Analysis of Car Crashes in Montgomery County, Maryland

ALY 6110

College of Professional Studies, Northeastern University

ALY 6110 Data Management & Big Data

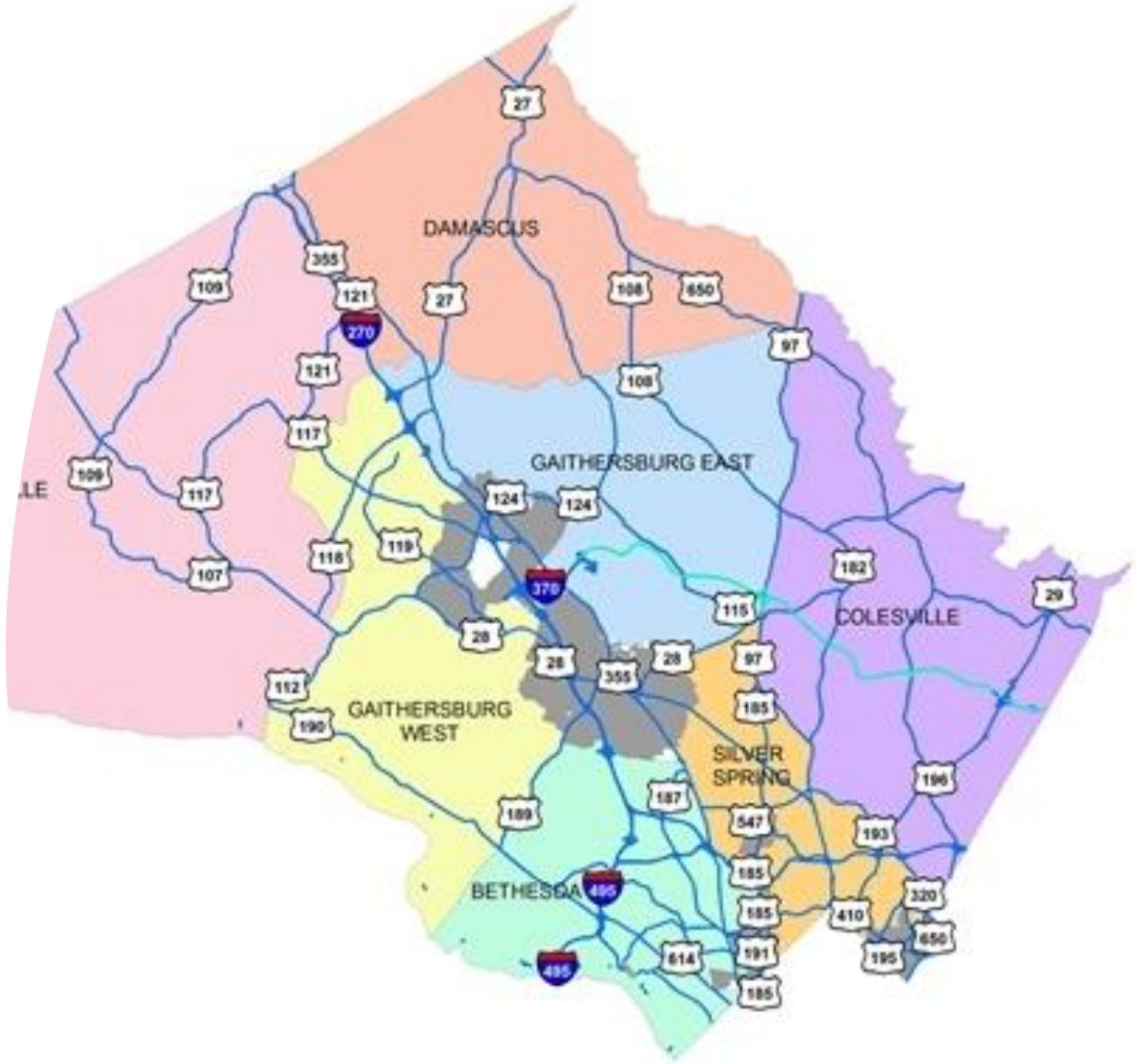
Dr. Donhoffner

February 13, 2025

Eunah Choi, Mallika Choudhari, Prabhakar Elavala, Nikhil Gade, Tejashwini Nagendra Singh

Introduction

- Montgomery County, Maryland
- Car Crashes collected by an Automated Crash Reporting System (ACRS) of the Maryland State Police
- we aim to identify actionable insights that can lead to significant improvements in road safety



Real World Problem

- How do lighting conditions impact traffic collisions in Montgomery County, and which locations require immediate streetlight installations to enhance road safety?
- **H_0 (Null Hypothesis)**: Lighting conditions have no significant impact on crash severity.
- **H_1 (Alternative Hypothesis)**: Lighting conditions significantly affect crash severity.
- By analyzing high-risk crash zones, we aim to determine how lighting contributes to collision severity and identify priority areas where streetlight installation can mitigate accident risks.

Dataset

- Crash Reporting – Drivers Data
- Format: CSV
- Structured Data
  - Numerical, Text, Location Variables
- 192k Rows, 39 distinctive columns
- Time: January 1st, 2015 – January 21st, 2025
- Location, Crash Severity, type of crash, weather conditions, traffic conditions, Driver at Fault(?)

Data Acquisition

- Montgomery County Police, Gaithersburg Police, Rockville Police, or the Maryland-National Capital Park Police
- Worked in Python
- **Feature Engineering:**
- Weather Road Condition Interaction
  - Combines weather conditions and road surface data
- Time of Day Classification
  - Morning, Afternoon, Evening, & Night
- Severity Index
  - Severity metric using injury levels, vehicle damage, and road conditions

Variable Name	Description	Unit of measurements
Weather	What type of weather condition at time of crash	Ex. Clear, Cloudy, Rain, Fog
Surface Condition	The state of the road	Ex. Dry, Ice, Wet
Light	Describing the time of day	Ex. Light, Dark, Dark (with lights on)
Traffic Control	What type of traffic control at site of crash	Ex. Stop Sign, Traffic Signal, no controls
Driver Substance Abuse	Was their substance evolved? What type?	Ex. None/None suspected, unknown, alcohol
Driver Distracted by	Was the driver distracted? What was the cause of the crash?	Ex. Not distracted, phone, unknown, inattentive

Methodology

- Python
  - Data Cleaning & Analysis
  - Critical Values: Lighting Conditions & Road Surface Conditions
- Clustering Techniques:
  - K-Means
  - DBSCAN
    - Utilized to identify high-risk crash zones
- Statistical Hypothesis Testing
  - Chi-Square Test
    - Significance of lighting conditions on crash severity

Methodology cont.

- K-Means clustering was implemented to segment crash-prone areas into high, moderate, and low-risk zones based on accident severity and frequency.
- DBSCAN clustering was used to detect high-density crash hotspots in poorly lit areas, identifying locations where street lighting improvements are most needed.
- Chi-Square testing confirmed a significant correlation between lighting conditions and crash severity, reinforcing the need for infrastructure upgrades in identified zones.
- Geospatial visualizations, including heatmaps and scatter plots, were generated to illustrate patterns in accident distribution and pinpoint areas requiring intervention.

K Means Clustering

Figure 2: Clustering

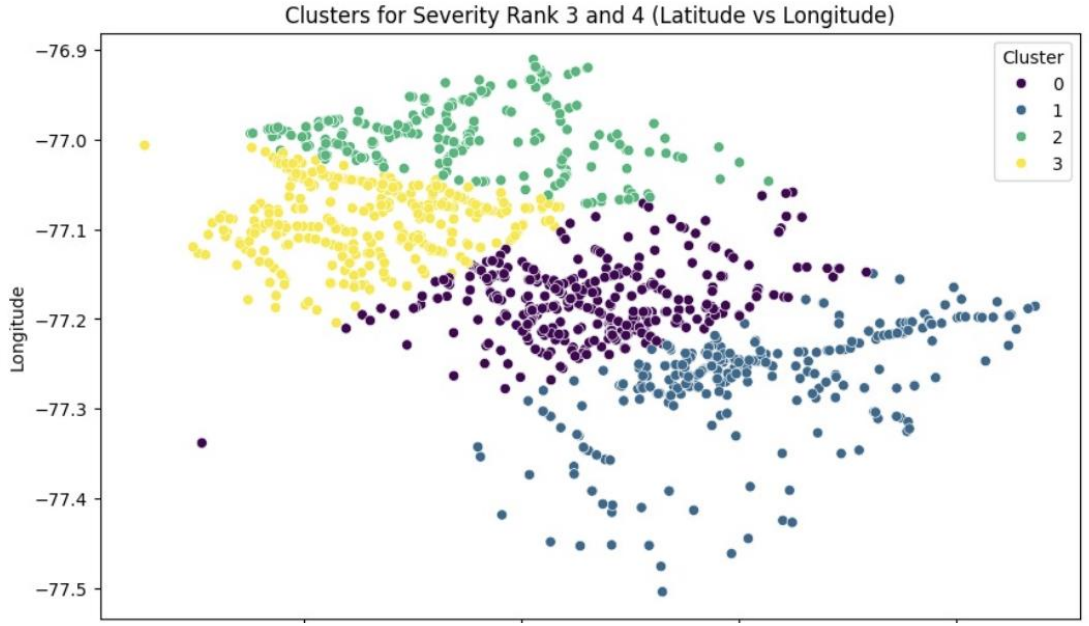
- K-Means & DBSCAN clustering identified high-risk zones based on severity and frequency.
- Dark purple clusters indicate severe crashes, guiding resource allocation for lighting, law enforcement, and infrastructure improvements.

Figure 4: Impact of Lighting Conditions on Crash Severity

- Dark (No Lights) zones show the highest crash severity, proving inadequate lighting worsens accidents.

Chi Square Test

- P-value = 4.6
- This extremely low p-value indicate statistically significant relationship between lighting conditions and crash severity.



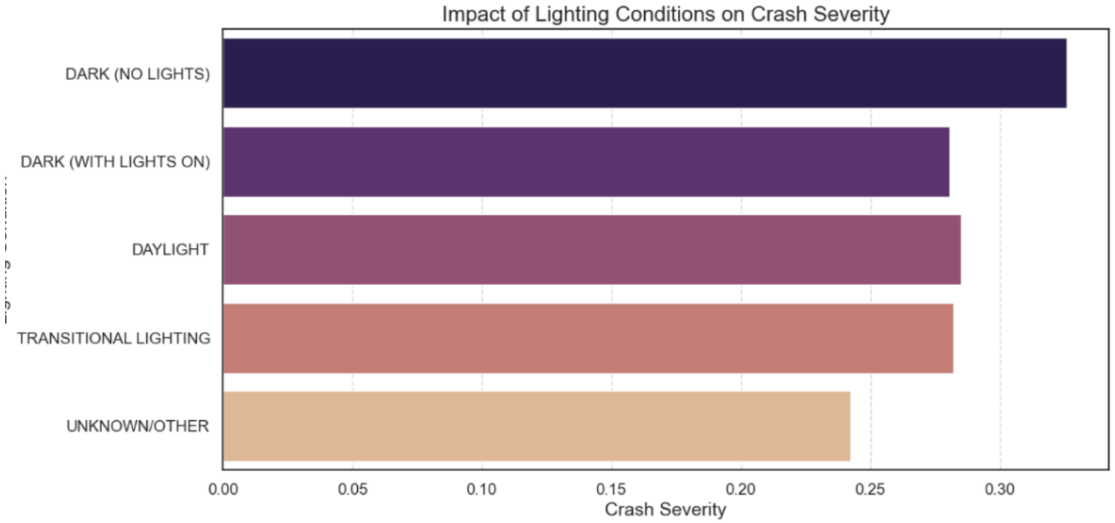


Figure 4: Impact of Lighting Conditions on Crash Severity

DBSCAN Clustering

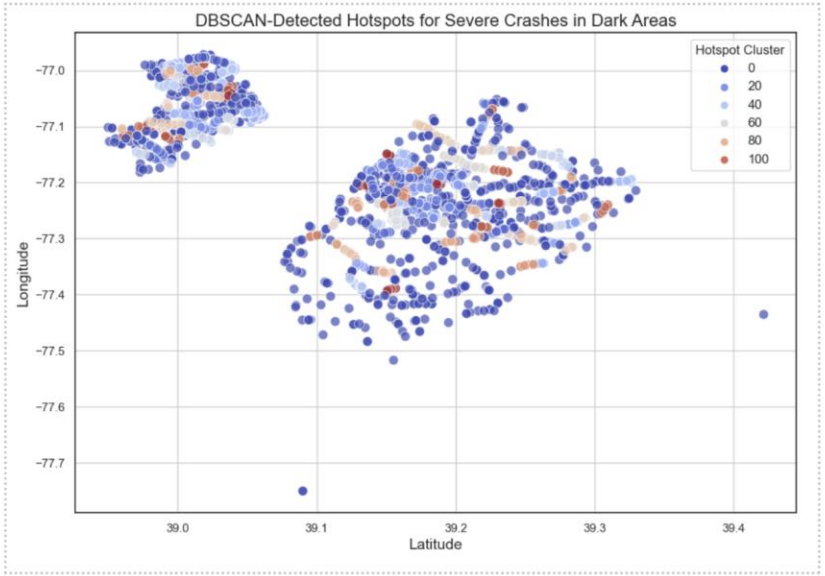


Figure 6: DBSCAN- Detected Hotspots for Severe Crashes in Dark Areas

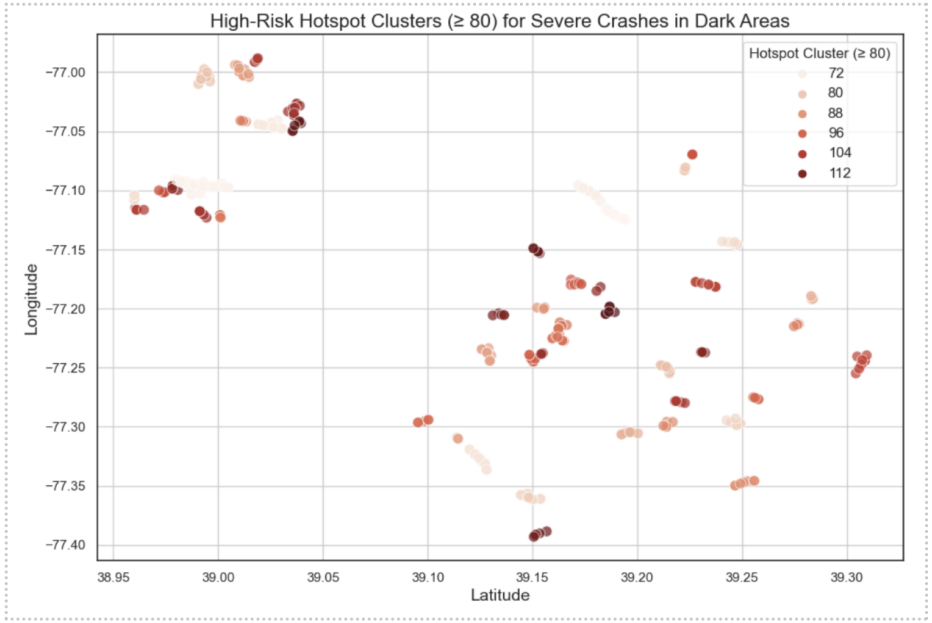


Figure 7: High Risk Hotspot Clusters

Figure 6: DBSCAN-Detected Hotspots

- DBSCAN clustering identified severe crash hotspots in dark areas, highlighting high-density accident zones requiring immediate intervention.
- Red and orange points indicate the most dangerous locations, reinforcing the need for improved street lighting and infrastructure upgrades.

Figure 7: High-Risk Hotspot Clusters (≥80)

- Filtered DBSCAN clusters (severity ≥ 80) pinpoint the most critical crash zones, ensuring targeted safety measures.
- Findings emphasize the urgency of strategic lighting installations in these high-risk areas to reduce severe crashes and improve road safety.

Top High-Risk Roads

- Woodfield Road (8)
- Clarksburg Road (8)
- Log House Road (7)
- Cattail Road (5)
- Damascus Road (5)

```
.spot Clusters Detected: [ 0  1 93  2  3  4  5 -1  6  7  8  9 10 88
. 23 24 25 26 31 27 28 29 30 32 33 34 35 36 37 38 39 40 57 41 48
.4 68 45 46 47 49 50 51 52 53 54 55 56 58 59 60 61 62 63 64 65 67
71 72 73 74 75 76 92 77 78 79 80 81 82 83 84 86 85 91 87 94 89 90]

'op High-Risk Roads for Immediate Streetlight Installation** 🚦
                                                  Road Name
                                                WOODFIELD RD
                                                CLARKSBURG RD
                                                LOG HOUSE RD
                                                  CATTAIL RD
                                                 DAMASCUS RD
.4                                                SENECA RD
10                                  EISENHOWER MEMORIAL HWY
1                                             BARNESVILLE RD
11                                             FIELDCREST RD
23                                             RIFFLE FORD RD
22                                                   RIDGE RD
18                                          OLD GEORGETOWN RD
2                                                 BRADLEY BLVD
25                                           SLIGO CREEK PKWY
17                                          MASSACHUSETTS AVE
28                                                SUNDOWN RD
19   RAMP 1 FR RAMP 4 (FR IS270) TO RIDGE RD
0                                                 ABERDEEN RD
16                                             MACARTHUR BLVD
13                                               FREDERICK RD
14                                            GREAT SENECA HWY
20  RAMP 8 FR IS 495 SB TO CLARA BARTON PKWY
21                                                RAYBURN RD
12                                              FOREST GLEN RD
9                                            DENNIS AVE (EB/L)
                                                COLESVILLE RD
5                                               STONEYBROOK DR
                                              STREAM VALLEY DR
                                             CLARA BARTON PKWY
                                          WEST OLD BALTIMORE RD
                                                CAPITAL BELTWAY
```

Results

Lighting conditions significantly impact crash severity; dark areas have the highest accident rates.

K-Means and DBSCAN clustering identified high-risk zones, prioritizing areas for intervention.

Key roads with severe crashes need urgent streetlight installation and traffic control improvements.

Findings validate targeted infrastructure upgrades, optimizing resource allocation for road safety enhancements.

Structured intervention plans for high, moderate, and low-risk zones will help reduce severe crashes and fatalities.

Conclusion

- Data-driven insights confirmed lighting conditions significantly impact crash severity.
- K-Means and DBSCAN clustering effectively identified high-risk zones for targeted interventions.
- Streetlight installations and traffic control improvements are critical to reducing severe crashes.
- Future work includes evaluating post-intervention crash data and integrating smart lighting solutions.
- Collaboration with authorities and urban planners ensures sustainable road safety enhancements.

References

Data Montgomery. *Crash Reporting - Drivers Data*. (2025, February 7). Crash Reporting Drivers Data | Open Data Portal. *Data Montgomery*. https://data.montgomerycountymd.gov/Public-Safety/Crash-Reporting-Drivers-Data/mmzv-x632/about_data

Data.Gov. (2025, February 7). Crash Reporting – Drivers Data. *Data.Gov*. *https://catalog.data.gov/dataset/crash-reporting-drivers-data*

Montgomery Planning. (2022, July). Predictive Safety Analysis Final Report. *Montgomery County*. https://montgomeryplanningboard.org/wp-content/uploads/2022/07/Attachment-A.-PSA-Final-Report-July-2022.pdf