Lab 3: States and Counties

JAKE JAMES

February 02, 2021

Introduction

The task is to explore the US census population estimates by county for 2015 from the package `usmap`. The data frame (`countypop`) has `3142` rows and `4` variables:

- `fips` is the 5-digit FIPS code corresponding to the county;
- `abbr` is the 2-letter state abbreviation; `county` is the full county name;
- `pop_2015` is the 2015 population estimate (in number of people) for the corresponding county.

Each row of the data frame represents a different county or a county equivalent. For the sake of simplicity, when we say a county, that also includes a county equivalent and when we say a state, that also includes the District of Columbia. Answer the following questions.

Excercises

Part 1: Length and Unique

a. How many unique 2-letter state abbreviations are there?

```
length(unique(countypop$abbr))
```

```
## [1] 51
```

There are 51 unique 2-letter state abbreviations.

b. What is the total number of counties in the US?

```
length(unique(countypop$fips))
```

```
## [1] 3142
```

There are 3142 counties in the US.

c. How many unique county names are there?

```
length(unique(countypop$county))
```

```
## [1] 1877
```

There are 1877 unique county names in the US.

Part 2: Count and Arrange

d. What are the top 10 most common county names?

```
countypop %>%
  count(county) %>%
  arrange(desc(n)) %>%
  head(10)
```

```
## # A tibble: 10 x 2
##    county               n
##    <chr>            <int>
##  1 Washington County    30
##  2 Jefferson County     25
##  3 Franklin County      24
##  4 Jackson County       23
##  5 Lincoln County       23
##  6 Madison County       19
##  7 Clay County          18
##  8 Montgomery County    18
##  9 Marion County        17
## 10 Monroe County        17
```

1. Washington

2. Jefferson

3. Franklin

4. Jackson

5. Lincoln

6. Madison

7. Clay

8. Montgomery

9. Marion

10. Monroe

e. Which state has the smallest number of counties?

```
countypop %>%
  count(abbr) %>%
  arrange(n) %>%
  head(1)
```

```
## # A tibble: 1 x 2
##   abbr      n
##   <chr> <int>
## 1 DC        1
```

Washington DC

f. Which state has the largest county in terms of population? How many people live in the largest county in terms of population?

```
arrange(countypop,desc(pop_2015))[1,]
```

```
## # A tibble: 1 x 4
##   fips  abbr  county             pop_2015
##   <chr> <chr> <chr>                 <dbl>
## 1 06037 CA    Los Angeles County 10170292
```

California (LA County)

Part 3 Group_by and Summarize

g. How many people live in each of the states?

```
## Displays each state and their total population.

countypop %>%
  group_by(abbr) %>%
  summarise(total_pop=sum(pop_2015))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 51 x 2
##    abbr  total_pop
##    <chr>     <dbl>
##  1 AK       738432
##  2 AL      4858979
##  3 AR      2978204
##  4 AZ      6828065
##  5 CA     39144818
##  6 CO      5456574
##  7 CT      3590886
##  8 DC       672228
##  9 DE       945934
## 10 FL     20271272
## # … with 41 more rows
```

h. What is the average population of a county in North Carolina?

```
countypop %>%
  filter(abbr== "NC") %>%
  summarise(mean(pop_2015))
```

```
## # A tibble: 1 x 1
##   `mean(pop_2015)`
##              <dbl>
## 1          100428.
```

The average population of a county in North Carolina is 100,428.

i. What is the largest county in terms of population of each of the states?

```
# Displays each state, their largest county, and that county's population.

countypop %>%
  group_by(abbr) %>%
  summarise(county=county[which.max(pop_2015)],max(pop_2015))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 51 x 3
##    abbr  county                  `max(pop_2015)`
##    <chr> <chr>                             <dbl>
##  1 AK    Anchorage Municipality           298695
##  2 AL    Jefferson County                 660367
##  3 AR    Pulaski County                   392664
##  4 AZ    Maricopa County                 4167947
##  5 CA    Los Angeles County             10170292
##  6 CO    Denver County                    682545
##  7 CT    Fairfield County                 948053
##  8 DC    District of Columbia             672228
##  9 DE    New Castle County                556779
## 10 FL    Miami-Dade County               2693117
## # … with 41 more rows
```