Lab 2 - Data wrangling

Jessica St. Jean

2026-02-10

```
library(tidyverse)
```

Questions

Part 1

```
?midwest
```

Question 1

```r
midwest |>
  count(state, sort = TRUE)
```

```
# A tibble: 5 x 2
  state     n
  <chr> <int>
1 IL      102
2 IN       92
3 OH       88
4 MI       83
5 WI       72
```

Comment: Illinois has the most amount of counties with 102, and Wisconsin has the least with 72.

Question 2

```r
midwest |>
  count(county,state) |>
  count(county, name = "n_states")|>
  filter(n_states == n_distinct(midwest$state))
```

```
# A tibble: 3 x 2
  county    n_states
  <chr>        <int>
1 CRAWFORD         5
2 JACKSON          5
3 MONROE           5
```

Question 3

```
midwest |>
  filter(popdensity > 25000) |>
  select(county,state,popdensity,poptotal,area) |>
  arrange(desc(popdensity))
```

```
# A tibble: 9 x 5
  county     state popdensity poptotal  area
  <chr>      <chr>      <dbl>    <int> <dbl>
1 COOK       IL        88018.  5105067 0.058
2 MILWAUKEE  WI        63952.   959275 0.015
3 WAYNE      MI        60334.  2111687 0.035
4 CUYAHOGA   OH        54313.  1412140 0.026
5 DU PAGE    IL        39083.   781666 0.02
6 MARION     IN        34659.   797159 0.023
7 HAMILTON   OH        34649.   866228 0.025
8 FRANKLIN   OH        28278.   961437 0.034
9 MACOMB     MI        25621.   717400 0.028
```

```
midwest |>
  filter(popdensity == max(popdensity)) |>
  select(county,state,popdensity,poptotal,area)
```

```
# A tibble: 1 x 5
  county state popdensity poptotal  area
  <chr>  <chr>      <dbl>    <int> <dbl>
1 COOK   IL        88018.  5105067 0.058
```

Question 4

Question 5

Question 6

Question 7

Question 8

Part 2

Question 9