Strawberries1

Keyu Feng

2024-10-05

#Set up

```
#Install the library
library(knitr)
library(kableExtra)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter()     masks stats::filter()
## x dplyr::group_rows() masks kableExtra::group_rows()
## x dplyr::lag()        masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(stringr)
setwd("C:/Users/16597/Downloads")
options(echo = FALSE, digits = 3,
        scipen = 999, warn = FALSE, message = FALSE)

#Quick overview of the data
strawberry <- read_csv("strawberries25_v3.csv", col_names = TRUE)
```

```
## Rows: 12669 Columns: 21
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (15): Program, Period, Geo Level, State, State ANSI, Ag District, County...
## dbl  (2): Year, Ag District Code
## lgl  (4): Week Ending, Zip Code, Region, Watershed
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(strawberry)
```

```
## Rows: 12,669
```

```
## Columns: 21
## $ Program            <chr> "CENSUS", "CENSUS", "CENSUS", "CENSUS", "CENSUS", "~
## $ Year               <dbl> 2022, 2022, 2022, 2022, 2022, 2022, 2022, 2022, 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> "COUNTY", "COUNTY", "COUNTY", "COUNTY", "COUNTY", "~
## $ State              <chr> "ALABAMA", "ALABAMA", "ALABAMA", "ALABAMA", "ALABAM~
## $ `State ANSI`       <chr> "01", "01", "01", "01", "01", "01", "01", "01", "01~
## $ `Ag District`      <chr> "BLACK BELT", "BLACK BELT", "BLACK BELT", "BLACK BE~
## $ `Ag District Code` <dbl> 40, 40, 40, 40, 40, 40, 40, 40, 40, 40, 40, 40, 40,~
## $ County             <chr> "BULLOCK", "BULLOCK", "BULLOCK", "BULLOCK", "BULLOC~
## $ `County ANSI`      <chr> "011", "011", "011", "011", "011", "011", "101", "1~
## $ `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 - ACRES BEARING", "STRAWBERRIES - ACR~
## $ Domain             <chr> "TOTAL", "TOTAL", "TOTAL", "TOTAL", "TOTAL", "TOTAL~
## $ `Domain Category`  <chr> "NOT SPECIFIED", "NOT SPECIFIED", "NOT SPECIFIED", ~
## $ Value              <chr> "(D)", "3", "(D)", "1", "6", "5", "(D)", "(D)", "2"~
## $ `CV (%)`           <chr> "(D)", "15.7", "(D)", "(L)", "52.7", "47.6", "(D)",~
```

#Remove columns with a single value in all rows

```
drop_one_value_col <- function(df){   ## takes whole dataframe
drop <- NULL

## test each column for a single value
for(i in 1:dim(df)[2]){
if((df |> distinct(df[,i]) |> count()) == 1){
drop = c(drop, i)
} }

## report the result -- names of columns dropped
## consider using the column content for labels
## or headers

if(is.null(drop)){return("none")}else{

   print("Columns dropped:")
   print(colnames(df)[drop])
   strawberry <- df[, -1*drop]
   }
}


## use the function

strawberry <- drop_one_value_col(strawberry)
```

```
## [1] "Columns dropped:"
## [1] "Week Ending"    "Zip Code"       "Region"         "watershed_code"
## [5] "Watershed"      "Commodity"
```

```
drop_one_value_col(strawberry)
```

```
## [1] "none"
```

#Load all states data instead of just California

```
# Remove the California-specific filter to explore all states
all_states <- strawberry

# Look at the unique values in the "Program" column for all states
unique(all_states$Program)
```

```
## [1] "CENSUS" "SURVEY"
```

```
# Split the data into CENSUS and SURVEY groups
all_states_census <- all_states |> filter(Program == "CENSUS")
all_states_survey <- all_states |> filter(Program == "SURVEY")

# Select specific columns (Year, Period, Data Item, Value) for SURVEY data
all_states_survey <- all_states |> select(Year, Period, `Data Item`, Value)
```

#Split columns from sorted data Looking at the data, I find that after splitting the columns, some information are false leading, my next goal is filling all valuables to their best-fit columns.

```
#Replace ' - ' (hyphen with spaces) with a comma.
strawberry <- strawberry |>
  mutate(`Data Item` = str_replace_all(`Data Item`, " - ", ","))

#Split 'Data Item' into 4 columns
strawberry <- strawberry |>
  separate_wider_delim(  cols = `Data Item`,
                         delim = ",",
                         names = c("Fruit",
                                   "Category",
                                   "Item",
                                   "Metric"),
                         too_many = "merge",
                         too_few = "align_start"
                       )
#Remove 'measured in' to metric columns
strawberry <- strawberry |>
  mutate(Metric = ifelse(grepl("MEASURED IN", Item), Item, Metric), # Move the 'Item' value to 'Metric'
    Item = ifelse(grepl("MEASURED IN", Item), NA, Item) # Set 'Item' to NA where we moved the value
  )
#Remove 'production' to its correct way.
strawberry <- strawberry |>
  mutate(
    Item = ifelse(grepl("PRODUCTION", Metric), "PRODUCTION", Item), # Move 'PRODUCTION' to 'Item'
    Metric = ifelse(grepl("PRODUCTION", Metric), sub("PRODUCTION, ", "", Metric), Metric)  # Remove 'PR
  )
#Remove 'utilized' from category to Item
```

```
strawberry <- strawberry |>
  mutate(
    Item = ifelse(grepl("UTILIZED", Category, ignore.case = TRUE),
                  paste("UTILIZED", Item, sep = " "), # Combine 'Item' with 'Utilized'
                  Item),  # Keep 'Item' unchanged if 'Utilized' not found
    Category = ifelse(grepl("UTILIZED", Category, ignore.case = TRUE), NA, Category)# Set 'Category' to
  )
#Consider a better waty to move items in one step.
movingitem<- c("ACRES BEARING", "ACRES NON-BEARING", "ACRES GROWN", "OPERATIONS WITH AREA BEARING", "YI

# Move terms from 'Metric' or 'Category' to 'Item' without replacing 'Metric' data
strawberry <- strawberry |>
  mutate(Item = ifelse(grepl(paste(movingitem, collapse = "|"), Category,
                             ignore.case = TRUE) & is.na(Item), Category,
      ifelse(grepl(paste(movingitem, collapse = "|"), Category, ignore.case = TRUE),
             paste(Item, Category, sep = ", "), Item)
    ),
    Category = ifelse(grepl(paste(movingitem, collapse = "|"), Category,
                            ignore.case = TRUE),
                      NA, Category)
  )
```

#Fixing the leading space

```
strawberry$Category[1]
```

```
## [1] NA
```

```
strawberry$Item[2]
```

```
## [1] "ACRES GROWN"
```

```
strawberry$Metric[6]
```

```
## [1] NA
```

```
strawberry$Domain[1]
```

```
## [1] "TOTAL"
```

```
#Trim the white space

strawberry$Category <- str_trim(strawberry$Category, side = "both")
strawberry$Item <- str_trim(strawberry$Item, side = "both")
strawberry$Metric <- str_trim(strawberry$Metric, side = "both")
```

#Split both 'Domain' and 'Domain Category' columns I find that in domain and domain category, the information is complicated, in this step I will split them into columns just like I did in 'Data Item'.

```r
# Split the Domain column into multiple categories
strawberry <- strawberry |>
  separate_wider_delim(
    cols = Domain,
    delim = " , ",
    names = c("Area Grown", "Fertilize", "Organic", "Chemical"),
    too_many = "merge",
    too_few = "align_start"
  )

#Loading variables to each column
strawberry <- strawberry |>
  mutate(
    Chemical = ifelse(grepl("CHEMICAL", `Area Grown`, ignore.case = TRUE), `Area Grown`, NA),
    Organic = ifelse(grepl("ORGANIC", `Area Grown`, ignore.case = TRUE), `Area Grown`, NA),
    Fertilize = ifelse(grepl("FERTILIZER", `Area Grown`, ignore.case = TRUE), `Area Grown`, NA),
    `Area Grown`= ifelse(grepl("CHEMICAL|ORGANIC|FERTILIZER", `Area Grown`, ignore.case = TRUE), NA, `A
  )

#Dealing with 'Domain Category' column
strawberry <- strawberry |>
  mutate(
    Chemical = ifelse(grepl("CHEMICAL", `Domain Category`, ignore.case = TRUE),
                      `Domain Category`,
                      Chemical),

    Organic = ifelse(grepl("ORGANIC", `Domain Category`, ignore.case = TRUE),
                     `Domain Category`,
                     Organic),

    Fertilize = ifelse(grepl("FERTILIZER", `Domain Category`, ignore.case = TRUE),
                       `Domain Category`,
                       Fertilize),

    `Area Grown` = ifelse(grepl("AREA", `Domain Category`, ignore.case = TRUE),
                          `Domain Category`,
                          `Area Grown`),

    `Domain Category` = ifelse(grepl("CHEMICAL|ORGANIC|FERTILIZER|AREA", `Domain Category`, ignore.case
  )

#Move 'Total' to its best place
strawberry <- strawberry |>
  mutate(Item = ifelse(grepl("Total", `Area Grown`, ignore.case = TRUE),
                       paste("Total", Item, sep = " "),
                       Item),

    `Area Grown` = ifelse(grepl("Total", `Area Grown`, ignore.case = TRUE), NA, `Area Grown`)
  )
```

#Split Chemical into three renamed columns Now we have splited chemical information, my next step is splitting it into 3 columns.

```r
table(strawberry$Chemical)
```

```
##
##                                        CHEMICAL, FUNGICIDE: (AZOXYSTROBIN = 128810)
##                                                                                  40
##                           CHEMICAL, FUNGICIDE: (BACILLUS AMYLOLIQUEFAC F727 = 16489)
##                                                                                  10
##                   CHEMICAL, FUNGICIDE: (BACILLUS AMYLOLIQUEFACIENS MBI 600 = 129082)
##                                                                                  15
##               CHEMICAL, FUNGICIDE: (BACILLUS AMYLOLIQUEFACIENS STRAIN D747 = 16482)
##                                                                                  20
##                                      CHEMICAL, FUNGICIDE: (BACILLUS PUMILUS = 6485)
##                                                                                  15
##                                 CHEMICAL, FUNGICIDE: (BACILLUS SUBT. GB03 = 129068)
##                                                                                   5
##                                     CHEMICAL, FUNGICIDE: (BACILLUS SUBTILIS = 6479)
##                                                                                  30
##                                                  CHEMICAL, FUNGICIDE: (BLAD = 30006)
##                                                                                  20
##                                    CHEMICAL, FUNGICIDE: (BORAX DECAHYDRATE = 11102)
##                                                                                  25
##                                           CHEMICAL, FUNGICIDE: (BOSCALID = 128008)
##                                                                                  30
##                           CHEMICAL, FUNGICIDE: (BT SUBSP KURSTAKI EVB-113-19 = 6544)
##                                                                                  15
##                                              CHEMICAL, FUNGICIDE: (CAPTAN = 81301)
##                                                                                  40
##                                       CHEMICAL, FUNGICIDE: (CHLOROTHALONIL = 81901)
##                                                                                   5
##                                 CHEMICAL, FUNGICIDE: (COPPER CHLORIDE HYD. = 23501)
##                                                                                  15
##                                     CHEMICAL, FUNGICIDE: (COPPER HYDROXIDE = 23401)
##                                                                                  25
##                                     CHEMICAL, FUNGICIDE: (COPPER OCTANOATE = 23306)
##                                                                                  15
##                                        CHEMICAL, FUNGICIDE: (CYFLUFENAMID = 555550)
##                                                                                  35
##                                           CHEMICAL, FUNGICIDE: (CYMOXANIL = 129106)
##                                                                                   5
##                                         CHEMICAL, FUNGICIDE: (CYPRODINIL = 288202)
##                                                                                  40
##                                      CHEMICAL, FUNGICIDE: (DIFENOCONAZOLE = 128847)
##                                                                                  40
##                                              CHEMICAL, FUNGICIDE: (DODINE = 44301)
##                                                                                   5
##                                         CHEMICAL, FUNGICIDE: (FAMOXADONE = 113202)
##                                                                                   5
##                                          CHEMICAL, FUNGICIDE: (FENHEXAMID = 90209)
##                                                                                  40
##                                         CHEMICAL, FUNGICIDE: (FLUDIOXONIL = 71503)
##                                                                                  40
##                                           CHEMICAL, FUNGICIDE: (FLUOPYRAM = 80302)
##                                                                                  35
```

```
##                                  CHEMICAL, FUNGICIDE: (FLUTOLANIL = 128975)
##                                                                           5
##                                CHEMICAL, FUNGICIDE: (FLUXAPYROXAD = 138009)
##                                                                          30
##                                  CHEMICAL, FUNGICIDE: (FOSETYL-AL = 123301)
##                                                                          35
##                                   CHEMICAL, FUNGICIDE: (IPRODIONE = 109801)
##                                                                          15
##                                  CHEMICAL, FUNGICIDE: (ISOFETAMID = 270000)
##                                                                          35
##                                     CHEMICAL, FUNGICIDE: (MANCOZEB = 14504)
##                                                                           5
##                                   CHEMICAL, FUNGICIDE: (MEFENOXAM = 113502)
##                                                                          40
##                             CHEMICAL, FUNGICIDE: (MONO-POTASSIUM SALT = 76416)
##                                                                          30
##                                  CHEMICAL, FUNGICIDE: (MYCLOBUTANIL = 128857)
##                                                                          25
##                                CHEMICAL, FUNGICIDE: (OXATHIAPIPROLIN = 128111)
##                                                                          10
##                                   CHEMICAL, FUNGICIDE: (PENTHIOPYRAD = 90112)
##                                                                          35
##                           CHEMICAL, FUNGICIDE: (POLYOXIN D ZINC SALT = 230000)
##                                                                          20
##                             CHEMICAL, FUNGICIDE: (POTASSIUM BICARBON. = 73508)
##                                                                          15
##                                  CHEMICAL, FUNGICIDE: (PROPICONAZOLE = 122101)
##                                                                          35
##                                  CHEMICAL, FUNGICIDE: (PYDIFLUMETOFEN = 90110)
##                                                                          15
##                                  CHEMICAL, FUNGICIDE: (PYRACLOSTROBIN = 99100)
##                                                                          35
##                                   CHEMICAL, FUNGICIDE: (PYRIMETHANIL = 288201)
##                                                                          40
##                                    CHEMICAL, FUNGICIDE: (PYRIOFENONE = 28828)
##                                                                           5
##                                      CHEMICAL, FUNGICIDE: (QUINOLINE = 55459)
##                                                                          20
##                             CHEMICAL, FUNGICIDE: (STREPTOMYCES LYDICUS = 6327)
##                                                                          15
##                                         CHEMICAL, FUNGICIDE: (SULFUR = 77501)
##                                                                          35
##                                  CHEMICAL, FUNGICIDE: (TETRACONAZOLE = 120603)
##                                                                          35
##                             CHEMICAL, FUNGICIDE: (THIOPHANATE-METHYL = 102001)
##                                                                          40
##                                         CHEMICAL, FUNGICIDE: (THIRAM = 79801)
##                                                                          40
##                                                  CHEMICAL, FUNGICIDE: (TOTAL)
##                                                                          16
##                              CHEMICAL, FUNGICIDE: (TRICHODERMA HARZ. = 119202)
##                                                                           5
##                                CHEMICAL, FUNGICIDE: (TRIFLOXYSTROBIN = 129112)
##                                                                          20
```

```
##                                 CHEMICAL, FUNGICIDE: (TRIFLUMIZOLE = 128879)
##                                                                           35
##                                     CHEMICAL, FUNGICIDE: (ZOXAMIDE = 101702)
##                                                                            5
##                                  CHEMICAL, HERBICIDE: (2,4-D, DIMETH. SALT = 30019)
##                                                                           10
##                                  CHEMICAL, HERBICIDE: (2,4-D, TRIISO. SALT = 30035)
##                                                                            5
##                                 CHEMICAL, HERBICIDE: (CARFENTRAZONE-ETHYL = 128712)
##                                                                           25
##                                           CHEMICAL, HERBICIDE: (CLETHODIM = 121011)
##                                                                           10
##                                  CHEMICAL, HERBICIDE: (COPPER ETHANOLAMINE = 24409)
##                                                                            5
##                                        CHEMICAL, HERBICIDE: (DIMETHENAMID = 129051)
##                                                                            5
##                                         CHEMICAL, HERBICIDE: (FLUMIOXAZIN = 129034)
##                                                                           35
##                                    CHEMICAL, HERBICIDE: (FLUROXYPYR 1-MHE = 128968)
##                                                                            5
##                                CHEMICAL, HERBICIDE: (GLUFOSINATE-AMMONIUM = 128850)
##                                                                            5
##                                CHEMICAL, HERBICIDE: (GLYPHOSATE ISO. SALT = 103601)
##                                                                           35
##                                CHEMICAL, HERBICIDE: (GLYPHOSATE POT. SALT = 103613)
##                                                                           20
##                                 CHEMICAL, HERBICIDE: (HALOSULFURON-METHYL = 128721)
##                                                                            5
##                                              CHEMICAL, HERBICIDE: (KANTOR = 129108)
##                                                                            5
##                                  CHEMICAL, HERBICIDE: (METSULFURON-METHYL = 122010)
##                                                                            5
##                                         CHEMICAL, HERBICIDE: (NAPROPAMIDE = 103001)
##                                                                           20
##                                         CHEMICAL, HERBICIDE: (OXYFLUORFEN = 111601)
##                                                                           25
##                                             CHEMICAL, HERBICIDE: (PARAQUAT = 61601)
##                                                                           25
##                                       CHEMICAL, HERBICIDE: (PENDIMETHALIN = 108501)
##                                                                           20
##                                          CHEMICAL, HERBICIDE: (PENOXSULAM = 119031)
##                                                                            5
##                                       CHEMICAL, HERBICIDE: (S-METOLACHLOR = 108800)
##                                                                            5
##                                       CHEMICAL, HERBICIDE: (SULFENTRAZONE = 129081)
##                                                                           10
##                                                    CHEMICAL, HERBICIDE: (TOTAL)
##                                                                           16
##                                         CHEMICAL, INSECTICIDE: (ABAMECTIN = 122804)
##                                                                           40
##                                         CHEMICAL, INSECTICIDE: (ACEQUINOCYL = 6329)
##                                                                           20
##                                        CHEMICAL, INSECTICIDE: (ACETAMIPRID = 99050)
##                                                                           40
```

```
##                                     CHEMICAL, INSECTICIDE: (AZADIRACHTIN = 121701)
##                                                                                 20
##                                  CHEMICAL, INSECTICIDE: (BEAUVERIA BASSIANA = 128924)
##                                                                                 15
##                                     CHEMICAL, INSECTICIDE: (BETA-CYFLUTHRIN = 118831)
##                                                                                  5
##                                           CHEMICAL, INSECTICIDE: (BIFENAZATE = 586)
##                                                                                 40
##                                        CHEMICAL, INSECTICIDE: (BIFENTHRIN = 128825)
##                                                                                 40
##                                  CHEMICAL, INSECTICIDE: (BT KURSTAK ABTS-1857 = 6523)
##                                                                                 25
##                                  CHEMICAL, INSECTICIDE: (BT KURSTAKI ABTS-351 = 6522)
##                                                                                 35
##                                    CHEMICAL, INSECTICIDE: (BT KURSTAKI EG7841 = 6453)
##                                                                                  5
##                                     CHEMICAL, INSECTICIDE: (BT KURSTAKI SA-11 = 6519)
##                                                                                 20
##                                  CHEMICAL, INSECTICIDE: (BT SUB AIZAWAI GC-91 = 6426)
##                                                                                 15
##                                        CHEMICAL, INSECTICIDE: (BUPROFEZIN = 275100)
##                                                                                 15
##                       CHEMICAL, INSECTICIDE: (BURKHOLDERIA A396 CELLS & MEDIA = 6534)
##                                                                                 15
##                                         CHEMICAL, INSECTICIDE: (CANOLA OIL = 11332)
##                                                                                 10
##                                           CHEMICAL, INSECTICIDE: (CARBARYL = 56801)
##                                                                                 15
##                                  CHEMICAL, INSECTICIDE: (CHLORANTRANILIPROLE = 90100)
##                                                                                 40
##                                       CHEMICAL, INSECTICIDE: (CHLORPYRIFOS = 59101)
##                                                                                 10
## CHEMICAL, INSECTICIDE: (CHROMOBAC SUBTSUGAE PRAA4-1 CELLS AND SPENT MEDIA = 16329)
##                                                                                 20
##                                     CHEMICAL, INSECTICIDE: (CYANTRANILIPROLE = 90098)
##                                                                                 35
##                                       CHEMICAL, INSECTICIDE: (CYCLANILIPROLE = 26202)
##                                                                                  5
##                                        CHEMICAL, INSECTICIDE: (CYFLUMETOFEN = 138831)
##                                                                                 25
##                                        CHEMICAL, INSECTICIDE: (CYFLUMETOFEN = 138831)
##                                                                                  5
##                                        CHEMICAL, INSECTICIDE: (CYPERMETHRIN = 109702)
##                                                                                  5
##                                           CHEMICAL, INSECTICIDE: (DIAZINON = 57801)
##                                                                                 20
##                                  CHEMICAL, INSECTICIDE: (EMAMECTIN BENZOATE = 122806)
##                                                                                  5
##                          CHEMICAL, INSECTICIDE: (ETHYL (2E;4Z)-DECADIENOATE = 144022)
##                                                                                  5
##                                        CHEMICAL, INSECTICIDE: (ETOXAZOLE = 107091)
##                                                                                 20
##                                        CHEMICAL, INSECTICIDE: (FENAZAQUIN = 44501)
##                                                                                  5
```

```
##                                CHEMICAL, INSECTICIDE: (FENBUTATIN-OXIDE = 104601)
##                                                                                20
##                                  CHEMICAL, INSECTICIDE: (FENPROPATHRIN = 127901)
##                                                                                25
##                                  CHEMICAL, INSECTICIDE: (FENPYROXIMATE = 129131)
##                                                                                30
##                                     CHEMICAL, INSECTICIDE: (FLONICAMID = 128016)
##                                                                                25
##                                 CHEMICAL, INSECTICIDE: (FLUPYRADIFURONE = 122304)
##                                                                                25
##                               CHEMICAL, INSECTICIDE: (HELICOVERPA ZEA NPV = 107300)
##                                                                                10
##                                    CHEMICAL, INSECTICIDE: (HEXYTHIAZOX = 128849)
##                                                                                30
##                                   CHEMICAL, INSECTICIDE: (IMIDACLOPRID = 129099)
##                                                                                35
##                                 CHEMICAL, INSECTICIDE: (LAMBDA-CYHALOTHRIN = 128897)
##                                                                                15
##                                      CHEMICAL, INSECTICIDE: (MALATHION = 57701)
##                                                                                40
##                                       CHEMICAL, INSECTICIDE: (METHOMYL = 90301)
##                                                                                 5
##                                  CHEMICAL, INSECTICIDE: (METHOXYFENOZIDE = 121027)
##                                                                                30
##                                     CHEMICAL, INSECTICIDE: (MUSTARD OIL = 4901)
##                                                                                 5
##                                         CHEMICAL, INSECTICIDE: (NALED = 34401)
##                                                                                40
##                                       CHEMICAL, INSECTICIDE: (NEEM OIL = 25006)
##                                                                                20
##                               CHEMICAL, INSECTICIDE: (NEEM OIL, CLAR. HYD. = 25007)
##                                                                                20
##                                     CHEMICAL, INSECTICIDE: (NOVALURON = 124002)
##                                                                                40
##                                        CHEMICAL, INSECTICIDE: (OXAMYL = 103801)
##                                                                                 5
##                                    CHEMICAL, INSECTICIDE: (PERMETHRIN = 109701)
##                                                                                 5
##                               CHEMICAL, INSECTICIDE: (PETROLEUM DISTILLATE = 63503)
##                                                                                 5
##                                CHEMICAL, INSECTICIDE: (PIPERONYL BUTOXIDE = 67501)
##                                                                                25
##                                  CHEMICAL, INSECTICIDE: (POTASSIUM SALTS = 79021)
##                                                                                15
##                                     CHEMICAL, INSECTICIDE: (PYRETHRINS = 69001)
##                                                                                25
##                                     CHEMICAL, INSECTICIDE: (PYRIDABEN = 129105)
##                                                                                15
##                                   CHEMICAL, INSECTICIDE: (PYRIPROXYFEN = 129032)
##                                                                                15
##                                     CHEMICAL, INSECTICIDE: (SOYBEAN OIL = 31605)
##                                                                                 5
##                                     CHEMICAL, INSECTICIDE: (SPINETORAM = 110007)
##                                                                                40
```

```
##                                  CHEMICAL, INSECTICIDE: (SPINOSAD = 110003)
##                                                                          25
##                                CHEMICAL, INSECTICIDE: (SPIROMESIFEN = 24875)
##                                                                          30
##                             CHEMICAL, INSECTICIDE: (SPIROTETRAMAT = 392201)
##                                                                           5
##                                CHEMICAL, INSECTICIDE: (SULFOXAFLOR = 5210)
##                                                                          20
##                                CHEMICAL, INSECTICIDE: (THIAMETHOXAM = 60109)
##                                                                          40
##                                                 CHEMICAL, INSECTICIDE: (TOTAL)
##                                                                          16
##                              CHEMICAL, INSECTICIDE: (ZETA-CYPERMETHRIN = 129064)
##                                                                           5
##                                 CHEMICAL, OTHER: (ACIBENZOLAR-S-METHYL = 61402)
##                                                                          30
##                                 CHEMICAL, OTHER: (ALKYL. DIM. BENZ. AM = 69105)
##                                                                           5
##                       CHEMICAL, OTHER: (AUREOBASIDIUM PULLULANS DSM 14940 = 46010)
##                                                                          10
##                       CHEMICAL, OTHER: (AUREOBASIDIUM PULLULANS DSM 14941 = 36010)
##                                                                          10
##                                CHEMICAL, OTHER: (BT KURSTAKI SA-12 = 6518)
##                                                                          10
##                                   CHEMICAL, OTHER: (CAPRIC ACID = 128955)
##                                                                          15
##                                 CHEMICAL, OTHER: (CAPRYLIC ACID = 128919)
##                                                                          15
##                          CHEMICAL, OTHER: (CAPSICUM OLEORESIN EXTRACT = 70704)
##                                                                          10
##                                   CHEMICAL, OTHER: (CHLOROPICRIN = 81501)
##                                                                          25
##                                CHEMICAL, OTHER: (CUPRAMMONIUM ACETATE = 36011)
##                                                                           5
##                                    CHEMICAL, OTHER: (CYTOKININS = 116801)
##                                                                          15
##                                 CHEMICAL, OTHER: (DECYLDIMETHYLOCTYL = 69165)
##                                                                           5
##                                   CHEMICAL, OTHER: (DICHLOROPROPENE = 29001)
##                                                                          25
##                                CHEMICAL, OTHER: (DIDECYL DIM. AMMON. = 69166)
##                                                                           5
##                              CHEMICAL, OTHER: (DIMETHYL DISULFIDE (DMDS) = 29088)
##                                                                           5
##                                   CHEMICAL, OTHER: (DIMETHYLDIOCTYL = 69149)
##                                                                           5
##                                 CHEMICAL, OTHER: (DODECADIEN-1-OL = 129028)
##                                                                           5
##                                         CHEMICAL, OTHER: (ETHEPHON = 99801)
##                                                                           5
##                                    CHEMICAL, OTHER: (FLUENSULFONE = 50410)
##                                                                           5
##                                    CHEMICAL, OTHER: (FLUTRIAFOL = 128940)
##                                                                          30
```

```
##                                           CHEMICAL, OTHER: (GARLIC OIL = 128827)
##                                                                               10
##                                         CHEMICAL, OTHER: (GIBBERELLIC ACID = 43801)
##                                                                                5
##                                       CHEMICAL, OTHER: (GLIOCLADIUM VIRENS = 129000)
##                                                                               10
##                                          CHEMICAL, OTHER: (HYDROGEN PEROXIDE = 595)
##                                                                               20
##                                        CHEMICAL, OTHER: (INDOLEBUTYRIC ACID = 46701)
##                                                                               10
##                                           CHEMICAL, OTHER: (IRON PHOSPHATE = 34903)
##                                                                               20
##                           CHEMICAL, OTHER: (ISARIA FUMOSOROSEA STRAIN FE 9901 = 115003)
##                                                                                5
##                                           CHEMICAL, OTHER: (METALDEHYDE = 53001)
##                                                                               15
##                                         CHEMICAL, OTHER: (METAM-POTASSIUM = 39002)
##                                                                               30
##                                            CHEMICAL, OTHER: (METAM-SODIUM = 39003)
##                                                                               15
##                                           CHEMICAL, OTHER: (MINERAL OIL = 63502)
##                                                                                5
##                                      CHEMICAL, OTHER: (PAECILOMYCES FUMOSOR = 115002)
##                                                                               15
##                                         CHEMICAL, OTHER: (PEROXYACETIC ACID = 63201)
##                                                                               20
##                                        CHEMICAL, OTHER: (POTASSIUM SILICATE = 72606)
##                                                                               15
##                       CHEMICAL, OTHER: (PSEUDOMONAS CHLORORAPHIS STRAIN AFS009 = 6800)
##                                                                               20
##                                       CHEMICAL, OTHER: (REYNOUTRIA SACHALINE = 55809)
##                                                                               30
##                                                                CHEMICAL, OTHER: (TOTAL)
##                                                                               16
##                             CHEMICAL, OTHER: (TRICHODERMA VIRENS STRAIN G-41 = 176604)
##                                                                                5
```

```
strawberry <- strawberry |>
  mutate(Chemical = str_replace_all(Chemical, "[,:=()]", ","))

#Split it into three columns
strawberryc<- strawberry |>
  separate_wider_delim(
    cols = Chemical,
    delim = ",",
    names = c("Type", "Ingredient", "Code"), #Separate Chemical into type, ingredient, and code.
    too_many = "merge",
    too_few = "align_start"
  )
#Filling in the columns
strawberryc<- strawberryc |>
  mutate(
    Type = ifelse(Type == "CHEMICAL" | is.na(Type), Ingredient, Type),
```

```
    Ingredient = ifelse(!is.na(Ingredient), str_extract(Code, "\\b[A-Za-z\\-\\.\\s]+\\b"), Ingredient),

    Code = str_replace(Code, "\\b[A-Za-z\\-\\.\\s]+\\b", "")
  )

#Clean 'Code' Column
strawberryc <- strawberryc |>
  mutate(
    Code = str_replace_all(Code, "^\\s*,+|,+\\s*$|\\s*,\\s*,+", ""),
    Code = str_trim(Code)
  )

head(strawberryc)
```

```
## # A tibble: 6 x 23
##   Program  Year Period `Geo Level` State   `State ANSI` `Ag District`
##   <chr>   <dbl> <chr>  <chr>       <chr>   <chr>        <chr>
## 1 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## 2 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## 3 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## 4 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## 5 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## 6 CENSUS   2022 YEAR   COUNTY      ALABAMA 01           BLACK BELT
## # i 16 more variables: `Ag District Code` <dbl>, County <chr>,
## #   `County ANSI` <chr>, Fruit <chr>, Category <chr>, Item <chr>, Metric <chr>,
## #   `Area Grown` <chr>, Fertilize <chr>, Organic <chr>, Type <chr>,
## #   Ingredient <chr>, Code <chr>, `Domain Category` <chr>, Value <chr>,
## #   `CV (%)` <chr>
```