

Crime Rates on College Campuses in Texas

By: Jeffrey Downs

Advisor: Dr. Kalanka Jayalath

Abstract

The purpose of this study was to determine whether or not there was a significant difference in the average crime rates between public and private universities in Texas, per 100 students from the years 2011 and 2013. Data was randomly compiled from an online database of 30 universities from both public and private universities. A two sample t-test was applied to see if there was a linear relationship between the crime rates and the year for both public and private universities separately from the years 2001 to 2013

Data

Histograms for public and private universities for years 2011 and 2013

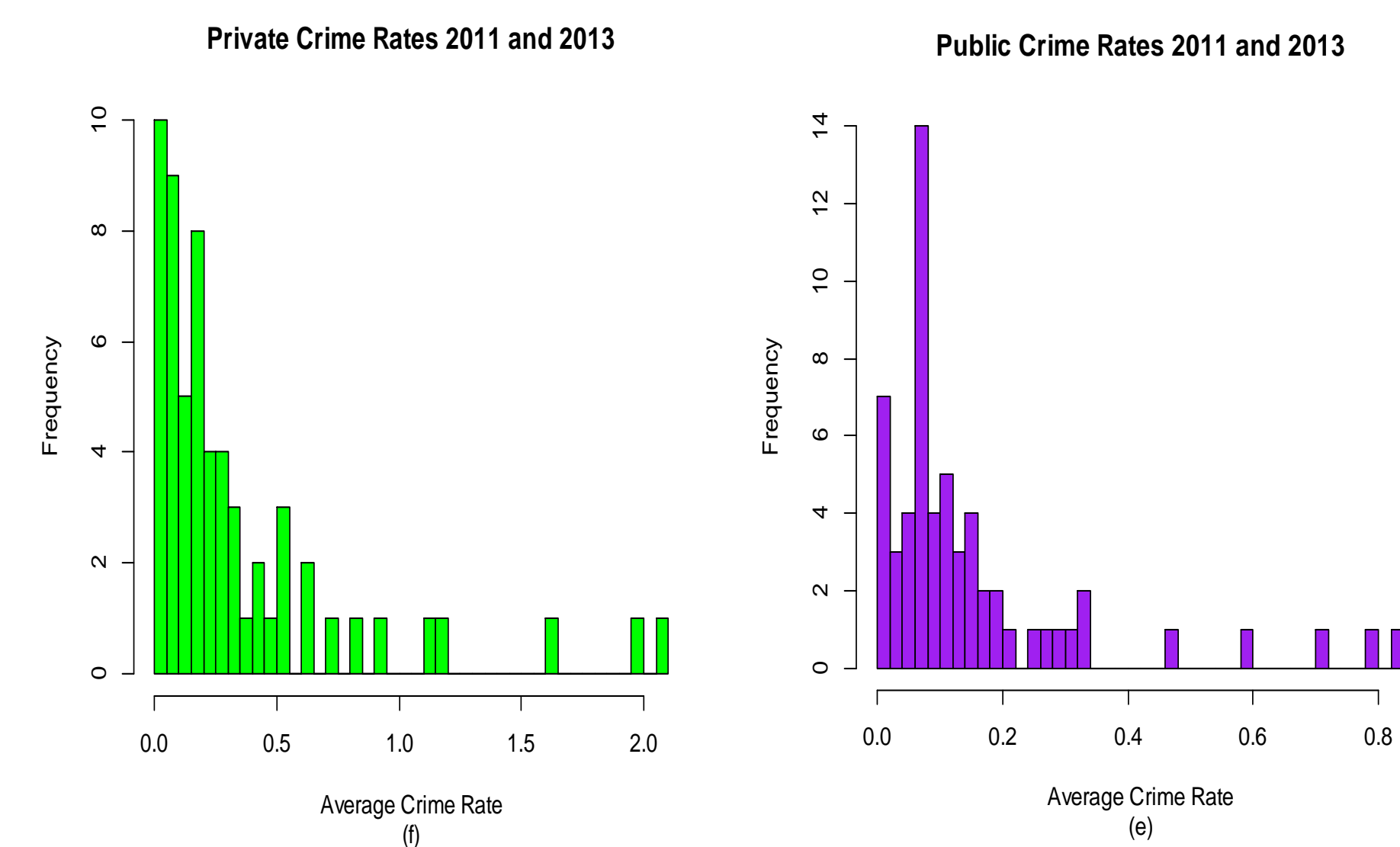


Table 1: Average Crime Rate Values of Public and Private Universities for 2011 and 2013, Per 100 Students

Public		Private	
2011	2013	2011	2013
0.08	0.15	0.31	0.11
0.27	0.11	0.83	0.00
0.00	0.00	0.52	0.33
0.47	0.34	0.00	0.71
0.59	0.34	0.09	0.09
0.04	0.07	0.00	0.00
0.14	0.10	0.30	0.08
0.71	0.32	0.22	0.17
0.18	0.20	0.61	0.63
0.07	0.09	0.05	0.00
0.07	0.07	1.65	2.09
0.29	0.00	0.30	0.54
0.20	0.00	0.00	0.45
0.79	0.84	0.25	0.26
0.17	0.13	0.11	0.41
0.11	0.08	1.96	1.15
0.05	0.07	0.10	0.00
0.15	0.12	0.20	0.20
0.13	0.11	1.18	0.00
0.04	0.05	0.19	0.20
0.07	0.07	0.22	0.15
0.26	0.15	0.13	0.39
0.09	0.08	0.18	0.20
0.06	0.15	0.20	0.30
0.07	0.11	0.10	0.31
0.03	0.06	0.24	0.46
0.00	0.01	0.92	0.51
0.02	0.07	0.07	0.08
0.08	0.09	0.03	0.12
0.22	0.07	0.07	0.08

Analysis

Statistical Test:

The data showed that the average crime rate for public universities was 0.16 with a standard deviation of 0.18, while the average crime rate for private universities was 0.35 with a standard deviation of 0.50.

The null and alternative hypothesis for the study is listed below:

$$H_0: \mu_1 = \mu_2 \text{ (average crime rate for public universities is the same as private universities)}$$

$$H_a: \mu_1 \neq \mu_2 \text{ (average crime rate for public universities is not the same as private universities)}$$

After determining the null and alternative hypothesis, the next step is to calculate the t-test statistic and p-value ($\alpha = 0.05$):

$$S_p = \sqrt{\frac{(60-1)0.18^2 + (60-1)0.50^2}{120-2}} \approx 0.358$$

$$t = \frac{(0.16 - 0.36) - 0}{0.358 \sqrt{\frac{1}{60} + \frac{1}{60}}} = -3.06$$

$$P\text{-value: } \text{tcdf}(-10^{99}, -3.06, 118) \approx 0.001 * 2 = .002$$

Based on the statistical testing procedure with significance level 0.05, we do have sufficient evidence to conclude that there is a significant difference in crime rates, per 100 students, between public and private universities for the years 2011 and 2013.

Confidence Interval Approach:

Since we have determined that there is a significant difference in crime rates amongst public and private universities for the years 2011 and 2013, we will obtain a 95% confidence interval to determine the possible values for the difference in population means μ_1 and μ_2 :

$$(\bar{X}_1 - \bar{X}_2 - t_{df, \alpha/2} * S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \bar{X}_1 - \bar{X}_2 + t_{df, \alpha/2} * S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$$

$$((0.16 - 0.36) - (1.982 * 0.065), (0.16 - 0.36) + (1.982 * 0.065))$$

$$(-0.328, -0.071)$$

With 95% confidence, we can conclude that the average difference in crime rates between public and private universities per 100 students for the years 2011 and 2013 is between -0.328 and -0.071.

This confidence interval helps confirm our findings in the t-test. Recall that we can say our null hypothesis is $\mu_1 = \mu_2$, and can be rejected because the confidence interval does not contain 0.

Linear Regression

The linear regression model was used to determine if there was a linear or non-linear relationship between the average crime rates of public and private universities between the years 2001 and 2013.

The correlation coefficient formula will help measure the direction and strength of the linear relationship:

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

The least squares regression line is the "best line" formula that minimizes errors as much as possible, where a is the intercept and b is the slope:

$$\hat{y} = a + bx$$

$$a = \bar{y} - b\bar{x}$$

$$b = r \frac{s_y}{s_x}$$

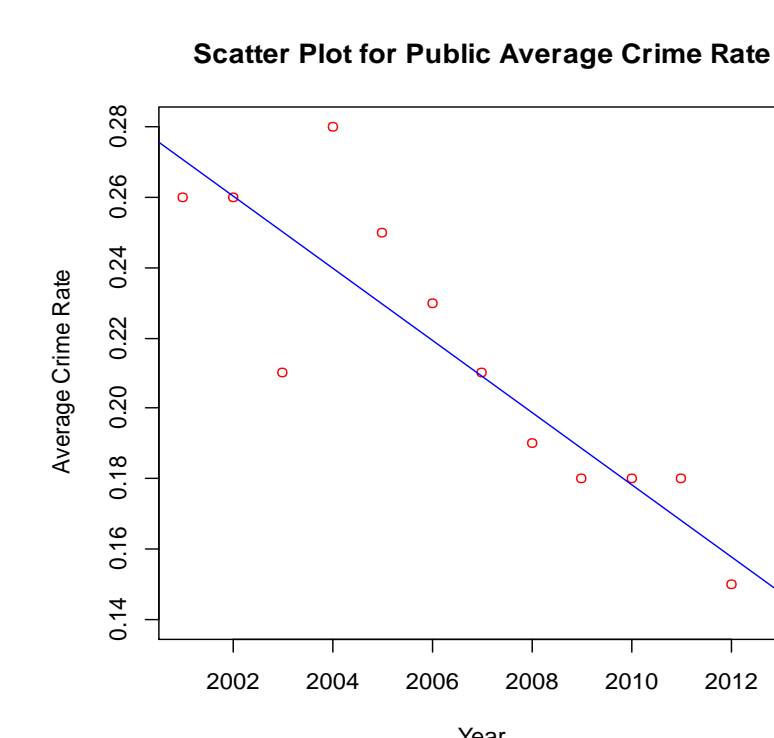
Table 2: Average Crime Rates for Public and Private Universities 2001-2013, Per 100 Students

Year	Public	Private
2001	0.26	0.51
2002	0.26	0.54
2003	0.21	0.63
2004	0.28	0.59
2005	0.25	0.67
2006	0.23	0.65
2007	0.21	0.75
2008	0.19	0.41
2009	0.18	0.54
2010	0.18	0.45
2011	0.18	0.37
2012	0.15	0.31
2013	0.14	0.33

The least squares regression line equations depicted on the graph below:

$$\hat{y} = a + bx$$

$$= 20.83 - 0.010275x$$



$$\hat{y} = a + bx$$

$$= 46.68 - 0.0230x$$

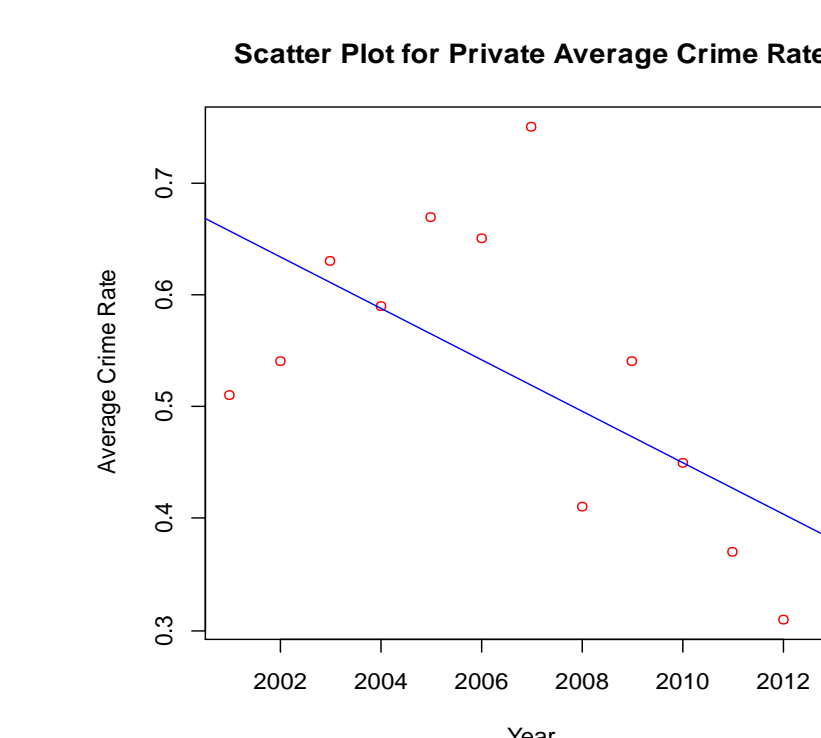


Figure 1: Scatterplots of average crime rates for public and private universities

Polynomial Regression

By looking at the scatterplots for public and private universities, the non-linear trend of private universities indicates an advanced regression model needs to be used.

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$$

$$\hat{y} = -23590 + 23.54 \text{ year} - 0.005869 (\text{year})^2$$

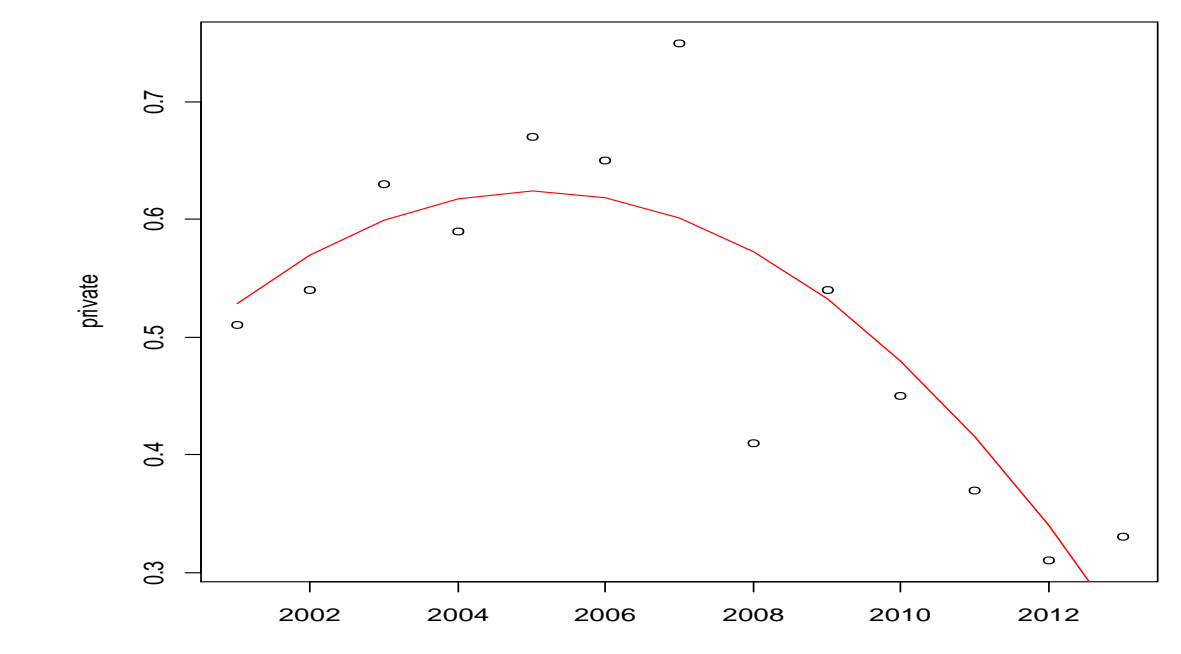


Figure 2: Polynomial Regression Model for Private Universities (2001-2013)

Table 3: Summary output of polynomial regression model from R-software

Coefficients	Estimate	Std. Error	t value	p-value
Intercept	-2.359e+04	7.228e+03	-3.264	0.00851
year	2.354e+01	7.203e+00	3.267	0.00847
year ²	-5.869e-03	1.794e-03	-3.271	0.00842

Residual standard error: 0.08029 on 10 degrees of freedom

Multiple R-squared: 0.7196, Adjusted R-squared: 0.6635
F-statistic: 12.83 on 2 and 10 DF, p-value: 0.001734

This graph shows a downward trend starting around year 2007. Prior to 2007, the average crime rate steadily increased but then started to decline each year after. There could be several contributing factors as to why the average crime rate would increase for the first five years and then start to decline after that. It does not follow the same general downward trend, on average, that public universities experienced during the same timeframe.

Conclusion

Based on the statistical testing procedure, there is a significant difference in average crime rate, per 100 students, for public universities versus private universities for the years 2011 and 2013.

The linear regression model showed that there was a negative linear relationship for public universities from the years 2001 to 2013, but there was a non-linear relationship for private universities over the same time span.

Hence, the polynomial regression model implemented to model the evident relationship.