Speed Monitoring Program Methodology

Introduction

Minnesota Speed Monitoring Program entails analysis and reporting of roadway speeds on five roadway types throughout Minnesota. These reports track any changes in roadway speeds overtime. These processes were once conducted in a Paradox program – a DOS based system that required considerable manual input. Beginning in 2009 a statistical software package has been used that has the capacity for complex analyses on very large datasets. Customized syntax was written for consistent and accurate data processing, analyses, and reporting while minimizing the risk of human errors. For any given site/s and any given period of time for which there are data, the OTST speed monitoring reports offer: the median (50^{th} percentile) speed, the 85^{th} percentile speeds, and a weighted average speed. These reports can be disaggregated at multiple levels for optimal understanding of speed trends on Minnesota roadways.

Methodology

Raw data format

Raw roadway speed data are provided by Traffic Data & Analysis (TDA) via multiple Excel files. Data are collected 24 hours a day seven days a week, and exported monthly from each operational site. Each file lists the total number of vehicles traveling within specified speed ranges or bins (see sample below for speed ranges). These data are then processed using the customized syntax to calculate 50^{th} percentile speed, 85^{th} percentile speed, and weighted average speed.

Table 1: Sample Raw Data Export

		Total Number of Vehicles Within Speed Range												
Date	**Hour**	**<=40 MPH**	**41-45 MPH**	**46-50 MPH**	**51-55 MPH**	**56-60 MPH**	**61-65 MPH**	**66-70 MPH**	**71-75 MPH**	**76-80 MPH**	**81-85 MPH**	**86-100 MPH**	**101-110 MPH**	**> 110 MPH**
1/1/2010	**00:00**	6	2	14	36	118	112	47	20	4	3	3	0	0
1/1/2010	**01:00**	9	6	13	36	94	69	32	12	3	3	1	0	0
1/1/2010	**02:00**	9	5	9	33	84	47	20	8	2	2	0	0	1
1/1/2010	**03:00**	9	5	8	28	57	65	23	5	1	1	0	1	0
…	**…**	…	…	…	…	…	…	…	…	…	…	…	…	…

There are 13 possible speed bins; however, vehicles count within the “>110 mph” category are assumed errors and removed from the data file. Additionally, outliers are filtered from aggregate reported trend data in order to afford an accurate picture of free-flow traffic. Please refer to *Outliers and Filters* section for further details.

50th Percentile or Median and 85th Percentile

The 50th percentile and 85th percentile are calculated using multiple steps. It is important to note that each hour's worth of data is processed independently for more robust reporting. This means that for any given hour the reported statistics, 50th percentile, 85th percentile, and weighted average, are available. For aggregate reporting, i.e. speed trends over a period of time, the hourly 50th percentile and 85th percentiles are averaged.

1. The total number of vehicles is calculated by adding the number of vehicles recorded in each bin. Represented by $\Sigma(X_{a..i})$ where X=total number of vehicles within a given bin at a given hour
2. The percent of distribution for each bin is then calculated: $X_{a,..i}/(\Sigma(X_{a..i}))*100$
3. The cumulative percent is calculated by sequentially summing the percent of distribution:
    a. For cumulative percent for $X_a=(X_a/(\Sigma(X_{a..i}))*100)$
    b. For cumulative percent $X_b=(X_a/(\Sigma(X_{a..i}))*100)+(X_b/(\Sigma(X_{a..i}))*100)$
    c. For cumulative percent $X_c=(X_a/(\Sigma(X_{a..i}))*100)+(X_b/(\Sigma(X_{a..i}))*100))+(X_c/(\Sigma(X_{a..i}))*100)$…
4. The bin at which the cumulative percent equals 50 is reported as the 50th percentile for the given hour
5. The bin at which the cumulative percent equals 85 is reported as the 85th percentile for the given hour

Weighted Average

Due to the volume of data available in these datasets and data storage limitations at the collection sites, individual vehicle records are not available. Without individual vehicle records, it is not possible to calculate a true statistical mean speed. As a proxy a weighted average is calculated in order to determine the weighted average vehicle speed.

1. Weighted average speed is calculated by multiplying the total count in each bin at a given hour by the midpoint of each bin, represented as *M*: $M*X$
    a. The midpoint for 1-40 mph is 20. Using the data sample in Table 1, at midnight there were six vehicles traveling within the 1-40 mph bin; therefore 20x6
2. The product is calculated for each speed bin, summed, then divided by the total number of vehicles within that given hour: $\Sigma(M*X)_{a...i}/\Sigma X$

Note. This method assumes that half of the vehicles are traveling within a given speed bin are evenly distributed within that speed bin.

Outliers and Filters

Outliers are extreme values either at the high or low end. While some analyses focus specifically on outlier data; however, these analyses focus on typical roadway speed and typical behaviors. Because these data are arranged as counts within speed bins, it is not possible to determine the true statistical mean; furthermore, it is not possible to identify true statistical outliers.

A proxy method of identifying outliers was previously used and then updated in 2009. Cases (rows of data) where the highest two bins, 85-100 mph and 101-110 mph *Or* lowest bin of <=40 mph, account for 10% or more than the general population for a given hour. This means if there were 100 vehicles recorded at a given hour: 15 were recorded as traveling 85-100 mph and 20 vehicles were recorded as traveling 101-110 mph, 35 of 100, 35% of vehicles exceed our cut-point of 10%. This row of data is flagged as having excessive outliers, does not represent normal conditions, and may have resulted from data collection failure/error. Rows of data that are flagged as having outliers are filtered so that they are not included in the analyses.