

# Kern Filter applied to original CDFA data

The purpose of this Rmarkdown file is to filter the CDFA data for Kern County, then write an output CSV of the filtered data

Notes:

The data being read in with the read\_csv() function is from the original CDFA data. Years 2013-2017 were separated by tabs in the same excel (.xls) sheet. Prior to being read in to this Rmarkdown document, each tab was converted into an individual .csv file in Excel because R does not deal with .xls sheets well, especially those with multiple tabs.

## 1. Read in CSVs

The data being read in with the read\_csv() function is from the original CDFA data. Years 2013-2017 were separated by tabs in the same excel (.xls) sheet. Prior to being read in to this Rmarkdown document, each tab was converted into an individual .csv file in Excel because R does not deal with .xls sheets well, especially those with multiple tabs

```
library(tidyverse)

# Specify years to be included in this filtering process
years = 2013:2017

# Read in CSVs of APNs for each year
for(i in years){

  # Read in specific year's CDFA CSV file
  tmp = read_csv(paste0("../R_input/CSV/CDFA/original/CDFA_OrgAPNs_",i,".csv")) %>%
    as.data.frame()

  # Assign a simple name to the dataframe
  assign(paste0("APN_",i),tmp)

  # Remove the temporary 'tmp' sheet to ensure there will be no accidental cross over between years
  rm(tmp)

}

head(APN_2013,5) # Below is a sample (n = 5) of the complete CDFA dataframe
```

##	Organic ID	Company	Site Name	Site APN	Site Acreage
## 1	33-001855	123 Farm	HSR	<NA>	814
## 2	33-001855	123 Farm	Black Bench	<NA>	1498
## 3	33-001855	123 Farm	HSR	<NA>	0
## 4	18-003983	3-C Hay & Grain	3-C Ranch	<NA>	1035
## 5	18-003983	3-C Hay & Grain	Johnston Ranch	<NA>	360
##	Site Address	Site City	Site State	Site Zip	
## 1	10600 Highland Springs Avenue	Cherry Valley	CA	92223	
## 2	9435 Bluff Road	Banning Bench	CA	92220	
## 3	10600 Highland Springs Avenue	Cherry Valley	CA	92223	
## 4	665--275 Lenard Road	Bieber	CA	<NA>	
## 5	10 Guttry Road	Bieber	CA	<NA>	
##	Site County	OpType	Address	Address 2	City

```
## 1  Riverside Producer 10600 Highland Springs Ave    <NA> Cherry Valley
## 2  Riverside Producer 10600 Highland Springs Ave    <NA> Cherry Valley
## 3  Riverside Handler 10600 Highland Springs Ave    <NA> Cherry Valley
## 4    Lassen Producer           2400 Pepper Rd      <NA>    Petaluma
## 5    Lassen Producer           2400 Pepper Rd      <NA>    Petaluma
##   State   Zip           Name Site APN_1
## 1    CA 92223   Jan Kielmann    <NA>
## 2    CA 92223   Jan Kielmann    <NA>
## 3    CA 92223   Jan Kielmann    <NA>
## 4    CA 94952 Robert Camozzi    <NA>
## 5    CA 94952 Robert Camozzi    <NA>
```

## 2. Filter each year for Kern County only'

The original CDFA data included information for the whole state

```
for(i in years){

  tmp = eval( # Evaluate what the name refers to -- In combination these commands are telling are to g
    as.name( # Tell R that the output of the paste0() command is the name of something
      paste0("APN_",i)) ) %>% # Paste elements together

    filter(`Site County` == "Kern") # Once the right dataframe has been called, filter for the county c

  assign(paste0("kern_APN_",i),tmp) # Assign a simple name to the output dataframe

  rm(tmp) # Remove the temporary dataframe

}

head(kern_APN_2013,5) # Below is a sample (n = 5) of the filtered dataframe
```

```
##   Organic ID           Company           Site Name Site APN
## 1  15-001012   Alvie Snow Farm           #5    <NA>
## 2  15-001018 Anthony Vineyards, Inc.    01 Mesa Ranch    <NA>
## 3  15-001018 Anthony Vineyards, Inc. 09 Central Ranch South    <NA>
## 4  15-001018 Anthony Vineyards, Inc. 10 North David Ranch    <NA>
## 5  15-001018 Anthony Vineyards, Inc. 11 Kovac Ranch    <NA>
##   Site Acreage   Site Address   Site City Site State Site Zip
## 1      32.00    16243 Magnolia    Wasco      CA    93280
## 2     299.65  5512 Valpredo Ave Bakersfield    CA    93307
## 3      81.00  5512 Valpredo Ave Bakersfield    CA    93307
## 4      77.00  5512 Valpredo Ave Bakersfield    CA    93307
## 5      71.00  5512 Valpredo Ave Bakersfield    CA    93307
##   Site County   OpType           Address Address 2           City State   Zip
## 1      Kern Producer 1907 7th Street    <NA>    Wasco      CA 93280
## 2      Kern Producer   P O Box 9578    <NA> Bakersfield    CA 93389
## 3      Kern Producer   P O Box 9578    <NA> Bakersfield    CA 93389
## 4      Kern Producer   P O Box 9578    <NA> Bakersfield    CA 93389
## 5      Kern Producer   P O Box 9578    <NA> Bakersfield    CA 93389
##           Name Site APN_1
## 1 Bev Rogers    <NA>
## 2 Tom Karle     <NA>
## 3 Tom Karle     <NA>
## 4 Tom Karle     <NA>
```

```
## 5 Tom Karle <NA>
```

### 3. Export filtered CSV

```
# Export Kern APNs for each year
for(i in years){

  write_csv(
    eval( # Evaluate what the name refers to -- In combination these commands are telling are to grab
      as.name( # Tell R that paste is a name
        paste0("kern_APN_",i)), # Paste elements together

    paste0("../R_output/CSV/CDFA/kern_filter/Kern_CDFA_APNs_",i,".csv")) # Write the output .csv file
}
```

The APNs in the exported CSVs went through 2 editing phases – 1 in Excel and 1 in Rstudio. See methods notes for more information.