Joaquin Basin, and Tulare Basin. The Sacramento Valley consists of respondents who farm almonds in Butte, Colusa, Glenn, or Tehama; the Delta region includes Sacramento, Solano, Yolo, and Yuba; the San Joaquin Basin includes San Joaquin, Stanislaus, Merced, and Madera; the Tulare Basin includes Tulare, Kings, Kern, and Fresno.

Figure 2 plots the location of the respondents’ almond orchards by region. We can see from this plot that a majority of the respondents farm almonds in the San Joaquin and Tulare Basins. It is important to note that the northern section of the Central Valley consists of the Sacramento Valley and Delta regions, while the southern section includes the San Joaquin and Tulare Basins. Therefore, a majority of the respondents farm almonds in the southern section of the Central Valley.

Figure 1: Location of Respondents' Almond Orchards by County in California

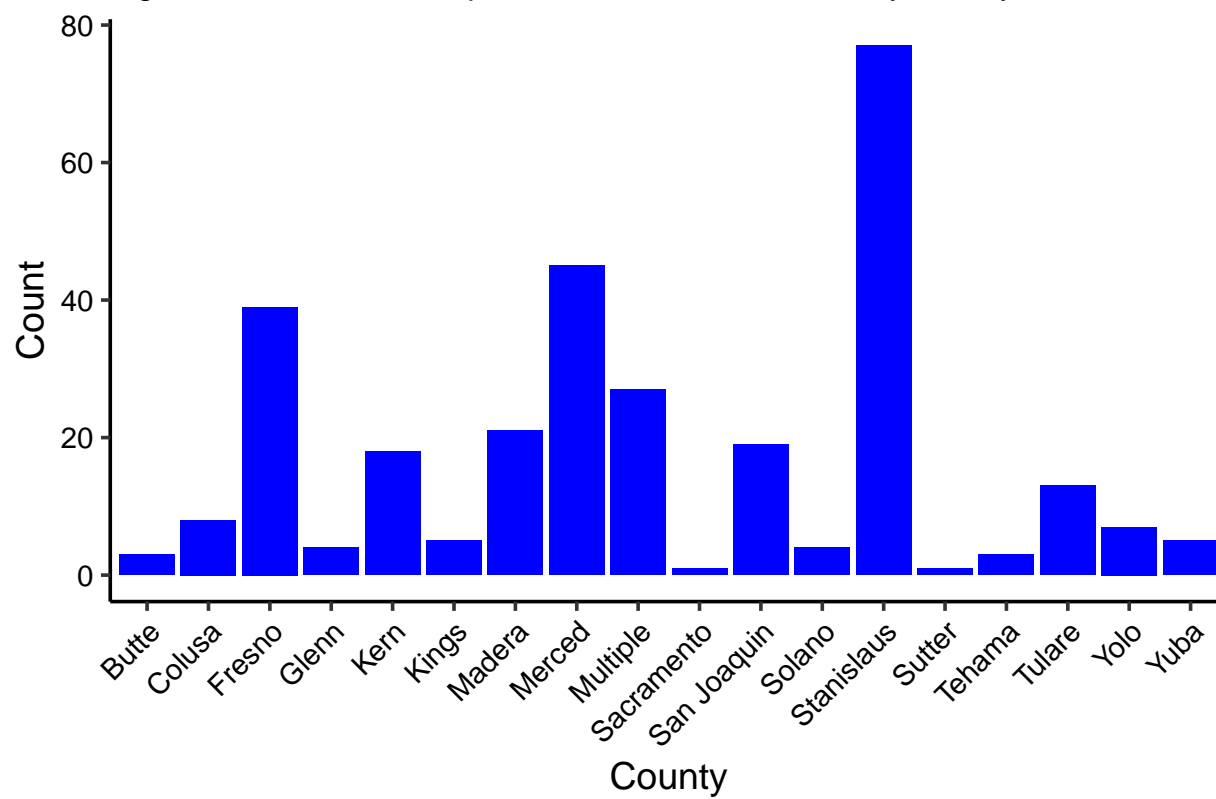


Figure 2: Location of Respondents' Almond Orchards by Region in California

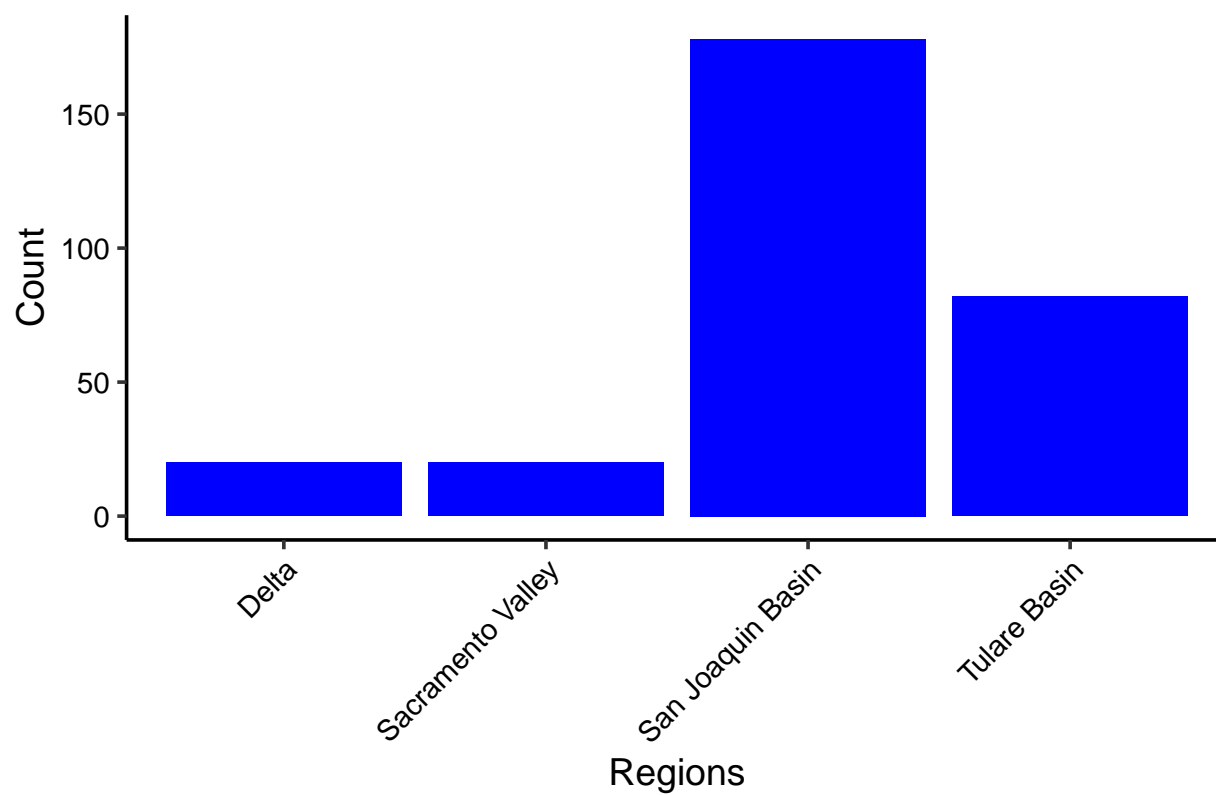
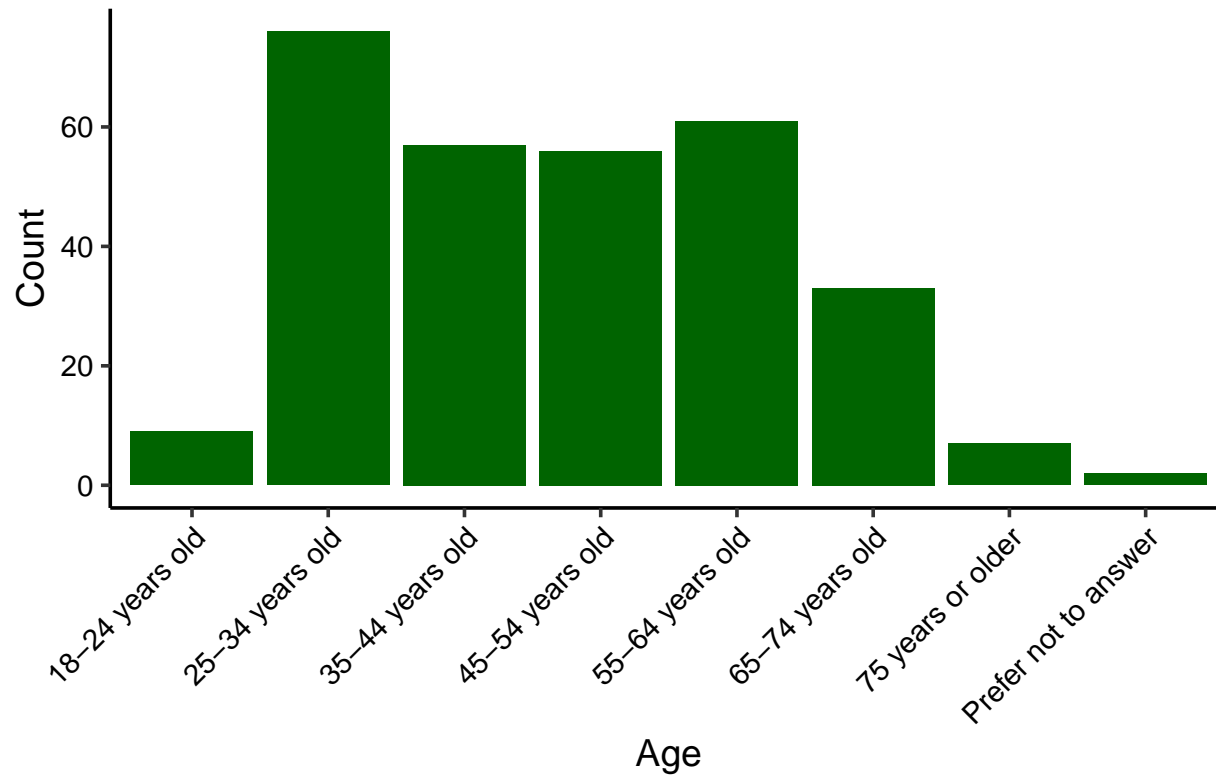


Figure 3: Age Ranges of Respondents



4 Analysis

4.1 Question 1: This part of analysis pertains to question 1, etc.
(helps organize relevant info for final deliverable)

4.2 Question 2:

5 Summary and Conclusions

6 References

- United States Department of Agriculture (USDA). (2019). 2018 California Almond Acreage Report. California Department of Food and Agriculture. Retrieved from https://www.nass.usda.gov/Statistics_by_State/California/Publications/Specialty_and_Other_Releases/Almond/Acreage/201904almac.pdf