6 Description of Data Plot Pages and Reported Statistics

Data Plot Organization

For every land use, statistics and data plots are presented for at least one independent variable and for at least one time period. For each land use, the data plots are organized in the following manner:

- First by setting
- Then by land use subcategory, if applicable
- Then by trip type
- Then by independent variable
- Then by time period

Data Plot Content

Data plots provide a fundamental display of the variance within the database. It is important to note that the data points on the plots represent **the observed number of trips**, not trip generation rates, plotted against the value of the independent variable.

Each data plot corresponds to a specific trip travel mode for a single combination of land use, land use subcategory, independent variable, time period, and setting. The standard data plot layout clearly identifies these defining factors, including the trip type. A lightly shaded watermark illustrating the travel mode is included in the background of the data plot.

For many land uses, the analyst can have strong confidence in the trip generation relationships presented in the 12th Edition. For others, where the data set is small, plots serve only as an initial indication of trip generation. Each plot with five or fewer data points includes the statement “Caution—Small Sample Size” above the plot area.

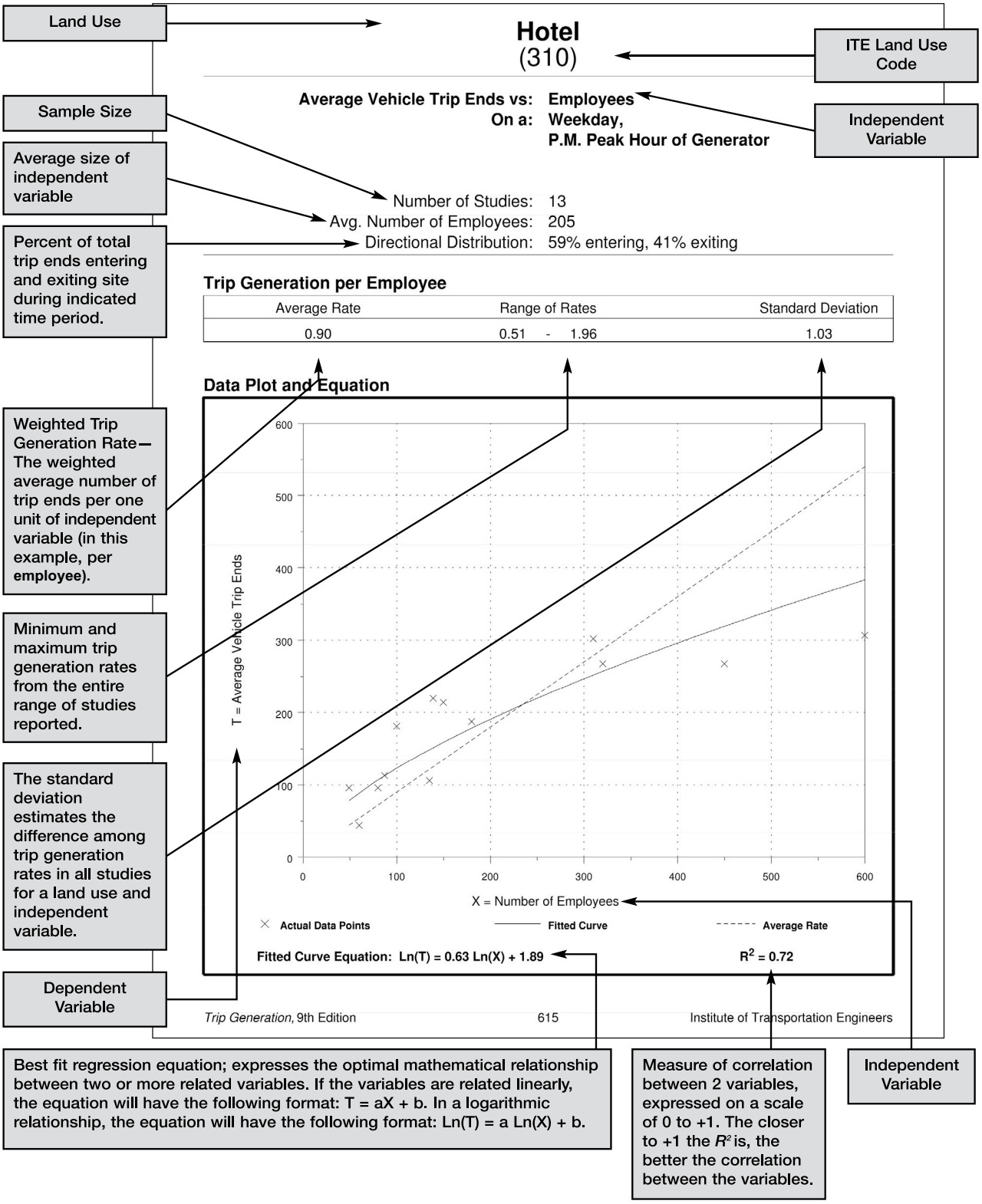
For some plots generated through the ITETripGen web app, the statement “Caution—Incomplete Data Set” may also appear above the plot area, indicating that the plot is based on a filtered subset rather than the entire database.

Data Plot Format

Figure 1 is a sample data plot illustrating the format for presenting statistical and descriptive information for a single land use, a specified independent variable, and a specified time period. The sample data page explains each element of the figure. Figure 1 is published in the *Trip Generation Handbook, 3rd Edition*; plots contained in TGM 12th Edition include some modifications to this figure such as inclusion of the setting/location and other minor formatting differences.

Each data point within a data plot represents the observed number of trips entering or exiting a single site, plotted against the value of the independent variable for the site.

Figure 1. Sample Data Page in *Trip Generation Manual*



Source: *Trip Generation Handbook, 3rd Edition*, Institute of Transportation Engineers, Washington, DC, 2017.

Reported Statistics

Average Trip Rate

The average trip generation rates displayed in the data plots are calculated using a **weighted average trip rate**, rather than the average of the individual rates. This approach minimizes the influence of outliers, as individual sites with large variances would otherwise skew the average.

Standard Deviation for the Weighted Average Trip Rate

The **standard deviation** measures how widely dispersed data points are around the calculated average. A lower standard deviation indicates a better fit to the average. Because the statistics are based on weighted averages—not simple arithmetic average—the reported standard deviation is an approximation and not statistically precise. The approximated standard deviations are provided for plots with three or more data points.

Regression Analysis

Each data plot (for every combination of land use, land use subcategory, independent variable, time period, setting, and trip type) shows a dashed line representing the average trip rate.

If there are at least four data points, the ITE TripGen web app evaluates whether a regression curve can be developed between the number of trips and the independent variable. The curve with the best (i.e., highest) coefficient of determination (R^2) is determined for each data plot. If the R^2 value is at least 0.50, the fitted curve equation and its coefficient of determination (R^2) are displayed on the plot.

The coefficient of determination is defined as the percent of the variance in the number of trips associated with the variance in the size of the independent variable. If the R^2 value is 0.75, then 75 percent of the variance in the number of trips is accounted for by the variance in the size of the independent variable. An R^2 value closer to 1.0 indicates a better fit; an R^2 value closer to 0 indicates a weaker fit. The fitted curve equation with the highest R^2 value is presented. Unlike the weighted average rate, the plotted fitted curve equation does not necessarily pass through the origin and may not be linear.

The following two general forms of fitted curve equations are considered:

Linear: $T = aX + b$

Logarithmic: $Ln(T) = aLn(X) + b$

Where X is the independent variable and T (the number of trips) is the dependent variable. The regression analysis determines the values of *a* and *b* that minimize the expected error in estimating the dependent variable.

In some cases, fitted equations yield a large or negative y-intercept, producing illogical trip-end estimates for small values of the independent variable. In such instances, caution should be used, and users are referred to Chapter 3, “Process for Estimating Trips Generated by a Study Site,” in the *Trip Generation Handbook*, 3rd Edition, for additional guidance.

Variations in the Statistics

Variation in trip generation characteristics for specific land uses is reflected in the range of rates, standard deviation, and coefficient of determination (R^2) value. These variations may be due to small sample sizes, site-specific marketing or economic factors, geographic differences, or unique site characteristics. Accordingly, judgment must be exercised in the use of the statistics in this manual.

Additional variation may result from differences in the duration of traffic counts; the season during which the traffic volumes were counted; and geographic location of the study site. The ITETripGen web app allows users to filter data by geographic region within the U.S. and Canada.

Cautions

The plots presented in the *Trip Generation Manual* cover only the ranges of independent variables for which data are available. Extrapolation beyond these ranges should be approached with caution, as no data exist to validate trip generation characteristics outside the documented bounds.

Even plots based on a single data point are included as reference. Extreme caution should be used in applying trip relationships derived from a single point. ITE hopes that inclusion of even the smallest dataset may stimulate new data collection and submission, further filling the gaps in the database.