



A review of Data and Information Visualization Methods

Sakshi

Assistant Professor Department of Computer Science and Applications, Guru Nanak College, Ferozpur Cantt, Punjab, India.

ABSTRACT

With availability of enough visualization techniques it can be very confusing to know what and when should be appropriate technique to use in order to convey maximum possible understanding. The basic purpose of visual representation is to efficiently interpret what is insight, as easy as possible. Different available visualization techniques are use for different situation which convey different level of understanding. This document is guide for the young researchers who wants to start work in visualization. The purpose of this piece of document is to collect all visualization techniques with their brief introduction. This paper deals with many definitions and aspects of visualization, how visualization take place i.e. different steps of visualization process, problems that are confront in visualization, categorization of visualization techniques on the bases of distinct perspective, typically known common data and information visualization techniques, basic interactive methods for visualization their advantages and disadvantages, interactivity process, and the scope of visualization up to some extent in different field of research. General Terms Visualization, Visualization techniques, Challenges, Interactive techniques.

Keywords: Visualization Categorization, Visualization Challenges, Data Visualization, Information Visualization, Visualization Interactivity.

1. INTRODUCTION

The information is now the essential part of human life, which encourage new ways to think and evolve new advancement. Huge amount of information or data are generated from different sources, these information are of diverse types and stored in various format. For example, the individual go through personal information i.e. email, chat, schedules, news etc, students access countless digital libraries for educational purposes [1]. Even from all discipline of life a lot of data being generated, access, and use for individual, collective benefits. Because of such

versatility of information heave significant issues that how to represent these information in more useful manner that the user can use it efficiently. The main issue is how represent these massive data in order to understand and extract knowledge from stored data or information, which is a ubiquitous task. In simple words, the collection of information is no longer a problem but extracting of valuable knowledge from the collected or available information is a big problem. Visualization methods are considered to be very important for the users' because it provides mental models of the information [1].



Visualization techniques make huge and complex information intelligible. Information visualization is a visual user interface that provides insight of information to the user [3]. The basic purpose of visualization is to create interactive visual representations of the information that exploit human's perceptual and cognitive capabilities of problem solving [2]. The goal of visualization is that the user can easily understand and interpret huge and complex set of information. There are a lot of issues in visualization, to tackle these issues many fields should be considered. Human Computer Interaction is one of the major fields, which make things easy to use and easy to understand and interpret. All usability issues are important to think about in order to fully achieving the visual representation objectives. The main objective of this document is to provide the basic knowledge about visualization for young minds want to work in visualization. First part of this document describe basic definitions of visualization and six step process of visualization, second part describe the challenges or problems confront by visualization and categories found in literature. Then in third part of the document briefly describe Data and Information visualization methods with examples. Second last part give the Interactivity techniques that are use for visualization and the major steps that are involve in visualization interactivity. Last section describes scope of visualization in few fields and field of research that requires urgent step of action.

2. VISUALIZATION

Visualization is a mental image or a visual representation of an object or scene or person or abstraction that is similar to visual perception [10]. Visualization has many definition but the

most referred one, which is found in literature is “the use of computer-supported, interactive, visual representations of data to amplify cognition”, where cognition means the power of human perception or in simple words the acquisition or use of knowledge [6] [7]. Visualization is a graphical representation that best conveys the complicated ideas clearly, precisely, and efficiently. These graphical depictions are easily understood and interpret effectively [8] [6]. The main goal of Visualization is to find out what insight. Visualization is the transformation of Symbolic representation to geometric representation. The goal of visualization is to analyze, explore, discover, illustrate, and communicate information in well understandable form. Visualization is use to present huge amount of information coherently, compactly, from different viewpoints, and provides several levels of details [65]. Visualization evaluation is using in parallel and it is use to determine how closely the visualization goal is achieved. In evaluation there are many dimensions to observe, involving Usability evaluation, how much visualization technique is effective, accurate, robust, efficacy, and easy to use etc [1]. The visualization is a powerful tool that can be use for different cognitive processes like exploratory, analytical and descriptive [9].

2.1 Process of Visualization

The approach of designing well disciplined visualization the process can be divided into different steps. The literature [11] “Chittaro, 2006” contains a list of six different steps i.e. mapping, selection, presentation, interactivity, usability, and evaluation, which identifies major activities involving visualization, to aim precise



and error less design. The following subsection explains briefly these six activities:

1. First step of visualization process is known is Mapping. Mapping means how to visualize information or how to encode information into visual form. In mapping data or information transform into graphical form under assumption of visual features. Good mapping produce accurate visual representation, and can achieve when there is accurate relationship between data objects and visual objects to be describe. (For this purpose a well define algorithm is used).

2. Second step of visualization process is called Selection. Selection means to select data among those data which is available according to the given task or job. Selection of data is directly dependant on the aim to get through visual graphics or pictorial representation. This task in the process is the most important task, because the selection of wrong data misleads the user to take crucial decisions and suffer through huge loss (financial, time etc), should avoid the inclusion of unnecessary data.

3. Third phase of the process is Presentation. In visualization perspective presentation means how to manage, organize information in the available space on the screen effectively. After intuitive mapping, clear and precise selection of data items it is really important to present it in more meaningful and understandable form.

4. Forth step of visualization process is called Interactivity. Interactivity means what are the provided facilities to organize, explore, and rearrange the visualization. User friendly interactivity enables a user to best explore, understand, and interpret the data or information, which improve their exploration capabilities.

5. Human factors are fifth visualization feature to be considered. Human factor involve to two broad categories, usability, and accessibility factors. The visualization is easy to use for the end user and people with special need should also use it. Visual perception knowledge and cognitive aspects make it very easy to design an effective visualization [2]. These factors are the common practice of Human Computer Interaction.

6. After creating usable visualization interface, the last step is to evaluate the created visual form. Evaluation is equally important, to find out whether the visualization method has effectiveness or not, the goal is achieved or not. The challenges confront visualization evaluation is proposed by Plaisant [12]. The usability evaluation were developed for graphical user interface and adapted and extended for information visualization interfaces. Alfredo et al, categorize evaluation methods into two broad categories i.e. analytical evaluation and empirical evaluation. Analytical evaluation includes expert reviews and cognitive walkthrough. Empirical evaluation includes controlled experiments, questionnaires, interviews and focus groups. Discovery of usability problems and guide lines for modification during system development is the part of analytic evaluation. Discovery of usability problems and modification suggestions after actual implementation or prototype is the part of empirical evaluation [7].

2.2 Challenges in Visualization

The creation or production of a perfect visualization method is a big challenge in order to fulfill all the requirement of users. Visualization is suffering through many problems. In this perspective Chaomei Chen introduce a very comprehensive document with the title “Top 10



Unsolved Information Visualization Problems” [15]. When someone wants to produce as effective visualization technique, he/she should try to consider all the discussed aspects. Usability issues are critical issues for visualization, which means how to make it easy to use and efficient. The visualization should offer enough information and satisfied user. Understanding elementary perceptual– cognitive tasks is the basic step regarding information visualization engineering, providing it according to human perception capability. Requires Prior knowledge about method, how to operate and use it, it should make it more generic, means the users have common understanding about techniques. Education and Training is requires for the researchers and practitioners to share the basic principles and skills about information visualization methods. The lack of intrinsic quality measures means there is no common evaluation and selection mechanism plus the unavailability of bench marks undermine advances in visualization methods. One of the long lasting problem is Scalability, how to manage huge visualization in available space, example [16] [15]. Researchers are much focus on scalability problem in data streams, which is explain by Wong et al, 2003 [17]. Scalability is one of the big problems in mobile devices environment, because the available screen is much smaller than computer screens. To make insight of data visually appealing in visualization is important, to understand insights of data with aesthetics scenes. Aesthetic mode of an image is very difficult to maintain and a big challenge for International visualization engineers. Aesthetic is a philosophical theory deals with the beauty and taste [18], to make picture or image pretty.

Information visualization changes over time, means there is dynamism in visualization. Auto shifting of Structure to dynamic paradigm is remains a challenge for engineers. This paradigm shift is comprehensively discussed by C Chen [19]. Visual thinking, reasoning, and analytics make information visualization more powerful to determine cause and its effect or Causality. The last one is Knowledge domain visualization, based on the previous nine problems. These problems categories as, the first three problems are based on user perspective, next four are technical challenges and last three are handled on disciplinary level. Chris Johnson point out top Scientific Visualization Research problems, this document comprehensively discussed the challenges faced by scientific visualization methods [20].

2.3 Categorization of Visualization Methods

Visualization techniques or methods are categorized differently by different authors. There is three categories of visualization i.e. information visualization, software visualization, and Scientific visualization. Scientific visualization helps to understand physical phenomena in data, mathematical models, in isosurfaces, volume rendering, and glyphs etc. Software visualization helps people to learn the use of computer software, program visualization helps programmers to handle complex software, and similarly algorithm animation support, encourage, and motivate student to learn the computation capability of an algorithm. Information visualization is “the depiction of information using spatial or graphical representations to facilitate comparison, pattern recognition, change detection, and other cognitive skills by

making use of visual system". In periodic table of visualization six main categories are mentioned, i.e. Data visualization visually represents quantitative data with or without axes in schematic or diagrammatic forms e.g. Table, Line chart, Pie chart, Histogram, and Scatter plot etc. Information visualization is an interactive interface of data to increase cognition or perception ability. Transform data into a changeable image, through which users can interact during manipulation, e.g. Data map, Tree map, Clustering, Semantic network, Time line, and Venn/ Euler diagram etc. Concept visualizations are methods use to elaborate ideas, plan, concepts, and analyze it easily, e.g. Mindmap, Layer chart, Concentric circle, Decision tree, Pert chart etc. Strategic visualization is a systematic approach in which an organization visually represent it strategies of development, formulation, communication, implementation, and some time its analysis, e.g. Organizational chart, Strategy map, Failure tree, and Portfolio diagram etc. Metaphor visualization organizes and structure information graphically. They convey insight of information through key characteristics of metaphor that is employed, e.g. Metro map, Story template, Funnel, and Tree etc. Compound visualization is the complementary use of different graphic representation formats in one single schema or frame, e.g. Cartoon, Rich picture, Knowledge map, and Learning map etc [14] [66]. This document is focus on two visualization categories Data Visualization and Information Visualization. These methods are alternatively use plus very important and contain useful and informative techniques.

These methods become more important when the data is massive and need to analyze and represent it in meaningful form, which is easy to understand and interpret.

3. VISUALIZATION METHODS

A number of visualization techniques have been developed in last few years due to representation of huge massive information and to analyze it. These techniques has many features like interactivity, usability, interface features etc, which make them easy to use and beneficial as discussed. These methods have evaluation mechanism to achieve the goal of visualization as briefly discussed. In this survey the focus are on conventional Data and information visualization techniques.

3.1 Data Visualization

Data visualization is the study of representing data in some systematic form, including attributes and variables for the unit of information [21]. Information visualization, scientific visualization, information graphics and statistical graphics are closely related to data visualization. There are many conventional data visualization techniques which are focused in this document because these techniques have generic features and common understanding. These data presentation should be beautiful, elegant, descriptive, and interpretable in order to convey message to the reader effectively [22]. There are new developed fascinating methods are introducing, but modern approaches have its own implementation problems and no commonality, so difficult to adopt. Data visualization represents data in the way that simplifies data interpretation and its relationship [30]. In the following section we going to discuss few known conventional data visualization

techniques and give its example in order to get initial idea about the technique.

3.1.1 Table

Table is the simple, easy to understand, and easy to interpret data representation technique. Table is a structured format, organized by rows and columns that convey relationships, is one of the most common definitions. Row has different synonyms like record, tuple, vector etc. Column has common synonyms like field, parameter, attribute, property etc. Rows represent variables and columns represent records with the set of values. Sometimes this arrangement can be altered, means columns represent variable and rows represent records. The entries may or may not be in ordered, it depends on the data and its attributes [6] [25]. According to Fink a table is the arrangement of data in rows and in columns. The role of table is essential part in research and in data analysis. It has significant in fields like print media, computer software, conveying concepts or ideas, and many other fields. The table conventions and basic terminology is depends on the context it has been used. Table has different varieties, flexibility, structure, notations, representation, and use [26] [27]. For example, the following table shows collected data for an experiment with three measurable units.

Course Names	Studied By	Total Marks obtained	Total Marks
Software Engineering	2	127	200
Requirement Engineering	3	255	300
Multivariate Calculus	3	157	300
Software Architecture	2	141	200
Relational DBMS	3	191	300
PHP development	3	188	300
Microsoft Dot Net Platform	2	154	200
HTML & Scripting	2	143	200
Data communication	2	177	200
Computer Networks	3	254	300

Fig 1: Shows Simple Table

Hoffman introduce a new term for table as “Table Visualization”. Table visualization is the representation of data in the form of table or data in table. Row is mention with the word “Dimension”, which represent independent variable in record or tuple. Whereas the column is refer as the dependent variable called with word “Variates” [29]. It has multi applications and uses in almost all fields.

3.1.2 Pie Chart

A pie chart is also called circle graph. Pie chart circle is divided into number of sectors, each circle describe a proportion in a whole quantity [34] Spence (2005)]. The pie chart control is use to determines the size of data wedge as compare to other data wedges. In pie chart a wedge represents the part of data that has common feature or characteristics. Wedges can be labeled to identify different data points. Most of the time is shown in percentage. There are two common variations i.e. Doughnut charts and Exploding pie charts. Doughnut charts are similar to standard pie charts except they have hollow center. While in Exploding charts wedges or segment or sector can be extracted from the rest of the wedges. These sectors or wedges provide good interactions [30] [32].

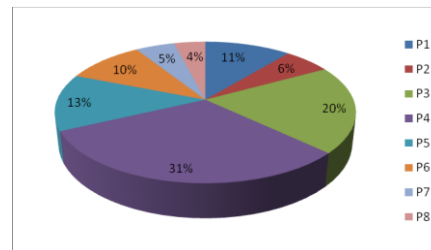


Fig 2: Standard Pie Chart

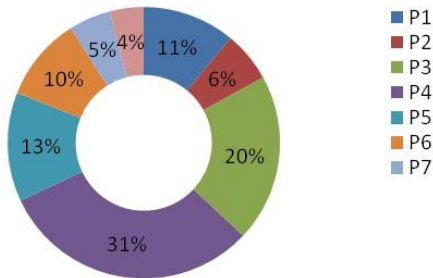


Fig 3: Doughnut Pie Chart

Multi-level Pie, Radial tree, or Ring chart is another variation of pie chart. This sort of chart is most of the time use for hierarchical data, illustrate by concentric circles. The centre circle represents the root or parent segment of the hierarchy and the upper or outer circles represent the child or down side data of the hierarchy [31] [32].

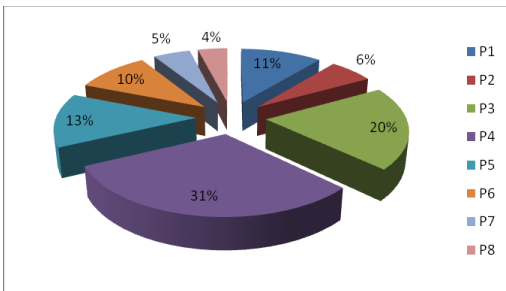


Fig 4: Exploding Pie Chart

Pie chart visualization is effective in some cases e.g. comparing a segment of pie chart to the rest of the segments of the pie chart, but it is difficult to compare different pie charts and different section of different pie charts among each other [32]. Color usability make it easier to understand and interpret, different color should use for different slice or segment.

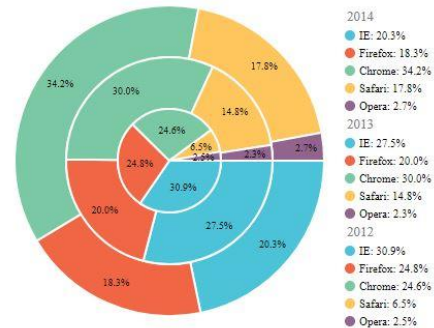


Fig 5: Multi-Level Pie Chart

The figure shows two level breakdowns, first level illustrate passengers and crew of the plan and the second level explain the gender in each portion of the first level. Multi level chart can be extend up to reasonable number of levels.

3.1.3 Bar Chart

One of the most commonly use data visualization method is bar chart, it is also called Bar Graph. In some literature bar chart is mention as column chart. Bar chart is most of the time use for discrete data not for continuous data. Bar Chart control has been use to represent data in horizontal bars, the vertical length of the bar represent the values. Bar chart is use to represent a single data series and related data points are group in one series. For example monthly salaries, it can be mutli bar graph i.e. percentage increase per month as shown in the following figures [30].

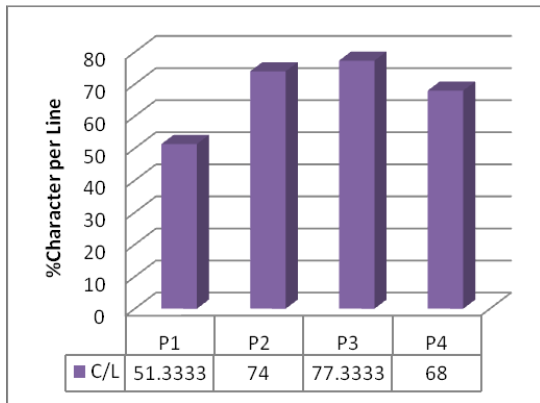


Fig 6: Single Bar Chart

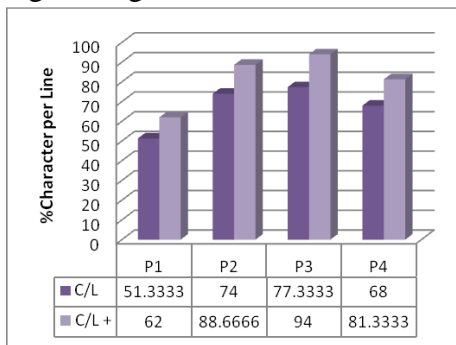


Fig 7: Multi Bar Chart

Another alternative of horizontal bar chart is vertical bar chart, which is use in the same way as horizontal bar chart. Floating column chart or High Low Open Close chart is another type of bar chart. In floating bar chart represent the low bound and high bound of the bar unlike simple bar chart where there is only high limit of the bar, as shown in figure 8. Candlestick Chart is another type of bar chart, in which top and bottom vertical line represents the low and high values, while the filled box represent the range of opening and closing values [30], as shown in figure 9.



Fig 8: Floating Bar Chart

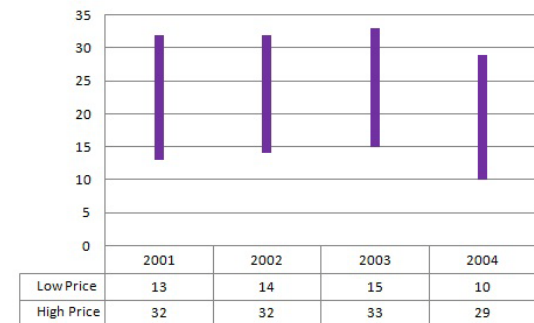


Fig 9: High Low Open Close Bar Chart

Different colors of the bar make it more visible and understandable. The need is to introduce some standard interactive mechanism for such type of visualization.

3.1.4 Histogram

Histogram is common and vital technique use in statistics and data analysis, was first introduce by Karl Pearson, and is a graphical representation which represents the distribution of data. It is use for the distribution of continuous data [34]. A histogram is the fall of specific items in class of variable, which consist of consecutive, non overlapping horizontal slabs. These slabs are mostly of the same size and adjacent, these may be of different size as well [35]. As shown in the following figures.

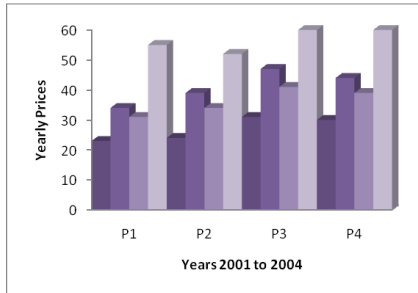


Fig 10: Histogram

3.1.5 Line Chart

Line chart is common well known graph in many fields, also label as line graph. It is a graph which is use to display information in connected points. These points are connected through continuous or straight line. Line graph is the extension of Scatter plot. Data points can be represented by icons or symbols, or can also draw simple line without icons [30] [36]. The line chart is often used to visualize a trend in data over time interval, means it use to show tendency in the data set and illustrate the data behavior with the passage of time or over a specific interval of time [37].

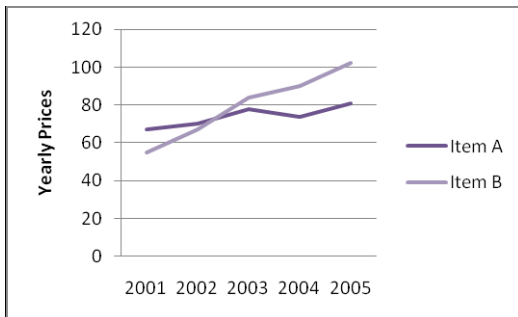


Fig 11: Simple Line Chart

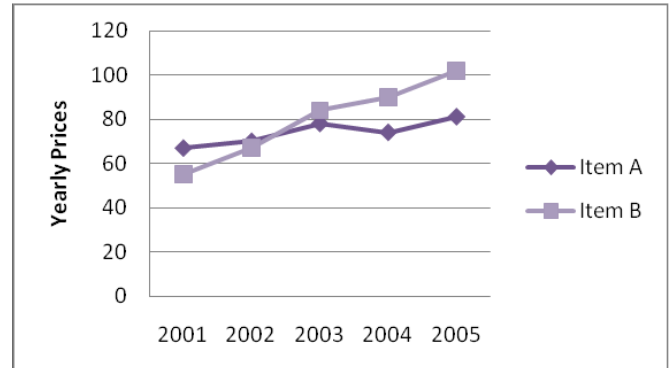


Fig 12: Line Chart with symbolic Data Point

There are many form or Variations of line chart or line graph, depends on the data points to be plot, for example; Step line chart, Reverse step line chart, vertical segment line chart, Horizontal segment line chart, Curve line chart [30].

3.1.6 Area Chart

Area chart is also called area graph, use to display quantitative data graphically. Area chart control is use represent data in bounded area. The bounded area is based on the line graph, the line is generated and the area below is shaded with colors, different texture and hatching, which produce area graph [30] as shown in the following figures 13 and 14.

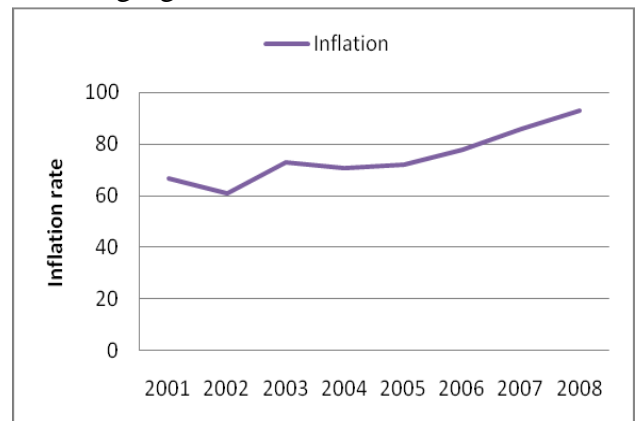


Fig 13: Line Graph

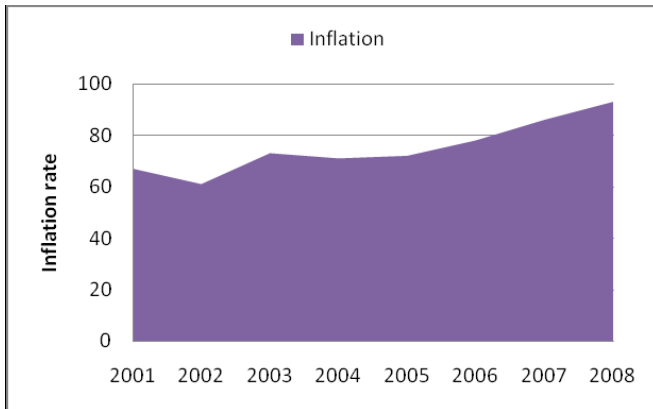


Fig 14: Area Chart

There are many form of Area chart or line graph, depends on the data points to be plot, for example; Step area chart, Curve area chart while segmented area chart is default area chart as shown in the following figure.

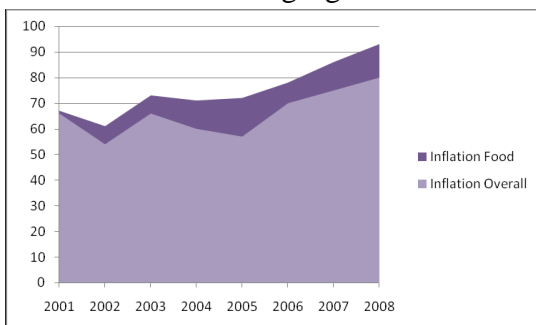


Fig 15: Multi Series Area Chart

3.1.7 Scatter Plot

Scatter Plot is also known as plot, plot chart, scatter chart, scattergram, scatter diagram or scatter graph. Scatter plot is graphical display of set of data in Cartesian coordinate, shows the relationship between two variables, one variable represent horizontal distance (independent variable) and second variable vertical distance (dependent variable) of data point from the coordinate axis [40]. Scatter plot shows the how strong the relationship are between the variables, and determines whether their exit any outlier in the data or not. It is use to look how the data is

dispersed [41]. Example of plot is shown in the following figure.

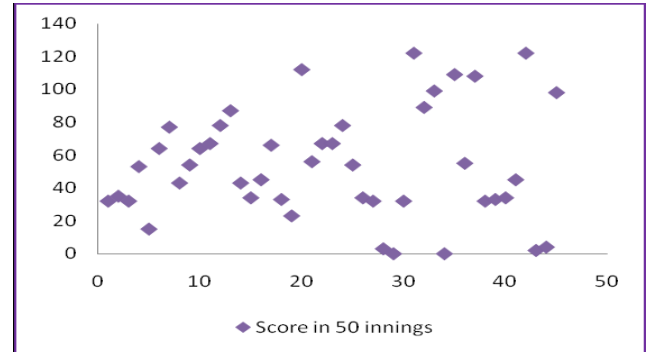


Fig 16: Scatter Plot

Scatter plot is useful to determine trend in the data and identify outliers easily. Outlier is an observation or data item which behaves differently from other members of the data set.

3.1.8 Bubble Chart

Plotted bubbles in bubble chart are defined by three different numeric parameters. These three numeric values represent data point i.e. one value determines its position along x-axis, one along y-axis, and third value represent the size of the bubble in the chart [30]. In bubble chart a bubble is differentiated from other bubbles in term of its size and in term of its position. As shown in the following figures 18 and 19.

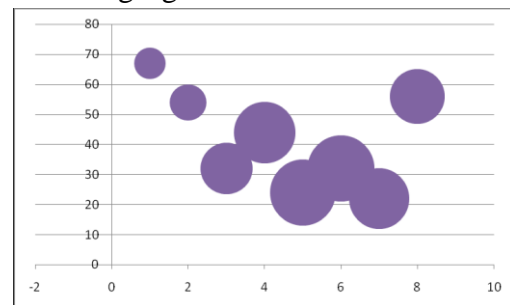


Fig 17: Single Series Bubble Chart

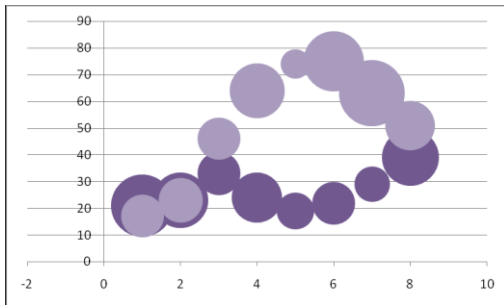


Fig 18: Multi Series Bubble Chart

A bubble chart is actually the variation of Scatter Plot, in which a data point is represented with a dot and in bubble chart dot is replaced with a bubble. It can be done if we have set of data points which contain three values for each data item [38]. According to Berman bubble graph can be useful in project management for the comparison of risk and success, where three values are net present value, probability of success and overall cost represent the size of the bubble [39].

3.1.9 Multiple Data Series

The combination of more than one charts can be use together to achieve some big purpose. Different data visualization techniques can be embedded into one chart control. A mix of column or bar chart control with line chart control over it shows trend of data series, as shown in the figure bellow [30]. Similarly different graph can be combining to create one graph. For example one can consider any combination from Area Series, Bar Series, Bubble Series, Candlestick Series, Line Series, and Plot Series etc [30].

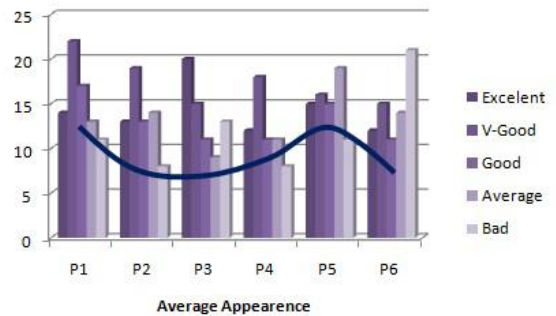


Fig 19: Multiple Data Series or Combination of Charts

3.2 Information Visualization

Information visualization is a research domain that concentrates on the use of visualization methods to assist people understanding data and evaluate or analyze data. Information visualization is the transmission of abstract data through the use of interactive visual interfaces. It enforces human visual cognition, to enable the user to get knowledge about the internal structure of the data, causal relationship, and dependences in it [Keim et al., 2006]. Information visualization is emerge field from several fields like Human Computer Interaction, Graphics, Psychology, and Computer Sciences etc. Its involvement is highly appreciated in scientific research, data mining, financial data analysis, product control studies, market analysis, and manufacture control studies etc [23]. According to Tomas et al information visualization is the “Visual representations and interaction techniques take advantage of the human eye’s broad bandwidth pathway into the mind to allow users to see, explore, and understand large amounts of information at once. Information visualization focused on the creation of approaches for conveying abstract information in intuitive ways” [23]. According to Robert et al, “Information Visualization deals with data that is

usually abstract, high-dimensional, and structured in a complex way” [53]. In the following section we going to discuss few known conventional data visualization techniques and give its example in order to get initial idea about the technique.

3.2.1 Parallel Coordinates

Parallel coordinates is a visualization technique used to plot individual data elements across many dimensions. Each dimension related to a vertical axis and each data object is displayed as a series

of connected points along the dimensions/axes [42]. Most standard visualization methods like line chart, bubble chart etc, fail to visualize multidimensional data. Parallel coordinate is very crucial when to display multidimensional data. Thousands of dimensions can be elaborate by this technique, parallel coordinates and few examples are discussed by Stephen Few [43]. As shown in the following figure 20.

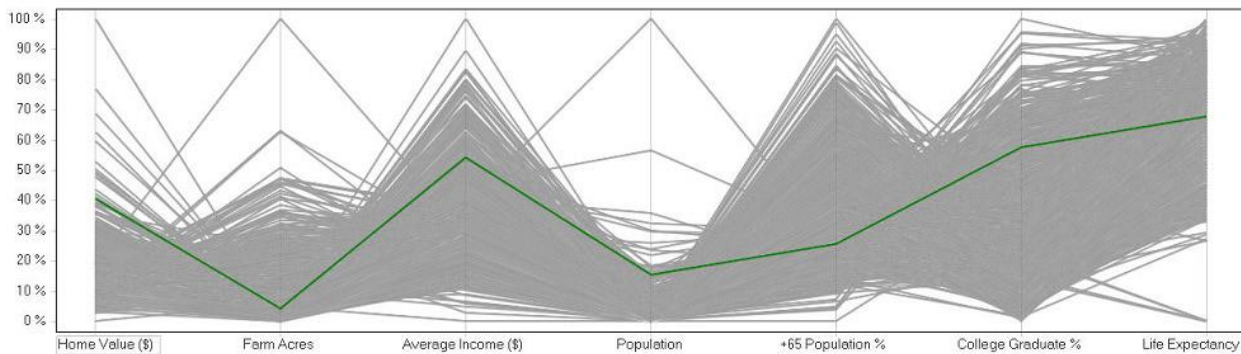


Fig 20: Parallel Coordinates

Alfred Inselberg wrote a comprehensive book about parallel coordinates, he mention the geometry of multidimensional visual data and its application in different fields, specially how parallel coordinates help one to recognize pattern in multiple set of huge data [44].

3.2.2 Tree Map

Tree map is also known as tree mapping. In information visualization, tree mapping is a technique is use to display hierarchical data in the form of nested or layered rectangles [45]. Tree map is use to visualize hierarchical structures. Tree map facilitate users to compare nodes and sub nodes at various depth and help to recognize patterns and expected results. Many data sets has hierarchical characteristics, the objects are

categorize into different categories, subcategories and so on. The following figure show tree map.

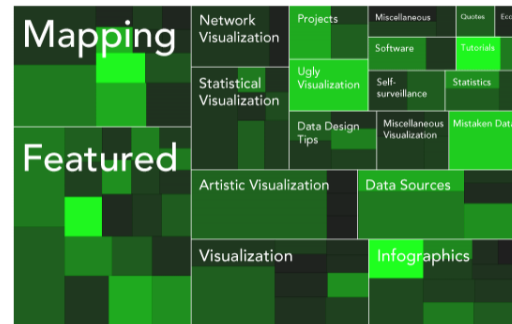


Fig 21: Tree Map

Each data object or item in data set is represented by a rectangle, which have different sizes recording user defined attributes or user defined properties. Color of each rectangle reflects the intensity of the defined attribute [46], as shown in the above figures.

3.2.3 Entity Relationship Diagram

Entity relationship diagram is constructed through Entity Relationship Modeling mechanism. Entity relationship modeling is an abstract and conceptual data representation technique use in software engineering. It is basically a database modeling methodology which is used to generate conceptual schema, semantic schema, and most of the time it is use to create relational database schema for a system. ER diagrams or ERDs are use as short form for entity relationship diagrams. Most of the time top-down approach is use to design a ER diagrams. Key elements of ER diagram are Entity; may be an object or concept about which information is need to be store, Relationship; means how two or more than two entities share information, and Attributes; is any unique characteristic of an entity [47]. As shown in the following example.

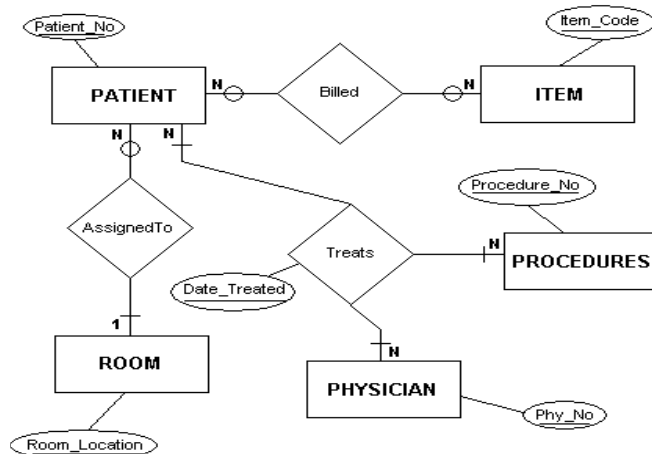


Fig 22: Hospital

Many other ERD's elements exits like strong entity, weak entity, key attribute, multivalued attribute, derived attribute, relationship and cardinality, all these elements are represented with different symbols or notations [48].

3.2.4 Cone Tree

Cone tree is another technique to display hierarchical data such as organizational body etc in three dimensions. The branches grow in the form of cone. Three dimensional cone trees supports more nodes as compare to two dimensional cone trees [50]. According to Andy & Bruce cone tree is an attractive interactive three dimensional visualization technique use to represent hierarchical data effectively utilize the available screen and human perceptual system. They describe the design, implementation, and evaluation of animated 3D cone tree. When a level is added in 3D cone tree system an inverted cone is created at bottom side. They further evaluate the cone system and provide the usability study of the tree interfaces [49].

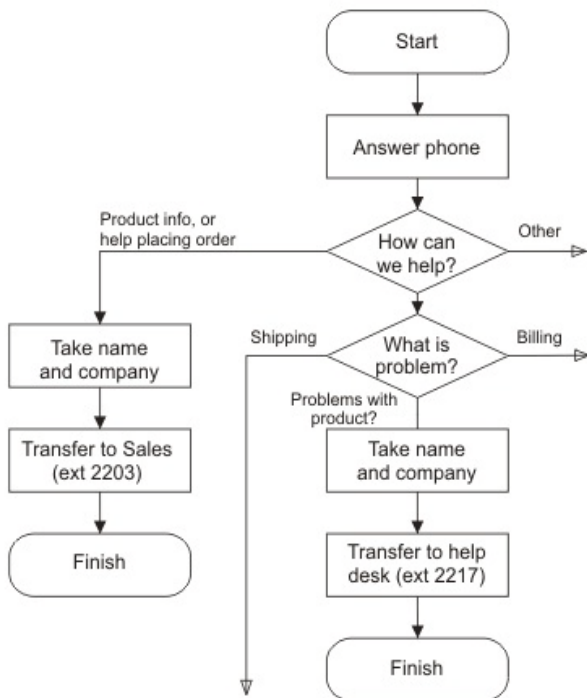
3.2.5 Time Line

Timeline is also termed sometime as chronology, because a timeline is a graphical representation of chronological sequence of events draw along straight line to enable users to understand the relationship between different events easily. The timeline may present in tabular, yearly paragraph or some other forms [51]. Simply timeline is an image or pictorial representation that shows the list of events that happened in the past or events that are the part of history. It may be of two types, i.e. linear timeline and comparative timeline. Linear timeline is an order of events that occurred in specific period of time, it can be represented as horizontally or vertically. Comparative timeline is two sets of events that occurred at a place or a set of events that occurred at two different places that may or may not be compared but can deduce for some sort of knowledge [52].

Figure 23 shows example of timeline.

3.2.5 Flow Chart

A flow chart or flow graph is graphical representation or symbolic representation of steps involve in a process. As there are different steps in a process represent different symbols. The flow chart is use to show the flow sequence of data, flow of direction and control flow in a process using arrows. Different symbols that has been use in the flow chart are, process, alternative process, delay, preparation, arrows, start, terminator, decision, and connector etc [54] [55]. The following figure is an example of flow chart.



Flow Chart Fig 24

Technology Advancement Time Line

3.2.6 Data Flow Diagram

Data flow diagram is shortening by DFD. DFD is important to find out how the data is process in a system, it is use to describe how the data or information is transform by the system. In other words DFD shows how the data is processed, how it is stored, and discover how the information flows through the processes. DFD is

a graphical representation of data flow through the system, and model its process aspects. It is use to visualize data processing, having different level. DFD of level 0 create overview of the overall system and DFD of level 1 elaborate each process of the system and provide detail view [55] [56], level 2 elaborate it further. DFD example is given bellow.

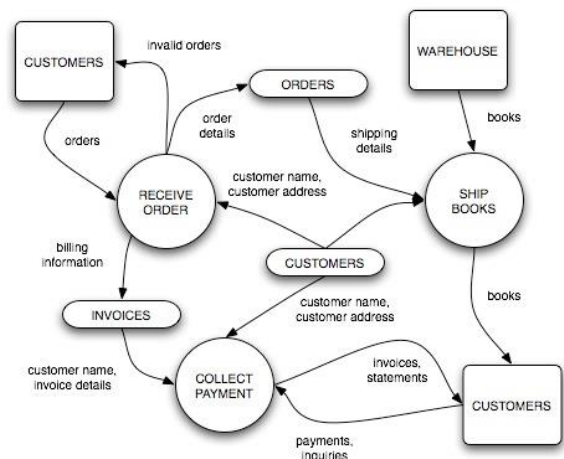


Fig 25: Data Flow Diagram

DFD illustrate what kind of input is required by the system and what type of output is generated by the system. What is the input of each module and where the data come from, where it is stored and for whom it will be the input. The DFD does not show any thing about the operational sequence, for this purpose flow chart has been used. Different notations are used for DFD, i.e. for process, input, output, file or database and flow etc.

3.2.7 Venn Diagram

Venn diagrams are introduced by John Venn in 1880. Venn diagrams are use to explain the relationship between two or many sets, that's why it is also called set diagram. It specified the relationship between objects of different sets. This means that to determine the common

characteristics in objects, represented by upside-down letter U, called intersection and determined which characteristics or not common, represented by X', called complement, total collection of items or objects, represented by letter U, called union in set theory [51]. Here are few examples in figure 26, one of them shows the interaction of Greek, Latin, and Russian alphabet.

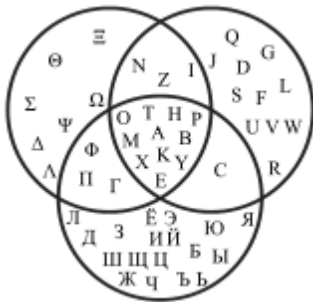


Fig 26: Examples of Venn Diagram

All comparative sets are the members of big set called universal set. Universal set is the set of objects or items under consideration. As Venn diagrams shows all logically possible relationships between two or more sets, so it is very helpful in many discipline. For examples this sort of diagram is used in scientific and engineering presentation, theoretical mathematics, probability, logic, statistics, linguistics and computer science and its application etc. Ruskey and Watson introduce a detail survey of Venn diagram [57].

3.2.8 Semantic Network

A semantic network is a graphical representation of logical relationship between different concepts. Semantic network generate directed graph, the combination of nodes or vertices, edges or arcs, and label over each edge. Nodes represent the individual or any concept and the relationship is represented by edge or arc, which is denoted by the label on the arc. More than one edge contains

the same label [58]. Semantic network is based on human cognition, and used to convey ideas about set of objects or concepts. Semantic network is knowledge representation schemes that involve nodes and links, produce directed graph containing nodes, links (arrows) and label on the links, which shows the relation between concepts or objects [59]. Semantic networks understandable knowledge of set of concepts that are related to each other. As shown by the following figure number 27. There are many advanced information visualization techniques are mentioned, e.g. Tablelens technique, TableLens, Node-link diagram etc.

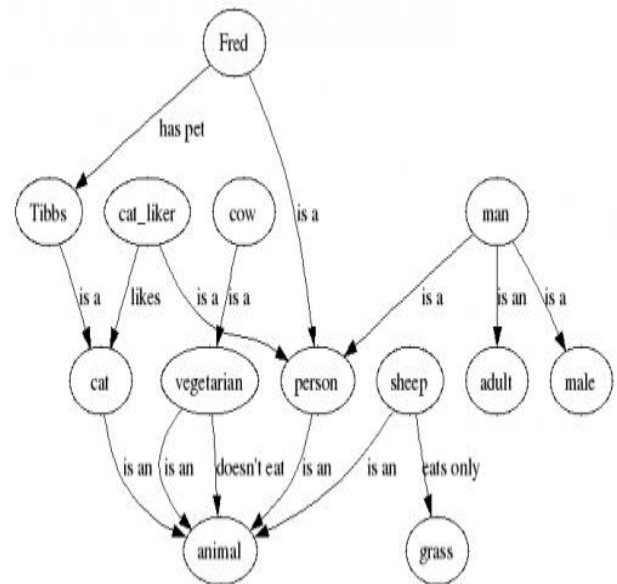


Fig 27: Semantic Network

4. INTERACTIVITY The users are interested in the abstract data about which they have desire to understand, the user don't have sufficient pre knowledge about that data. Hence for the exploration, analysis, and for the representation of data or information visualization interactive techniques are exceptionally momentous. The challenge in information visualization is to

provide data visually in order to the user effectively understand the information for which the user is looking for, for this purpose provide interaction mechanism that make is possible to manipulate visualization effectively and effortlessly as probable [53]. Users can interact with interfaces or visualization in different ways by means of mouse over, single click, double click, or can add multiple interactive options by mouse right button click. There are many interactive techniques available to interact with charts or graphical representation to understand the drill down details. Card et al introduce the interactive mechanism of visualization in 1999. This mechanism describes human interaction in different phases as shown in the following diagram.

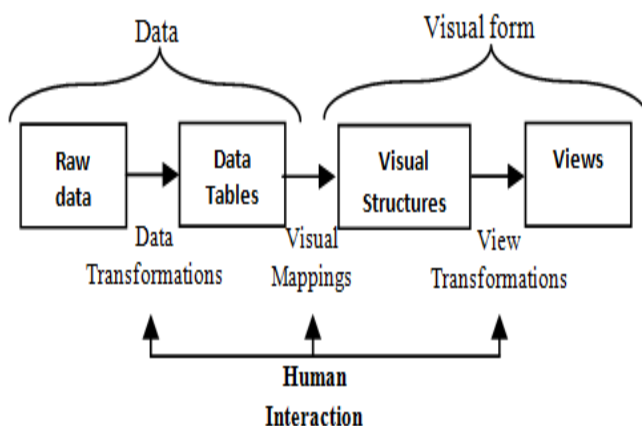


Fig 28: Interactive Mechanism

4.1 Zoom in and Zoom out or Zooming

Sometimes user interfaces needs to bring it closer in order to view of contents clear or in details. The content display of user interface must not be each user’s choice, so required fair scaling techniques which enable user to make it according to their wish. In computing, zooming user interface or zoomable user interface “ZUI” in graphical environment, enable user to change

the scale of viewed area of interfaces up to different level of details. When the contents required maximizing in size for detail view zoom in technique is use, and when adjust to smaller size done with zoom out technique. This technique is more useful for mobile device because the screen size is very limited as compare to the content to be display. The zoomable user interface provides more productive interactions. Different strategies have been introduced to produce broader view of information, and for navigation via it. These navigation supports broader view and detail view of the entire information.

4.2 Zoom + Pan

Zoomable visual interface starts with the overview and then enable the users to zoom in to achieve the level of detail up to some extent in information of interest. Zoomed visual interface can be pan across the rest of the interface without zooming out the interface. Alternatively the interface can zoom in, zoom out to return to the overview, and again zoom in to another portion of detail, and pan zoom in state and in zoom out state. Two advantages of this strategy are; it uses the screen efficiently and provides unlimited scalability. It has two very costly disadvantages i.e. when zoom in it loses overview and navigation is slower.

Steps for Interactive Visualization Strategies

Interaction strategies support further scalability and complexity of visual information. There are some major steps that should be considered for interaction strategy in every visualization design. These steps are briefly explain in the following section

1. Selecting

One of the most important and fundamental requirements in visualization is interactive selection of data entities or subset or part of whole data or whole data set, that is of interest to the user. This is useful to view detail information about the selected entities, highlight entities that are covered or hidden, cluster entities that are related and extracting entities that may be use in future. There are two probable conditions through which users can specify selections. First: users can select data entities directly; visualization should enable the users to select entities directly. For the selection of data entities variety of techniques has been use discuss by Wills [Wills, 1996], such as pointing at individual entities or pointing to group of entities, etc. Second: users can select data entities indirectly, for example, XmdvTool [Ward, 1994] enables users to make selections in tabular data in parallel coordinates by specifying range criteria on data attributes.

2. Linking

Linking and brushing are the most common form of linking. Linking is useful to relate information among multiple views. Information can be mapped differently to different views to reveal or expose different perspectives (viewpoint) or different portions of the information. On the bases of selection criteria the user can select entities in one structure, which then shows the distribution the selected entities in another structure. Fig 31: Interactive brushing and linking between histogram plots (top) and geographic map (bottom) of dataset.

3. Filtering

Filtering enables users to dynamically adjust amount of information to display, means to decrease or increase information quantity that need to display, and focus on information of interest. For this purpose need some dynamic query values to manipulate, in this regard visual widgets can play vital role e.g. slider can be use in different ranges, field box can be use to put attribute value in specified range etc. these widgets enables one from the current query parameters and enables to quickly adjust query parameters and instantaneously view filtered results in the visualization in real time.

4. Rearranging or Remapping

As a single mapping or plotting visual form of information may be not enough, so the users must empower to customize mapping among many maps. To enable users to customize maps by its own choice provide the way to better understand the information. As the spatial layout is the most important visual mapping, rearranging the spatial layout of the information is the most effective for producing different insights i.e. TableLens [1][30][66]

CONCLUSION AND SCOPE OF VISUALIZATION

The main idea behind this document is to initiate the basic of visualization for future generation, keeping in mind the significance of visualization. The business sector still mostly ignore, misunderstood, used in inefficiently data and information visualization, which leads to the overall less production. So the researcher needs to concentrate on visualization techniques in



order to get maximum benefits offer by visualization. Another important aspect is to clearly specify what data type needs to represent which visualization method, that interpret maximum understandings. As Data mining is one of the emerging field of computer science, largely use in the business sector produce results in the form of visual data mining, which required urgent attentions to produce interactive visualization that easily describe what insight. Because interactive and animative visualization is more beneficial to convey the conclusion of huge or very huge data set that is not easy to understand under traditional approaches. Another focus area is mobile devices visualization or visualization of mobile devices, where it is more important to identify visualization methods for it with keep in mind the limitations of mobile devices. Data mining techniques, mining results visualization and interactive mechanisms on mobile devices is also a significant field of research which is agenda of my present and future research work. The today world demands the standardization of data and information visualization techniques with respect to data to be present, interactive mechanisms with respect to visualization methods, mobile visualization with respect to different aspects of data and information visualization.

6. REFERENCES

[1]. Chris North, "Information Visualization", Center for Human-Computer Interaction, Department of Computer Science Virginia Polytechnic Institute and State University Blacksburg, VA 24061 USA

- [2]. Ware, C. (2004). "Information Visualization: Perception for Design", Morgan Kaufmann.
- [3]. Spence, R. (2001). "Information Visualization", AddisonWesley.
- [4]. Beatriz Sousa Santos, "Evaluating Visualization techniques and tools: what are the main issues?".
- [5]. North, C. (2005). "Toward Measuring Visualization Insight. IEEE Computer Graphics and Applications", vol. 11, no. 4, pp. 443-456
- [6]. S. Card, J. MacKinlay, and B. Shneiderman, (1998). "Readings in Information Visualization: Using Vision to Think". Morgan Kaufmann.
- [7]. Alfredo R. Teyseyre and Marcelo R. Campo, (2009). "An Overview of 3D Software Visualization", IEEE Transactions on Visualization and Computer Graphics, vol.15, No.1.
- [8]. E.R. Tufte, (1997). "Visual Explanations: Images and Quantities, Evidence and Narrative", Graphics Press, 1997.
- [9]. D.M. Butler, J.C. Almond, R.D. Bergeron, K.W. Brodlie, and A.B. Haber, (1993). "Visualization Reference Models", Proc. Fourth IEEE Conf. Visualization, G.M. Nielson and D. Bergeron, eds., pp. 337-342.
- [10]. WebSter Dictionary: www.websterdictionary.org/definition/visualization
- [11]. L. Chittaro, (2006). "Visualizing Information on Mobile Devices", ACM Computer, v.39 n.3, p.40-45.
- [12].Plaisant C, (2004). "The Challenge of Information Visualization Evaluation", Proceedings of AVI 2004: 6th International



Conference on Advanced Visual Interfaces, ACM Press, New York, 2004, pp. 109-116

[13]. Chittaro L. (2004). "Special Issue on HCI Aspects of Mobile Devices and Services", Personal and Ubiquitous Computing Journal, Vol. 8, No.2.

[14]. A Periodic Table of Visualization Methods.

http://www.visualliteracy.org/periodic_table/periodic_table.html

[15]. Chaomei Chen, (2005). "Top 10 Unsolved Information Visualization Problems", IEEE Computer Graphics and Applications, 2005.

[16]. D. Harel and Y. Koren, (2000). "A Fast Multiscale Method for Drawing Large Graphs", Proc. 8th Int'l Symp. Graph Drawing, 2000, pp. 183-196

[17]. P.C. Wong et al. (2003). "Dynamic Visualization of Transient Data Streams", Proc. IEEE Symp. Information Visualization, IEEE CS Press, 2003, pp. 97-104.

[18]. Webster Dictionary: www.websterdictionary.org/definition/aesthetic

[19]. C. Chen, (2004). "Information Visualization: Beyond the Horizon", 2nd ed., Springer, 2004

[20]. C. Johnson, (2004). "Top Scientific Visualization Research Problems", IEEE Computer Graphics and Applications, vol. 24, no. 4, pp. 13-17, July/Aug. 2004.

[21]. Michael Friendly, (2008). "Milestones in the history of thematic cartography, statistical graphics, and data visualization", <http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf>.

[22]. Smashing Magazine: www.smashingmagazine.com/2007/08/02/datavisualization-modern-approaches/comment-page1/#comment-559828

[23]. Bederson, B. B., & Shneiderman, B. (2003). "The craft of information visualization: readings and reflections", San Francisco, CA: Morgan Kaufmann Publishers

[24]. J. J. Thomas and K. A. Cook, (2006). "A Visual Analytics Agenda", IEEE Transactions on Computer Graphics and Applications, 26(1):12-19, January/February 2006

[25]. Oliveira, M.C., Levkowitz, H. (2003). "Visual Data Exploration to Visual Data Mining: A Survey", IEEE Transaction on Visualization and Computer Graphs 9(3), 378-394 (2003)

[26]. A Fink, (2008). "How to conduct surveys: A step-by-step guide", Sage Publications, Inc.

[27]. Morgan, George (2004). "SPSS for Introductory Statistics". Hillsdale: Lawrence Erlbaum

[28]. Muzammil Khan, Mushtaq Raza, and Nasir Rashid (2010). "Appropriate Length of Text Line with Special Relationship to Eye Blink to Reduce Maximum Focus Loss". International Conference on Internet Computing 2010: 97-102

[29]. P.E. Hoffman, (1999). "Table Visualizations: A Formal Model and Its Applications", doctoral dissertation, Computer Science Dept., Univ. of Massachusetts at Lowell, 1999.

[30]. Adobe Flex, Advanced Data Visualization Developer Guide

[31]. Clark Jeff, (2006). Multi-level Pie Charts.



- [32].Stephen Few, (2007). "Save the Pies for Dessert", Perceptual Edge Visual Business Intelligence Newsletter August 2007.
- [33].Spence, I. (2005). "The origins and usage of a statistical chart", Journal of Educational and Behavioral Statistics, 30, 353-368
- [34].Pearson, K. (1895). "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 186: 343-326
- [35] Howitt, D. and Cramer, D. (2008). "Introduction to Statistics in Psychology", Prentice Hall.
- [36].Burton G . Andreas (1965). "Experimental psychology" . page.186
- [37].Neil J. Salkind (2006). "Statistics for People who (think they) Hate Statistics: The Excel Edition". Page 106.
- [38].Online Microsoft Excel Help: <http://office.microsoft.com/en-us/excel-help/creating-abubble-chart-HA001117076.aspx>
- [39].Jeff Berman (2007). "Maximizing project value: defining, managing, and measuring for optimal return", AMACOM Div American Mgmt Assn, 2007. Page .63-64.
- [40]. Utts, Jessica M. (2005). "Seeing Through Statistics", 3rd Edition, Thomson Brooks/Cole, 2005, pp 166-167
- [41].What is Scatter plot? www.psychwiki.com/wiki/What_is_a_scatter_plot%3F
- [42]. Zach Gemignani, (2010). "Better Know a Visualization: Parallel Coordinates",
- [43].Stephen Few, (2006). "Multivariate Analysis Using Parallel Coordinates", September 12, 2006.
- [44]. Alfred Inselberg, (2009). "Parallel Coordinates: Visual Multidimensional Geometry and Its Applications", Springer, 2009.
- [45].Shneiderman, Ben; Plaisant, Catherine (2009). "Treemaps for space-constrained visualization of hierarchies".
- [46]. IBM Research and the IBM Cognos software group, www958.ibm.com/software/data/cognos/maneyes/page/Treema.p.html
- [47].Paul Beynon-Davies (2004). "Database Systems". Houndmills, Basingstoke, UK: Palgrave
- [48].Jeffrey A.Hoffer, Mary B. Prescott, and Fred R. McFadden. Modern Database Management, 7th Edition.
- [49].Andy Cockburn and Bruce McKenzie, (2000). "An evaluation of cone trees". In People and Computers XIV: British Computer Society Conference on Human Computer Interaction, pages 425-436. Springer Verlag, 2000
- [50].Usability First: www.usabilityfirst.com/glossary/cone-tree/
- [51].Search CIO Mid market: <http://searchciomidmarket.techtarget.com/definition/timeline>, <http://searchcio-midmarket.techtarget.com/definition/Venndiagram>
- [52].Student Guide to Timeline: www.lkwdpl.org/study/timeline/
- [53].Robert Kosara, Helwig Hauser, and Donna L. Gresh, (2003). "An Interaction View on Information Visualization". Published in 2003.



- [54]. Flow Breeze: Flowcharting and Process Diagramming Tools, www.breezetre.com/articles/what-is-a-flowchart.htm
- [55]. Roger S. Pressman. Software Engineering, A Practitioner's Approach. Sixth Edition, Page 348, 227, 309.
- [56]. Bruza, P. D., Van der Weide, Th. P. (1993), "The Semantics of Data Flow Diagrams", University of Nijmegen, 1993.
- [57]. F. Ruskey and M. Weston (2005). "Venn Diagram Survey", The electronic journal of combinatorics, June 2005.
- [58]. Amaryllis Deliyanni, Robert A. Kowalski. (1977). Logic and Semantic Networks. Proceedings of the Workshop On Logic And Data Bases, Toulouse, November 1977.
- [59]. Generating of Marketing insight, Semantic networks.
- [60]. Becker, R. A., Eick, S. G., and Wilks, A. R. (1995). "Visualizing network data", IEEE Transactions on Visualization and Graphics, 1(1):16-28.
- [61]. Stoakley, R., Conway, M. J., and Pausch, R. (1995). "Virtual reality on a WIM: Interactive worlds in miniature", Proceedings of the Conference on Human Factors in Computing Systems 1995, pp. 265--272
- [62]. Bederson, B.B., Hollan, J., Perlin, K., Meyer, J., Bacon, D., and Furnas, G. (1996). Pad++: "A Zoomable Graphical Sketchpad for Exploring Alternate Interface Physics". Journal of Visual Languages and Computing, 7(1):3-32
- [63]. Leung, Y. K., and Apperley, M. D. (1994). "A review and taxonomy of distortion-oriented presentation techniques". ACM Transactions on Computer-Human Interaction, 1(2):126—160
- [64]. Robbins D.C., Cutrell E., Sarin R., Horvitz E (2004). "ZoneZoom: Map Navigation for Smart phones with Recursive View Segmentation", Proceedings of AVI 2004: 6th International Conference on Advanced Visual Interfaces, ACM Press, New York, 2004, pp. 231-234
- [65]. Gerald J. Kowalski and Mark A. Maybury. Information Storage and Retrieval System, Theory and Implementation. Second Edition.
- [66]. O. Kulyk, R. Kosara, J. Urquiza, and I. Wassink, (2007). "Human-Centered Aspects," Human-Centered Visualization Environments, Gi-Dagstuhl Research Seminar, A. Kerren, A. Ebert, and J. Meyer, eds., chapter 2, pp. 10-75, 2007". 7.