

STAT 201: Practice Midterm

Solutions

Question 1

```
gun_violence |>
  count(city_or_county) |>
  arrange(-n) |>
  mutate(prop = n / sum(n)) |>
  slice(1:5) |>
  kable()
```

| city_or_county | n | prop |
|----------------|-------|-----------|
| Chicago | 10298 | 0.0462661 |
| Baltimore | 3703 | 0.0166366 |
| New Orleans | 2939 | 0.0132041 |
| Philadelphia | 2824 | 0.0126875 |
| Jacksonville | 2372 | 0.0106567 |

Question 2

```
gun_violence |>
  mutate(n_shot = n_killed + n_injured) |>
  arrange(-n_shot) |>
  slice(1:6) |>
  select(date, state, city_or_county, n_shot) |>
  kable()
```

| date | state | city_or_county | n_shot |
|------------|---------|--------------------|--------|
| 2016-06-12 | Florida | Orlando | 103 |
| 2017-11-05 | Texas | Sutherland Springs | 47 |

| date | state | city_or_county | n_shot |
|------------|------------|----------------|--------|
| 2015-12-02 | California | San Bernardino | 35 |
| 2015-05-17 | Texas | Waco | 27 |
| 2017-07-01 | Arkansas | Little Rock | 25 |
| 2016-07-25 | Florida | Fort Myers | 21 |

Question 3

```
state_incidents <- gun_violence |>
  count(state, year)

state_incidents |>
  group_by(state) |>
  summarise(avg_n = mean(n)) |>
  arrange(-avg_n) |>
  slice(c(1:3, 48:50)) |>
  kable()
```

| state | avg_n |
|------------|---------|
| Illinois | 4173.75 |
| California | 3792.75 |
| Florida | 3549.25 |
| Wyoming | 120.00 |
| Vermont | 113.00 |
| Hawaii | 67.50 |

Question 4

```
state_incidents2 <- state_incidents |>
  left_join(census, by = c("state", "year")) |>
  left_join(laws, by = "state")
```

Question 5

```
state_incidents2 <- state_incidents2 |>
  mutate(rate = n/population * 100000)

state_incidents2 |>
  group_by(state) |>
  summarise(avg_rate = mean(rate)) |>
  arrange(-avg_rate) |>
  slice(c(1:3, 48:50)) |>
  kable()
```

| state | avg_rate |
|-----------|-----------|
| Alaska | 43.556789 |
| Delaware | 41.743675 |
| Louisiana | 41.026640 |
| Utah | 8.424150 |
| Arizona | 7.900990 |
| Hawaii | 4.743822 |

Question 6

```
state_incidents2 |>
  filter(year == 2017) |>
  mutate(log_population = log10(population)) |>
  ggplot(aes(x = gun_laws2017, y = rate, col = log_population)) +
  geom_point() +
  theme_classic() +
  scale_color_viridis_c() +
  labs(x = "Number of gun laws", y = "Rate of gun violence (per 100000)",
       title = "Rate of gun violence and number of laws",
       subtitle = "2017",
       color = "Population (log10)")
```

