Maxton Lam's Portfolio of Water Use Data

Maxton Lam

3/10/2021

The data for this are available from the United States Geological Survey for years 1950-2015.

Step 1: Load libraries

```
library(tidyverse)
library(here)
library(janitor)
library(readxl)
library(sf)
library(USAboundaries)
library(scales)
library(ggthemes)
library(gghighlight)
library(ggnewscale)
```

Step 2: Load data

```
#Read in 1950 data, skip 3 lines, join by "Area", clean names, coerce into numerics, replace NA with 0
#Use (.) in replace. The dot represents the respective dataframe being called.
#Deselect extra column so that d_wu_1950 has 8 columns
d_wu_1950 <- lapply(excel_sheets(here("data/us1950.xlsx")),
function(x) read_excel(here("data/us1950.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
```

```r
  clean_names() %>%
  mutate_at(1:9, as.numeric) %>%
  replace(is.na(.), 0) %>%
  select(1:6, 8, 9)

#Read in 1955 data, follow 1950 procedures
#Deselect extra column so that d_wu_1955 has 10 columns
d_wu_1955 <- lapply(excel_sheets(here("data/us1955.xlsx")),
function(x) read_excel(here("data/us1955.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:11, as.numeric) %>%
  replace(is.na(.), 0) %>%
  select(1:8, 10, 11)

#Read in 1960 data, follow 1950 procedures
d_wu_1960 <- lapply(excel_sheets(here("data/us1960.xlsx")),
function(x) read_excel(here("data/us1960.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:35, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1965 data, follow 1950 procedures
d_wu_1965 <- lapply(excel_sheets(here("data/us1965.xlsx")),
function(x) read_excel(here("data/us1965.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:33, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1970 data, follow 1950 procedures
d_wu_1970 <- lapply(excel_sheets(here("data/us1970.xlsx")),
function(x) read_excel(here("data/us1970.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:34, as.numeric) %>%
  replace(is.na(.), 0)
```

```r
#Read in 1975 data, follow 1950 procedures
d_wu_1975 <- lapply(excel_sheets(here("data/us1975.xlsx")),
function(x) read_excel(here("data/us1975.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:34, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1980 data, follow 1950 procedures
d_wu_1980 <- lapply(excel_sheets(here("data/us1980.xlsx")),
function(x) read_excel(here("data/us1980.xlsx"), skip = 3, sheet = x)) %>%
  reduce(left_join, by = "Area") %>%
  clean_names() %>%
  mutate_at(1:34, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1985 data, clean names, coerce into numerics, replace NA with 0
#Use read_tsv which already accounts for a tab delimited file.
d_wu_1985 <- read_tsv(here("data/us1985.txt")) %>%
  clean_names() %>%
  mutate_at(4:163, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1990 data, clean names, coerce into numerics, replace NA with 0
#Filter out "NA" because this could throw off the data when replaced with 0
d_wu_1990 <- read_excel(here("data/us1990.xls")) %>%
  clean_names() %>%
  filter(state != "NA") %>%
  mutate_at(4:163, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 1995 data, clean names, coerce into numerics, replace NA with 0
#Use mutate(across()) to select certain columns
#Turn FIPS to numeric because it is a number and to convert the notes at the end to 0
#Keep state and county_name as characters
d_wu_1995 <- read_excel(here("data/us1995.xls")) %>%
  clean_names() %>%
  mutate(across(c(1, 3:4, 6:252), as.numeric)) %>%
```

```
  replace(is.na(.), 0)

#Read in 2000 data, clean names, coerce into numerics, replace NA with 0
d_wu_2000 <- read_excel(here("data/us2000.xls")) %>%
  clean_names() %>%
  mutate_at(2:70, as.numeric) %>%
  replace(is.na(.), 0)

#Read in 2005 data, clean names, coerce into numerics, replace NA with 0
d_wu_2005 <- read_excel(here("data/us2005.xls")) %>%
  clean_names() %>%
  mutate(across(c(2:4, 6:108), as.numeric)) %>%
  replace(is.na(.), 0)

#Read in 2010 data, clean names, coerce into numerics, replace NA with 0
d_wu_2010 <- read_excel(here("data/us2010.xlsx")) %>%
  clean_names() %>%
  mutate(across(c(2, 4:117), as.numeric)) %>%
  replace(is.na(.), 0)

#Read in 2015 data, skip 1 line, clean names, coerce into numerics, replace NA with 0
d_wu_2015 <- read_excel(skip = 1, here("data/us2015.xlsx")) %>%
  clean_names() %>%
  mutate(across(c(2, 4:141), as.numeric)) %>%
  replace(is.na(.), 0)
```

Step 3: Organize data by sector and FIPS

(Federal Information Processing Standard)

```
#Create wu_1950 with,
#state, public supply, irrigation, rural, industrial, thermoelectric, and year.
#Pivot long with 4 columns state, year, sectors, withdrawals
wu_1950 <- d_wu_1950 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = NA, Industrial = inpt_wgw_fr + inpt_wsw_fr,
         Thermoelectric = NA, Year = 1950, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
```

```
        Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1955 with same procedures as wu_1950
wu_1955 <- d_wu_1955 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
          ir_wsw_fr, Rural = NA, Industrial = inpt_wgw_fr + inpt_wsw_fr,
        Thermoelectric = NA, Year = 1955, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
        Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1960 with same procedures as wu_1950
#1. Original tags are ir_wgw_fr + ir_wsw_fr,
#these are not ideal because the values are 0 when in reality,
#total withdrawals should be > 0.
#2. New column tag is ir_w_fr_to *Caution, the tag includes deliveries
wu_1960 <- d_wu_1960 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_w_fr_to,
        Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
        Industrial = oi_wgw_fr + oi_wsw_fr, Thermoelectric =
          pt_wgw_fr + pt_wsw_fr, Year = 1960, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
        Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1965 with same procedures as wu_1950
wu_1965 <- d_wu_1965 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation =
          ir_wgw_fr + ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr +
          ls_wsw_fr, Industrial = oi_wgw_fr + oi_wsw_fr,
        Thermoelectric = pt_wgw_fr + pt_wsw_fr, Year = 1965, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
        Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1970 with same procedures as wu_1950
wu_1970 <- d_wu_1970 %>%
```

```
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = oi_wgw_fr + oi_wsw_fr, Thermoelectric = pt_wgw_fr +
           pt_wsw_fr, Year = 1970, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
         Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1975 with same procedures as wu_1950
wu_1975 <- d_wu_1975 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = oi_wgw_fr + oi_wsw_fr, Thermoelectric = pt_wgw_fr +
           pt_wsw_fr, Year = 1975, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
         Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1980 with same procedures as wu_1950
wu_1980 <- d_wu_1980 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = oi_wgw_fr + oi_wsw_fr, Thermoelectric = pt_wgw_fr +
           pt_wsw_fr, Year = 1980, State = area) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric,
         Year) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1985 with same procedures as wu_1950
#Starting from here, data is sectioned off into counties.
#Summarize across State (scode) so that state-wide consumption is calculated
wu_1985 <- d_wu_1985 %>%
  mutate("Public Supply" = ps_wgwfr + ps_wswfr, Irrigation = ir_wgwfr +
           ir_wswfr, Rural = do_ssgwf + do_ssswf + ls_gwtot + ls_swtot,
         Industrial = in_wgwfr + in_wswfr + mi_wgwfr + mi_wswfr,
         Thermoelectric = pt_wgwfr + pt_wswfr, State = scode) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
```

```
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 1985) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1990 with same procedures as wu_1985
wu_1990 <- d_wu_1990 %>%
  mutate("Public Supply" = ps_wgwfr + ps_wswfr, Irrigation = ir_wgwfr +
           ir_wswfr, Rural = do_ssgwf + do_ssswf + ls_gwtot + ls_swtot,
         Industrial = in_wgwfr + in_wswfr + mi_wgwfr + mi_wswfr,
         Thermoelectric = pt_wgwfr + pt_wswfr, State = scode) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 1990) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_1995 with same procedures as wu_1985
wu_1995 <- d_wu_1995 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = in_wgw_fr + in_wsw_fr + mi_wgw_fr + mi_wsw_fr,
         Thermoelectric = pt_wgw_fr + pt_wsw_fr, State = state_code) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 1995) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_2000 with same procedures as wu_1985
wu_2000 <- d_wu_2000 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = it_wgw_fr +
           it_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = in_wgw_fr + in_wsw_fr + mi_wgw_fr + mi_wsw_fr,
         Thermoelectric = pt_wgw_fr + pt_wsw_fr, State = statefips) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 2000) %>%
```

```r
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_2005 with same procedures as wu_1985
wu_2005 <- d_wu_2005 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + ls_wgw_fr + ls_wsw_fr,
         Industrial = in_wgw_fr + in_wsw_fr + mi_wgw_fr + mi_wsw_fr,
         Thermoelectric = pt_wgw_fr + pt_wsw_fr, State = statefips) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 2005) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_2010 with same procedures as wu_1985
wu_2010 <- d_wu_2010 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + li_wgw_fr +
           li_wsw_fr, Industrial = in_wgw_fr + in_wsw_fr + mi_wgw_fr +
           mi_wsw_fr, Thermoelectric = pt_wgw_fr + pt_wsw_fr, State = statefips) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 2010) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new data set wu_2015 with same procedures as wu_1985
wu_2015 <- d_wu_2015 %>%
  mutate("Public Supply" = ps_wgw_fr + ps_wsw_fr, Irrigation = ir_wgw_fr +
           ir_wsw_fr, Rural = do_wgw_fr + do_wsw_fr + li_wgw_fr + li_wsw_fr,
         Industrial = in_wgw_fr + in_wsw_fr + mi_wgw_fr + mi_wsw_fr,
         Thermoelectric = pt_wgw_fr + pt_wsw_fr, State = statefips) %>%
  select(State, "Public Supply", Irrigation, Rural, Industrial, Thermoelectric) %>%
  group_by(State) %>%
  summarize(across(c(1:5), sum)) %>%
  mutate(Year = 2015) %>%
  pivot_longer(cols = 2:6, names_to = "Sectors", values_to = "Withdrawals")
```

Step 4: Combine data and organize for plotting

```
#Create new object wu_all that combines all 14 objects
#Then filter out FIPS that do not denote one of the United States
#Will have 3500 observations because it combines all data points from the 14
#data frames besides the non-states.
#This number represents 1950-2015 data for freshwater withdrawals per sector.
wu_all <- rbind(wu_1950, wu_1955, wu_1960, wu_1965, wu_1970, wu_1975, wu_1980,
                wu_1985, wu_1990, wu_1995, wu_2000, wu_2005, wu_2010, wu_2015,
                by = "State") %>%
  filter(!State%in%c("78", "72", "53", "0", "State"))

#Create wu_all_plot to show change in total freshwater withdrawals from sectors
#over time
#replace NA with 0 (some data not available in certain time periods)
#Group by Year and summarize withdrawals
wu_all_plot <- wu_all %>%
  mutate_at(4, as.numeric) %>%
  replace(is.na(.), 0) %>%
  group_by(Year) %>%
  summarize(across(c(3), sum)) %>%
  mutate_at(1, as.numeric)

#Create plot
ggplot() +
  geom_line(data = wu_all_plot, aes(x = Year, y = Withdrawals),
            color = "dodgerblue", size = 2) +
#geom_label adds a withdrawals label to each year
#label = scales::comma(signif(Withdrawals, 3))), size = 1.5)
#makes the values in the label have 3 significant figures, have commas,
#and have 1.5 size font.
  geom_label(data = wu_all_plot, aes(x = Year, y = Withdrawals,
                                     label = scales::comma(signif(Withdrawals,
                                                                  3))),
             size = 1.5) +
  labs(x = "Year", y = "Withdrawals (Mgal/day)",
       caption = "Figure 1: Total Fresh Withdrawals in the USA 1950-2015.
       Plot created by Maxton Lam. Data from USGS.") +
```

```
  theme(axis.text = element_text(color = "grey", size = 8),
        axis.ticks.x = element_blank(),
        axis.ticks.y = element_blank(),
        panel.background = element_blank(),
        panel.grid.major = element_line(size = 0.1, color = "grey"),
        plot.caption = element_text(hjust = 0, size = 10, face = "bold"),
#legend.position = "" removes the legend
        legend.position = "") +
#This code extends the x axis continuously while breaking off the axis in 20
#(more or less) intervals
 scale_x_continuous(breaks = scales::pretty_breaks(n = 20)) +
#Code sets the y axis interval from (0, 380000) and breaks off the axis in 10
#(more or less) intervals
 scale_y_continuous(breaks = scales::pretty_breaks(n = 10),
                    limits = c(0, 380000))
```

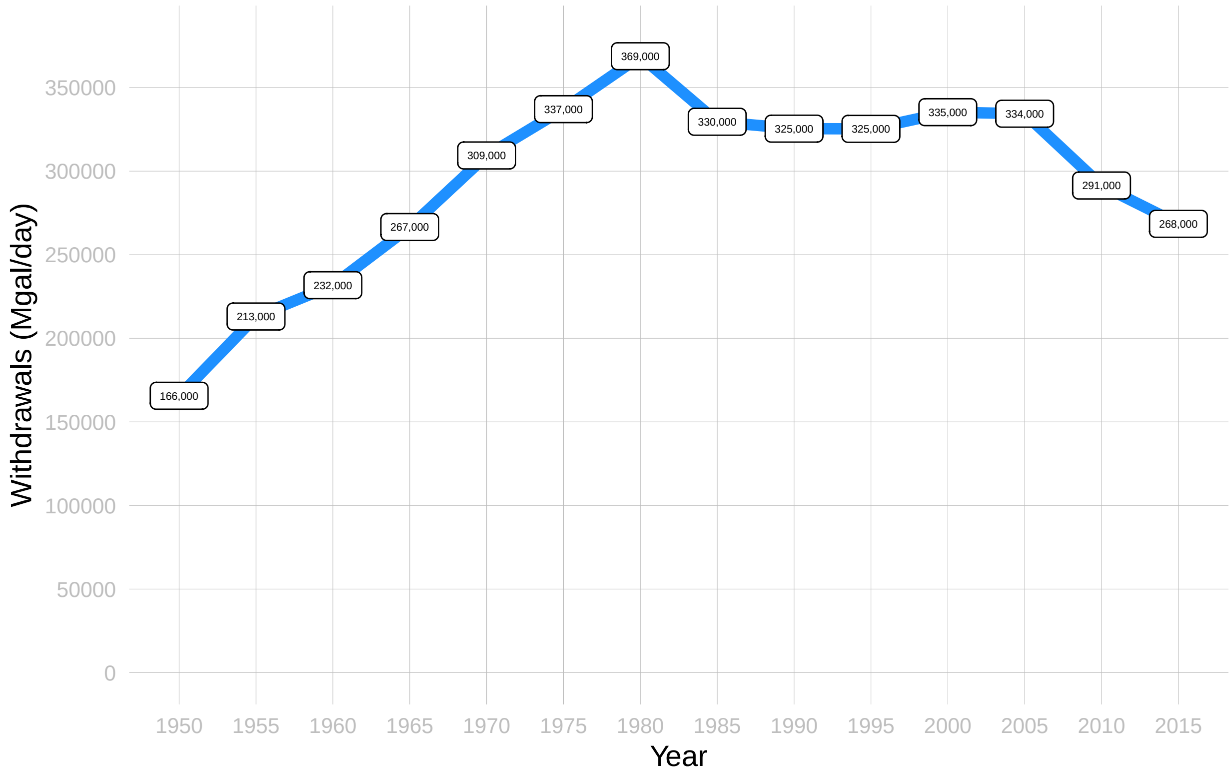

Figure 1: Total Fresh Withdrawals in the USA 1950−2015. Plot created by Maxton Lam. Data from USGS.

Step 5: Update code - implementing feedback from prelab 07 and lab 07

Prelab 07 Code and Data check

1. The code follows best practices.

Lab 07 Code and Data check

1. The code follows best practices.

Step 6: Organize data for plotting timeseries of sectoral withdrawals

```
#Make new object wu_all_sectors from wu_all
#Remove NA and replace with 0
#Group by Sectors and year
#Get sum of withdrawals
wu_all_sectors <- wu_all %>%
  mutate_at(4, as.numeric) %>%
  replace(is.na(.), 0) %>%
  group_by(Sectors, Year) %>%
  summarise(across(c(2), sum)) %>%
  mutate_at(2, as.numeric)

#Test
wu_all_check <- wu_all_sectors %>%
  summarise(across(c(2), sum))

#Create ggplot that combines wu_all_sectors and wu_all_plot
ggplot() +
  geom_col(data = wu_all_sectors, aes(x = Year, y = Withdrawals,
                                      fill = reorder(Sectors, Withdrawals)),
                                      position = position_dodge(3.5),
                                      width = 4) +
  scale_fill_manual(values = c("Turquoise", "dodgerblue3", "Red", "Orange",
                               "Darkgreen")) +
#Need line to indicate total withdrawals across the years
  geom_line(data = wu_all_plot, aes(x = Year, y = Withdrawals/2),
            color = "grey", size = 1) +
#Need point to indicate the year for each data set
  geom_point(data = wu_all_plot, aes(x = Year, y = Withdrawals/2), color = "grey",
             size = 2, fill = "grey") +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 14),
                     expand = c(0,0),) +
  scale_y_continuous(labels = scales::comma,
                     breaks = scales::pretty_breaks(n = 10),
#expand adjusts the bars to the x axis so that it is not hovering above it
                     expand = c(0,0),
                     limits = c(0, 200000),
```

```r
#Trans multiplies the left axis by 2 to produce the right axis.
#This is needed so that the line is appropriately and aesthetically graphed
                     sec.axis = sec_axis(trans = ~.*2,
                                         breaks = scales::breaks_pretty(n = 10),
                                         name = "Total Withdrawals (Mgal/day)",
                                         labels = label_comma())) +
  labs(x = "Year", y = "Sector Withdrawals (Mgal/day)",
       caption = "Figure 2: Fresh Water Withdrawals in the USA 1950-2015.
       Plot created by Maxton Lam. Data from USGS.") +
#This makes the graph more presentable by adding pre-set aesthetics in the code
  theme_few() +
  theme(legend.title = element_blank(),
        legend.position = "top",
        legend.text = element_text(size = 8, color = "black"),
        axis.text = element_text(color = "black", size = 8),
#Use vjust to adjust spacing between axis title and ticks
        axis.title.y = element_text(vjust = 3),
        axis.title.y.right = element_text(color = "grey", vjust = 3),
        axis.text.y.right = element_text(color = "grey"),
        plot.caption = element_text(size = 10, face = "bold", hjust = 0))
```

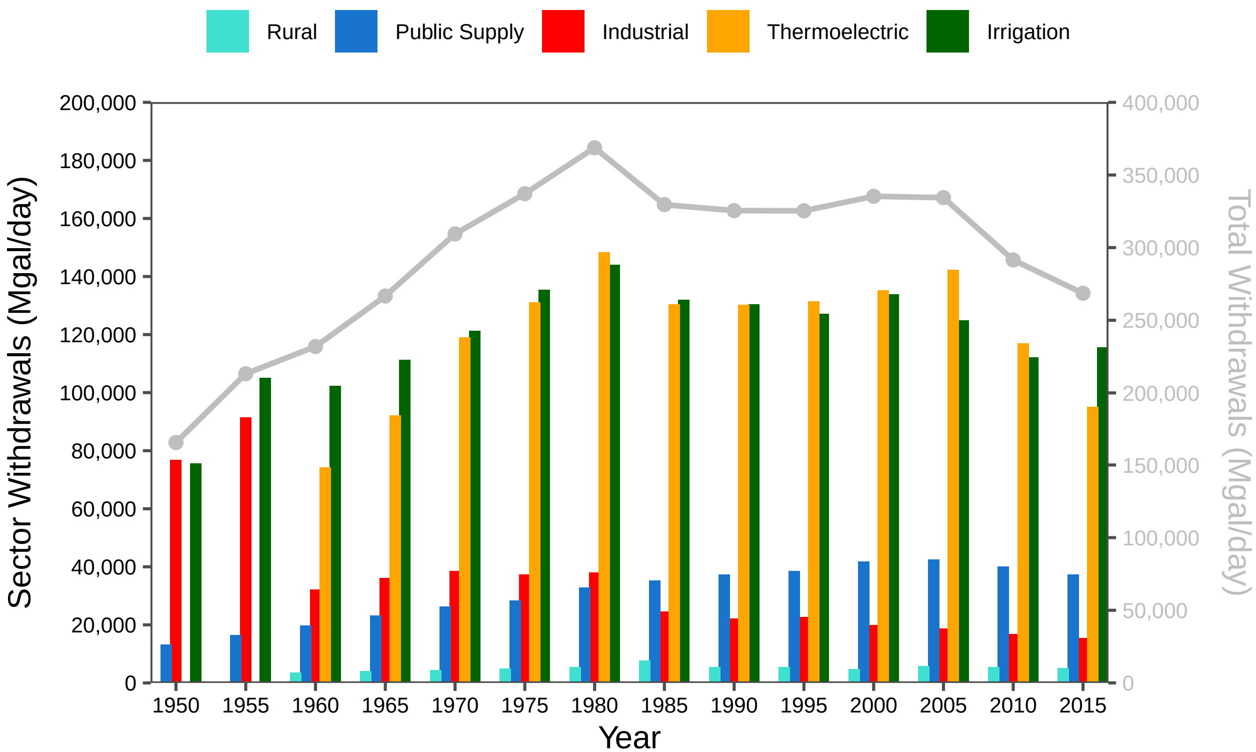

Figure 2: Fresh Water Withdrawals in the USA 1950−2015.
Plot created by Maxton Lam. Data from USGS.

The take-home message from the plot, with respect to water use by sector, over time.
Throughout the years, total freshwater withdrawals continued to until it hit its peak in 1980. After that, total withdrawals have generally decreased each year. This is potentially due to a growing concern over resource conservation in which the country tries to conserve water. This trend can be seen in the irrigation and thermoelectric sector. One difference, however, is that data for the thermoelectric sector in 1950-1955 is not available. These are the two sectors with the highest withdrawals because demand increases with a increase in population. Industrial withdrawals are high from 1950-1955 but significantly decreases from then. This occurs because thermoelectric was once included in industrial withdrawals but was counted separately after 1955. Public supply withdrawals have been steadily increasing over the time period while rural withdrawals have stayed fairly consistent.

My plot's data appear to be different than the USGS plot. The bars in the USGS plot show different patterns than my plot.
The USGS plot shows different patterns because their sectors may include some other types of withdrawals. For example, the "Other" sector in the USGS plot correlates to industrial, mining, and aquaculture. I did not include aquaculture in my "Industrial" sector. The plot also differs in that it separates thermoeletric and industrial from 1950-1955 while my plot combines the two. The USGS plot also uses total withdrawals as compared to

the freshwater withdrawals I used for my plot. This major difference distinguishes between the two plots.

Step 7: Organize data for plotting timeseries of sectoral withdrawals for Arizona

```
#Create same sector plot but for Arizona
az_wu <- wu_all %>%
  filter(State == "4") %>%
  mutate(across(c(1:2, 4), as.numeric)) %>%
  replace(is.na(.), 0) %>%
  group_by(Sectors, Year) %>%
  summarise(across(c(2), sum))

#Create same total plot but for Arizona
az_wu_all <- wu_all %>%
  filter(State == "4") %>%
  mutate(across(c(1:2, 4), as.numeric)) %>%
  replace(is.na(.), 0) %>%
  group_by(Year) %>%
  summarise(across(c(3), sum))

#Plot
ggplot() +
  geom_col(data = az_wu, aes(x = Year, y = Withdrawals,
                                    fill = reorder(Sectors, Withdrawals)),
                                    position = position_dodge(3.5),
                                    width = 4) +
  scale_fill_manual(values = c("Turquoise", "Orange", "Red", "dodgerblue3",
                              "Darkgreen")) +
  geom_line(data = az_wu_all, aes(x = Year, y = Withdrawals),
            color = "grey", size = 1) +
  geom_point(data = az_wu_all, aes(x = Year, y = Withdrawals), color = "grey",
             size = 2, fill = "grey") +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 14),
                     expand = c(0,0),) +
  scale_y_continuous(labels = scales::comma,
                     breaks = scales::pretty_breaks(n = 10),
                     expand = c(0,0),
```

```r
                     limits = c(0, 9000),
#Use a tranformation of 1 because total withdrawals and sector withdrawals data are similar
                     sec.axis = sec_axis(trans = ~.*1, breaks = scales::breaks_pretty(n = 10),
                                         name = "Total Withdrawals (Mgal/day)",
                                         labels = label_comma())) +
  labs(x = "Year", y = "Sector Withdrawals (Mgal/day)",
       caption = "Figure 3: Fresh Water Withdrawals in Arizona 1950-2015.
       Plot created by Maxton Lam. Data from USGS.") +
  theme_few() +
  theme(legend.title = element_blank(),
        legend.position = "top",
        legend.text = element_text(size = 8, color = "black"),
        axis.text = element_text(color = "black", size = 8),
        axis.title.y = element_text(vjust = 3),
        axis.title.y.right = element_text(color = "grey", vjust = 3),
        axis.text.y.right = element_text(color = "grey"),
        plot.caption = element_text(size = 10, face = "bold", hjust = 0))
```

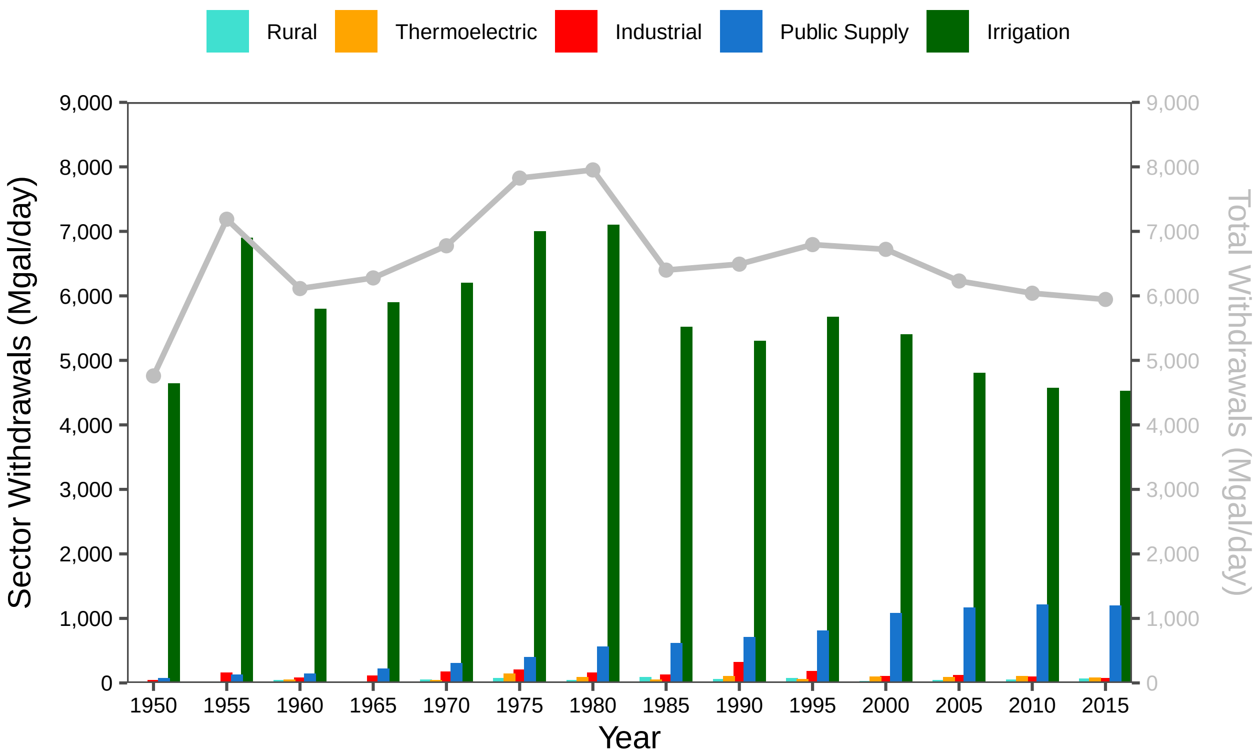

Figure 3: Fresh Water Withdrawals in Arizona 1950−2015. Plot created by Maxton Lam. Data from USGS.

The take-home message from the plot, with respect to water use by sector, over time in Arizona.
The take-home message is that irrigation withdrawals makes up the majority of Arizona's freshwater withdrawals. This makes sense because Arizona typically has a dry climate, so more water is used to ensure plants receive enough water. The next largest sector, public supply, has been continuously increasing over the years. This is due to more people populating the state, which results in higher demand. There is a notable decline in freshwater withdrawals after 1980 which can be attributed to water conservation. Climate change causes many impacts, including drought, so water is conserved and efficiently withdrawn to be more sustainable.

Comparing the USA portfolio and Arizona's portfolio.
A similarity is that both plots show a peak in 1980 and decreases from then. Both plots also indicate that irrigation is one of the largest sectors for freshwater withdrawals. One difference, however, is that the USA plot shows the thermoelectric sector as a large sector but it is almost non-existent in the Arizona plot. Another difference is that the USA total withdrawals gradually increases until 1980, but there are two peaks in Arizona. The first peak is in 1955 and the second peak, which is the maximum, is in 1980.

Step 8: Organize data for plotting timeseries of state fresh water use

```
fips <- d_wu_2015 %>%
  select(state, State = statefips) %>%
  unique() %>%
  filter(!State%in%c(0, 11, 72, 78))

#Public supply
ps_all <- wu_all %>%
  mutate_at(4, as.numeric) %>%
  replace(is.na(.), 0) %>%
  filter(Sectors == "Public Supply") %>%
  mutate_at(1, as.numeric) %>%
  mutate_at(2, as.numeric) %>%
  inner_join(fips, by = "State")

ir_all <- wu_all %>%
  mutate(across(c(1:2, 4), as.numeric)) %>%
  replace(is.na(.), 0) %>%
  filter(Sectors == "Irrigation") %>%
  inner_join(fips, by = "State")

data1 <- c("ps_all", "ir_all")

cap <- c(
"Figure 4: Public supply by state. Plot created by Maxton Lam. Data from USGS (2015).",
"Figure 5: Irrigation by state. Plot created by Maxton Lam. Data from USGS (2015).")

myplot <- function(data1, cap) {
  ggplot() +
    geom_line(data = data1,
              aes(x = Year, y = Withdrawals, color = state), size = 1) +
    scale_color_manual(values = rep(c("darkblue", "royalblue4", "steelblue3", "dodgerblue2", "skyblue"), 10)) +
    gghighlight(max(Withdrawals), max_highlight = 5L,
                unhighlighted_params = list(size = .5, color = alpha("grey", 0.4)),
                label_params = list(size = 3, fill = "white", fontface = "bold",
                                    color = c("darkblue", "royalblue4", "steelblue3", "dodgerblue2", "skyblue"))) +
    scale_x_continuous(breaks = breaks_pretty(n = 15),
```

```
                      expand = c(0, 1)) +
    scale_y_continuous(breaks = breaks_pretty(n = 7),
                      expand = c(0, 0),
                      labels = label_comma()) +
    labs(caption = cap,
         x = "\nYear\n",
         y = "Withdrawals (Mgal/day)\n") +
    theme(axis.text = element_text(color = "black", size = 10),
          panel.background = element_blank(),
          axis.line = element_line(color = "black"),
          plot.caption = element_text(hjust = 0, size = 10, face = "bold"),
          legend.position = "")
}

for(i in seq_along(data1)){
  print(myplot(get(data1[i]), cap[i]))
}
```

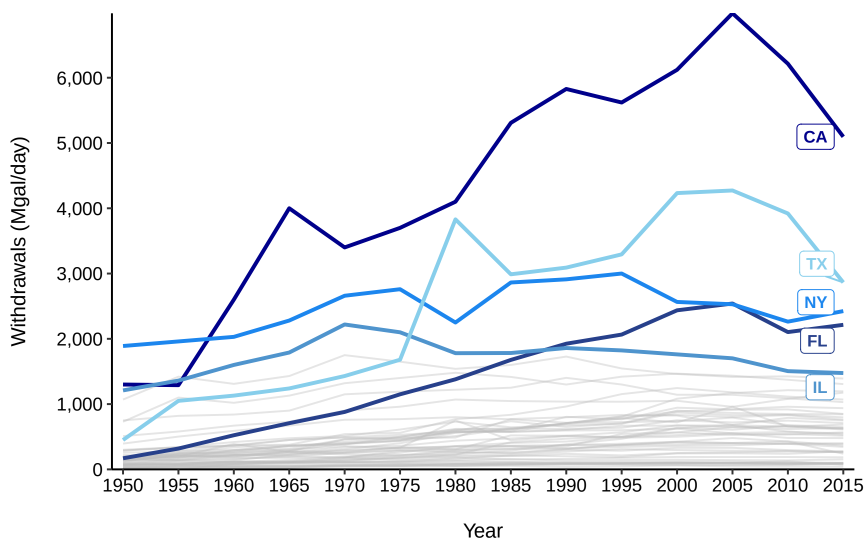

Figure 4: Public supply by state. Plot created by Maxton Lam. Data from USGS (2015).

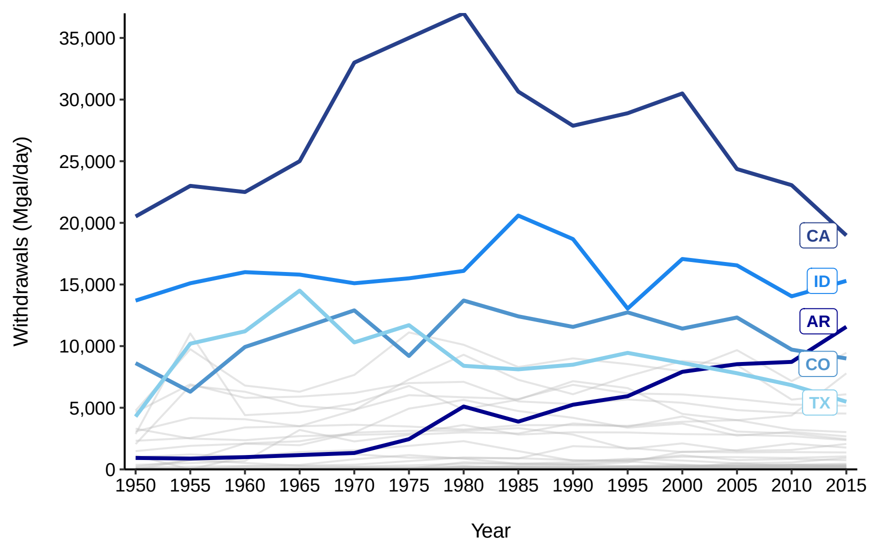

Figure 5: Irrigation by state. Plot created by Maxton Lam. Data from USGS (2015).

The take-home message from the plot.

The take-home message for the public supply plot shows that total withdrawals are now decreasing in big states such as California and Texas. In other states, there has not been a significant change in public supply withdrawals over the years. This decreasing trend can indicate that states are being more efficient with water withdrawals even though population continues to rise. The take-home message for the irrigation plot is that western states tend to withdraw more water for irrigation purposes. In California, Colorado, and Texas, irrigation withdrawals are decreasing due to improved technologies in water conservation.

Step 9: Organize data for plotting timeseries of gw and sw withdrawals for USA

```r
#Create new object source_wu_1965 with 5 new columns: State, Year, Population, Surface Water, Groundwater, and Total
#Pivot long to show sector and withdrawals
source_wu_1965 <- d_wu_1965 %>%
  mutate(State = area, Year = 1965, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + oi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr +
           pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           ls_wgw_fr + oi_wgw_fr + pt_wgw_fr) %>%
  select(State, Year, Population, SW, GW, Total) %>%
  pivot_longer(cols = 4:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1970 with same directions as 1965
source_wu_1970 <- d_wu_1970 %>%
  mutate(State = area, Year = 1970, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + oi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr +
           pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           ls_wgw_fr + oi_wgw_fr + pt_wgw_fr) %>%
  select(State, Year, Population, SW, GW, Total) %>%
  pivot_longer(cols = 4:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1975 with same directions as 1965
source_wu_1975 <- d_wu_1975 %>%
  mutate(State = area, Year = 1975, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + oi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr +
```

```
          pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
          ls_wgw_fr + oi_wgw_fr + pt_wgw_fr) %>%
  select(State, Year, Population, SW, GW, Total) %>%
  pivot_longer(cols = 4:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1980 with same directions as 1965
source_wu_1980 <- d_wu_1980 %>%
  mutate(State = area, Year = 1980, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + oi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + oi_wsw_fr +
           pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           ls_wgw_fr + oi_wgw_fr + pt_wgw_fr) %>%
  select(State, Year, Population, SW, GW, Total) %>%
  pivot_longer(cols = 4:6, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1985 with same directions as 1965
#Need to group by state and then summarize to receive state data
source_wu_1985 <- d_wu_1985 %>%
  mutate(State = scode, Population = po_total,
         SW = ps_wswfr + ir_wswfr + do_ssswf + ls_swtot + in_wswfr + mi_wswfr + pt_wswfr,
         GW = ps_wgwfr + ir_wgwfr + do_ssgwf + ls_gwtot + in_wgwfr + mi_wgwfr + pt_wgwfr,
         Total = ps_wswfr + ir_wswfr + do_ssswf + ls_swtot + in_wswfr + mi_wswfr +
           pt_wswfr + ps_wgwfr + ir_wgwfr + do_ssgwf +
           ls_gwtot + in_wgwfr + mi_wgwfr + pt_wgwfr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 1985) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1990 with same directions as 1985
source_wu_1990 <- d_wu_1990 %>%
  mutate(State = scode, Population = po_total,
         SW = ps_wswfr + ir_wswfr + do_ssswf + ls_swtot + in_wswfr + mi_wswfr + pt_wswfr,
         GW = ps_wgwfr + ir_wgwfr + do_ssgwf + ls_gwtot + in_wgwfr + mi_wgwfr + pt_wgwfr,
         Total = ps_wswfr + ir_wswfr + do_ssswf + ls_swtot + in_wswfr + mi_wswfr +
           pt_wswfr + ps_wgwfr + ir_wgwfr + do_ssgwf +
```

```r
          ls_gwtot + in_wgwfr + mi_wgwfr + pt_wgwfr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 1990) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_1995 with same directions as 1985
source_wu_1995 <- d_wu_1995 %>%
  mutate(State = state_code, Population = total_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr + mi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr +
           mi_wsw_fr + pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 1995) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_2000 with same directions as 1985
source_wu_2000 <- d_wu_2000 %>%
  mutate(State = statefips, Population = tp_tot_pop,
         SW = ps_wsw_fr + it_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr + mi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + it_wgw_fr + do_wgw_fr + ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + it_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr +
           mi_wsw_fr + pt_wsw_fr + ps_wgw_fr + it_wgw_fr + do_wgw_fr +
           ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 2000) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_2005 with same directions as 1985
source_wu_2005 <- d_wu_2005 %>%
  mutate(State = statefips, Population = tp_tot_pop,
```

```
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr + mi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + ls_wsw_fr + in_wsw_fr +
           mi_wsw_fr + pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           ls_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 2005) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_2010 with same directions as 1985
source_wu_2010 <- d_wu_2010 %>%
  mutate(State = statefips, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + li_wsw_fr + in_wsw_fr + mi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + li_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + li_wsw_fr + in_wsw_fr +
           mi_wsw_fr + pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           li_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 2010) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")

#Create new object source_wu_2015 with same directions as 1985
source_wu_2015 <- d_wu_2015 %>%
  mutate(State = statefips, Population = tp_tot_pop,
         SW = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + li_wsw_fr + in_wsw_fr + mi_wsw_fr + pt_wsw_fr,
         GW = ps_wgw_fr + ir_wgw_fr + do_wgw_fr + li_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr,
         Total = ps_wsw_fr + ir_wsw_fr + do_wsw_fr + li_wsw_fr + in_wsw_fr +
           mi_wsw_fr + pt_wsw_fr + ps_wgw_fr + ir_wgw_fr + do_wgw_fr +
           li_wgw_fr + in_wgw_fr + mi_wgw_fr + pt_wgw_fr) %>%
  select(State, Population, SW, GW, Total) %>%
  group_by(State) %>%
  summarise(across(c(1:4), sum)) %>%
  mutate(Year = 2015) %>%
  pivot_longer(cols = 3:5, names_to = "Sectors", values_to = "Withdrawals")
```

```r
#Create wu_all_source to combine above data
#Filter out irrelevant FIPS
wu_all_source <- rbind(source_wu_1965, source_wu_1970, source_wu_1975,
                       source_wu_1980, source_wu_1985, source_wu_1990,
                       source_wu_1995, source_wu_2000, source_wu_2005,
                       source_wu_2010, source_wu_2015, by = "State") %>%
  filter(!State%in%c("78", "72", "53", "0", "State"))

#Create pop_plot from wu_source_all to filter out population data
pop_plot <- wu_all_source %>%
  filter(!Sectors%in%c("SW", "GW")) %>%
  select(State, Year, Population) %>%
  mutate(across(c(1:3), as.numeric)) %>%
  group_by(Year) %>%
  summarise(across(c(2), sum))

#Create source_plot from wu_source_all to filter out source data
source_plot <- wu_all_source %>%
  select(State, Year, Sectors, Withdrawals) %>%
  mutate(across(c(1:2, 4), as.numeric)) %>%
  group_by(Sectors, Year) %>%
  summarise(across(c(2), sum))

#Create plot
ggplot() +
  geom_col(data = source_plot, aes(x = Year, y = Withdrawals, fill = reorder(Sectors, Withdrawals)),
         position = position_dodge(3), width = 3, color = "black") +
  scale_fill_manual(values = c("lightblue", "deepskyblue2", "darkblue")) +
  geom_line(data = pop_plot, aes(x = Year, y = Population, color = "Population"), size = 2) +
  scale_color_manual(name = "", values = c("Population" = "hotpink")) +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 14),
                     expand = c(0,0)) +
  scale_y_continuous(labels = scales::comma,
                     breaks = scales::pretty_breaks(n = 10),
                     expand = c(0,0),
                     limits = c(0, 400000),
                     sec.axis = sec_axis(trans = ~./1000, breaks = scales::pretty_breaks(n = 10),
                                         name = "Population (Millions)")) +
```

```
labs(x = "Year", y = "Withdrawals (Mgal/day)",
     caption = "Figure 6: Trends in population and fresh water withdrawals by source in USA.
     Plot created by Maxton Lam. Data from USGS (1965-2015).") +
theme_few() +
theme(legend.title = element_blank(),
      legend.position = "top",
      legend.text = element_text(size = 10, color = "black"),
      axis.text = element_text(size = 10, color = "black"),
      axis.title.y = element_text(vjust = 3),
      axis.title.y.right = element_text(color = "hotpink", vjust = 3),
      axis.text.y.right = element_text(size = 10, color = "hotpink"),
      plot.caption = element_text(size = 10, face = "bold", hjust = 0))
```

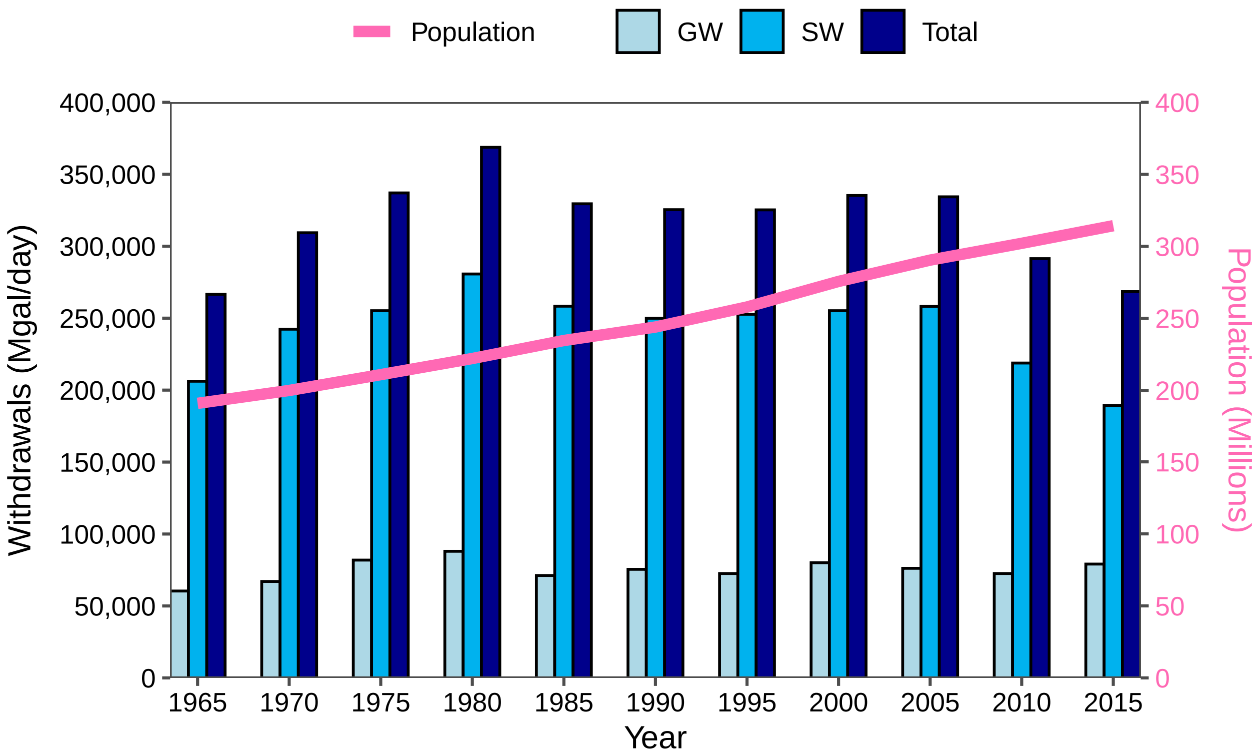

Figure 6: Trends in population and fresh water withdrawals by source in USA. Plot created by Maxton Lam. Data from USGS (1965–2015).

The take-home message from the plot, with respect to water use by water source (GW and SW), over time.
The take-home message is that there is no significant increase in fresh water withdrawals over time. Groundwater and surface water withdrawals appear to peak in 1980 and has been showing a decreasing trend since. This is expected because technologies have improved to the point where society is more efficient with water withdrawals.

The take-home message from the plot, with respect to TOTAL water use and POPULATION over time.
The take-home message is that as population increases, total withdrawals appear to increase and then decrease slightly by year. Total withdrawals peak in 1980 and has been decreasing since. This negative correlation can be attributed to improved technology that makes water withdrawals more efficient. However, it cannot be implied that total withdrawals will continue to decrease if population increases because no matter how efficient water withdrawal is, water will always be used to a certain point.

Step 10: Organize data for plotting timeseries of gw and sw withdrawals for Arizona

```
#Create az_source to plot freshwater withdrawals by source in AZ
az_source <- wu_all_source %>%
  filter(State == "4") %>%
  select(State, Year, Sectors, Withdrawals) %>%
  mutate(across(c(1:2, 4), as.numeric)) %>%
  group_by(Sectors, Year) %>%
  summarise(across(c(2), sum))

#Create az_pop to plot population over time in AZ
az_pop <- wu_all_source %>%
  filter(State == "4") %>%
  filter(!Sectors%in%c("SW", "GW")) %>%
  select(State, Year, Population) %>%
  mutate(across(c(1:3), as.numeric)) %>%
  group_by(Year) %>%
  summarise(across(c(2), sum))

#Create same plot but for Arizona
ggplot() +
  geom_col(data = az_source, aes(x = Year, y = Withdrawals, fill = reorder(Sectors, Withdrawals)),
           position = position_dodge(3), width = 3, color = "black") +
  scale_fill_manual(values = c("deepskyblue2", "lightblue", "darkblue")) +
  geom_line(data = az_pop, aes(x = Year, y = Population, color = "Population"), size = 2) +
  scale_color_manual(name = "", values = c("Population" = "hotpink")) +
  scale_x_continuous(breaks = scales::pretty_breaks(n = 14),
                     expand = c(0,0)) +
  scale_y_continuous(labels = scales::comma,
                     breaks = scales::pretty_breaks(n = 10),
                     expand = c(0,0),
                     limits = c(0, 9000),
                     sec.axis = sec_axis(trans = ~./1000, breaks = scales::pretty_breaks(n = 10),
                                         name = "Population (Millions)")) +
  labs(x = "Year", y = "Withdrawals (Mgal/day)",
       caption = "Figure 7: Trends in population and fresh water withdrawals by source in Arizona.
       Plot created by Maxton Lam. Data from USGS (1965-2015)") +
  theme_few() +
```

```r
  theme(legend.title = element_blank(),
        legend.position = "top",
        legend.text = element_text(size = 10, color = "black"),
        axis.text = element_text(size = 10, color = "black"),
        axis.title.y = element_text(vjust = 3),
        axis.title.y.right = element_text(color = "hotpink", vjust = 3),
        axis.text.y.right = element_text(size = 10, color = "hotpink"),
        plot.caption = element_text(size = 10, face = "bold", hjust = 0))
```

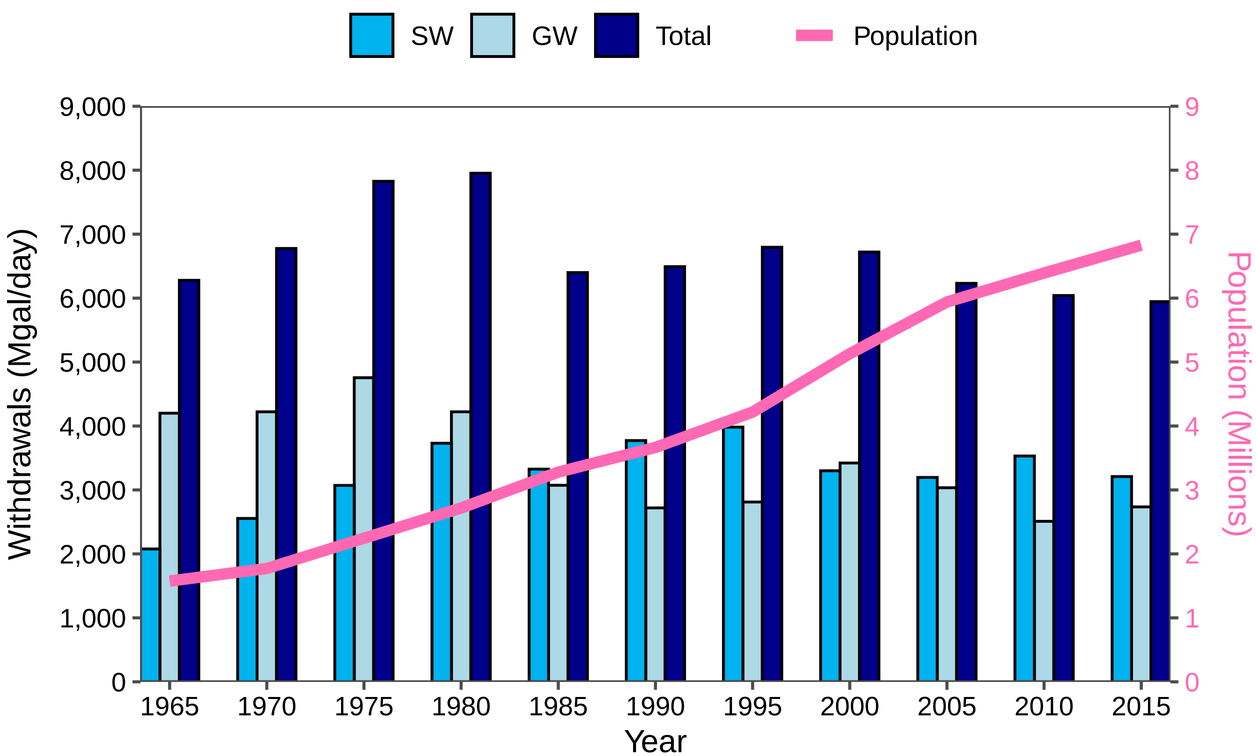

Figure 7: Trends in population and fresh water withdrawals by source in Arizona. Plot created by Maxton Lam. Data from USGS (1965–2015)

The take-home message from the plot, with respect to water use by each water source (GW and SW), over time.
The take-home message is that groundwater withdrawals have decreased since its peak in 1975 while there is no clear trend for surface water. This

decrease in groundwater withdrawals can be attributed to climate change and drought. Surface water withdrawals does not show a clear decreasing tred because Arizona has access to a stable supply in the form of the Colorado River.

The take-home message from the plot, with respect to TOTAL water use and POPULATION over time.
As population increases, total withdrawals also increases until 1980, where total withdrawals significantly decrease and have remained relatively constant since. Though population increases, it can be implied that Arizona has found a way to be more efficient and sustainable with water withdrawals.

Comparing the USA plot to the Arizona plot.
In both plots, total withdrawals peak in 1980 and have decreased since then. Another similarity is that both plots share the same trend with respect to total withdrawals and population over time. One difference is that surface water makes up the majority of total withdrawals in the USA plot but contributes, more or less, the same amount as groundwater for Arizona. Another difference is that the population percent growth in Arizona is much larger than the USA plot.