```r
1   ---
2   title: "state-change"
3   output: html_document
4   ---
5
6   ## Load libs
7   ```{r, echo=FALSE, warning=FALSE}
8   library(mapproj)
9   library(choroplethrMaps)
10  library(ggplot2)
11  library(dplyr)
12  library(choroplethr)
13  library(stringr)
14  library(tidyr)
15  library(purrr)
16  library(readr)
17  ```
18
19  ## Load data
20  ```{r}
21  source('H:/2014-working-files/parse-widows/combine.r')
22  ```
23
24  ## Clean data
25  ```{r}
26  # helper functions
27  clean <- function(x) { as.numeric(gsub(" .*|\\(|\\)", "", as.character(x))) }
28  makeDate <- function(x) { as.Date(as.character(x), format = "%Y%m%d") }
29
30  ### drop descriptions on state/county codes
31  admits$state_fips <- clean(admits$STATE_CODE)
32  admits$county_fips <- clean(admits$PRISON_COUNTY)
33  admits <- admits %>% mutate(stateFIPS = floor(county_fips/1000))
34
35
36  ### sub_dates
37  admits$addate <- makeDate(admits$PRISON_ADMISSION_DATE)
38  admits$adyear <- format(admits$addate, "%Y")
39
40  # groups offenses codes as drugs (or not)
41  admits$BJS_OFFENSE_1 <- clean(admits$BJS_OFFENSE_1)
42  admits$BJS_OFFENSE_2 <- clean(admits$BJS_OFFENSE_2)
43  admits$BJS_OFFENSE_3 <- clean(admits$BJS_OFFENSE_3)
44
45  admits <- admits %>% mutate(type1 = 340 <= BJS_OFFENSE_1 & BJS_OFFENSE_1 <= 450)
46  admits <- admits %>% mutate(type2 = 340 <= BJS_OFFENSE_2 & BJS_OFFENSE_2 <= 450)
47  admits <- admits %>% mutate(type3 = 340 <= BJS_OFFENSE_3 & BJS_OFFENSE_3 <= 450)
48
49  admits <- admits %>% mutate(drugs = type1 | type2 | type3)
50  admits <- admits %>% mutate(drugs = drugs & !is.na(drugs))
51  ```
52
53  ## add in state exceptions
54
55  ## NC and SC include people with six month sentences
56  ```{r}
```

```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)) %>%
```

```
109      rename(region = county_fips)
110
111  counties <- left_join(counties, countPerCounty2006, by='region')
112  counties$perCapitaAdmissions2006 <- counties$admissions2006/counties$population2006
113  counties$perCapitaDrugs2006 <- counties$percentDrugs2006*counties$perCapitaAdmissions2006
114
115  counties = counties %>% mutate(stateFIPS = floor(region/1000))
116  ```
117
118  ## Overwrite ohio2006
119  ```{r}
120  ohio2006Counties = read_csv('H:/5-24-2016/ohio2006.csv') %>% transmute(region = region,
     ohioAdmissions06 = ohioAdmissions06)
121
122  counties = counties %>%
123    #select(-ohioAdmissions06) %>%
124    left_join(ohio2006Counties, by='region') %>%
125    mutate(admissions2006     = ifelse(stateFIPS == 39, ohioAdmissions06, admissions2006),
126           perCapitaDrugs2006 = ifelse(stateFIPS == 39, NA, perCapitaDrugs2006),
127           perCapitaAdmissions2006 = admissions2006/total_population)
128  ```
129
130  ## Create data for export
131  ```{r}
132  exportCounties <- counties %>%
133    mutate(admissions2006         = ifelse(admissions2006 >= 5, admissions2006, NA),
134           admissions2013         = ifelse(admissions2013 >= 5, admissions2013, NA),
135           admissions2014         = ifelse(admissions2014 >= 5, admissions2014, NA),
136
137           perCapitaAdmissions2006 = ifelse(admissions2006 >= 5, perCapitaAdmissions2006,
     NA),
138           perCapitaAdmissions2013 = ifelse(admissions2013 >= 5, perCapitaAdmissions2013,
     NA),
139           perCapitaAdmissions2014 = ifelse(admissions2014 >= 5, perCapitaAdmissions2014,
     NA),
140
141
142           percentDrugs2006 = ifelse(percentDrugs2006*admissions2006 >= 5,
     percentDrugs2006, NA),
143           percentDrugs2013 = ifelse(percentDrugs2013*admissions2013 >= 5,
     percentDrugs2013, NA),
144           percentDrugs2014 = ifelse(percentDrugs2014*admissions2014 >= 5,
     percentDrugs2014, NA),
145
146           perCapitaDrugs2006 = ifelse(percentDrugs2006*admissions2006 >= 5,
     perCapitaDrugs2006, NA),
147           perCapitaDrugs2013 = ifelse(percentDrugs2013*admissions2013 >= 5,
     perCapitaDrugs2013, NA),
148           perCapitaDrugs2014 = ifelse(percentDrugs2014*admissions2014 >= 5,
     perCapitaDrugs2014, NA))
149
150
151  write_csv(exportCounties, 'H:/export_5-24-2016/export-counties.csv')
152  ```
153
154  ## Join counties with last of valid states/years
155  ```{r}
```

```r
156  validStates = read_csv('H:/2014-working-files/state-change/valid-state.csv')
157  counties = counties %>%
158    #select(-validState2006) %>% select(-validState2013) %>% select(-validState2014) %>%
159    left_join(validStates, by='stateFIPS')
160  ```
161
162  ## Slope chart data
163  ```{r}
164  validCounties = filter(counties, validState2006 & validState2013)
165
166  tempFct = transform(validCounties,
167                      fctpop=cut(total_population, c(0, 100000, 300000, 10000000)),
168                      fctadmit=cut(perCapitaAdmissions2014, c(0, .00125, .0025, .1)),
169                      state = ifelse(stateFIPS == 36 | stateFIPS == 48 | stateFIPS == 6,
     state2014, 'other'),
170                      fctdrug=cut(perCapitaDrugs2014, c(0, .0005, .0015, .1))) %>%
171    filter(!is.na(total_population))
172  ```
173
```