Investigation of Temperature and Traffic Volume Relationship in Chicago

Olivia Rueschhoff and Dr. Feridun Tasdan
Department of Mathematics and Philosophy, Western Illinois University

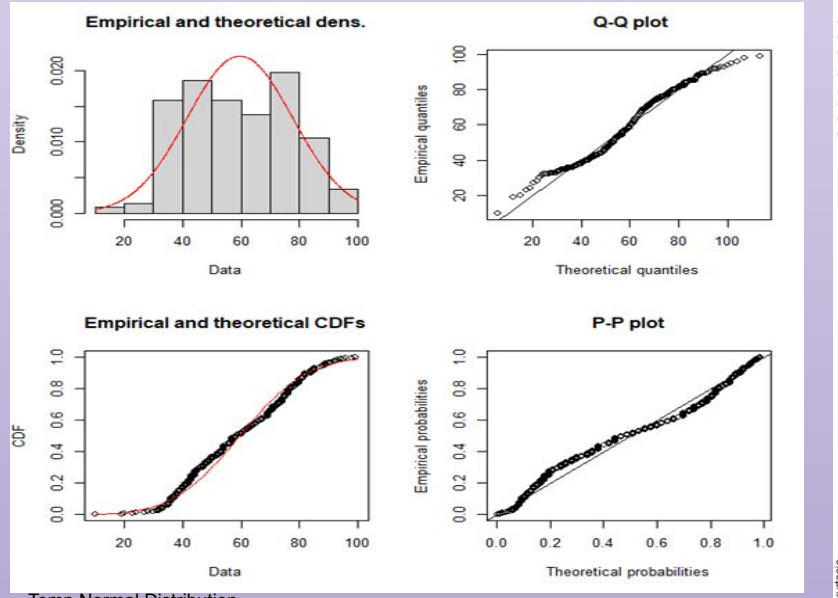
Abstract

A bivariate distribution represents the relationship between two data sets, aiding in predicting new data points. This study examines the correlation and underlying distributions of daily maximum temperatures and traffic volume in Chicago using 2006 data from the National Centers for Environmental Information and Chicago.gov. Although temperature and traffic are not typically linked, extreme conditions can influence traffic flow.

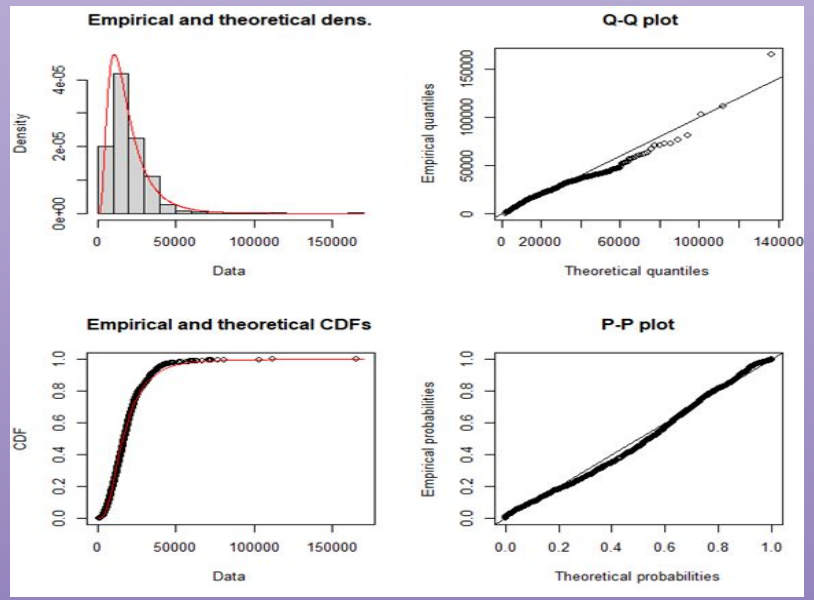
Analysis was conducted using Excel and R, employing graphical methods, correlation analysis, and goodness-of-fit tests. Scatterplots, histograms, and boxplots were used for visualization, while advanced methods like Cullen and Frey graphs, Kolmogorov-Smirnov tests, and convex hull trimming refined distribution identification.

Results revealed a minimal correlation ($r = 0.12$, $R^2 = 0.0144$), improving slightly after outlier removal ($r = 0.1412$). Maximum temperature data followed a uniform or beta distribution, while vehicle volume aligned with gamma or log-normal distributions. Spearman's Rank Correlation Test confirmed a weak but statistically significant relationship ($\rho = 0.165$, $p < 0.05$).

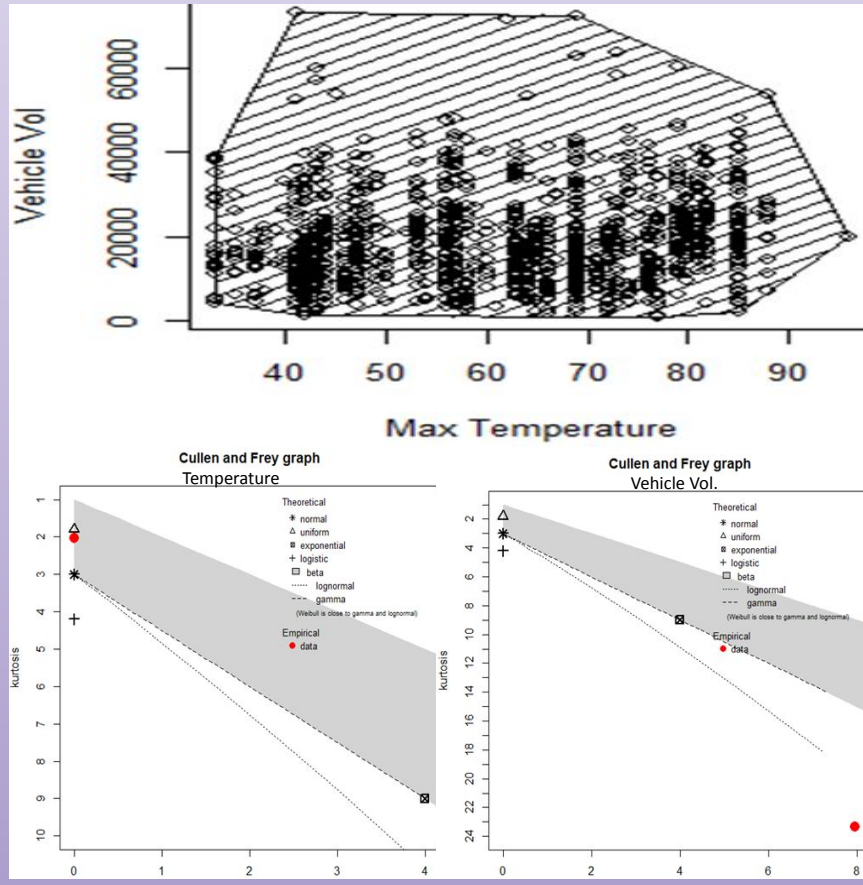
Despite identifying possible bivariate distributions, the weak correlation suggests temperature and traffic volume are not strongly interrelated, limiting predictive reliability. Limitations include potential data inaccuracies and missing summer entries. Expanding the dataset beyond 2006 could provide further insights. While not useful for precise predictions, this study enhances understanding of bivariate distributions and the effects of outlier treatment on correlation analysis as well as statistical modeling and data-driven decision-making.



Temp Normal Distribution



Vehicle Volume Log-normal Distribution



Why this Matters

This research is important because it contributes to a deeper understanding of how environmental factors, such as temperature, may influence urban mobility and transportation patterns. While temperature alone may not be a strong predictor of traffic volume, identifying underlying distributions and assessing correlations help refine statistical modeling techniques used in traffic forecasting and urban planning. By applying advanced data analysis methods, this study highlights the importance of proper data treatment, and drawing meaningful conclusions from real-world datasets. These findings can inform policymakers, city planners, and transportation agencies in optimizing infrastructure, managing congestion, and improving public safety during extreme weather conditions. Ultimately, this research enhances knowledge of bivariate distributions and their role in data-driven decision-making across multiple fields.

References

Comparison of P-P plots and Q-Q plots. (1999, September). Retrieved from Simon Fraser University: https://www.sfu.ca/sasdoc/sashtml/qc/chap8/sect9.htm

Everitt, B., & Hothorn, T. (2011). *An Introduction to Applied.* (R. Gentleman, K. Hornick, & G. Parmigiani, Eds.) London, Munchen: Springer.

Ford, C. (2015, August 26). *Understanding QQ Plots.* Retrieved from University of Virginia Library: https://library.virginia.edu/data/articles/understanding-q-q-plots

Hogg, R. V., Tanis, E. A., & Zimmerman, D. L. (2021). In *Probability and Statistical Inference* (10th ed.). Pearson.

How to Perform Spearman Correlation in R. (2024). Retrieved from OnlineSPSS: https://www.onlinespss.com/spearmen-correlation-in-r/

Kolmogorov-Smirnov Test in R Programming. (2023, March 10). Retrieved from GeeksforGeeks: https://www.geeksforgeeks.org/kolmogorov-smirnov-test-in-r-programming/

Quaresma, D. F., Pereira, T. E., & Fireman, D. (2021). *Validation of a simulation model for FaaS performance benchmarking using predictive validation.* Campina Grande: Federal University of Campina Grande. Retrieved from https://www.researchgate.net/publication/353060353_Validation_of_a_simulation_model_for_FaaS_performance_benchmarking_using_predictive_validation

Acknowledgements
Special thanks to Dr. Feridun Tasdan, the Department of Mathematics, and the Centennial Honors College for their guidance and support.