1. **Introduction:** At the beginning of your paper, you must describe the data, in a paragraph. Note the following, and any similar information:

* What is the source of the data? Where and when was it created?
* If it is a sample, from what population was it drawn, and how was the sample selected?
* Do you suspect any sampling bias?
* Was it an experiment or an observational study?
* How were measurements taken, or questions asked?
* Do you suspect any bias in the questions or measurements?
* Why is this data of interest to you, and why should the class find it interesting?
* What kind of data cleaning was necessary (R code for this must show…)

The data set we are analyzing today is based on data from the Burlington Fire Department. This data includes info on incident types and response times from Aug 1, 2018, through November 30, 2019. Data was pulled from the fire department RMS database. There is no sample bias as the data is strictly incident responses of the fire department. This data is an observational study to see the time taken from call to response of the fire department and locations of incidents. The only bias that is seen is possibility in rounding times to the nearest minute, but this is understandable and would not affect the results.

We believe this data could be very interesting to understand for what reasons the population of Burlington most frequently call the fire department, as well as, the efficiency and speed that they have. Because this city has a large population of students we are also intrigued to see if the large influx of college students during the fall and spring semester change the data, for example: the reasons for the calls and the times these emergencies occur. Another factor we believe could strong affect our data would be the winter season, as Vermont is known for being very cold, this could change the reasons to why the fire department was called.

To be able to use this data set we had to do some cleaning: the most important part was to change the names of our variables due to them being very long and incomprehensible. We also modified the geopoint column by separating and creating two new variables, Longitude and Latitude. This was done to be able to use Leaflet to plot coordinates.

1. **Data Analysis:** Write R code to create some relevant graphs, using techniques that we’ve used in class (ggplot, maybe dplyr). About **4 or 5 graphs** should be plenty, depending on complexity. Include some numerical summaries as well. If possible and appropriate, include **a bootstrap confidence interval.**

For each graph and numerical summary, write a paragraph or two summarizing what you see, and suggesting some implications. For example, describe patterns that you observe in a graph, and suggest why they make sense, given what you know about the subject, or if they are unexpected. Do you think there is a cause-effect relationship between any variables? Explain your reasoning.

1. **Conclusions:** Write some overall conclusions – an overall summary of what you learned from your analysis. Summarize in one paragraph.
2. **Limitations / Recommendations:** Write a paragraph describing some of the limitations that are inherent in your study. Also discuss ideas for future research that might build on the work you did in this project. Summarize in one paragraph.