Cartography M.Sc.

An Interactive Visual Interface for Geodata Query with Space-time Cube

Mengyao Gao

Outline

- Introduction and Motivation
- Related Work
- Methodology
- Prototype Implementation
- Evaluation
- Conclusion

Introduction and Motivation

Background:

- Data portals are developed to enable data available to public to generate useful information and deliver services (Davies, 2010).
- However, the use of these data are far reaching the expected level and visual interface has a great impact on the users' behaviours of searching and success rate of finding ideal datasets (Safarov et al., 2017)(Xiao et al., 2020).

Introduction and Motivation

Research gap:

- Most existing visualization platforms for data query only use simple symbols to visualize data content directly, which couldn't give an overview of the various kinds of data in the entire database and their important metadata.
- Important aggregated metadata for geodata query in those existing visualization platforms are not intuitively available to users.
- Those visualization platforms tend to visualize spatiotemporal datasets but only time slider in the 2D map has been used. It is not clear whether space time cube, a popular time visualization method recently, is feasible here or not.

Introduction and Motivation

Aim of the thesis:

- This research tries to use space time cube with 3D symbols to visualize time of the dataset, data quality, data size, and data category of all available datasets to check if they are feasible to assist users in performing usual tasks effectively.
- It also tries to visualize aggregated metadata which are important in data query to provide spatial overviews of all available datasets.

Introduction and Motivation

Research Objective:

The overall objective of this research is to develop an interactive visual interface to visualize aggregated metadata using space time cube to enable users to get good overviews and find needed datasets effectively.

- To understand the components of data quality.
- To gain an overview of related works in visualizing semantic information of spatiotemporal data using space time cube.
- To develop the conceptual design for visualization of metadata.
- To implement a prototype based on the conceptual design.
- To evaluate the effectiveness and satisfaction of this visual interface.

Introduction and Motivation

Research Questions:

- How to apply the space time cube visualization method to visualize metadata of spatiotemporal data to enable users to get overviews and find needed datasets effectively?
  a) Which visual variables should be chosen to visualize aggregated metadata and applied to 3D symbols in the whole space time cube?
  b) Which functionalities should be designed in this space time cube to enable users to get a good overview and find needed datasets effectively?
- How to evaluate the proposed visualization method?
  a) How is the effectiveness of the proposed visualization method using space time cube?
  b) How is the satisfaction of users towards the proposed visualization method using space time cube?

Introduction and Motivation

Innovations:

- This visual interface would firstly visualize the aggregated metadata of the datasets which could give users intuitively overviews of important information of all datasets.
- Especially, the visualization of data quality, one of selected metadata here, may give intuitively users data quality information, which may help improve efficiency and accuracy of the query process.
- Additionally, the application of space time cube with 3D symbols in visualization platforms of geodata query may provide a new feasible way to visualize spatiotemporal datasets besides time slider.

Related Work

There are mainly two branches of visualization platforms for geodata query:

- The first branch is developed to process interlinked datasets. It focuses more on improving the querying algorithm, providing the possibility for exploring the relations between interlinked datasets, and only using points, lines, and polygons to visualize the data content. Some typical platforms are Map4rdf (de León et al., 2012), Facete (Stadler et al., 2014), and Sextant platform (Nikolaou et al., 2015). Visualizing time using time slider has become one development trend of this branch, such as SexTant (Bereta et al., 2013).
- The second branch focuses on adapting different visualization methods to show good overview and help query. For example, BDTD platform uses choropleth map to visualize density of publications (Ramalho & Segundo, 2020).

Related Work

- Space time cube is a popular time visualization method recently. It uses height to represent time information. It has been applied to visualize three kinds of spatiotemporal information, namely spatiotemporal path, station, and spatiotemporal prism (Kveladze, 2015).
- The symbolization in 3D visualization is the variant of 2D symbols with more visual variables such as lighting, shadow and perspective. Point symbol in 3D visualization can be rendered as billboards, solid charts or custom symbols with different shapes (Haeberling et al., 2008).
- To assess data quality is a complicated task. It has different components in different data standards and researches. Among those components, completeness almost exists in all data standards and is independent of application. So, it is used in this research to represent data quality information.

Methodology

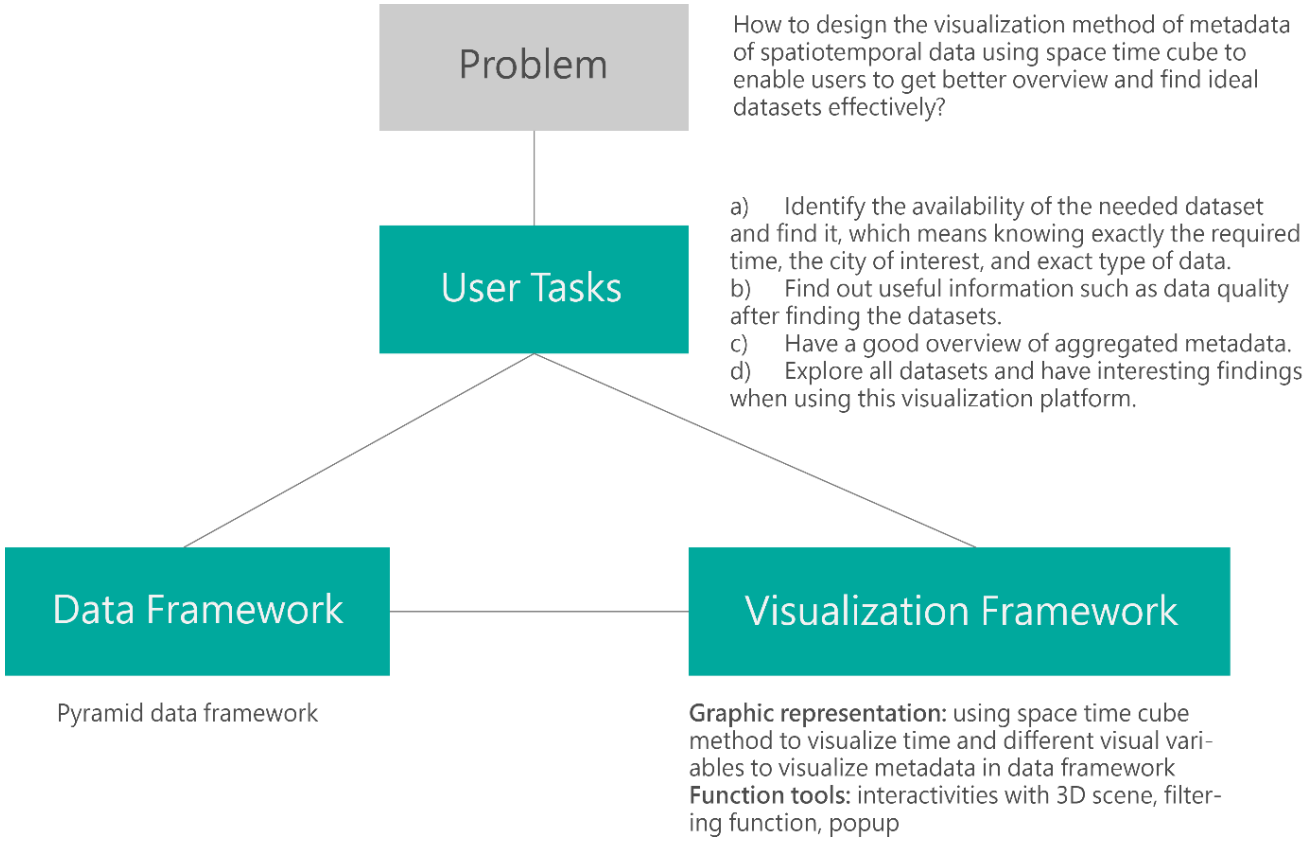

Fig. 1 A concept framework for visualization of multivariable spatiotemporal data adapted in this research. (Li & Kraak, 2010)

Methodology: Data framework

Table. 1 The pyramid data framework in the concept framework of this visualization

1st level	2nd level	3rd level	Length	Level of organization
Dataset	Space	Latitude	-	Quantitative
		Longitude	-	Quantitative
	Time	Day	Day	Quantitative
		Week	Week	Quantitative
		Month	Month	Quantitative
		Year	Year	Quantitative
	Attribute	Data Size	KB	Quantitative
		Data Category	Geographic data, social data, economic data, nighttime light	Qualitative
		Completeness	-	Quantitative

Methodology: Data framework

Four chosen types of data:

- Geographic data are commonly used data in spatial analysis and mapping.
- Economic data are quite important in studying the economic development of a region. They are also downloaded frequently from data portals.
- Social data, newly emerged type of data, are very useful in researches studying social behaviours, social events, and society. They are really popular in many data portals.
- Night-time light data belong to the category of secondary data calculated from remote sensing images. There are also various types of data got from remote sensing, such as the DEM, satellite cloud images and so on. They are really important in geography, hydrology, forestry, and many other related fields. In this research, night-time light data is chosen as the test dataset.

Methodology: Data framework

econ_id	postcode	factor	year	value
1	3201	A1	2010	7.48
54	3201	A2	2010	204183
119	3201	A3	2010	266.32
171	3201	B1	2010	647562
236	3201	B2	2010	931.3
301	3201	B3	2010	147.7
358	3201	C1	2010	206.2
423	3201	C2	2010	9
488	3201	C3	2010	286199
553	3201	D1	2010	65.2
670	3201	E10	2010	19577
709	3201	E1	2010	2211
774	3201	E2	2010	179
878	3201	E4	2010	31090
995	3201	E7	2010	60044
1060	3201	E8	2010	48300

Fig. 2 The screenshot of the economic test dataset of Nanjing in 2010.

- The first level is the economic dataset of Nanjing in 2010.
- In the second level, the spatial information is the coordinates of Nanjing. The time is 2010.
- As for the attribute aspect, the data size is the size of this dataset. The data category is economic data.
- As for the calculation of completeness, since there are totally 129 economic factors according to the standard of Bureau of Statistics, the completeness value is equal to the total number of available factors divided by 129.

Methodology: Visualization framework

Design Goals:

- Visualize spatiotemporal distribution of datasets with space time cube.
- Incorporate visualization of data quality with space time cube.
- Incorporate visualization of data category of datasets with space time cube.
- Incorporate visualization of data size of datasets with space time cube.

Methodology: Visualization framework

Layout Design:

- Good layout design would make the visual interface clear and help users understand the map. It organizes all components of this interface.
- There are generally five components here, namely title, map, legend, filter panels, and information panel of data and contributors. The title could describe the main idea of this visual interface. The 3D map is the main component which includes all datasets. Legend is an important component in helping users better understand the map. Filter panels can allow users to perform filter functions, which can not only help users find datasets quickly but also help remove unnecessary information to give better overview. The information panel is used to give detailed description of each type of data and indicate resources.

Methodology: Visualization framework

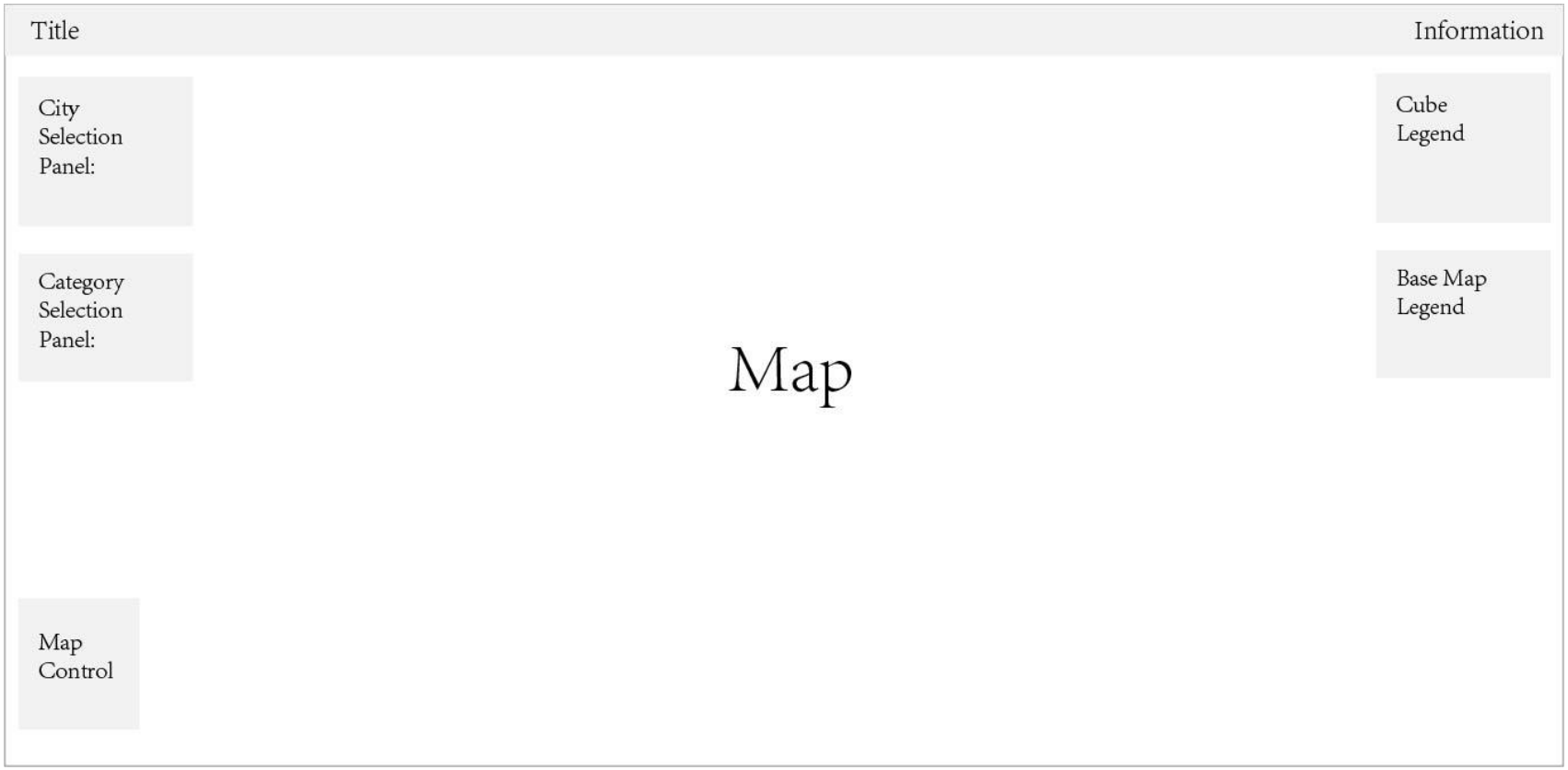

Fig. 3 Layout of the visual interface of geodata browsing using space time cube

Methodology: Visualization framework

Table. 2 Visualization of each dimension in the data framework

Dimension in the data framework	Visualization
Dataset	A cube symbol
Space	Geometric centre of the city
Time	The z-axis in the whole 3D scene represents time. The height of each cube symbol represents the time range. Each symbol is located at the corresponded position in the whole space time cube.
Data category	Colour hue
Completeness	Colour saturation
Data size	The length of the bottom edge of the dataset cube symbol.

Methodology: Visualization framework

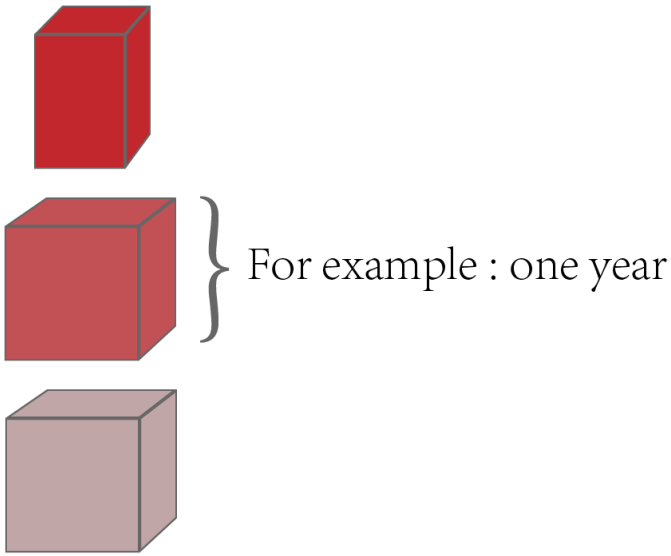

Fig. 4 Cube symbols representing datasets in the whole space time cube with size, colour hue, colour saturation representing data size, data category, and data quality separately

- The red colour indicates different data category.
- The saturation of each dataset cube in this figure is equal to the corresponding colour with full saturation multiply the corresponding completeness value of each dataset.
- The length of the bottom edge of each cuboid is proportional to the data size.
- The height of each dataset cube represents the time range and each dataset cube is located at the corresponded time position.

Methodology: Visualization framework

Base Map Design:

- Base map is necessary to show basic geographic information and can be used to show supplementary information.
- It is designed to visualize the number of available datasets to give a good overview and help users to compare the open data work of each city.
- Choropleth map with different colour values applied to different cities representing available number of datasets is used.

Methodology: Visualization framework

Functionalities:

- The first set of functionalities is related to interactivities with 3D scene, such as zooming in, zooming out, and rotation.
- The second set aims to provide more related information when user exploring this interface. When users hover a space time cube, a popup will be shown giving information about data category, data size, data quality, and time.
- The last set of functionalities is the filtering function, which is designed to help users better perform queries. Users can filter those space time cubes by city and category.

Prototype Implementation

- Jiangsu province, one of the most economically active provinces in China, is used as the test area.
- Geographic, social, economic, and night-time light data of each city in Jiangsu province are used as test datasets.

Prototype Implementation

Technologies Adapted:

- Web-related technologies: HTML, CSS, JavaScript
- JavaScript libraries: d3, mapbox gl js, threebox.js, three.js
- Moodle bundler: webpack
- For deployment: netlify
- Services: mapbox, github

Link: https://unruffled-dijkstra-9403b9.netlify.app/

Evaluation

- In order to know if the proposed visual interface could assist users finish user tasks proposed in the concept framework and the attitude of users towards using this novel visualization method here, the effectiveness and satisfaction of this visual interface were evaluated.
- The effectiveness refers to accuracy and completeness users could finish usual tasks mentioned in the concept framework.
- The satisfaction here refers to the comfort and satisfaction when users use this visual interface.
- Benchmark tasks and user attitude investigation using well-structured questions in the form of online questionnaire were used.
- There were totally 32 participants with various cultural backgrounds.

Evaluation: Benchmark tasks

Table. 3 Benchmark tasks for evaluating effectiveness of this visualization platform

Tasks	T1	T2	T3	T4	T5
Aim	To test if this platform can enable users find needed datasets	To test if this platform can give an effective overview of time	To test if this platform can give an effective overview of data category.	To test if users can identify and compare data quality, which is one of this project's innovation.	To test if this platform can enable users explore spatial distribution.
State ment	In 2009, the size of economic data in Nanjing is 13kb.	Night-time light data from 2012 to 2020 in Wuxi is available.	There are two kinds of data available in Lianyungang.	The completeness of night-time light data in Suzhou in 2019 is better than that in 2017.	Institutions in the southern part of Jiangsu province published more paper than that in the northern part.

Evaluation: Result

The overall performance of this visual interface is not bad, which indicates using space time cube is a feasible way to visualize spatiotemporal datasets for geodata query. By using size, color, and saturation to visualize data size, data category, and data quality respectively, it could assist users in finding needed datasets and having good overviews of important information.

What can be improved:

- Add more data content description in the popup so that user could have a rough impression on the data content when hovering data cubes, which may help explore spatial distribution.
- Add instructions of interaction with 3D scene right before users begin to use this visual interface. It is especially useful for the people who are not familiar with web-based 3D visualization platform.
- Using adaptive design to adapt different screens and devices.
- Providing tutorial of how to use this visual interface to users may help users use this visual interface better.

Conclusion

- This research designed a visual interface for geodata query which combines the novel time visualization method, space time cube, and 3D symbols using visual variables, namely size, color, and saturation, to visualize data size, data category, and data quality respectively.
- The difficulties are choosing different visual variables for the 3D symbols in the whole space time cube to make the 3D map clear and easily understandable by users and how to use interactive functionalities to overcome the critical issues of 3D visualization such as occlusion problem.
- The utilization of space time cube and visualization of aggregated metadata give users good overview of time and other important information of all datasets, which are the innovations of this research.
- The evaluation suggested that the combination of space time cube and 3D symbols could assist users in finishing user tasks proposed in the concept framework, which lays the foundation of further researches.

Future Work

- Evaluate and compare the efficiency of using space time cube and other visualization methods to visualize datasets, e.g. using time slider to display datasets.
- Include more heterogenous datasets.
- This research only shows that the visual variables and functionalities selected here could give good overviews and assist users in finishing usual tasks. More researches about which novel visual variables or exact different values of each visual variable could improve the efficiency of this visual interface can be carried out.
- More researches on whether this visual interface can provide possibility for data exploration and interesting finds can be carried out.

References

Bereta, K., Nikolaou, C., Karpathiotakis, M., Kyzirakos, K., & Koubarakis, M. (2013). SexTant: Visualizing Time-Evolving Linked Geospatial Data. International Semantic Web Conference (Posters & Demos), 177–180.

Davies, T. (2010). Open data, democracy and public sector reform. A Look at Open Government Data Use from Data. Gov. Uk, 1–47.

de León, A., Wisniewki, F., Villazón-Terrazas, B., & Corcho, O. (2012). Map4rdf-faceted browser for geospatial datasets. Proceedings of the First Workshop on USING OPEN DATA. W3C.

Haeberling, C., Bär, H., & Hurni, L. (2008). Proposed cartographic design principles for 3D maps: A contribution to an extended cartographic theory. Cartographica: The International Journal for Geographic Information and Geovisualization, 43(3), 175–188.

Kveladze, I. (2015). Space-time cube, design and usability [PhD Thesis]. University of Twente.

Li, X., & Kraak, M.-J. (2010). A temporal visualization concept: A new theoretical analytical approach for the visualization of multivariable spatio-temporal data. 2010 18th International Conference on Geoinformatics, 1–6.

Nikolaou, C., Dogani, K., Bereta, K., Garbis, G., Karpathiotakis, M., Kyzirakos, K., & Koubarakis, M. (2015). Sextant: Visualizing time-evolving linked geospatial data. Journal of Web Semantics, 35, 35–52.

Ramalho, L. de F., & Segundo, W. R. (2020). R-Shiny as an Interface for Data Visualization and Data Analysis on the Brazilian Digital Library of Theses and Dissertations (BDTD). Publications, 8(2), 24.

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. Information Polity, 22(1), 1–24.

Stadler, C., Martin, M., & Auer, S. (2014). Exploring the web of spatial data with facete. Proceedings of the 23rd International Conference on World Wide Web, 175–178.

Xiao, F., Wang, Z., & He, D. (2020). Understanding users' accessing behaviors to local Open Government Data via transaction log analysis. Proceedings of the Association for Information Science and Technology, 57(1), e278.


ITC
UNIVERSITY OF TWENTE.
TECHNISCHE
UNIVERSITÄT
DRESDEN
Technical
University
of Munich
TU
WIEN
TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Appendix 1: Result of Each Question in User Attitude Investigation

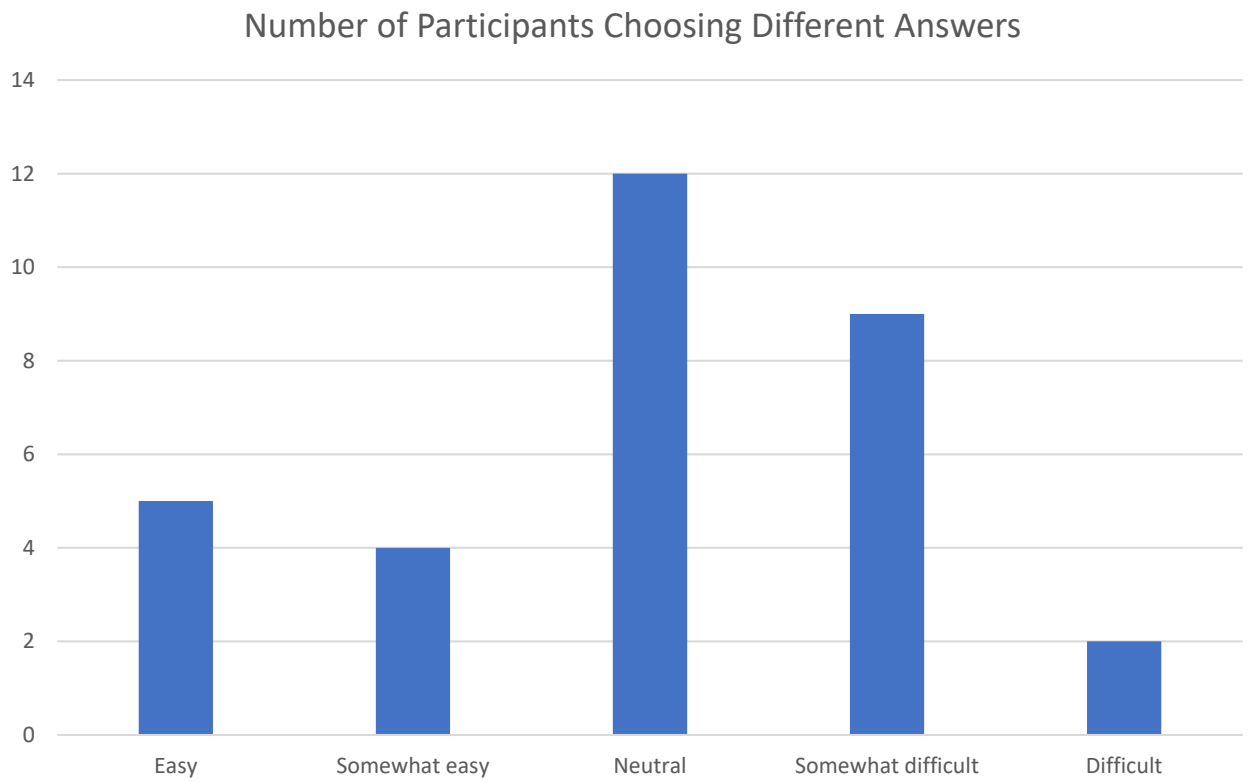

Fig. 3 The result of question *"How do you rank the overall ease of usability of this visualization interface?"*

Appendix 1: Result of Each Question in User Attitude Investigation

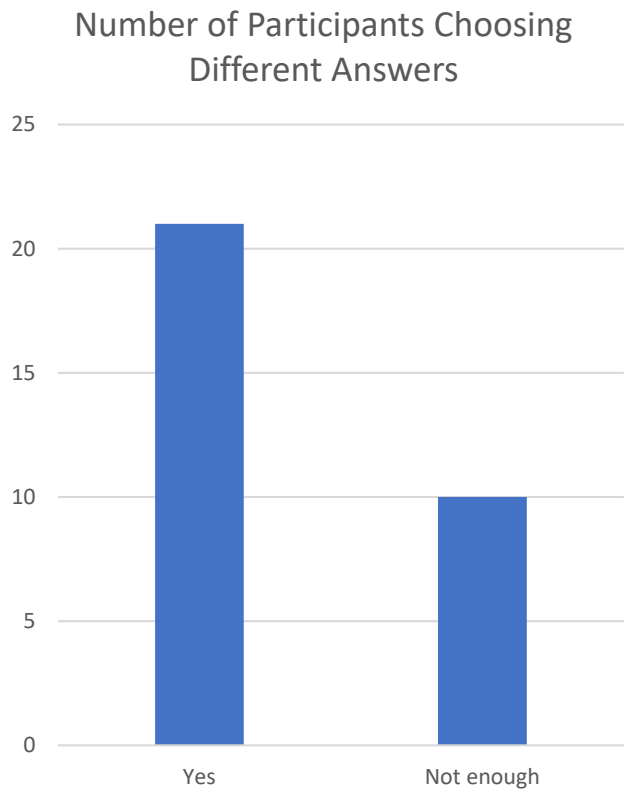

Fig. 4 The result of question “Do you think the map elements are sufficient enough to assist you in fulfilling all tasks?”

Appendix 1: Result of Each Question in User Attitude Investigation

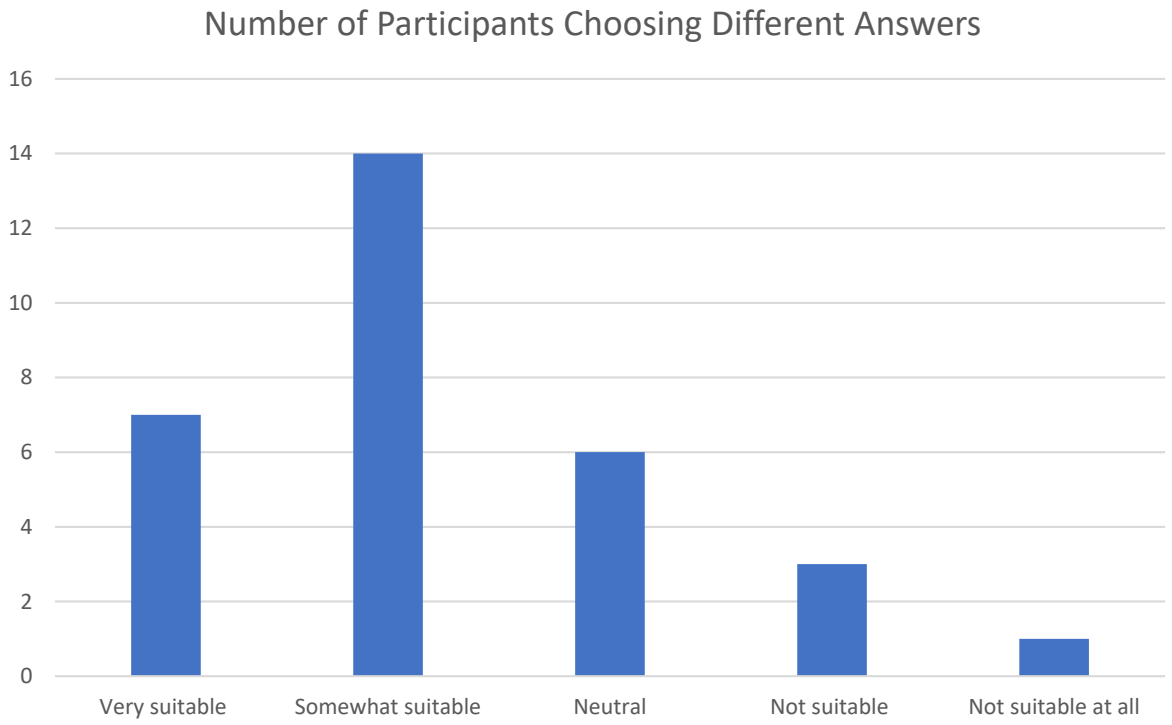

Fig. 5 The result of question "Do you think the space time cube is suitable to visualize spatiotemporal data in this platform?"

Appendix 2: Evaluation Result

Constructive feedbacks from participants

Number of participants	Feedback
Five	It would be better to allow users to rotate the scene, which could solve the occlusion problem and make it easier to navigate.
Two	It would be better to implement adaptive design.
Two	The colour scheme could change since the colours used to represent economic data and social data are quite similar.
One	2D visualization of time such as time line may be more suitable here.
One	It would be better to add more information in the popup such as city, detail description of data.
One	It would be clearer to add labels to different years.
One	In the city selection panel, it would be better if these cities are ordered alphabetically.