Milestone 2

Meredith Klashman and Ruben Vargas

October 4, 2021

Scenario 1: Hospital Funding and Health Equity

Description of dataset:

The 3 datasets (oshpd hospital cost, mortality by county, & CA county demographics) that we will be utilizing in this project come from the PHW251 Box Folder. For the OSHPD hospital cost we have data dating back to Septemeber 2013, the mortality data from 2014-2019 and the county demographics from 2012.

How does the dataset relate to the group problem statement and question?

The CA county demographics dataset will allow us to locate the 5 counties with the lowest population for square mile, highest median range, and highest proportion of renters vs. homeowners. We will merge this data with information on OSHPD funding projects and mortality data to create a table to advocate for development funding partnerships in these counties.

Import

Use appropriate import function and package based on the type of file. Utilize function arguments to control relevant components (i.e. change column types, column names, missing values, etc.)

```
hospital_cost <- read.csv("/home/rstudio/PHW251_MK_RV/oshpd_hospital_cost.csv")
mort_county <- read.csv("/home/rstudio/PHW251_MK_RV/mort_by_county.csv")
county_demo <- read.csv("/home/rstudio/PHW251_MK_RV/ca_county_demographics.csv")
```

Identify data types for 5+ data elements/columns/variables

```
pop12_sqmi <- pull(county_demo, pop12_sqmi)
class(pop12_sqmi)
```

```
## [1] "numeric"
```

```
med_age <- pull(county_demo, med_age)
class(med_age)
```

```
## [1] "numeric"
```

```
renter_occ <- pull(county_demo,renter_occ)
class(renter_occ)
```

```
## [1] "integer"
```

```
owner_occ <- pull(county_demo,owner_occ)
class(owner_occ)
```

```
## [1] "integer"
```

```
Cause_Desc <- pull(mort_county,Cause_Desc)
class(Cause_Desc)
```

```
## [1] "character"
```

Each variable is in its desired format.

Provide a basic description of the 5+ data elements

```
#mean
mean(pop12_sqmi)
```

```
## [1] 665.0612
```

```
mean(med_age)
```

```
## [1] 38.48966
```

```r
mean(renter_occ)
```

```
## [1] 95553.91
```

```r
mean(owner_occ)
```

```
## [1] 121299.5
```

```r
#median
median(pop12_sqmi)
```

```
## [1] 103.4236
```

```r
median(med_age)
```

```
## [1] 37.05
```

```r
median(renter_occ)
```

```
## [1] 25140
```

```r
median(owner_occ)
```

```
## [1] 39306
```

```r
#range
range(pop12_sqmi)
```

```
## [1]     1.543841 17398.353740
```

```r
range(med_age)
```

```
## [1] 29.6 51.0
```

```r
range(renter_occ)
```

```
## [1]     140 1696455
```

```r
range(owner_occ)
```

```
## [1]     357 1544749
```

```r
#unique values
unique(Cause_Desc)
```

```
##  [1] "All causes (total)"
##  [2] "Alzheimer's disease"
##  [3] "Malignant neoplasms"
##  [4] "Chronic lower respiratory diseases"
##  [5] "Diabetes mellitus"
##  [6] "Assault (homicide)"
##  [7] "Diseases of heart"
##  [8] "Essential hypertension and hypertensive renal disease"
##  [9] "Accidents (unintentional injuries)"
## [10] "Chronic liver disease and cirrhosis"
## [11] "Nephritis, nephrotic syndrome and nephrosis"
## [12] "Parkinson's disease"
## [13] "Influenza and pneumonia"
## [14] "Cerebrovascular diseases"
## [15] "Intentional self-harm (suicide)"
```