INFERENTIAL STATISTICAL ANALYSIS OF LAW ENFORCEMENT SURVEILLANCE ON CONSTRUCTION ZONE SAFETY

Tulio Sulbaran, Ph.D [1], David Marchman [2]

Abstract – Every year several hundred people are killed and several thousand people are injured nationwide in roadway construction zones. Construction zones are inherently more hazardous for drivers and workers than non-construction zones. Therefore, states DOTs are continuously implementing measurements to improve the safety of drivers and workers in construction zones. Law enforcement surveillance is one of the special measurements implemented in construction zones by several departments of transportation throughout the United States (U.S.) to reduce the number of crashes. Unfortunately across the U.S., there are limited numbers of statistical analyzes documenting the impact of law enforcement surveillance in construction zones. This paper focuses on the inferential statistical analysis that was performed to measure the impact of law enforcement surveillance in sample construction zones in Mississippi. Six specific statistical analyses were established to determine if there was any correlation between the studied variables. The six analyses were as follows: Analysis 1 - Law Enforcement Presences Vs Number of Citations; Analysis 2 - Law Enforcement Over Time Vs Number of Citations; Analysis 3- Number of Citations per Week Vs Number of Crashes Per Week; Analysis 4- Distribution of Volume Vs Distribution of Crashes; Analysis 5- Time of The Day Vs Number of Crashes and Analysis 6- Law Enforcement Presences Vs Number of Crashes. The analyzes presented in this paper can also be used to support case studies for class discussion and therefore build bridges to make connections between real data and educational experiences that facilitate engineering education. The results presented in this paper reveal the safety impact of law enforcement in construction zones. Additionally, it emphasizes the importance of sharing real life statistical analysis (obtained from applied research experiences) with students to facilitate meaningful learning experiences.

Keywords: Law Enforcement Surveillance, Construction Zone, Statistics, Inferential

INTRODUCTION

Construction zone's fatalities are on the rise and are likely to continue climbing across the nation as departments of transportation continue repairing and upgrading the United States' aging roadways [Safe Roads 2003]. This is particularly compounded in Mississippi due an all time peek volume of construction zones as well as the 1987 four-lane highway program and TEA-21 [Young 2001]. Therefore, it is very important to implement programs such as the law enforcement surveillance that aim to improve the safety records of construction zones. Unfortunately, in many cases these programs are implemented without the proper collection, processing, archiving and analysis of the data to evaluate their impact on safety.

[1] Assistant Professor – School of Construction at the University of Southern Mississippi, Box 5138, Hattiesburg, MS, 39406. E-mail: Tulio.Sulbaran@usm.edu.

[2] Professor – School of Construction at the University of Southern Mississippi, Box 5138, Hattiesburg, MS, 39406. E-mail: David.Marchman@usm.edu.

Several agencies contributed to the data collection including District 5 Office of the Mississippi Department of Transportation, Planning Division of the Mississippi Department of Transportation, Mississippi Highway Patrol, Ridgeland Police Department, and Traffic Engineering Division of Mississippi Department of Transportation. Upon receiving the data from the different agencies, the data was restructured and consolidated to server as the foundation for descriptive and inferential statistics which are presented in this paper.

Overview of Statistical Analysis

Statistic analysis pertains to collection, analysis, interpretation, and presentation of data as well as drawing valid conclusions and making reasonable decisions on the basis of such analysis [Wikipedia 2006]. In most research projects the statistical analysis involves three major steps, done in roughly this order: Cleaning and organizing the data for analysis (Data Preparation), Describing the data (Descriptive Statistics), Testing Hypotheses and Models (Inferential Statistics)

Data Preparation. It involves checking or logging the data in; checking the data for accuracy; entering the data into the computer; transforming the data; and developing and documenting a database structure that integrates the various measures.

Inferential Statistics: Focus on trying to reach conclusions that extend beyond the raw data. Inferential statistics are used to make inferences from the descriptive statistics to more general conditions; where the descriptive statistics simply is used to describe what's going on with the data. The inferential statistical "tools" available for use within SPSS are Chi-square, *T* test, Regression, General Linear Model, and Correlation. [SPSS 2006]

Statistical Software

A statistical package is a computer application that is specialized for statistical analysis. It enables the research teams to obtain the results of standard statistical procedures and statistical significance tests, without requiring low-level numerical computations or programming. In addition to provide the results of standard statistical procedures, statistical packages provide facilities for data management [Wikipedia 2006]. There several commercially available statistical packages in the market, the following is a list of some of the packages available: AM Software, Bascula, CENVAR, CLUSTERS, Epi Info, Generalized Estimation System (GES), IVEware, PCCARP, R survey package, SAS/STAT, Stata, SUNDANN, VPLX, WesVar, and SPSS

The statistical package used for the statistical analysis in this project was SPSS. The version of SPSS used 13.0 for Windows. The primary reason for using this software was that it met the needs of the project was the statistical software most commonly used at the University of Southern Mississippi. The following section provides a description of SPSS.

SPSS Description

SPSS for Windows is a statistical and data management package for analysts and researchers. SPSS for Windows provides a broad range of capabilities for the entire analytical process. SPSS Inc. is a leading worldwide provider of predictive analytics software. They have been in business for more than 37 years, and have more than 120,000 customers (academic institutions, healthcare providers, market research companies and government agencies) [SPSS 2006].

Inferential Statistic

Based on the main objective of the project (evaluate the safety impact of increased law enforcement surveillance on construction zones), the available data and the results from the descriptive statistics. Six specific statistical analyses were established aiming to determine if there was any correlation between the studied variables. The six analyses were as follows:

- Analysis 1 - Law Enforcement Presences Vs Number of Citations:
- Analysis 2 - Law Enforcement Over Time Vs Number of Citations:
- Analysis 3- Number of Citations per Week Vs Number of Crashes Per Week
- Analysis 4- Distribution of Volume Vs Distribution of Crashes
- Analysis 5- Time of The Day Vs Number of Crashes
- Analysis 6- Law Enforcement Presences Vs Number of Crashes

Analysis 1 - Law Enforcement Presences Vs Number of Citations. This analysis focused on how much the presence of law enforcement (either MHP or RPD) impacted the number of citations issued in the studied area. The analysis was based on the data obtained from MHP, RPD and MDOT. Figure 2, shows a linear chart plotting number of citations (count) vs. time (week date) in the studied area during the time that law enforcement was present on the studied area. This chart was created in SPSS using the sequence of steps shown in Figure 3.

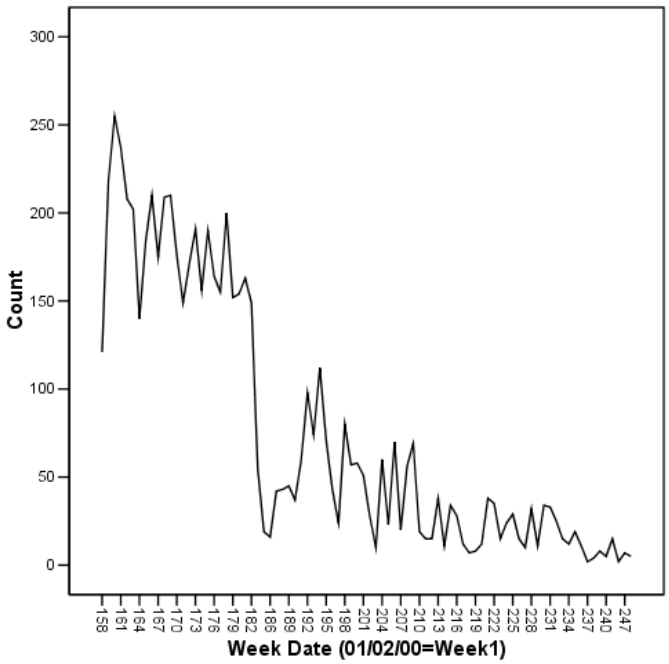


Figure 2. Number of Citations Overtime during

Law Enforcement Presence in the Studied Area

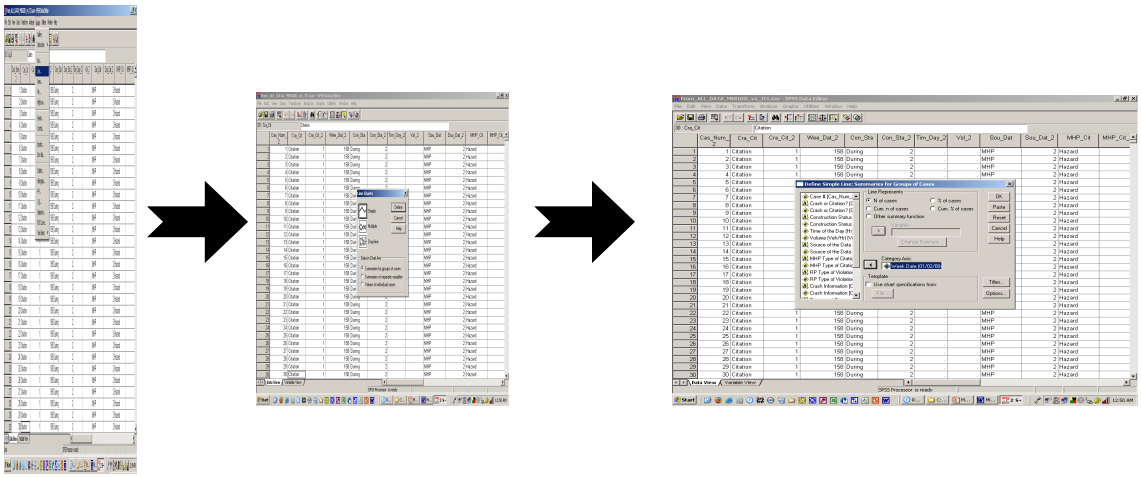


Figure 3. SPPS Screen Shoots of Chart Steps

In order to measure the impact of law enforcement on the number of citations, data prior to the presence of law enforcement present was also collected, plotted and analyzed. Figure 4, shows a linear chart plotting number of citations (count) vs. time (week date) in the studied area from January 2000 till the last day of law enforcement presence in the studied area. This chart was created in SPSS using the sequence of steps shown in Figure 3.

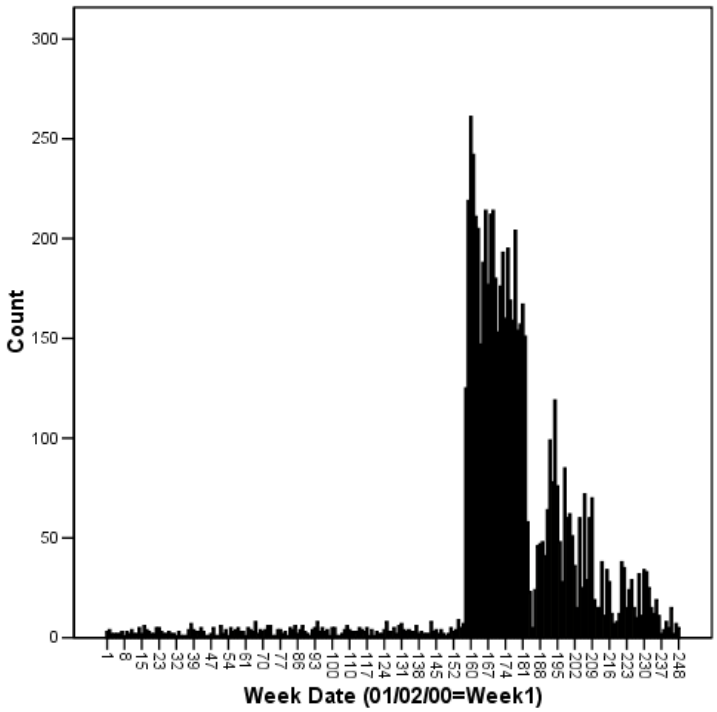


Figure 4. Number of Citations Overtime in the Studied Area

A t-test was performed to measure the correlation between number of citation issued and the presence of law enforcement. The variables included in the analysis were: Week Date, Number of Citation and Law Enforcement presence as show in Figure 5. This analysis was performed in SPSS using the sequence of steps shown in Figure 6.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Wee_Dat_2	Numeric	11	0	Week Date (01/02/00=Week	None	None	11	Right	Ordinal
2	Cas_Num_	Numeric	8	2		None	None	9	Right	Scale
3	Num_Case	Numeric	7	0	Number of Cases	None	None	7	Right	Scale
4	Law_Enf	Numeric	8	0	Permanent Law Enforcement	{0, No Perman	None	8	Right	Nominal
5										
6										
7										

Figure 5. Variable used in the Analysis

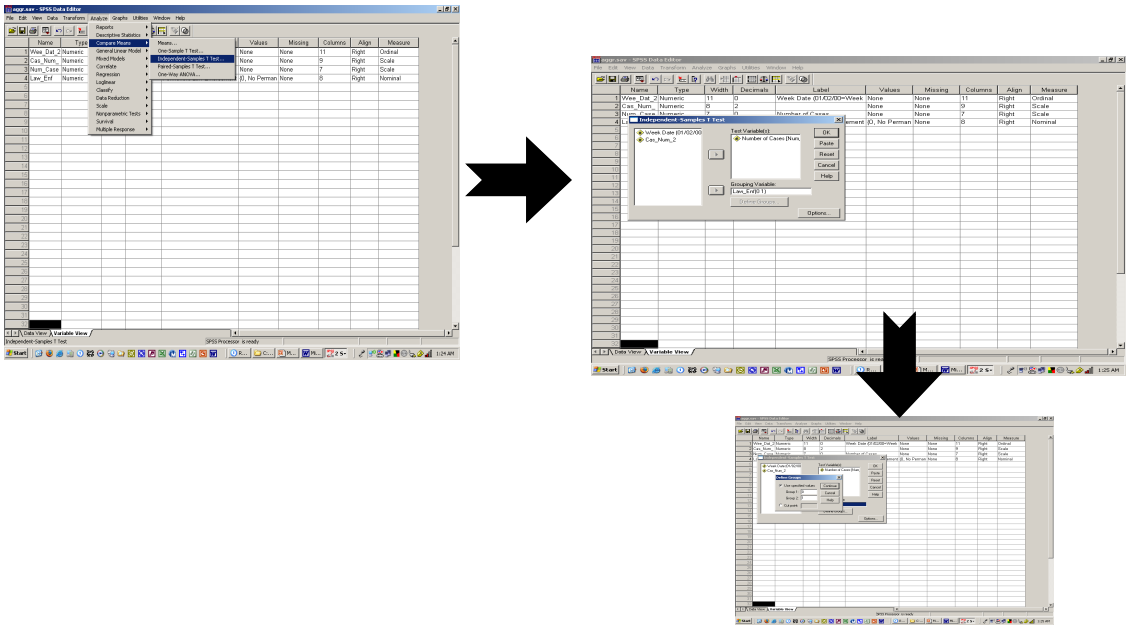


Figure 6. SPPS Screen Shoots of Analysis Steps

Table 1 shows a summary of the data analyzed. As shown in Table 1, there was data for 148 weeks of no permanent law enforcement presence and 84 weeks of law enforcement (either MHP or RP) permanent presence. The analysis of the data indicated that there was an average of 3.45 citations per week issued during the non permanent law enforcement period and 79.11 average citations per week issued during the law enforcement permanent presence. The standard deviation regarding number of citation issues were 2.015 and 74.919 respectively and the standard error mean were .166 and 8.174 respectively.

Table 1. Summary of Data Analyzed Number of Citations

	Permanent Law Enforcement	N	Mean	Std. Deviation	Std. Error Mean
Number of Cases	No Permanent Law Enforcement	148	3.45	2.015	.166
	Either MHP or RP Permanent	84	79.11	74.919	8.174

The t-test shown in Table 2 showed a significance level below 0.1% that the compared groups/conditions (No Permanent Law Enforcement Vs. Permanent Law Enforcement) were different. As stated by Glenberg, values of test statistics that occur with a relative frequency (Sig.) of less than 5% are in the rejection region [Glenberg 1996]. The rejection region means that the null hypothesis (no difference between groups/conditions) can be rejected, thus there is a difference between groups/conditions. Therefore, with a significance level of less than 0.1%, the null hypothesis (which is that there is no difference between the groups/conditions) is rejected. This less than 0.1% means that only in less than 1/1000 cases in which the true means (number of citations) were the same; the sample will show results as extreme as the one observed here. Thus, it can was concluded that there is statistically significant difference in the number of citations between the period with no permanent law enforcement and the period with permanent law enforcement in the studied area.

Table 2. t-test Law Enforcement Presence Vs. Number of Citations

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number of Cases	Equal variances assumed	434.407	.000	-12.297	230	.000	-75.654	6.152	-87.776	-63.533
	Equal variances not assumed			-9.253	83.068	.000	-75.654	8.176	-91.916	-59.393

Analysis 2 - Law Enforcement Over Time Vs Number of Citations. This analysis focused on determining the effect of the number of weeks that the law enforcement agencies (MHP and RP) stayed the in the studied area on the number of citations issued per week. Figure 5-7 shows the number of citations issues per week the area. The square 1 corresponds to one of the law enforcement agencies and the area in the square 2 corresponds to the other law enforcement agency.

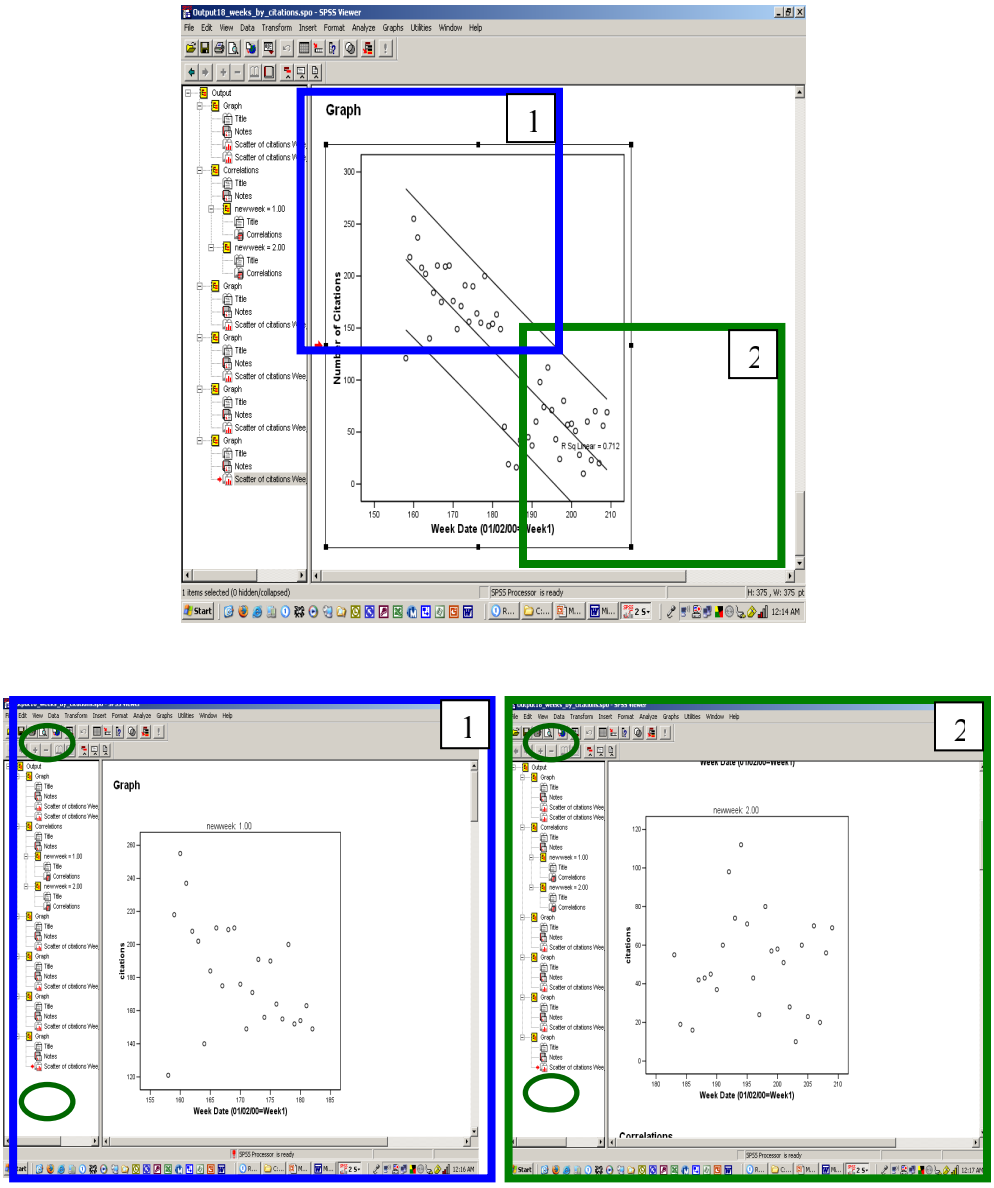


Figure 7. Number of Citations Issued in the Studied over time

A Bivariate Correlations was performed between number of citations over time to computes Pearson's correlation coefficient with its significance levels. As previously described, Pearson's correlation coefficient is a measure of linear association. The variables included in the analysis were: Week Date and Number of Citations. This analysis was performed in SPSS using the sequence of steps show in Figure 5-8.

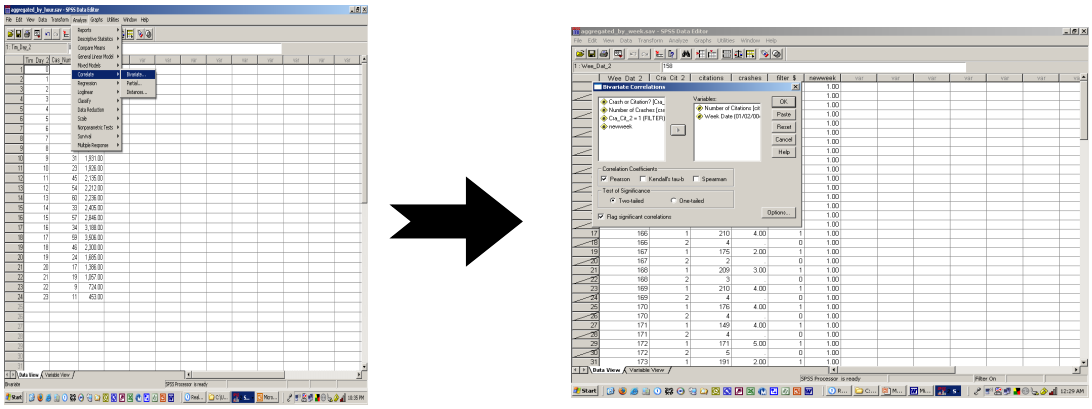


Figure 5-8. SPPS Screen Shoots of Analysis Steps

The correlation analysis shown in Table 3 showed a significance of less than 0.01% that the compared variables (number of citation and week date) were related with a pearson correlation factor of -.844 which indicates that the longer that a law enforcement agency is in the studied area the lower the number of citations issued.

Table 3. Correlation of Citations Over Time

Correlations

		Number of Citations	Week Date (01/02/00= Week1)
Number of Citations	Pearson Correlation	1	-.844**
	Sig. (2-tailed)	.	.000
	N	51	51
Week Date (01/02/00=Week1)	Pearson Correlation	-.844**	1
	Sig. (2-tailed)	.000	.
	N	51	51

**. Correlation is significant at the 0.01 level (2-tailed).

In addition to the overall bivariate correlation, a bivariate correlation was performed for each law enforcement agency (MHP and RP). The correlation analysis shown in Table 4 showed a significance of 0.02% that the compared variables (number of citation and week date) were related with a pearson correlation factor of -.449 which indicates that the longer that this law enforcement agency was in the studied area the lower the number of citations issued. The correlation analysis shown in Table 5 showed that there is not a strong pearson correlation factor between the number of weeks that this second law enforcement agency was in the studied area and the number of citations issued. It is worth noting that this different in correlation between the two law enforcement agency could be attributed to many factors not linked to the enforcement agency such as: number of hours, days of operation, sequence of law enforcement agencies in the studied area, just to mention a few of the possible factors that were not considered in this research project.

Table 4. Correlation of Citations Over Time

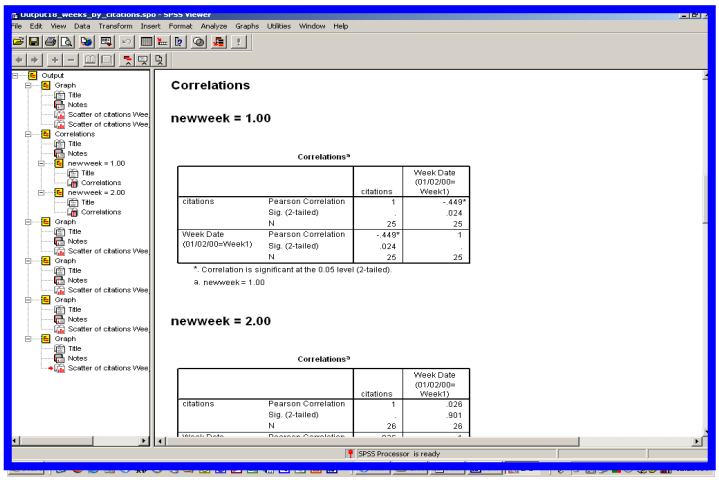
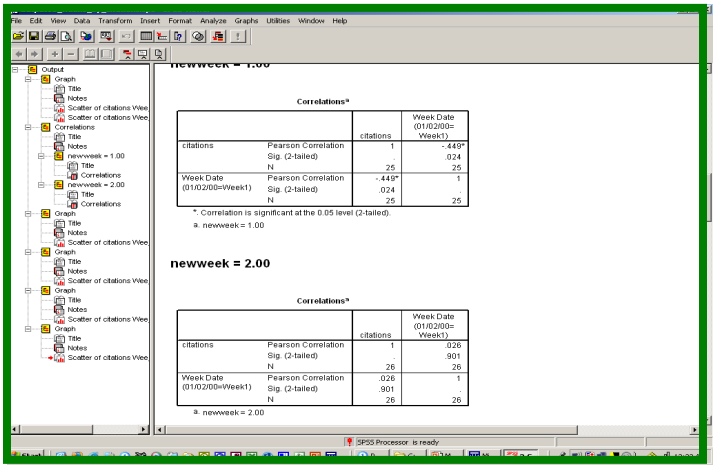


Table 5. Correlation of Citations Over Time



Analysis 3- Number of Citations per Week Vs Number of Crashes Per Week. This analysis focused on determining the effect of the number citations per week on the number of crashes per week on the studied area. Figure 9 shows graphs that plots the number of citations issues per week vs. number of crashes.

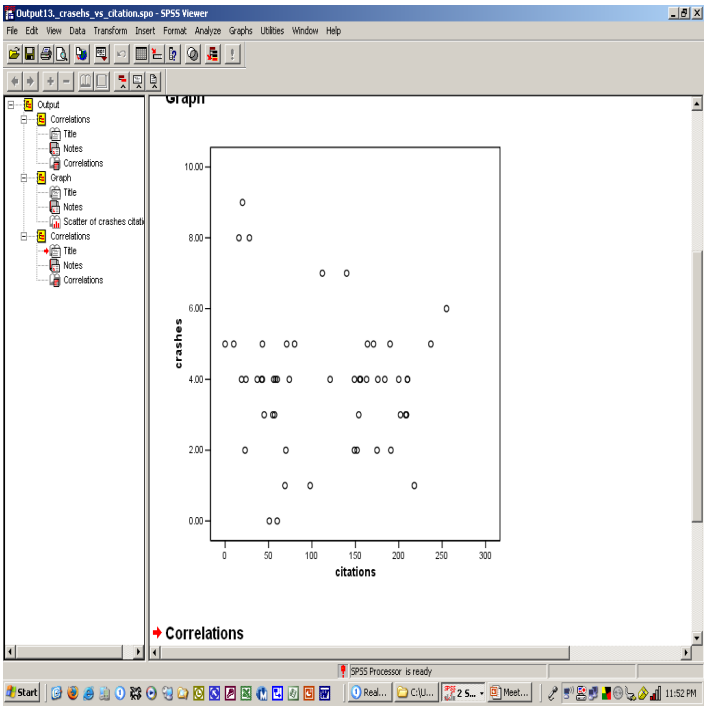


Figure 9. Number of Crash Vs Number of Citations in the Studied Overtime

A Bivariate Correlations was performed between number of citations per week and the number of crashes per week to computes Pearson's correlation coefficient with its significance levels. This analysis was performed in SPSS using the sequence of steps show in Figure 10.

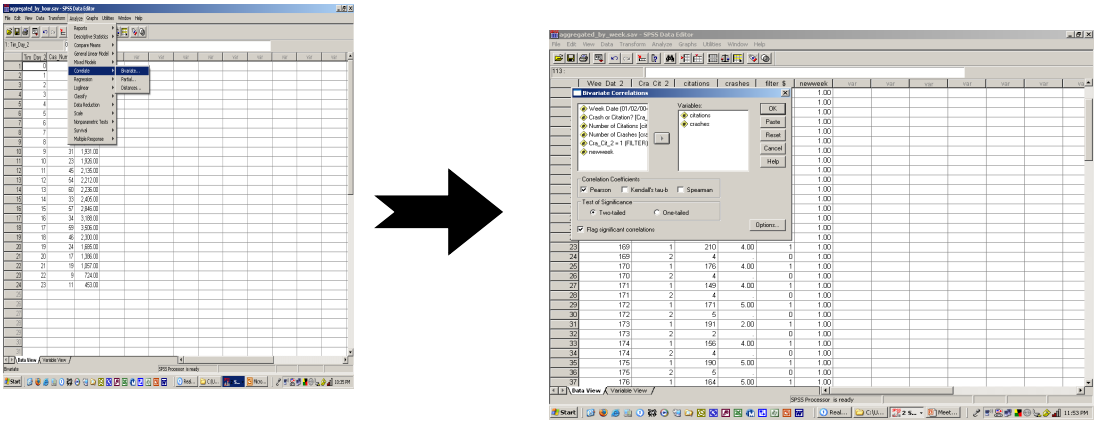


Figure 10. SPPS Screen Shoots of Analysis Steps

The correlation analysis shown in Table 6 showed that there is not a strong pearson correlation factor between the number citations issued by the law enforcement agencies and the number of crashes in the studied area

Table 6. Correlation of Citations Overtime

Correlations

		citations	crashes
citations	Pearson Correlation	1	-.125
	Sig. (2-tailed)	.	.379
	N	101	52
crashes	Pearson Correlation	-.125	1
	Sig. (2-tailed)	.379	.
	N	52	52

Analysis 4- Distribution of Volume Vs Distribution of Crashes. This analysis focused on determining the effect of the traffic volume on the number of crashes in the studied area. The crashes reported in the studied area during the period of the studied were grouped by the time of the day that they occurred and then compared against the average hourly volume for that particular time of the day. Figure 11 shows the number of crashes compared to the traffic volume at the moment of the crash.

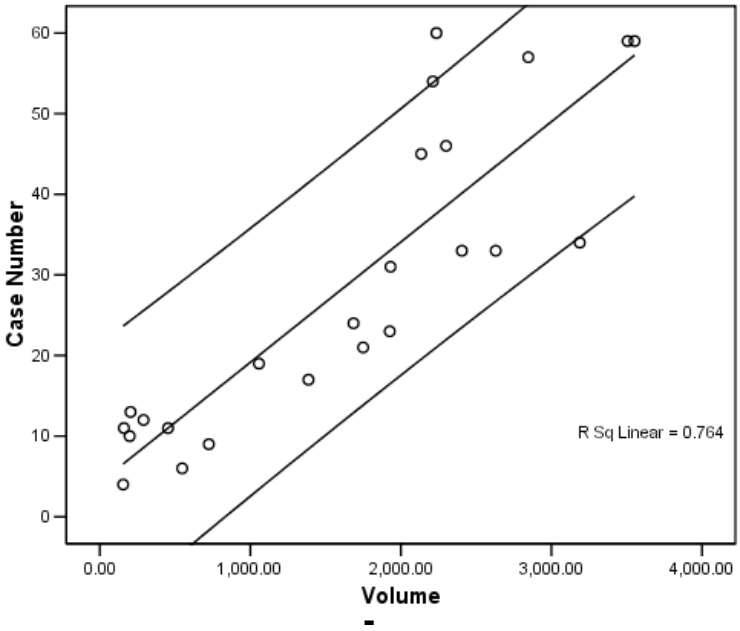


Figure 11. Number of Crashes Vs Traffic Volume

A Bivariate Correlations was performed between number of crashes and traffic volume to computes Pearson's correlation coefficient with its significance levels. This analysis was performed in SPSS using the sequence of steps show in Figure 12.

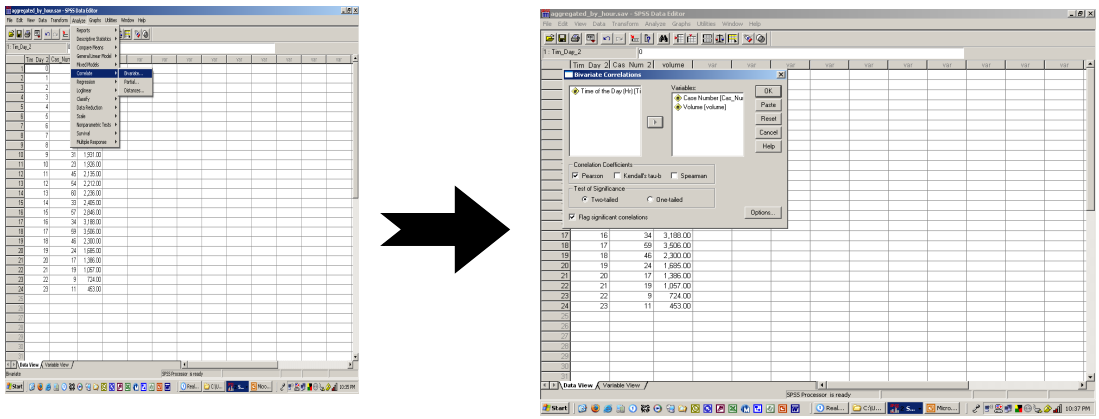


Figure 12. SPPS Screen Shoots of Analysis Steps

The correlation analysis shown in Table 7 showed a significance of less than 0.01% that the compared variables (number of crashes and traffic volume) were related with a Pearson correlation factor of .874 which indicates that the higher the traffic volume higher the number of crashes in the studied area

Table 7. Correlation of Number of Crashes (Cases) Vs. Traffic Volume

Correlations

		Case Number	Volume
Case Number	Pearson Correlation	1	.874**
	Sig. (2-tailed)	.	.000
	N	24	24
Volume	Pearson Correlation	.874**	1
	Sig. (2-tailed)	.000	.
	N	24	24

**. Correlation is significant at the 0.01 level (2-tailed).

Analysis 5- Time of the Day Vs Number of Crashes. This analysis focused on identifying the variation of traffic volume over the time of the day and it implication on the number of crashes. The previous analysis established that a there was a direct correlation between traffic volume and number of crashes in the studied area. By plotting the traffic volume and number of crashes, it was observed that this direct correlation held constant through the day with exceptions of the period following the lunch hour until the end of the normal business day as show in Figure 13.

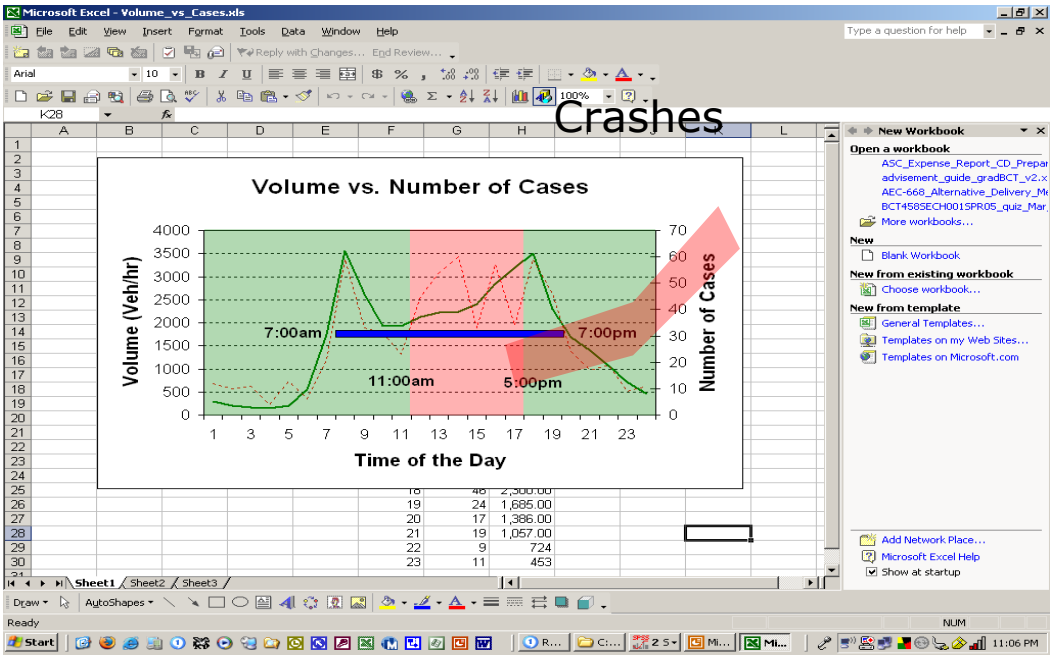


Figure 13. Time of the Day Vs. Traffic Volume and Number of Crashes

Analysis 6- Law Enforcement Presences Vs Number of Crashes. A t-test was performed to measure the correlation between number of crashes and the presence of law enforcement. The variables included in the analysis were: Week Date, Number of crashes and Law Enforcement presence as show in Figure 514. This analysis was performed in SPSS using the sequence of steps shown in Figure 15.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Wee_Dat_2	Nume	11	0	Week Date (01/02/00=Week1)	None	None	11	Right	Ordinal
2	Cra_Cit_2	Numeric	1	0	Crash or Citation?	{1, Citation}...	None	9	Right	Nominal
3	Per_MHP_	Numeric	1	0	Permanent Presence of MHP	{1, No}...	None	9	Right	Nominal
4	Per_RP_2	Numeric	1	0	Permanent Presence of RP	{1, No}...	None	8	Right	Nominal
5	Con_Sta_2	Numeric	1	0	Construction Status	{1, After}...	None	9	Right	Ordinal
6	Number_of_	Numeric	7	0	Number of Crashes	None	None	28	Right	Scale
7	Presence	Numeric	8	2	Law Enforcement Permanent Presence	{1.00, No}...	None	10	Right	Scale
8	filter_$	Numeric	1	0	Cra_Cit_2 = 2 (FILTER)	{0, Not Selecte	None	10	Right	Scale
9										

Figure 14. Variable used in the Analysis

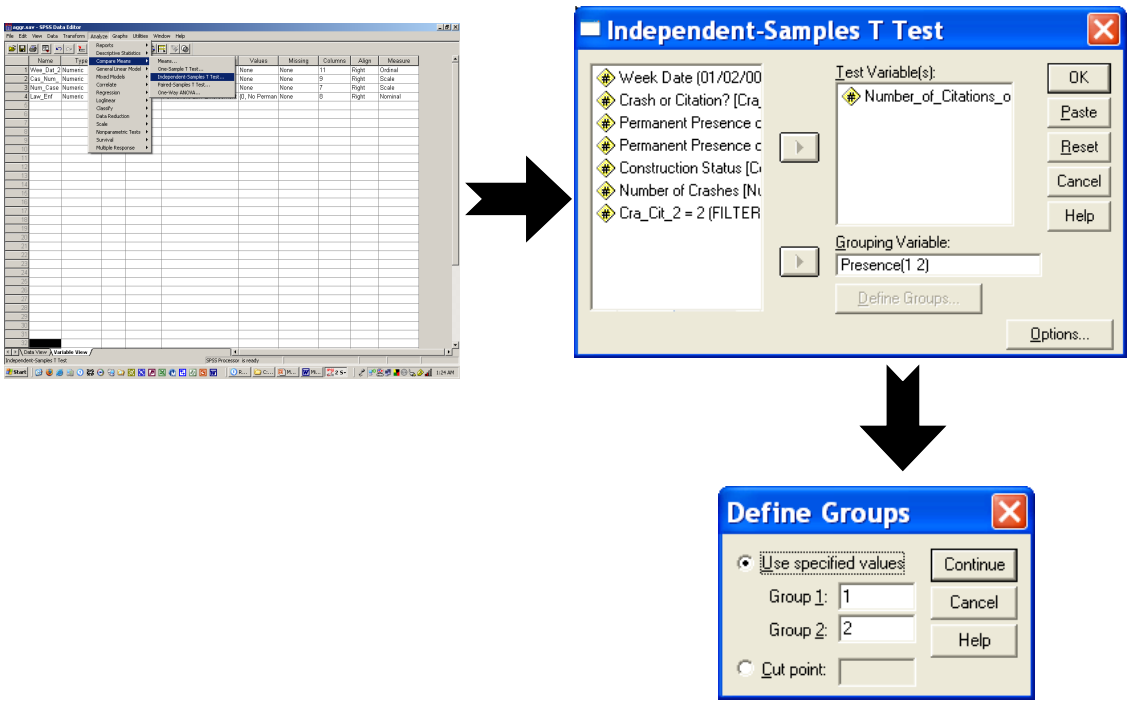


Figure 15. SPPS Screen Shoots of Analysis Steps

An assumption that was made during the analysis was that any date without crash information was considered to have 0 crashes. Table 8 shows a summary of the data analyzed. As shown in Table 8, there was crash data for 155 weeks of no permanent law enforcement presence and 53 weeks of law enforcement (either MHP or RP) permanent presence. The analysis of the data indicated that there was an average of 3.11 crashes per week during the non permanent law enforcement presence period and an average of 3.94 crashes per week during the law enforcement permanent presence period. The standard deviations regarding number of crashes were 1.905 and 1.895 respectively and the standard error mean were .153 and .260 respectively.

Table 8 Summary of Data Analyzed Number of Citations

Group Statistics

	Law Enforcement Permanent Presence	N	Mean	Std. Deviation	Std. Error Mean
Number of Crashes	No	155	3.11	1.905	.153
	Yes	53	3.94	1.895	.260

The t-test shown in Table 9 showed a significance level of .347 which it is above 5%. As stated by Glenberg, values of test statistics that occur with a relative frequency (Sig.) of less than 5% are in the rejection region [Glenberg 1996]. The rejection region means that the null hypothesis (no difference between groups/conditions) can be rejected, thus there is a difference between groups/conditions. Therefore, with a significance level of .347 which is more than 0.05, the null hypothesis (which is that there is no difference between the groups/conditions) is not rejected.

Thus, it can was concluded that there is no statistically significant difference in the number of crashes between the period with no permanent law enforcement and the period with permanent law enforcement in the studied area.

Table 9. t-test Law Enforcement Presence Vs. Number of Citations

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Number of Crashes	Equal variances assumed	.888	.347	-2.753	206	.006	-.834	.303	-1.431	-.237
	Equal variances not assumed			-2.760	90.496	.007	-.834	.302	-1.434	-.234

LESSONS LEARNED

It is worth noting that this first project from the Mississippi Department of Transportation to quantitatively document the safety impact of increased law enforcement surveillance on highway construction was a success. It provided quantitative evidences of the program effectiveness. It also helps develop a sample process to evaluate other programs in the future and identify the data required for those evaluations. Some of the data required for similar evaluations include: timeframe of the intervention (in this case the law enforcement presence in the studied area), number of citations issued by law enforcement agencies over time, number of crashes in the studied area over time and hourly traffic volume in the studied area over time. It was also evident (based on the statistical analysis) that the most useful inferential statistical analysis for the intended analysis were the t-test and bivariate correlation. Furthermore, histograms, line charts, and scatter plots seems to be the most practical type of chart to present the gathered data.

SUMMARY

One of the special measures implemented, in construction zones by several departments of transportation around the United States, to reduce the number of crashes is the increase of law enforcement surveillance. This chapter focuses on the descriptive and inferential statistical analysis to quantify the impact of law enforcement in construction zones.

The results presented in this chapter indicate that the permanent presence of law enforcement agencies in the studied area significantly increased the number of citations issued. This number of citations is reduced as law enforcement agencies stay in the studied area over time. It was also determined that there was not a direct correlation between the number of citations issued and the number of crashes. The number of crashes however was directly related to the traffic volume in the studied area.

It is also expected that the results and process presented in this paper could be used by other research teams to perform similar analysis of law enforcement surveillance or others methods implemented around the U.S. to reduce the deaths and injuries in road construction zones.

References

Safe Roads (2003) “Safe Roads for Road Workers”, http://www.saferoadsforworkers.org/ Last visited: June 2003

SPSS, “Descriptive and Inferential Statistics, (2006) http://www.utexas.edu/its/rc/tutorials/stat/spss/spss2/#Inferential, Last visited: September 2006

StatSoft, (2006) “Basic Statistics”, http://www.statsoft.com/textbook/stbasic.html#Descriptive%20statisticsb, Last visited: September 2006

Young, T., (2001). “Work Zone Safety Awareness Week Designated April 9-12, 2001”, http://www.mdot.state.ms.us/news/2001releases/release033001.htm. Last visited: May 2003.