Number of violations, percentage of violations (each of these metrics by TYPE)

Number of facilities inspected, or violations at nearby facilities (use location data to aggregate)

Aggregate NAICS code into a smaller number of categories to give us some information without being too granular

What entity are you aggregating?

What aggregation functions? (Max, min, median, count, etc.)

What time and space are you aggregating over?

Final deliverable of report: final list generated by best “selected” model (of facilities to intervene on, etc.)