

# Process Book

## Visualizing Bilateral Trade Between Countries



Andrew Lee, Franklin Yuan

<https://github.com/candrewlee14/data-viz-project>

# Table of Content

<b>Team Information</b>	<b>3</b>
<b>Background and Motivation</b>	<b>3</b>
<b>Project Objectives</b>	<b>3</b>
<b>Data</b>	<b>4</b>
<b>Visualization Design</b>	<b>4</b>
<b>Must-Have Features</b>	<b>11</b>
<b>Optional Features</b>	<b>12</b>
<b>Project Schedule</b>	<b>12</b>
<b>Peer Feedback</b>	<b>13</b>
<b>TA Proposal Feedback (Pranav Rajan)</b>	<b>13</b>
<b>Data Processing and Models</b>	<b>14</b>
<b>Design Evolution</b>	<b>14</b>
First Iteration (Nov 10, 2022)	14
Second Iteration (Nov 11, 2022)	16
<b>Evaluation</b>	<b>18</b>
<b>Related Work</b>	<b>18</b>

# Project Proposal

## Team Information

Project Title: Visualizing Bilateral Trade Between Countries

Project Repository: <https://github.com/candrewlee14/data-viz-project>

Team Members:

- Andrew Lee, [candrewlee14@gmail.com](mailto:candrewlee14@gmail.com), u1256173
- Franklin Yuan, [franklin.yuan@utah.edu](mailto:franklin.yuan@utah.edu), u1369155

## Background and Motivation

When China first joined the World Trade Organization (WTO) in 2002, few people expected how the world's economic dynamics would be so fundamentally impacted by this event. In a span of just two decades, China's exports rose from \$333 billion in 2002 to \$3,548 billion (or \$3.5 trillion) in 2021, a whopping 965% increase. By 2020, China had overtaken the US as the world's biggest trade partner. But as China's trade relations with the world grew, so did tensions between the US and China. In 2018, the Trump administration launched its "trade war" against China by slapping tariffs on two-thirds of Chinese exports to the US. One of the US's main goals was to reduce its deep trade deficit with China which stood at \$481 billion before Trump took office.

Franklin is from China and Andrew is from the US. We both are very interested in the trade relations between the US and China. We are particularly interested in gaining more insight into the countries' bilateral export and import data both in gross volume and by product category over the years. This project is our attempt to create a visualization that makes the analysis on any two countries' bilateral trade data more accessible to the general public.

## Project Objectives

The visualization allows a user to select two countries in the world to study their bilateral trade relations. Through exploring the visualization, a user will be able to answer:

- How much did country A import from and export to country B over the years?
- What products are country A's top exports and imports with country B?
- How much of a specific product did country A export to country B in a given year?
- How much is the trade balance between country A and country B in a given year?
- How has the bilateral trade balance changed over the years?

The goal is to help the user develop a deeper understanding and intuition of the trade relations between two specific countries.

# Data

The dataset we are using will likely come from the World Trade Organization (<https://stats.wto.org/>). Countries periodically report their trade information with other countries. These trades are broken down into product/sector categories (Animal, Dairy, Chemical, Manufacturing, etc).

The bilateral import data they publicly offer includes the following information and more:

- Importing country
- Exporting country
- Product category
- Volume in US dollars
- Year

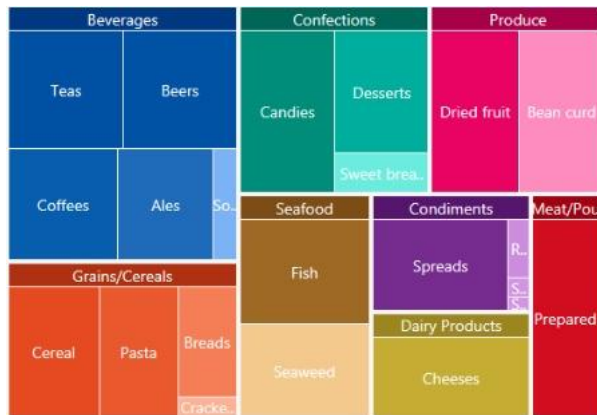
There shouldn't be a lot of data processing and wrangling involved in this project. The data from the World Trade Organization is highly structured, tabular data in CSV files with clearly defined types, so it should be relatively straightforward to use that data when making our visualizations. Most of the difficulty of this project will be on the frontend design and visualizations.

## Visualization Design

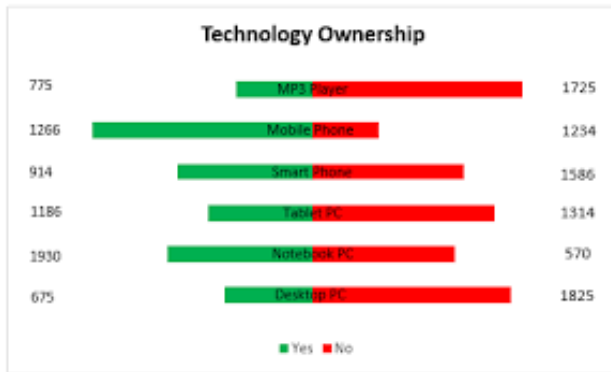
### *Brainstorming*

We first come up with many different types of visualization we can create to present the data.

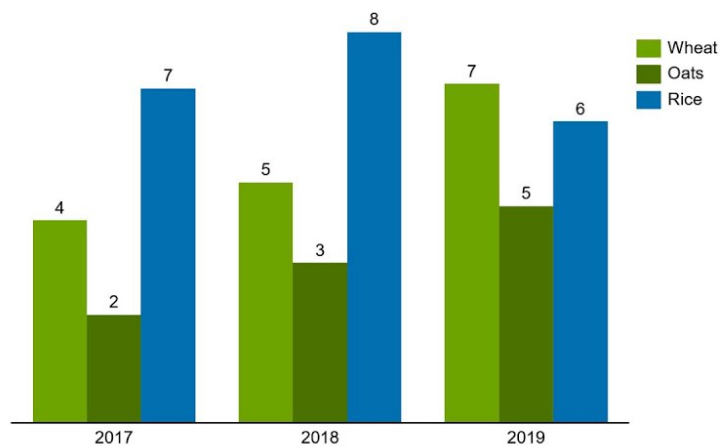
- A tree map can be used to show a given country's exports based on the volumes of product/sector categories.



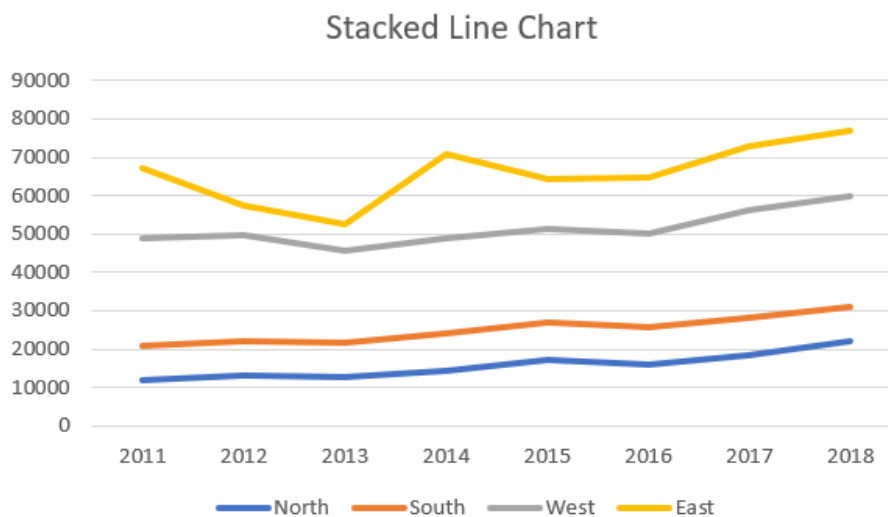
- The trade balance of different product categories between two countries can be represented by a centered stacked bar chart. We can rank these bars based on various criteria such as the trade surplus towards either country.



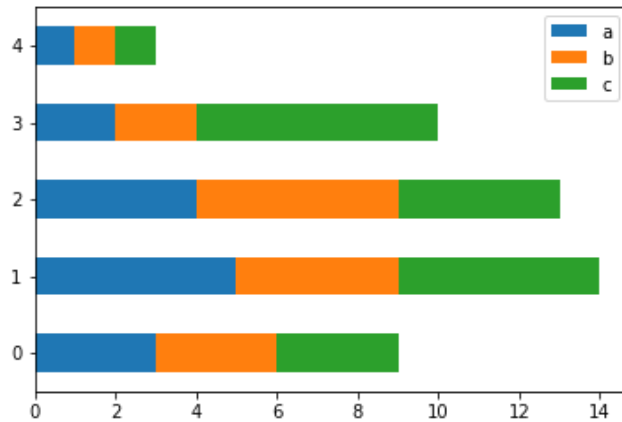
- Clustered bar chart can also be used to show the change in export volumes across different product categories over the years.



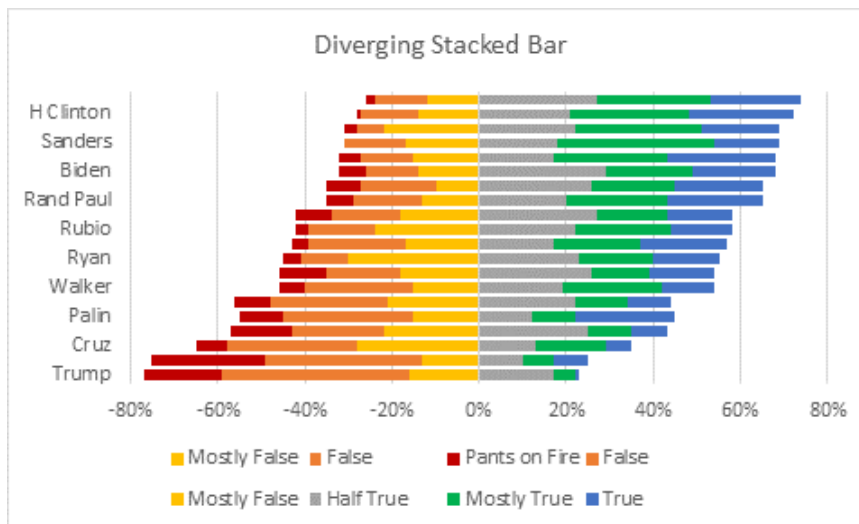
- We can also use a line chart to show how a country's exports/imports change over the years.



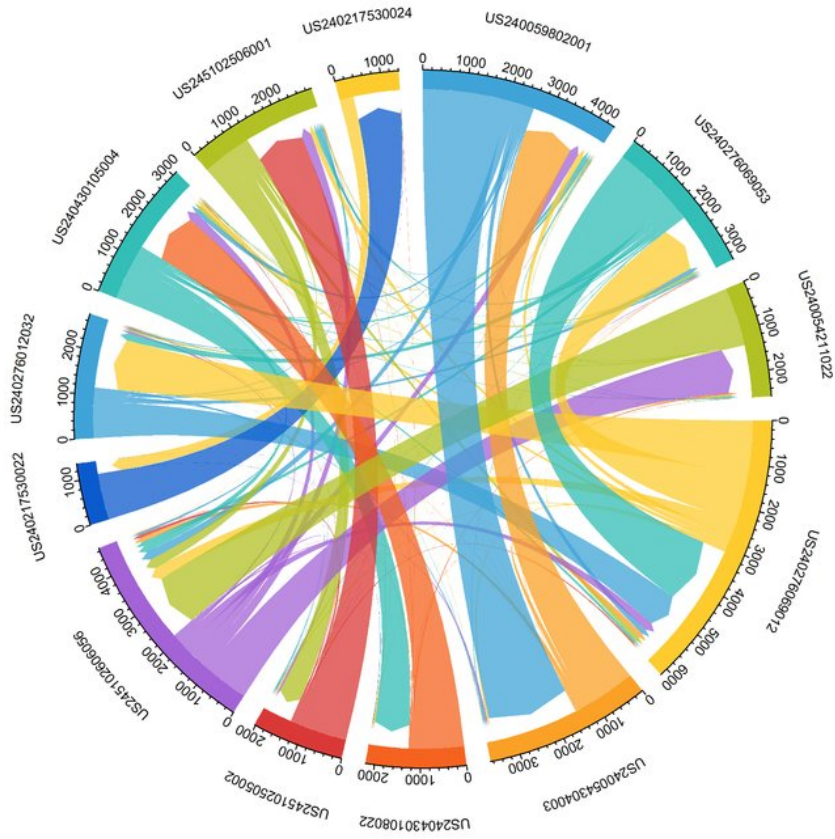
- A stacked bar chart can be used to compare two countries' exports/imports broken down by product category.



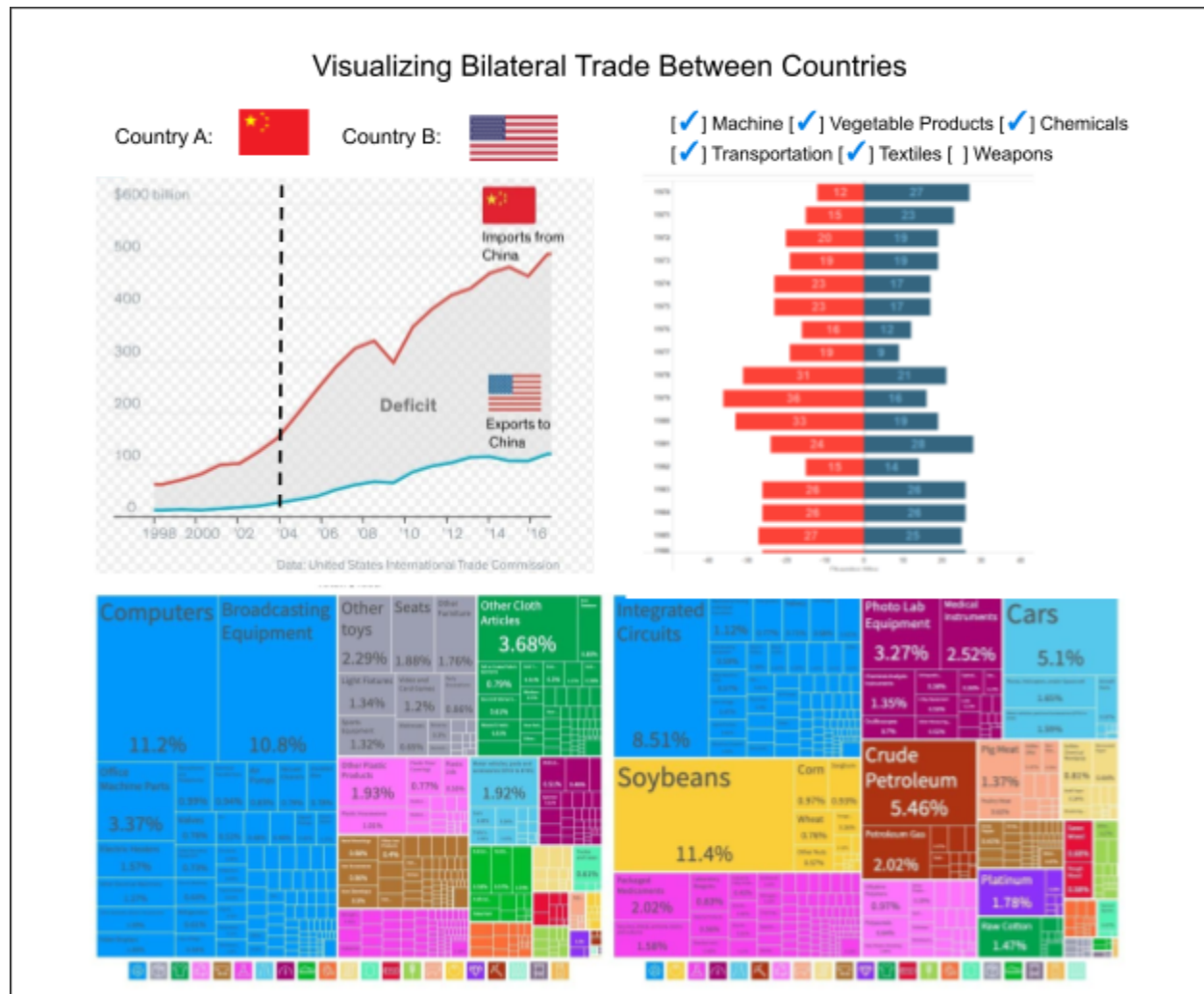
- A diverging stacked bar can be used to view multiple product's sub-categories.



- A country-to-country export/import chord graph (arrow indicates trade direction) if we want to see the trade relations between more than two countries. This can get a bit overwhelming with too many categories or countries.



## Prototype 1: Line chart, Centered Stacked Bar Chart, Tree Map



- A user can change country A and country B
- In this version, the main visualization will be the line chart where both countries' exports over the years are compared. If a user can select a specific year on the line chart which will automatically update the other views.
- Centered stacked bar chart: a user can select any number of product categories to show the trade balance. The bars can be reordered based on various criteria.
- Tree maps: one for country A on the left and one for country B on the right, showing each country's major export to the other country grouped by product categories. When the mouse hovers over a rectangle, more details will appear.

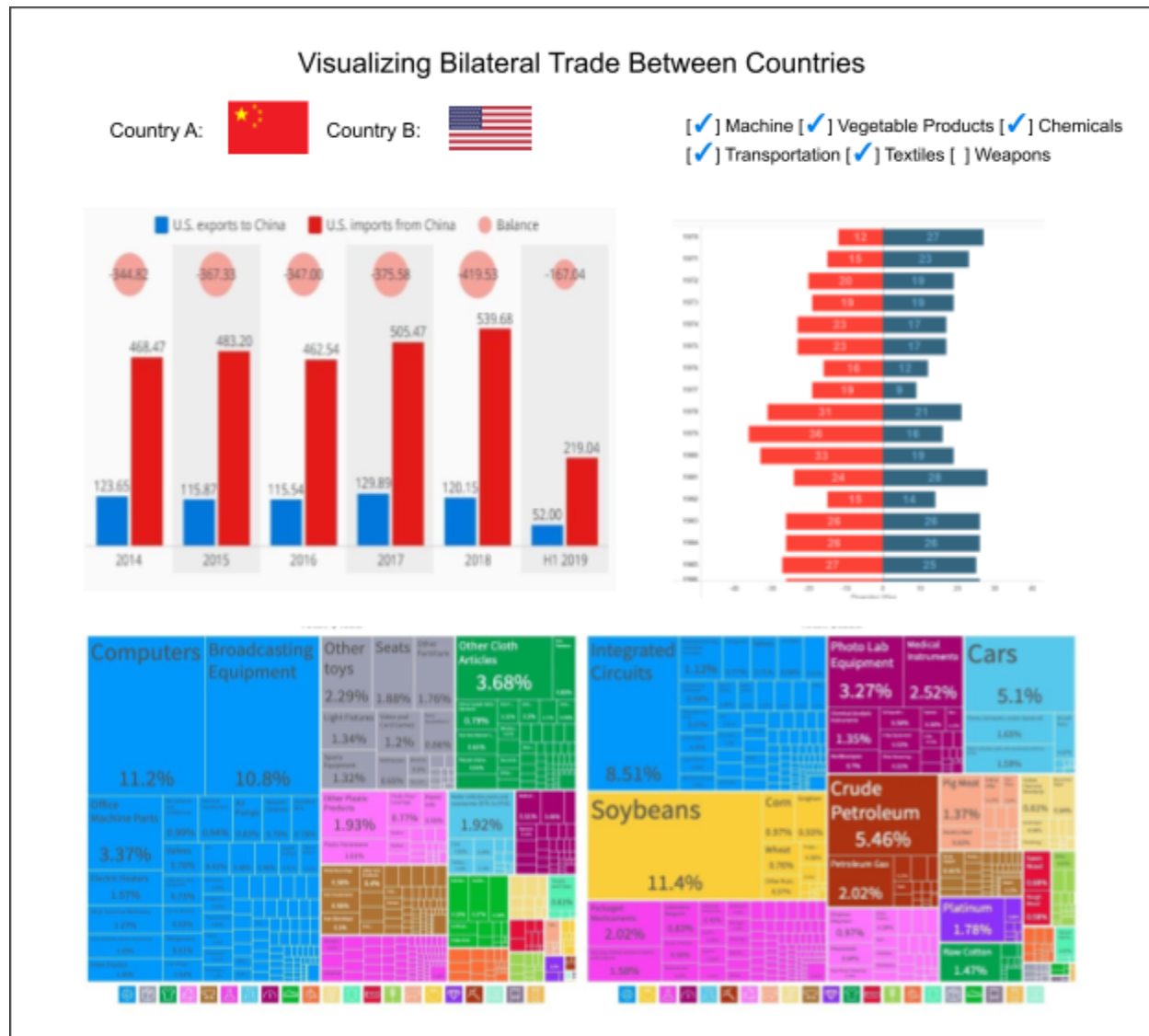


## Prototype 2: Line Chart, Sankey Diagram, Tree Map



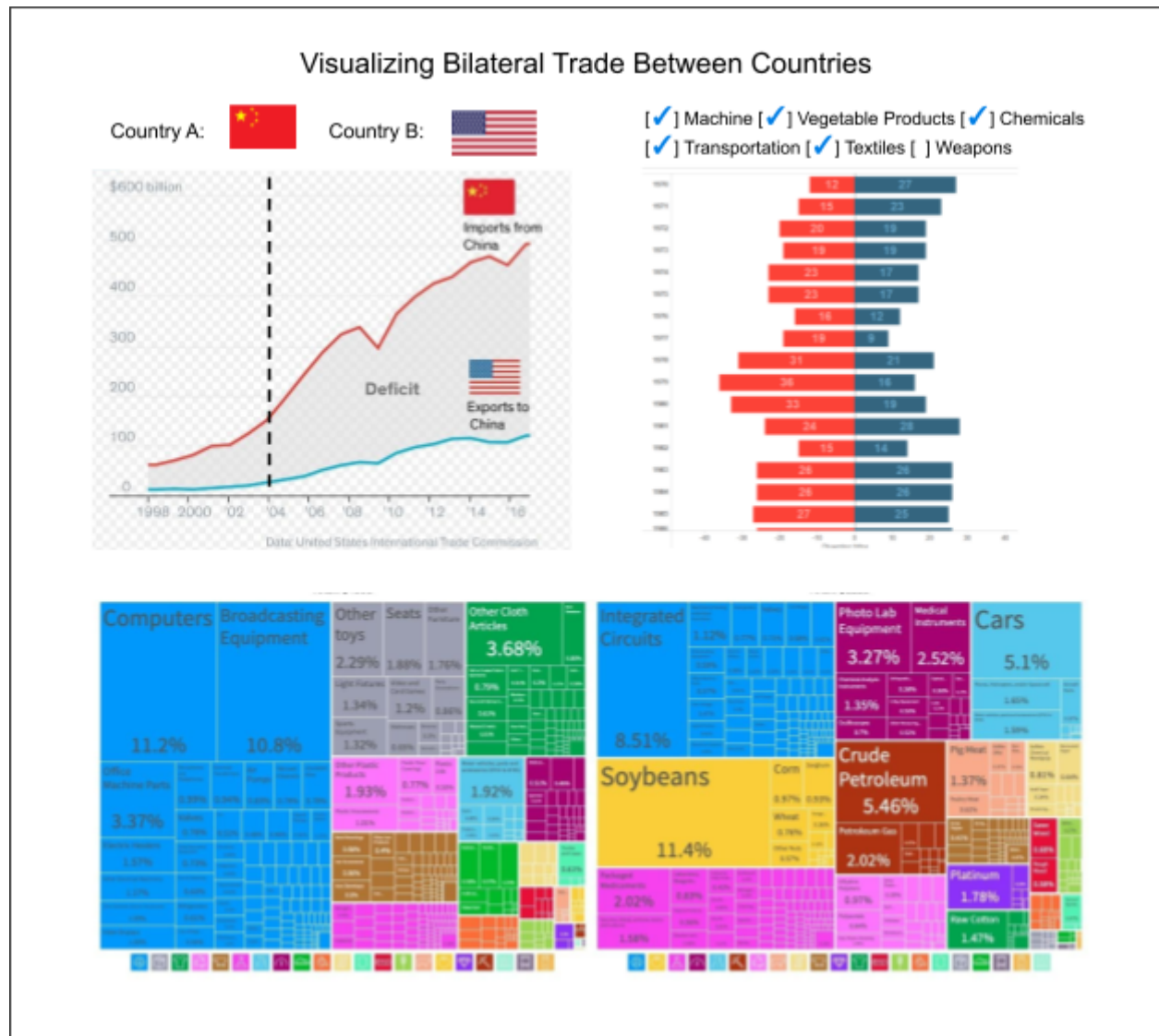
- The main difference between Prototype 1 and Prototype 2 is the visualization on the top right. Instead of a bar chart, Prototype 2 has a Sankey diagram that shows the export flows between the two countries grouped by product categories. A user can select and unselect product categories to show on the diagram.

### Prototype 3: Clustered Bar Chart, Centered Stacked Bar Chart, Tree Map



- The main difference between Prototype 1 and Prototype 3 is the visualization on the top left. Instead of a line chart, Prototype 3 has a clustered bar chart that shows the export changes across different product categories over the years. Like the line chart, a user can use the mouse to select which year to reflect in the other views.

## Final Design



- A time slider will be added to allow users to change the year of the data which will automatically update the Bar Chart and the Tree Maps.

## Must-Have Features

- Dropdown menu for choosing country A and country B
- A line chart
  - X-axis: year
  - Y-axis: trade volume
  - When hovered over, the other views will update automatically based on selected year in the line chart
- A centered stacked bar chart
  - Diverging color

- Filtering based on product categories
  - Option to reorder the bars based on certain criteria
- Two tree maps
  - Color encoding for different product categories
  - Hover to show details
- A time slider to adjust the year

## Optional Features

- Playable animations that show the change in bilateral trading over time. The boxes in the treemap could change shape, and the bars in the centered bar chart may shift sizes and rankings.

## Project Schedule

### Week 0 (October 16–October 22)

- Explore different datasets
- Prepare Project Proposal
- **October 21, 2022:** Project Proposal Due

### Week 1 (October 23–October 29)

- Data acquisition and wrangling
- Plan overall technical architecture and the website's framework
- **October 25, 2022:** Mandatory class meeting to discuss our project

### Week 2 (October 30–November 5)

- Implementation part 1

### Week 3 (November 6–November 12)

- Implementation part 2
- **November 11, 2022:** Project Milestone Due (a functional project prototype)
  - Code in its current state
  - Process book in its current state
  - Working visualization prototype

### Week 4 (November 13–November 19)

- Implementation part 3
- Mandatory project review with TA Mentor

### Week 5 (November 20–November 26)

- Debugging
- Refine website design
- Any other catch up

### Week 6 (December 27–December 3)

- Finish process book write-up
- Record screen-cast
- **December 2, 2022:** Final Project Due
- Finish peer evaluation

# Peer Feedback

*Peers: Kaden Hendrickson, Ethan Ramos*

## **Objectives:**

Both Kaden and Ethan were really intrigued by our project, especially by our [motivation](#) in doing it. They considered the project's scope appropriate; however, they did raise their concern in us putting too much data on the website. Our split between must-have and optional features seemed reasonable to them. As we showed them the dataset we found and the web framework we were planning to use (Svelte), they thought our week-by-week project schedule was well-thought-out.

## **Visualization:**

They really liked how we planned to use three different kinds of charts to visualize the data. Our main visual encodings include lengths and areas which represent the trading volumes and colors which match the major product categories. One suggestion they gave us was to only have one treemap and give the users more ways to interact with it and explore. Having two treemaps side-by-side might be too much.

## **Features:**

The main features we presented to them include a country selector, tooltips, a time slider that may play animations that show the change over time, and a product category filter. They considered our designed interaction meaningful and asked us whether all of the views are coordinated which we confirmed.

## **Our response:**

We appreciate the feedback Kaden and Ethan gave us and definitely want to make sure that we present the right amount of data to the users. We will consider trimming the number of treemaps from two to one as we implement our design.

# TA Proposal Feedback (Pranav Rajan)

This is a very cool project and I'm excited to see what insights you find in your final project. Your background and motivation is solid and I think you have enough metrics and variety of metrics to create different kinds of visualizations and a dashboard. I really liked your visualization designs and the final visualization layout and the timeline of completion is well documented. As you build your project you might want to think about what kind of interactions and widgets you want your audience to use for updating and interaction with your visualizations to gain insight about the data you have gathered. Some questions that might help are: who is the audience you are designing for (policy and economic experts, people curious about China-US trade relations etc) and what kinds of widgets you want people to use to interact with your visualization tool to find interesting insights. I think you guys are well on your way and I'm excited to see your progress at the next milestone. Good job and good luck with implementation!

# Implementation

## Data Update

We initially planned to use data from the World Trade Organization, but found a cleaner data set from Harvard University's Dataverse (<https://dataverse.harvard.edu/dataverse/atlas>). The trade data has already been cleaned by Harvard researchers; any inconsistent reporting has been resolved which is why we decided to go with this one. It uses a trade classification system called Harmonized Systems (HS) which groups all products into 10 major categories. These main categories also break down to a more detailed level which will give the users more insight into the bilateral trade.

The original data set is stored in several dta files and contains data fields, such as ISO country code and Product Complexity Index, that we don't need. After downloading the dta files, we wrote a Python script to convert the dta files into csv format and dropped all needless columns in order to reduce the data size. Though the trade data is available from 1995 to 2020, to keep it to a reasonable size and limit the website's startup time, we decided to only use trade data between 2010 and 2020.

## Design Evolution

### Framework and Language

We chose to use the Svelte framework because it speeds up development of good component-driven UIs (compared to vanilla HTML & Javascript) without adding too much bloat (compared to React/Angular). Reactive variables and blocks make it easy to re-run code whenever data changes, which allows components that share data to refresh dynamically on filters, sorts, brushes, etc.

We also chose to use Typescript within Svelte. Typescript allows us to confidently write code that modifies and maps data, knowing the output types are what we would expect when passed into other functions and components as properties. This makes it easier to do the complex data manipulation necessary in a visualization project.

### First Iteration

#### *Country Selector:*

- We wanted an easy way to select two different countries, and settled on a searchable dropdown list. We were able to use a packaged Svelte component for this and style it to our liking.



**Country 1**

United States of America ×

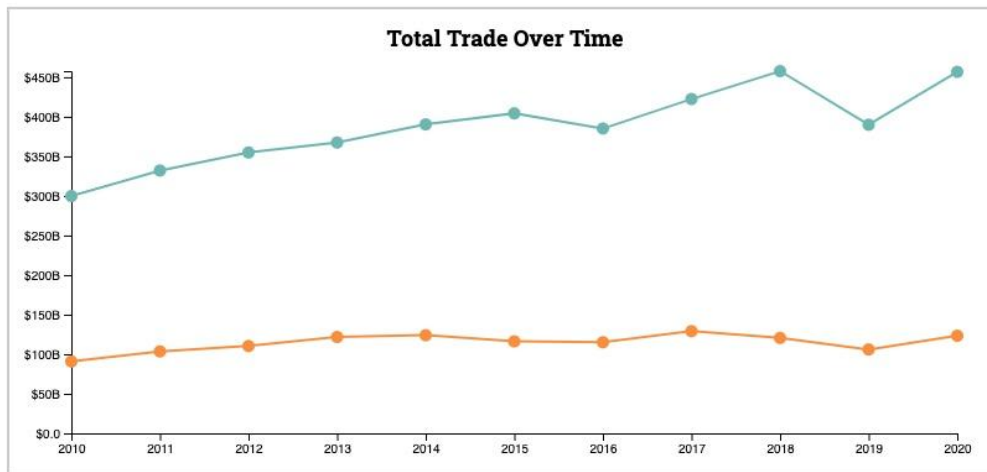


**Country 2**

China ×

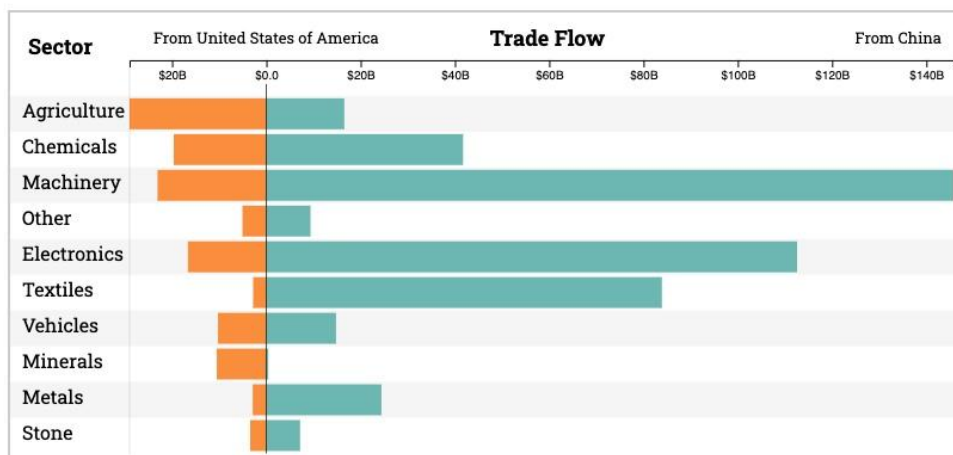
#### Line Chart:

- We created a line chart to give the user a chronological view of the trade between the two selected countries. The line color matches the color we assigned to the country in the selector.
- The x-axis ranges from the year 2010 to 2020; the y-axis represents the trade volume and updates when we select different countries.



#### Stacked Bar Chart:

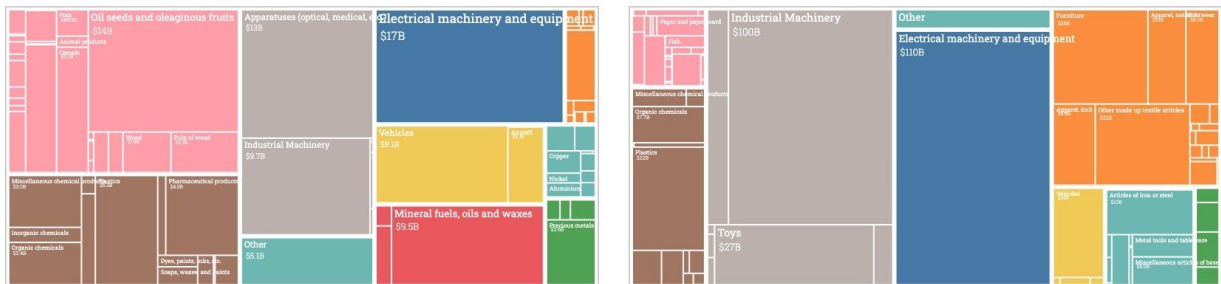
- We created a stacked diverging bar chart to help users compare the bi-directional trade volume across all major product categories.
- For now, we were only using the trade data from 2020.
- The dividing line moves as data changes to allow the best use of space.



#### Treemaps:



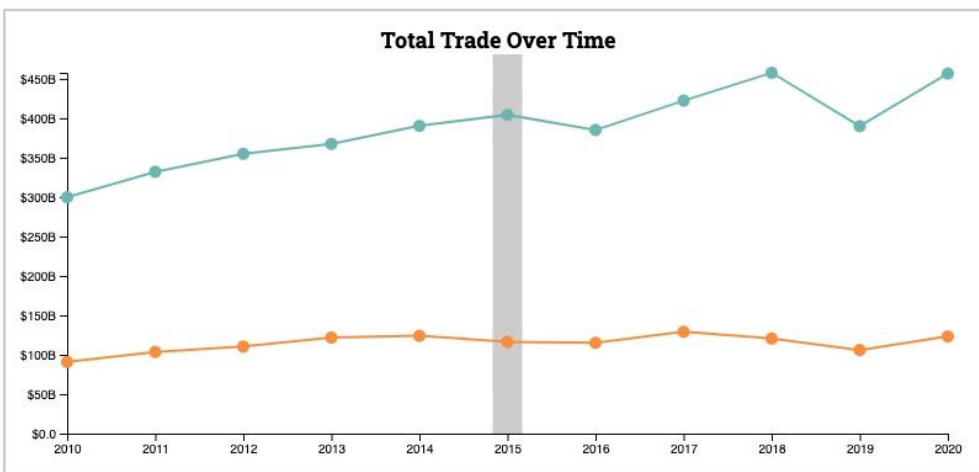
- The side by side comparison allows the users to see the composition of a country's exports to and imports from another country.
- It took us a while to get the treemaps right as we are using a more drilled down version of the trade categories.
- We hadn't gotten the size of the texts quite right yet.



## Second Iteration (Project Milestone)

### Line Chart:

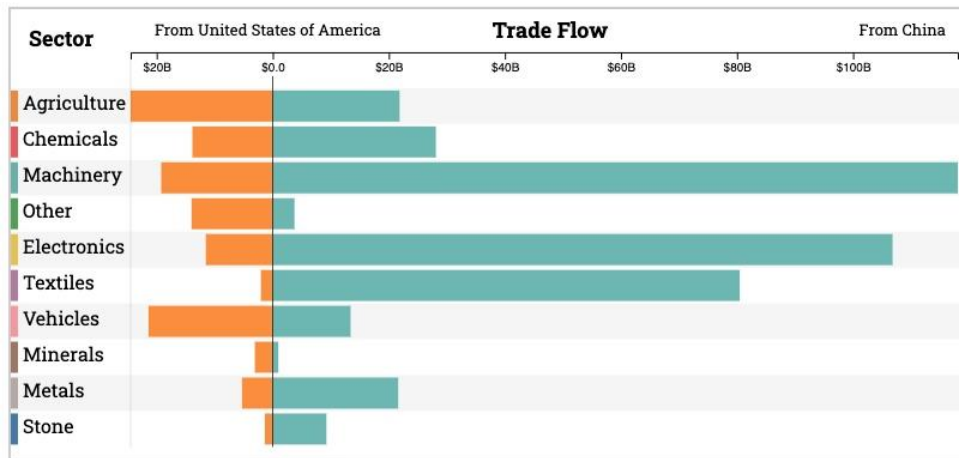
- We added an interactive feature to the line chart. A click on the data points moves the gray bar to a specific year. This updates the other views (stacked bar chart and tree maps) to the data of that specific year.



### Stacked Bar Chart:

- We color coded the product categories and rendered the legend on the left side of the chart.





### Tree Maps:

- The color encoding from the stacked bar chart is also reflected by the rectangles in the tree maps.



### Next Steps:

- All: add a tooltip and hovering feature to all charts to show details
- Line Chart: add an option for users to choose a layer either in total \$ value or share of the overall export/import volume
- Stacked Bar Chart: add ordering that enables the users to order the product categories alphabetically (asc-dsc), by net trade value (asc-dsc), by gross export value (asc-dsc), and by gross import value (asc-dsc).
- Tree Maps: add filter to allow the users to exclude and only show certain product categories on the tree maps

## Evaluation

*What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?*

## Related Work

1. Harvard University's Atlas of Economic Complexity website (<https://atlas.cid.harvard.edu/explore>)
2. OEC World (<https://oec.world/>)