

Project 2 Report:

Creating a methodological tool to aid learning of basics of data visualization.

Parth Kapadia
Interaction Design
M.Des (2017-19)

Guide: **Prof. Venkatesh Rajamanickam**

IDC School of Design
Indian Institute of Technology, Bombay.

Declaration

I declare that this written document represents my ideas in my own words and where other' ideas or words have been included, I have adequately cited and referenced the original sources. I also declared that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Parth Kapadia

176330011

Industrial Design Centre

Indian Institute of Technology, Bombay.

Approval sheet

The Interaction Design Project II entitled “ *Creating a methodological tool to aid learning of basics of data visualization*” by Parth Kapadia, roll number 176330011 is approved, in partial fulfilment of the Masters in Design Degree in Interaction Design at the Industrial Design Centre, Indian Institute of Technology, Bombay.

Internal -

Extrenal -

Guide -

Chairperson -

Table of contents

1. Abstract	1
2. Introduction	2
3. Initial Inspirations	3
3.1 Drawing dynamic visualizations by Bret Victor	3
3.2 Interactive Exploration of a Dynamical System by Bret Victor	3
3.3 3.3 Playful interaction by Nicky Case	4
3.4 Linkki - Planar linkage based kinetic toy	4
4. Secondary Research	5
4.1 How visualization courses have changed over past 10 years	5
4.2 Design activity framework for Visualization design	5
4.3 Worksheets for guiding novices through the visualization design process	6
4.4 Sketching designs using the Five Design-Sheet methodology	6
4.5 Lecture slides by Enrico Bertini	6
4.6 Guide to information Graphics	6
5. Primary Research	7
5.1 Visualization Analysis and Design by Tamara Munzner	7
5.1.1 Why this book?	7
5.1.2 About this book	8
5.2 Ethnographic study	8
6. Key Learnings	10

Table of contents continued...

7. Design consideration space	11
8. Design explorations	12
8.1 Magnetic board viz	12
8.2 Suggestive documentation tool	12
8.3 Vizphernalia	13
8.4 Method cards	14
9. Detailed design explorations	15
9.1 Study on cards in data visualization domain	15
9.2 Study on few other miscellaneous deck of cards	18
9.3 Way we interact with cards	18
9.4 Detailed description of VizSuit	19
9.5 When and how to use the cards	22
10. Workshop protocol	23
11. Evaluation & conclusion	24
12. Bibliography	25
13. Appendix I	26
13. Appendix II	29

1. Abstract

Data visualization is a general term used to describe any attempt made to help people understand the significance of data by placing it in visual context. Creating an effective visualization is a challenging task which requires the designer of the visualization to take mindful decisions at various phases of the process of creation.

Through this project I attempt to make a methodological tool to help students understand the basics of data visualization. The tool developed is in the form of a deck of cards (VizSuit) designed to guide students through the linear - multi iterative process of creation of any data visualization.

2. Introduction

Data visualization is being considered as a relatively modern field however it has very deep roots dating back to the earliest map making and visual depiction in the caves. Although along this entire period there has been constant developments in technologies, mathematical theories, empirical observation and recording of data has enabled wider use of graphics in different form and content. This has led to profound changes in this field of data visualization.[1]
[2]

There has been remarkable advancements in visualization algorithms, techniques, methodologies and applications. Literature regarding data visualization is available but is vast, textual and very much in static form. This content is very difficult for the students to consume as it is spread out across several seminal papers and books[3]. Predominantly the field of data visualization had been a part of computer science engineering however due to recent increased use of visualization in industries and business a large number of non-technical students have started to show interest in viz courses [2]. Owing to its predominance in a technical field most of the content present so far caters to students from a technical background and uses a lot of technical jargons. The field being too vast as well as students enrolling with different backgrounds there is considerable amount of challenge that exists for tutors to decide upon the depth and spread of topics which can be encompassed in the course of data visualization.

3. Initial inspirations

This section documents relevant work in the two directions in which I envisioned my tool to be like at beginning of my project timeline. Firstly, an interactive digital tool. Interactive here refers to a way in which a user can directly manipulate the variables of the subject of study with some interaction with the computer. Second direction was more towards tangibility to see if user can manipulate objects in the physical world itself to learn the principles of data visualization.

3.1 Drawing dynamic visualizations by Bret Victor

Bret Victor has developed a tool to create dynamic pictures that are data driven like D3 however it also facilitates direct manipulation of this picture by interacting with it like Illustrator.

Throughout the history of science, diagrams and graphs have been essential thinking tools. In the past, such visualizations were drawn with pen on paper, and could embrace the directness, freedom, and expressiveness of hand drawing. Most modern visualizations are programmed instead, where a single description can dynamically generate a unique picture for any dataset.

Today's tools offer the benefits of one or the other -- either directness or dynamics -- but not both. Photoshop and Illustrator allow direct-manipulation drawing of static pictures. D3, R, and Processing allow indirect-manipulation coding of dynamic pictures.

Bret Victor. (2013, February 1). Drawing dynamic visualizations [Video file]. Retrieved from <https://vimeo.com/66085662>

3.2 Interactive Exploration of a Dynamical System by Bret Victor

A user interface for exploring systems of differential equations. Every variable is shown as a plot; every parameter has a knob that can be adjusted in real time. This ubiquitous visualization and in-context-manipulation helps the user develop a sense for how the parameters of the system influence its behavior.

Bret Victor. (2011, May 16). Interactive exploration of a dynamical system [Video file]. Retrieved from <https://vimeo.com/23839605>

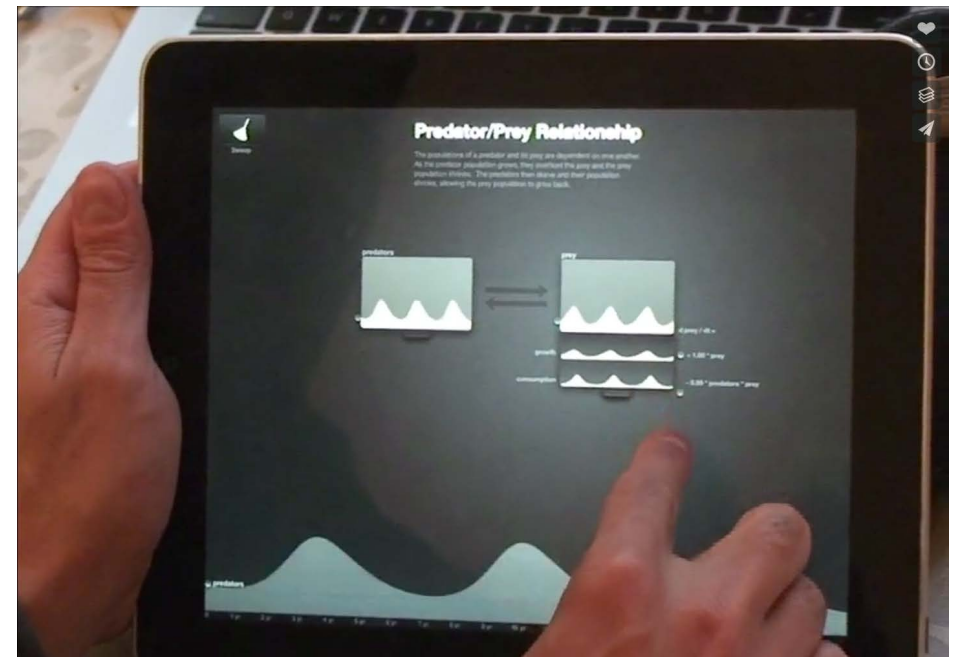


Fig 1. Screenshot from the video Interactive exploration of a dynamical system

3.3 Playful interaction by Nicky Case

Nicky case tells stories about systems. He takes the systems of the world around us - social, physical, political, environmental, economic and explains them through a system of games. He does not exactly gamify things however he gives the flexibility of decision making at various points to the user which indulges the user in the story and makes the user more involved with the interactions which are facilitated. He makes the complex simple through the power of interactivity and play.

Ref - <https://ncase.me/>

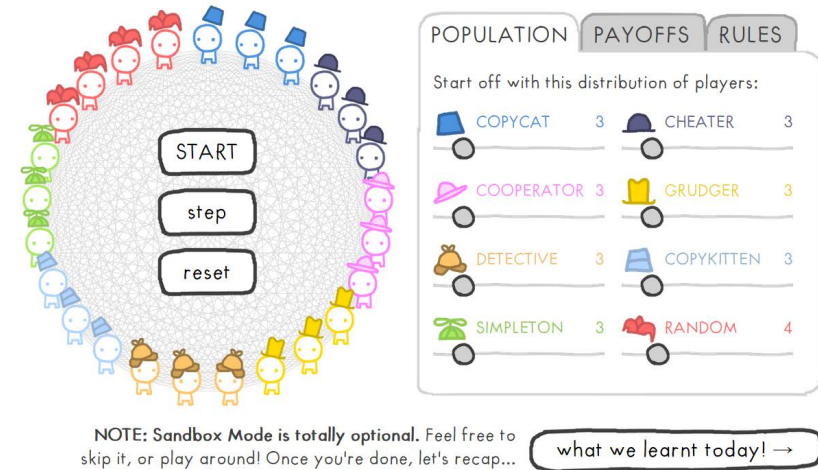


Fig 2. Screenshot from explorable on N.Case's website

3.4 Linkki - Planar linkage based kinetic toy

It is a kinetic construction toy based on a planar linkage mechanism with which users can design movements, make kinetic arts, and learn basic STEM subjects through hands-on play. Consisting of a modular box, bars and circles, and active/interactive modules, LINKKI simplifies the technical construction toy to such an extent as to look minimal but still retains the versatility of construction blocks as an educational tool and designer's prototyping tool. The simplicity of LINKKI not only allows intuitive play, but also enables users to easily tweak and create custom parts for themselves, thus making it easy to incorporate raw material in addition to the existing parts for prototyping use. [4]

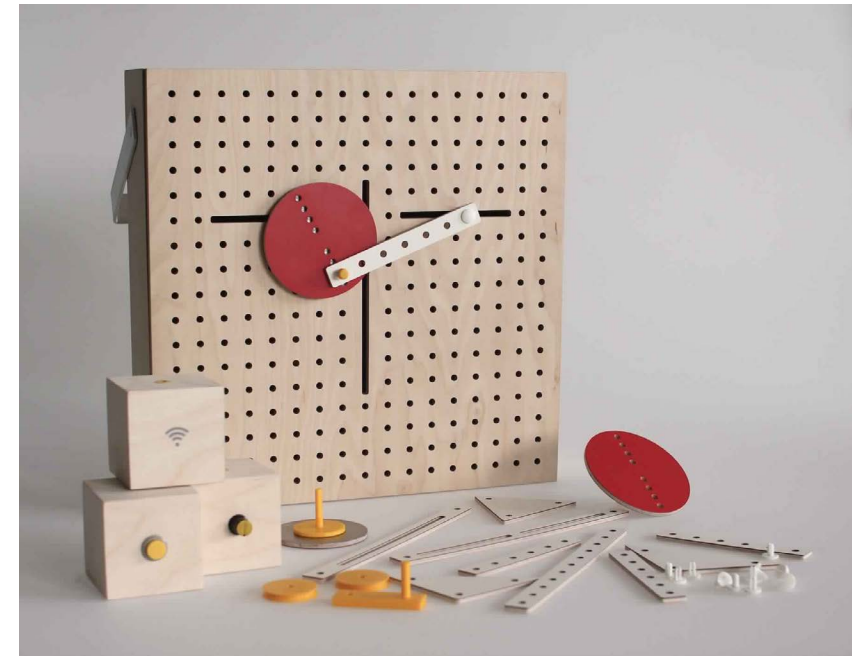


Fig 3. Linkki prototype

4. Secondary research

4.1 How visualization courses have changed over past 10 years

This article mainly examines and documents the insights gained at the panel, called “The Changes We Have Made to our Visualization Courses over the Last 10 Years,” Holly Rushmeier, Jörn Kohlhammer, David Ebert, Beatriz Sousa Santos, and Daniel Weiskopf discussed how they’ve changed their courses to reflect current problems and practical solutions. Each of these panelist have many years’ of experience teaching courses covering topics such as scientific visualization, data visualization, information visualization, visualization techniques, and visual analytics.[2]

4.2 Design activity framework for Visualization design

This paper talks about the importance of the connection between what a designer does and the decisions the designer makes while creating visualizations. So broadly this paper sets up the link between the process and the design decisions at various stages of creation of visualization. The proposed design activity framework enables the process model to explicitly link with the nested model (renowned visualization design decision model) by T.Munzner. The framework includes four overlapping activities that characterize the design process, with each activity mapping outcomes related to the nested model.[5]

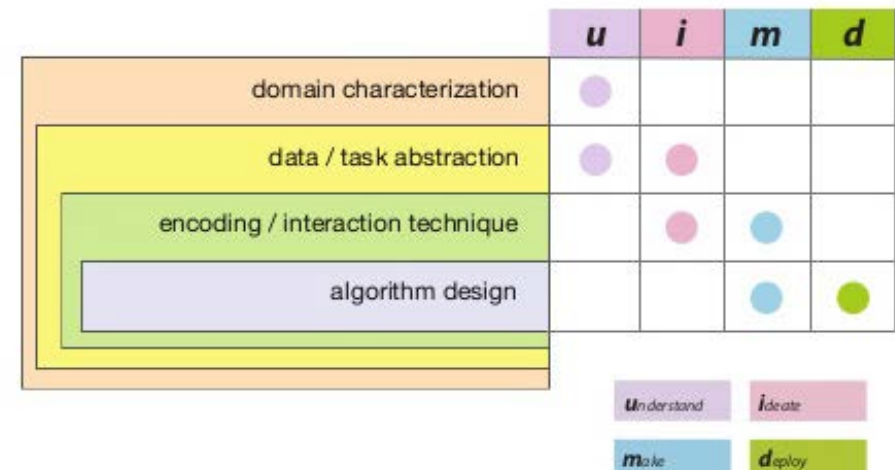


Fig 4. Design activity framework

4.3 Worksheets for guiding novices through the visualization design process

While the design activity framework was a broader look at the entire process of visualization creation this paper talks about the finer nuances of teaching, creating and evaluating a data visualization. In this paper a more detailed structure is given to the design activity framework so as to guide novices through their own real world design process of creation of visualization. To provide students with such concrete guidelines, they created worksheets for each design activity: understand, ideate, make, and deploy. Each worksheet presents a high-level summary of the activity with actionable, guided steps for a novice designer to follow. [6]

4.4 Sketching designs using the Five Design-Sheet methodology

This paper talks about sketching as being a tool of quick low fidelity prototyping and an effective way of exploring alternative solutions at no cost at all. It explores how this informal way can enable people to think divergently at first and then finally converge onto a thoughtful decision of final visualization choices. Sketching serves as an effective way of managing various thoughts and ideas by getting them out on paper and quickly re iterating its better forms. Sketching on five sheets helps user to think through their idea of solution and consider their potential effectiveness.[7]

4.5 Lecture slides by Enrico Bertini

Enrico Bertini is a professor at NYU Tandon School of Engineering. Going through his slides from Information Visualization course (Spring 2016) gave me a rough overview of the course which takes place in this college of NYU. It was insightful in a way to get a general idea about how an information visualization course could possibly vary when taught to students from technical background. [8]

4.6 Guide to information Graphics

This book rightly points out at how rarely making infographics is a point of teaching in schools or job training but how repeatedly it is used. It serves as a wonderful compilation of helping us learn ways to express ourselves through language of graphics effectively and eloquently. It describes the practices which works really well and also points out the common pitfalls which people which encounter.

5. Primary research

Being new to this domain I needed to get a better hold of this field of data visualization. My research was mainly two folds, one was through detailed study of existing literature from the book Visualization Analysis and Design by Tamara Munzner.

Secondly it was through sort of quick ethnographic study by taking up interactive data visualization course by Prof. Venkatesh Rajamanickam at IDC (2018). I also participated in a Geo-visualization workshop conducted by Mr. Rasagy Sharma, designer at Mapbox, India.

5.1 Visualization Analysis and Design by Tamara Munzner

5.1.1 Why this book?

I choose to refer this book firstly because Tamara Munzner, author of the book herself is a notable person in the community and data visualization. She is an expert in this domain.

Secondly in this book Munzner has beautifully provided a systematic and comprehensive framework for thinking about visualization in terms of principles and design choices. While writing this book the author has assumed that the readers do not have prior experience in programming, mathematics, human computer interaction or graphic design and is supposed to be used as an introductory level reading for visualization course - them being my targeted audience for the project.

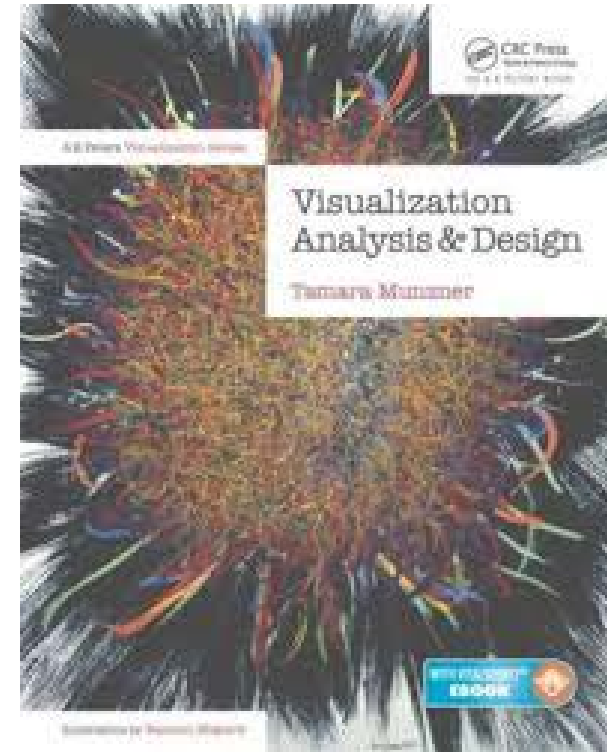


Fig 5. Cover page of the book Visualization Analysis and Design by Tamara Munzner

5.1.2 About this book

The book begins with a definition of visualizations and walks through its many implications in Chapter 1, which ends with a high-level introduction to an analysis framework of breaking down vis design according what-why-how questions that have data-task-idiom answers. Chapter 2 addresses the what question with answers about data abstractions, and Chapter 3 addresses the why question with task abstractions, including an extensive discussion of deriving new data, a preview of the framework of design choices for how idioms can be designed, and several examples of analysis through this framework. Chapter 4 extends the analysis framework to two additional levels: the domain situation level on top and the algorithm level on the bottom, with the what/why level of data and task abstraction and the how level of visual encoding and interaction idiom design in between the two. This chapter encourages using methods to validate your design in a way that matches up with these four levels. Chapter 5 covers the principles of marks and channels for encoding information. Chapter 6 presents eight rules of thumb for design. Whereas the rest of this book gives a detailed analysis of various idioms with suitable examples. It also walks readers through nuances of use of space and color.[3]

5.2 Ethnographic study

The course conducted by Prof. Venkatesh Rajamanickam was broken down into two major parts. It was a studio based approach of teaching where open discussions are highly encouraged. Critiquing also plays a major role in such environments. This studio based approach is a major difference in data visualization course when taken up in design schools as compared to technical schools.

First major part of the course was contributing to open source maps (OSM). It involved, on field data collection as well as arm chair mapping. This activity exposed us to various ways in which data could be found. After the first phase of contributing to OSM, we underwent a project based learning approach (again a widely adopted method). We were given a scenario on which visualization was supposed to be created. There was no defined approach for us to reach the final results. But my personal approach to this was first collecting the data from various resources. Then from the data collected I thought of creating various forms of narrative out of it. Then as per the narrative I sought out to again look if I had sufficient data for the same. If not is it possible to derive it from the existing data? Or is it possible to find the missing data from other sources?

Post data collection, the data was further cleaned. Post creation of a narrative and finding out what I wanted to show, my focus was shifted to imagining the way I wanted my final output to look like. I started off by creating some quick paper sketches. Me being mediocre at coding I quickly jumped onto prototyping the viz to meet the time constraints. I wanted to ensure I allot sufficient time in making the visualization. Due to this reason I restricted myself to lesser exploration of various design idioms.

Apart from this project of creating a data visualization, we also followed another activity in class in which each student was supposed to present a different prototyping tool. These presentations enlightened me about the various tools which are present and various limitations of each prototyping tool. Also the extent of expertise and skills each tool requires to create certain

kind of visualization. Each tool had different set of design idioms which could be prototyped through it. Due to these presentations we were exposed to various kinds of visualization idioms which already exists.

Personally I took a tool which required medium level of coding skills and had a moderate learning curve attached to it. This made me spend some more time on figuring out the technical details which was a good learning experience in itself but my project was compromised when it came to final results. I could not spend enough time in actually getting my visualization iterated after testing it, whether it actually sufficed the goal I had intended it to achieve. At the end of the entire module we were supposed to document our entire process in the form of a medium article.

Alongside as I was progressing with my literature survey on reading the book by Tamara Munzner I realized that there was an intended process/ a designed framework which I should have followed to achieve better results. This made me realize of many shortcomings in my approach which could be bridged.

The experience of attending the Geo-viz workshop opened door to a completely new domain of viz to me. There was a distinctive approach followed in this workshop where we performed short tasks on daily basis. This approach led us to a not so deep but a broader understanding of this field as a whole. We used two sorts of learning methods during this workshop. One way of learning is by creating a viz from the data. While the other was completely opposite where we breakdown an already created viz to actually understand the fundamentals and thought process which went behind the viz. The first one seems to be more concrete but the second approach quickly escalates our understanding of the basics of the visualization and some of its core principles.

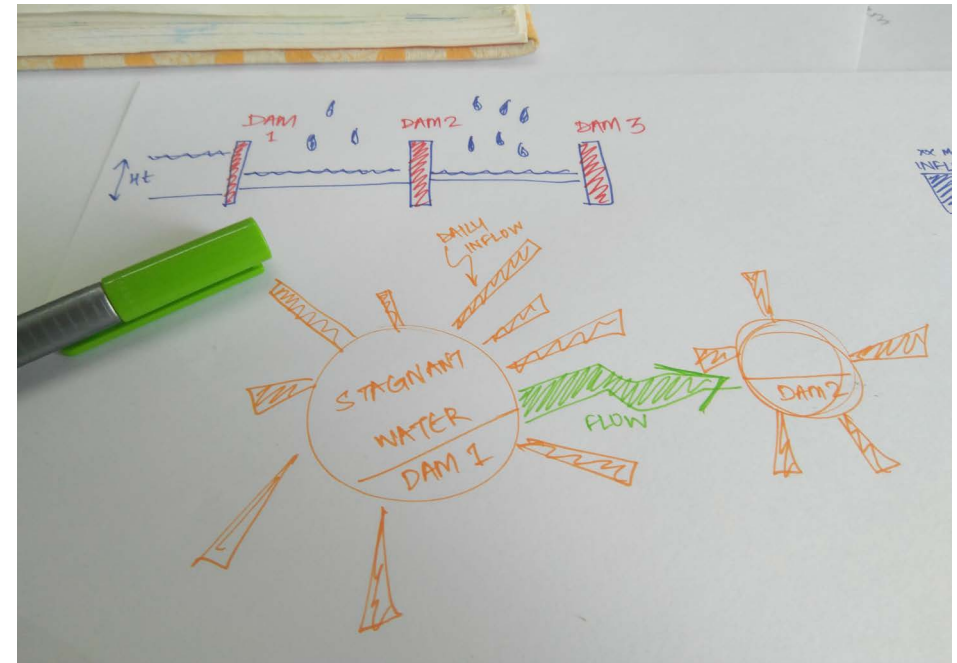


Fig 6. Initial quick paper prototypes

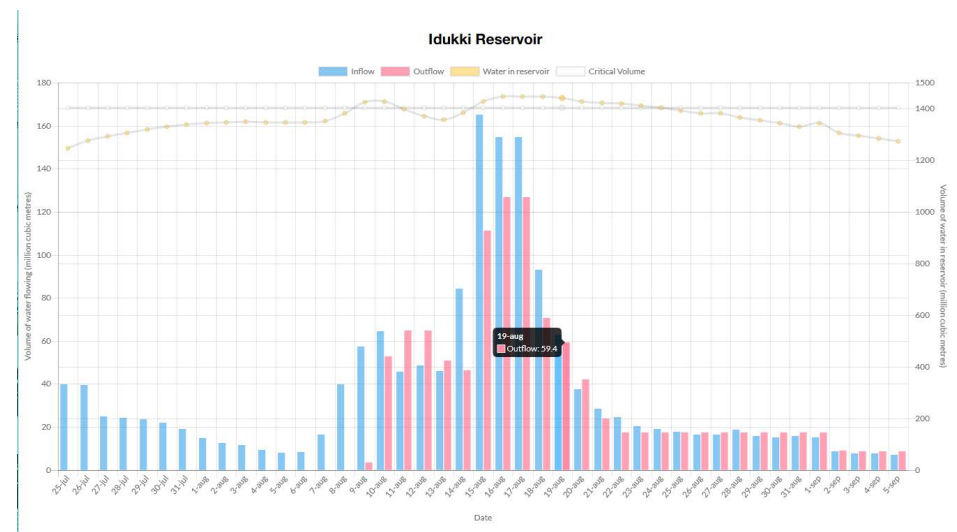


Fig 7. Screenshot from the final Data Viz created

6. Key learnings

This section documents the key learnings from my study and few identified pitfalls which students tend to get trapped into as novices

1. Students end up thinking of their end visualization first without having a correct rationale for what actually they intend to show.
2. Interaction idioms/ Layouts are not explored up to the full extent
3. Not enough emphasis is laid on information density
4. Students directly tend to jump to one kind of solution and start implementing it.
5. They do not spend time on iterating the visualization after testing of the viz. primarily could be because of the time constraints.
6. Writing about things helps students think in a deeper manner about the design decisions previously made.
7. Paper prototyping serves as a quick handy tool to think through our iterations and get some early feedback.
8. Paper prototyping can also enable students to quickly think of the effectiveness a particular visualization may serve or not.
9. Every prototyping tool has certain limitations. They usually may act as a tool bias for a designer.

7. Design consideration space

Students must develop the **data vizualization vocabulary**

Exploration of various **design idioms** is important

Ability to identify the **type of data** is required

An **iterative and multi-linear approach** is required to be developed

Learning can happen in two ways namely-

Making (Constructive)

Critiquing (Destructive/Breaking down)



Fig 8. Design consideration space

8. Design explorations

8.1 Magnetic board viz

This concept as the name suggests has a magnetic board onto which various visualizations can be formed by placing magnetic tokens onto it. These tokens will be a bunch of various marks and channels with which a student can quickly pick, place, create and iterate their basic plots.

Pro -

Easy to construct visualizations

Easier to iterate

Cons -

The visualizations cannot be interactive

Visualizations created are limited by the tokens which are present

8.2 Suggestive documentation tool

This concept is about creating a suggestive documentation tool. It has a structured set of questions which every student has to fill while creating a visualization. These set of questions will serve as triggers in the mind of students to think through the visualization they are creating and also design decisions they make. This questionnaire is created in a way that it guides students through an incremental step by step fundamental breakdown of the data visualization creation process. As a student writes into this tool, other probing questions will also turn up on the screen. It will prompt suggestions in real time as per the content that is being fed into it.

Pro -

Suggestions might lead to more exploration

Writing things down makes a person think through their entire process more deeply, this will enhance the quality of the output that is produced.

Cons -

An intelligence will be difficult to be built in

There is less room for exploration of ideas as everything will in textual form

8.3 Vizphernalia

A kit to construct physical visualizations. It follows a pedagogy of learning by constructing. Vizphernalia is an assorted collection of objects which helps to students to quickly create visualizations in physical world. This concept has been inspired from the Froebel's gift.

The kit comprises of following things:

1. Activity & Evaluation cards – The card will be having certain problem statement and relevant data to it. Evaluation cards will basically be useful in the end when visualization has been created and it requires to be evaluated.
2. VizDeck – A deck of cards which will serve as a repository for various visualization idioms. Each card has a short description about the type of visualization and an indicative image for the same.
3. A4 sheets – They will be blank pages which will serve as a plane for creation of visualization on it.
4. Colored thread – They will help to create viz having relationships between various nodes. They will serve as to form links between the various marks.
5. Lego blocks – A mixed set of size, shape and colors of Lego blocks which have their affordance of their own and can be used as per the will and creativity of students.
6. Board pins – Pins will be used to mark various points on visualization
7. Straw – Another element to vary the thickness of links between the nodes.
8. Colored tiles – Tokens in shape of squares of varied size and color to be used as a stack or for grouping.

9. Colored clay – Clay is mainly for the purpose of exploring 3d visualizations.

Most of these materials will be available in varied color and sizes. And none of them have a prescribed way of using. The above stated ways of use are just suggestive but students are at liberty to express their creativity and create the best kind of visualization out of it.

Pro –

Physical visualization do not have any prototyping tool bias

Students get started with hand on very quickly

It's easy to iterate the visualization

Cons –

The visualizations created could possibly be limited because of fixed set of objects provided in the kit.

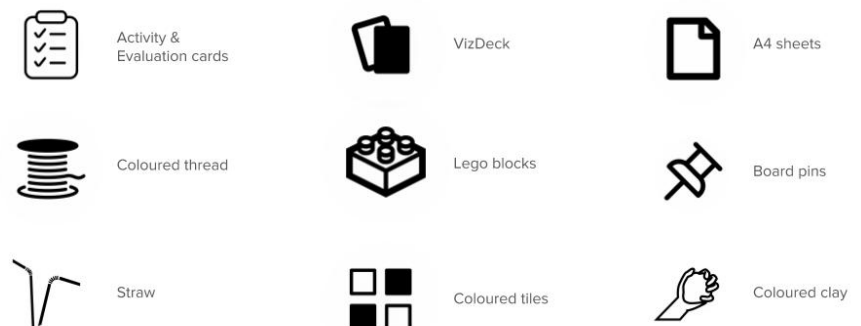


Fig 9. Contents of Vizphernalia

8.4 Method cards

A card based design tool that informs the decisions while designing a data visualization. It guides students with a methodological approach and prompts suggestions and inspiration at every stage during the process of viz creation. This suit of cards makes students think through every choice they make and also help them self-critique and validate their own viz. It is a constructive way of learning by taking one step at a time. The deck of cards also serves the purpose of being a repository of the knowledge of visualization that already exists i.e the visualization idioms a.k.a the different layout of visualization.

Pro -

Tangibility helps students quickly compare and think of various solutions

Does not involve any form of prototyping tool bias

It guides through the entire process of visualization creation

Expands the visualization vocabulary

Cons -

If deck get too huge in number then it gets difficult to explore everything

The information that can fit onto the card is limited

9. Detailed design explorations

9.1 Study on cards in data visualization domain

Graphic continuum -

It is a deck of 52 cards 3" x 4.5" containing image on one face and the flip side has a short description about the image of the type of visualization. The different visualizations are split up following 6 categories of -

- Distribution
- Time
- Comparing
- Geo-spatial
- Part to whole.
- Relationship

Each category has been allotted a different color for quick segregation.

This deck is a sort of general repository for various visualization types and comes into use when you only need to pick the visualization idiom from variety of options. This deck is useful only at the end of visualization design process. This deck has covers a limited set of visualizations.

Pro-

Colors help is quick segregation

Serves as a good repository

Con-

Image is on the flip side of description

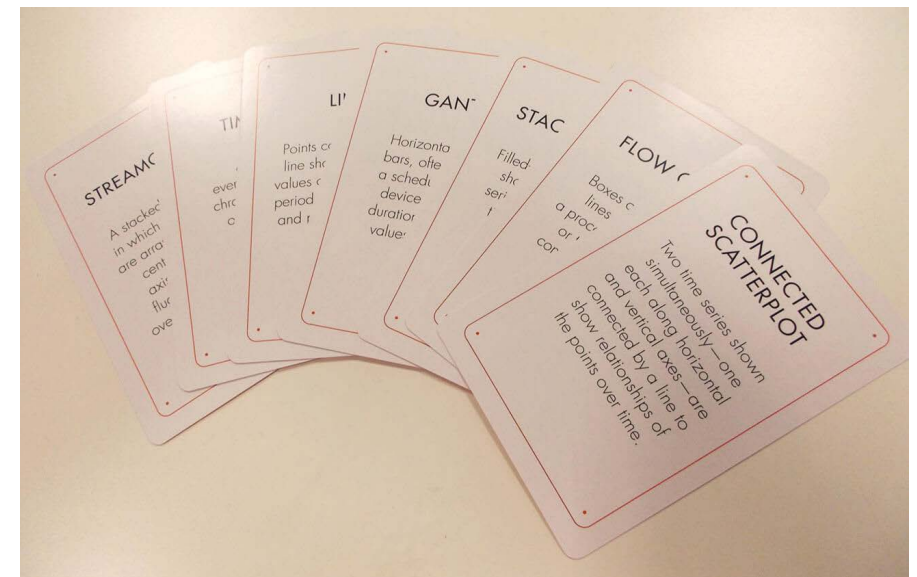


Fig 10 . Graphic Continuum deck

Chart chooser

Chart chooser developed by Stephanie Evergreen is a deck of 51 cards which encompasses 33 visualization types falling under these six categories of-

- Small number, percentages, frequencies
- Time
- Survey responses
- Comparison
- Place
- It's complicated

This deck also has 16 other bonus cards which gives prompts on how to make the visualization more effective by avoiding certain common blunders. Chart chooser again is a sort of repository of visualization types however it has a better suggestive system of which chart type to pick to create the final visualization depending upon the kind of data we have. Although the entire deck has been built from the perspective of displaying data. It does not encompass the factors in which the viz designer would like to do an exploratory data analysis (EDA) and discover things for self.

Pro-

The suggestive system is quite user friendly

Categories are quite well sorted from a display point of view or showing the end results.

Cons -

Chart choosers scope is limited to displaying data (Probably journalists/ presentations POV)

It encompasses a very limited set of visualization types

Useful only at the stage of picking the right kind of chart from the repository

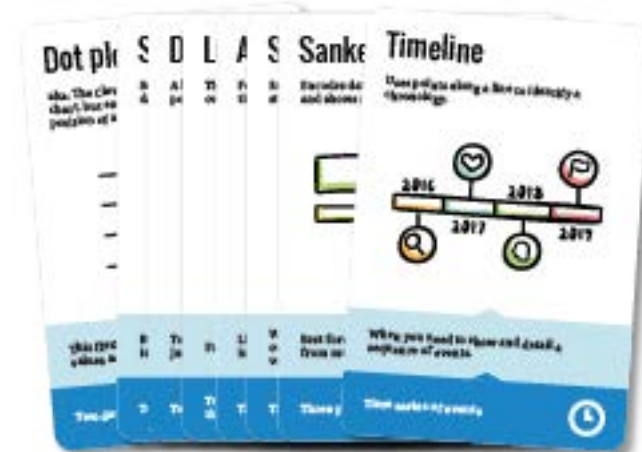
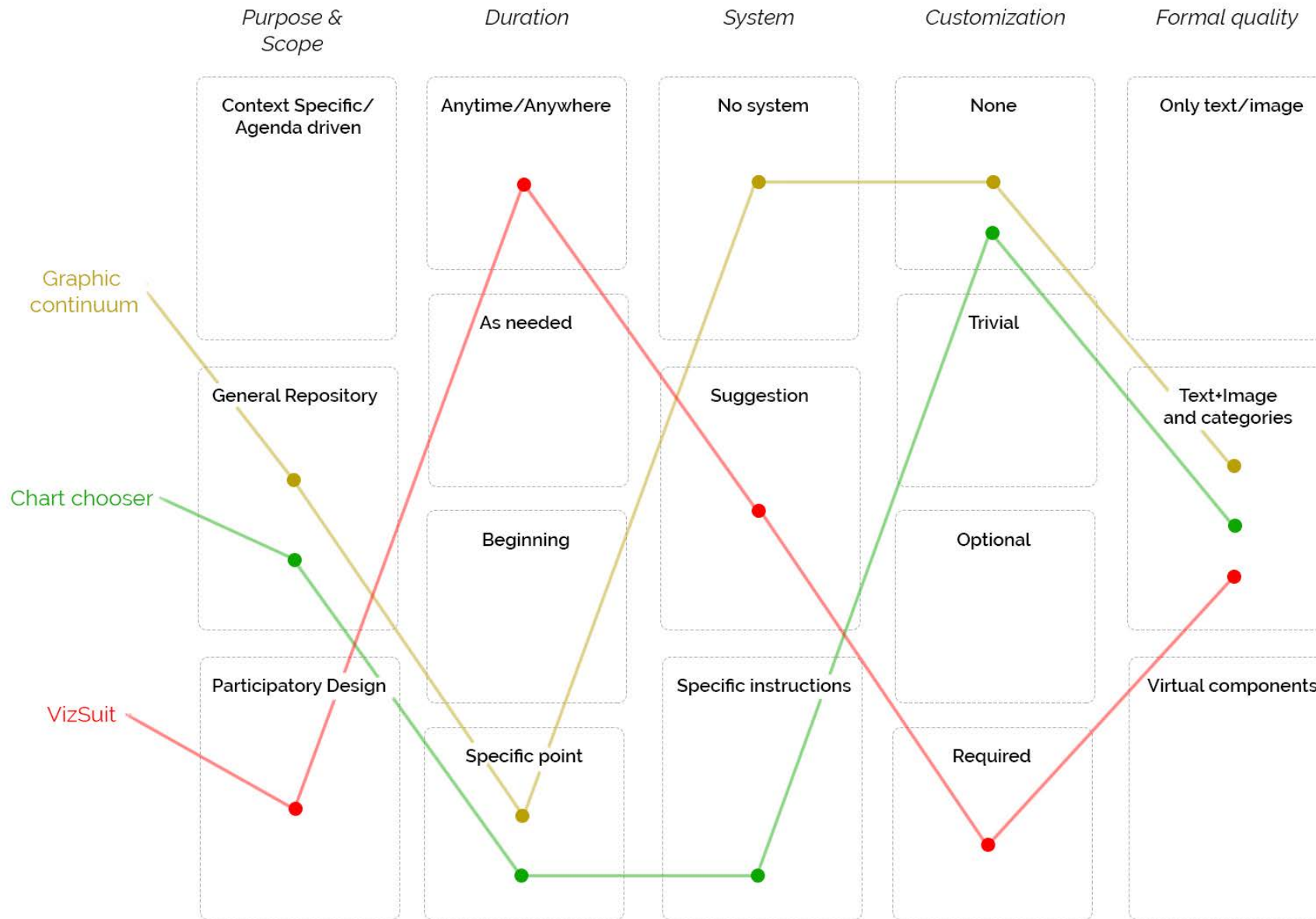


Fig 11 . Chart chooser deck



This framework has been adopted from a paper which reviews eighteen card based design tools in order to understand how they might benefit designers.[9]

9.2 Study on few other miscellaneous deck of cards [10]

Documented below is the list of few deck of cards which I studied. These deck are not directly relevant to learning data visualization however for the inspiration purpose they came in handy and informed multiple decision of mine.

The list of deck are as follows:

- Habit former
- First Spanish flash card
- Slang
- Corporate flash
- Group works
- 50 ways to go green
- Raw food designers deck
- The design deck
- DSD cards
- Worldview
- Powerful facilitation cards
- Should you eat cookies in bathtub
- Practicia

Few generalized patterns I derived from the above decks are-

- How many printed sides does the deck have?
- Does it have illustrations?
- How frequently are the cards supposed to be used?
- Does it have a combination of multiple decks?
- The physical design of the cards - Usual two sided, opens like a book and various dimensions and colors
- How to interconnect multiple cards?
- The purpose of cards is to ideate, diagnose, learn, play or present?

9.3 Way we interact with cards [10]

Given below documents a short study which describes the ways in which we interact with the cards and how we can play around with the nodal nature of individual cards. This short study gave me a sense of the way in which individual suits were to be designed in the entire deck. This study informed my decisions of how physically each individual card should be designed. (Dimensions/ Shape/ Color)

4 randomizers-

- Shuffle - Rearranging by sliding one over the other.
- Deal - Distribute in an orderly rotation
- Draw - Pull or drag from a pile
- Flip - Turn over to expose the other side

3 Sense makers-

- Card sort (Group/stack) - They help in forming categories
- Sequence/ Ranking - This way we give the cards some order of preference
- Compare/ Combine - The cards placed adjacent are evaluated against each other or a combination of multiple can be used to generate some new ideas.

Stephanie Gioia. (2014, Mar 25). 7 motion of cards [Video file].

Retrieved from https://www.youtube.com/watch?v=PSDFsDK_xJQ

9.4 Detailed description of VizSuit

Nested model

This is the core model on the basis of which the VizSuit is developed upon.

The top level is to characterize the problems and data of a particular domain, the next level is to map those into abstract operations and data types, the third level is to design the visual encoding and interaction to support those operations, and the innermost fourth level is to create an algorithm to carry out that design automatically and efficiently. [11]

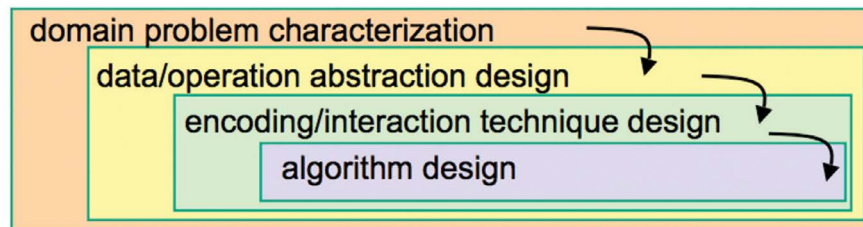


Fig 12 . Nested model

Details on individual suit

The entire deck of the VizSuit can be split into different categories of cards which have been closely derived from the nested model.

The different categories that emerged are as follows -

Data types/categories - Nominal, Ordinal, Quantitative (Discrete/Continuous)

Domain tasks - The task which the user of the visualization would intend to do. This task usually involves terminology pertaining to the domain of the data (i.e Biology, sports etc.)

Abstract tasks - This task is usually written in abstract form. These tasks are devoid of any terms or jargons from the domain of the data. They use a vocabulary which is more generic and native to data visualization

Keyword cards - These are a few identified common purpose or intentions for creating a visualization. These keywords are useful in identifying which idiom is most suitable for the task in hand.

Trends, Patterns, Max/Min, Outlier, Discover, Identify, Locate, Search, Lookup, Compare, Summarize are to name a few

Cheat sheet which maps keywords to idiom cards

Effectiveness card - This card ranks different encodings which is suitable for different kind of data.

Idiom cards - These are the repository of the various forms of visualization techniques which are known to us. They have a short description and icon to depict the form of visualization. There were many categories which have been identified from various sources like -

- Deviation
- Correlation
- Ranking
- Distribution
- Change over time
- Part to whole
- Magnitude
- Spatial
- Proportion
- Range
- Pattern
- Hierarchy
- Location
- Relationship
- Flow
- Comparison

However on further studying each type of visualization idiom, there seemed to be considerable overlaps in their functions and purpose they serve. The six categories which are adopted by the deck Graphic Continuum seemed to be comprehensive enough, and rather than forming categories of my own I chose to stick to those itself. Finally every card fell under these six categories -

- Comparison
- Distribution
- Relationship
- Part to whole
- Geo-spatial
- Data over time

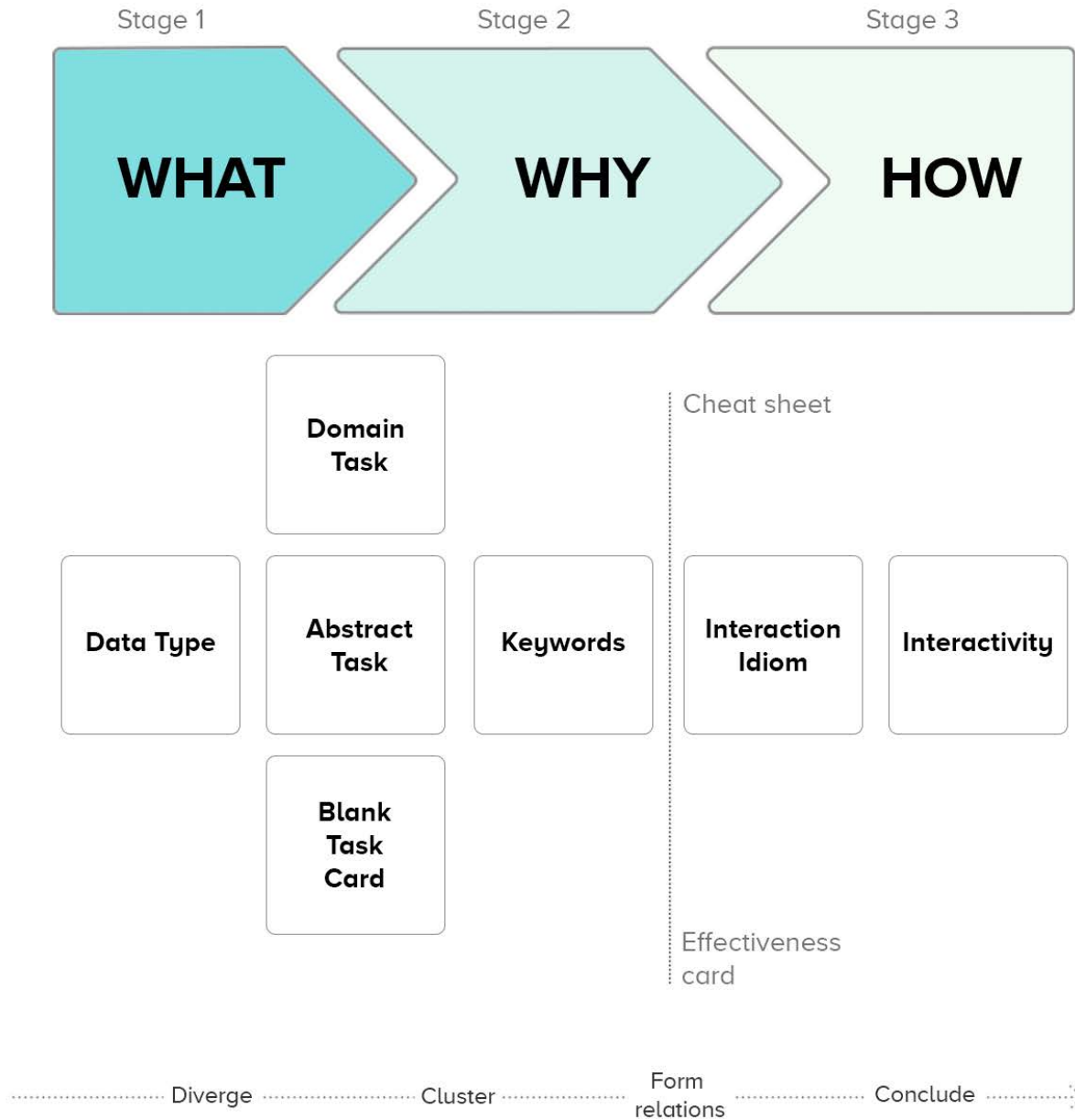
Firstly I found out an extensive list of visualization idioms which are already existing and using the technique of open card sort put every visualization under the above mentioned categories.

Interactivity cards - These are few inspiration cards which can make a visualization richer in case we have the liberty to make the viz interactive. Interactive here refers to the freedom for the user to manipulate the visualization by changing variables.

Filter, Arrange, Align, Order, Change, Select, Navigate, Superimpose, Partition



9.5 When and how to use the cards



10. Workshop protocol

Introduction

This workshop is aimed at taking students through the design experience of a visualization. This workshop is supposed to be a part on in class activity to serve as in introductory hands on session for the students. Participants are expected to work in small groups (2-3) and carry out a task of constructing visualization using this set of cards.

Goal of the workshop

The goal of this workshop is to introduce participants to basics of data visualization and to help them develop a mental model of how to steer through, when faced with a challenge of creating a visualization.

Product goals/ Expected outcomes

- Students improve on their viz vocab
- Students form a mental model of the process of viz design
- They are exposed to larger set of design idioms
- Cards serve as a guiding tool and help participants make appropriate decisions during the design process

Participants Persona

Mainly B.Des 2nd year or M.Des 1st year students

Venue - IDC Studio

Method/ Flow of the workshop

1. Pre-workshop questionnaire
2. Introduction on data visualization
3. Explaining the cards broadly - What/Why/How
4. Introduction to types of data
5. Short exercise to categorize data
6. Brief about marks and channels
7. Brief about effectiveness and expressiveness principles
8. Actual data given and asked to sort the categories
9. Inspiration card of deriving new data is handed over
10. Categorization of derived data
11. Domain and Abstract cards given so as to choose their expected viz outcomes, also blank domain cards are given to encourage participants to come up with their own set of outcomes.
12. Looking back if there is sufficient data or can new data be derived
13. Keyword tokens are handed over for participants to breakdown their decided task to understand the goal which they choose to address through their viz
14. Through keywords they can explore various viz idioms with help of cheat sheet.
15. Sketch out their viz
16. Explain and critique their viz
17. Interactivity cards are handed out
18. Sketch second iteration of viz
19. Explain and critique
20. Final qualitative evaluation

11. Evaluation & Conclusion

The evaluation is still under progress.

This section will be detailed out post analysis of multiple workshops I intend to conduct.

12. Bibliography

- [1] M. Friendly, "A Brief History of Data Visualization," *Handb. Data Vis.*, pp. 15-56, 2006.
- [2] G. S. Owen et al., "How visualization courses have changed over the past 10 years," *IEEE Comput. Graph. Appl.*, vol. 33, no. 4, pp. 14-19, 2013.
- [3] T. Munzner, *Visualization Analysis & Design Visualization Analysis & Design*.
- [4] E. Y. Park, "LINKKI: A planar linkage-based kinetic toy as a tool for education and design," *Proc. 6th Annu. Conf. Creat. Fabr. Educ. - FabLearn '16*, pp. 86-89, 2016.
- [5] N. McCurdy, J. Dykes, and M. Meyer, "Action Design Research and Visualization Design," *Proc. Beyond Time Errors Nov. Eval. Methods Vis. - BELIV '16*, pp. 10-18, 2016.
- [6] S. Mckenna, A. Lex, and M. Meyer, "Worksheets for Guiding Novices through the Visualization Design Process."
- [7] J. C. Roberts, C. Headleand, and P. D. Ritsos, "Sketching Designs Using the Five Design-Sheet Methodology," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 1, pp. 419-428, 2016.
- [8] Enrico Bertini, "Lecture Slides - Google Drive," 2016.
[Online]. Available: <https://drive.google.com/drive/folders/0B-9uY9BLNUVFajg1bGg5YWp3V0k>. [Accessed: 26-Oct-2018].
- [9] C. Wölfel et al., "Method Card Design Dimensions : A Survey of Card-Based Design Tools To cite this version : HAL Id : hal-01497456," 2017.
- [10] "DECKAHOLIC." [Online]. Available: <http://www.deckaholic.com/>. [Accessed: 17-Nov-2018].
- [11] T. Munzner, "Nested Models.pdf."

APPENDIX I

Summary of visualization idioms

Comparison

Bar chart - Numerical comparison between categories

Marimekko - Compare categories and subcategories on basis of width and height of each segment

Multi-set Bar - Multiple data series plotted together under same category

Radial bar - Numerical comparison between categories mainly preferred over bar for aesthetics

Radial column - Numerical comparison between categories and can include negative value

Heatmap - Shows variance between multiple variables

Proportional area chart - Used for quick overall comparison in sizes of data

Nightingale rose chart - Uses area as an indicator for the value of individual categories

Pictogram chart - Small discrete data sets represented as icons placed side by side to compare them

Dot matrix - Used to give a quick overview of the distribution and proportions of each category in a data set and also to compare distribution and proportion across other datasets, in order to discover patterns

Stacked bar - Variation of bar chart to show how a larger category is divided

into smaller categories and what the relationship of each part has on the total amount

Bubble - Helpful in comparing upto 4 variables to draw relationships, patterns and correlations using position, proportion and color

Radar chart - Helpful in showing highest and lowest scoring variables within a dataset

Span - Used to display dataset ranges between a minimum value and a maximum value

Bullet - Has extra visual elements to pack more information over bar graph. It has a target and a background qualitative scale

Time

Line - Used to show trends and how data changed over time

Area graph - Used to convey development of quantitative values across a period

Spiral - The development of a quantity over a large period of time is shown on a Archimedian spiral in form of dots/bars/lines.

Stacked area - Used to convey multiple variables' quantitative development across a period

Steam - Used to convey variation of different variables across a central baseline

Gantt chart - Organization tool for project management to visualize parallelly running activities

Timeline - Helpful in displaying list of events in chronological order

Part to whole

Donut - Used to show the proportions and percentages between categories but has more information density over pie because center is cut out

Pie chart - Used to show the proportions and percentages between categories

Sunburst - This type of visualization shows hierarchy through a series of rings, that are sliced for each category node

Circle packing - Uses packing of circles to show the hierarchy

Treemap - Quantity of each category displayed via its area size

Relationship

Parallel Coordinates Plot - Used for plotting multivariate numerical data

Scatterplot - Mainly used to find correlation between two variables

Arc diagram - Nodes are placed on a single line and arc are used to make connections

Chord - Chord diagram visualizes the inter-relationships between entities

Non ribbon chord - Emphasis on connections within the data

Venn - Displays possible logical relationships between collection of sets

Tree diagram - Way to visually represent hierarchy with a tree like structure

Sankey - They show flow and their quantities in proportion to one another

Distribution

Histogram - A Histogram visualizes the distribution of data over a continuous interval or certain time period. Each bar in a histogram represents the tabulated frequency at each interval/bin

Population pyramid - It's a back to back histogram to display distribution of population amongst different age groups and gender

Stem & leaf plot - Stem & Leaf Plots are a way of organizing data via their place value to show the distribution of data

Parallel sets - They show the flow and proportions like Sankey but they don't use arrows and they divide the flow-path at each displayed line-set.

Geo-spatial

Bubble map - Area of circle at geographical regions represents the value in dataset

Choropleth - Individual geographical regions have color, shade or pattern as per the value in dataset

Connection map - Connection Maps are drawn by connecting points placed on a map by straight or curved lines.

Dot map - Equisized dots are plotted over geographical region to bring out spatial patterns or distribution of data

Flow map - Flow Maps geographically show the movement of information or objects from one location to another and their amount.

Hex/Tile map - The data is formatted to bins and size of each bin represents the value at that geographic location.

Dorlings - Abstract representation of spatial pattern. Area of circle usually depicts the value of data of different regions and they are collectively packed together.

APPENDIX II

Visualization vocabulary

1. Data visualization - It is a general term that describes any effort to help people understand the significance of data by placing it in a visual context.
2. Data type - The type of the data is its structural or mathematical interpretation. At the data level, what kind of thing is it: an item, a link, an attribute?
3. Attribute - Some specific property that can be measured, observed and logged
Synonyms - Variable/Data dimensions
4. Item - An individual entity that is discrete, such as row in a simple table or a node in a network
For example, items may be people, stocks, coffee, shops, genes, or cities.
5. A link is a relationship between items, typically within a network.
6. A grid specifies the strategy for sampling continuous data in terms of both geometric and topological relationships between its cells.
7. A position is spatial data, providing a location in two-dimensional (2D) or three-dimensional (3D) space.
For example, a position might be a latitude-longitude pair describing a location on the Earth's surface or three numbers specifying a location within the region of space measured by a medical scanner.
8. Dataset - Any collection of information that is target of analysis.
9. Tables - Dataset in form of rows and columns
Each row represents the item and each column represents attribute.
10. Networks - The dataset type well suited for showing relations between two or more items.
An item is usually called node and the relation between nodes is called as link.
Synonym Node - Vertex, Link - edge
11. Trees - Network with hierarchical structure are called trees.
12. Set - Unordered group of items
13. List - A group of items with a specific ordering
14. Categorical data - It has two or more categories but there does not exist an intrinsic ordering to these categories.
Synonym - Nominal Data
15. Ordered data - This kind of data has an implicit ordering.
16. Ordinal - The data which has an ordering but arithmetic operations cannot be applied to it.
17. Quantitative data - Quantitative data is anything that can be measured or counted. A subset of ordered data which supports arithmetic operation on it. It is the measurement of magnitude.
18. Sequential data - Ordered data having a homogeneous range from a minimum to a maximum value.
19. Diverging data - Ordered data having two sequences pointing in opposite direction and meeting at a common point zero point.

20. Cyclic data - Ordered data where the values wrap around back to the starting point rather than continuing to increase indefinitely. Example - Time measurements, day, hours

21. Algorithm - An algorithm is defined as a specified process for solving a problem - often written by a human and performed by a computer.

22. Analytics - A broad definition of analytics is the review of data to discover, understand, and communicate meaningful patterns. Or more simply, analytics is the useful insights from raw data.

23. Bias - In general, bias can be defined as an inclination towards a thing, group, or person, often in a way that is inaccurate or unfair. Bias in statistics or data analytics can affect your insights, leading to poor, and potentially, costly business decisions.

24. Correlation - A simple definition of correlation is the relationship between two or more variables (or data sets).

25. Data processing - Data processing is simply a sequence for collecting raw data and turning it into meaningful information.

26. Outlier - An outlier is a piece of data that is distant from the remaining set of data. Think of it as a straggler.