HUDM 5026 - Introduction to Data Analysis and Graphics in R
POTD 03 – Data Transformation

General instructions for POTDs:

- Write up your solutions and examples in a .Rmd file and knit it and submit an html or pdf file.
- Clearly label each part by number and letter, if applicable.
- Include plenty of comments in your code.
- The file should run without any errors from top to bottom.
- The write-up is due before the next class meeting.
- Although you may work collaboratively with others in class, each individual will turn in their own assignment.

Run the following code to create a tibble using the `state.x77` data.

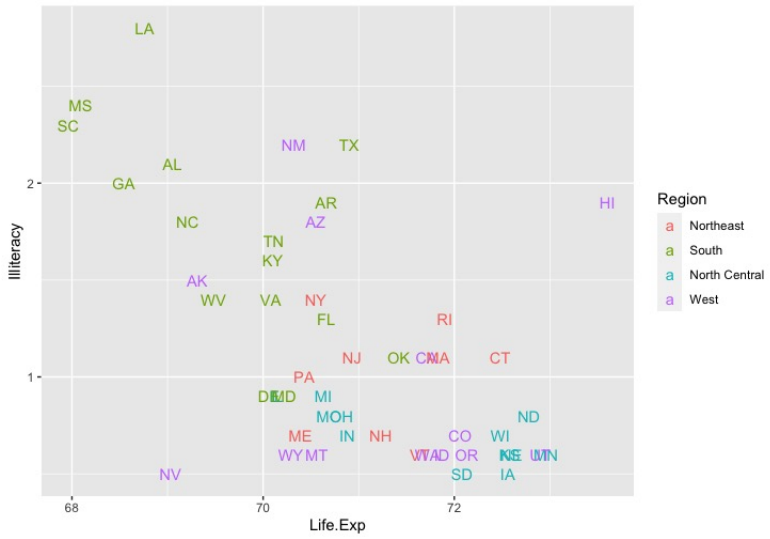
```
library(tidyverse)

# Create a tibble called dat using state.x77
dat <- data.frame(state.x77)
dat <- tibble(dat)

# Add the region and division and abbreviation information
dat$Region <- state.region
dat$Division <- state.division
dat$abbrev <- state.abb
print(dat, n = 50, width = 100)
```

Task 1 *Create a scatterplot of life expectancy on the horizontal and illiteracy rate on the vertical and color the point by state region.*

Task 2 *Now, instead of using `geom_point()`, use `geom_text` to create the same plot as above but with the state abbreviations plotted for each point instead of circles. It should look like this:*



Task 3 *Filter the data set to include only states that have illteracy rates higher than 1.6 and life expectancies lower than 75. Verify visually that your results agree with your plot.*

Task 4 *Create a data set that only contains Hawaii by filtering on region and life expectancy.*

Task 5 *Arrange the data, first in ascending then in descending order, based on state area. Which are the three states with most area? Which are the three states with least area?*

Task 6 *Use* `mutate()` *to add a variable to the data frame that gives state area in square kilometers (instead of square miles).*

Task 7 *Create a data frame with four variables: state population, state area, the number of people per square mile and the number of people per square kilometer.*

Task 8 *This one is open-ended; creativity is encouraged. Create some plots or demonstrate some transformations with these data.*