Walk. Bike. Safe. Texas:
E-Bike Crash Analysis Report

Project #: 2026-TTI-G-1YG-0032

Authors:
Myunghoon Ko
Neal Johnson
Joan Hudson
Marcie Perez

Prepared for the Texas Department of Transportation

April 2026

Texas A&M Transportation Institute

E-Bike Crash Analysis

Electric Bicycle–Involved Crash Identification and Data Preparation

The research team extracted crash data from the Texas Department of Transportation's Crash Records Information System (TxDOT CRIS) on March 10, 2026. The dataset included all reportable crashes occurring between 2016 and 2025.

As an initial screening step, the team applied a set of filtering criteria to identify crashes potentially involving electric bicycle (e-bike) and similar micromobility devices. Due to current limitations in explicitly identifying e-bikes within the CRIS database, a proxy-based approach was used to capture relevant cases. Specifically, crashes were selected if they met at least one of the following conditions:

- Unit Description = 1 or 5 (Motor Vehicle or Other Roadway Unit) and Vehicle Body Style = 92 (coded as e-scooter/micromobility device), resulting in 33,766 crashes
- Unit Description = 3 (Pedalcyclist), resulting in 25,349 crashes
- E-scooter flag = 1 (Yes), resulting in 420 crashes (including two crashes also identified under the conditions above)

Based on these criteria, a total of 59,533 crashes were identified as potential e-bike–involved crashes for further review and analysis. The resulting dataset included:

- 124,281 units involved in crashes
- 113,366 primary persons associated with the crashes
- 149,636 contributing factors recorded

It is important to note that this screening process was designed to capture a broad set of candidate crashes involving e-bikes and related micromobility devices. Due to coding limitations and potential inconsistencies in reporting, the identified crashes may include non–e-bike cases (e.g., traditional bicycles or other devices). In particular, the code, *Vehicle Body Style = 92*, corresponds to an "other" category that requires additional description in the crash narrative field, which limits its specificity for identifying e-bikes.

To address these limitations, additional refinement and validation steps were conducted in subsequent phases of the analysis using crash narrative fields to improve the accuracy of e-bike crash identification.

Keyword Pattern Approach on Crash Narratives for Identifying E-Bike Crashes

To improve the identification of e-bike–involved crashes beyond structured crash data fields, the research team developed a keyword-based approach using crash narratives. This approach was designed to supplement the initial screening process by leveraging descriptive text fields, which often contain critical information not captured in coded variables.

Using crash narrative data extracted from the CRIS database, the research team implemented a systematic keyword pattern search to identify potential e-bike–related crashes. The analysis focused on the *investigator narrative* field, which provides a textual description of crash circumstances.

Keyword Pattern Development

A comprehensive set of keyword patterns was developed to capture variations in how e-bikes and related micromobility devices are described in crash narratives. The patterns were designed to ensure strict matching, minimizing false positives from generic terms such as "bike" or "bicycle." The keyword set included:

- Variations of e-bike terminology (e.g., *e-bike, e bike, ebike, electric bike, e bicycle, electric bicycle, motorized/motor bicycle, battery-powered bike, pedal assist, pedelec*)
- Variations of e-scooter terminology (e.g., *e-scooter, e scooter, escooter, electric scooter, battery-powered scooter)*
- Other types (e.g., *moped*)

Regular expression (regex) techniques were used to account for variations in spacing, hyphenation, and capitalization (e.g., "e-bike," "e bike," "ebike"). All keyword matching was conducted in a case-insensitive manner. Appendix A provides the full python code used for the narrative analysis including the keywords.

Keyword Matching and Flagging

Each crash narrative was evaluated against the defined keyword patterns. Two key variables were generated:

- **Keyword Match Flag**: A binary indicator identifying whether any e-bike/e-scooter/moped–related keyword was present in the narrative
- **Matched Keywords Field**: A list of specific keywords detected within each narrative (e.g., e-bike when matching e-bike terminology, e-scooter when matching e-scooter terminology, or moped)

This process enabled the identification of crashes where e-bike involvement was explicitly or implicitly described in narrative text, even when not captured in structured data fields.

Keyword Analysis Results

Table 1 summarizes the results of the keyword-based narrative screening applied to the 59,533 crashes identified from TxDOT CRIS (2016–2025). Using a strict set of regular-expression keyword patterns on the investigator narrative field, the process flagged 797 narratives as containing terms related to e-bikes, e-scooters, or mopeds. Among the flagged narratives, the team identified 337 e-bike crashes from keyword matches, along with 335 e-scooter crashes and 125 moped crashes.

Table 1. Results of Keyword Analysis using Crash Narratives

Total narratives reviewed (2016 to 2025)	59,533
Narratives flagged by keyword	797
E-Bike	337
E-Scooter	335
Moped	125

Figure 1 to Figure 3 highlight how unit coding and body-style coding are distributed among crashes classified as mopeds, e-scooters, and e-bikes.

- E-bikes: E-bike crashes are overwhelmingly coded as unit_desc_id = 3 (pedalcyclist) (303 crashes, 90%), while a smaller share appears under unit_desc_id = 1 (motor vehicle) with veh_body_styl_id = 92 (Explain in Narrative) (26 crashes, 8%) and unit_desc_id = 3 (pedalcyclist) with veh_body_styl_id = 92 (Explain in Narrative) (4 crashes, 1%). This indicates that e-bike involvement is typically embedded within the general pedalcyclist category, reinforcing why narrative keyword screening and supplemental rules for unit_desc_id = 3 (pedalcyclist) are necessary for identifying e-bike crashes.

- E-scooters: E-scooter crashes are most often identified by unit_desc_id = 5 (motorized conveyance) alone (248 crashes, 74%), with a notable portion coded as unit_desc_id = 3 (pedalcyclist) (47 crashes, 14%). Smaller shares are associated with other unit types (e.g., unit_desc_id = 1 or 8), and only a small fraction explicitly use the combined code unit_desc_id = 5 with veh_body_styl_id = 92. This distribution supports the need for using the e_scooter_id flag and not relying only on the unit/body-style pairing to identify e-scooter crashes.

- Mopeds: Most moped-involved crashes are consistently coded using the expected combination unit_desc_id = 1 and veh_body_styl_id = 92, which accounts for 111 crashes (89%). The remaining cases are small and scattered across other unit description values (e.g., unit_desc_id = 3, 4, or 7) while still showing body style 92, suggesting occasional inconsistencies in unit coding even when the body style indicates a moped.

Table 2. Variable Description

Variable	Code	Description
UNIT_DESC_ID	1	Motor Vehicle
	2	Train
	3	Pedalcyclist
	4	Pedestrian
	5	Motorized Conveyance
	7	Non-Contact
	8	Other (Explain in Narrative)
BODY_STYL_ID	92	Other (Explain in Narrative)

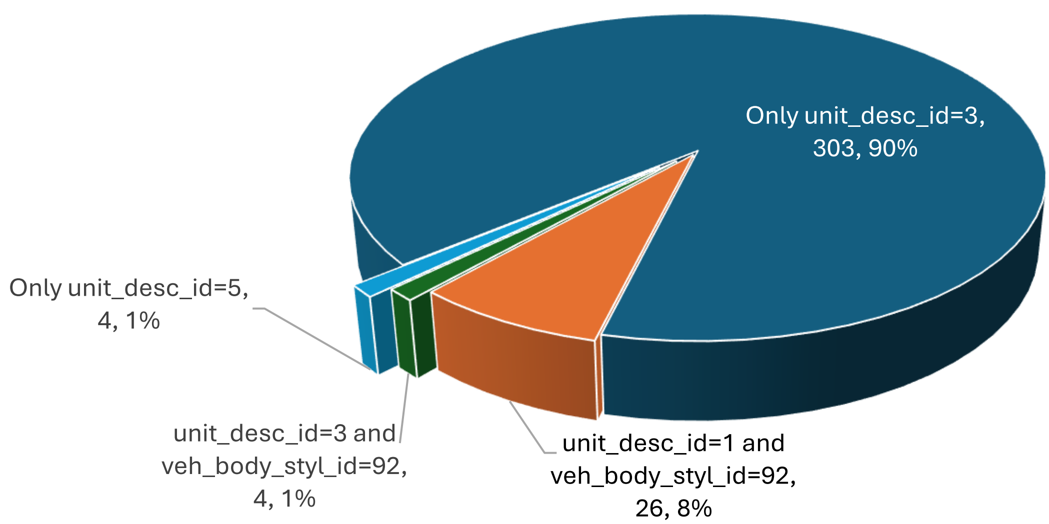

Figure 1. Unit description and body style of e-bikes involved in the crashes

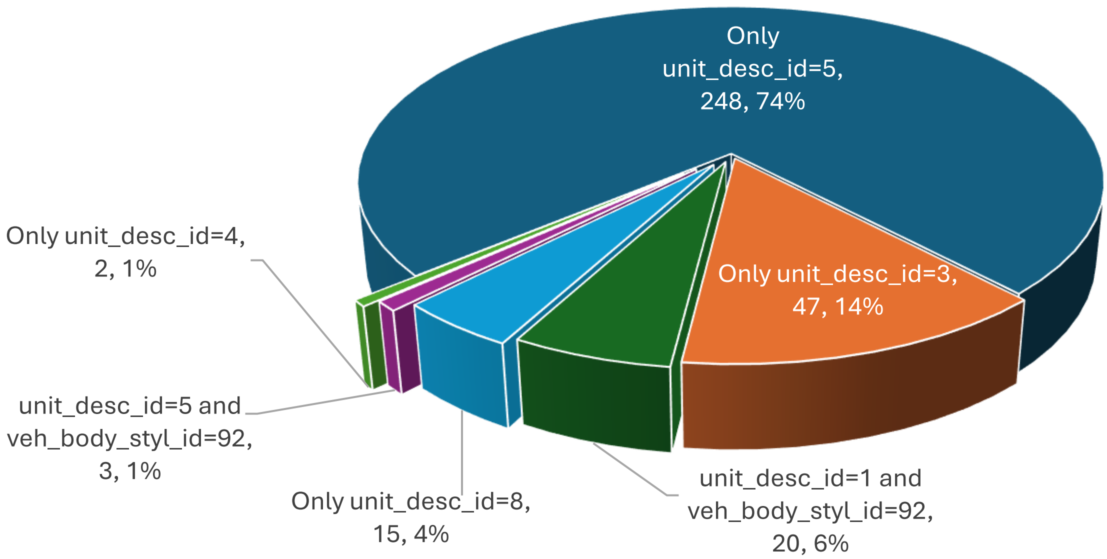

Figure 2. Unit description and body style of electric scooters involved in the crashes

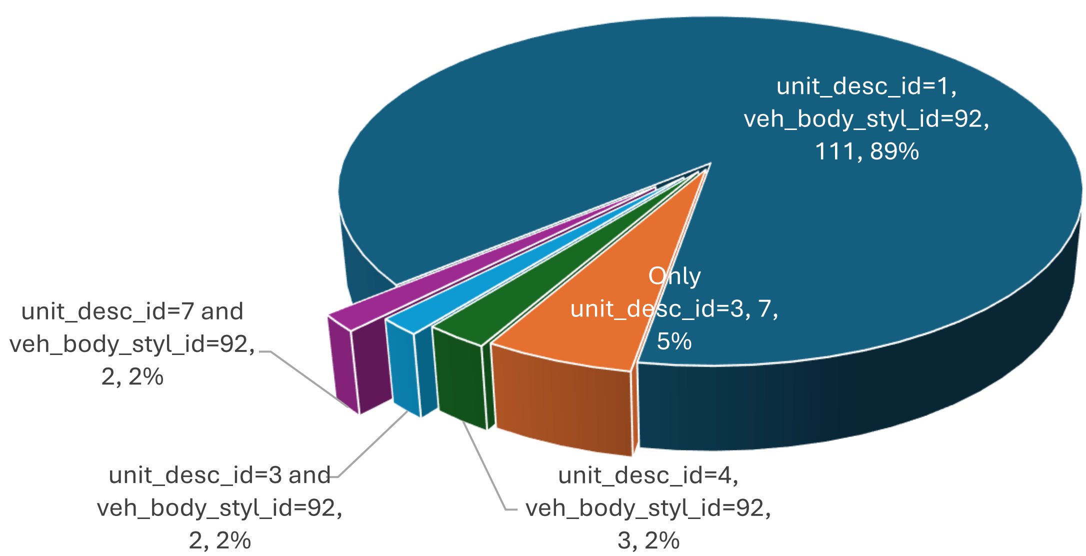

Figure 3. Unit description and body style of mopeds involved in the crashes

Limitations and Considerations

While using a keyword pattern approach improves the detection of e-bike–related crashes, there are several important limitations to consider. The narrative descriptions in crash reports can vary in both detail and terminology, which may result in some cases being missed or incorrectly classified. Additionally, certain crashes involving e-bikes may not contain explicit keywords, while others might use ambiguous language. To address these challenges, the team made an effort to include a wide range of terms related to e-bikes in the keyword pattern approach, aiming to capture as many relevant crashes as possible.

Crash Analysis

The objective of this analysis is to identify the characteristics of e-bike crashes—specifically crash severity, roadway and environmental conditions, contributing factors, and demographic patterns—by comparing e-bike crashes with non-electric bicycle (non-e-bike) crashes.

The TTI team assembled a dataset of all bicycle-related crashes from 2016–2025, consisting of 25,081 non-e-bike crashes and 393 electric-powered micromobility crashes. Of the 393 electric-powered cases, 337 e-bike crashes were identified through a systematic review of crash narratives using e-bike–related keywords (e.g., “e-bike,” “electric bicycle,” “battery-assisted bicycle”) as mentioned in the previous sections.

Because e-bikes, mopeds, and e-scooters are not always consistently coded in vehicle/unit fields, the team also included an additional 56 moped and e-scooter crashes based on unit description coding patterns that suggested pedalcyclist involvement. These cases met one of the following conditions:

- Unit Description = 3 and Vehicle Body Style = missing (54 crashes)
- Unit Description = 3 and Vehicle Body Style = 92 (2 crashes)

Yearly Trends

Figure 4 shows that non-e-bike crashes remained relatively stable from 2016 to 2025 (roughly 2,200–2,700 crashes per year), with a noticeable dip in 2020 followed by a gradual increase through 2024–2025. In contrast, e-bike crashes were rarely recorded prior to 2022 (generally fewer than 15 per year), but the number rises sharply beginning in 2022 (35 crashes) and increases each year thereafter—65 in 2023, 70 in 2024, and 165 in 2025. Overall, the trend suggests that e-bike crashes began appearing more consistently in the crash system starting around 2022–2023, with substantial growth in reported cases by 2025.

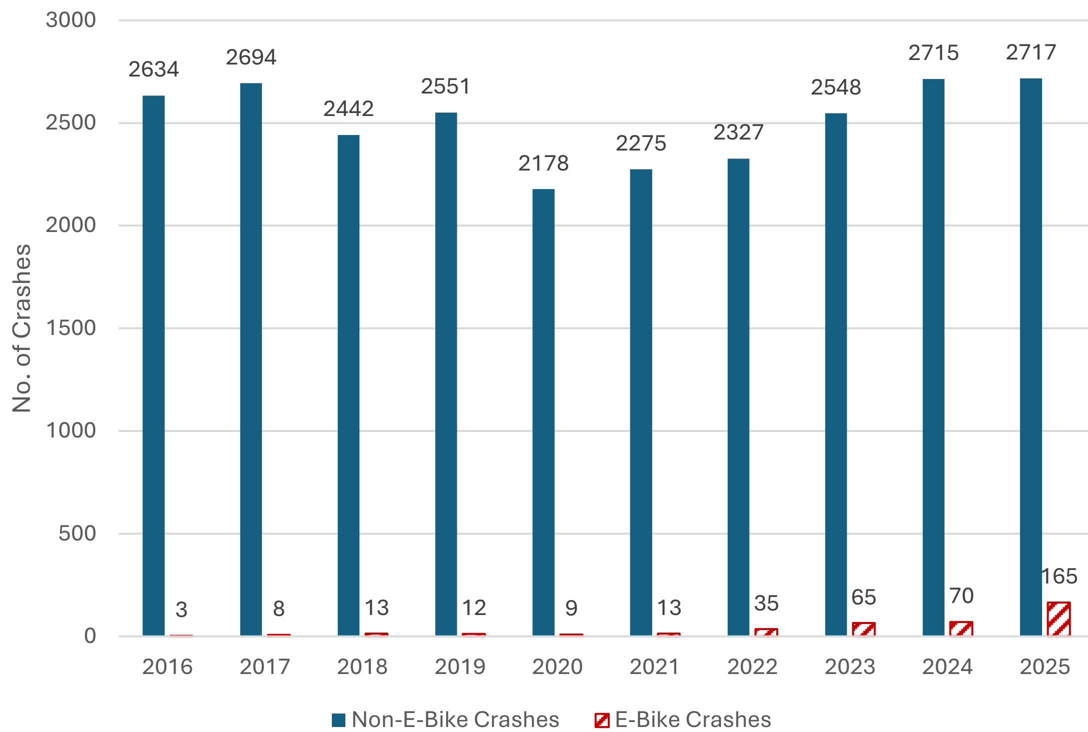

Figure 4. Yearly trends in e-bike and non-e-bike crashes, 2016 to 2025

Crash Distribution by County

Table 3 indicates that bicycle crashes (both non-electric and electric) are concentrated in a small number of large urban counties. Harris County has the highest number of crashes for both categories, accounting for 5,264 non-e-bike crashes (21%) and 47 e-bike crashes (12%). Bexar County represents an equal share of e-bike and non-e-bike crashes (11% each), while Travis County shows a higher proportion of e-bike crashes (11%) than non-e-bike crashes (9%). Overall, the top 10 counties account for a substantial share of crashes in both groups; however, e-bike crashes are relatively less concentrated in these counties (60%) compared with non-e-bike crashes (68%). The information for all counties is provided in Appendix A.

Table 3. Cashes in the Top 10 Counties

County	Non-E-bike Crashes	E-Bike Crashes
Harris	5264 (21%)	47 (12%)
Bexar	2802 (11%)	43 (11%)
Travis	2304 (9%)	43 (11%)
Dallas	2009 (8%)	29 (7%)
Tarrant	1499 (6%)	28 (7%)
Collin	730 (3%)	12 (3%)

Denton	687 (3%)	16 (4%)
Hidalgo	632 (3%)	2 (<1%)
El Paso	598 (2%)	10 (3%)
Galveston	596 (2%)	5 (1%)

Crash Severity

E-bike crashes show a higher share of minor injuries (57%) than non-e-bike crashes (46%) as shown in Figure 5. In contrast, non-e-bike crashes have a higher proportion of possible injuries (30%) compared to e-bike crashes (20%). The distributions for the most severe outcomes are similar across groups, with fatal crashes representing 3% of non-e-bike crashes and 2% of e-bike crashes, and serious injury crashes accounting for 13% and 14%, respectively. No-injury crashes are also comparable (8% for non-electric vs. 7% for electric).

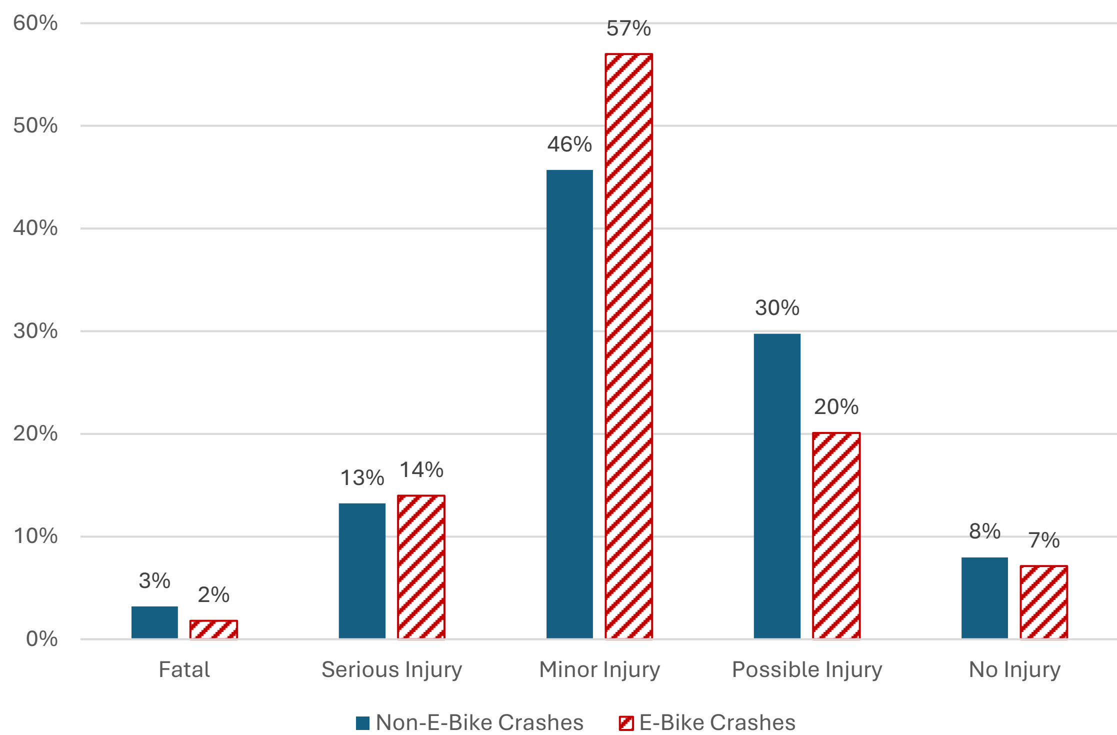

Figure 5. Crash severity distribution of e-bike and non-e-bike crashes

Roadway and Environmental Characteristics

This section compares e-bike and non-e-bike crashes across selected roadway and environmental characteristics, including posted speed limits and highway functional classification.

Posted Speed Limits

Figure 6 compares the distribution of posted speed limits for non-e-bike and e-bike crashes. Approximately three out of four crashes occurred on roadways with posted speed limits of 40 mph or lower. For both crash types, the largest share occurred on roadways posted at 30 mph (31% for non-e-bike crashes and 33% for e-bike crashes). The next most common posted speed limit was 35 mph (24% non-electric vs. 22% electric). Crashes on roadways posted at 40 mph and 45 mph each accounted for similar proportions across the two groups (approximately 13–14% in both).

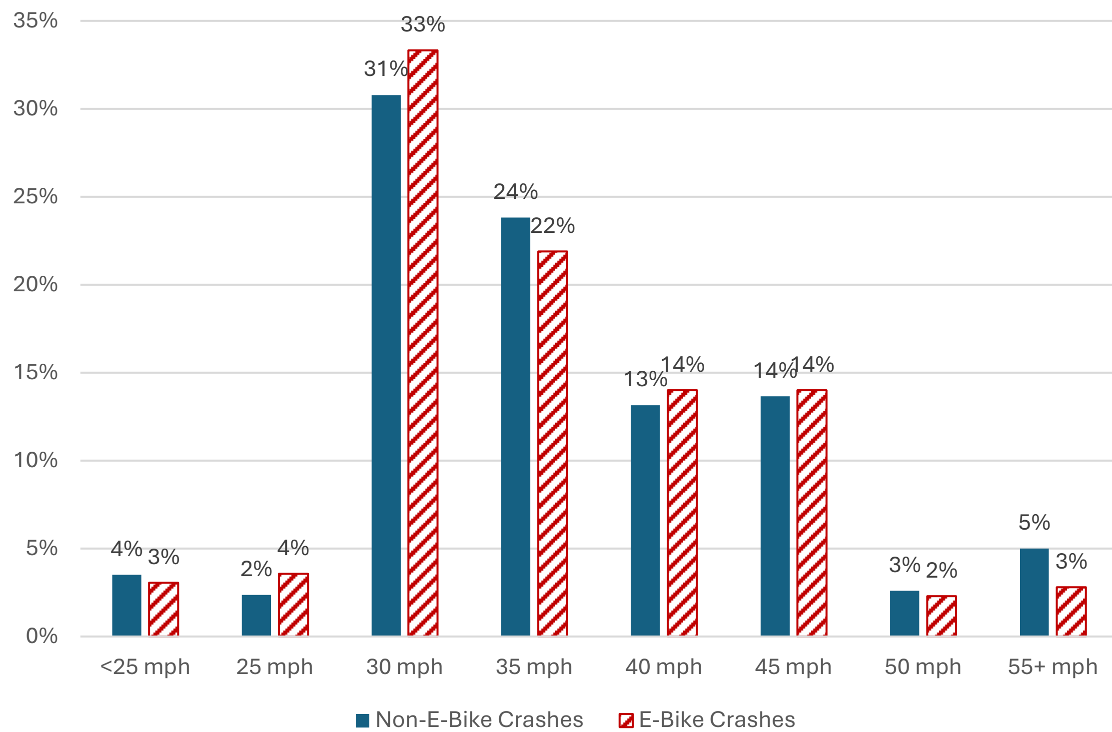

Figure 6 . Posted speed limits of e-bike and non-e-bike crashes

Highway Functional Classification

Figure 7 shows that bicycle crashes are concentrated on city streets for both e-bike and non-e-bike, accounting for 67% and 66% of crashes, respectively. U.S. and State highways represent the second-largest share for both groups (14% each). Crashes on farm-to-market roads and county roads are less common but similar across the two crash types (non-electric: 7% each; electric: 8% each). Although interstates account for the smallest proportion of crashes, it is notable that bicycle crashes occurred on these facilities given the limited accessibility for bicyclists in many locations. In addition, a small number of crashes occurred on tollways and other road types (approximately 0.1% to 0.3%), but these categories are not shown in Figure 7 due to their very low frequency.

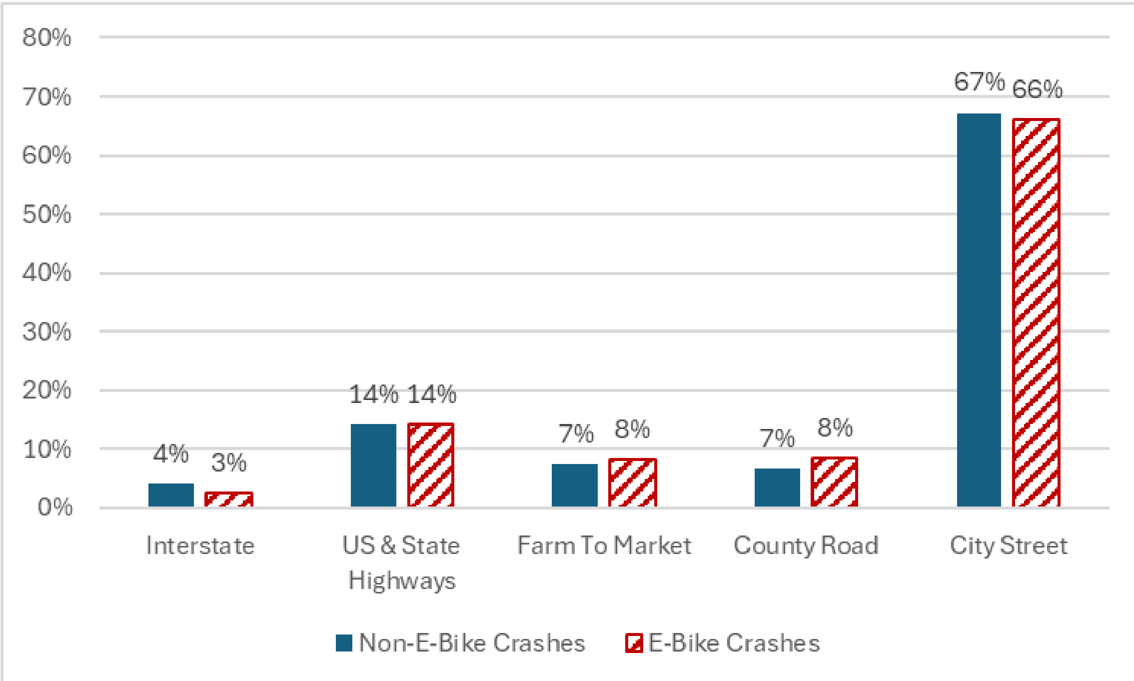

Figure 7. Crash distribution by highway functional classification

Collision Type by Location Context

Table 4 indicates that, for both e-bikes and non-e-bikes, the most common crash scenario involves one motor vehicle going straight, particularly at intersections (37% for non-e-bikes and 27% for e-bikes) and at non-intersection locations (27% for non-e-bikes vs. 20% for e-bikes).

Turning-movement crashes are also concentrated at intersections. Right-turn crashes account for 13% of intersection-related crashes for non-e-bikes and 14% for e-bikes, while left-turn crashes account for 11% and 12%, respectively.

Crashes at driveway access points represent a smaller share overall, but e-bike crashes show a higher proportion in turning scenarios. For example, driveway-related right-turn crashes are higher for e-bikes (9%) than for non-e-bikes (4%), and driveway-related left-turn crashes are also slightly higher for e-bikes (4% vs. 3%).

Table 4. Crash Distribution of Selected Collision Types by Location Context (Intersection-Related, Non-Intersection, and Driveway Access)

Collision Type	Intersection/-related		Not Intersection		Driveway Access	
	Non-E-Bike	E-Bike	Non-E-Bike	E-Bike	Non-E-Bike	E-Bike
One Motor Vehicle - Going Straight	37%	27%	27%	20%	4%	3%
One Motor Vehicle - Turning Right	13%	14%	<1%	<1%	4%	9%
One Motor Vehicle - Turning Left	11%	12%	<1%	2%	3%	4%

Day of Week

Figure 8 shows that crash occurrence is fairly evenly distributed across weekdays for both non-e-bike and e-bike crashes, with each weekday generally accounting for about 15–17% of crashes. E-bike crashes peak on Wednesday (17%) and are also relatively high on Monday and Tuesday (16% each). Non-e-bike crashes remain consistent from Tuesday through Friday (15% each) and are slightly lower on Monday (15%).

Both crash types decline over the weekend. Non-e-bike crashes drop to 13% on Saturday and 11% on Sunday, while e-bike crashes decrease to 12% on Saturday and Sunday. Overall, the lowest proportions for both groups occur on Sunday.

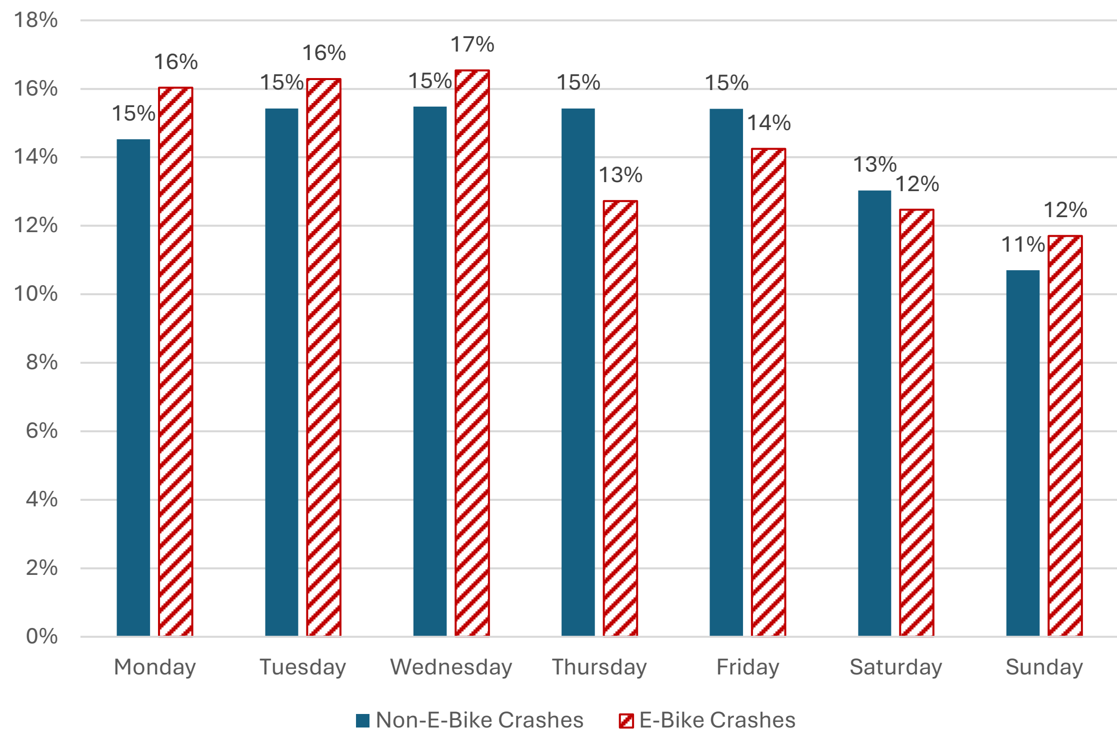

Figure 8. Crash Distribution by day of week

Time of Day

Time of day was broken down into two-hour intervals. Figure 9 shows that crashes for both e-bikes and non-e-bikes are concentrated during daytime and early evening hours. Crash proportions are lowest overnight (midnight to 5:59 a.m.), generally at 1–3% per two-hour interval. Crash involvement increases beginning in the morning (6:00–11:59 a.m.), remains elevated through the afternoon, and peaks during the late afternoon commute period (4:00–5:59 p.m.) at 18% for non-e-bikes and 19% for e-bikes. After 6:00 p.m., crash proportions gradually decline into the late evening.

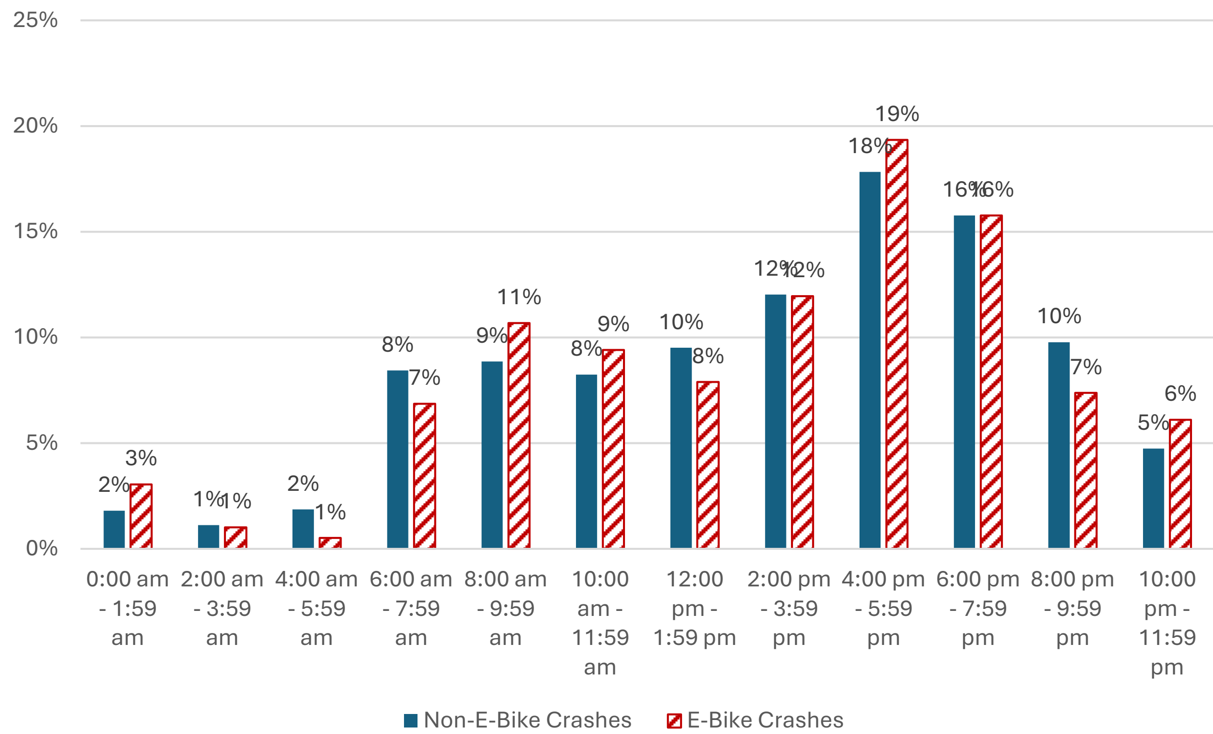

Figure 9. Crash distribution of e-bike and non-e-bike crashes by time of day (two-hour intervals)

Lighting Conditions

Figure 10 shows similar distribution between e-bike and non-e-bike crashes by lighting conditions. Most bicycle crashes occurred during daylight conditions for both non-e-bikes and e-bikes (about 70% each). Crashes occurring in darkness were less common but still notable: approximately 9–10% occurred in dark, not lighted conditions and about 16–17% occurred in dark, lighted conditions for both groups. Very few crashes were recorded under dark (unknown lighting) conditions (around 1% or less).

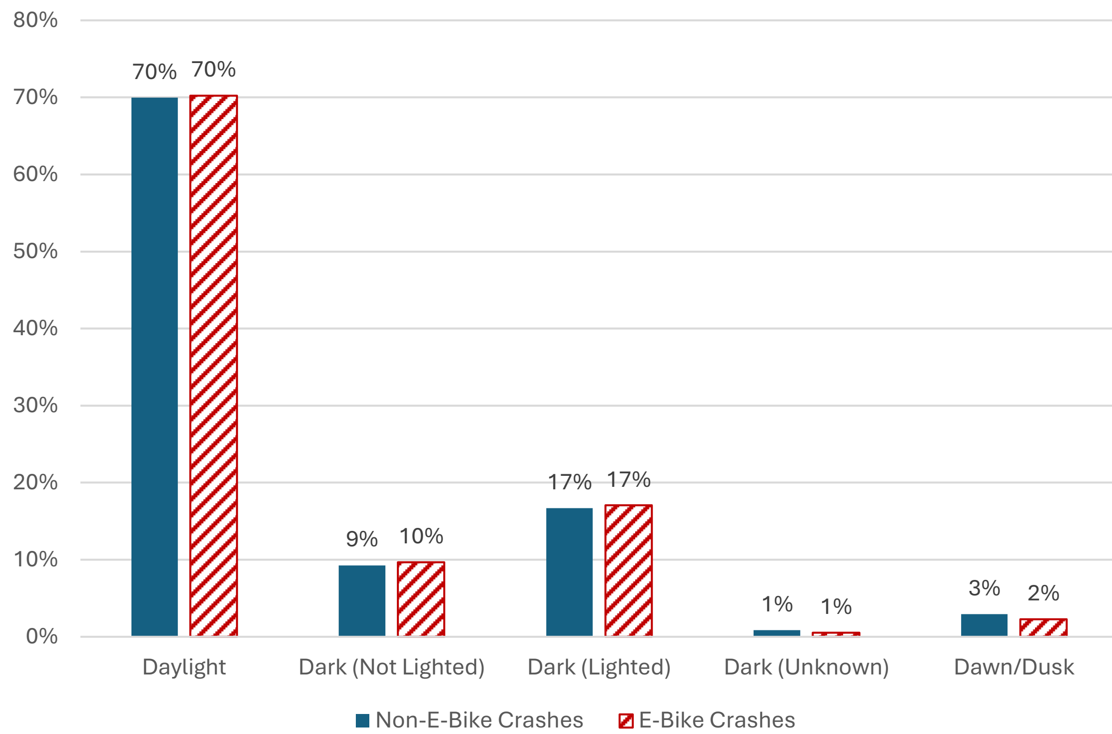

Figure 10. Crash distribution of e-bike and non-e-bike crashes by lighting conditions

Urban/Rural Area Type

The distribution of crashes by area type is very similar for e-bikes and non-e-bikes, indicating no meaningful difference between urban and rural settings (Figure 11). In both groups, the vast majority of crashes occurred in urban areas, with only a small share occurring in rural areas; however, e-bike crashes show a slightly higher proportion in rural areas.

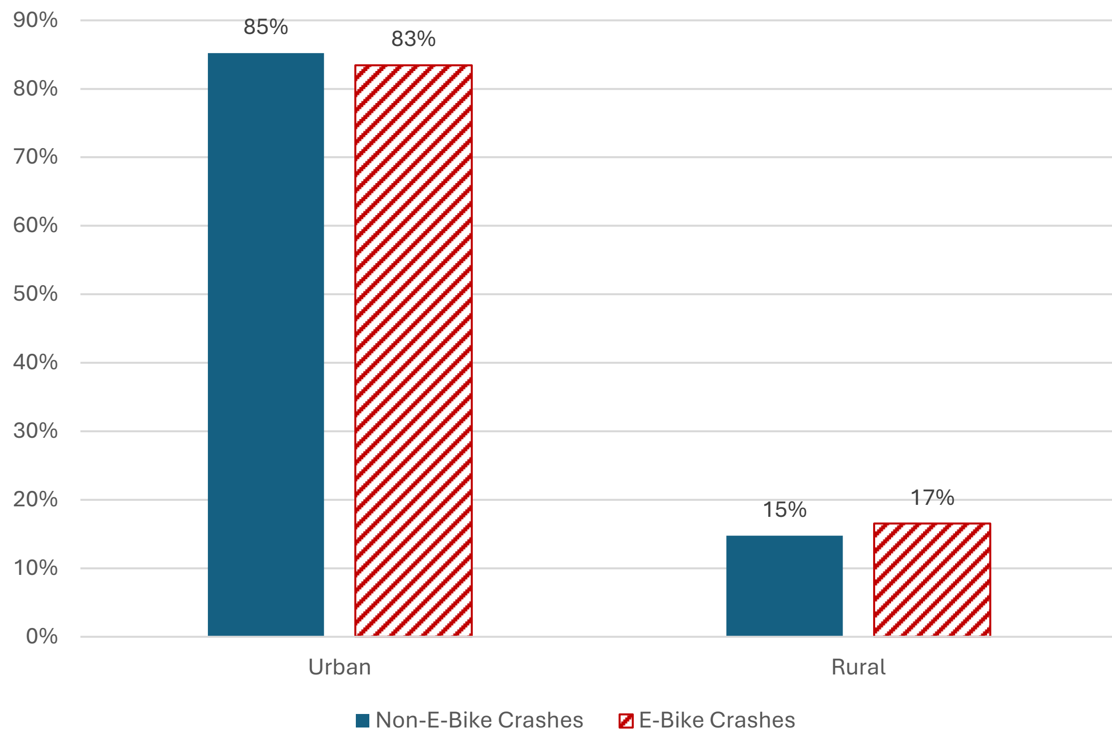

Figure 11. Crash distribution of e-bike and non-e-bike crashes by area type

Helmet Use

Figure 12 shows that helmet use was uncommon among both e-bike and non-e-bike users involved in crashes. Helmets were reported as worn in 19% of non-e-bike crashes and 14% of e-bike crashes, while helmets were not worn in 72% and 64%, respectively. E-bike users records also have a higher share of unknown helmet use (22%) compared with non-e-bike users (10%), indicating more missing or unreported helmet information for e-bike crashes.

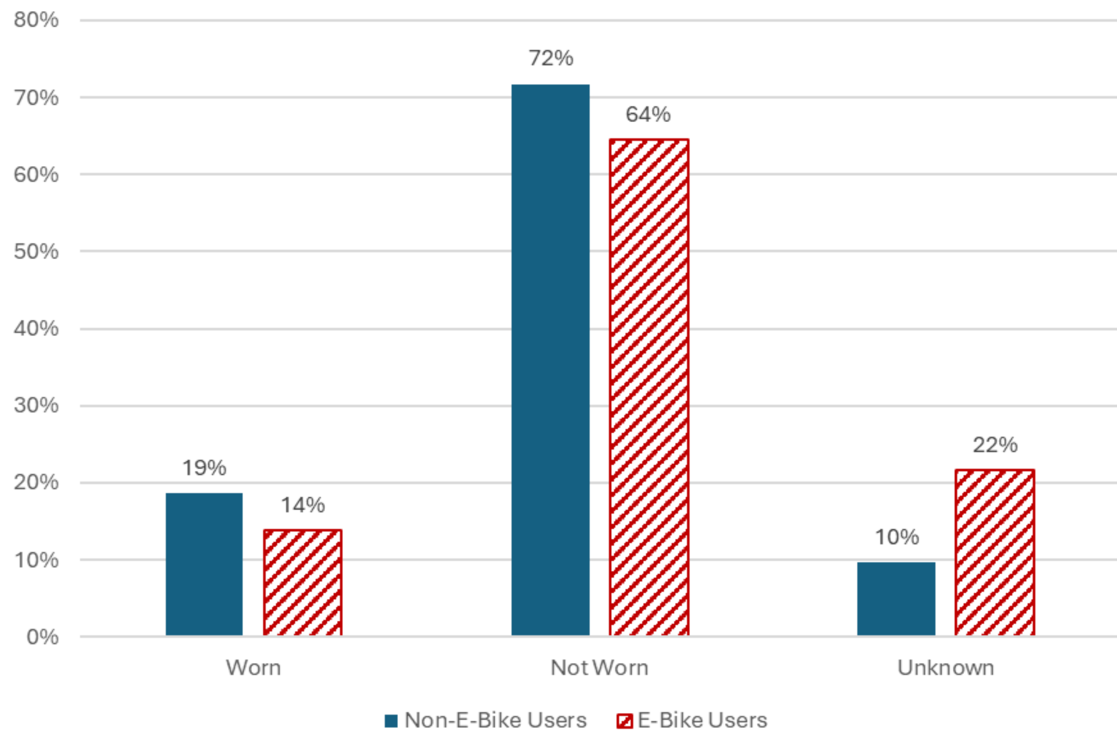

Figure 12. Helmet use of e-bike and non-e-bike users at the time of the crash.

Demographic Characteristics

The demographic analysis in this section is based on individual bicyclists involved in crashes, rather than crash records alone. Between 2016 and 2025, the dataset includes 25,081 non-e-bike crashes and 393 e-bike crashes. When the team extracted person-level records from CRIS, 24,805 non-e-bike users and 383 e-bike users were identified. The number of bicyclists is slightly lower than the number of crashes because some crashes do not contain complete or usable bicyclist/person information in CRIS—most commonly due to hit-and-run incidents and cases where the bicyclist could not be identified or did not provide sufficient information to the other party or responding officers.

Bicyclist Injury Severity

While Figure 5 presents overall crash severity based on the most severe injury among all individuals involved, Figure 13 focuses specifically on the injury severity of the bicyclist (e-bike and non-ebike users). The distributions in both figures are expected to be similar, as bicyclists—being more vulnerable than motor vehicle occupants—are typically the most severely injured in bicycle-motor vehicle crashes. As a result, the overall crash severity shown in Figure 5 often reflects the bicyclist's injury severity depicted in Figure 13, leading to closely aligned patterns across the two figures.

Consistent with bicyclists' vulnerability, most outcomes involve some level of injury for both groups (Figure 13). In both distributions, minor injury is the most common outcome, accounting for 45% of non-e-bike users and 56% of e-bike users. Serious injuries are similar between the two groups (13% non-electric vs. 12% electric), and fatal injuries are relatively rare (3% vs. 2%). A key difference is observed in possible injury, which represents a larger share among non-e-bike users (30%) than e-bike users (21%).

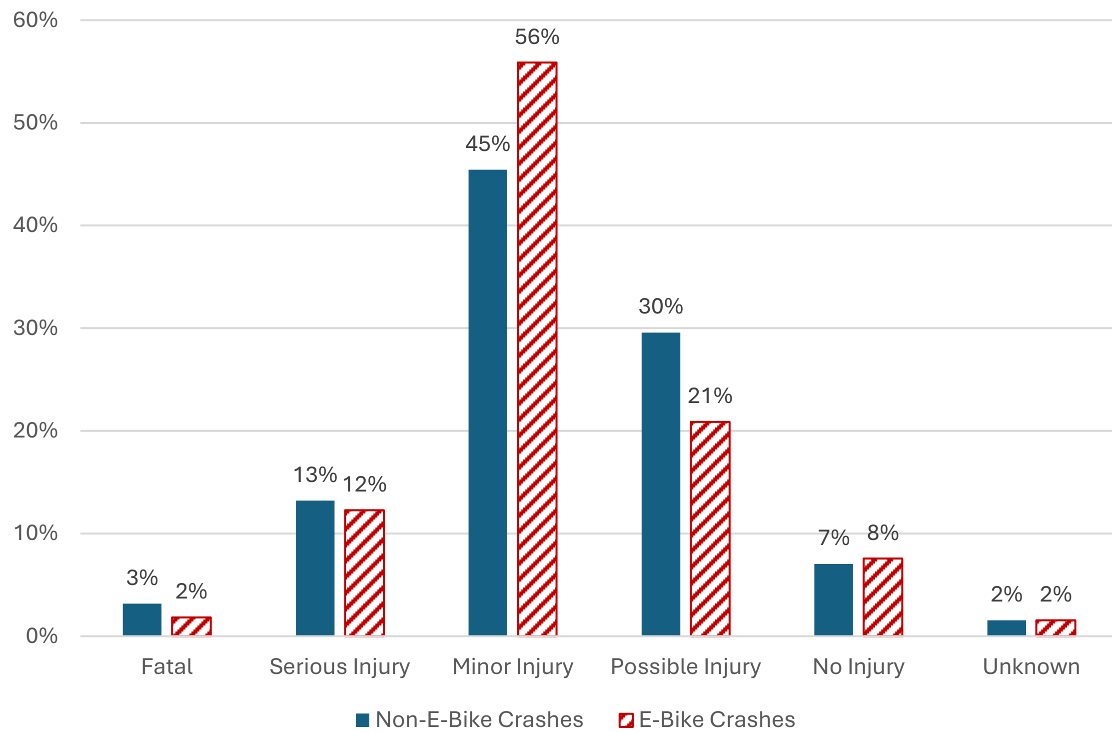

Figure 13. Distribution of bicyclist injury severity of e-bike and non-e-bike crashes

Bicyclist Age

Figure 14 indicates that age distributions differ between rider groups. E-bike crashes are relatively more concentrated among younger riders (ages 16–30), whereas non-e-bike crashes account for a higher proportion of older riders (age 51 and above). The largest share for both groups occurs among riders ages 21–30 (17% for non-e-bike users and 22% for e-bike users). E-bike users also represent higher proportions in the 10–15 (15% vs. 12%) and 16–20 (18% vs. 10%) age groups. In contrast, non-e-bike users represent larger shares in older age groups, particularly 51–60 (14% vs. 9%) and >60 (12% vs. 5%).

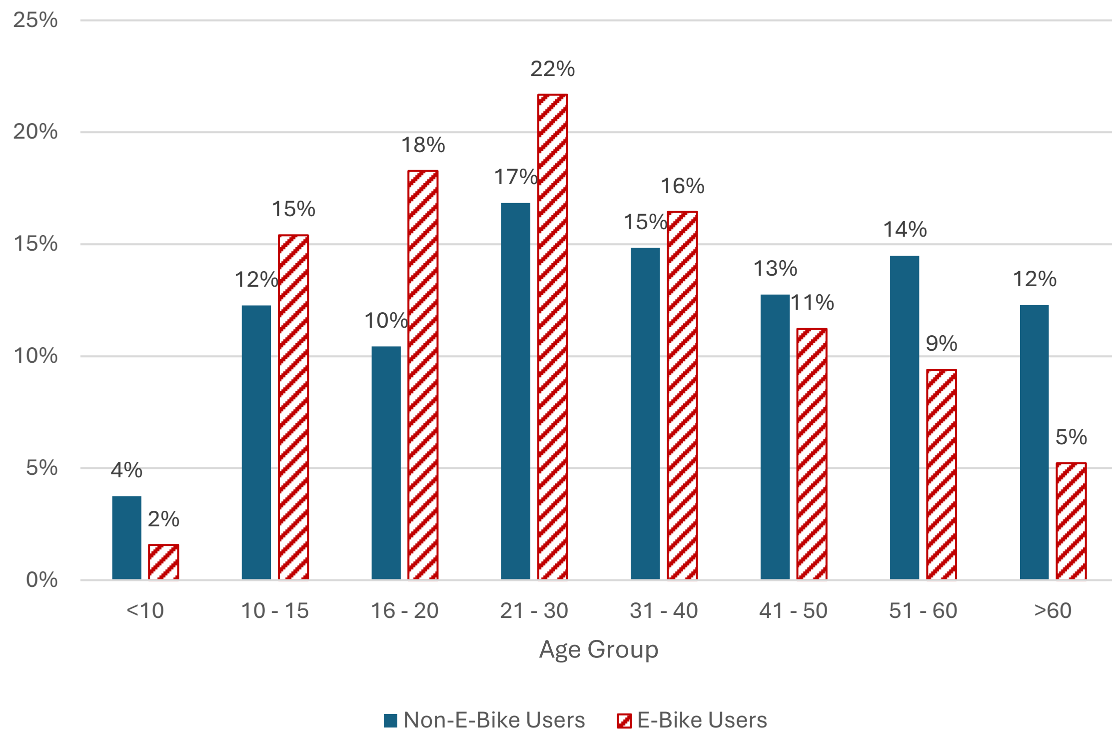

Figure 14. Distribution of bicyclist age group of e-bike and non-e-bike crashes

Figure 15 suggests that e-bike crashes are associated with a higher share of more severe bicyclist outcomes across most age groups. Figure 14 shows that e-bike users are more concentrated in younger age groups (particularly ages 16–30); however, Figure 15 indicates that across age groups, e-bike users generally experience higher proportions of fatal/serious/minor injury outcomes than non-e-bike users.

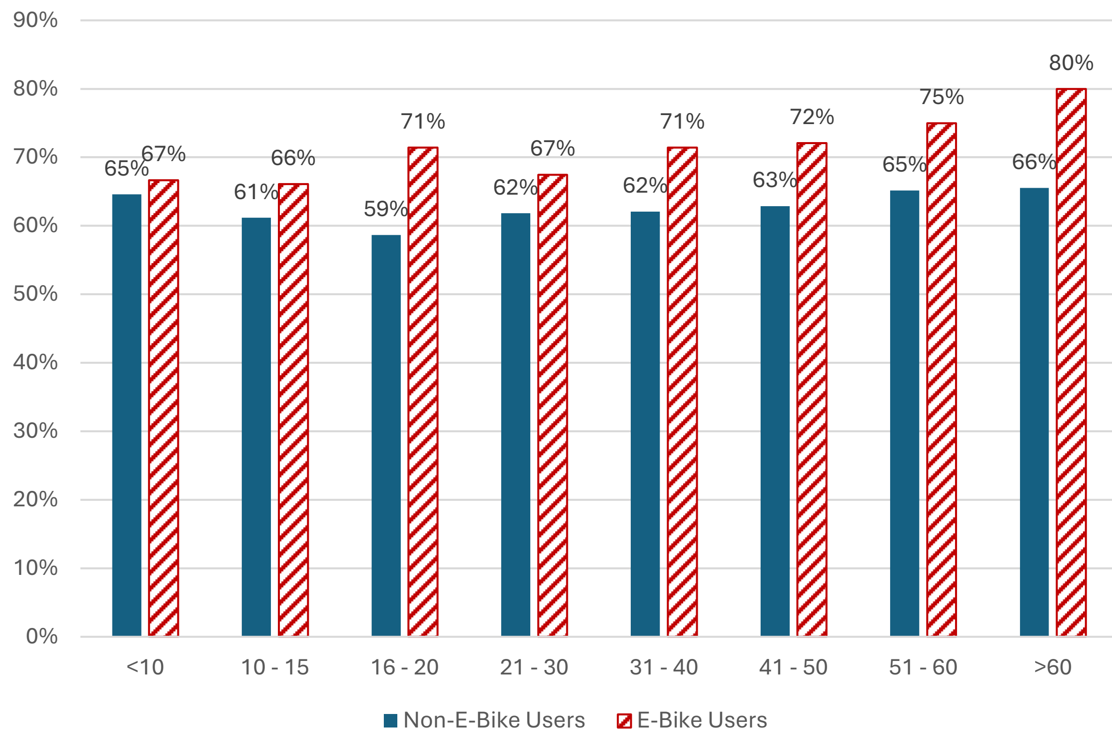

Figure 15. Percent of bicyclists with fatal, serious, and minor injuries by age group

Helmet non-use is common among injured bicyclists. Across age groups, a majority of bicyclists who sustained fatal, serious, or minor injuries were not wearing a helmet at the time of the crash (Figure 16). For non-e-bike users, the percent not wearing a helmet ranges from 62% (age >60) to 81% (ages 10–15 and 16–20). For e-bike users, non-use is also high in most age groups (generally 63%–77%), with particularly notable values for age <10 (100%). Overall, the results suggest consistently low helmet use among injured bicyclists, with some variation by age group and bicycle type.

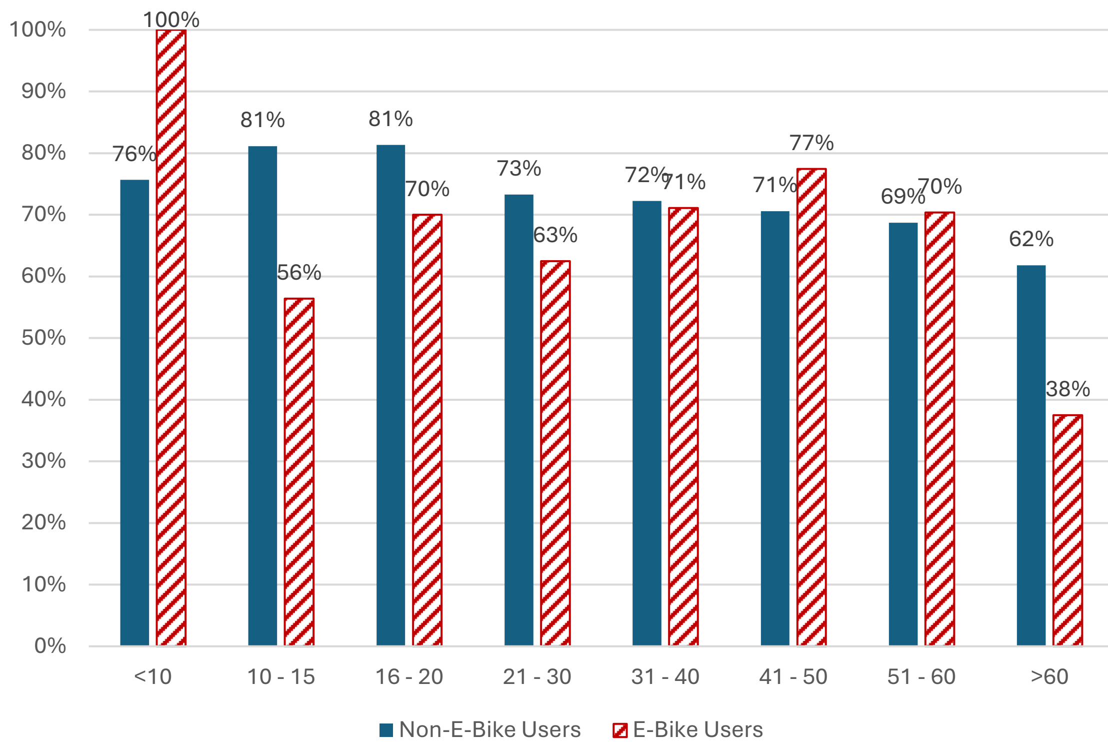

Figure 16. Percent of Injured bicyclists not wearing a helmet by age group

Bicyclist Gender

Bicyclists involved in crashes are predominantly male for both e-bikes and non-e-bikes (Figure 17). Male riders account for 83% of non-e-bike users and 86% of e-bike users, while female riders account for 17% and 14%, respectively. Overall, the gender distributions are similar between the two groups, with e-bike users showing a slightly higher proportion of males and a slightly lower proportion of females.

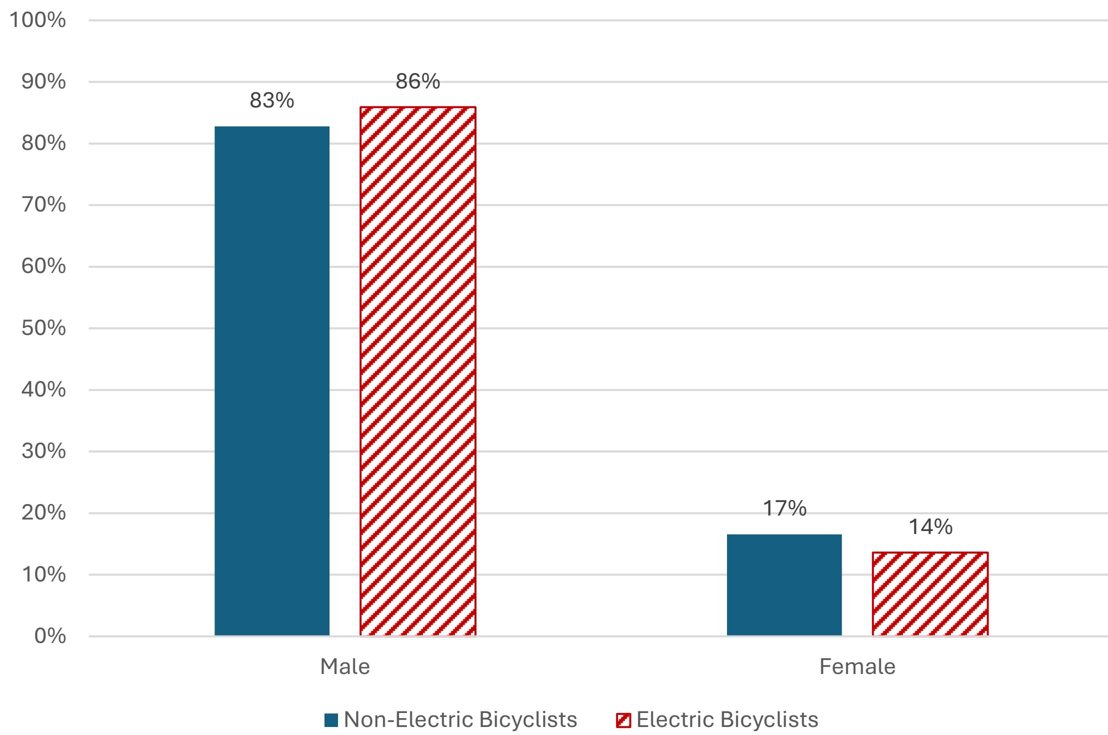

Figure 17. Distribution of bicyclist gender of e-bike and none-e-bike crashes

Crash Contributing Factors

Table 5 summarizes the most frequently reported crash contributing factors assigned to drivers, non-e-bike and e-bike users. Among drivers, the leading factors were driver inattention (n=2,434; 9%) and failure to yield right of way (n=2,010; 8%), followed by animal on road—wild (n=1,496; 6%). For non-e-bike users, the most common factor was "Other (Explain in Narrative)" (n=3,199; 13%), while other frequently cited factors included driver inattention (n=1,089; 4%) and violations related to traffic control devices (e.g., disregard stop sign/light and stop-and-go signal, each at about 3%). For e-bike users, "Other (Explain in Narrative)" was also the most common factor (n=60; 15%), followed by driver inattention (n=26; 7%) and failed to control speed (n=21; 5%). Overall, driver inattention appears as a prominent contributing factor across all three groups, while the relatively high share of "Other" for both bicyclist categories suggests that many bicyclist-related contributing circumstances may be captured primarily through narrative descriptions rather than predefined factor codes.

In addition, Table 5 compares how often selected contributing factors appear in all-severity crashes versus KAB (fatal + suspected serious + suspected minor injury) crashes, using the last column to indicate whether a factor is overrepresented (ratio higher than the overall KAB share for that group) or underrepresented (ratio lower than the overall KAB share).

For drivers, 58% of all driver-involved records are KAB crashes (0.58 overall). Most listed driver factors have ratios near this baseline (0.58–0.67), suggesting they occur in KAB crashes at roughly expected levels. "Animal on road—wild" (0.67) and "failed to control speed" (0.66) are modestly overrepresented in KAB crashes compared with the overall driver baseline, while "failed to yield—stop sign" and "failed to yield—to pedestrian" align closely with the baseline (both 0.58).

For non-e-bike users, the overall KAB share is 0.59. Several factors show ratios above this baseline—particularly "disregard stop sign or light" (0.67) and "disregard stop-and-go signal" (0.65)—indicating these violations are more common among KAB crashes than would be expected given the overall KAB proportion. "Other (explain in narrative)" is also slightly elevated (0.63), while "driver inattention" aligns with the baseline (0.58).

For e-bike users, the overall KAB share is higher (0.67), meaning a larger portion of e-bike users are involved in KAB crashes. Within this group, several factors are notably overrepresented relative to the 0.67 baseline, including driver inattention (0.81), "animal on road—wild" (0.79), "failed to yield—stop sign" (0.80), and "disregard stop sign or light" (0.79). In contrast, "failed to control speed" is underrepresented (0.48), indicating it appears less frequently in KAB crashes than would be expected given the overall KAB share for electric bicyclists.

Table 5. Selected Crash Contribution Factors Assigned to Drivers, Non-E-Bike and E-Bike Users

Category	Factor	All Severity Crashes		KAB Severity Crashes		Proportion of Frequency in KAB to all Severity
		Frequency	%	Frequency	%	
Drivers	Driver Inattention	2434	9%	1473	10%	0.61
	Failed To Yield Right Of Way - To Pedestrian	2010	8%	1165	8%	0.58
	Animal On Road - Wild	1496	6%	995	7%	0.67
	Other (Explain In Narrative)	1276	5%	804	5%	0.63
	Failed To Yield Right Of Way - Stop Sign	1103	4%	637	4%	0.58
	Failed To Control Speed	1094	4%	723	5%	0.66
Total Drivers		25876		14939		0.58
Non-E-Bike Users	Other (Explain In Narrative)	3199	13%	2017	13%	0.63
	Driver Inattention	1089	4%	633	4%	0.58
	Pedestrian Failed To Yield Right Of Way To Vehicle	839	3%	508	3%	0.61
	Disregard Stop Sign Or Light	829	3%	559	4%	0.67
	Disregard Stop And Go Signal	793	3%	516	3%	0.65

	Animal On Road - Wild	750	3%	486	3%	0.65
	Total Bicyclists	25322		14951		0.59
E-Bike Users	Other (Explain In Narrative)	60	15%	42	16%	0.70
	Driver Inattention	26	7%	21	8%	0.81
	Failed To Control Speed	21	5%	10	4%	0.48
	Animal On Road - Wild	19	5%	15	6%	0.79
	Failed To Yield Right Of Way - Stop Sign	15	4%	12	5%	0.80
	Disregard Stop Sign Or Light	14	4%	11	4%	0.79
	Total Bicyclists	393		265		0.67

Conclusions on E-Bike vs. Non-E-Bike Crash Comparison

The goal of this analysis was to better understand how e-bike crashes compare to traditional bike crashes by a number of different variables. While some variables did not vary much from e-bikes to traditional bikes, others did. From this data, there are a few key takeaways:

- Across age groups, e-bike users generally experience higher proportions of fatal/serious/minor injury outcomes than non-e-bike users.
- Helmet use is higher among e-bike users aged 10-40 involved in a crash compared to traditional bike riders, however there was also higher percentage of unknown helmet use among e-bike users involved in crashes.
- Helmet non-use for e-bike riders under the age of 10 injured in a crash was 100%.
- E-bike demographics do differ from traditional bike riders with E-bike users involved in crashes tending to be younger (under the age of 40) and the percentage of males using e-bikes being slightly higher compared to females.
- The most common contributing factors for drivers were driver inattention, failure to yield right-of-way and animal on road
- For all bikes, “Other (explain in narrative)” was the most common contributing factor as was driver inattention. However, for e-bikes failed to control speed also showed up.

One of the most challenging tasks of this analysis was identifying e-bike crashes in the CRIS dataset. While there could have been some e-bikes missed in this process, the project team felt the method chosen provided a good dataset and balanced accuracy with efficient use of resources.

Crash Typing

In addition to the crash analysis described above, the project team used the Pedestrian and Bicycle Crash Analysis Tool (PBCAT), Version 3 to categorize a small sample of crashes to better understand the circumstances leading to e-bike crashes. To do this analysis, the team selected the crashes involving the younger e-bike users, 40 or younger, who were involved in crashes resulting in a KAB injury. From the dataset, 100 crashes were identified as crashes involving an e-bike which resulted in a fatality, suspected serious injury or suspected minor injury. Using the matrix of crash types from the PBCAT User Guide[1] as shown in Figure 18, the project team selected the crash type for each of the 100 e-bike crashes based on the motorist and non-motorist maneuvers.

As seen in the figure, a shorthand code is determined for each crash indicating the motorist maneuvers first and the non-motorist maneuver second following the dash. The left-most column is the motorist maneuver and the top row is the non-motorist maneuver. If the motorist was going straight when the crash occurred, the first letter would be S. If the non-motorist was crossing the path of the motorist the first letter after the dash would be C. The final letter indicates whether the crossing non-motorist was coming from the left or right of the motorist. If they were coming from the motorist's right in this scenario, the crash type would be S-CR. The user guide provides detailed depictions of each crash type.

Non-Motorist Maneuver / *Motorist Maneuver*	CR: Crossing Path from Motorist's Right	CL: Crossing Path from Motorist's Left	CU: Crossing Path, Unknown Direction	PS: Parallel Path Same Direction	PO: Parallel Path Opposite Direction	PU: Parallel Path Unknown Direction	MU: Moving in Unknown Path/ Direction	ST: Stationary	OU: Other/ Unusual	UN: Unknown	FC: Non-motorist Fall or Crash
S: Going Straight	S-CR	S-CL	S-CU	S-PS	S-PO	S-PU	S-MU	S-ST	S-OU	S-UN	n/a[2]
R: Turning Right	R-CR	R-CL	R-CU	R-PS	R-PO	R-PU	R-MU	R-ST	R-OU	R-UN	
L: Turning Left	L-CR	L-CL	L-CU	L-PS	L-PO	L-PU	L-MU	L-ST	L-OU	L-UN	
P: Parked	P-CR	P-CL	P-CU	P-PS	P-PO	P-PU	P-MU	n/a	P-OU	P-UN	
E: Entering Traffic Lane	E-CR	E-CL	E-CU	E-PS	E-PO	E-PU	E-MU	E-ST	E-OU	E-UN	
B: Backing	B-CR	B-CL	B-CU	B-PS	B-PO	B-PU	B-MU	B-ST	B-OU	B-UN	
O: Other Maneuver	O-CR	O-CL	O-CU	O-PS	O-PO	O-PU	O-MU	O-ST	O-OU	O-UN	
U: Unknown Maneuver	U-CR	U-CL	U-CU	U-PS	U-PO	U-PU	U-MU	U-ST	U-OU	U-UN	
N: Non-Collision	Not Applicable – No Crash type returned										N-FC

Figure 18. Matrix of crash type detailed from motorist and non-motorist maneuver selections

[1] https://highways.dot.gov/turner-fairbank-highway-research-center/pbcat

The results of the PBCAT crash typing of the 100 e-bike crashes involving people ages 40 and younger showed that a slight majority of e-bike crashes occurred when the motorist was traveling straight (52 crashes, 52%). As seen in Table 6, one quarter of these 100 e-bike crashes involved a motorist turning right (25 crashes, 25%) and 16% of the crashes involved a left turning motorist. Other motorist maneuvers included parked (not moving), backing, an other/unusual.

Table 6. Motorist maneuvers as determined from the crash narrative review

Motorist Maneuver	Crash Frequency
Going straight	52
Turning right	25
Turning left	16
Parked	4
Backing	2
Other/unusual	1
Total	100

As for the non-motorist maneuver and direction of travel in relation to the motor vehicle, 50 crashes involved a non-motorist traveling in a parallel direction to the motorist and 48 of the crashes involved a crossing non-motorist. As seen in Table 7, the top crash types using the PBCAT codes are:

- Straight motorist colliding with an ebike on the right (S-CR) – 19 crashes
- Straight motorist colliding with an ebike on the left (S-CL) – 17 crashes
- Straight motorist colliding with an ebike going the same direction (e.g., rear end, sideswipe) (S-PS) – 13 crashes
- Right turning motorist colliding with an ebike going the same direction (e.g., right hook crash) (R-PS) – 13 crashes
- Left turning motorist colliding with an ebike going in the opposite direction (left turn crash) – 13 crashes

Table 7. PBCAT crash type totals

PBCAT Crash Typing Code	Total
S-CR	19
S-CL	17
L-PO	13
R-PS	13
S-PS	13
R-CR	5
P-PS	4

R-CL	3
B-CL	2
L-PS	2
R-PO	2
R-PU	2
L-CR	1
O-OU	1
S-CU	1
S-PO	1
S-UN	1
Total	100

In addition, as of 2021 the TxDOT CRIS data included actions on the part of both the motorist and non-motorist prior to the crash. These codes, as found in Appendix C, reflect more detailed contributing factors that are not in the police reported crashes. For the 100 e-bike crashes involving riders 40 and younger and a fatal or injury crash (KAB), the top crash typing codes applied in CRIS were determined.

The top pedalcyclist actions, as shown in Table 8, are riding cycle while crossing road at intersection or crosswalk (marked or unmarked) (57) and riding in roadway (vehicle lane - unknown if bike markings present) - with traffic (12).

Table 8. Pedalcyclist action by frequency

Pedalcyclist Action	Frequency
Riding cycle while crossing road at intersection or crosswalk (marked or unmarked)	57
Riding in roadway (vehicle lane - unknown if bike markings present) - with traffic	12
Riding in bike lane - with traffic	7
Riding on sidewalk	6
Riding cycle while crossing road at driveway	5
Riding cycle while crossing road not at intersection or crosswalk	3
Data not available at data entry	3
Riding in shoulder - with traffic	2
In parking lot	1
Not in roadway	1
Riding in bike lane - against traffic	1
Riding in roadway (vehicle lane without bike markings) - with traffic	1
Unknown	1

The top PBCAT crash types, as shown in Table X, include bicyclist ride through - signalized intersection (25), motorist right turn - same direction (10), and motorist left turn - opposite direction (9).

Table 9. Pedalcyclist PBCAT crash type by frequency

Pedalcyclist PBCAT Crash Type	Frequency
Bicyclist ride through - signalized intersection	25
Motorist right turn - same direction	10
Motorist left turn - opposite direction	9
Crossing paths - intersection - other/unknown	6
Motorist right turn - opposite direction	5
Parallel paths - other/unknown	5
Bicyclist ride through - sign-controlled intersection	3
Motorist drive out - commercial driveway/alley	3
Motorist left turn - same direction	3
Motorist overtaking - other/unknown	3
Data not available at data entry	3
Backing vehicle	2
Bicyclist left turn - same direction	2
Bicyclist ride out - sign-controlled intersection	2
Motorist drive out - sign-controlled intersection	2
Motorist drive out - signalized intersection	2
Motorist turning error - left turn	2
Bicyclist overtaking - passing on right	1
Bicyclist ride out - midblock - unknown	1
Bicyclist right turn - same direction	1
Crossing paths - uncontrolled intersection	1
Motorist drive out - midblock - unknown	1
Motorist drive out - residential driveway	1
Motorist drive through - sign-controlled intersection	1
Motorist lost control - other/unknown	1
Motorist right turn on red - same direction	1
Motorist turning error - right turn	1
Multiple threat - sign-controlled intersection	1
Non-roadway	1
Unusual circumstances	1

Conclusion on PBCAT Crash Analysis

Therefore, the analysis of PBCAT codes and crash typing shows that most of the e-bike crashes involving people 40 years old and younger with a KAB severity are at intersections and more likely at intersections controlled by a traffic signal. They are either with a motor vehicle going straight or turning right or left. Overall, this analysis was helpful in helping to guide messaging and outreach efforts surrounding e-bikes specifically. The analysis does point to demographic differences, severity differences and some behaviors such as helmet use that can be a focus for outreach, but also shows that overall e-bikes are involved in crashes mostly on the same roads, in the same areas, and during the same times as traditional bikes.

Appendix A. Python Code for Keyword Pattern Analysis

E-Bike / Moped / E-Scooter Crash Identification & Precision Analysis

==

This script:

1. Loads ALL crash narratives from bikecrashes.csv filtered to year == 2024
2. Searches investigator_narrative for e-bike related keywords (strict matching)
3. Adds evaluation columns (keyword_match, vehicle classification) to ALL records
4. Matches ALL cases with vehicle classifications using:
   - moped   : unit_desc_id == 1 AND veh_body_styl_id == 92
   - e-scooter : (unit_desc_id == 5 AND veh_body_styl_id == 92) OR e_scooter_id == 1
   - unclassified e-bike (unit_desc_id == 3): pedalcyclist with keyword match
   - motorcycle: excluded from e-bike classification
   - unclassified e-bike (other): keyword match but no matching condition
5. Captures ALL units involved in each crash (not just one)
6. Calculates precision of keyword-based classifications
7. Outputs complete dataset with evaluation columns in CSV and Excel

"""

```
import pandas as pd

import re

from pathlib import Path

from openpyxl.styles import PatternFill, Font, Alignment

from openpyxl.utils import get_column_letter
```

```
# ==================================================================================

# CONFIGURATION
# ==================================================================================


NARRATIVE_FILE   = "bikecrashes.csv"      # Narrative CSV file
VEHICLE_FILE     = "bikeunits.csv"        # Vehicle CSV file


# Column names — narrative file
NARRATIVE_CASE_ID_COL  = "crash_id"
NARRATIVE_TEXT_COL     = "investigator_narrative"
NARRATIVE_YEAR_COL     = "year"         # Set to None if already filtered to 2024


# Column names — vehicle file
VEHICLE_CASE_ID_COL    = "crash_id"
VEHICLE_UNIT_DESC_COL  = "unit_desc_id"
VEHICLE_BODY_STYLE_COL = "veh_body_styl_id"
VEHICLE_ESCOOTER_COL   = "e_scooter_id"   # e-scooter flag variable


# Vehicle body style code for motorcycles (adjust based on your data)
MOTORCYCLE_BODY_STYLE_ID = 71  # Update this with the actual motorcycle code from
your data


# Output file paths
```

```
OUTPUT_CSV   = "all_crashes_with_ebike_evaluation.csv"

OUTPUT_EXCEL = "all_crashes_ebike_report.xlsx"


#
===========================================================================
====

# KEYWORD PATTERNS - STRICT MATCHING ONLY

#
===========================================================================
====


EBIKE_PATTERNS = [

   r"\be[\s\-]?bike\b",            # e-bike, e bike, ebike (NOT standalone "bike")

   r"\be[\s\-]?bicycle\b",         # e-bicycle, e bicycle (NOT standalone "bicycle")

   r"\belectric[\s\-]?bike\b",      # electric bike, electric-bike

   r"\belectric[\s\-]?bicycle\b",    # electric bicycle

   r"\belectric[\s\-]?cycle\b",     # electric cycle

   r"\belectric[\s\-]?scooter\b",    # electric scooter

   r"\be[\s\-]?scooter\b",         # e-scooter, escooter

   r"\bmotor[\s\-]?bicycle\b",      # motor bicycle (NOT motorbike)

   r"\bmotorized[\s\-]?bicycle\b",   # motorized bicycle

   r"\bbattery[\s\-]?assisted[\s\-]?bike\b",     # battery-assisted bike

   r"\bbattery[\s\-]?assisted[\s\-]?bicycle\b",   # battery-assisted bicycle

   r"\bbattery[\s\-]?powered[\s\-]?bike\b",      # battery-powered bike

   r"\bbattery[\s\-]?powered[\s\-]?bicycle\b",    # battery-powered bicycle

   r"\bpedal[\s\-]?assist\b",       # pedal assist

   r"\bpedelec\b",               # pedelec
```

```
    r"\bmoped\b",                   # moped
]

COMBINED_PATTERN = re.compile(
    "|".join(EBIKE_PATTERNS),
    flags=re.IGNORECASE
)

#
==========================================================================
====
# HELPER FUNCTIONS
#
==========================================================================
====

def load_file(filepath: str, label: str, sheet: int = 0) -> pd.DataFrame:
    """Load a CSV or Excel file into a DataFrame."""
    path = Path(filepath)
    if not path.exists():
        raise FileNotFoundError(
            f"[ERROR] {label} file not found: '{filepath}'\n"
            f"        Please confirm the file is in the same folder as this script."
        )
    size_kb = path.stat().st_size / 1024
    print(f"  Loading {label}: {path.name}  ({size_kb:.1f} KB)")
```

```
    if path.suffix.lower() in [".xlsx", ".xls"]:
        return pd.read_excel(path, sheet_name=sheet)

    # Try multiple encodings for CSV files
    encodings = ['utf-8', 'latin-1', 'iso-8859-1', 'cp1252']

    for encoding in encodings:
        try:
            return pd.read_csv(path, low_memory=False, encoding=encoding)
        except UnicodeDecodeError:
            if encoding == encodings[-1]:  # Last encoding attempt
                raise
            continue

    return pd.read_csv(path, low_memory=False)



def validate_columns(df: pd.DataFrame, required_cols: list, file_label: str):
    """Raise a clear error if any expected columns are missing."""
    missing = [c for c in required_cols if c not in df.columns]
    if missing:
        raise KeyError(
            f"[ERROR] The following columns were not found in {file_label}:\n"
            f"    {missing}\n"
            f"    Available columns: {list(df.columns)}"
        )
```

```
def find_matched_keywords(text: str) -> str:
    """Return a comma-separated string of all unique keyword matches found in text."""
    if not isinstance(text, str):
        return ""
    matches = COMBINED_PATTERN.findall(text)
    return ", ".join(sorted(set(m.lower() for m in matches)))


def has_keyword_match(text: str) -> bool:
    """Return True if any e-bike keyword is found in text."""
    if not isinstance(text, str):
        return False
    return bool(COMBINED_PATTERN.search(text))


def classify_vehicle(row) -> str:
    """
    Apply classification rules based on unit_desc_id, veh_body_styl_id, and e_scooter_id.
     - moped     : unit_desc_id == 1 AND veh_body_styl_id == 92
     - e-scooter : (unit_desc_id == 5 AND veh_body_styl_id == 92) OR e_scooter_id == 1
     - Unclassified E-Bike (Pedalcyclist): unit_desc_id == 3
     - Motorcycle (excluded): veh_body_styl_id == MOTORCYCLE_BODY_STYLE_ID
     - No vehicle record found
     - Unclassified E-Bike (Other): keyword match, but no matching vehicle code
```

```
  - Not E-Bike Related: no keyword match and no e-bike vehicle code
"""
uid  = row.get(VEHICLE_UNIT_DESC_COL)
bsid = row.get(VEHICLE_BODY_STYLE_COL)
escooter = row.get(VEHICLE_ESCOOTER_COL)
has_keyword = row.get("ebike_keyword_flag", False)

if pd.isna(uid) or pd.isna(bsid):
    if has_keyword:
        return "No Vehicle Record (keyword match)"
    return "No Vehicle Record"

# Exclude motorcycles
if bsid == MOTORCYCLE_BODY_STYLE_ID:
    return "Motorcycle (excluded)"

# Moped classification
if uid == 1 and bsid == 92:
    return "Moped (unit_desc_id=1, veh_body_styl_id=92)"

# E-Scooter classification - TWO CONDITIONS
if (uid == 5 and bsid == 92) or (escooter == 1):
    return "E-Scooter (unit_desc_id=5, veh_body_styl_id=92 OR e_scooter_id=1)"

# Unclassified E-Bike - Pedalcyclist (unit_desc_id = 3)
if uid == 3:
```

```
        if has_keyword:
            return "Unclassified E-Bike - Pedalcyclist (unit_desc_id=3, keyword match)"
        return "Pedalcyclist (unit_desc_id=3, no keyword)"

    # Other cases with keyword match
    if has_keyword:
        return "Unclassified E-Bike - Other (keyword match, no matching vehicle code)"

    # No e-bike indicators
    return "Not E-Bike Related"



def style_excel(writer: pd.ExcelWriter, sheet_name: str, df: pd.DataFrame):
    """Apply header formatting and auto column widths to an Excel sheet."""
    ws = writer.sheets[sheet_name]
    header_fill = PatternFill("solid", fgColor="1F4E79")
    header_font = Font(color="FFFFFF", bold=True)

    for col_idx in range(1, len(df.columns) + 1):
        cell = ws.cell(row=1, column=col_idx)
        cell.fill = header_fill
        cell.font = header_font
        cell.alignment = Alignment(horizontal="center", wrap_text=True)

    for col_idx, col_name in enumerate(df.columns, start=1):
        max_len = max(
```

```
            len(str(col_name)),
            df.iloc[:, col_idx - 1].astype(str).str.len().max() if len(df) > 0 else 0
        )
        ws.column_dimensions[get_column_letter(col_idx)].width = min(max_len + 4, 60)



# 
==============================================================================
====
# MAIN ANALYSIS
# 
==============================================================================
====

def main():
    print("\n" + "=" * 65)
    print("  E-BIKE / MOPED / E-SCOOTER CRASH IDENTIFICATION & PRECISION")
    print("  (KEEPING ALL RECORDS WITH EVALUATION COLUMNS)")
    print("=" * 65)

    # -----------------------------------------------------------------
    # 1. Load data
    # -----------------------------------------------------------------
    print("\n[1] Loading data...")
    narratives = load_file(NARRATIVE_FILE, "Crash Narratives (bikecrashes.csv)", sheet=0)
    vehicles   = load_file(VEHICLE_FILE,   "Vehicle Data (bikeunits.csv)")
```

```
    # Normalize column names to lowercase to avoid case-sensitivity issues
    narratives.columns = narratives.columns.str.lower().str.strip()
    vehicles.columns   = vehicles.columns.str.lower().str.strip()

    # Validate required columns
    validate_columns(narratives, [NARRATIVE_CASE_ID_COL, NARRATIVE_TEXT_COL],
NARRATIVE_FILE)
    validate_columns(
        vehicles,
        [VEHICLE_CASE_ID_COL, VEHICLE_UNIT_DESC_COL, VEHICLE_BODY_STYLE_COL,
VEHICLE_ESCOOTER_COL],
        VEHICLE_FILE
    )

    # Filter to 2024
    if NARRATIVE_YEAR_COL and NARRATIVE_YEAR_COL in narratives.columns:
        before = len(narratives)
        narratives = narratives[narratives[NARRATIVE_YEAR_COL] == 2024].copy()
        print(f"  Filtered to year 2024: {before:,} -> {len(narratives):,} records")

    print(f"  Narrative records : {len(narratives):,}")
    print(f"  Vehicle records   : {len(vehicles):,}")

    # ---------------------------------------------------------------
    # 2. Keyword search on ALL investigator_narratives
    # ---------------------------------------------------------------
    print("\n[2] Searching ALL investigator_narratives for e-bike keywords...")
```

```
    narratives["matched_keywords"] =
narratives[NARRATIVE_TEXT_COL].apply(find_matched_keywords)
    narratives["ebike_keyword_flag"] =
narratives[NARRATIVE_TEXT_COL].apply(has_keyword_match)

    flagged = narratives[narratives["ebike_keyword_flag"]].copy()
    pct = len(flagged) / len(narratives) * 100 if len(narratives) > 0 else 0
    print(f"  Total cases      : {len(narratives):,}")
    print(f"  Keyword flagged   : {len(flagged):,}  ({pct:.2f}% of all narratives)")
    print(f"  Not flagged      : {len(narratives) - len(flagged):,}")

    # ------------------------------------------------------------------
    # 3. Merge ALL narratives with vehicle data
    #    Keep ALL vehicle units, then aggregate them per crash
    # ------------------------------------------------------------------
    print("\n[3] Matching ALL cases with vehicle records...")

    vehicles_subset = vehicles[
        [VEHICLE_CASE_ID_COL, VEHICLE_UNIT_DESC_COL, VEHICLE_BODY_STYLE_COL,
VEHICLE_ESCOOTER_COL]
    ].copy()

    merged = narratives.merge(
        vehicles_subset,
        left_on=NARRATIVE_CASE_ID_COL,
        right_on=VEHICLE_CASE_ID_COL,
        how="left",
```

```
  suffixes=("", "_veh")
)

merged["vehicle_classification"] = merged.apply(classify_vehicle, axis=1)

# Priority: Moped > E-Scooter > Pedalcyclist (keyword) > Other Unclassified >
#         Motorcycle (excluded) > Pedalcyclist (no keyword) > Not E-Bike > No Record
PRIORITY = {
  "Moped (unit_desc_id=1, veh_body_styl_id=92)"                            : 1,
  "E-Scooter (unit_desc_id=5, veh_body_styl_id=92 OR e_scooter_id=1)"          : 2,
  "Unclassified E-Bike - Pedalcyclist (unit_desc_id=3, keyword match)"          : 3,
  "Unclassified E-Bike - Other (keyword match, no matching vehicle code)"        : 4,
  "Motorcycle (excluded)"                                                 : 5,
  "Pedalcyclist (unit_desc_id=3, no keyword)"                               : 6,
  "Not E-Bike Related"                                                    : 7,
  "No Vehicle Record (keyword match)"                                       : 8,
  "No Vehicle Record"                                                     : 9,
}
merged["_priority"] = merged["vehicle_classification"].map(PRIORITY)

# Get the best classification per crash (for summary stats)
best_classification = (
  merged.sort_values("_priority")
     .drop_duplicates(subset=[NARRATIVE_CASE_ID_COL], keep="first")
     [[NARRATIVE_CASE_ID_COL, "vehicle_classification"]]
     .rename(columns={"vehicle_classification": "best_classification"})
```

```
)

# Aggregate all units per crash
all_units = (
    merged.groupby(NARRATIVE_CASE_ID_COL)
    .agg({
        VEHICLE_UNIT_DESC_COL: lambda x: ", ".join(map(str,
x.dropna().astype(int).unique())) if x.notna().any() else "No Record",
        VEHICLE_BODY_STYLE_COL: lambda x: ", ".join(map(str,
x.dropna().astype(int).unique())) if x.notna().any() else "No Record",
        VEHICLE_ESCOOTER_COL: lambda x: ", ".join(map(str,
x.dropna().astype(int).unique())) if x.notna().any() else "No Record",
        "vehicle_classification": lambda x: " | ".join(x.dropna().unique()),
        "_priority": "min"  # Keep the best priority
    })
    .reset_index()
    .rename(columns={
        VEHICLE_UNIT_DESC_COL: "all_unit_desc_ids",
        VEHICLE_BODY_STYLE_COL: "all_veh_body_styl_ids",
        VEHICLE_ESCOOTER_COL: "all_e_scooter_ids",
        "vehicle_classification": "all_vehicle_classifications"
    })
)

# Merge back with original narratives (ALL records)
final = narratives.merge(all_units, on=NARRATIVE_CASE_ID_COL, how="left")
final = final.merge(best_classification, on=NARRATIVE_CASE_ID_COL, how="left")
```

```
    # Fill NaN for cases with no vehicle records
    final["best_classification"] = final["best_classification"].fillna("No Vehicle Record")
    final["all_unit_desc_ids"] = final["all_unit_desc_ids"].fillna("No Record")
    final["all_veh_body_styl_ids"] = final["all_veh_body_styl_ids"].fillna("No Record")
    final["all_e_scooter_ids"] = final["all_e_scooter_ids"].fillna("No Record")
    final["all_vehicle_classifications"] = final["all_vehicle_classifications"].fillna("No Vehicle
Record")

    # Add a simplified flag for filtering
    final["is_potential_ebike"] = final["ebike_keyword_flag"] | final["best_classification"].isin([
        "Moped (unit_desc_id=1, veh_body_styl_id=92)",
        "E-Scooter (unit_desc_id=5, veh_body_styl_id=92 OR e_scooter_id=1)",
        "Unclassified E-Bike - Pedalcyclist (unit_desc_id=3, keyword match)",
        "Unclassified E-Bike - Other (keyword match, no matching vehicle code)"
    ])

    # Drop internal priority column
    final = final.drop(columns=["_priority"], errors="ignore").reset_index(drop=True)

    # Calculate statistics using keyword-flagged cases only
    keyword_flagged = final[final["ebike_keyword_flag"]]
    unmatched     = keyword_flagged["best_classification"].str.contains("No Vehicle Record",
na=False).sum()
    motorcycle_count = keyword_flagged["best_classification"].eq("Motorcycle
(excluded)").sum()
    matched       = len(keyword_flagged) - unmatched - motorcycle_count
```

```
    moped_count    = keyword_flagged["best_classification"].str.contains("Moped",
na=False).sum()
    escooter_count = keyword_flagged["best_classification"].str.contains("E-Scooter",
na=False).sum()
    pedalcyclist_keyword_count =
keyword_flagged["best_classification"].str.contains("Pedalcyclist.*keyword",
na=False).sum()
    unclass_other_count  =
keyword_flagged["best_classification"].str.contains("Other.*keyword", na=False).sum()


    print(f"  All cases processed        : {len(final):,}")
    print(f"  Keyword-flagged cases      : {len(keyword_flagged):,}")
    print(f"  Matched with vehicle record : {matched:,}")
    print(f"  Motorcycles (excluded)     : {motorcycle_count:,}")
    print(f"  No vehicle record found    : {unmatched:,}")


    # ----------------------------------------------------------------
    # 4. Precision calculation (based on keyword-flagged cases)
    # ----------------------------------------------------------------
    print("\n[4] Calculating precision rates (keyword-flagged cases only)...")


    correctly_classified = moped_count + escooter_count
    precision = (correctly_classified / matched * 100) if matched > 0 else 0


    print(f"\n  +------------------------------------------------+")
    print(f"  |  PRECISION RATE SUMMARY (KEYWORD-FLAGGED)      |")
    print(f"  +------------------------------------------------+")
```

```
print(f"  | Total narratives reviewed (2024) : {len(narratives):>6,}      |")
print(f"  | Narratives flagged (keyword)     : {len(keyword_flagged):>6,}      |")
print(f"  | Matched to vehicle record        : {matched:>6,}      |")
print(f"  | Motorcycles (excluded)           : {motorcycle_count:>6,}      |")
print(f"  +-----------------------------------------------+")
print(f"  | Correctly classified             : {correctly_classified:>6,}      |")
print(f"  |   -> Moped                       : {moped_count:>6,}      |")
print(f"  |   -> E-Scooter                   : {escooter_count:>6,}      |")
print(f"  | Unclassified - Pedalcyclist      : {pedalcyclist_keyword_count:>6,}      |")
print(f"  | Unclassified - Other             : {unclass_other_count:>6,}      |")
print(f"  | No vehicle record found          : {unmatched:>6,}      |")
print(f"  +-----------------------------------------------+")
print(f"  | Precision rate                   : {precision:>6.1f}%     |")
print(f"  +-----------------------------------------------+")

# ---------------------------------------------------------------
# 5. Build summary tables
# ---------------------------------------------------------------

precision_summary = pd.DataFrame({
    "Metric": [
        "Total narratives reviewed (2024)",
        "Narratives flagged by keyword",
        "Flagged rate (%)",
        "Matched to vehicle record",
        "Motorcycles (excluded)",
```

```
        "Correctly classified — Moped",
        "Correctly classified — E-Scooter",
        "Total correctly classified",
        "Unclassified — Pedalcyclist (unit_desc_id=3)",
        "Unclassified — Other (no vehicle code)",
        "No vehicle record found",
        "Precision rate (%)",
    ],
    "Value": [
        len(narratives),
        len(keyword_flagged),
        round(pct, 2),
        matched,
        motorcycle_count,
        moped_count,
        escooter_count,
        correctly_classified,
        pedalcyclist_keyword_count,
        unclass_other_count,
        unmatched,
        round(precision, 2),
    ]
})

classification_breakdown = (
    final.groupby("best_classification", dropna=False)
```

```
        .size()
        .reset_index(name="count")
        .sort_values("count", ascending=False)
)
classification_breakdown["pct_of_all"] = (
    classification_breakdown["count"] / len(final) * 100
).round(2)

# Keyword flagged cases breakdown
keyword_classification_breakdown = (
    keyword_flagged.groupby("best_classification", dropna=False)
        .size()
        .reset_index(name="count")
        .sort_values("count", ascending=False)
)
keyword_classification_breakdown["pct_of_flagged"] = (
    keyword_classification_breakdown["count"] / len(keyword_flagged) * 100
).round(2)

keyword_breakdown = (
    final[final["matched_keywords"] != ""]["matched_keywords"]
    .str.split(", ")
    .explode()
    .value_counts()
    .reset_index()
)
```

```
    keyword_breakdown.columns = ["keyword", "count"]

    # All unit_desc_id x veh_body_styl_id combinations seen
    unit_body_breakdown = (
        merged.groupby(
            [VEHICLE_UNIT_DESC_COL, VEHICLE_BODY_STYLE_COL,
VEHICLE_ESCOOTER_COL], dropna=False
        )
        .size()
        .reset_index(name="count")
        .sort_values("count", ascending=False)
    )

    # -----------------------------------------------------------------
    # 6. Save outputs
    # -----------------------------------------------------------------
    print("\n[5] Saving outputs...")

    # Reorder columns for better readability
    key_cols = [
        NARRATIVE_CASE_ID_COL,
        "ebike_keyword_flag",
        "is_potential_ebike",
        "matched_keywords",
        "best_classification",
        "all_unit_desc_ids",
```

```
    "all_veh_body_styl_ids",
    "all_e_scooter_ids",
    "all_vehicle_classifications"
]
other_cols = [col for col in final.columns if col not in key_cols]
final_reordered = final[key_cols + other_cols]

final_reordered.to_csv(OUTPUT_CSV, index=False)
print(f"  Complete dataset saved (CSV) -> {OUTPUT_CSV}")
print(f"    Total records: {len(final_reordered):,}")
print(f"    Potential e-bike cases: {final_reordered['is_potential_ebike'].sum():,}")

with pd.ExcelWriter(OUTPUT_EXCEL, engine="openpyxl") as writer:

    precision_summary.to_excel(writer, sheet_name="Precision Summary", index=False)
    style_excel(writer, "Precision Summary", precision_summary)

    # All cases
    final_reordered.to_excel(writer, sheet_name="All Cases", index=False)
    style_excel(writer, "All Cases", final_reordered)

    # Potential e-bike cases only (for easier review)
    potential_ebike = final_reordered[final_reordered["is_potential_ebike"]]
    potential_ebike.to_excel(writer, sheet_name="Potential E-Bike Cases", index=False)
    style_excel(writer, "Potential E-Bike Cases", potential_ebike)
```

```
    # Classification breakdowns
    classification_breakdown.to_excel(writer, sheet_name="All Cases - Classification",
index=False)
    style_excel(writer, "All Cases - Classification", classification_breakdown)

    keyword_classification_breakdown.to_excel(writer, sheet_name="Keyword Flagged -
Classification", index=False)
    style_excel(writer, "Keyword Flagged - Classification",
keyword_classification_breakdown)

    unit_body_breakdown.to_excel(writer, sheet_name="Unit x Body x E-Scooter",
index=False)
    style_excel(writer, "Unit x Body x E-Scooter", unit_body_breakdown)

    keyword_breakdown.to_excel(writer, sheet_name="Keyword Frequency", index=False)
    style_excel(writer, "Keyword Frequency", keyword_breakdown)

  print(f"  Excel report saved        -> {OUTPUT_EXCEL}")
  print("\n  Sheets included:")
  print("    * Precision Summary          -- key metrics for keyword-flagged cases")
  print("    * All Cases                  -- COMPLETE dataset with evaluation columns")
  print("    * Potential E-Bike Cases     -- filtered view of likely e-bike crashes")
  print("    * All Cases - Classification   -- classification breakdown for all records")
  print("    * Keyword Flagged - Classification -- classification breakdown for keyword-
flagged only")
  print("    * Unit x Body x E-Scooter     -- all unit/body/escooter combinations")
  print("    * Keyword Frequency          -- which keywords triggered the most matches")
```

```
    print("\n[OK] Analysis complete!\n")
    print(f"  Use 'ebike_keyword_flag' column to filter keyword-matched cases")
    print(f"  Use 'is_potential_ebike' column to filter all potential e-bike cases")
    print(f"  Use 'best_classification' column to see vehicle type determination")


if __name__ == "__main__":
    main()
```

Appendix B. Non-E-Bike and E-Bike Crashes by County

County	Non-E-Bike Crashes	E-Bike Crashes
Harris	5264 (21%)	47 (12%)
Bexar	2802 (11%)	43 (11%)
Travis	2304 (9%)	43 (11%)
Dallas	2009 (8%)	29 (7%)
Tarrant	1499 (6%)	28 (7%)
Collin	730 (3%)	12 (3%)
Denton	687 (3%)	16 (4%)
Hidalgo	632 (3%)	2 (<1%)
El Paso	598 (2%)	10 (3%)
Galveston	596 (2%)	5 (1%)
Nueces	558 (2%)	13 (3%)
Fort Bend	547 (2%)	13 (3%)
Cameron	451 (2%)	4 (1%)
Montgomery	436 (2%)	10 (3%)
Williamson	417 (2%)	10 (3%)
Brazos	416 (2%)	15 (4%)
Jefferson	382 (2%)	4 (1%)
Lubbock	340 (1%)	8 (2%)
Brazoria	264 (1%)	6 (2%)
Bell	253 (1%)	5 (1%)
Webb	224 (<1%)	2 (<1%)
Mclennan	219 (<1%)	1 (<1%)
Hays	161 (<1%)	7 (2%)
Potter	135 (<1%)	3 (<1%)
Taylor	131 (<1%)	1 (<1%)
Wichita	124 (<1%)	2 (<1%)
Tom Green	103 (<1%)	0 (0%)
Smith	101 (<1%)	3 (<1%)
Gregg	93 (<1%)	0 (0%)
Comal	91 (<1%)	3 (<1%)
Midland	85 (<1%)	2 (<1%)
Orange	82 (<1%)	1 (<1%)
Guadalupe	81 (<1%)	2 (<1%)
Johnson	77 (<1%)	2 (<1%)
Ellis	75 (<1%)	1 (<1%)
Ector	73 (<1%)	1 (<1%)
Bowie	68 (<1%)	2 (<1%)
Randall	67 (<1%)	1 (<1%)
Angelina	58 (<1%)	0 (0%)

County	Non-E-Bike Crashes	E-Bike Crashes
Grayson	58 (<1%)	2 (<1%)
Victoria	58 (<1%)	1 (<1%)
Kerr	52 (<1%)	2 (<1%)
San Patricio	51 (<1%)	1 (<1%)
Kaufman	46 (<1%)	0 (0%)
Walker	46 (<1%)	2 (<1%)
Val Verde	43 (<1%)	1 (<1%)
Parker	42 (<1%)	2 (<1%)
Matagorda	40 (<1%)	0 (0%)
Nacogdoches	40 (<1%)	0 (0%)
Waller	39 (<1%)	0 (0%)
Coryell	38 (<1%)	4 (1%)
Liberty	38 (<1%)	2 (<1%)
Maverick	35 (<1%)	1 (<1%)
Aransas	33 (<1%)	0 (0%)
Hunt	32 (<1%)	0 (0%)
Lamar	31 (<1%)	4 (1%)
Bastrop	30 (<1%)	0 (0%)
Caldwell	30 (<1%)	0 (0%)
Chambers	28 (<1%)	1 (<1%)
Hardin	28 (<1%)	0 (0%)
Rockwall	26 (<1%)	1 (<1%)
Jim Wells	25 (<1%)	0 (0%)
Kleberg	24 (<1%)	0 (0%)
Navarro	23 (<1%)	0 (0%)
Burnet	22 (<1%)	1 (<1%)
Kendall	22 (<1%)	0 (0%)
Palo Pinto	22 (<1%)	1 (<1%)
Wharton	20 (<1%)	0 (0%)
Atascosa	19 (<1%)	0 (0%)
Harrison	19 (<1%)	0 (0%)
Henderson	19 (<1%)	0 (0%)
Medina	18 (<1%)	0 (0%)
Polk	18 (<1%)	0 (0%)
Lampasas	17 (<1%)	0 (0%)
Starr	17 (<1%)	0 (0%)
Gillespie	16 (<1%)	0 (0%)
Gray	16 (<1%)	1 (<1%)
Hood	16 (<1%)	1 (<1%)
Rusk	16 (<1%)	1 (<1%)

County	Non-E-Bike Crashes	E-Bike Crashes
Washington	15 (<1%)	0 (0%)
Brown	14 (<1%)	0 (0%)
Nolan	14 (<1%)	0 (0%)
Titus	14 (<1%)	0 (0%)
Wood	14 (<1%)	0 (0%)
Bee	13 (<1%)	0 (0%)
Cherokee	13 (<1%)	0 (0%)
Fannin	13 (<1%)	0 (0%)
Uvalde	13 (<1%)	0 (0%)
Colorado	12 (<1%)	0 (0%)
Erath	12 (<1%)	0 (0%)
Hill	12 (<1%)	0 (0%)
Jasper	12 (<1%)	0 (0%)
Panola	12 (<1%)	0 (0%)
Pecos	12 (<1%)	0 (0%)
Anderson	11 (<1%)	1 (<1%)
Calhoun	11 (<1%)	0 (0%)
Cooke	11 (<1%)	1 (<1%)
Milam	11 (<1%)	0 (0%)
Moore	11 (<1%)	0 (0%)
San Jacinto	11 (<1%)	0 (0%)
Upshur	11 (<1%)	0 (0%)
Austin	10 (<1%)	0 (0%)
Hale	10 (<1%)	0 (0%)
Hockley	10 (<1%)	0 (0%)
Howard	10 (<1%)	0 (0%)
Wise	9 (<1%)	1 (<1%)
Andrews	8 (<1%)	0 (0%)
Dewitt	8 (<1%)	0 (0%)
Hutchinson	8 (<1%)	1 (<1%)
Morris	8 (<1%)	0 (0%)
Scurry	8 (<1%)	0 (0%)
Willacy	8 (<1%)	0 (0%)
Wilson	8 (<1%)	0 (0%)
Cass	7 (<1%)	0 (0%)
Hopkins	7 (<1%)	0 (0%)
Mcculloch	7 (<1%)	0 (0%)
Young	7 (<1%)	0 (0%)
Fayette	6 (<1%)	0 (0%)
Robertson	6 (<1%)	0 (0%)

County	Non-E-Bike Crashes	E-Bike Crashes
Stephens	6 (<1%)	0 (0%)
Trinity	6 (<1%)	0 (0%)
Dallam	5 (<1%)	0 (0%)
Grimes	5 (<1%)	0 (0%)
Karnes	5 (<1%)	0 (0%)
Lee	5 (<1%)	1 (<1%)
Montague	5 (<1%)	0 (0%)
Refugio	5 (<1%)	0 (0%)
Archer	4 (<1%)	0 (0%)
Burleson	4 (<1%)	1 (<1%)
Culberson	4 (<1%)	0 (0%)
Deaf Smith	4 (<1%)	0 (0%)
Eastland	4 (<1%)	0 (0%)
Frio	4 (<1%)	0 (0%)
Houston	4 (<1%)	0 (0%)
Lavaca	4 (<1%)	0 (0%)
Leon	4 (<1%)	0 (0%)
Limestone	4 (<1%)	0 (0%)
Llano	4 (<1%)	0 (0%)
Ochiltree	4 (<1%)	0 (0%)
Shelby	4 (<1%)	0 (0%)
Somervell	4 (<1%)	0 (0%)
Van Zandt	4 (<1%)	0 (0%)
Bailey	3 (<1%)	0 (0%)
Blanco	3 (<1%)	0 (0%)
Bosque	3 (<1%)	0 (0%)
Brooks	3 (<1%)	0 (0%)
Coleman	3 (<1%)	0 (0%)
Crosby	3 (<1%)	0 (0%)
Dawson	3 (<1%)	0 (0%)
Gonzales	3 (<1%)	0 (0%)
Rains	3 (<1%)	0 (0%)
San Augustine	3 (<1%)	0 (0%)
Yoakum	3 (<1%)	0 (0%)
Zapata	3 (<1%)	0 (0%)
Brewster	2 (<1%)	1 (<1%)
Camp	2 (<1%)	0 (0%)
Carson	2 (<1%)	0 (0%)
Childress	2 (<1%)	0 (0%)
Clay	2 (<1%)	0 (0%)

County	Non-E-Bike Crashes	E-Bike Crashes
Comanche	2 (<1%)	0 (0%)
Delta	2 (<1%)	0 (0%)
Dimmit	2 (<1%)	0 (0%)
Falls	2 (<1%)	0 (0%)
Freestone	2 (<1%)	0 (0%)
Gaines	2 (<1%)	0 (0%)
Garza	2 (<1%)	0 (0%)
Hartley	2 (<1%)	0 (0%)
Kimble	2 (<1%)	0 (0%)
Lamb	2 (<1%)	0 (0%)
Live Oak	2 (<1%)	0 (0%)
Madison	2 (<1%)	0 (0%)
Mills	2 (<1%)	0 (0%)
Parmer	2 (<1%)	0 (0%)
Real	2 (<1%)	0 (0%)
Red River	2 (<1%)	0 (0%)
Reeves	2 (<1%)	0 (0%)
Terry	2 (<1%)	0 (0%)
Ward	2 (<1%)	0 (0%)
Winkler	2 (<1%)	0 (0%)
Bandera	1 (<1%)	0 (0%)
Callahan	1 (<1%)	0 (0%)
Collingsworth	1 (<1%)	0 (0%)
Concho	1 (<1%)	0 (0%)
Crockett	1 (<1%)	0 (0%)
Duval	1 (<1%)	0 (0%)
Fisher	1 (<1%)	0 (0%)
Floyd	1 (<1%)	0 (0%)
Franklin	1 (<1%)	0 (0%)
Hamilton	1 (<1%)	0 (0%)
Hardeman	1 (<1%)	0 (0%)
Haskell	1 (<1%)	0 (0%)
Hudspeth	1 (<1%)	0 (0%)
Irion	1 (<1%)	0 (0%)
Lynn	1 (<1%)	0 (0%)
Marion	1 (<1%)	0 (0%)
Martin	1 (<1%)	0 (0%)
Mason	1 (<1%)	0 (0%)
Mitchell	1 (<1%)	0 (0%)
Newton	1 (<1%)	0 (0%)

County	Non-E-Bike Crashes	E-Bike Crashes
Reagan	1 (<1%)	0 (0%)
Sabine	1 (<1%)	0 (0%)
San Saba	1 (<1%)	0 (0%)
Schleicher	1 (<1%)	0 (0%)
Sterling	1 (<1%)	0 (0%)
Swisher	1 (<1%)	0 (0%)
Tyler	1 (<1%)	0 (0%)
Upton	1 (<1%)	0 (0%)
Wilbarger	1 (<1%)	0 (0%)
Zavala	1 (<1%)	0 (0%)
Total	25081 (100%)	393 (100%)

Appendix C: PBCAT Code in TxDOT CRIS Database

Pedalcyclist Action Codes

ID	Description
1	IN PARKING LOT
2	NOT IN ROADWAY
3	OTHER IN ROADWAY (EX. STANDING OR PLAYING)
4	OTHER RIDING WHILE DISTRACTED (TALKING, EATING, ELECTRONIC DEVICE, ETC.)
5	RIDING CYCLE WHILE CROSSING ROAD AT DRIVEWAY
6	RIDING CYCLE WHILE CROSSING ROAD AT INTERSECTION OR CROSSWALK (MARKED OR UNMARKED)
7	RIDING CYCLE WHILE CROSSING ROAD NOT AT INTERSECTION OR CROSSWALK
8	RIDING IN BIKE LANE - AGAINST TRAFFIC
9	RIDING IN BIKE LANE - WITH TRAFFIC
10	RIDING IN ROADWAY (VEHICLE LANE - UNKNOWN IF BIKE MARKINGS PRESENT) - AGAINST TRAFFIC
11	RIDING IN ROADWAY (VEHICLE LANE - UNKNOWN IF BIKE MARKINGS PRESENT) - WITH TRAFFIC
12	RIDING IN ROADWAY (VEHICLE LANE WITH BIKE MARKINGS IN THE LANE) - AGAINST TRAFFIC
13	RIDING IN ROADWAY (VEHICLE LANE WITH BIKE MARKINGS IN THE LANE) - WITH TRAFFIC
14	RIDING IN ROADWAY (VEHICLE LANE WITHOUT BIKE MARKINGS) - AGAINST TRAFFIC
15	RIDING IN ROADWAY (VEHICLE LANE WITHOUT BIKE MARKINGS) - WITH TRAFFIC
16	RIDING IN SHOULDER - AGAINST TRAFFIC
17	RIDING IN SHOULDER - WITH TRAFFIC
18	RIDING ON SIDEWALK
19	WORKING IN ROADWAY OR SHOULDER (INCIDENT RESPONSE)
95	DATA NOT AVAILABLE AT DATA ENTRY
97	NOT APPLICABLE
99	UNKNOWN

Pedalcyclist Identification

ID	Description
1	BACKING VEHICLE
2	BICYCLIST FAILED TO CLEAR - MULTIPLE THREAT
3	BICYCLIST FAILED TO CLEAR - TRAPPED
4	BICYCLIST FAILED TO CLEAR - UNKNOWN
5	BICYCLIST INTENTIONALLY CAUSED
6	BICYCLIST LEFT TURN - OPPOSITE DIRECTION
7	BICYCLIST LEFT TURN - SAME DIRECTION
8	BICYCLIST LOST CONTROL - MECHANICAL PROBLEMS
9	BICYCLIST LOST CONTROL - OTHER/UNKNOWN
10	BICYCLIST LOST CONTROL - OVERSTEERING, IMPROPER BRAKING, SPEED
11	BICYCLIST LOST CONTROL - SURFACE CONDITIONS
12	BICYCLIST OVERTAKING - EXTENDED DOOR
13	BICYCLIST OVERTAKING - OTHER/UNKNOWN
15	BICYCLIST OVERTAKING - PASSING ON LEFT

16 BICYCLIST OVERTAKING - PASSING ON RIGHT
17 BICYCLIST RIDE OUT - COMMERCIAL DRIVEWAY/ALLEY
18 BICYCLIST RIDE OUT - MIDBLOCK - UNKNOWN
19 BICYCLIST RIDE OUT - OTHER MIDBLOCK
20 BICYCLIST RIDE OUT - PARALLEL PATH
21 BICYCLIST RIDE OUT - RESIDENTIAL DRIVEWAY
22 BICYCLIST RIDE OUT - SIGN-CONTROLLED INTERSECTION
23 BICYCLIST RIDE OUT - SIGNALIZED INTERSECTION
24 BICYCLIST RIDE THROUGH - SIGN-CONTROLLED INTERSECTION
25 BICYCLIST RIDE THROUGH - SIGNALIZED INTERSECTION
26 BICYCLIST RIGHT TURN - OPPOSITE DIRECTION
27 BICYCLIST RIGHT TURN - SAME DIRECTION
28 BICYCLIST TURNING ERROR - LEFT TURN
29 BICYCLIST TURNING ERROR - OTHER
30 BICYCLIST TURNING ERROR - RIGHT TURN
31 BICYCLIST LOST CONTROL - ALCOHOL/DRUG IMPAIRMENT
32 BUS/DELIVERY VEHICLE PULLOVER
33 CROSSING PATHS - INTERSECTION - OTHER/UNKNOWN
34 CROSSING PATHS - MIDBLOCK - OTHER/UNKNOWN
35 CROSSING PATHS - UNCONTROLLED INTERSECTION
36 HEAD-ON - BICYCLIST
37 HEAD-ON - MOTORIST
38 HEAD-ON - UNKNOWN
39 MOTORIST DRIVE IN/OUT PARKING
40 MOTORIST DRIVE OUT - COMMERCIAL DRIVEWAY/ALLEY
41 MOTORIST DRIVE OUT - MIDBLOCK - UNKNOWN
42 MOTORIST DRIVE OUT - OTHER MIDBLOCK
43 MOTORIST DRIVE OUT - RESIDENTIAL DRIVEWAY
44 MOTORIST DRIVE OUT - RIGHT TURN ON RED
45 MOTORIST DRIVE OUT - SIGN-CONTROLLED INTERSECTION
46 MOTORIST DRIVE OUT - SIGNALIZED INTERSECTION
47 MOTORIST DRIVE THROUGH - SIGN-CONTROLLED INTERSECTION
48 MOTORIST DRIVE THROUGH - SIGNALIZED INTERSECTION
49 MOTORIST INTENTIONALLY CAUSED
50 MOTORIST LEFT TURN - OPPOSITE DIRECTION
51 MOTORIST LEFT TURN - SAME DIRECTION
52 MOTORIST LOST CONTROL - ALCOHOL/DRUG IMPAIRMENT
53 MOTORIST LOST CONTROL - MECHANICAL PROBLEMS
54 MOTORIST LOST CONTROL - OTHER/UNKNOWN
55 MOTORIST LOST CONTROL - OVERSTEERING, IMPROPER BRAKING, SPEED
56 MOTORIST LOST CONTROL - SURFACE CONDITIONS
57 MOTORIST OVERTAKING - BICYCLIST SWERVED
58 MOTORIST OVERTAKING - MISJUDGED SPACE
59 MOTORIST OVERTAKING - OTHER/UNKNOWN
60 MOTORIST OVERTAKING - UNDETECTED BICYCLIST

61 MOTORIST RIGHT TURN - OPPOSITE DIRECTION
62 MOTORIST RIGHT TURN - SAME DIRECTION
63 MOTORIST RIGHT TURN ON RED - OPPOSITE DIRECTION
64 MOTORIST RIGHT TURN ON RED - SAME DIRECTION
65 MOTORIST TURN/MERGE - OTHER/UNKNOWN
66 MOTORIST TURNING ERROR - LEFT TURN
67 MOTORIST TURNING ERROR - OTHER
68 MOTORIST TURNING ERROR - RIGHT TURN
69 MULTIPLE THREAT - MIDBLOCK
70 MULTIPLE THREAT - SIGN-CONTROLLED INTERSECTION
71 NON-ROADWAY
72 PARALLEL PATHS - OTHER/UNKNOWN
73 SIGN-CONTROLLED INTERSECTION - OTHER/UNKNOWN
74 SIGNALIZED INTERSECTION - OTHER/UNKNOWN
75 UNKNOWN APPROACH PATHS
76 UNKNOWN LOCATION
77 UNUSUAL CIRCUMSTANCES
95 DATA NOT AVAILABLE AT DATA ENTRY
97 NOT APPLICABLE