Mining, Eh?

1. Basic Info

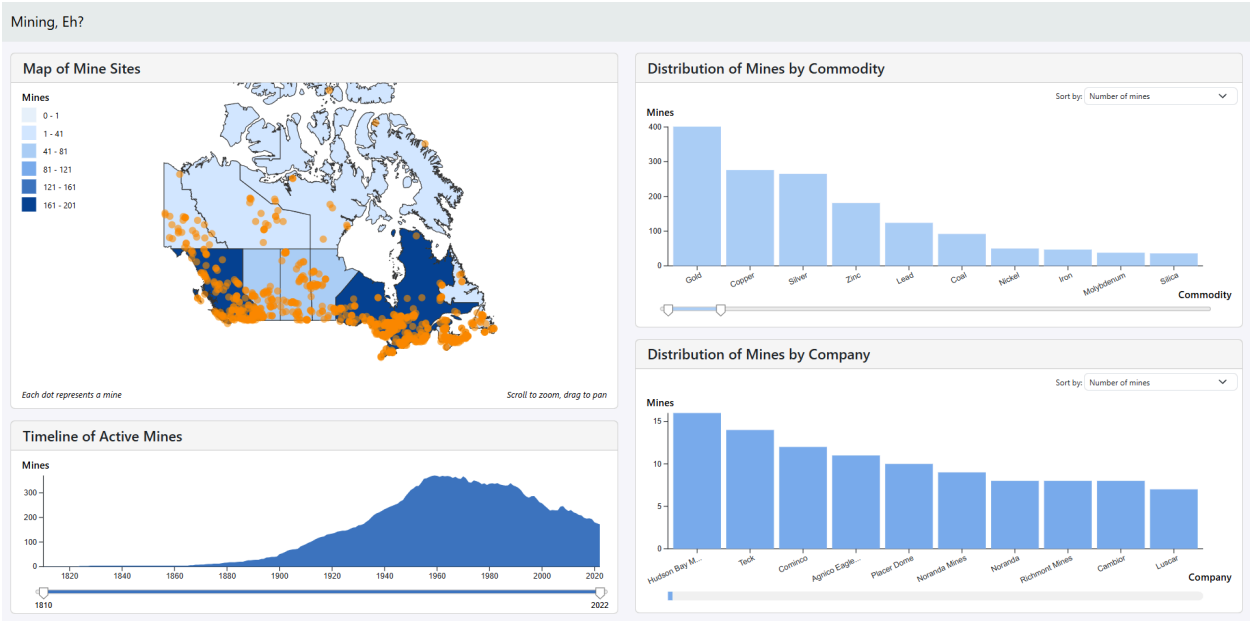
Group Members:

- Adwait Kulkarni
- Jamie Ma
- Simon Liang
- Sparsh Trivedy

Live demo: https://jamiema1.github.io/Mining-Eh/

Video Demo: https://youtu.be/XKTumtTmjnw

2. Overview



Canada's mining sector has been pivotal in driving economic growth, resource extraction, and technological innovation for over two centuries. However, understanding the key aspects of this sector, such as the distribution of mines across the country and whether their geographical locations correlate to the commodities they produce, remains a challenge due to the dispersed nature of the data.

Our initiative is to serve mining history enthusiasts by providing an exploratory tool to assist users in developing a deeper understanding of Canada's mining history. By visualizing the distribution of mines and the commodities they produced within Canada's provinces and territories from 1810-2022, we aim to uncover temporal and geospatial trends in commodity production. By providing users with the option to filter mining companies and see how many mines each company has worked on, we offer users information about how mine ownership has evolved through the years and which companies lead the way with the most mines worked on. Through a dashboard-style visualization, users can explore mine distribution data at both country-wide and province-specific levels, including the ability to filter by commodities produced, company ownerships, and operational periods.

3. Data

The Principle Productive Mines of Canada dataset (source) contains tabular data for 947 mines in Canada from 1810 to 2022. We will leverage 27 of the 35 attributes (see table below) with a focus on exploring topics related to mine ownership, geospatial information, operational timelines, and resource information.

We conducted our data preprocessing using a Python script with Pandas (see `data/preprocess.py`):

1. Convert the raw `.xlsx` file into a `.csv` file of just the data. Then, read and convert the `.csv` into a Pandas DataFrame.
2. Remove 8 attributes that contained either information that was not useful for our purposes or sparse textual information that cannot be easily visualized.
3. Add an `id` column corresponding to the mines index within the dataset and an `active_status` column, which is `True` if the mine is `open`, `False` otherwise.
4. Replace any `open` values within the `close1`, `close2`, and `close3` attributes with the last year of the dataset (`2022`). This ensures those attributes are homogeneous and only contain a number type, apart from any `N/A` values.
5. Remove any hyphens from province names to ensure they are consistent.
6. Replace any empty cells with `N/A` for easier parsing during later stages.
7. Trim any leading and trailing whitespace.
8. Normalize company names by removing special characters, suffixes, etc.
9. Fix any data quality issues by doing a manual sanity check and correcting any values that appear incorrect and could pose problems when visualizing.
10. Convert the DataFrame back into a `.csv` file (see `data/mines.csv`).

The resulting preprocessed tabular dataset contains 947 items and 27 attributes.

Attribute	**Type**	**Cardinality / Range**	**Notes**
id	Categorical	947	Unique identifier for a mine.
company1 company2 company3 company4 company5 company6	Categorical	1185	A mine has 1-6 names of the mining companies that operated the mine at any point in the history of the mine.
namemine	Categorical	601	Optional name of the mine.
town	Categorical	446	Optional name of the town the mine is located in.
province	Categorical	12	Full name of the province or territory the mine is located in. No mines are located in Prince Edward Island.
latitude	Quantitative	[42.10, 75.39]	Latitude coordinate of the mine.
longitude	Quantitative	[-140.72, -52.95]	Longitude coordinate of the mine.
open1 close1 open2 close2 open3 close3	Ordinal	148	A mine has 1-3 pairs of years corresponding to when a mine started and suspended production. Mines that are still active have a final close date of 2022 (the last year of this dataset) and have an active_status value of True.
active_status	Categorical	2	Whether a mine is currently active or not.
commodity1 commodity2 commodity3 commodity4 commodity5 commodity6 commodity7 commodity8	Categorical	103	A mine has 1-8 commodities that are produced by the mine (ex. Gold, Silver, Copper, etc).

4. Tasks

#	Domain-Specific	Abstract	Attributes
1	How have mining operations changed over time in Canada? Mining history enthusiasts may want to explore how mining activity has changed over time, including the rise and decline of operations in different regions. By examining historical trends in mining expansions and closures, users can gain insights into broader patterns of Canada's mining industry.	Browse mining operation open/close times.	open(1-3) close(1-3)
2	Where are specific resources mined, and how are mining activities distributed across Canada over time? Mining history enthusiasts may be interested in exploring how the extraction of specific resources such as gold, copper, or nickel has shifted over time within different regions in Canada. By tracing the rise and fall of mining activities for specific commodities, users can uncover historical peaks in production and regional concentrations. This exploration can reveal insights into economic booms, regional development, and changes in the importance of certain resources throughout Canada's mining history.	Explore spatial and temporal distributions of commodity extraction.	commodity(1-8) open(1-3) close(1-3) latitude longitude province
3	Which companies are most active in the Canadian mining sector, and how do their operations and partnerships vary across time and space? Mining history enthusiasts may be interested in exploring how individual companies contributed to the expansion of mining across Canada and their activities over time. By analyzing company-specific mining operations across regions and commodities, along with identifying which companies collaborated with each other, users can uncover deeper insights into industry shifts and corporate relationships.	Compare company operations spatially and temporally.	company(1-6) open(1-3) close(1-3) latitude longitude province commodity(1-8)

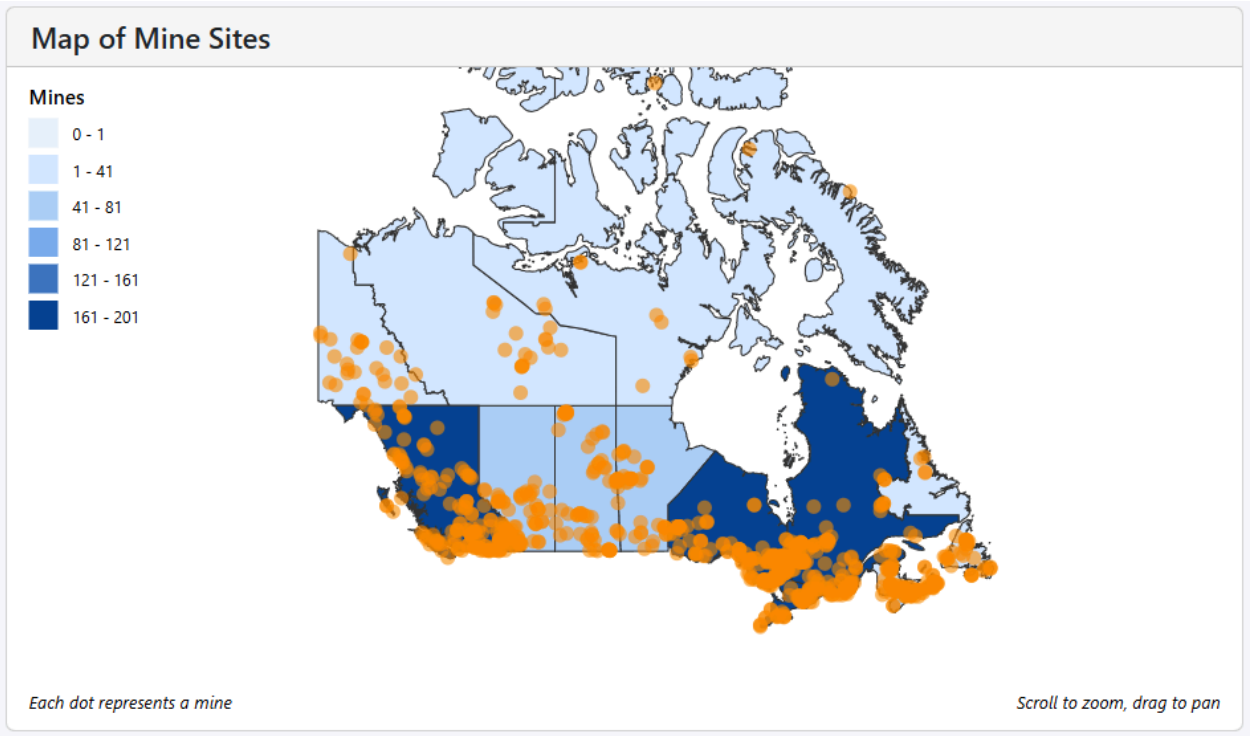
4	Where exactly are mining operations located? Mining history enthusiasts may want to explore the exact location of the historical mining operations, including their latitude and longitude, for specific mining operations across Canada. This information can help them discover old mining sites they might want to visit. Plan trips to explore historic mining regions.	Identify geospatial locations of mining sites.	latitude longitude namemine

5. Visualizations

We created a dashboard-style visualization with 4 views: a spatial view (choropleth dot map), a temporal view (area chart), a commodity view (bar chart), and a company view (bar chart). The spatial view explores the distribution of mines at the country and provincial levels. The temporal view explores changes in the number of mines over time. The commodity and company views show the distribution of specific commodities and how many mines a certain company has worked on, respectively. Each pair of views is bi-directionally linked through interactions, allowing the user to slice and dice the data during exploration.

The colour palette chosen uses shades of blue with secondary colours of grey, orange, and red. This palette was chosen because it is colourblind safe and has the benefit of the primary colour blue contrasting well with the accent colour of orange. Moreover, each view's primary colour is a particular shade of blue to signify the linkage between all four views. In addition, the accent colour (orange) is used to denote a mine within all views, further demonstrating the connection between views.

View 1: Choropleth Dot Map (Spatial View)



Marks:

- Poly mark (level 1) represents a province
- Point mark (level 2) represents a mine

Channels:

- *Horizontal position/order*: longitude of a province
- *Vertical position/order*: latitude of a province
- *2D boundaries*: 2D boundaries are shared between provinces in the map
- *Color hue and saturation*: Darker shades represent a larger number of mines, and lighter shades represent a smaller number of mines

Description/Rationale:

The main view will be a choropleth map of Canada visualizing the distribution of mines across different provinces. This view allows users to see geospatial trends in mine distribution and supports filtering through both interactions within the view and linkages with other views.

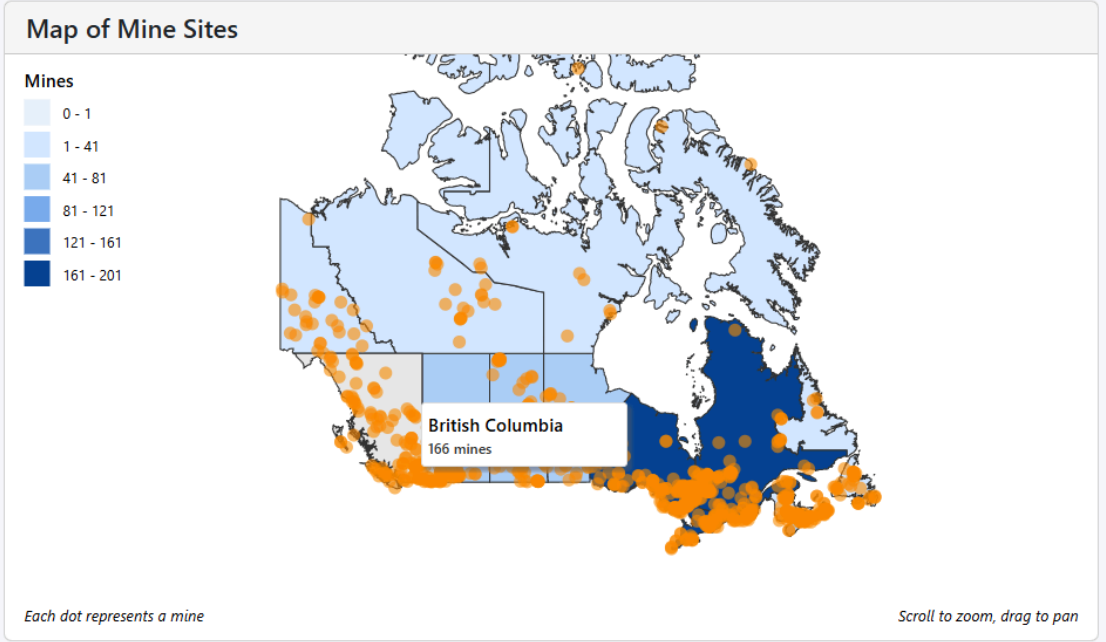
The choropleth map was chosen due to its ability to represent spatial data across geographical regions and to easily recognize trends through the use of colour. By providing a high-level overview of the number of mines across Canada, users can easily see trends and make comparisons between provinces (T2, T3, and T4). Compared to a dot density map, where the size of a singular dot encodes the number of mines in a province, the choropleth map avoids issues with circles overlapping or exceeding borders and provides a more intuitive way of visualizing the data within regional boundaries.

Following the principle of expressiveness, we are using the magnitude-ordered saturation channel to encode the number of mines, a quantitative attribute, and the identity color hue channel to encode individual provinces, a categorical attribute. Using the hue and saturation channels to show the number of mines within each province allows for the discovery of geospatial trends and attracts attention to provinces with a larger number of mines that may be of interest. With different saturation levels, it is also easy to compare which province has a higher or lower number of mines compared to some other province, so relative comparisons are made easier with the saturation channel, following Weber's Law. Compared to other encodings, such as the shape/symbol channel, the color hue channel better addresses our tasks, since colors are more effective than shapes for comparison purposes.

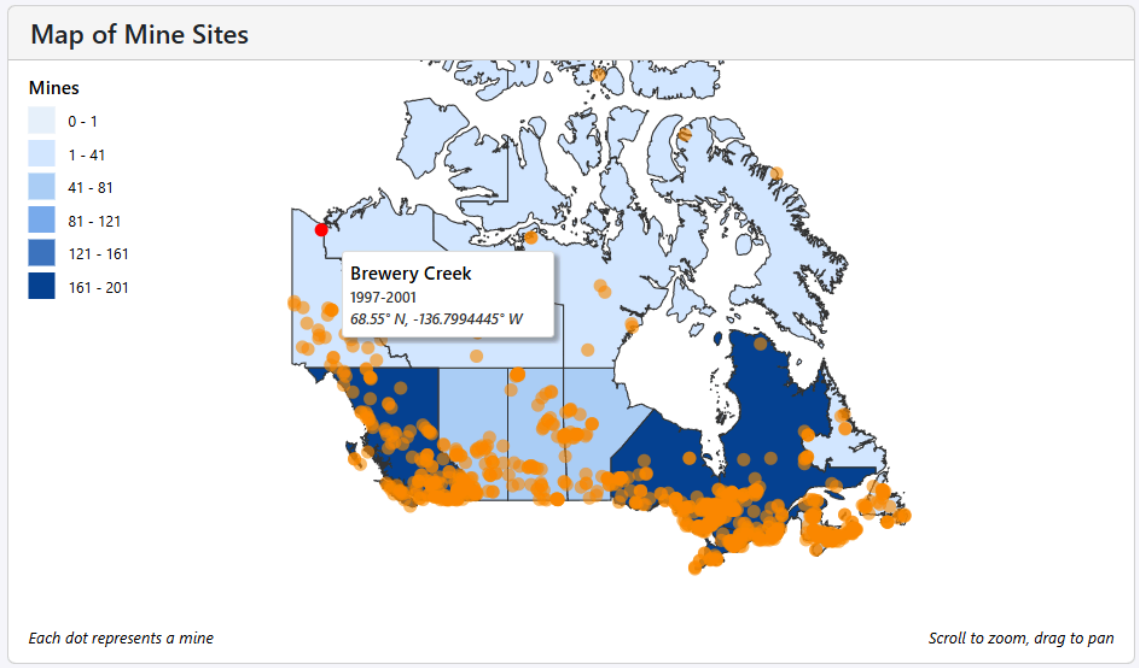
Interactivity:

Tooltip

- On hover of a province, the province will become highlighted and a tooltip will appear displaying the province name and the number of mines.



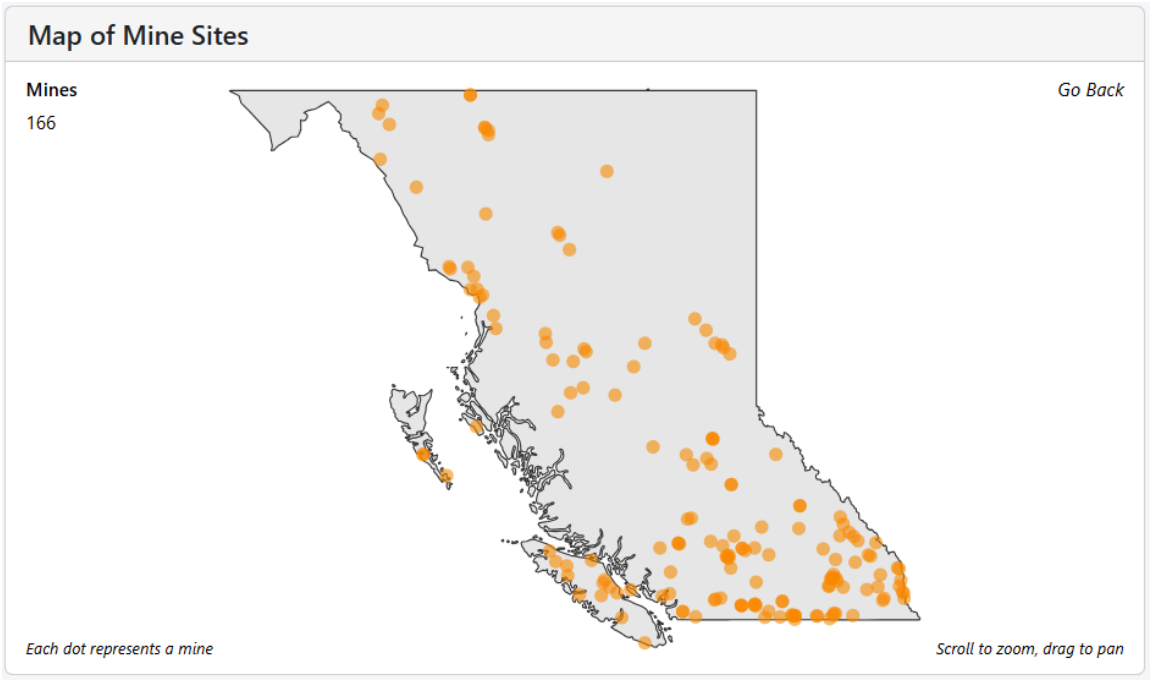
- On hover of a mine, the dot will become highlighted and a tooltip will appear displaying the mine name and longitude and latitude coordinates.



Through redundant encoding, tooltips enhance the user's ability to gather precise data.

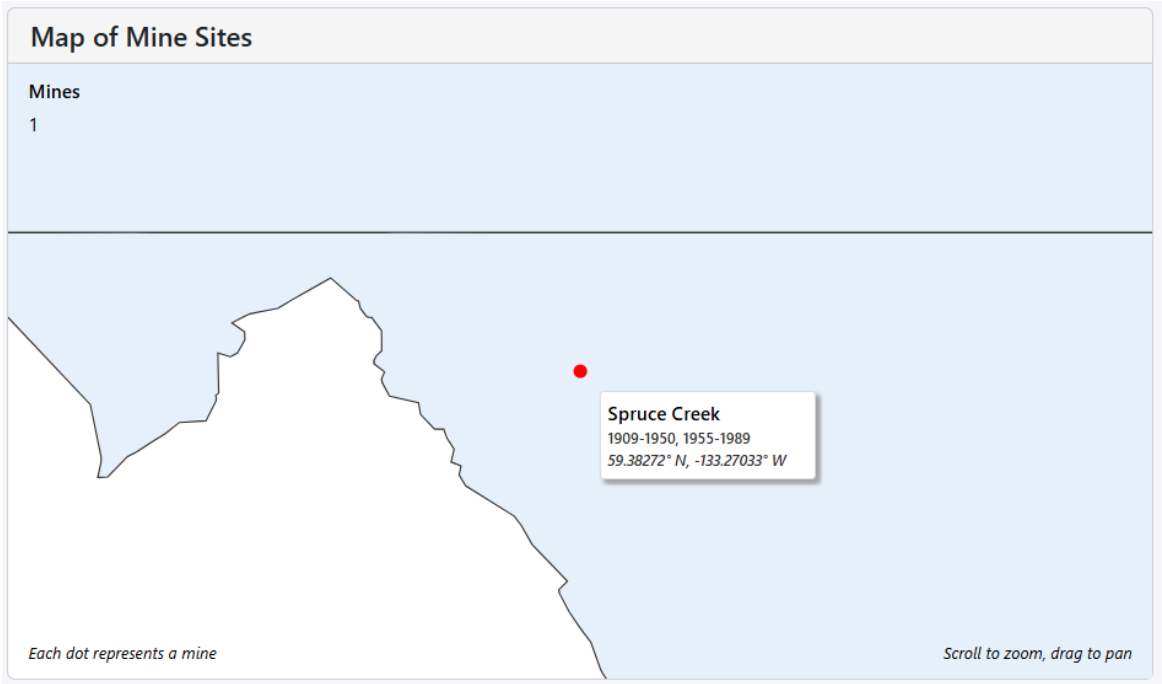
Province Filter

- On a click of a province, the main map will zoom in to show the drill-down view of the selected province. Additionally, all other views will only show data for mines within the selected province. Clicking the “Go Back” button will restore the top-level view and remove the province filter. This interaction addresses T2 and T3 by allowing users to view province-specific details, such as locating mines that produce a specific commodity or finding mines that a given company operates.



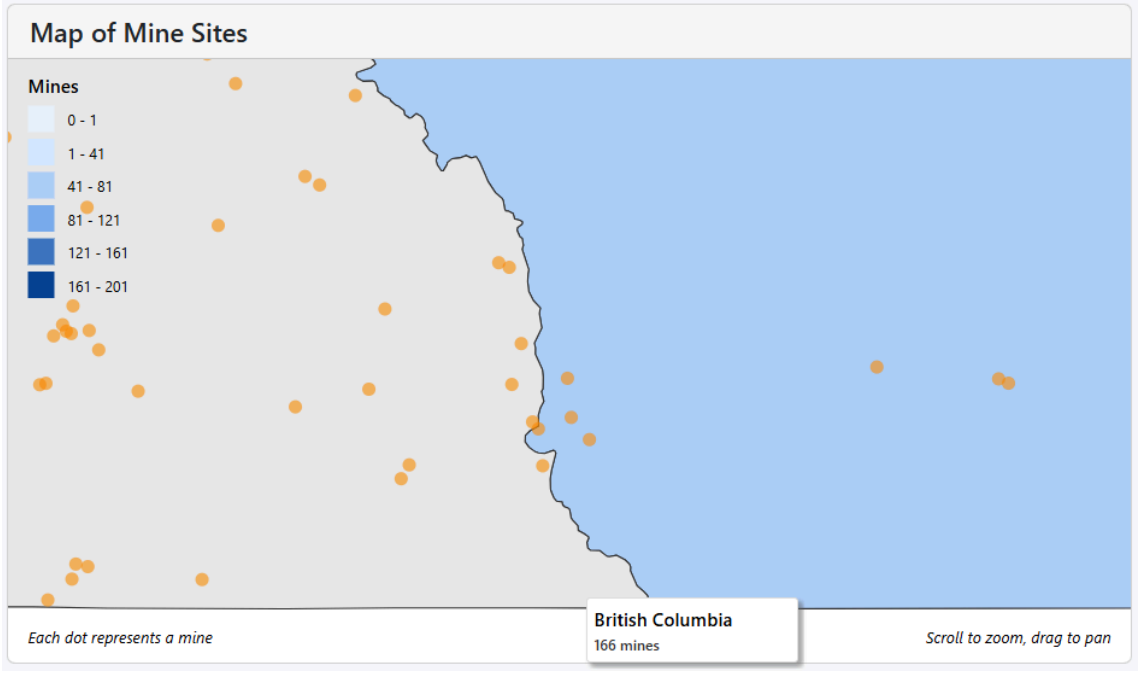
Mine Filter

- On a click of mine, the point becomes selected and the map zooms in on it. Additionally, all other points are hidden, and all other views will only show properties of the selected mine. Clicking the same mine deselects it and removes the mine filter. T4 includes finding the opening and closing times of individual mines along with their commodities and companies working on them, so this feature allows users to gather these details on a per mine basis.

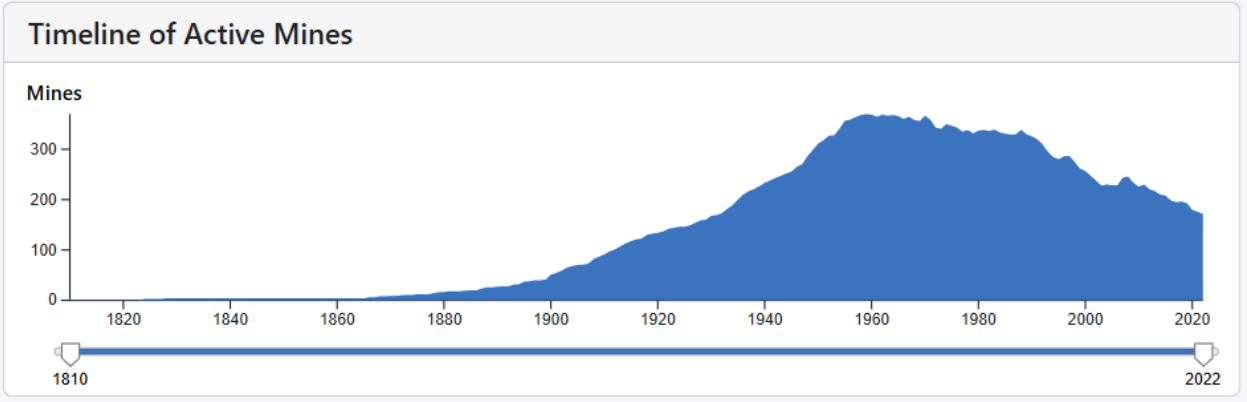


Zoom/Pan

- The map also has zoom and pan features, so users are not limited to viewing province-specific data. Certain regions may be more interesting than provinces, such as the Rocky Mountains that straddle the border between British Columbia and Alberta, so adding zoom and pan features allows users to explore those regions. This interaction addresses T2, T3, and T4 by allowing users to discover trends in geospatial regions that are not bounded by provincial borders.



View 2: Area Chart (Temporal View)



Marks:

- Poly mark represents the number of active mines within a time range.

Channels:

- *Horizontal position/order*: Year
- *Vertical position/order*: Number of mines
- *Vertical size:* Number of mines

Description/Rationale:

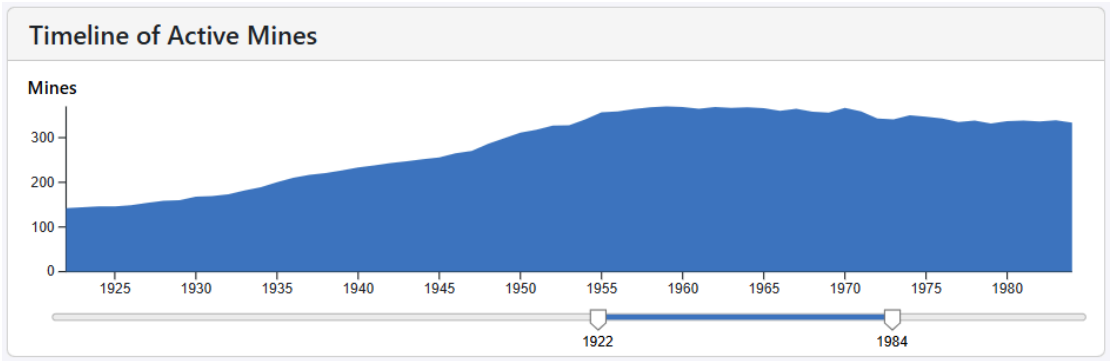
The second view is an area density chart that shows the number of mines that were active in a given year for a specific time window. This view allows users to see temporal trends in mine distribution and supports filtering through both interaction with the range slider and linkages with other views.

This view primarily supports the task of browsing the opening and closing times of mines (T1), but through linking with other views, it also serves to address T2 and T3. This chart was chosen over other options, such as histograms, because it allows users to analyze subtle changes in the number of mines and offers a continuous picture over multiple years instead of a bar that goes up and down. Following the principle of expressiveness, our choice of channels to use positions on a common scale to encode each numeric attribute (year and number of mines) on its respective axis allows for accurate comparisons and identification of trends. Although other idiom choices, such as line charts, can convey the same information, by using an area chart that has an inherent line chart built into it (top boundary is a line), we can add additional encoding (vertical size).

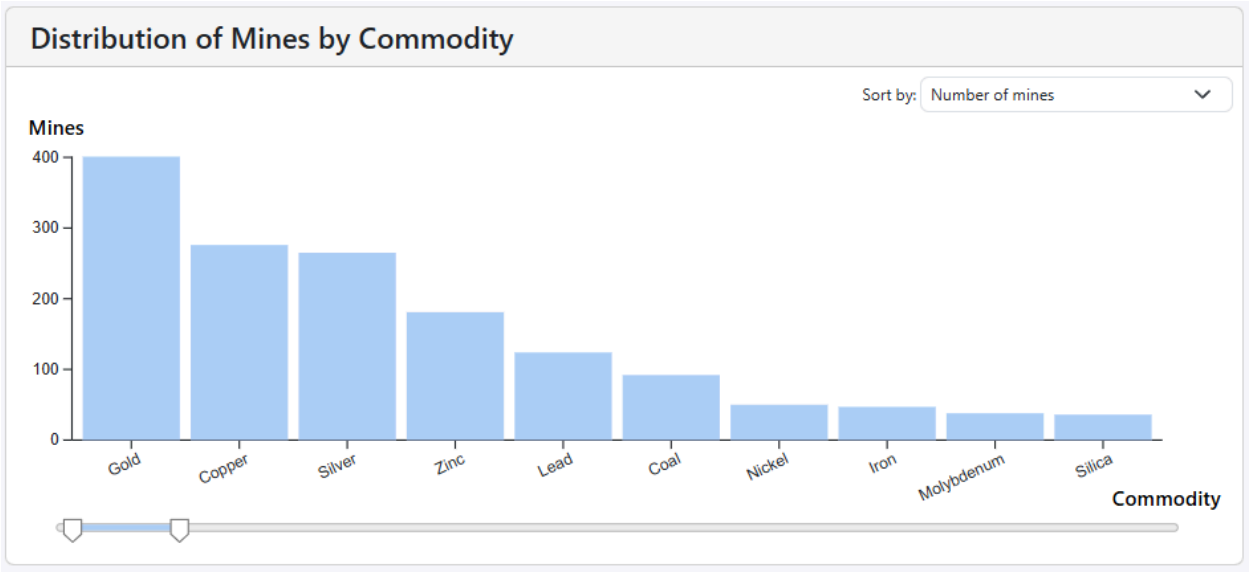
Interactivity:

Time Filter

- On change of the range slider, all views will only show data for mines that were active during that time range. The minimum and maximum values (year) dynamically update based on the earliest and latest year a mine was active, based on any other filters being applied. Through this interaction, users can address T1, T2, and T3 by viewing both macro- and microscopic trends in the temporal data.



View 3: Bar Chart (Commodity View)



Marks:

- Bar (segment) represents a commodity

Channels:

- *Horizontal position*: Commodity
- *Horizontal order*: Descending order based on the number of mines producing the commodity (or alphabetical by commodity name)
- *Vertical position/order*: Number of mines
- *Vertical size:* Number of mines

Description/Rationale:

The third view is a bar chart displaying the number of mines grouped by the commodities produced. This view allows users to see what commodities are most commonly produced, and through interaction, it allows for further exploration into mines producing a specific commodity.

The bar chart was chosen due to its ability to allow users to easily compare values across different bars by using length on the same scale. Compared to a scatterplot or line chart, the bar chart is a better fit given the data we are trying to visualize: one quantitative attribute (number of mines) and one categorical attribute (commodity), and it better addresses the task of comparison and lookup (T2 and T3).

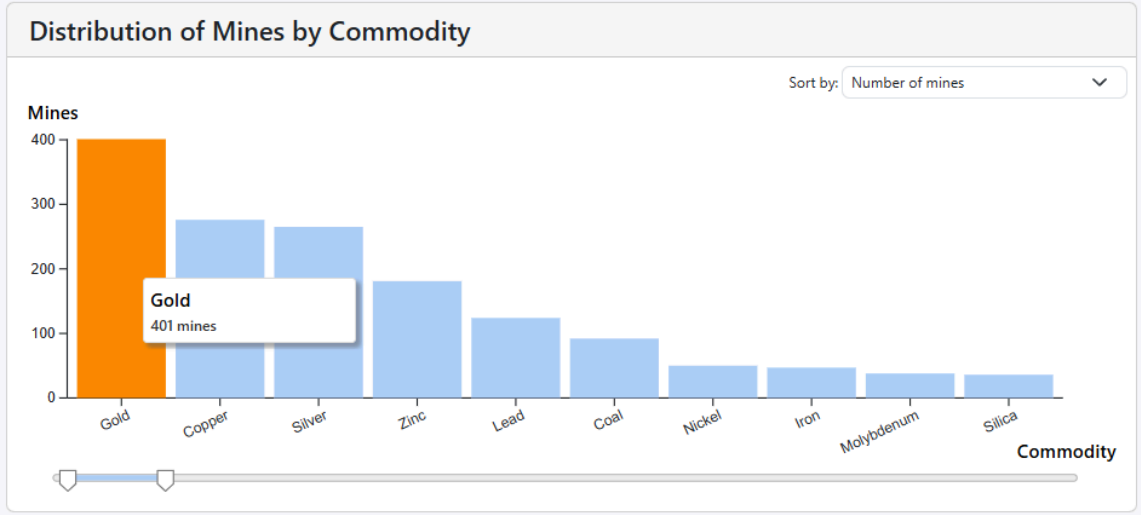
While we do provide a default value for the number of bars based on exploratory data analysis, we allow users to change the number of bars shown using a sliding window. In this fashion, users can either see the big picture (all commodities) or hone in on either a small subset of commodities or a specific commodity of interest. The bar chart can accommodate this variability, as it scales well (up to a reasonable limit), and due to the compactness of the chart, it can provide necessary information even while the number of bars increases.

By providing a default ordering of the bars in descending order of number of mines, we bring attention to commodities that are produced more frequently, as there are likely to be more interesting trends and information to be discovered through follow-up exploration. Furthermore, the use of ordering allows for easier comparison, as in addition to the relative lengths of the bars, the horizontal position can be used to associate relative ranks for each commodity.

Interactivity:

Tooltip

- On hover of a bar, the bar will become highlighted, and a tooltip will appear displaying the commodity name and the number of mines that produce it.

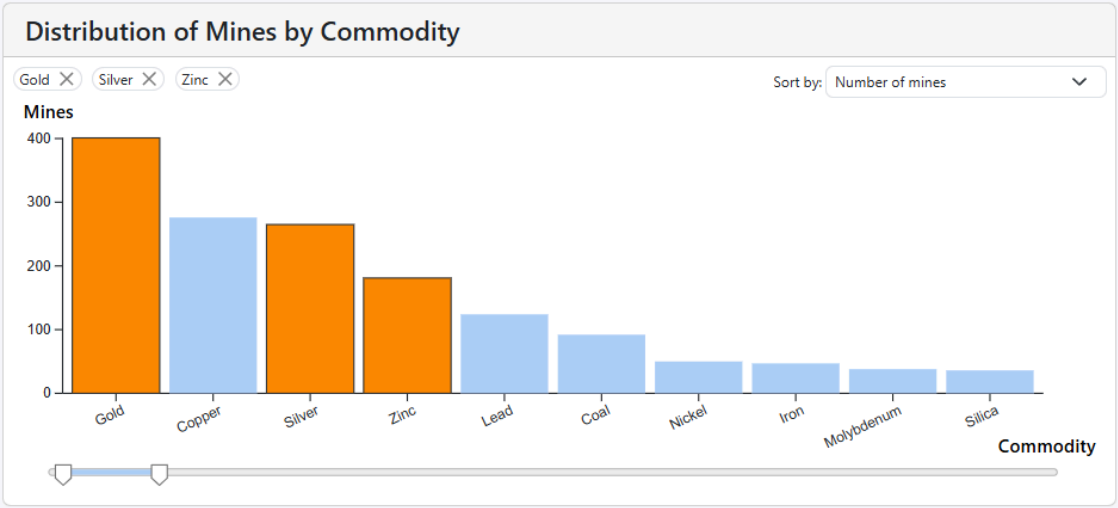


  Through redundant encoding, tooltips enhance the user's ability to gather precise data.

Commodity Filter

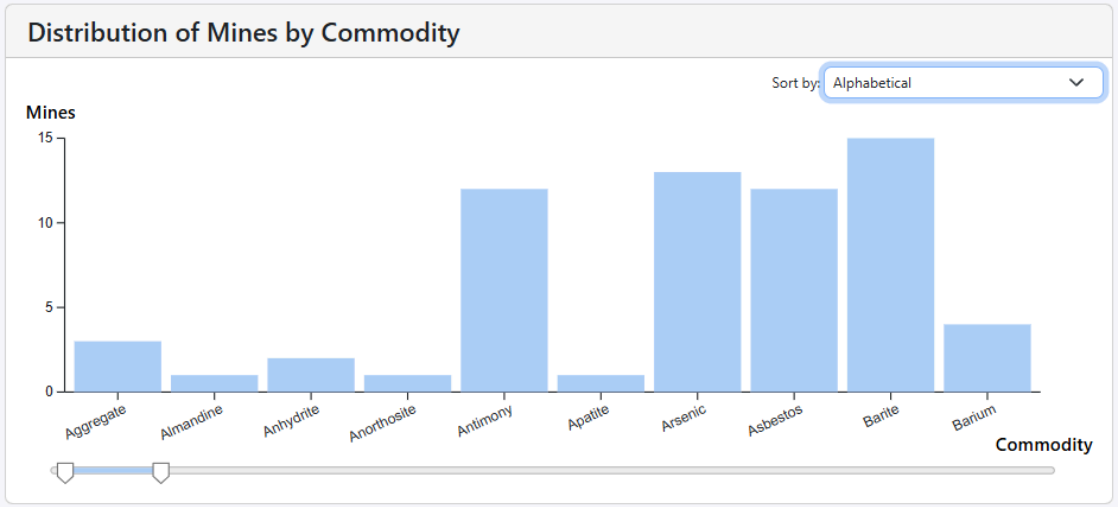
- The selected bar will become highlighted on a click of a bar, and all other views will be filtered to only show data for mines that produce the selected commodity. Selecting multiple bars allows users to find mines that produce any of the selected commodities. This interaction addresses T2 by allowing users to explore

the spatial distribution of specific commodities through the combined use of interactions in this view and View 1: Choropleth Dot Map.



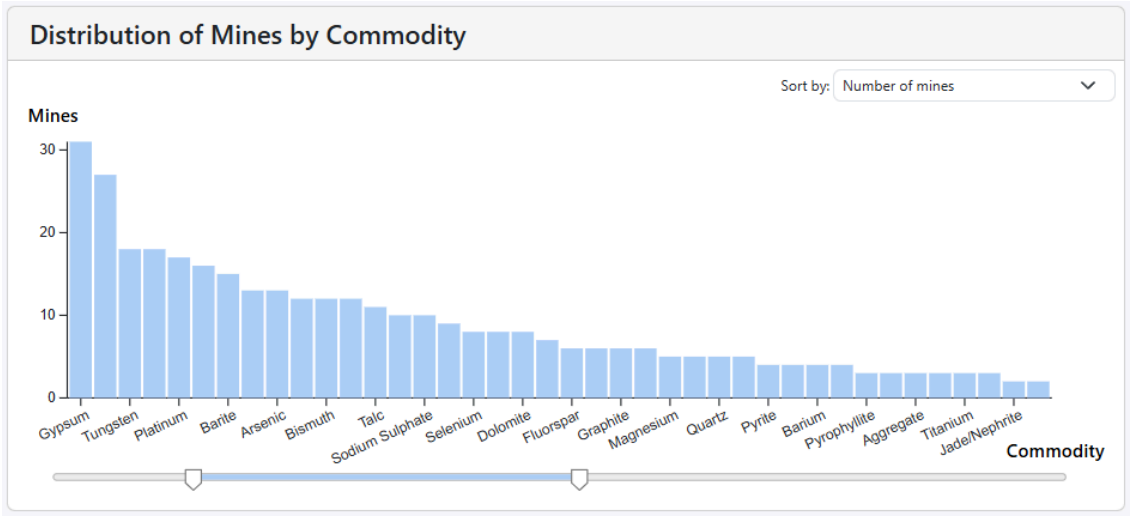
Sorting

- On selection of the “Sort by” dropdown, the chart will change the horizontal ordering of the bars. By default, the chart shows the number of mines in descending order. This interaction addresses T2 by allowing users to easily identify commodities of interest either by name or by the number of mines.

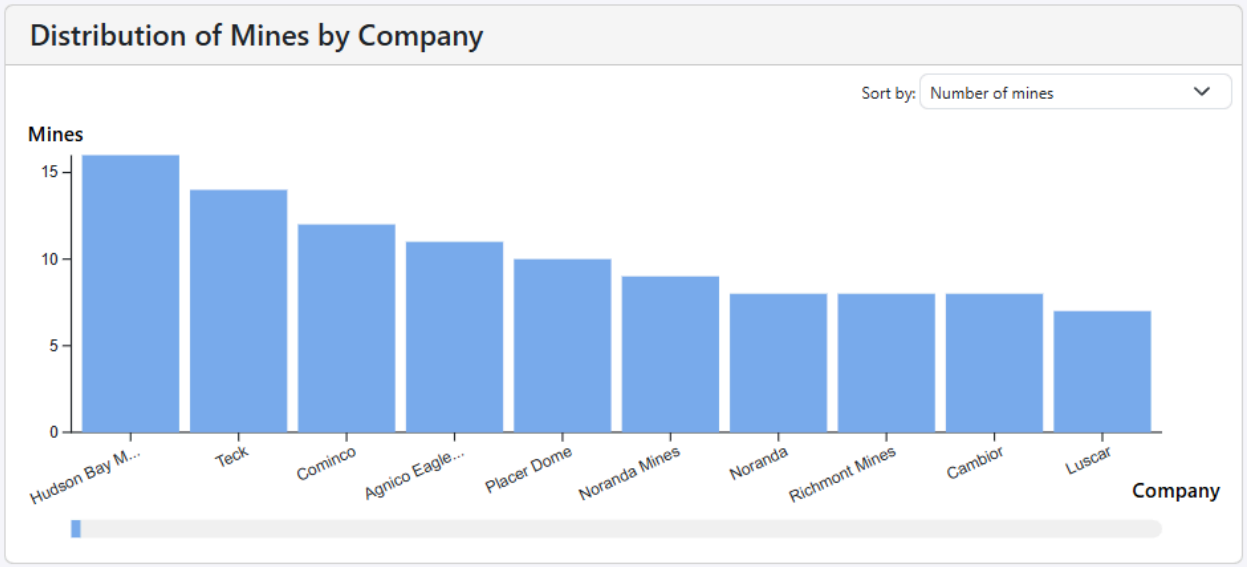


Commodity Visibility Slider

- On change of the range slider, this view will only show a subset of all commodities based on the start and end indices provided. More distance between the handles results in more commodities to be displayed and vice versa. This interaction addresses T2 by allowing users to explore different subsets of commodities that they may be curious to investigate further.


Distribution of Mines by Commodity
Sort by:
Number of mines
Mines
30
20
10
0
Gypsum
Tungsten
Platinum
Barite
Arsenic
Bismuth
Talc
Sodium Sulphate
Selenium
Dolomite
Fluorspar
Graphite
Magnesium
Quartz
Pyrite
Barium
Pyrophyllite
Aggregate
Titanium
Jade/Nephrite
Commodity

View 4: Bar Chart (Company View)



Marks:

- Bar (segment) represents a company

Channels:

- *Horizontal position*: Company
- *Horizontal order:* Descending order based on the number of mines operated by a company (or alphabetical by company name)
- *Vertical position/order*: Number of mines
- *Vertical size:* Number of mines

Description/Rationale:

The last view is a bar chart displaying the number of mines grouped by the companies operated by. This view allows users to see which companies own the most mines, and through interaction, it allows for further exploration into mines owned by specific companies.

Most of the arguments for this view are the same as the other bar chart for commodities, so we will only provide descriptions and rationales for any design decisions that are different. As mentioned in the earlier argument, compared to a scatterplot or line chart, the bar chart is a better fit given the data we are trying to visualize: one quantitative attribute (number of mines) and one categorical attribute (company), and it better addresses the task of comparison and lookup (T3).

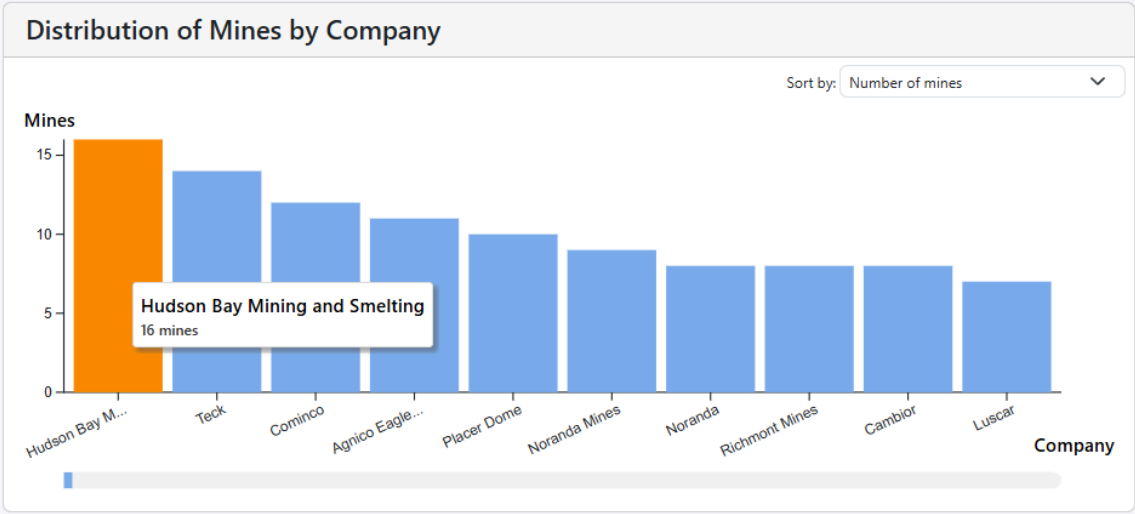
Since there are too many companies to be shown in a single chart (~1000), based on exploratory data analysis, we restricted the chart to only show at most 10 companies at once. By limiting the number of bars, we ensure that the chart stays compact and information-dense without losing out on interpretability and overcrowding. Moreover, by making the y-axis static, we allow users to scroll to see different subsets of companies while maintaining the same scale to ensure accurate comparisons.

Furthermore, by providing a default ordering of the bars in descending order of the number of mines, we bring attention to companies that operate the most mines, as they are likely to contain interesting trends. However, allowing users to change the order to instead be alphabetical order aids in the task of lookup (T3) for users who are trying to find a specific company.

Interactivity:

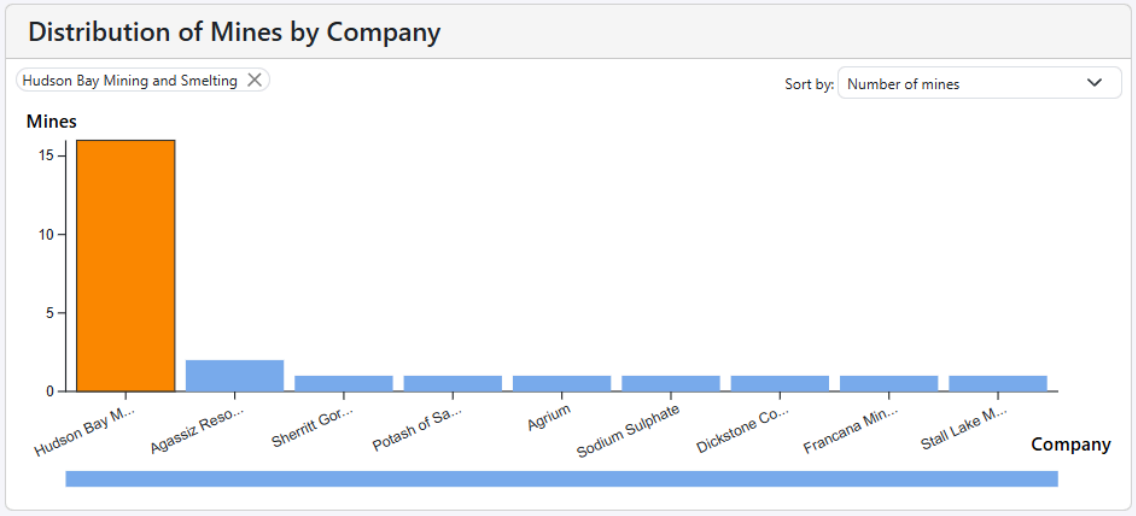
Tooltip

- On hover of a bar, a tooltip will appear displaying the company name and the number of mines that it was the owner of.



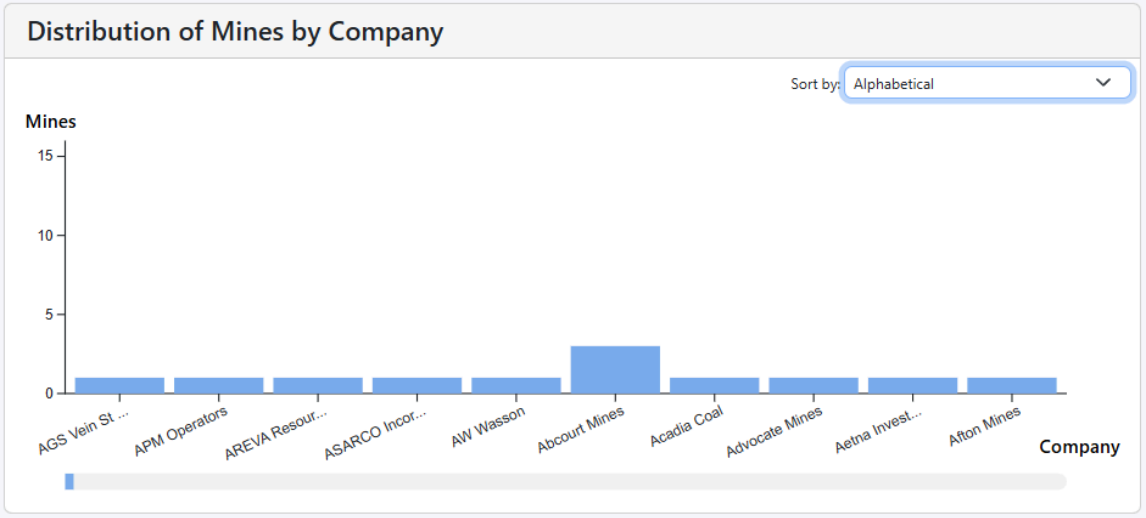
Company Filter

- On a click of a bar, the selected bar will become highlighted, and all views will be filtered to only show data for mines owned by that company. This interaction addresses T3, as through linking, users can explore other views and learn more about a specific company's geospatial, temporal, and commodity production trends.



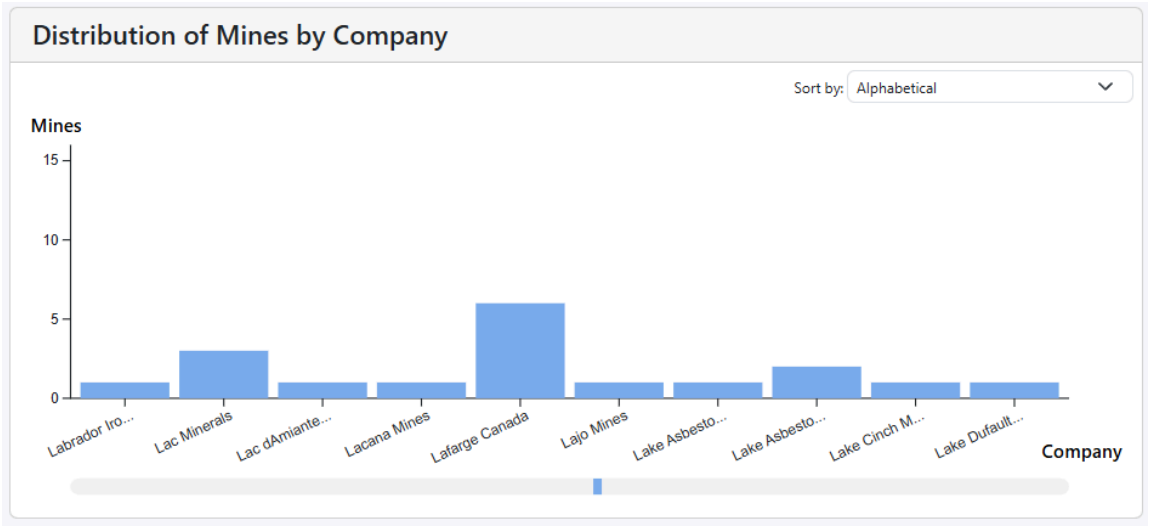
Sorting

- On selection of the “Sort by” dropdown, the chart will change the horizontal ordering of the bars. By default, the chart shows the number of mines in descending order. However, users can order the bars by alphabetical company names using a dropdown. The ability to sort alphabetically aids in the task of company lookup.



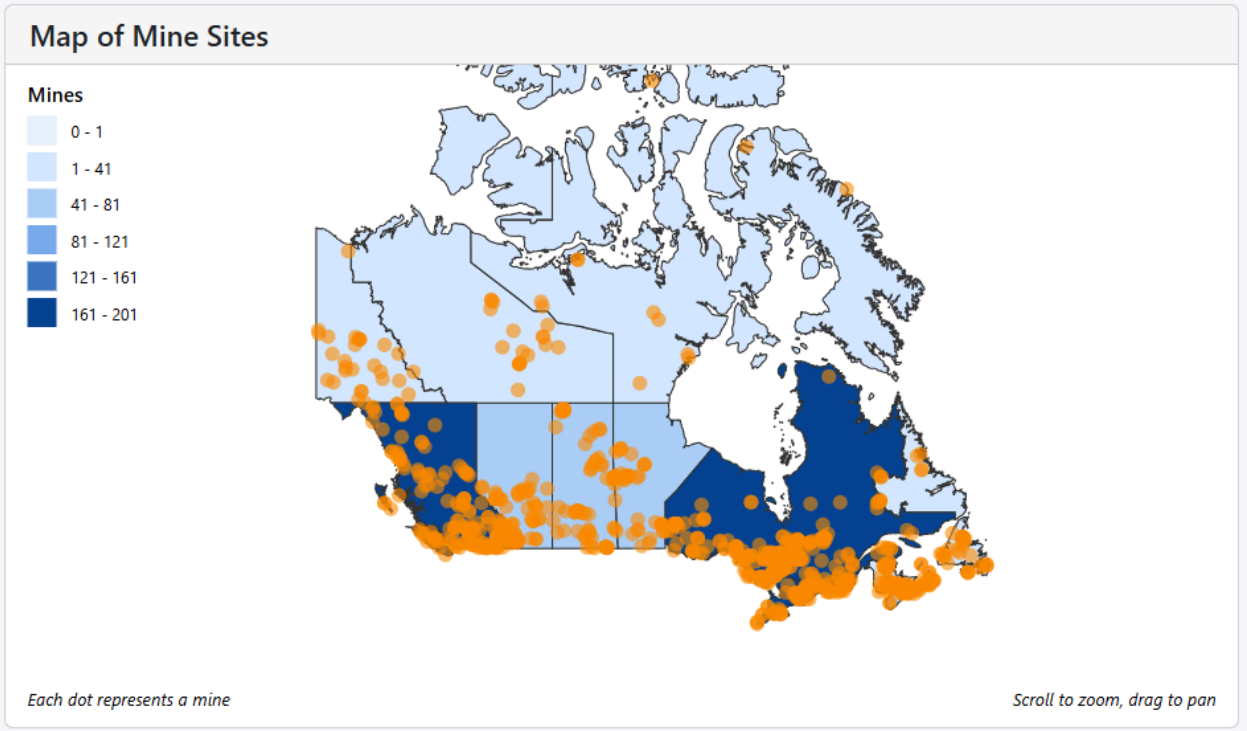
Commodity Visibility Slider

- On change of the range slider, this view will only show a subset of all commodities based on the horizontal position of the slider. The chart will show a sliding window of at most 10 companies at a time. The vertical axis remains static regardless of the companies shown, allowing users to accurately make comparisons based on previously seen data within the chart.

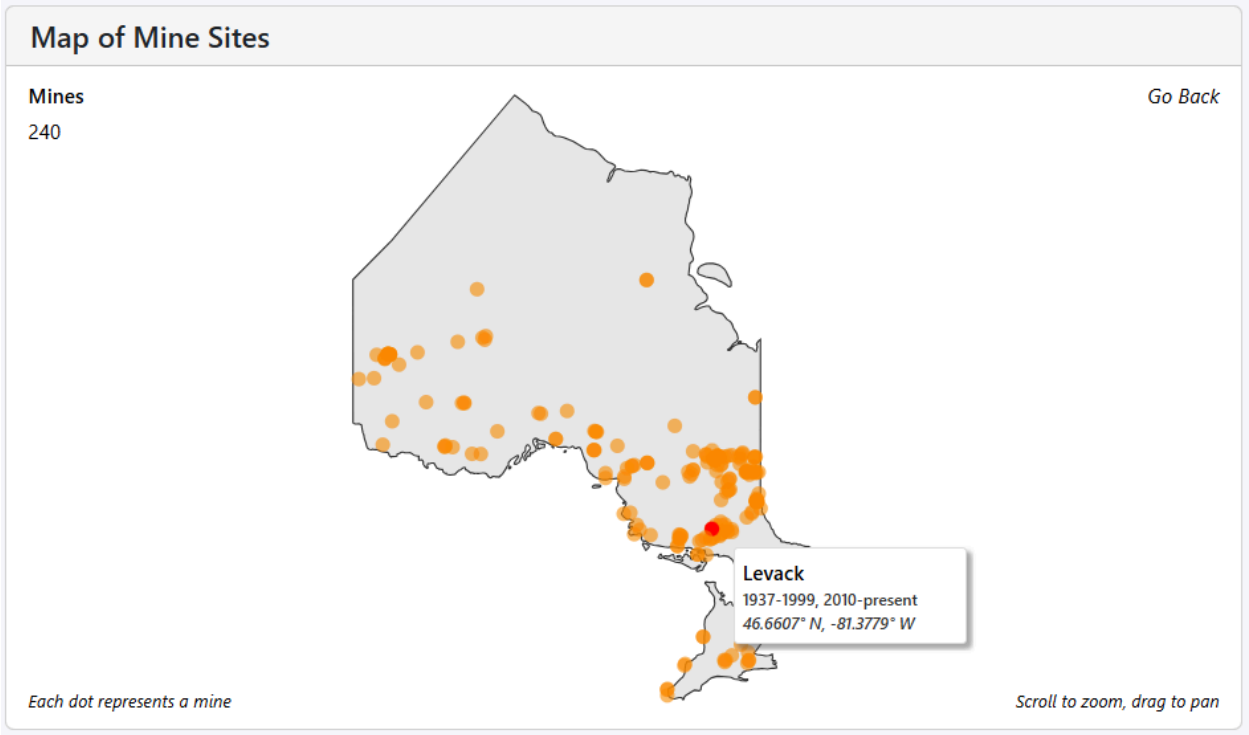

Distribution of Mines by Company
Sort by: Alphabetical
Mines
15
10
5
0
Labrador Iro...
Lac Minerals
Lac dAmiante...
Lacana Mines
Lafarge Canada
Lajo Mines
Lake Asbesto...
Lake Asbesto...
Lake Cinch M...
Lake Dufault...
Company

6. Usage Scenarios

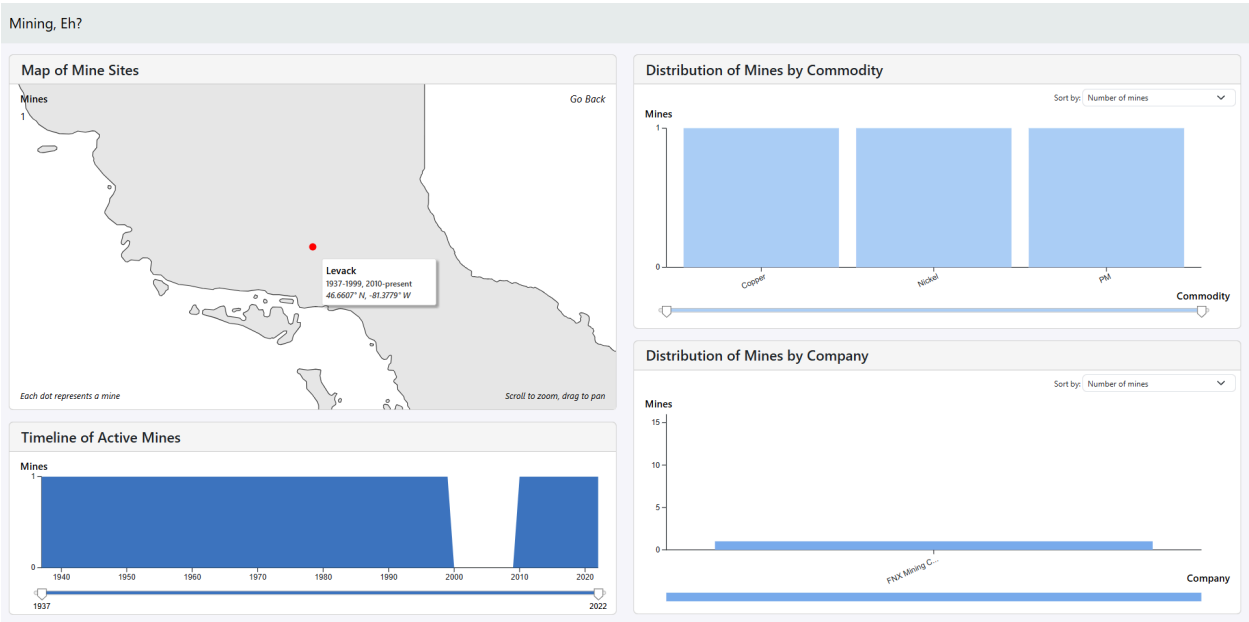
Mike is a mining engineering student and a casual mining enthusiast who enjoys spending his free time learning about natural resources and how the industry has evolved. He is looking for a tool that allows him to explore how mines are distributed across Canada, how this distribution has changed over time, which commodities are or were mined at these locations, and which companies have been involved in these activities.



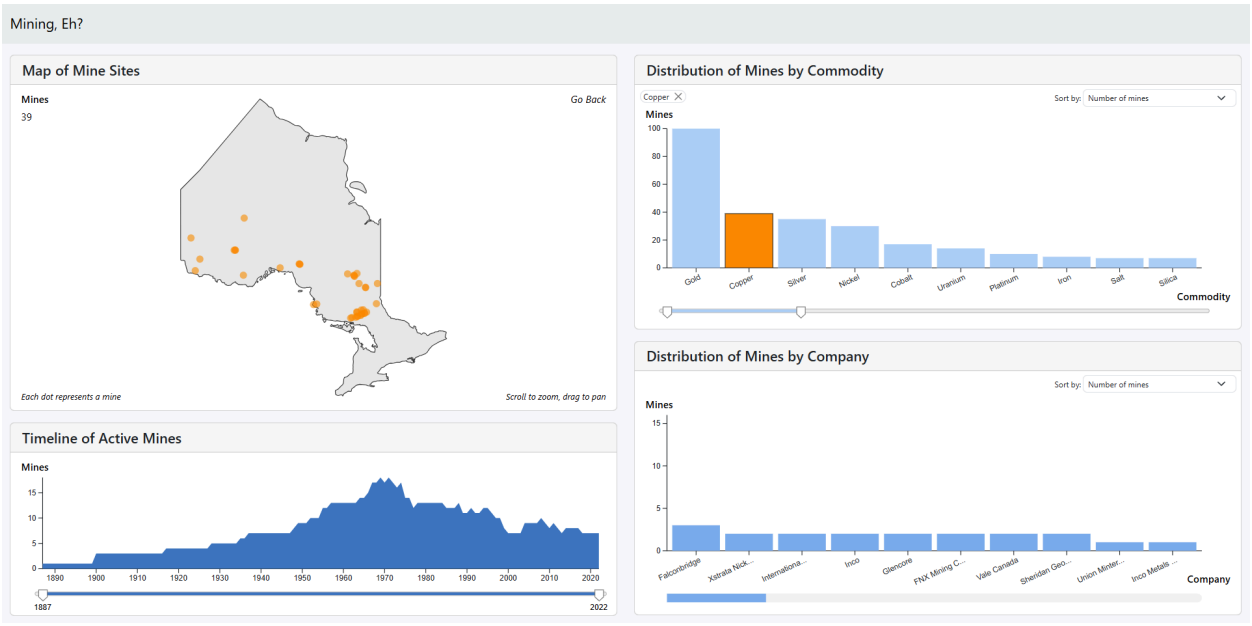
When Mike opens our tool, he is greeted by the choropleth map of Canada, “Map of Mine Sites”. This high-level view helps Mike quickly understand which provinces have had the most or least mining activity and how mine sites are geographically distributed. He can also zoom in and out and pan to better distinguish mines within high-density regions.



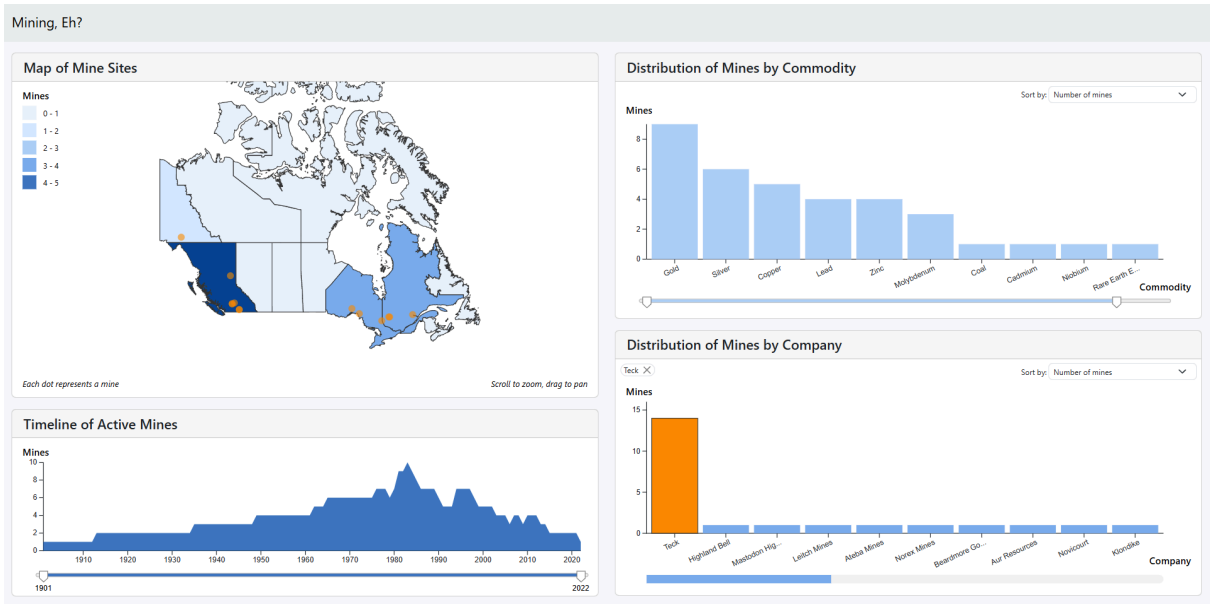
Ontario, the most saturated province, catches his attention. Curious, he clicks on it. The map zooms in to display only Ontario, showing just the mines within the province along with the number of mines in the top left corner of the view, plotted by their latitude and longitude. As he continues to explore, he hovers over various mines and learns more detailed information by reading the tooltips.



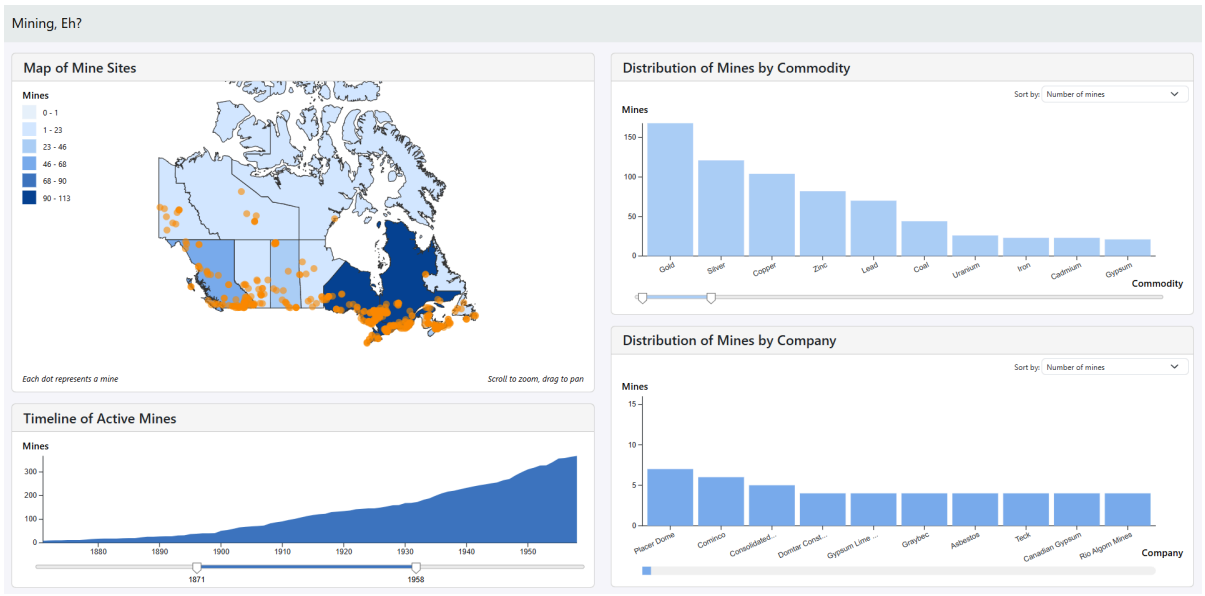
Intrigued to learn more about a particular mine (“Levack”), Mike clicks on it. From the bidirectional linking with the other views, he can learn who owned the mine (“FNX Mining Company”), what commodities it produced (Copper, Nickel, and PM), and interestingly, that it was closed between 2000-2009.



Satisfied with his findings, he deselects the mine and shifts his attention back to mining within Ontario. Curious about what commodities are produced in these mines, Mike turns his attention to the bidirectionally linked “Distribution of Mines by Commodity” bar chart, which shows the number of mines producing each commodity. Looking to learn more about copper mines in Ontario, he clicks on the copper bar, which highlights it, and further filters the map to display only copper-related mines in Ontario. This helps him understand where copper is typically found in Ontario. Repeating this process with other commodities allows him to see how various natural resources are distributed across the province.



Mike has completed a few internships at mining companies and notices some of them in the “Distribution of Mines by Company” bar chart. He wonders, “What mines are operated by my previous employers?” and “Which other companies have they worked with?”. By default, the bars are sorted in descending order of number of mines by default, which allows Mike to observe which companies are most active in the mining industry in the selected period. Spotting his former employer, he clicks on their bar. The map and commodity chart update to show only the mines and commodities associated with that company, while the company bar chart only shows the selected company and other companies it has collaborated with. This gives Mike a deeper understanding of company relationships and their mining operations.



Eager to explore how mining operations have changed over time, Mike resets all filters and turns to the ‘Timeline of Active Mines’ area chart. This timeline shows how the number of active mines has changed over time. As he adjusts the time slider, the map and both bar charts update dynamically, letting him observe shifts in mining activity across different periods of resource booms, regional expansions, and the rise or decline of specific commodities.

Through this interactive exploration, Mike can satisfy his curiosity and uncover meaningful insights about Canada’s mining history. He learns how mining has evolved, where specific resources have been concentrated, which companies have led the industry, and how commodity trends have risen and fallen across time and place.

7. Credits

During the development of our project, we drew inspiration from past hall of fame projects to benchmark what views and interactions were possible, and the level of detail required for the report. Moreover, we re-used boilerplate code covered in the D3 tutorials within the course for topics such as project and file structure, as well as facilitating interactions and bi-directional linking. We also used this publicly available GeoJSON file for our map of Canada (GitHub) and used Color Hunt to assist with the selection of our colour palette.