**Group seven midterm project**

SMART question

**Question** – Within our data set of 5,181 observations of police shootings from 2015 to 2020 in the United States, is there a correlation between the U.S. state of observation and whether or not a body camera was **on** during the shooting?

**S** – While this data set has a lot more information than simply the U.S. state where the shooting took place and whether the police body camera was on or off, we are only looking only at these variables to maintain specificity. The EDA will only seek to understand whether there is a correlation between these two variables.

**M** – We will use basic statistical analysis to identify whether there is a correlation between the two variables. Correlation is easily measurable with data reshaping.

**A** – This EDA requires only basic data analysis in R. The amount of data is manageable with only 5,181 observations spanning over a five year period. We also believe finding a correlation within this data set is achievable given the differences in state policy on the use of police body cameras.

**R** – Studying the use of body cameras in police work is an important topic of study for data-driven policy research in the United States. We may be able to apply this correlation between the U.S. state of observation and whether the body camera was on or off during the shooting to state policy on body cameras during police work.

**T** – This project, given that there are three people working collaboratively, will require a couple of weeks to perform the data analysis and another couple of weeks to write the report. This corresponds well the time constraints of the midterm project.