

Strawberries_HW2

Jie Fei

```
# read the cleaned strawberry and chemical data
strawberry <- read.csv("/Users/jie/Library/CloudStorage/OneDrive-BostonUniversity/Main Folder/02 Courses/PhD (BU)/2024 Fall/MA615 Data Science in R/HW/Strawberry HW/cleaned_strawberry.csv")
```

```
# load required package
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(tidyr)
library(reader)
```

```
## Loading required package: NCmisc
```

```
##
## Attaching package: 'reader'
```

```
## The following objects are masked from 'package:NCmisc':
##
##   cat.path, get.ext, rmv.ext
```

```
library(ggplot2)
library(tidyverse)
```

```
## — Attaching core tidyverse packages ————— tidyverse 2.0.0 —
## ✓ forcats   1.0.0   ✓ readr      2.1.5
## ✓ lubridate 1.9.3   ✓ stringr   1.5.1
## ✓ purrr     1.0.2   ✓ tibble    3.2.1
```

```
## — Conflicts — tidyverse_conflicts() —  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ⓘ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts  
to become errors
```

```
# install.packages('PubChemR')  
library(PubChemR)
```

```

#### clean and organize data set as needed

# remove null value where the value column is (D) and (NA)
strawberry <- strawberry[strawberry$Value != "(D)" & strawberry$Value != "(NA)", ]

# keep data only from California and Florida
# the analysis will focus on these two states
strawberry2 <- subset(strawberry, State %in% c("CALIFORNIA", "FLORIDA"))

# create subset for specific domains
strawberry_total <- subset(strawberry2, Domain == "TOTAL")
strawberry_area <- subset(strawberry2, Domain == "AREA GROWN")
strawberry_organic <- subset(strawberry2, Domain == "ORGANIC STATUS")
strawberry_chemical <- subset(strawberry2, !(Domain %in% c("TOTAL", "AREA GROWN", "ORGANIC STATUS")))

# in chemical data set, split the domain column and create two new columns (new domain and sub domain)
strawberry_chemical2 <- strawberry_chemical %>%
  mutate(
    New_Domain = ifelse(grepl(",", Domain), trimws(sapply(strsplit(as.character(Domain), ","), `[, 1])), NA),
    Sub_Domain = ifelse(grepl(",", Domain), trimws(sapply(strsplit(as.character(Domain), ","), `[, 2])), Domain)
  )

# split the domain category column and create two new columns (domain category and code)
strawberry_chemical2 <- strawberry_chemical2 %>%
  mutate(
    Domain_Category = ifelse(grepl("=", `Domain.Category`),
                             trimws(sub(".*\\(((^=)+)=.*\\)", "\\1", `Domain.Category`)),
                             trimws(sub(".*\\(((^)+)\\).*", "\\1", `Domain.Category`))),
    Code = ifelse(grepl("=", `Domain.Category`),
                  trimws(sub(".*=\\s*([^]+)\\).*", "\\1", `Domain.Category`)),
                  NA)
  )

# remove old columns
strawberry_chemical2 <- strawberry_chemical2 %>% select(-Domain)
strawberry_chemical2 <- strawberry_chemical2 %>% select(-`Domain.Category`)

# display 5 sample rows to double check
strawberry_chemical2_display <- strawberry_chemical2 %>%
  head(5)

print(strawberry_chemical2_display)

```

```
##      Program Year Period Geo.Level      State State.ANSI Ag.District
## 8763  SURVEY 2023   YEAR     STATE CALIFORNIA           6      <NA>
## 8767  SURVEY 2023   YEAR     STATE CALIFORNIA           6      <NA>
## 8768  SURVEY 2023   YEAR     STATE CALIFORNIA           6      <NA>
## 8770  SURVEY 2023   YEAR     STATE CALIFORNIA           6      <NA>
## 8772  SURVEY 2023   YEAR     STATE CALIFORNIA           6      <NA>
##      Ag.District.Code County County.ANSI      Fruit      Category
## 8763              NA  <NA>          NA STRAWBERRIES BEARING - APPLICATIONS
## 8767              NA  <NA>          NA STRAWBERRIES BEARING - APPLICATIONS
## 8768              NA  <NA>          NA STRAWBERRIES BEARING - APPLICATIONS
## 8770              NA  <NA>          NA STRAWBERRIES BEARING - APPLICATIONS
## 8772              NA  <NA>          NA STRAWBERRIES BEARING - APPLICATIONS
##      Item Metric  Value CV.... New_Domain Sub_Domain
## 8763 MEASURED IN LB  <NA>  3,300  <NA>    CHEMICAL  FUNGICIDE
## 8767 MEASURED IN LB  <NA>  2,800  <NA>    CHEMICAL  FUNGICIDE
## 8768 MEASURED IN LB  <NA>  6,600  <NA>    CHEMICAL  FUNGICIDE
## 8770 MEASURED IN LB  <NA> 603,100  <NA>    CHEMICAL  FUNGICIDE
## 8772 MEASURED IN LB  <NA> 30,300  <NA>    CHEMICAL  FUNGICIDE
##      Domain_Category  Code
## 8763      AZOXYSTROBIN 128810
## 8767      BORAX DECAHYDRATE 11102
## 8768      BOSCALID 128008
## 8770      CAPTAN 81301
## 8772      CYPRODINIL 288202
```

```
### Analysis 1
```

```
### What are the total usage pattern of fertilizers and chemicals in each state across different years (2018-2023)?
```

```
# read data
```

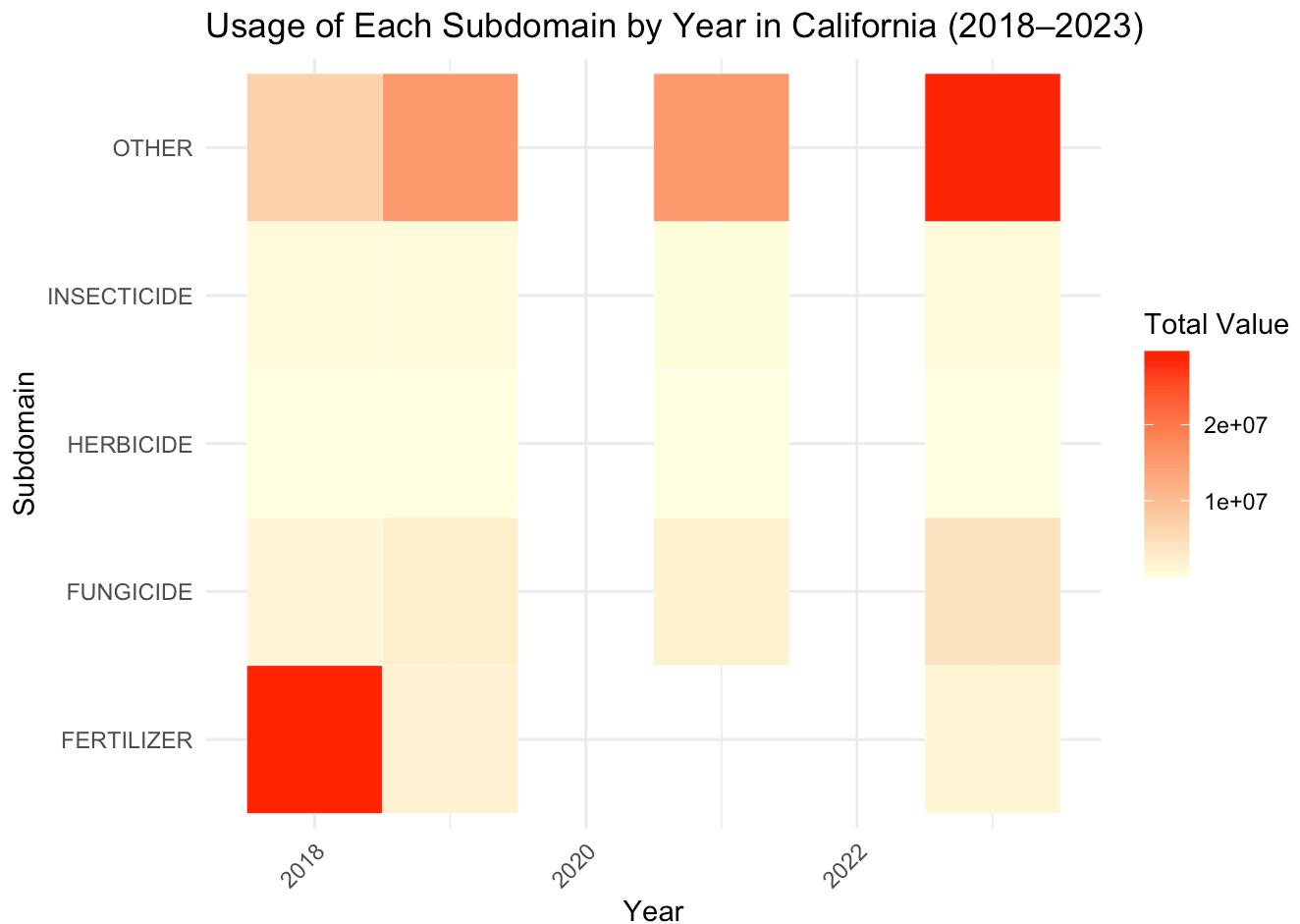
```
strawberry_chemical2_item_mib <- read.csv("strawberry_chemical2_item_mib.csv")
```

```
# filter data only for California
```

```
california_usage <- strawberry_chemical2_item_mib %>%
  filter(State == "CALIFORNIA") %>%
  group_by(Year, Sub_Domain) %>%
  summarize(Total_Value = sum(Value, na.rm = TRUE)) %>%
  ungroup()
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `.groups` argument.
```

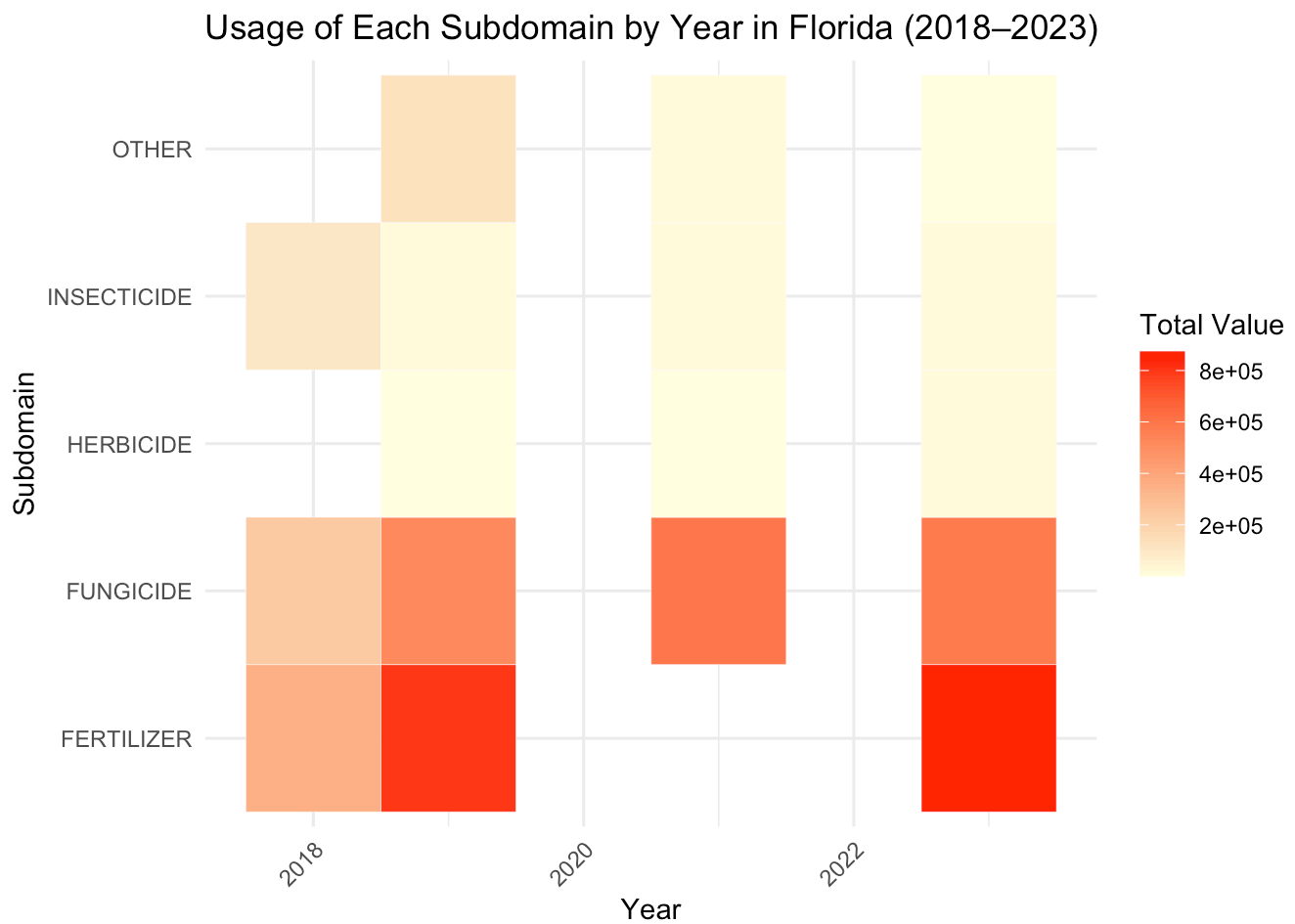
```
# plotting the heat map for California
ggplot(california_usage, aes(x = Year, y = Sub_Domain, fill = Total_Value)) +
  geom_tile(color = "white") +
  scale_fill_gradient(low = "lightyellow", high = "red") +
  theme_minimal() +
  labs(
    title = "Usage of Each Subdomain by Year in California (2018–2023)",
    x = "Year",
    y = "Subdomain",
    fill = "Total Value"
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```
# filter data only for Florida
florida_usage <- strawberry_chemical2_item_mib %>%
  filter(State == "FLORIDA") %>%
  group_by(Year, Sub_Domain) %>%
  summarize(Total_Value = sum(Value, na.rm = TRUE)) %>%
  ungroup()
```

```
## `summarise()` has grouped output by 'Year'. You can override using the
## `.groups` argument.
```

```
# plotting the heat map for Florida
ggplot(florida_usage, aes(x = Year, y = Sub_Domain, fill = Total_Value)) +
  geom_tile(color = "white") +
  scale_fill_gradient(low = "lightyellow", high = "red") +
  theme_minimal() +
  labs(
    title = "Usage of Each Subdomain by Year in Florida (2018–2023)",
    x = "Year",
    y = "Subdomain",
    fill = "Total Value"
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



The figures compare agricultural chemical usage by sub domain (fertilizers, fungicides, herbicides, insecticides, and others) between California and Florida from 2018 to 2023. In California, fertilizers and "other" chemicals show the highest usage, with a consistent trend over the years, while fungicides, herbicides, and insecticides have relatively low and stable usage. The color intensity indicates a broader range of total values, with California's overall chemical usage exceeding Florida's.

In Florida, fertilizer usage is also prominent, especially in 2019 and 2023. Fungicides also show significant usage but at lower levels than fertilizers, while herbicides and insecticides remain minimal across the years. The color scale for Florida highlights a lower total usage range than California, indicating that California uses more agricultural chemicals overall. These differences underscore regional variations in agricultural practices and chemical dependency.

Analysis 2

What are the differences in the usage of each chemical or fertilizer between Florida and California each year?

filter data for California and Florida, and aggregate by Year, State, and Domain_Category

```
domain_comparison <- strawberry_chemical2_item_mib %>%
  filter(State %in% c("CALIFORNIA", "FLORIDA")) %>%
  filter(Domain_Category != "TOTAL") %>%
  group_by(Year, State, Domain_Category) %>%
  summarise(Total_Value = sum(Value, na.rm = TRUE)) %>%
  arrange(Year, Domain_Category, State)
```

*## `summarise()` has grouped output by 'Year', 'State'. You can override using the
`.groups` argument.*

pivot the data to compare California and Florida side by side

```
domain_comparison_wide <- domain_comparison %>%
  pivot_wider(names_from = State, values_from = Total_Value, values_fill = 0) %>%
  rename(California_Usage = CALIFORNIA, Florida_Usage = FLORIDA)

print(domain_comparison_wide)
```

```
## # A tibble: 241 × 4
## # Groups:   Year [4]
##   Year Domain_Category      California_Usage Florida_Usage
##   <int> <chr>                <int>          <int>
## 1  2018 ABAMECTIN              200              0
## 2  2018 ACEQUINOCYL          1400              0
## 3  2018 ACETAMIPRID          1200              0
## 4  2018 AZADIRACTIN           600              0
## 5  2018 AZOXYSTROBIN         1100              0
## 6  2018 BIFENAZATE           5100              0
## 7  2018 BIFENTHRIN           1100              0
## 8  2018 BOSCALID              1800              0
## 9  2018 CAPTAN                94800            87400
## 10 2018 CHLORANTRANILIPROLE   600              0
## # i 231 more rows
```

Our results were similar to those in Analysis 2, showing that each state has its own unique patterns and trends in chemical and fertilizer usage.

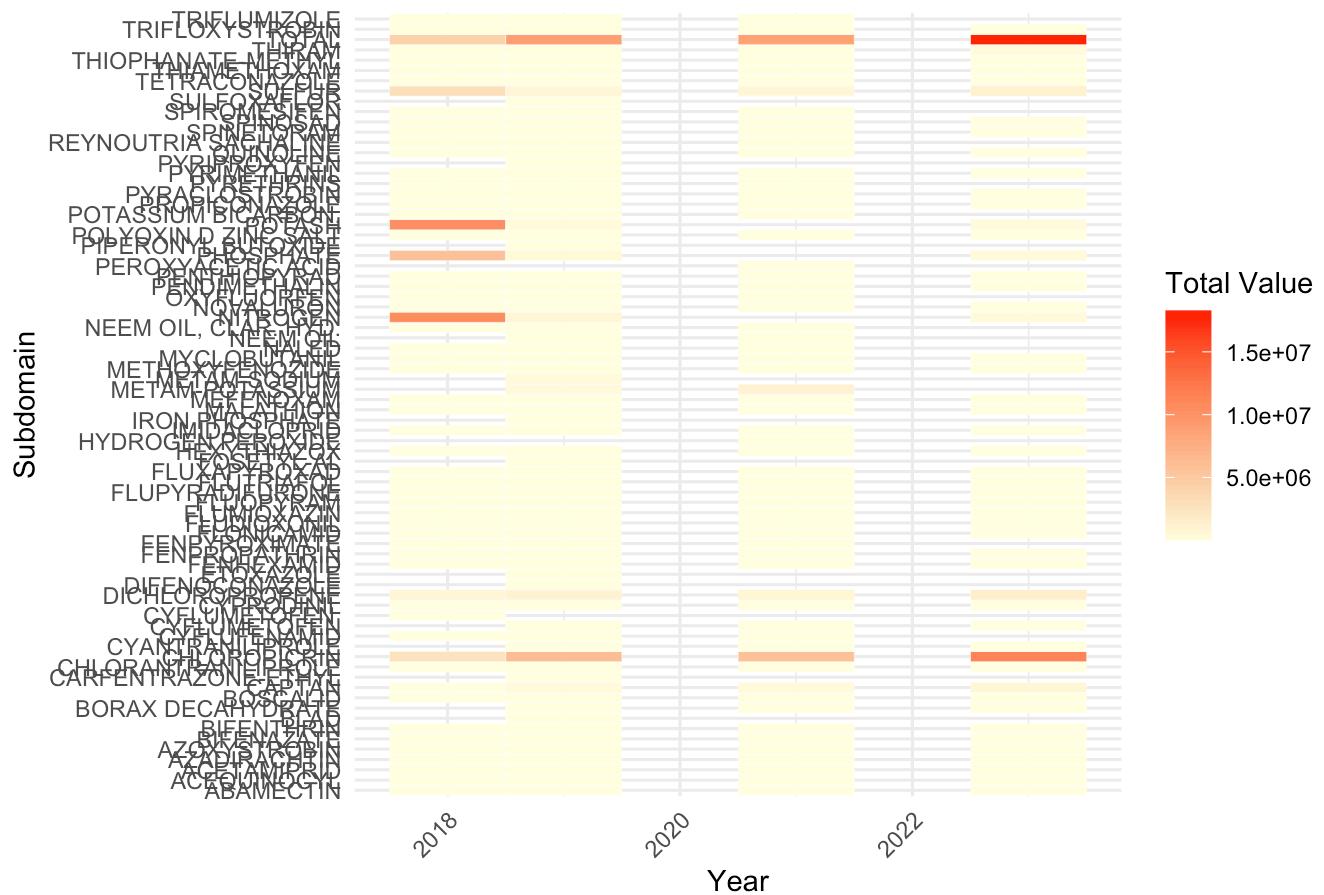
Analysis 3
What are the detailed total usage pattern of fertilizers and chemicals in California across different years (2018–2023)?

```
# filter data only for California
california_usage <- strawberry_chemical2_item_mib %>%
  filter(State == "CALIFORNIA") %>%
  group_by(Year, Domain_Category) %>%
  summarize(Total_Value = sum(Value, na.rm = TRUE)) %>%
  ungroup()
```

`summarise()` has grouped output by 'Year'. You can override using the
`.groups` argument.

```
# plotting the heat map for California
ggplot(california_usage, aes(x = Year, y = Domain_Category, fill = Total_Value)) +
  geom_tile(color = "white") +
  scale_fill_gradient(low = "lightyellow", high = "red") +
  theme_minimal() +
  labs(
    title = "Usage of Each Subdomain by Year in California (2018–2023)",
    x = "Year",
    y = "Subdomain",
    fill = "Total Value"
  ) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```


Usage of Each Subdomain by Year in California (2018–2023)



The heat map illustrates the usage of various chemicals and fertilizers sub domains in California from 2018 to 2023, with each row representing a different sub domain and each column corresponding to a specific year. The intensity of color (from light yellow to deep red) reflects the total usage amount, as indicated by the color scale on the right; darker shades represent higher usage levels, with the highest concentrations close to 15 million units shown in dark red. The chart highlights trends in chemicals and fertilizer s usage, with some chemicals and fertilizers showing significant spikes in specific year s, marked by intense red blocks, while others have consistently low or no usage across t he years, represented by lighter yellow or white blocks. Gaps or lighter colors between years for certain chemicals and fertilizers suggest either low, inconsistent, or non-con tinuous application of these substances, indicating variable demand or regulatory change s over time.

Analysis 4

What are the top 3 most used substances in each sub domain each year in California?

filter out rows where Domain_Category is "TOTAL" and find top 3 substances in each Sub_Domain (chemicals and fertilizers)

```
top_substances <- strawberry_chemical2_item_mib %>%
  filter(Domain_Category != "TOTAL" & State == "CALIFORNIA") %>%
  group_by(Year, Sub_Domain, Domain_Category) %>%
  summarise(Total_Value = sum(Value, na.rm = TRUE)) %>%
  arrange(Year, Sub_Domain, desc(Total_Value)) %>%
  group_by(Year, Sub_Domain) %>%
  slice_max(order_by = Total_Value, n = 3)
```

`summarise()` has grouped output by 'Year', 'Sub_Domain'. You can override
using the `.groups` argument.

display the result
print(top_substances)

```
## # A tibble: 56 × 4
## # Groups:   Year, Sub_Domain [19]
##   Year Sub_Domain Domain_Category Total_Value
##   <int> <chr>      <chr>          <int>
## 1  2018 FERTILIZER NITROGEN      10676000
## 2  2018 FERTILIZER POTASH       10583000
## 3  2018 FERTILIZER PHOSPHATE    5745000
## 4  2018 FUNGICIDE SULFUR        298200
## 5  2018 FUNGICIDE CAPTAN         94800
## 6  2018 FUNGICIDE THIRAM        20100
## 7  2018 HERBICIDE PENDIMETHALIN   2700
## 8  2018 HERBICIDE OXYFLUORFEN     800
## 9  2018 HERBICIDE FLUMIOXAZIN     100
## 10 2018 INSECTICIDE NEEM OIL, CLAR. HYD. 90500
## # i 46 more rows
```

Analysis 5

What are structure/composition/function/potential hazards of the top 3 most used substances in each sub domain for each year in California?

function 1

```
GHS_searcher<-function(result_json_object){
  result<-result_json_object
  for (i in 1:length(result[["result"]][["Hierarchies"]][["Hierarchy"]])){
    if(result[["result"]][["Hierarchies"]][["Hierarchy"]][i][["SourceName"]]=="GHS Classification (UNECE)"){
      return(i)
    }
  }
}
```

function 2

```
hazards_retriever<-function(index,result_json_object){
  result<-result_json_object
  hierarchy<-result[["result"]][["Hierarchies"]][["Hierarchy"]][index]
  i<-1
  output_list<-rep(NA,length(hierarchy[["Node"]]))
  while(str_detect(hierarchy[["Node"]][i][["Information"]][["Name"]],"H") & i<length(hierarchy[["Node"]])){
    output_list[i]<-hierarchy[["Node"]][i][["Information"]][["Name"]]
    i<-i+1
  }
  return(output_list[!is.na(output_list)])
}
```

function to safely retrieve information if it exists

```
safe_get <- function(x, ...) {
  result <- tryCatch({
    Reduce(function(x, name) if (!is.null(x) && name %in% names(x)) x[[name]] else NULL,
list(x, ...))
  }, error = function(e) NULL)
  result
}
```

extract unique Domain_Category values

```
unique_categories <- unique(top_substances$Domain_Category)
```

initialize a list to store results

```
result_list <- list()
```

loop through each unique category

```
for (category in unique_categories) {
  # retrieve data using get_pug_rest for each category
  result_d <- get_pug_rest(
    identifier = category,
    namespace = "name",
    domain = "compound",
    operation = "classification",
```

```
    output = "JSON"
  )

  # check if result_d contains expected data structure
  if (!is.null(safe_get(result_d, "result", "Hierarchies", "Hierarchy"))) {
    # process the retrieved data
    hazard_info <- hazards_retriever(GHS_searcher(result_d), result_d)

    # store the results in the list
    result_list[[category]] <- hazard_info
  } else {
    # handle cases where data is missing
    result_list[[category]] <- "Data not available"
  }
}
```

```
## Request failed [404]. Retrying in 3.5 seconds...
```

```
## Request failed [404]. Retrying in 2.3 seconds...
```

```
## Request failed [404]. Retrying in 2.5 seconds...
```

```
## Request failed [404]. Retrying in 7.4 seconds...
```

```
## Request failed [404]. Retrying in 1.8 seconds...
```

```
## Request failed [404]. Retrying in 7.3 seconds...
```

```
## Request failed [404]. Retrying in 2.4 seconds...
```

```
## Request failed [404]. Retrying in 5.2 seconds...
```

```
# view result_list
print(result_list)
```

```
## $NITROGEN
## [1] "H280: Contains gas under pressure; may explode if heated [Warning Gases under pressure]"
## [2] "H200: Physical Hazards"
## [3] "Hazard Statement Codes"
## [4] "H281: Contains refrigerated gas; may cause cryogenic burns or injury [Warning Gases under pressure]"
## [5] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [6] "H300: Health Hazards"
## [7] "H319: Causes serious eye irritation [Warning Serious eye damage/eye irritation]"
## [8] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [9] "H400: Environmental Hazards"
## [10] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $POTASH
## [1] "Data not available"
##
## $PHOSPHATE
## logical(0)
##
## $SULFUR
## [1] "H228: Flammable solid [Danger Flammable solids]"
## [2] "H200: Physical Hazards"
## [3] "Hazard Statement Codes"
## [4] "H315: Causes skin irritation [Warning Skin corrosion/irritation]"
## [5] "H300: Health Hazards"
## [6] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [7] "H319: Causes serious eye irritation [Warning Serious eye damage/eye irritation]"
## [8] "H370: Causes damage to organs [Danger Specific target organ toxicity, single exposure]"
## [9] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [10] "H413: May cause long lasting harmful effects to aquatic life [Hazardous to the aquatic environment, long-term hazard]"
## [11] "H400: Environmental Hazards"
##
## $CAPTAN
## [1] "H303: May be harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H315: Causes skin irritation [Warning Skin corrosion/irritation]"
## [5] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [6] "H318: Causes serious eye damage [Danger Serious eye damage/eye irritation]"
## [7] "H330: Fatal if inhaled [Danger Acute toxicity, inhalation]"
## [8] "H331: Toxic if inhaled [Danger Acute toxicity, inhalation]"
## [9] "H340: May cause genetic defects [Danger Germ cell mutagenicity]"
## [10] "H351: Suspected of causing cancer [Warning Carcinogenicity]"
## [11] "H361: Suspected of damaging fertility or the unborn child [Warning Reproductive]
```

```
toxicity]"
## [12] "H370: Causes damage to organs [Danger Specific target organ toxicity, single exposure]"
## [13] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Specific target organ toxicity, repeated exposure]"
## [14] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [15] "H400: Environmental Hazards"
## [16] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $THIRAM
## [1] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H302+H332: Harmful if swallowed or if inhaled [Warning Acute toxicity, oral; acute toxicity, inhalation]"
## [5] "H315: Causes skin irritation [Warning Skin corrosion/irritation]"
## [6] "H316: Causes mild skin irritation [Warning Skin corrosion/irritation]"
## [7] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [8] "H319: Causes serious eye irritation [Warning Serious eye damage/eye irritation]"
## [9] "H320: Causes eye irritation [Warning Serious eye damage/eye irritation]"
## [10] "H330: Fatal if inhaled [Danger Acute toxicity, inhalation]"
## [11] "H332: Harmful if inhaled [Warning Acute toxicity, inhalation]"
## [12] "H340: May cause genetic defects [Danger Germ cell mutagenicity]"
## [13] "H341: Suspected of causing genetic defects [Warning Germ cell mutagenicity]"
## [14] "H361: Suspected of damaging fertility or the unborn child [Warning Reproductive toxicity]"
## [15] "H370: Causes damage to organs [Danger Specific target organ toxicity, single exposure]"
## [16] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Specific target organ toxicity, repeated exposure]"
## [17] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [18] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [19] "H400: Environmental Hazards"
## [20] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $PENDIMETHALIN
## [1] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [5] "H351: Suspected of causing cancer [Warning Carcinogenicity]"
## [6] "H361: Suspected of damaging fertility or the unborn child [Warning Reproductive toxicity]"
## [7] "H361d: Suspected of damaging the unborn child [Warning Reproductive toxicity]"
## [8] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Specific target organ toxicity, repeated exposure]"
```

```

## [9] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment,
acute hazard]"
## [10] "H400: Environmental Hazards"
## [11] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous t
o the aquatic environment, long-term hazard]"
##
## $OXYFLUORFEN
## [1] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment,
acute hazard]"
## [2] "H400: Environmental Hazards"
## [3] "Hazard Statement Codes"
## [4] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous t
o the aquatic environment, long-term hazard]"
## [5] "Hazardous to the aquatic environment, acute hazard"
## [6] "Environmental Hazards"
## [7] "Hazard Classes"
## [8] "Hazardous to the aquatic environment, long-term hazard"
## [9] "<img src=\"https://pubchem.ncbi.nlm.nih.gov/images/ghs/GHS09.svg\" style=\"widt
h:40px;height:40px\"/> GHS09"
## [10] "Hazard Pictograms"
##
## $FLUMIOXAZIN
## [1] "H360: May damage fertility or the unborn child [Danger Reproductive toxicity]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H361d: Suspected of damaging the unborn child [Warning Reproductive toxicity]"
## [5] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment,
acute hazard]"
## [6] "H400: Environmental Hazards"
## [7] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous t
o the aquatic environment, long-term hazard]"
## [8] "Hazardous to the aquatic environment, acute hazard"
## [9] "Environmental Hazards"
## [10] "Hazard Classes"
## [11] "Hazardous to the aquatic environment, long-term hazard"
##
## $NEEM OIL, CLAR. HYD.`
## [1] "Data not available"
##
## $MALATHION
## [1] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [5] "H320: Causes eye irritation [Warning Serious eye damage/eye irritation]"
## [6] "H331: Toxic if inhaled [Danger Acute toxicity, inhalation]"
## [7] "H341: Suspected of causing genetic defects [Warning Germ cell mutagenicity]"
## [8] "H350: May cause cancer [Danger Carcinogenicity]"
## [9] "H370: Causes damage to organs [Danger Specific target organ toxicity, single ex
posure]"
## [10] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Sp
ecific target organ toxicity, repeated exposure]"

```

```
## [11] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [12] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [13] "H400: Environmental Hazards"
## [14] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $BIFENAZATE
## [1] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H319: Causes serious eye irritation [Warning Serious eye damage/eye irritation]"
## [5] "H320: Causes eye irritation [Warning Serious eye damage/eye irritation]"
## [6] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Specific target organ toxicity, repeated exposure]"
## [7] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [8] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [9] "H400: Environmental Hazards"
## [10] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
## [11] "Hazardous to the aquatic environment, acute hazard"
## [12] "Environmental Hazards"
## [13] "Hazard Classes"
## [14] "Hazardous to the aquatic environment, long-term hazard"
##
## $CHLOROPICRIN
## [1] "H301: Toxic if swallowed [Danger Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [5] "H314: Causes severe skin burns and eye damage [Danger Skin corrosion/irritation]"
## [6] "H315: Causes skin irritation [Warning Skin corrosion/irritation]"
## [7] "H318: Causes serious eye damage [Danger Serious eye damage/eye irritation]"
## [8] "H319: Causes serious eye irritation [Warning Serious eye damage/eye irritation]"
## [9] "H330: Fatal if inhaled [Danger Acute toxicity, inhalation]"
## [10] "H335: May cause respiratory irritation [Warning Specific target organ toxicity, single exposure; Respiratory tract irritation]"
## [11] "H370: Causes damage to organs [Danger Specific target organ toxicity, single exposure]"
## [12] "H372: Causes damage to organs through prolonged or repeated exposure [Danger Specific target organ toxicity, repeated exposure]"
## [13] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [14] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [15] "H400: Environmental Hazards"
```



```
## [16] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $DICHLOROPROPENE
## [1] "Data not available"
##
## $`REYNOUTRIA SACHALINE`
## [1] "Data not available"
##
## $`METAM-POTASSIUM`
## [1] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H312: Harmful in contact with skin [Warning Acute toxicity, dermal]"
## [5] "H314: Causes severe skin burns and eye damage [Danger Skin corrosion/irritation]"
## [6] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [7] "H332: Harmful if inhaled [Warning Acute toxicity, inhalation]"
## [8] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [9] "H400: Environmental Hazards"
## [10] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
##
## $ACEQUINOCYL
## [1] "H317: May cause an allergic skin reaction [Warning Sensitization, Skin]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H370: Causes damage to organs [Danger Specific target organ toxicity, single exposure]"
## [5] "H373: May causes damage to organs through prolonged or repeated exposure [Warning Specific target organ toxicity, repeated exposure]"
## [6] "H400: Very toxic to aquatic life [Warning Hazardous to the aquatic environment, acute hazard]"
## [7] "H400: Environmental Hazards"
## [8] "H410: Very toxic to aquatic life with long lasting effects [Warning Hazardous to the aquatic environment, long-term hazard]"
## [9] "Hazardous to the aquatic environment, acute hazard"
## [10] "Environmental Hazards"
## [11] "Hazard Classes"
## [12] "Hazardous to the aquatic environment, long-term hazard"
##
## $FLUTRIAFOL
## [1] "H302: Harmful if swallowed [Warning Acute toxicity, oral]"
## [2] "H300: Health Hazards"
## [3] "Hazard Statement Codes"
## [4] "H312: Harmful in contact with skin [Warning Acute toxicity, dermal]"
## [5] "H332: Harmful if inhaled [Warning Acute toxicity, inhalation]"
## [6] "H411: Toxic to aquatic life with long lasting effects [Hazardous to the aquatic environment, long-term hazard]"
## [7] "H400: Environmental Hazards"
```

```
## [8] "H412: Harmful to aquatic life with long lasting effects [Hazardous to the aquatic environment, long-term hazard]"
```

```
# The structure/composition/function/potential hazards of the top 3 most used substances in each sub domain for each year in California were showed below to get more details about these substances.
```