**Author:** ﻿Bostock, Michael and Ogievetsky, Vadim and Heer, Jeffrey

**Title:** ﻿D3 data-driven documents

**Venue:** *IEEE transactions on visualization and computer graphics* 17.12 (2011): 2301-2309.

**Year:** 2011

**Number Citations:** 2161

**Aim:** When building visualizations, designers often employ multiple tools simultaneously. This is particularly true on the web, where interactive visualizations combine varied technologies: HTML for page content, CSS for aesthetics, JavaScript for interaction, SVG for vector graphics, and so on. The great successes of the web as a platform is the (mostly) seamless cooperation of such technologies, enabled by a shared representation of the page called the document object model (DOM). Unfortunately, this blissful interoperability is typically lost with visualization toolkits due to encapsulation of the DOM with more specialized forms.

**Conclusion:** Moreover, our comparison to Protovis is conservative, as in our benchmarks the majority of the scene must be redrawn on each frame. This provides a useful bound on performance, but obscures the common case of more localized updates. By limiting updates to the changing parts of a scene, D3 transforms provide greater scalability than Protovis. D3 also allows more control over document structure, allowing further optimization; for example, SVG's use element efficiently replicates shapes, while CSS3 provides hardware acceleration of certain animated transitions.

**How this informs some of the design interventions/feedback?**

Based on this paper, we decided to use d3 to make map and other function. And we also find that post-hoc manipulation of visualizations through the developer console is a unique and compelling benefit of D3's design.

**Author:** ﻿Esteban-Guitart, Moises and Moll, Luis C

**Title:** ﻿Lived experience, funds of identity and education

**Venue:** *Culture & Psychology* 20.1 (2014): 70-81.

**Year:** 2014

**Number Citations:** 64

**Aim:** The aim of this paper is to contribute to the discussion on funds of identity. First, we emphasize the extent to which history and time are constitutive dimensions of culture and experience and the affect this has on funds of identity and meaning construction. This implies seeing the students’ funds of identity as resources for learning and seeing schools as a context that must also be linked to other practices and activities in which people are involved.

**Conclusion:** The intention behind the concept of funds of identity is to overcome certain limitations in the funds of knowledge approach. The first limitation is the primacy and exclusivity that is given to families as the focus of attention when documenting a student’s funds of knowledge. The second limitation of the funds of knowledge approach is the fact that the ethnographic research carried out by teachers in visits made to their students’ homes has been based almost exclusively on the use of interviews.

**How this informs some of the design interventions/feedback?**

According to this paper, we understand the impact of funds and student achievement or family, and we have a more profound impact on the relationship between education and funding, which makes us see the importance of our project.

**Author:** ﻿Chen, Baizhu and Feng, Yi

**Title:** ﻿Determinants of economic growth in China: Private enterprise, education, and openness

**Venue:** *China Economic Review* 11.1 (2000): 1-15.

**Year:** 2000

**Number Citations:** 441

**Aim:** This essay investigates the source of cross-provincial variations of economic growth in China. A statistical analysis of data on 29 provinces, municipalities, and autonomous regions from 1978 through 1989 confirms the findings in the literature of empirical studies of economic growth based upon cross- We find that private and semi-private enterprises, higher education and international trade all lead to an increase in economic growth in China. We also find that high fertility, high inflation, and the Presence of state-owned enterprises (SOE) reduce growth rates among the provinces.

**Conclusion:** The economic reform that started in 1978 has brought across-the-board benefits to all provinces in China. However, the coastal provinces have experienced much higher economic growth than the inner provinces. The uneven economic growth has resulted in a wider income gap between the coastal and inner provinces. Local and central governments should emphasize education in the inner provinces. Not only should a larger share of government expenditures be spent on local school facilities and teachers' compensation, but also a labor system should be established to reward education. The system that retains the best and brightest brains and provides incentives for young people to be educated is one based on market mechanisms that link social status and financial incentives to educational attainment.

**How this informs some of the design interventions/feedback?**

According to this paper, we understand the important impact of increased education on economic development. Therefore, all provinces in China have increased their education funds year by year and improved policies to attract talents.

**Author:** ﻿Yu, Chong Ho

**Title:** ﻿Exploratory data analysis

|  |
| --- |
| **Venue:** *Methods* 2 (1977): 131-160. |

**Year:** 1977

**Number Citations:** 18439

**Aim:** This research tradition was founded by John Tukey, who often relates EDA

To detective work. In EDA, the role of the researcher is to explore the data in as many ways as

Possible until a plausible “story” emerges. A detective does not collect just any information.

He or she collects clues related to the central question of the case. By the same token, EDA is not

"fishing" or "torturing" the data set until it confesses. Rather, it is a systematic way to investigate

Relevant information from multiple perspectives. Tukey emphasizes the role of data analysis in

Research, rather than mathematics, statistics, and probability.

**Conclusion:** Tukey is well aware of this potential pitfall in confirmatory data analysis, though he didn’t explicitly name the term “conformation bias.” As a remedy, Tukey proposes an exploratory approach to urge researchers consider the otherwise. Confirmation bias is related to another psychological weakness: false sense of certainty. The traditional statistical approach that presents the finding in a confirmatory tone is embraced by the audience who prefers certainty to ambiguity. Tukey creates a paradigm shift by asserting that progress of statistics can only be made when analysts move away from certainty.

**How this informs some of the design interventions/feedback?**

According to this paper, we realize that in doing data analysis, we must get rid of our own bias against the data results, in order to correctly view the results of the analysis and draw correct conclusions.

**Author:** ﻿Pedregosa, Fabian and Varoquaux, Ga{\"e}l and Gramfort, Alexandre and Michel, Vincent and Thirion, Bertrand and Grisel, Olivier and Blondel, Mathieu and Prettenhofer, Peter and Weiss, Ron and Dubourg, Vincent and others

**Title:** ﻿Scikit-learn: Machine learning in Python

**Venue:** *Journal of machine learning research* 12.Oct (2011): 2825-2830.

**Year:** 2011

**Number Citations:** 16044

**Aim:** Scikit-learn is a Python module integrating a wide range of state-of-the-art machine learning algorithms for medium-scale supervised and unsupervised problems. This package focuses on bringing machine learning to non-specialists using a general-purpose high-level Language. Emphasis is put on ease of use, performance, documentation, and API consistency. It has minimal dependencies and is distributed under the simplified BSD license, encouraging its use in both academic and commercial settings.

**Conclusion:** Scikit-learn exposes a wide variety of machine learning algorithms, both supervised and unsupervised, using a consistent, task-oriented interface, thus enabling easy comparison of methods for a given application. Since it relies on the scientific Python ecosystem, it can easily be integrated into applications outside the traditional range of statistical data analysis. Importantly, the algorithms, implemented in a high-level language, can be used as building blocks for approaches specific to a use case, for example, in medical imaging (Michel et al., 2011). Future work includes online learning, to scale to large data sets

**How this informs some of the design interventions/feedback?**

At the beginning of the project, we intend to use machine learning to analyze our data to get the relationship between education funds and the number of students to expand the background of our project.

**Author:** ﻿ ﻿Rong, Xue Lan and Shi, Tianjian

**Title:** ﻿ ﻿Inequality in Chinese education

**Venue:** *Journal of Contemporary China* 10.26 (2001): 107-124.

**Year:** 2001

**Number Citations:** 168

**Aim:** This article investigates the status of educational equality in China in the context of the reform of major Chinese economic, political, and social institutions at the turn of the twenty-first century. In the first part of this article, the authors address the importance of the theoretical issue of equality in education and explore the relationship between theories of human capital, modernization, and political culture. They begin with a general introduction of the theme, and then explore the linkages between basic education (literacy and numeracy), the increasing productivity of the labor force, and the essential guarantee of basic human rights—necessary for survival. They then discuss how advanced education provides people with the necessary adaptability and creativity to perform in a world characterized by rapidly developing technology and a complex management system.

**Conclusion:** Despite the claims of Chinese officials that inequality in education has been eliminated, we found that disadvantaged groups in China have not shared in the recent economic prosperity enjoyed by the urban population in coastal regions. The educational gap among different segments may reflect either the widespread disparity in the level of economic development between different geographic locations or the historical, geographic, and sociocultural differences between provinces. Whatever the explanation however, lac k of educational investment, long identified by scholars inside and outside of China, has been a major factor in China’s current educational problems.

**How this informs some of the design interventions/feedback?**

According to this paper, we have obtained a very large education gap between provinces in China, which is not only related to the geographical location and historical reasons of each province, but also has a huge relationship with the investment of education funds. This paper, together with the previous two, adds a background to our project.

**Author:** ﻿ DiBiase, D., MacEachren, A. M., Krygier, J. B., & Reeves, C.

**Title:** Animation and the Role of Map Design in Scientific Visualization

**Venue:** *The Journal of Machine Learning Research* 14.1 (2013): 2349-2353.

**Year:** 1992

**Number Citations:** 490

**Aim:** Scientists visualize data for a range of purposes, from exploring unfamiliar data sets to communicating insights revealed by visual analyses. As the supply of numerical environmental data has increased, so has the need for effective visual methods, especially for exploratory data analysis. Map animation is particularly attractive to earth system scientists who typically study large spatio-temporal data sets. In addition to the "visual variables" of static maps, animated maps are composed of three basic design elements or "dynamic variables"–scene duration, rate of change between scenes, and scene order. The dynamic variables can be used to emphasize the location of a phenomenon, emphasize its attributes, or visualize change in its spatial, temporal, and attribute dimensions. In combination with static maps, graphs, diagrams, images, and sound, animation enhances analysts' ability to express data in a variety of complementary forms.

**Conclusion:** Visualization provides cartographers with an opportunity to play a creative role in the "grand challenge" (National Research Council 1988) of global change research. In so doing we will collaborate with expert, highly motivated scientists who rely on graphics as windows on complex computational models and voluminous model-produced data sets. Cartographers' ability to express data and concepts in multiple complementary forms will be more valuable than our suggestions on how to select a single optimal form. Seeking higher levels of abstraction is likely to be a more fruitful approach than seeking greater realism in devising alternative representations. Map animation can be used to visualize spatio-temporal data in both realistic (chronological) and abstract (reordered and paced) forms. The effectiveness of temporally and spatially abstract animated maps in prompting scientific insight should be judged not only in isolation, but also in combination with static maps, graphs, diagrams, and sound.

**How this informs some of the design interventions/feedback?**

According to this paper, we have an understanding of map design. And we know the rules of making a map. So we can use this rules and ways to make a great map for the further visualization.

**Authors:** Tsang, Mun C

**Title:** Financial reform of basic education in China

**Venue:** Economics of Education Review, 1996, 15(4): 423-444.

**Year:** 1966

**Number Citations:** 209

**Aim:** Since the early 1980s, the financing of basic education in China has moved rapidly away from a centralized system with a narrow revenue base to a decentralized system with a diversified revenue base. This paper provides a critical assessment of the impacts of the financial reform of basic education in China, focusing on issues of structure, resource mobilization, inequality, and inefficiency.

**Conclusion:** the reform in the financing of (basic) education is necessary and inevitable. In moving away from a centralized financing system, the education reform gives lower levels of government the power, responsibility and incentives to mobilize government resources for education. It also encourages and develops non -government mobilization as an increasingly important secondary source to finance education. Based on the objectives for structural change and overall resource mobilization alone, the financial reform in basic education has been a success in China.

**How the work informs your project?**

This paper is the basic inspire our project, it makes us to do the data visualization for the education funding and undergraduate population in China.

**Authors:** Harper, Jonathan and Agrawala, Maneesh

**Title:** Deconstructing and restyling D3 visualizations

**Venue:** Proceedings of the 27th annual ACM symposium on User interface software and technology. ACM, 2014: 253-262.

**Year:** 2014

**Number Citations:** 27

**Aim:** The D3 JavaScript library has become a ubiquitous tool for developing visualizations on the Web. Yet, once a D3 visualization is published online its visual style is difficult to change. This paper is a tool for deconstructing and restyling existing D3 visualizations.

**Conclusion:** They have presented a pair of tools for deconstructing and restyling D3 visualizations. The tools empower viewers to modify the visual look of existing D3 visualizations without having to understand or examine the underlying code. We believe there are several open directions for future work.

**How the work informs your project?**

It is an useful tools for d3 and it is use for our making map, the tools can analyze a D3 visualization to extract the data, the marks and the mapping between them. The tools allow users to easily modify D3 visualizations without examining the underlying code and we show how they can be used to deconstruct and restyle a variety of D3 visualizations.

**Authors:** Graves, Alvaro and Hendler, James

**Title:** Visualization tools for open government data

**Venue:** Proceedings of the 14th Annual International Conference on Digital Government Research. ACM, 2013: 136-145.

**Year:**2013

**Number Citations:** 43

**Aim:** In recent years many government organizations have implemented Open Government Data (OGD) policies to make their data publicly available. This data usually covers a broad set of domains, from financial to ecological information. While these initiatives often report anecdotal success regarding improved efficiency and governmental savings, the potential applications of OGD remain a largely uncharted territory. In this paper, we claim that there is an important portion of the population who could benefit from the use of OGD, but who cannot do so because they cannot perform the essential operations needed to collect, process, merge, and make sense of the data.

**Conclusion:** In this paper they have shown that there is a gap between current Open Data initiatives and an important part of the stakeholders of the Open Government Data Ecosystem. It showed that this gap prevents or limits stakeholders in making extensive use of publicly available data, due to lack of technical knowledge.

**How the work informs your project?**

We try to find some ways to deal with the data we download from National Bureau of Statistics of China. The paper showed us a good way to deal with the data from OGD, although we did not use it in the end but it is a good paper for other project

**Authors:** Bao, Fan and Chen, Jia

**Title:** Visual framework for big data in d3. Js

**Venue:** 2014 IEEE Workshop on Electronics, Computer and Applications. IEEE, 2014: 47-50.

**Year:** 2014

**Number Citations:** 18

**Aim:** The concepts of Data Warehouse, Cloud Computing and Big Data have been proposed during the era of data flood. By reviewing current progresses in data warehouse studies, this paper introduces a framework to achieve better visualization for Big Data. This framework can reduce the cost of building Big Data warehouses by divide data into sub dataset and visualize them respectively.

**Conclusion:** This framework for data visualization, data mining and statistical methods is not comprehensive and may not meet all the needs of users. However, based on the idea in this article, the user can build a framework according to his demand and then develop visual constraints, choose different data dimensions to statistic data.

**How the work informs your project?**

Based on the findings in this paper, we should find a nice concept deal with big data in d3, we can build a framework according to his demand and then develop visual constraints.

**Authors:** Satyanarayan, Arvind and Wongsuphasawat, Kanit and Heer, Jeffrey

**Title:** Declarative interaction design for data visualization

**Venue:** Proceedings of the 27th annual ACM symposium on User interface software and technology. ACM, 2014: 669-678.

**Year:** 2014

**Number Citations:** 43

**Aim:** Declarative visualization grammars can accelerate development, facilitate retargeting across platforms, and allow language-level optimizations. However, existing declarative visualization languages are primarily concerned with visual encoding, and rely on imperative event handlers for interactive behaviors. In response, we introduce a model of declarative interaction design for data visualizations.

**Conclusion:** The model contributes a substantive step towards enabling declarative interaction design for data visualization. An important next step is to assess the language’s accessibility through user evaluations. Are new users able to learn this model? Can experts accustomed to callback-driven programming quickly transition to a reactive model? Their work here primarily focuses on the design of model primitives, and they only implemented cursory optimizations of our reactive dataflow graph. While existing performance is adequate for many common visualization scenarios, more extensive performance optimization is possible. For example, insights from the streaming data literature might be applied to optimize interactive queries.

**How the work informs your project?**

Based on the findings in this paper, it showed a good model of declarative interaction design for data visualizations.

**Authors:** Healey, Christopher G

**Title:** Choosing effective colours for data visualization

**Venue:** Proceedings of Seventh Annual IEEE Visualization'96. IEEE, 1996: 263-270.

**Year:** 1996

**Number Citations:** 342

**Aim:** In this paper we describe a technique for choosing multiple colors for use during data visualization. Our goal is a systematic method for maximizing the total number of colors available for use, while still allowing an observer to rapidly and accurately search a display for any one of the given colors.

**Conclusion:** This paper has presented a simple method for choosing effective colors for use during data visualization. We accomplish this by controlling color distance, linear separation of colors, and color category similarity during color selection.

**How the work informs your project?**

Based on the findings in this paper, it is useful for us to choose colors for our map.

**Authors:** Zhu, Nick Qi

**Title:** Data visualization with D3. js cookbook

**Venue:** Zhu N Q. Data visualization with D3. js cookbook[M]. Packt Publishing Ltd, 2013.

**Year:** 2013

**Number Citations:** 50

**Aim:** Packed with practical recipes, this is a step-by-step guide to learning data visualization with D3 with the help of detailed illustrations and code samples. If you are a developer familiar with HTML, CSS, and JavaScript, and you wish to get the most out of D3, then this book is for you. This book can also serve as a desktop quick-reference guide for experienced data visualization developers.

**How the work informs your project?**

Based on this book, it is very useful for our d3 skill, we have learned a lot of knowledge for our project by this book, This is a very detailed book for d3.

**Authors:** Wang, Chaoli and Yu, Hongfeng and Ma, Kwan-Liu

**Title:** Importance-driven time-varying data visualization

**Venue:** Wang C, Yu H, Ma K L. Importance-driven time-varying data visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2008, 14(6): 1547-1554.

**Year:** 2008

**Number Citations:** 149

**Aim:** Time-dependent simulations and time-varying data can be found in almost every major scientific discipline. Time-varying data are dynamic in nature and can be categorized by different temporal behaviors they exhibit. The first category of time-varying data is regular, which usually involves a certain phenomenon that grows, persists, and declines in several (distinct) stages.

**Conclusion:** In this paper, they introduce an approach to characterize the dynamic temporal behaviors exhibited by time-varying volume data. We derive an importance measure for each spatial block in the joint feature temporal space of the data.

**How the work informs your project?**

Based on this paper, their approach is general in the sense that it encompasses all three categories of time-varying data (regular, periodic, and turbulent). In a quantitative manner, they show that different spatial blocks have varying importance values over time and different time steps may not be equally important either. We show that there are several interesting and more cost-effective ways to visualize and understand large time-varying volume data by utilizing their importance measures.