

Overdose R script

2025-06-30

To start we read in the data and put the tidyverse and other packages in the library. Remember you must install each package once and only once on your machine.

```
# install.packages("tidyverse")
# install.packages("reshape")
# install.packages("RColorBrewer")
# install.packages("scales")

library(tidyverse)
library(reshape)
library(RColorBrewer)
library(scales)
```

First attempt at bars and lines did not have a legend. Figured out how to use scales package to add in legend.

```
odmf <- read.csv("Overdose deaths M_F.csv")
View(odmf)

odbar <- ggplot(data = odmf, show.legend = T) +
  geom_bar(aes(x = Year, y = Total.Overdose.Deaths),
           fill = "blue", stat = "identity") +
  ggtitle("US Overdose Deaths by Sex 1999-2022")

# Now add lines
odbar + geom_line(aes(x = Year, y = Male),
                  color = "yellow", stat = "identity", linewidth = 2,
                  show.legend = T) +
  geom_line(aes(x = Year, y = Female),
            color = "orange", stat = "identity", linewidth = 2,
            show.legend = T)

# To add legend and include totals need another line for total
ggplot(odmf, aes(x = Year)) +
  geom_bar(aes(y = Total.Overdose.Deaths), stat = "identity", fill = "blue") +
  geom_line(aes(y = Total.Overdose.Deaths, color = "Total.Overdose.Deaths"), size = 0.1) +
  geom_line(aes(y = Female, color = "Female"), size = 1.25) +
  geom_line(aes(y = Male, color = "Male"), size = 1.25) +
  scale_color_manual(values = c("Male" = "gold2", "Female" = "orange",
                                "Total.Overdose.Deaths" = "blue")) +
  labs(title = "Figure 1. U.S. Overdose Deaths by Sex, 1999-2022",
       y = "Deaths", x = "Year", color = "Sex") +
  theme_bw() +
```

```
theme(legend.position=c(0.2,0.8)) +
scale_y_continuous(labels = comma)
```

Make the same graphic using the base barplot function and manually create the legend.

```
#try base barplot
#can't add lines onto barplot I think.....ah need x-coords of bars maybe 1-24??
# need to find x-coords of bars

xcoords <- seq(0.75, length = 24, by = 1.2)

barplot(odmf$Total.Overdose.Deaths, names.arg = odf$Year,
       ylab = "Total Overdose Deaths", axes = F, col = "blue",
       main = "US Overdose Deaths by Sex 1999 - 2022",
       cex.names = 0.7, las = 2)
axis(2, at = c(0,20000,40000,60000,80000,100000,120000),
     labels = c("0","20k","40k","60k","80k","100k","120k"), las =2,
     cex.axis = 0.7)
lines(xcoords,odmf$Female, col = "orange", lwd = 2)
points(xcoords,odmf$Female, col = "orange")
lines(xcoords,odmf$Male, col = "yellow", lwd = 2)
points(xcoords,odmf$Male, col = "yellow")
legend(2, 100000, legend = c("Total","Male", "Female"),
      fill = c("blue","yellow", "orange"), cex = 0.8)

#Perfect! This worked!
```

For the second target graphic will need to melt the data.

```
#OK let's try next graphic
oddrug <- read.csv("Overdoses by drug.csv")
View(oddrug)

colnames(oddrug)[2] <- ("Synth2")
colnames(oddrug)[3] <- c("Meth")

odmelt2 <- melt(oddrug, "Year")
View(odmelt2)
dim(odmelt2)
colnames(odmelt2) <- c("Year","Drug","Deaths")

ggplot(odmelt2, aes(x = Year, y = Deaths, color = Drug)) +
  geom_line(linewidth=2) +
  scale_color_brewer(palette = "Set1") +
  theme_bw() +
  theme(legend.position=c(0.2,0.8),legend.justification=c(0.2,0.8)) +
  scale_y_continuous(label = comma, limits = c(0,100000))
```

Next graphic we will use base R plot.

```

#Figure 3 let's try graphing both y-ranges and adding in second axis...
odrates <- read.csv("OD Rates.csv")
View(odrates)
dim(odrates)

plot(odrates$Year, odrates$Total.Overdose.Deaths, col = "black",
     pch = 19, xlab = "Year", main = "Overdose Deaths",
     ylab = "Deaths per 100,000", ylim = c(-5,35), axes = F)
points(odrates$Year, odrates$Percent.Change*100, col = "red",
       pch = 19, cex = 0.75)
lines(odrates$Year, odrates$Total.Overdose.Deaths, col = "black",
      lwd = 2)
lines(odrates$Year, odrates$Percent.Change*100, col = "red",
      lwd = 2)
axis(1,cex.axis = 0.7)
axis(2,cex.axis = 0.7)
axis(4, at = seq(-5, length = 9, by = 5),
      labels = c("-5%", "0%", "5%", "10%", "15%", "20%", "25%", "30%", "35%"), las = 2,
      cex.axis = 0.7)
legend(2000, 35, legend = c("Rate", "Annual % Change"),
      fill = c("black", "red"), cex = 0.8)

```

The basic bar chart showing change will be straightforward.

```

#Now for fourth target graphic.
odchg <- read.csv("Chg 2015 to 22.csv")
View(odchg)
colnames(odchg)[2] <- c("Change")

ggplot() + geom_bar(data = odchg, aes(x = Demographic, y = Change),
                   fill = "blue", stat = "identity") +
  ggtitle("Fold Change 2015-22 in Overdose Death Rates per 100,000")

```

And the fifth graphic will require some spacing in the csv file to get the proper order of bars.

```

od22 <- read.csv("OD rates 2022.csv")
View(od22)

#needed to rename all the labels and fix spacing
od22_3 <- read.csv("OD rates 2022_v2.csv")
View(od22_3)

colvec <- rep(c("black", "hotpink", "navy"), 7)
colvec

ggplot() + geom_bar(data = od22_3, aes(x = Demographic, y = Rate),
                   fill = colvec, color = "white",
                   stat = "identity", width = 0.75) +
  ggtitle("Drug Overdose Death Rates 2022") +
  ylab("Deaths per 100,000") +
  xlab("Demographic Groups Non-Hispanic for each Race") +
  theme(axis.text.x = element_text(angle = 55, hjust = 1, vjust = 1,
                                    size = 8, color = "gray23"))

```

The End!