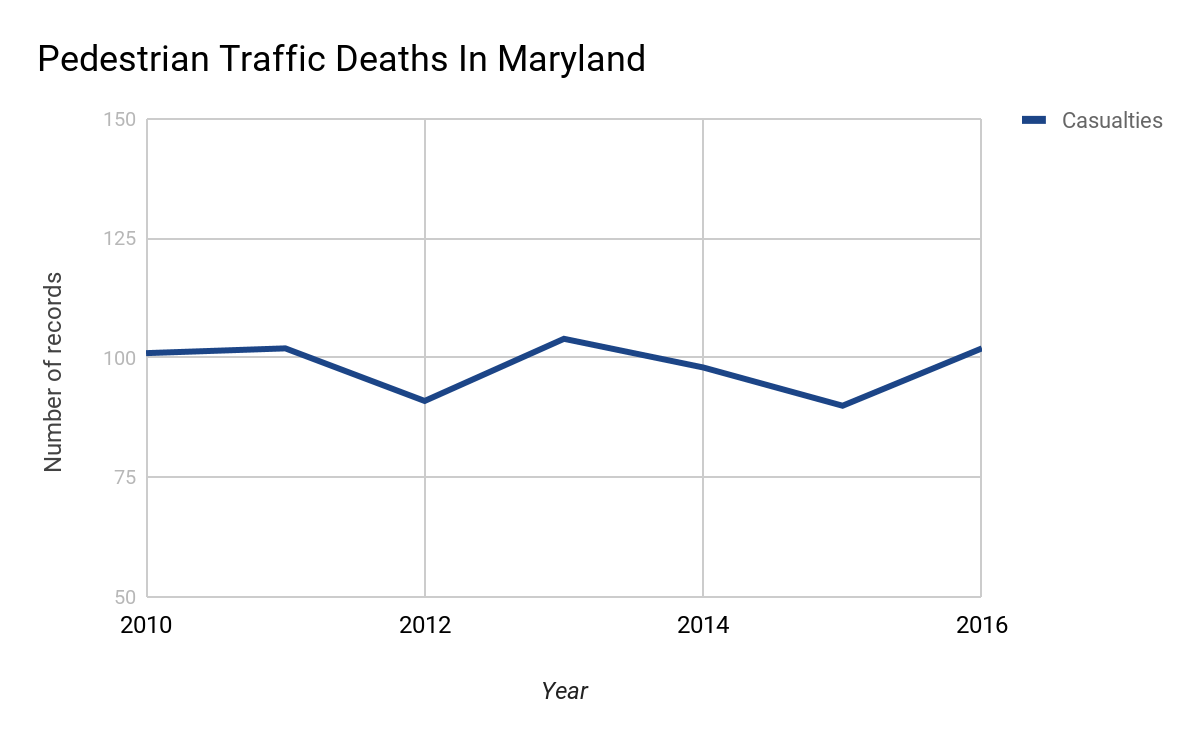
Camila Velloso

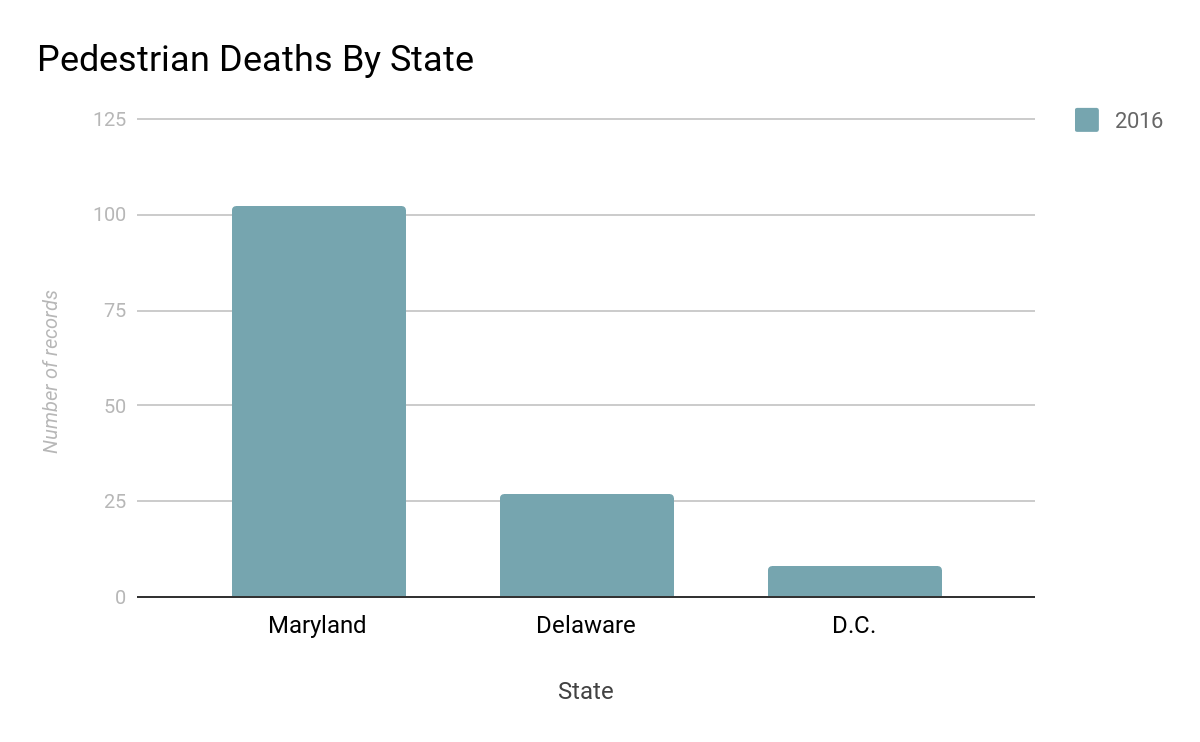
05/14/19

JOUR772: Data Analysis Project Final Memo

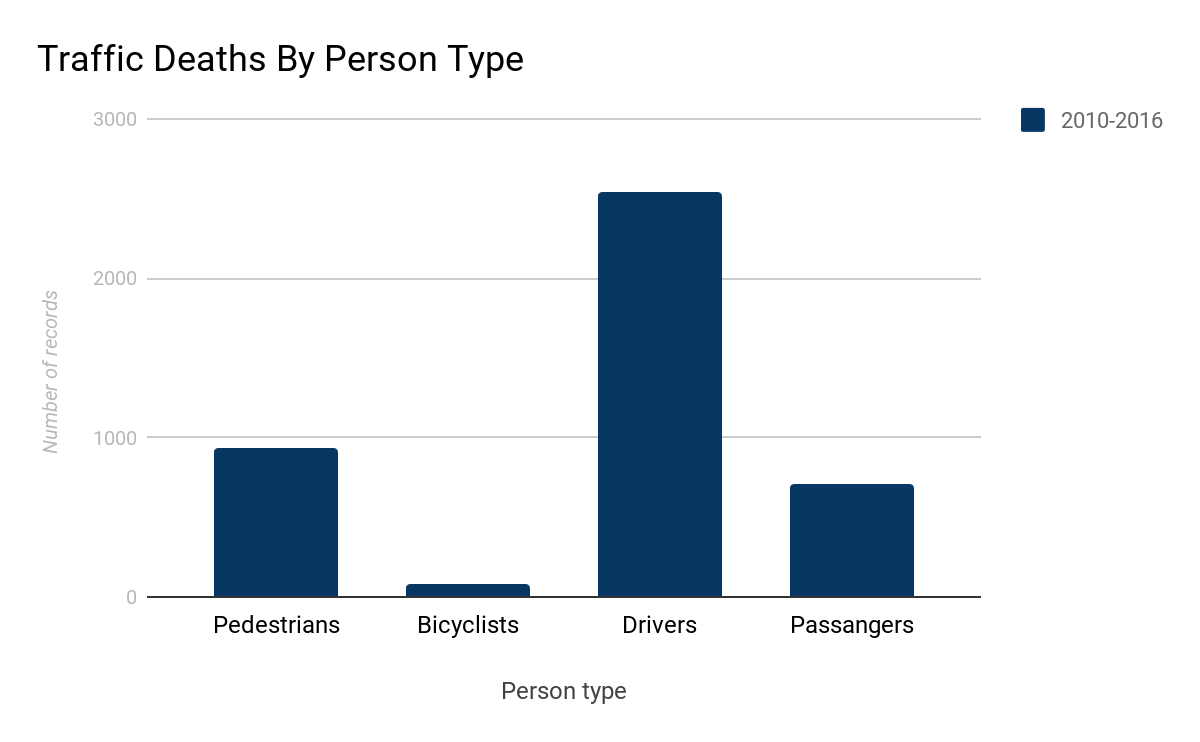
In 2016, 102 pedestrians died in traffic accidents in Maryland - a trend that has remained relatively stagnant since 2010.



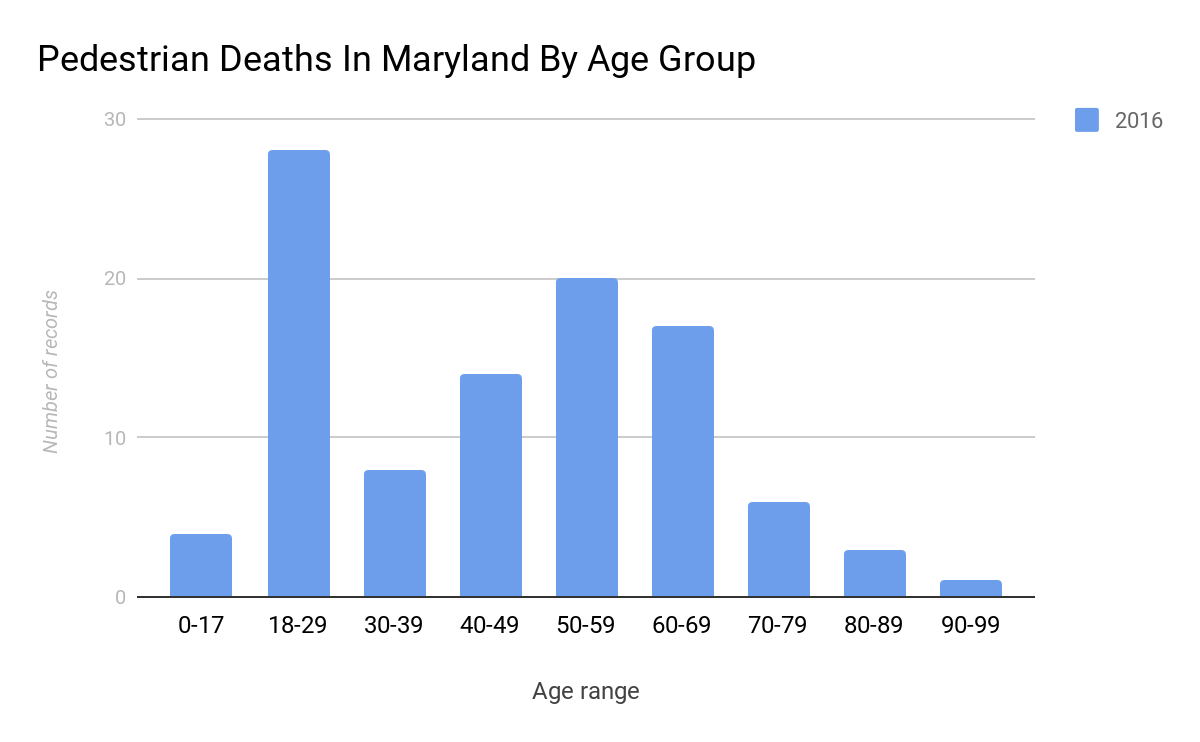
The findings from the Fatality Analysis Reporting System (FARS) reveal that among the states of Maryland, Delaware, and the District of Columbia, Maryland had a significantly higher rate of pedestrian deaths in 2016.



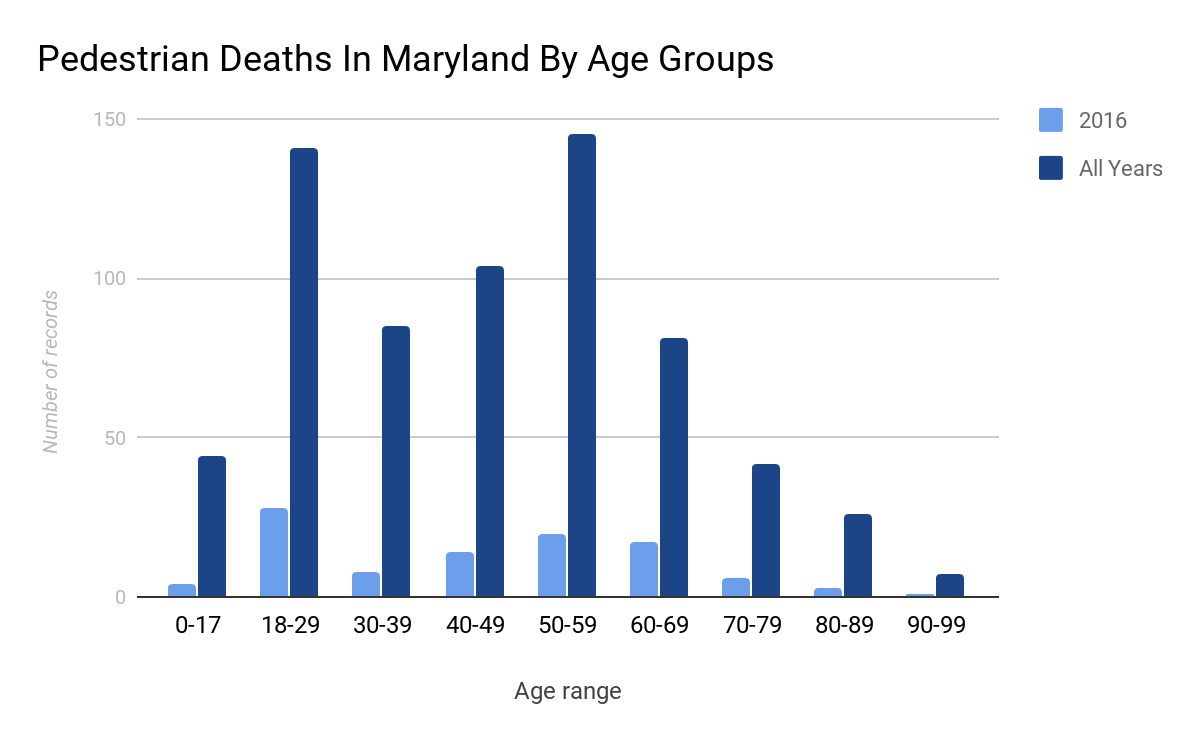
In comparison to other “Person Type” in the PERSON Data File, drivers died more frequently in traffic accidents, followed by pedestrians, passengers and bicyclists.



People between the ages of 18-29 make up the largest share of pedestrian deaths in Maryland in 2016, followed by 50-59-year-olds and 60-69-year-olds. That is, 18-29-year-olds comprised of 27.7 percent of pedestrian deaths in Maryland in 2016, while 50-59-year-olds and 60-69-year-olds comprised of 19.8 percent and 16.8 percent, respectively.

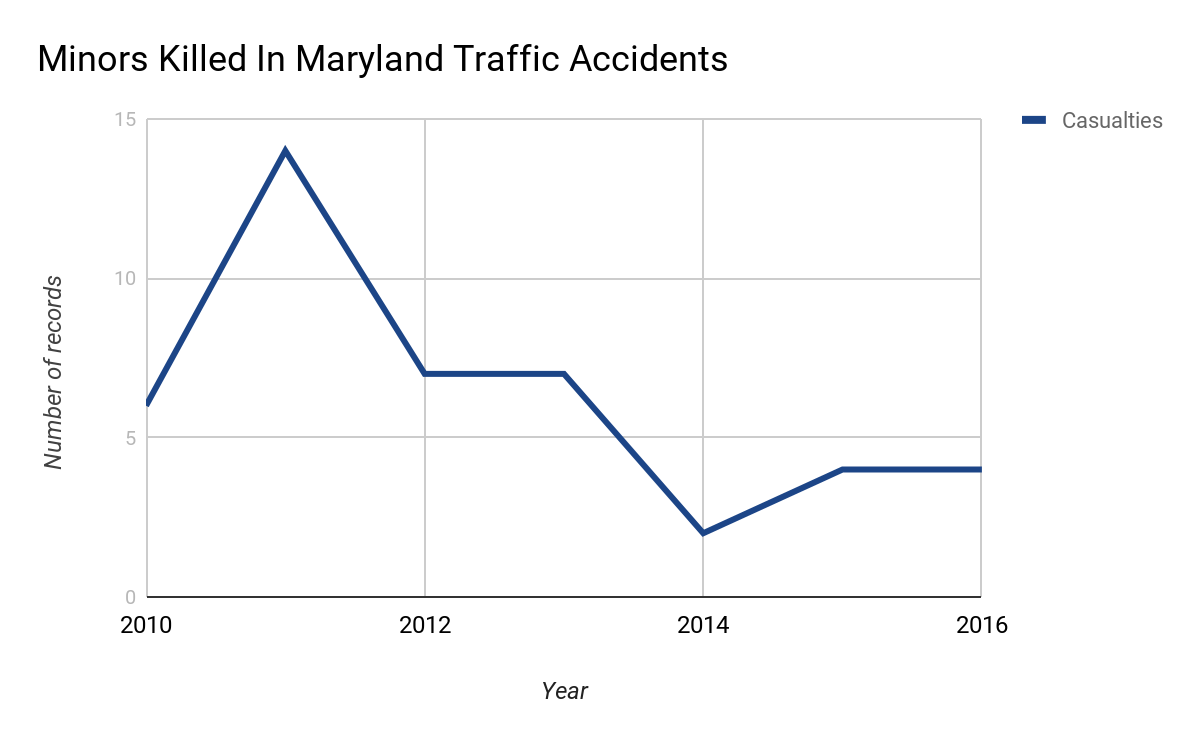


The trend holds true for the years between 2010 and 2016. The age groups of 18-29 and 50-59 made up the two largest shares of pedestrian deaths in Maryland.

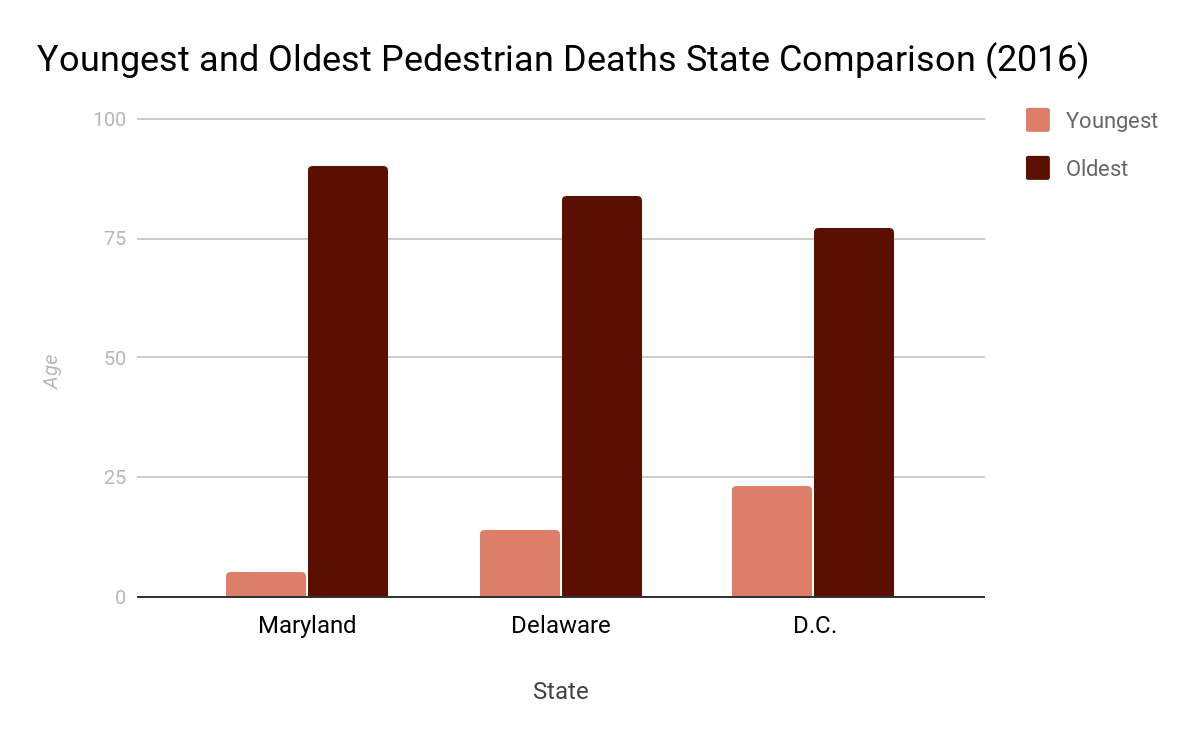


*(“All years” refers to 2010, 2011, 2012, 2013, 2014, 2015 and 2016 combined)*

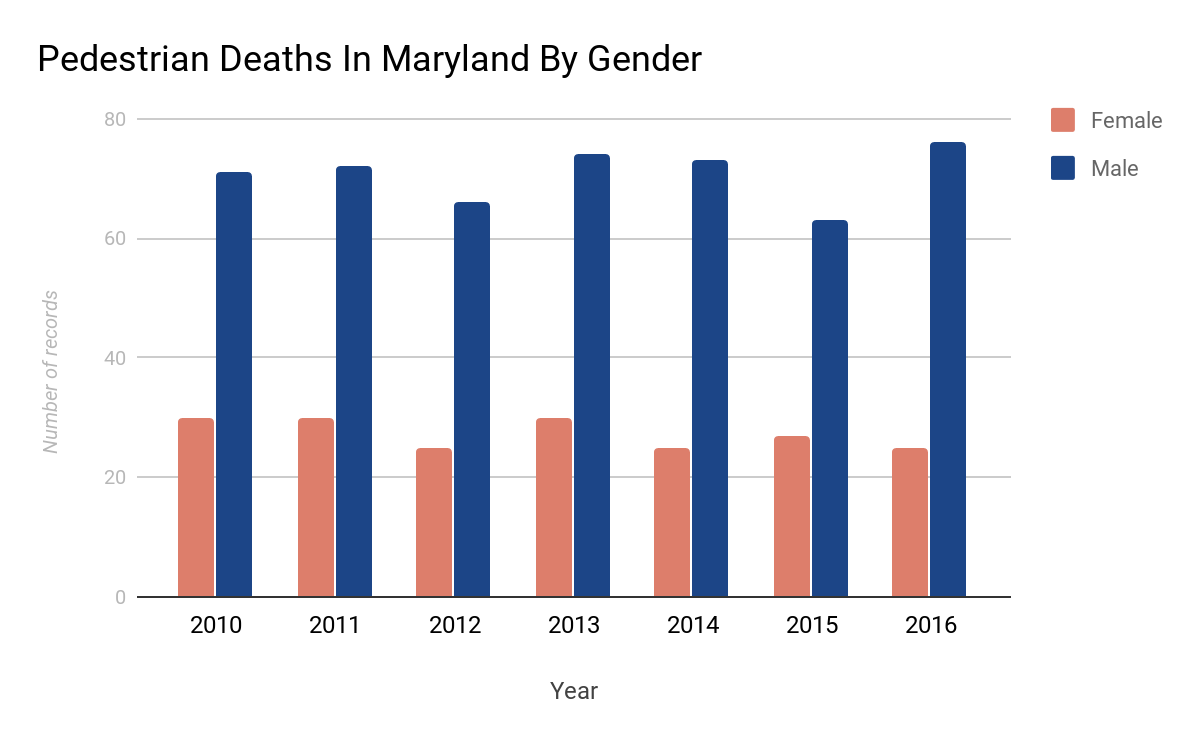
Moreover, 44 pedestrian minors were killed in fatal traffic accidents in Maryland between 2010 and 2016. However, the right-skewed distribution of data suggests that there is a downward trend of minors involved in pedestrian traffic deaths in Maryland.



Still keeping with age, the data revealed the youngest and pedestrian death in each of the three states in 2016. As shown below, the extremities are greatest in Maryland, where the youngest person killed was 5-years-old and the oldest 90-years-old. In Delaware, the youngest person killed was 14-years-old and the oldest 84-years-old. Finally, the youngest person killed in the district was 23-years-old and the oldest 77-years-old.



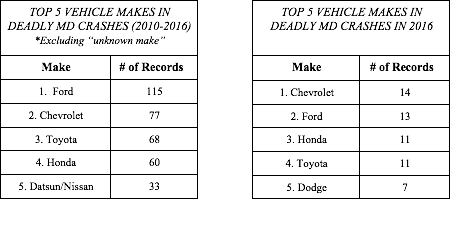
In terms of gender, male pedestrians died at least twice as frequently as female pedestrians in Maryland between 2010 and 2016. For instance, 25 female pedestrians were killed in Maryland in 2016 compared to 76 male pedestrians. In total, 495 men and 192 women died in Maryland traffic accidents from 2010 to 2016.



Besides looking at the attributes of the victims, the data also pointed to specific locations and vehicle makes that could help explain the high rate of pedestrian deaths in traffic accidents.

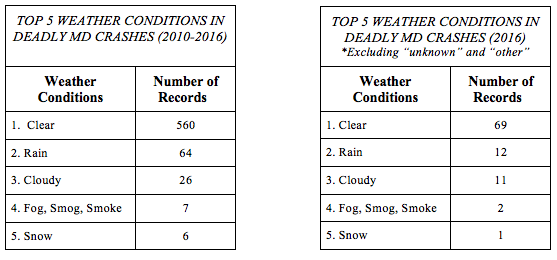
Ford and Chevrolet topped the top five list of vehicle makes involved in deadly pedestrian traffic accidents in Maryland in the years between 2010 and 2016, followed Toyota and Honda. The number of records drops significantly between the first and second ranked vehicles and then again between the fourth and fifth ranked vehicles.

Table 1. Table 2.



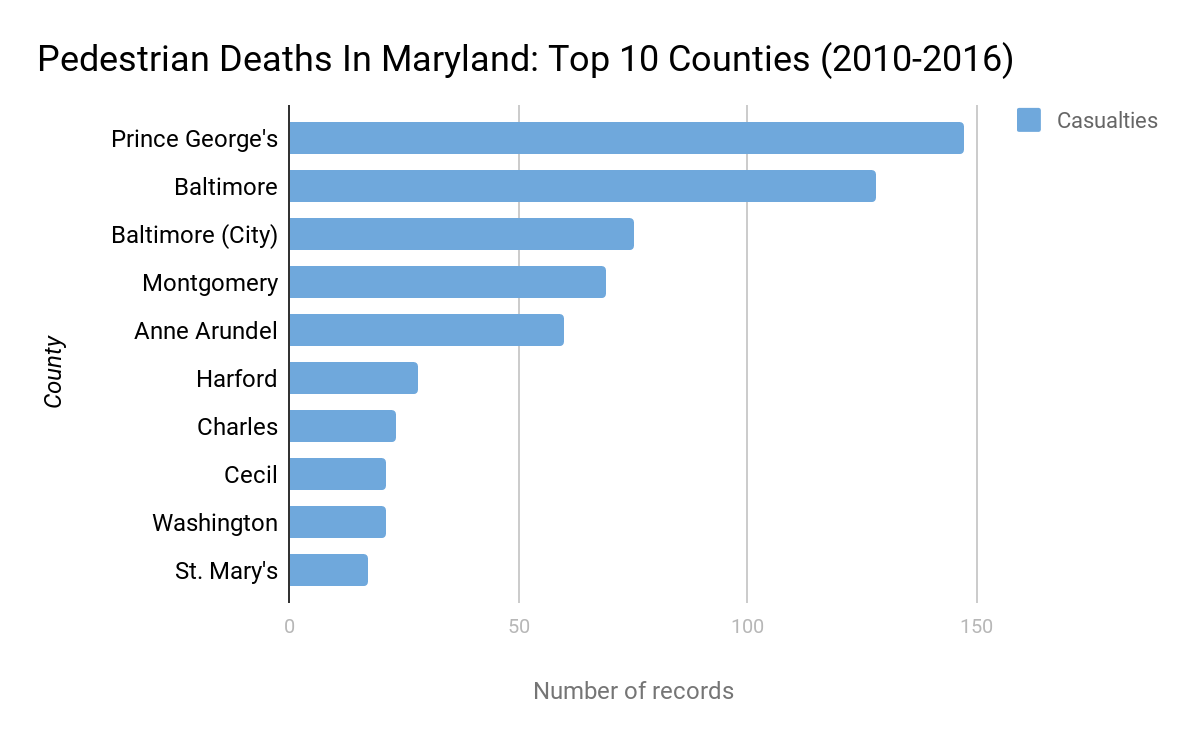
Surprisingly, most deadly pedestrian traffic incidents occurred in clear weather conditions. In fact, “clear” skies made up 81 percent of the top five weather conditions in fatal pedestrian accidents in Maryland, followed by “rain,” “cloudy,” “fog, smog, smoke” and “snow.”

Table 3. Table 4.

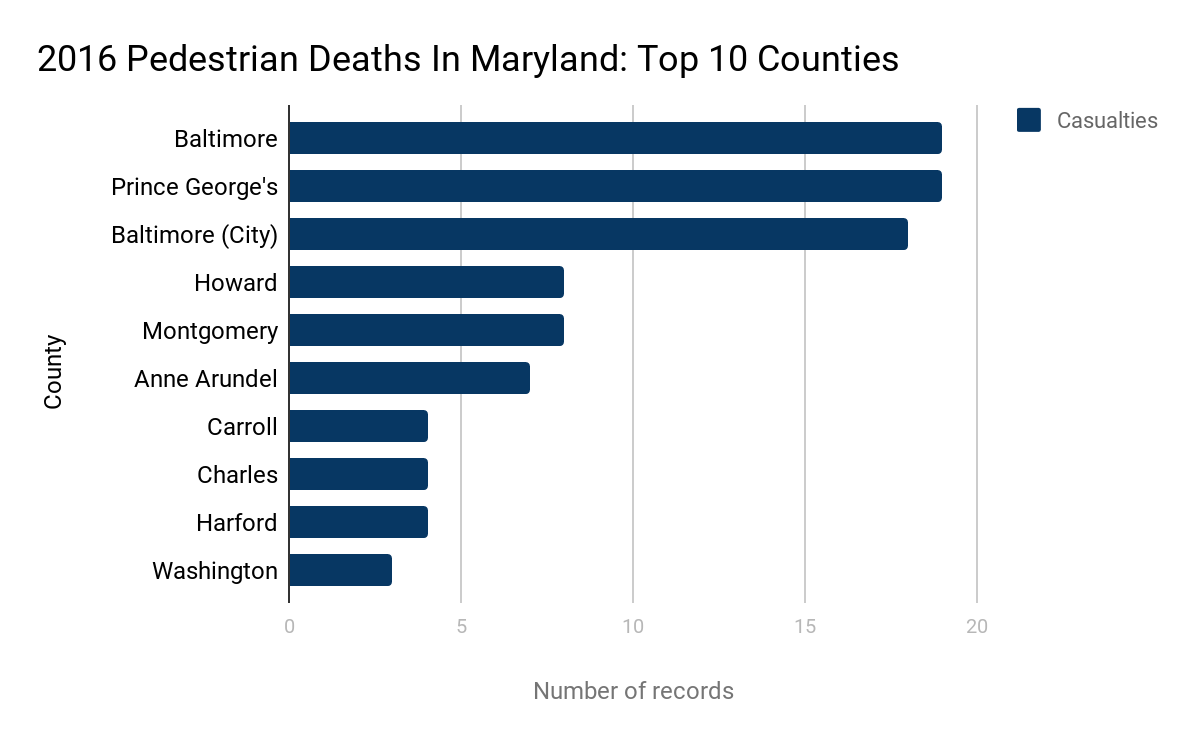


Finally, location played an important role in determining the deadliness of traffic incidents involving pedestrians. Using the RELJCT2 field in the ACCIDENT Data File, I found that most pedestrian fatalities occurred in “Non-Junctions,” that is, not at an intersection. This is confirmed by looking at the LOCATION field in the PERSON Data File, in which most accidents happened in “Not At Intersection-On Roadway, Not in Marked Crosswalk.”

In Maryland, Prince George’s County and Baltimore County stood out for having the highest number of pedestrian fatalities.



Both between 2010 and 2016 and in 2016 alone, Prince George’s County, Baltimore County and Baltimore City topped the list of deadliest pedestrian jurisdictions in Maryland.



When I first began this project, I set out to study fatalities linked to distracted driving in Maryland, Delaware and Washington D.C. using the FARS dataset. But one morning on my daily commute to the university, I heard an interesting [NPR report](https://www.npr.org/sections/thetwo-way/2018/02/28/589453431/pedestrian-fatalities-remain-at-25-year-high-for-second-year-in-a-row) on the radio that found an increase in pedestrian fatalities in the US, with children and elderly people being “especially vulnerable” to this type of incident.

With this in the back of my mind, I wanted to test the assumption that (a) there’s been an increase in pedestrian fatalities and (b) that children and the elderly are most vulnerable. As I began running functions in R Studio using the ACCIDENTS, PERSONS, and VEHICLES datasets, I found that in Maryland between 2010 and 2016, pedestrian fatalities in traffic accidents remained relatively stable over the past years. Moreover, the data showed that people between the ages of 18-29 make up the largest share of pedestrian deaths in Maryland in 2016, followed by 50-59-year-olds. Most functions were created using the filters “INJ\_SEV” as “4,” denoting a fatal Injury, and “PER\_TYP” as “5,” denoting a pedestrian.

From this preliminary analysis, I already gathered that there were limitations to the FARS data. At first, null values in many of the fields complicated my analysis. I learned to use specific filters to avoid producing null values in the categories I wanted to analyze. The same is true for values indicating human error in inputting data. For example, some years were listed as “9999” or ages were designated as “998,” values that would significantly alter the distribution of data and compromise the analysis. Therefore, if I wanted to look at data between 2010 and 2016 using “DEATH\_YR” in the PERSONS Data File, I would write a function that included the following statement: filter(DEATH\_YR != 9999, DEATH\_YR != 2017).

Another limitation of the FARS data is the scope. The most recent data is from 2016 and only includes cases from Maryland, Delaware and Washington D.C. Accordingly, I can’t apply my findings to a cross-national analysis. I would need verify my findings by interviewing sources, looking at other news reports and cross-referencing figures with other datasets and publications. For example, I could double check the number of pedestrian deaths in reports by the Governors Highway Safety Association, the National Highway Traffic Safety Administration, or the CDC’s Motor Vehicle Safety. Additionally, I could contact the Maryland Department of Transportation for help putting figures into context or simply obtaining more information. I would also speak to government officials to learn more about their plans to combat fatal traffic incidents and make roads safer.

There are many positives to the FARS dataset, however. For instance, it was relatively simple to find information about a specific case across datasets. If I were looking at a specific accident or death, I would use the “st\_case\_year” within each data set to gather information. I could learn details about the victim, such as age, sex and ethnicity using the PERSONS Data File, learn about the vehicle make and model using the VEHICLES Data File, or learn about the location of the crash and whether alcohol or drugs were reported using the ACCIDENTS Data File. With this knowledge, I could potentially find sources and write the lead of a story.

The story I would pitch around this data would focus on what the state of Maryland is currently doing to combat fatal pedestrian traffic accidents, citing the high rate of pedestrian fatalities. The Baltimore Sun [reported](https://www.baltimoresun.com/news/maryland/bs-md-traffic-fatalities-20190417-story.html) a similar story few weeks ago. In 2018, “the number of pedestrians killed in traffic rose to 133 last year from 117 in the prior year, a nearly 14% increase.” The Sun outlined Maryland officials’ five-year plan to “cut the number of traffic fatalities and serious injuries in half by 2030.”

However, the article provides few clues as to why so many fatalities were occurring on Maryland roads, besides “not wearing a seat, driving impaired or distracted, and speeding.” I feel more context, such as listing the counties with the highest pedestrian death rate or analyzing the most dangerous roads or intersections in the state, would benefit readers and help them understand the causes of fatal traffic accidents.

As a story lead, I would reel readers in with the case of the youngest pedestrian death in Maryland in 2016. Something along the lines of:

At 10:20 a.m. on Monday, Jan. 25, 2016, a 5-year-old girl was struck and killed by a Kia vehicle in Carroll County, Maryland. She died less than an hour later at 11:09 a.m.

The driver, a 49-year-old woman, was not reported to be distracted, speeding or under the influence of alcohol or drugs at the time of the collision.

I would look through local news reports about fatal accidents that day to find more information about the victim and driver. I would also reach out to the Carroll County government to gather information about new programs or legislation enacted in the past years to increase road safety and prevent pedestrian deaths. I would contact road safety organizations to learn more about trends in data and whether the number of pedestrian deaths in the area have been increasing in recent years, for a more complete analysis.

*Link to Github repository with the R script:*

<https://github.com/camilavelloso/fars_exercise>