

Data Analysis

The data used to solve this problem is the collision data collected from NYC OpenData. An explanation and discussion with examples of the data is following. Data is a single dataframe containing the Location, weather, light condition, road condition, junction type, no of people involved, no of vehicles involved etc.

SEVERITYCODE	LOCATION	WEATHER	SPEEDING	LIGHTCOND	ROADCOND	JUNCTIONTYPE	PERSONCOUNT	VEHCOUNT
2	RAINIER AVE S AND S BRANDON ST	Clear	Y	Daylight	Dry	At Intersection (intersection related)	3	2
2	OLSON PL SW BETWEEN 2ND AVE SW AND 3RD AVE SW	Raining	Y	Dark - Street Lights On	Wet	Mid-Block (not related to intersection)	1	1
1	35TH AVE SW BETWEEN 37TH AVE SW AND MARINE VIE...	Clear	Y	Daylight	Dry	Mid-Block (but intersection related)	4	4
1	MARION ST BETWEEN 2ND AVE AND 3RD AVE	Raining	Y	Daylight	Wet	Mid-Block (not related to intersection)	2	2
2	HARVARD AVE AND E DENNY WAY	Clear	Y	Daylight	Dry	At Intersection (intersection related)	2	2