Strawberries in the United States: Exploratory Data Analysis

GRS 615: Data Science in R

Nicole Kingdon

October 18, 2023

Introduction

Strawberries & Positive Health

Strawberries are a fruit that holds several minerals, vitamins, and phytonutrients (CRAIG 1997), which have positive implications on human health (Afrin et al. 2016). Specifically, strawberries have been found to help reduce likelihood of cancer, diabetes, obesity, neurodegeneration, cardiovascular disease, and metabolic syndrome (see Figure 1) (Afrin et al. 2016). Although strawberries as a healthy food is the norm, pesticides appear to be harming the beneficial factors of this fruit.

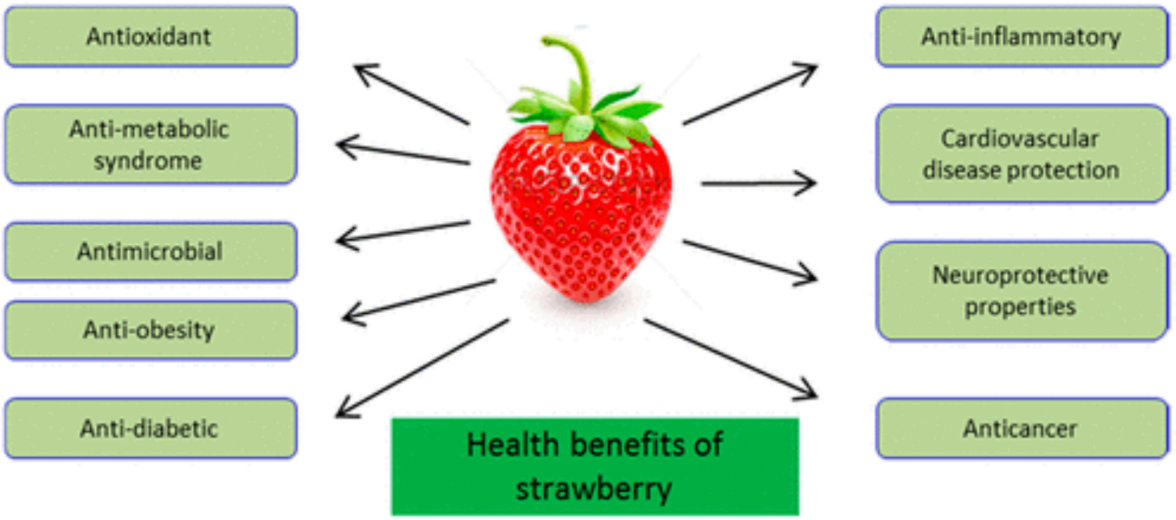

Figure 1: Health benefits of strawberries (Afrin et al., 2016)

Strawberries & Pesticides

Pesticides are used on fruit and vegetable crops, including strawberries, with hopes to increase the quantity (Fenik, Tankiewicz, and Biziuk 2011). Pesticides are made of chemical compounds to reduce or completely eliminate pests from impacting crops (Afrin et al. 2016). These chemical compounds may increase the yield of the crop, but may have a large risk on human health. Additionally, they may contaminate bodies of water and soil with the chemicals, help pests develop resistance to the chemicals, and impact helpful organisms from persisting in areas where pesticides are used. Overall, there are positive and negative impacts of using pesticides (see Figure 2), but it is important to further examine these impacts, specifically on strawberries, to understand the implications of using such.

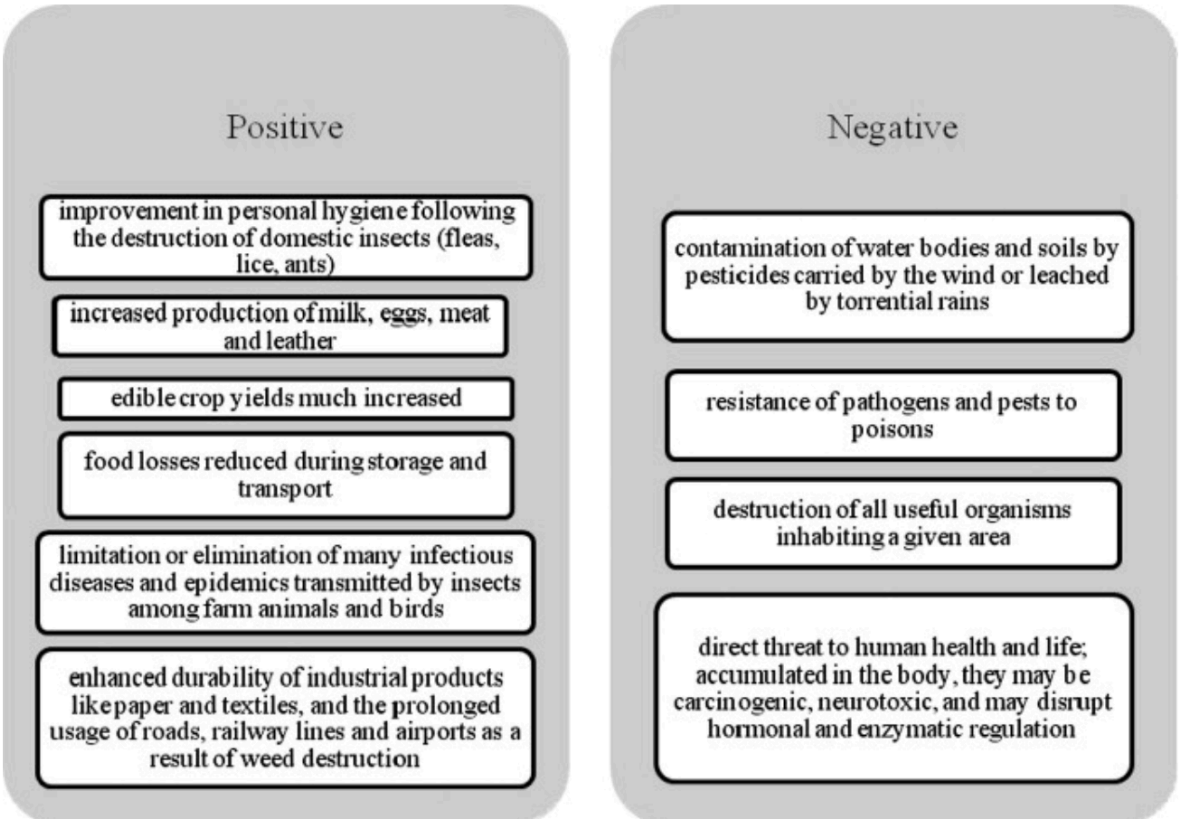

Figure 2: Positive and negative implications of pesticides on fruit and vegetable crops (Fenik et al., 2011)

Organic Strawberries

Strawberries that use alternative and safer methods to production are considered organic (Verteramo Chiu and Gomez 2023). Along with these alternative methods comes a higher cost

of production and care, which increases the price to the consumer.

Analyzing Strawberry Production

The literature varies on if organic or non-organic is environmentally better (Afrin et al. 2016). This exploratory data analysis on production of strawberries, which encompasses both processed (non-organic) and fresh (organic) market data, will help us better understand production of strawberries in the United States.

Data Acquisition & Assessment

USDA-NASS Data

The data was acquired from U.S. Department of Agriculture (USDA) and the National Agricultural Statistics Service (NASS). The data was uploaded for data cleaning and organizing and exploratory data analysis by Professor Haviland Wright, who chose the following data: USDA-NASS.

The data frame uploaded to R is titled `strawberry` (see below).

```
strawberry <- read_csv("strawberry.csv", col_names = TRUE)
```

```
Rows: 4,314
Columns: 21
$ Program            <chr> "CENSUS", "CENSUS", "CENSUS", "CENSUS", "CENSUS", "~
$ Year               <dbl> 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 202~
$ Period             <chr> "YEAR", "YEAR", "YEAR", "YEAR", "YEAR", "YEAR", "YE~
$ `Week Ending`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Geo Level`        <chr> "STATE", "STATE", "STATE", "STATE", "STATE", "STATE~
$ State              <chr> "ALASKA", "ALASKA", "ALASKA", "ALASKA", "ALASKA", "~
$ `State ANSI`       <chr> "02", "02", "02", "02", "02", "02", "02", "06", "06~
$ `Ag District`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Ag District Code` <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ County             <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `County ANSI`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Zip Code`         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Region             <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ watershed_code     <chr> "00000000", "00000000", "00000000", "00000000", "00~
$ Watershed          <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Commodity          <chr> "STRAWBERRIES", "STRAWBERRIES", "STRAWBERRIES", "ST~
$ `Data Item`        <chr> "STRAWBERRIES, ORGANIC - OPERATIONS WITH SALES", "S~
```

```
$ Domain              <chr> "ORGANIC STATUS", "ORGANIC STATUS", "ORGANIC STATUS~
$ `Domain Category`   <chr> "ORGANIC STATUS: (NOP USDA CERTIFIED)", "ORGANIC ST~
$ Value               <chr> "2", "(D)", "(D)", "(D)", "2", "(D)", "(D)", "142",~
$ `CV (%)`            <chr> "(H)", "(D)", "(D)", "(D)", "(H)", "(D)", "(D)", "1~
```

Census Data

The data offers census data based on state that represents fresh market (organic) and process market (non-organic) sales.

```
Rows: 864
Columns: 21
$ Program             <chr> "CENSUS", "CENSUS", "CENSUS", "CENSUS", "CENSUS", "~
$ Year                <dbl> 2021, 2021, 2021, 2021, 2021, 2021, 2021, 2021, 202~
$ Period              <chr> "YEAR", "YEAR", "YEAR", "YEAR", "YEAR", "YEAR", "YE~
$ `Week Ending`       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Geo Level`         <chr> "STATE", "STATE", "STATE", "STATE", "STATE", "STATE~
$ State               <chr> "ALASKA", "ALASKA", "ALASKA", "ALASKA", "ALASKA", "~
$ `State ANSI`        <chr> "02", "02", "02", "02", "02", "02", "02", "06", "06~
$ `Ag District`       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Ag District Code`  <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ County              <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `County ANSI`       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Zip Code`          <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Region              <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ watershed_code      <chr> "00000000", "00000000", "00000000", "00000000", "00~
$ Watershed           <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Commodity           <chr> "STRAWBERRIES", "STRAWBERRIES", "STRAWBERRIES", "ST~
$ `Data Item`         <chr> "STRAWBERRIES, ORGANIC - OPERATIONS WITH SALES", "S~
$ Domain              <chr> "ORGANIC STATUS", "ORGANIC STATUS", "ORGANIC STATUS~
$ `Domain Category`   <chr> "ORGANIC STATUS: (NOP USDA CERTIFIED)", "ORGANIC ST~
$ Value               <chr> "2", "(D)", "(D)", "(D)", "2", "(D)", "(D)", "142",~
$ `CV (%)`            <chr> "(H)", "(D)", "(D)", "(D)", "(H)", "(D)", "(D)", "1~
```

Survey Data

Additionally, it holds survey information for each state, specifically indicating pesticides and bacterium used to preserve strawberry crop yield. In addition, it offers fresh and process market data.

```
Rows: 3,450
Columns: 21
$ Program            <chr> "SURVEY", "SURVEY", "SURVEY", "SURVEY", "SURVEY", "~
$ Year               <dbl> 2022, 2022, 2022, 2022, 2022, 2022, 2022, 2022, 202~
$ Period             <chr> "MARKETING YEAR", "MARKETING YEAR", "MARKETING YEAR~
$ `Week Ending`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Geo Level`        <chr> "STATE", "STATE", "STATE", "STATE", "STATE", "STATE~
$ State              <chr> "CALIFORNIA", "CALIFORNIA", "CALIFORNIA", "FLORIDA"~
$ `State ANSI`       <chr> "06", "06", "06", "12", "12", "12", NA, NA, NA, "06~
$ `Ag District`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Ag District Code` <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ County             <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `County ANSI`      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ `Zip Code`         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Region             <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ watershed_code     <chr> "00000000", "00000000", "00000000", "00000000", "00~
$ Watershed          <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
$ Commodity          <chr> "STRAWBERRIES", "STRAWBERRIES", "STRAWBERRIES", "ST~
$ `Data Item`        <chr> "STRAWBERRIES - PRICE RECEIVED, MEASURED IN $ / CWT~
$ Domain             <chr> "TOTAL", "TOTAL", "TOTAL", "TOTAL", "TOTAL", "TOTAL~
$ `Domain Category`  <chr> "NOT SPECIFIED", "NOT SPECIFIED", "NOT SPECIFIED", ~
$ Value              <chr> "108", "(D)", "(D)", "169", "(D)", "(D)", "0", "135~
$ `CV (%)`           <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,~
```

States

There were 47 states (c("ALASKA", "CALIFORNIA", "CONNECTICUT", "FLORIDA", "GEORGIA", "IDAHO", "ILLINOIS", "INDIANA", "IOWA", "KENTUCKY", "LOUISIANA", "MAINE", "MARYLAND", "MASSACHUSETTS", "MICHIGAN", "MINNESOTA", "MONTANA", "NEBRASKA", "NEW HAMPSHIRE", "NEW JERSEY", "NEW YORK", "NORTH CAROLINA", "OHIO", "OKLAHOMA", "OREGON", "PENNSYLVANIA", "RHODE ISLAND", "SOUTH CAROLINA", "SOUTH DAKOTA", "TENNESSEE", "VERMONT", "WASHINGTON", "WEST VIRGINIA", "WISCONSIN", "ALABAMA", "ARIZONA", "COLORADO", "KANSAS", "MISSOURI", "NEVADA", "NEW MEXICO", "VIRGINIA", "ARKANSAS", "NORTH DAKOTA", "TEXAS", "UTAH", "OTHER STATES")) with two states considered as "other states".

Years

The data was from the years c(2021, 2019, 2016, 2022, 2020, 2018, 2017).

Assumptions & Motivations

Census Data

The `census` data was a nation-wide collection of data about the fresh and process markets related to strawberries. This data has values that are indicated as `(D)`, which are data that was withheld upon request by the strawberry market in that particular state. This could leave out important information in the data.

Survey Data

The `survey` data was collected via a survey sent out to each state in the United States. There were only 11 out of 47 states who returned the survey (c("CALIFORNIA", "FLORIDA", "OTHER STATES", "NEW YORK", "NORTH CAROLINA", "OREGON", "WASHINGTON", "MICHIGAN", "OHIO", "PENNSYLVANIA", "WISCONSIN")), which includes the "other states". The "other states" did not have any data relating to pesticides and bacterium. This is only a 23% response rate, which is not comprehensive of all the states and the entire United States process market. The states that did return the survey will still be able to show a report of pesticide and bacterium usage on their processed strawberry crops.

Data Cleaning & Organizing

R Packages

The following R packages were used to clean and organize the data:

```
library(knitr)
library(kableExtra)
library(tidyverse)
library(stringr)
library(dplyr)
```

Organization

The data was organized into two data frames: `census` and `survey`. The `census` data frame was cleaned and organized to show fresh and process market sales, and the `survey` data frame was prepared to show pesticide and bacterium data.

Cleaning

Initial Cleaning

The following initial data cleaning derived from Professor Wright.

Removed columns with a single value in all columns

```
## define function
drop_one_value_col <- function(df){
col_name <- NULL
col_val <- NULL
suppressWarnings({
for(i in 1:dim(df)[2]){
if((df |> distinct(df[,i]) |> count()) == 1){
  col_name = c(col_name, colnames(df[i]))
  col_val = c(col_val, df[1,i])
} }
})

if(is.null(col_name)){return("No Columns to drop")}else{
   col_val = unlist(col_val)
   attributes(col_val) = NULL
   drp = data.frame(col_name, col_val)
   return(drp)
   }
}

str <- drop_one_value_col(strawberry)

# str |> kable(caption = "Dropped Single-Value Columns: names and values")

str <- str$col_name

strawberry <- strawberry |> select(!all_of(str))
```

Is every line associated with a state?

```
## state_all contains the number of rows containing data
## for each of the 47 strawberry-growing states.
state_all <- strawberry |> group_by(State) |> count()

## test if every row is associated with a state by summing the
```

```
## counts and testing for equality with the total rows in the
## data frame

if(sum(state_all$n) == dim(strawberry)[1]){print("Every row has value in the State column.
```

```
[1] "Every row has value in the State column."
```

The data is organized by state. The state with the most rows is CALIFORNIA.

Examine the California data

```
## filter rows of California data from the CENSUS data
calif_census <- strawberry |> filter((State=="CALIFORNIA") & (Program=="CENSUS"))


## ## filter rows of California data from the SURVEY data
calif_survey <- strawberry |> filter((State=="CALIFORNIA") & (Program=="SURVEY"))

census_col <- colnames(calif_census)

survey_col <- colnames(calif_survey)
```

List of the composite columns

Census: Data Item, Domain Category

Survey: Data Item, Domain, Domain Category

Separating Data Frames

The following separation of data frames derived from Professor Wright.

The two new data frames are as follows: `strwb_census`, which holds all the `CENSUS` rows, and `strwb_survey`, which holds all the `SURVEY` rows.

Census

After splitting `CENSUS` and `SURVEY` rows into two data frames, Professor Wright has first organized the CENSUS data.

Separated composite columns and cleaned the Value column.

Composite columns in the `strwb_census`: Data Item, Domain category

Column separators in CENSUS: ",", "-", ":"

Separated `Data Item` into columns by ",".

```
## This will be done in stages --

#####################################################
## split `Data Item` into "Fruit", "temp1","temp2","temp3"
## then test the columns created for numer of distinct values
## split the columns until you have columns of
## subjects, properties, values, and metrics (where metrics
## are the units defined for the values)

## In this case, the subject is State/Strawberries --
## strawberries grown reported by state.

## When using separate_wider_delim() when you don't know the
## number of columns the function will return,
## use the "too_many" and "too_few" parameters to set up
## the function.  Generally, setting both parameters
## to "error" will produce helpful error messages.

  strwb_census <- strwb_census |>
  separate_wider_delim(  cols = `Data Item`,
                         delim = ",",
                         names = c("Fruit",
                                 "temp1",
                                 "temp2",
                                 "temp3"),
                         too_many = "error",
                         too_few = "align_start"
                       )

## Test the columns for the number of distinct values.
## for example:
##
# a <- strwb_census |> distinct(Fruit)
## The Fruit column only has one value: STRAWBERRIES the
## subject under investigation.
##
## Remember - the value in single-value columns
## are often needed for Labels on tables and plots.
```

```
##
## Testing the temp1 column guides the next step.
# a <- strwb_census |> distinct(temp1)
## The "temp1" column has 4 distinct values
##
##    " ORGANIC - OPERATIONS WITH SALES"
##    " ORGANIC - PRODUCTION"
##    " ORGANIC - SALES"
##    " ORGANIC"
##
##  (Note the leading space in each string --
##       which is fixed below.)
##
##  You can see that this column needs to be split between
##  "organic" and the properties "OPERATIONS WITH SALES",
##  "PRODUCTION" and "SALES",
##    using " - " as the column delimiter.
##
##  The column "prop_acct" contains the properties,
##   which are are accounting metrics related to
##   strawberry growing operations.


################################################
## split temp1 into crop_type, Prop_acct

strwb_census <- strwb_census |>
  separate_wider_delim(  cols = temp1,
                         delim = " - ",
                         names = c("crop_type",
                                   "prop_acct"),
                         too_many = "error",
                         too_few = "align_start"
                       )

## Once again, test the columns to plan your next step.
##
# a <- strwb_census |> distinct(crop_type)
## Column "crop_type' has single value  "organic"

# a <- strwb_census |> distinct(prop_acct)
```

```
##
## The stringss in the "prop_acct" column are row labels
## for values reported in the "Values" column.

##    "OPERATIONS WITH SALES"
##    "PRODUCTION"
##    "SALES"
##    "NA"

## Note that the NA is in a row where the value
## is labeled in another column.
##

###############################################
## trim the strings
## you can see which columns contain string values that need
## to have leading or trailing spaces that need to be trimmed.


# glimpse(strwb_census)

strwb_census$crop_type <- str_trim(strwb_census$crop_type, side = "both")

strwb_census$temp2 <- str_trim(strwb_census$temp2, side = "both")

strwb_census$temp3 <- str_trim(strwb_census$temp3, side = "both")



###############################################
## split temp2 into market_type, measure

##
## The temp2 column requires a different logic.
##

## start by looking at the unique entries in the temp2 column.

# a <- strwb_census |> distinct(temp2)
#
# temp2
# 1  NA
```

```
# 2 " MEASURED IN CWT"
# 3 " MEASURED IN $"
# 4 " FRESH MARKET - OPERATIONS WITH SALES"
# 5 " FRESH MARKET - SALES"
# 6 " PROCESSING - OPERATIONS WITH SALES"
# 7 " PROCESSING - SALES"

## temp2 contains data for three separate columns
##
##   All Strawberries  (is this a Total?)
##   Fresh Market
##   Processing
##
##  To understand these labels see
##     "Strawberries: An Economic Assessment of the Feasibility
##      of Providing Multiple-Peril Crop Insurance",
##        prepared by Economic Research Service, USDA
##             for the Federal Crop Insurance Corporation
##                  October 31, 1994
##
```

```
## make a copy of the temp2 column named `Fresh Market`.
strwb_census <- strwb_census |> mutate(`Fresh Market` = temp2, .after = temp2)

## Remove cells in `Fresh Market` column
##   that begin "MEASURED"
strwb_census$`Fresh Market` <- strwb_census$`Fresh Market` |> str_replace( "^MEA.*", "")

## Remove cells in `Fresh Market` column
##   that begin "PROCESSING"
strwb_census$`Fresh Market` <- strwb_census$`Fresh Market` |> str_replace( "^P.*", "")

## substitute a space for NA in `Fresh Market` column
strwb_census$`Fresh Market`[is.na(strwb_census$`Fresh Market`)] <- ""

## in temp2 column, remove cells that begin "FRESH"
 strwb_census$temp2 <- strwb_census$temp2 |> str_replace("^F.*", "")

## Now fix the entries in the `Fresh Market` column
##   Remove "FRESH MARKET - " from the cells
strwb_census$`Fresh Market` <- strwb_census$`Fresh Market` |> str_replace("^FRESH MARKET -
```

Created a “Process Market” column.

```
## Make a copy of temp2 named `Process Market`
strwb_census <- strwb_census |> mutate(`Process Market` = temp2, .after = temp2)

## remove `Process Market` cells beginning "MEASURED"
strwb_census$`Process Market` <-  strwb_census$`Process Market` |> str_replace("^MEA.*", "

## substitute space for NA in `Process Market` column
strwb_census$`Process Market`[is.na(strwb_census$`Process Market`)] <- ""

## In temp2, remove cells that begin "PROCESSING"
strwb_census$temp2 <- strwb_census$temp2 |> str_replace("^P.*", "")

## In `Processing Market`, remove "PROCESSING - " from cells
strwb_census$`Process Market` <-  strwb_census$`Process Market` |> str_replace("PROCESSING
```

Removed NA’s from prop_acct, temp2, and temp3.

```
## substitute a space for NA in prop_acct column
strwb_census$prop_acct[is.na(strwb_census$prop_acct)] <- ""

## substitute a space for NA in temp2 column
strwb_census$temp2[is.na(strwb_census$temp2)] <- ""

## substitute a space for NA in temp2 column
strwb_census$temp3[is.na(strwb_census$temp3)] <- ""
```

Combined temp2 with temp3 to create Metric column. Removed parts of string that did not matter. Relocated columns.

```
strwb_census <- strwb_census |> unite(temp2, temp3, col="Metric", sep="")

## Now fix the entries in the Metric column
##   Remove "MEASURED IN " from the cells
strwb_census$Metric <- strwb_census$Metric |> str_replace("MEASURED IN ", "")

## move Metric to the end
strwb_census <- strwb_census |> relocate(Metric, .before = Domain)

strwb_census <- strwb_census |> relocate(`Process Market`, .before = Metric)
```

```r
strwb_census <- strwb_census |> rename(Totals = prop_acct)

#drop_one_value_col(strwb_census)
```

The Value column was transformed.

```r
## remove commas from numbers
## fix footnotes

## basic tools

## start by getting the Values column so you can work on it

vals <- strwb_census$Value


c <- vals |> str_replace_all(",", "")
# vals[1:20]
# c[1:20]

## Now notice what happens when the
## the strings of digits are cast to numerics.

## for example
c <- as.numeric(c)
# c[1:20]


### remove commas from Value entries
dcomma <- function(c){
  x_new <- as.numeric(gsub(",", "", c))
  return(x_new)
}



########################################  footnotes

## finds single uppor case Character in parens in s2
## e.g. "(D)"
```

```r
## To fine the location and value of the footnotes

v <- strwb_census$Value


## find the footnote locations
## fn_i: locations
fn_i <- v |> str_detect("^\\([:upper:]\\)$") ## returns


## dcomma returns numbers and NA's
v1 <- dcomma(v)

## locations of NA's
na_i <- is.na(v1)

## Demonstration that the locations of the footnotes
## are the same as the locations of the NA's

# length(v) == sum(na_i == fn_i)

## update dcomma()
## Integrate transformation of the values column and
## reporting the footnote values.


dcomma <- function(c){
  suppressWarnings({
  xnew = as.numeric(gsub(",", "", c))
  fns = unique(c[is.na(xnew)])
  vtran = list("new_vec" = xnew, "footnotes" = fns)
  return(vtran)
  })
}


v_trns <- dcomma(v)


 a <- v_trns$new_vec
 # a[1:20]
```

```
# v_trns$footnotes
```

I finished cleaning and organizing the `strwb_census` data frame, which is detailed below.

First, I selected particular columns that had necessary data.

```
strwb_census <- strwb_census |>
  select(1:2, 4, 6:11, 14:15)
```

Next, I removed the "," from the `Value` column and transformed them into numeric values. This introduced rows with `NA` values.

```
strwb_census$Value <- as.numeric(str_replace_all(strwb_census$Value,pattern = ",", replace
```

After that, I cleaned up the `CV (%)` column by changing the values to numbers, instead of strings. This also introduced rows with `NA` values.

```
strwb_census$`CV (%)` <- as.numeric(strwb_census$`CV (%)`)
```

Futhermore, I am going to omit all rows in the `Value` and `CV (%)` columns with `NA` values. These are being omitted because they hold no meaning. Only the `Value` column had to be adjusted, as the `CV (%)` column did not have any values if the `Value` column also did not.

```
strwb_census <- na.omit(strwb_census[strwb_census$Value, ])
```

Finally, to complete the `strwb_census` cleaning and organizing, I am going to arrange the `State` column to be in ascending order.

```
strwb_census <- strwb_census |>
  arrange(State)
```

Professor Wright had organized the SURVEY data frame splitting the marketing, and production data from the chemical application data. In the strawberry data frame, The `CENSUS` rows contain marketing, sales, and production data. The `SURVEY` rows contain rows which may be redundant with the CENSUS rows and chemical application rows. These rows contain fresh and process market sales data, which have been removed.

References

Afrin, Sadia, Massimiliano Gasparrini, Tamara Y. Forbes-Hernandez, Patricia Reboredo-Rodriguez, Bruno Mezzetti, Alfonso Varela-López, Francesca Giampieri, and Maurizio Battino. 2016. "Promising Health Benefits of the Strawberry: A Focus on

Clinical Studies." *Journal of Agricultural and Food Chemistry* 64 (22): 4435–49. https://doi.org/10.1021/acs.jafc.6b00857.

CRAIG, WINSTON J. 1997. "Phytochemicals." *Journal of the American Dietetic Association* 97 (10): S199–204. https://doi.org/10.1016/s0002-8223(97)00765-7.

Fenik, Jolanta, Maciej Tankiewicz, and Marek Biziuk. 2011. "Properties and Determination of Pesticides in Fruits and Vegetables." *TrAC Trends in Analytical Chemistry* 30 (6): 814–26. https://doi.org/10.1016/j.trac.2011.02.008.

Verteramo Chiu, Leslie J., and Miguel I. Gomez. 2023. "A Tale of Two Strawberries: Conventional and Organic Open-Field Production in California." *Sustainability* 15 (19): 14363. https://doi.org/10.3390/su151914363.