

Behavior-based Predictive Safety Analytics - Pilot Study (02-020) Dataset

Dataset available at: <https://doi.org/10.15787/VTTI/464GB9>

(This dataset supports report **Behavior-based Predictive Safety Analytics – Pilot Study**)

This U.S. Department of Transportation-funded dataset is preserved by the Virginia Tech Transportation Institute (VTTI) in their data repository (<https://dataverse.vtti.vt.edu/>), and is available at <https://doi.org/10.15787/VTTI/464GB9>

The related final report **Behavior-based Predictive Safety Analytics – Pilot Study**, is available from the National Transportation Library’s Digital Repository at <https://rosap.ntl.bts.gov/view/dot/49167>.

Metadata from the VTTI Repository record:

Description:

- **Project Description:** A subset of the SHRP2 data was used to construct the datasets used for the analyses in the project. The general aim of this project was to investigate and develop statistical models predicting individual driver crash involvement based on driving style, demographic and behavioral history data. For each individual driver, six consecutive calendar months were extracted beginning from the second month of data collection (study period, months 2-7). This six-month data interval was used to calculate driving style measures and crash/near crash involvement. In addition, questionnaire data for each participant, collected prior to the start of the SHRP2 data collection was retrieved. In addition, a range of further inclusion criteria were applied. In particular, drivers selected for the present analysis were required to have participated in SHRP2 data collection for at least seven months, and to have driven more than 1,000 miles in the six-month study period. This resulted in a dataset of 2,800 and 3.91 million trips, amounting to a total of 27.16 million miles driving distance and 0.69 million driving hours.
- **Data Scope:** This dataset contains 2800 drivers from the SHRP2 data collection. Data include questionnaire factors on driver behaviors and risk perception, exposure metrics based on time, hours, and trips, crash-related data, and driver behavior variable mined from the six-month study period. All data is at the driver-level and is continuous.
- **Data Specification:** Accompanying PDF titled “BPSA Data Dictionary” includes data specification. “NA” is used to denote missing values.

Subject:

Engineering; Other

Keyword:

Crash, Near Crash, Driver Behavior Questionnaire, Crash Rate, Driver Behaviors

Recommended citation:

Andrew Miller; Wenyan Huang; Susan Soccolich; Felix Dreger, 2018, "Behavior-based Predictive Safety Analytics - Pilot Study (02-020)", <https://doi.org/10.15787/VTTI/464GB9>, VTTI, V1

Dataset description:

This dataset contains 1 .zip file collection described below.

Behavior-based Predictive Safety Analytics – Pilot Study (02-020) Data.zip:

This collection contains 1 .xlsx file and 1 .pdf file listed below.

- BPSA Data Dictionary.pdf
- BPSA – SAFE-D Data Upload_DUL_excluded.xlsx

The .xlsx file is a Microsoft Excel file, which can be opened with Excel, and other free available software, such as OpenRefine, and the .pdf file format is an Adobe Acrobat Portable Document Format (PDF) file and can be opened with the Adobe Acrobat software.

National Transportation Library (NTL) Curation Note:

As this dataset is preserved in a repository outside U.S. DOT control, as allowed by the U.S. DOT's Public Access Plan (<https://doi.org/10.21949/1503647>) Section 7.4.2 Data, the NTL staff has performed *NO* additional curation actions on this dataset.

NTL staff last accessed this dataset at <https://doi.org/10.15787/VTTI/464GB9> on 2021-02-01.

If, in the future, you have trouble accessing this dataset at the host repository, please email NTLDataCurator@dot.gov describing your problem. NTL staff will do its best to assist you at that time.