# Homework 2: Testing

1. In your lecture notes, you will find "treatment" and "control" for a simple A/B test carried out by the New York Times Movies section. Recall from lecture that A/B testing is a kind of randomized trial in which visitors are divided into two groups. As members of the control group browse the web site, they are shown the site's typical design, the pages are unchanged. The treatment group, on the other hand, sees an experimental version of the site, with pages typically containing one or more changes. Our data for this lab come from one such test. Recall from lecture that the treatment for this experiment involved removing a navigation bar from the middle of the page heading. The experiment was restricted to pages in the Movies Section (and not the other sections like Science or Health or Opinion) and took place between February 7 and March 10 in 2008.

   The experimenters had reason to believe that under the treatment, the page without the navigation bar, visitors will be forced to make greater use the search box. To study this, the experiment randomized visitors into two groups, the treatment and a control. The question of interest: Do visitors receiving the treatment use the search box more often than visitors reveiving the control? To answer this, you will look at 1,000 visitors who were new to the site, visitors that had never been to `nytimes.com` before. Remember from lecture that this determination is based on so-called cookie data stored by the visitor's browser. The details are not particularly important here; what is important is that these 1,000 visitors had never seen `nytimes.com` much less the Movies Section before. Once they arrived, they were randomized to receive either the treatment or the control.

   You can load the data into R (RStudio) with the command

   ```
   source("http://www.stat.ucla.edu/~cocteau/stat105/data/movies.R")
   ```

   You should see an object in your workspace called `movies`. There are 1,000 rows in this data set, each being a different visit (by a different visitor). The variables are `id`, the visitors ID (set via a cookie  see your lecture notes); `treatment`, whether the visitor recieved pages that were unchanged or were missing a navigation bar; `count`, the number of times this visitor clicked on the go button in the search box, initiating a search; `outcome` a 0/1 variable that is 0 if the visitor did not use the search box and 1 if they did at least once; `day` indicating the day the visit took place, an integer where 0 means the rst day of the study and 31 means the last; and `hour`, the hour of the day the visit took place, with 0 being midnight and 12 being noon.

The study was designed to examine if removing the navigation bar at the top of the page led to increased use of the search box. In advance of the trial, the experimenters do not believe that this treatment would cause less use of the search box. (This is similar to the set-up for the examples of "one-sided tests" given in lecture.) Let's cast this into the hypothesis testing framework we have developed. First, you can have a look at the experimental results with the following command.

```
table(movies$outcome,movies$treatment)
```

We now want to assess whether the differences we see (do we see any?) are the result of our study design (and hence "noise") or if there is a real effect from removing the navigation bar, prompting visitors to make greater use of the search box.

(a) Produce a graphic representing the experimental outcomes in the table above and interpret what you see.

(b) State the null hypothesis associated with the experimenter's study and link this to the graphic in (a) and the table above.

(c) Identify a test statistic to evaluate this hypothesis and write code to compute it from your table of results.

(d) Describe how you would simulate values of the test statistic under the null hypothesis and write code to do this (the `sample` command mentioned in class will prove very helpful here).

(e) Generate the null distribution and compare it to the exerimenter's results in the table above both informally (graphically) and via a P-value.

(f) (Bonus) What can you say about the average number of search box uses per visit?

2. The data from the Travel Section experiment can be loaded with the command

```
source("http://www.stat.ucla.edu/~cocteau/stat105/data/nyt.R")
```

We did a fair bit of analysis in lecture, but now you have an opportunity to dig a little. Perform some EDA on the data, identifying some aspect that is of interest to you. This could either be some formal test or a collection of graphics. In short, tell a story from these data. Open territory includes the referral fields, the agent (i.e. the browser and operating system they are on), time of day, etc. It is also acceptable to repeat analysis given in class, although it's probably more fun to try something new.