**Methodology**

**Objectives**

Through this project, we sought to better understand three overarching themes related to the ongoing civic unrest at universities across the United States. Those themes include:

* Differences in how law enforcement treat pro-Palestine student activists at public school and private schools.
* Regional differences in law enforcement responses to pro-Palestine protests.
* New York and, more crucially, its universities place in all of this, as the media has portrayed New York and its major cities as the epicenter of protests and arrests.

**Methodology**

To establish the number of arrests and detainments, we used a dataset publicly available on the New York Times website. We limited our data from April 18 to May 18.

To normalize the arrest data, we established how large the crowds were at each campus when the arrests took place.

We attempted to find three estimates of crowd sizes as described in print/digital news articles, broadcast clips, audio clips. Additionally, we sought to only include crowd size estimates of protests that resulted in arrests.

In cases where this information was not available using the mediums above, we found the data via social media posts published by reputable sources with photos/videos that could also illustrate the number being described.

Where only clips using the word “hundreds” could be found, that qualitative measure was replaced with the median between 100 and 999, which equals 549.5.

Where only clips using a range of numbers (such as 45-50) could be found, the median between those two numbers was used.

In any case where a “crowd size” estimate could not be found, that data was left blank in the cleaned excel sheet.

Where no quantitative or qualitative data could be found about the number of protesters at a scene prior to arrests, those universities were left out of the cleaned dataset.

We then gauged how many protesters were arrested as compared to the actual crowd sizes at the time of the police response. By using pandas, we created an additional column of data. This additional column, denoted as PCT-ARREST in the jupyter notebook, stores the number of arrests/detainments divided by the average crowd size for each individual university.

Pandas was also used to interrogate the data on a series of questions, which included:

* Total number of arrests by university type (public or private)?
* Percentage of protesters arrested as compared to crowd size, by university type (public or private)?
* Which universities saw the most arrests?
* Which schools arrested the highest percentage of protesters, based on the ratio of arrests to crowd size?
* Which schools arrested the smallest percentage of crowd members?

A series of graphs were made using Datawrapper and Rawgraphs.

**Discussion/Conclusion**

We were unable to find at least three credible estimates of crowd sizes for many of the universities where arrests have taken place since April 18. As a result, the average crowd size (denoted as “AVG-CROWD” in the excel sheet and jupyter notebook) is not as accurate or consistent across all university entries.

There are also other factors unrelated to the data itself that, without context, could lead to a misrepresentation of the issue.

For example, some of the schools with high arrests-to-crowd ratios experienced early morning encampment sweeps. Of course, it’d be difficult for students on the ground to attempt to mobilize hundreds of outside protesters on such short notice.

One outlier is the University of Mary Washington because the ratio was 100 (the entire crowd was arrested), this would be a matter that would require more reporting.

In addition, the legal parameters that govern a demonstrator at an encampment are much different than those that govern a student who participates in more traditional, “acceptable” forms of activism, like marches or walk-outs. In other words, it may be easier, legally, to arrest someone for setting up an encampment than it is when that person participates in other forms of protest.

Nevertheless, we found that public universities have arrested more students – with and without data normalization. We also found that New York City universities’ arrest-to-crowd ratios are not that significant – in fact, smaller cities with lesser known schools have been arresting students at higher rates. But again, for the reasons listed above, this insight should be looked at with caution.