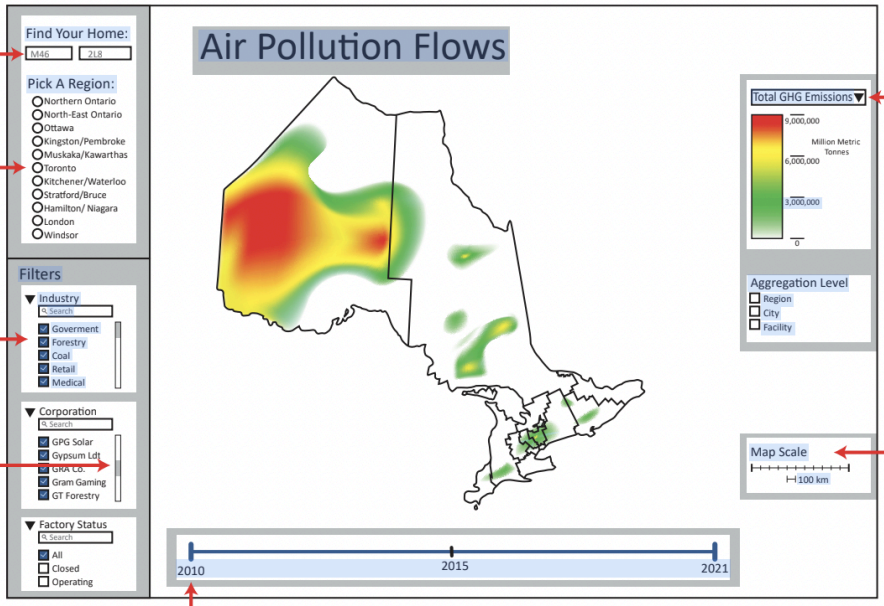
Ontario Environmental Polluters: Keeping Companies Accountable

Zeeshan Gazi, Shay Houseknecht, Ever Hughes, Khushil Nagda, Ivanka Tkachenko
Group 20

1. Landing Page



Enter your zip code to see heat map around your home

Or click to zoom in to a specific region

View heat map by specific industry (toggle off)

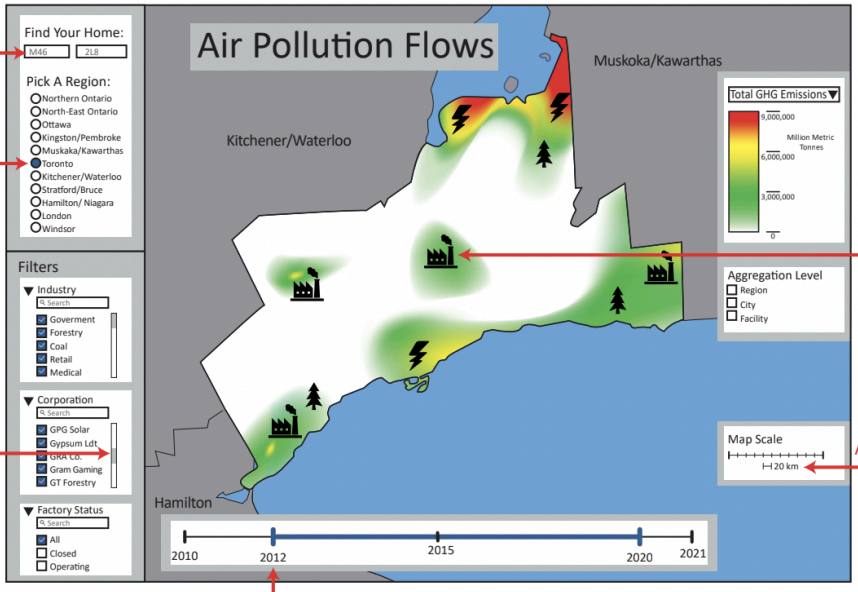
Sliding list to view more

Drop down menu for alternate polluters

Can zoom manually instead of picking a region

Show emissions for specific years only by sliding bars

2. Zoomed In View



Enter your zip code to see heat map around your home

Or click to zoom in to a specific region

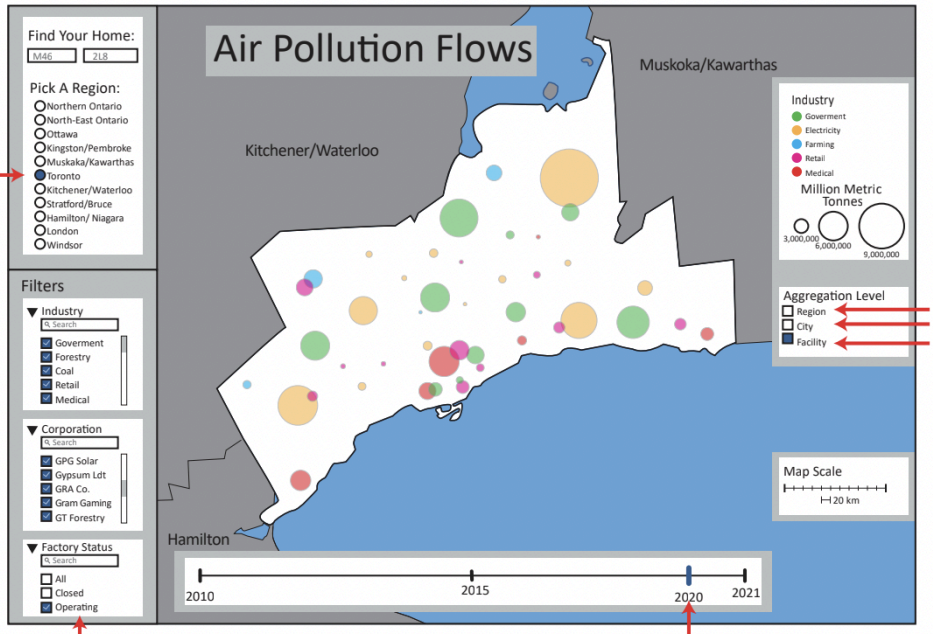
Sliding list to view more

Zooming in will add industry icons

Adjusts automatically when zoomed in

Additional years annotated when bar slides and lands

3. Facility Aggregation



Still zoomed into Toronto

Generate Choropleths

Generate Dot Map

Can toggle to examine emmisons from still operating facilities only

Can converge both sliders to a single year

The first visualization (1. Landing Page) is an interactive heat map of Ontario with regulated facilities who report their CO2 emissions. Our goal is to display which locations are emission hotspots for an audience of Ontario citizens. Facility location and density of CO2 emissions will be encoded through area (2D size) and color hue. In the zoomed out view, areas with a high concentration of emitters will have larger coloured areas on the map compared to areas of low concentration. A diverging colour scale, based on traditional heat map colours, will encode emission density. Warmer colours will naturally correspond to a higher concentration of emissions. A green to red scale is not the most effective choice, but our audience will be familiar with the heat map scale. A sequential scale varying colour saturation may be used instead for higher saliency. Due to pre-attentive processing, areas with high emissions will pop-out from the map. Using the Gestalt principles of similarity and proximity, the bird's eye view will group facilities together into broader regions of emissions on the map. A tradeoff of this approach is that the most effective magnitude channel, position on a scale, will not be used to encode density of emissions or facilities. A barplot with the number of emitters or emissions per city would use the highest rank channels, but the spatial aspect would be missing. Using the effectiveness principle, spatial channels should be used to encode the most important attributes, which is facility location.

The interactive elements will mainly use view manipulation, allowing for free exploration of the map. Viewers can use navigation to zoom in to various levels of detail, click and drag to move around the map, and take different perspectives (2. Zoomed In View). Selection will be available so that viewers can click on specific facilities to get more information on them. Depending on the reader's interests, they can explore the city they live in or any areas of their choosing through a search feature. As the reader zooms in, there is also a data transformation and aggregation of the heat map spots. Clusters will break apart into multiple spots as the reader hones in on a specific area. Viewers can filter by region, industry, and facility status for a reader-driven story.

Readers will be able to navigate to a dot map with changed visual mappings (3. Facility Aggregation). The goal of this view is to focus more on the individual facilities and trends among them. The heat map view is best for understanding overall trends in GHG emissions, but the heat spots can obscure smaller patterns and facility locations. In this map, location will be encoded with a simple mark on the map, and CO2 emissions will be encoded with size of the mark (area). The biggest CO2 emitters will be immediately evident to the viewer through pop-out, drawing them in to investigate further. We plan to use the identity channel of colour to encode industry groups, a categorical and nominal variable. Due to Gestalt similarity, facilities of the same colour and industry will be perceptually grouped together. A limitation is that industry type has a high cardinality, which would lead to poor discriminability. To address this, we will group industries into broader categories and reduce cardinality so fewer colours are used. A qualitative colour scale will be used since industry is a discrete attribute. Simple images may be imposed on the marks to represent each industry, easing semantic understanding for our audience.

Interactive elements will include navigation and selection. Viewers will also have the option to transform the data through filtering. The map can be filtered by year via a slider to see how emissions change live over time. Selecting a certain facility will bring up an additional plot with facility-specific information. The base maps will be simple for optimal separability between the facility data and the map. A tradeoff is that some important features of the map may be missing, like street names, landmarks, and public spaces. More detailed backgrounds could be used, but this would lead to more

interference of the most important attributes. A shape-preserving Mercator projection will be used, which viewers will be the most familiar with. This map does distort areas, but distortion will be minimized due to the zoom level of the map. Additional facility information will be displayed via labels, like aggregations of CO2 emissions, since comparing areas can be deceiving.

Development Plan

The technology we plan to use is Tableau. We plan to create an interactive dashboard of different map views, charts, and data that will be available online. Users will be able to interact with data through filters, sliders (like the time slider for emissions), and other interactive capabilities, making it easier to explore and discover patterns or insights. We cleaned our data and added extra variables using Python web scraping (Requests library), Excel, and R (tidyverse). We used Google Maps to source coordinates for the facilities in the dataset, and the Canadian government for NAICS code industries.

The work is going to be split evenly among the members, to the best of our ability. This involves the creation of the product and the video demo as well. As of right now, the plan is to have Ivanka do the data processing and cleaning, and having Ever, Zeeshan, and Khushil work on most of the Tableau visualizations, but all of us plan to get involved making edits. Ever will start by building the base of the dashboards and getting the interactions up and running, and then Khushil will come in after to add in extra details and finalize dashboard elements.

Zeeshan will act as support with Tableau and will also take lead on the video script. Shay will take lead on the video editing and putting together the variety of clips to make one cohesive video. We plan to all be involved in the recording/speaking engagement part of the demo video. We will try to get everyone involved in making key design decisions for the visualization and in finalizing their own part of the script for the video. Balancing workload will be a priority due to everyone’s busy schedule at this time of year.