Visualization Concepts

Introduction:

Visualization: Computer-based visualization systems provide ***visual representations*** of ***datasets*** included to help ***people*** carry out some ***task*** more effectively.

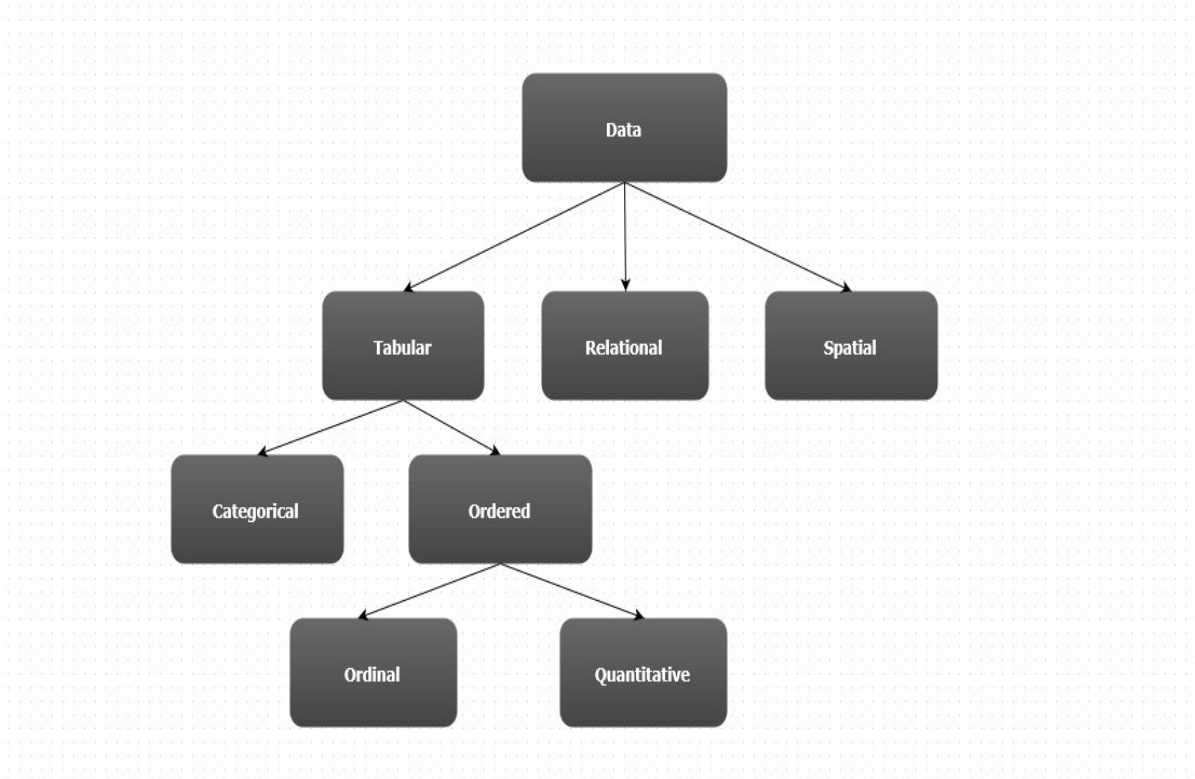
The keywords in the above definition represent the whole essence of visualization. In real time people need some visual representation of the dataset they have to gain greater understanding of the task. Humans have tendency to learn more by seeing something than raw data. We have huge number of design alternatives in this space. This is one of the reason that many people started adopting visualizations for representing the data.

Note: Please do watch the video posted before learning concepts[4].

Principles of Visualization:

1. Know your visual channel types and ranks
2. Categorical color constraints
3. Power of the plane.
4. Danger of Depth
5. Resolutions beat immersion
6. Eyes beat memory
7. Validate against right threat

.



***Datasets and Data Types*:**

DataSets can take different forms. Data types are the fundamental units that combine to make up datasets. Tables, Networks, Fields, Trees etc are few examples. There are three primary dataset types.

1. Tabular : Tabular data is divided into two types.
   a. Categorical
   b. Ordered
      i. Quantitative
      ii. Ordinal
1. Relational
2. Spatial

Categorical Data:

This data doesn't have an implicit order. We can only distinguish between two categories whether if they are same or different. Suppose there are 10 items which are fruits, we can can distinguish which is apple, which is orange etc.

Eg: Fruits like apples, orange etc.,

Ordinal Data:

This data have particular ordering but cannot be converted into exact or definite numerical values. Example is shirt sizes. A shirt can be small, medium or Large. Medium comes between the other two. But what is the exact difference between large and medium?

Quantitative Data:

Any data that have numerical values are Quantitative data. Age, Height, Distance etc.,

Relational Data:

In this kind of data graphs are connected by links. Networks, Trees are examples of this data. Trees are best used for hierarchical data. Usually tables used in databases are relational data. Rows represent the item and column represent the attributes.

Spatial Data:

Data about a location or a point in 3D space can be referred as Spatial Data. Examples are Geographic locations, Tensors etc.,

Visual Encoding:

Analyze the data as combination as Marks and Channels

Now the Question is what are marks and channels.

Marks: Marks are the basic graphical elements in a visualization. These are nothing but geometric primitives. Used to represent items and links.

- Points - Zero Dimensional
- Lines - One Dimensional
- Areas - Two Dimensional
- Volume - Three Dimensional(rarely used).

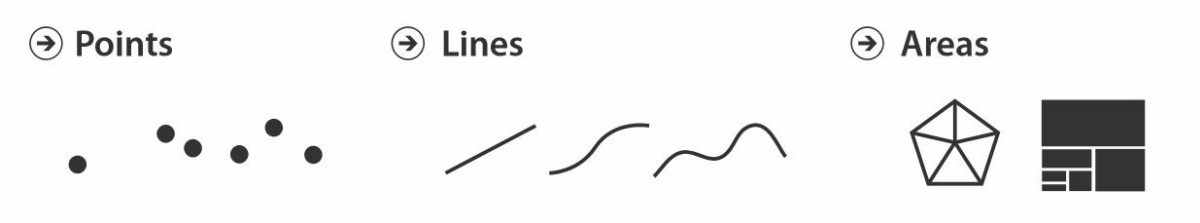


Image Source[1]

Visual Channels: Visual channels control the appearance based on marks.

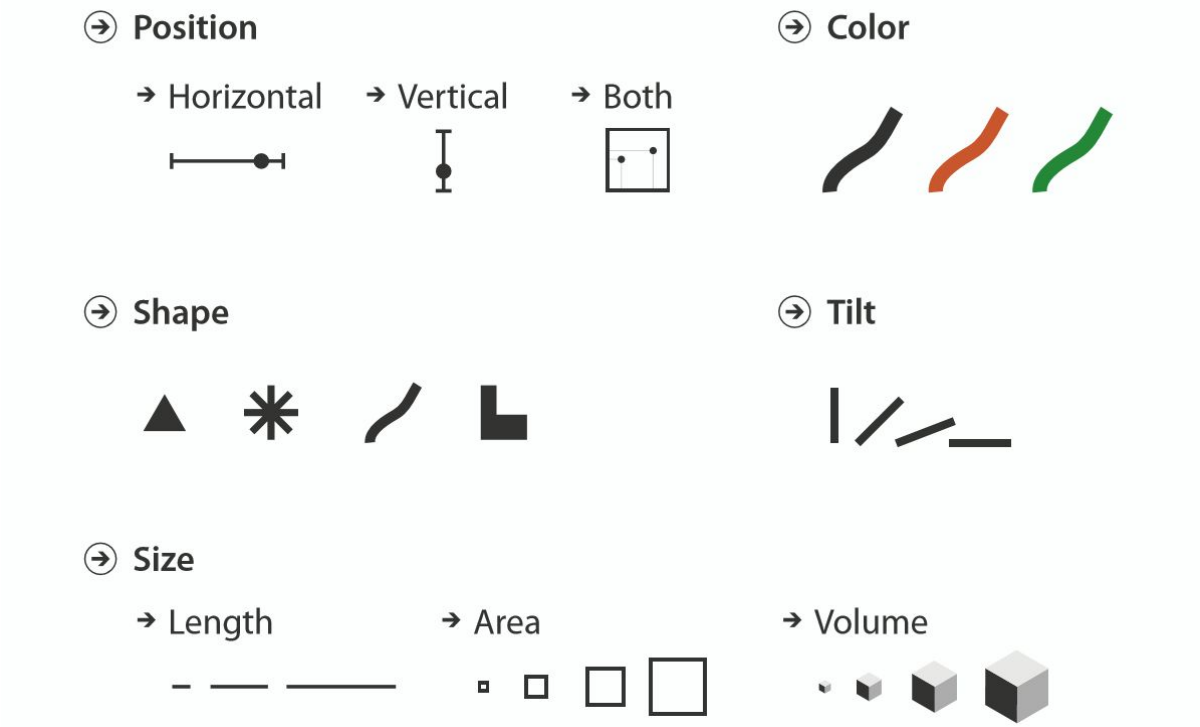


Image Source[1]

Types of Channels:

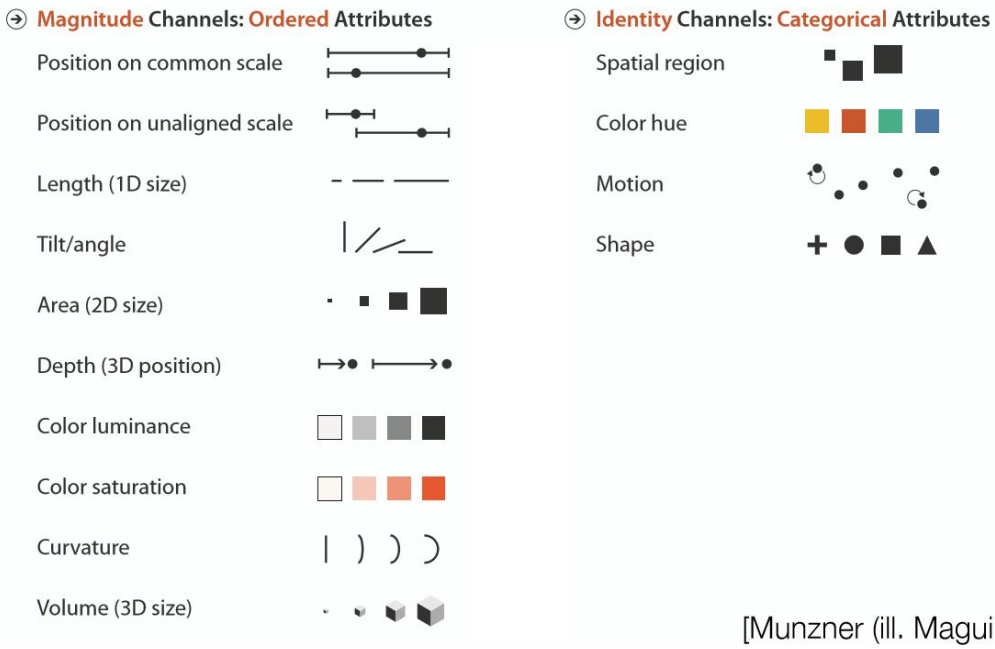
Magnitude Channels: Used for Ordinal and Quantitative data. If you can put a question on data, 'How much?', then use magnitude channels.

Identity Channels: Used for Categorical data. If you can put a question on data, 'What? Where?', then use Identity channels.

Grouping:

Please refer to video for more details about the following channels.

- Containment
- Connection
- Similarity
- Proximity
- Identity => what or where, Magnitude => how much



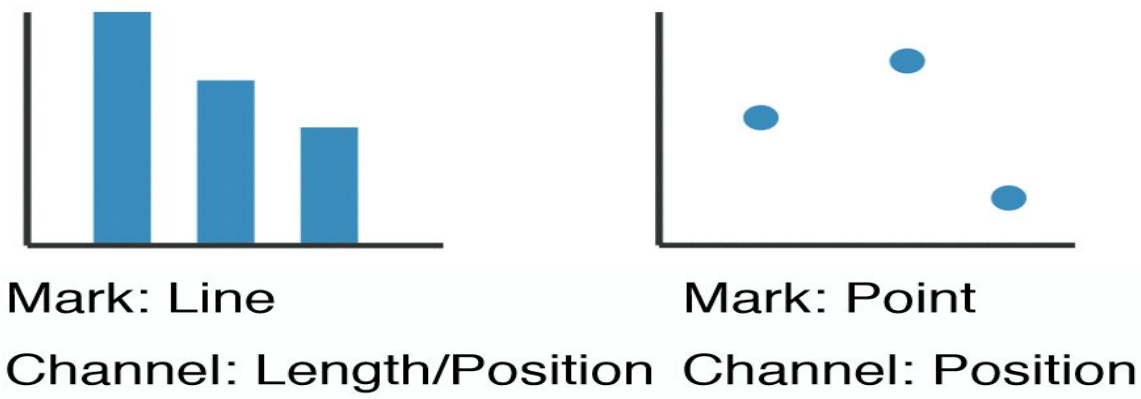
[Munzner (ill. Maguire), 2014]

***Hue*:** A Hue is somewhat synonymous to what we usually refer as colors. Different hues have different wavelengths in spectrum. The value is the measurement of the brightness of a color. Red, Green, Yellow, Blue etc are different hues.

How do we use these Marks and Channels?

Usually we map one attribute a channel. Also it is possible that we can map one attribute to multiple channels.

Examples:



Expressiveness and Effectiveness are two most important principles for visualization.

Expressiveness:

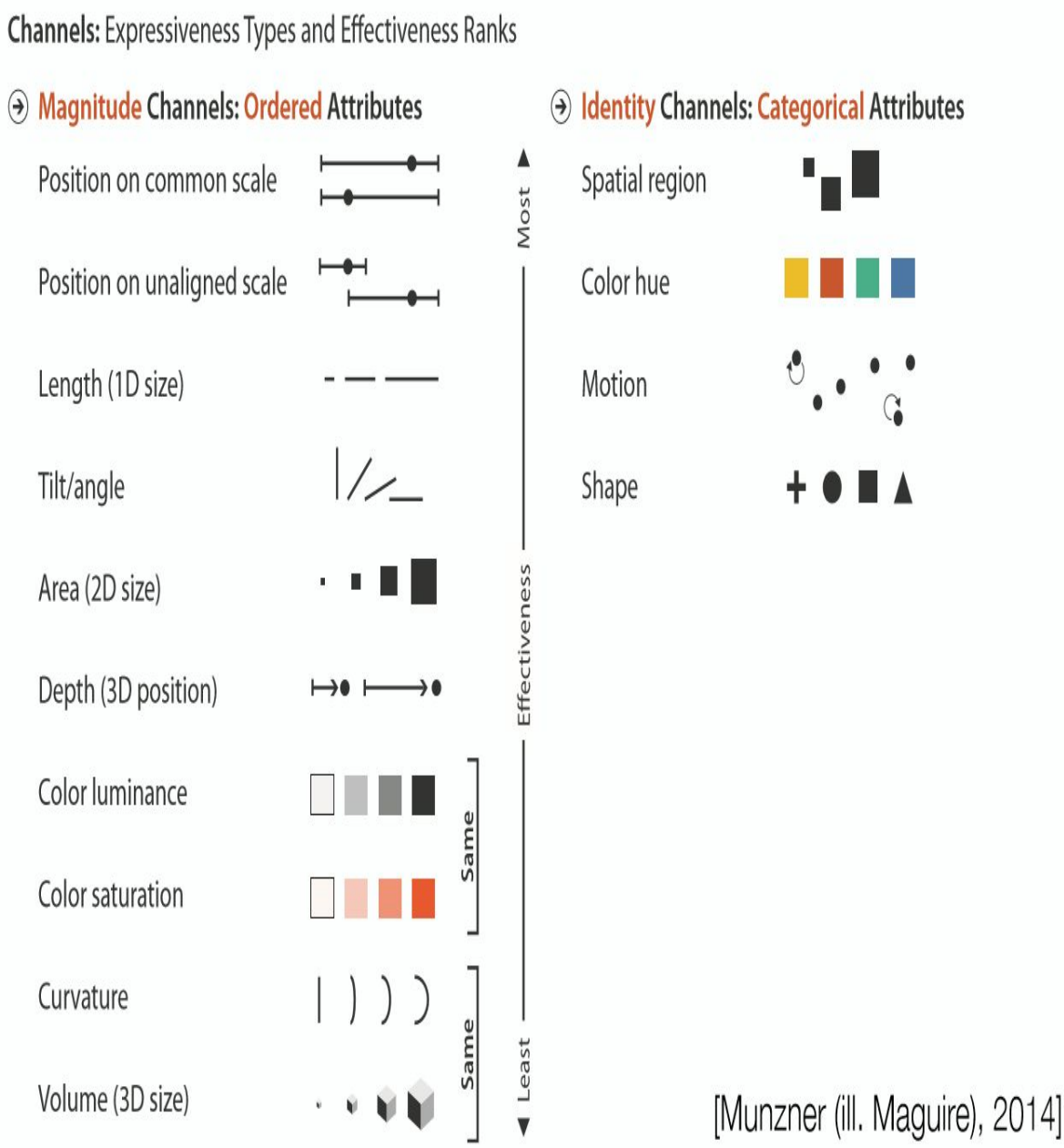
All the data from the dataset and nothing more should be shown.

- This principle is represented by identifying ‘Types’ (categories) in dataset.
- Do encode the data in an ordered fashion.
- Don’t encode categorical data in way that implies ordering.

Effectiveness:

The most important attributes should be the most salient.

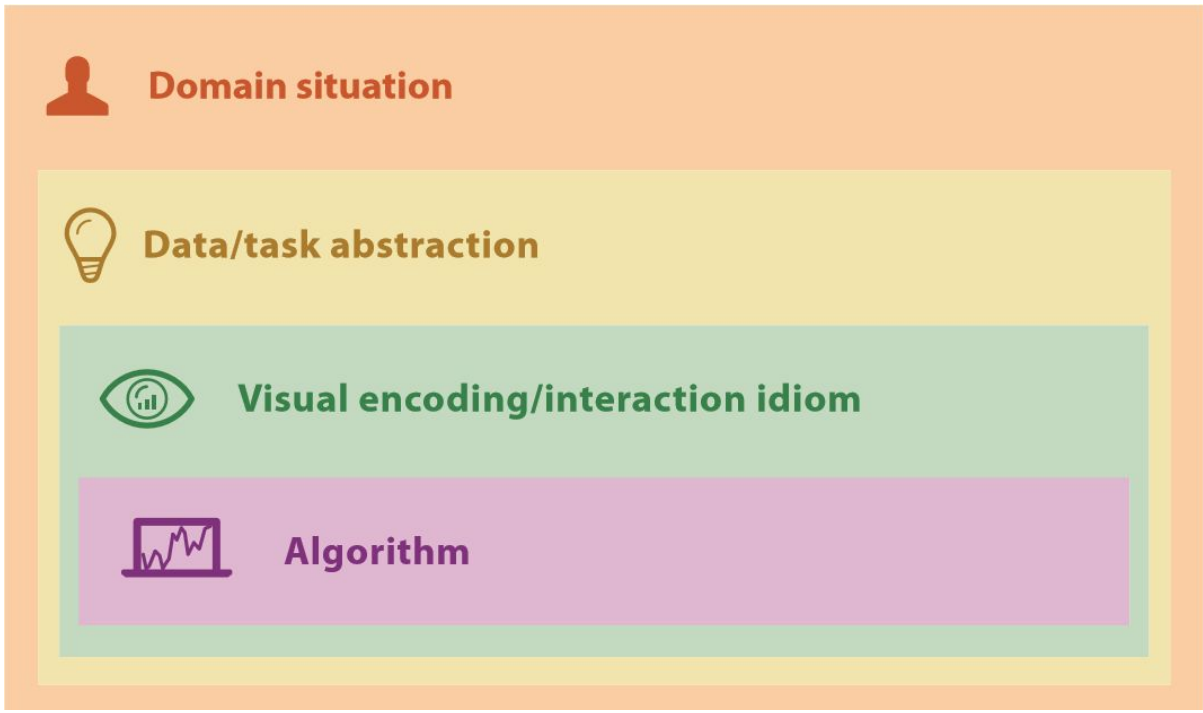
- Saliency - How noticeable something is
- This will be achieved by ranking
- Rankings come from accuracy, discriminability, separability and pop out.



Design:

The process of creating a visualization involves,

- Understanding real world problems
- Mapping problems to abstract task and data types
- Visually encoding the data(through marks and channels)
- coding up the algorithm for displaying visualization.



Conclusion:

Visualizations provides an external cognition tool that helps users analyze, explore and understand data and make well informed decisions[2]

References & Resources:

[1] Lecture by Prof. D. Koop at Umass Dartmouth. Click Here.
[2] Lecture by Prof. Edith Law and Mike Terry. Click Here.
[3] Lecture by Prof. Tamara Munzner. Click Here.
[4] “Visualization Analysis and Design” by Prof. Tamara Munzner.