

Case Study: Forecasting NYC Motor Vehicle Collisions for Vision Zero

Motor vehicle collisions are a leading cause of injury and death in major cities, making transportation safety a critical priority for public officials. In response, New York City launched Vision Zero in 2014, a city-wide initiative committed to eliminating traffic fatalities through targeted interventions, data-driven planning, and proactive resource allocation. A core component of Vision Zero is accurate forecasting: anticipating when and where collisions are most likely to occur so that agencies can strategically deploy law enforcement, adjust traffic patterns, and implement preventative safety measures. To support this mission, the NYC Open Data portal provides over a decade of detailed collision records under the Motor Vehicle Collisions dataset, documenting every NYPD-reported crash at both daily and hourly resolution.

It's your first day on the job. You've just joined NYC's Vision Zero analytics team as the new Data and Forecasting Lead, and city leadership has given you a high-stakes assignment. With the holiday season approaching, you must determine which NYC borough is expected to experience the highest number of motor vehicle collisions on January 1, 2026. New Year's Day is historically marked by heavy travel and unpredictable traffic patterns, and your forecast will guide how police officers, EMS teams, and roadway safety crews are deployed across the city. The accuracy of your predictions will support city preparedness, faster response times, and safer holiday travels.

With the tools, scripts, and data provided in the GitHub repository, you will run a daily time-series forecasting model to identify the borough at greatest risk on January 1, 2026. You will then use an hourly forecasting model to generate an hour-by-hour breakdown for that borough. Your job is to produce a clear summary of when and where NYC faces the greatest collision risk on January 1, 2026.

Details on formatting, expectations, and evaluation criteria are provided in the rubric. All code, data, and starter materials are available here: <https://github.com/evelynmaxwell/DS-4002-CS3>