Topological Equivariant Artist Model

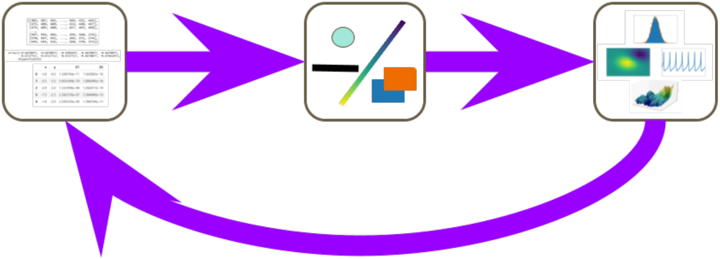
December 27, 2021

Hannah Aizenman, Tom Caswell, Michael Grossberg

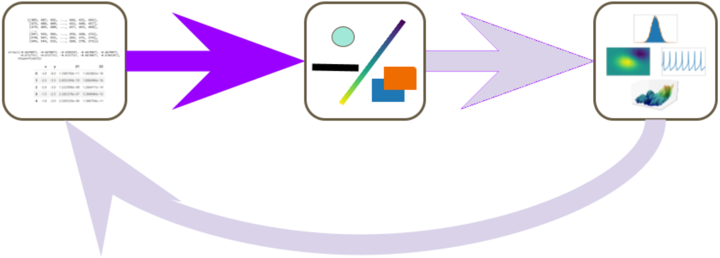
What are we doing?

- develop a model for describing data to graphic transformations
- specify a visualization library architecture based on this model
- implement functional(ish) components based on this model using ideas from functional programming

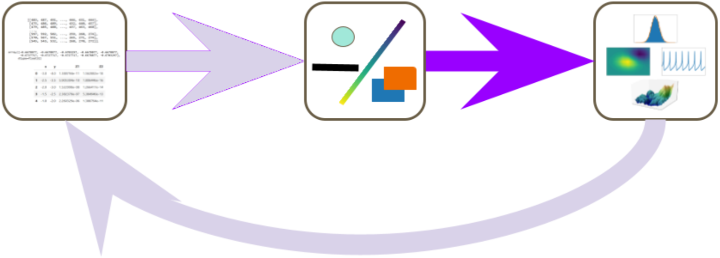
What do visualization libraries do?



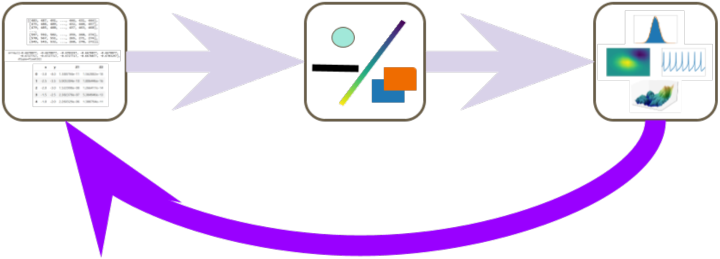
What do visualization libraries do?



What do visualization libraries do?



What do visualization libraries do?



Case Study: Matplotlib

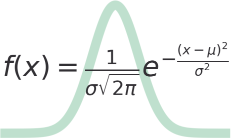
Visualizations Preserve Structure

continuity topological properties [1], i.e. how elements in a dataset are organized, e.g. discrete rows in a table, networked nodes, pixels in an image, points on a line

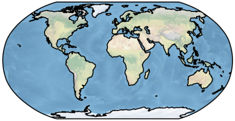
equivariance data and visual encodings are matched such that transformations have an equivalent effect on data and graphical representations, e.g. rotating a matrix and image, shifting points on a line and a line graphic

Continuity

NAME	TEMP (°F)	PRCP (in.)
NEW YORK LAGUARDIA AP	61.00	0.4685
BINGHAMTON	-12.00	0.0315
NEW YORK JFK INTL AP	49.00	0.7402
ISLIP LI MACARTHUR AP	11.00	0.0709
SYRACUSE HANCOCK INTL AP	13.00	0.0118



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{\sigma^2}}$$



(a) (b) (c)

Expressed in container

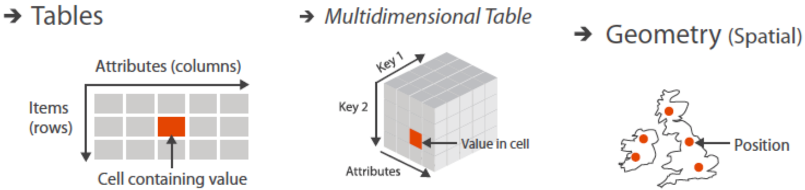


Figure: Figure 2.8 in Munzner's Visualization Analysis and Design[2]

Why is container type not enough?

Equivariance

Retinal Variables & Marks visual encodings should match properties of the data [3]

Graphical Integrity graphs show **only** the data[4]

Naturalness easier to understand when properties match[5]

Expressiveness which structure preserving mappings can a tool implement[6]]

Frameworks for Expressing Visual Equivariance

linguistically visualization has syntax, semantics, grammar
expresses how to design structure preserving visualizations [6, 7, 1]

algebraically transformations on data and graphics are equivalent
symmetric [8]

D data

R data representation

V visualization

$$\begin{tikzcd} D \arrow[r, "r_1"] \arrow[d, "\alpha"'] & R \arrow[r, "\nu"] & V \arrow[d, "\omega"] \\ D \arrow[r, "r_2"'] & R \arrow[r, "\nu"'] & V \end{tikzcd}$$

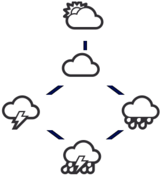
categorically $understanding = read \circ render$ [9]

Domain specific libraries know their structure[22]

DATE	LATITUDE	LONGITUDE	PRCP (in.)	NAME
2021-01-01	43.1167	-77.6767	0.2205	ROCHESTER GTR INTL AP
2021-01-01	41.5	-73.9	0.0000	STONYKILL NEW YORK
2021-01-01	42.7431	-73.8092	0.2598	ALBANY AP
2021-01-01	43.8	-73.7	0.0000	SCHROON LAKE NEW YORK
2021-01-01	43.0078	-73.6511	0.0000	SARA NEW YORK
2021-01-01	40.7794	-73.8803	0.6614	NEW YORK LAGUARDIA AP
2021-01-01	40.6386	-73.7622	0.6299	NEW YORK JFK INTL AP
2021-01-01	43.1111	-76.1039	0.4094	SYRACUSE HANCOCK INTL AP
2021-01-01	40.7939	-73.1017	0.5984	ISLIP LI MACARTHUR AP
2021-01-01	43.35	-73.6167	0.1181	GLENS FALLS AP

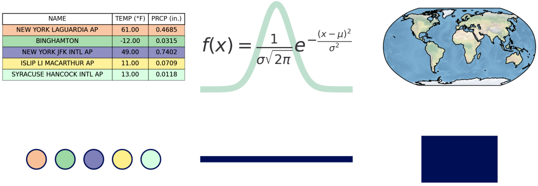
ggplot[10]
Vega[13]
Altair[16]
Tableau [19]
[20, 21]

ImageJ[11]
ImagePlot[14]
Napari[17]



Gephi[12]
Graphviz[15]
Networkx[18]

General purpose libraries generally can't[32]



(a) (b) (c)

1. Matplotlib[23] →Seaborn[24], xarray [25]
2. D3 [26]
3. VTK [27, 28](MayaVi[29])→ Titan[30], ParaView[31]

Mathematical Data Abstraction

Fiber Bundles "unified, dimension-independent framework" that expresses data as the mapping between continuity and fields [33, 34]

Simplicial Databases systemic way to apply a schema to a fiber, i.e. provides a way to name fields and bind them to types system (e.g. `int`, `float`) [35, 36]

Sheaves on Bundles "algebraic data structure" for representing data over topological spaces [37]

Fiber Bundles

$$\begin{tikzcd} F \arrow[r, hook] & E \arrow[d, "\pi"'] \\ & K \arrow[u, bend right, "\tau\in\Gamma(K,E)"'] \end{tikzcd}$$

Base Space $(K, \mathcal{T})$ with

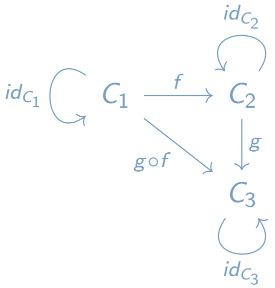
- open sets $U \subset K$
- points $k \in U \in K$

Fiber Space $F_k = \pi^{-1}(k)$

Total Space $E{\restriction_U} = K{\restriction_U} \times F{\restriction_U}$
for all $U \subset K$

Sections
$$\Gamma(U, E{\restriction_U}) \coloneqq \left\{\tau : U \to E{\restriction_U} \;\middle|\; \pi(\tau(k)) = k \text{ for all } k \in U\right\}$$

Category $\mathcal{C}$



associativity

if $f : C_1 \to C_2$, $g : C_2 \to C_3$ and $h : C_3 \to C_4$ then $h \circ (g \circ f) = (h \circ g) \circ f$

identity

for every $f : C_1 \to C_2$ there exists identity morphisms $f \circ id_{C_1} = f = id_{C_2} \circ f$

Opposite Category $\mathcal{C}^{op}$

$$
\begin{tikzcd}
C_1 \arrow[loop left, "id_{C_1}"] \arrow[r, "f"] \arrow[dr, "g \circ f"'] & C_2 \arrow[loop above, "id_{C_2}"] \arrow[d, "g"] \\
& C_3 \arrow[loop below, "id_{C_3}"]
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
C_1 \arrow[loop left, "id_{C_1}"] & C_2 \arrow[loop above, "id_{C_2}"] \arrow[l, "f"'] \\
& C_3 \arrow[loop below, "id_{C_3}"] \arrow[u, "g"'] \arrow[ul, "f \circ g"]
\end{tikzcd}
$$

Functor $F : \mathcal{C} \to \mathcal{D}$

$$
\begin{tikzcd}
c \arrow[r, "F"] \arrow[d, "f"'] & F(c) \arrow[d, "F(f)"] \\
c' \arrow[r, "F"'] & F(c')
\end{tikzcd}
$$

composition

$$F(g) \circ F(f) = F(g \circ f)$$

identity

$$F(id_c) = id_{F(c)}$$

Natural Transformation $\alpha : F \to G$

$$\begin{tikzcd}[column sep=huge] \mathcal{C} \arrow[r, bend left=50, "F"] \arrow[r, bend right=50, "G"'] \arrow[r, phantom, "\Downarrow\alpha"] & \mathcal{D} \end{tikzcd}$$

Natural Transformation $\alpha : F \to G$

$$
\begin{tikzcd}
F(c) \arrow[d, "F(f)"'] \arrow[rr, bend left, "\alpha_c"] & c \arrow[l, "F"'] \arrow[r, "G"'] \arrow[d, "f"] & G(c) \arrow[d, "G(f)"] \\
F(c') \arrow[rr, bend right, "\alpha_{c'}"'] & c' \arrow[l, "F"] \arrow[r, "G"] & G(c')
\end{tikzcd}
$$

Presheaf: $\mathcal{O} : U \to \Gamma(U, E\restriction_U)$

$$
\begin{tikzcd}
\Gamma(U_1, E\restriction_{U_1}) & & \Gamma(U_2, E\restriction_{U_2}) \arrow[ll, hook', "\iota^*"'] \\
U_1 \arrow[u, "\mathcal{O}_K"] \arrow[rr, hook, "\iota"'] \arrow[r, phantom, ""] & \arrow[u, dotted, "\mathcal{O}_K"] & U_2 \arrow[u, "\mathcal{O}_K"']
\end{tikzcd}
$$

Sheaf

Stalk & Germ

$$\mathcal{O}(K)\restriction_k \coloneqq \lim_{U \ni k} \Gamma(U, E\restriction_U)\tau(k) \in \mathcal{O}(K)$$