

Stat 133, Spring 2011  
Homework 1: Introduction to R  
Due Friday, Feb. 11 at 11:59 PM

## Traffic Flow on Highways in California

Data was collected by loop detectors at one particular location of eastbound Interstate 80 in Sacramento, California. There are six columns and 1740 rows in the data set. The rows correspond to successive five minute intervals from March 14 to 20, 2003, where the data values in a row report the flow (number of cars) and occupancy (the proportion of time there was a car over the loop) in each of three lanes on the freeway. Lane 1 is the leftmost lane, lane 2 is in the center, and lane 3 is the rightmost. The data set can be found at <http://www.stat.berkeley.edu/classes/s133/data/flow-occ.csv> The original data are from the Freeway Performance Measurement System (PEMS) website (Freeway Performance Measurement System (PEMS) website)

1. Read the data in R. Note that you don't need to actually download the data in order to read it.
2. Is there a particular lane for which the flow of traffic is consistently higher than the other lanes?
3. Is there a relationship between flow and occupancy?

The data for this problem was provided by John Rice of the UC Berkeley statistics department.

## Ages of United States Senators

A lot of excitement was generated when Barack Obama was elected president because of his youth. Does the same trend in age hold in the United States Senate?

At <http://www.stat.berkeley.edu/classes/s133/data/senate.txt> is a tab-separated file with the following columns, recorded for each past and present United States senator: **Name**, the name of the senator; **start**, the first year that the senator served in office; **end**, the last year they served; **State**, the state from which they were elected; **Party**, their political party, and **Birthday**, when they were born.

Plot the average age of the senators against time, for all the years of available data. The start and end years are recorded so that a senator is active if the year in question is greater than or equal to the start year and less than or equal to the end year. Are senators getting younger over time?

## Temperatures in Berkeley in January

In my backyard, I have a temperature sensor attached to a computer which records the temperature every fifteen minutes. I've extracted the minimum, mean and maximum temperature for each day in January over the last seven years into <http://www.stat.berkeley.edu/classes/s133/data/january.tab>. The columns are **min**, the minimum temperature for the day; **mean**, the average of all 96 temperatures for the day; **max**, the maximum temperature for the day; **year**; and **day**. Remember that this data is only for the last seven Januaries.

1. Over the last seven years, which year had the hottest January? Which year had the coldest?
2. The computer to which the sensor is connected was broken for a few days in January one year. Which year was it? Explain how you found the answer – simply examining the data by eye is not acceptable.