

## City Team Data Task

The following is a project that has crossed the desk of a member of one of our city teams.

### Program description

Failure to appear in court costs the City \$200 per missed appearance. A recent randomized control trial was aimed at reducing the number of people who fail to appear at a case's first arraignment. A two arm trial was implemented for the first arraignment for every case from January 1, 2015 to September 30, 2015. The intervention was a re-design of the court summons that police officers gave out. The court summons in the treatment group emphasized loss aversion — there are significant additional financial penalties for failing to appear, including a warrant for arrest.

Randomization was at the level of the court summons. Police officers gave a new court summons half of the time and the old court summons the other half of the time. The City keeps track of each interaction a defendant has with the court, including whether or not the defendant failed to appear in court.

### Overview of the data task

There are six components to this data task. Note that you may use any software that you prefer, as long as it allows for some scripting capabilities since you will be asked to submit your script. We prefer STATA or R. This task requires you to make decisions about how to treat the data; please feel free to use the data to gain as much insight as as possible about how we should regard this program. Note your decisions next to the tables.

1. From the various sheets in the ResearchTask.xlsx file, please create a dataset that includes: each case; demographic information; and court data on offenses, actions, and action status.

2. Clean the data as needed. Please note the decisions you make.
3. Present one table assessing whether the randomization was done properly.
4. Present one table assessing whether the new court summonses had an effect on failures to appear at first arraignment. The variable that indicates whether or not each case had the treatment is “treatment.”
5. The recently elected mayor has publicly committed to equitable justice. Her team is particularly interested in how this program may affect specific subgroups. Present one table to them that you think is relevant.
6. This year the mayor has set aside substantial funds for another randomized control trial to test a new message and any related intervention costs. Please perform power calculations and determine how many months would be required to detect a reasonable effect at the 0.05 level between the treatment and control. Feel free to offer up to 3 different suggestions for the number of months this should run.
7. Given the results of the first trial, please provide any recommendations you have in designing or implementing the next trial and intervention. Recommendations should not take up more than half a page.