U.S. Border Crossing Entry Visualization

WU Zhiwei XU Xiaohan Cai hairui Liu ran

Computer Science Engineering Department

Hong Kong University of Science and Technology, Hong Kong

Abstract

Import trade is an important part of national development concerns. How to visualize as comprehensive information as possible in a visualization project is a very valuable task. Thus, we present the ImportNetwork to provide users an overall view, including the value, trend and transport etc. Besides, users can get the corresponding information through different time and place of the mouse. A case study is conducted to demonstrate its accessibility and expressiveness.

**Keywords**: Import, Border Crossing Entry, visualization, interactivity.

# Introduction

## Background

In recent years, research has shown that imported goods can have multiple effects on a country. For example, imported food generally improve nutritional, health and demographic outcomes[1]. And it will affect the prices of related products in the entrance area and even the regional economy[2]. Besides, protective import tariffs generally lead to economic-environmental trade-offs. Therefore, research on imported goods is particularly important. In this project, we combined the characteristics and types of data to design a complete and comprehensive visualization project. The inner relationship of data about the United States can be displayed intuitively.

## Data

The Border Crossing Entry data set we used is from Kaggle.com [3]. The Border Crossing Entry data set takes 35MB of memory size and contains 346734 listing data samples with 8 columns, which spans from 1996 to 2019. These columns in the data set including Port Name, State, Boeder, Measures, Value, etc. The data set contains many different type of data, including integer, float, string, date, list, etc.

## Contributions

This project is therefore proposed for more comprehensive visual analysis of Border Crossing Entry. Specifically, we designed a fully functional system, which shows the changes in the value of imported goods on each states over the years and the relationship between the states and ports that import goods in different years. Our contributions could be summarized as follows:

• A new interactive visual design for Border Crossing Entry analysis, the information displayed by which will change according to the user's hover point. Users can see the change of data very intuitively.

• A comprehensive understanding of the Border Crossing Entry. Our system can not only show the value of imported goods, the port of goods, the method of transportation, the trend, but also compare and display data between different states.

# Related work

First of all, we did the data cleaning and worked on Feature Engineering. Data processing is one of the most important jobs before the visualization of data. We had spared no effort to finish this step. We used a full range of feature engineering to process the data set, including but not limited to the selection of data columns and multiple encoding methods of data columns. Secondly, we designed our project by reviewing the content in class and draw lessons from the methods of excellent visualization projects over the years.

## Feature Engineering

### Drop Useless Columns

We removed two columns: Port\_Code, Location. There is a one-to-one correspondence between the Port\_Name and the Port\_Code. Considering that we need to show the name of the Port in the visualization project, we dropped the column of the gateway number. In addition, the Location shows the latitude and longitude of each state in the United States. In the coding of this project, we use D3.js to display the various continents in the United States. We don't need to code the latitude and longitude, so we dropped this column out.

### Normalize

Due to the large time span of our data set, accurate date display is meaningless and not feasible. So we convert the date in the A column of the dataset into a year. In the following visualization process, we will show the Border Crossing Entry data of the US according to the year.

## Feature Encoding

For the Measure whose feature values do not have an ordered relationship, we use one-hot encoding to construct a vector with the same size with the number of unique value in these columns.

# Design goals

Import trade has become an important factor affecting the economic and political of a country. From a lot of research on import trade, we find that the main concerns are as follows:

### T1: For a given state, where are the goods imported from? By studying the continents and the ports from which the cargo originates, we can grasp the distribution of ports connecting to a certain continent. Another task of this project is to visually show the changes in the value of imported goods on different continents.

### T2: How total value of imported goods changed by state over time. Another task of this project is to visually show the changes in the value of imported goods on different continents, includes the geographical location and the comparison between different states, etc.

### T3: How are imported goods imported? What percentage of each method? Visually showing the import methods and proportion of imported goods by each methods is also a very meaningful design for this project.

# Visualization design

We propose to build a visual system, named ImportNetwork, to complete these tasks mentioned in Section 3. Our ImportNetwork contains four views. The Network View intuitively visualizes the changes in the import volume of the states with larger imports over time. The Map View directly shows the distribution of the value of import on the map. Measure Statistics View, which serves as an assistive view for analyze the method and proportion of them. The last is a Sankey View, which vividly shows the relationship between imported ports and destination states. The rationales we follow and the concrete designs for the four views and the embedded interactions are introduced below.

## Design Rationale

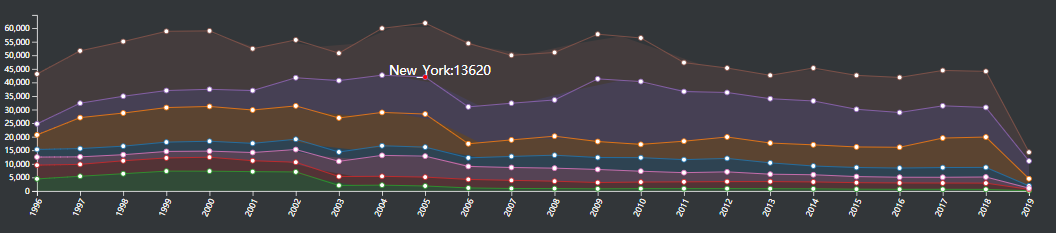
Based on task analysis, we have compiled the following design rationales:

### R1: Show the value and trends. As a research on network analysis of imported goods, we need to first determine and present the quantity of goods and the changes over time.

### R2: Present the relationship of ports and destination. According to the T1, for a continent, our project needs to show the origin of the goods.

### R3: Improve interactivity and intuitiveness. The system we design needs to provide user interaction, not just a single graph which shows statistical values. And on this basis, use appropriate coding methods and diagram types to display different types of data to improve the intuitiveness of the project.

## Network View



1. Network View for Stacked View, different states are represented by stack areas of different colors.

Stacked View is the first view in the network view, mainly for showing the trend of the total imported goods’ value, and making comparisons between different continents. We use the area of different color to represent the total value of imported good of different state in USA. And use the trend of the line to show the trend of how total imported value of different state change. When the mouse hovers over the point of the stack area, the name of the state as well as the total export value of the state will be displayed dynamically.

The years on the x-axis in this view can be interactive. When the selected years are different, other views will change accordingly

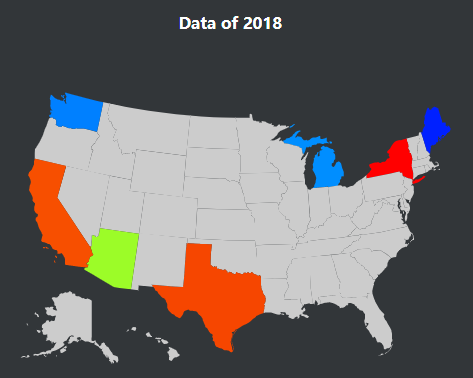


Figure 2: Network View for Map View, different colors represent different extent of total imported value. The gray areas are the states that are not covered in the dataset

Map View is also an interactive view. When the mouse hovers over a continent, it will change the color of the continent to white to highlight the selected effect and display the name of the continent at the same time. When the selected year in the Stacked View in Figure 1 changes, Map View will change the color of each state in the map to the data of Corresponding year. For example, the picture shows the data of year 2018 because in the figure one we chose hover the mouse on the data of 2018.

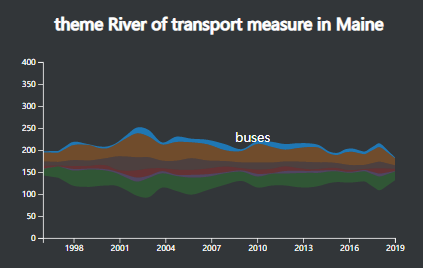


Figure 3: Network View for Measure Statistics View, different colors represent different measure of transportation.

The Measure Statistics View is a theme river chart, when the mouse hovers in a specific state in the Map View(Figure 2), the Measure Statistics View will show the composition of the measure of transportation of imported goods in that state. This figure shows the data of Maine State. We can see that buses account for a small part of the import transportation measures.

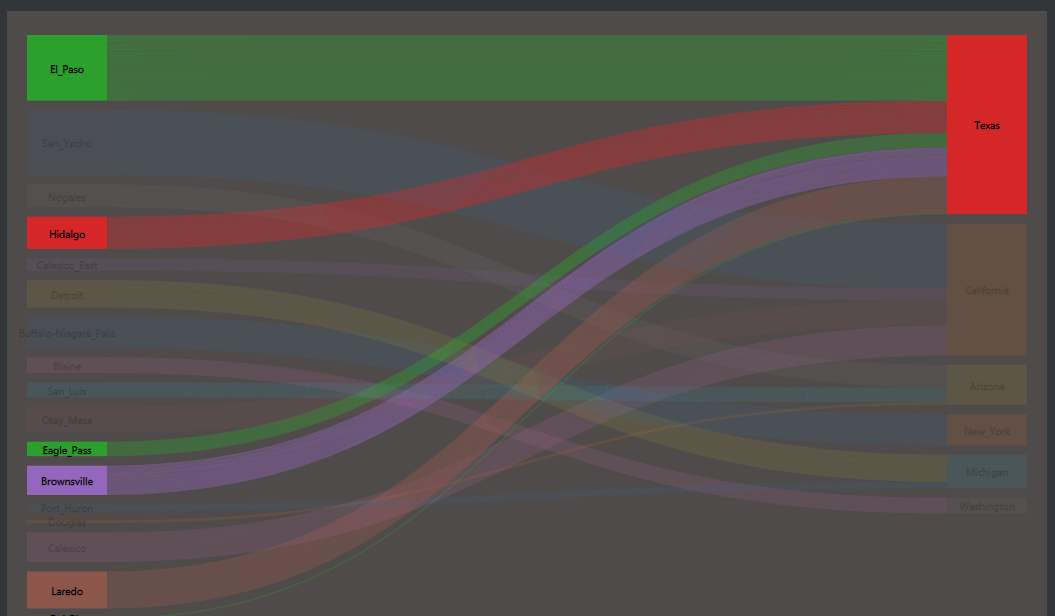


Figure 4: Network View for Sankey View

The Sankey View encodes the source of imported goods in each state, when hovering over a specific state, we can see the main port of origin of imported goods in this state.

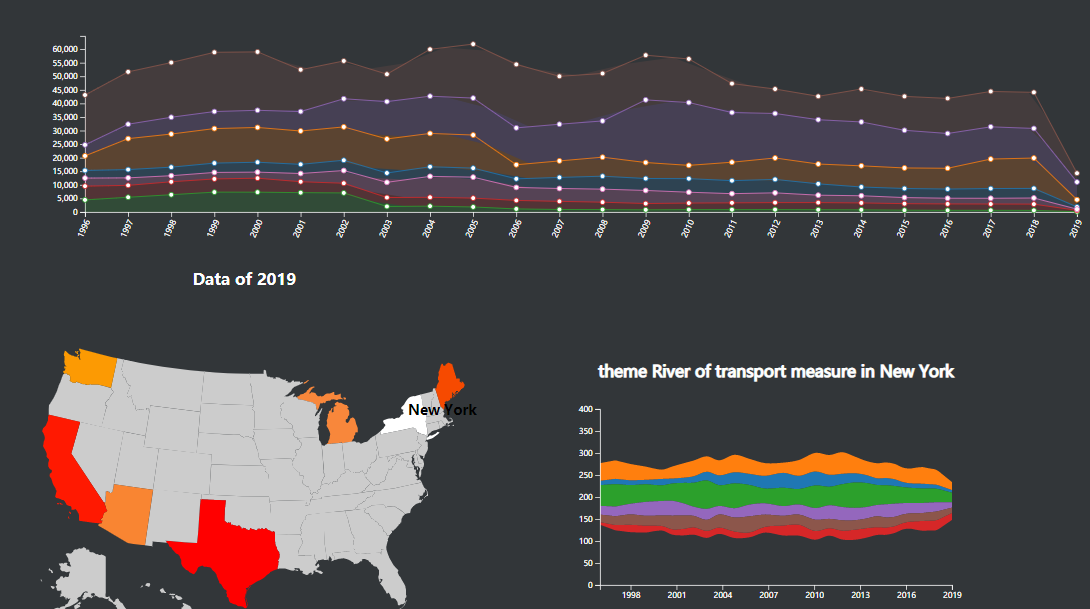


Figure 4: Combined Network View of Stacked View, Map View and Measure Statistics View

The three views of Stacked View, Map View and Measure Statistics View are reactive. Map view in the combination view will change according to the selected year, while measure statistics view will show the transportation measures composition of corresponding state according to the selected state in Map View. For example, 2019 is selected in the Stacked View, so the Map View shows the value of imported goods in each state in 2019. New York is selected in Map View, so Measure Statistics View will display the composition of New York's import measures.

# Evaluation

Firstly, we determined our declaration -D3.js and JavaScript to make our project runs in browser. In this way, the geospatial can be showed in a clear way and make it as an interactive system. Our system can present comprehensive information to users and allow users to choose information and display it. Besides them, we verify the expressiveness and efficiency of our system. We completed this system in a complete and comprehensive way(using D3.js). In addition, we use reasonable comments and appropriate hues to improve the accessibility of the entire project.

Through our ImportNetwork which shows the import information, we summarized our findings below.

First of all, in the overview (Fig. 1), we notice that we found that New York、Texas and California imported goods from 1996 to 2019 always ranked in the top 5 in all cities. Although the import volume of some states changed significantly, such as Washington, but their proportion in the total import volume is still relatively small.

Besides them, different cities have different forms of entry measure. For example, Michigan 's mode of inbound transportation is mainly bus and chunk. Except for this city, most of states consists of passengers.

During browsing, we further find that most ports only correspond to one city, but a city can be accessed by many ports.

# Conclusion

In this paper, we propose ImportNetwork, an interactive and intuitive visual system to help users understand the overall import information. We summarize the tasks and principles of project design, and design 4 views to comprehensively display import information from different aspects and latitudes. The case study exemplifies the effectiveness, expressiveness, accessibility and the usefulness of our project design.

ImportNetwork still has room to improve. A primal improvement needed is to display import information for most states in U.S. It can be found in our project that most of the data in the data set are coastal cities, and the data in inland areas are missing, which has affected the completeness of the visualization of the entire project to some extent. Thus, we should find a more complete dataset for our project. In addition to mining the inherent relationships of the data, we can spread our concerns to other aspects. For example, local economic level, import tax, etc. But this requires more professional knowledge and analysis, which may take a long time to learn.

References

1. Henning Tarp Jensen,Marcus R. Keogh-Brown,Bhavani Shankar,Wichai Aekplakorn,Sanjay Basu,Soledad Cuevas,Alan D. Dangour,Shabbir H. Gheewala,Rosemary Green,Edward Joy,Nipa Rojroongwasinkul,Nalitra Thaiprasert,Richard D. Smith. International trade, dietary change, and cardiovascular disease health outcomes: Import tariff reform using an integrated macroeconomic, environmental and health modelling framework for Thailand[J]. SSM - Population Health,2019,9.
2. Maria Bas,Vanessa Strauss-Kahn. Input-trade liberalization, export prices and quality upgrading[J]. Journal of International Economics,2015,95(2).
3. Akhil. Border Crossing Entry Data, 2019.

https://www.kaggle.com/akhilv11/border-crossing-entry-data

1. Imports of Certain Worsted Wool Fabric: Implementation of Tariff Rate Quota Established Under Title V of the Trade and Development Act of 2000: Removal of Regulations[J]. The Federal Register / FIND,2019,84(136).