

Chinese Students Studying in U.S.

Group Members

Qian Yu | Xuancheng Fan

INFO 247: Final Project Report

7 May 2018

Project Goals

While the phenomenon of increasing numbers of Chinese students studying in the U.S. has captured great attention of people in both China and U.S., there is few comprehensive illustration of it with diverse data and rich information, according to the existing news coverage or academic reports.

The goal of this project is to display three meaningful aspects of how Chinese students study in the U.S. by adopting a set of information and data visualization techniques:

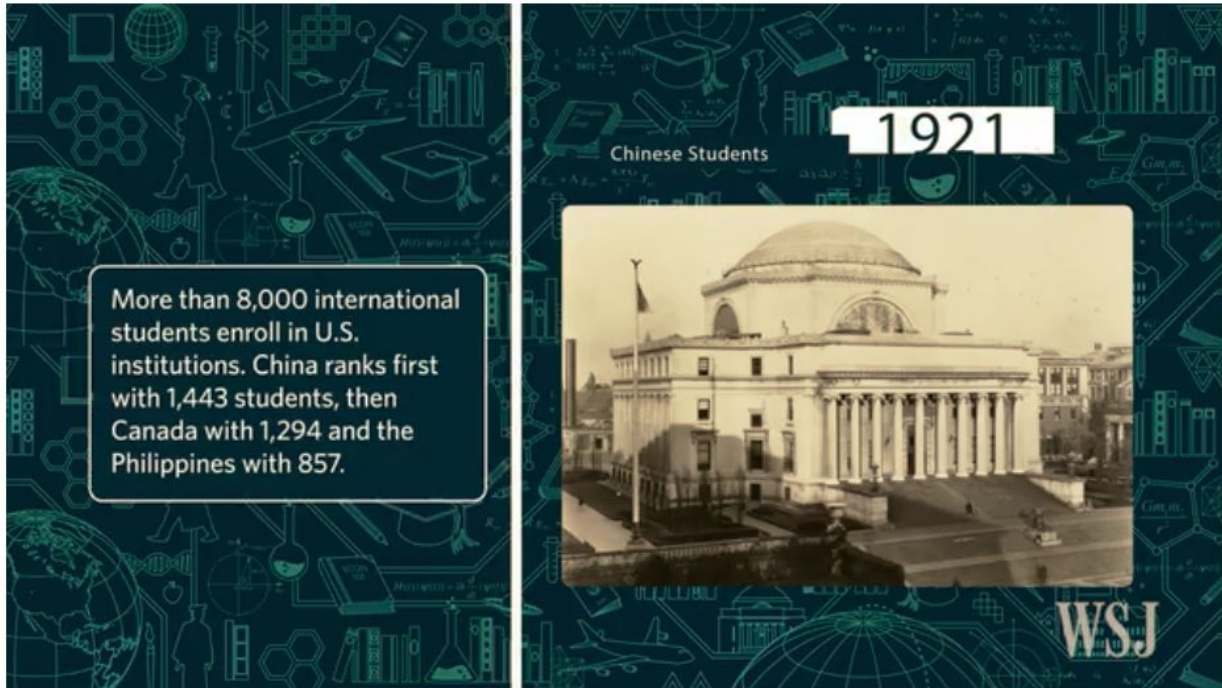
- Chinese students studying in the U.S. from a historically temporal perspective
- Changes of Chinese students studying in the U.S. based on social and economic events and development in both Chinese and American societies.
- Characteristics of Chinese students in the U.S. including study fields, education levels, OPT approval, origin cities, and their distribution in the U.S..

Discussion of Related Work

Data visualization project from the WSJ

The holistic storyline of the project and the historical timeline section are inspired by the data visualization project - *How International Students Are Changing U.S. Colleges* (the URL can be found in Appendix) from The Wall Street Journal. Its animated video displaying the history of international students in America (as shown in

Figure 1) provides us with reference in designing the historical timeline of Chinese students studying in U.S. in this project.



A look at the history of international students in America.

Figure 1: The screenshot of history of international students in America from WSJ

Infographic from Open Doors 2017

The infographic on displaying U.S. destinations for international students (as shown in *Figure 2)* provides this project with reference in designing the “Top 5 States” in the studying path section.

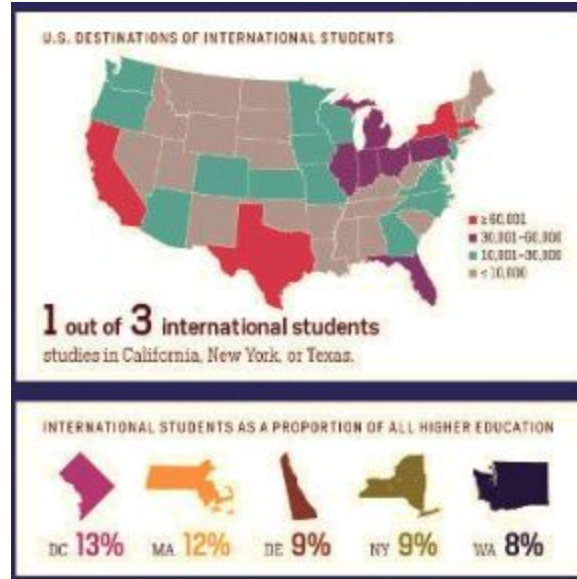


Figure 2: The screenshot of infographic from Open Doors 2017

Sketching & planning storyboards

Following “before you think about style, you must think about style” highlighted by Cairo Alberto (2013) in the *Functional Art: An introduction to information graphics and visualization*, a preliminary sketch (as shown in Figure 3) displaying the structure of this project is done with Adobe XD. This sketch is very useful for us to keep a clear idea of the information and data visualization focus in mind during the whole project design process. At the meantime, it effectively acts as a guide for us to gather information and data to accomplish the goal of the project. Also, this sketch makes us think about the visual elements such as typefaces, color palettes and the layout.

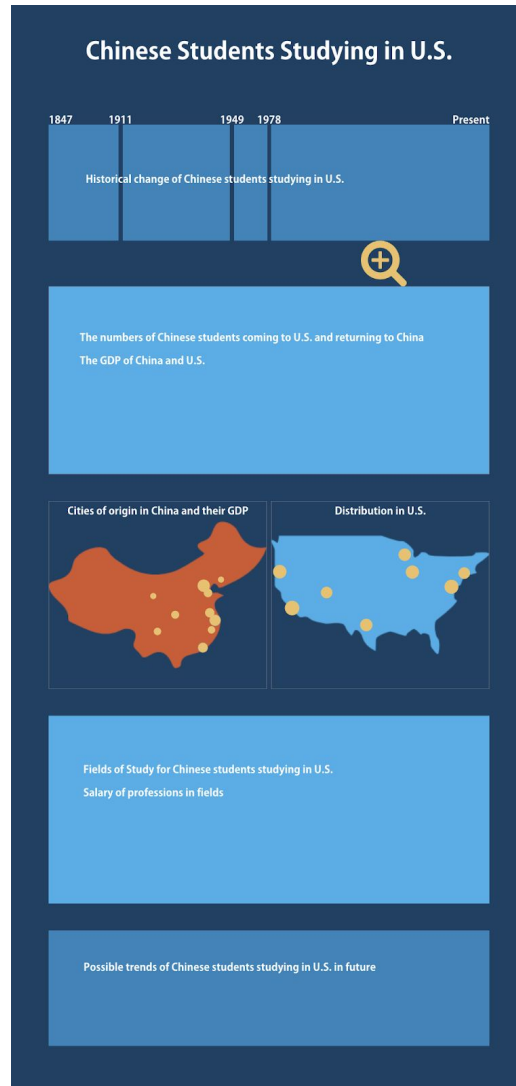


Figure 3: The preliminary sketch

D3.js for the flying map

To visualize the top 5 areas in U.S. where the students from different cities in China study, we decided to make the animation to present the moving from one China city to one U.S. area. Thus, we referred to three projects to solve the issues we met during development:

- Referred to the project, interactive map with d3.js (as shown in *Figure 4*), we learned how to gather the geo data and visualize the world map as we needed.



Figure 4: Example of world map

- Referred to the project, flight animation with d3.js (as shown in *Figure 5*), we find a way to implement the flight animation between two points on the map.

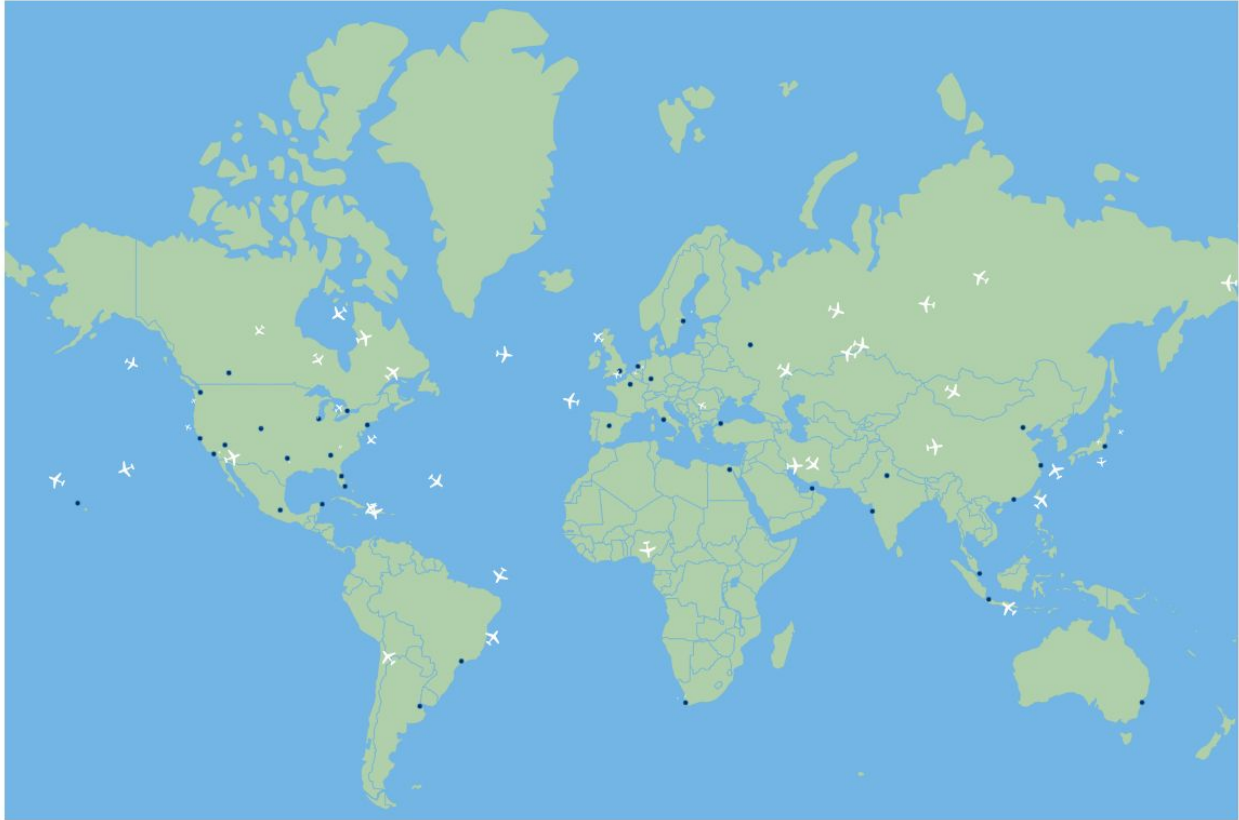


Figure 5: Flight animation

- Referred to projection transitions v4 (as shown in *Figure 6*), we compare the results of different projection style and choose the most suitable one.

Projection Transitions v4

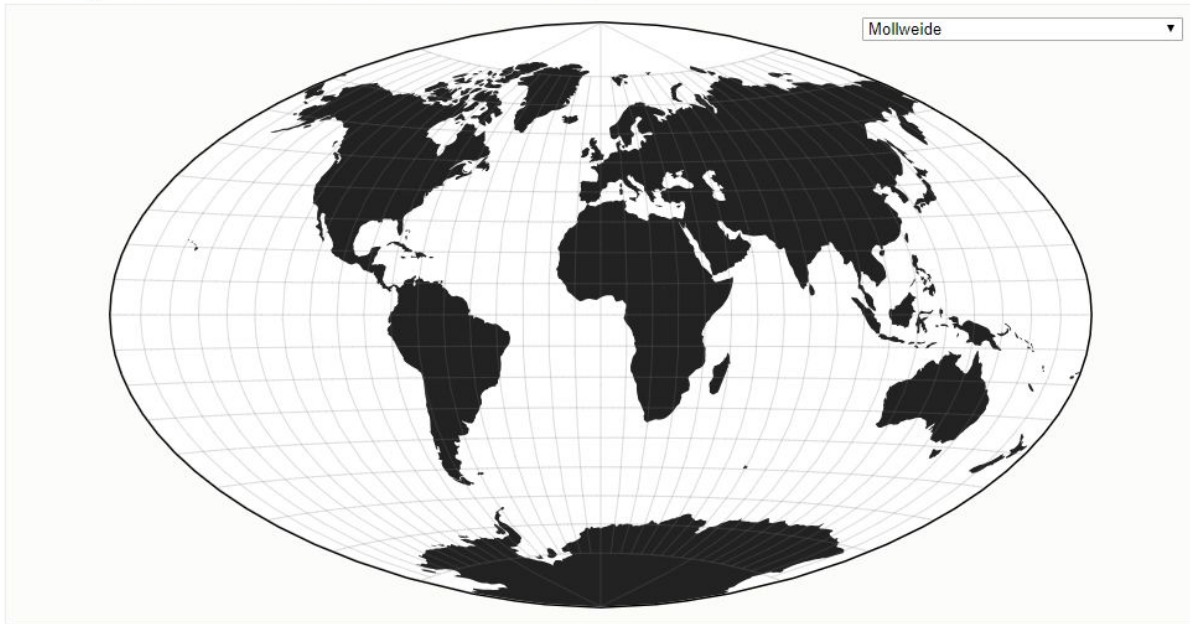


Figure 6: The result of Mollweidd projection

Articles support data visualizations choice and strategy

According to our design, we need to integrate the change of the total number of student, the number of students in the graduate and undergraduate education level, and the GDP of U.S. and China into one graph. To make two types of data more distinct and prioritize the trend of the number of student, we decided to use a line graph for the trend of the number of student and two area graphs for the GDP. Referred to Area Charts (reference URI can be found in Appendix), the purpose of our GDP data matches the ideal use-cases for area charts as below:

1. You need a simple comparison between the trend/proportion of each category.

2. You need to compare a small number of categories.
3. You need a simple comparison between trends and not exact values.

Description of Visualization

- Landing page with an overview of storyline

A landing page (as shown in *Figure 7*) with five buttons accordingly linked to sections is created to allow the audience to have a quick overview of the topic and the storyline. Meanwhile, to highlight the topic, a quote about the Chinese students in U.S. from the *Open Doors 2017* is placed on the landing page.



Figure 7: Landing page

- Infographic

The first section of the project is to display the historical timeline of how Chinese students studying in U.S. by the use of infographic (as shown in *Figure 8*). The infographic narrative consists of four phases (i.e., *the Late Qing Period, the Republican*

Period, Early Days of P.R.C, and The Post 1978 Period) which are marked by critical events, such as westernization movement, P.R.C was founded, and the reform and opening up during the history.

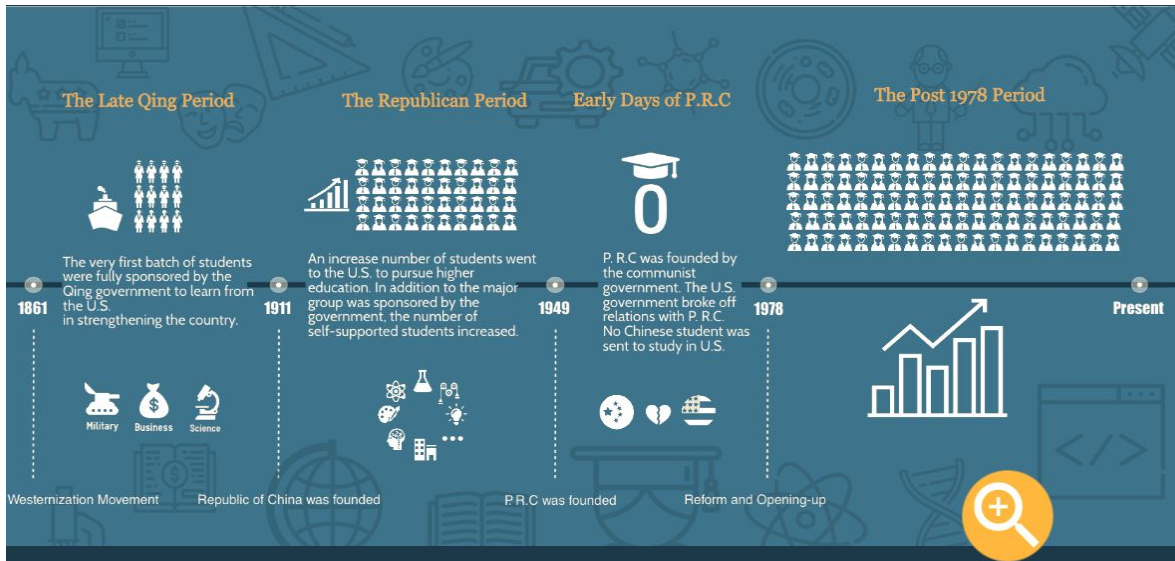


Figure 8: Infographic

Isotype and flat icons (as shown in Figure 9) are used to visualize the elements in the infographic narrative.

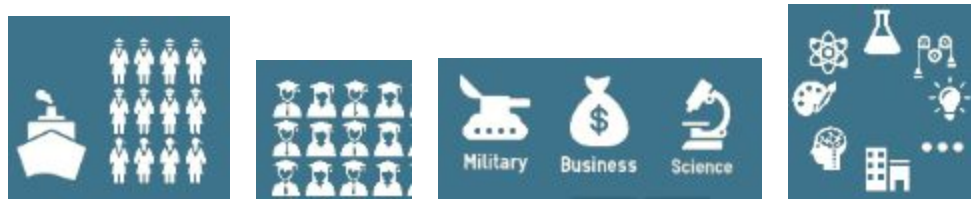


Figure 9: Isotype and icons

- Changes

The second section of the project is to display the changes of Chinese students studying in U.S. from 2000 to 2016. As this time period belongs to the *Post 1978* period of the historical timeline in the infographic section, a zoom-in button styled with

CSS plus effect animation (as shown in *Figure 10*) is created to provide the audience with an interactive navigation to the second section - *Changes in Modern Time*.



Figure 10: Zoom-in button

With regard to the changes in modern time, both line and area graphs in Highcharts (as shown in *Figure 11*) are used to display the changes in the total number of student, the number of students in the graduate and undergraduate education level, and the GDP between U.S. and China. At the same time, four critical social and economic events are marked with the dash-dot line in the graph to provide the audience with contextual information in interpreting the data.

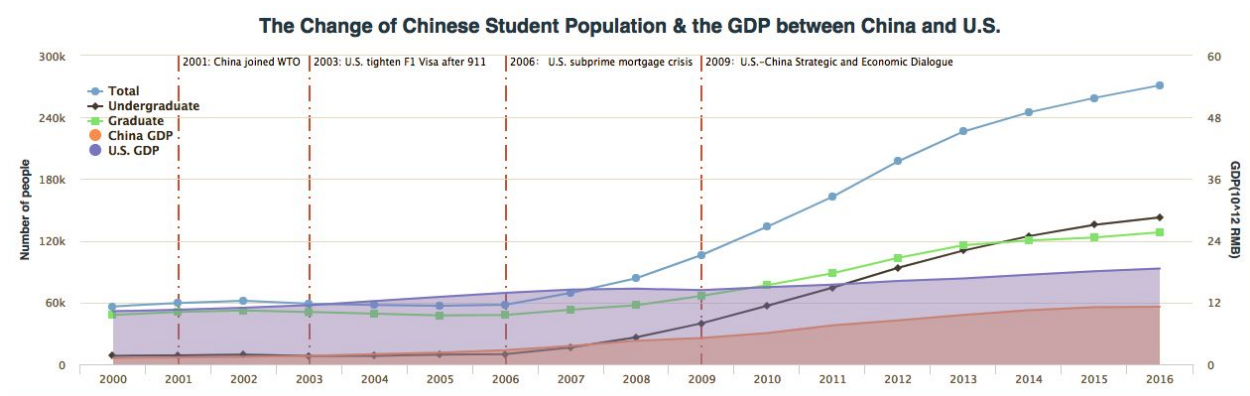


Figure 11: The change of Chinese student population & the GDP between China and U.S.

Also, another line graph in Highcharts (as shown in *Figure 12*) is adopted to display the changes of optional practical training (OPT) got by the Chinese students and its percentage of the total number. Meanwhile, as the OPT approval is a meaningful part of the change of Chinese students population, *Figure 12* is vertically aligned with *Figure 11* by the year label of the x-axis.

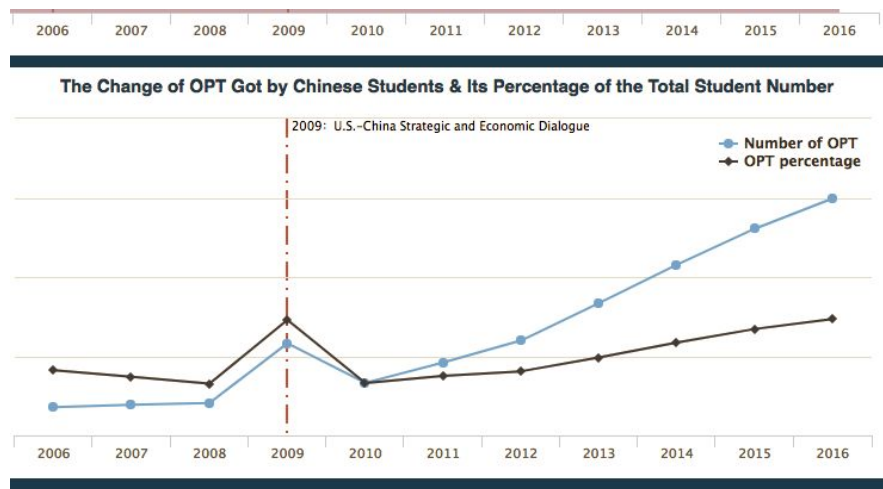


Figure 12: The change of OPT got by Chinese students & its percentage of the total student number

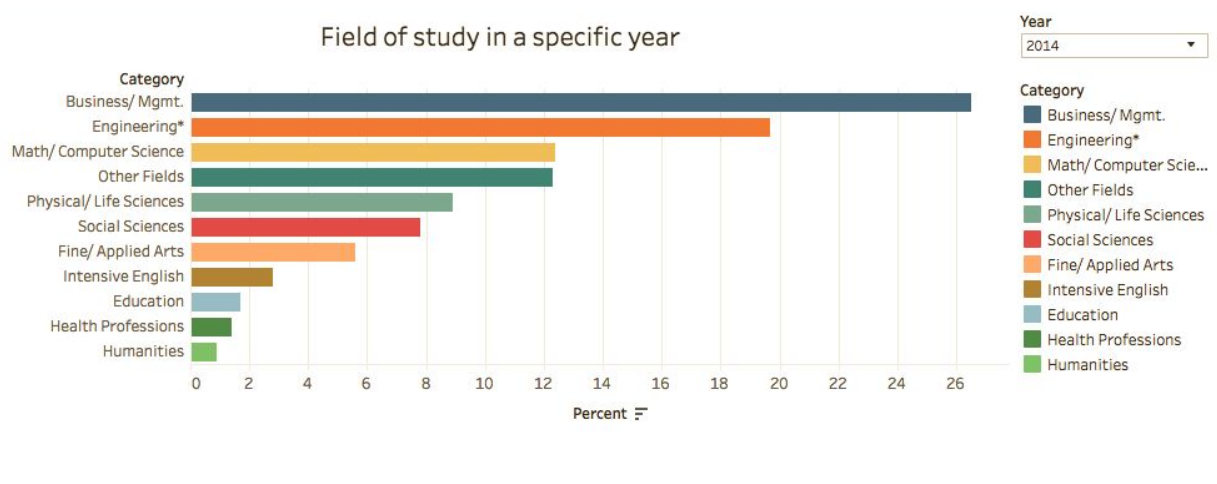
- Field

As the third section of the project is to visualize the study fields of the Chinese students studying in U.S., a study fields wheel button styled with CSS rotate effect animation (as shown in *Figure 13*) is placed at the end of the second section, providing the audience with an interactive navigation to the following section.



Figure 13: Study fields wheel button

With regard to the study fields, a horizontal bar graph and a line graph with filters in Tableau (as shown in Figure 14) are adopted to display what majors/disciplines that Chinese students choose to pursue in U.S. from 2009 to 2014.



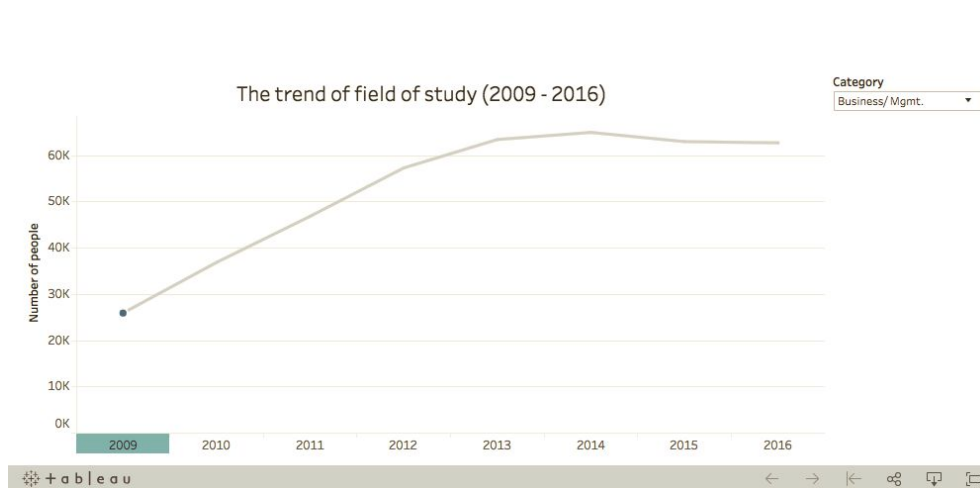


Figure 14: Study fields and the trend from 2009 to 2016

- Distribution

The fourth section of the project is to visualize the geographical path that students coming from different cities in China choose when they pursue education in U.S.. To provide the audience with an interactive navigation from the third section to the fourth one, an Earth button styled with CSS spinning effect animation (as shown in *Figure 15*) is added in the third section.



Figure 15: Earth button

To effectively visualize the studying path, a geo-map graph based on the D3.js (as shown in *Figure 16*) is adopted. The audience can view the top five areas of students

coming from different cities in China choose in U.S. by playing with the drop-down menu on the top of the graph. Meanwhile, an infographic graph displaying the top five states where Chinese students choose in U.S. is horizontally placed next to the geo-map graph to provide comprehensive information of the studying path.

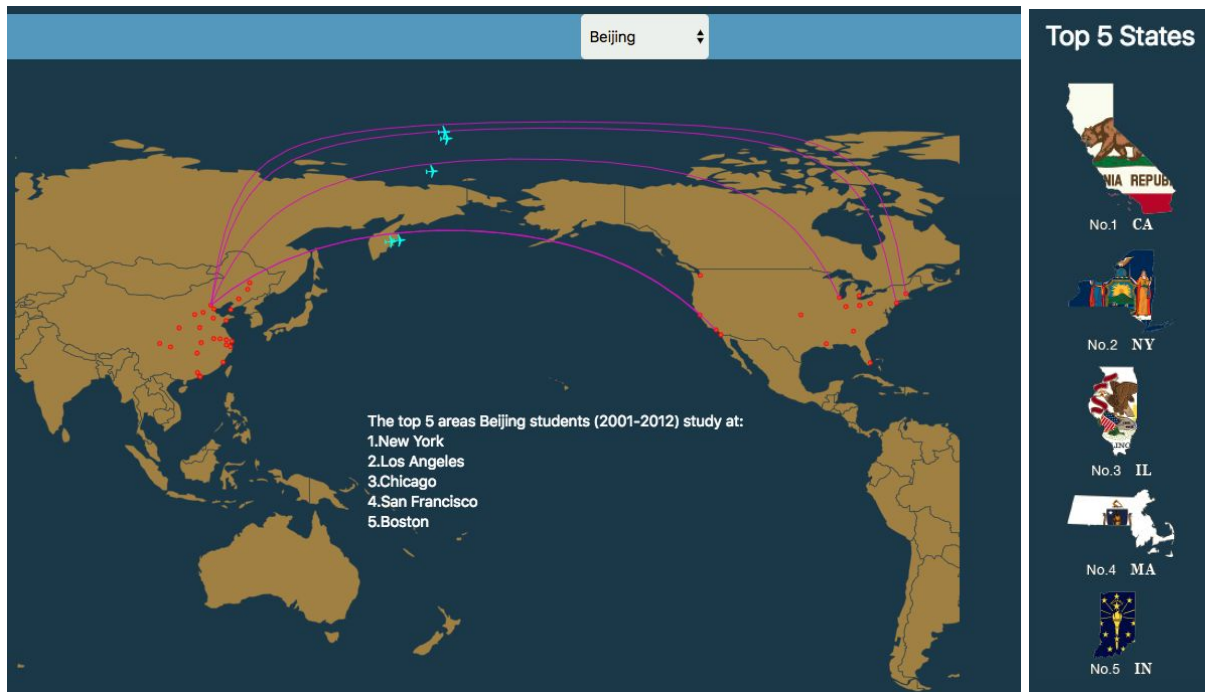


Figure 16: Studying path

In addition to the geo-map graph, a stacked bar graph (as shown in *Figure 17*) displaying the average GDP of different cities in China and their numbers of students studying in U.S. from 2001 to 2012 is added to assist audience in interpreting the data. Generally, the city GDP doesn't directly associate with the number of students studying in U.S.. A potential reason might be that Chinese students nowadays choose other countries such as the UK, Australia and Germany to pursue higher education.

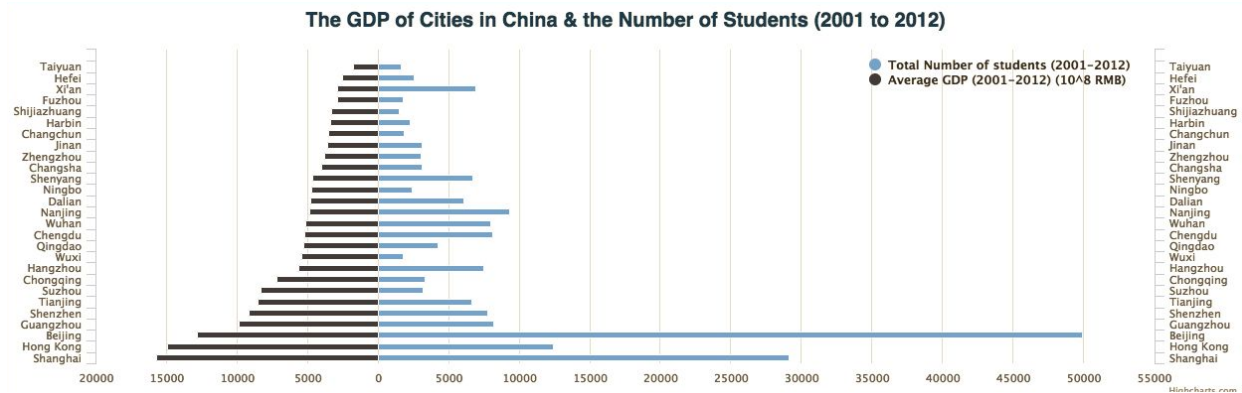


Figure 17: The GDP of cities in China & the number of students studying in U.S. (2001 to 2012)

- Future

The last section of the project is to illustrate the trends of Chinese students studying in U.S. based on the data and information of the above several sections. To keep a consistent experience of the interactive navigation, a rock button (as shown in Figure 18) styled with CSS bouncing effect animation is added to guide the audience to the future section.



Figure 18: Rocket button

The future section consists of three major trends of Chinese students studying in U.S. in the future. They are “student diversity”, “study field diversity” and “distribution diversity”. To provide the audience with an engaging experience in viewing this information, three interactive cards (as shown in *Figure 19*) styled with CSS flipping effect animation are created where the front-page of the card display the key word of the trend while its end-page provides more details about the keyword (as shown in *Figure 20*).

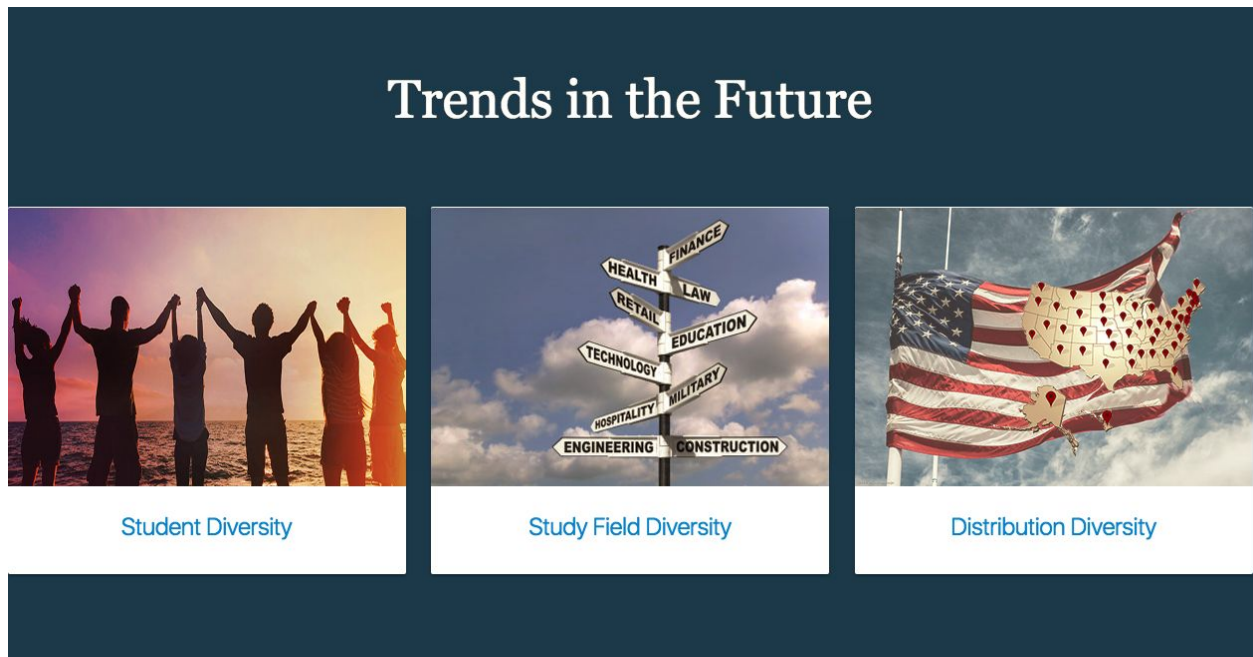


Figure 19: Trends in the future



Figure 20: The front-page and end-page of the card

Data Source

Data used for this project are from the following five source:

- Open Doors data from 2000 to 2016

Open Doors of the Institute of International Education (IIE) is a comprehensive resource on international students studying at higher education institutions in the United States. Data on the total number of Chinese students and the numbers in undergraduate and graduate level from 2000 to 2016 can be accessed from the Open Doors website. At the meantime, the number of Chinese students' Optional Practical Training (OPT) approval from 2006 to 2016 can be accessed. These data are used to display the second and third sections of this project.

(data source URL can be found in Appendix)

- World Bank Open Data

The GDP of China and U.S. from 2000 to 2016 is collected from the World Bank Open Data to display the changes in the gap between these two countries. The display of GDP gap between China and U.S. can provide the changes in the number of Chinese students studying in U.S. with more contextual information.

(data source URL can be found in Appendix)

- Data from U.S. Immigration and Customs Enforcement obtained through a Freedom of Information Act (FOIA) Request

The Brookings institution released a report on the geography of foreign students in U.S. higher education: origins and destinations in 2014. Its data source has cities in China and the numbers of students studying in U.S. from 2018 to 2012. These data are used to visualize the fourth section of this project.

(data source URL can be found in Appendix)

- National Bureau of Statistics of China Database

The GDP of cities in China from 2018 to 2012 can be accessed from the National Bureau of Statistics of China database. These data are used to visualize the fourth section of this project.

(data source URL can be found in Appendix)

- Natural Earth Data

The world map data is collected from Natural Earth. Natural Earth is a public domain map dataset available at 1:10m, 1:50m, and 1:110 million scales. Featuring tightly integrated vector and raster data, with Natural Earth you can make a variety of visually pleasing, well-crafted maps with cartography or GIS software.

(data source URL can be found in Appendix)

- Articles about the history of Chinese students in U.S.

News articles and academic thesis about the history of Chinese students in U.S. are used to get information and data to do the infographic narrative of this project. An article from the ChinaDaily USA and the thesis: the historical evolution of Chinese students' overseas education in the United States since modern China are used for the infographic section of this project.

(references can be found in Appendix)

Visualization Tools & Techniques

The tools and some specific techniques we used are:

- Infographic with Adobe Illustrator
- Highcharts - Two Y axes, Bar with negative stack
- Tableau
- D3.js: gdal, topojson, projection method

Here are the reasons why we choose this tools and the techniques:

- Infographic

The Adobe Illustrator was adopted to create the historical timeline in a vector image which can be displayed in a pixel perfect level. Meanwhile, the layers mechanism of Illustrator allows graphics to be added as a background for the infographic.

- Highcharts

We used highcharts to build three figures: 1) the trend of the number of students with the change of GDP, 2) the trend and the percentage of the OPT, and 3) the GDP of China cities compared with the number of students of this city. We used two axes for figure 1) and 2) because there are two categories of data in one figure and we would like to show the trend within one category and the comparison between two categories while both of them are changing with time. Within two Y axes, we expect audience to

have a basic understanding the relationship between the the number of student and the GDP.

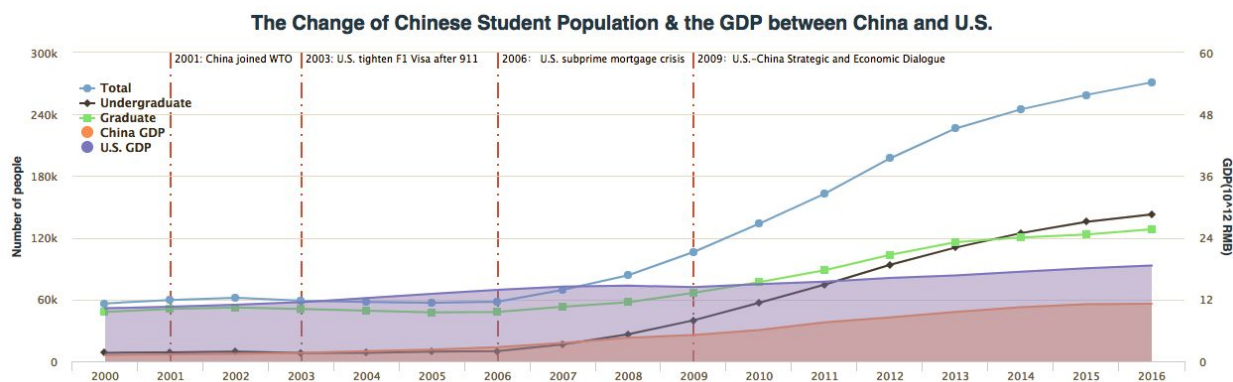


Figure 11: The change of Chinese student population & the GDP between China and U.S.

We used bar with negative stack for figure 3) because we need to compare two categories of data (GDP of cities and number of students) but each category has about 20 same sub-categories (cities). Meanwhile, we just need to compare the data of one specific time and there is no change of time. Bar with negative stack could make the comparisons more legible and easy to understand.

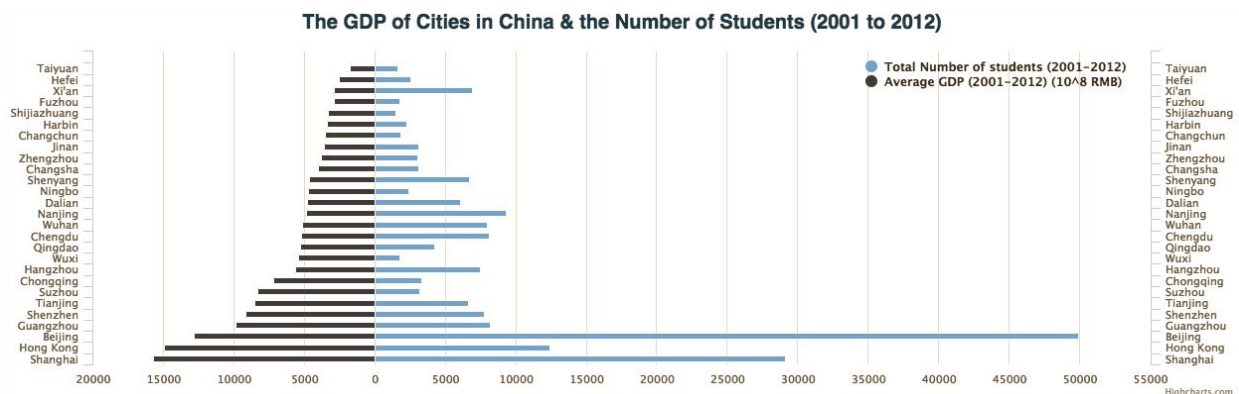


Figure 17: The GDP of cities in China & the number of students studying in U.S. (2001 to 2012)

- Tableau

We used Tableau to present the comparison of the field of study of one specific year and the trend of a specific field of study (as shown in *Figure 11*). Since we have about 10 fields of study and more than 10 years to compare, it would be confusing to have all data in one figure. Instead, using Tableau enables us to let the audience choose which specific year or field of study to explore and make the visualization more concise.

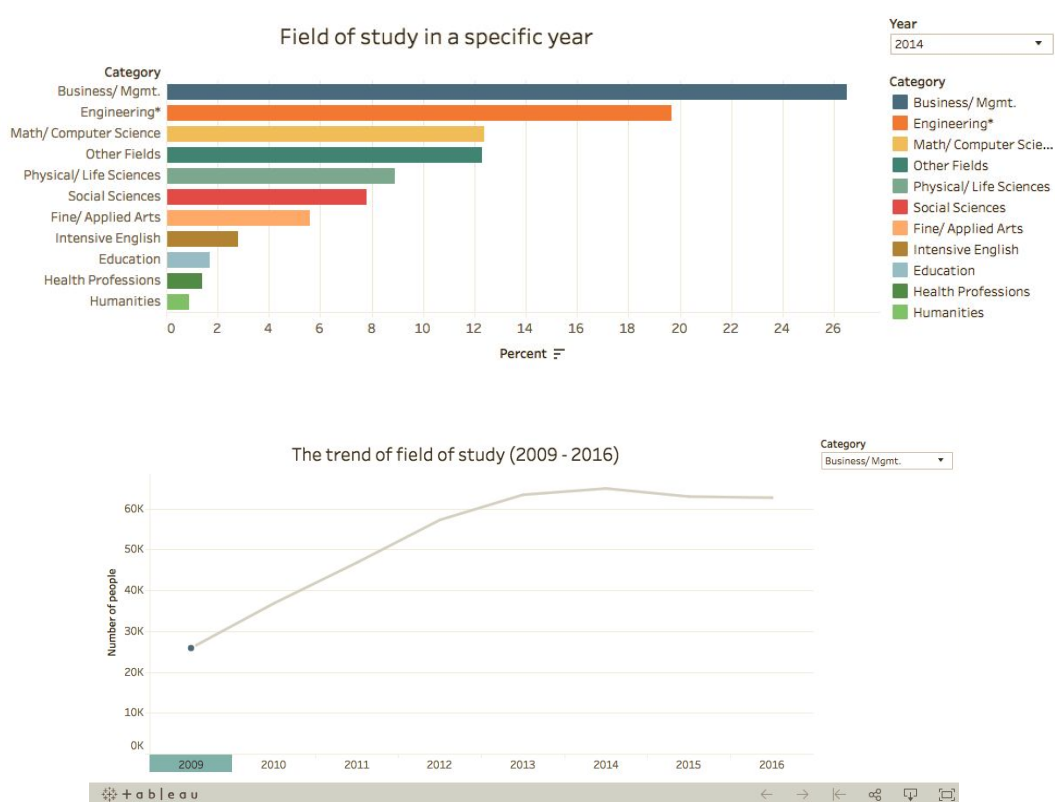


Figure 14: Study fields and the trend from 2009 to 2016

- D3.js

We used D3 to visualize the trip from one China city to the U.S. area that students from this city prefer to study at. With the animation of D3, we would like to make a

deeper impression of the long journey and, furthermore, show the preference of students according to the distribution and its potential relationship between the origin of students and the destination, for example, the students from the east part of China may be less likely to go to the central part of U.S.

To implement the animation, we first download the geo data from Natural Earth and we used below techniques or tools to advance:

- gdal

It will create GeoJson file from the raw files we download from Natural Earth

- topojson

The GeoJson is very big since it contains a lot of attributes we don't need. So we used topojson to remove unnecessary attributes. As a result, the size of the final json files are less than 10% of GeoJSON file.

- Projection method

To make the flight animation more vivid, we would like to make curve path for the plane, not the straight lines. Thus, we tried several projection methods but have some issues. For example, `geoMercator()` (as shown in *Figure 21*) would make it difficult to contain all the flight route in the figure and `geoAlbers()` (as shown in *Figure 22*) is not instinctive and hard to understand based on the result of user testing.

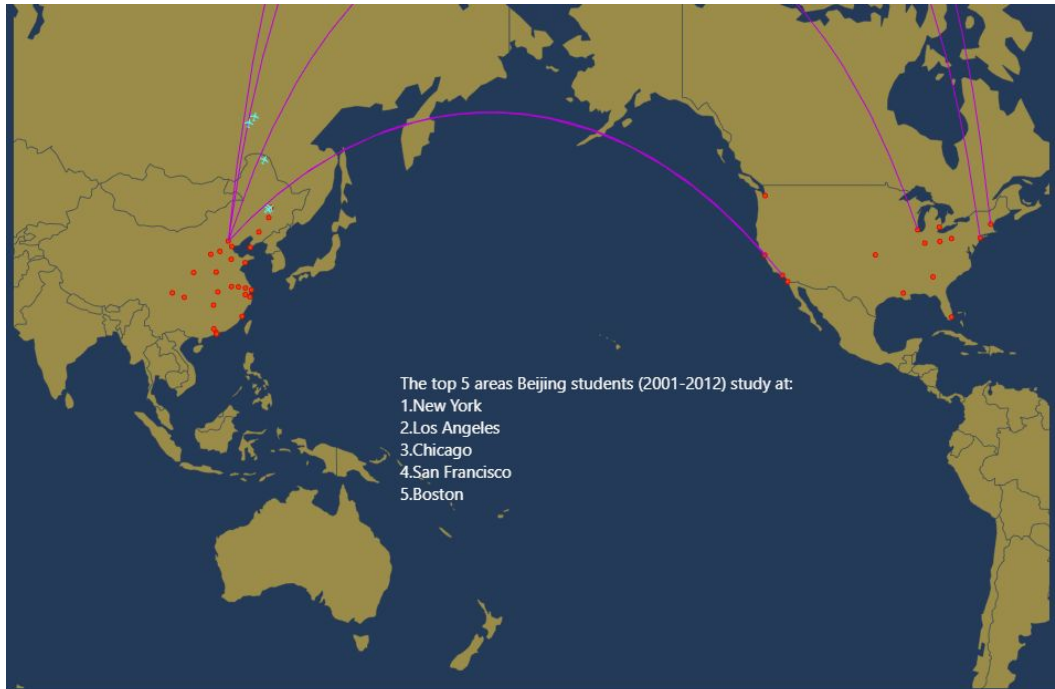


Figure 21: An example of geoMercator()

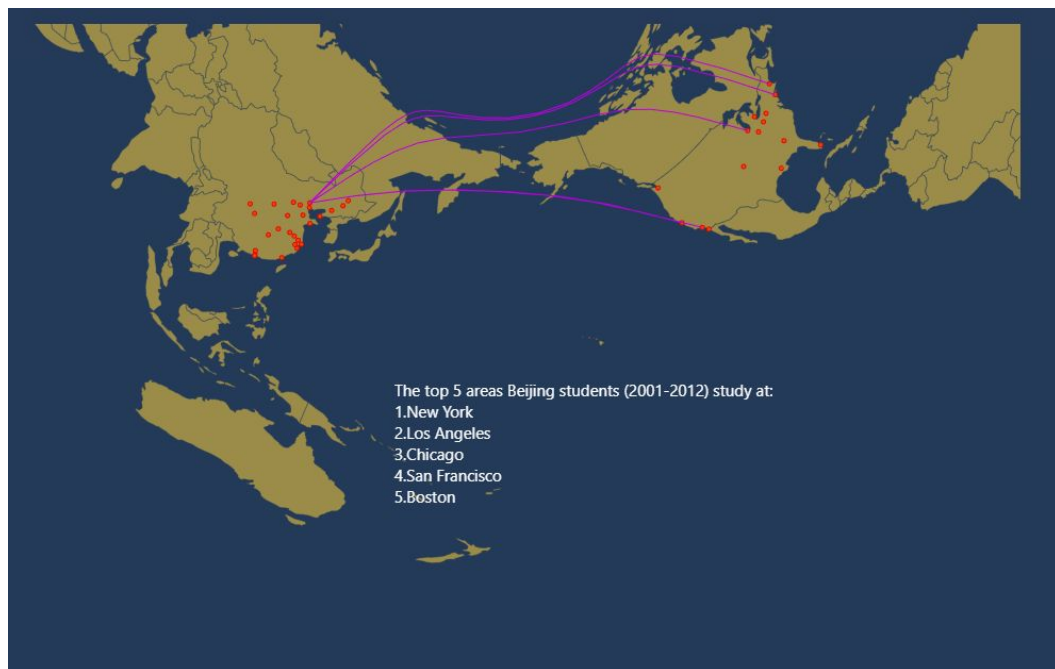


Figure 22: An example of geoAlbers()

After several attempts and comparison, we chose geoEquirectangular() (as shown in Figure 22) as the best one to fit our requirement.

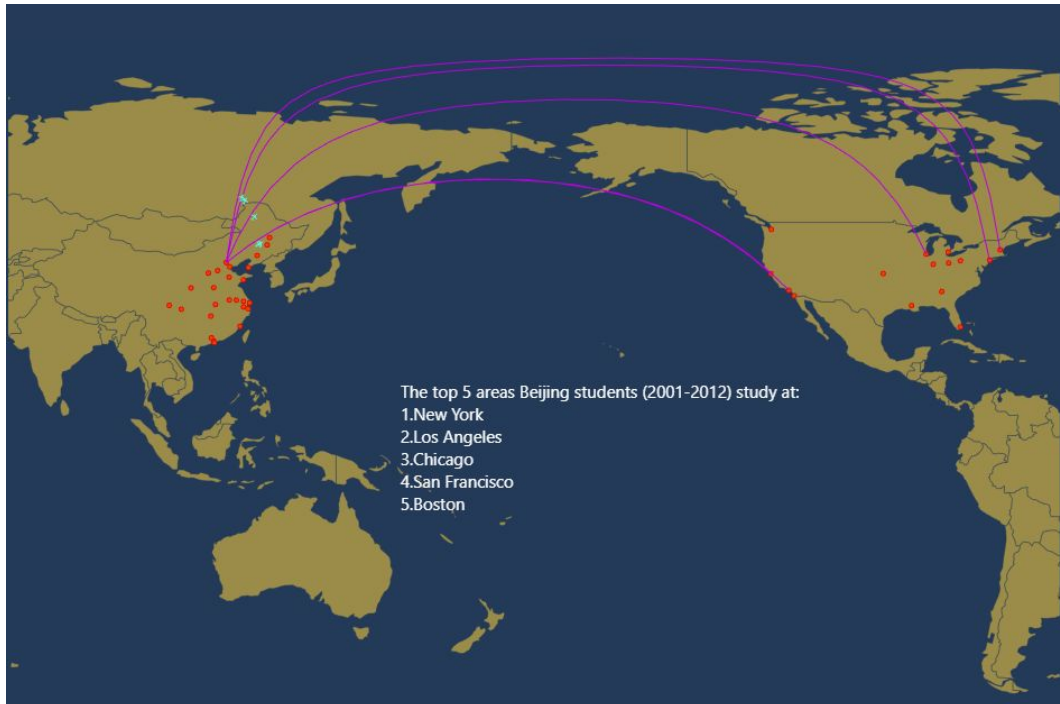


Figure 23: An example of geoEquirectangular()

Results

According to the usability done with six representative users including five are Chinese students pursuing education in American institutions and one has great interest in this topic, we find results as follows:

- The storyline design is effective

To all representative users in the usability test, the landing page with a highlighted quote and five buttons linked to the following sections, effectively provides them with an overview of the storyline of this project. Meanwhile, the alignment of these buttons can effectively guide the user to firstly interact with the *Historical Timeline* section of the project.

- The infographic is informative

Even though most Chinese students in the representative users have some knowledge about the historical background of Chinese students studying in U.S., the infographic of this project effectively delivers a comprehensive view of the historical timeline and critical events to them.

- The animated buttons can trigger interactions but distracted

The four animated buttons styled with CSS effects do trigger interactions from the representative users when they view the motions. However, some representative users said that these animated buttons caused distraction when they viewed the data and information visualization.

“I was firstly attracted by this animated yellow circle with a zoom-in icon when coming to the first section. This fantastic motion makes me really want to click it without finishing viewing the infographic.” — A user from the usability test

With regard to this result, the navigation design with animated buttons need to be improved in the further step.

- An effective use of different data visualization techniques

The combination of different graphs made by Highcharts and Tableau effectively display the data about Chinese students studying in U.S. to most representative users. Also, a year label has been added to the Tableau charts to deliver the information in a more straightforward way.

- Animation with D3.js makes the data visualization amazing

The animated planes flying out from different cities in China to different places in U.S. on the map is the part received most “WOW” from the representative users. All

users chose their hometown cities in China to take a look at the popular destinations in U.S.. Nevertheless, the interactive hint with the drop-down menu might be strengthened to provide the user with more clear guide of interaction.

- Animated flipping cards is appealing and engaging

Most representative users felt boring when they were exposed to the static cards displaying three major trends about the Chinese students studying in U.S. in future. According to feedback from the users, flipping animation has been added to these cards in which, the front-pages focus on the keywords of the trend while the end-pages provides more details about those keywords when the user hover mouse on the cards.

The Final Deliverable

A website has been hand-coded to display the information and data visualization. All

files can be forked from the following Github repository:

<https://github.com/hoperose/InfoViz-Final-Project/tree/master>

Responsibilities of Group Members

Sections in the project	Workload	
Ideation & brainstorming	Qian Yu (60%)	Fred Fan (40%)
Infographic design & creation	Qian Yu (100%)	

Data collection	Qian Yu (70%)	Fred Fan (30%)
Data visualization design & creation	Qian Yu (10%)	Fred Fan (90%)
Narrative	Qian Yu (90%)	Fred Fan (10%)
Web design and development	Qian Yu (100%)	
Heuristic evaluation	Qian Yu (50%)	Fred Fan (50%)
Report writing	Qian Yu (60%)	Fred Fan (40%)

Appendix

Data source URLs

- Open Doors data from 2000 to 2016: <https://www.iie.org/opendoors>
- World Bank Open Data: <https://data.worldbank.org/>
- Data from U.S. Immigration and Customs Enforcement obtained through a Freedom of Information Act (FOIA) Request:
<https://www.brookings.edu/interactives/the-geography-of-foreign-students-in-u-s-higher-education-origins-and-destinations/>
- National Bureau of Statistics of China Database:
<http://data.stats.gov.cn/easyquery.htm?cn=E0105>
- Nature Earth Data:
<http://www.natureearthdata.com/downloads/>

References

- Zhou, W. (2017). The history of Chinese studying abroad. *China Daily USA*. Retrieved from http://usa.chinadaily.com.cn/epaper/2017-04/21/content_29029959.htm
- The Wall Street Journal. (2015). How international students are changing U.S. colleges. Retrieved from <http://graphics.wsj.com/international-students/>
- Cairo, A. (2013). *The Functional Art: An introduction to information graphics and visualization*. Berkeley, CA: New Riders.
- Luo, D. (2013). *Seeking Modernity, Brain Gain, And Brain Drain: The historical evolution of Chinese students' overseas education in the United States since modern China*. Retrieved from https://ecommons.luc.edu/cgi/viewcontent.cgi?article=2853&context=luc_theses
- Noda, T. (2013). Interactive Map with d3.js. Retrieved from <http://www.tnoda.com/blog/2013-12-07>
- Noda, T. (2014). Flight Animation with d3.js. Retrieved from <http://www.tnoda.com/blog/2014-04-02>
- Macy, A. (2017). Projection Transitions v4. Retrieved from <https://bl.ocks.org/alexmac/6700d44240d2b6d3ec9767a5a5854e42>
- FusionCharts. Area Charts. Retrieved from <https://www.fusioncharts.com/chart-primers/area-chart/>

Screenshots of the project

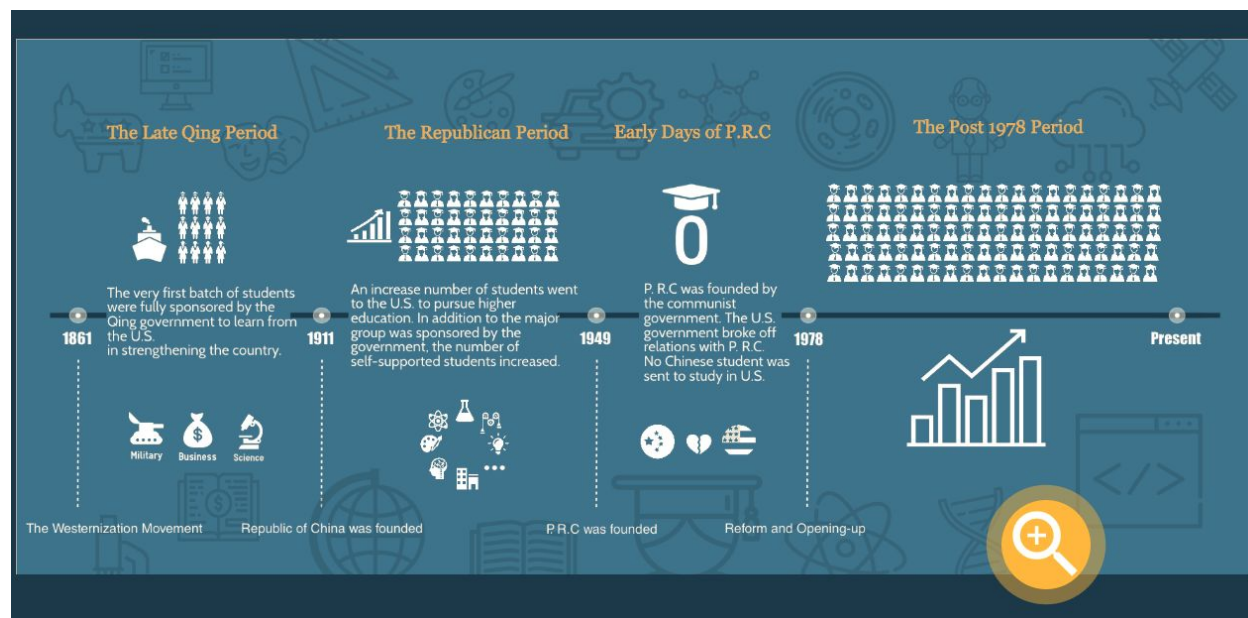
Final Project of InfoViz_247, UC Berkeley

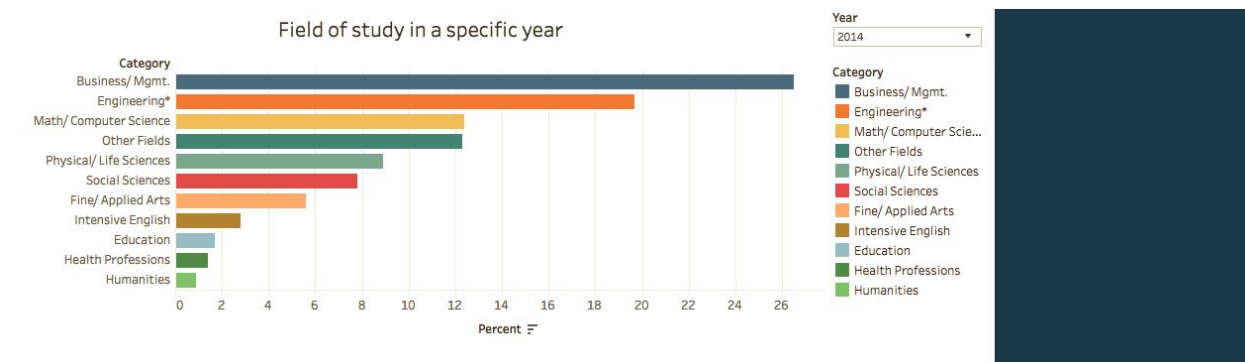
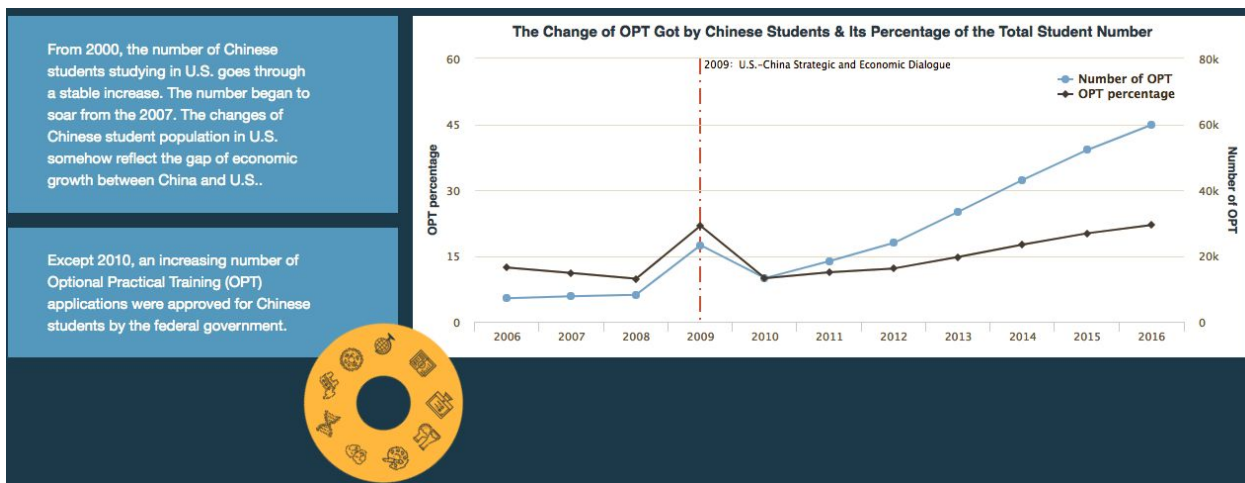
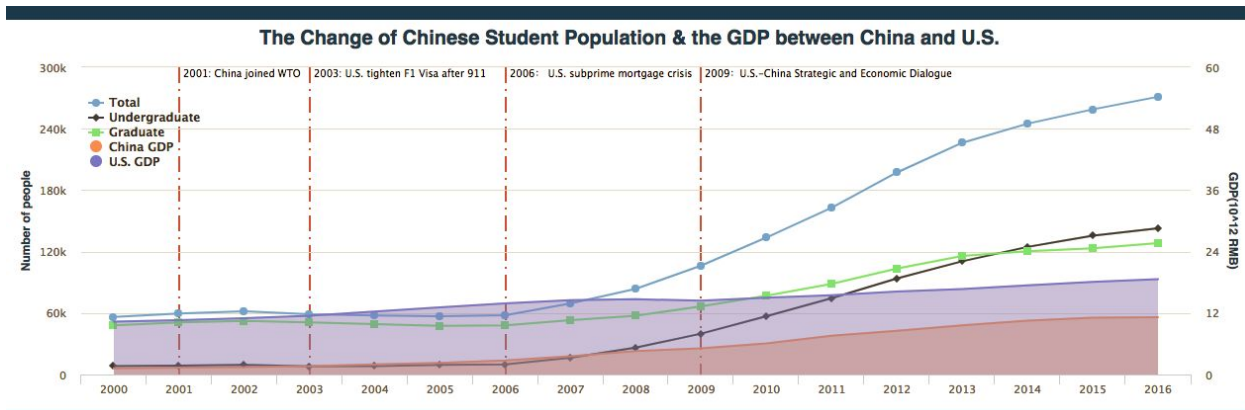
Chinese Students Studying in U.S.

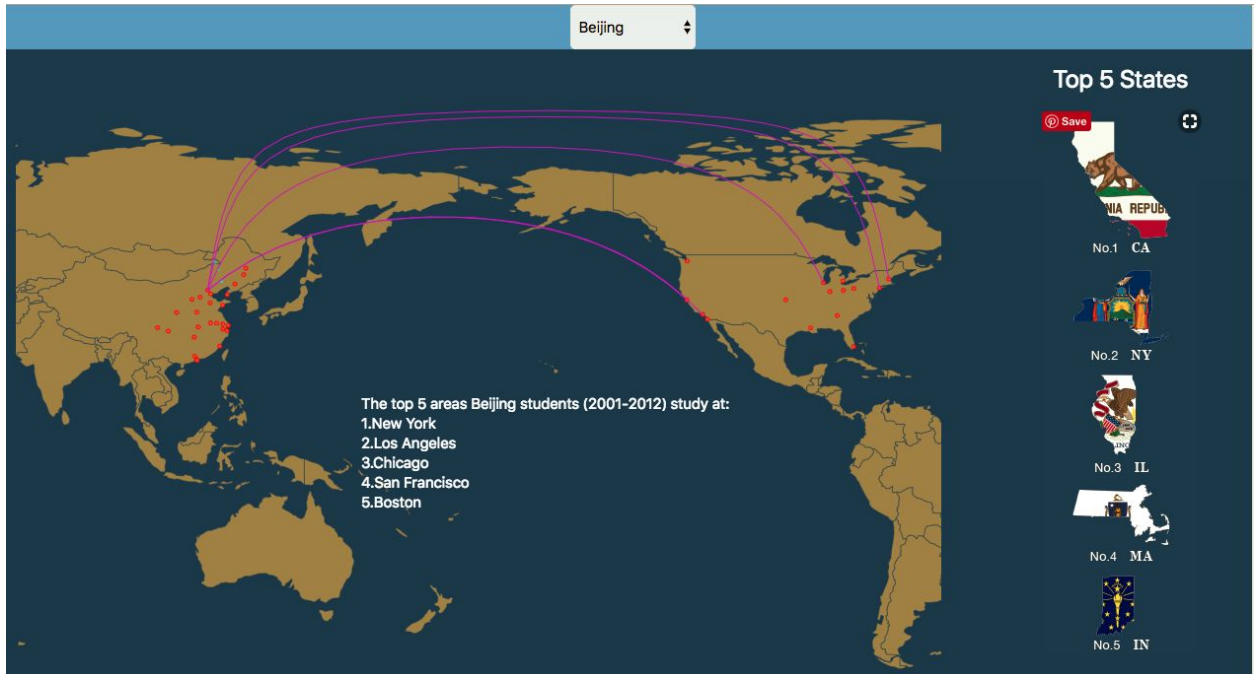
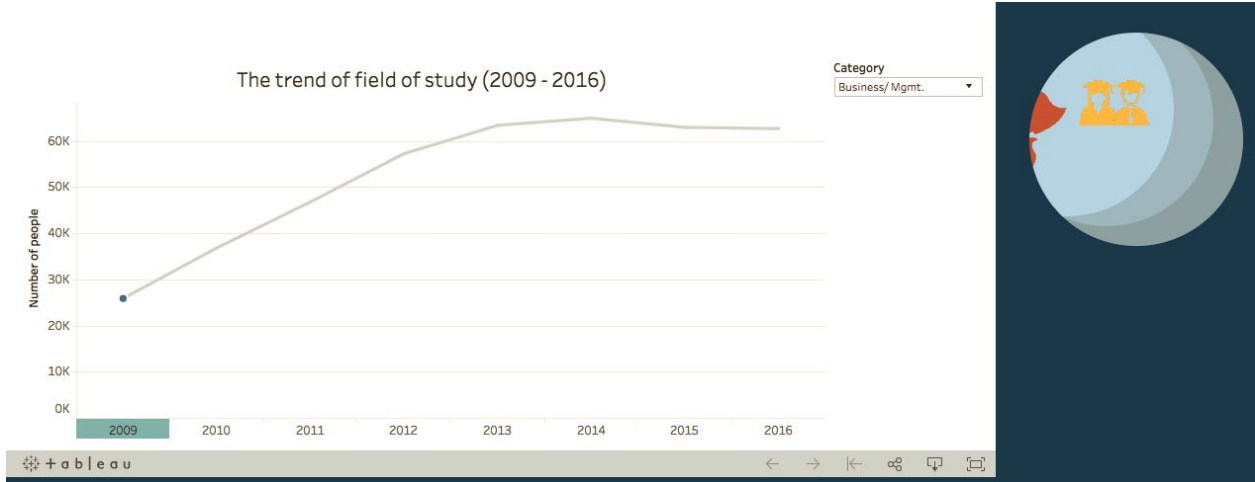
“China is the largest sender of international students to the U.S., with Chinese students accounting for 32.5% (350,755) of all international enrollments (1,078,822) in the U.S..”

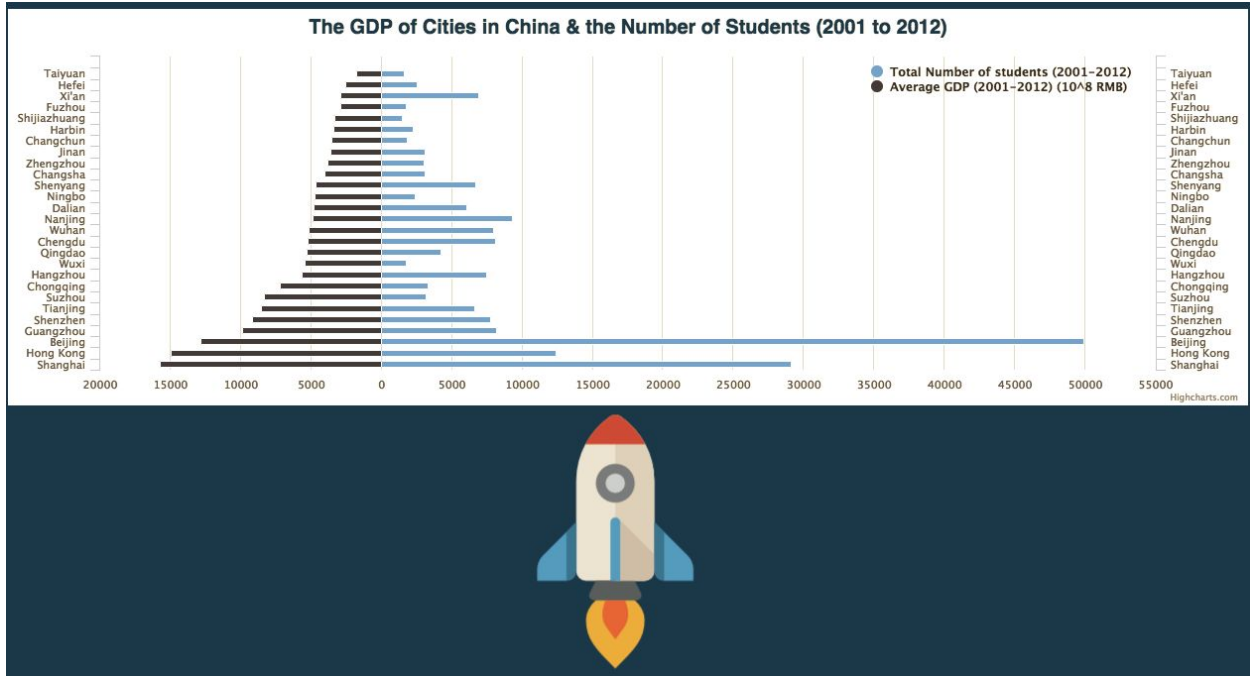
— *Open Doors 2017*, Institute of International Education (IIE)

HISTORICAL TIMELINE CHANGES IN MODERN TIME STUDY FIELD STUDYING PATH FUTURE









Trends in the Future



Student Diversity



Study Field Diversity



Distribution Diversity

Visualization tools

Highcharts

Tableau

D3.js

Adobe Creative Suite

Data source

Open Doors 2000-2017, *Institute of International Education (IIE)*

Data from U.S. Immigration and Customs Enforcement obtained through a Freedom of Information Act (FOIA) Request, *Brookings Institution*

The history of Chinese studying abroad, *China Daily USA*

GDP of cities in China 2001-2012, *National Bureau of Statistics of China*

Designers: *Fred Fan & Qian Yu*

© 2018 Copyright