

## Homework 2

MSCS 6520 Business Analytics

Spring 2018

Assigned: February 5, 2018

Due: February 12, 2017 (by beginning of class)

### Readings

Chapters 5, 7, 10-11, 15-16 in [R for Data Science](#).

[Two Cultures](#) by Leo Breiman

### Exercises

Download the milwaukee\_weather.csv file from D2L (R Tutorials and Data).

1. Load the data using the read\_csv() function. Look at the resulting tibble using the head() function. How many columns are there? What are they and what are their types? Use the nrow() function to find the number of rows.
2. Plot the avg\_temp over time. What two columns do you need? What type of plot is appropriate? Do you see a repeating pattern?
3. Let's identify the timescales (e.g., year, month, day of the week) which best predict the variation in the temperature data. Extract the year, month, and day of the week into new columns. (Hint: Review how we manipulated date/times in the slides and reference the lubridate manual on D2L.) Save the output of head() on the resulting tibble. What types should they be? Generate plots for temperature vs the year, month, and day of the week. What type of plots are appropriate? Which timescales are most predictive of the variations?
4. Plot the snowfall over time. What two columns do you need? Try plotting the data with both the geom\_line() and geom\_point() plot types – which one do you think is better suited?
5. Snow only occurs at lower temperatures. Create a new column called total\_precipitation that contains the sums of the snowfall and rainfall columns. Then, create another column called has\_snowed which is a factor or logical (Boolean) type indicating whether it snowed that day. (Hint: Think about how we used ifelse() to replace NAs with the string “none”.) Save the output of head() on the resulting tibble. Lastly, plot total\_precipitation vs has\_snowed and color it by has\_snowed. What type of plot is most appropriate?
6. Plot the avg\_wind\_speed over time. Do you see a pattern?
7. Plot the snowfall vs the avg\_wind\_speed. Do you see a pattern?

Prepare a document containing the answers to the above questions, the plots, and the output of head() in exercises 1, 3, and 5 to demonstrate that you loaded and transformed the data correctly, respectively. Submit the document as a PDF to D2L. You may work in pairs, in which case, you should only submit one PDF per group.