hw_4_rmd

Alyssa Melvin

October 24, 2018

```
library(rmarkdown)
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")
```

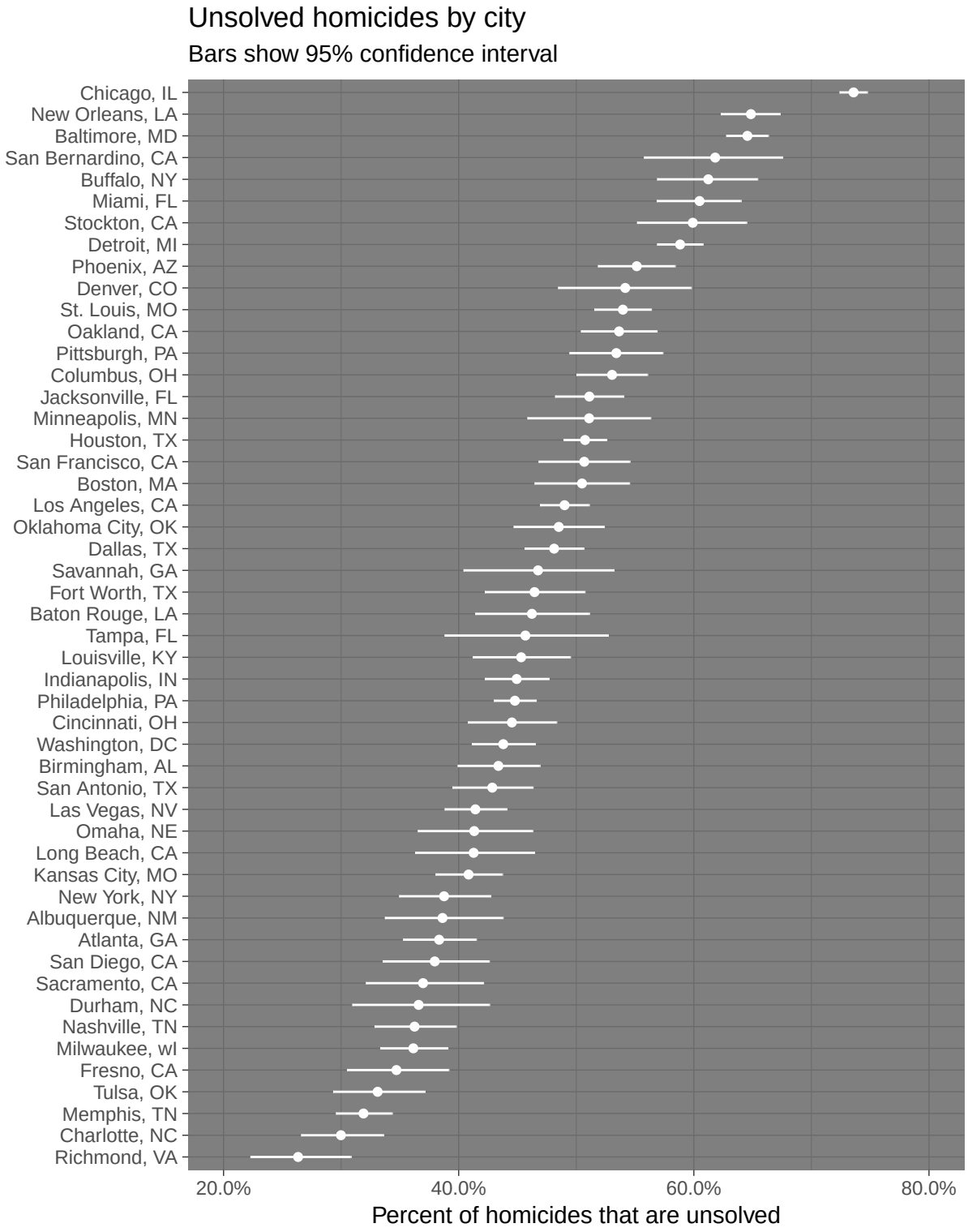
```
## Parsed with column specification:
## cols(
##   uid = col_character(),
##   reported_date = col_integer(),
##   victim_last = col_character(),
##   victim_first = col_character(),
##   victim_race = col_character(),
##   victim_age = col_character(),
##   victim_sex = col_character(),
```

```
##   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")) %>%
  filter(city_name != "Tulsa, AL") %>%
  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)
```

```
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
```


Unsolved homicides by city
Bars show 95% confidence interval
Chicago, IL
New Orleans, LA
Baltimore, MD
San Bernardino, CA
Buffalo, NY
Miami, FL
Stockton, CA
Detroit, MI
Phoenix, AZ
Denver, CO
St. Louis, MO
Oakland, CA
Pittsburgh, PA
Columbus, OH
Jacksonville, FL
Minneapolis, MN
Houston, TX
San Francisco, CA
Boston, MA
Los Angeles, CA
Oklahoma City, OK
Dallas, TX
Savannah, GA
Fort Worth, TX
Baton Rouge, LA
Tampa, FL
Louisville, KY
Indianapolis, IN
Philadelphia, PA
Cincinnati, OH
Washington, DC
Birmingham, AL
San Antonio, TX
Las Vegas, NV
Omaha, NE
Long Beach, CA
Kansas City, MO
New York, NY
Albuquerque, NM
Atlanta, GA
San Diego, CA
Sacramento, CA
Durham, NC
Nashville, TN
Milwaukee, wI
Fresno, CA
Tulsa, OK
Memphis, TN
Charlotte, NC
Richmond, VA
20.0%
40.0%
60.0%
80.0%
Percent of homicides that are unsolved