

Investigating Problem of Distracted Drivers on Louisiana Roadways Dataset

Dataset available at: https://digitalcommons.lsu.edu/transet_data/29/

(This dataset supports report **Investigating Problem of Distracted Drivers on Louisiana Roadways**)

This U.S. Department of Transportation-funded dataset is preserved by the Transportation Consortium of South-Central States (TRAN-SET) in the LSU Digital Commons Repository (<https://digitalcommons.lsu.edu>), and is available at https://digitalcommons.lsu.edu/transet_data/29/

The related final report **Investigating Problem of Distracted Drivers on Louisiana Roadways**, is available from the National Transportation Library's Digital Repository at <https://rosap.ntl.bts.gov/view/dot/61460>

Metadata from the LSU Digital Commons Repository record:

Document Type: Data Set

Publication Date: 12-2018

Abstract: Corresponding data set for Tran-SET Project No. 17SALSU10. Abstract of the final report is stated below for reference: "Due to the escalating usage of cellphone and social networking, distracted driving is and will remain as one of the most serious problems faced by Departments of Transportation (DOTs) and law enforcement agencies. Under the aim of in-depth investigation of distracted driving crashes in Louisiana, the specific objectives of this study are: (1) reviewing the crash reports for the quality of distracted driving crash reporting, (2) analyzing distracted driving-related crashes through regression model and data mining algorithm to link the severity of distracted driving crashes with the contributing factors collected in crash data, (3) investigating the observable characteristics of distracted driving roadside and video survey, and (4) recommending the countermeasures utilizing the analysis results and reviews. About 60,000 crashes from ten-year crash data, three types of distracted driving related crashes are modeled: Fatal (K) and Severe (A) Injury; Moderate (B) and Complaint (C) Injury; and Property-damage only (PDO). One statistical method was used for prediction, multinomial logistic regression, and one data mining algorithms was used, random forest. Higher speed limit, curved road, head-on crashes were identified among the key factors. Data mining algorithms performed better in prediction compared to the multinomial logistic regression when sensitivity and specificity were used to compare the predicted results. Fisher's exact tests of roadside manual observation data shows that gender has no significant influence in cellphone distraction (regardless of distraction type), however age can be influential and associated with driver distraction. Association rule mining of observation data shows that the most predominant type of cellphone use is manipulating mainly occurs at intersections, whereas talking is more associated with segments. In-vehicle video data were coded by the software FaceReader, which captures facial expressions of drivers while driving. Initial results do suggest valence in emotion can be attributed to timing before, during, and after cellphone calls and texting. Physical countermeasure development towards reducing the distraction-related crash severity should be targeted at preventing lane departure crashes. Physical countermeasure development towards reducing the distraction-related crash severity should be targeted at preventing lane departure crashes. Strict enforcement of texting ban with awareness campaign are also expected to prevent distracted driving."

Comments: Tran-SET Project No. 17SALSU10

Recommended citation:

Sun, X. (2018). Investigating Problem of Distracted Drivers on Louisiana Roadways. Retrieved from https://digitalcommons.lsu.edu/transet_data/29

Dataset description:

This dataset contains 1 file described below.

TranSET_17SALSU10_Data.zip:

- TranSET_17SALSU10_Data Folder
 - Data_Notes.docx
 - Accessibility_Notes.docx
 - 4_R_Studio Files Folder
 - Rscript.txt
 - Observation_Data.dat
 - All_Data_Analysis_17SALSU10.R
 - 3_Data_Files Folder
 - 3_Crash_Data.xlsx
 - 2_Louisiana_Crash_Data.xlsx
 - 1_FaceReader_Participation_Info.xlsx
 - 2_Tables Folder
 - Table 8_F-test for emotions before, during, and after texting.xlsx
 - Table 7_Results of Logistic Regression.xlsx
 - Table 6_Apriori generated 4-itemset rules for_cellphone_use = talking.xlsx
 - Table 5_Apriori generated 4-itemset rules for_cellphone_use = manipulating.xlsx
 - Table 4_Multinomial logistic regression results.xlsx
 - Table 3_Distraction-related attributes in various crash databases and reports.xlsx
 - Table 2_Percentages of each items of observed variables by cellphone use type.xlsx
 - Table 1_Frequency and percentage.xlsx
 - 1_Figures Folder
 - Image_Descriptions.docx
 - Figure 9.jpg
 - Figure 8.jpg
 - Figure 7.jpg
 - Figure 6_Relative frequency of the items observed.xlsx
 - Figure 6_Relative frequency of the items observed.csv
 - Figure 6.jpg
 - Figure 5.jpg
 - Figure 4.jpg
 - Figure 3.jpg
 - Figure 21c.jpg

- Figure 21b.jpg
- Figure 21a.jpg
- Figure 20.jpg
- Figure 2.jpg
- Figure 19.jpg
- Figure 18.jpg
- Figure 17b.jpg
- Figure 17a.jpg
- Figure 16b.jpg
- Figure 16a.jpg
- Figure 15.jpg
- Figure 14.jpg
- Figure 13.jpg
- Figure 12.jpg
- Figure 11.jpg
- Figure 10.jpg
- Figure 1.jpg

File Type Descriptions:

- The .docx file is a Microsoft Word file, which can be opened with Word and other free word processor programs, such as Kingsoft Writer, OpenOffice Writer, and ONLYOFFICE.
- The .txt file type is a common text file, which can be opened with a basic text editor. The most common software used to open .txt files are Microsoft Windows Notepad, Sublime Text, Atom, and TextEdit (for more information on .txt files and software, please visit <https://www.file-extensions.org/txt-file-extension>).
- The file extension .dat is traditionally used by many various applications or programs for their data or resource files (for more information on .dat files and software, please visit <https://www.file-extensions.org/dat-file-extension>).
- The .r file type is related to R programming language. R is a language and environment for statistical computing and graphics (for more information on .r files and software, please visit <https://www.file-extensions.org/r-file-extension>).
- The .xlsx and .xls file types are Microsoft Excel files, which can be opened with Excel, and other free available software, such as OpenRefine.
- The .csv, Comma Separated Value, file is a simple format that is designed for a database table and supported by many applications. The .csv file is often used for moving tabular data between two different computer programs, due to its open format. The most common software used to open .csv files are Microsoft Excel and RecordEditor, (for more information on .csv files and software, please visit <https://www.file-extensions.org/csv-file-extension>).
- The .jpg file extension is associated with JPEG (Joint Photographic Experts Group) file format. JPEG is a lossy image compression algorithm that significantly reduces the file size of the original image at the cost of quality. The higher the compression ratio the lower the quality of the .jpg file (for more information on .jpg files and software, please visit <https://www.file-extensions.org/jpg-file-extension>).

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://ntl.bts.gov/public-access>) 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://digitalcommons.lsu.edu/transet_data/29/ on 2022-04-19. 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.