

Factors affecting the validity of the traffic stop and race data:

Number of people stopped vs number of tickets: The traffic stop counts in the data as submitted by departments are for tickets not people. A police officer may give out more than one ticket per stop which can be miscalculated as number of people stopped. State law requires that data associated with officer's interactions with a vehicle operator be collected at all roadside stops.

Non-discretionary stops: Tickets and warnings issued as the result of crashes, marijuana tickets, under-age drinking tickets, and externally generated stops may be included in the data sets as if they were discretionary (e.g. a crash may have been coded as a moving violation).

Numbers too small for analysis: In many of the spreadsheets, the numbers of stops for operators of color are too small for valid analysis of the data.

Non-standardization of data entry and coding: Data entry and coding in the law enforcement data collection systems have not been standardized, and officer training on how to enter data has not been available therefore some percentage of the data may be inaccurate.

Timeframe: Not all data was extracted for the same timeframe.

Missing data: Data was frequently missing from tickets (e.g. race was not coded), or more than one box was checked in a category. Tickets or warnings issued to a company (i.e. commercial motor vehicle stops) were missing data, including age and gender of the driver.

Non-audited data: Data include improperly coded tickets by police officers and have not been checked for accuracy.

Data missing from some police departments: Some police departments were unable to extract the data from their records management systems, so that data is not available.

Analysts should be cautious when comparing results from different jurisdictions because of lack of data, missing data, differing timeframes for data extractions, differences in coding, etc.