

1)

Decision: data, audience, message

Exploratory: quickly, many

Explanatory: not one-shot, be careful craft multiple drafts

Visualization: beginning process understand data, middle debug, end communicate

Reason use visualization: human visual + info density in image + convey pattern difficult

Unaligned baseline (length/distance baseline > no base > area > color)

Nominal (equality): discrete, quantitative, categorical

Ordinal: interval scaled, comparison, subtraction, no division

Numerical: quantitative, continuous, ratio-scale, comparison, subtraction, division

Categorical: unordered discrete values 无序离散

Ordinal: discrete with a natural order 具有自然顺序离散

Interval: natural order with even steps 自然顺序均匀

Gestalt psy: human brain find structure in visual stimuli

Weber's Law: relative. Perception: perceive percentage increase. Pre-attentive: 200ms

Steven's Power Law – NOT linear. actual and perceived value $S=I^N$

2)

Scatter Plot 散点图(多 Y — X, sampled data, 不必有 0) / slopegraph 坡度图 / Dot Chart (horizontal bar/point) / Stacked Bar Graphs 一柱多块(subgroup, individual pie, poor on compare) / Area Chart 折线下面积(proportional bar graph w/ connected bars, area perception channel require caution) / packed bubble 大大卷 / Bar chart(ratio scaled)

Line: Continuous. Trendy. Connection

Contingency Table - two catevar, at least one numerical (not original data, aggregation func) -> lattice plots 点阵图

Hierarchical Data 分层(split layer tree): Space Filling, Node-Link (graph/tree): directed, acyclic, ordering. (Run out space, encode structure, racial layout space for deeper node)

Node-Link Diagram 八爪鱼状

Hyperbolic Trees 圆形视野八爪鱼树状

Tree Map: hierarchy (bar chart), space filling.

Melting/pivoting: new var tell data from original col, data in pivot col get row in trans data

3)

clarity 清晰 clutter(data-ink)杂乱 trim 修剪

grid 网格 distortion 变形 Tick Marks 刻度线

Data to ink ratio: prominent 突出 Remove clutter

non-data ink: ink NOT represent data, redundant & removable, change as data change

chart-ink: visual element, graphical element refer to theme.

Bar charts must have 0!

Graph mislead 注意: change Y data fit tight, scale area, change X aggregation

Aspect ratio: shape 45°斜率

Guide decision: audience & task try to complete

4)

Divergent 发散/不同的: var 2 separate extremes & neutral ground in between

Sequential: continuous vary low to high

Qualitative: map var nominal type, not continuous. Clutter too many colors.

-Color scales: Discrete/Categorical, Sequential, Divergent

-Variables: Categorical, Ordinal, numerical

Panel plot: data dense set timeseries lines. Break lines into separate, confuse/mislead

Smooth timeseries data: moving average, weighted moving-average, LOESS

Dichotomy 二分法 绿色最敏感

Hue: distinguish 色相 -> color wheel (similar shade adjacent 相邻, 对面相反)

Saturation: how pure, white/gray 饱和度

Lightness 反射: sequential color palette Brightness 自发性

RGB 彩色正方体: Red, green, blue, 3D cube

HSV 钻石型, 白->黑(尖): cone

CIE-lab: match human perception color

Color blind: unaware data distinction w/ certain palette

Color: value continuous data, diff type terrain, encode topological info, categorical code info

Circle show data as location: not strong encoding comparison, circle pitfall (encode area not r), overplot (hard to read)

Topographical 地形

Choropleth(heatmap)染色地图: individual colored correspond

Cartogram 染色分块图: match landmass

Radial/Rose Plots 地球两极状 r square root 半径平方根

5)

Histogram 直方图 Univariate Scatterplot 单变量散点图 Jittering 抖动不会失真

Beeswarm/Dot Plot 蜂巢点图: discretizes y axis 离散 y 轴

Tukey Box-Plot/Box and Whiskers 单个方形糖葫芦状: outliers plotted as points, one quartile either side of median

Box Plot 箱形图 95% confidence interval

Quantile-Quantile Plots 点状轨迹图: two distributions comparison, plot 2 vars value in order <-> perfect normal 完美法线

Kernel Density Estimation (KDE) Curve -> sum normal distributions 正态分布求和

Distribution Curve: discreteness 离散性

Histo/Density 多峰状 Raincloud Plots 山丘云状 Ridgeline Plot 多座山脊状

Violin plot/Bean plot 纺锤状 median 中位数

Axis Transformation 轴变换: **Shift scale by offset**

Exponential func 指数: $a \cdot n^x$

Logarithms 对数 -> (Inverse of exponent)指数的倒数 -> $b^p = x$ (logarithm x, base b, power p)

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$
 Base only affects the scale, same graph

Natural logarithm e: irrational -> logex (ln)