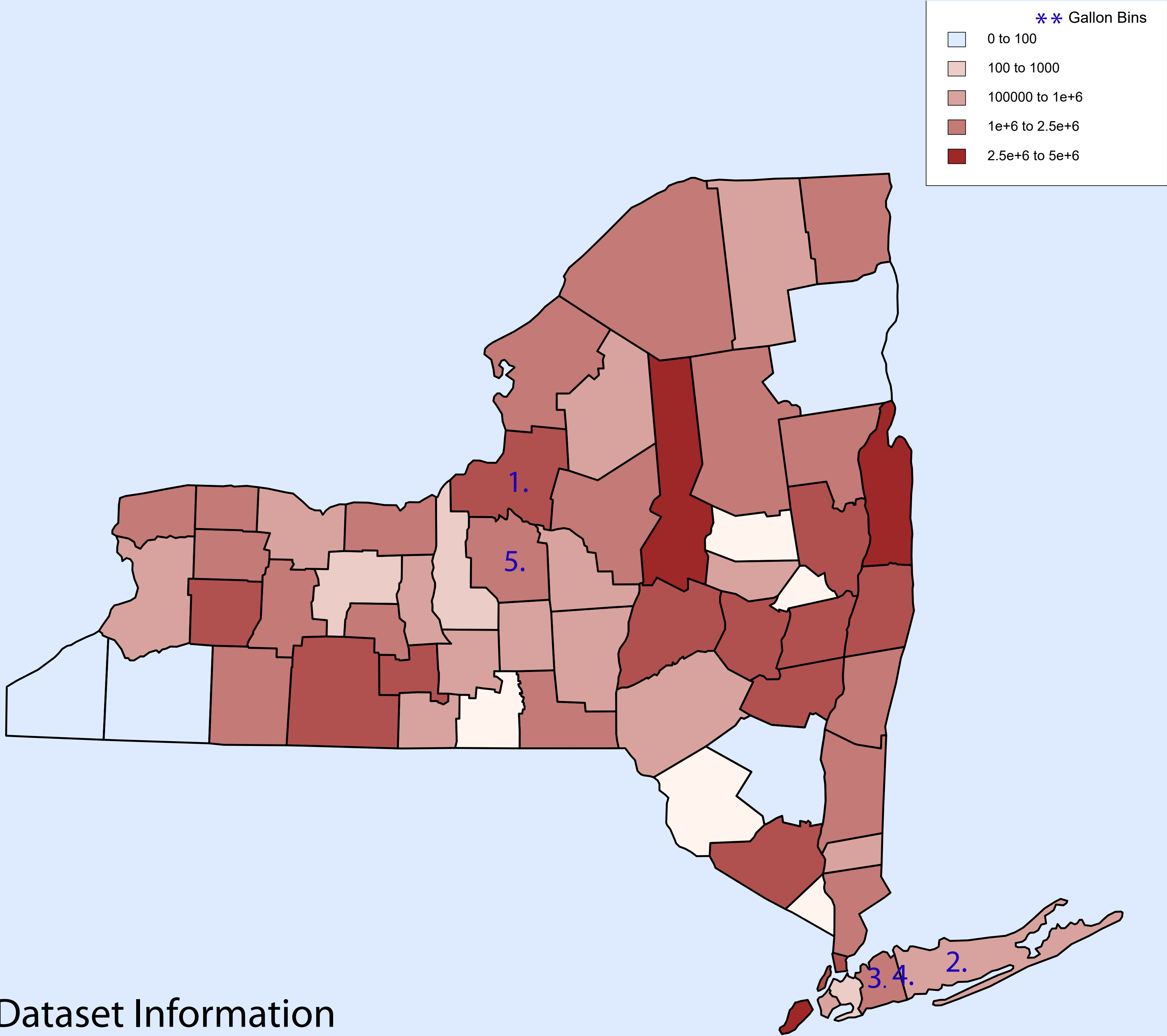


Hazardous Spills - 23 Years in New York

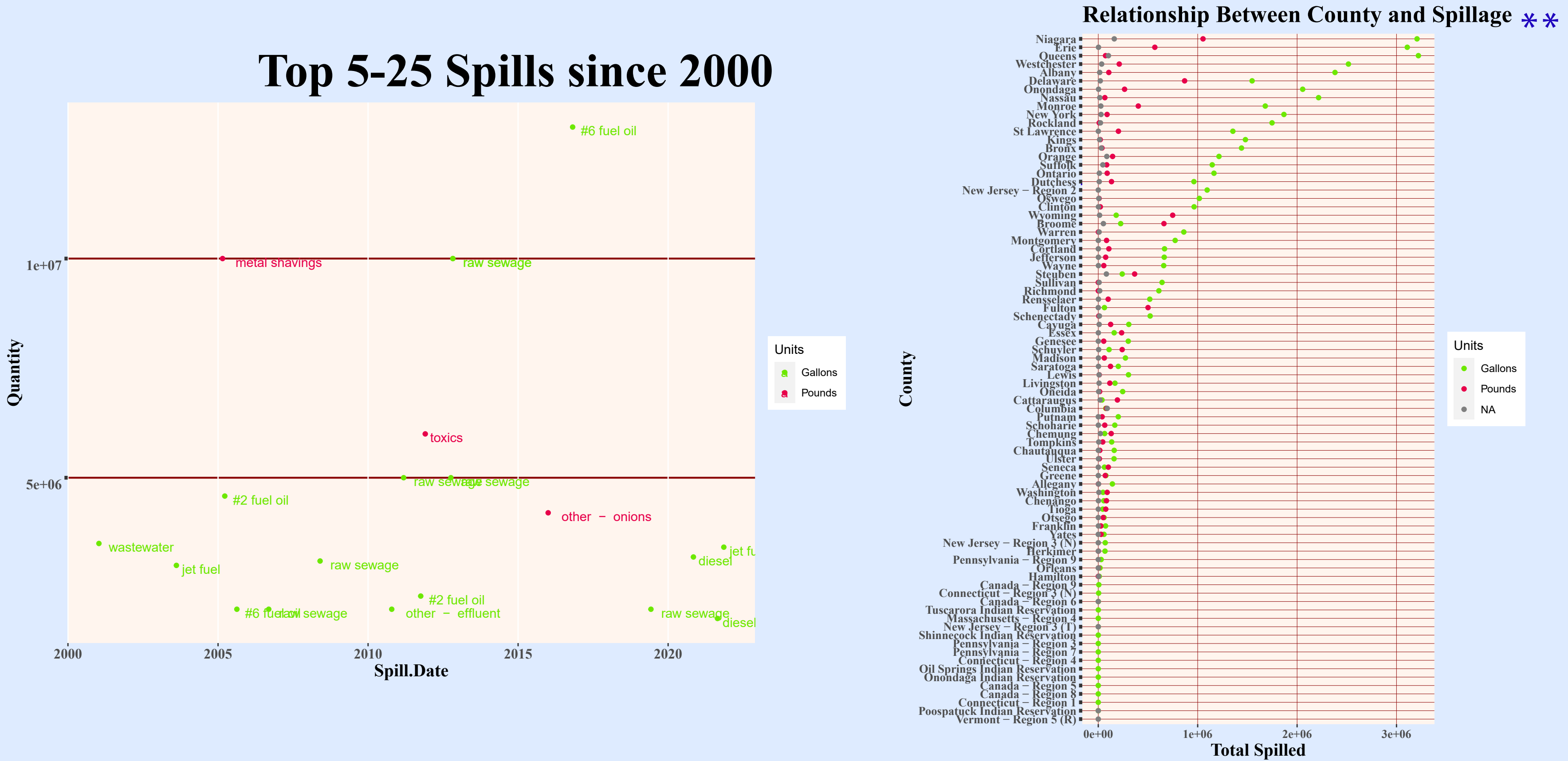


How do Spill Outliers Contribute to the County Aggregate?

- 1.6e+07 Gallons of Petroleum in 2009, in Oswego County (Gardner Creek) caused by Vandalism
- 1.3e+07 Gallons of Petroleum in 2016, In Suffolk County (Long Island Sound) caused by Storm
- 1.0e+07 Gallons of Raw Sewage in 2012, In Nassau County (Hewlett Harbor) caused by Storm
- 1.0e+07 Gallons of Raw Sewage in 2012, In Nassau County (Rockaway Creek) caused by Storm
- 1.0e+07 Pounds of Metal Shavings in 2005, In Onondaga County (Syracuse) caused by Storm

How does removal of outliers affect County Aggregates?

**indicates the removal of the Top 25 outliers



Dataset Information

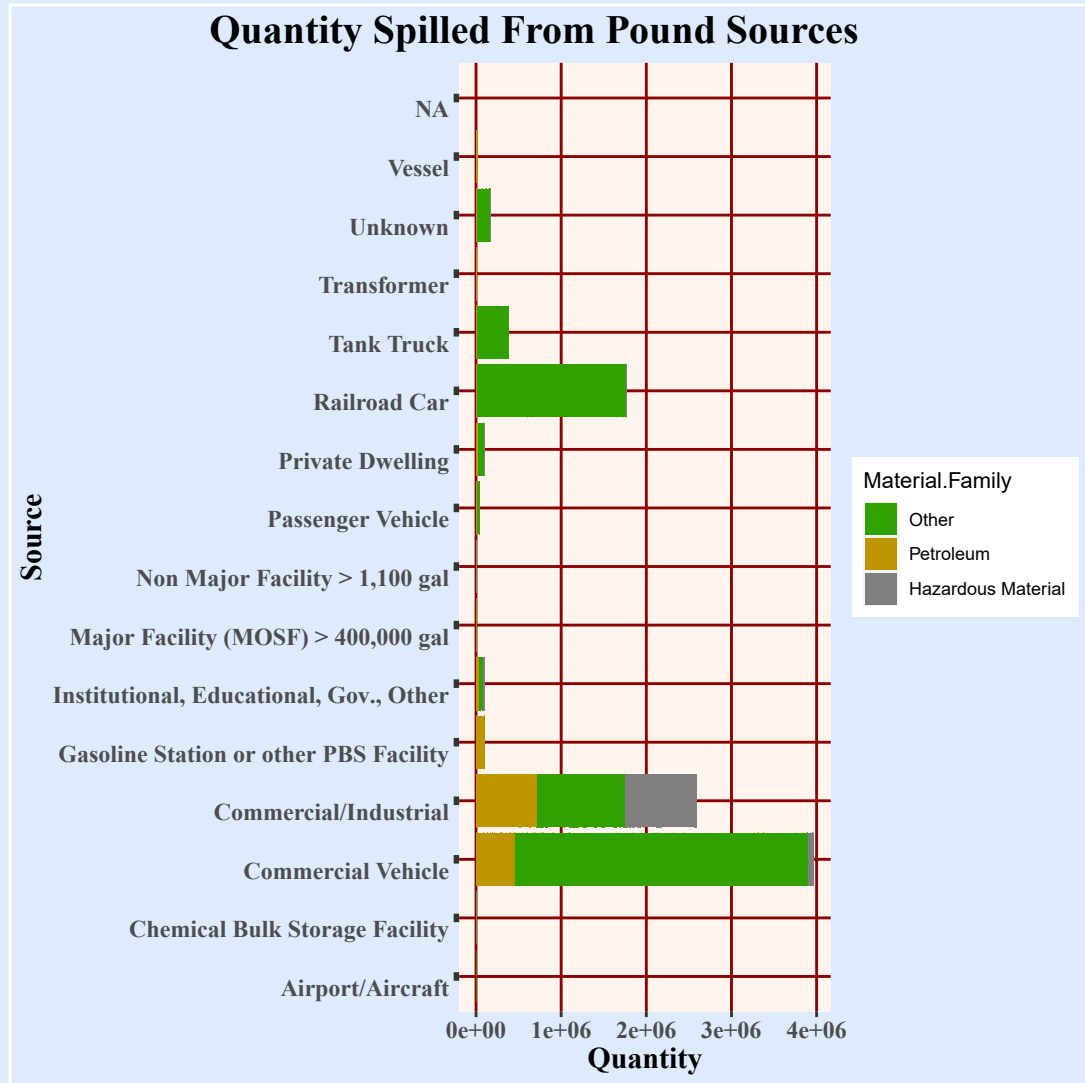
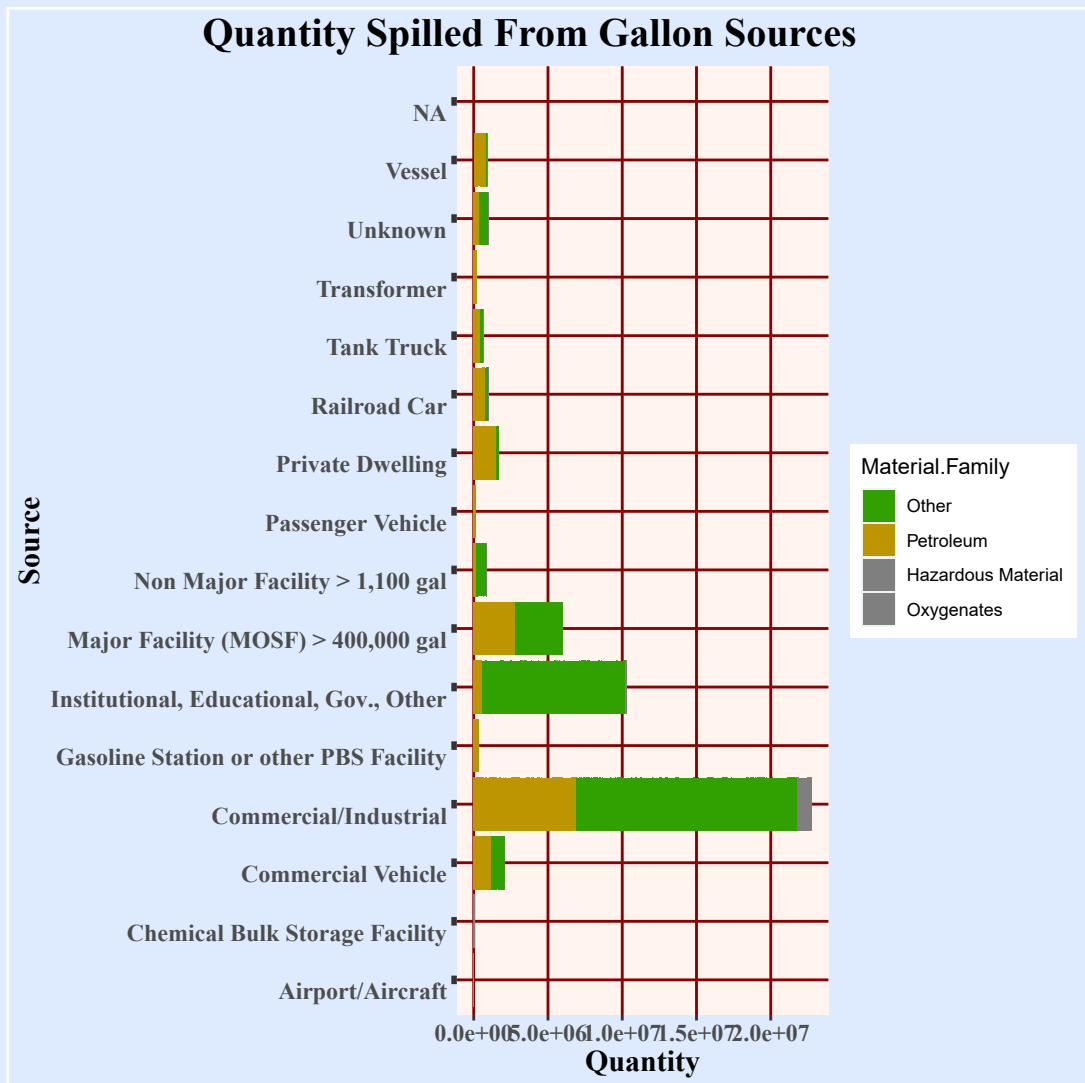
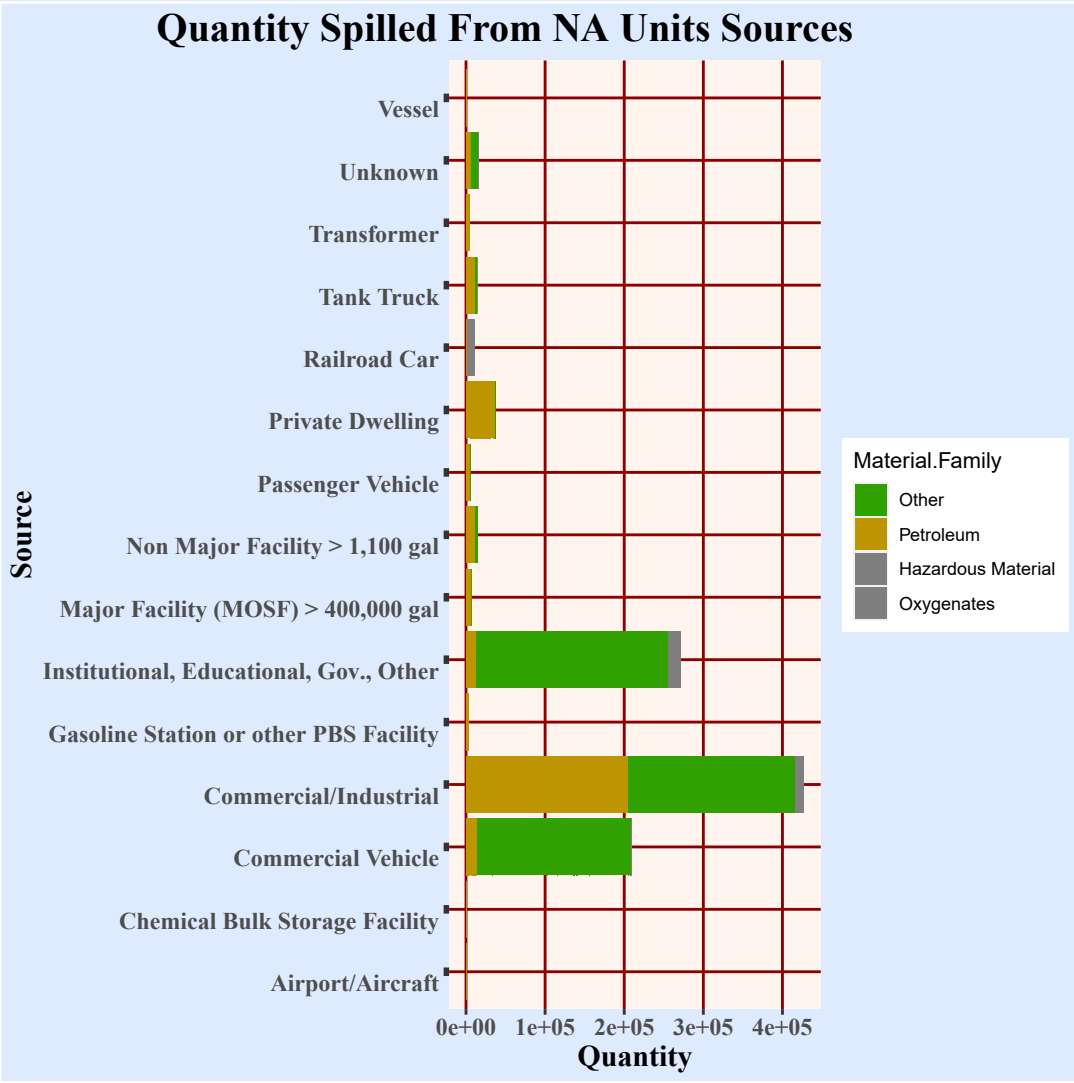
The Dataset was obtained from <https://data.ny.gov/Energy-Environment/Spill-Incidents/u44d-k5fk/data>

The original Dataset had over 500,000 rows dating from 1900 to present, Cleaning performed to segment out from year 2000 onward to present Which resulted in around 320,000 rows

20 Variables, including columns of address data, Material Family, Source, Waterbody, Quantity, Units, Contributing Factor, Open / Close Date, DEC region, and amount recovered

Context of Study

The comparison of removing outliers from data is important for attempting to find patterns, these outliers could cause critical information to become distorted in the eyes of state lawmakers, who may improperly allocate hasmat resources from biased or inaccurate reporting.



Distribution of Spills by Units

