International Journal of Advances in Electronics and Computer Science, ISSN(p): 2394-2835 Volume-4, Issue-11, Nov.-2017
http://iraj.in

TYPES OF ADVANCED DATA VISUALISATION: STRUCTURE AND PURPOSE

GULLU YAKAR

Graphic Department Necmettin Erbakan University Turkey
E-mail: gyakar@konya.edu.tr

Abstract: Data visualisation is graphical edition in order to reveal important points of data with intensive content and to present them as easily perceivable. In the analysis of scientific researches, statistical findings and results of market research, to provide the discovery of data structure is one of the purposes of data visualisation. Another purpose of it is to make communication easier by using results of selection which is targeted to be easily understood by audience in such fields as developments as to agenda. Data visualisation is common product of the fields of statistics, information technologies, broadcasting and visual communication design. This research uses the method of literature compilation with the aim of representing factors in the choice of data visualisation methods. In the direction of development of technology and change of communication forms, it is foreseen that applications of visualisations will change and the research will form a basis for new studies in that sense.

Keywords: Data visualisation; chart; diagram; thematic map; infographic

I. INTRODUCTION

Because of creating easily and quickly perceivable and recollective solutions for long-term, visual communication is more effective compared to other types of communication. Since individuals frequently encounter color and marker coding, they have mental schemes as to this kind of expression. As reader needs to conduct mental reconsideration as to text (Becer, 1999; p.198), they use current mental schemes for visual expressions. It is one of the advantages that visual communication is independent from the notion of language and also universal.

Visual communication draw its strength from the capability of simplification. In order to emphasize the impact simplified presentation of information on audience, Baer, who cited from Horn (2008, p. 20), includes the expression "What we need is not more information but the ability to present the right information to the right people at the right time, in the most effective and efficient form". Simple, functional and attractive presentation of information presents a solution to overcome informational convergence exposed in social and occupational lives.

1.1. Basic notions in data visualisation

The first subject to be handles about determining the type of data visualisation is data and characteristics belonging to data. Data, which is recognized as "unprocessed information", is quantitative or qualitative, multivariate and abstract. Editing by charts, diagrams and illustrations is called data visualisation to make data understandable and be able to analyze its structural specifications, distribution and relations. This editing can be developed by such additions as photograph, typography, animation and sound. Because it takes advantage of theories and practices of statistics, information technologies, broadcasting and visual design, data visualisation is interdisciplinary.

Likely applications of visualisation can send message by itself, it can be used as supporting element to increase the clarity of written communication. Such visualisations as illustration and chart is more quickly perceived compared to text. If information to be wanted to be sent is multi-dimensional, statistical or abstract, visualisation makes cognition easier. "Data visualizations can be very space efficient by visualization a large set of numbers in a small space. By designing a visualization that displays all of the data within the readers' field of view, this enables us to see the entire data set with minimal eye movement without scrolling or flipping between pages" (Krum, 2013, s. 4).

Such design elements as point, line, color, tissue, shape, dimension and location in applications of visualisation are edited by such criteria as integrity, emphasis, alignment and closeness are edited. Criteria and elements to be used are determined according to type of data. Nominal datum are categorized by using location, color and forms or sets-limits. Location, dimension, shades of a color and forms are used so that consecutive relation can be emphasized in ordinal data. By editing elements with the help of such criteria of design as location, dimension, shades of color and direction in quantitative data, the amount/frequency of variables is shown.

Table 1. Factors in method selection

Data Types	***Categorical***	- Nominal - Ordinal
	Quantitave	- Interval Scale - Ratio Scale
Database	***- Online*** ***- Fixed***	
Format	***Static***	- Printed - Digital
	Dynamic	
	Interactive	- Web - Touch Screen

Because there are more than one variable in an application of visualisation, the method to be chosen changes with respect to purpose (Table 2) and format (Figure 1). Visualisation can be in static, dynamic or interactive types. Static visualisations don't change in the sense of time-interaction. It is possible to monitor in digital platforms as it can be used in such printed materials as newspaper-magazine-book. Dynamic visualisation is editing by such additions as video and animation. The visualisations, which allow users to interfere, are called interactive. Interactive visualisation must be designed with respect to the platform where it will be watched. Kiosks to be used in the systems of guidance are large touch screens. A visualisation designed to be used on smart phones will not be suitable for touch screens.

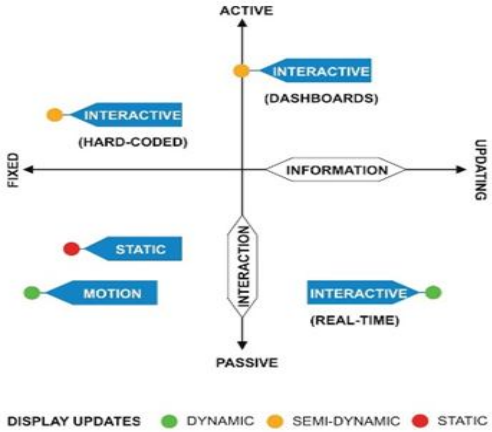

Figure 1. Visualisation formats (Lankow, Ritchie, Crooks, 2012)

Figure 1 whose original name is "infographic format quadrant" is adapted with the name "visualisation formats". This chart indicates the formation of visualisation in the sense of database-interaction variables. If data which is received from database at certain intervals is presented on design pattern, it is updating/real-time and if it was collected and processed beforehand, it is fixed/hard-coded. If visualisation can be specialized by using such specifications as choosing and filtering, it is called active and if it was prepared beforehand or updated data is watched without any interference, it is called passive.

Another classification is toward handling visualisation as abstract or pictorial (Quispel & Maes, 2013). While pictorial visualisation means schematization of the appearances of creatures, such design factors as point, line, color and tissue are used in abstract visualisation. The visualisations which photograph, pictogram and illustration are used are the pictorial kind(Figure 2).

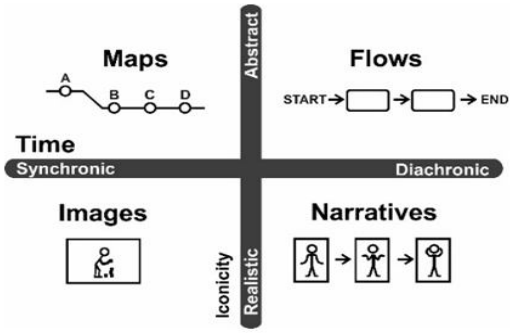

Figure 2. The Visual Archetype approach (Li, Tiwari, Alcock&Bermell-Garcia, 2016, s. 89)

1.2. Types of data visualisation

Because applications of data visualisation are common product of the fields of statistics, information Technologies and visual design, various factors can be mentioned in classification. These factors are data structure, type of visualisation, function and design specifications.

1.2.1. Types of visualisation in terms of their structures

Types of visualisation are reviewed in terms of the possibility to handle and monitor data in this section.

1.2.1.1. Basic visualisations

This type, which is used to recognize, summarize data, show tendencies, emphasize important parts and make a connection between parts, is prepared for unidimensional datum. Interaction mostly recognizes static visualisations which have no the capacity of animation or a little bit capacity.

Because it is prepared with such aims as answering a question, proving or supporting an idea, transferring information and summarizing a known subject, it is possible to recognize this type as explanatory visualisation. These visualisations, which are the product of deductive system of thinking, reviews data in a limited field because of focusing on approving the validity of a certain hypothesis.

1.2.1.2. Advanced visualisation

Real-time, interactive or dynamic visualisations, which are used in multi-dimensional data sets, are in this type. Interaction presents the opportunity to question and visualize personally to be able to discover new relations. Thus, audience can review the relation between the variables in which he is only interested in a large dataset. The fact that user can specialize visualisation by filtering in the direction of his fields of interest provides that he can deeply review tendencies and relations within dataset and access to findings he need. That's why, the type of advanced visualisation can also be named "visualisation of discovery". Visualisation of discovery has a circular structure. It allows to ask new questions, discover different dimensions of the interested subject and reveal unknown points of data instead of presenting a final outcome. Because of its open-end structure, it can be said that this type of

visualisation is the summarization of inductive thinking.

If visualisation includes both explanatory and exploratory specifications, it is named mixed model. Mixed visualisations summarize data and enable that unknown points of data can be reached for designer or audience.

1.2.2. Types of visualisation in terms of their purposes

In this section, types of visualisation are reviewed in terms of their purposes, types of variables and specifications of design. In the review, common samples are given. Few (2004, p. 5) classifies quantitative datum as nominal comparison, time-series, ranking, part-to-whole, deviation, frequency distribution, correlation in terms of the messages they send. With the aim of decreasing the number of categories, the taxonomy by Meirelles (2013) and Kirk (2012, p. 120) is taken as a reference. Common points of taxonomic findings are shown in Table 2.

Table 2. Types of visualisation in terms of their purposes

Type	Function	Visualisation	
Categorical	Comparing categorical values	bar chart, floating bar chart, histogram, dot plot, Sankey diagrams, radial chart, word cloud	
Hierarchical	Assessing hierarchies and part-of-a-whole relationships	pie chart, square pie chart, tree, treemap	
Relational	Plotting connections and relationships	bubble plot, scatter plot, network, radial network, force directed network	
Temporal	Showing changes over time	line chart, area chart, stacked area chart, stream graph, flow map, timeline, Sankey diagram	
Geo-spatial	Mapping geo-spatial data	**Color based map**	**Size-position based map**
		choropleth map, isaritmic map	dot plot, bubble plot, graduated dot symbolization, network connection

1.2.2.1. Comparing categorical values

The fact that every one of categorical data is dimensioned with respect to its numerical value provides that datum can be compared to each other in terms of amount. *Bar chart* compares quantitative categories with the help of bars whose lengths and heights change with respect to their amount. *Floating bar* or *Gantt chart*, a different type of this chart, is different from bar chart with not starting from zero point. *Histograms* are a version confused with bar chart. While bar chart compares categorical data, histogram suggests distribution by recognizing the frequency of quantitative values on vertical axis (y) and their intervals on horizontal axis (x). Unlikely bar chart, there is no gap between bars and data is read with the difference of height. Visualisations composed of the symbols, which are perceived as dot or like dot, are named *dot plot*. Categories can be separated with colors and symbols in this type orders and distributions of numerical values are shown on x-y axes. *Sankey diagrams* represent flow and relation between data. Amount is indicated with lines which have about strip thickness. Colors separate different categories, the direction and width of line show divisions within the group. *Radial charts* are based on the principle that a circle is divided into sectors and more minor and concentric circles.

1.2.2.2. Assessing hierarchies and part-of-a-whole relationships

It provides the comparison of the values of categorical datum to each other and the whole part. *Pie chart* is the most frequently-used type in monitoring the relation of parts with the whole part. *Square pie chart* or *waffle chart* is editing of quantitative datum in units within a grid with the essence of emptiness-fullness. While more than one value can be suggested with different colors, it is possible to distribute every one of value to equal-size grids. *Treemap* is based on the essence that a rectangular area is set organizedly with each other and divided into rectangular which is dimensioned in the direction of their values. A different color for every set and shades of a color can be used to emphasize hierarchical value of elements in the same set.

1.2.2.3. Plotting connections and relationships

It provides connection, distribution between values and accession to the structure in multiple data set. Plot-type charts line up two quantitative datum on x-y axes and show whether the amount of relation is high or low with respect to frequency-infrequency of units. Typical example is *scatter plot* which uses dot as unit. When a third quantitative value is added to chart, circles are used instead of dots. In this type called *bubble plot*, circles whose sizes are proportionate to the amount of quantitative data are used. In case categorical values are added, it is possible to indicate circles with different colors. Network-type charts use lines to show connections of categorical data. In *radial network*, categorical datum, which are ordered by being grouped around the circle, are connected to other datum they are correlated with lines. The degree of relation is shown by thickness of the arcs of circles which connect datum to each other by traversing the circle and categories are shown by arcs of circles in different colors. *Force-directed network (hairball)* is to connect categorical datum shown by circles to other categories they are correlated with them with lines. Circular distance of a category increases with the number of categories connected to it. In that case, the category which has lots of connections should be at

center. Every one of category is shown by different colors or color groups are created.

1.2.2.4. Showing changes over time

It is the visualisation of time-dependent quantitative change of data on a constant chart. Time is shown on horizontal axis. *Line chart* is based on the essence that continuous variable is shown on x-axis (time) and values are shown on y-axis. Dots are combined with lines in the intersection of time-value axes, therefore increases and decreases can be easily seen. If there is more than one value, different line colors can be given to every one of value. *Area chart* is the representation of areas where are denoted with colorful lines with color fill in line chart. It is sufficient that y-axis starts from the least data in line chart but y-axis of area chart is started from zero point. *Stacked area chart*, which is similar to area chart, is used in multiple data sets. Datum are lined up through up by starting from down with respect to their amounts, therefore a hierarchical relation can be also seen. Likely area chart, *stream graph* indicates multiple quantitative data as stacked. However, there is no x-axis in the form of ground in stream graph, this situation means that there is no the value of zero and chart doesn't indicate a negative value. *Flow map*, which is similar to Sankey diagrams, is a flow chart indicating the change of quantitative data over time.

1.2.2.5. Mapping geo-spatial data

It is marking of datum with respect to their geographical distributions on map. The maps, which indicate where such specific cases as social, political, cultural and economic data are seen with frequency-infrequency, the differences of colors etc., are named thematic map. Thematic maps can be grouped in terms of the elements of design used to denote quantitative data. They are maps with the specifications of colors and dimension-location. Maps with the feature of color are a type whose examples are frequently seen in such subjects as election results and weather forecasts. *Choropleth maps* are the indication of quantitative datum which belong to units with chroma, lightness-darkness or tissues in a map divided into provinces/regional units. Colors as many as the number of groups can be used by showing redundancy of amount with dark color and fewness with light color. *Isarithmic maps* use different colors for categories and saturation and lightness-darkness for the amount of every one of category. There is no contour between colors indicating categories. The mixture of more than one color on a region indicates that the categories representing colors are combined in this region. Dot plot map is marking of the distribution of data on map with respect to their geographical locations. Amount is indicated by dots which become more frequent and infrequent on map. *Bubble plot* and *graduated dot symbolization* are visualisation of quantitative data with overlapped circles on map. Bubble plot provides the separation of overlapped circles by using transparent circles. The size of circle indicates redundancy of the amount of data. *Network connection map* connects the regions correlated with each other on map with lines or arrows. The amount of relations of the regions whose number of lines is a lot is high and as number decreases, relation proportionately decreases.

1. 2.3. Other kinds of data visualisation

Textual structures and infographics, which are not included in the classification in the study, are presented in this section. In textual structures, the kind named *word clouds* or *clusteredword diagrams* can be mentioned. Books, articles, poems, contents of talk/conversation or tweets are unstructured texts. Distributions or frequencies of words, sentences or subjects in text provides the discovery of structure and the analysis of document. Dimension or line-up can be edited with respect to the frequency of words. Meirelles (2013, p. 185) presents interactive samples by classifying textual structures under a separate title, Kirk (2012, p .131) handles as static chart about the comparison of categories. Because there are samples of textual structures at the levels of both basic visualisation and advanced visualisation, it is found suitable that it is presented under an independent title in the research. *Infographics*, which are not included in any one of the types mentioned in the research, are also applications of data visualisation. Infographics can combine more than one visualisation in different types with typographic elements, illustration and graphs. It compiles the results of qualitative and/or quantitative studies within the scope of their subjects. Infographics are "illustration of information, not visualization" according to Ryan (2016, p. 152). Mainstay of this expression is explained as that "infographics express certain points of data, are aesthetically rich, but shallow and mostly static". An opponent approach in this subject is defended by Krum (2013, p.53). Krum, who presents "static-zooming-clickable-animated-video-interactive" types under the title of "infographic levels of complexity", names all the samples called application of data visualisation as infographic. Probable reason for this naming is that the full name of "infographic" is "information graphics". The term information graphics is inclusive. The fact that infographics are not only in static type is also determined by Davis and Quinn (2013, p. 16): "The infographic can be represented in a static format, designed for print, or a dynamic medium, allowing for interactivity". This situation indicates that the notions of "data visualisation" and "information graphics" in the literature have equivalent contents.

CONCLUSION

The fact that the amount of information faced in social and professional fields gradually increases and necessary time to internalize information is relatively limited increases the significance of applications of

data visualisation. Uyan Dur (2012, p. 282) handles the subject in the axis of audience-designer: "The examples of data visualization in the newspapers lead the designers to a research for more effective and accurate designs in terms of functionality and esthetics while encouraging the reader to open up her/his perception for visual analysis". As data visualisation, one of the relatively newest sub-fields of the design of visual communication, receives positive feedbacks from users, designers will increase their activities in this field. Because the discipline of visual design closely follows modern applications and innovations, this situation can be seen as an inevitable result of technological development. It is thought that the presented classification and qualifications are developed with new findings and constitute a source in the education of design in addition to that the research is reference for Turkish literature and next studies.

REFERENCES

[1] Baer, K. (2008). *Information design workbook: graphic approaches, solutions, and inspiration + 30 case studies.* Massachusetts: Rockport, (Chapter 1)

[2] Becer, E. (1999). *İletişim ve grafik tasarım*. Ankara: Dost.

[3] Davis, M. & Quinn, D. (2013). Visualizing text: the new literacy of infographics. *Reading Today*, 31, 16-18.

[4] Few, S. (2004). Eenie, meenie, minie, moe: selecting the right graph for your message. *Perceptual Edge Intelligent Enterprise*, 1-7.

[5] Katz, J. (2012). *Designing information: human factors and common sense in information design*. New Jersey: Wiley.

[6] Kirk, A. (2012). *Data visualization: a successful design process*. Birmingham: Packt, (Chapter 5).

[7] Krum, R. (2013). *Cool infographics: effective communication with data visualization and design*. New Jersey: Wiley, (Chapter 1).

[8] Lankow, J., Ritchie, J., & Crooks, R. (2012). *Infographics: the power of visual storytelling.* New Jersey: Wiley.

[9] Li, K., Tiwari, A., Alcock, J. & Bermell-Garcia, P. (2016). Categorisation of visualisation methods to support the design of human-computer interaction systems. *Journal of Applied Ergonomics*, 55, 85-107.

[10] Meirelles, I. (2013). *Design for information:.an introduction to the histories, Theories and Best Practices Behind Effective Information Visualizations.* Massachusetts: Rockport.

[11] Ryan, L. (2016). *The visual imperative: creating a visual culture of data discovery*. Massachusetts: Morgan Kaufmann.

[12] Quispel, A. & Maes, A. (2013). Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design. *Journal of Visual Languages and Computing*, 25, 107-116.

[13] Uyan Dur, B. İ. (2012). Analysis of data visualizations in daily newspapers in terms of graphic design. *Procedia-Social and Behavioral Sciences*, 51, 278-283

★★★