LAB REPORT

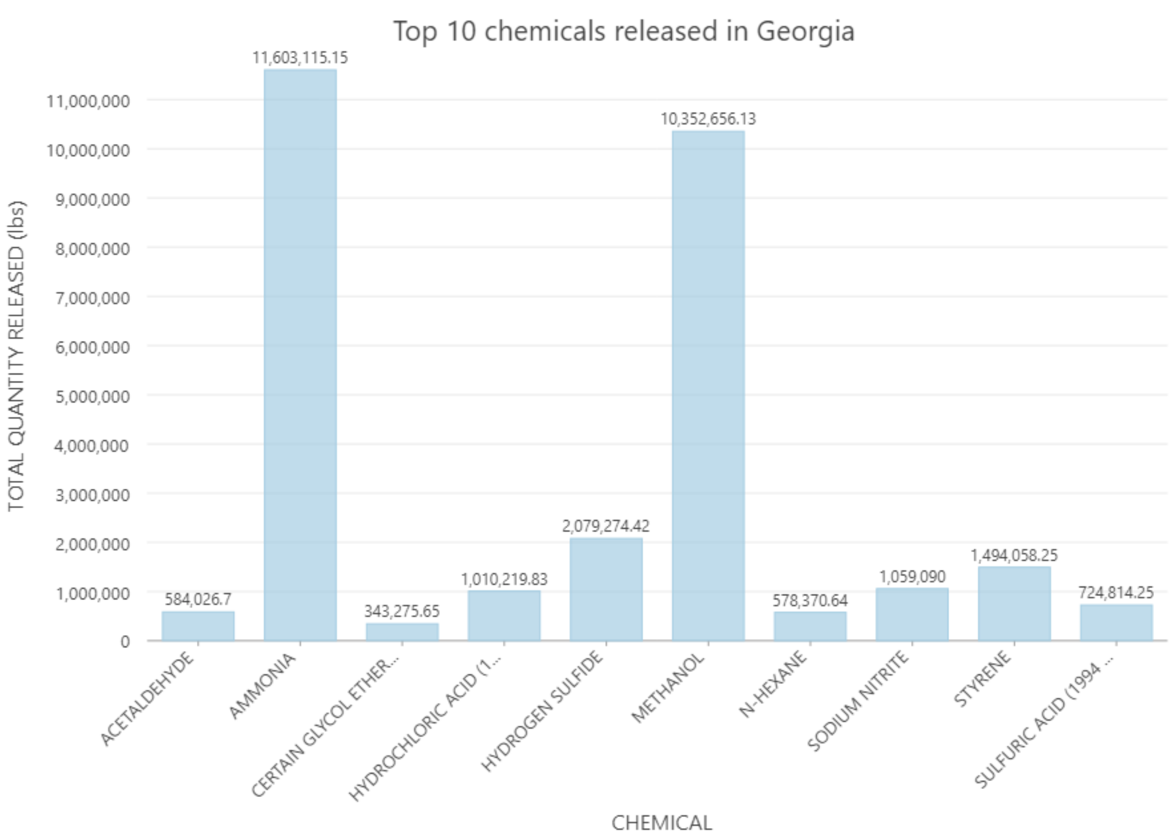
Question 1: What is the *written* definition of the other two columns that the table includes? Please phrase your answer like 1) and 2) above.

Use the words *distinct* and *total* once each in your answers and include *units* if necessary.

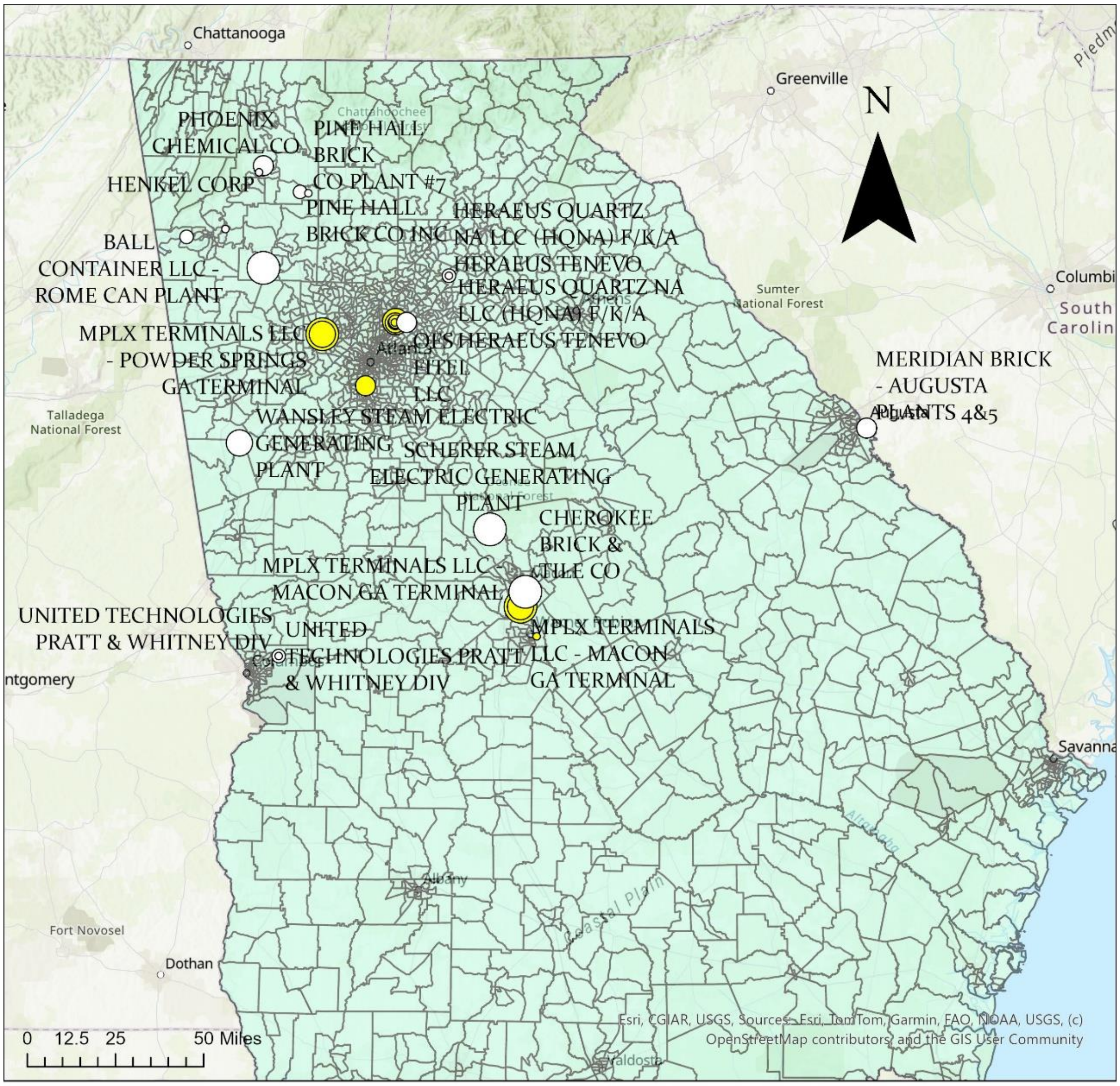
1a. Colname [fill in UNIQUE_GATRIS_CHE] Description : This column denotes the total number of distinct/unique chemicals that is used by a particular parent company

1b. Colname [[SUM_GATris_REL] Description : This column denotes the total amount in pounds released for those distinct chemicals on facility in 2018

Graphic 1: On your own, make a nice chart of the *top 10 chemicals in Georgia* by how much is produced. Note that you will need to create a new summary table to answer the question, don't use *TRIs_Aero_2018_Statistics*. Give both axes labels and give the chart a nice heading. A former example is given below. For full points, ensure your labeling doesn't have underscores and abbreviations that are difficult to understand. Ensure units are communicated.



Graphic 2: Make a map with the SoVI data and TRI release of two chemicals that you select (your choice). Use graduated / proportional circles (or rings) to show the amount of each chemical being released. Your map should ensure that the SoVI data categories make sense vis-à-vis your color choice. You can use some transparency for your graduated circles. Label by Facility Name (not the chemical—your legend should take care of that). Remember your legend shouldn't have underscores, etc. Remember to include a scale bar.



Legend

Hydrogen Fluoride

- 5.00 - 309.00
- 309.01 - 1184.00
- 1184.01 - 2534.00
- 2534.01 - 5300.00
- 5300.01 - 50220.00

Benzene

- 60.00 - 64.00
- 64.01 - 120.00
- 120.01 - 180.00
- 180.01 - 250.00
- 250.01 - 330.00
- Social Vulnerability - Georgia

This map illustrates the amount of Hydrogen Fluoride and Benzene released to air, water and land on-site at the facility (in pounds). The white circles are used to represent Hydrogen Fluoride and the yellow circles represent Benzene, with the circles proportionally representing the amount released. The facility names are labelled to indicate factories from which these particular chemicals are discharged.

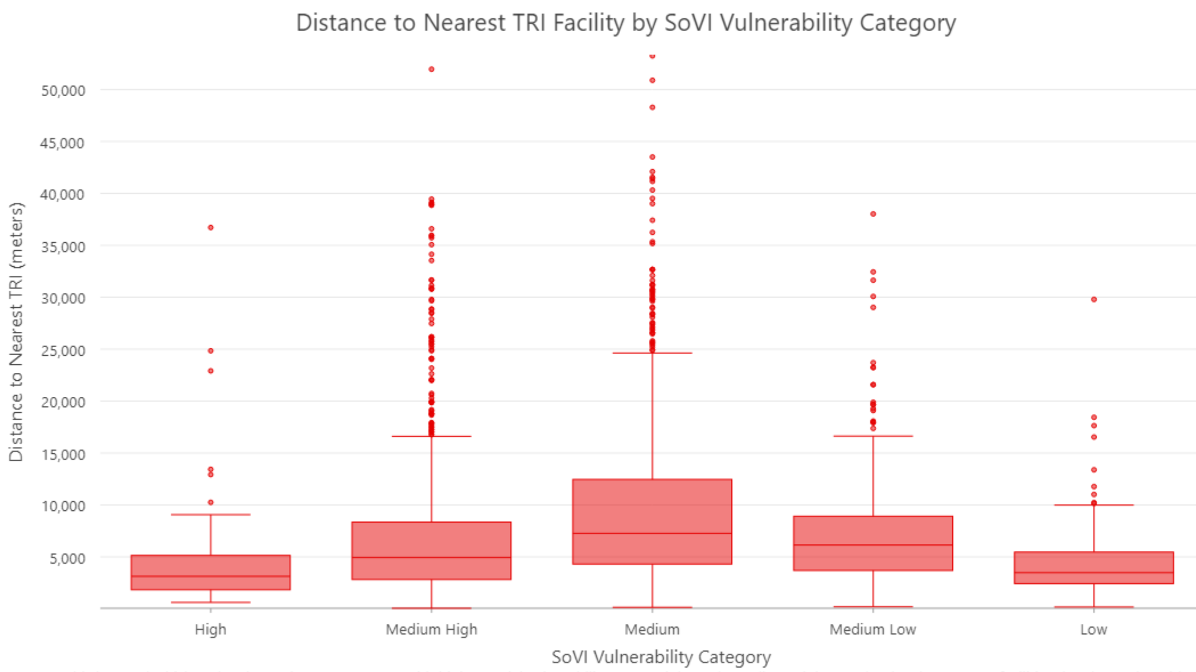
Question 2: Name three columns in the SoVI dataset comprised from *composite data* that can be used to assess social vulnerability. How are they different? Use the word 'ordinal' in your answer.
Answer 2: Three columns in the SoVI dataset that are derived from composite data include SOVI0610GA, SOVI0610_1, and SOVI0610_2. The SOVI0610GA field provides a continuous numerical score for each census tract, reflecting its level of social vulnerability. These scores typically range from approximately -13 (indicating low vulnerability) to +13 (indicating high vulnerability). In contrast, SOVI0610_1 and SOVI0610_2 are ordinal scales that categorize vulnerability into ranked groups based on the same underlying factors. SOVI0610_1 includes five levels: Low, Medium Low, Medium, Medium High, and High; while SOVI0610_2 simplifies this to three categories: Low, Medium, and High.

Question 3: Describe 3 (or more) variations on the traditional circular buffer that we can use, as discussed in class.
Answer 3: The variations on the traditional circular buffer that can be used are 1. Overlapping (discs) vs non-overlapping (rings), 2. Single-distance vs multi-distance buffer, 3. Graduate distance based on variable in dataset

Question 4: Fill out the table below.

Distance from TRI	Total Population (P0010001)	Minimum SoVI (SOVI0610GA)	Average SoVI (SOVI0610GA)	Maximum SoVI (SOVI0610GA)
Within tract (1 m)	1461873	-9.576	0.706	8.156
3 KM	5137755	-12.687	0.161	12.870
10 KM	8892194	-12.687	-0.045	12.870
All of Georgia (Use original SOVI table)	9687651	-12.687	0	12.870

Graphic 3: Create a box and whisker plot of the distance to the nearest facility by the SoVI category of the tract (there should be 5 categories). If you can change the colors by the category, that would be great (or reorder them by high to low vulnerability). You can use kilometers or meters as the distance. You can use miles if you'd like.



This box and whisker plot shows that census tracts with higher social vulnerability (e.g., "High" and "Medium High") tend to be closer to TRI facilities than less vulnerable areas. The median distance decreases as vulnerability increases, suggesting a potential environmental justice concern where more vulnerable communities face greater exposure to toxic release sites.

Question 5: They recently have closed the Environmental Justice EPA monitoring sites. The US has also dropped a lawsuit against a chemical facility in Louisiana that was producing chemical responsible for cancer in this "cancer alley".

You may be needed in the future for map support to prove that some EJ issues are arising. IN YOUR OWN WORDS, please discuss the different subjective decisions you would want to make in order to show that there are issues. What modeling decisions could result in a good, better, or excellent model? To receive points on this question, please be specific as possible—no general answers like "data should be up to date", "data should be from an authoritative source", or "methods should be good for the data at hand", etc. Name actual datasets, actual parameters, etc. and use GIS terms.

Answer 5:

1) Input Data

Good: Use EPA's TRI (Toxic Release Inventory) data for a single year (e.g., 2018), and basic ACS population data at the county level.

Better: Use multi-year TRI data (e.g., 2010–2020) to track long-term emissions trends and SoVI tract-level vulnerability data from the University of South Carolina. We could also try to include Census block group level demographics to improve spatial resolution.

Excellent: Combine TRI data with EJScreen indicators (e.g., cancer risk, air toxics respiratory hazard index), CDC Wonder data which analyses and collects cancer occurrences across the United States and also collects the different types of cancer occurences, and plume modeling outputs. Use land parcel data to model population exposure more precisely.

2) GIS Combination / Analysis Methods

Good: Use simple buffer analysis (e.g., 1 km, 3 km, 10 km) around TRI sites and calculate population exposure using spatial joins.

Better: Incorporate spatial joins with SoVI data, use zonal statistics to calculate vulnerable population within buffers, and apply kernel density estimation to assess TRI concentration.

Excellent: Use areal interpolation or dasymetric mapping to allocate population within plume areas. Perform spatial regression to identify clusters of high exposure and high vulnerability. Consider network-based accessibility for evacuation or exposure routes.

3) Output Media (Communication) and Statistics

Good: Maps with proportional symbols showing TRI amounts and population data; basic tables of summary stats.

Better: Interactive web maps using ArcGIS Online or StoryMaps, charts showing trends in emissions by demographic groups, and boxplots comparing distances to TRI by SoVI category.

Excellent: A dashboard with real-time updates, dynamic filters (e.g., by chemical, race, income), infographics, and ANOVA or regression analysis showing statistically significant EJ disparities. Include downloadable data layers and interactive variables for transparency.