

Researcher Dictionary
for
Time Stamp Data

Version 1.0

August 17, 2010

REVISION HISTORY

[illegible]

INTRODUCTION

The following data dictionary describes time stamp variables available in the naturalistic driving data for use by the research community. In addition to this introduction, the data dictionary includes four parts:

Revision History – This data dictionary should be considered a working document that will evolve over time. The revision history shown on the previous page provides a table which describes updates to the document.

Related Reading – A list of related subject areas and specific documents of value to users of the data set described in this data dictionary.

Data Description – This section describes what data is available and how the data are stored.

List of Variables – A list of the entries (variables) in the dictionary which can be used as a table of contents to locate specific variables in the document.

RELATED READING

Individuals working with these data are encouraged to become familiar with them, the method in which they were collected, and literature in the area of secondary data analyses. The following references are provided as starting points to assist the researcher in his or her efforts.

100-Car Study Overview

The 100-Car Naturalistic Driving Study was an instrumented vehicle study conducted in the Northern Virginia / Washington, D.C. area over a two-year period. The primary purpose of the study was to collect large-scale naturalistic driving data. To this end the instrumentation was designed to be unobtrusive, study participants were given no special instructions, and experimenters were not present. Approximately 100 vehicles were instrumented with a suite of sensors including forward and rearward radar, lateral and longitudinal accelerometers, gyro, GPS, access to the vehicle CAN, and five channels of compressed digital video. Collection rates for the various sensors ranged from 1Hz to 10Hz. This collection effort resulted in approximately 2,000,000 vehicles miles and 43,000 hours of driving data.

Methods

100-Car Methods

The methods used for collecting the data are described in:

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R. R. (2006). The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car Field Experiment DOT HS 810 593.

Secondary Data Analysis

Use of data collected by other organizations is becoming increasingly common in this digital age. In some fields, such as the social sciences or business, the use of previously collected data is more common than, for example, in psychology or product development. The primary benefit of this approach is cost savings. There are also risks that can threaten the validity of analyses conducted in this manner. The following references include discussion and recommendations for secondary analysts.

Akerstrom, M., Jacobsson, K., Wasterfors, D. (2004). "Reanalysis of previously collected material" in Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman (eds), *Qualitative Research Practice*, Thousand Oaks, CA. Sage Publications Ltd.

Corti, L. Thompson, P. (2004). "Secondary analysis of archived data", in Clive Seale, Giampietro Gobo, Jaber Gubrium, and David Silverman (eds), *Qualitative Research Practice*, Thousand Oaks, CA. Sage Publications Ltd.

Dale, A. Arber, S., and Procter, M. (1988). *Doing Secondary Analysis*, Unwin Hyman Ltd., London.

Hyman, H. (1972). *Secondary Analysis of Sample Surveys*, Wesleyan University Press, Middletown, Connecticut.

Kiecolt, K. and Nathan, L. (1985). *Secondary Analysis of Survey Data – Sage University Paper Series on Quantitative Applications in the Social Sciences*, 53. Sage Publications, Beverly Hills, CA.

DATA DESCRIPTION

This dataset consists of two tab-delimited files, one for events and one for baseline epochs. Each row in a given file represents one event or baseline epoch.

LIST OF VARIABLES

The following variables are included in the text files.

| Variable Number | Variable Name | Variable Description | Example |
|-----------------|----------------|--|--|
| 1 | Webfileid | Unique identifying value for each event or epoch. | |
| 2 | Number of days | The number of days a vehicle has been in the study at the time of the event or epoch selection. This timestamp is reported in vehicle days as the proportional baseline sampling was completed on a per vehicle basis. | A value of 14 in this field would indicate that an event was observed or a baseline epoch was sample from a file collected on the 14 th day the vehicle was in the study. |