Data Analytics & Visualization Project:

An In-Depth Analysis of Seattle's Collision Data

BY VARUN GREWAL

E-MAIL: VGREWAL.TECH@GMAIL.COM

LINKEDIN | GITHUB

Table of Contents

Introduction:
Key Information

- **Project Goal:**
  - Take a deep dive into the "Collisions" historical dataset for the city of Seattle
  - Better understand the underlying causes behind the **"Injury"** and **"Property Damage"** collisions
  - Discover the statistical relationships between different variables and identify hidden insights
  - Produce meaningful recommendations based on gained insights

- **Data Description:**

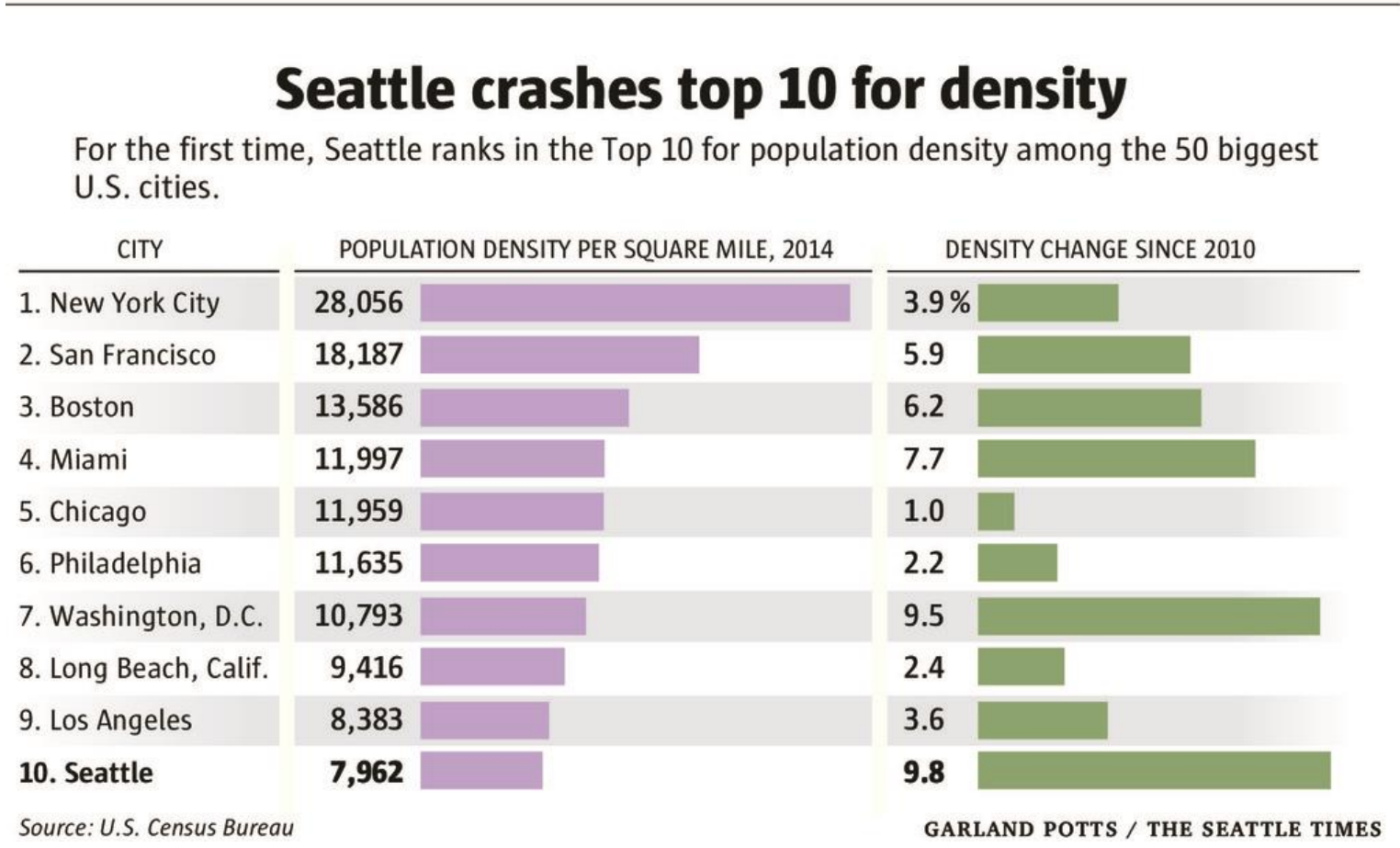
Source: Seattle Police Department | **Year:** 2004-2020 | **Dataset Size:** 195K rows X 38 columns

- **Tools Used:**

Tool: Jupyter Notebook **| Language:** Python
Packages: NumPy | Pandas | Matplotlib | Seaborn | Folium

Introduction: City of Seattle

- Seattle is a growing city. In 2019, It had a net gain of about 11,400 people (+1.5% growth), reaching a total population of 753,700.
- Population Density of 7.9K/mile which makes it the no. 10 spot among the 50 most populous cities in the US
- High density results in "Traffic" which in turn creates ideal conditions for "Collisions"



Introduction:
Need for Analysis?

"Traffic Collisions" has multiple negative consequences for society:

- It can result in loss of human life or a life altering serious injury
- It can result in property damage/financial loss to the people involved and the city
- It can cause traffic jams lasting hours which can result in billons of dollars in lost productivity
- It creates unsafe road conditions for other drivers

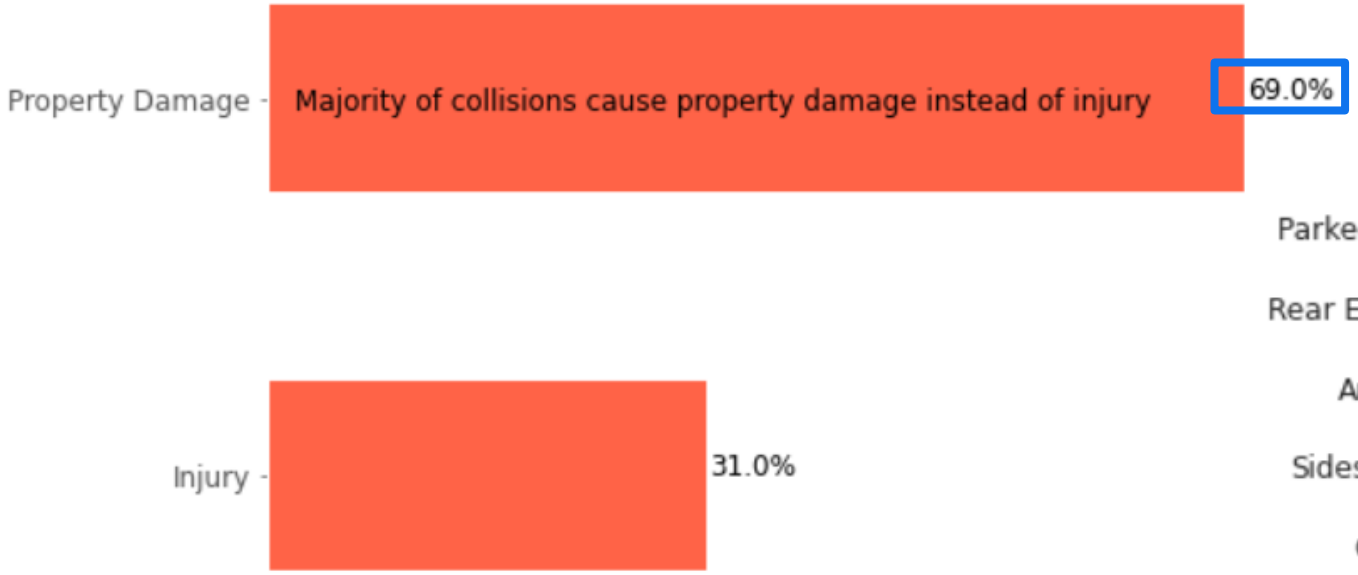
Introduction:
Audience for the Analysis?

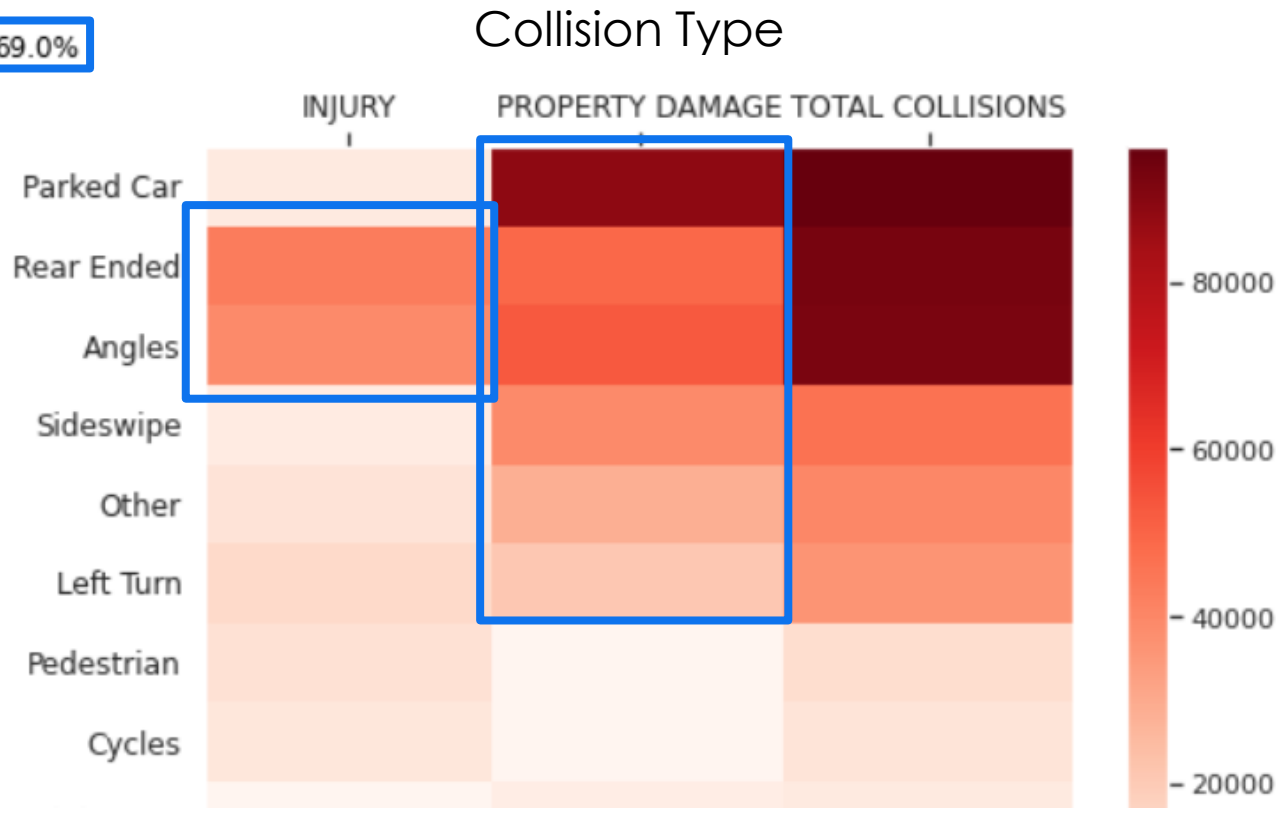
- Potential Employers
- Fellow Data Scientists/ Enthusiasts
- Public Office Holders (Mayor, City Council Members, etc.)
- City/Transportation Planners
- Emergency service providers such as Police, Fire and Medical Technicians
- Non-Profit Organizations
- Residents of the City

Insight 1:

Big picture – Property Damage (69%) is much bigger problem, but Injury (31%) is also quite material.

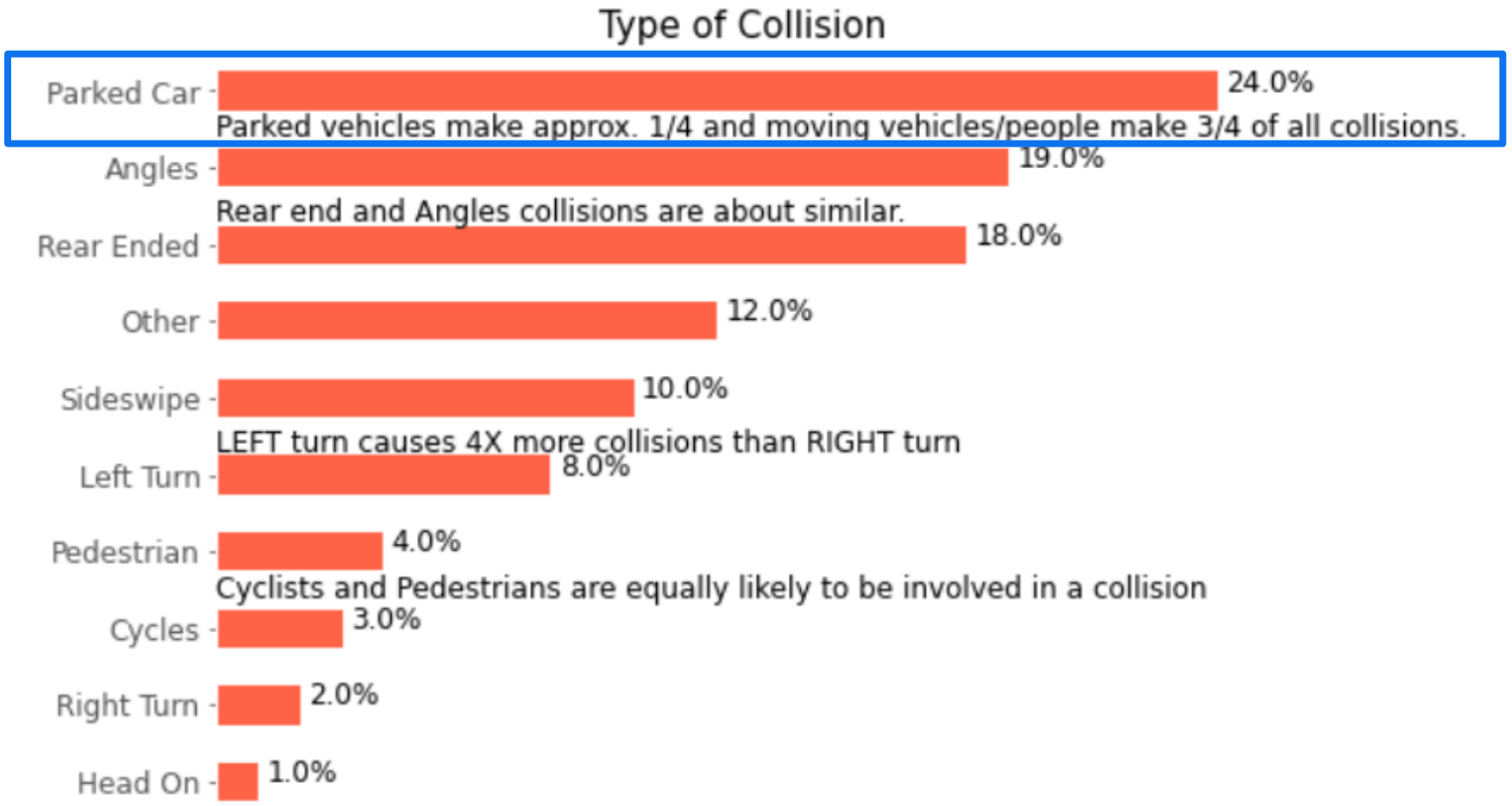
Key collision types – Parked Car, Rear Ended, Angles, Sideswipe and Left turn





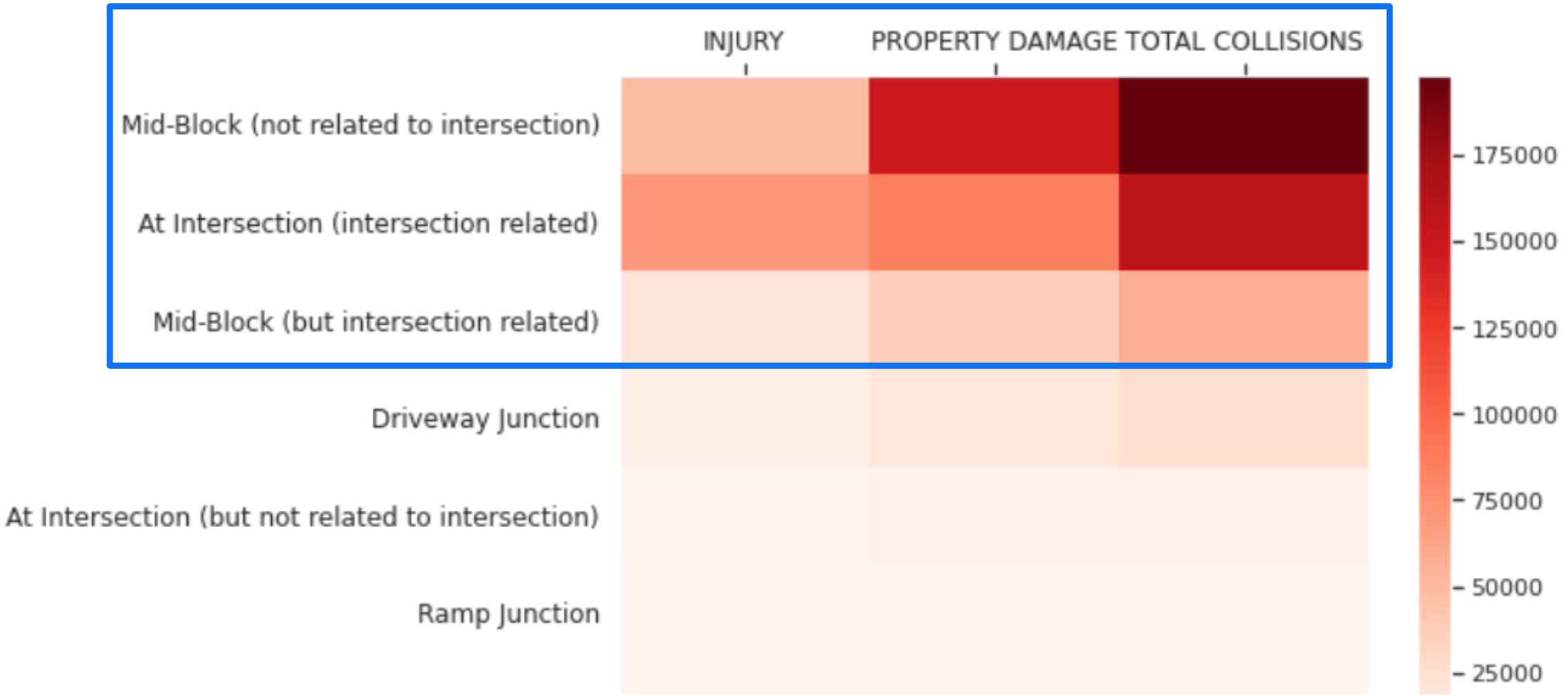
Insight 1 (contd.):

Key collision types (Parked Car, Rear Ended, Angles, Sideswipe and Left turn) make up about 79% of all collisions. Pedestrian and Cycles only make up about 7%.

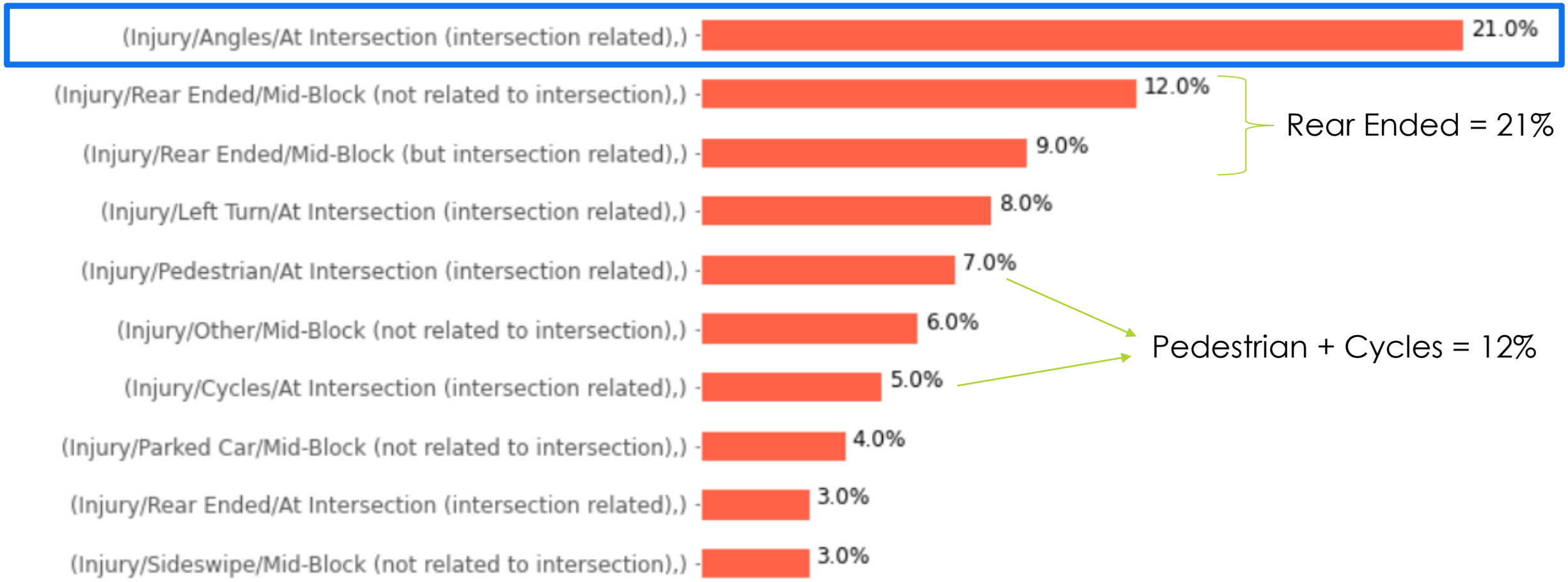


Insight 2:

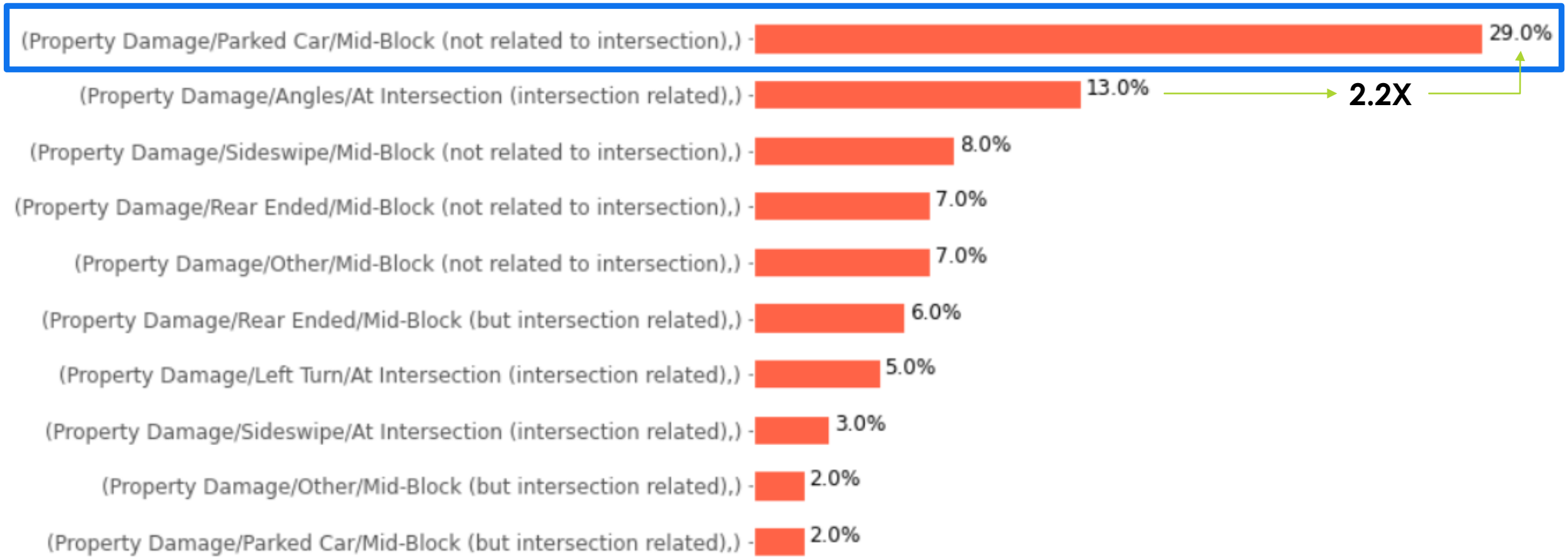
Mid-block and Intersections make bulk of collisions. Makes sense since key collision types are more likely to occur there.



Insight 3:
Top 10 - Injury/Collision/Junction type



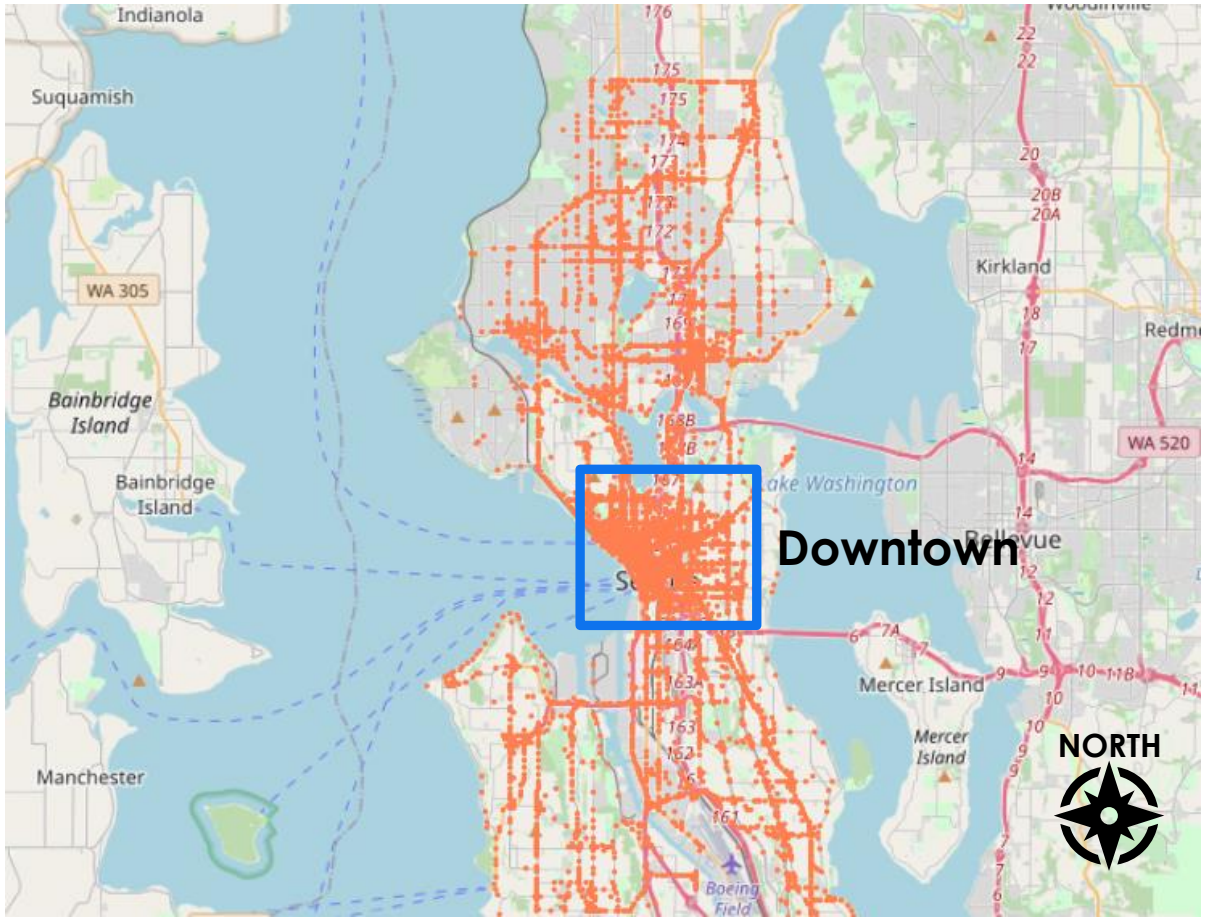
Insight 4:
Top 10 - Property Damage/Collision/Junction type



Insight 5:
Top 5000 locations for Collisions

Area around Seattle Downtown is quite active in collisions

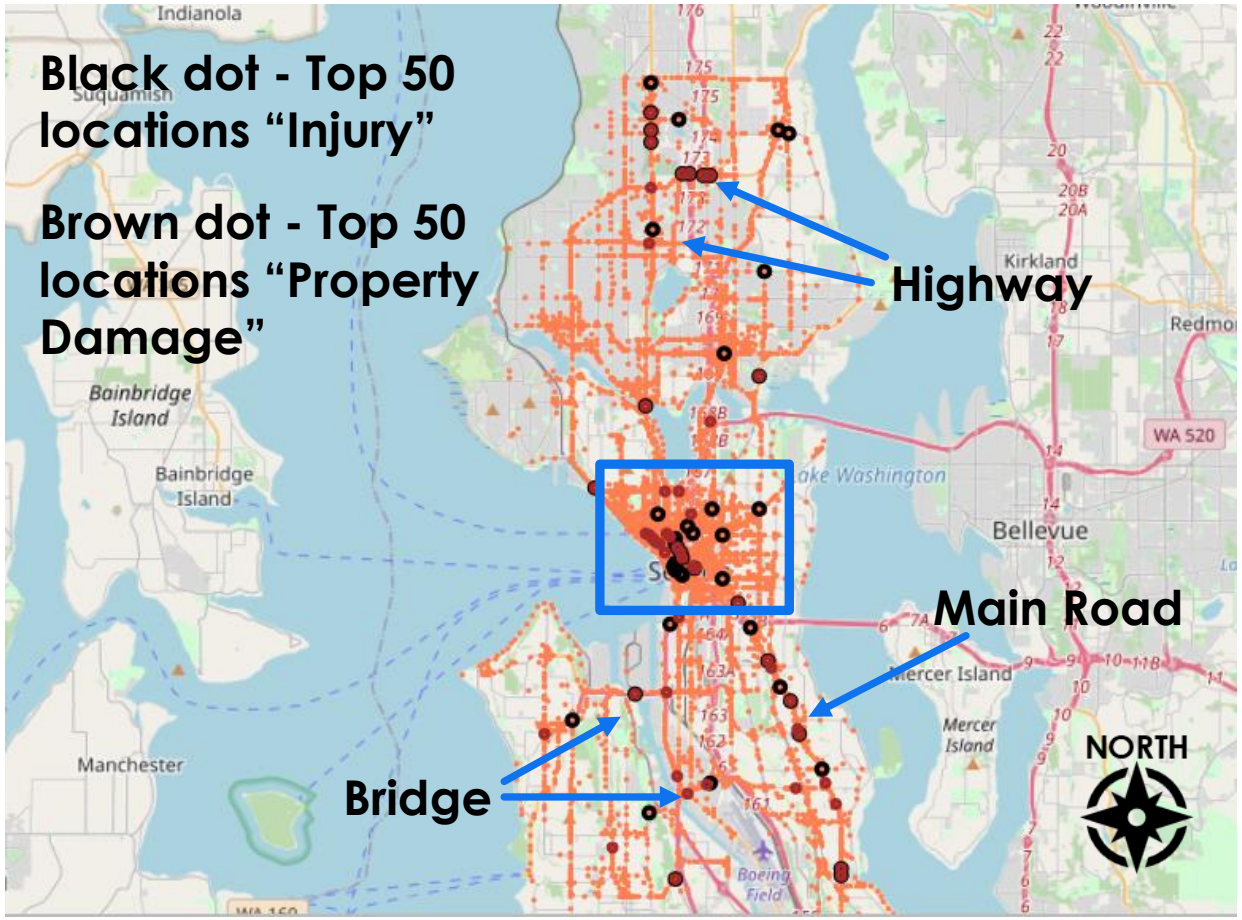
As expected, so are major roads leading to the city



Insight 6:
Top 50 locations superimposed on Top 5000 locations

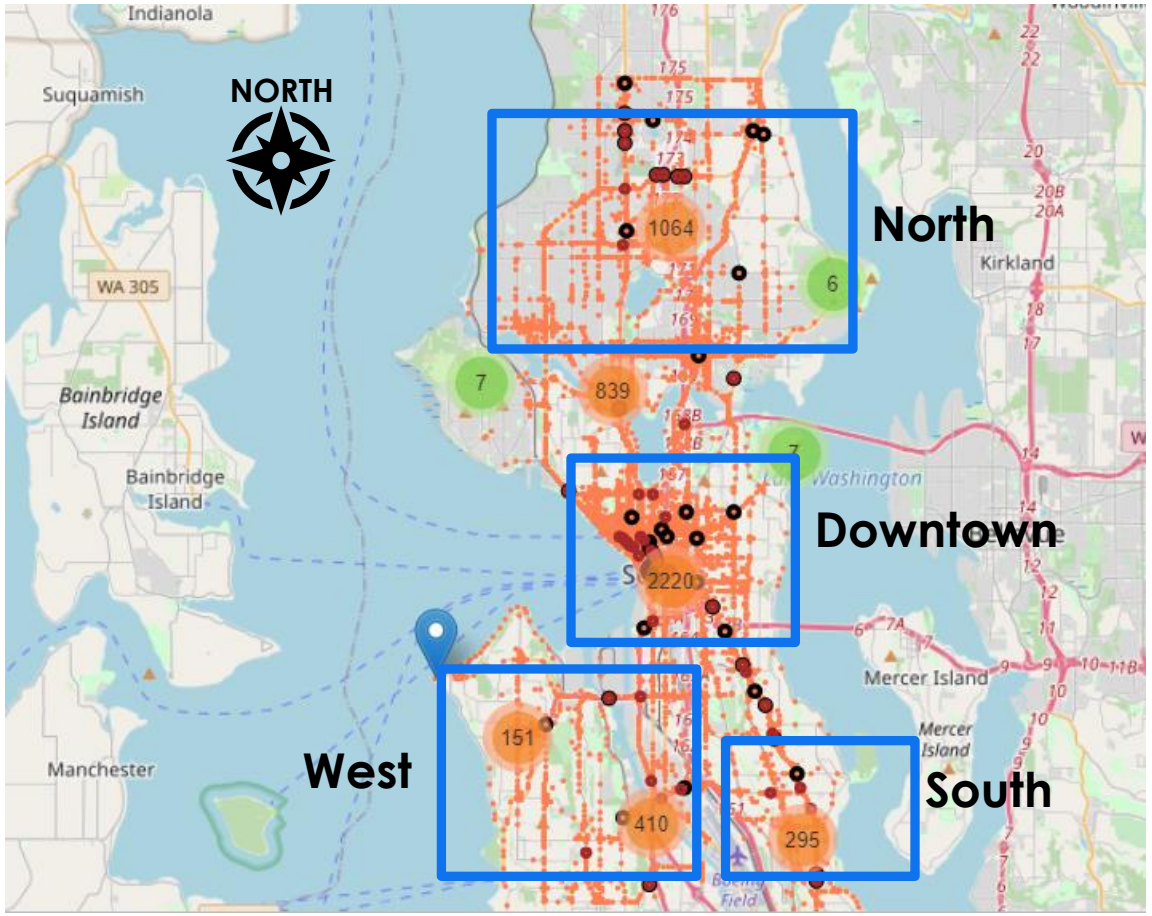
Except for Seattle downtown, top 50 locations seem to be mostly located on major roads, highways or bridges.

Also, some locations have overlap between Injury and Property Damage

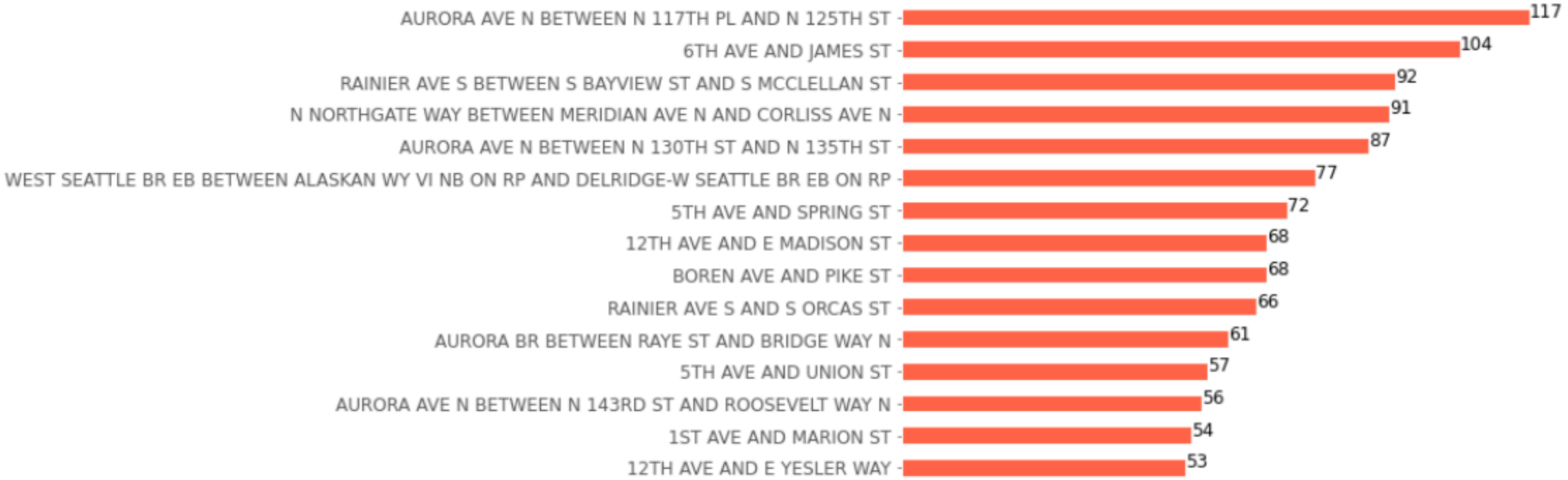


Insight 7:
Number of collisions grouped by location superimposed on Top 5000 and Top 50 locations

- Majority of collisions occur in and around Seattle Downtown which exceeds:
  - North Seattle by **2X**
  - West Seattle by **4X**
  - South Seattle by **7.5X**
- Why Downton Seattle? Because it is dense and crowded with pedestrians and vehicles

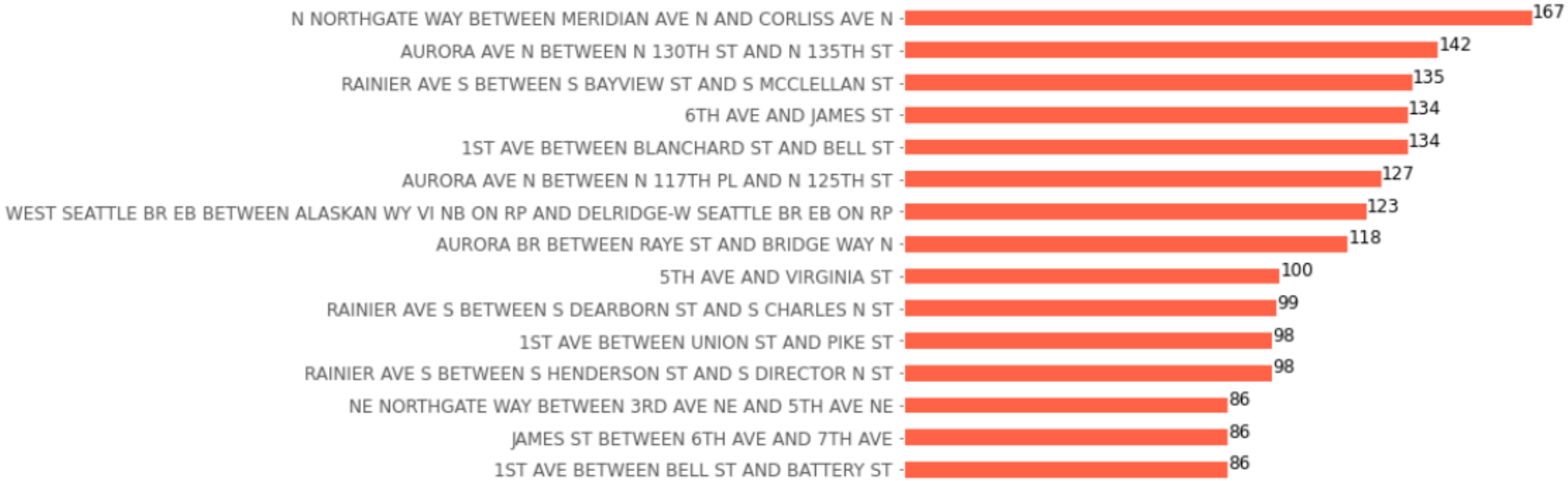


Insight 8:
Top 15 - Most active locations for "Injury"



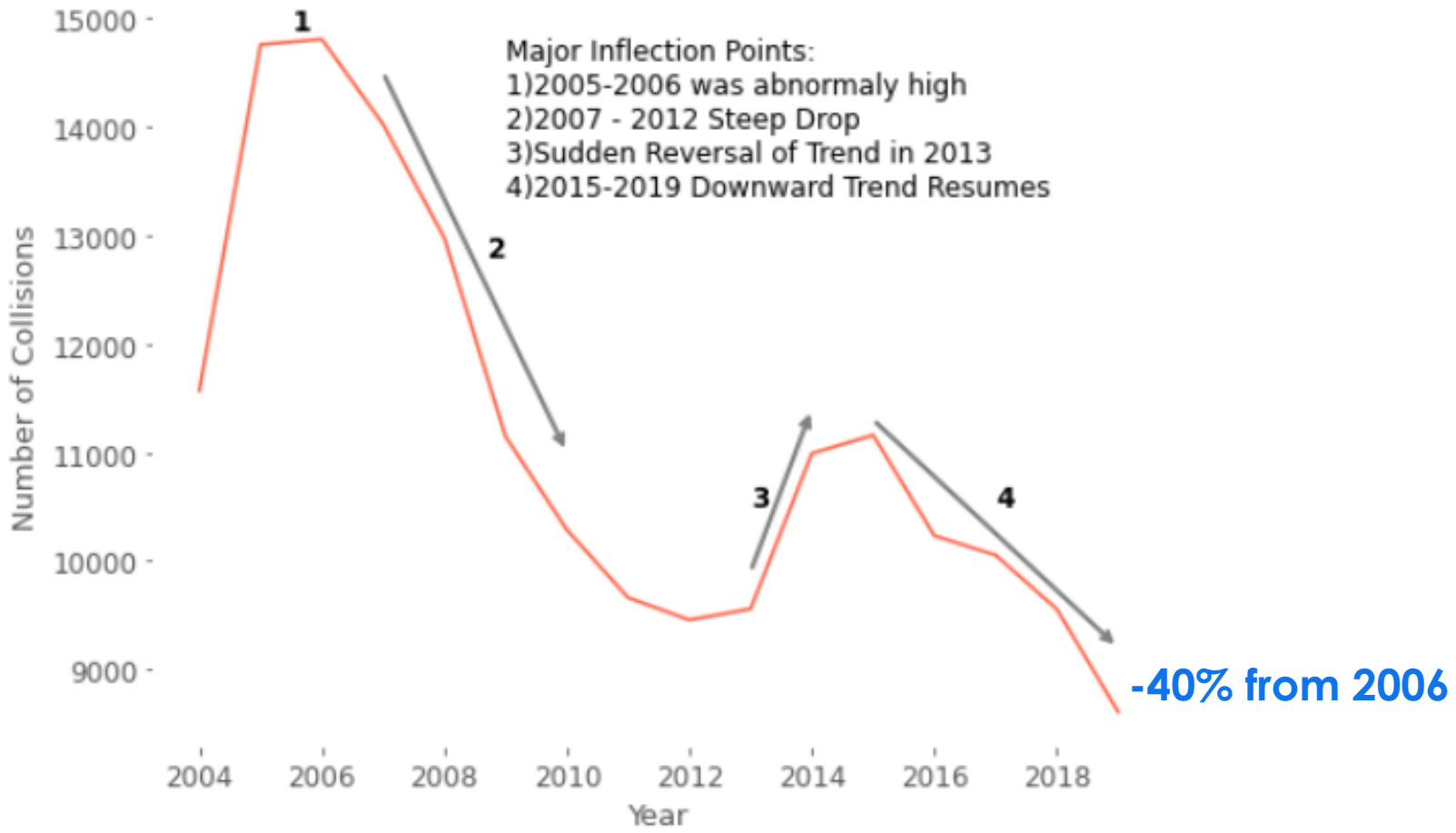
Insight 9:

Top 15 - Most active locations for "Property Damage"



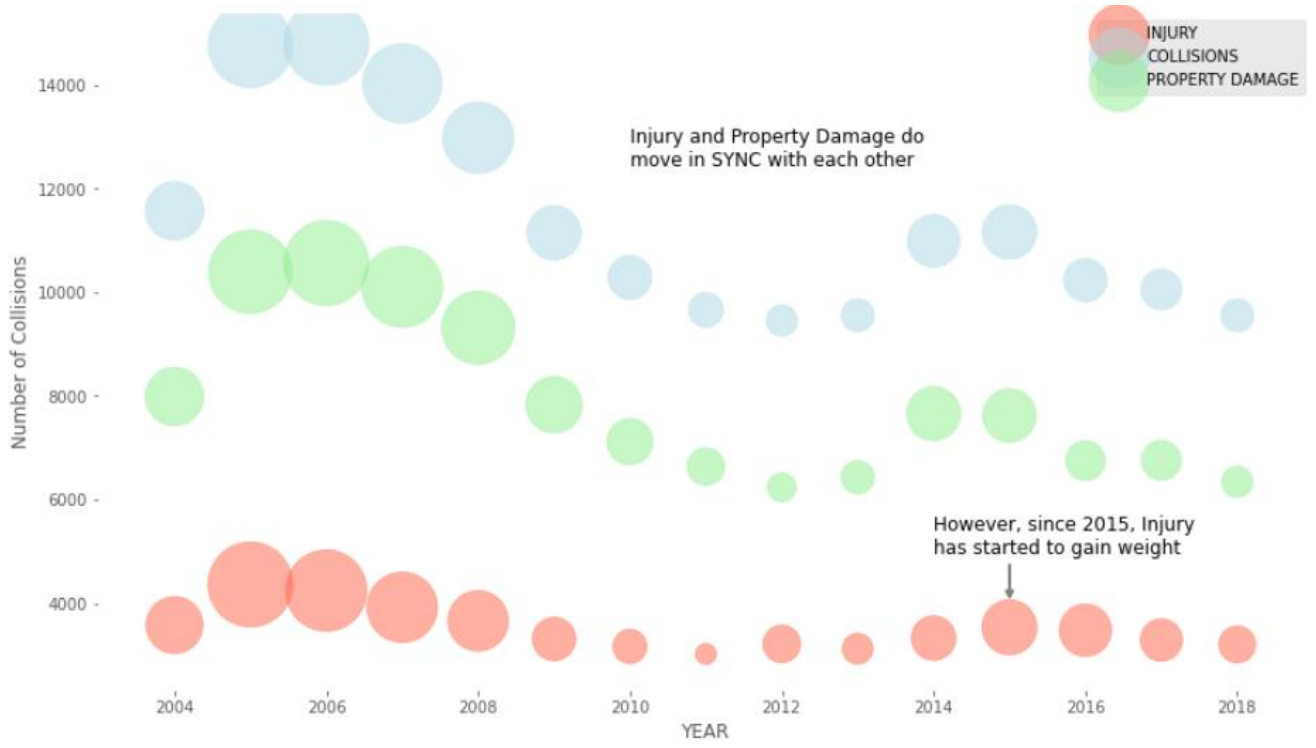
Insight 10:

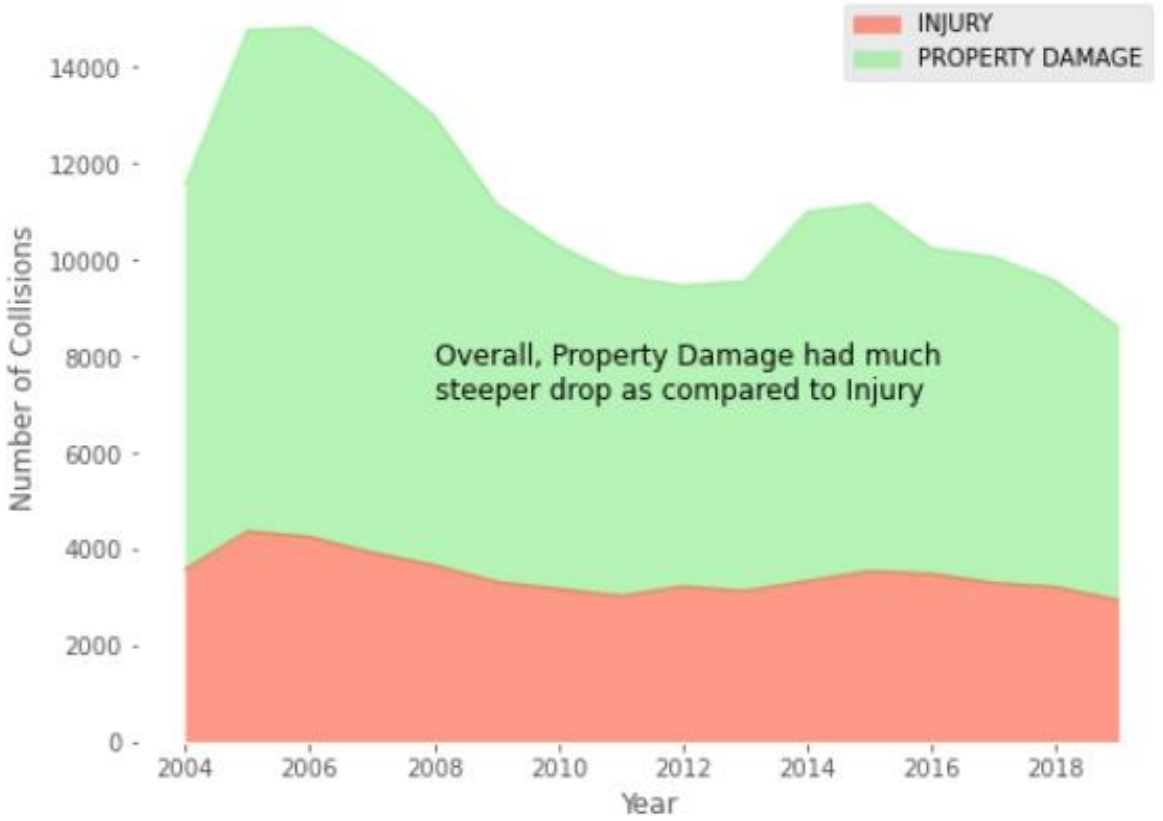
Trend – Total collisions have fallen by approx. 40% from its peak in 2006
It is a very material improvement – Suggests Government efforts are successful



Insight 11:

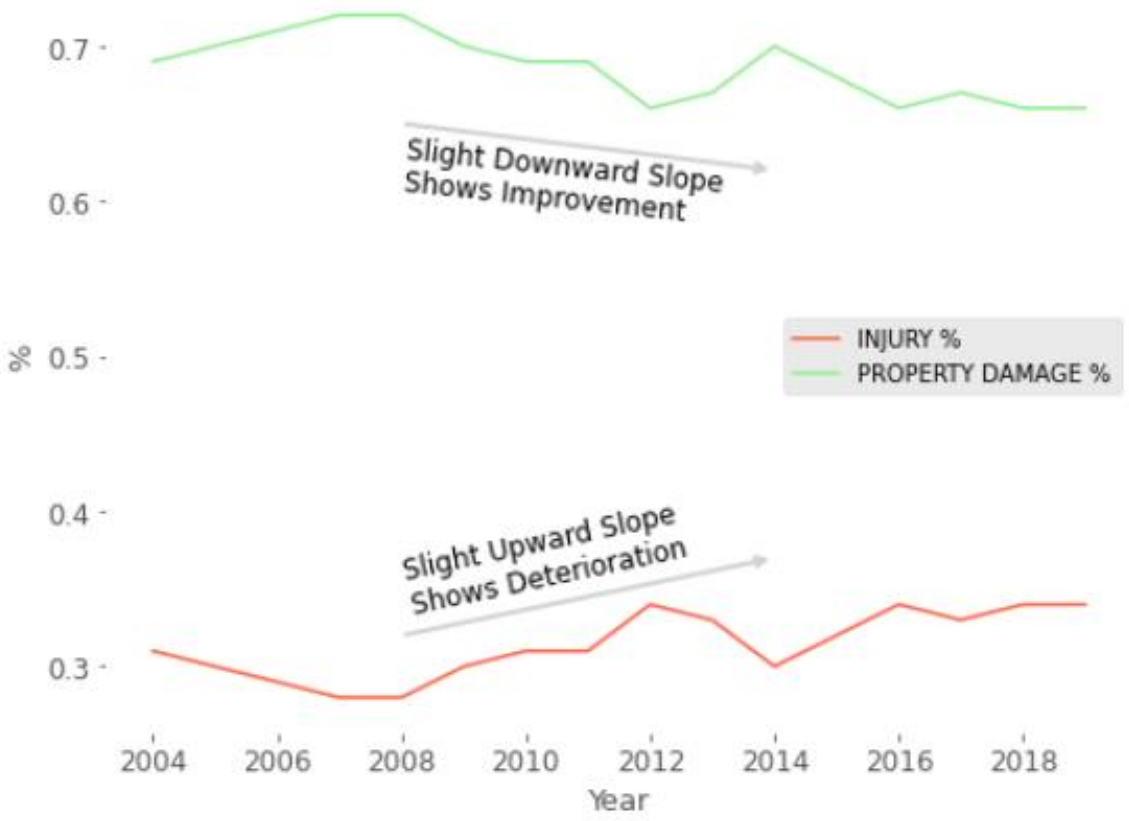
Weight: Bubble chart is hinting at change in weight of "Injury' collisions. But it is not easy to discern. Area chart gives us a better signal. In absolute terms, "Property Damage" collisions dropped by a significant number.

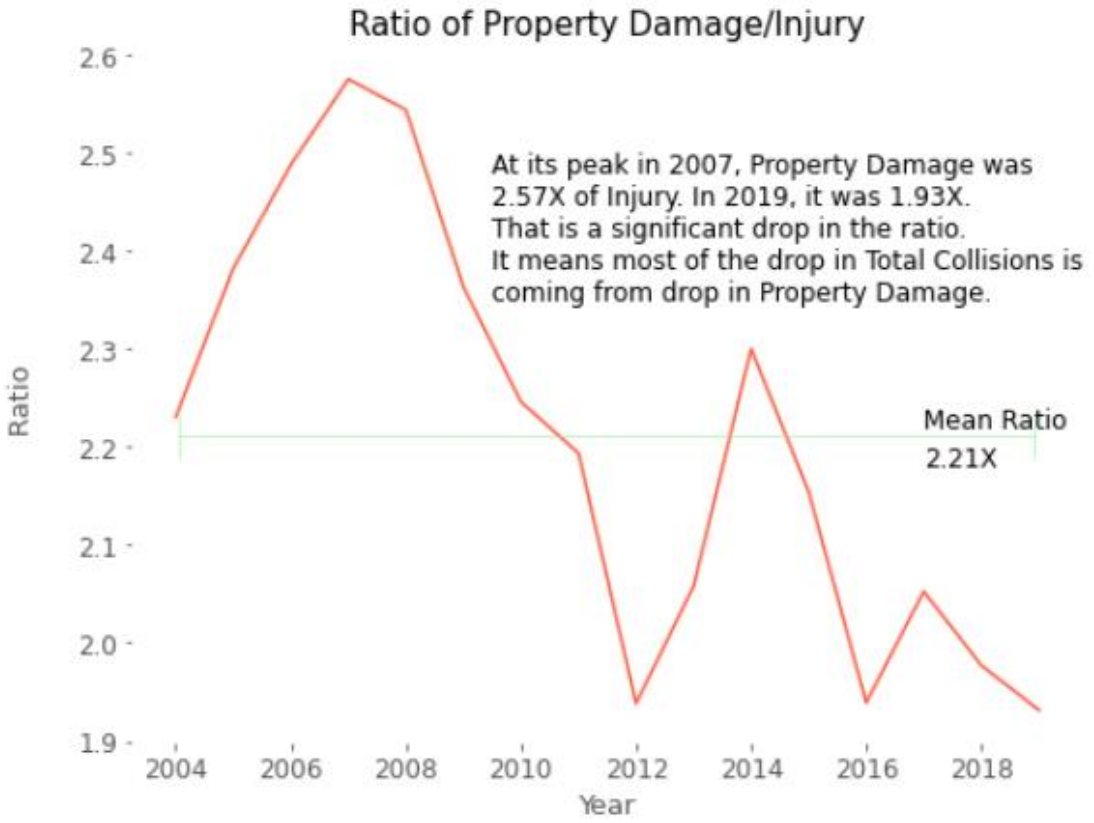




Insight 11 (contd.):

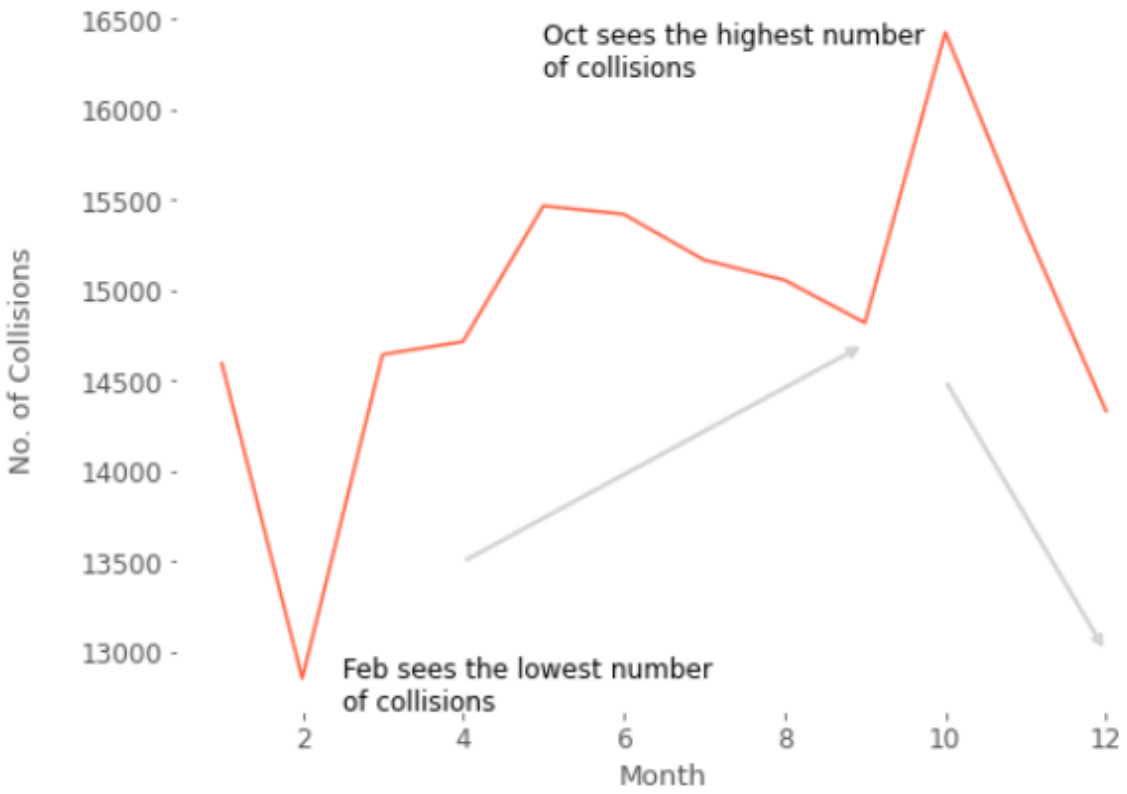
Weight and Ratio chart, makes it clear that drop in "Property Damage" collisions is driving the overall reduction in Total collisions.

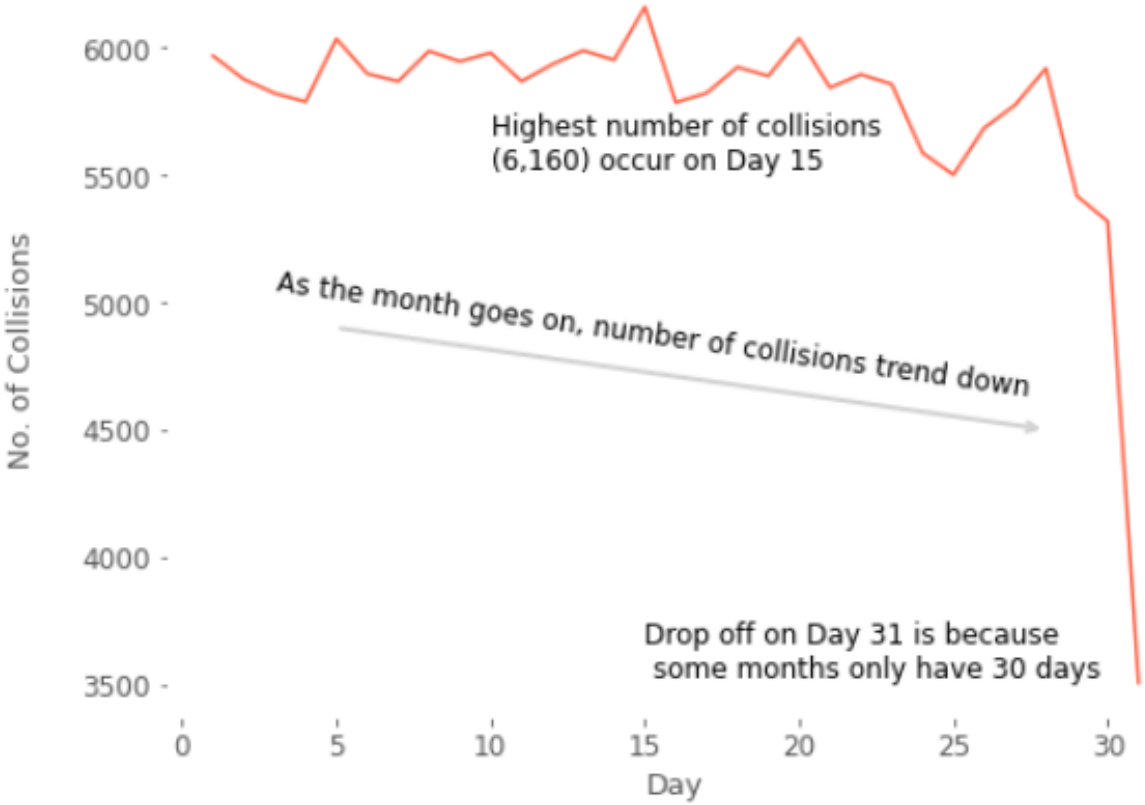




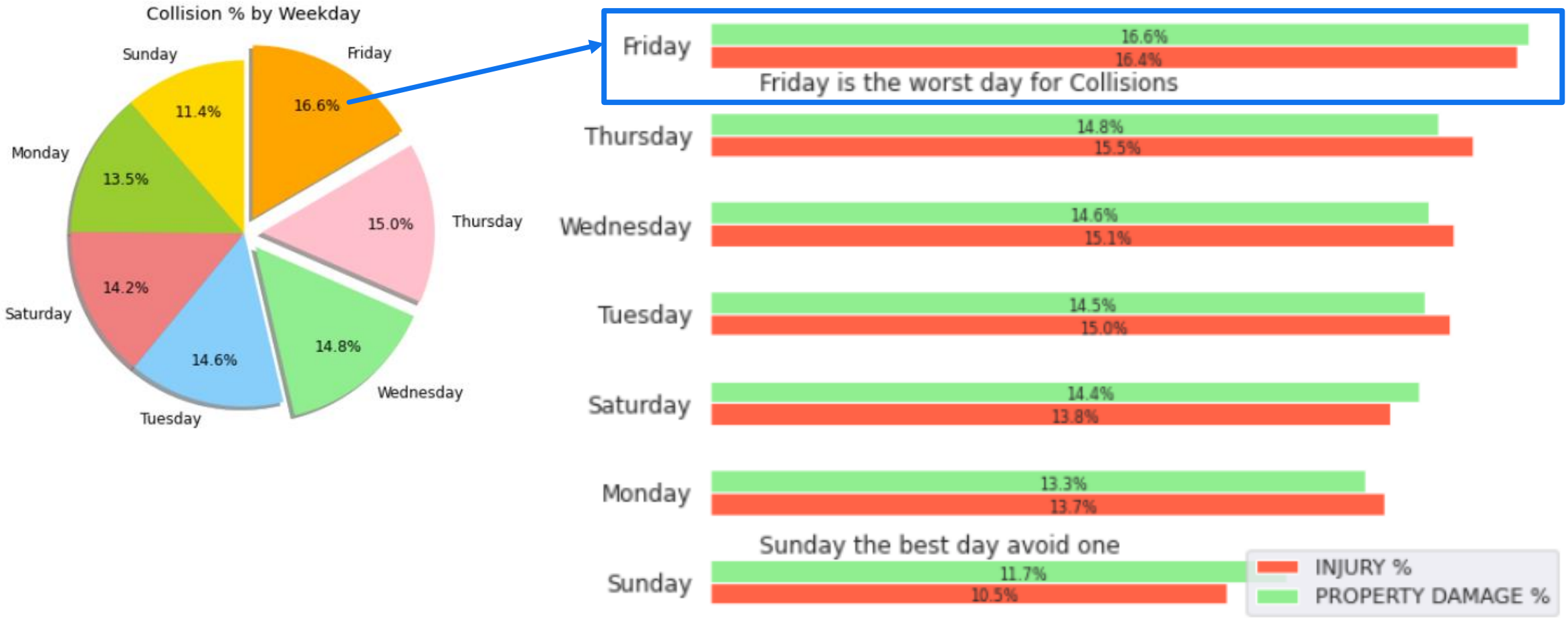
Insight 12:

Trend 2: There is a significant variance in number of collisions depending on the month and day.



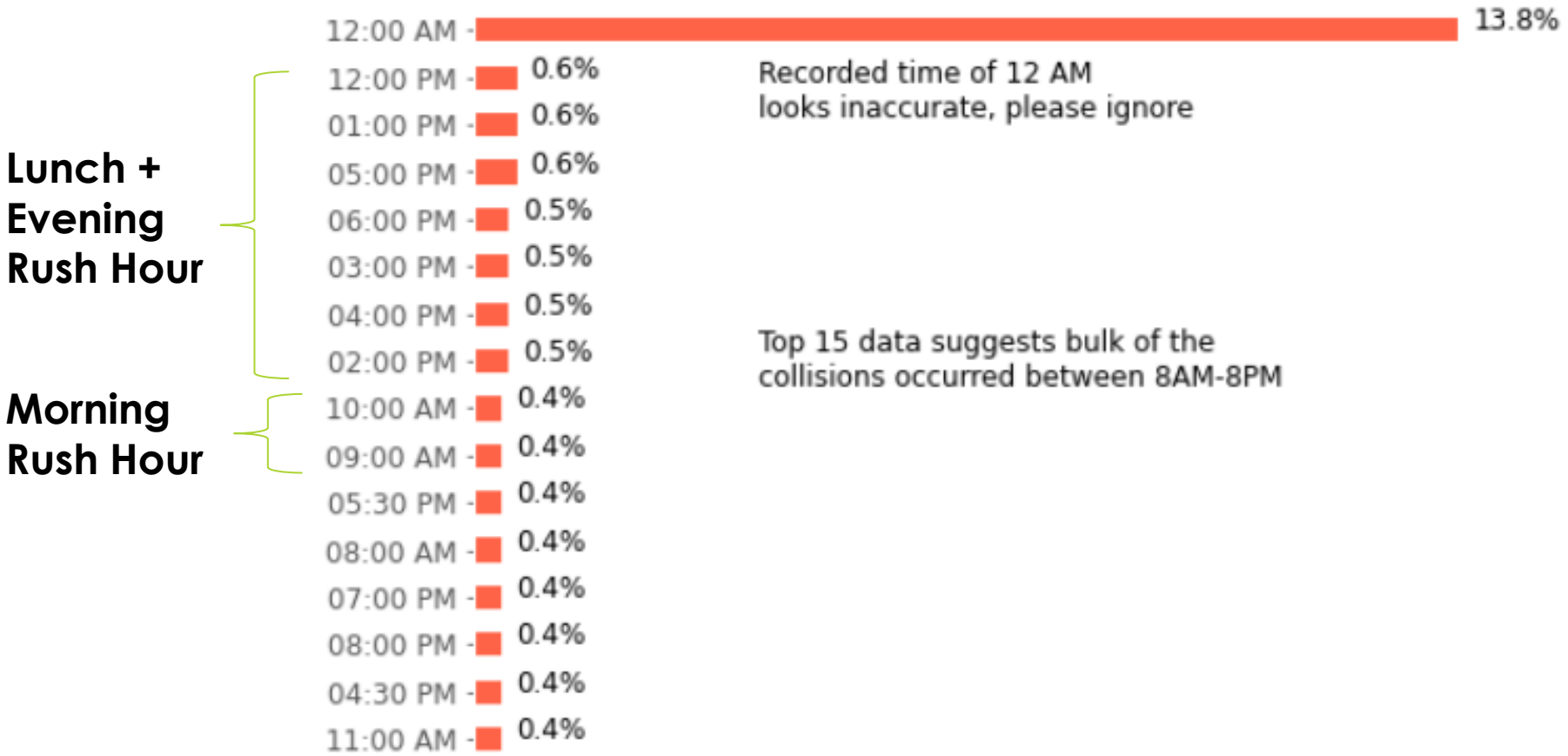


Insight 13:
Trend 3: Friday seems to be worst day for collisions

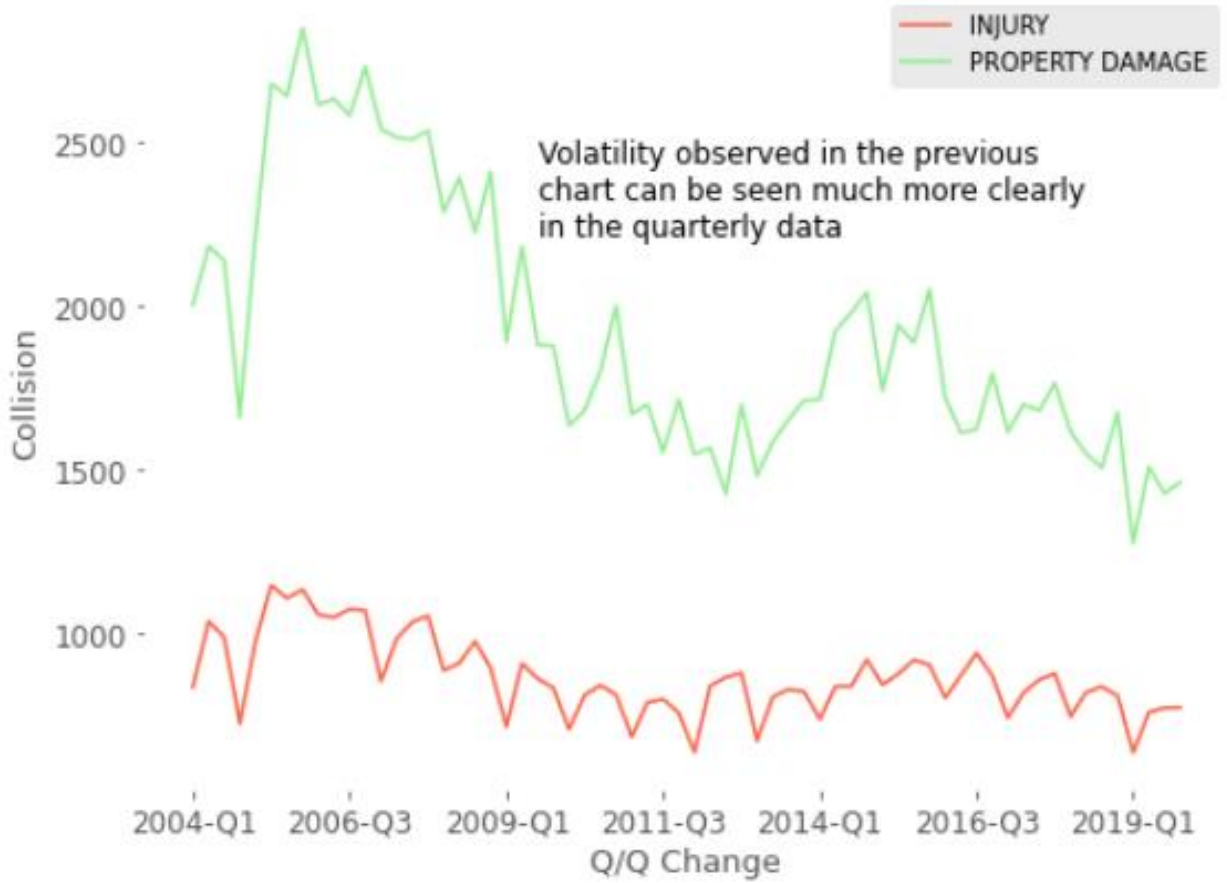


Insight 14:

Top 15 - Most active time slots for Collisions
Lunch time, Evening rush hour and Morning rush hour stands out

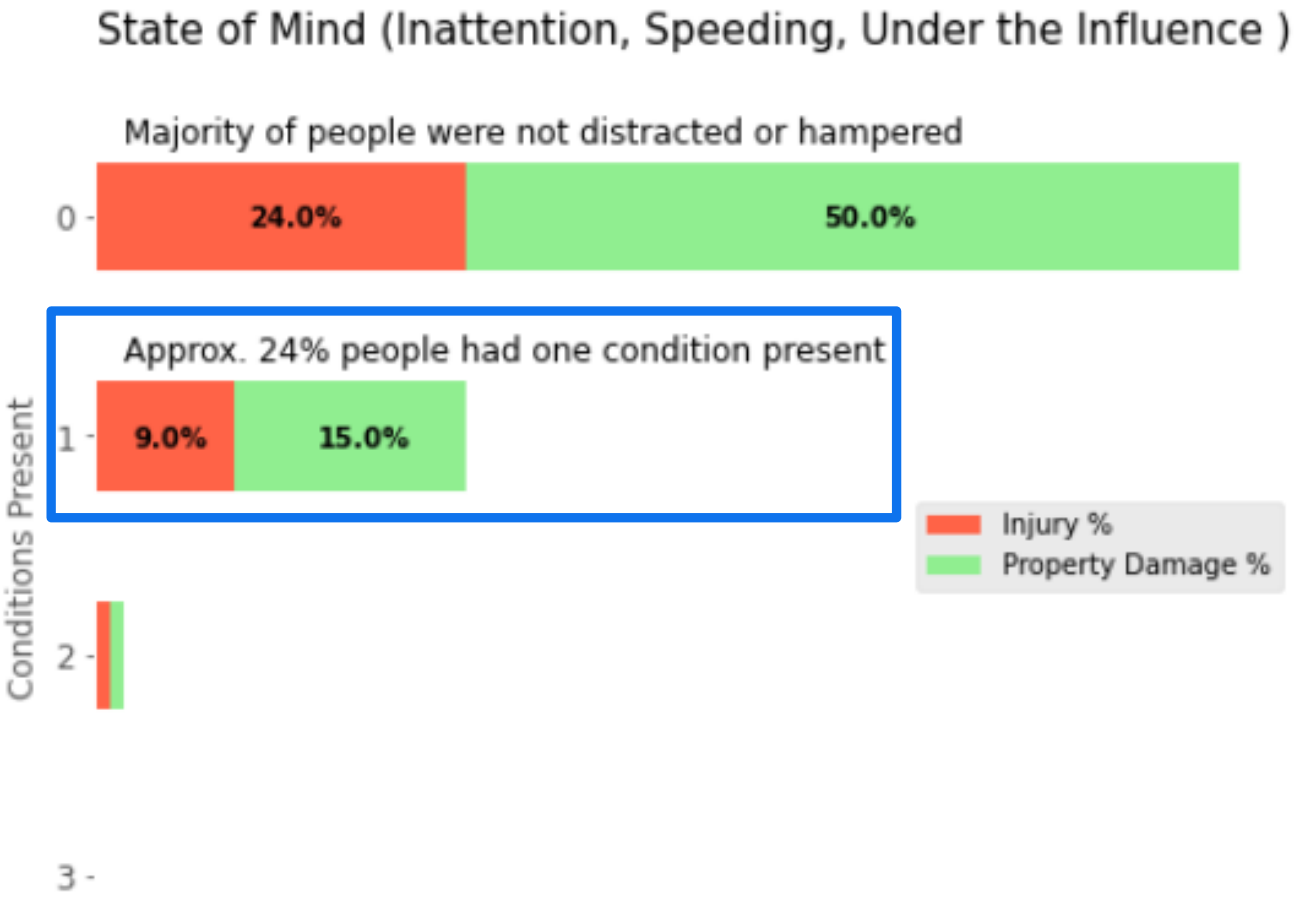


Insight 15:
"Property Damage" collisions are more sensitive to change which means it will be easier to reduce than "Injury"

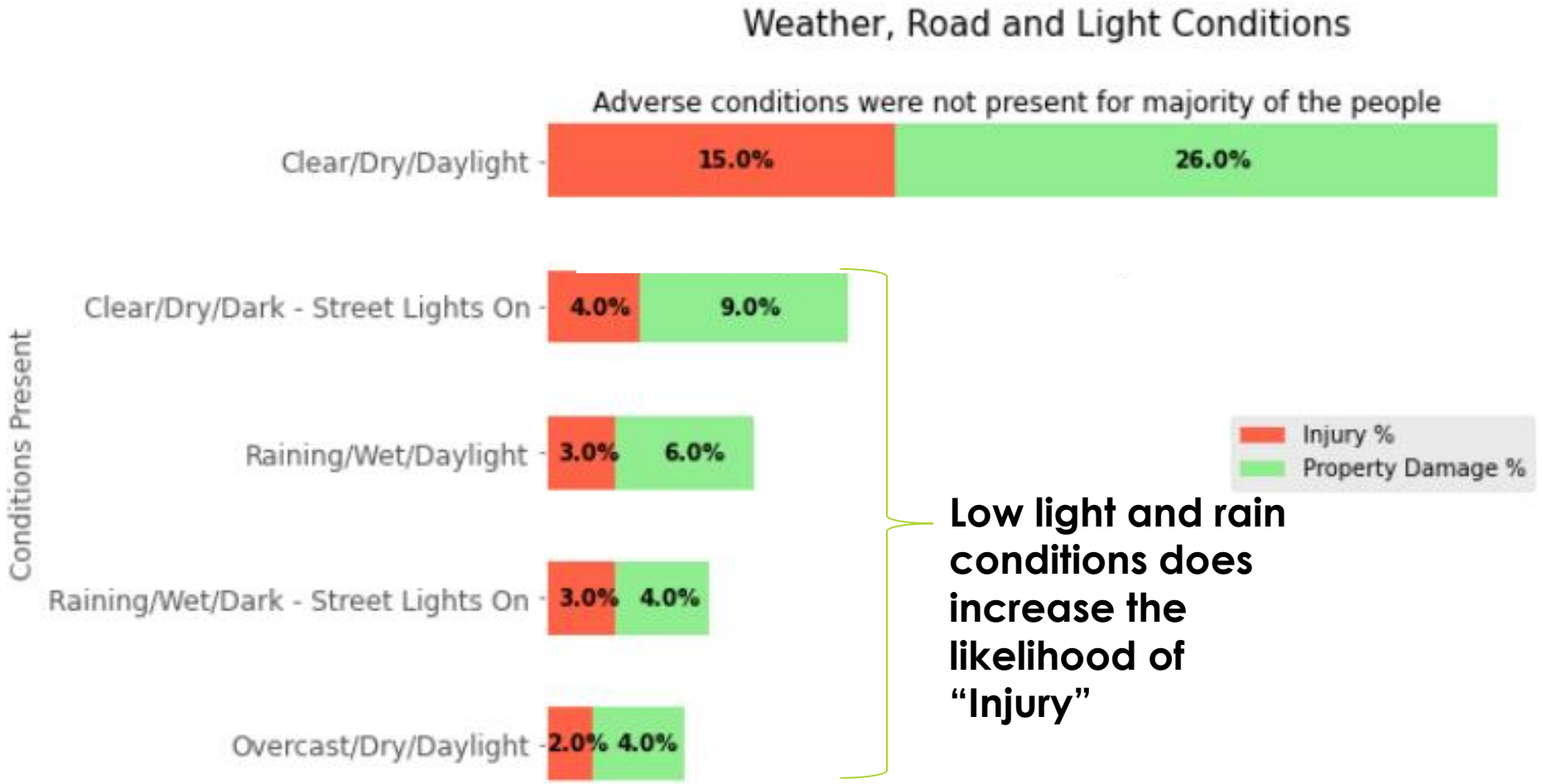


Insight 16:

24% of the Total collisions and 9% of the "Injury" collisions could have been easily avoided if the driver was not impaired

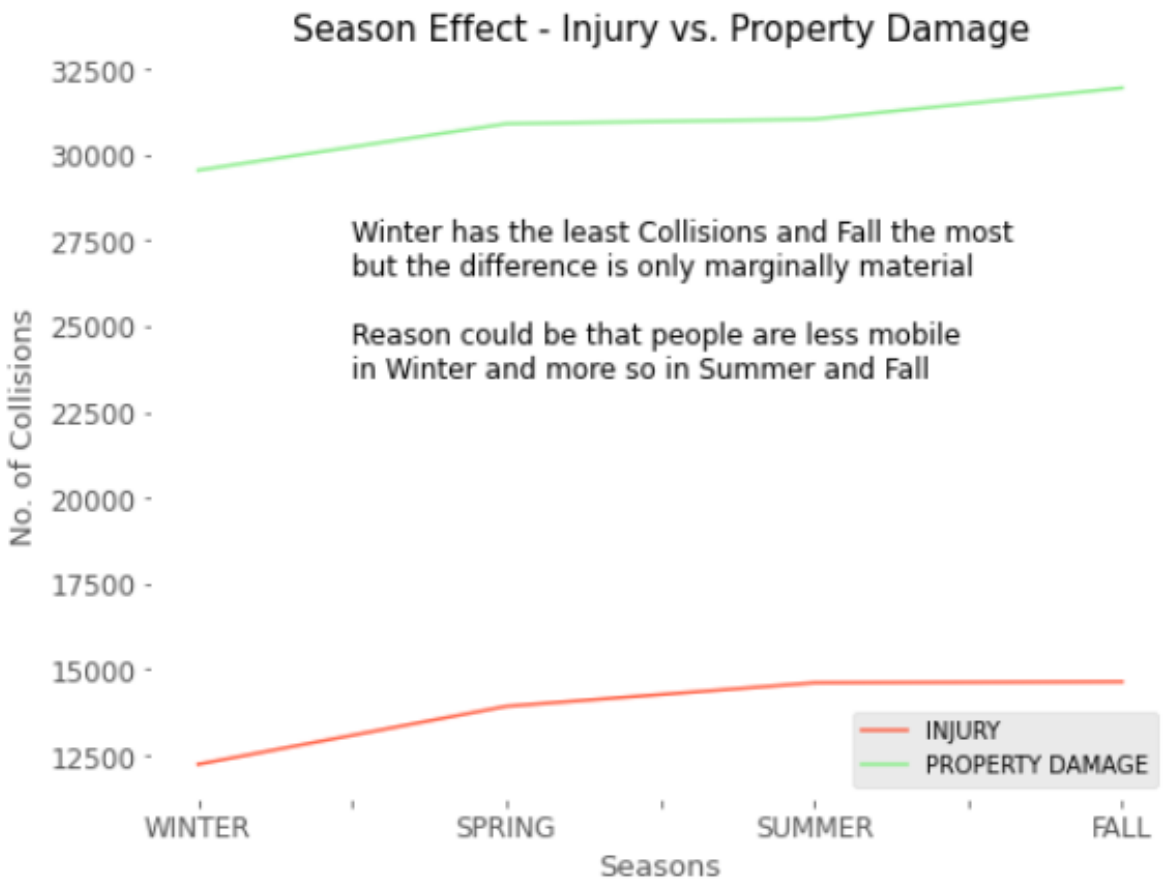


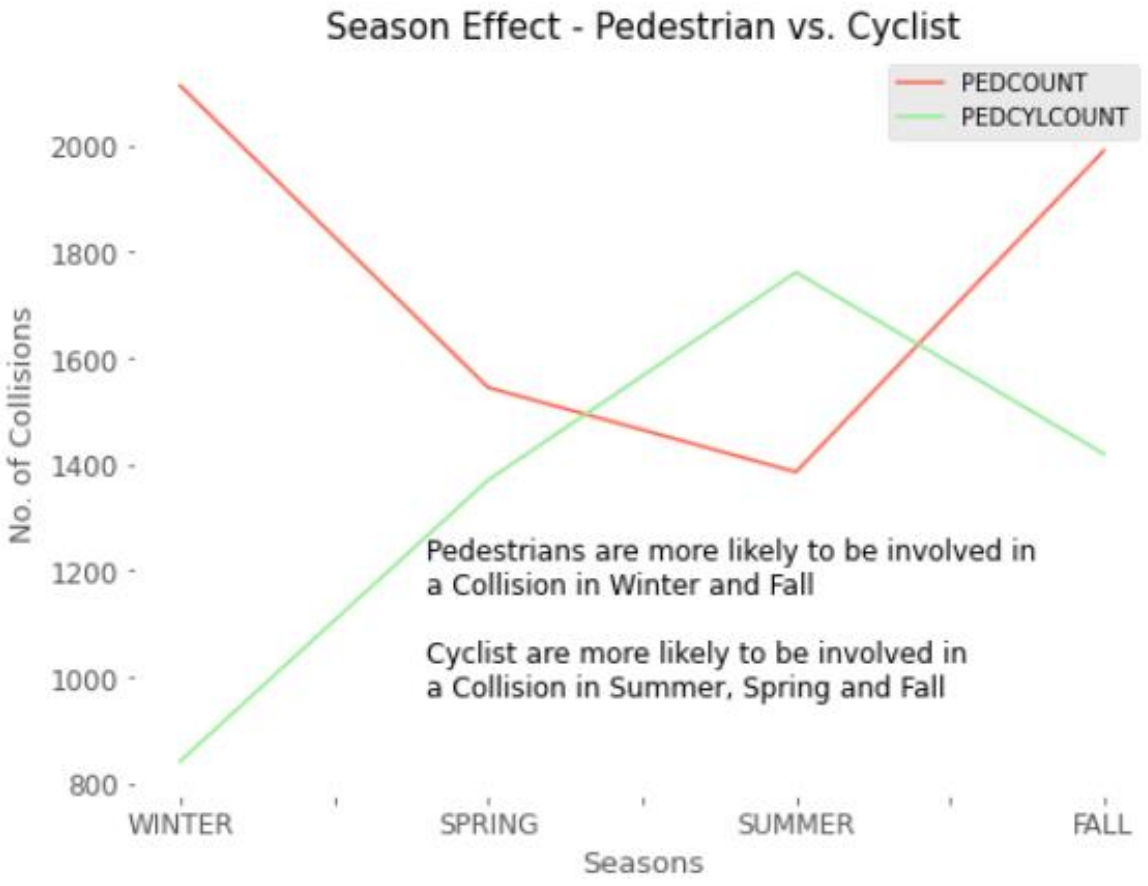
Insight 17:
Rain and low light conditions does play a material role in collisions



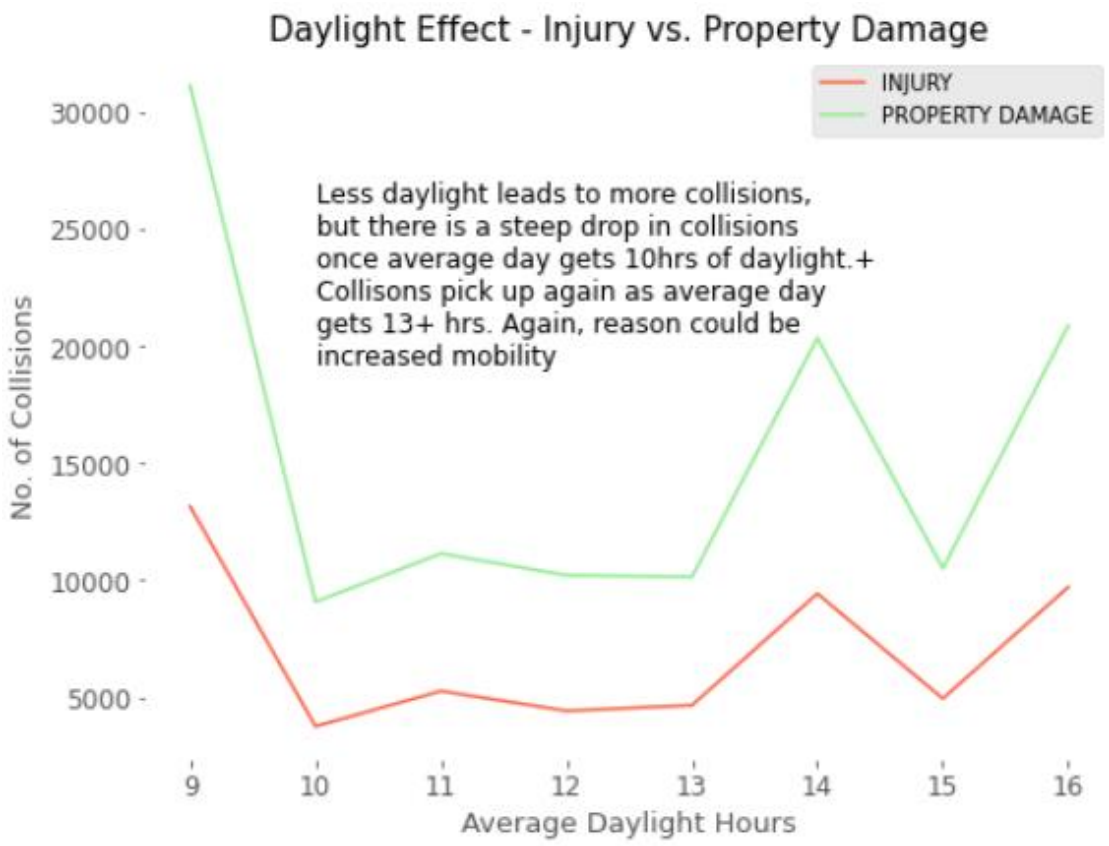
Low light and rain conditions does increase the likelihood of "Injury"

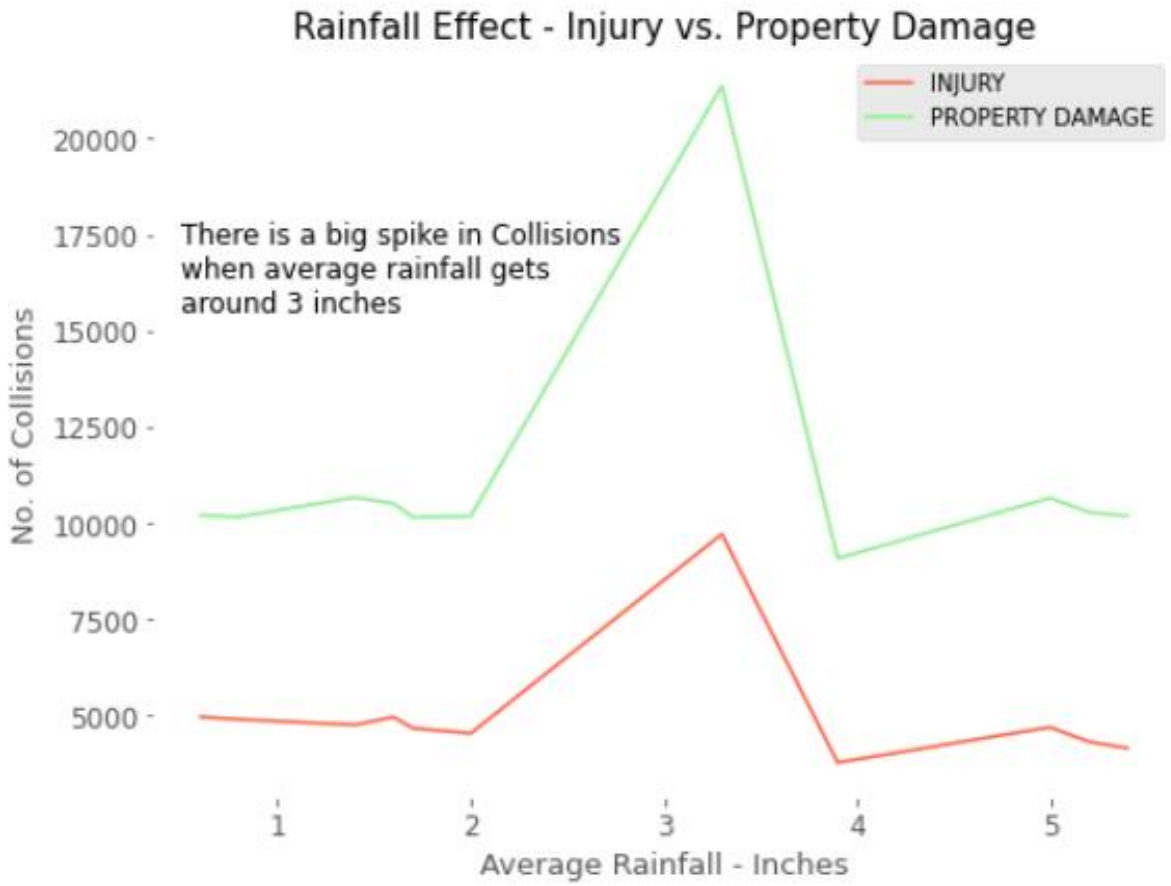
Insight 18:
Effect of Season on the collisions



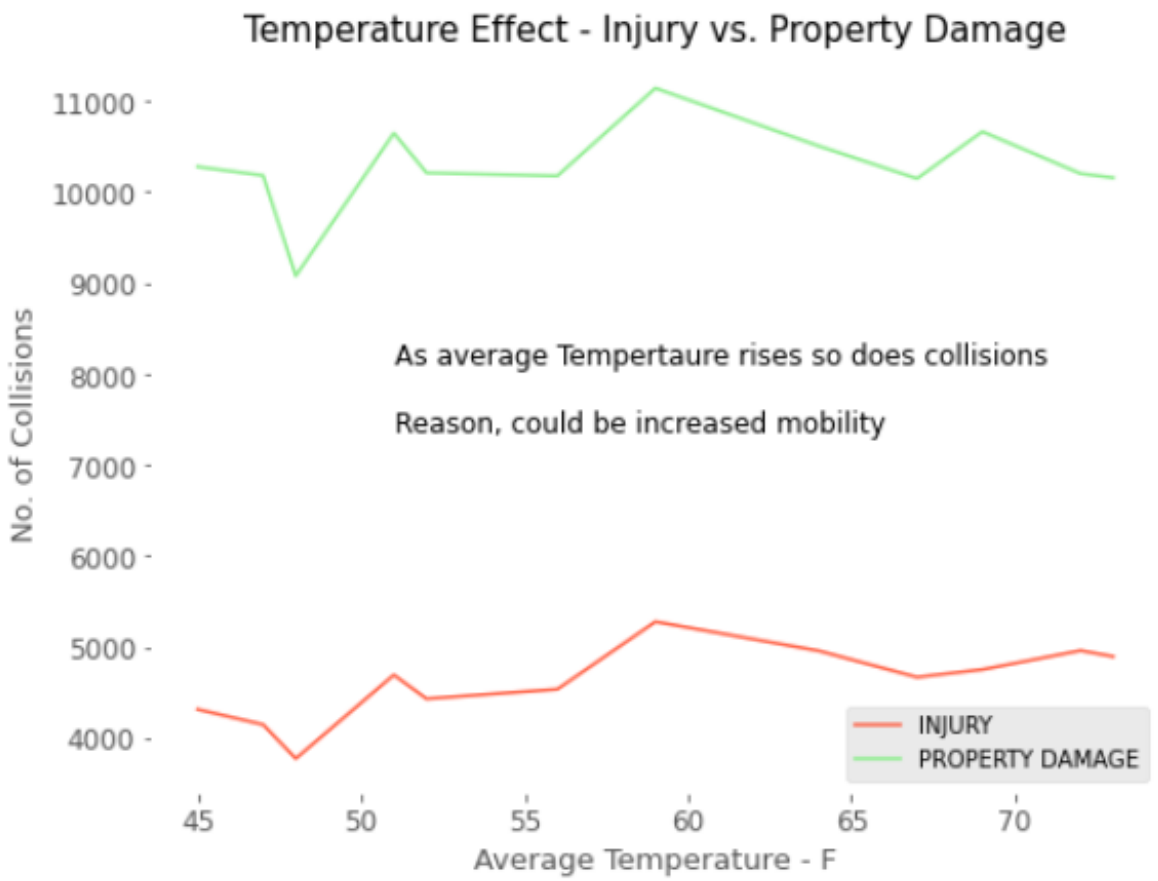


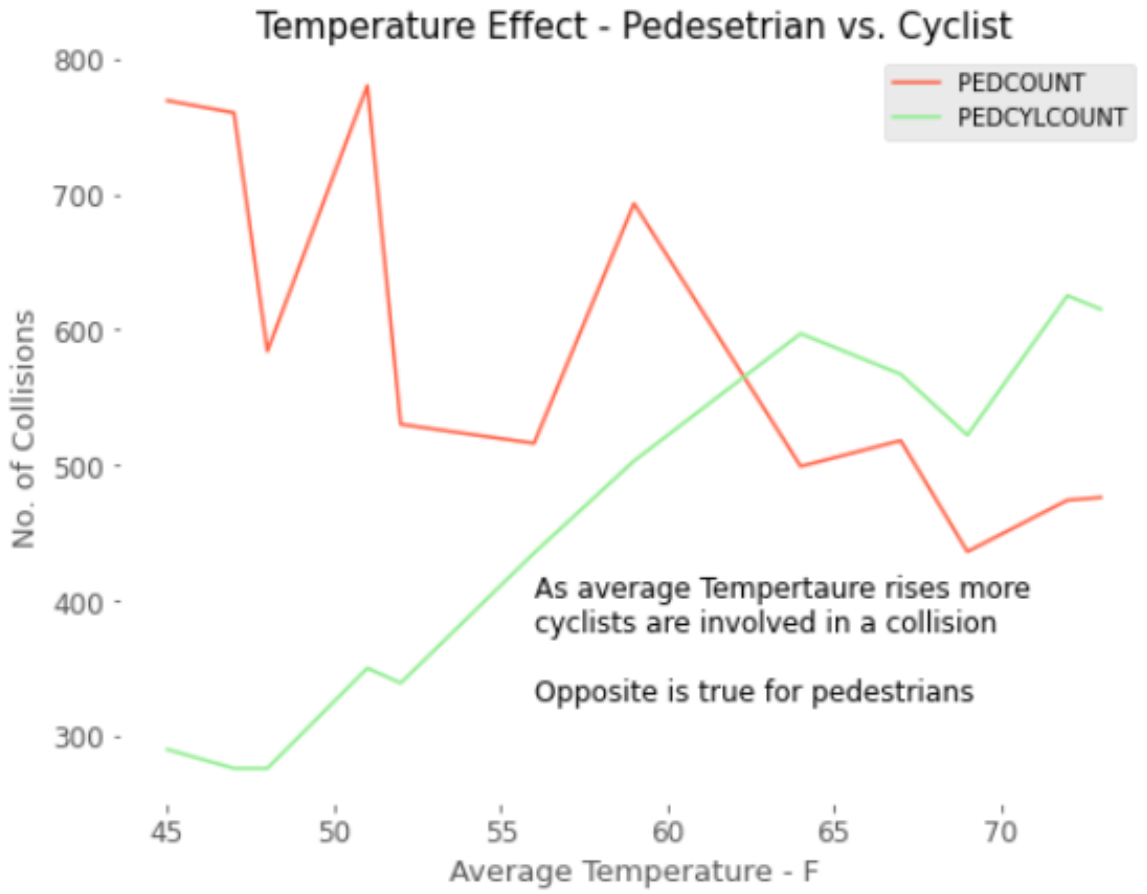
Insight 19:
Effect of Daylight and Rainfall on the collisions





Insight 20:
Effect of Temperature on the collisions





Recommendation:
For Government

- **The "Big Picture" observation is that overall collisions are trending down. The trend is quite material, and it suggests whatever actions government authorities have taken seems to have achieved the desired results. Government should keep investing in these methods. That said, it will be easier for them to reduce Property Damage as compared to Injury collisions. To reduce injuries, government will have to get creative with their methods.**

- **Low Hanging Fruit: 24% of the Total collisions and 9% of the "Injury" collisions could have been easily avoided if the driver was not impaired. Government needs to invest more in education of general public so that residents of the city can take ownership of their own safety.**

- **Most of the collisions seem to be concentrated in certain areas. For example, area in and around Seattle Downtown particularly stands out. I will recommend to prioritize these areas for further on-ground investigation. May be there is some external reason that is causing high rate of collisions in these areas which is not getting picked up in data.**

Recommendation:
For Government (contd.)

- **To reduce "Property Damage" collisions, increase Parking Enforcement in Top 15 "Property Damage" collision areas.**

- **To reduce "Injury" collisions, reduce speed limit and increase traffic police presence for Top 15 "Injury" collision areas.**

- **In general, increase Parking Enforcement and Traffic Surveillance:**
  - **Daily between 12PM-6PM**
  - **On Thursday and Friday**
  - **Between Day 1-15**
  - **September, October and November**
  - **Low light or rainy conditions**

- **Have a dedicated LEFT turn signal for most intersections in high density areas.**

Recommendation:
For Drivers

Drivers can decrease the probability of collision if they pay **extra attention** in following situations:

- **During Parking**
- **At an intersection or signal**
- **When making a LEFT turn**
- **Dark, low light or rainy conditions**
- **High density neighborhoods such as Seattle Downtown**
- **Around pedestrians and cyclists**
- **Thursday and Friday**
- **Month of October, May and June**
- **Between 12PM-6PM**

Recommendation:
For Pedestrians

Pedestrians can decrease the probability of collision if they pay **extra attention** in following situations:

- **At an intersection or signal**
- **Winter and Fall (or low temperature conditions)**
- **Dark or low light conditions**
- **High density neighborhoods such as Seattle Downtown**
- **Month of November, December and January**

Recommendation:
For Cyclists

Cyclists can decrease the probability of collision if they pay **extra attention** in following situations:

- **At an intersection or signal**
- **Summer (especially on days when temperature is high)**
- **Month of May, July and August**

Thank You! ☺

QUESTIONS/COMMENTS/FEEDBACK:

E-MAIL: VGREWAL.TECH@GMAIL.COM

LinkedIn | GITHUB