Crashspot – Week 8 Report

■ Focus

Model interpretation + spatial integration

■ Accomplishments

- Loaded the tuned Week 7 Random Forest model
- Computed Permutation Importance to identify key predictors of multi-vehicle crashes
- Generated and saved:
- docs/figures/week8_permutation_importance.png
- (Optional, if SHAP installed) docs/figures/week8_shap_beeswarm.png
- docs/maps/week8_predicted_hotspots.html (interactive Folium map of predicted crash risks)
- Confirmed end-to-end workflow with message: “Week 8 completed cells ready.”

■ Results

- Most important predictor: ve_total (number of vehicles)
- This strongly drives the classification of multi-vehicle crashes in the model.
- Other predictors (hour, month, weekday, is_weekend, is_night, persons, peds) showed minimal added predictive value in the current feature set.
- Interactive hotspot map highlights where higher predicted crash risks cluster spatially in Monroe, LA.

■ Outputs

- Figures:
- docs/figures/week8_permutation_importance.png
- (optional) docs/figures/week8_shap_beeswarm.png
- Maps:
- docs/maps/week8_predicted_hotspots.html

■ Significance

- Concludes the research phase (Weeks 1–8):
Raw crash data → cleaned features → predictive modeling → evaluation → interpretation → geospatial mapping.
- Demonstrates the integration of machine learning with GIS visualization.
- Provides actionable insight: vehicle count (ve_total) is the dominant risk factor for multi-vehicle crashes, with spatial clusters visible in Monroe.

■ Next Steps

- Transition from research → application:
- Build a web/app interface to host the hotspot map and allow dynamic crash risk exploration.
- Enrich features (road type, weather, lighting) for improved predictive accuracy.
- Explore advanced interpretability (SHAP, partial dependence plots).