

02__visualizations

Isabel García Valdivia

4/25/2018

PART II. Data Tables and Plots.

Constructing data tables and plots with new legal status variable (legalstat) in (cps.dta) data set to look at the data numerically and visually.

Load Libraries/Packages

```
library(vcd)           #three-way crosstabs
library(descr)         #weighted data and cross tabs
library(xtable)        #export to LaTeX
library(ggplot2)       #visuals
library(ggthemes)      #changing visual themes
library(gridExtra)     #packages to export plots
library(cowplot)       #same as above
```

Tabulations of the Data: Older Adults by Legal Status and Age

```
#Legal status and age
table1 <- crosstab(cps$age_chr, cps$legalstat_chr, #DV, IV
  weight = cps$asecwt, #weight
  digits = list(prop = 3, percent = 1), #Named list with number of digits after the decimal point
  prop.c = TRUE, #column percents
  xlab = "Legal Status", #x-axis label
  ylab = "Age", #y-axis label
  main = "Old Age by Legal Status", #overall title
  plot = F) #Logical: if TRUE (default), a mosaic plot is produced.
table1
xtable(table1) #export to LaTeX
```

Boxplot: Older Adults by Legal Status and Age

Note: I commented-out the titles of the graphs due for my presentation, but the code is provided here. This is applicable for all the graphs.

```
#Graph 1: boxplot legal status by age
boxplot1 <- ggplot(cps, aes(x= legalstat_chr, y= age_num, fill=legalstat_chr)) +
  geom_boxplot(varwidth = TRUE) #boxplot #varwidth - changes according to the number of observations
  theme_classic() + #theme
  theme(plot.title = element_text(hjust = 0.5),
    legend.position='none',
    legend.title = element_blank()) + #Center title and move legend
  #labs(title = "Distribution of Old Age (50+ years old) in the U.S.",
    labs(y = "Age (years)") + #labels for title and y-axis
  scale_x_discrete(name= "Legal Status", #re-label the x-axis labels
```

```

labels=c("legal immigrants","undocumented", "us-born"))

boxplot1 #check graph
ggsave("g1_boxplot.png",
       boxplot1,
       dpi = 600) #save the file as .png

```

Graphs 1-3: Display graphs of older adults (50+ years) by legal status and gender.

Graph 1: Legal Immigrants by Age and Gender

```

pop_tb1 <- as.data.frame(table(legalfor_df$age_chr, legalfor_df$female_log))
names(pop_tb1)[1] <- paste("Age") #change the label name
names(pop_tb1)[2] <- paste("Female") #change the label name
sum(pop_tb1$Freq) #foreign-born and legal

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g1_legalfor <- ggplot(pop_tb1, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb1, Female == F), stat = "identity", mapping = aes(y = (Freq/8058)*100)) +
  geom_bar(data = subset(pop_tb1, Female == T), stat = "identity", position = "identity", mapping = aes(
  scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
                      name="", labels=c("Men", "Women"))) + #Change the names of legend labels for "fill"
  coord_flip() + # Turn to horizontal barplots
  #ggtitle('Legal Immigrants by Age and Gender (3.8 million)') +
  #theme(plot.title = element_text(hjust = 0.5)) + #Center title
  theme_classic() +
  scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
                    limits = c(-30, 30), # Define/set the limits on scales
                    labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5)))) + # Label the lim
  xlab("Age (years)") +
  ylab("% Within Legal Foreign-Born")

g1_legalfor #check graph
ggsave("g1_legalimmigrants.png",
       g1_legalfor,
       dpi = 600) #save the file as .png

```

Graph 2: Undocumented Immigrants by Age and Gender

```

pop_tb2 <- as.data.frame(table(undocfor_df$age_chr, undocfor_df$female_log))
names(pop_tb2)[1] <- paste("Age") #change the label name
names(pop_tb2)[2] <- paste("Female") #change the label name
sum(pop_tb2$Freq) #foreign-born and undoc

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g2_undocfor <- ggplot(pop_tb2, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb2, Female == F), stat = "identity", mapping = aes(y = (Freq/1075)*100)) +

```

```

geom_bar(data = subset(pop_tb2, Female == T), stat = "identity", position = "identity", mapping = aes
scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
                    name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
coord_flip() + # Turn to horizontal barplots
#ggtitle('Undocumented Immigrants by Age and Gender (1.8 million)') +
#theme(plot.title = element_text(hjust = 0.5)) + #Center title
theme_classic() +
scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
                    limits = c(-30, 30), # Define/set the limits on scales
                    labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5))))) + # Label the lim
xlab("Age (years)") +
ylab("% Within Undocumented Foreign-Born")

g2_undocfor #check graph
ggsave("g2_undocimmigrants.png",
       g2_undocfor,
       dpi = 600) #save the file as .png

```

Graph 3: U.S. Born Immigrants by Age and Gender

```

pop_tb3 <- as.data.frame(table(usborn_df$Age_chr, usborn_df$female_log))
names(pop_tb3)[1] <- paste("Age") #change the label name
names(pop_tb3)[2] <- paste("Female") #change the label name
sum(pop_tb3$Freq) #usborn

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g3_usborn <- ggplot(pop_tb3, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb3, Female == F), stat = "identity", mapping = aes(y = (Freq/49323)*100))
  geom_bar(data = subset(pop_tb3, Female == T), stat = "identity", position = "identity", mapping = aes
  scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
                    name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
  coord_flip() + # Turn to horizontal barplots
  #ggtitle('U.S. Born by Age and Gender (95.6 million)') +
  #theme(plot.title = element_text(hjust = 0.5)) + #Center title
  theme_classic() +
  scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
                    limits = c(-30, 30), # Define/set the limits on scales
                    labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5))))) + # Label the lim
  xlab("Age (years)") +
  ylab("% Within U.S. Born")

g3_usborn#check graph
ggsave("g3_usborn.png",
       g3_usborn,
       dpi = 600) #save the file as .png

```

Export the tables to a single .pdf file

```
#Export Graphs 1-3 titled olderadultsinUS.png
olderadultsinUS <- plot_grid(g1_legalfor, g2_undocfor, g3_usborn, align = "v", nrow = 3, rel_heights = c(1/2, 1/2))
png(file = "olderadultsinUS.png", width = 1500, height = 1000, res = 250)
plot_grid(g1_legalfor, g2_undocfor, g3_usborn, align = "v", nrow = 3, rel_heights = c(1/2, 1/2))
dev.off() #end save
```

Graphs 4-6: Display graphs of Mexican older adults (50+ years) by legal status and gender.

Graph 4: Mexican U.S. Born Immigrants by Age and Gender

```
pop_tb4 <- as.data.frame(table(usbornmex_df$age_chr, usbornmex_df$female_log))
names(pop_tb4)[1] <- paste("Age") #change the label name
names(pop_tb4)[2] <- paste("Female") #change the label name
sum(pop_tb4$Freq) #usborn mex

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g4_usbornmex <- ggplot(pop_tb4, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb4, Female == F), stat = "identity", mapping = aes(y = (Freq/2033)*100)) +
  geom_bar(data = subset(pop_tb4, Female == T), stat = "identity", position = "identity", mapping = aes(y = (Freq/2033)*100)) +
  scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
    name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
  coord_flip() + # Turn to horizontal barplots
  #ggtitle('U.S. Born Mexicans by Age and Gender (2.9 million)') +
  #theme(plot.title = element_text(hjust = 0.5)) + #Center title
  theme_classic() +
  scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
    limits = c(-30, 30), # Define/set the limits on scales
    labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5)))) + # Label the limits
  xlab("Age (years)") +
  ylab("% Within U.S. Born Mexicans")

g4_usbornmex #check graph
ggsave("g4_usbornmex.png",
  g4_usbornmex,
  dpi = 600) #save the file as .png
```

Graph 5: Mexican Legal Immigrants by Age and Gender

```
pop_tb5 <- as.data.frame(table(legalmex_df$age_chr, legalmex_df$female_log))
names(pop_tb5)[1] <- paste("Age") #change the label name
names(pop_tb5)[2] <- paste("Female") #change the label name
sum(pop_tb5$Freq) #usborn mex

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g5_legalmex <- ggplot(pop_tb5, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb5, Female == F), stat = "identity", mapping = aes(y = (Freq/1955)*100)) +
  geom_bar(data = subset(pop_tb5, Female == T), stat = "identity", position = "identity", mapping = aes(y = (Freq/1955)*100)) +
  scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
    name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
  coord_flip() + # Turn to horizontal barplots
  #ggtitle('U.S. Born Mexicans by Age and Gender (2.9 million)') +
  #theme(plot.title = element_text(hjust = 0.5)) + #Center title
  theme_classic() +
  scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
    limits = c(-30, 30), # Define/set the limits on scales
    labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5)))) + # Label the limits
  xlab("Age (years)") +
  ylab("% Within U.S. Born Mexicans")

g5_legalmex #check graph
ggsave("g5_legalmex.png",
  g5_legalmex,
  dpi = 600) #save the file as .png
```

```

scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
                  name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
coord_flip() + # Turn to horizontal barplots
#ggtitle('Legal Mexican Immigrants by Age and Gender (3 million)') +
#theme(plot.title = element_text(hjust = 0.5)) + #Center title
theme_classic() +
scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
                  limits = c(-30, 30), # Define/set the limits on scales
                  labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5))))) + # Label the lim
xlab("Age (years)") +
ylab("% Within Legal Mexican Immigrants")

g5_legal_mex #check graph
ggsave("g5_legal_mex.png",
      g5_legal_mex,
      dpi = 600) #save the file as .png

```

Graph 6: Mexican Undocumented Immigrants by Age and Gender

```

pop_tb6 <- as.data.frame(table(undocmex_df$age_chr, undocmex_df$female_log))
names(pop_tb6)[1] <- paste("Age") #change the label name
names(pop_tb6)[2] <- paste("Female") #change the label name
sum(pop_tb6$Freq) #477 undoc

#Plot Age Range with respective frequencies and label with mexican_bpl by age
g6_undocmex <- ggplot(pop_tb6, aes(x = Age, y = Freq, fill = Female)) +
  # Define each legal status in barplot distribution
  geom_bar(data = subset(pop_tb6, Female == F), stat = "identity", mapping = aes(y = (Freq/477)*100)) +
  geom_bar(data = subset(pop_tb6, Female == T), stat = "identity", position = "identity", mapping = aes
  scale_fill_manual(values=c("#CC6666", "#9999CC", "#66CC99"), #Change colors
                    name="", labels=c("Men", "Women")) + #Change the names of legend labels for "fill"
coord_flip() + # Turn to horizontal barplots
#ggtitle('Undocumented Mexican Immigrants by Age and Gender (700 Thousands)') +
#theme(plot.title = element_text(hjust = 0.5)) + #Center title
theme_classic() +
scale_y_continuous(breaks = seq(-30, 30, 5), # Set interval for breaks for y axis
                  limits = c(-30, 30), # Define/set the limits on scales
                  labels = paste0(as.character(c(seq(30, 0, -5), seq(5, 30, 5))))) + # Label the lim
xlab("Age (years)") +
ylab("% Within Undocumented Mexicans")

g6_undocmex #check graph
ggsave("g6_undocmex.png",
      g6_undocmex,
      dpi = 600) #save the file as .png

```

Export the tables to a single .pdf file

```

#Export Graphs 4-6 titled olderadultsinUS.png
mexolderadultsinUS <- plot_grid(g5_legal_mex, g6_undocmex, g4_usbornmex, align = "v", nrow = 3, rel_heights = c(1, 1, 1))
png(file = "mexolderadultsinUS.png", width = 1500, height = 1000, res = 250)

```

```
plot_grid(g5_legal_mex, g6_undoc_mex, g4_usborn_mex, align = "v", nrow = 3, rel_heights = c(1/2, 1/2))  
dev.off() #end save
```