4:16CV180-CDP

Racial Disparity of Traffic Stops in St. Louis

Presented by Keith Kallstrom

Racial Disparity of Traffic Stops in St. Louis

This paper will summarize racial disparities of traffic stops within a random sample of 15 local St. Louis municipalities. The disparity index is a measure of the likelihood that a driver of a specific race or ethnic group would be stopped in comparison to the residential population over the age of 16. The disparity index is determined by the following equation:

$$\text{Disparity Index} = \frac{\text{Proportion of black stops}}{\text{Proportion of blacks in the population}}$$

A value greater than 1 signifies an over-representation and values less than 1 signify an under-representation in traffic stops. The ideal disparity index value would be 1, which would indicate an equal representation of blacks and whites being stopped in comparison to the overall demographic makeup of the area. The information used in this study was gathered from the Missouri Attorney General's website.

Police Municipality	2013 Disparity
Beverly Hills Police Dept.	1.01
Breckenridge Hills Police Dept.	1.15
Clarkson Valley Police Dept.	2.65
Edmundson Police Dept.	1.61
Glendale Police Dept.	10.72
Maplewood Police Dept.	2.59
Maryland Heights Police Dept.	2.58
Northwoods Police Dept.	0.54
Riverview Police Dept.	1.36
Rock Hill Police Dept.	0.52
Shrewsbury Police Dept.	4.19
St. Ann Police Dept.	1.69
Town and Country Police Dept.	5.02
Uplands Park Police Dept.	0.8
Winchester Police Dept.	2.98

Each aforementioned municipality name contains an active link to the data source. Information on the site reflects data as reported by each individual law enforcement agency. As noted by the Attorney General's office, when reported numbers appeared incomplete or inaccurate, every effort was made to verify the data with the reporting agency in question. The following tables represent frequency distribution. Below are frequency tables, histograms, an ogive, and a stem-and-leaf plot to visualize the data.

Frequency Table

2013 Disparity Rates for black drivers	
Disparity value	Frequency
0.5 - 1.74	8
1.75 - 2.99	4
3 - 4.24	1
4.25 - 5.49	1
5.5 - 6.74	0
6.75 - 7.99	0
8 - 9.24	0
9.25 - 10.75	1

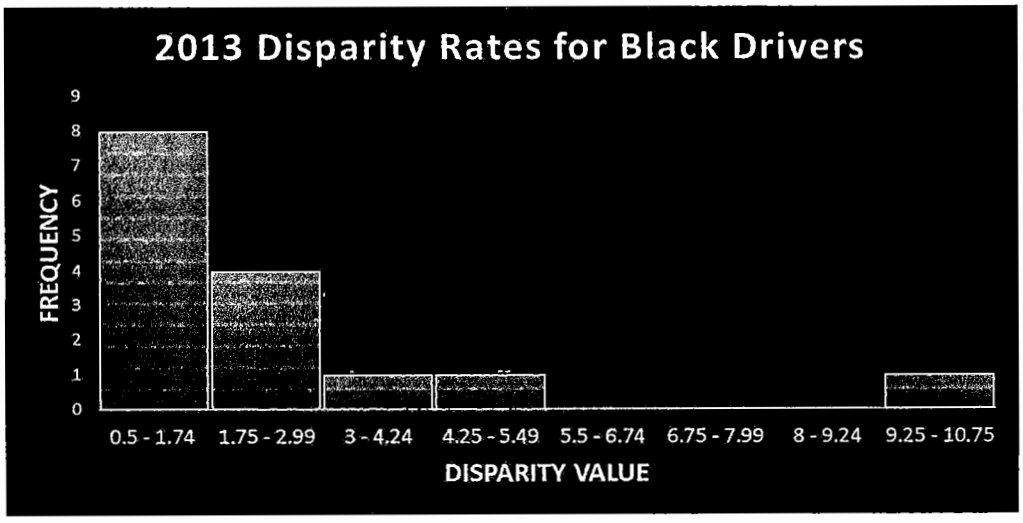
Relative Frequency Table

2013 Disparity Rates for black drivers	
Disparity value	Relative Frequency
0.5 - 1.74	0.53
1.75 - 2.99	0.27
3 - 4.24	0.07
4.25 - 5.49	0.07
5.5 - 6.74	0.00
6.75 - 7.99	0.00
8 - 9.24	0.00
9.25 - 10.75	0.07

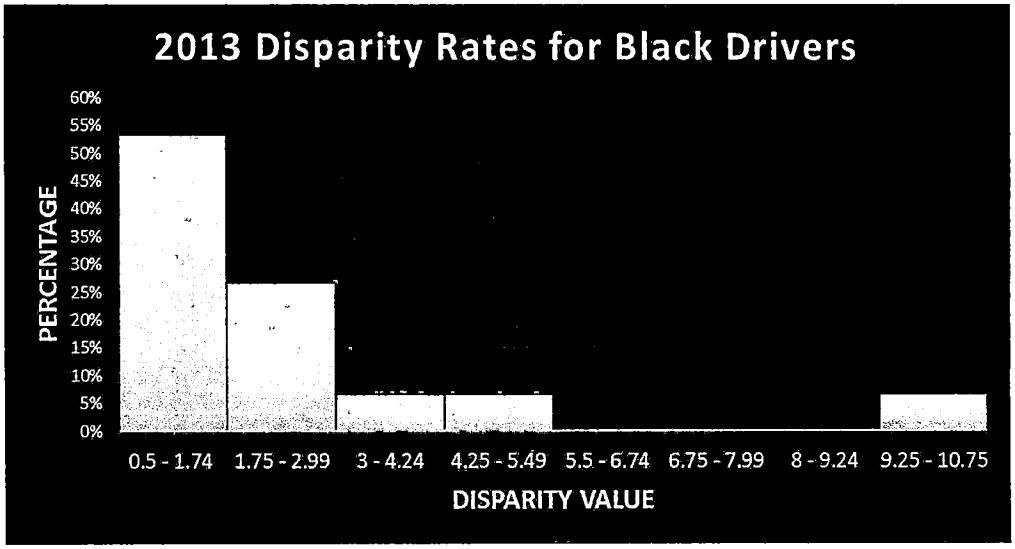
Cumulative Frequency Table

2013 Disparity Rates for black drivers	
Disparity value	Cumulative Frequency
0.5 - 1.74	8
1.75 - 2.99	12
3 - 4.24	13
4.25 - 5.49	14
5.5 - 6.74	14
6.75 - 7.99	14
8 - 9.24	14
9.25 - 10.75	15

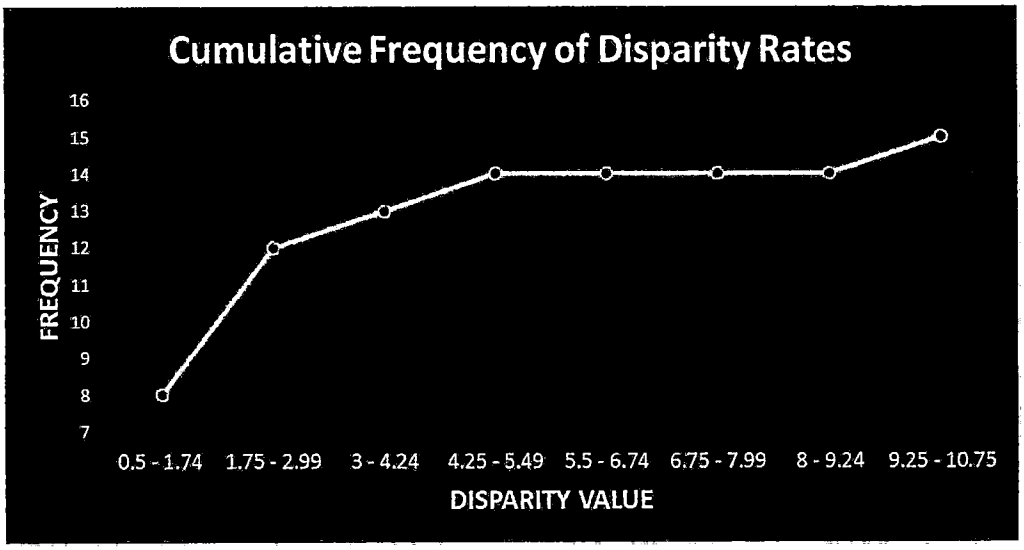
Histogram



Relative Frequency Histogram



Ogive



Stem and Leaf Plot

Stem and Leaf Plot Disparity Value

0	52	54	80		
1	01	15	36	61	69
2	58	59	65	98	
3					
4	19				
5	02				
6					
7					
8					
9					
10	72				

Legend:

0 52 = 0.52

1 01 = 1.01

Based on this data, we can assume the following central tendencies are correct:

Mean	2.627
Median	1.690
Mode	N/A
Midrange	5.880

The preceding graphs do not indicate that the results followed normal distribution. Normal distribution will illustrate a bell-shaped histogram. This histogram was skewed drastically to the right, or *positively skewed.* Distributions skewed to the right are more common than those to the left, because it is more common to see exceptionally larger values than smaller ones. Another consideration is the single outlier (sample value that lies very far away from the vast majority of the other sample values) of 10.72, from the Glendale Police Department.

The range, mean absolute deviation, and standard deviation are as follows:

Range	10.200
Mean Absolute Deviation	1.656
Standard Deviation	2.595

The mean absolute deviation of this data was 1.656. On average, the disparity index was 1.656 above and below the average of 2.63. This result was determined by the following equation:

$$\frac{\Sigma|x-\bar{x}|}{n}=\frac{24.845}{15}=1.656$$

Based on these calculations, the empirical rule asserts that roughly 68% of the data should fall between 0.037 and 5.217, 95% between -2.553 and 7.807, and roughly 99.7% between -5.143 and 10.397.

$$\bar{x}=2.627$$
$$s=2.59$$
$$\bar{x}-s=.037 \quad x+s=5.217$$
$$\bar{x}-2s=-2.553 \quad \bar{x}+2s=7.807$$
$$\bar{x}-3s=-5.143 \quad \bar{x}+3s=10.397$$

Estimating the standard deviation, 2.59, shows a result very close to the actual calculated standard deviation, which was 2.55. The range rule of thumb was also a good estimate for the standard deviation. Range Rule of Thumb is a summary based on the principle that for many data sets, the vast majority (approximately 95%) of sample values lie within 2 standard deviations of the mean.

$$s=\frac{\text{range}}{4}=\frac{10.2}{4}=2.55$$

The z-score of the highest and lowest data elements was 3.12 and -.81, respectively. These z-scores give a good indication as to whether or not there are unusually high or low values. The highest data value of 10.72, from the Glendale Police Department, is unusual. It is

more than 2 standard deviations above the mean. The lowest data value of .52, from the Rock Hill Police Department, is not unusual because it is well within 2 standard deviations of the mean. These scores suggest that while Rock Hill pulls over black drivers at a lower average than white drivers, Glendale pulled over a much higher percentage of black drivers than any other municipality selected for the sample.

Glendale

$$z = \frac{x - \bar{x}}{s} = \frac{10.72 - 2.63}{2.59} = \frac{8.09}{2.59} = 3.12$$

Rock Hill

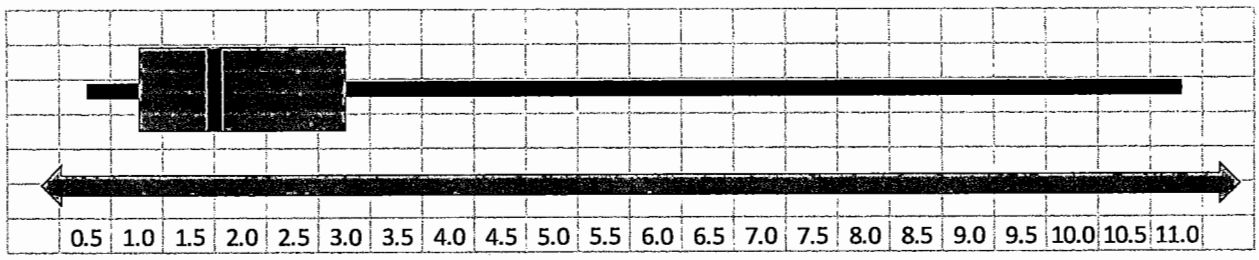
$$z = \frac{x - \bar{x}}{s} = \frac{.52 - 2.63}{2.59} = \frac{-2.11}{2.59} = -.81$$

The z-score of the mean of the data is zero. Naturally, zero falls in the exact center of -2 and +2. Subsequently, a score of zero represents the mean score.

$$\mathbf{z} = \frac{x - \bar{x}}{s} = \frac{2.627 - 2.627}{2.59} = \frac{0}{2.59} = 0$$

Below is interpretation of the 5-number summary expressed through a boxplot (or box and whisker diagram). This graph shows the lowest value, first quartile, median (or second quartile), third quartile, and the highest value.

Lowest value	0.52
Q_1	1.01
Median	1.69
Q_3	2.98
Highest value	10.72



Summary

An overall conclusion based on the sample data of these 15 municipalities suggests that there is a greater than average disparity among the rates of black drivers pulled over for traffic stops. Only one, or 7%, of the municipalities had a disparity rate close to the desired value of 1. This was the Beverly Hills Police Department. Three of the municipalities had rates below significantly lower than desired: Northwoods, Uplands Park, and Rock Hill. 11 had rates significantly higher than desired: Breckenridge Hills, Clarkson Valley, Edmundson, Maplewood, Maryland Heights, Riverview, Shrewsbury, St. Ann, Town and Country, Winchester, and especially Glendale, with a grossly highest disparity of 10.72. Further study of the demographic makeup of each of these municipalities would further quantify the data, but it is suggestive that blacks are pulled over much more frequently per capita, than whites. I believe that these results can be logically extrapolated to the entire St. Louis County region, because the municipalities represent a wide geographic sample of the population. There is a relatively equal distribution of western, southern, and northern municipalities. When comparing these samples to the disparity index of Ferguson (1.37) it becomes obvious that there are other municipalities with far higher indications of racial profiling during traffic stops.

References

Missouri Attorney General Chris Koster. (2014). Retrieved January 31, 2015, from http://ago.mo.gov/VehicleStops/Results.php?coId=95

Triola, M. (2015). Statistics for Describing, Exploring, and Comparing Data. In *Essentials of statistics* (5th ed., pp. 43-141). Boston: Pearson Addison Wesley.