SEE Conf 3rd
蚂蚁金服体验科技大会

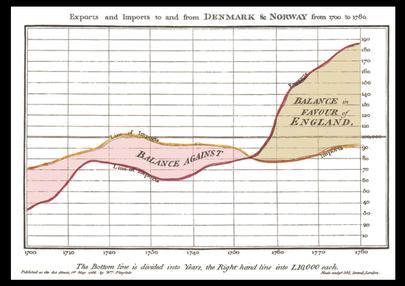
数据可视化的智能研发流程
智能可视化体系 AVA

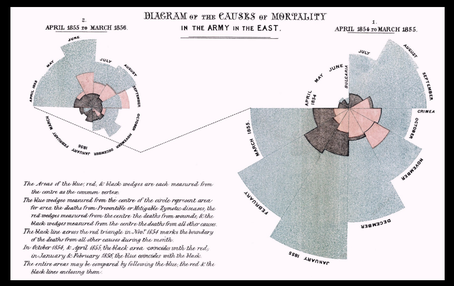
步茗
AntV 核心开发者

璆鸣
AntV 核心开发者

数据可视化实现方式的发展

从手绘到代码化、工具化，研发与设计成本逐渐降低





研发成本

时间

2009 HIGHCHARTS

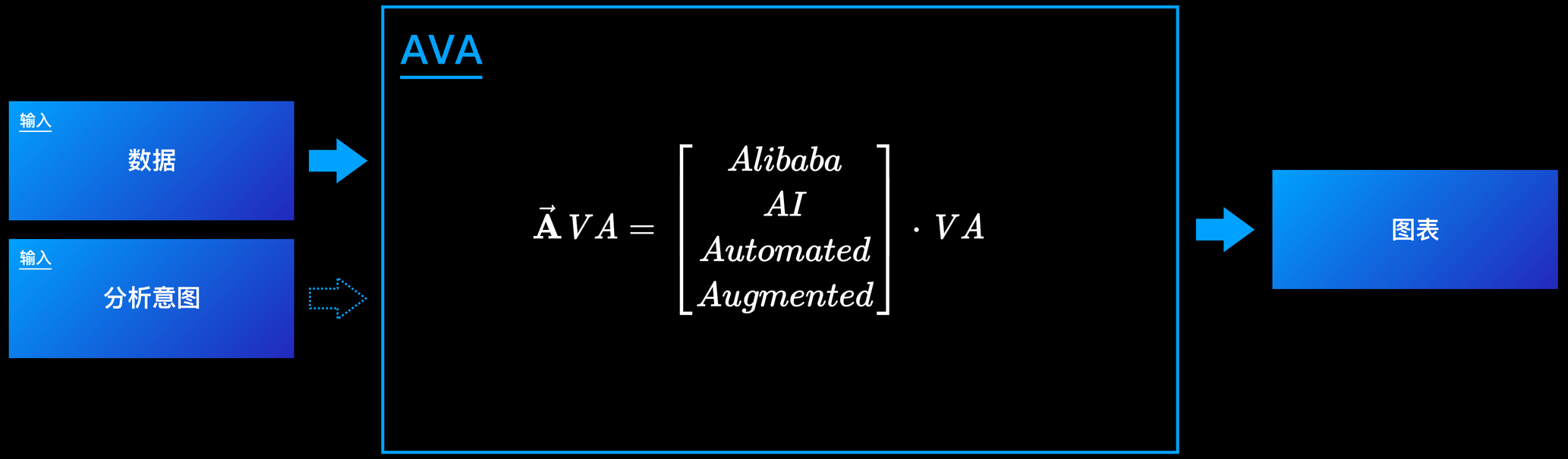
2011 D3 tableau

2013 ECHARTS

2015 AntV

$\vec{\mathbf{A}}VA$

$\vec{\mathbf{A}}$ Visual Analytics



可视化研发之痛

痛苦的研发体验各有各的痛苦

理想场景

人员齐备 接口到位 需求定好 设计定稿

后端

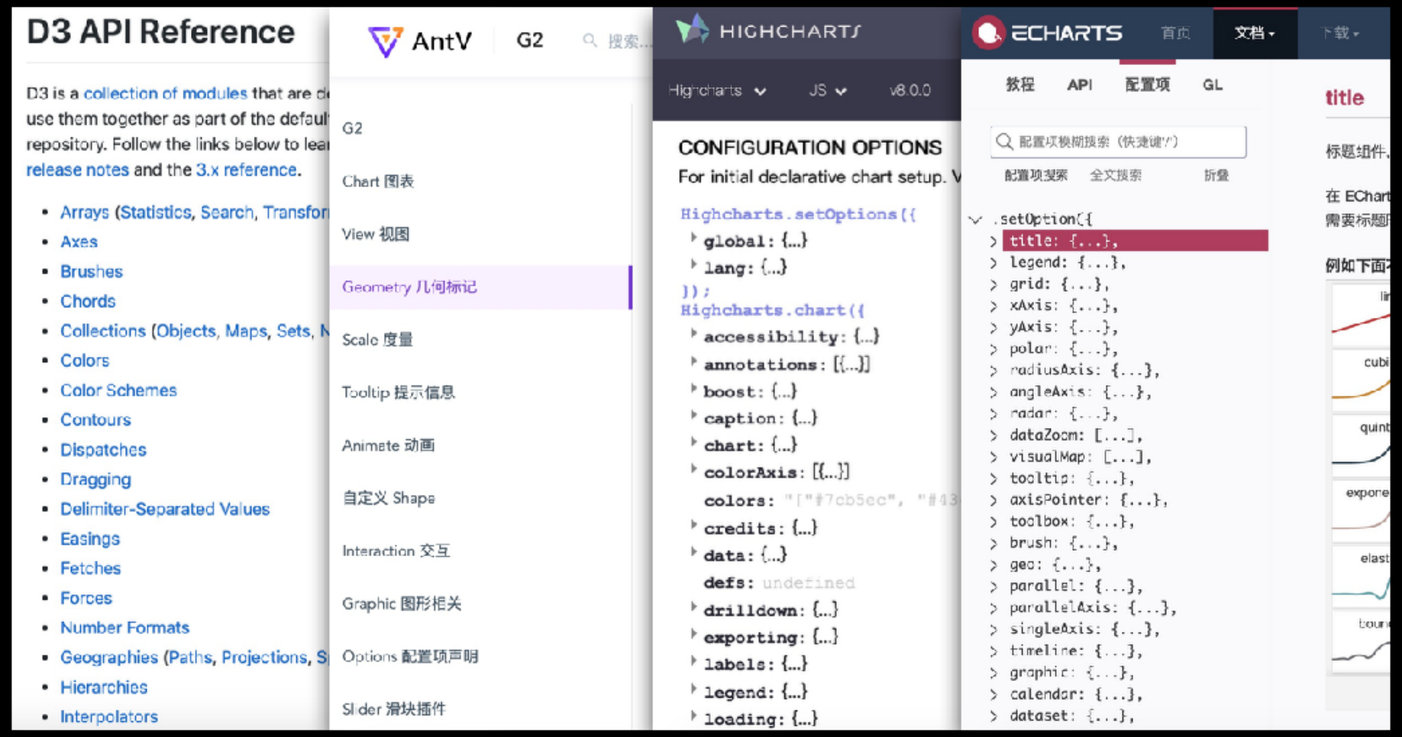
前端

设计

理想场景 – 文档之痛

怎么画呢

在文档丛林中探险



专业知识多

数据处理	图形学	可视化	设计学
统计学	几何计算	视觉通道	设计原理
线性代数	图形算法	图形语法	色彩空间
特征工程	渲染原理	图表分类学	人机交互
……	……	……	……

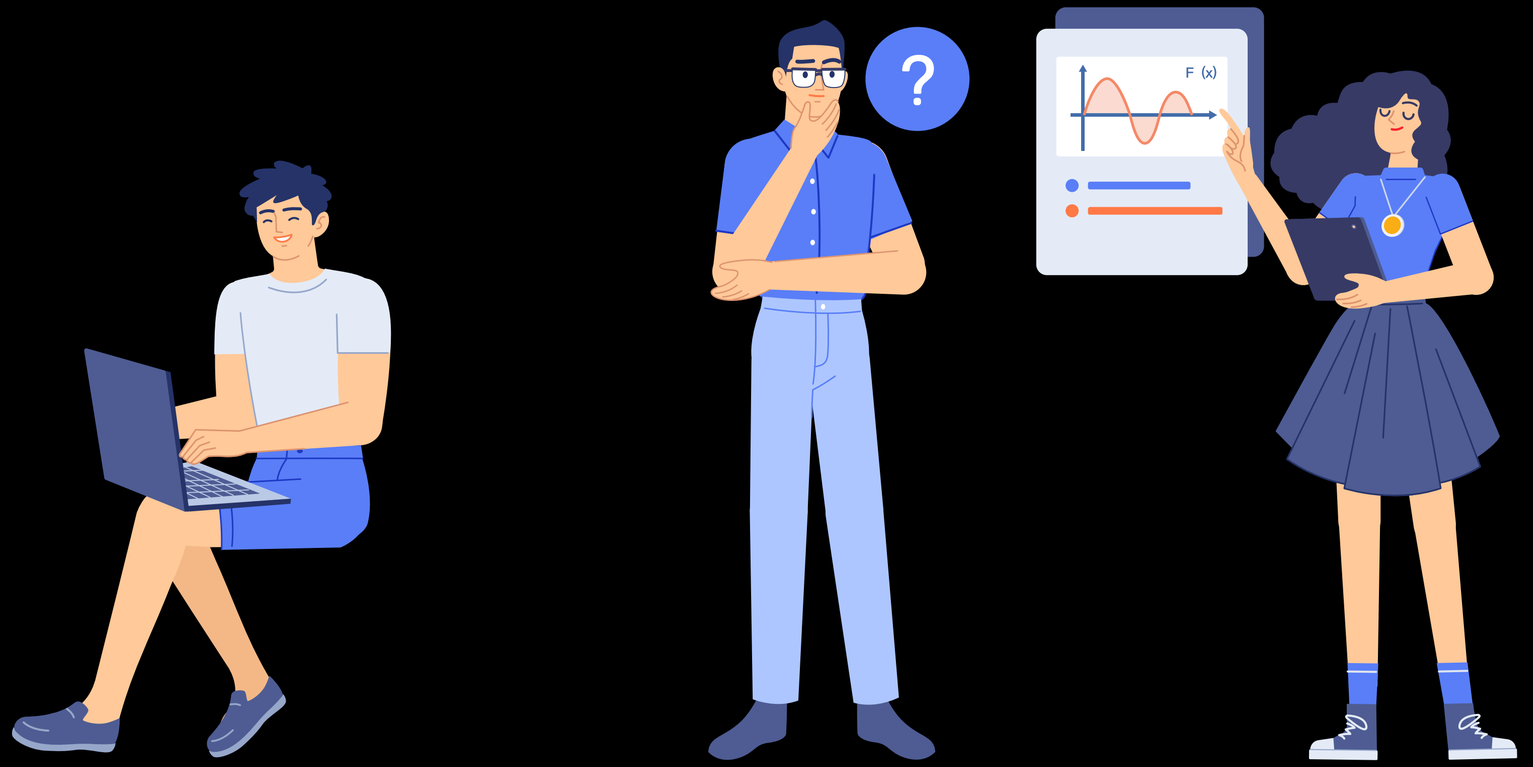
难道没有 可视化辅助 的 可视化研发 吗?

AVA.autoChart(container, data, {...});

视频链接：https://www.bilibili.com/video/av82223118/

高端场景 – 沟通之痛

程序员追着打的设计



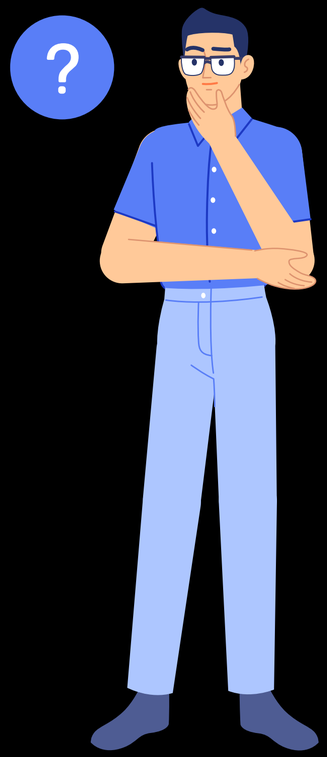
画 得出来就一定能 开发 出来吗?

ChartCube

视频链接：https://www.bilibili.com/video/av82223419/

没数场景 – 造数之痛

图形背后的数据长啥样



autoChart(container, []);

初始化 > 图表类型

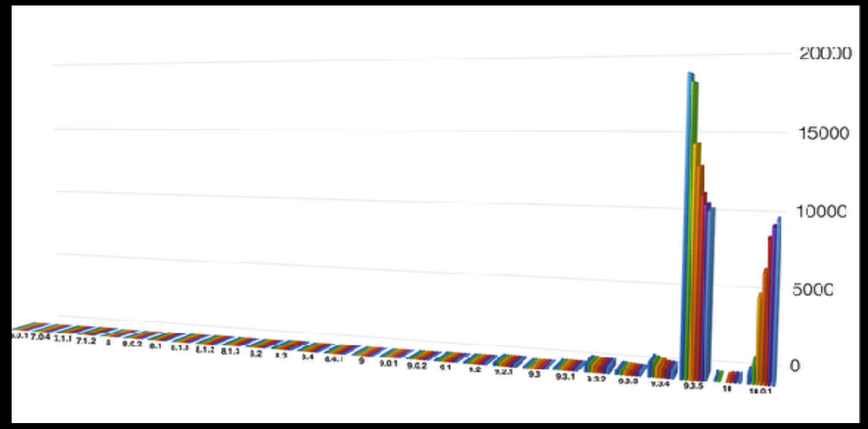
视频链接：https://www.bilibili.com/video/av82223498/

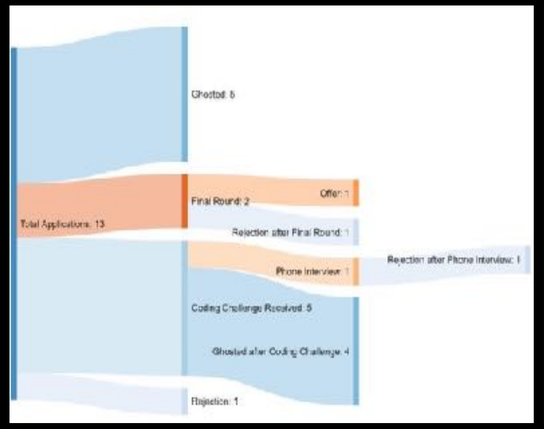
修罗场景

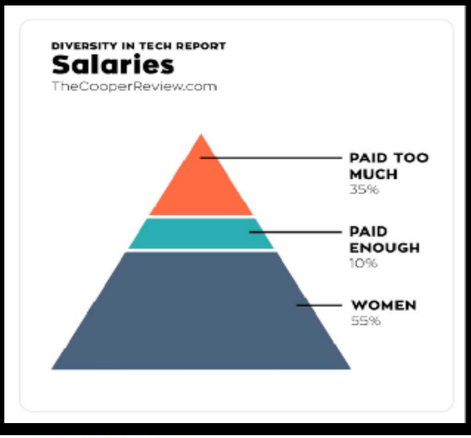
我接的都是一句话需求.....

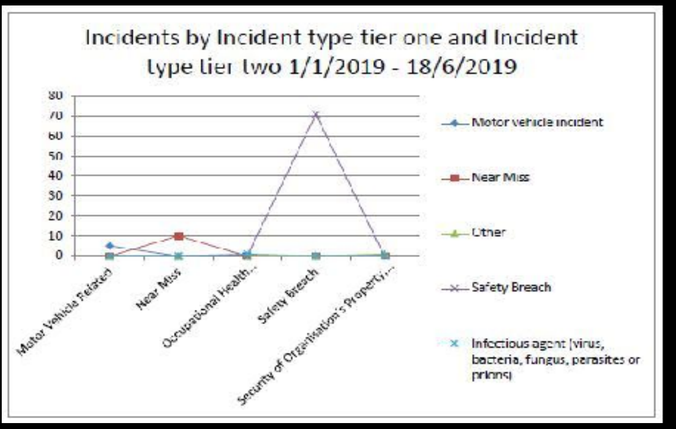
autoChart(container, []);

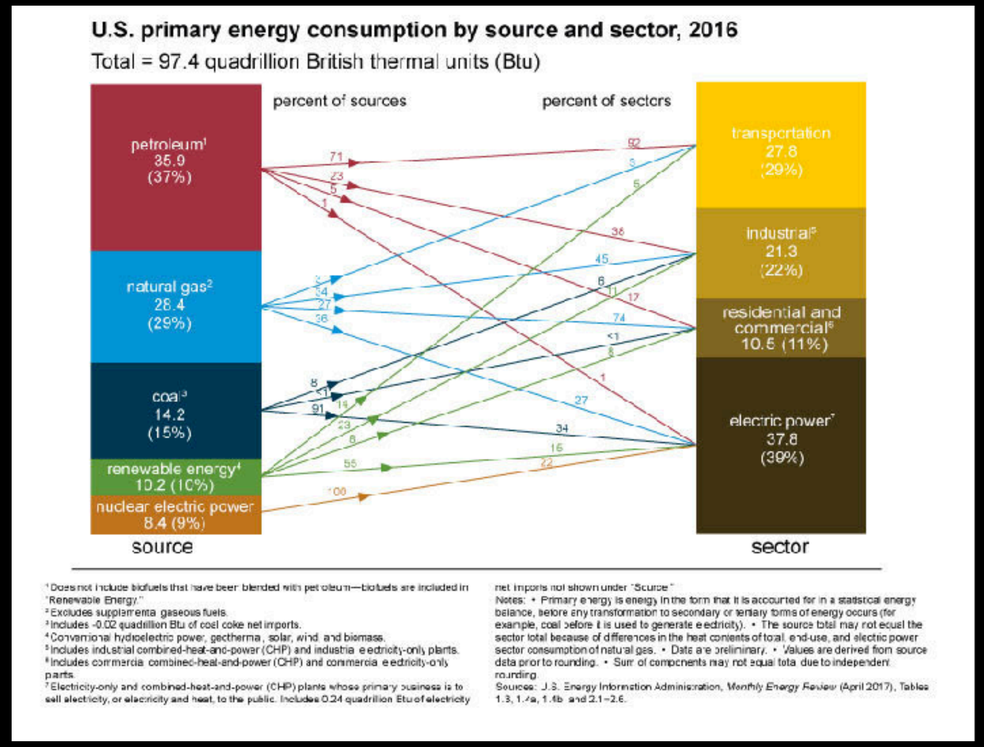
初始化 > 模拟数据

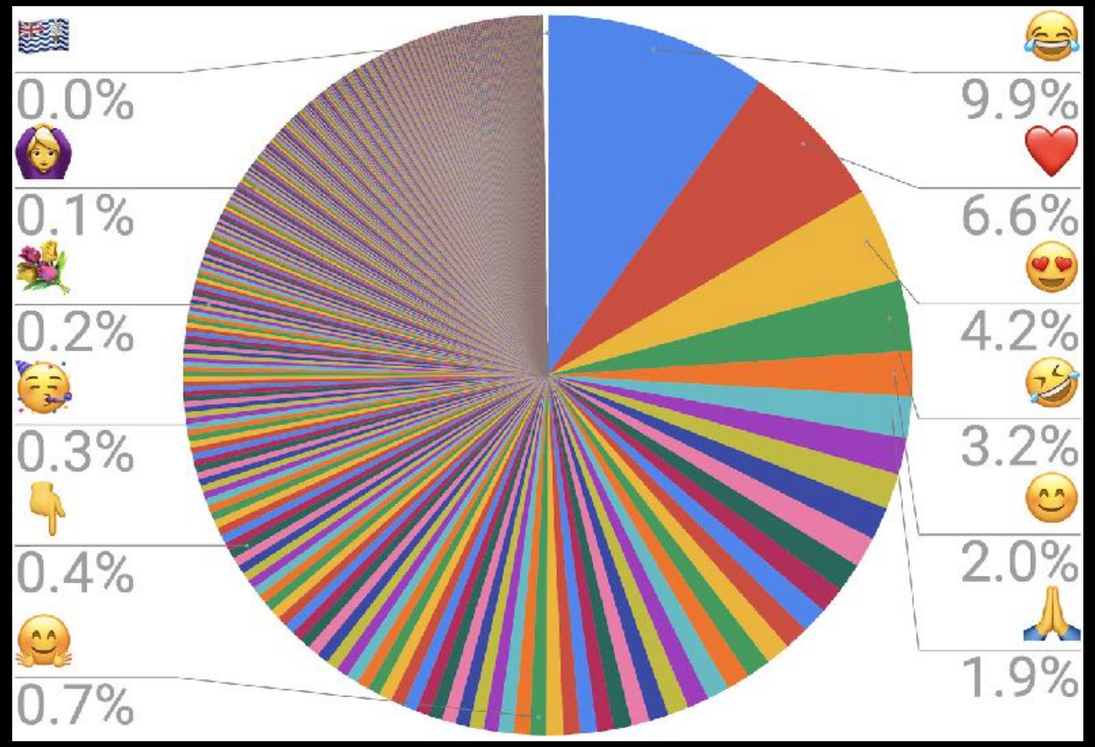


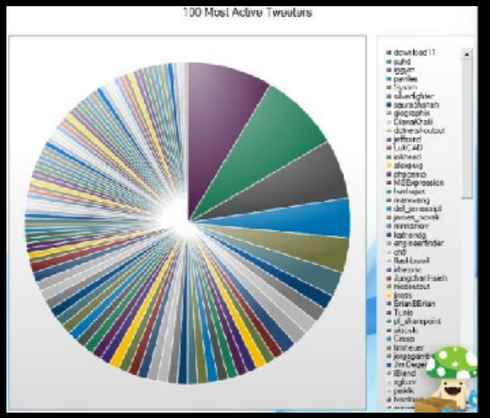


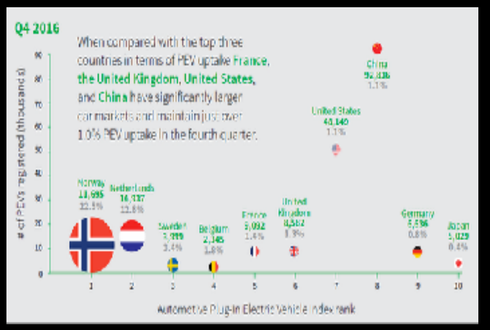










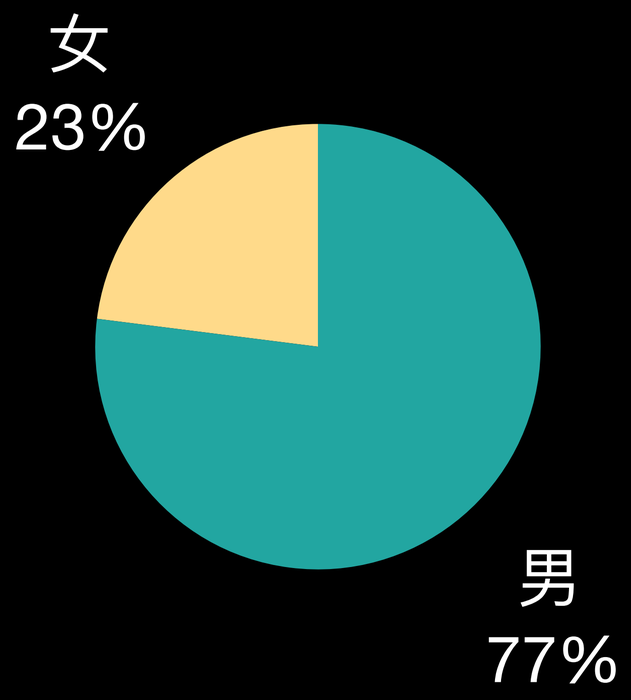


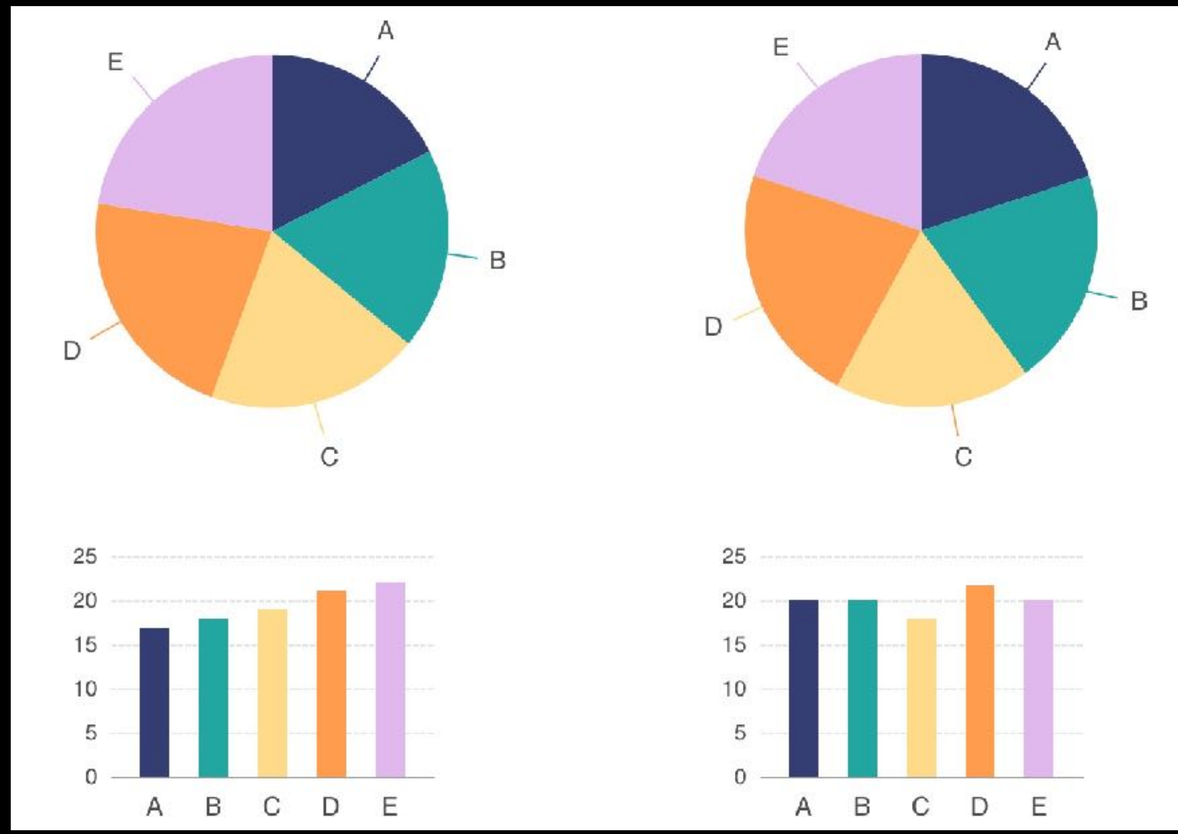
✗ 错误案例

每种图表都有她适用的场景

饼图

- 突出占比
- 分类有限
- 差异明显





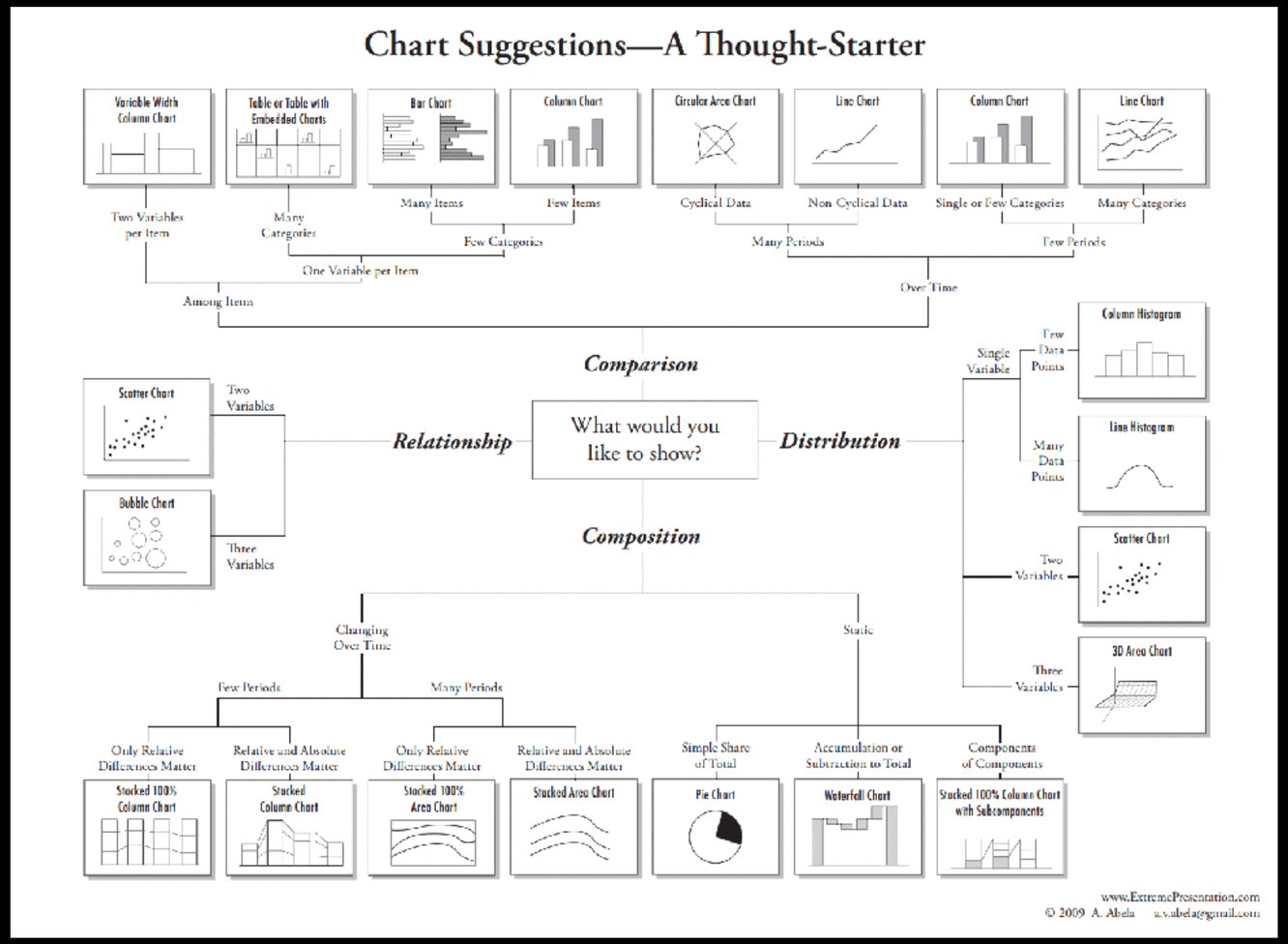
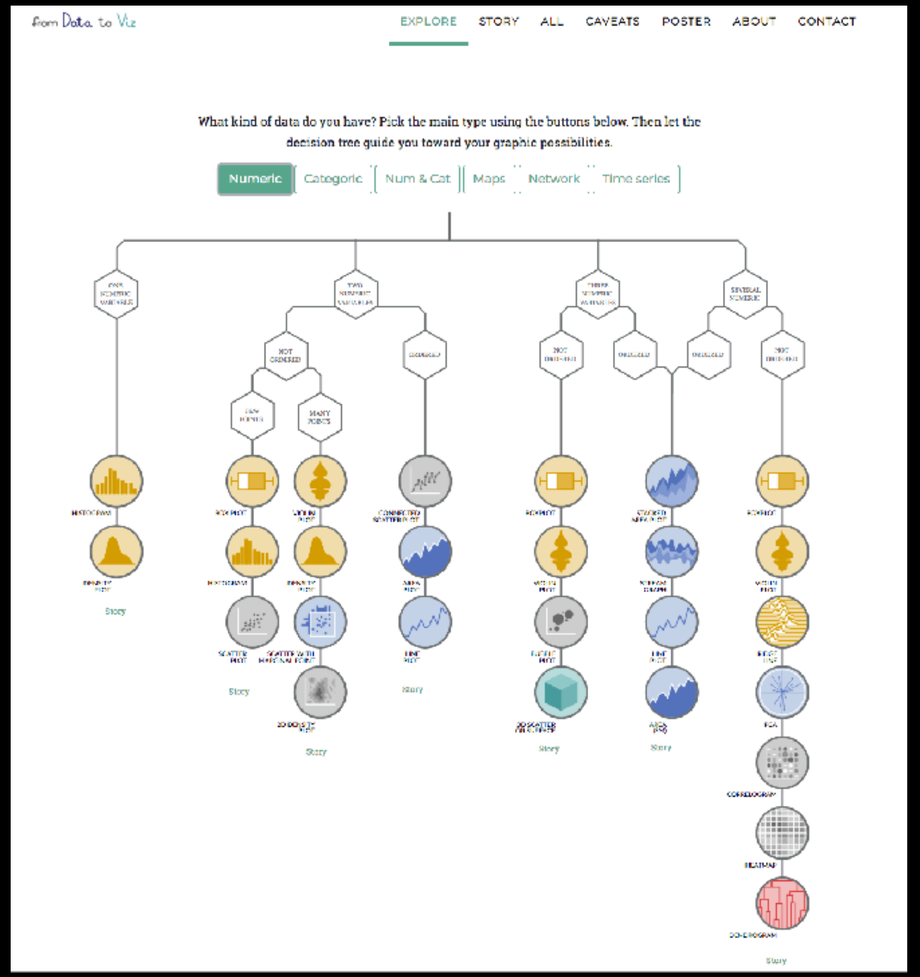
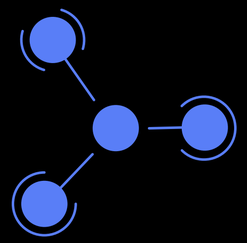


Chart Suggestions—A Thought-Starter
Variable Width Column Chart
Table or Table with Embedded Charts
Bar Chart
Column Chart
Circular Area Chart
Line Chart
Column Chart
Line Chart
Two Variables per Item
Many Categories
Many Items
Few Items
Cyclical Data
Non Cyclical Data
Single or Few Categories
Many Categories
Few Categories
Many Periods
Few Periods
One Variable per Item
Over Time
Among Items
Comparison
Scatter Chart
Two Variables
Relationship
What would you like to show?
Distribution
Bubble Chart
Three Variables
Composition
Single Variable
Few Data Points
Many Data Points
Two Variables
Three Variables
Column Histogram
Line Histogram
Scatter Chart
3D Area Chart
Changing Over Time
Static
Few Periods
Many Periods
Only Relative Differences Matter
Relative and Absolute Differences Matter
Only Relative Differences Matter
Relative and Absolute Differences Matter
Simple Share of Total
Accumulation or Subtraction to Total
Components of Components
Stacked 100% Column Chart
Stacked Column Chart
Stacked 100% Area Chart
Stacked Area Chart
Pie Chart
Waterfall Chart
Stacked 100% Column Chart with Subcomponents
www.ExtremePresentation.com
© 2009 A. Abela a.v.abela@gmail.com

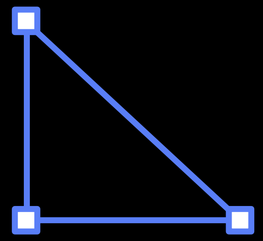

from Data to Viz
EXPLORE
STORY
ALL
CAVEATS
POSTER
ABOUT
CONTACT
What kind of data do you have? Pick the main type using the buttons below. Then let the decision tree guide you toward your graphic possibilities.
Numeric
Categoric
Num & Cat
Maps
Network
Time series

视频链接：https://www.bilibili.com/video/av82223577/

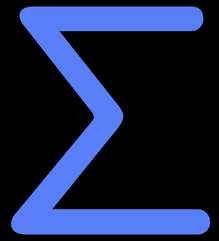
原理



knowledge
知识库



rule
推荐规则



advisor
推荐器

```
{
  "Line" : {
    "name" : "Line"
    "definition" : "Use line segments to show changes in data in a ordinal dimension."
    "purpose" : [
      0 : "Comparison"
      1 : "Trend"
    ]
    "coordinateSystem" : "Cartesian2D"
    "graphicType" : ""
    "shape" : [
      0 : "Lines"
    ]
    "dataPrerequisites" : [
      0 : {
        "minQty" : 1
        "maxQty" : 1
        "fieldConditions" : [
          0 : "Time"
          1 : "Ordinal"
```


knowledge

知识库

图表类型

- 名称
- 定义
- 分析目的
- 所需数据
- 涉及通道
- ……

diff-pie-sector SOFT

> *Difference should be big enough for pie sectors.*

饼图类（或，以角度或弧长映射数值变量的）图表，每个区块的数值大小应有较为明显的差异，差异性越大越好。

打分算法

> 公理：当 n 个数字的和相等时，数字越接近，乘积越大。
> 举例：1+9=10; 5+5=10; 9=1*9<5*5=25

设饼图由 n 个区块（扇形）组成，每个区块对应的数值为 $v_i(i=1,..,n)$。

已知饼图所有区块的和最终换算100%或 360 度，统一归为 1。等比例缩放后每个区块对应的数值为 $s_i = \frac{v_i}{\sum_{i=1}^{n} v_i}$

且有 $\sum_{i-1}^{n} s_i = 1$ 。

计算这些缩放后数值的乘积，记为 $P_s = \prod_{i=1}^{n} s_i$ 。

根据公理，当这 n 个区块的数值完全一样（$s_i = \frac{1}{n}$）时，他们的乘积最大。最大乘积为 $P_{max} = \prod_{i-1}^{n} \frac{1}{n} = (\frac{1}{n})^n$。

则本规则对应分数 r 为：

$$r = \frac{|P_{max} - |P_s||}{P_{max}}$$

当各区块的数值完全一样（区分度极端小）时，$r = 0$ ；

当各区块的数值差异越大时，P_s 越趋向于 0， r 也越大（越趋向于 1）。

rule
推荐规则

HARD

- 字段属性检查
- 分析目的检查
- 冗余字段检查
- ……

SOFT

- 分类数量评估
- 数值差异评估
- 相关性评估
- ……

rule

推荐规则

以柱状图为例

- M(v) 是数据和图表类型的匹配质量分数
- d(X)表示 x 轴字段中的不重复值个数（有几个柱子）

$$\frac{\pi^2}{6} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^2}$$

$$v_i (i = 1, .., n$$

$$P_{max} = \prod_{i=1}^{n} \frac{1}{n} = (\frac{1}{n})^n$$

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{2\pi i \xi x}\, d\xi$$

$$\mathbf{M}(v) = \begin{cases} 0 & |d(X)| = 1 \\ 1 & 2 \le |d(X)| \le 20 \\ \frac{20}{|d(X)|} & |d(X)| > 20 \end{cases}$$

$$X_k = \frac{1}{N} \sum_{n=0}^{N-1} x_n e^{i 2\pi k \frac{n}{N}}$$

$$r = \frac{|P_{max} - |P_s||}{P_{max}}$$

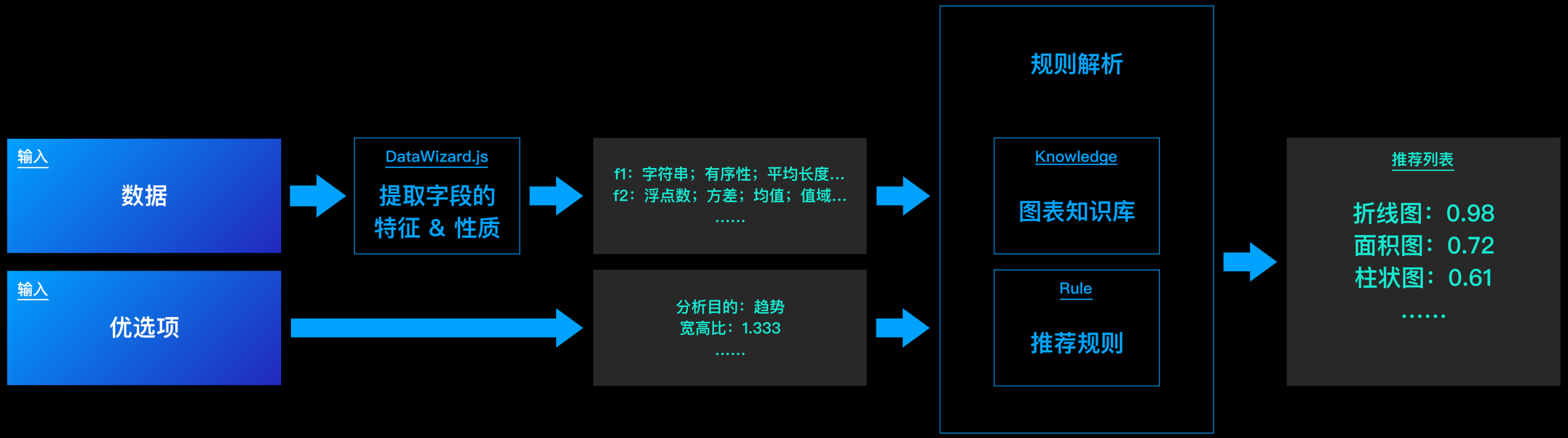
$$s_i = \frac{v_i}{\sum_{i=1}^{n} v_i}$$

$$\sum_{i=1}^{n} s_i$$

Luo, Yuyu, et al. "DeepEye: Towards Automatic Data Visualization." 2018 IEEE 34th International Conference on Data Engineering (ICDE). IEEE, 2018.

advisor – 解析器

工作流程



智能可视化的未来

研发态：全智能图表推荐，代码极简

100+ 行原始代码

50+ 行半自动代码

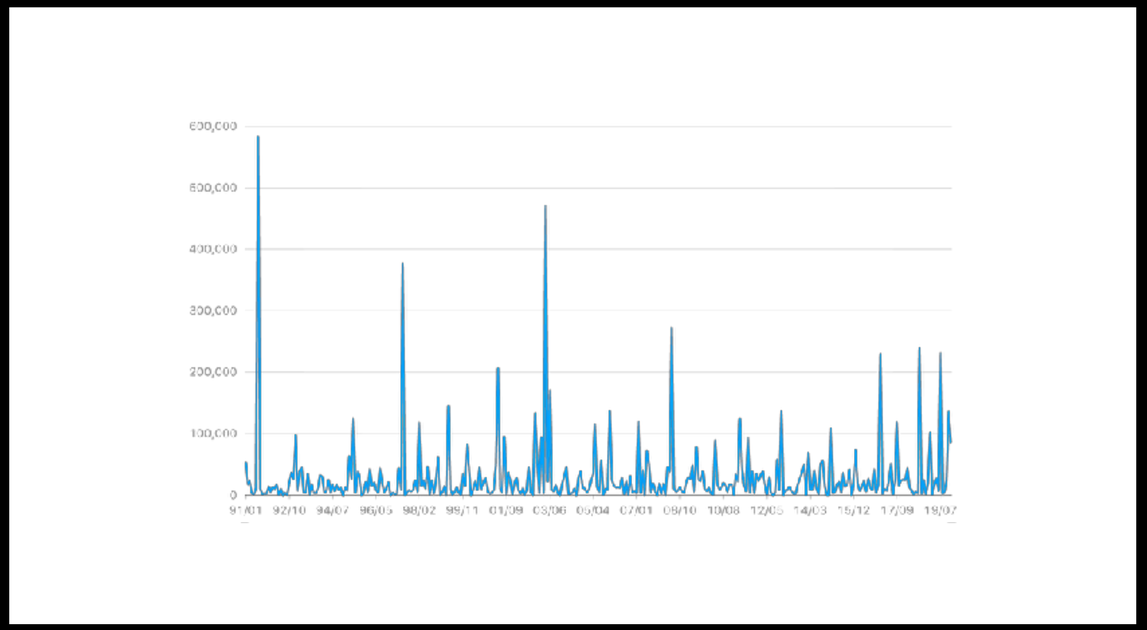
1 行智能代码

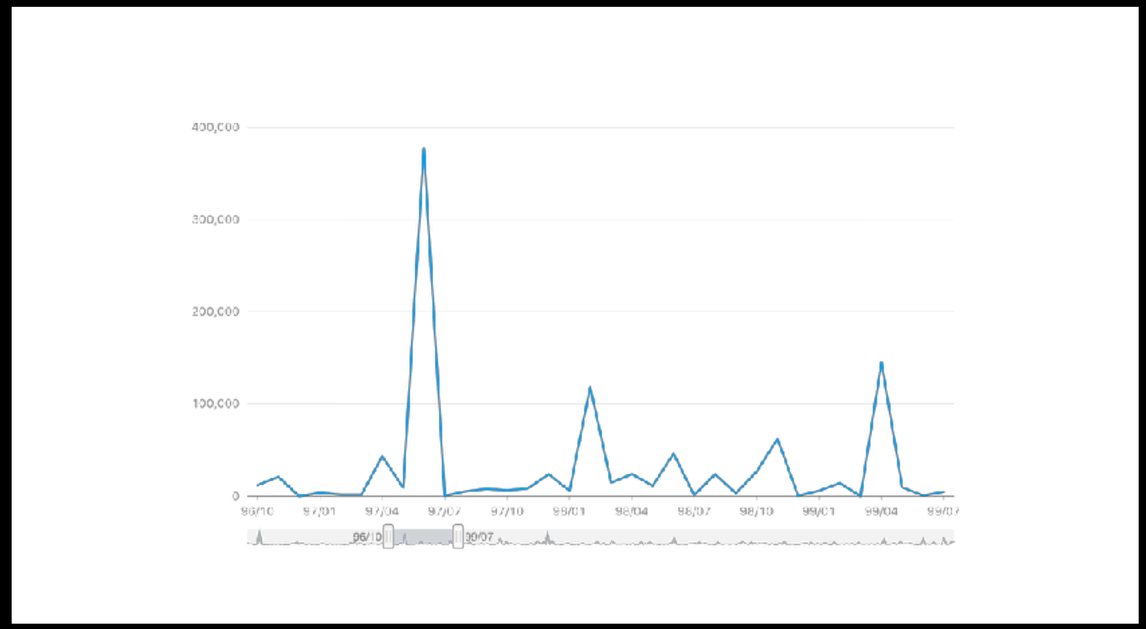
```
fetch('../data/baby-names.json')
  .then(res => res.json())
  .then(data => {
    const { DataView } = DataSet;
    const dv = new DataView().source(data)
      .transform({
        type: 'fill-rows',groupBy: [ 'name' ], orderBy: [ 'year' ]
      })
      .transform({
        type: 'impute', field: 'n', method: 'value', value: 0
      })
      .transform({
        type: 'aggregate', fields: [ 'n' ], operations: [ 'sum' ],
        groupBy: [ 'year', 'name' ], orderBy: [ 'year' ],
        as: [ 'count' ]
      });
    const chart = new G2.Chart({
      container: 'container',
      forceFit: true,
      height: 500,
      animate: false,
      padding: [ 20, 140, 60, 50 ]
    });
    chart.source(dv, {
      year: {tickInterval: 10}
    });
    chart.axis('count', {
      line: {
        lineWidth: 1,
        stroke: '#BFBFBF'
      },
      tickLine: {
        length: 8,
        stroke: '#ddd'
      },
      grid: null
    });
    chart.area()
      .position('year*count')
      .adjust([ 'stack', 'symmetric' ])
      .color('name')
      .shape('smooth')
      .opacity(1);
    chart.render();
  });
```

```
fetch('../data/baby-names.json')
  .then(res => res.json())
  .then(data => {
    const container = document.getElementById('mountNode');
    autoChart(container, data, {
      toolbar:true,
      development: true,
      config: {
        configs: {
          title: {
            visible: true,
            text: '销售趋势'
          },
          legend: { position: 'right-top' },
          smooth: true,
          point: {
            visible: true,
            size: 7,
            shape: 'diamond'
          },
          xField: 'f1',
          yField: 'f2',
        },
        type: 'Line',
      }
    });
  });
```

```
fetch('../data/baby-names.json')
  .then(res => res.json())
  .then(data => {
    autoChart(document.getElementById('mountNode'), data);
  });
```

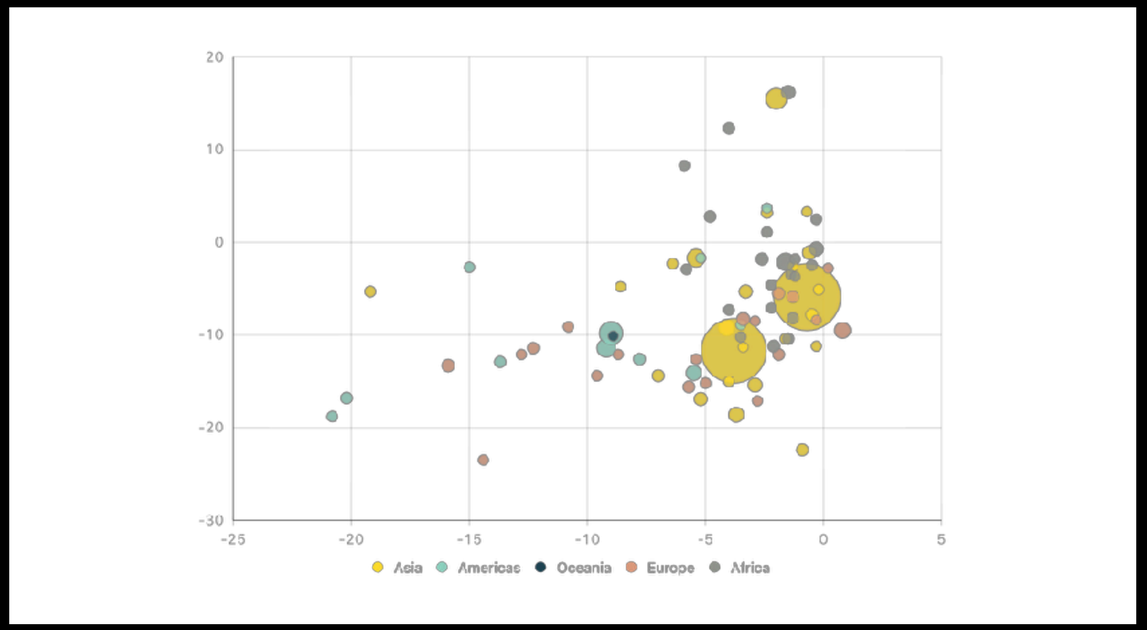
阅读态：交互分析组件智能增强

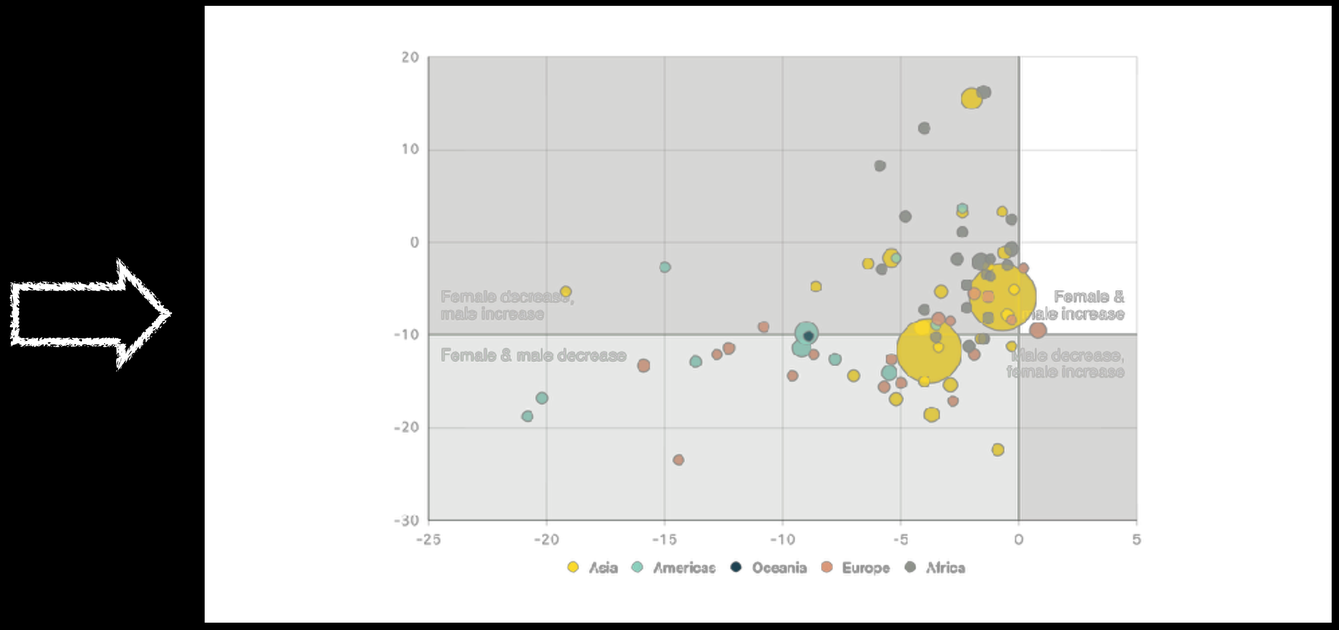




富交互组件动态按需启用

分析态：运行时的可视分析智能辅助





自动识别出趋势、回归、奇异点等统计特征，自动演示模式的 Storytelling 能力

AntV

让人们在数据世界里获得视觉化思考能力

$$\vec{\mathbf{A}} = \begin{bmatrix} Alibaba \\ AI \\ Automated \\ Augmented \end{bmatrix}$$

$$\vec{A}VA$$

AVA 钉钉交流群

即将开源！抢先 star，快人一步： https://github.com/antvis/AVA

蚂蚁金服体验科技
ANT EXPERIENCE TECHNOLOGY

3rd SEE Conf

蚂蚁金服体验科技大会，初衷是希望设计与技术能在碰撞中彼此融合，SEE 是 Seeking Experience and Engineering 的缩写，同时 SEE 也代表着“看见”，希望技术能看见设计的价值，也希望设计能看见技术的力量，在彼此看见中互相融合成长，一起让世界更美好。

语雀专栏 | https://www.yuque.com/seeconf
SEE Conf 官网 | https://seeconf.antfin.com/
参与知乎互动，赢下届门票 | https://www.zhihu.com/question/363807174

SEE Conf

更多议题

蚂蚁金服体验科技
ANT EXPERIENCE TECHNOLOGY

体验科技与好的产品
玉伯（蚂蚁金服 体验技术部负责人）

基于地域文化的设计创新
何人可（湖南大学设计艺术学院院长）

Ant Design 4.0：创造快乐工作
林外（蚂蚁金服 高级体验设计专家）
线丝（蚂蚁金服 高级创意设计师）

决策机构体验科技：数字驾驶舱
逸达（蚂蚁金服 前端技术专家）
可言（蚂蚁金服 高级产品经理）
十喜（蚂蚁金服 高级体验设计师）

智能可视化体系 AVA
步茗（蚂蚁金服 数据技术专家、AVA 负责人）
璆鸣（蚂蚁金服 前端技术专家、DataWizard 负责人）

使用 React 开发小程序 - Remax
边柳（蚂蚁金服 高级前端工程师、Ant Design 核心贡献者）

Evolution: Serverless For Frontend - 探索下一代 Node 研发模式
天猪（蚂蚁金服 高级前端专家、Egg.js 核心开发者）

云凤蝶可视化搭建的推导与实现
江木（蚂蚁金服 高级前端工程师、antd-mobile 核心开发者）

“云”端的语雀 —— 用 JavaScript 全栈打造商业级应用
不四（蚂蚁金服 高级前端技术专家，语雀产品技术负责人）

蚂蚁金服 Web 3D 技术探索之路
烧鹅（蚂蚁金服 前端技术专家、Oasis 3D 引擎负责人）

精雕细琢，打造极致可视化体验
逍为（蚂蚁金服 高级前端工程师、AntV 核心贡献者）

蚂蚁海外本地化设计
竹摇（蚂蚁金服 高级体验设计专家）

让价值被发现：如何在 B 端产品做增长？
覃一（蚂蚁金服 高级体验设计师）
瀚雅（蚂蚁金服 高级体验设计师）

围绕应用生命周期的企业级产品设计策略
壹乐（蚂蚁金服 高级体验设计师）

普惠金融体验设计创新思路：参与感对话设计
姚维（蚂蚁金服 体验设计专家）

JCD 思维如何驱动复杂系统设计
今辰（蚂蚁金服 体验设计专家）

资产的秩序之美：通过模式化的方法构建设计资产的内在一致性
吾笙（蚂蚁金服 高级体验设计师）

解放图形化设计生产力 — HiTu
线丝（蚂蚁金服 高级创意设计师）