**RUT-VIRT-DATA-PT-04-2023-U-LOLC  
PROJECT PROPOSAL** – module 7 project 1

**GROUP 7**  
Abraham, Susan J  
Glantz, Adam  
Kabir, Debbie  
Mysllinj, Andi

1. **RESEARCH QUESTIONS**

Most significant vehicular crashes in New York City during the period of study are caused by *inattentive driving*. We propose to investigate how this phenomenon interacts with other variables:

* What is the relative percentage of significant crashes caused by inattentive driving vs. other causes? Adam Barplot, data cleaning, come up with a minimum threshold or “top 10” to includeDoes this relationship hold over time of day: e.g., does the same general percentage of crashes caused by inattentive driving hold during the weekday rush hours? (8-9am and 5-6pm EST). TBD (team): Bin the data into rush hours vs non, scatter plot over time,
* Do crashes caused by inattentive driving also result in the most fatalities, the most injuries, or both? Andi
* Debbie – presentation look and feel

1. **IMPORTANCE OF THE STUDY**

* It is uncontroversial and in everyone’s interest to understand when and why vehicular crashes happen in the hope of eliminating them.
* Analysis like this may be used as an input to “train” self-driving vehicles in the future.
* New York City has particular relevance as the largest city in the United States, as well as the closest major city to Rutgers University’s main campus

1. **SOURCES**

* The City of New York provides a JSON on vehicular accidents that result in an injury, death, or at least $1,000 in damage here: <https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95>. It can also be downloaded as a csv file. At this writing, the data is current within the last 24 hours.
* The data was collected by the New York City Police Department and was first made available in 2014.