Investigating Regional Disparities in Car Accident Locations: A Comparison of Rainy and Clear Days

Changwoo Kim

University of Michigan | Master of Data Science

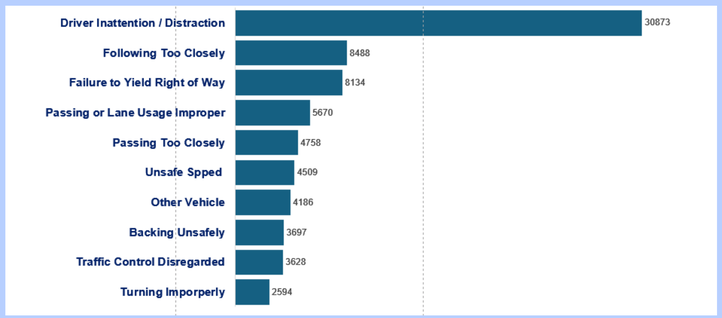
✓ ***"Keep your eyes on the road, your hands on the wheel, and your mind on driving."***

1 Research Question

Are the locations of vehicle accidents significantly different on rainy days compared to clear days?

- While various factors contribute to automobile accidents, driver negligence remains the most significant cause. Rainy conditions, in particular, can disrupt a driver's focus and lead to overlooking pedestrians. Through this study, **we aim to identify specific locations where driving in rainy weather presents greater challenges compared to clear days.** Additionally, we seek to explore potential measures to reduce accidents in such environments.

2 Top 10 Reason for Accident (2021)



- **Driver negligence is the predominant cause of accidents**, and rainy conditions can significantly contribute to such negligence. Therefore, it is essential to carefully consider the impact of rain on driving behavior and ***explore various measures to ensure safer driving in rainy weather.***

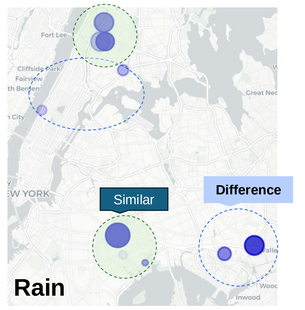
3 Data + [1] external weather Data

Using a total of **196,099** data points, we analyzed factors such as the cause of accidents, vehicle types, accident locations, and the scale of damage. External data, such as **New York State's weather information**, was utilized to extrapolate the weather conditions on the days of accidents using rainfall data, allowing us to verify the hypothesis.

- Some data points lacked location information, which was subsequently **verified through secondary validation using street-level details.** Additionally, the scale of vehicle accident damage was quantified based on the number of injured individuals, and the differences in scale were visualized on a map.

Premise #1 Comparison of Regional Differences Based on Conditions

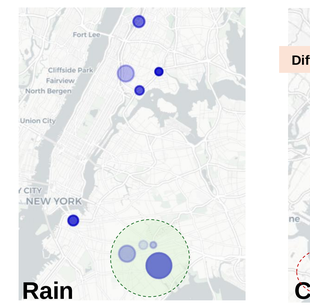
Total number of collisions in 2021: **110,557** | in 2022: **85,542**



< Radius : Sum of Causality, Color : Accident Rate Ranking >

2021

	Latitude	Longitude	Accident Rate
Rain	Atlantic Ave / Pennsylvania Ave.		0.000710
	95 / GrandConsourse	40.845190 -73.911200	0.000543
	Belt Parkway @23A Exit	40.667313 -73.739260	0.000543
Clear	Atlantic Ave / Pennsylvania Ave.		0.000566
	University Height Brg / Major Deegan Exp	-73.912750	0.000509
	Linden Bldv / Pennsylvania Ave.		0.000509



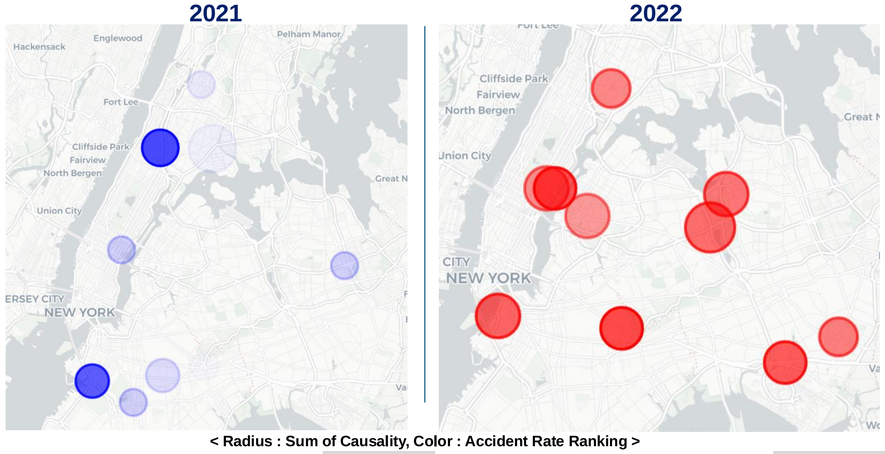
< Radius : Sum of Causality, Color : Accident Rate Ranking >

2022

	Latitude	Longitude	Accident Rate
Rain	Linden Bldv / Pennsylvania Ave.		0.000852
	Bruckner Blvd / Hunts Point Ave		0.000660
	Flatbush Ave / Tillary St		0.000660
Clear	University Height Brg / Major Deegan Exp	-73.912750	0.000899
	Atlantic Ave / Pennsylvania Ave.		0.000699
	Atlantic Ave / Logan St.		0.000599

Premise #2 How significant are the actual regional differences, and what is the scale of these differences?

Locations with Increased Accident Rates During Rain



< Radius : Sum of Causality, Color : Accident Rate Ranking >

Latitude	Longitude	Δ Delta	Δ Percent	Latitude	Longitude	Δ Delta	Δ Percent
40.820435	-73.936230	0.000232	1228.85%	40.684937	-73.91014	0.000222	890.78%
40.667076	-73.995080	0.000211	1118.11%	40.761180	-73.95791	0.000222	890.78%
40.753624	-73.969440	0.000171	453.68 %	40.666306	-73.79171	0.000225	450.43%

- **The patterns of increased accident rates showed completely distinct trends, revealing that the accident rate could be up to 12 times higher on rainy days. It was observed that the increase in accidents was not merely concentrated in areas with generally high accident rates but rather in specific locations where rain led to a higher frequency of accidents.**

Key Findings

- It was confirmed that locations where accidents frequently occur on rainy days differ from those on clear days.
- Certain areas experience significantly higher accident rates on rainy days, showing up to a tenfold increase.
- The trend of frequently occurring accidents was found to differ between the datasets from 2021 and 2022.

Suggestion

▣ Issues *(Preliminary)*

- High Pedestrian Population
- Too many transportation facilities (subway stations, parking lots, bus stops, etc.)
- Downhill road from a bridge
- Road Parking lots

▣ Suggestions

- **Relocation of bus stops:** The bus stop is located right in front of the intersection, and the nearby parking lot narrows the road to a single lane, potentially disrupting traffic flow.
- **Strengthening enforcement** of illegal parking near the intersection: Illegally parked vehicles obstruct the visibility of turning vehicles and disrupt traffic flow.
- **Removal of parking spaces in front of the park** and lane expansion: This would facilitate smoother traffic flow for buses and shuttle buses during their movement.

Limitations

- **The study faced limitations due to the lack of traffic volume data. If traffic volume data had been available, a more accurate analysis of cases during rainy and non-rainy conditions could have been conducted with combined data.**
- **Additionally, more statistical approaches could have been applied through quantitative analysis, but time constraints did not allow for this. For instance, hypothesis testing using trends in monthly accident counts and t-tests could have been employed for validation of other assumptions.**

1) Data Link : Kaggle New York Weather Data Set (Weather data of New York City from 2016 to **Oct 25, 2022**)