```
---
title: "state-change"
output: html_document
---

## Load libs
```{r, echo=FALSE, warning=FALSE}
library(mapproj)
library(choroplethrMaps)
library(ggplot2)
library(dplyr)
library(choroplethr)
library(stringr)
library(tidyr)
library(purrr)
library(readr)
```

Load data
```{r}
source('H:/2014-working-files/parse-widows/combine.r')
```

Clean data
```{r}
# helper functions
clean <- function(x) { as.numeric(gsub(" .*|\\(|\\)", "", as.character(x))) }
makeDate <- function(x) { as.Date(as.character(x), format = "%Y%m%d") }

### drop descriptions on state/county codes
admits$state_fips <- clean(admits$STATE_CODE)
admits$county_fips <- clean(admits$PRISON_COUNTY)
admits <- admits %>% mutate(stateFIPS = floor(county_fips/1000))


### sub_dates
admits$addate <- makeDate(admits$PRISON_ADMISSION_DATE)
admits$adyear <- format(admits$addate, "%Y")

# groups offenses codes as drugs (or not)
admits$BJS_OFFENSE_1 <- clean(admits$BJS_OFFENSE_1)
admits$BJS_OFFENSE_2 <- clean(admits$BJS_OFFENSE_2)
admits$BJS_OFFENSE_3 <- clean(admits$BJS_OFFENSE_3)

admits <- admits %>% mutate(type1 = 340 <= BJS_OFFENSE_1 & BJS_OFFENSE_1 <= 450)
admits <- admits %>% mutate(type2 = 340 <= BJS_OFFENSE_2 & BJS_OFFENSE_2 <= 450)
admits <- admits %>% mutate(type3 = 340 <= BJS_OFFENSE_3 & BJS_OFFENSE_3 <= 450)

admits <- admits %>% mutate(drugs = type1 | type2 | type3)
admits <- admits %>% mutate(drugs = drugs & !is.na(drugs))
```

add in state exceptions

NC and SC include people with six month sentances
```{r}
```
```

```
admits <- admits %>%
  filter(stateFIPS != 37 | OFFENSE_1_SENTENCE > 11) %>%
  filter(stateFIPS != 45 | OFFENSE_1_SENTENCE > 11)
```

Load correct population numbers for 06, 13 and 14
```{r}
pop2010 = read_csv('H:/5-24-2016/pop2010.csv') %>%
  filter(AGEGRP == 0) %>%
  filter(YEAR == 6 | YEAR == 7) %>%
  transmute(region = paste(STATE, COUNTY, sep=''), population = TOT_POP, year = paste('population', YEAR + 2007, sep='')) %>%
  spread(year, population)

pop2000 = read_csv('H:/5-24-2016/pop2000.csv')  %>%
  filter(SEX == 0 & ORIGIN == 0 & RACE == 0) %>%
  transmute(region = paste(STATE, COUNTY, sep=''), population2006 = POPESTIMATE2006)


pop = full_join(pop2000, pop2010, by='region') %>% mutate(region = as.double(region))
```

Look at change since 2006
```{r}
admit2014 <- subset(admits, adyear == '2014')
admit2013 <- subset(admits, adyear == '2013')
admit2006 <- subset(admits, adyear == '2006')

data(df_county_demographics)
df_county_demographics <- df_county_demographics %>%
  full_join(pop, by='region') %>%
  mutate(region = ifelse(region == 12086, 12025, region)) ## FIPS updated in 90s; not changed in NCRP data

countPerCounty2014 <- admit2014 %>% group_by(county_fips) %>%
  summarise(admissions2014 = length(drugs), percentDrugs2014 = mean(drugs), state2014 = first(STATE_CODE), state_fips2014 = first(state_fips)) %>%
  rename(region = county_fips)

counties <- left_join(df_county_demographics, countPerCounty2014, by='region')
counties$perCapitaAdmissions2014 <- counties$admissions2014/counties$population2014
counties$perCapitaDrugs2014 <- counties$percentDrugs*counties$perCapitaAdmissions

#2013
countPerCounty2013 <- admit2013 %>% group_by(county_fips) %>%
  summarise(admissions2013 = length(drugs), percentDrugs2013 = mean(drugs), state2013 = first(STATE_CODE), state_fips2013 = first(state_fips)) %>%
  rename(region = county_fips)

counties <- left_join(counties, countPerCounty2013, by='region')
counties$perCapitaAdmissions2013 <- counties$admissions2013/counties$population2013
counties$perCapitaDrugs2013 <- counties$percentDrugs2013*counties$perCapitaAdmissions2013

#2006
countPerCounty2006 <- admit2006 %>% group_by(county_fips) %>%
  summarise(admissions2006 = length(drugs), percentDrugs2006 = mean(drugs), state2006 = first(STATE_CODE), state_fips2006 = first(state_fips)) %>%
```

```
  rename(region = county_fips)

counties <- left_join(counties, countPerCounty2006, by='region')
counties$perCapitaAdmissions2006 <- counties$admissions2006/counties$population2006
counties$perCapitaDrugs2006 <- counties$percentDrugs2006*counties$perCapitaAdmissions2006

counties = counties %>% mutate(stateFIPS = floor(region/1000))
```

Overwrite ohio2006
```{r}
ohio2006Counties = read_csv('H:/5-24-2016/ohio2006.csv') %>% transmute(region = region, ohioAdmissions06 = ohioAdmissions06)

counties = counties %>%
  #select(-ohioAdmissions06) %>%
  left_join(ohio2006Counties, by='region') %>%
  mutate(admissions2006     = ifelse(stateFIPS == 39, ohioAdmissions06, admissions2006),
         perCapitaDrugs2006 = ifelse(stateFIPS == 39, NA, perCapitaDrugs2006),
         perCapitaAdmissions2006 = admissions2006/total_population)
```

Create data for export
```{r}
exportCounties <- counties %>%
  mutate(admissions2006          = ifelse(admissions2006 >= 5, admissions2006, NA),
         admissions2013          = ifelse(admissions2013 >= 5, admissions2013, NA),
         admissions2014          = ifelse(admissions2014 >= 5, admissions2014, NA),

         perCapitaAdmissions2006 = ifelse(admissions2006 >= 5, perCapitaAdmissions2006, NA),
         perCapitaAdmissions2013 = ifelse(admissions2013 >= 5, perCapitaAdmissions2013, NA),
         perCapitaAdmissions2014 = ifelse(admissions2014 >= 5, perCapitaAdmissions2014, NA),


         percentDrugs2006 = ifelse(percentDrugs2006*admissions2006 >= 5, percentDrugs2006, NA),
         percentDrugs2013 = ifelse(percentDrugs2013*admissions2013 >= 5, percentDrugs2013, NA),
         percentDrugs2014 = ifelse(percentDrugs2014*admissions2014 >= 5, percentDrugs2014, NA),

         perCapitaDrugs2006 = ifelse(percentDrugs2006*admissions2006 >= 5, perCapitaDrugs2006, NA),
         perCapitaDrugs2013 = ifelse(percentDrugs2013*admissions2013 >= 5, perCapitaDrugs2013, NA),
         perCapitaDrugs2014 = ifelse(percentDrugs2014*admissions2014 >= 5, perCapitaDrugs2014, NA))


write_csv(exportCounties, 'H:/export_5-24-2016/export-counties.csv')
```

Join counties with last of valid states/years
```{r}
```

```
validStates = read_csv('H:/2014-working-files/state-change/valid-state.csv')
counties = counties %>%
  #select(-validState2006) %>% select(-validState2013) %>% select(-validState2014) %>%
  left_join(validStates, by='stateFIPS')
```

Slope chart data

```{r}
validCounties = filter(counties, validState2006 & validState2013)

tempFct = transform(validCounties,
                    fctpop=cut(total_population, c(0, 100000, 300000, 10000000)),
                    fctadmit=cut(perCapitaAdmissions2014, c(0, .00125, .0025, .1)),
                    state = ifelse(stateFIPS == 36 | stateFIPS == 48 | stateFIPS == 6, state2014, 'other'),
                    fctdrug=cut(perCapitaDrugs2014, c(0, .0005, .0015, .1))) %>%
  filter(!is.na(total_population))
```