A Systematic Review of Visualization Recommendation Systems: Goals, Strategies, Interfaces, and Evaluations

Other titles in Foundations and Trends® in Databases

Learned Query Optimizers
Bolin Ding, Rong Zhu and Jingren Zhou
ISBN: 978-1-63828-382-9

More Modern B-Tree Techniques
Goetz Graefe
ISBN: 978-1-63828-372-0

Data Structures for Data-Intensive Applications: Tradeoffs and Design Guidelines
Manos Athanassoulis, Stratos Idreos and Dennis Shasha
ISBN: 978-1-63828-184-9

Consensus in Data Management: From Distributed Commit to Blockchain
Faisal Nawab and Mohammad Sadoghi
ISBN: 978-1-63828-160-3

A Systematic Review of Visualization Recommendation Systems: Goals, Strategies, Interfaces, and Evaluations

Zehua Zeng
University of Maryland, College Park
zhzeng@terpmail.umd.edu

Leilani Battle
University of Washington
leibatt@uw.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Databases

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

Z. Zeng and L. Battle. *A Systematic Review of Visualization Recommendation Systems: Goals, Strategies, Interfaces, and Evaluations*. Foundations and Trends® in Databases, vol. 14, no. 1, pp. 1–71, 2024.

ISBN: 978-1-63828-403-1

Foundations and Trends® in Databases

Volume 14, Issue 1, 2024

Editorial Scope

Foundations and Trends® in Databases publishes survey and tutorial articles in the following topics:

- Data Models and Query Languages
- Query Processing and Optimization
- Storage, Access Methods, and Indexing
- Transaction Management, Concurrency Control and Recovery
- Deductive Databases
- Parallel and Distributed Database Systems
- Database Design and Tuning
- Metadata Management
- Object Management
- Trigger Processing and Active Databases
- Data Mining and OLAP
- Approximate and Interactive Query Processing
- Data Warehousing
- Adaptive Query Processing
- Data Stream Management
- Search and Query Integration
- XML and Semi-Structured Data
- Web Services and Middleware
- Data Integration and Exchange
- Private and Secure Data Management
- Peer-to-Peer, Sensornet and Mobile Data Management
- Scientific and Spatial Data Management
- Data Brokering and Publish/Subscribe
- Data Cleaning and Information Extraction
- Probabilistic Data Management

Information for Librarians

Foundations and Trends® in Databases, 2024, Volume 14, 4 issues. ISSN paper version 1931-7883. ISSN online version 1931-7891. Also available as a combined paper and online subscription.

Contents

1 Introduction **2**

2 Background: Key Considerations in Designing Visualization Recommendation Systems **7**
2.1 Target Users 7
2.2 Assessing the Effectiveness of Visualization Recommendations 8
2.3 Adapting Different Datasets 8
2.4 Adapting Different Tasks 9

3 Common Architectures for Visualization Recommendation Systems **10**
3.1 Algorithm 10
3.2 Interface 11
3.3 Data Processing System 12

4 Inputs to and Outputs From Visualization Recommendation Algorithms **13**
4.1 Input Corpus 13
4.2 Output Recommendations 18

5 Learning Methods **25**
5.1 Rule-Based Methods 25
5.2 Model-Based Methods 26

5.3 Programming Language-Based Methods 29
5.4 Takeaways: Selecting Appropriate Learning Methods 30

6 Common User Interface Designs 32
6.1 Classic Interfaces 32
6.2 Natural Language Interfaces 36
6.3 Interactions/Sketch 38
6.4 Takeaways: Guidelines on Selecting Intuitive Interface Designs 41

7 Evaluation Methods 43
7.1 Self-Evaluated Visualization Recommendation Systems . . 43
7.2 Evaluation Framework 44

8 Open Challenges 50
8.1 Modular Recommendation Design 50
8.2 Automated Benchmarking of Visualization Recommendation Systems 51
8.3 Refining Recommendations 52
8.4 Agency Versus Automation 53

9 Conclusion 55

References 57

A Systematic Review of Visualization Recommendation Systems: Goals, Strategies, Interfaces, and Evaluations

Zehua Zeng[1] and Leilani Battle[2]

[1] *University of Maryland, College Park, USA; zhzeng@terpmail.umd.edu*
[2] *University of Washington, Seattle, USA; leibatt@cs.washington.edu*

ABSTRACT

Visualization recommendation systems help data analysts navigate large, complex datasets by generating visualizations of meaningful patterns, outliers, and insights that could influence downstream decision-making. However, recommendations can easily mislead or confuse analysts when they are not developed with care. In this survey, we review how visualization recommendation systems have been designed over the last 25 years and classify them by their underlying recommendation goals and high-level implementation strategies, including the user interfaces provided for navigating and interpreting the recommended visualizations. To understand their efficacy, we also review how visualization recommendation systems are evaluated in the literature. Given these observations, we present several open challenges and promising directions for future work in designing effective visualization recommendation systems.

Zehua Zeng and Leilani Battle (2024), "A Systematic Review of Visualization Recommendation Systems: Goals, Strategies, Interfaces, and Evaluations", Foundations and Trends® in Databases: Vol. 14, No. 1, pp 1–71. DOI: 10.1561/1900000088.

1

Introduction

Data is essential to understanding how to best support individuals, organizations, and societies. It helps us make smart decisions in many areas of life, such as whether to bring an umbrella to work, plan for retirement, or distribute vaccinations and disaster relief resources. However, humans struggle to make sense of large, complex data in its raw form (Fry, 2008). We need an intuitive representation of these data that we can observe and reason about, one popular method being to create interactive data visualizations. Specifically, we mean computer-generated images of data that people can view and manipulate through a graphical user interface (Card *et al.*, 1999). That being said, an alternative approach is to generate aggregate statistics (mean, median, standard deviation, etc.) or fit statistical models to the data (e.g., linear regression, logistic regression, etc.). However, statistical models may fail to capture the shape and nuance of the original dataset. A classic example is Anscombe's quartet (see Figure 1.1), which showcases four datasets with the same mean and variance but produce completely different scatterplots. In this case, simply looking at the statistics would lead one to believe all four datasets are the same. In contrast, the visualizations reveal drastic differences between them. Furthermore, it

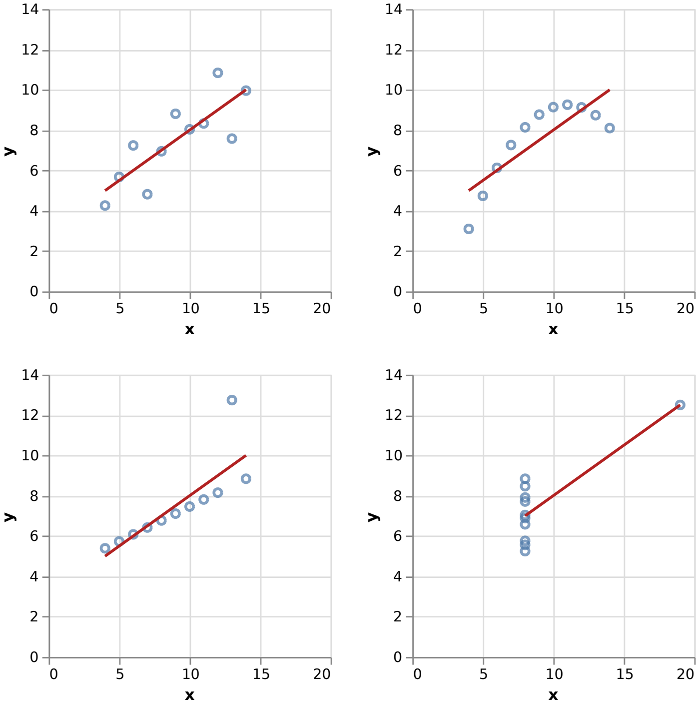

Figure 1.1: Anscombe's quartet is a group of datasets (x, y) that share identical mean, standard deviation, and regression line but which are qualitatively different. This set underscores the significance of visualizing data alongside statistical properties to gain comprehensive insights.

may be unclear what distributional characteristics to calculate without seeing the data first.

Visualization tools address this problem by generating intuitive image representations of data that allow us to ascertain the fundamental characteristics of a large dataset. In this way, we can exploit the human vision system to help analysts interpret data quickly and easily. For example, we can render summaries of key data attributes (i.e., columns) for a massive (e.g., billions of records) dataset using just a few bars in a bar chart, allowing us to better utilize screen real estate to maximize a person's interpretation of large, complex data. That being said, visualizations over large data can also be laborious to generate and can inadvertently bias or mislead users when designed incorrectly. For example, the significance of Anscombe's quartet is lost when x and y are plotted as univariate histograms rather than as a bivariate scatterplot. It can be challenging to select the right attributes to analyze, transformations to apply, and visual encodings to best visualize a dataset.

1.1 What are Visualization Recommendation Systems?

Visualization recommendation systems aim to streamline this process by automating some or even all of the choices an analyst must make when creating effective visualizations (Zeng and Battle, 2023). This can involve recommending attributes, transformations, and/or visual encodings. When recommending attributes and transformations, the goal is to select which data to display. When recommending visual encodings, the goal is to generate visualizations that maximize the viewer's ability to perceive the selected data accurately. When designed well, visualization recommendation systems can lower the barrier to creating effective data visualizations. In other words, they increase the usability of visualization tools as well as the utility of the output visualizations. Further, they can help analysts avoid drawing biased or erroneous conclusions from the data by steering analysts away from common analysis and perceptual pitfalls. However, when designed poorly, visualization recommendation systems can perpetuate the problems they intended to solve (Correll, 2019). We and others observed many scenarios under which visualization recommendation systems generate sub-optimal results (Zeng *et al.*, 2022), and even worse, promote questionable practices such as p-hacking (Zgraggen *et al.*, 2018) or cognitive biases (Wall *et al.*, 2019).

1.2 What is Interesting About Visualization Recommendation Systems?

We observe that the key challenge in designing visualization recommendation systems lies in appropriately considering the user's analysis context, i.e., their analysis goals, data expertise, and potential biases. Knowing what domain context to include and when to include it can enable visualization recommendation systems to prune as well as contextualize generated recommendations. For example, when the user has little to no experience with a specific dataset or domain, then introductory visualizations such as summary histograms of individual attributes could be useful (Wongsuphasawat *et al.*, 2017). When the user has extensive knowledge of a topic or domain, these same recommendations may seem redundant or generic (Bao *et al.*, 2022).

A second challenge in designing effective visualization recommendation systems lies in how to incorporate best practices when extracting meaningful insights from data and visualizing them for various audiences. For example, Zeng and Battle (2023) found 59 different papers from the graphical perception literature that provides guidance on selecting effective visual encodings, yet very few of these papers were ever used in the design of visualization recommendation systems (Zeng *et al.*, 2023). Similarly, Zeng *et al.* (2022) reviewed many different strategies for extracting insights from data in the literature; however, most visualization recommendation systems are only built around one or two of these methods. Ideally, the design of a visualization recommendation system should take all three factors into account (user analysis context, insight extraction methods, and graphical perception guidelines). However, the vast literature across these three topics can be difficult to navigate, given how they span human-computer interaction, data visualization, data management, and data mining research.

1.3 What Do We Aim for Readers to Take Away From This Survey?

In this survey, we review the state of the art in visualization recommendation systems and algorithms. We discuss how user analysis context is currently encoded within visualization recommendation systems. We review how user context is used to inform the detection and extraction of insights from large, complex data. Further, we consider how existing techniques translate these insights into perceptually effective visualizations to be recommended to an analyst. Finally, we discuss key challenges and opportunities for future research. Through this survey, we hope to provide readers with:

- A firm grasp of the key considerations in designing visualization recommendation systems (Section 2).
- Awareness of the high-level strategies for implementing visualization recommendation systems (Section 3–Section 5).
- An understanding of what it means for a visualization recommendation system to be "effective" or not (Section 7).
- Actionable research problems/questions that one could work on in future research (Section 8).

References

Amar, R., J. Eagan, and J. Stasko. (2005). "Low-Level Components of Analytic Activity in Information Visualization". In: *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization. INFOVIS '05*. USA: IEEE Computer Society. 15. DOI: 10.1109/INFOVIS.2005.24.

Bach, B., E. Freeman, A. Abdul-Rahman, C. Turkay, S. Khan, Y. Fan, and M. Chen. (2023). "Dashboard Design Patterns". *IEEE Transactions on Visualization and Computer Graphics*. 29(1): 342–352. DOI: 10.1109/TVCG.2022.3209448.

Bako, H., A. Varma, A. Faboro, M. Haider, F. Nerisse, B. Kenah, and L. Battle. (2022). "Streamlining Visualization Authoring in D3 Through User-Driven Templates". In: *(to appear in) IEEE TVCG. IEEE VIS '22 Short Papers*.

Bako, H. K., A. Varma, A. Faboro, M. Haider, F. Nerrise, B. Kenah, J. P. Dickerson, and L. Battle. (2023). "User-Driven Support for Visualization Prototyping in D3". In: *Proceedings of the 28th International Conference on Intelligent User Interfaces. IUI '23*. Sydney, NSW, Australia: Association for Computing Machinery. 958–972. DOI: 10.1145/3581641.3584041.

Bao, C. S., S. Li, S. G. Flores, M. Correll, and L. Battle. (2022). "Recommendations for Visualization Recommendations: Exploring Preferences and Priorities in Public Health". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. CHI '22*. New Orleans, LA, USA: Association for Computing Machinery. DOI: 10.1145/3491102.3501891.

Battle, L., R. Chang, and M. Stonebraker. (2016). "Dynamic Prefetching of Data Tiles for Interactive Visualization". In: *Proceedings of the 2016 International Conference on Management of Data. SIGMOD '16*. San Francisco, California, USA: Association for Computing Machinery. 1363–1375. DOI: 10.1145/2882903.2882919.

Battle, L., P. Duan, Z. Miranda, D. Mukusheva, R. Chang, and M. Stonebraker. (2018a). "Beagle: Automated Extraction and Interpretation of Visualizations from the Web". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18*. New York, NY, USA: Association for Computing Machinery. 1–8. DOI: 10.1145/3173574.3174168.

Battle, L., P. Duan, Z. Miranda, D. Mukusheva, R. Chang, and M. Stonebraker. (2018b). "Beagle: Automated Extraction and Interpretation of Visualizations from the Web". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18*. Montreal QC, Canada: ACM. 594:1–594:8. DOI: 10.1145/3173574.3174168.

Battle, L., P. Eichmann, M. Angelini, T. Catarci, G. Santucci, Y. Zheng, C. Binnig, J.-D. Fekete, and D. Moritz. (2020). "Database Benchmarking for Supporting Real-Time Interactive Querying of Large Data". In: *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data. SIGMOD '20*. Portland, OR, USA: Association for Computing Machinery. 1571–1587. DOI: 10.1145/3318464.3389732.

Battle, L. and J. Heer. (2019). "Characterizing exploratory visual analysis: A literature review and evaluation of analytic provenance in tableau". In: *Computer graphics forum*. Vol. 38. No. 3. Wiley Online Library. 145–159.

Battle, L. and A. Ottley. (2023). “What Do We Mean When We Say "Insight"? A Formal Synthesis of Existing Theory”. *(To appear in) IEEE TVCG*.

Battle, L. and C. Scheidegger. (2021). “A Structured Review of Data Management Technology for Interactive Visualization and Analysis”. *IEEE Transactions on Visualization and Computer Graphics*. 27(2): 1128–1138. DOI: 10.1109/TVCG.2020.3028891.

Bieliauskas, S. and A. Schreiber. (2017). “A Conversational User Interface for Software Visualization”. In: *2017 IEEE Working Conference on Software Visualization (VISSOFT)*. 139–143. DOI: 10.1109/VISSOFT.2017.21.

Bostock, M., V. Ogievetsky, and J. Heer. (2011). “D^3 Data-Driven Documents”. *IEEE Transactions on Visualization and Computer Graphics*. 17(12): 2301–2309. DOI: 10.1109/TVCG.2011.185.

Card, S. K., J. D. Mackinlay, and B. Shneiderman. (1999). “Using vision to think”. In: *Readings in information visualization: using vision to think*. 579–581.

Casner, S. M. (1991). “Task-Analytic Approach to the Automated Design of Graphic Presentations”. *ACM Trans. Graph.* 10(2): 111–151. DOI: 10.1145/108360.108361.

Ceneda, D., T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. (2017). “Characterizing Guidance in Visual Analytics”. *IEEE Transactions on Visualization and Computer Graphics*. 23(1): 111–120. DOI: 10.1109/TVCG.2016.2598468.

Cetintemel, U., M. Cherniack, J. A. DeBrabant, Y. Diao, K. Dimitriadou, A. Kalinin, O. Papaemmanouil, and S. B. Zdonik. (2013). “Query Steering for Interactive Data Exploration.” In: *CIDR*.

Chan, S.-M., L. Xiao, J. Gerth, and P. Hanrahan. (2008). “Maintaining interactivity while exploring massive time series”. In: *2008 IEEE Symposium on Visual Analytics Science and Technology*. 59–66. DOI: 10.1109/VAST.2008.4677357.

Chen, J., M. Ling, R. Li, P. Isenberg, T. Isenberg, M. Sedlmair, T. Möller, R. S. Laramee, H.-W. Shen, K. Wünsche, and Q. Wang. (2021a). "VIS30K: A Collection of Figures and Tables From IEEE Visualization Conference Publications". *IEEE Transactions on Visualization and Computer Graphics*. 27(9): 3826–3833. DOI: 10.1109/TVCG.2021.3054916.

Chen, X., W. Zeng, Y. Lin, H. M. AI-maneea, J. Roberts, and R. Chang. (2021b). "Composition and Configuration Patterns in Multiple-View Visualizations". *IEEE Transactions on Visualization and Computer Graphics*. 27(2): 1514–1524. DOI: 10.1109/TVCG.2020.3030338.

Cleveland, W. S. and R. McGill. (1984). "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods". *Journal of the American Statistical Association*. 79(387): 531–554. DOI: 10.2307/2288400. (Accessed on 03/31/2023).

Cleveland, W. S. and R. McGill. (1985). "Graphical Perception and Graphical Methods for Analyzing Scientific Data". *Science*. 229(4716): 828–833. DOI: 10.1126/science.229.4716.828.

Correll, M. (2019). "Ethical Dimensions of Visualization Research". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19*. Glasgow, Scotland Uk: Association for Computing Machinery. 1–13. DOI: 10.1145/3290605.3300418.

Dabek, F. and J. J. Caban. (2017). "A Grammar-based Approach for Modeling User Interactions and Generating Suggestions During the Data Exploration Process". *IEEE Transactions on Visualization and Computer Graphics*. 23(1): 41–50. DOI: 10.1109/TVCG.2016.2598471.

Demiralp, Ç., P. J. Haas, S. Parthasarathy, and T. Pedapati. (2017). "Foresight: Recommending Visual Insights". *Proceedings of the VLDB Endowment*. 10(12): 1937–1940. DOI: 10.14778/3137765.3137813.

Deng, D., A. Wu, H. Qu, and Y. Wu. (2023). "DashBot: Insight-Driven Dashboard Generation Based on Deep Reinforcement Learning". *IEEE Transactions on Visualization and Computer Graphics*. 29(1): 690–700. DOI: 10.1109/TVCG.2022.3209468.

Dibia, V. and Ç. Demiralp. (2019). "Data2Vis: Automatic Generation of Data Visualizations Using Sequence-to-Sequence Recurrent Neural Networks". *IEEE Computer Graphics and Applications*. 39(5): 33–46. DOI: 10.1109/MCG.2019.2924636.

Dimitriadou, K., O. Papaemmanouil, and Y. Diao. (2014). "Explore-by-example: an automatic query steering framework for interactive data exploration". In: *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*. *SIGMOD '14*. Snowbird, Utah, USA: Association for Computing Machinery. 517–528. DOI: 10.1145/2588555.2610523.

Ding, R., S. Han, Y. Xu, H. Zhang, and D. Zhang. (2019). "QuickInsights: Quick and Automatic Discovery of Insights from Multi-Dimensional Data". In: *Proceedings of the 2019 International Conference on Management of Data*. *SIGMOD '19*. Amsterdam, Netherlands: Association for Computing Machinery. 317–332. DOI: 10.1145/3299869.3314037.

Ehsan, H., M. A. Sharaf, and P. K. Chrysanthis. (2016). "MuVE: Efficient Multi-Objective View Recommendation for Visual Data Exploration". In: *2016 IEEE 32nd International Conference on Data Engineering (ICDE)*. 731–742. DOI: 10.1109/ICDE.2016.7498285.

Eichmann, P., E. Zgraggen, C. Binnig, and T. Kraska. (2020). "IDEBench: A Benchmark for Interactive Data Exploration". In: *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. *SIGMOD '20*. Portland, OR, USA: Association for Computing Machinery. 1555–1569. DOI: 10.1145/3318464.3380574.

Epperson, W., D. Jung-Lin Lee, L. Wang, K. Agarwal, A. G. Parameswaran, D. Moritz, and A. Perer. (2022). "Leveraging Analysis History for Improved In Situ Visualization Recommendation". In: *Computer Graphics Forum*. Vol. 41. No. 3. Wiley Online Library. 145–155.

Fast, E., B. Chen, J. Mendelsohn, J. Bassen, and M. S. Bernstein. (2018). "Iris: A Conversational Agent for Complex Tasks". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. *CHI '18*. New York, NY, USA: Association for Computing Machinery. 1–12. DOI: 10.1145/3173574.3174047.

Fry, B. (2008). *Visualizing data: exploring and explaining data with the Processing environment*.

Gathani, S., S. Monadjemi, A. Ottley, and L. Battle. (2022). "A Grammar-Based Approach to Applying Visualization Taxonomies to Interaction Logs". In: *(to appear in) Computer Graphics Forum. EuroVis '22*.

Gilson, O., N. Silva, P. W. Grant, and M. Chen. (2008). "From web data to visualization via ontology mapping". In: *Computer graphics forum*. Vol. 27. No. 3. Wiley Online Library. 959–966.

Gonzalez, H., A. Halevy, C. S. Jensen, A. Langen, J. Madhavan, R. Shapley, and W. Shen. (2010). "Google Fusion Tables: Data Management, Integration and Collaboration in the Cloud". In: *Proceedings of the 1st ACM Symposium on Cloud Computing. SoCC '10*. Indianapolis, Indiana, USA: Association for Computing Machinery. 175–180. DOI: 10.1145/1807128.1807158.

Gotz, D. and Z. Wen. (2009). "Behavior-Driven Visualization Recommendation". In: *Proceedings of the 14th International Conference on Intelligent User Interfaces. IUI '09*. Sanibel Island, Florida, USA: Association for Computing Machinery. 315–324. DOI: 10.1145/1502650.1502695.

Grammel, L., M. Tory, and M.-A. Storey. (2010). "How Information Visualization Novices Construct Visualizations". *IEEE Transactions on Visualization and Computer Graphics*. 16(6): 943–952. DOI: 10.1109/TVCG.2010.164.

Harper, J. and M. Agrawala. (2018). "Converting Basic D3 Charts into Reusable Style Templates". *IEEE Transactions on Visualization and Computer Graphics*. 24(3): 1274–1286. DOI: 10.1109/TVCG.2017.2659744.

Hearst, M. and M. Tory. (2019). "Would You Like A Chart With That? Incorporating Visualizations into Conversational Interfaces". In: *2019 IEEE Visualization Conference (VIS)*. 1–5. DOI: 10.1109/VISUAL.2019.8933766.

Heer, J. (2019). "Agency plus automation: Designing artificial intelligence into interactive systems". *Proceedings of the National Academy of Sciences*. 116(6): 1844–1850. DOI: 10.1073/pnas.1807184115.

Hoque, E., P. Kavehzadeh, and A. Masry. (2022). "Chart Question Answering: State of the Art and Future Directions". *Computer Graphics Forum*. 41(June): 555–572. DOI: 10.1111/cgf.14573.

Horvitz, E. (1999). "Principles of mixed-initiative user interfaces". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '99*. Pittsburgh, Pennsylvania, USA: Association for Computing Machinery. 159–166. DOI: 10.1145/302979.303030.

Hu, K., M. Bakker, S. Li, T. Kraska, and C. Hidalgo. (2019a). "VizML: A machine learning approach to visualization recommendation". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '19*. New York, NY, USA: Association for Computing Machinery. 1–12. DOI: 10.1145/3290605.3300358.

Hu, K., S. '. S. Gaikwad, M. Hulsebos, M. A. Bakker, E. Zgraggen, C. Hidalgo, T. Kraska, G. Li, A. Satyanarayan, and Ç. Demiralp. (2019b). "VizNet: Towards A Large-Scale Visualization Learning and Benchmarking Repository". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19*. Glasgow, Scotland Uk: Association for Computing Machinery. 1–12. DOI: 10.1145/3290605.3300892.

Hu, K., S. '. S. Gaikwad, M. Hulsebos, M. A. Bakker, E. Zgraggen, C. Hidalgo, T. Kraska, G. Li, A. Satyanarayan, and Ç. Demiralp. (2019c). "VizNet: Towards A Large-Scale Visualization Learning and Benchmarking Repository". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19*. Glasgow, Scotland Uk: Association for Computing Machinery. 1–12. DOI: 10.1145/3290605.3300892.

Hu, K., D. Orghian, and C. Hidalgo. (2018). "DIVE: A Mixed-Initiative System Supporting Integrated Data Exploration Workflows". In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics. HILDA'18*. Houston, TX, USA: Association for Computing Machinery. DOI: 10.1145/3209900.3209910.

John, R. J. L., N. Potti, and J. M. Patel. (2017). "Ava: From Data to Insights Through Conversations". In: *Conference on Innovative Data Systems Research*. URL: https://api.semanticscholar.org/CorpusID:1519997.

Kandel, S., A. Paepcke, J. M. Hellerstein, and J. Heer. (2012). "Enterprise Data Analysis and Visualization: An Interview Study". *IEEE Transactions on Visualization and Computer Graphics*. 18(12): 2917–2926. DOI: 10.1109/TVCG.2012.219.

Kandel, S., A. Paepcke, J. Hellerstein, and J. Heer. (2011). "Wrangler: interactive visual specification of data transformation scripts". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11*. New York, NY, USA: Association for Computing Machinery. 3363–3372. DOI: 10.1145/1978942.1979444.

Key, A., B. Howe, D. Perry, and C. Aragon. (2012). "VizDeck: Self-Organizing Dashboards for Visual Analytics". In: *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. SIGMOD '12*. Scottsdale, Arizona, USA: Association for Computing Machinery. 681–684. DOI: 10.1145/2213836.2213931.

Kim, Y. and J. Heer. (2018). "Assessing Effects of Task and Data Distribution on the Effectiveness of Visual Encodings". *Comput. Graph. Forum*. 37(3): 157–167. DOI: 10.1111/cgf.13409.

Kim, Y., K. Wongsuphasawat, J. Hullman, and J. Heer. (2017). "GraphScape: A Model for Automated Reasoning about Visualization Similarity and Sequencing". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17*. Denver, Colorado, USA: Association for Computing Machinery. 2628–2638. DOI: 10.1145/3025453.3025866.

Kristiansen, Y. S., L. Garrison, and S. Bruckner. (2022). "Semantic Snapping for Guided Multi-View Visualization Design". *IEEE Transactions on Visualization and Computer Graphics*. 28(1): 43–53. DOI: 10.1109/TVCG.2021.3114860.

Lee, D. J.-L., J. Lee, T. Siddiqui, J. Kim, K. Karahalios, and A. Parameswaran. (2020). "You can't always sketch what you want: Understanding Sensemaking in Visual Query Systems". *IEEE Transactions on Visualization and Computer Graphics*. 26(1): 1267–1277. DOI: 10.1109/TVCG.2019.2934666.

Lee, D. J.-L., V. Setlur, M. Tory, K. Karahalios, and A. Parameswaran. (2022). "Deconstructing Categorization in Visualization Recommendation: A Taxonomy and Comparative Study". *IEEE Transactions on Visualization and Computer Graphics*. 28(12): 4225–4239. DOI: 10.1109/TVCG.2021.3085751.

Lee, D. J.-L., D. Tang, K. Agarwal, T. Boonmark, C. Chen, J. Kang, U. Mukhopadhyay, J. Song, M. Yong, M. A. Hearst, and A. G. Parameswaran. (2021). "Lux: Always-on Visualization Recommendations for Exploratory Dataframe Workflows". *Proc. VLDB Endow.* 15(3): 727–738. DOI: 10.14778/3494124.3494151.

Lee-Robbins, E. and E. Adar. (2022). "Affective learning objectives for communicative visualizations". *IEEE Transactions on Visualization and Computer Graphics*. 29(1): 1–11.

Li, H., Y. Wang, S. Zhang, Y. Song, and H. Qu. (2022). "KG4Vis: A Knowledge Graph-Based Approach for Visualization Recommendation". *IEEE Transactions on Visualization and Computer Graphics*. 28(01): 195–205. DOI: 10.1109/TVCG.2021.3114863.

Lin, H., D. Moritz, and J. Heer. (2020). "Dziban: Balancing Agency & Automation in Visualization Design via Anchored Recommendations". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20*. Honolulu, HI, USA: Association for Computing Machinery. 1–12. DOI: 10.1145/3313831.3376880.

Luo, Y., X. Qin, N. Tang, and G. Li. (2018). "DeepEye: Towards Automatic Data Visualization". In: *2018 IEEE 34th International Conference on Data Engineering (ICDE)*. 101–112. DOI: 10.1109/ICDE.2018.00019.

Mackinlay, J. (1986). "Automating the Design of Graphical Presentations of Relational Information". *ACM Transactions on Graphics*. 5(2): 110–141. DOI: 10.1145/22949.22950.

Mackinlay, J., P. Hanrahan, and C. Stolte. (2007). "Show Me: Automatic Presentation for Visual Analysis". *IEEE Transactions on Visualization and Computer Graphics*. 13(6): 1137–1144. DOI: 10.1109/TVCG.2007.70594.

Mannino, M. and A. Abouzied. (2018). "Expressive Time Series Querying with Hand-Drawn Scale-Free Sketches". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18*. Montreal QC, Canada: Association for Computing Machinery. 1–13. DOI: 10.1145/3173574.3173962.

McNutt, A. M. and R. Chugh. (2021). "Integrated Visualization Editing via Parameterized Declarative Templates". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21*. Yokohama, Japan: Association for Computing Machinery. DOI: 10.1145/3411764.3445356.

Moritz, D., C. Wang, G. L. Nelson, H. Lin, A. M. Smith, B. Howe, and J. Heer. (2019). "Formalizing Visualization Design Knowledge as Constraints: Actionable and Extensible Models in Draco". *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 438–448. DOI: 10.1109/TVCG.2018.2865240.

Oppermann, M., R. Kincaid, and T. Munzner. (2021). "VizCommender: Computing Text-Based Similarity in Visualization Repositories for Content-Based Recommendations". *IEEE Transactions on Visualization and Computer Graphics*. 27(2): 495–505. DOI: 10.1109/TVCG.2020.3030387.

Pandey, A., A. Srinivasan, and V. Setlur. (2023). "MEDLEY: Intent-based Recommendations to Support Dashboard Composition". *IEEE Transactions on Visualization and Computer Graphics*. 29(1): 1135–1145. DOI: 10.1109/TVCG.2022.3209421.

Purich, J., H. Mahmood, D. Chou, C. Udeze, and L. Battle. (2022). "An Adaptive Benchmark for Modeling User Exploration of Large Datasets". *arXiv preprint arXiv:2203.15748*.

Qian, X., R. A. Rossi, F. Du, S. Kim, E. Koh, S. Malik, T. Y. Lee, and N. K. Ahmed. (2022). "Personalized Visualization Recommendation". *ACM Trans. Web*. 16(3). DOI: 10.1145/3538703.

Qu, Z. and J. Hullman. (2018). "Keeping Multiple Views Consistent: Constraints, Validations, and Exceptions in Visualization Authoring". *IEEE Transactions on Visualization and Computer Graphics*. 24(1): 468–477. DOI: 10.1109/TVCG.2017.2744198.

Raasveldt, M. and H. Mühleisen. (2019). "DuckDB: An Embeddable Analytical Database". In: *Proceedings of the 2019 International Conference on Management of Data. SIGMOD '19.* Amsterdam, Netherlands: Association for Computing Machinery. 1981–1984. DOI: 10.1145/3299869.3320212.

S. Kristiansen, Y., L. Garrison, and S. Bruckner. (2022). "Content-Driven Layout for Visualization Design". In: *Proceedings of the 15th International Symposium on Visual Information Communication and Interaction. VINCI '22.* New York, NY, USA: Association for Computing Machinery. DOI: 10.1145/3554944.3554950.

Saket, B., A. Endert, and Ç. Demiralp. (2019). "Task-WBased Effectiveness of Basic Visualizations". *IEEE Transactions on Visualization and Computer Graphics.* 25(7): 2505–2512. DOI: 10.1109/TVCG.2018.2829750.

Saket, B., D. Moritz, H. Lin, V. Dibia, C. Demiralp, and J. Heer. (2018). "Beyond heuristics: Learning visualization design". *arXiv preprint arXiv:1807.06641.*

Sarikaya, A., M. Correll, L. Bartram, M. Tory, and D. Fisher. (2019). "What Do We Talk About When We Talk About Dashboards?" *IEEE Transactions on Visualization and Computer Graphics.* 25(1): 682–692. DOI: 10.1109/TVCG.2018.2864903.

Satyanarayan, A., D. Moritz, K. Wongsuphasawat, and J. Heer. (2017). "Vega-Lite: A Grammar of Interactive Graphics". *IEEE Transactions on Visualization and Computer Graphics.* 23(1): 341–350. DOI: 10.1109/TVCG.2016.2599030.

Satyanarayan, A. and J. Heer. (2014). "Lyra: An interactive visualization design environment". In: *Computer graphics forum.* Vol. 33. No. 3. Wiley Online Library. 351–360.

Seo, J. and B. Shneiderman. (2004). "A Rank-by-Feature Framework for Unsupervised Multidimensional Data Exploration Using Low Dimensional Projections". In: *IEEE Symposium on Information Visualization.* 65–72. DOI: 10.1109/INFVIS.2004.3.

Setlur, V., S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang. (2016). "Eviza: A Natural Language Interface for Visual Analysis". In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology. UIST '16*. Tokyo, Japan: Association for Computing Machinery. 365–377. DOI: 10.1145/2984511.2984588.

Setlur, V., M. Correll, A. Satyanarayan, and M. Tory. (2023). "Heuristics for Supporting Cooperative Dashboard Design". *IEEE Transactions on Visualization and Computer Graphics*: 1–11. DOI: 10.1109/TVCG.2023.3327158.

Shen, L., E. Shen, Y. Luo, X. Yang, X. Hu, X. Zhang, Z. Tai, and J. Wang. (2023). "Towards Natural Language Interfaces for Data Visualization: A Survey". *IEEE Transactions on Visualization and Computer Graphics*. 29(6): 3121–3144. DOI: 10.1109/TVCG.2022.3148007.

Shen, L., E. Shen, Z. Tai, Y. Song, and J. Wang. (2021). "TaskVis: task-oriented visualization recommendation". In: *Proc. EuroVis*. Vol. 21. 1–5.

Siddiqui, T., A. Kim, J. Lee, K. Karahalios, and A. Parameswaran. (2016). "Effortless Data Exploration with zenvisage: An Expressive and Interactive Visual Analytics System". *Proceedings of the VLDB Endowment*. 10(4).

Siddiqui, T., P. Luh, Z. Wang, K. Karahalios, and A. Parameswaran. (2020). "ShapeSearch: A Flexible and Efficient System for Shape-Based Exploration of Trendlines". In: *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data. SIGMOD '20*. Portland, OR, USA: Association for Computing Machinery. 51–65. DOI: 10.1145/3318464.3389722.

Srinivasan, A., S. M. Drucker, A. Endert, and J. Stasko. (2019). "Augmenting Visualizations with Interactive Data Facts to Facilitate Interpretation and Communication". *IEEE Transactions on Visualization and Computer Graphics*. 25(1): 672–681. DOI: 10.1109/TVCG.2018.2865145.

Stolte, C., D. Tang, and P. Hanrahan. (2002). "Polaris: a system for query, analysis, and visualization of multidimensional relational databases". *IEEE Transactions on Visualization and Computer Graphics*. 8(1): 52–65. DOI: 10.1109/2945.981851.

Tian, Y., W. Cui, D. Deng, X. Yi, Y. Yang, H. Zhang, and Y. Wu. (2024). "ChartGPT: Leveraging LLMs to Generate Charts from Abstract Natural Language". *IEEE Transactions on Visualization and Computer Graphics*: 1–15. DOI: 10.1109/TVCG.2024.3368621.

Vartak, M., S. Rahman, S. Madden, A. Parameswaran, and N. Polyzotis. (2015). "SEEDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics". *Proceedings of the VLDB Endowment*. 8(13).

Viegas, F. B., M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. (2007). "ManyEyes: A Site for Visualization at Internet Scale". *IEEE Transactions on Visualization and Computer Graphics*. 13(6): 1121–1128. DOI: 10.1109/TVCG.2007.70577.

Voigt, M., S. Pietschmann, L. Grammel, and K. Meißner. (2012). "Context-aware recommendation of visualization components". In: *The fourth international conference on information, process, and knowledge management (eKNOW)*. Vol. 1. 101–109.

Wall, E., J. Stasko, and A. Endert. (2019). "Toward a Design Space for Mitigating Cognitive Bias in Vis". In: *2019 IEEE Visualization Conference (VIS)*. 111–115. DOI: 10.1109/VISUAL.2019.8933611.

Wang, C., Y. Feng, R. Bodik, I. Dillig, A. Cheung, and A. J. Ko. (2021). "Falx: Synthesis-Powered Visualization Authoring". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. *CHI '21*. Yokohama, Japan: Association for Computing Machinery. DOI: 10.1145/3411764.3445249.

Wang, C., J. Thompson, and B. Lee. (2024). "Data Formulator: AI-Powered Concept-Driven Visualization Authoring". *IEEE Transactions on Visualization and Computer Graphics*. 30(1): 1128–1138. DOI: 10.1109/TVCG.2023.3326585.

Wang, Y., Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, and D. Zhang. (2020). "DataShot: Automatic Generation of Fact Sheets from Tabular Data". *IEEE Transactions on Visualization and Computer Graphics*. 26(1): 895–905. DOI: 10.1109/TVCG.2019.2934398.

Wesley, R., M. Eldridge, and P. T. Terlecki. (2011). “An Analytic Data Engine for Visualization in Tableau”. In: *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data. SIGMOD ’11.* Athens, Greece: Association for Computing Machinery. 1185–1194. DOI: 10.1145/1989323.1989449.

Wongsuphasawat, K., D. Moritz, A. Anand, J. Mackinlay, B. Howe, and J. Heer. (2016a). “Towards a General-Purpose Query Language for Visualization Recommendation”. In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics. HILDA ’16.* San Francisco, California: Association for Computing Machinery. DOI: 10.1145/2939502.2939506.

Wongsuphasawat, K., D. Moritz, A. Anand, J. Mackinlay, B. Howe, and J. Heer. (2016b). “Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations”. *IEEE Transactions on Visualization and Computer Graphics.* 22(1): 649–658. DOI: 10.1109/TVCG.2015.2467191.

Wongsuphasawat, K., Z. Qu, D. Moritz, R. Chang, F. Ouk, A. Anand, J. Mackinlay, B. Howe, and J. Heer. (2017). “Voyager 2: Augmenting Visual Analysis with Partial View Specifications”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI ’17.* Denver, Colorado, USA: Association for Computing Machinery. 2648–2659. DOI: 10.1145/3025453.3025768.

Wu, A., Y. Wang, M. Zhou, X. He, H. Zhang, H. Qu, and D. Zhang. (2022). “MultiVision: Designing Analytical Dashboards with Deep Learning Based Recommendation”. *IEEE Transactions on Visualization and Computer Graphics.* 28(1): 162–172. DOI: 10.1109/TVCG.2021.3114826.

Wu, A., L. Xie, B. Lee, Y. Wang, W. Cui, and H. Qu. (2021). “Learning to Automate Chart Layout Configurations Using Crowdsourced Paired Comparison”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI ’21.* Yokohama, Japan: Association for Computing Machinery. DOI: 10.1145/3411764.3445179.

Yang, J., P. F. Gyarmati, Z. Zeng, and D. Moritz. (2023). “Draco 2: An Extensible Platform to Model Visualization Design”. In: *(to appear in) IEEE TVCG. IEEE VIS 2023 Short Papers.*

Zeng, Z. and L. Battle. (2023). "A Review and Collation of Graphical Perception Knowledge for Visualization Recommendation". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. CHI '23*. Hamburg, Germany: Association for Computing Machinery. DOI: 10.1145/3544548.3581349.

Zeng, Z., P. Moh, F. Du, J. Hoffswell, T. Y. Lee, S. Malik, E. Koh, and L. Battle. (2022). "An Evaluation-Focused Framework for Visualization Recommendation Algorithms". *IEEE Transactions on Visualization and Computer Graphics*. 28(1): 346–356. DOI: 10.1109/TVCG.2021.3114814.

Zeng, Z., J. Yang, D. Moritz, J. Heer, and L. Battle. (2023). "Too Many Cooks: Exploring How Graphical Perception Studies Influence Visualization Recommendations in Draco". *IEEE Transactions on Visualization and Computer Graphics*: 1–11. DOI: 10.1109/TVCG.2023.3326527.

Zgraggen, E., Z. Zhao, R. Zeleznik, and T. Kraska. (2018). "Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18*. Montreal QC, Canada: Association for Computing Machinery. 1–12. DOI: 10.1145/3173574.3174053.

Zhang, S., H. Li, H. Qu, and Y. Wang. (2023a). "AdaVis: Adaptive and Explainable Visualization Recommendation for Tabular Data". *IEEE Transactions on Visualization and Computer Graphics*. (01): 1–14. DOI: 10.1109/TVCG.2023.3316469.

Zhang, S., H. Li, H. Qu, and Y. Wang. (2023b). "AdaVis: Adaptive and Explainable Visualization Recommendation for Tabular Data". *IEEE Transactions on Visualization and Computer Graphics*: 1–14. DOI: 10.1109/TVCG.2023.3316469.

Zong, J., D. Barnwal, R. Neogy, and A. Satyanarayan. (2021). "Lyra 2: Designing Interactive Visualizations by Demonstration". *IEEE Transactions on Visualization & Computer Graphics (Proc. IEEE InfoVis)*. DOI: 10.1109/TVCG.2020.3030367.