

Stormy Streets: Decoding NYC's Weather-Crash Connection

Yash Gupta

Statistics and Data Science | University of Texas at Austin

Results

I created a couple of visualizations/graphs to draw upon, mostly bar and line charts as I was conducting both comparison analyses and trend analyses.

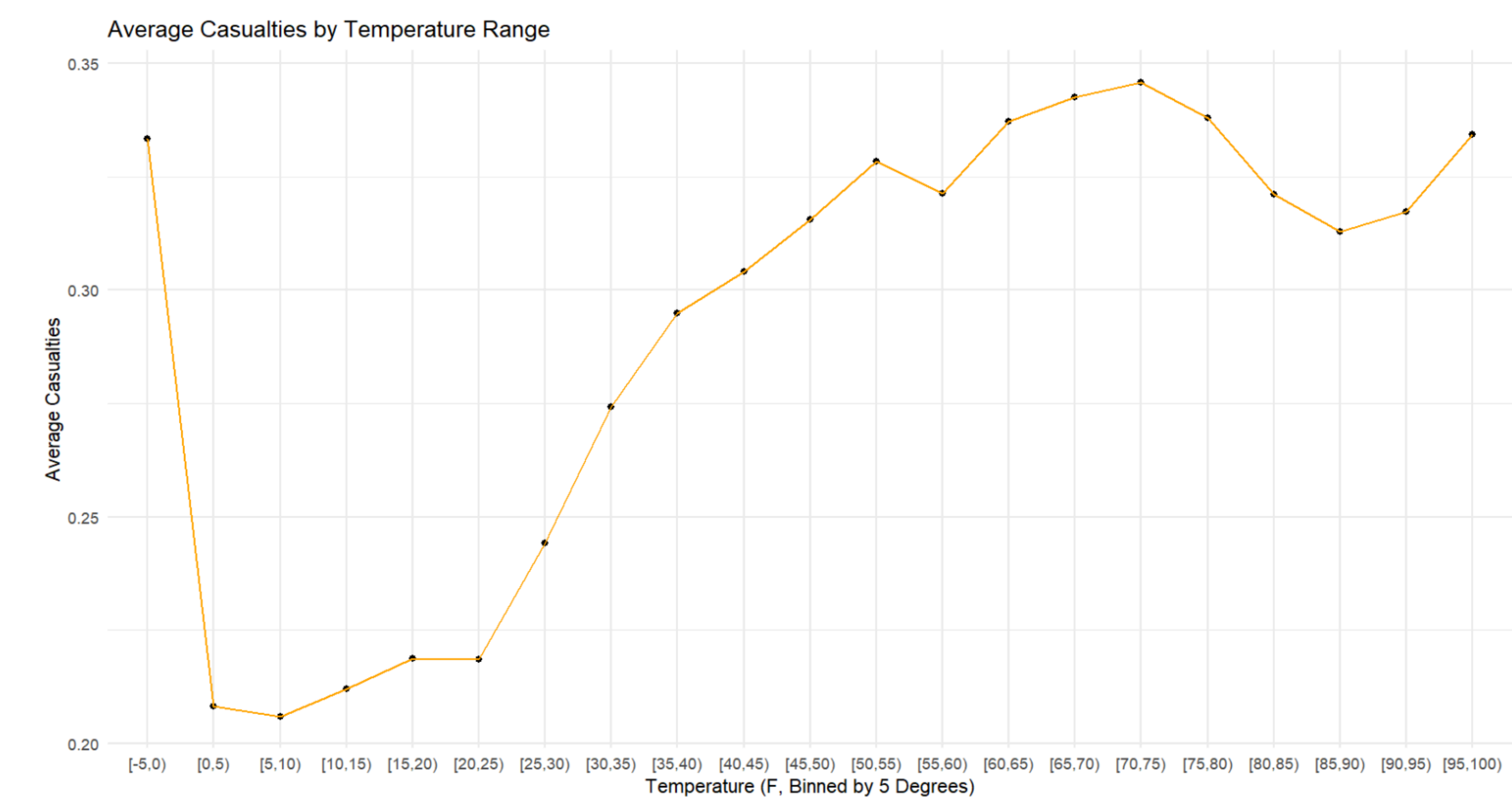
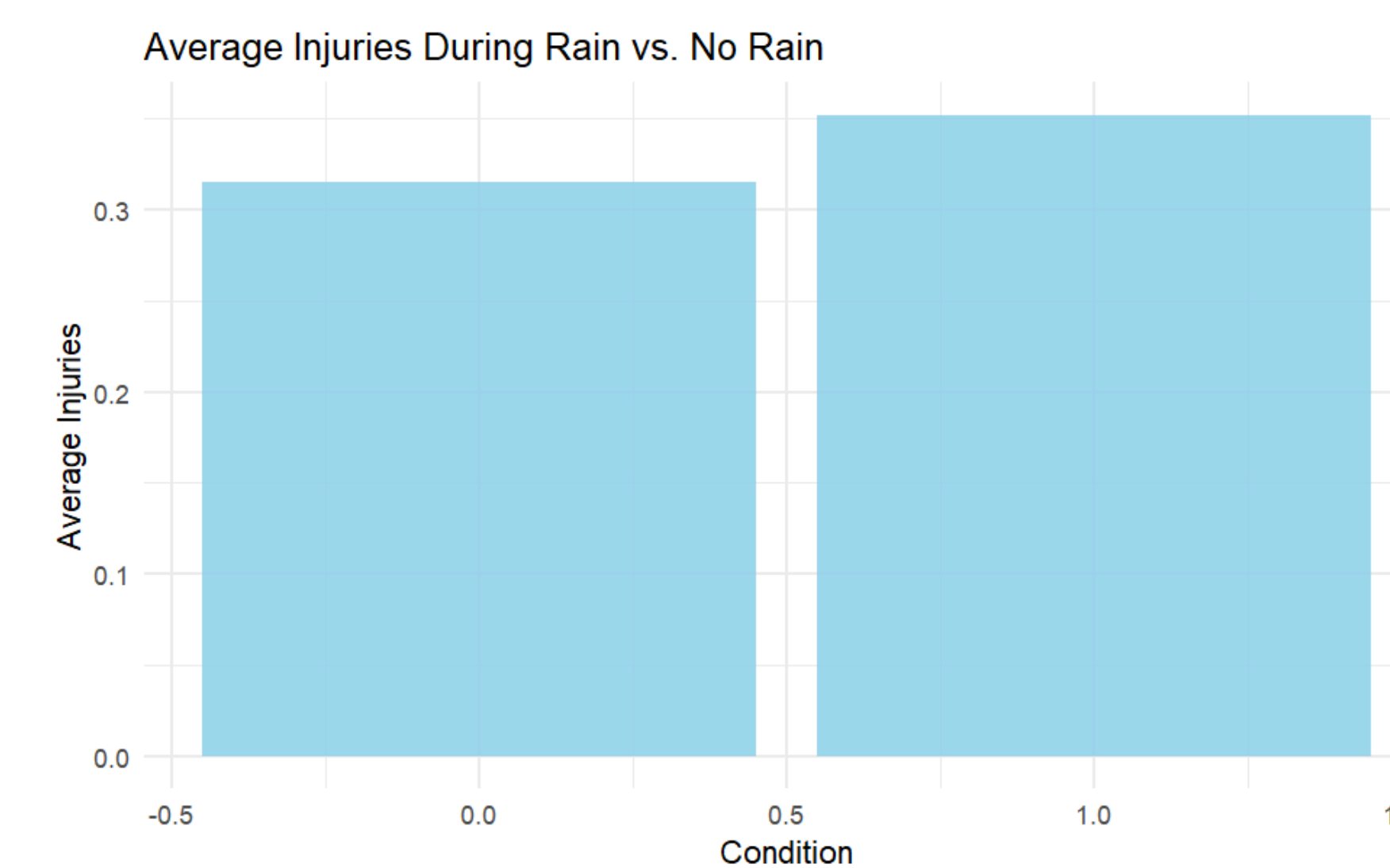
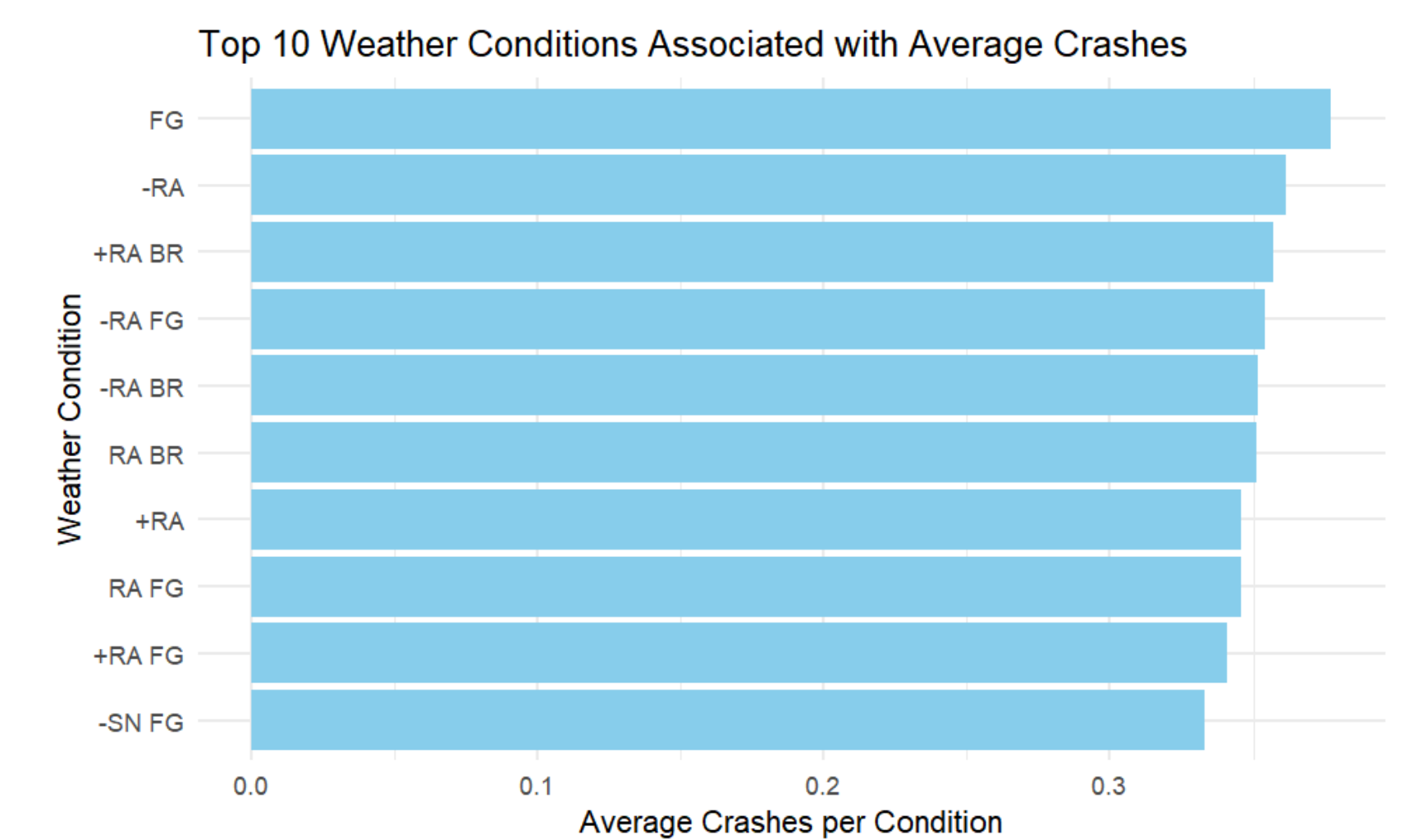
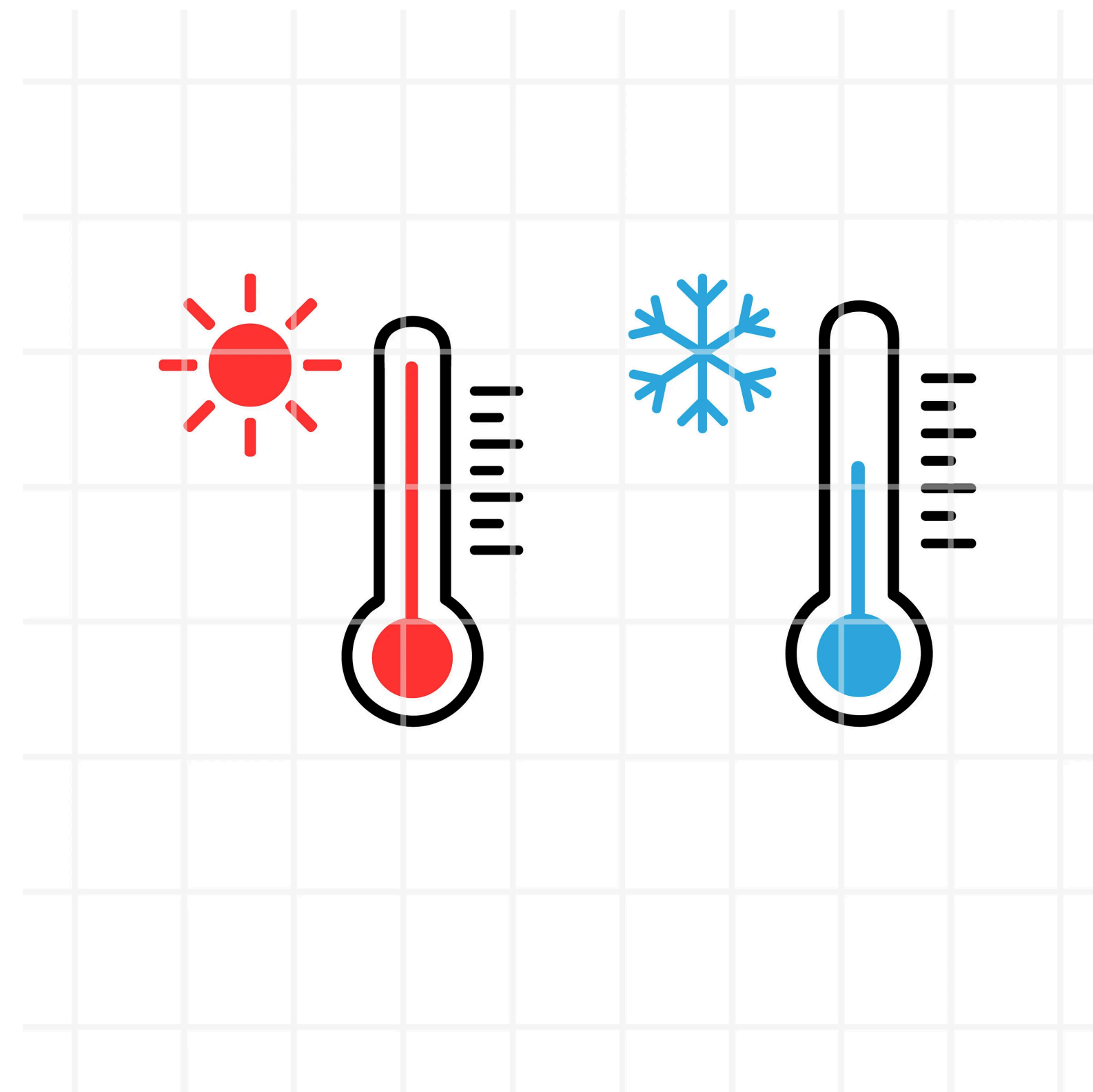
We can identify that conditions such as Fog (FG), Rain (RA) and Mist (BR) lead to a higher incidence of crashes. We can additionally analyze that while rain is occurring, the rate of injury is slightly higher given that there is a crash at about 4% of a difference. A graph tracking the affect on temperature indicates that temperature does not tend to have too high of an effect on the amount of casualties unless it falls in a very low temperature. Other factors I measured tended to have little to no effect on crashes, such as visibility, wind speed, etc.

Weather's role in crashes

Through analyzing the NYC Open Data Crashes dataset, I decided to look and see if I could find correlations with weather conditions at the time in order to merge with my data. This could be useful in helping to identify what conditions lead to more crashes, including those resulting in casualties (injuries and fatalities), and guidance could be built around those conditions.



Weather clip-art



3 visualizations/graphs of different variables for crashes

Data Collection and Cleaning

I utilized the R programming language and RStudio to create my analyses. I first captured data from the NYCOpenData set for crashes in NYC and IOWA's IEM weather station in NYC as well, and I then merged the datasets. This was done through first using a full join to keep my data synchronized through time and date (to the nearest hour). I then conducted some data cleaning through dropping NA values, and recategorizing missing data, which then led into me conducting experiments, through creating graphs, tables, and charts to determine what factors lead to a higher crash incidence rate, especially for injuries/fatalities.



Logos for data/software used

Conclusions/Recommendations

Through this analysis, factors such as rain and fog tend to lead to higher incidence of crashes/casualties.

Some key changes that could be made to the NY DoT / FHA for Vulnerable Road users would be to

- Install enhanced lighting systems for fog/rain conditions
- Weather-responsive traffic management systems
- Increasing following distance to 8-10 seconds in poor weather for vehicles, and reducing speed by 1/3 on wet roads and 1/2 on snowy/icy services
- Mandating use of low beams in fog, along with headlights in precipitation.

Acknowledgments

Thanks to Emily Rothenberg from the Columbia University Data Science Institute for providing me with this opportunity and introducing me to the TDSP.

References (MLA 9)

- (NYPD), Police Department. "Motor Vehicle Collisions - Crashes: NYC Open Data." Motor Vehicle Collisions - Crashes | NYC Open Data, 17 Jan. 2025, data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95/about_data.
- Akrherz@iastate.edu, Daryl Herzmann. "IEM:: Download Daily Summary Data." Iowa Environmental Mesonet, mesonet.agron.iastate.edu/request/daily.phtml. Accessed 20 Jan. 2025.
- Strauss, Jon. "Navigating Smart Roads: Safeguarding Vulnerable Road Users through Technology." Navigating Smart Roads: Safeguarding Vulnerable Road Users Through Technology | FHWA, highways.dot.gov/public-roads/spring-2024/innovation. Accessed 20 Jan. 2025.