Table II: Overview of Scenario Components and Elements' Hierarchy along with their Availability in $D1$ (US) and $D2$ (EU)

Scenario Comp.	Elements	Sub-elements	Sub-sub elements	Details	Datasets	
					$D1$	$D2$
Static Env.	Road Layout	Junction	-	-	×	✓
	Static Object	-		-	✓	×
Dynamic Env.	Actors	Vehicles	Passenger cars	Coupe,Wagon, etc	✓	✓
			Light Trucks	Utility, Light Pickup, etc	✓	×
			Large Trucks	Cargo Van, Large Pickup, etc	✓	×
			Buses	School Bus,Transit Bus, etc	✓	✓
			Heavy Good Vehicles	Trams, etc	×	✓
			Other/Unknown Vehicles	Limousine,Terrain, etc	✓	✓
		Pedestrians	-	-	✓	✓
		Motorcyclists	2-Wheel Motorcycle		✓	✓
			Moped		✓	✓
			3-Wheel Motorcycle		✓	×
			Unenclosed 3-Wheel Motorcycle		✓	×
			Scooter		✓	×
			Other		✓	×
			Unknown		✓	×
		Pedalcyclists	-	-	✓	✓
		Unknown	-	-	✓	×
	Activity	Vehicle Lat. Activity	Turning Left	-	✓	✓
			Turning Right	-	✓	✓
			Negotiating a Curve	-	✓	✓
			Backing Up	-	✓	✓
			U-turn	-	✓	✓
			Merge/Lane Change	-	✓	✓
		Vehicle Long. Activity	Driving Straight	-	✓	✓
			Decelerating in Road	-	✓	×
			Overtaking	-	✓	×
			Starting in Road	-	✓	×
			Leaving a Parking Position	-	✓	×
			Stopped in Roadway	-	✓	×
			Parked in Travel Lane	-	✓	×
			Other Maneuver	-	✓	×
			Unknown	-	✓	×
Conditions	Weather	Dry	-	-	✓	✓
		Rain	-	-	✓	✓
		Snow/Sleet	-	-	✓	✓
		Fog/Mist	-	-	×	✓
		Strong Wind	-	-	×	✓
		Other	-	-	✓	✓
		Unknown	-	-	✓	✓
	Light	Dark	-	-	✓	✓
		Daylight	-	-	✓	✓
		Dark, but Lighted	-	-	✓	✓
		Dawn/Twilight	-	-	✓	✓
		Street Lights	Light lit	-	×	✓
			Light unlit	-	×	✓
			No lights	-	×	✓
			None/unlit	-	×	✓
		Other		-	✓	×
		Unknown		-	✓	×

(Comp.= Component, Env.=Environment)

Comparison D1 vs. D2:

We compared both datasets by identifying the scenario components and element hierarchies. Our goal of this comparison was to see if we could find fatality statistics for scenario components and elements, as shown in Figures A to E.

We also marked whether the elements and sub-elements were found in one or both datasets, as shown in Table II. For example, in the scenario element "Actors", the sub-elements are 'Vehicles', 'Pedestrians', and 'Pedalcyclists', etc. Under 'Vehicles', further sub-elements are 'Passenger cars', 'Buses', etc. The data for most of these elements are available in both datasets. Additionally, we observed differences in how certain elements were grouped. For example, 'Passenger Cars' in the $D1$ dataset only included cars, but in $D2$, it also included vans and pickups, which were separately grouped as 'Light trucks' in the $D1$ dataset. We looked at the distribution of sub-elements and sub-sub-elements to understand variations in the two datasets.

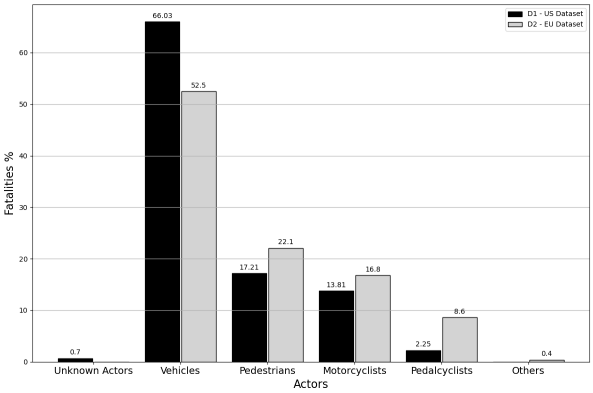

Figure A: Actors

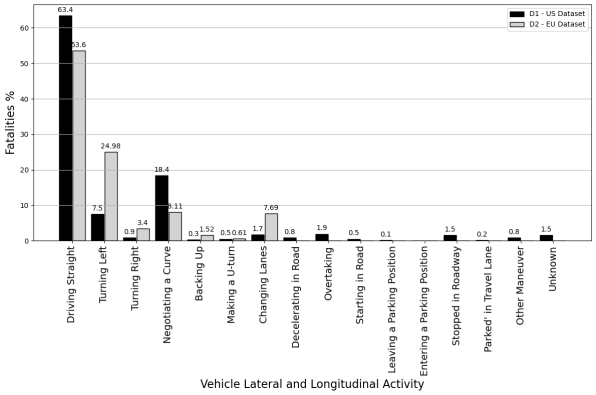

Figure B: Vehicle Lateral and Longitudinal Activity

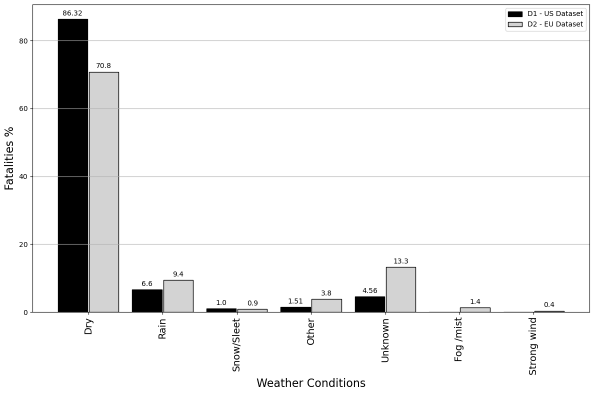

Figure C: Weather Conditions

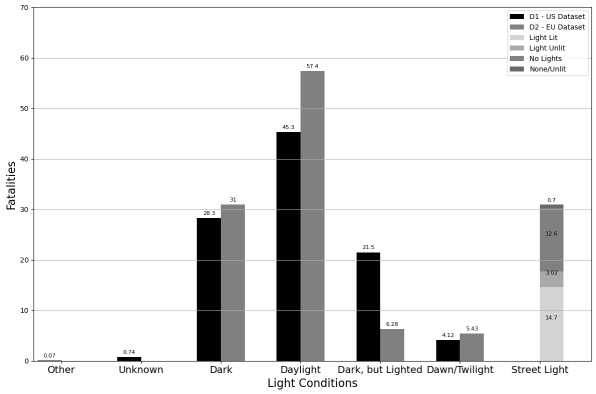

Figure D: Light Conditions

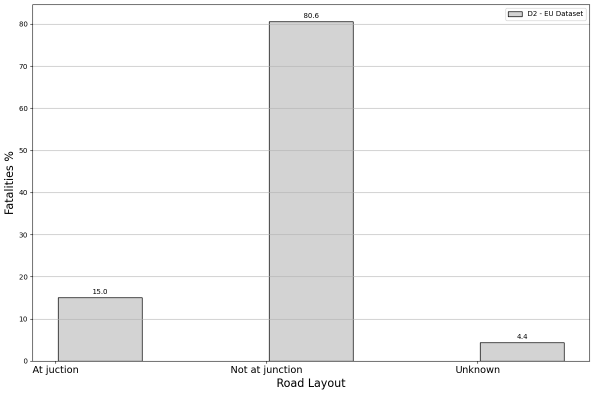

Figure E: Road Layout

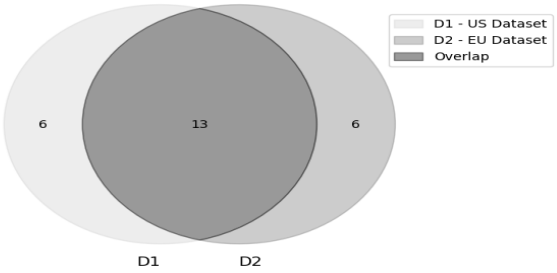

Figure F: Overlap between the number of sub-elements in $D1$ (US) vs. $D2$ (EU)

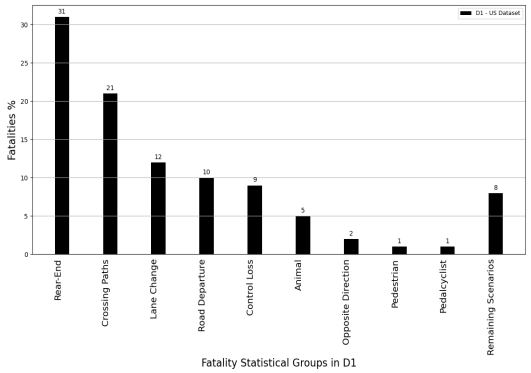

Figure G: Number of fatalities per scenario group in $D1$

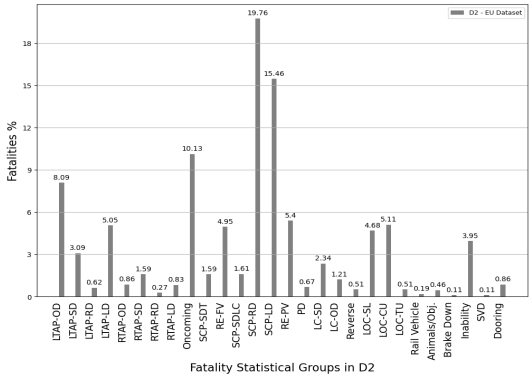

Figure H: Number of fatalities per scenario group in $D2$

Figure A shows the distribution of fatality (in percentage) associated with "Actors" in traffic incidents from both datasets. The x-axis shows the sub-elements of the element "Actor", such as 'Vehicles',' Pedestrians', and 'Motorcyclists', 'Pedalcyclists', 'Others' and 'Unknown'. The y-axis represents the cumulative fatality percentage attributed to each actor sub-element and its sub-sub elements. For example, The fatality counts shown for 'Vehicles' represent the cumulative sum of all its sub-sub elements such as (large trucks, buses, passenger cars, etc.). In $D1$ and $D2$, the most pronounced fatality rate is observed in incidents involving vehicles, accounting for 66.03% and 52.5% of the fatalities. Pedestrians and motorcyclists also represent significant proportions of fatalities. In $D1$, pedestrians account for approximately 17.21% of fatalities, which is lower than the percentage in $D2$ (22.1%). However, a reverse trend is observed for motorcyclists, with $D1$ showing about 13.81% fatalities compared to $D2$'s (16.8%). Pedalcyclists represent a higher fatality rate in $D2$ (8.6%) compared to $D1$ (2.25%). In $D1$, the fatality rate for unknown actors is 0.7, while in $D2$, it is 0, representing missing data. In $D2$, the fatality rate for unknown actors is 0.4, with 0 representing missing data in $D1$ for other actors.

Figure B shows the distribution of fatalities (in percentage) of sub-sub element 'Vehicle Lateral activity' and 'Vehicle Longitudinal activity' in traffic incidents from both datasets. The 'Vehicle Lateral activity' and 'Vehicle Longitudinal activity' are driving maneuvers. In both datasets, most accidents occur in the 'Going straight' maneuver. Figure C and Figure D show the distribution of sub-elements of "Weather" and "Light" conditions in both datasets. In both datasets, accidents mainly occur in dry and daylight conditions. Figure E shows the distribution of sub-elements of Road Layout. There is no data available related to the sub-elements of "Road Layout" in $D1$. While in $D2$, 15% of accidents occur at junctions, and 80.6% of accidents occur when not at a junction. Additionally, a small percentage of accidents in $D2$ (4.4%) have an unknown road layout. Figure F concisely shows the overlap between the number of sub-elements in $D1$ and $D2$. D1 consists of 19 sub-elements, and $D2$ consists of 19 sub-elements, with 13 sub-elements common to both datasets. However, for sub-sub elements, $D1$ has more data than $D2$.

Furthermore, each dataset contains scenario groups that are based on similar sub-elements and sub-sub-elements. These scenario groups show the relative number of fatalities as a percentage. In $D1$, there are nine scenario groups: Rear-End, Crossing Path, Lane Change, Road Departure, Animal, Opposite Direction, Pedestrian, Pedalcyclist, and Remaining Scenarios. In $D2$, there are 28 scenario groups. The distribution of fatalities across each scenario group is shown in Figure G and Figure H.