



Background

In January 2012, the LA County Sheriff's Department began efforts to divert people with substance abuse or mental health issues from the criminal justice system by allowing individuals arrested for misdemeanor offenses during a substance abuse or mental health crisis to be released on bail rather than held in jail pre-trial. In addition, they also started a pilot program that ran through October 2013 to provide caseworkers and treatment to some of these individuals. The Sheriff has asked you to evaluate the effectiveness of the pilot program, and provided the data described below. The program's manager has left the Department, so we know little about how people were assigned to it, but we do know assignment was done independently when an individual was arrested. This means that it's possible that someone with multiple arrests could have been referred to the program on some occasions but not on others. No information is provided on take-up of services.

Data

There are three data sets: case.csv, demo.csv, and prior_arrests.csv. Case.csv is the main dataset showing the dates of arrest and disposition (trial or court appearance) during the time the pilot program ran, as well as an indicator of whether the arrestee was referred to the pilot program for that arrest. Demo.csv contains some demographic information about the arrestee, and prior_arrests.csv contains information on individuals' arrest histories from 2008-2011.

Dataset Descriptions

	N	Level
Case.csv	25000	Case (one row per case)
Demo.csv	14353	Person (one row per person)
Prior_arrests.csv	43602	Arrest (one row per arrest)

Case.csv

	Caseid	Person_id	Arrest_date	Dispos_date	Treat
Type	Numeric	Numeric	Character	Character	Numeric
Range	(4, 100000)	(1, 20000)	2012-01-02, 2013-10-28	2012-01-05, 2018-03-04	(0, 1)
Description	A unique identifying number for the arrest	A unique identifying number for the individual	Arrest Date	Date of Trial or Court Appearance	1 = treated, 0 = not treated

Demo.csv

	Person_id	race	gender	bdate
Type	Numeric	Character	Character	Character
Range	(1, 20000)	ASIAN, BLACK, HISPANIC, WHITE	M = male, F = Female	Birth Date
Description	A unique identifying number for the individual	The individual's racial category	The individual's gender	The individual's date of birth

Prior_arrests.csv

	Person_id	Arrest_date
Type	Numeric	Character
Range	(1, 20000)	2008-01-01, 2011-12-31
Description	A unique identifying number for the individual	Arrest Date

Part 1: Data Management

- a. The main outcome of interest is whether the defendant was arrested again prior to his/her disposition (trial or court appearance) date. Create a variable called 're_arrest' that equals one if the defendant was arrested prior to trial and zero otherwise. E.g., imagine person 1 was arrested three times.

Caseid	Person_id	Arrest_date	Dispos_date	Re_arrest
123	1	2012-01-01	2012-04-05	0
1344	1	2012-06-05	2012-09-13	1
1451	1	2012-07-09	2012-11-21	0

- b. Create a variable called prior_arrests that equals the number of arrests prior to the current arrest. Assume that all of the individual's arrests prior to the study period are contained in prior.csv. For example, if person z has five arrests in prior.csv and two in case.csv, her first arrest in case.csv should have prior_arrests = 5 and the second should have prior_arrests = 6. If someone is not included in prior_arrests.csv, assume they had zero arrests at the start of the study period.
- c. Create an 'age' variable that equals the defendant's age at the time of each arrest.

Part 2: Statistical Analysis

Help determine whether the program should be continued/expanded by estimating its effect on re-arrests prior to disposition. Using whatever methods you feel will best answer the question, prepare a short memo (no more than two pages) with your recommendation, and your justification for it.