7 REFERENCES

Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of Safety Research*, *34*(5), 597–603. https://doi.org/10.1016/j.jsr.2003.05.009

Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, *32*(5), 633–642. https://doi.org/10.1016/S0001-4575(99)00094-9

Abdulhafedh, A. (2017). Road Crash Prediction Models: Different Statistical Modeling Approaches. *Journal of Transportation Technologies*, *7*(2), 190–205. https://doi.org/10.4236/JTTS.2017.72014

Agafonkin, V. (2016). *Clustering millions of points on a map with Supercluster | by Mapbox | maps for developers*. Mapbox Blog. https://blog.mapbox.com/clustering-millions-of-points-on-a-map-with-supercluster-272046ec5c97

Agafonkin, V. (2020). *A Web Map from Scratch / Vladimir Agafonkin / Observable*. https://observablehq.com/@mourner/simple-web-map

Asalor, J. O. (1984). A general model of road traffic accidents. *Applied Mathematical Modelling*, *8*(2), 133–138. https://doi.org/10.1016/0307-904X(84)90066-0

Assi, K., Rahman, S. M., Mansoor, U., & Ratrout, N. (2020). Predicting Crash Injury Severity with Machine Learning Algorithm Synergized with Clustering Technique: A Promising Protocol. *International Journal of Environmental Research and Public Health 2020, Vol. 17, Page 5497*, *17*(15), 5497. https://doi.org/10.3390/IJERPH17155497

Bertrand, F. (2021). *SweetViz* (2.1.2). https://github.com/fbdesignpro/sweetviz

Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S., & Schaub, T. (2016). *The GeoJSON Format, RFC 7946*. https://datatracker.ietf.org/doc/html/rfc7946

Chang, L. Y., & Wang, H. W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis & Prevention*, *38*(5), 1019–1027. https://doi.org/10.1016/J.AAP.2006.04.009

Devasurendra, K., Perera, L., & Bandara, S. (2016). An insight to motorized two and three wheel crashes in developing countries: A case study in Sri Lanka. *Journal of Transportation Safety & Security*, *Vol 9*, 204–215. https://doi.org/10.1080/19439962.2016.1236052

Dissanayake, S., & Amarasingha, N. (2012). Effects of Geometric Design Features on Truck Crashes on Limited-Access Highways. *Final Reports & Technical Briefs from Mid-America Transportation Center*.

Dissanayake, S., & Perera, L. (2011). A Survey Based Study of Factors Related to Older driver Highway Safety. *Journal of Transportation Safety & Security*, *3*(2), 77–94. https://doi.org/10.1080/19439962.2010.537437

Esri. (2020). *ArcGIS Pro* (2.6). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Feurer, M., & Hutter, F. (2019). Hyperparameter Optimization. In F. Hutter, L. Kotthoff, & J. Vanschoren (Eds.), *Automated Machine Learning: Methods, Systems, Challenges* (pp. 3–33). Springer International Publishing. https://doi.org/10.1007/978-3-030-05318-5_1

Flynn, D., Gilmore, M., & Sudderth, E. (2018). *Estimating Traffic Crash Counts Using Crowdsourced Data: Pilot analysis of 2017 Waze data and Police Accident Reports in Maryland*. https://rosap.ntl.bts.gov/view/dot/37256

Global Status Report on Road Safety 2018. (2018). In *World Health Organization* (Issue 1). https://www.who.int/publications/i/item/9789241565684

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jiménez-Mejías, E., Martínez-Ruiz, V., Amezcua-Prieto, C., Olmedo-Requena, R., Luna-Del-Castillo, J. D. D., & Lardelli-Claret, P. (2016). Pedestrian- and driver-related factors associated with the risk of causing collisions involving pedestrians in Spain. *Accident Analysis and Prevention*, *92*. https://doi.org/10.1016/j.aap.2016.03.021

Jordahl, K., Bossche, J. van den, Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., … Leblanc, F. (2020). *geopandas/geopandas: v0.8.1*. https://doi.org/10.5281/ZENODO.3946761

Joshua, S. C., & Garber, N. J. (1990). Estimating truck accident rate and involvements using linear and poisson regression models. *Transportation Planning and Technology*, *15*(1), 41–58. https://doi.org/10.1080/03081069008717439

Katchova, A. (2020). *Econometrics Academy - Probit and Logit Models*. Econometrics Academy. https://sites.google.com/site/econometricsacademy/masters-econometrics/probit-and-logit-models

Ke, G., Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017, December). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Advances in Neural Information Processing Systems 30 (NIP 2017)*. https://www.microsoft.com/en-us/research/publication/lightgbm-a-highly-efficient-gradient-boosting-decision-tree/

Kearney, M. (2017). *Cramér's V*. https://doi.org/10.4135/9781483381411.n107

Kidando, E., Kitali, A. E., Kutela, B., Ghorbanzadeh, M., Karaer, A., Koloushani, M., Moses, R., Ozguven, E. E., & Sando, T. (2021). Prediction of vehicle occupants injury at signalized intersections using real-time traffic and signal data. *Accident Analysis & Prevention*, *149*, 105869. https://doi.org/10.1016/J.AAP.2020.105869

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction Policy Problems. *American Economic Review*, *105*(5), 491–495. https://doi.org/10.1257/AER.P20151023

Komorowski, M., Marshall, D. C., Salciccioli, J. D., & Crutain, Y. (2016). Exploratory data analysis. In *Secondary Analysis of Electronic Health Records* (pp. 185–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-43742-2_15/FIGURES/18

Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, *2017-December*, 4766–4775. https://doi.org/10.48550/arxiv.1705.07874

Malmasi, S., & Zampieri, M. (2017). Detecting hate speech in social media. *International Conference Recent Advances in Natural Language Processing, RANLP*, *2017-September*, 467–472. https://doi.org/10.26615/978-954-452-049-6-062

Mapbox. (2021). *Mapbox*. Mapbox Inc. https://www.mapbox.com/

Mehdizadeh, A., Cai, M., Hu, Q., Yazdi, M. A. A., Mohabbati-Kalejahi, N., Vinel, A., Rigdon, S. E., Davis, K. C., & Megahed, F. M. (2020). A review of data analytic applications in road traffic safety. Part 1: Descriptive and predictive modeling. *Sensors (Switzerland)*, *20*(4). https://doi.org/10.3390/S20041107

Morgenthaler, S. (2009). Exploratory data analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 33–44. https://doi.org/10.1002/WICS.2

National Council for Road Safety, Sri Lanka. (2020). https://www.transport.gov.lk/web/index.php?option=com_content&view=article&id=29&Itemid=149&lang=en

Paredes, M., Hemberg, E., O'Reilly, U. M., & Zegras, C. (2017). Machine learning or discrete choice models for car ownership demand estimation and prediction? *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings*, 780–785. https://doi.org/10.1109/MTITS.2017.8005618

Perera, L., & Dissanayake, S. (2012). Contributing Factors to Older-Driver Injury Severity in Rural and Urban Areas. *Journal of the Transportation Research Forum*, *49*(1), 5–22. https://doi.org/10.5399/OSU/JTRF.49.1.2506

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, *2018-December*, 6638–6648. https://github.com/catboost/catboost

QGIS. (2021). *QGIS* (3.22.0). Open Source Geospatial Foundation Project. https://qgis.org/en/site/index.html

Reback, J., jbrockmendel, McKinney, W., den Bossche, J. van, Augspurger, T., Cloud, P., Hawkins, S., Roeschke, M., gfyoung, Sinhrks, Klein, A., Hoefler, P., Petersen, T., Tratner, J., She, C., Ayd, W., Naveh, S., Darbyshire, J., Garcia, M., … Seabold, S.

(2022). *pandas-dev/pandas: Pandas 1.4.1*. Zenodo. https://doi.org/10.5281/zenodo.6053272

Rokach, L. (2019). *Ensemble Learning : Pattern Classification using Ensemble Methods* (Second). World Scientific.

Roland, J., Way, P. D., Firat, C., Doan, T. N., & Sartipi, M. (2021). Modeling and predicting vehicle accident occurrence in Chattanooga, Tennessee. *Accident Analysis & Prevention*, *149*, 105860. https://doi.org/10.1016/J.AAP.2020.105860

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention*, *43*(5), 1666–1676. https://doi.org/10.1016/j.aap.2011.03.025

Senaratna, N. (2021). *LK Geo Data*. https://github.com/nuuuwan/geo-data

Shajith, S. L. A., Pasindu, H. R., & Ranawaka, R. K. T. K. (2019). Evaluating the Risk Factors in Fatal Accidents involving Motorcycle – Case Study on Motorcycle Accidents in Sri Lanka. *Engineer: Journal of the Institution of Engineers, Sri Lanka*, *52*(3), 33–42. https://doi.org/10.4038/engineer.v52i3.7363

Shneiderman, B. (2014). The big picture for big data: Visualization. *Science*, *343*(6172), 730. https://doi.org/10.1126/SCIENCE.343.6172.730-A/ASSET/B23773B5-0932-4174-8E8C-071CCF844B85/ASSETS/SCIENCE.343.6172.730-A.FP.PNG

Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*, 484–489. https://doi.org/10.1038/nature16961

Singh, A. (2017). Deep Learning Will Radically Change the Ways We Interact with Technology. *Harvard Business Review*. https://hbr.org/2017/01/deep-learning-will-radically-change-the-ways-we-interact-with-technology

Sri Lanka's Road Safety Country Profile. (2021). https://www.roadsafetyfacility.org/country/sri-lanka

Tay, R., & Rifaat, S. M. (2007). Factors contributing to the severity of intersection crashes. *Journal of Advanced Transportation*, *41*(3), 245–265. https://doi.org/10.1002/atr.5670410303

Thangamani, B. (2019). The Economic Impact of Road Accidents: The Case of Sri Lanka. *South Asia Economic Journal*, *20*, 124–137. https://journals.sagepub.com/doi/abs/10.1177/1391561418822210

Tjoa, E., & Guan, C. (2019). A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, *14*(8), 1–21. https://doi.org/10.1109/tnnls.2020.3027314

Uber. (2021). *Kepler.gl* (2.5.5). Uber Inc. https://kepler.gl/

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Williams, A. F. (1999). *The Haddon matrix: its contribution to injury prevention and control*.

Yee, A., & Alvarado, M. (2012). Pattern recognition and monte-carlo tree search for go gaming better automation. *Lecture Notes in Computer Science*, *7637 LNAI*, 11–20. https://doi.org/10.1007/978-3-642-34654-5_2

Yuan, Z., Zhou, X., & Yang, T. (2018). Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, *18*. https://doi.org/10.1145/3219819

Zychlinski, S. (2018, February 24). *The Search for Categorical Correlation*. Towards Data Science. https://towardsdatascience.com/the-search-for-categorical-correlation-a1cf7f1888c9