Homework_4

Grant Erlandson

October 24, 2018

```r
library(tidyr)
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
```

```r
library(broom)
library(purrr)
library(ggplot2)
library(ggfortify)
library(ggthemes)
library(ggmap)
library(stringr)
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
```

```r
list.files("../Data/")
```

```
## [1] "homicide-data.csv"
```

```r
homicides <- read.csv("../Data/homicide-data.csv")

homicides2 <- homicides %>%
  unite(city_name, city, state, sep = ", ")
```

```r
unsolved <- homicides2 %>%
  select(city_name, disposition) %>%
  mutate(solved = disposition == "Closed by arrest") %>%
  group_by(city_name) %>%
  summarize(n_homicides = n(), n_unsolved = sum(!solved))

baltimore <- unsolved %>%
  filter(city_name == "Baltimore, MD")
baltimore_prop <- prop.test(x = baltimore$n_unsolved,
          n = baltimore$n_homicides)
tidy(baltimore_prop)

## # A tibble: 1 x 8
##   estimate statistic  p.value parameter conf.low conf.high method
##      <dbl>     <dbl>    <dbl>     <int>    <dbl>     <dbl> <chr>
## 1    0.646      239. 6.46e-54         1    0.628     0.663 1-sam~
## # ... with 1 more variable: alternative <chr>

unsolved <- unsolved %>%
  mutate(test = map2(n_unsolved, n_homicides, ~ prop.test(.x, n = .y))) %>%
  mutate(test = map(test, ~ tidy(.x))) %>%
  unnest(.drop = TRUE) %>%
  select(city_name, estimate, conf.low, conf.high)

## Warning in prop.test(.x, n = .y): Chi-squared approximation may be
## incorrect

unsolved_plot <- unsolved %>%
  mutate(city = (fct_reorder(.f = city_name, .x = estimate, .desc = FALSE)))

ggplot(unsolved_plot, aes(x = estimate, y = city))+
  geom_point(color = "White")+
  geom_errorbarh(aes(xmin = conf.low, xmax = conf.high), height = 0, color =
"White")+
  scale_x_continuous(name = "Percent of homicides that are unsolved", labels
= percent)+
  labs(y = "")+
  theme_dark()
```

Y-axis (top to bottom): Chicago, IL; New Orleans, LA; Baltimore, MD; San Bernardino, CA; Buffalo, NY; Stockton, CA; Phoenix, AZ; Denver, CO; St. Louis, MO; Pittsburgh, PA; Columbus, OH; Jacksonville, FL; Minneapolis, MN; Houston, TX; San Francisco, CA; Los Angeles, CA; Boston, MA; Oklahoma City, OK; Dallas, TX; Savannah, GA; Baton Rouge, LA; Louisville, KY; Indianapolis, IN; Cincinnati, OH; Philadelphia, PA; Washington, DC; Birmingham, AL; San Antonio, TX; Las Vegas, NV; Omaha, NE; Long Beach, CA; Kansas City, MO; Albuquerque, NM; San Diego, CA; Atlanta, GA; Sacramento, CA; Durham, NC; Nashville, TN; Milwaukee, WI; Fresno, CA; Tulsa, OK; Memphis, TN; Charlotte, NC; Richmond, VA; Tulsa, AL

X-axis: Percent of homicides that are unsolved — 0.0%, 25.0%, 50.0%, 75.0%