hw\_4\_rmd

Alyssa Melvin

October 24, 2018

library(readr)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(tidyr)   
library(purrr)  
library(broom)  
library(ggplot2)  
library(forcats)  
library(scales)

##   
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':  
##   
## discard

## The following object is masked from 'package:readr':  
##   
## col\_factor

homicides <- read\_csv("../data/homicide-data.csv")

## Parsed with column specification:  
## cols(  
## uid = col\_character(),  
## reported\_date = col\_integer(),  
## victim\_last = col\_character(),  
## victim\_first = col\_character(),  
## victim\_race = col\_character(),  
## victim\_age = col\_character(),  
## victim\_sex = col\_character(),  
## city = col\_character(),  
## state = col\_character(),  
## lat = col\_double(),  
## lon = col\_double(),  
## disposition = col\_character()  
## )

homicides <- homicides %>%   
 unite(city\_name, "city", "state", sep = ", " )  
  
baltimore <- homicides %>%   
 group\_by(city\_name) %>%   
 filter(city\_name == "Baltimore, MD") %>%   
 mutate(unsolved = disposition %in% c("Closed without arrest",   
 "Open/No arrest")) %>%  
 summarise(total = n(), unsolved = sum(unsolved))  
   
baltimore\_prop\_results <- prop.test(x = baltimore$unsolved, n = baltimore$total)  
  
unsolved <- homicides %>%   
 mutate(unsolved = disposition %in% c("Closed without arrest",   
 "Open/No arrest")) %>%   
 select(city\_name, unsolved) %>%   
 group\_by(city\_name) %>%   
 summarise(N = n(), unsolved = sum(unsolved)) %>%   
 ungroup() %>%   
 mutate(prop\_results = map2(.x = unsolved, .y = N, .f = prop.test)) %>%   
 mutate(prop\_results = map(.x = prop\_results, .f = tidy)) %>%   
 unnest(prop\_results, .drop = TRUE) %>%   
 select(city\_name, estimate, conf.low, conf.high)

## Warning in .f(.x[[i]], .y[[i]], ...): Chi-squared approximation may be  
## incorrect

unsolved %>%   
 ggplot() +  
 theme\_dark() +  
 geom\_point(mapping = aes(x = estimate, y = reorder(city\_name, estimate)),   
 color = "white") +  
 geom\_errorbarh(mapping = aes(x = estimate, y = city\_name,   
 xmin = conf.low, xmax = conf.high), height = 0,  
 color = "white") +  
 scale\_x\_continuous(labels = percent,  
 limits = c(0.2, 0.8)) +  
 labs(y = "", x = "Percent of homicides that are unsolved") +  
 ggtitle("Unsolved homicides by city", subtitle = "Bars show 95% confidence interval")

## Warning: Ignoring unknown aesthetics: x

## Warning: Removed 1 rows containing missing values (geom\_point).

## Warning: Removed 1 rows containing missing values (geom\_errorbarh).

