

TUNING UP MUSIC HIGHWAY CHECKPOINT 2

In this project, we analyze serious crashes along the I-40 corridor in Madison and Henderson counties, Tennessee. Our goal is to explore how various roadway features contribute to crash severity, and to estimate the potential safety improvements that could result from targeted interventions.

We began by filtering crash data using precise GPS locations to isolate incidents that occurred directly on I-40. We then manually classified different segments of the highway according to a range of geospatial features — such as the presence of guardrails, cable barriers, rumble strips, pavement condition, road curvature, shoulder width, nearby ramps, and others. These features were assigned based on visual inspection of satellite imagery, engineering records, and field observations.

With this manually enriched dataset, we plan to train predictive models that estimate the likelihood of serious injury or fatality when a crash occurs, given the surrounding road conditions at that location. After validating model performance, we will simulate the impact of specific safety improvements by modifying the geospatial feature inputs (for example, adding cable barriers to sections where none exist). This allows us to estimate how much different interventions might reduce the number of severe injuries or fatalities, even without needing to predict crashes themselves.

Ultimately, this approach gives us a way to model not just where crashes happen, but how roadway design influences crash severity, providing a data-driven framework for prioritizing cost-effective safety improvements along this critical highway.

For the exploratory data analysis we have done so far, we have noticed the following trends:

- **Guardrails:** Crash severity decreases as guardrail protection increases; segments with no guardrails have the highest severe crash rates.
- **Cable barriers:** The presence of cable barriers, particularly on the shoulder, is associated with lower severity rates.
- **Pavement condition:** Poor and fair pavement conditions show higher severe crash rates compared to good or recently resurfaced pavement.
- **Proximity to entrances/exits:** Crashes near highway entrances or exits show slightly elevated severity risk.
- **Surrounding terrain:** Crashes occurring near forested or complex terrain have lower severe crash rates than those in open or plain areas.
- **Number of lanes:** Two-lane segments show slightly higher severity rates than wider segments.
- **Shoulder type and width:** Wider paved shoulders are associated with higher severity, possibly reflecting locations with higher baseline risk.
- **Speed limits:** Higher posted speed limits correspond to higher severe crash rates.
- **Medians and dividers:** Guardrail or cable dividers may be concentrated in higher-risk areas, showing higher severity rates where they are present.
- **Lane markings:** Poor or faded markings show a strong association with higher severity crashes.
- **Nighttime lighting:** Lighted sections show lower severity crash rates compared to unlit segments.