

可视化叙事

信息可视化

石洋 副教授

<https://idvxlab.com>

同济大学

课程大纲

- 为什么用数据来讲述故事?
- 什么是数据故事?
- 数据：数据故事的基础
- 叙事：数据故事的结构
- 视觉：数据故事的场景

课程大纲

- **为什么用数据来讲述故事?**
 - 人是叙事性动物
 - 讲故事比摆数据更有效
 - 数据故事的三种关键元素
 - 用数据故事促进行动
 - 数据故事中叙事和可视化的力量

人是叙事性动物

- 人类是叙事性动物，讲故事是人类的天性
 - 讲故事是传授知识、文化标准、灌输道德价值观和建立社会纽带的有效方式，对人类群体生活十分重要
 - 讲故事的方法历经千年变化，从篝火故事的口口相传、洞穴壁画的精美描绘到书面语言文字记录再到如今的数字时代的大众媒体、社交平台
 - 向人类有效传达新的见解，需要学习利用讲故事的力量



讲故事比摆数据更有效

- Storytelling is the most powerful way to put ideas into the world today.
讲故事是将想法带入当今世界最有力的方式。 — Robert McKee , screenwriting expert and author
- You can't beat a story with fact. You can only beat it with a better story.
事实无法打败故事，比故事更好的是更好的故事。 — Shawn Callahan, "Putting Stories to Work" author



故事更具有情绪感染力



故事更容易理解



故事更容易记忆



故事更具有说服力

讲故事比摆数据更有效

- 2004年，卡内基梅隆大学开展“拯救儿童”慈善捐款实验
 - 宣传册版本一：基于统计学的方法。展示关于非洲各个国家的人受食物短缺和干旱影响的统计数据。比如，强调大暴雨导致赞比亚玉米产量下降42%，使得三百万赞比亚人面临饥荒。
 - 宣传册版本二：基于故事的方法。讲述来自非洲马里的7岁小姑娘Rokia的故事，她的家庭非常贫困，常常吃了上顿没有下顿。
 - 实验结果：第二组收到的捐款是第一组的一倍多，一个代表性的非洲女孩的故事比数百万非洲受苦受难者的统计数据更能号召捐款。
- 这项研究表明，一个简单的故事在情感上与观众建立联系，比统计数据更有说服力。

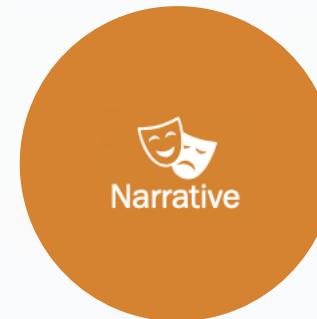
数据故事的三种关键元素

- 数据故事包括三种元素：数据、叙事、可视化

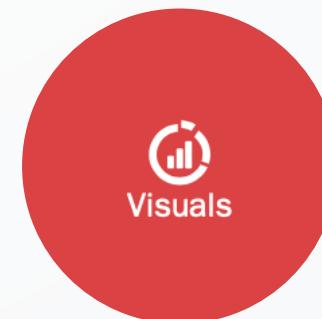
- **数据 (Data)** : 数据故事的基础和起源，也是叙事和视觉的基础。
- **叙事 (Narrative)** : 是利用各种方法对在时间及逻辑关系上相互关联的事件进行再现的过程，是人们相互沟通、传递信息的重要手段。
- **可视化 (Visuals)** : 视觉效果呈现。



发人深省的统计数据



引人注目的叙事



令人惊叹的可视化效果

数据故事的三种关键元素

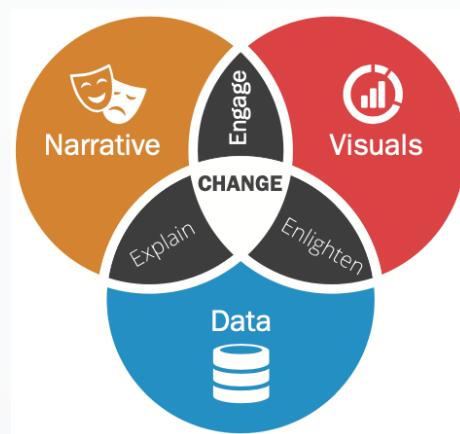
- **数据、叙事、视觉的组合模式**

- **解释：数据+叙事**，叙事元素为数据赋予了结构，向观众解释数据中发生了什么。
- **阐明：数据+可视化**，可视化可帮助呈现隐藏在数据表中的有趣的数据模式和异常值，向观众展示没有图表或图形就看不到的见解。
- **吸引注意：叙事+可视化**，可以吸引观众注意并帮助他们理解，就像插图故事书、电影一样吸引观众兴趣并沉浸于此。



数据故事的三种关键元素

- 数据、叙事、可视化的组合模式
 - 建立有效的数据故事需要融合**数据+叙事+可视化**，帮助观众建立新的认知，说服他们采取行动！



有效
的数据
故
事

建立新的认知

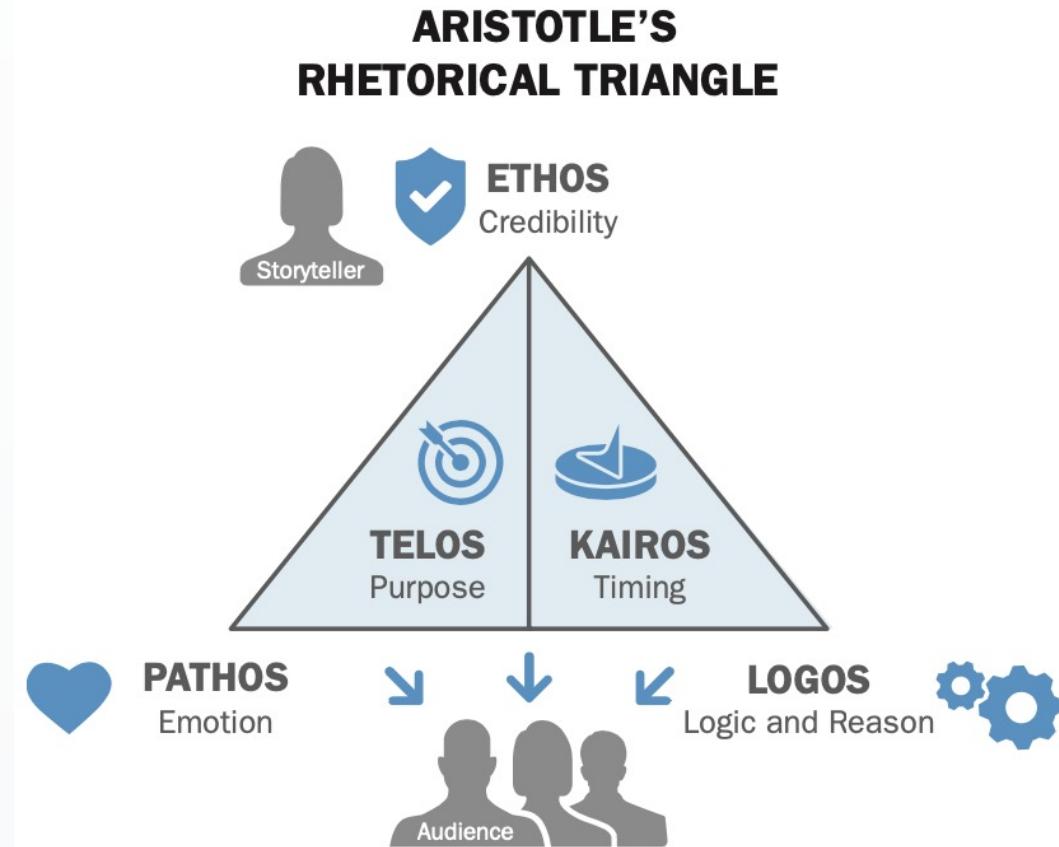


采
取
行
动

数据故事的三种关键元素

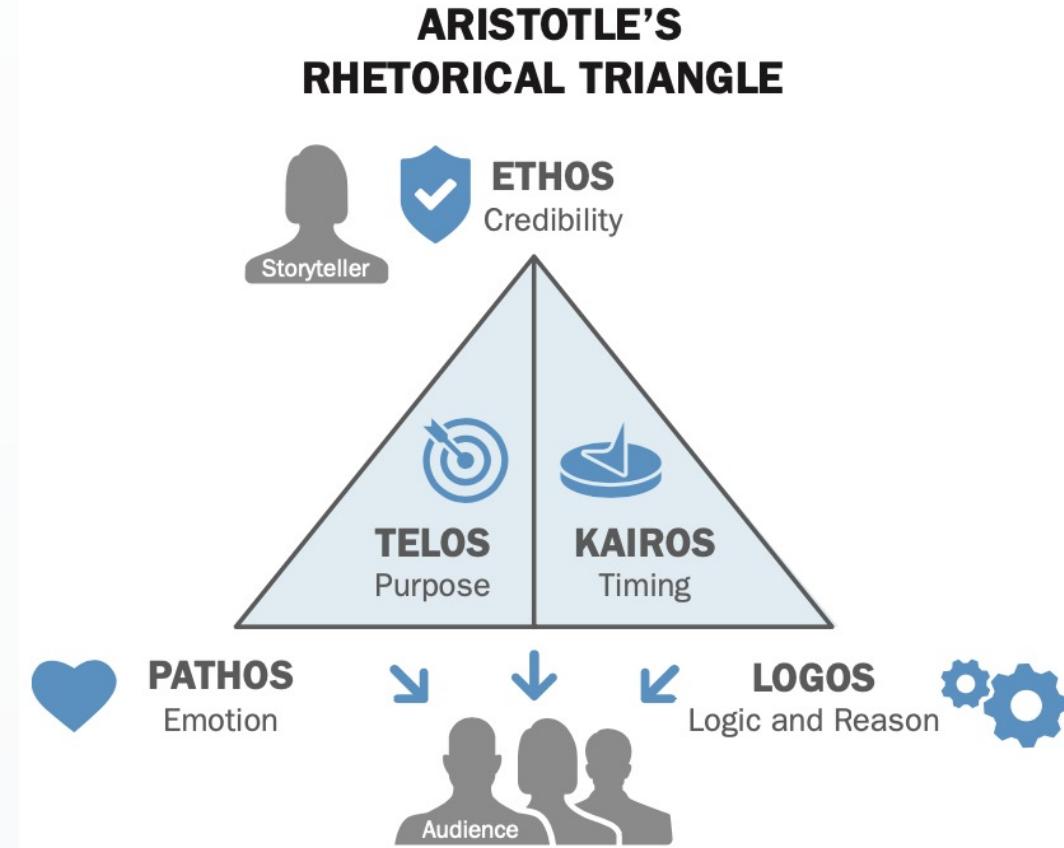
- **亚里士多德的修辞三角**

- 人格说服 (Ethos喻德) : 说服者自身的可信度说服他人
- 情感说服 (Pathos喻情) : 情感上说服他人
- 逻辑说服 (Logos喻理) : 逻辑上说服他人
- 说服的目的 (Telos) : 明确沟通的重点和目的
- 说服的时间 (Kairos) : 在正确的时间说服他人



数据故事的三种关键元素

- 数据故事需要这五个说服方式，使其成为可用的最强大的沟通形式之一。
 1. 首先从人格说服的角度，说服者本身和数据都是可信的（ETHOS）。
 2. 其次，因为数据故事是基于事实和数字的，传达的信息需要有逻辑性，使观众得以理解（LOGOS）。
 3. 当将数据编织成令人信服的故事时，情感上的共鸣让信息更触动人心（PATHOS）。
 4. 在数据故事中加入沟通的重点和目的（TELOS）。
 5. 在正确的时间（KAIROS）与正确的受众分享相关数据故事，让信息可以成为变革的强大催化剂，让人们做出行动。



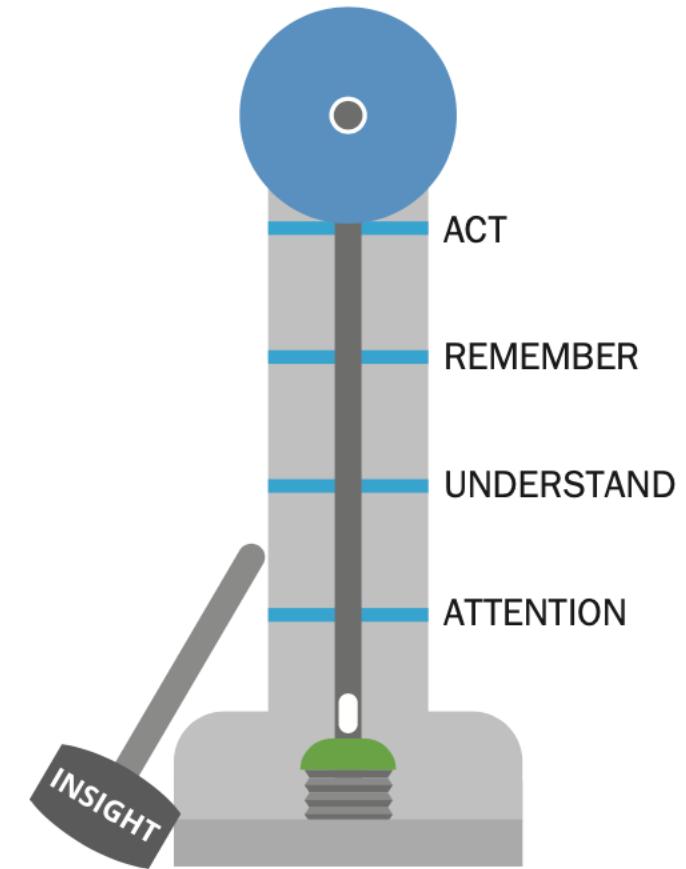
用数据故事促进行动

- 当你想和他人分享新见解时，有4个关键的沟通目标

- 吸引注意力 (Attention)：新的见解可以在有限的时间和繁杂的信息环境中吸引观众注意。
- 理解内容 (Understand)：确保观众清晰且实在的理解了你传达的新见解的内容。
- 记住内容 (Remember)：新的见解与观众有足够的共鸣并保留在观众的脑海中。
- 采取行动 (Act)：激励观众根据新的见解采取行动。

如何敲响顶部的动作钟？

HOW STRONG IS YOUR DATA COMMUNICATION?



数据故事中叙事和可视化的力量

- 叙事的力量
 - ✓ 加深记忆

单词

merciless
reckless
endless
hopeless
doubtless
helpless
useless
priceless

故事

纣王被伐

纣王生性merciless（残忍的），迷恋妲己reckless（不计后果的），引发后患endless（无穷尽的），百姓感到hopeless（无望的），揭竿起义doubtless（无疑的），纣王被困helpless（无助的），悔恨愧疚useless（无用的），民心才是priceless（无价的）。

数据故事中叙事和可视化的力量

- **叙事的力量**

- ✓ 加深记忆
- ✓ 叙事可以产生强大说服作用

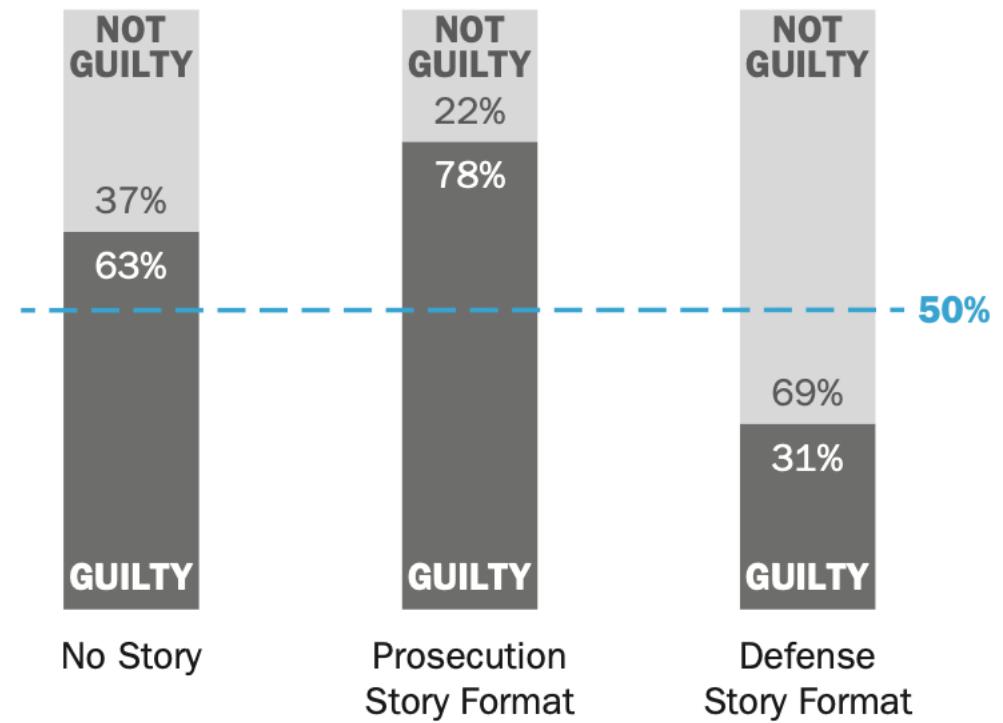
1983年陪审团对一家酒吧斗殴致死案存在争议。

研究者在实际审判中听取了双方证据事实后，
63%的人认为被告犯谋杀罪。

当检察官以故事形式提供相同证据时，**78%**的
人认为被告犯谋杀罪。

当被告辩护律师以故事形式分享证据时，**31%**
的人认为被告犯谋杀罪。

**STORY FORMAT INFLUENCES
DECISION MAKING**



数据故事中叙事和可视化的力量

- **可视化的力量**

- ✓ 图像比文字更难忘——图优效应

1870年罗切斯特大学的研究人员通过一次性接触2500多张图片的实验验证图优效应，每张图片接触时间5-10秒。长达三天后，参与者可以回忆起90%以上的图像。



图片来源<https://primeyourpump.com/2019/05/14/picture-superiority-effect/>

数据故事中叙事和可视化的力量

• 可视化的力量

- ✓ 图像比文字更难忘
- ✓ 图像比文字更容易理解

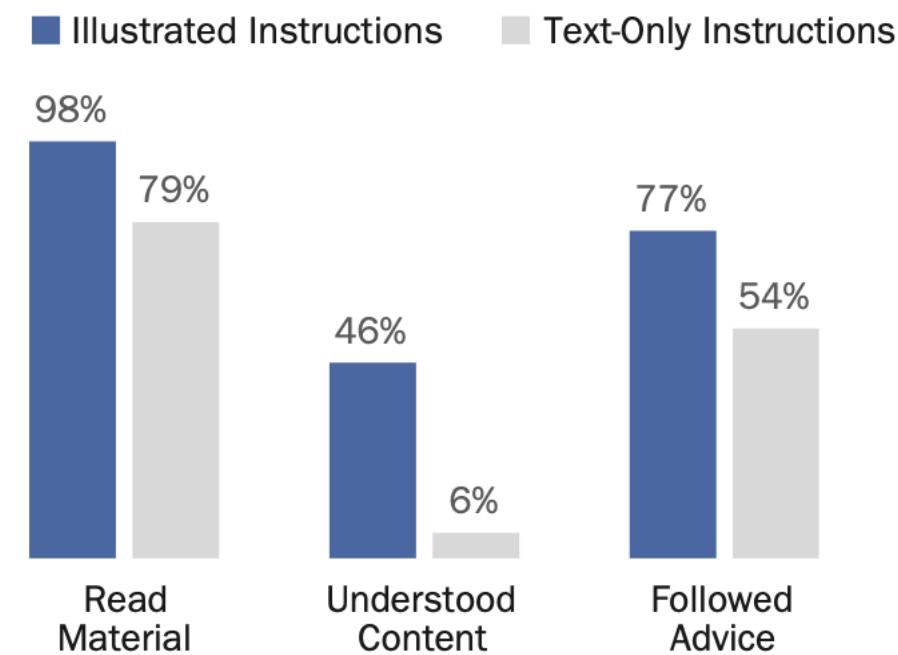
1996年，密歇根州立大学的研究人员向400名前往急诊室治疗伤口的患者提供了两种关于伤口护理的说明形式，文字说明或插图说明。

收到插图说明的患者相比收到文字说明的患者
更有可能阅读说明 (98%: 79%)

收到插图说明的患者相比收到文字说明的患者
更理解相关内容 (46%: 6%)

收到插图说明的患者相比收到文字说明的患者
更按照伤口护理建议采取行动 (77%: 54%)

ILLUSTRATED INFORMATION OUTPERFORMS TEXT-ONLY INFORMATION



数据故事中叙事和可视化的力量

• 可视化的力量

- ✓ 图像比文字更难忘
- ✓ 图像比文字更容易理解
- ✓ 图像比文字更有说服力

2014年，康奈尔大学的研究人员研究了图像对信息说服力的影响。他们测试了人们对一种可以增强免疫功能，减少感冒出现的药物是否会起作用的态度，分别用纯文字、文字+图表的形式去说明。

97%看到图表的人认为药物会起作用

68%看到纯文本的人会觉得药物会起作用。



图片来源<https://valnd.com/can-a-vitamin-be-a-drug/>

数据故事中叙事和可视化的力量

- 叙事和可视化的结合有助于构建有效的数据故事，为新见解提供被注意、被理解、被记住和激励人们采取行动的最佳机会，推动仅靠事实无法实现的行动。

2015年理查德·潘为了说服立法者通过强制接种麻疹疫苗法案，利用FRED Measles模拟器模拟麻疹爆发时的传播，州参议员马蒂·布洛克最初对该法案持怀疑态度，但在观看了他的麻疹爆发模拟故事后，认识到该法案对保护该州的公共卫生很重要。该法案最后顺利通过。

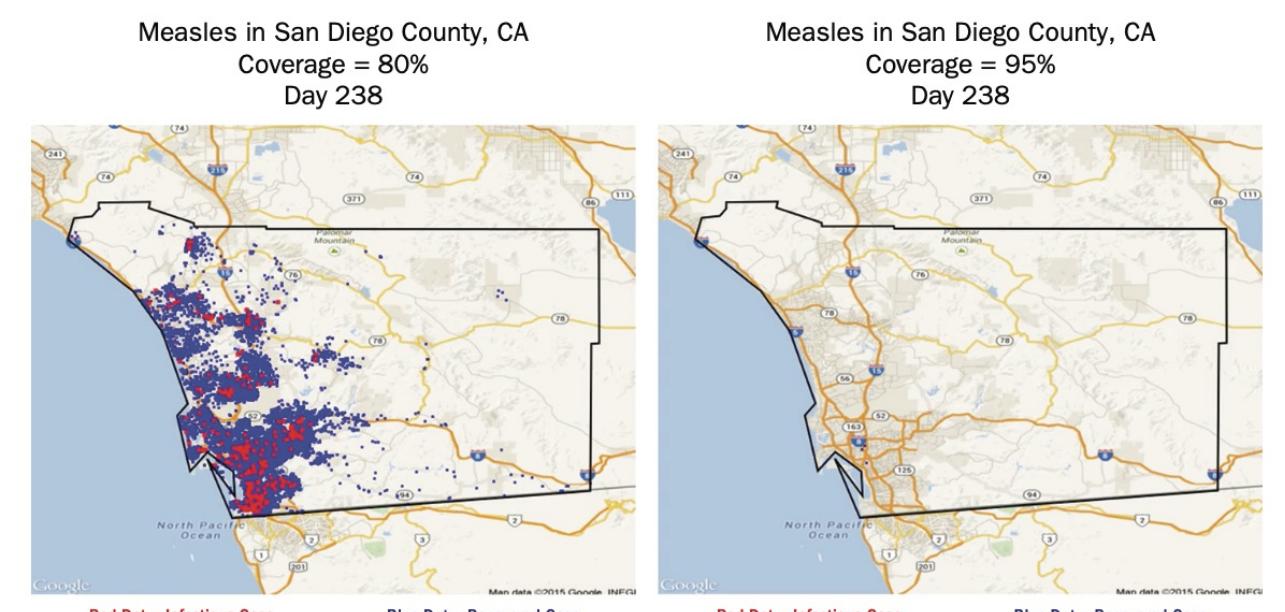


Figure 2.8 FRED Measles shows how a measles outbreak will expand exponentially within a county if there is only an 80% vaccination rate.
Source: Public Health Dynamics Laboratory, University of Pittsburgh.

课程总结

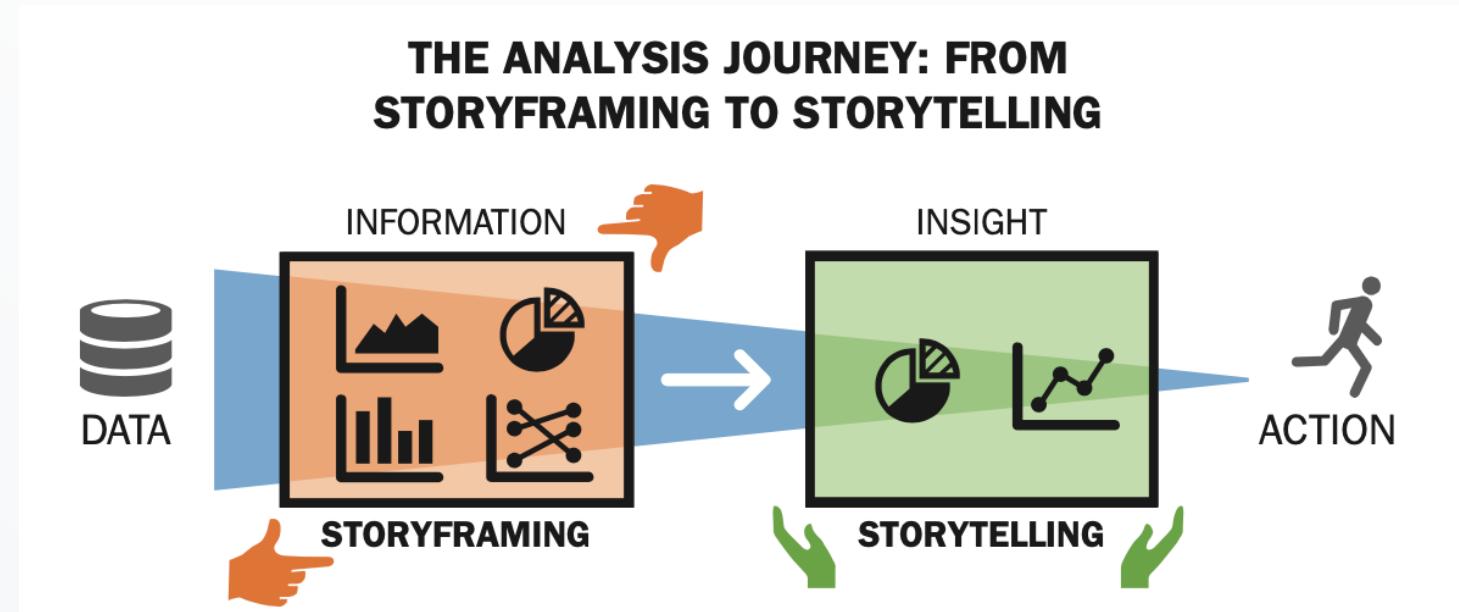
- 本节课
 - 人是叙事性动物
 - 讲故事比摆数据更有效
 - 数据故事的三种关键元素
 - 用数据故事促进行动
 - 数据故事中叙事和可视化的力量
- 下节课
 - 什么是数据故事？

课程大纲

- **什么是数据故事？**
 - 数据故事的六要素
 - 数据故事的体裁
 - 历史上具有叙事特征的可视化
 - 数据故事需要讲故事的人
 - 在讲述故事之前了解你的受众
 - 数据故事的应用领域
 - 数据故事的制作流程

数据故事

- **数据故事：使用叙事元素和解释性视觉来交流数据见解的一种结构化方法。**
 - **故事框架 (Storyframing) :** 通过故事框架将大量数据提炼成更有针对性的关键指标和维度，侧重于向观众提供有用的信息——这些信息不一定能转化为有意义的发现。
 - **讲故事 (Storytelling) :** 通过讲故事将重要发现或见解解释给他人，让目标受众对相关见解有扎实的了解，并让他们有所行动。



数据故事的六要素

• 新闻的倒金字塔模型

- 最重要的信息放在最前面，自上而下地从最重要的信息到不重要的细节。
- 优点：短时间快速传递信息。
- 缺点：放弃了叙事的情感力量。

• 文学故事的传统叙事模型

- 通过高潮逐步构建最重要的信息。
- 优点：保留了叙事的情感力量，观众可以更好地保留信息，并更有可能采取行动。

COMPARISON OF NEWS AND LITERARY STORY FORMATS

INVERTED PYRAMID



NARRATIVE STRUCTURE

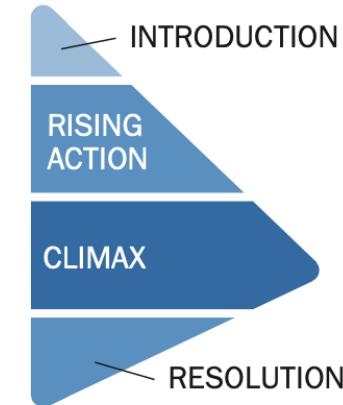


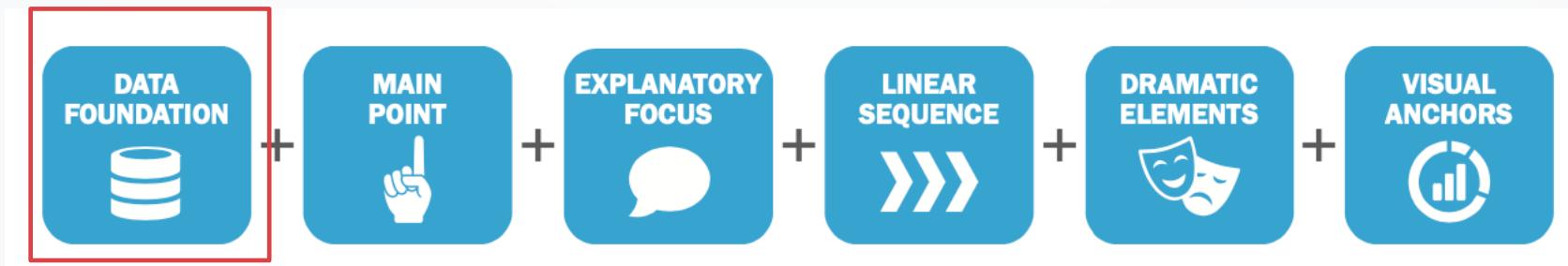
Figure 4.3 The inverted pyramid approach features the most important information at the beginning while a traditional narrative structure builds up to the most important information with a climax.

当目标是向观众解释见解并让观众沉浸其中，文学故事的叙事模型更适合作为数据故事的原型。

数据故事的六要素

• 要素一：数据基础

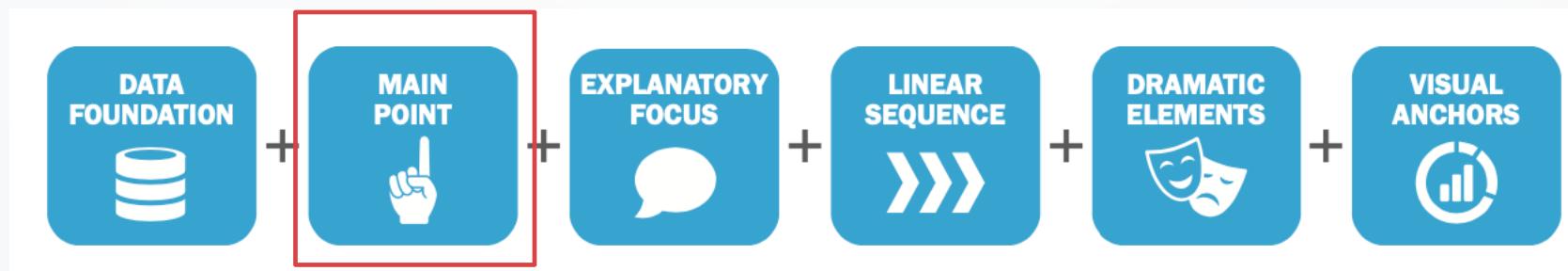
- 数据是数据故事的基石，数据故事的事实完全来自于数据，这是区分数据故事与其他类型的故事的一个独特特征。
- 每个数据故事都是由定量或定性的数据组成的，这些数据通常是分析或深入观察的结果。一些创造性内容可以用于数据故事的结构和传达，但真实的数据故事不会偏离数据事实太远的。
- 数据的质量和可信度将决定数据故事的可信度和影响力。



数据故事的六要素

- **要素二：主要观点**

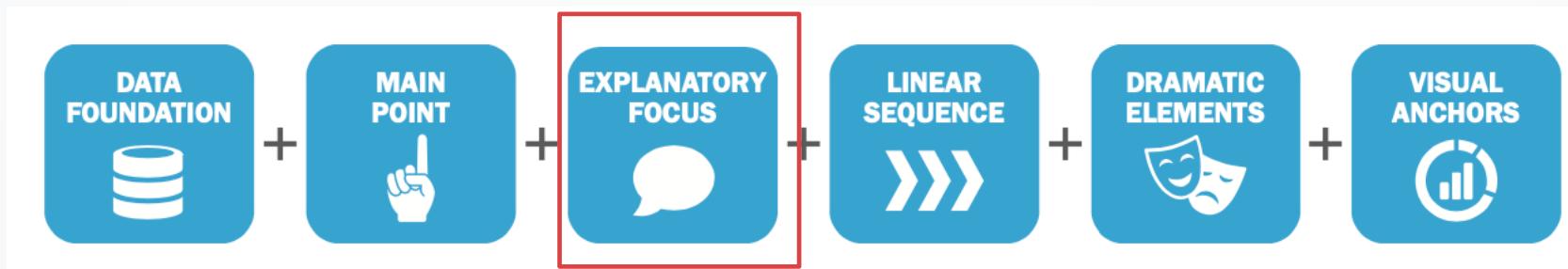
- 数据故事必须有一个主要观点，数据故事中不同的事实应该支持主要观点，确保数据故事有明确的目的。
- 数据故事是为了引导受众更好地理解观点或见解并引发讨论、去行动和改变。为了不让受众感到困惑，一个有凝聚力的数据故事，必须优先展示主要观点并且观点不能过多。



数据故事的六要素

• 要素三：解释性重点

- 每个数据故事必须有一个解释性重点。
- 数据故事需要基于更深入的分析和推理，努力阐明某件事发生的方式和原因来帮助受众理解数据，而不是仅仅是呈现表明细节。
- 解释性比描述性更具有挑战，因为我们对事情为什么或如何发生会有一定程度的不确定性和猜测，作为数据故事讲述者，需要根据可用的最佳信息来分享观点。



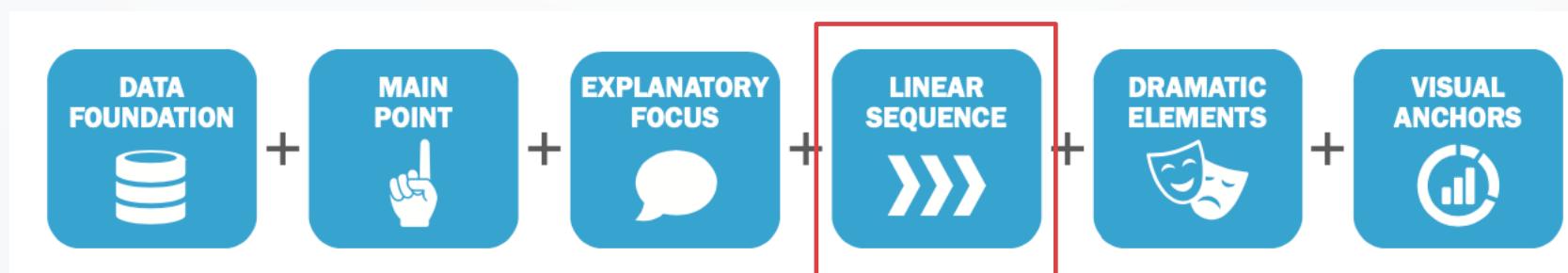
数据故事的六要素

• 要素四：线性序列

- 每个数据故事都遵循一个线性序列，以顺序的方式引入关键数据点，直到得出中心要点或结论。
- 分阶段暴露给观众信息，通过这个顺序的过程，观众逐步获得对主要观点或结论的理解。

以《绿野仙踪》为例，主人公多萝西·盖尔遇到了一系列有趣的事件。

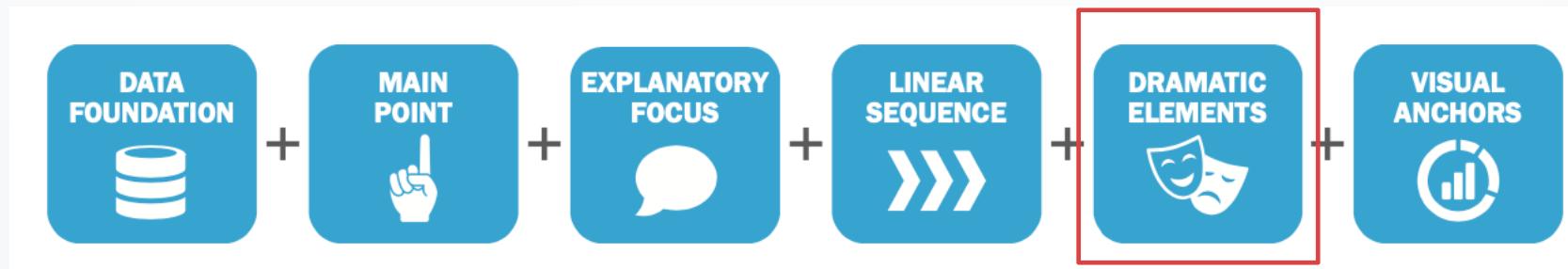
- ①龙卷风将她带到了一个陌生的地方；
- ②她得到了一双神奇的鞋子；
- ③她在黄砖路上旅行时遇到了一些旅伴；
- ④巫师派她和她的朋友执行一项任务；
- ⑤他们用一桶水融化了一个邪恶的女巫。



数据故事的六要素

• 要素五：戏剧性元素

- 数据故事与文学和电影中使用的故事有着相似的戏剧性基础——背景、情节、人物等。
 - 从背景来看，必须提供足够的背景细节，以便你的观众正确理解你的见解。例如，上个季度200万美元的销售额可能是好是坏，具体取决于去年同一季度的销售结果。
 - 从情节来看，每个事件都有一个推进故事和发展人物的目的。同样，如何构建和排序信息构成了数据故事的支柱。
 - 从人物来看，正在分析的数据通常是关于人的：客户、潜在客户、员工、合作伙伴、学生、患者、选民等。通过突出显示数据背后的人员，您可以为数据故事添加更相关的人类视角。
- 将戏剧性元素融入到数据故事中的次数越多，它在情感层面上对观众的吸引力就越大。

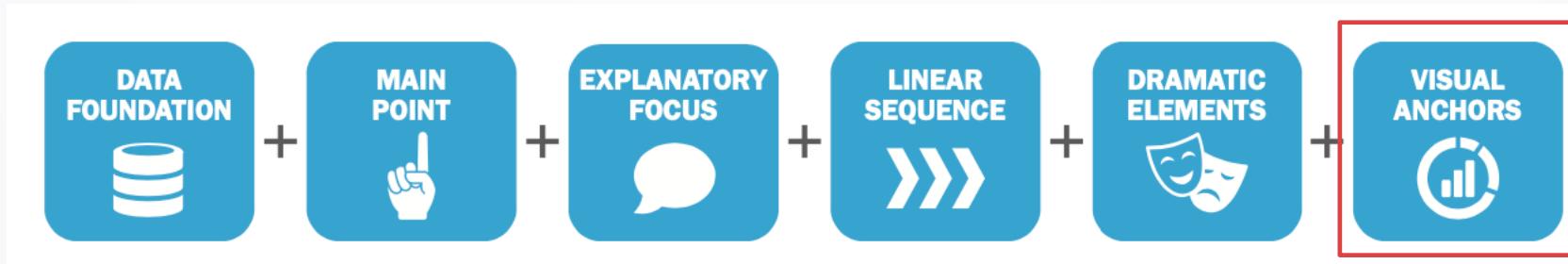
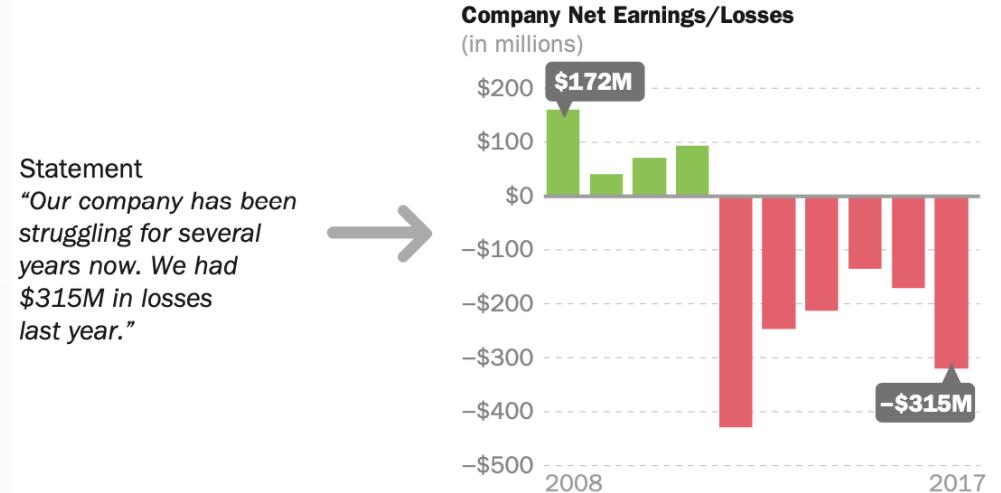


数据故事的六要素

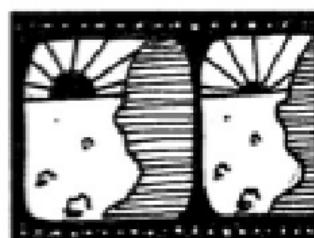
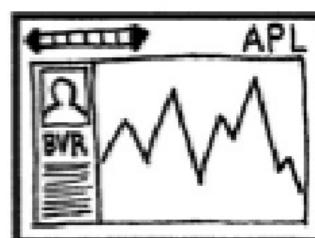
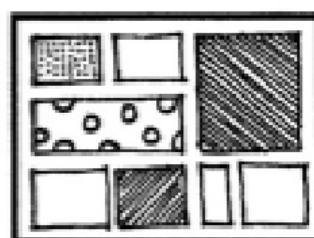
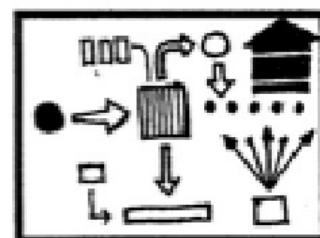
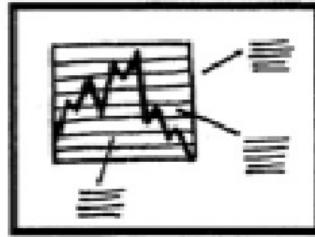
- **要素六：视觉锚定**

- 数据可视化可以帮助人们看到数据中的模式、趋势和异常。
- 数据可视化可在分析过程中可用于数据发现阶段中的发现见解。
- 其他形式的图像也可帮助加强讲故事。例如，图标可用于为观众创建心理快捷方式，或者照片可以为关键数据点添加情感强调。

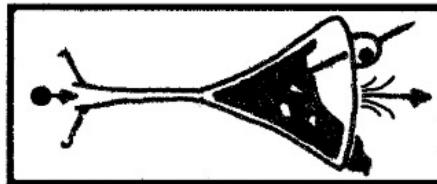
VISUALS CAN HELP YOUR DATA SPEAK (MAYBE EVEN SCREAM)



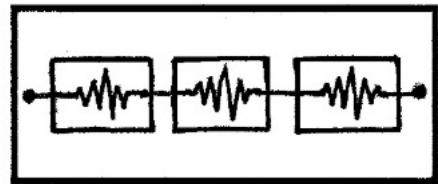
数据故事的体裁



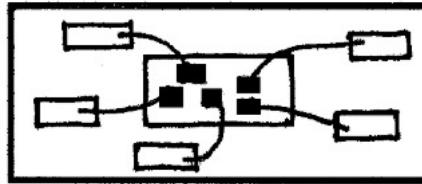
Martini Glass Structure



Interactive Slideshow



Drill-Down Story

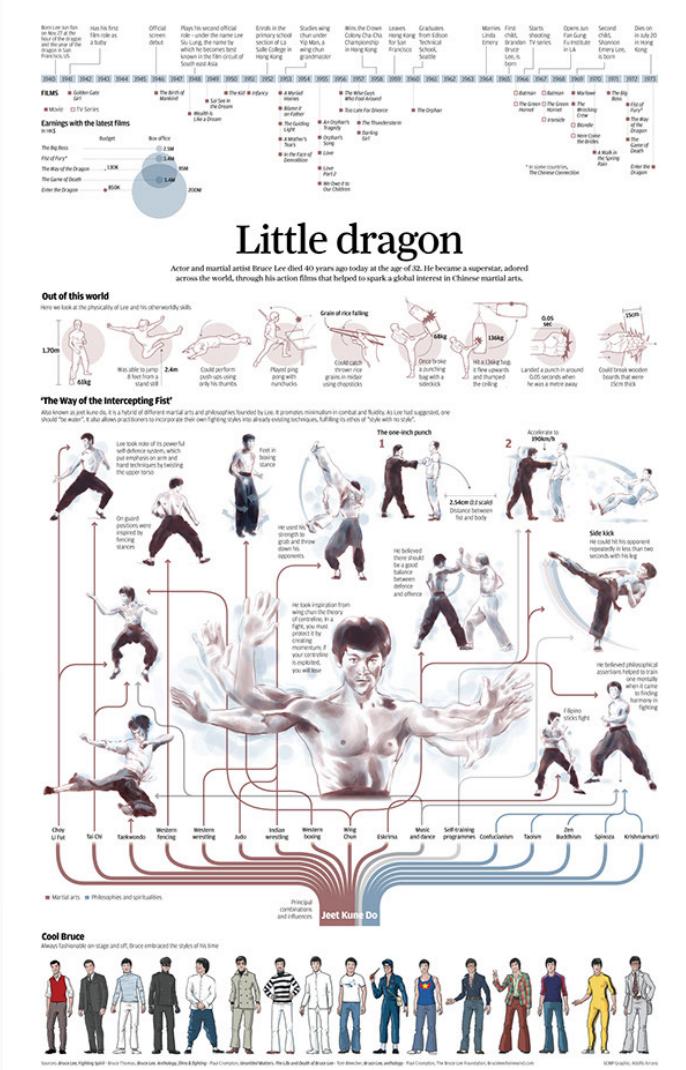
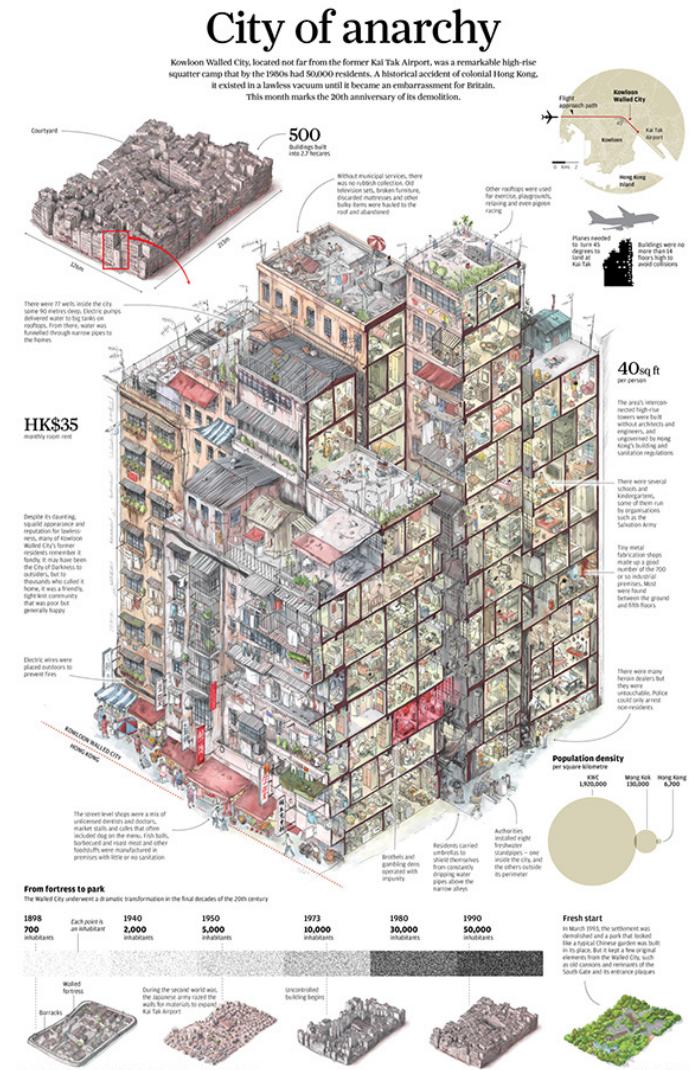


- Segal and Heer 在2010年提出了7种不同的数据故事体裁
- 当然，这一分类并非一成不变。随着计算机和多媒体技术的进步，数据故事的形式也在不断丰富

可视化叙事的体裁

信息图 Infographics

一种为高效传播信息而生的视觉表现形态。一般体现为一张将多个图表/信息进行组合排版的静态图片。



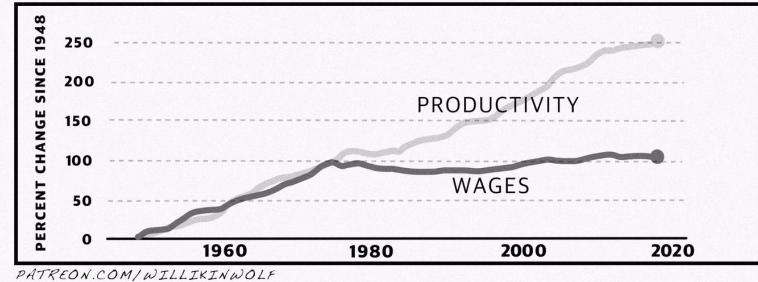
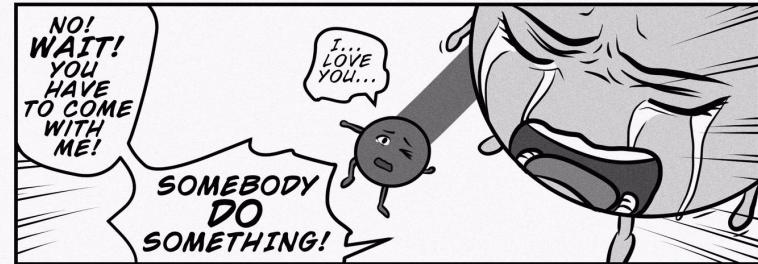
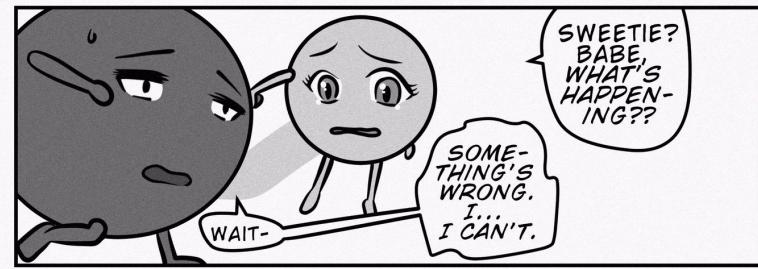
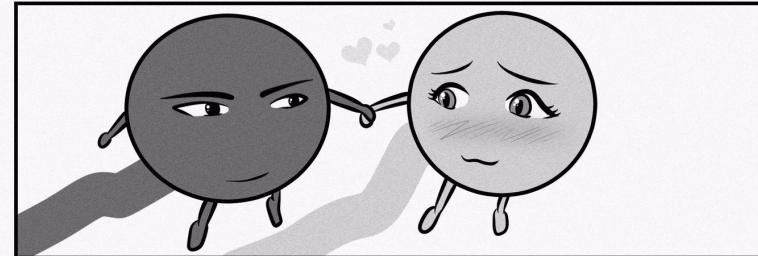
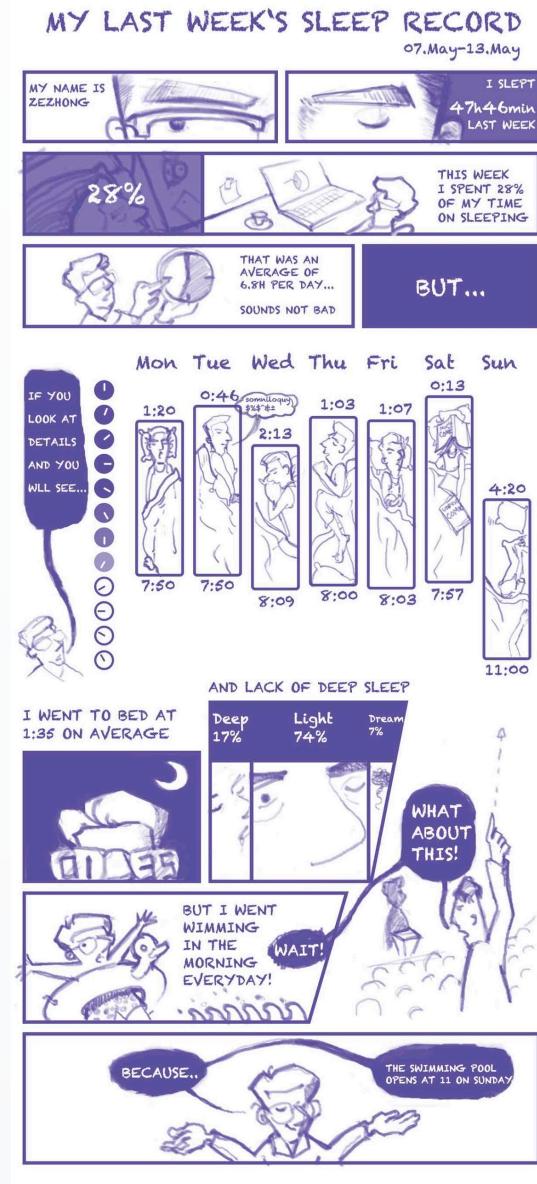
可视化叙事的体裁

数据漫画 Data Comics

一种继承了漫画艺术特征的信息设计，如使用漫画的分格布局、夸张的视觉风格等。

<https://www.datacomics.net/>

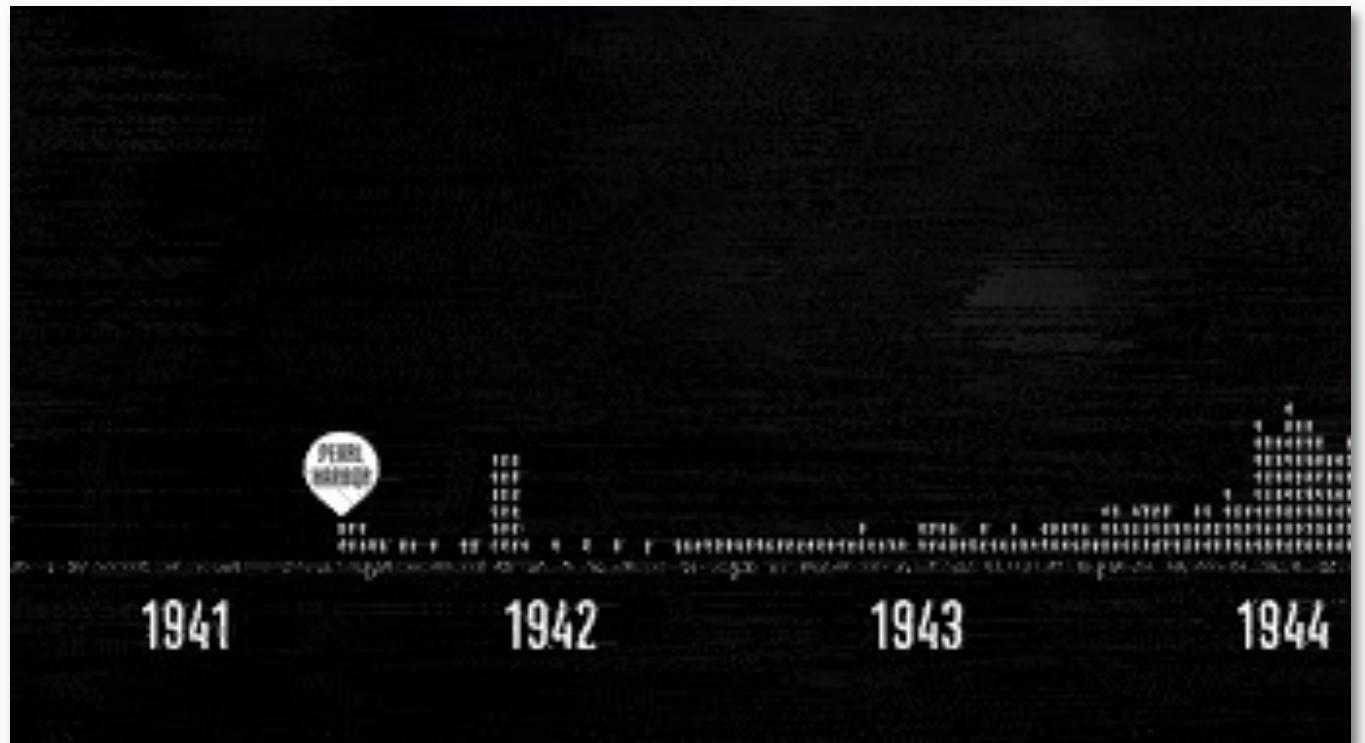
<https://twitter.com/WillikinWolf/status/1176006515968241665>



可视化叙事的体裁

数据视频 Data Video

一种动态的数据故事体裁，常常借鉴了影视化的叙事手法，如镜头语言、场面调度、配音配乐等。



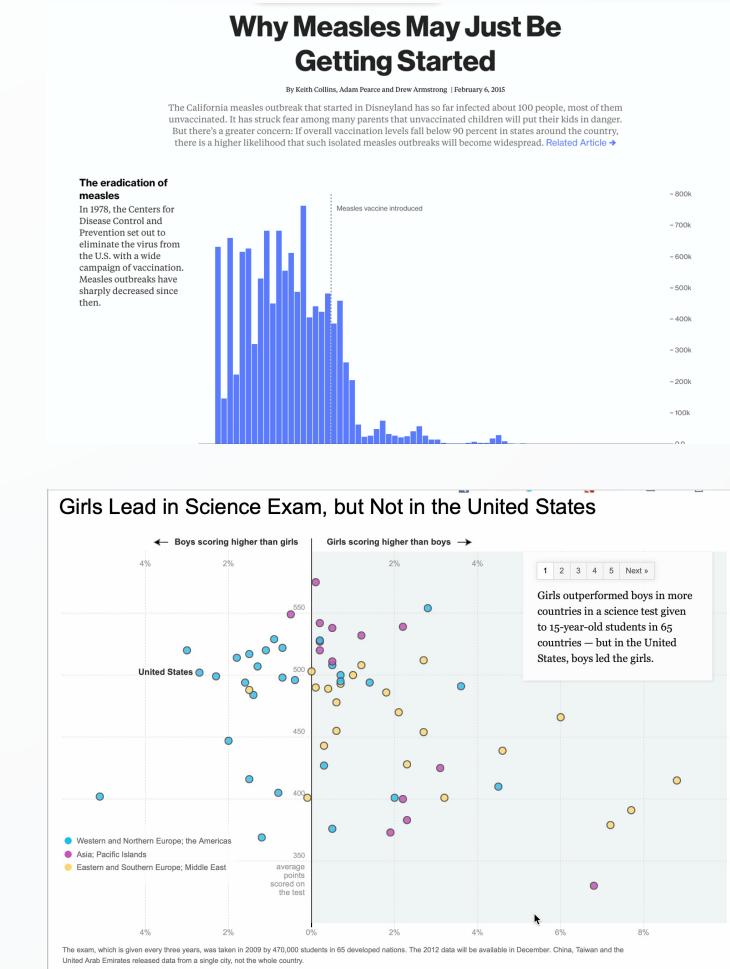
<https://www.youtube.com/watch?v=DwKPFT-RioU>

可视化叙事的体裁

交互网页 Website

运用网页的各种特性，使得用户可以与数据可视化进行互动，如滚动、点击、鼠标悬停，筛选，搜索等

- 滚动叙事：在用户向下滚动页面时显示内容
- 步进叙事：用户需要点击故事的步骤和场景



图片来源

<https://archive.nytimes.com/www.nytimes.com/interactive/2013/02/04/science/girls-lead-in-science-exam-but-not-in-the-united-states.html>
https://blog.csdn.net/weixin_26727597/article/details/108915797

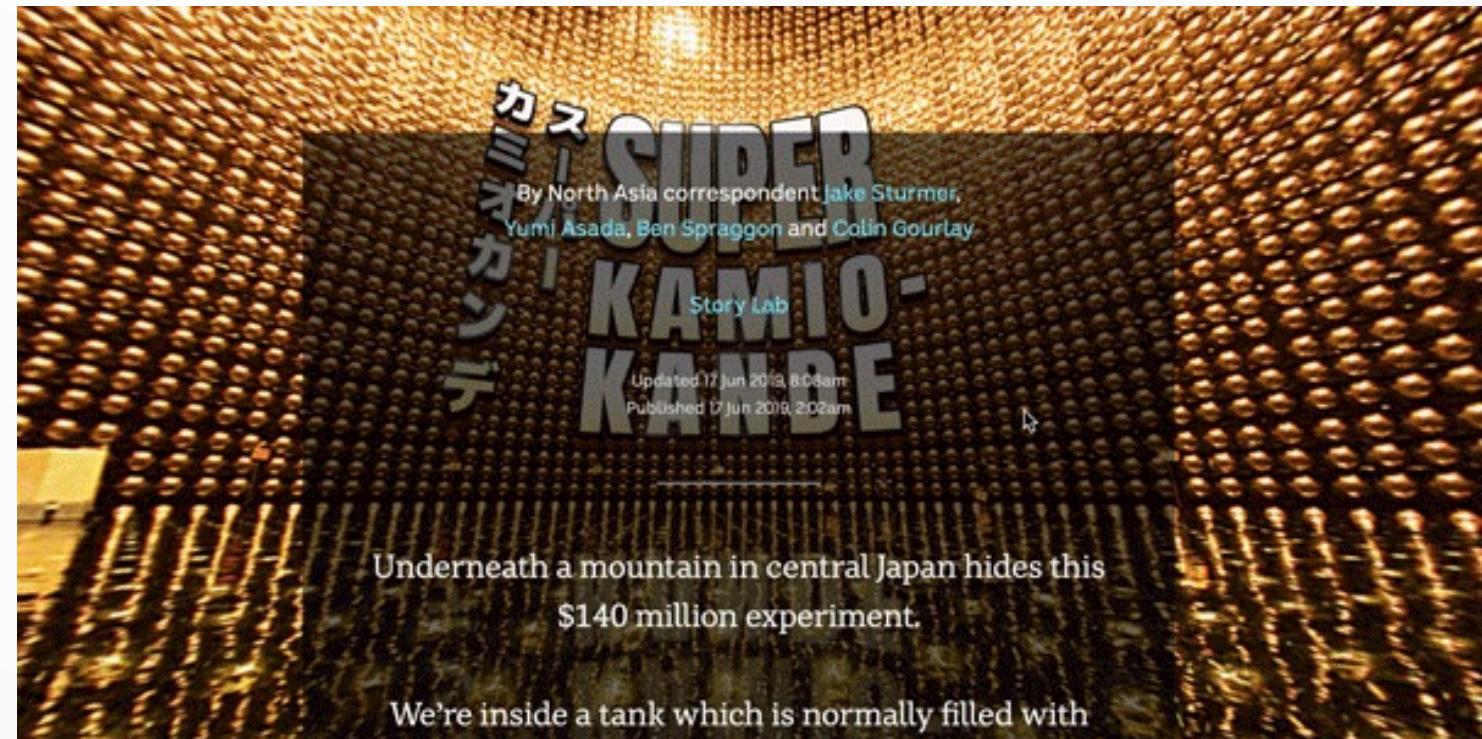
可视化叙事的体裁

沉浸式故事 Immersive Story

VR/AR

多重感官

仿真设计



<https://www.abc.net.au/news/2019-06-17/inside-super-kamiokande-360-tour/11209104>

可视化叙事的体裁

实体可视化 Physicalization

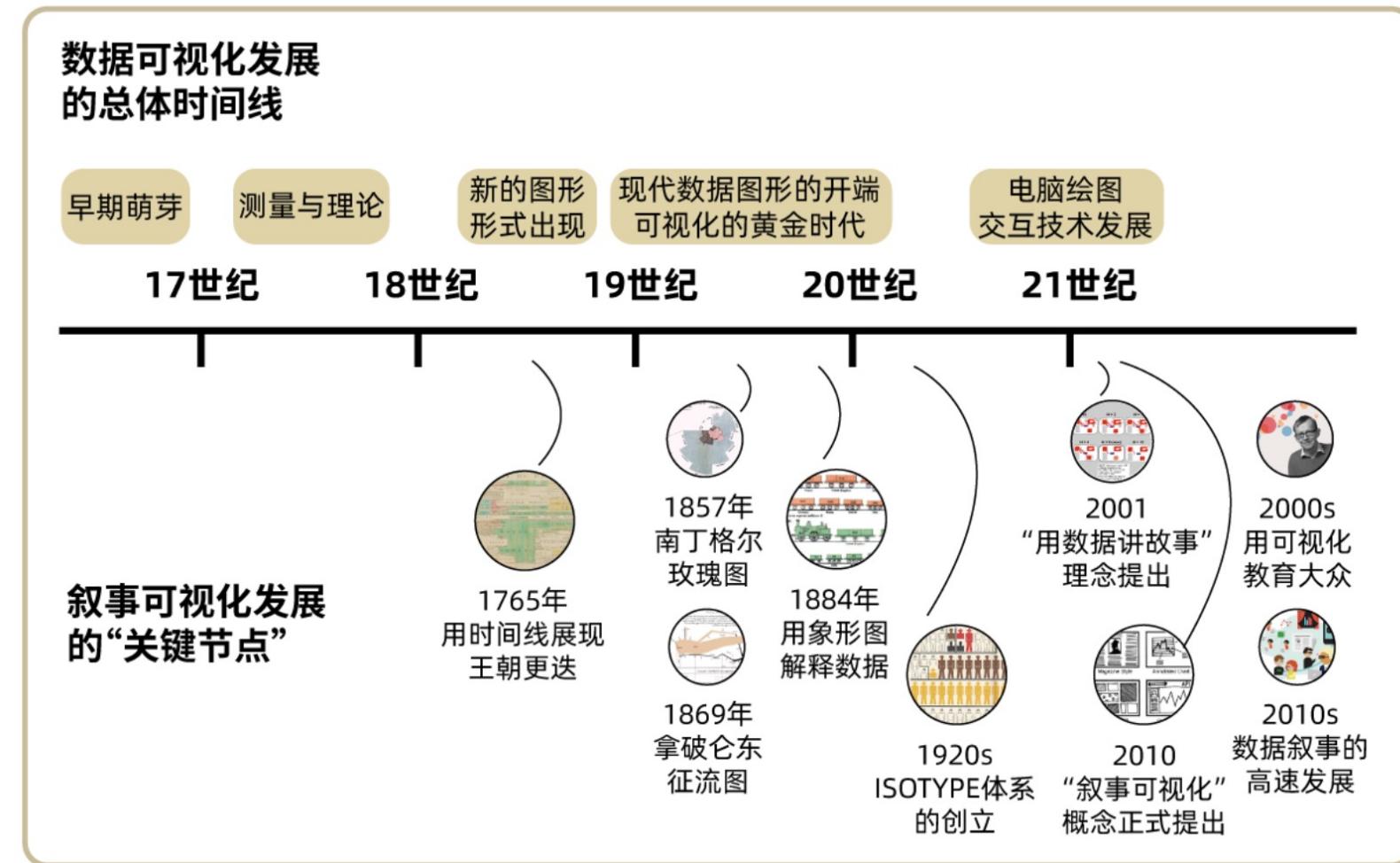
实景布展、3D打印等



<https://www.fastcompany.com/90171107/printing-out-the-privacy-policies-of-facebook-snap-and-others>



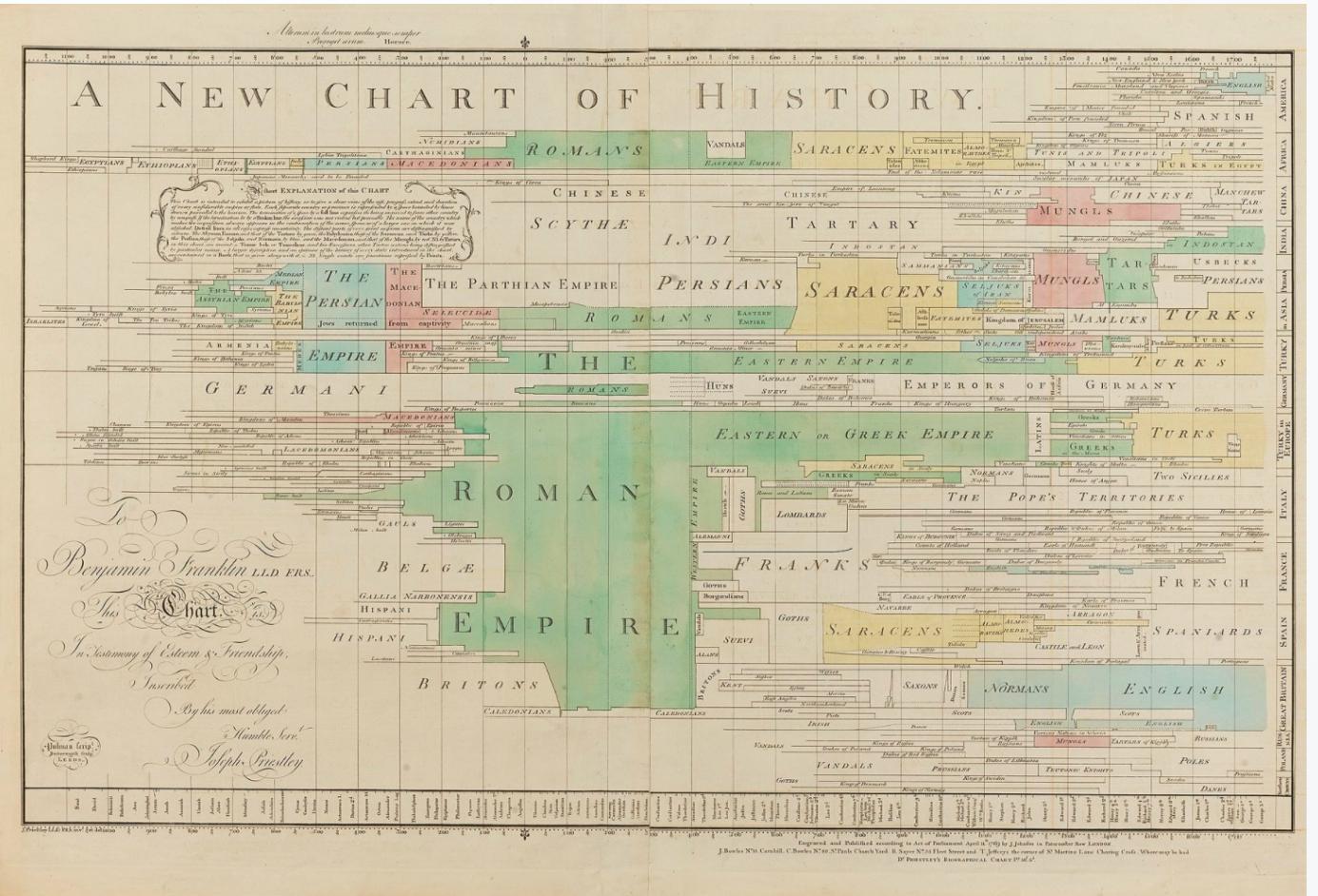
历史上具有叙事特征的可视化



用可视化展现帝国更迭

历史新图 1769年

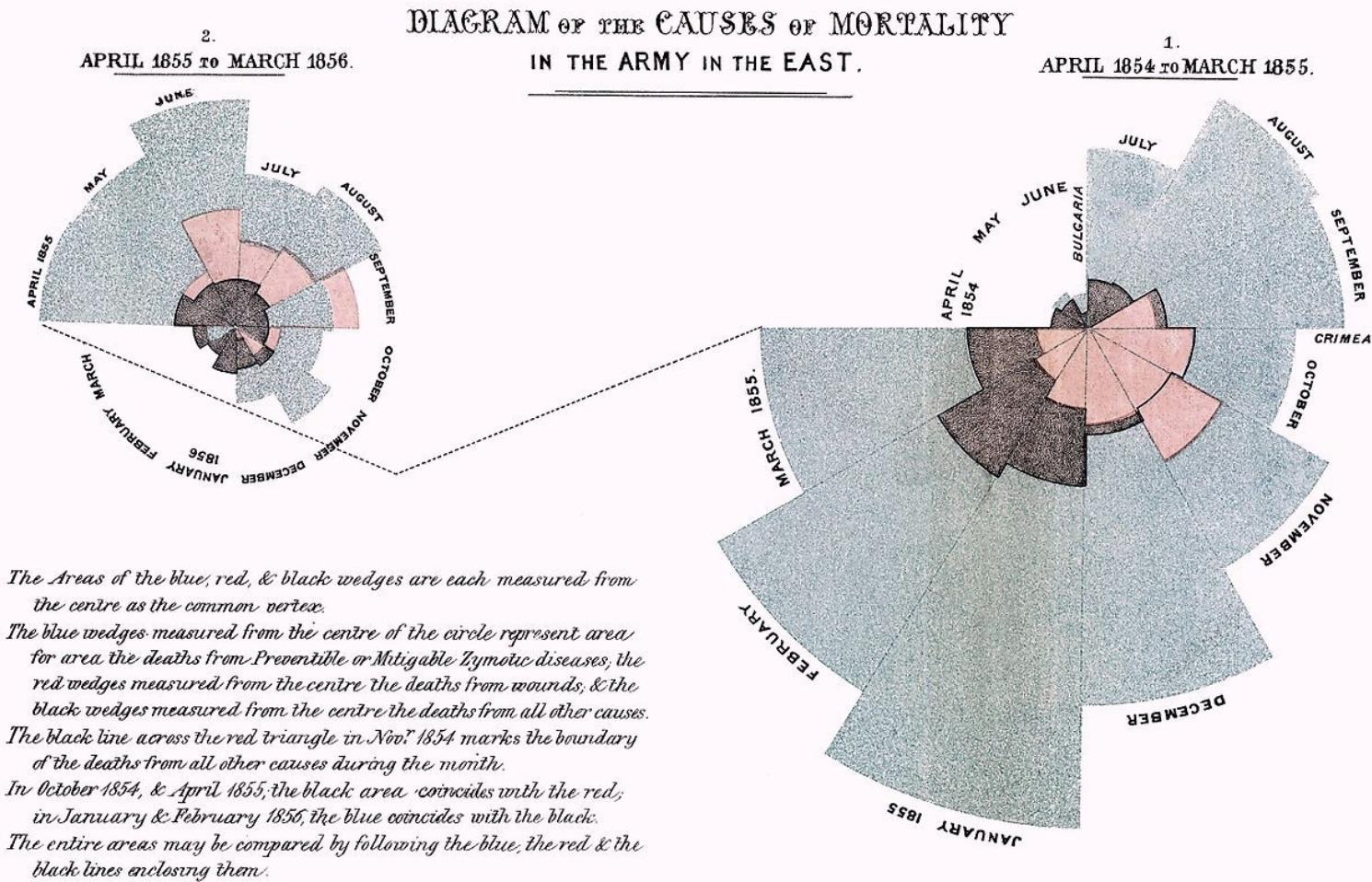
用填色时间表的形式，展示了世界各地区(如英国、俄罗斯、中国等)被不同帝国(如罗马、波斯、蒙古等)统治的时间，由此也展示了一幅巨大的帝国更迭史。



用玫瑰图推动公共卫生改革

南丁格尔玫瑰图 1858年

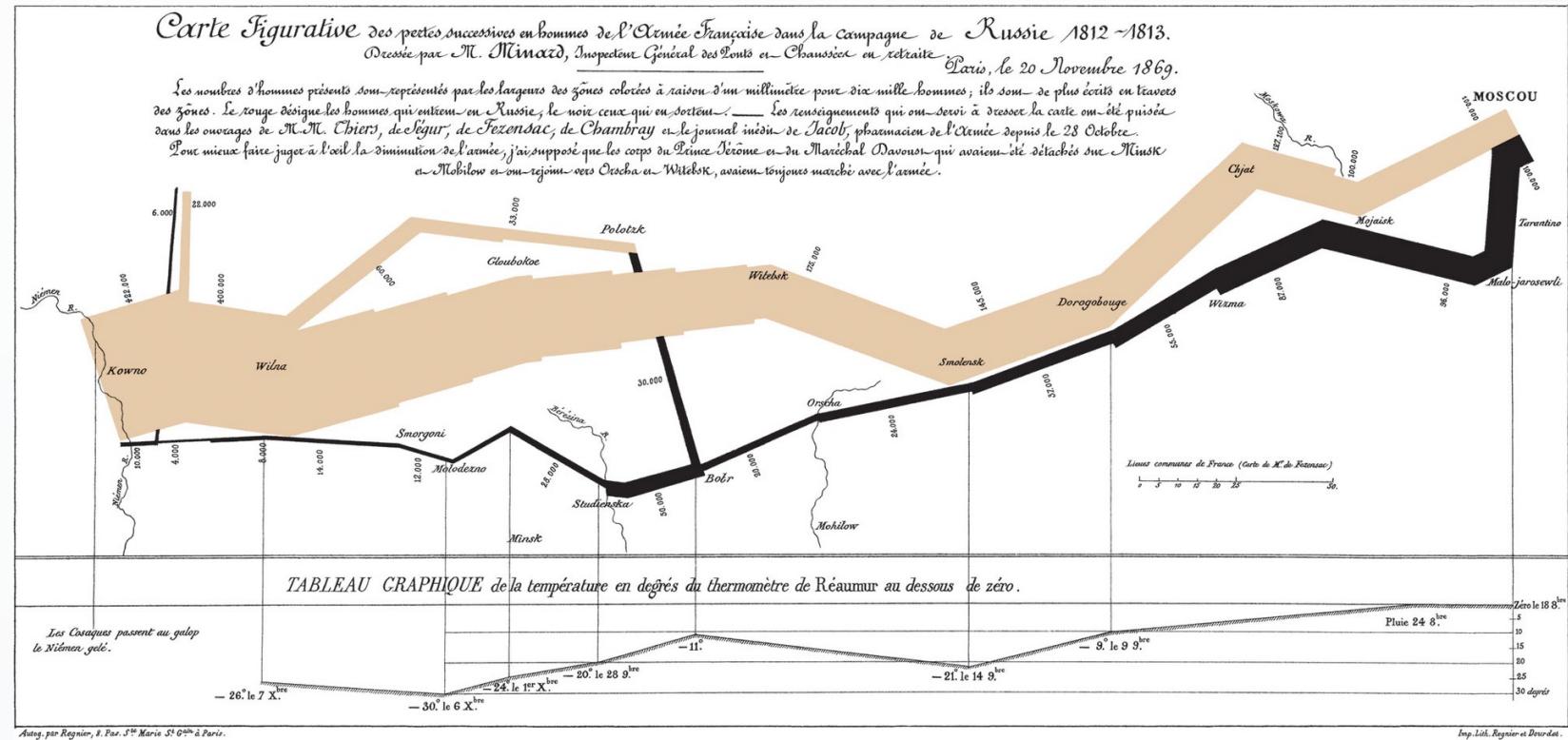
图中共有 12 个扇区，分别代表 12 个月。各扇区的角度相同，用半径来表示死亡人数，蓝色表示死于可避免的感染的人数，粉色表示受伤过重而死亡的人数，黑色表示死于其他原因的人数。



用流向地图反思战争

法国军队的持续溃败 1869年

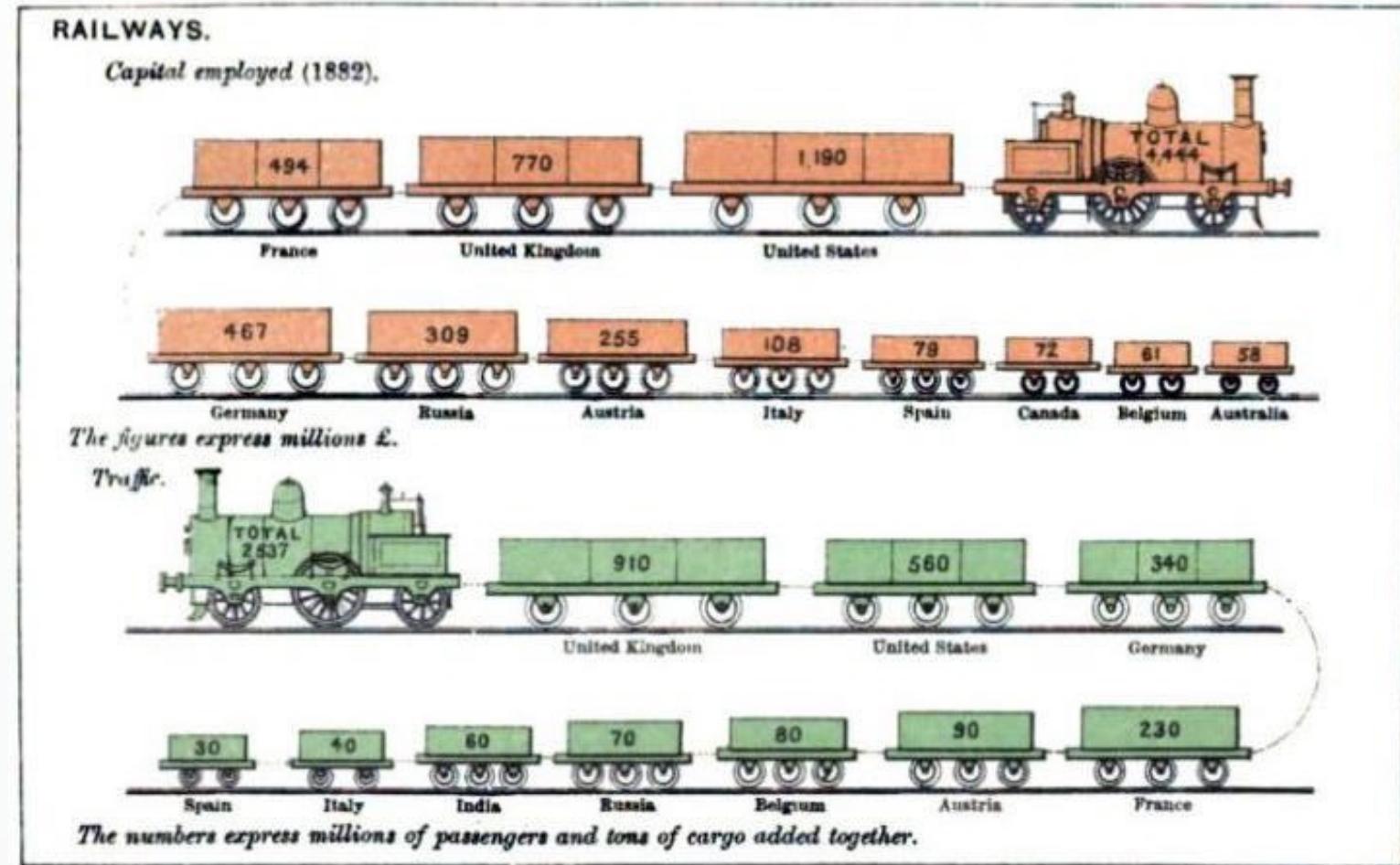
这幅图描绘了拿破仑 1812 年东征俄国时，法国士兵的行军方向、缩减情况，以及对应的时间节点。其中，黄色部分代表东征时士兵人数的变化，而黑色部分代表军队从俄国撤回时人数的变化。每一毫米的缩减，就代表 1 万名士兵的死亡。



用装饰性可视化启发民智

铁路 1884年

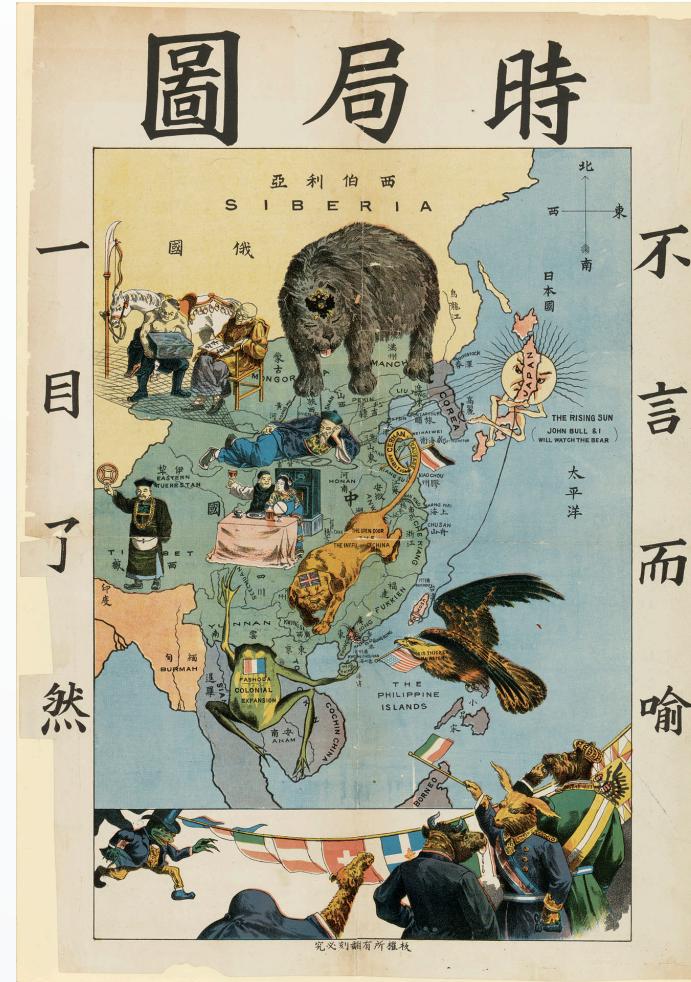
这幅图用与数字成正比的图形符号来表示数据。比起抽象化的统计图表，趣味性更强、对于民众来说更容易理解



用装饰性可视化启发民智

时局图 1898年

用一幅地图形象地描绘了列强虎视眈眈、瓜分我国领土的现实情况。装饰性，文字注解，以及较为夸张的画风，都使得这幅作品具备了数据漫画的特色。



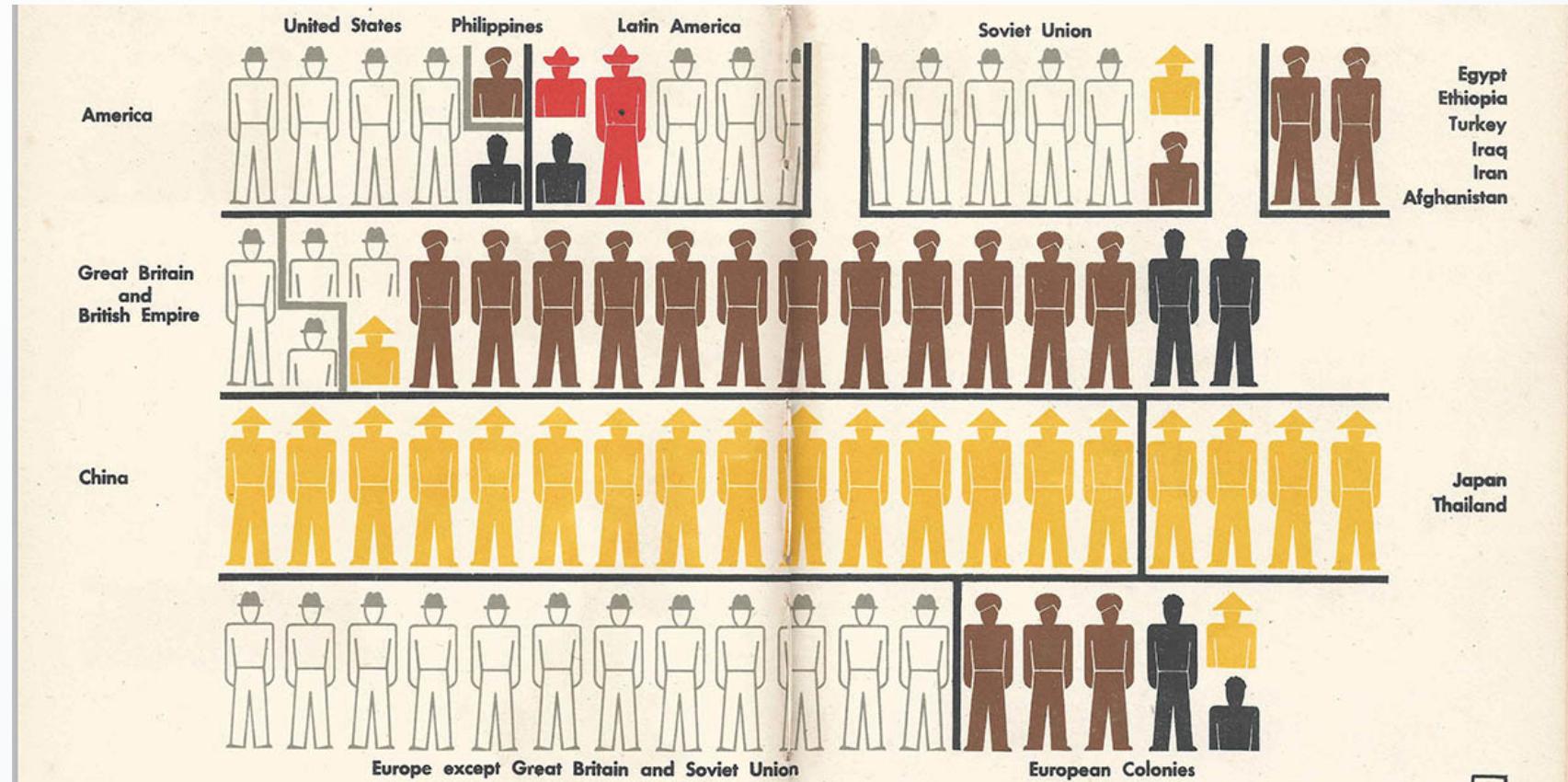
用装饰性可视化启发民智

国际图形字体教育体系

ISOTYPE的发明

20世纪20年代

利用图形化的统计数据呈现方式，在保持娱乐性、吸引注意力的同时，对观众进行“视觉教育”。



数据故事需要讲故事的人

- **Hans Rosling教授**

- 著名公共卫生学家Hans Rosling于2010年在TED发表的演讲，就是数据故事用于教育科普的经典案例。
- 这篇题为《我们正在赢得与儿童死亡率的战争》的演讲，通过数据向大众揭示了过去几十年中全人类在降低儿童死亡率上取得的长足进步。Hans教授还呼吁人们放下成见，正视“发展中国家”在降低儿童死亡率的战斗中做出的突出贡献——毕竟在这张气泡图里，许多亚洲国家气泡的位置比欧美国家要更加左下角。



The Good News of the Decade? We're Winning the War Against
Child Mortality (Hans Rosling, 2010)

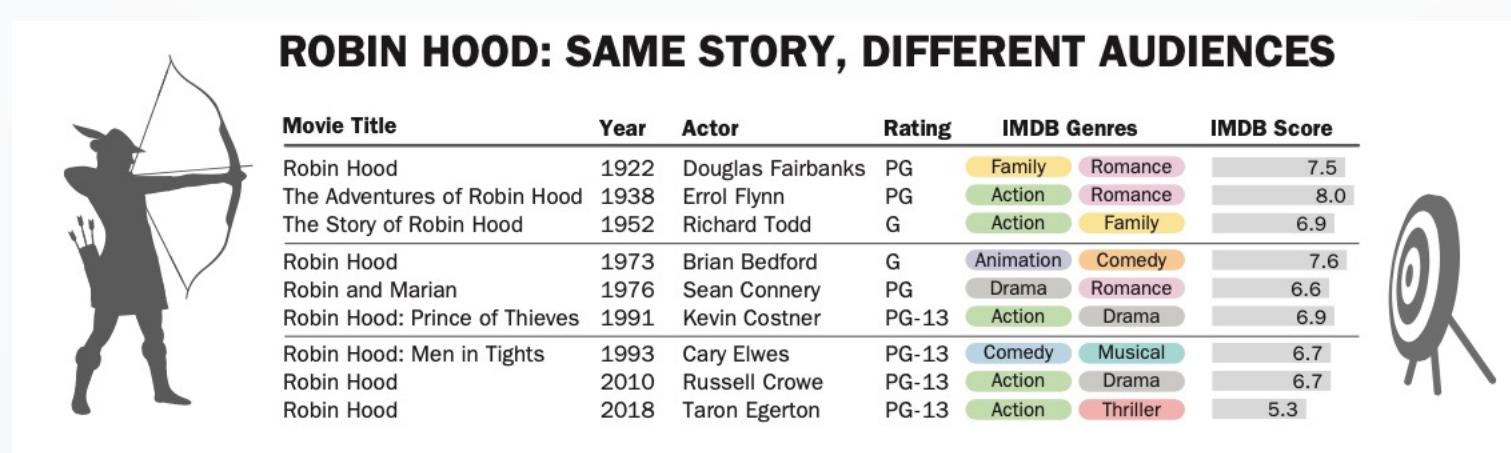
数据故事需要讲故事的人

- **数据故事的成功取决于讲故事的人的9种能力。**
 - **确定关键见解：**可以找到有意义的数据见解，并决定是否需要将其作为数据故事进行分享。
 - **减少/消除偏见：**保证讲述数据故事客观性。
 - **充足背景知识：**讲述数据故事之前有足够的背景知识，确保见解是有意义的并能和受众产生共鸣。
 - **了解受众：**为目标受众定制合适的内容。
 - **策划信息：**策划故事中应包含的信息。太多的信息会让受众不知所措，但太少可能无法引起注意。
 - **整合故事：**确定故事的流程以及不同元素组合方式。
 - **提供旁白：**讲故事的人充当信息的引导者，帮助观众理解和解释数据。
 - **选择视觉效果：**可视化设计将决定受众如何感知和理解见解。不同方式可视化可传达截然不同的信息。
 - **增加可信度：**讲故事的人的声誉和专业知识可以带来可信度和权威。不值得信任或过于偏见的叙事者会破坏原本合理的数据故事的有效性。

讲故事前了解受众

- 受众多会影响数据故事塑造的重点和方向

➤ 例如，下图每一版本罗宾汉电影用其独特的方式讲述同一个故事，吸引特定的受众。当你要向受众推荐一部罗宾汉电影时，如果不了解目标受众，很难选到合适的电影版本。比如面向6岁的孩子推荐Russell crowe的2010 PG-13版本。



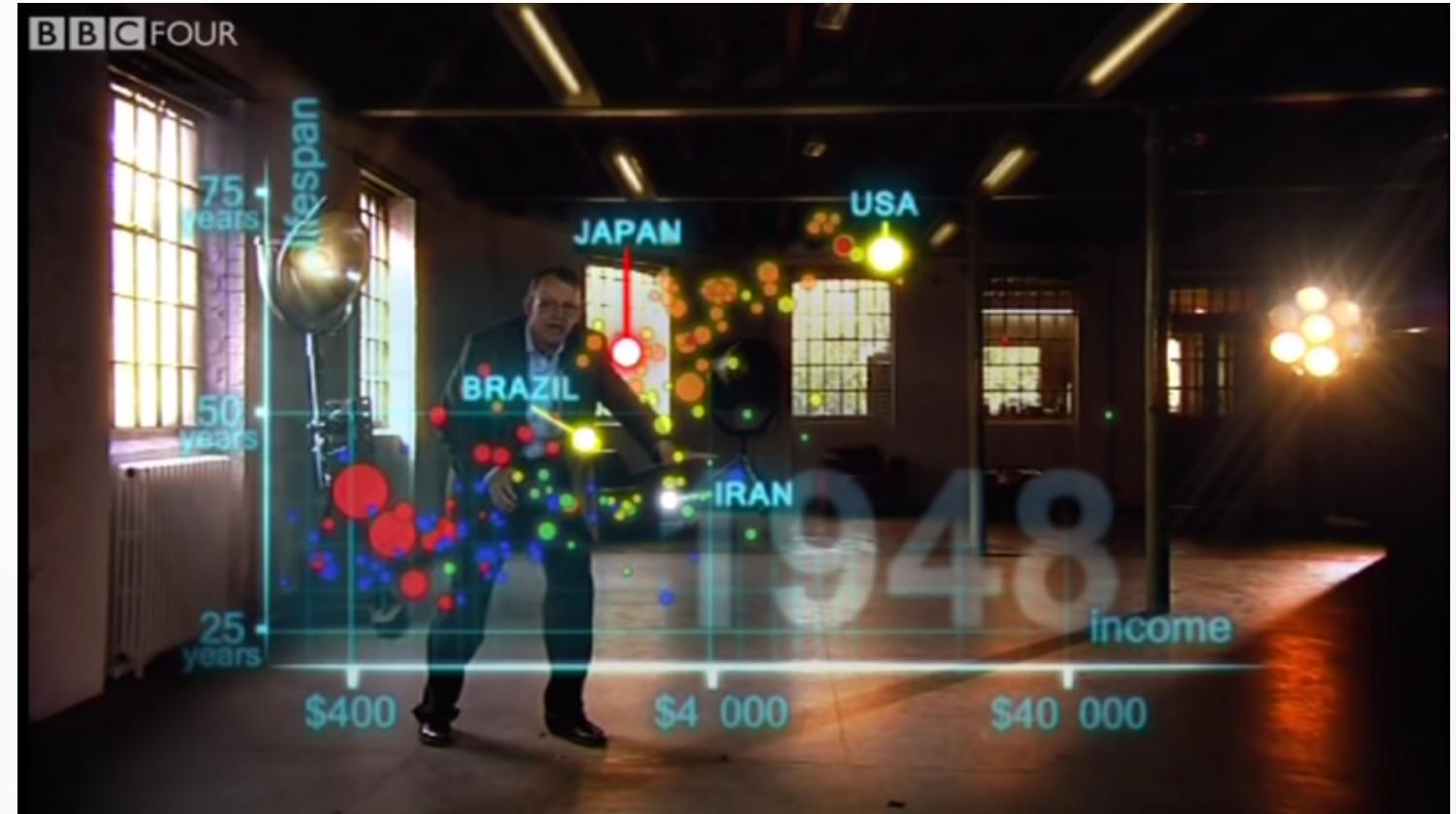
讲故事前了解受众

- 数据故事应该分享给正确的受众，理想情况下分享给有能力带来改变的受众群体或个人。
- **如何为目标受众量身定制数据故事。以下有8个考虑因素：**
 1. 关键目标和优先项：与受众保持一致的关键目标和优先项，确保数据故事对受众是有意义的。
 2. 态度和喜好：了解受众对见解的态度或叙事方式的喜好，可以帮助指导如何准备和讲述数据故事。
 3. 特定的预料：尽可能预测受众的问题，将答案编织到你的数据故事中。
 4. 合适的时机：在合适的时机分享数据故事。
 5. 主题熟悉度：受众涉及不同水平的知识背景，根据受众对主题熟悉度的不同进行调整。
 6. 数据素养：精通分析数据的受众和不精通数据的受众，调整细节数量和行业术语等。
 7. 资历水平：比如高管通常没有耐心或带宽来坐着看整个数据故事。
 8. 多样化受众：平衡不同背景、兴趣和议程的多样化受众之间的冲突需求。

数据故事应用领域

汇报与演讲

用生动的形式讲述研究/企业故事

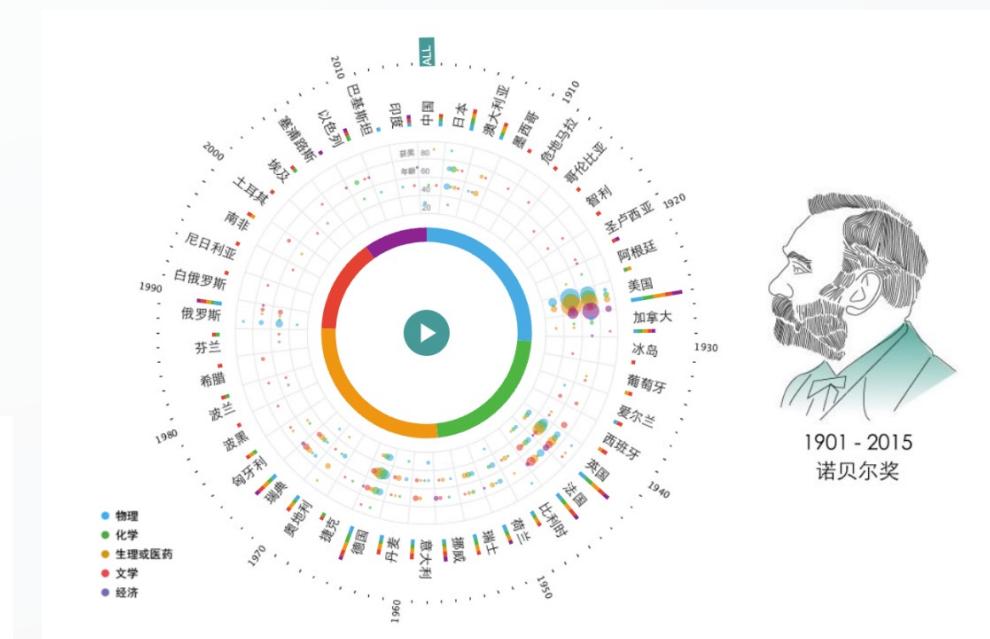
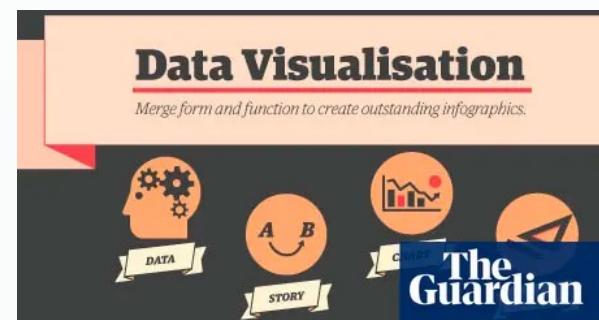


200 Countries, 200 Years, 4 Minutes. BBC (2010)

数据故事应用领域

新闻传播

数据新闻 (Data Journalism) 的崛起，专门做数据可视化报道的新闻编辑室出现。

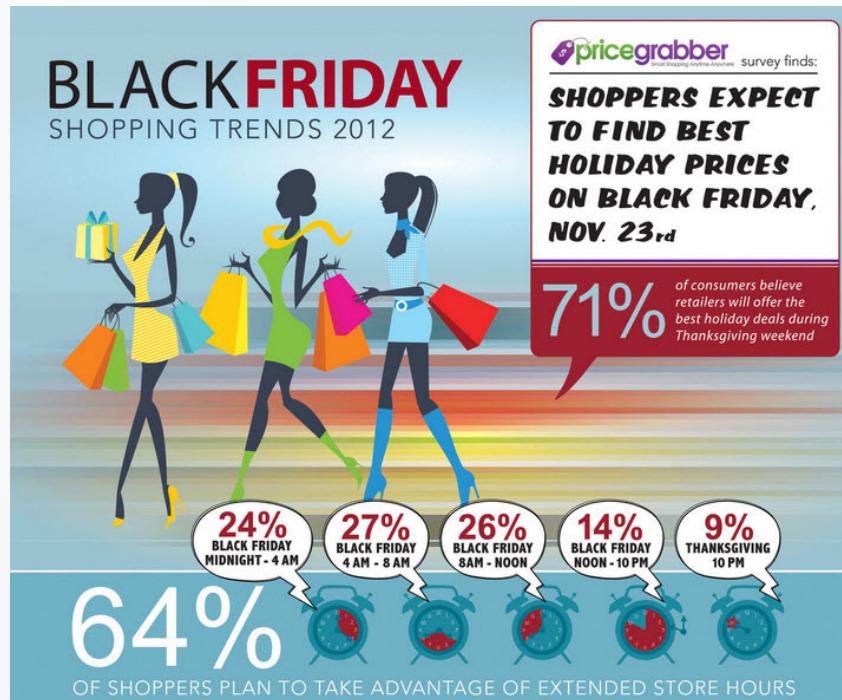


星空彩绘诺贝尔，发表于 2013 年

数据故事应用领域

广告、营销、公关

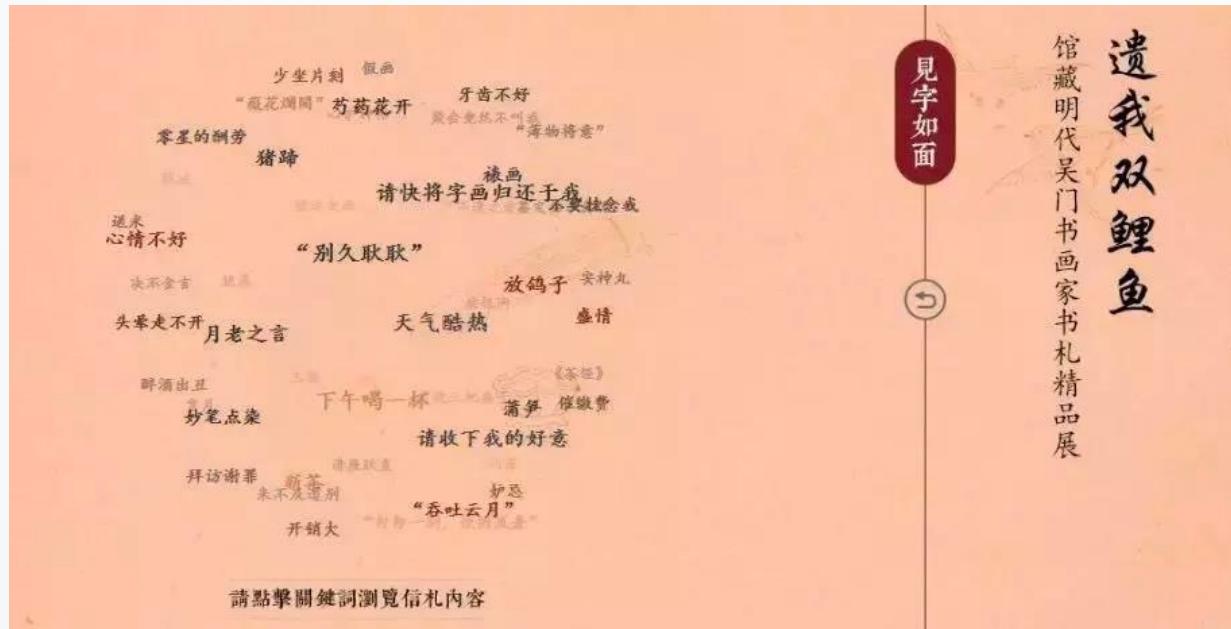
用数据叙事达到刺激、说服、形象塑造等目的



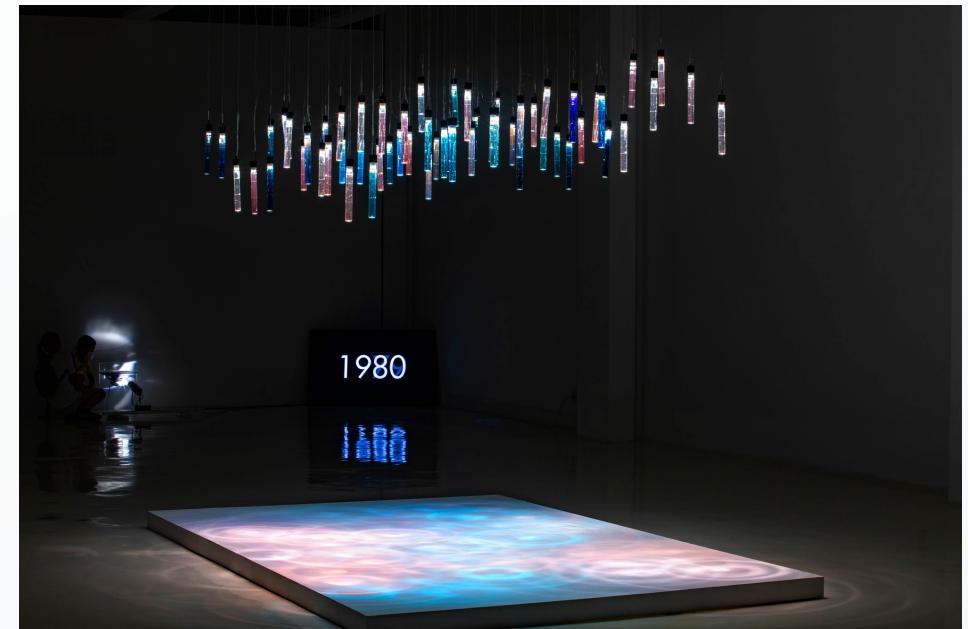
数据故事应用领域

公共教育/人文艺术

在公共空间，利用数据可视化叙事来传递知识和态度

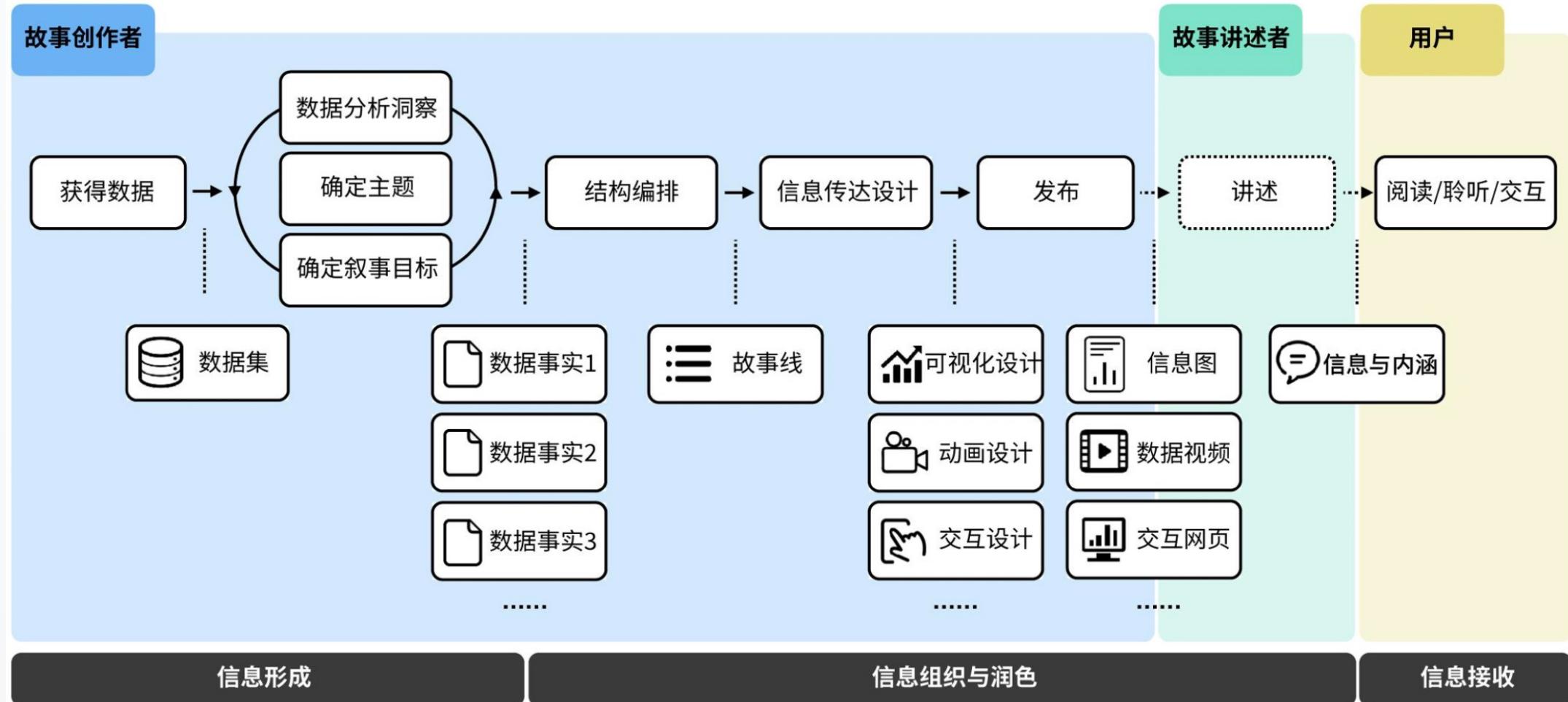


上海博物馆
https://www.sohu.com/a/327883533_426335

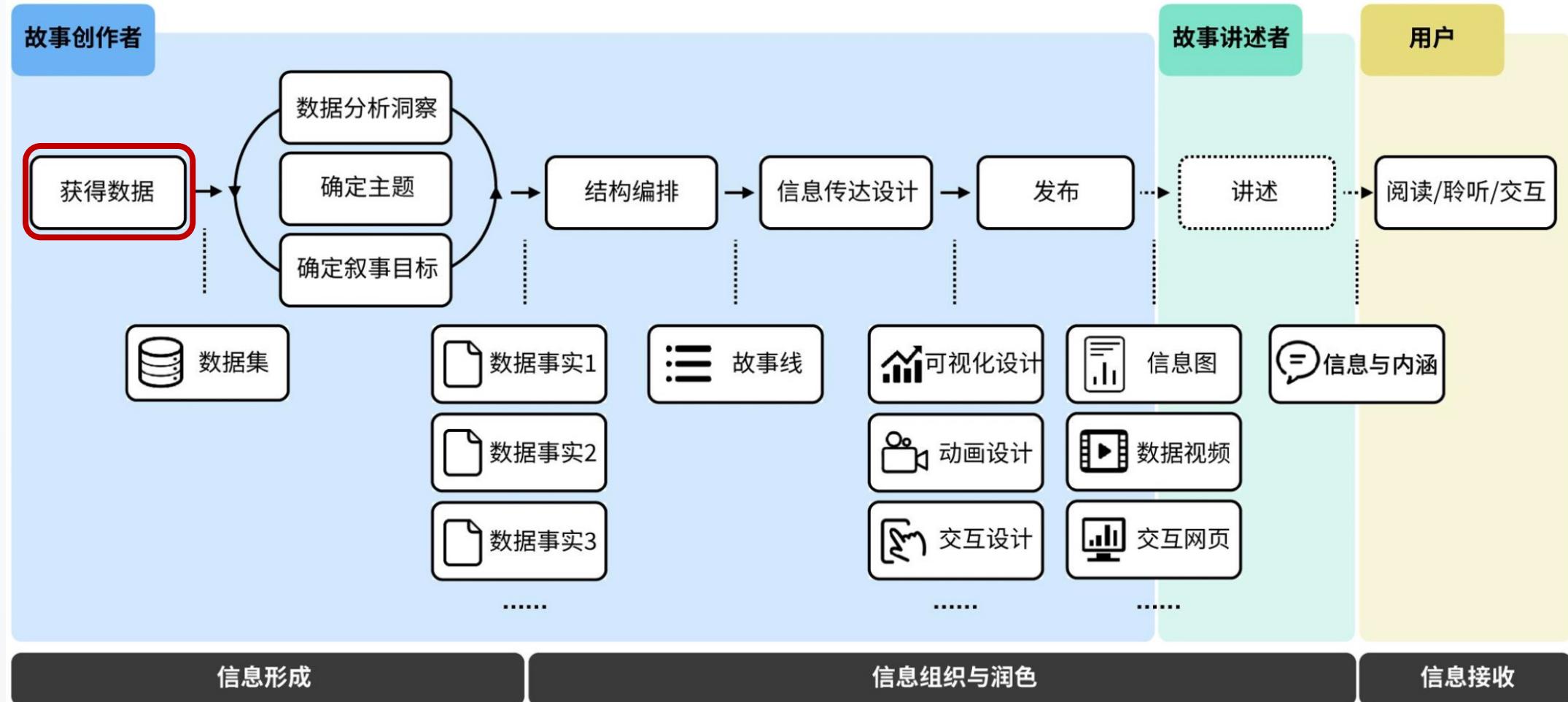


VIS艺术项目
<https://youtu.be/CiohDzBNNMM>

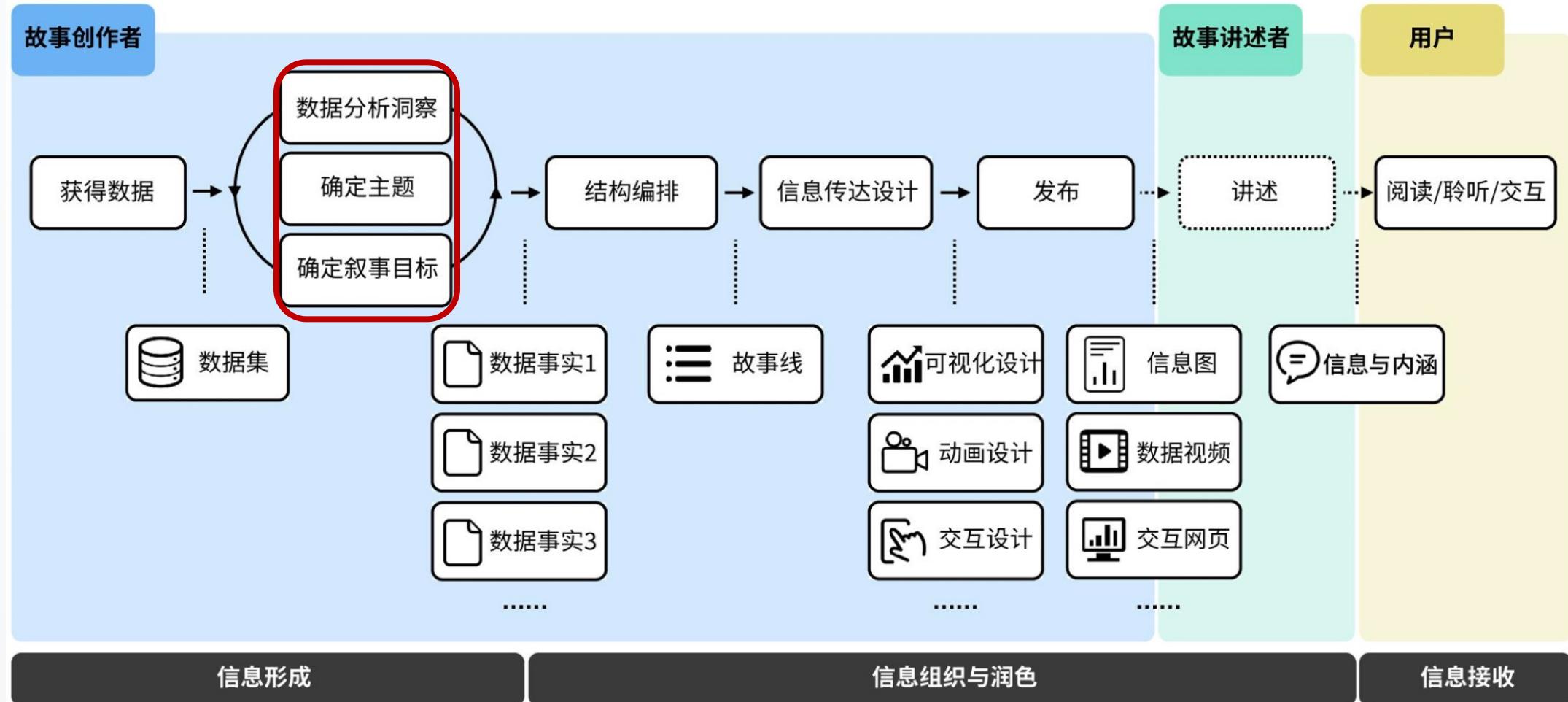
数据故事的制作流程



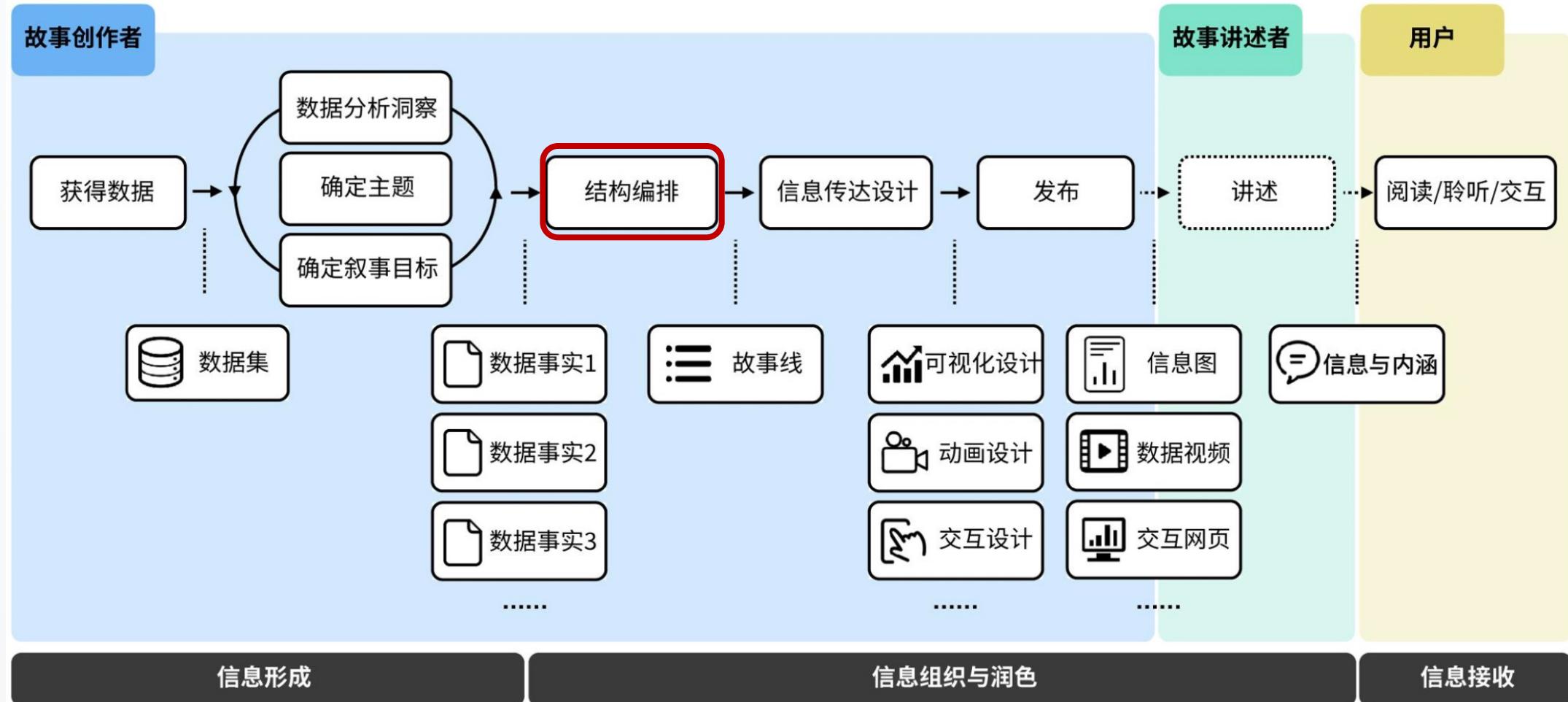
数据故事的制作流程



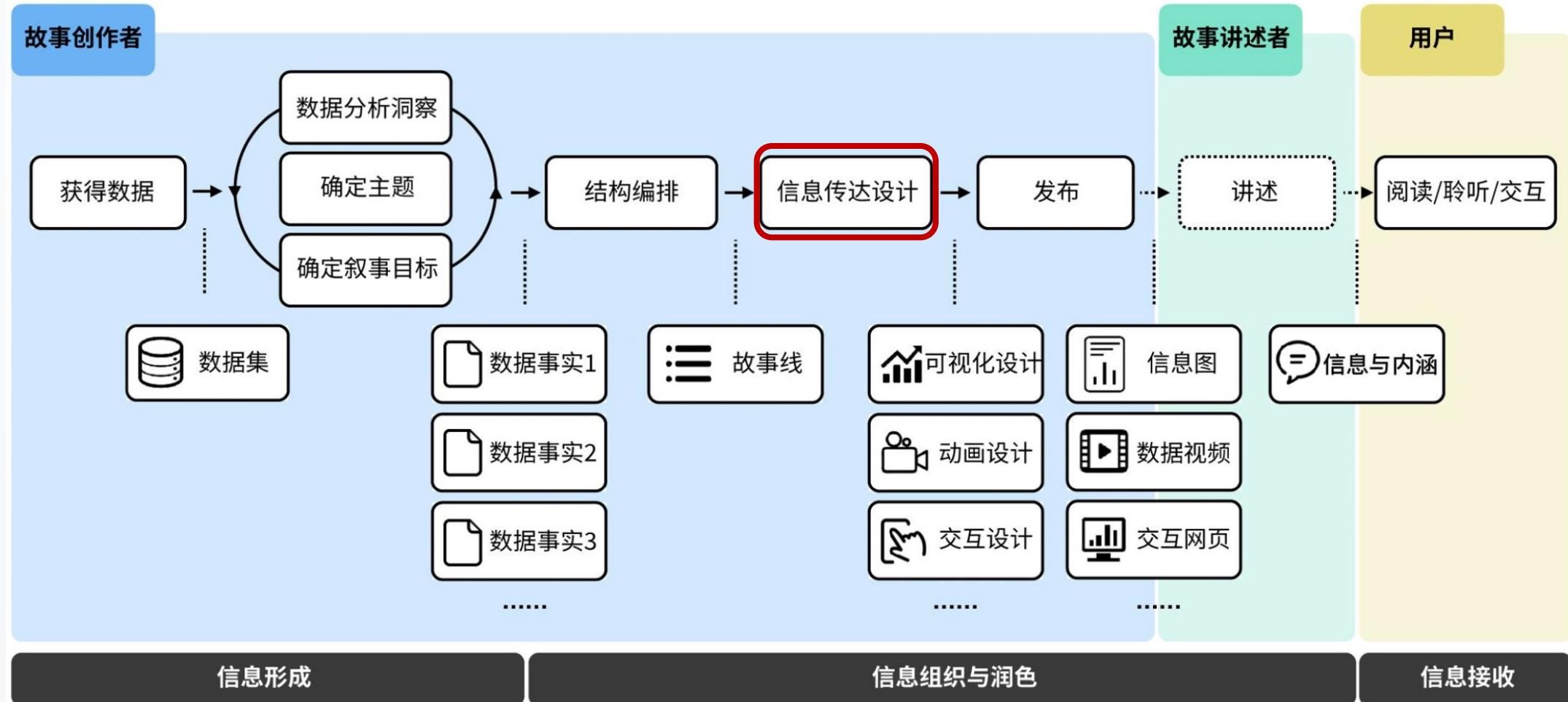
数据故事的制作流程



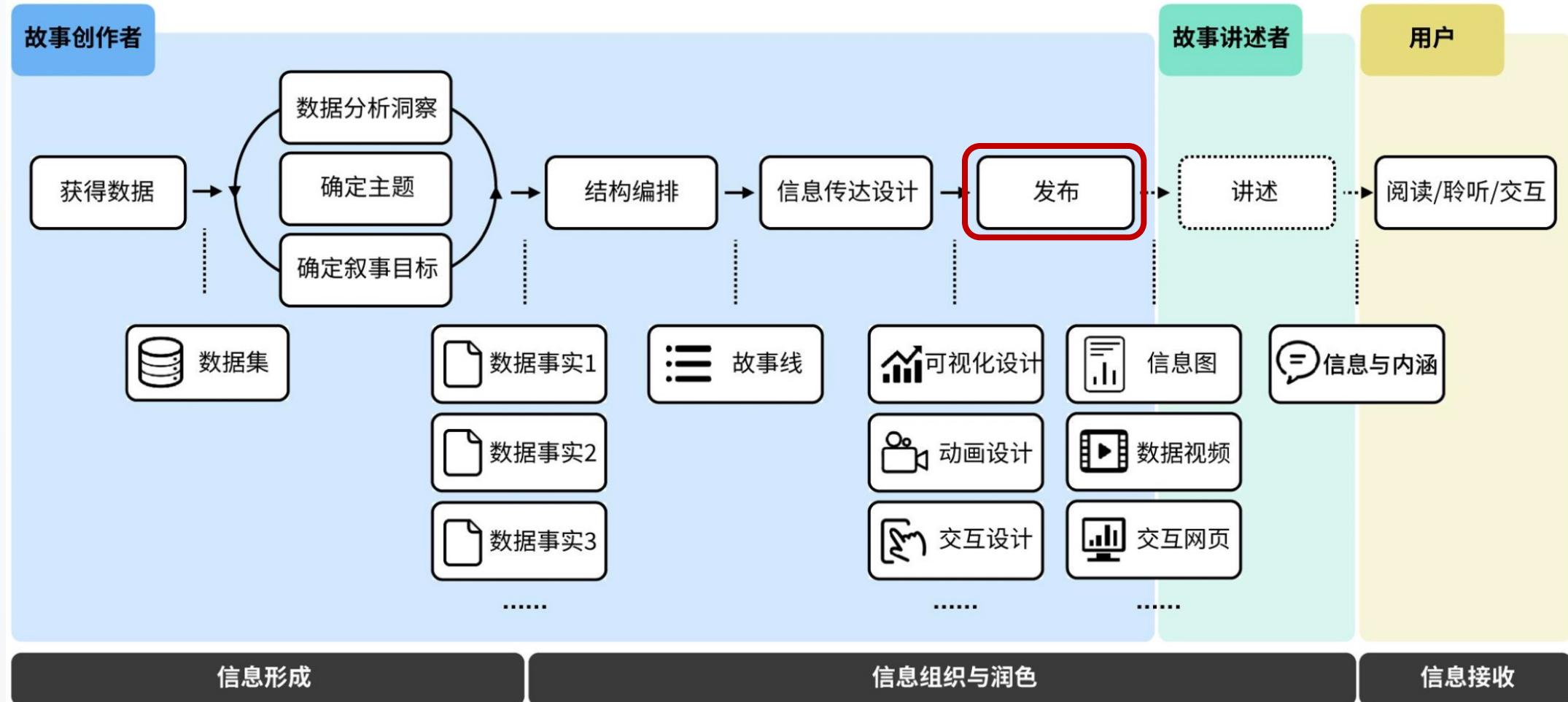
数据故事的制作流程



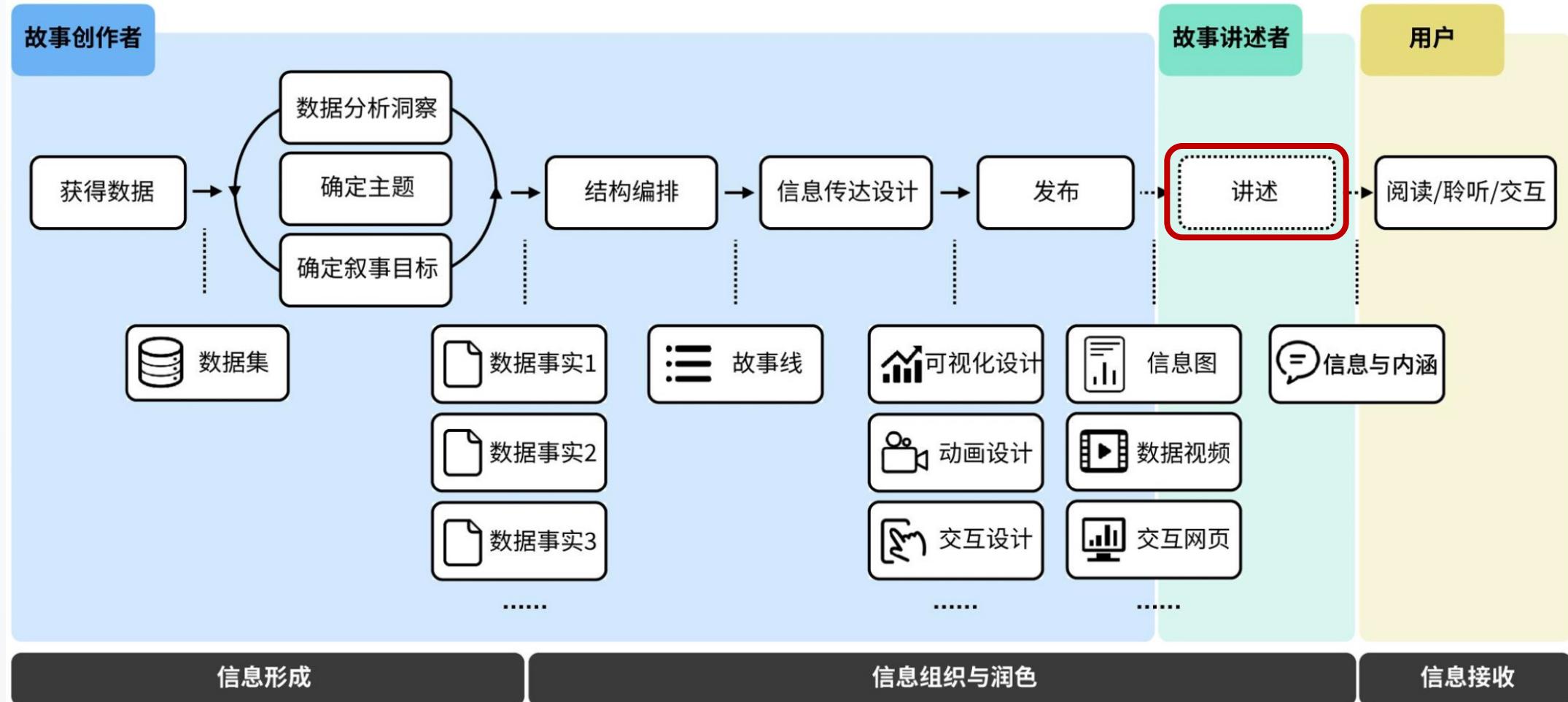
数据故事的制作流程



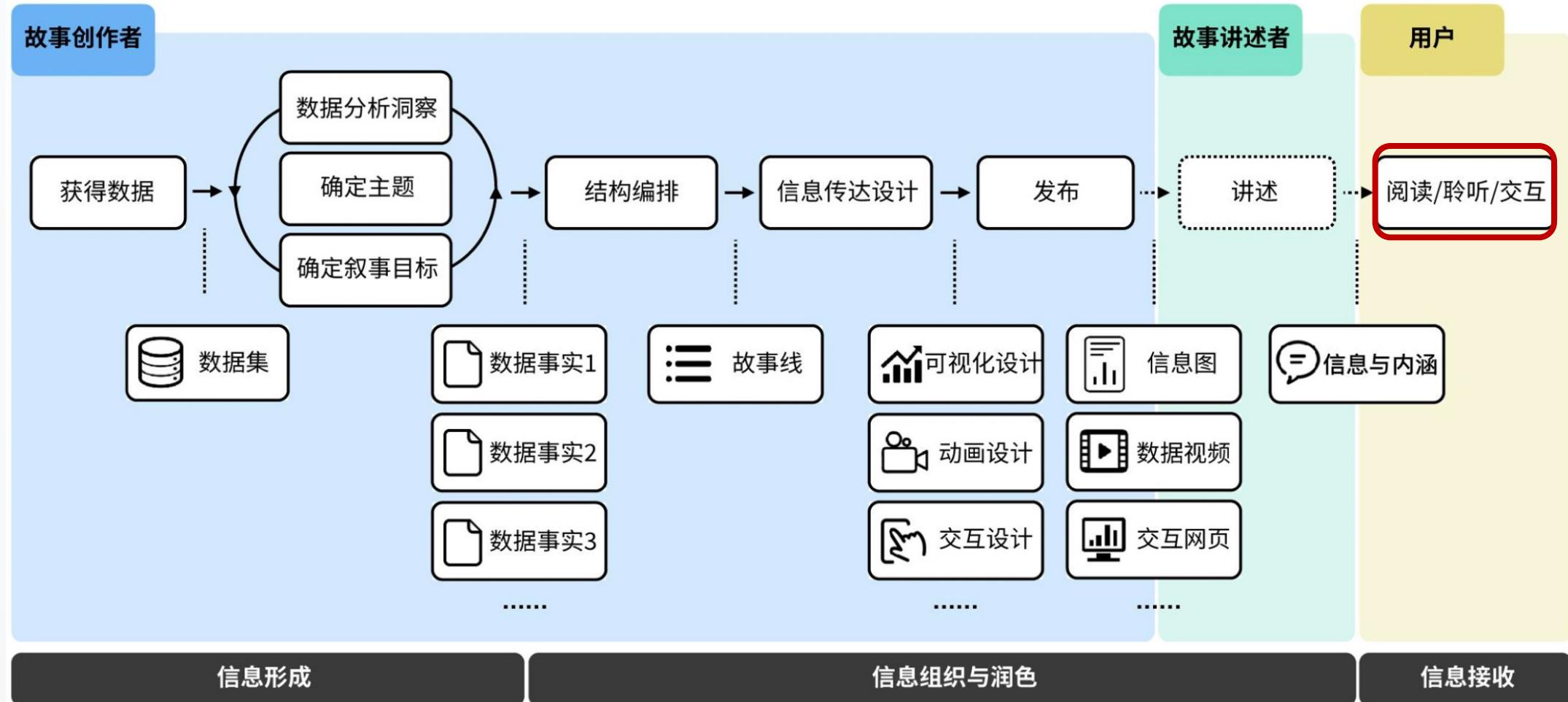
数据故事的制作流程



数据故事的制作流程



数据故事的制作流程



课程总结

- 本节课
 - 数据故事的六要素
 - 数据故事的体裁
 - 历史上具有叙事特征的可视化
 - 数据故事需要讲故事的人
 - 在讲述故事之前了解你的受众
 - 数据故事的应用领域
 - 数据故事的制作流程
- 下节课
 - 数据：数据故事的基础

课程大纲

- **数据：数据故事的基础**

- 数据故事的基石
- 找到有意义的数据洞察
- 分析的过程：从数据发现到数据解释
- 分析和传达需要遵循准则

数据故事的基石

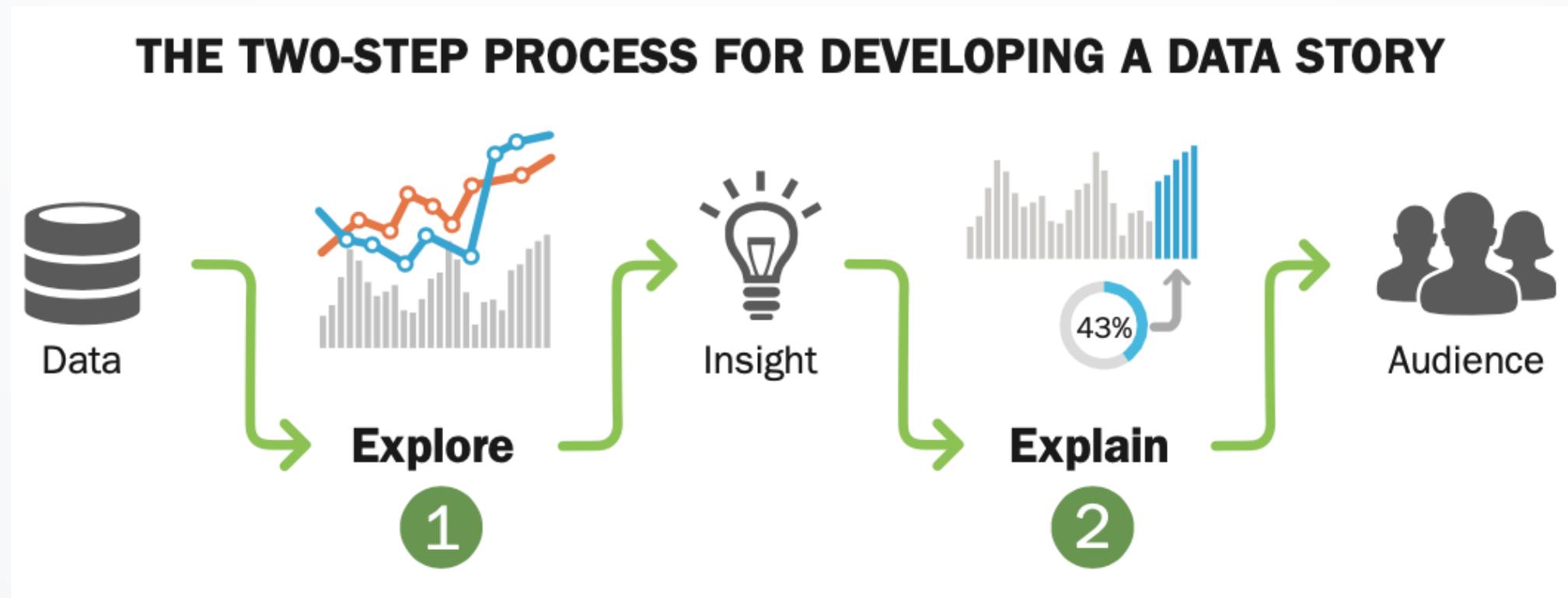
- 数据是每个数据故事的基本组成部分，从不相关或不可靠的数据中提取的见解可能会破坏一个精心制作的数据故事。
- **检查数据相关性**
 - 评估数据是否能够支撑起数据故事的宽度和高度。
 - 评估数据的新鲜度和及时性。
- **检查数据可靠性**
 - 评估数据源是否正确收集、处理和维护数据。
 - 评估分析过程中数据处理方式是否保证数据完整性、真实性。

找到有意义的数据洞察

- Avinash Kaushik提出三种方式评估每个数据洞察：
 - 为什么受众应该关心？
 - 价值性。洞察价值回报越大，受众采取行动越坚定。
 - 相关性。洞察与目标受众越相关，就越有可能得到关注，并让受众采取行动。
 - 受众应该怎么做？
 - 实用性。洞察是可行的，受众有能力采取行动。
 - 精确性。洞察可以明确告诉受众怎么去操作和行动。
 - 潜在的影响效果是什么？
 - 具体性。洞察是否能创造具体的效果。
 - 情境化。是否有背景信息或基准信息，帮助受众去理解洞察的意义和效果。需要伴随着足够的背景信息，以便受众能够充分理解其意义或独特性。

分析的过程：从数据发现到数据解释

- 数据分析是一个两步走的过程：数据发现和数据解释。数据故事的形成始于使用探索性数据可视化来发现见解。一旦发现了有意义的见解，就会使用解释性数据可视化来讲述故事。



分析的过程：从数据发现到数据解释

• 数据发现阶段

- 你就像一位大胆、好奇的考古学家在探索数据。
- 在这个阶段，您是这些探索性数据可视化的唯一受众。
- 你对数据有深入的了解。
- 可视化方面，可视化效果不必漂亮因为只是作为功能性工具帮助发现有价值的内容，更专注于它的速度和灵活性。

• 数据解释阶段

- 你就像一位考古学老师向学生解释你的发现。
- 你不再是受众，因为你试图与他人分享你的数据发现。
- 您的目标受众很有可能不会像您一样了解数据。
- 可视化方面，更专注于它的简单、清晰和凝聚力，以便更有效地将您发现的内容传达给不了解的受众。

THE TWO SIDES OF INDIANA JONES

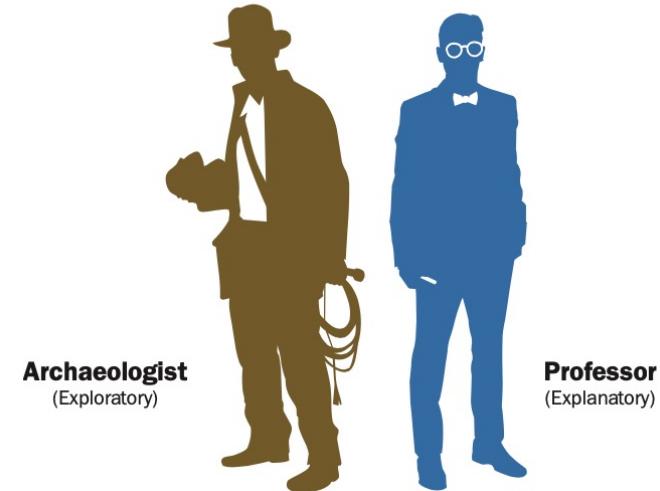


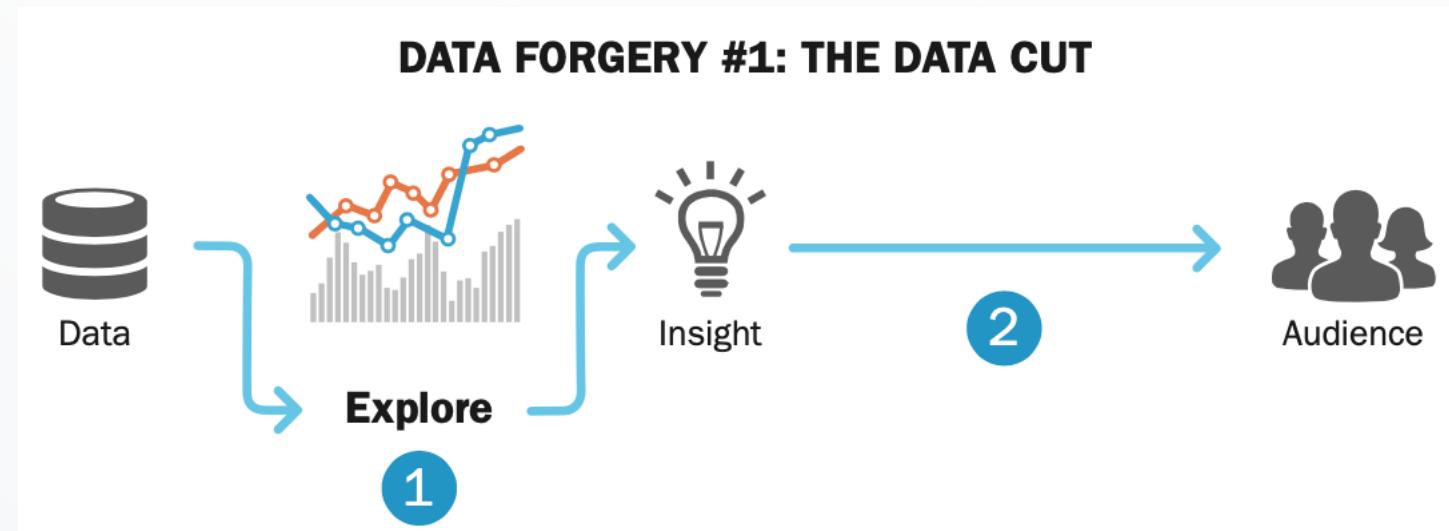
Table 5.1 Differences between the Two Stages of the Analysis Process

	Exploratory	Explanatory
Goal	Understand	Communicate
Audience	You	Other people
Data familiarity	Very familiar (you)	Less familiar (others)
Visualization focus	Flexibility and speed	Simplicity, clarity, and cohesion
Narrative	Unknown	Known
Outcome	Insight	Action

分析的过程：从数据发现到数据解释

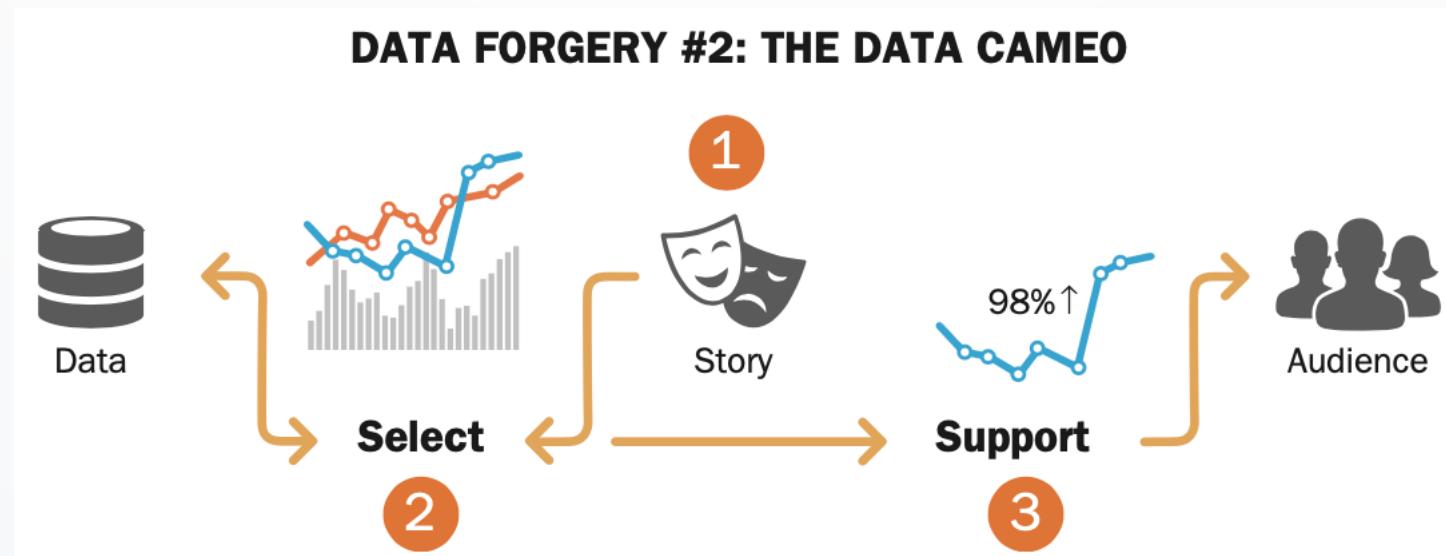
- **问题1：数据削减 (The Data Cut)**

- 数据专家容易遇到这样的问题。
- 发现数据有意义的见解后，没有任何解释地直接打包给受众，忽视了受众没有那么了解数据，忽视了拥有叙事和解释性视觉效果的重要性，以帮助他人更好地理解洞察力的意义。



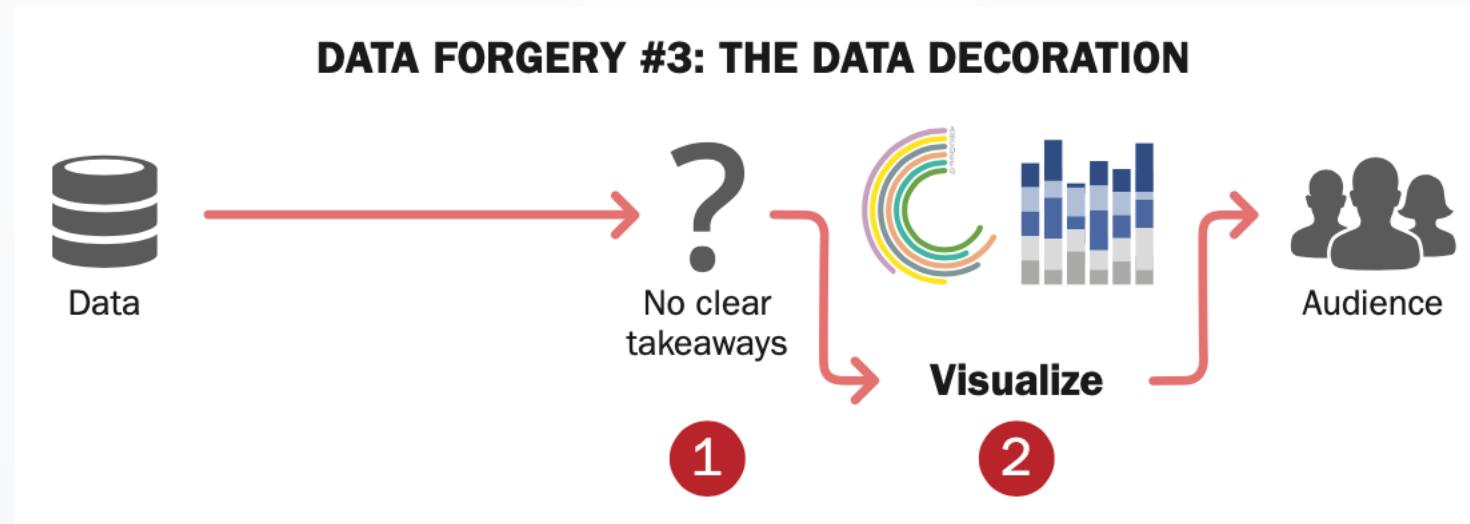
分析的过程：从数据发现到数据解释

- 问题2：数据客串（The Data Cameo）
 - 商业用户容易遇到这样的问题。
 - 叙事先入为主，只有支持叙事的数据点才被“选择”，而与叙事冲突的数据点被忽略。



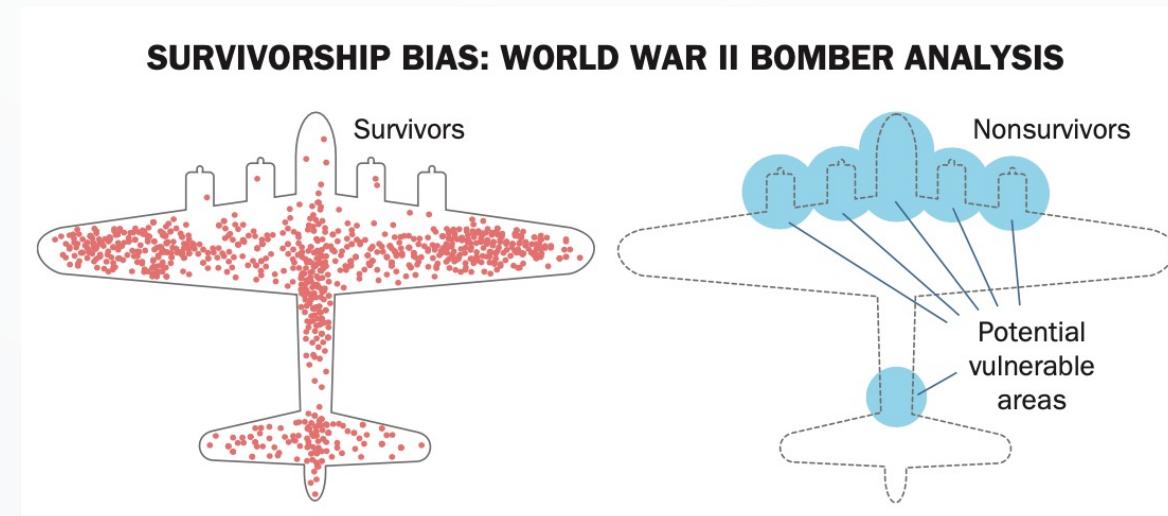
分析的过程：从数据发现到数据解释

- 问题3：数据装饰（The Data Decoration）
 - 简单的分享数据可视化效果而回避真正的分析工作。
 - 跳过形成洞察的数据发现阶段，直接形成数据可视化但缺乏数据故事的阶段。



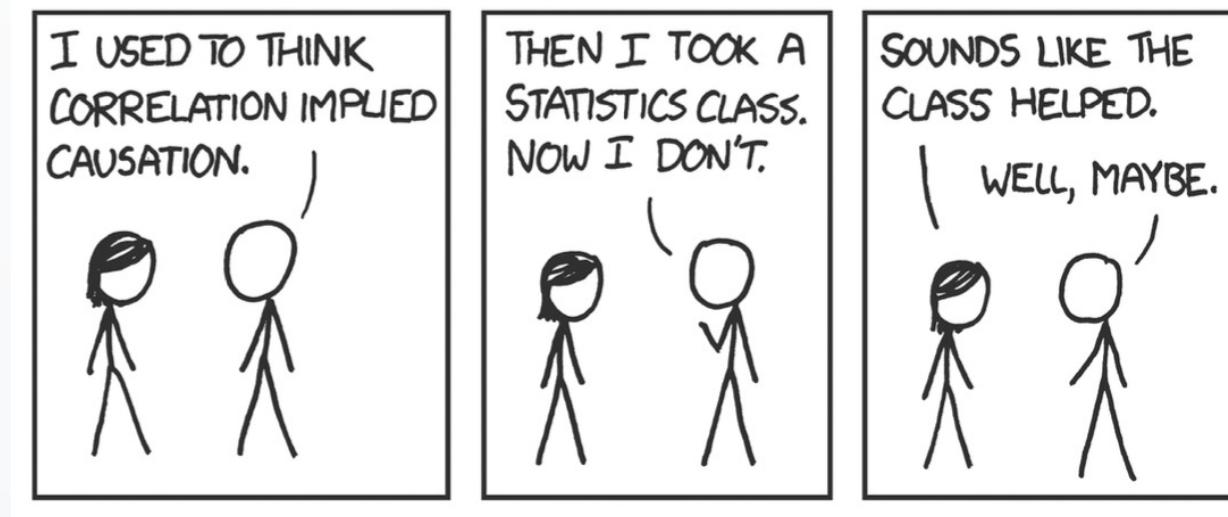
分析和传达数据需要遵循准则

- 三种认知偏差扭曲数据故事
 - **证实偏差**: 是一种只寻找和接受支持现有观点的证据的倾向，它会导致忽略与现有观点相左的相反信息。
 - **幸存者偏差**: 只关注成功或幸存的东西，忽略那些没有幸存的，造成错误的结论。
 - **知识的诅咒**: 和受众沟通内容时倾向于假设受众已经知道相关必要的背景或知识。



分析和传达数据需要遵循准则

- 三种逻辑谬误侵蚀叙事的数据基础
 - **相关性谬误**: 错误地将相关事件视为因果关系的一部分。
 - **德克萨斯神枪手谬误**: 为了支持自己的论点, 在大量的数据或证据中小心地选出对自己的观点有利的证据, 而不使用那些对自己不利的数据或证据。
 - **草率概括谬误**: 提出没有经过充分证明的广泛主张。



课程总结

- 本节课
 - 数据故事的基石
 - 找到有意义的数据洞察
 - 分析的过程：从数据发现到数据解释
 - 分析和传达需要遵循准则
- 下节课
 - 叙事：数据故事的结构