

Groups_Milestone4

PhuongTseng & Carri Beshears

2022-11-22

Scenario 1: Hospital Funding and Health Equity

You are a researcher in the California Department of Public Health Office of Health Equity (OHE). A policy has just been created to fund a public-private partnership for healthcare facility improvement in rural areas of California that have received minimal funding from the Department of Health Care Access and Information (HCAI) over the past 5 years. You are tasked with exploring and evaluating which 5 counties are the best targets for the development fund proposals. There are multiple components to this request.

First, OHE would like you to focus on rural areas, non-homeowners, and aging individuals as populations of interest in your analysis. Your task is to explore the California county census demographic dataLinks to an external site and begin to identify counties that share three common attributes:

- 1) low population per square mile `pop12_sqmi1`,
According to the U.S. Census on population density <https://www.census.gov/newsroom/blogs/random-samplings/2015/03/understanding-population-density.html>, we will define low population per square mile as less 100 population per sq mile.
- 2) high median age `med_age`,
According to <https://www.ppic.org/publication/californias-population/> and the U.S. Census, California's median age in 2020 is 37.3, therefore, anything above 37 will be considered as above the median age.
- 3) a high proportion of renters vs. homeowners (you may need to create a new variable for this third criteria). We're defining high proportion of renters vs. homeowners as renters over the total population of renters and owners occupied household

```
ca_county_demographic2<- ca_county_demographic %>%
  mutate("rent_pro" = ca_county_demographic$renter_occ/
          ca_county_demographic$households *100)

ca_county_mortality2 <- ca_county_mortality %>%
  filter(Strata_Name == "Total Population") %>%
  filter(Cause_Desc == "All causes (total)") %>%
  filter(Year == 2020) %>% filter(Geography_Type == "Occurrence") %>%
  rename("name" = "County") %>% select(name, Count)

ca_county_dem_mort <- merge(ca_county_demographic2, ca_county_mortality2, by = "name")

hcai_closure <- hcai_healthcare_construction %>%
  filter(`OSHPD Project Status` == "In Closure")

hcai_closure <- hcai_closure %>% filter(rownames(hcai_closure) == 13341:13398)
hcai_closure <- hcai_closure %>% select(County, `Total Costs of OSHPD Projects`)
hcai_closure$name <- c("Alameda", "Alpine", "Amador", "Butte", "Calaveras",
  "Colusa", "Contra Costa", "Del Norte", "El Dorado",
  "Fresno", "Glenn", "Humboldt", "Imperial", "Inyo", "Kern",
  "Kings", "Lake",
  "Lassen", "Los Angeles", "Madera", "Marin", "Mariposa",
  "Mendocino", "Merced",
  "Modoc", "Mono",
  "Monterey", "Napa", "Nevada", "Orange", "Placer", "Plumas",
  "Riverside", "Sacramento", "San Benito", "San Bernardino",
  "San Diego", "San Francisco", "San Joaquin",
  "San Luis Obispo", "San Mateo", "Santa Barbara",
  "Santa Clara", "Santa Cruz", "Shasta",
  "Sierra", "Siskiyou", "Solano", "Sonoma", "Stanislaus", "Sutter", "Tehama", "Trinity",
  "Tulare", "Tuolumne", "Ventura", "Yolo", "Yuba")
hcai_closure <- hcai_closure %>% select(name, `Total Costs of OSHPD Projects`)

ca_county_dem_mort_prog <- merge(ca_county_dem_mort, hcai_closure, by = "name")
ca_county_dem_mort_prog <- ca_county_dem_mort_prog %>%
```

```

mutate("Mortality Rate" = ca_county_dem_mort_prog$Count/
      ca_county_dem_mort_prog$pop2012*100)

working_dataset <- ca_county_dem_mort_prog %>% filter(pop12_sqmi < 100) %>%
  filter(med_age > 37) %>% filter(rent_pro > 35) %>%
  select(name, pop12_sqmi, med_age, rent_pro, `Mortality Rate`,
        `Total Costs of OSHPD Projects`)

working_dataset <- working_dataset %>%
  rename("County Name" = "name", "Pop/mi^2" = "pop12_sqmi")
working_dataset <- working_dataset %>%
  rename("Median Age" = "med_age", "Renter Proportion" = "rent_pro")

```

Visualizations (minimum 2) -One print quality table as requested in scenario -One print quality plot or chart as requested in scenario -Label visual as needed for presentation -1-2 sentence interpretation

```
knitr::kable(working_dataset, digits = 2)
```

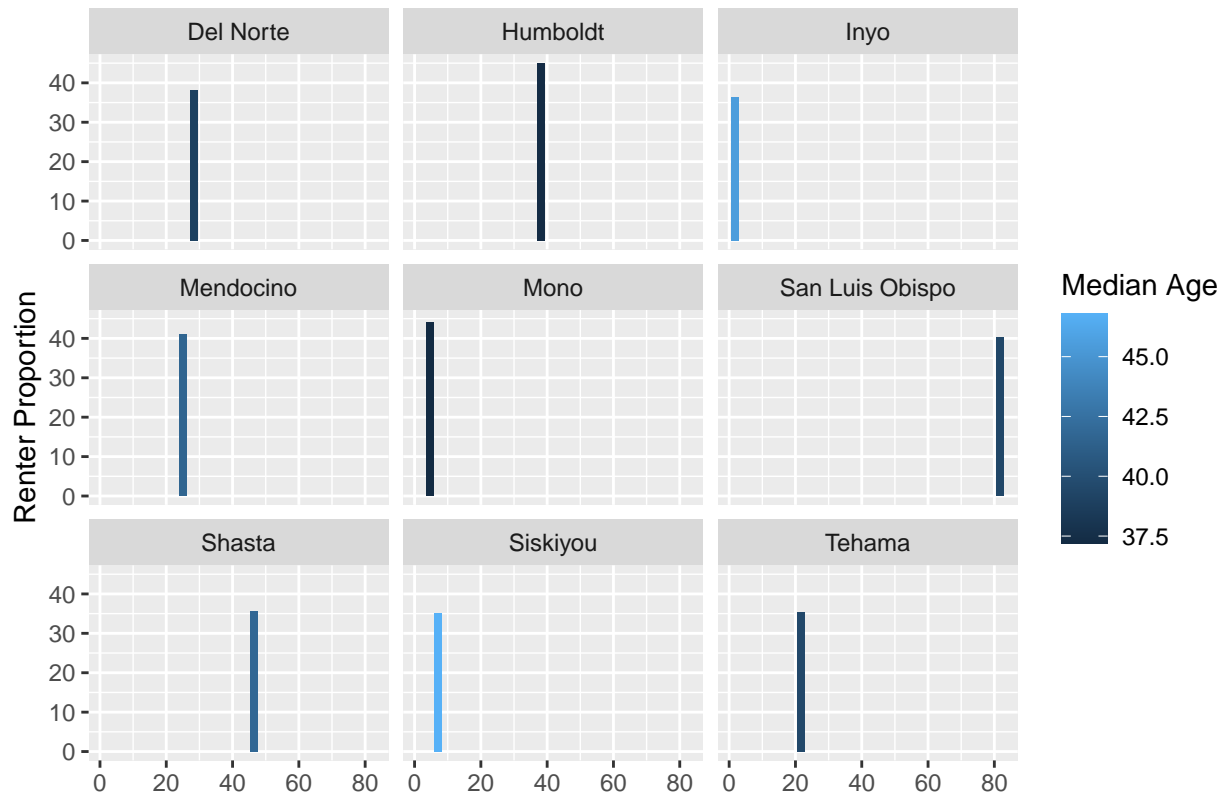
County Name	Pop/mi ²	Median Age	Renter Proportion	Mortality Rate	Total Costs of OSHPD Projects
Del Norte	28.30	39.0	38.29	1.01	0
Humboldt	38.06	37.3	44.99	0.96	0
Inyo	1.82	45.5	36.38	1.06	0
Mendocino	25.08	41.6	41.05	1.03	\$34,803.00
Mono	4.60	37.2	44.04	0.29	0
San Luis	81.82	39.4	40.28	0.88	\$89,105.00
Obispo					
Shasta	46.48	41.8	35.64	1.50	\$505,710.00
Siskiyou	7.12	46.8	35.25	1.18	0
Tehama	21.52	39.5	35.36	0.93	0

```
library(ggplot2)

df <- working_dataset %>%
  mutate(median_age =
    round(`Median Age`,0),
    pop_mi = round(`Pop/mi^2`,2))
one <- ggplot(df,
  aes(x = pop_mi,
    y = `Renter Proportion`, fill = `Median Age`)) +
  geom_col(position = "dodge") +
  theme(axis.title = element_blank(),
    axis.title.y = element_text()) +
  labs(fill = "Median Age",
    title = "Renter Proportion by Median Age",
    y = "Renter Proportion") + facet_wrap(~ `County Name`)

one
```

Renter Proportion by Median Age

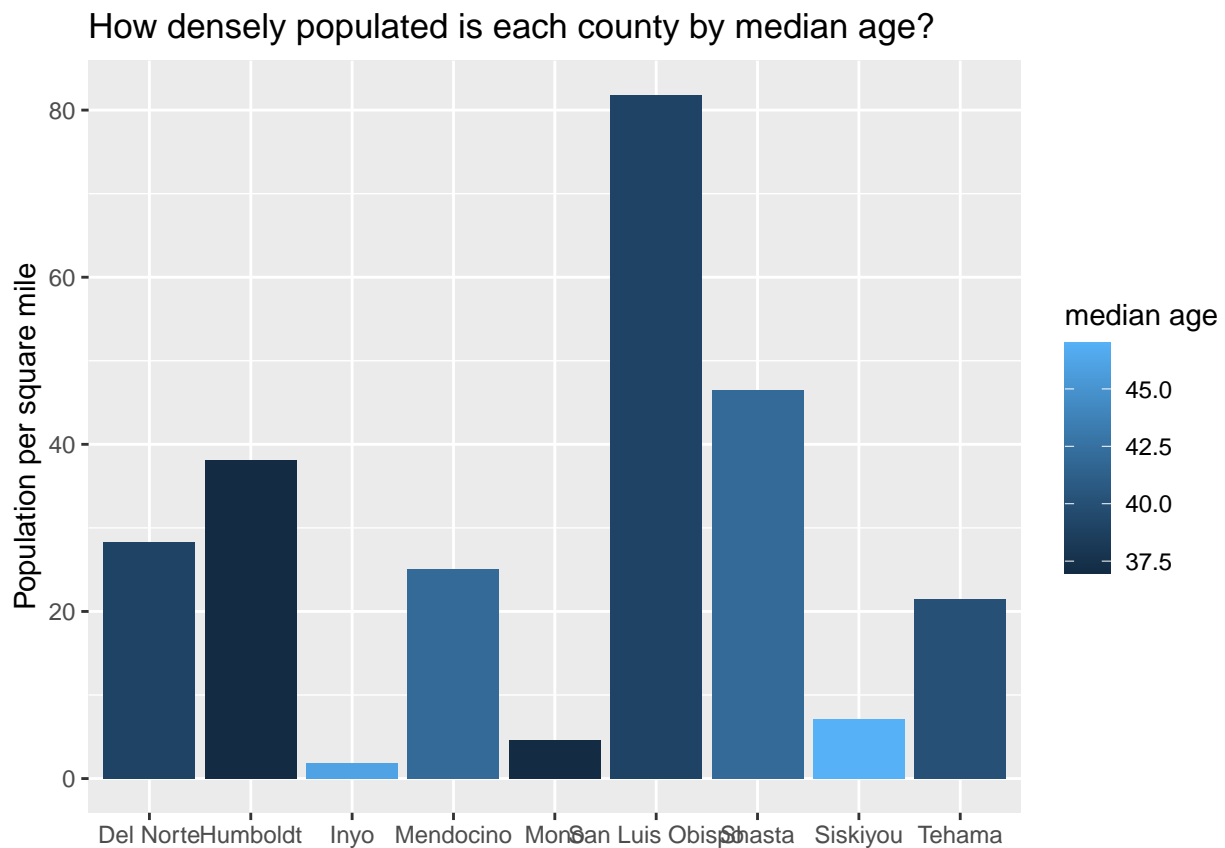


Visual one shows Renter Proportion by County with median age as the fill/gradient so that we visibly see if any of these counties would meet the three common criteria. Humboldt and Mono populations are younger than other counties but more renters than owners. Inyo and Siskiyou have the lowest population per square mile, high median age, and a high proportion of renters making these two ideal for the development fund proposals.

```
library(tidyverse)

two <- df %>%
  ggplot( aes(x = `County Name`,
              y = pop_mi, fill = median_age)) +
  geom_col(position = "dodge") +
  theme(axis.title = element_blank(),
        axis.title.y = element_text()) +
  labs(fill = "median age",
       title = "How densely populated is each county by median age?",
       y = "Population per square mile")

two
```



For visual two, we are interested in looking at density of each county and we noticed that San Luis Obispo is highly populated or denser than other counties with median age of 40 followed by Shasta as the second dense county making them less ideal for the development proposal than Inyo and Siskiyou.