

# Project3

---

Team ID: Team 6

NAME: Connor Rosenberg

NAME: Rongkui Han

NAME: Yuqing Yang

NAME: Nassim Ali-Chaouche

---

## 1 Introduction

### 1.1 Background

### 1.2 Questions of Interest

## 2 Analysis Plan

### 2.1 Population and study design

### 2.2 Statistical Analysis

#### 2.2.1 Descriptive Analysis

The data set contains the US traffic fatalities panel data for the 48 contiguous states from 1982 to 1988, where each state is observed in seven time periods. There are 336 rows and 34 columns in the data set. Each row represents an observation for a state in a time point and each column represents a variable for the observation.

To compare the traffic deaths across the different states and times, we use the fatality rate here, which is the number of annual traffic death in every 10,000 people of that state.

First, we found out there is two missing values in the data set, which are the variable that whether the state had a mandatory jail sentence or not and the variable that whether the state had a mandatory community service or not for California in 1988. Besides, there are 41 states did not change their policy in the mandatory jail sentence from 1982 to 1988. Considering that the majority of the states had the same policy in the mandatory jail sentence across the time, and California did not have the mandatory jail sentence from 1982 to 1987, we assume that California did not have the mandatory jail sentence in 1988. The relationship between the fatality rate and the policy in the mandatory jail sentence for the other seven states is shown in the Figure 1(a). From Figure 1(a), both South Carolina and Nevada changed their policy in the mandatory jail sentence from no to yes in 1983, but their fatality rates changed in the opposite direction after that. Similarly, we could also observe this situation in Oregon, Utah, and Connecticut. Given this plot, we may also need to consider the effects of other variables in the data set or the effects of unobserved variables.

Then, we explore the distribution of fatalities rate for all the observations. Figure 1(b) shows that overall, observations that had the mandatory jail sentence had a higher fatality rate than observations that did not have the mandatory jail sentence had a higher fatality rate, but the difference in medians is not significant. Figure 1(c) shows that distributions of fatalities rate for different years are roughly the same.

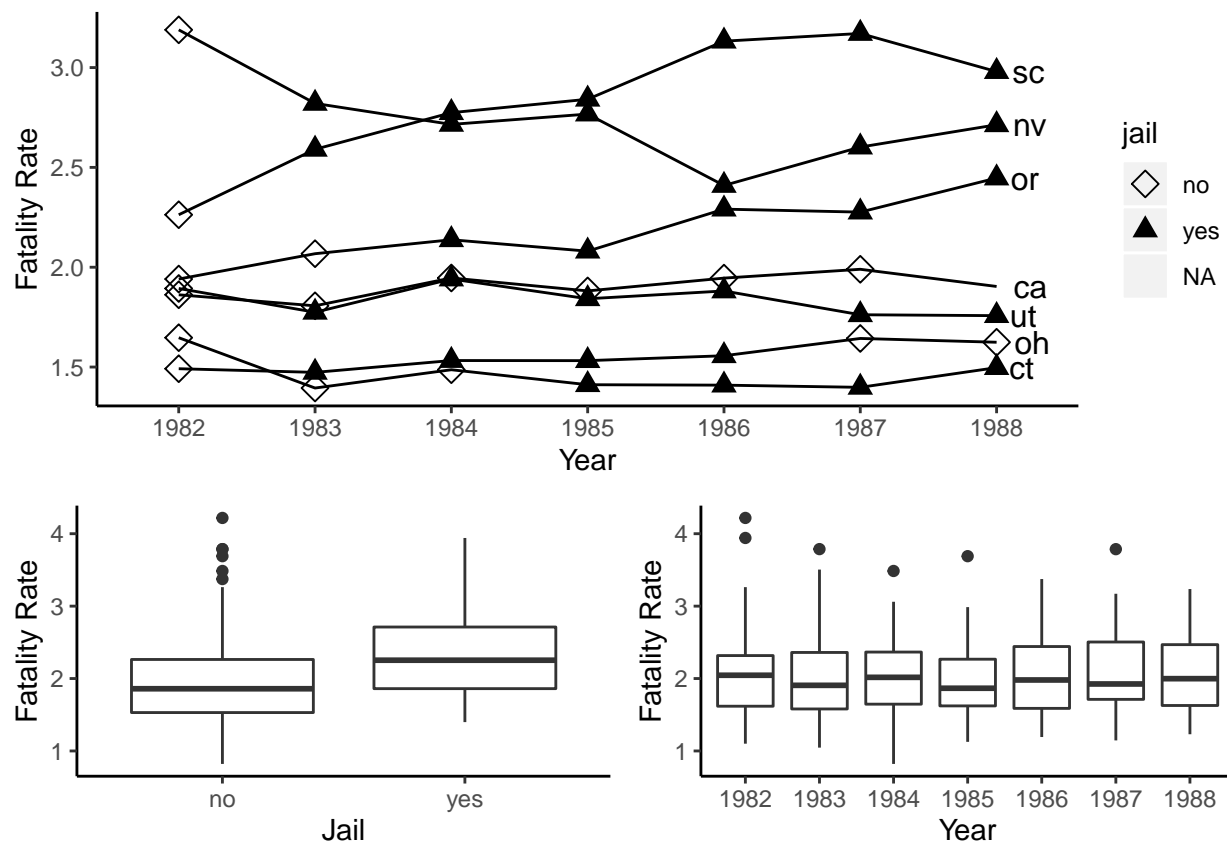


Figure 1: (a) Fatality rate for states that changed their policy in the mandatory jail sentence from 1982 to 1988 (b) Boxplot of the fatality rate for different policies in the mandatory jail sentence (c) Boxplot of the fatality rate for different years