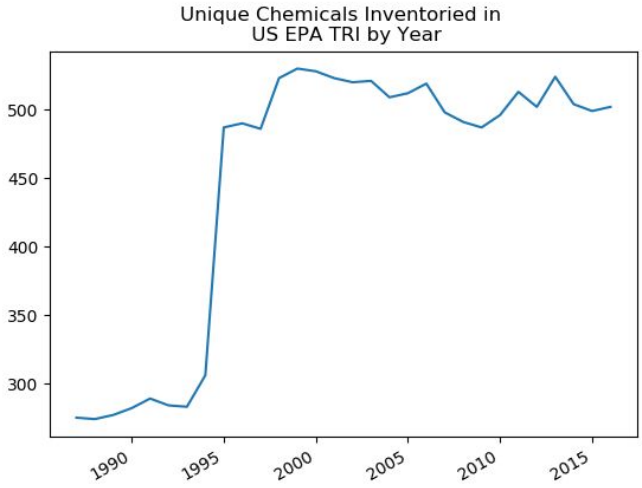
CAPSTONE I PROPOSAL
Kevin_Handel_proposal_1.pdf
October 2018

1. Do changes in the fate of specific chemicals used within US industry markedly change at specific times across the whole of the reporting entities?
Ho: Concerning the reported fate of specific chemicals in use by US-based companies, no industry-wide pivots, that are independent of reporting company, occur within marked time intervals.
H1: Changes in fate for specific chemicals occur at identifiable times across the reporting industries.

2. The US EPA maintains a database for chemicals used by US-based industries. Companies are required to self-report on the "fate" of chemicals brought in-house, and the results have been compiled as part of the Toxics Release Inventory (TRI) program. Data stored in csv form and compiled by Year for 1987 - 2016 (inclusive) is available from the US EPA website, while Kaggle.com maintains the full repository in a single 5,000,000+ line csv file. As of 2016, approximately 500 unique chemicals are tracked in the TRI program.



3. MVP: Data will be analyzed by specific chemical for changes in fate. If identified, statistics will be performed to quantify the significance of these changes, such as the percentage of companies reporting such changes and the modified direction or stream of the chemicals. Results will be reported both numerically and graphically.

MVP+: Since practices regarding the use of and disposal streams for chemicals deemed toxic by the EPA is governed by EPA guidelines and/or regulations, numeric modeling may reveal predictive company behavior which may be used to more effectively manage these chemicals. In addition, as some toxic chemicals used are specific to an industrial sector, the data may reveal trends toward recycling and reuse and away from air/water/landfill disposal.

MVP++: Through correlation with data regarding rule-making for environmental regulation (secondary data sources will be required), the instantiation of more restrictive rules may be identified by the changes in fate that are revealed from the primary data source.