

# 经济数据解读 02（丽湖校区）

## 第二次线下课

黄嘉平

中国经济特区研究中心

<https://huangjp.com/teaching/econdata>

2026.4.9



# 小组汇报

**汇报内容：**选择一个经济、金融、民生或其他社会领域的热门话题进行分享。要**利用数据**分析现象，注重**历年数据的变化**，挖掘隐藏在现象**背后的故事**。注意所选数据的**真实性和实效性**，并选择**合适的方式展示**数据和分析结果。

- 例如：美国的对华出口管制措施及其影响。此题目可以包括政策解读、实际影响、中国的应对措施、今后的发展趋势等内容。
- 最终选课人数为 **23** 人，共组成 **12** 队。后三周每周有 4 队进行汇报，每队可用时间为 20 分钟，其中汇报时间为 17 分钟，留出 3 分钟点评时间。
- 队员间的分工形式由全队商议决定。原则上以队为单位打分，每个人的得分相同。如果出现全程不参与的人，其他队员可汇报给助教，经助教确认事实后，取消不参与者的汇报成绩。1 人队会在基础得分上适当加分。

# 关于小组汇报的分组

- **分组第一阶段**：自由组队。在 4月17日（星期五）前报给助教姜宏卓（通过私聊或其他方式）。这一阶段只能报 2 人队。
- **分组第二阶段**：第一阶段没有组队的同学，将由助教随机组队。随机组队会尽量考虑同一校区的同学。
- **最终分组结果和汇报顺序**将在 4月23日 第三次线下课上公布。
- 在最终分组公布后原则上不再接受人员和汇报时间调整，确实需要调整的小组可先在课下和其他组自行协调，双方同意后将调整结果报给助教。确实存在沟通困难情况的可以和助教协商。

# 数据的可视化

# 统计数据可视化

Edward R. Tufte 在其著作 *The Visual Display of Quantitative Information* 中写道：

优秀的统计数据可视化应当将复杂的内容通过清晰、准确、有效的方式呈现。

统计图表需要

- **展示数据**
- 使读者关注事实，而不是方法、图形设计、绘图技术等其他方面
- **避免扭曲数据包含的信息**
- 在小空间内展示很多数值
- 在大数据集上体现一致性
- 鼓励读者比较数据的不同部分
- 从整体到细节展示数据的不同层面
- 有清晰合理的目的：解释、探索、展示、或装饰
- 与数据的统计学或语言描述机密结合

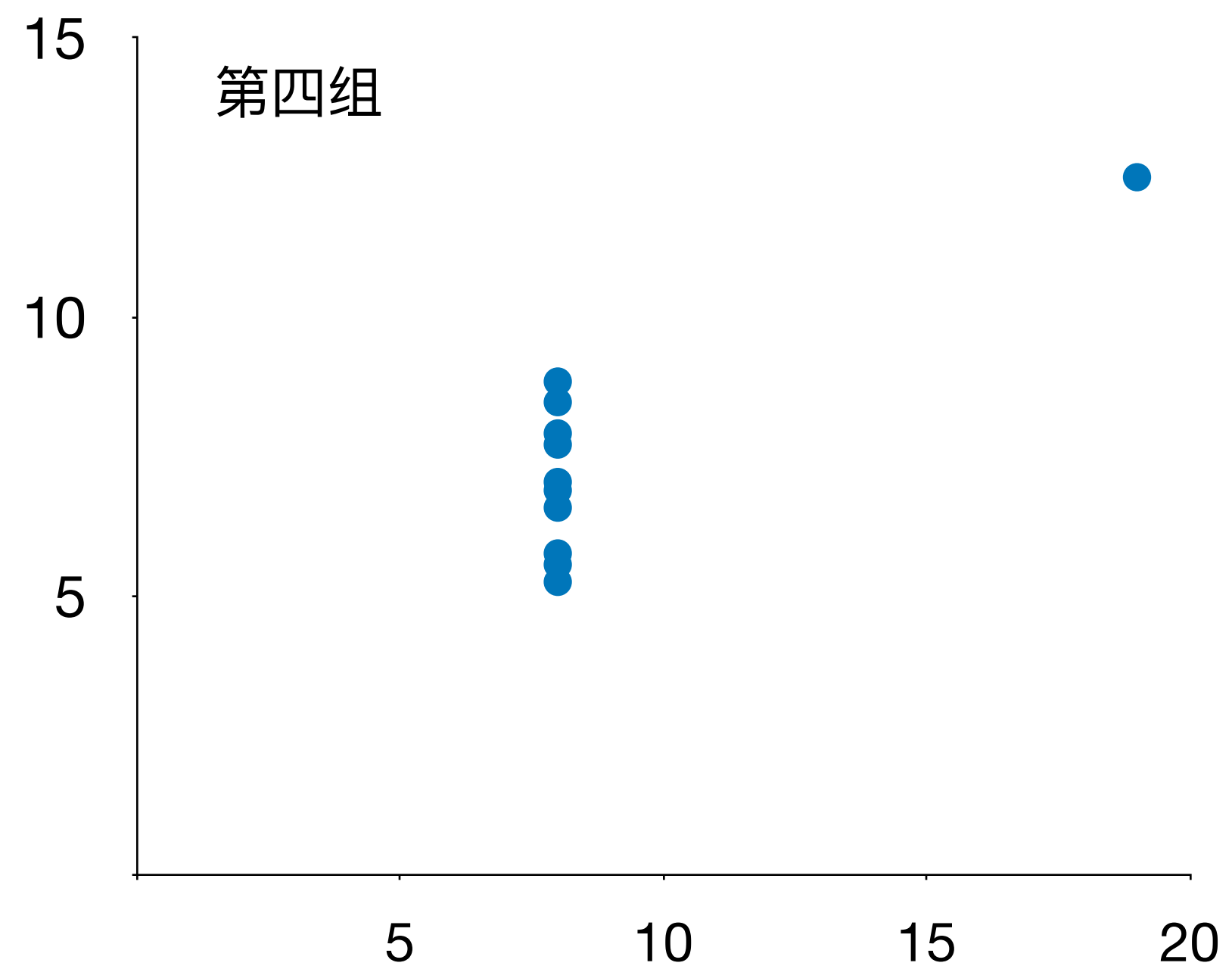
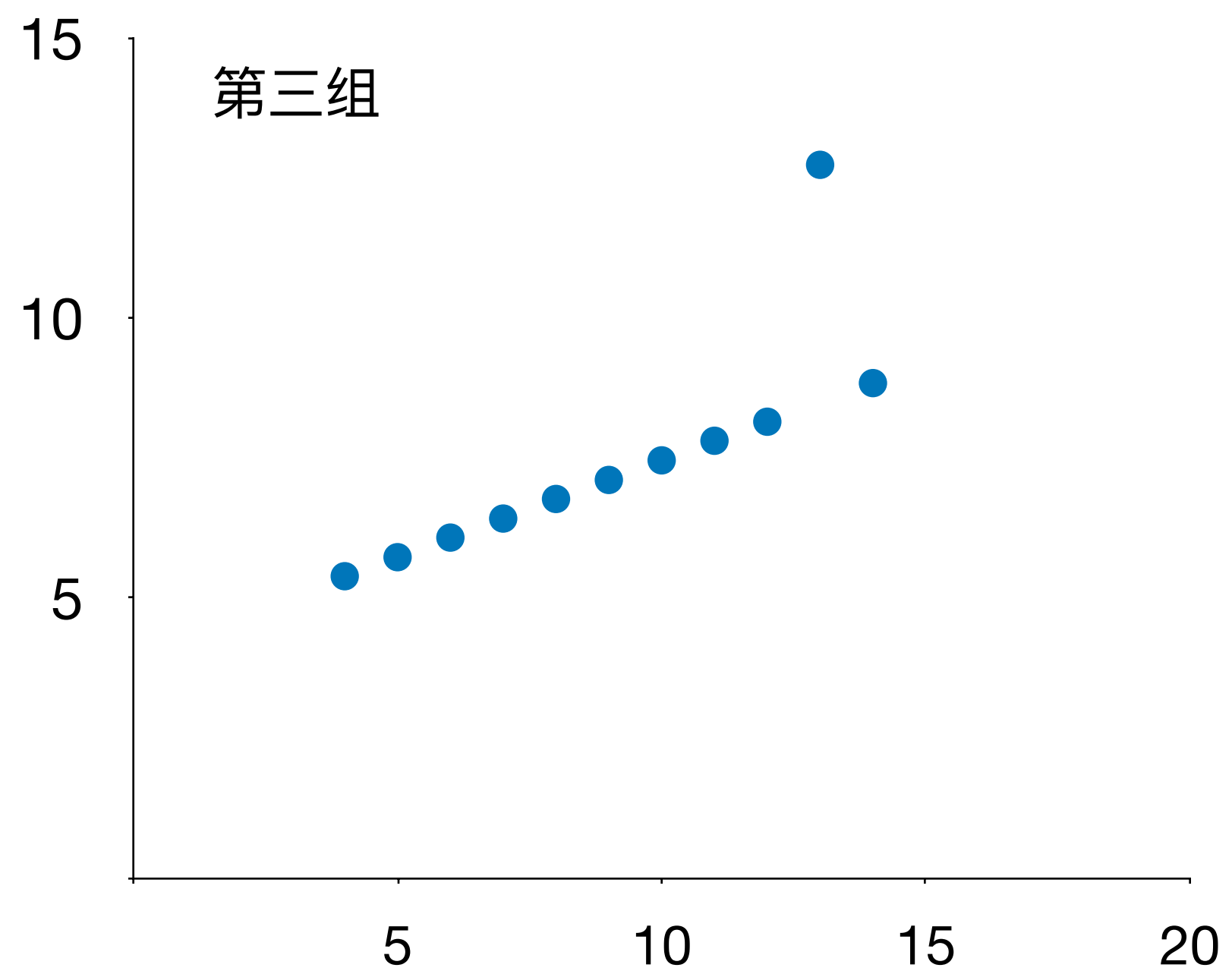
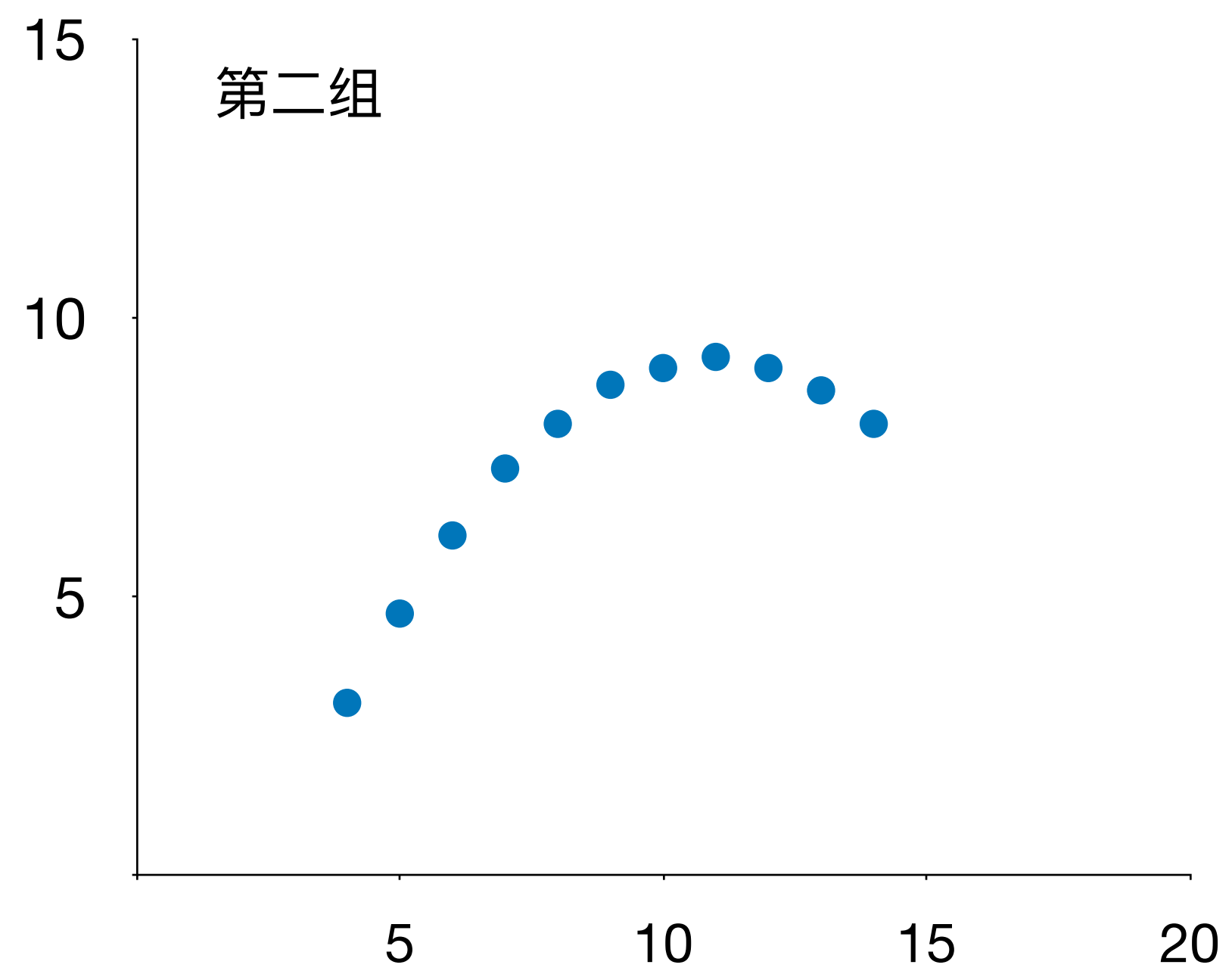
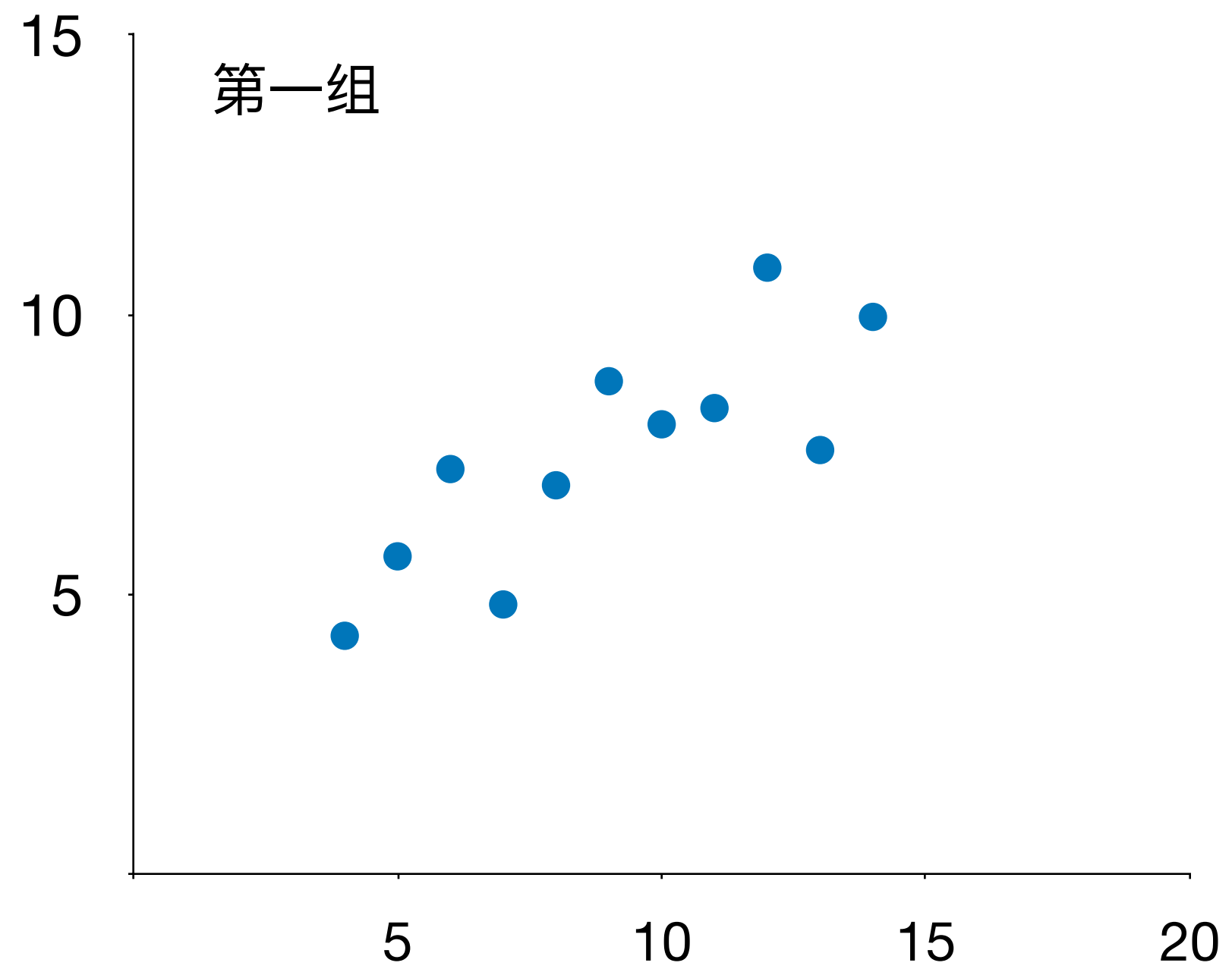
# 图表有时候比表格更有效

Anscombe, F. J. (1973). Graphics in Statistical Analysis. *American Statistician*, 27:17-21.

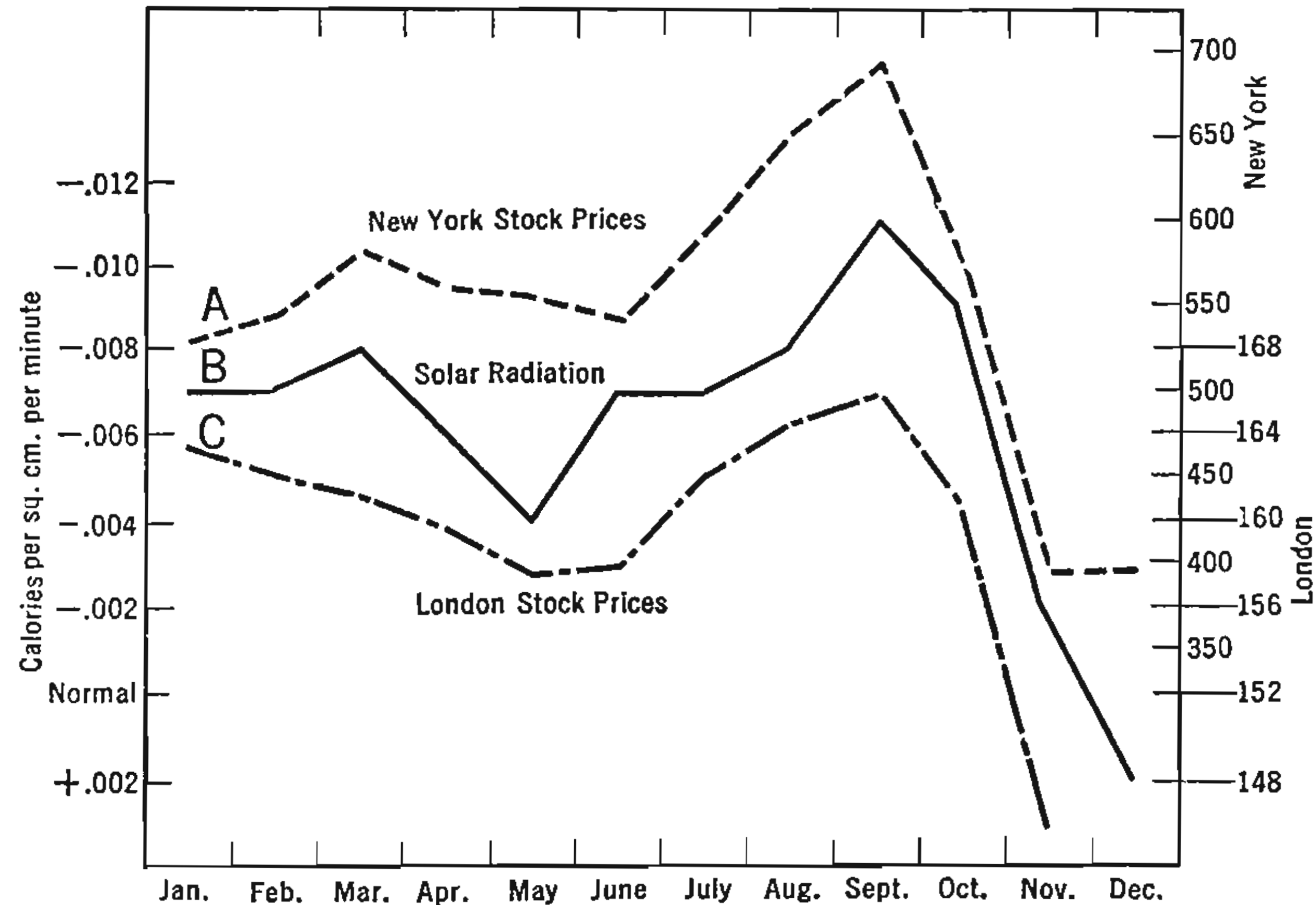
第一组		第二组		第三组		第四组	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.1	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.1	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.7	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.8	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.3	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.1	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.1	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.1	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.1	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.3	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.7	5.0	5.73	8.0	6.89

四组数据的描述性统计一致：

- 样本量 = 11
- $X$  的均值 = 9.0,  $Y$  的均值 = 7.5
- 线性回归函数  $Y = 3 + 0.5X$
- 斜率的估计标准误 = 0.118 ( $t = 4.24$ )
- $X - \bar{X}$  的平方和 = 110.0
- 回归平方和 = 27.50
- 残差平方和 = 13.75
- 相关系数 = 0.82
- $R^2 = 0.67$



# 图表无法填补数据或模型本身的不足



SOLAR RADIATION AND STOCK PRICES

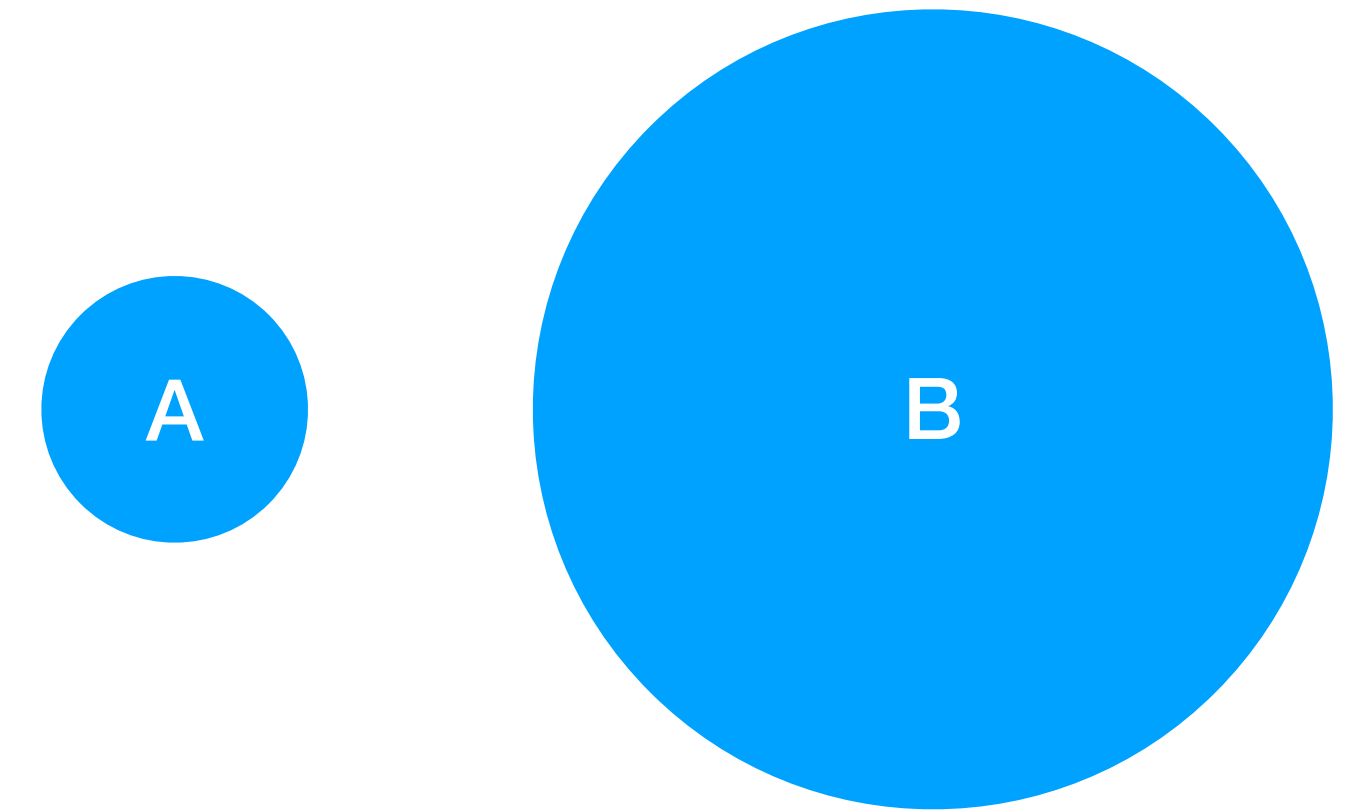
A. New York stock prices (Barron's average). B. Solar Radiation, inverted, and C. London stock prices, all by months, 1929 (after Garcia-Mata and Shaffner).

Dewey, E. R. & Dakin, E. F. (1947). *Cycles: The Science of Prediction*, p144. New York.

图片取自 Tufte, E. R. (2001). *The Visual Display of Quantitative Information*, 2nd Edition, p15.

# 不能用数据骗人

- 我们都知道数据会“骗人”，可视化后的数据同样会“骗人”。
- 图表不应当扭曲数据中的信息，但是不同的人对同一图形的感知可能存在差异。
- 那么我们应该为每一个人量身定做一幅图吗？或是以读者的平均感知水平为标准？
- Tufte 认为，可视化应该满足两个原则：
  1. 图中的物理测度（长度、面积等）应当和它所展示的数值成正比
  2. 应当有清晰详细全面的标注，以减少读者对数据的误解。例如在图中加入对数据的说明，或标注重要的事件等。
- “欺骗系数” = 图中展示的效果大小 / 数据中的效果大小



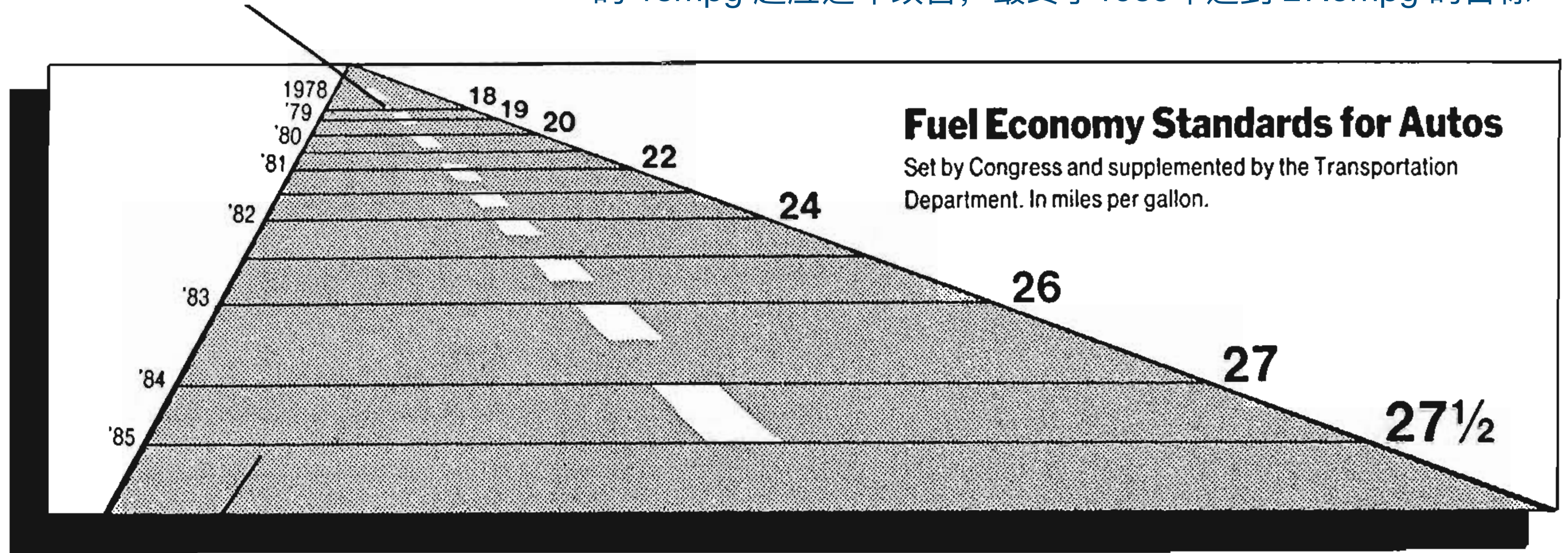
B 的面积是 A 的几倍？

- 8 倍
- 9 倍
- 10 倍
- 11 倍

*New York Times*, August 9, 1978, p. D-2.

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.

美国国会和交通部对新车的燃油经济性曾作出规定，从1978年的 18mpg 起应逐年改善，最终于1985年达到 27.5mpg 的目标



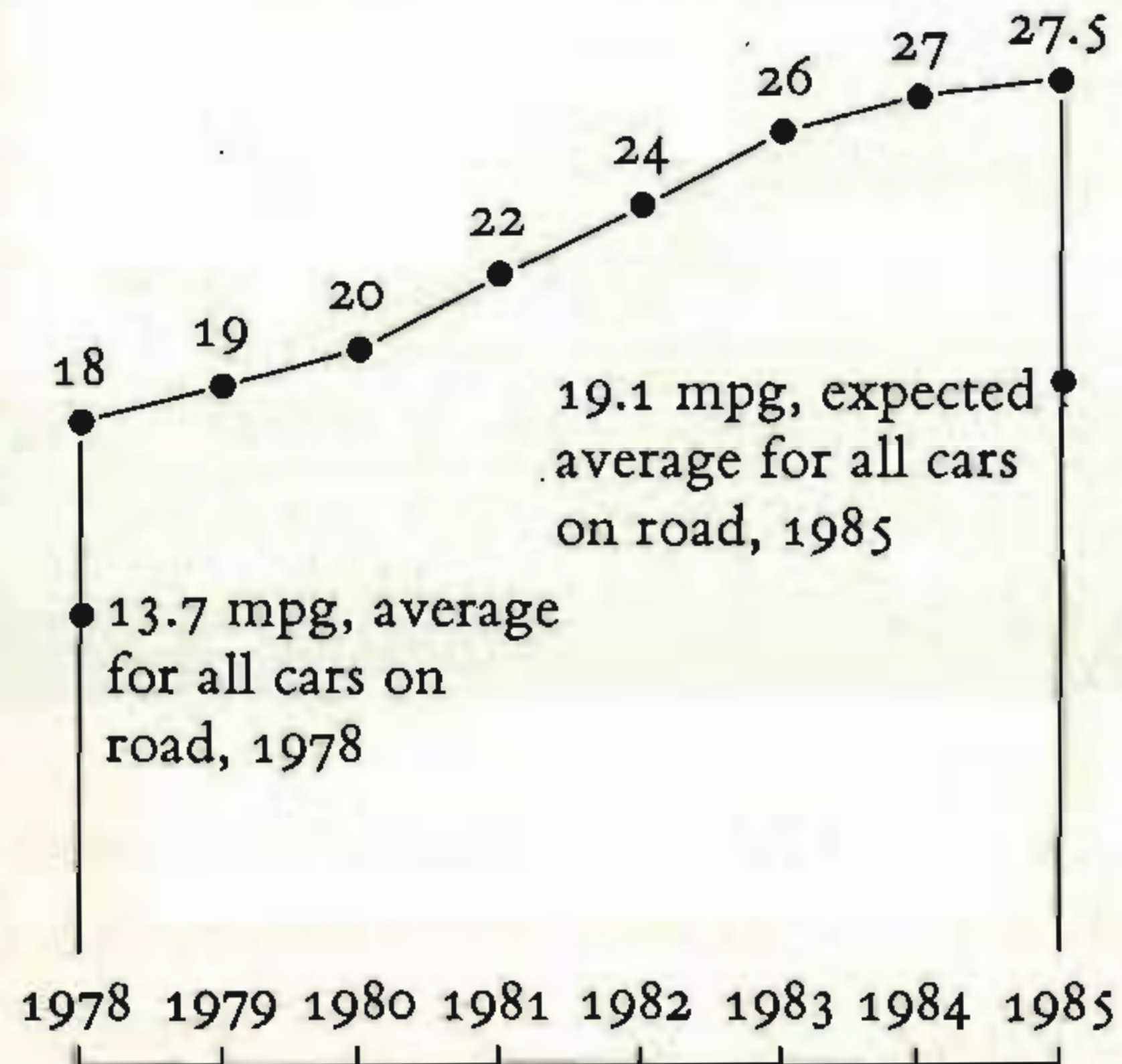
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

真实增长率为  $(27.5 - 18) / 18 = 0.53$

图中展示的增长率为  $(5.3 - 0.6) / 0.6 = 7.83$

欺骗系数 =  $7.83 / 0.53 = 14.8$

REQUIRED FUEL ECONOMY STANDARDS:  
NEW CARS BUILT FROM 1978 TO 1985



### Tufte 对前图的改善（无欺骗版）

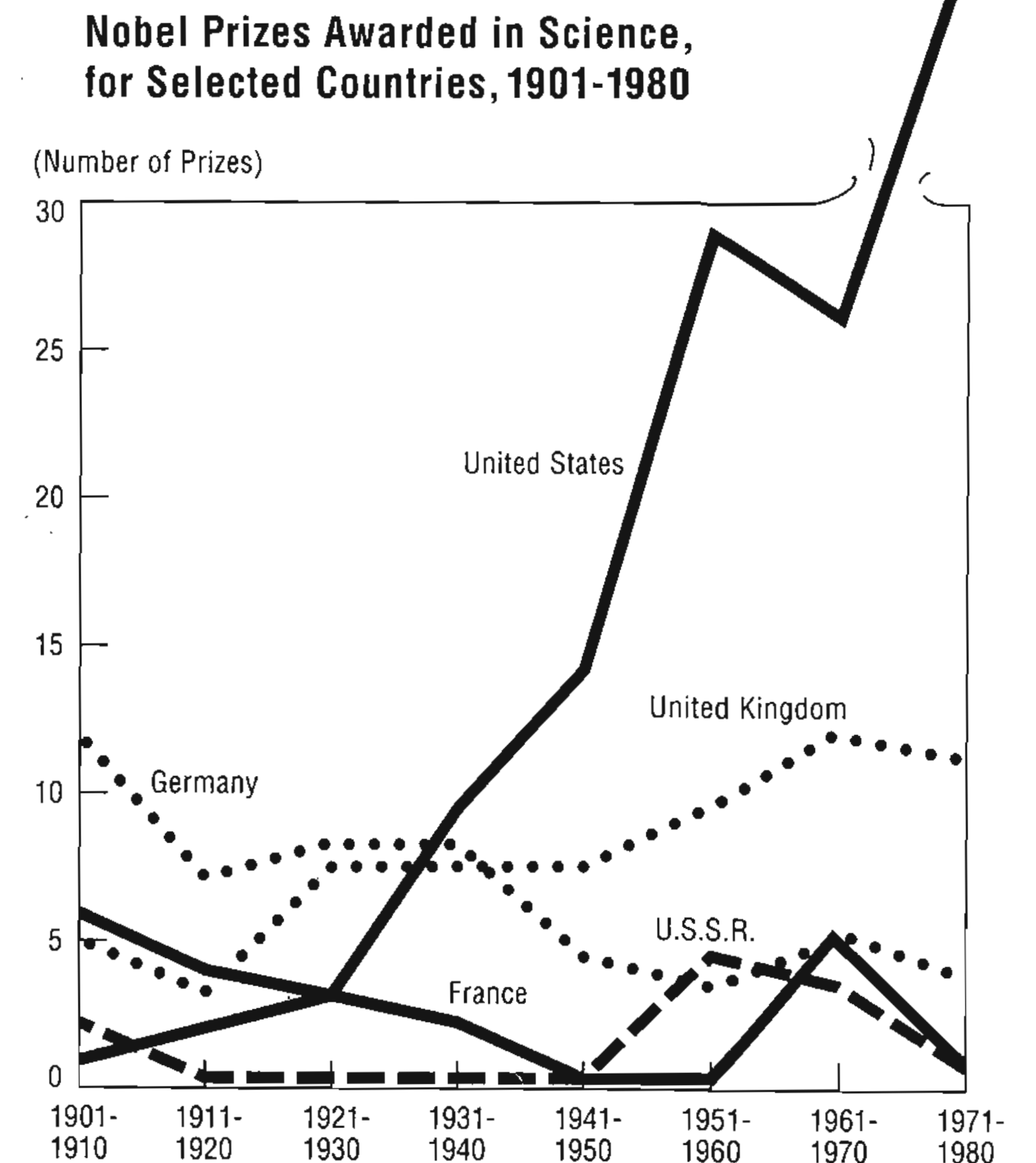
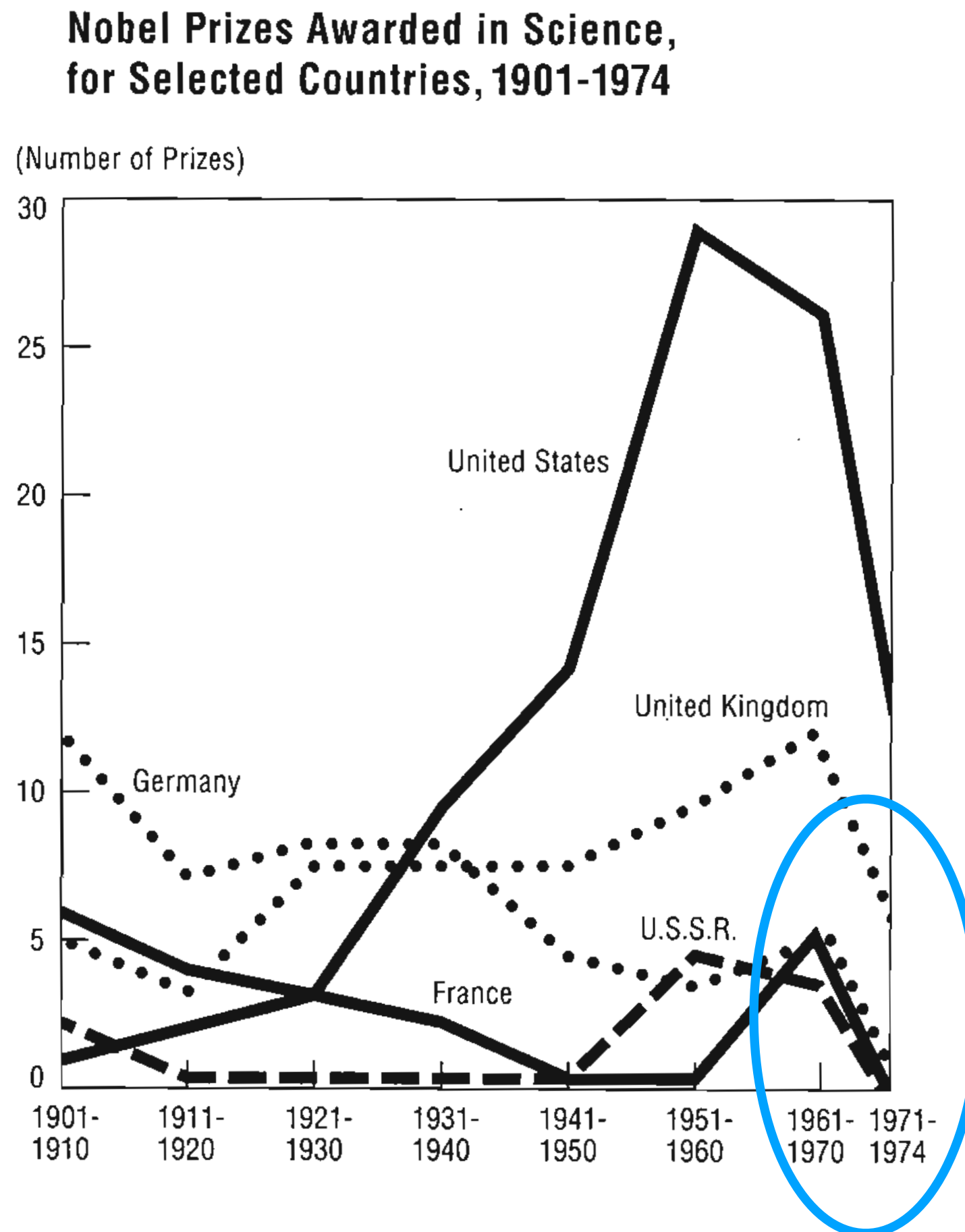
除了准确展示了数字，此图的改进点还包括

- 加入了对所有车辆平均燃油经济性的估计值（标准只针对新车）
- 时间发展从左到右，符合阅读习惯
- 年份间隔相同
- 折线图更好的展示了1980-1983年间标准变化率的增大

# 展示数据的变化，而不是图形设计的变化

- 人们会根据图的一部分预测其他部分，因此图表的各部份之间应该保持一致
- 左图的横轴最右侧的跨度为4年，与其他部分的10年不一致
- 右图是将数据补充完整后展示的正确信息（即使没有准确的数据，也应将4年间的的数据乘以2.5进行估计）
- 左图的问题完全来自图形设计上的变化，而不是数据本身的变化

National Science Foundation, *Science Indicators, 1974* (Washington, D.C., 1976), p. 15.



**应该避免的错误**

**从《每日经济新闻》汲取教训**

# 媒体中的可视化实践

- *Economist, Financial Times, Wall Street Journal* 等报纸/杂志在数据可视化实践上走在前列。

- 《每日经济新闻》是成都传媒集团主管主办的报刊，在数据可视化方面也非常积极，但缺乏对细节的追求，存在很多具有代表性的错误。

## 每日经济新闻

NATIONAL BUSINESS DAILY

2024.09.27 星期五 今日8版 第4803期

网址 www.nbd.com.cn 服务热线 400-868-0008

### 中共中央政治局召开会议 分析研究当前经济形势和经济工作 中共中央总书记习近平主持会议

新华社北京9月26日电 中共中央政治局9月26日召开会议，分析研究当前经济形势，部署下一步经济工作。中共中央总书记习近平主持会议。

会议认为，今年以来，以习近平同志为核心的中央团结带领全党全国各族人民攻坚克难、沉着应对，加大宏观调控力度，着力深化改革开放，扩大内需战略、优化经济结构、推动经济高质量发展取得积极成效。要全面贯彻党的二十届三中全会部署，坚持稳中求进工作总基调，完整、准确、全面贯彻新发展理念，着力扩大内需、深化改革开放、促进高质量发展，推动经济高质量发展取得积极成效。

会议提出，我国经济发展的基本面没有变，要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。

会议强调，要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。

### 涨3.61%! 沪指突破3000点 本周内沪指已上涨近10%，距离5月高点还差约6%

本周内沪指已上涨近10%，距离5月高点还差约6%

### 股市持续上涨有两方面原因

光大证券研报指出，9月26日，A股市场放量上涨，上证指数、创业板指、万得全A分别上涨3.61%、4.24%、4.01%。值得注意的是，国际油价大幅回落，带动大宗商品价格上涨。近期股市上涨主要是受以下两方面原因驱动：一是政策预期超预期。其次，市场预期分化加剧，市场已经处于超买阶段，个股处于估值洼地，具有较好的投资价值。

沪深两市行业涨幅前十	涨幅
银行	9.40%
非银金融	8.88%
食品饮料	8.59%
医药生物	5.40%
家用电器	5.32%
通信	5.29%
国防军工	5.19%
有色金属	5.02%
水利电力	4.99%
农林牧渔	4.83%

### 以市值管理“硬约束” 推动上市公司治理变革

19.23万亿元，如果这些个股净资产归一，则市值将增加至约2600亿元。

《指引》提出，长期股权投资公司应加强对被投企业的市值管理，推动其提升市值管理水平。要加强对被投企业的市值管理，推动其提升市值管理水平。要加强对被投企业的市值管理，推动其提升市值管理水平。

### 做好“基层”“三保”工作

会议强调，要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。

## 每日经济新闻

02

### 中央定调促进房地产市场止跌回稳 对商品房建设要严控增量、优化存量、提高质量

新华社北京9月26日电 中共中央政治局9月26日召开会议，分析研究当前经济形势，部署下一步经济工作。中共中央总书记习近平主持会议。

会议认为，今年以来，以习近平同志为核心的中央团结带领全党全国各族人民攻坚克难、沉着应对，加大宏观调控力度，着力深化改革开放，扩大内需战略、优化经济结构、推动经济高质量发展取得积极成效。要全面贯彻党的二十届三中全会部署，坚持稳中求进工作总基调，完整、准确、全面贯彻新发展理念，着力扩大内需、深化改革开放、促进高质量发展，推动经济高质量发展取得积极成效。

会议提出，我国经济发展的基本面没有变，要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。

会议强调，要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。要坚持以高质量发展为首要任务，着力扩大内需、深化改革开放、促进高质量发展。

### 去年全国地表水优良水质断面比例达到89.4% 生态环境部答每经问：已超出“十四五”目标4.4个百分点

生态环境部答每经问：已超出“十四五”目标4.4个百分点

### 奋进强国路 阔步新征程

奋进强国路 阔步新征程

### 水环境治理持续深化

水环境治理持续深化

### 丰巢探索更丰富盈利模式

丰巢探索更丰富盈利模式

### Rokid AR眼镜面临“淘汰赛”

Rokid AR眼镜面临“淘汰赛”

### 银行信用卡业务，何时改观？

银行信用卡业务，何时改观？

### 和府捞面门店总数不减反增

和府捞面门店总数不减反增

### “蓝天保卫战”取得成效

“蓝天保卫战”取得成效

### 奋进强国路 阔步新征程

02

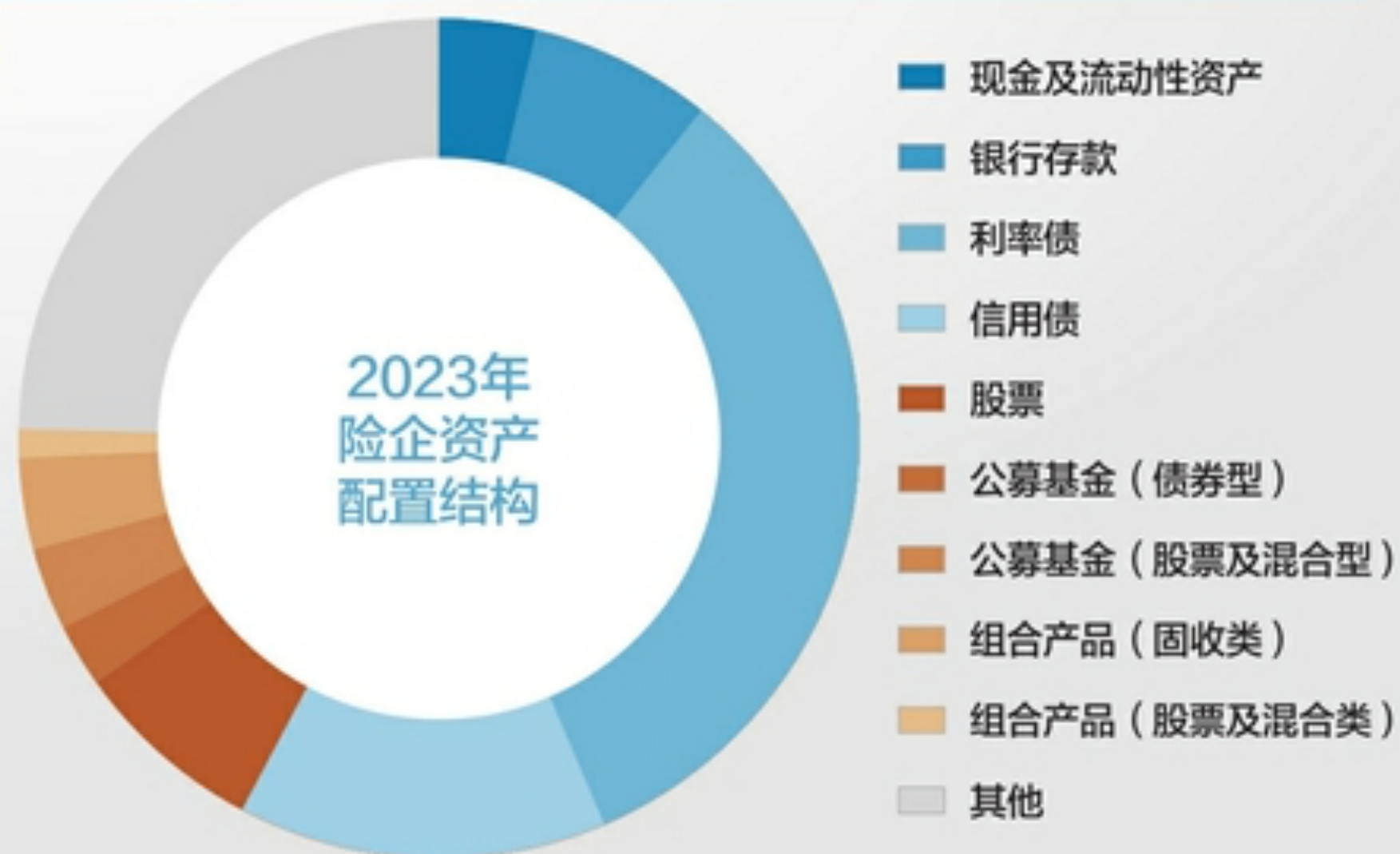
69243658000109

2024年9月30日

# 保险资金运用余额达31.8万亿元

9月27日，国新办举行国务院政策例行吹风会。金融监管总局人身保险监管司司长罗艳君表示，保险资金是比较典型的中长期资金。到今年8月末，保险资金运用余额31.8万亿元，同比增长10.4%，保险业通过债权、股权等多种方式为实体经济提供资金支持28.8万亿元，同比增长12.2%。

## 保险资金运用规模持续增长

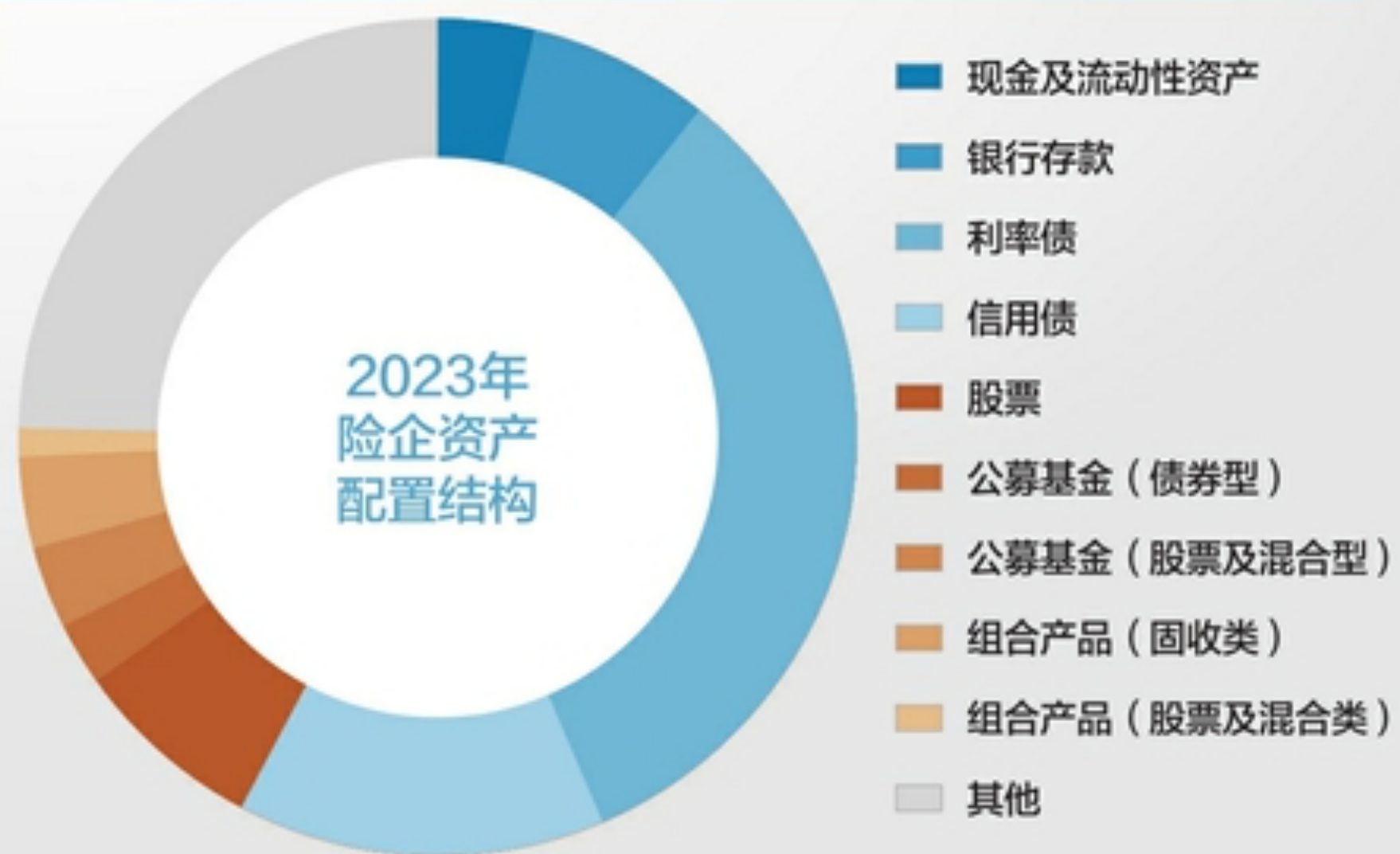


2024年9月30日

# 保险资金运用余额达31.8万亿元

9月27日，国新办举行国务院政策例行吹风会。金融监管总局人身保险监管司司长罗艳君表示，保险资金是比较典型的中长期资金。到今年8月末，保险资金运用余额31.8万亿元，同比增长10.4%，保险业通过债权、股权等多种方式为实体经济提供资金支持28.8万亿元，同比增长12.2%。

保险资金运用规模持续增长



金额宜用柱形图表示

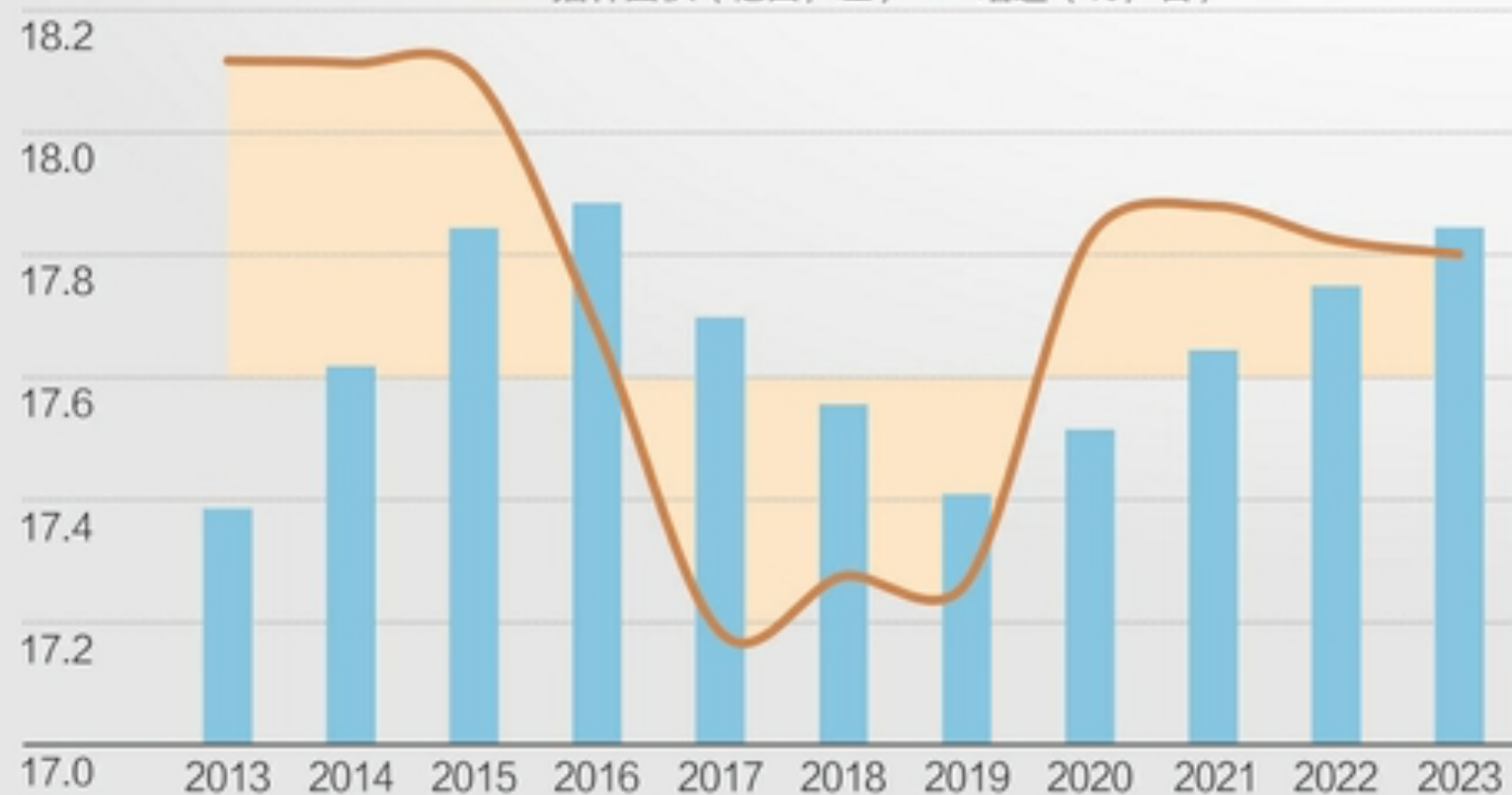
没有标注数据值

# 落实大食物观工作取得积极成效

9月19日，国家发展改革委有关负责人在回答《每日经济新闻》记者提问时表示，国家发展改革委坚持以保障粮食安全为底线，多途径开发食物来源，落实大食物观工作取得积极成效。比如，粮食生产能力持续提高。2023年粮食总产量达到1.39万亿斤，连续9年稳定在1.3万亿斤以上；食物供给更加丰富多元，水产品产量7116万吨，稳居全球首位。

### 我国粮食播种面积及同比增速

■ 播种面积 (亿亩, 左) ■ 增速 (% , 右)

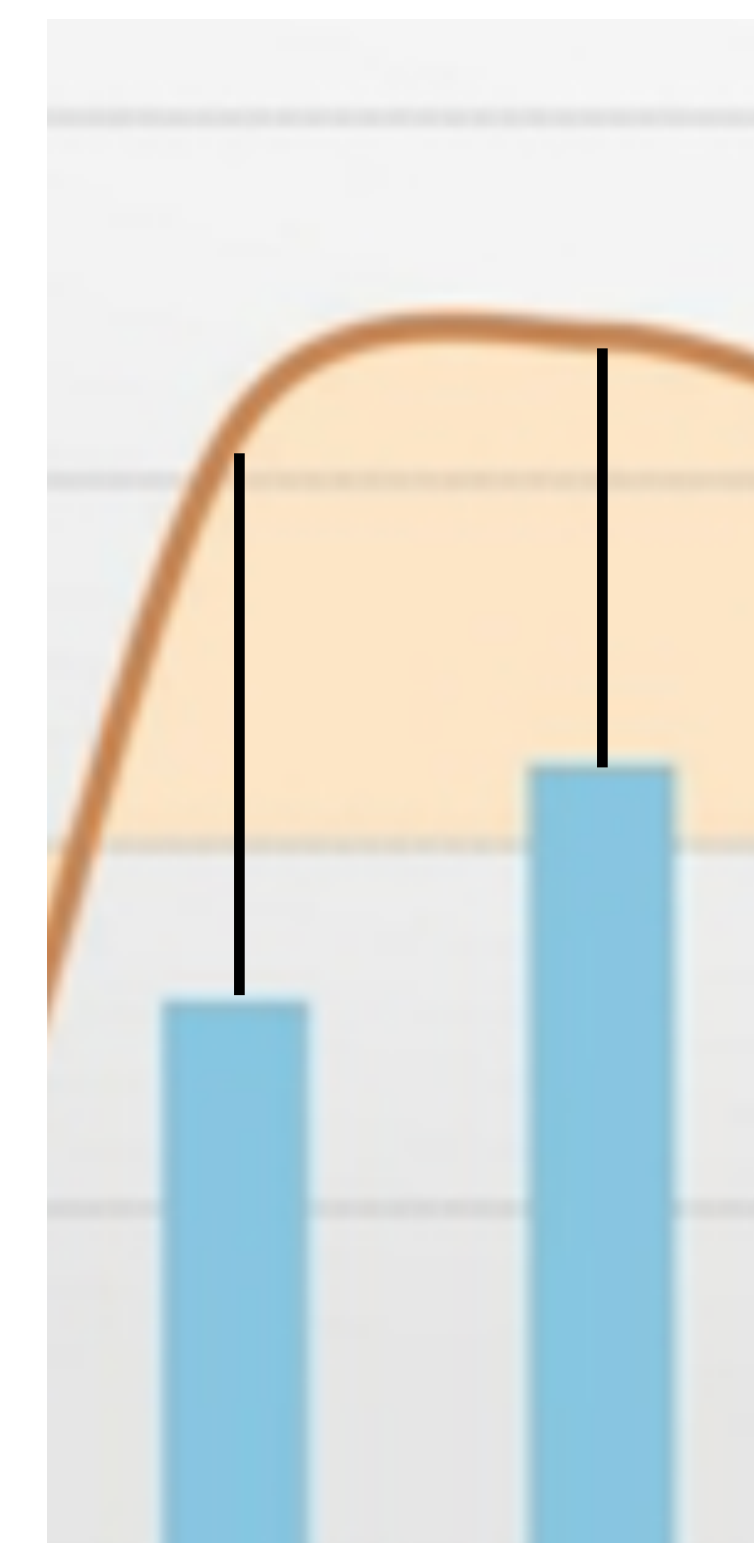
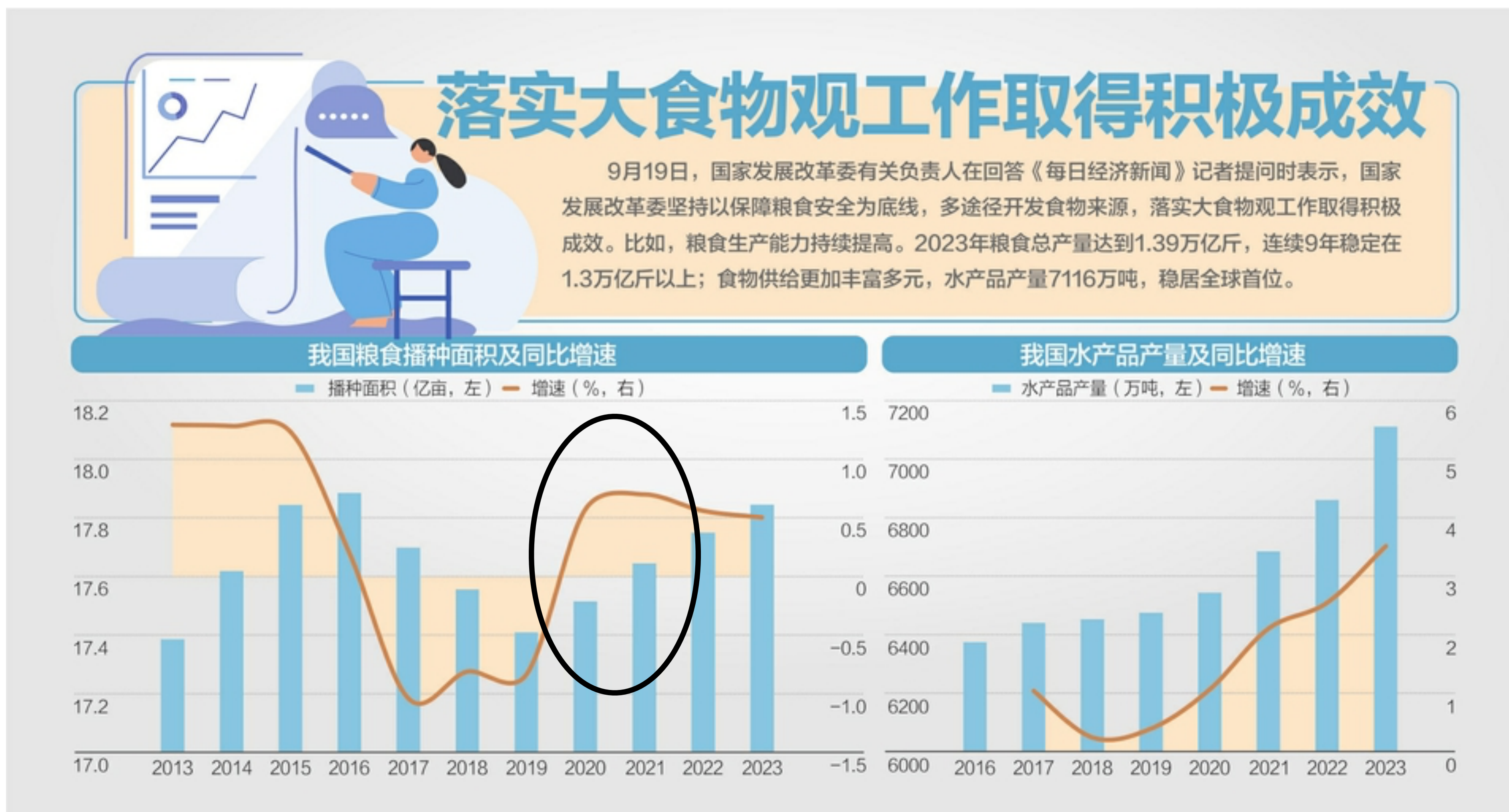


### 我国水产品产量及同比增速

■ 水产品产量 (万吨, 左) ■ 增速 (% , 右)



2024年9月20日



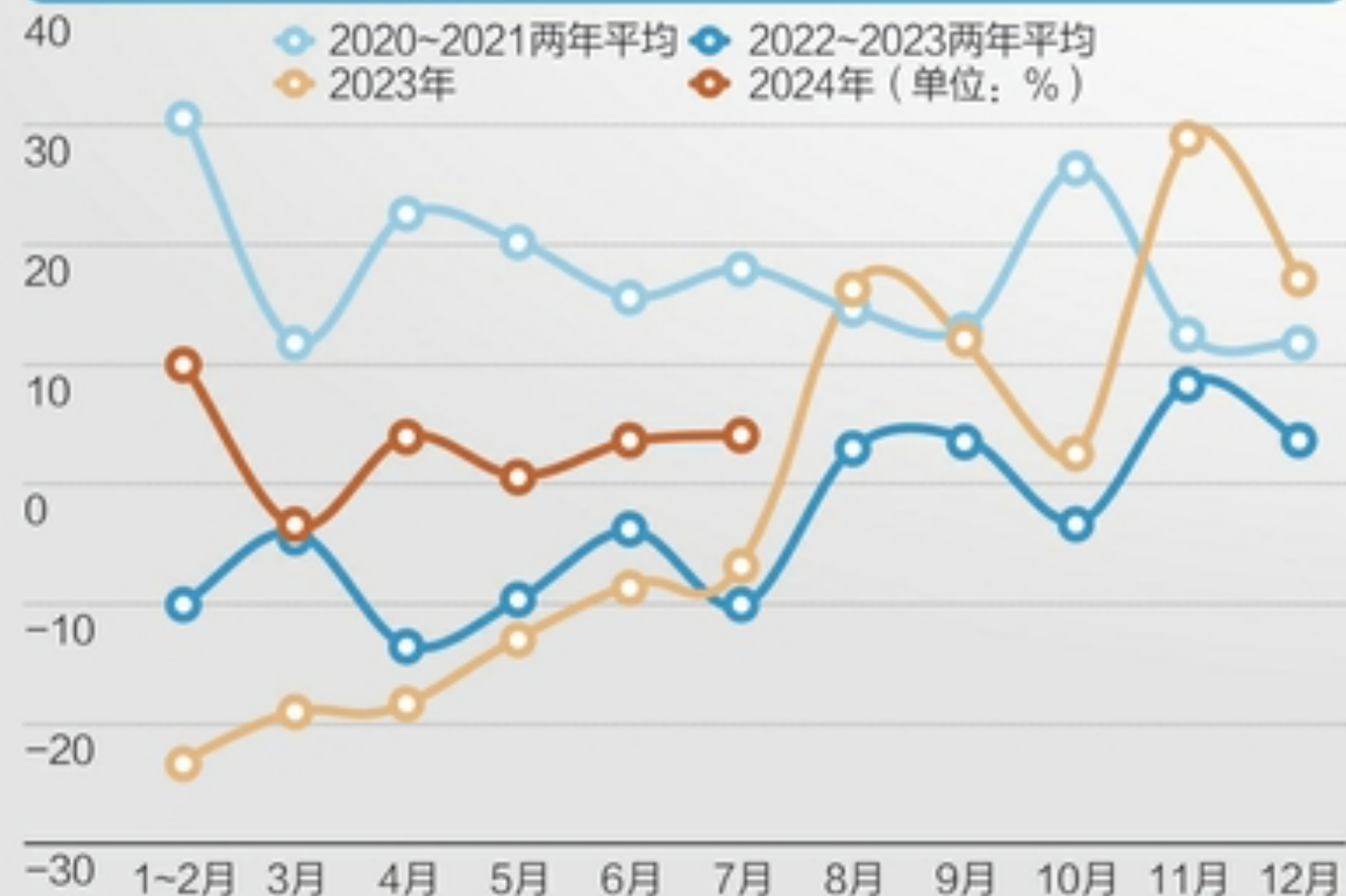
1. 平滑的折线图只会起到误导的作用
2. 柱形图的纵坐标应从 0 点开始

# 有效需求是工业企业利润能否企稳关键

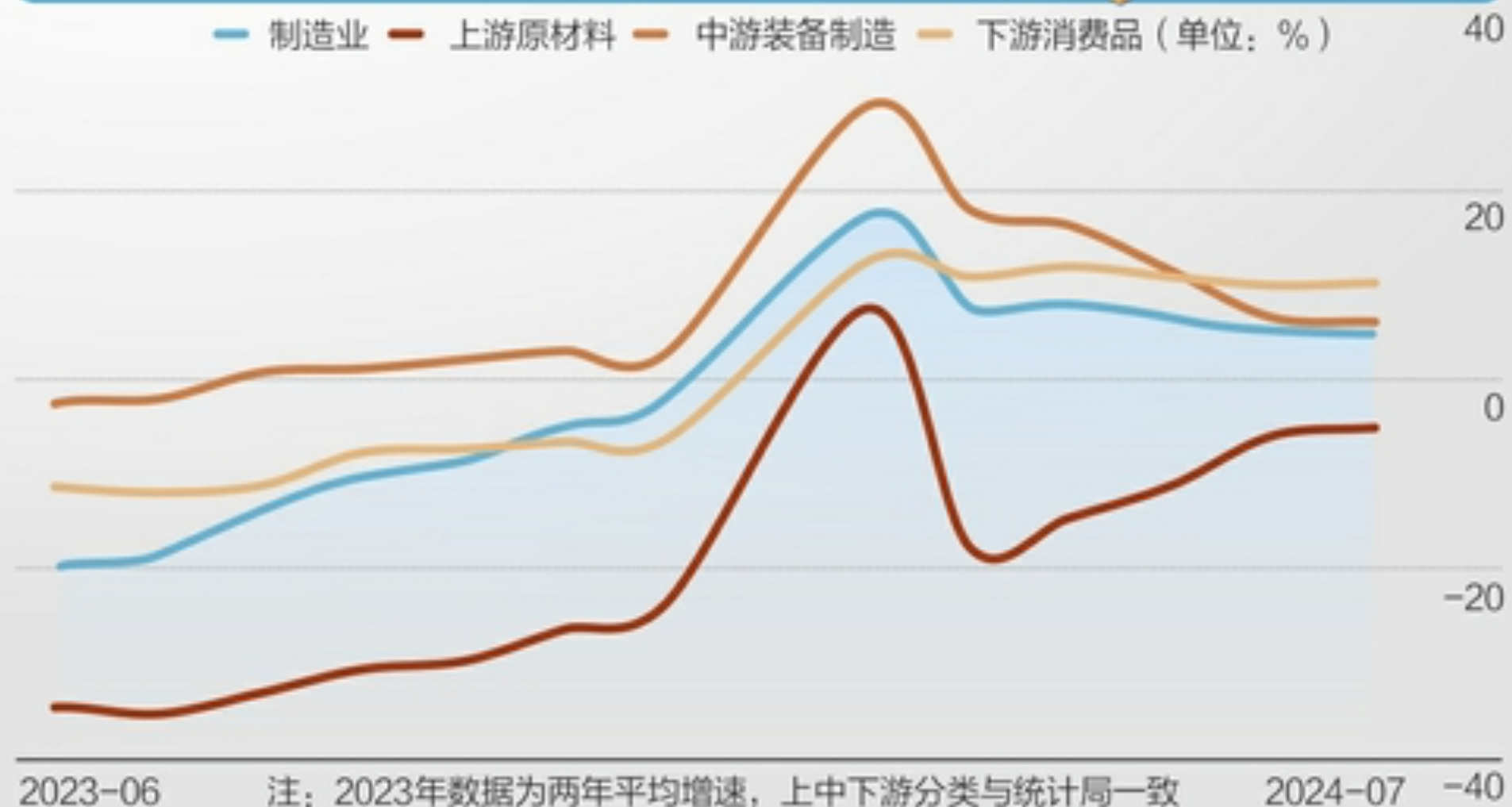
2024年1~7月规上工业企业利润平稳增长，较前值略有回升。浙商证券研报指出，规上工业企业利润能否企稳的关键在于有效需求，当前以价换量特征仍然明显，供给快于需求使得工业企业库存继续被动冲高。在内需方面，大规模设备更新、消费品以旧换新政策有助于继续支撑需求。



### 规上工业企业利润增速有所回升



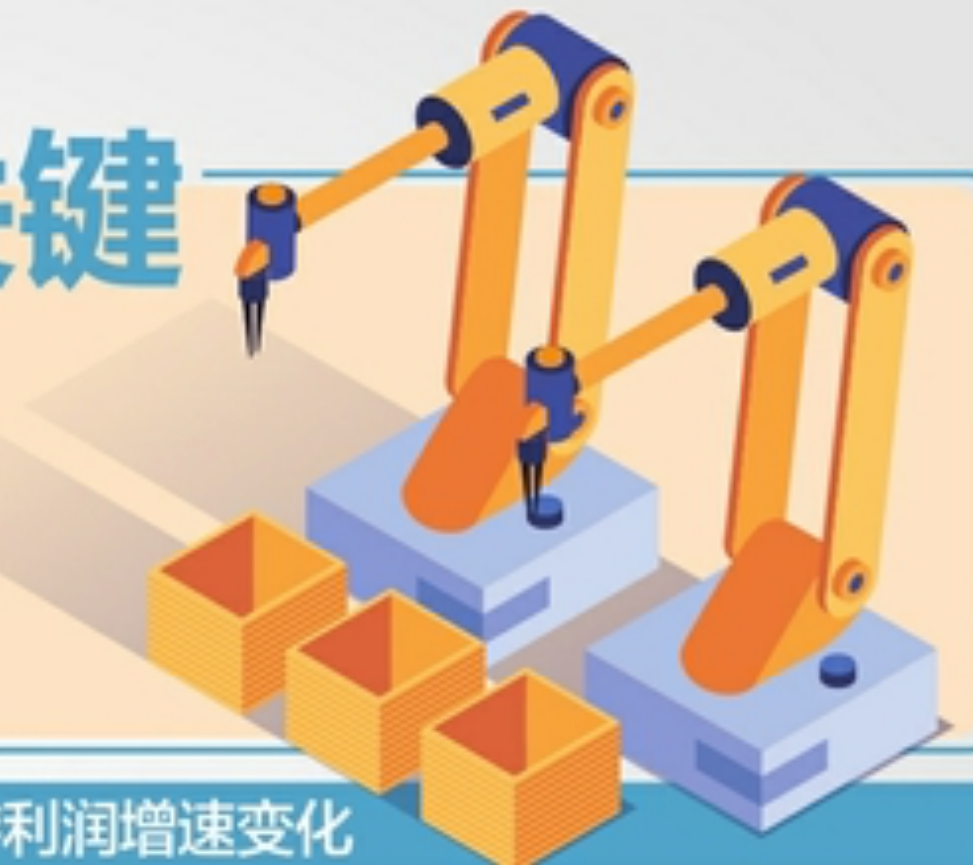
### 制造业上中下游利润增速变化



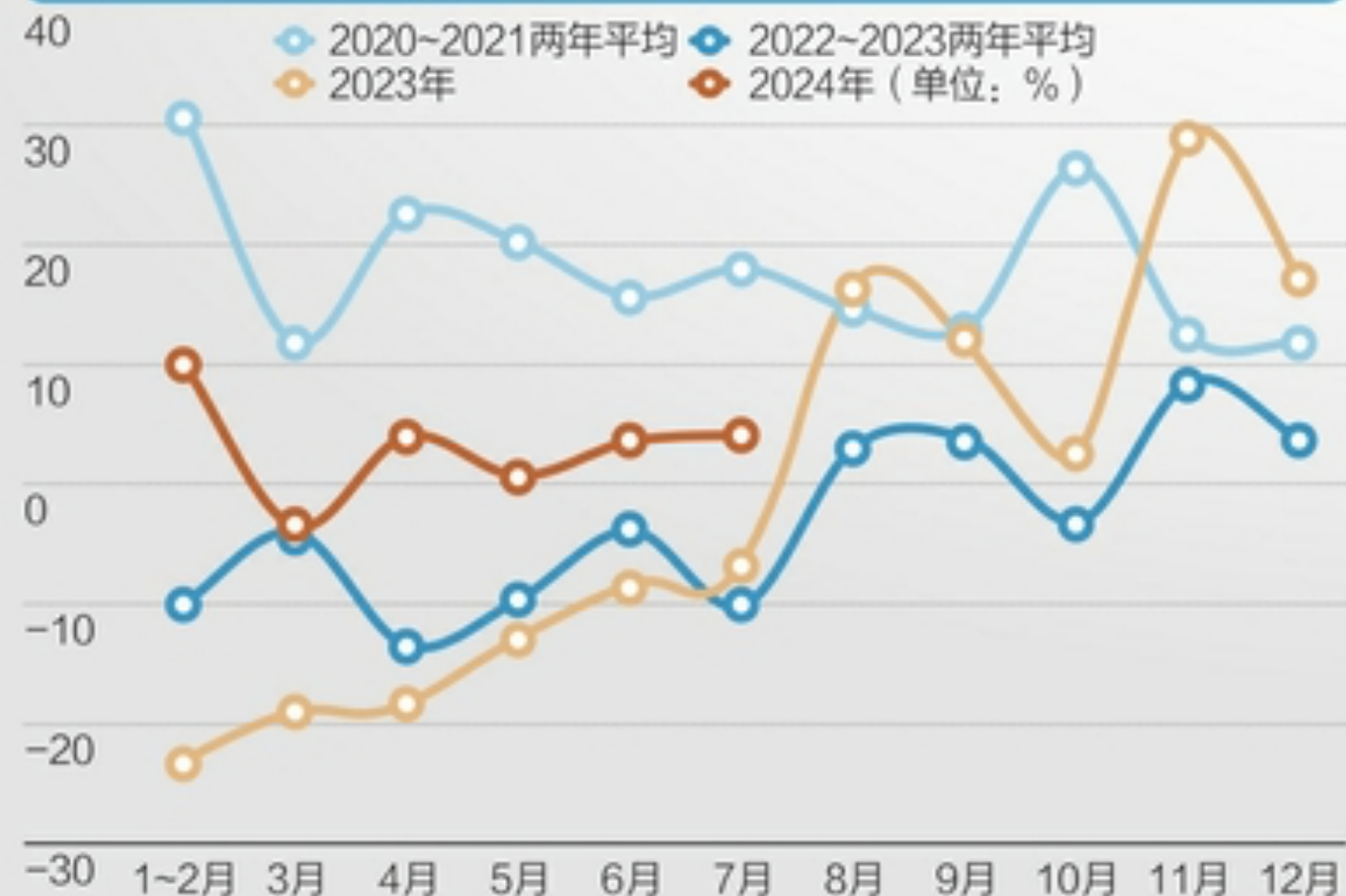
注：2023年数据为两年平均增速，上中下游分类与统计局一致

# 有效需求是工企利润能否企稳关键

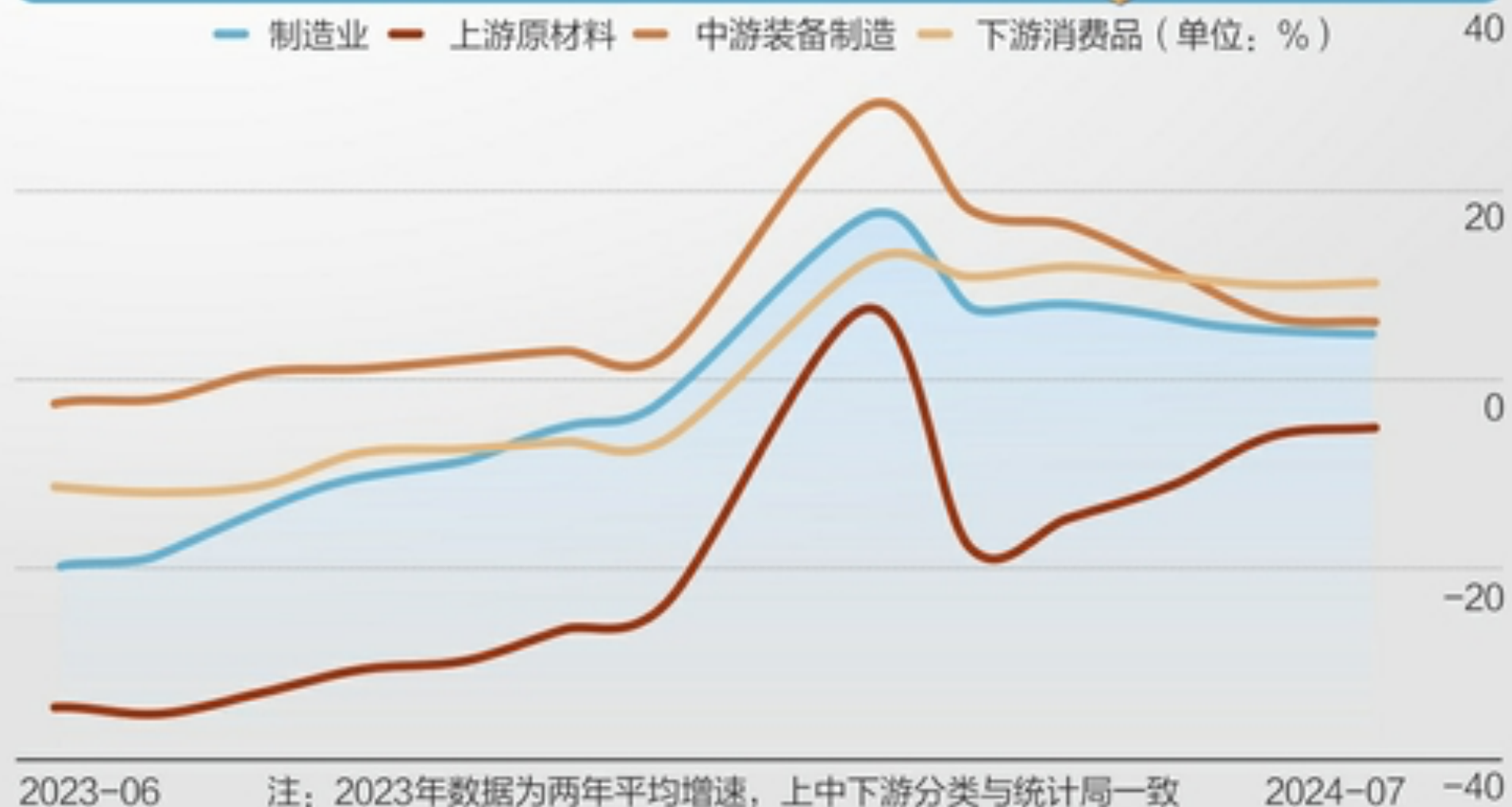
2024年1~7月规上工企利润平稳增长，较前值略有回升。浙商证券研报指出，规上工企利润能否企稳的关键在于有效需求，当前以价换量特征仍然明显，供给快于需求使得工业企业库存继续被动冲高。在内需方面，大规模设备更新、消费品以旧换新政策有助于继续支撑需求。



### 规上工业企业利润增速有所回升



### 制造业上中下游利润增速变化



另一个平滑折线图的例子

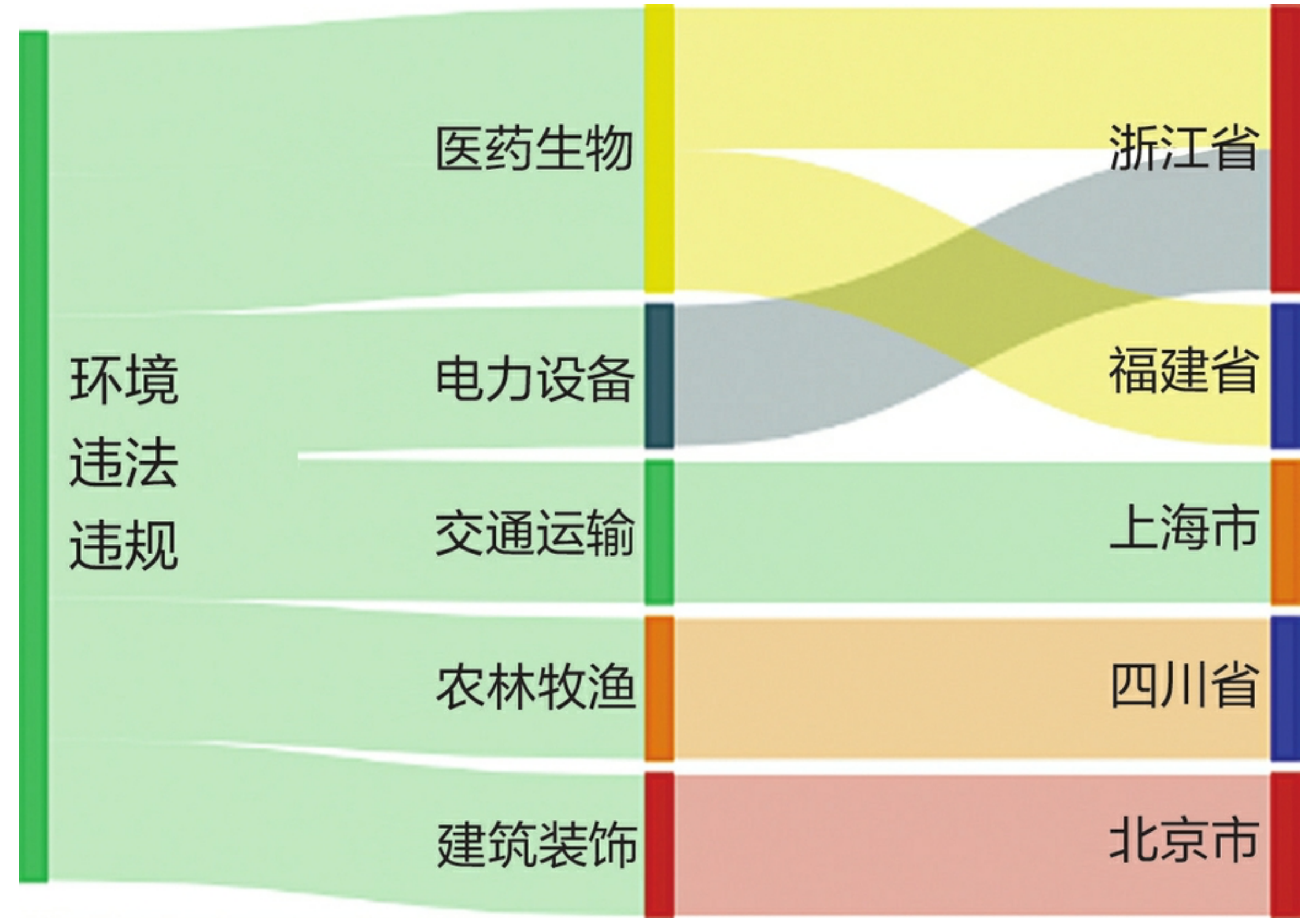
横坐标缺失（这是月度数据吗？）

2024年8月19日

每日经济新闻联合环保领域知名NGO（非政府组织）公众环境研究中心（IPE），自2020年9月起，基于31个省区市、337个地级市政府发布的环境质量、环境排放和污染源监管记录等权威数据来源，每周收集剖析中国数千家上市公司及其旗下数万家分公司（包括分公司、参股公司和控股公司）的环境信息数据，发布“A股绿色周报”，旨在借助环境数据库及专业解析、传播能力，让上市公司经营活动中的环境信息更加阳光透明。根据8月第三周收集到的数据，记者发现，共有6家上市公司在近期暴露了环境风险。

根据8月第三周收集到的数据，记者发现，共有6家上市公司在近期暴露了环境风险。

### 8月第三周环境风险榜涉及公司分布情况

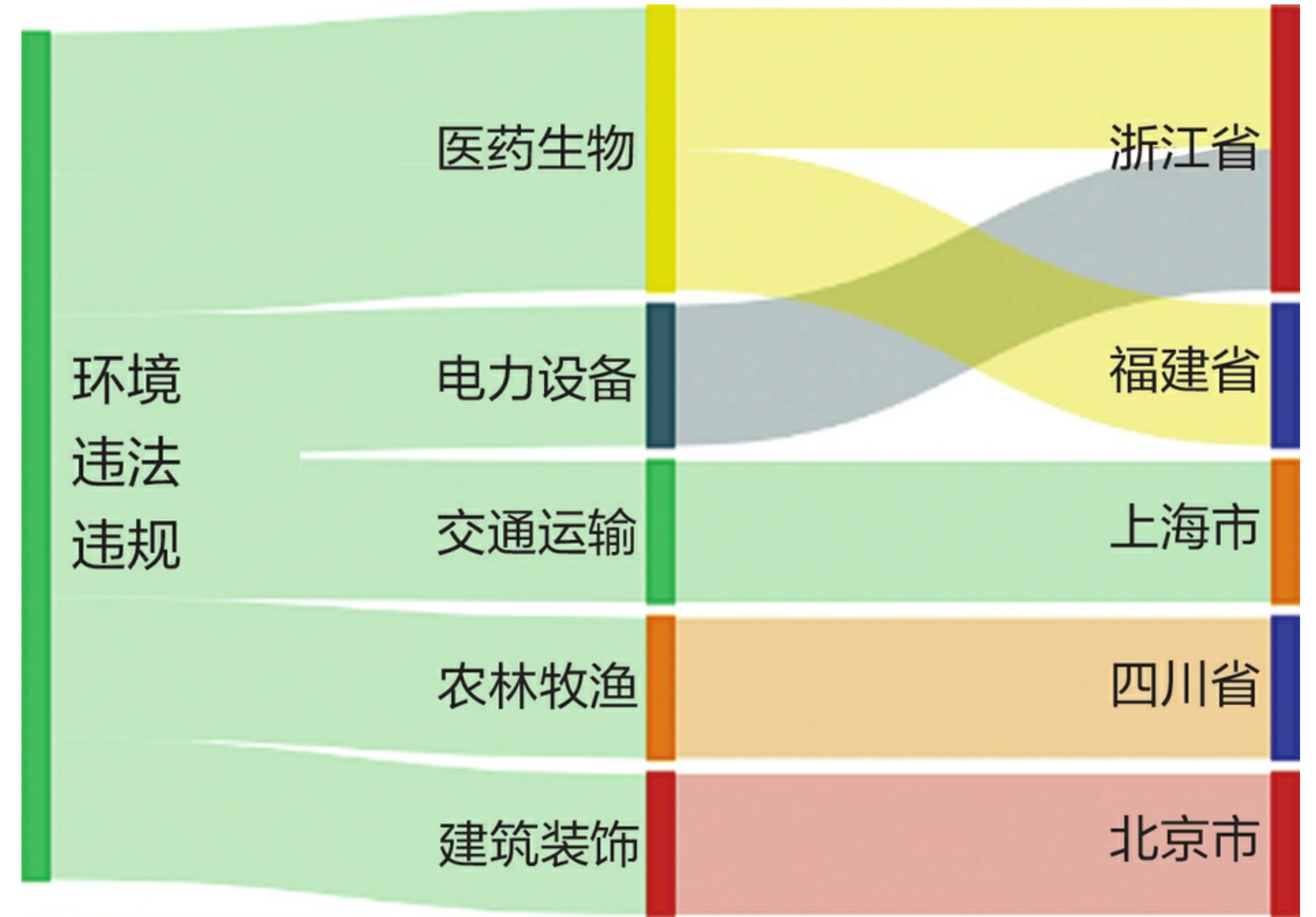


2024年8月19日

每日经济新闻联合环保领域知名NGO（非政府组织）公众环境研究中心（IPE），自2020年9月起，基于31个省区市、337个地级市政府发布的环境质量、环境排放和污染源监管记录等权威数据来源，每周收集剖析中国数千家上市公司及其旗下数万家分公司（包括分公司、参股公司和控股公司）的环境信息数据，发布“A股绿色周报”，旨在借助环境数据库及专业解析、传播能力，让上市公司经营活动中的环境信息更加阳光透明。根据8月第三周收集到的数据，记者发现，共有6家上市公司在近期暴露了环境风险。

根据8月第三周收集到的数据，记者发现，共有6家上市公司在近期暴露了环境风险。

8月第三周环境风险榜涉及公司分布情况



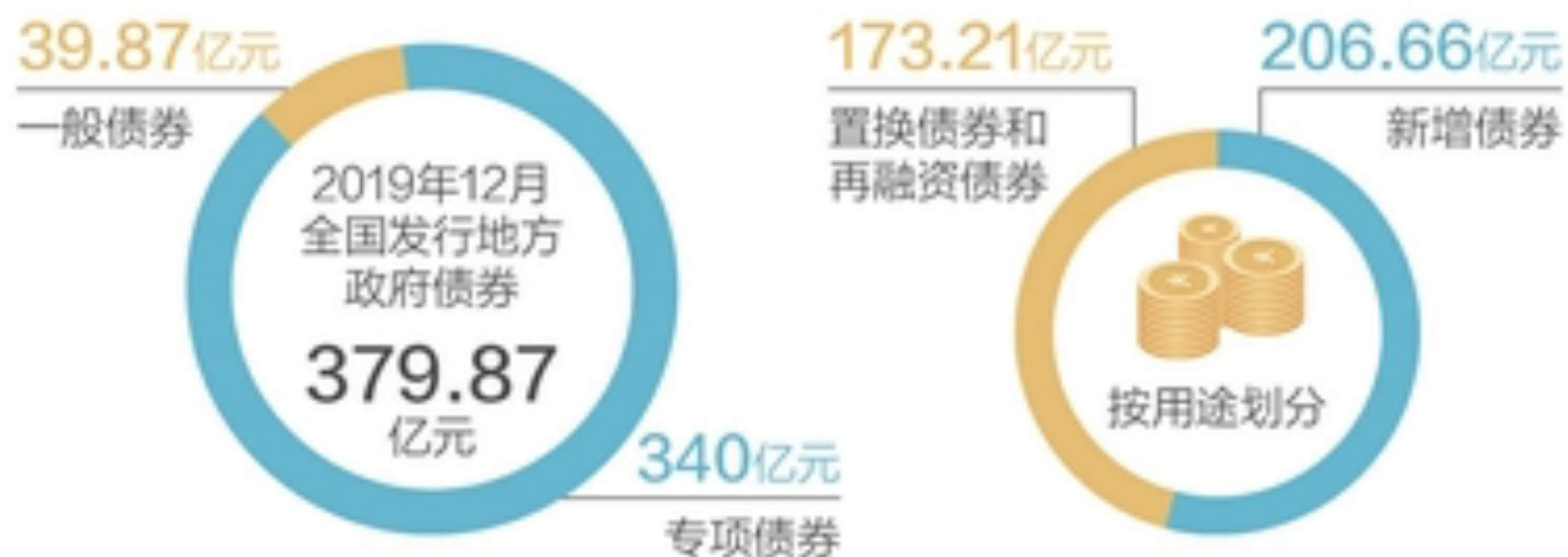
为了画图而画图？

只有六家公司，为什么不用列表？

2020年1月23日

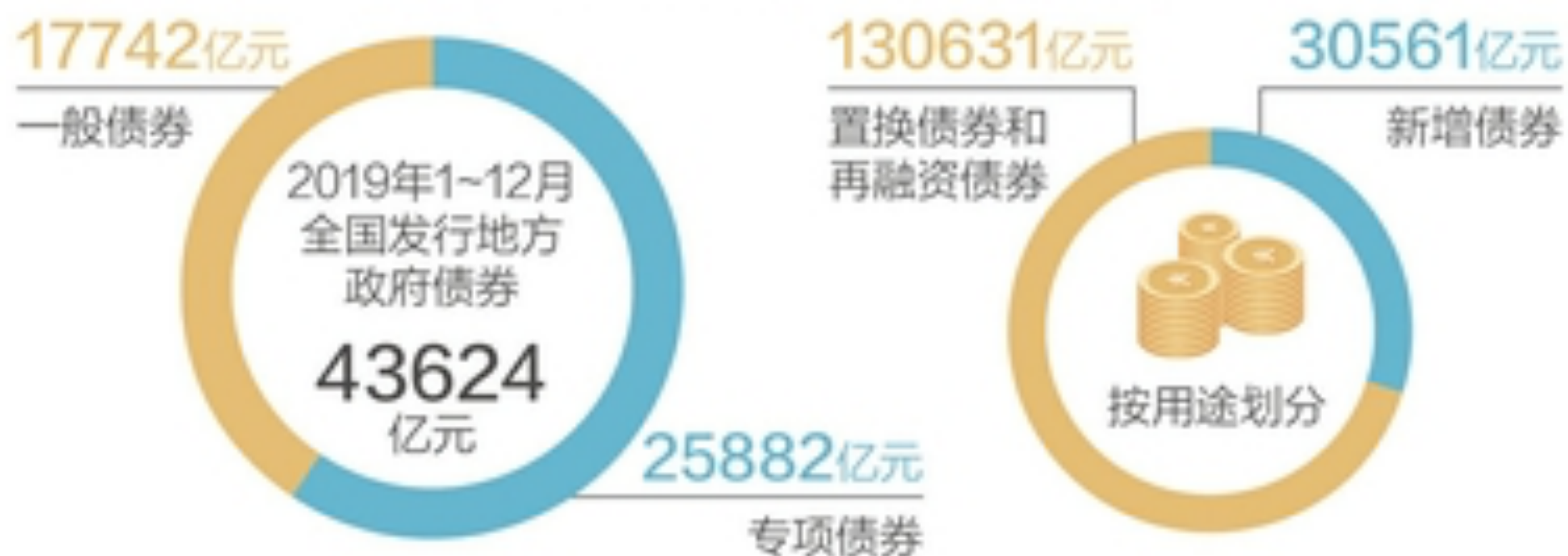
## 2019年全国发行地方政府债券43624亿元

▶ 上月全国发行地方政府债券379.87亿



2019年12月 | 地方政府债券平均发行期限 **18.67年** | 地方政府债券平均发行利率 **3.71%**

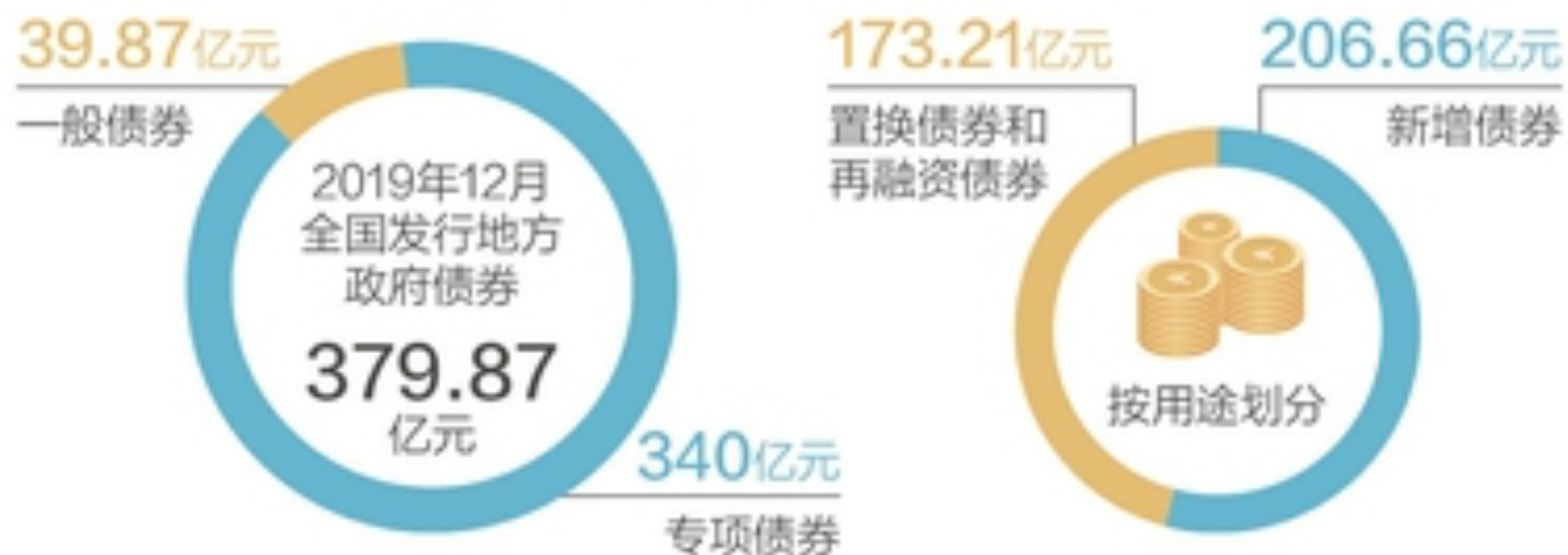
▶ 2019年全国地方新增债券30561亿



2020年1月23日

## 2019年全国发行地方政府债券43624亿元

▶ 上月全国发行地方政府债券379.87亿

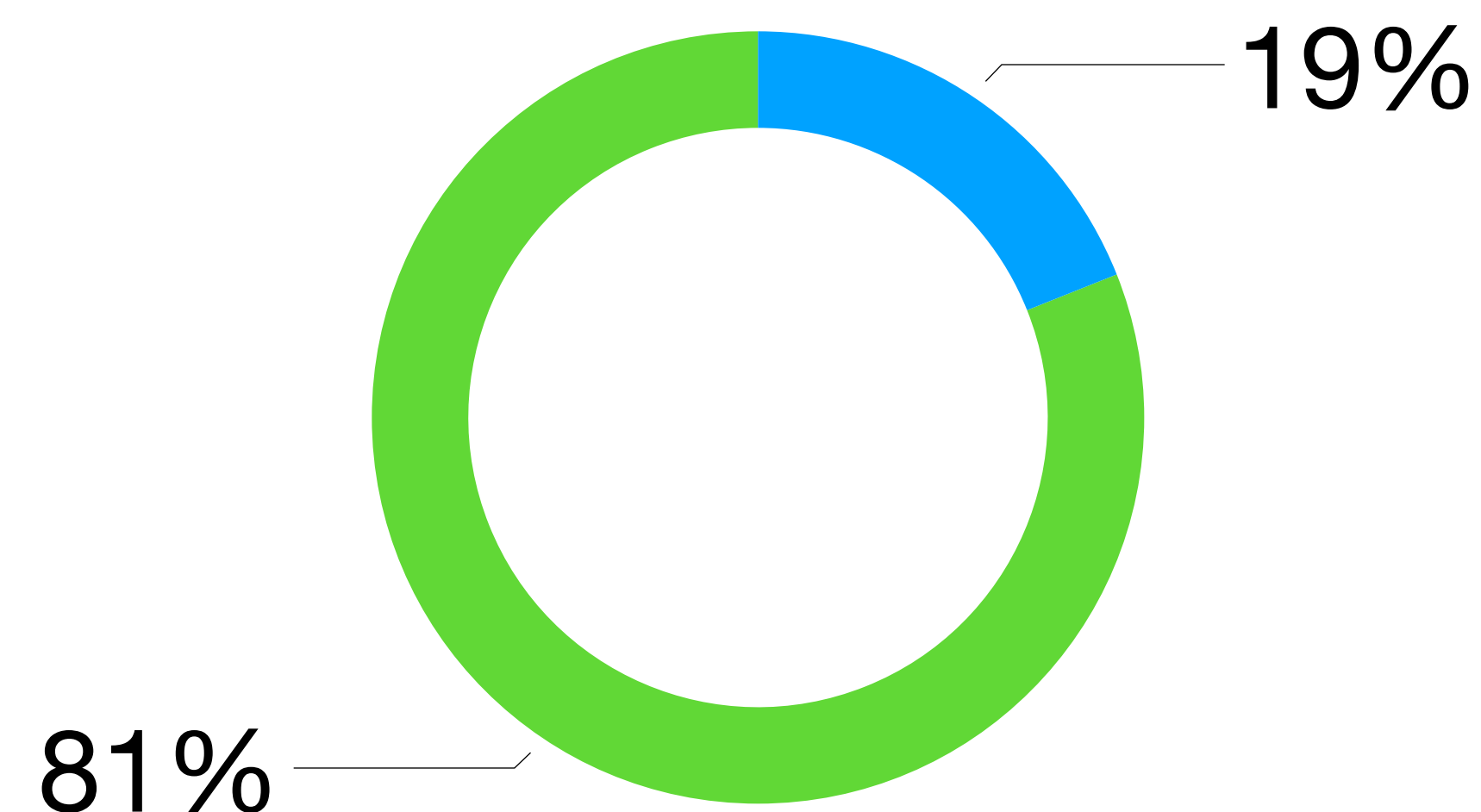


2019年12月 | 地方政府债券  
平均发行期限 **18.67年** | 地方政府债券  
平均发行利率 **3.71%**

▶ 2019年全国地方新增债券30561亿



$$\frac{30561}{130631 + 30561} \approx 0.1896$$

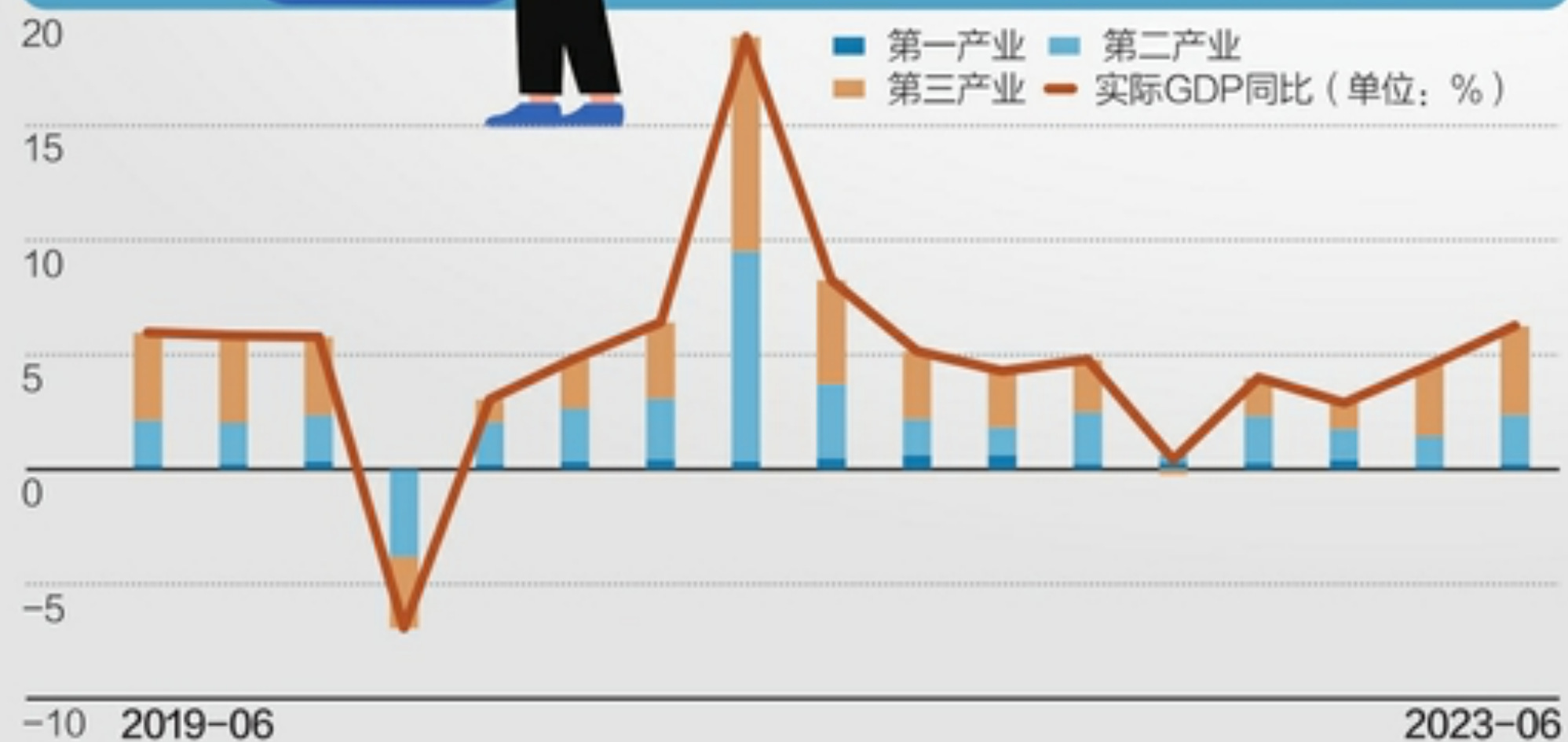


图中的数字和角度不匹配  
(饼图的常见错误)

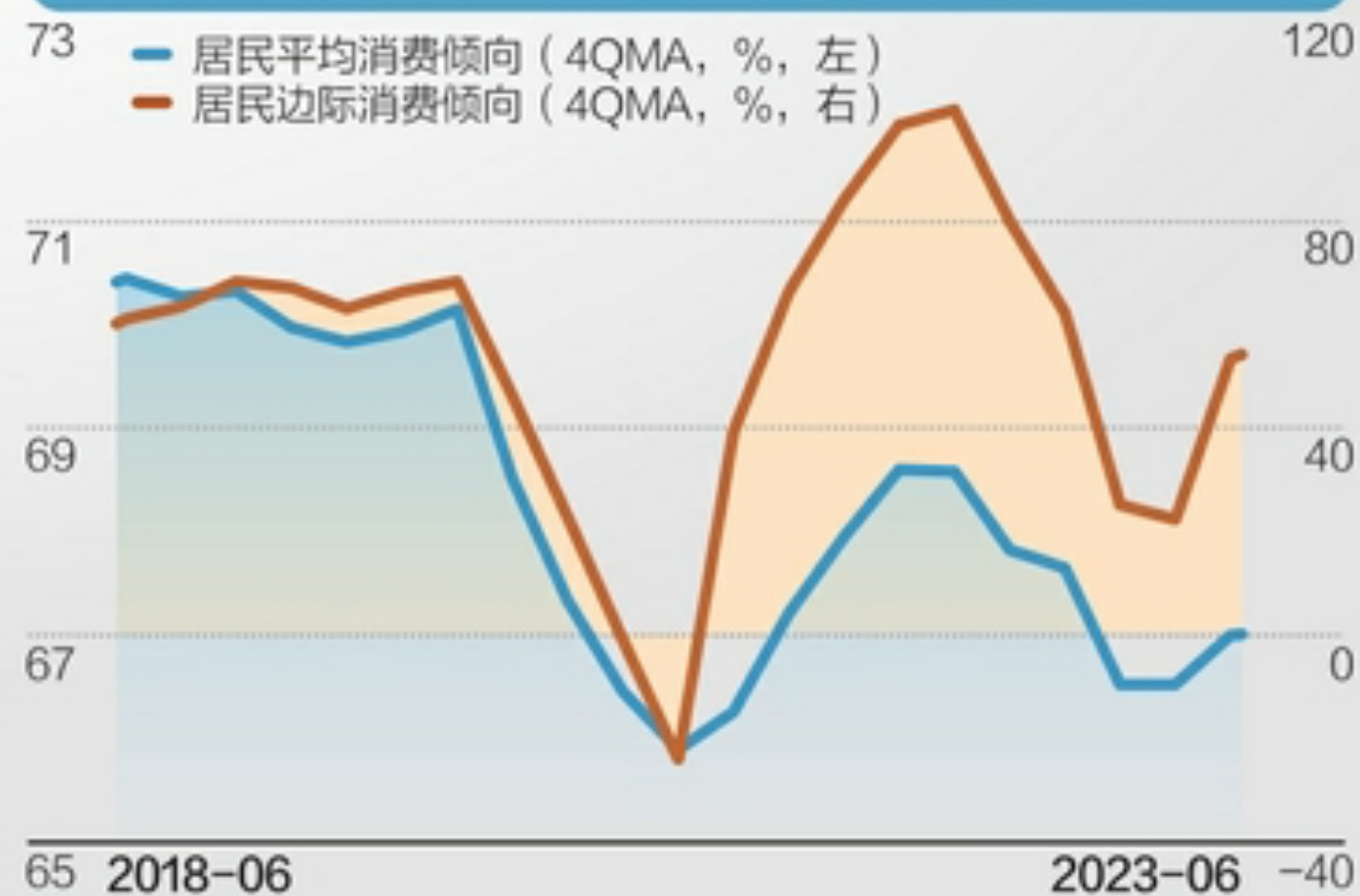
## 券商：经济存在被数据掩盖的“四大改善”

申万宏源研报指出，从边际变化来看，6月经济数据呈现见底回升的情况，结构上存在被数据掩盖的“四大改善”：其一，劳动参与率继续回升，促使居民总收入改善大于人均收入与失业率指标表现，居民消费倾向回暖。其二，地产后周期消费改善强于零售，6月强劲竣工让这一趋势有望延续。其三，市场未跟踪的居民全口径服务消费企稳反弹。其四，剔除价格因素后固定投资增速仍高，基建仍强、地产渐见底。

### 实际GDP分季度同比变化



### 居民平均消费倾向二季度进一步回暖

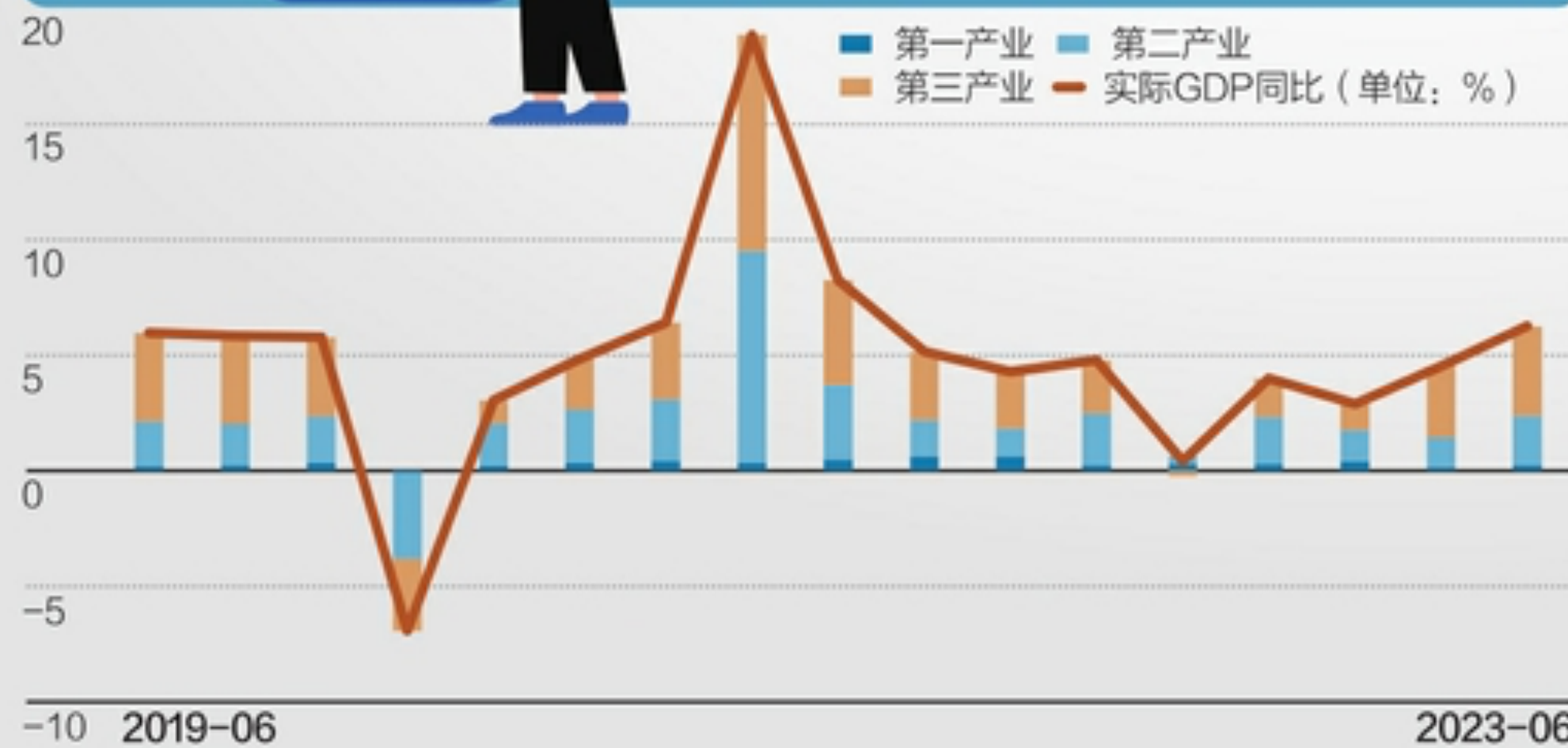


2023年7月18日

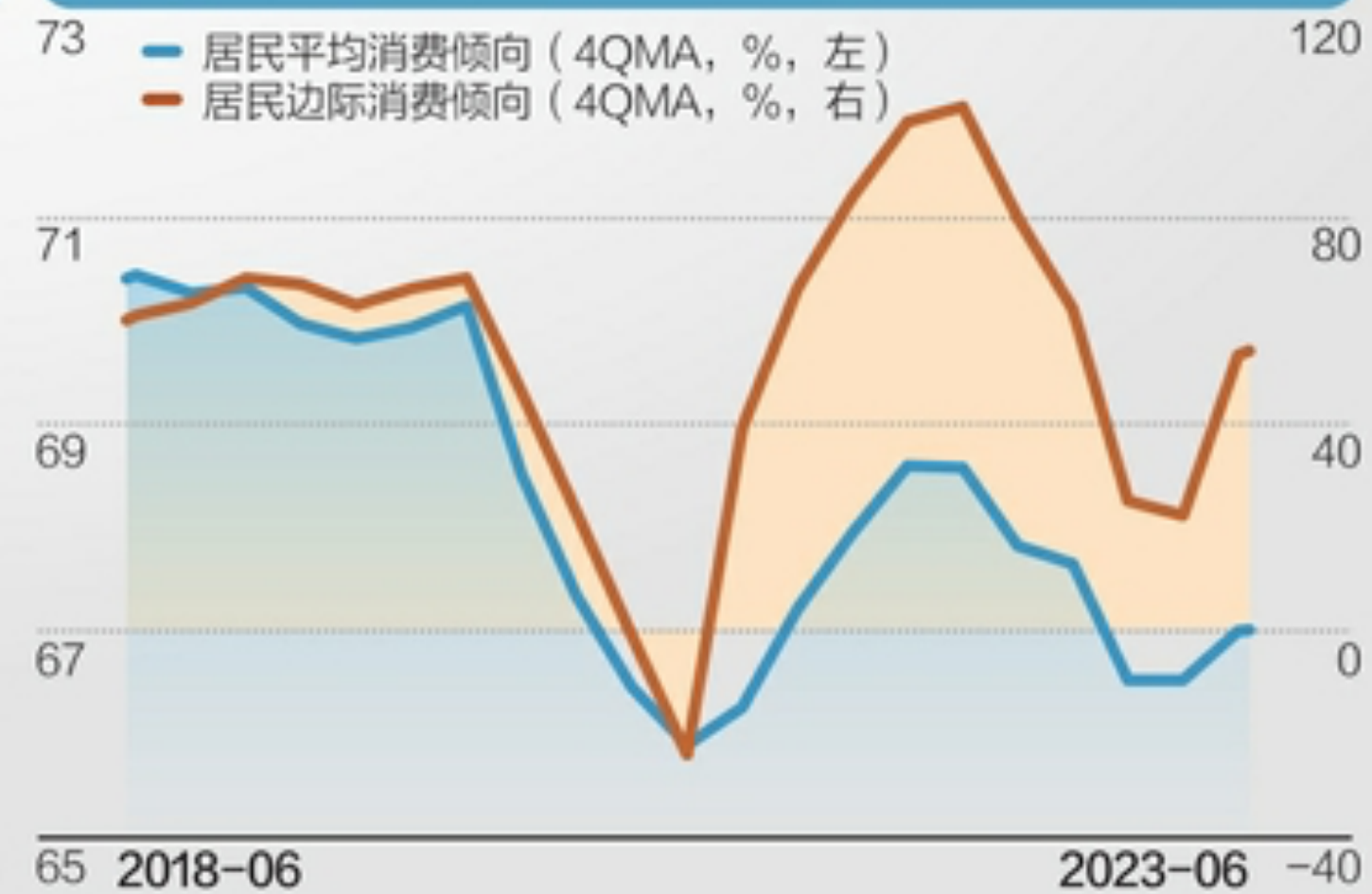
## 券商：经济存在被数据掩盖的“四大改善”

申万宏源研报指出，从边际变化来看，6月经济数据呈现见底回升的情况，结构上存在被数据掩盖的“四大改善”：其一，劳动参与率继续回升，促使居民总收入改善大于人均收入与失业率指标表现，居民消费倾向回暖。其二，地产后周期消费改善强于零售，6月强劲竣工让这一趋势有望延续。其三，市场未跟踪的居民全口径服务消费企稳反弹。其四，剔除价格因素后固定投资增速仍高，基建仍强、地产渐见底。

实际GDP分季度同比变化



居民平均消费倾向二季度进一步回暖



左图中的堆叠柱形图显示的量叫做“三次产业对国内生产总值增长的拉动”，是GDP增速和三次产业贡献率的乘积。三者之和为实际GDP增速。

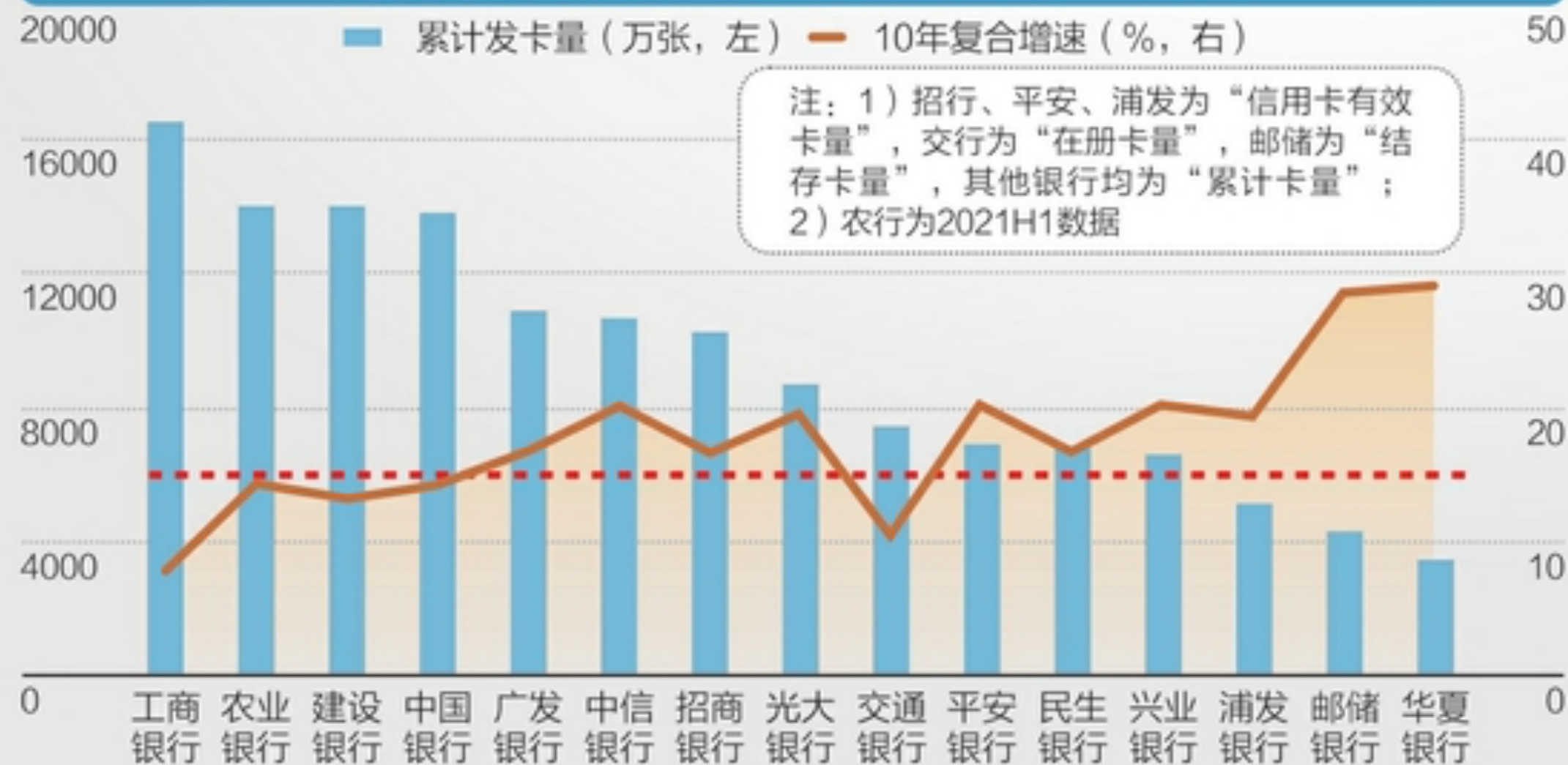
1. 季度数据的时间格式应为 2019Q1

2. 右图中的两个量可以进行比较吗？平均消费倾向 =  $C/Y$ ，边际消费倾向 =  $\Delta C/\Delta Y$

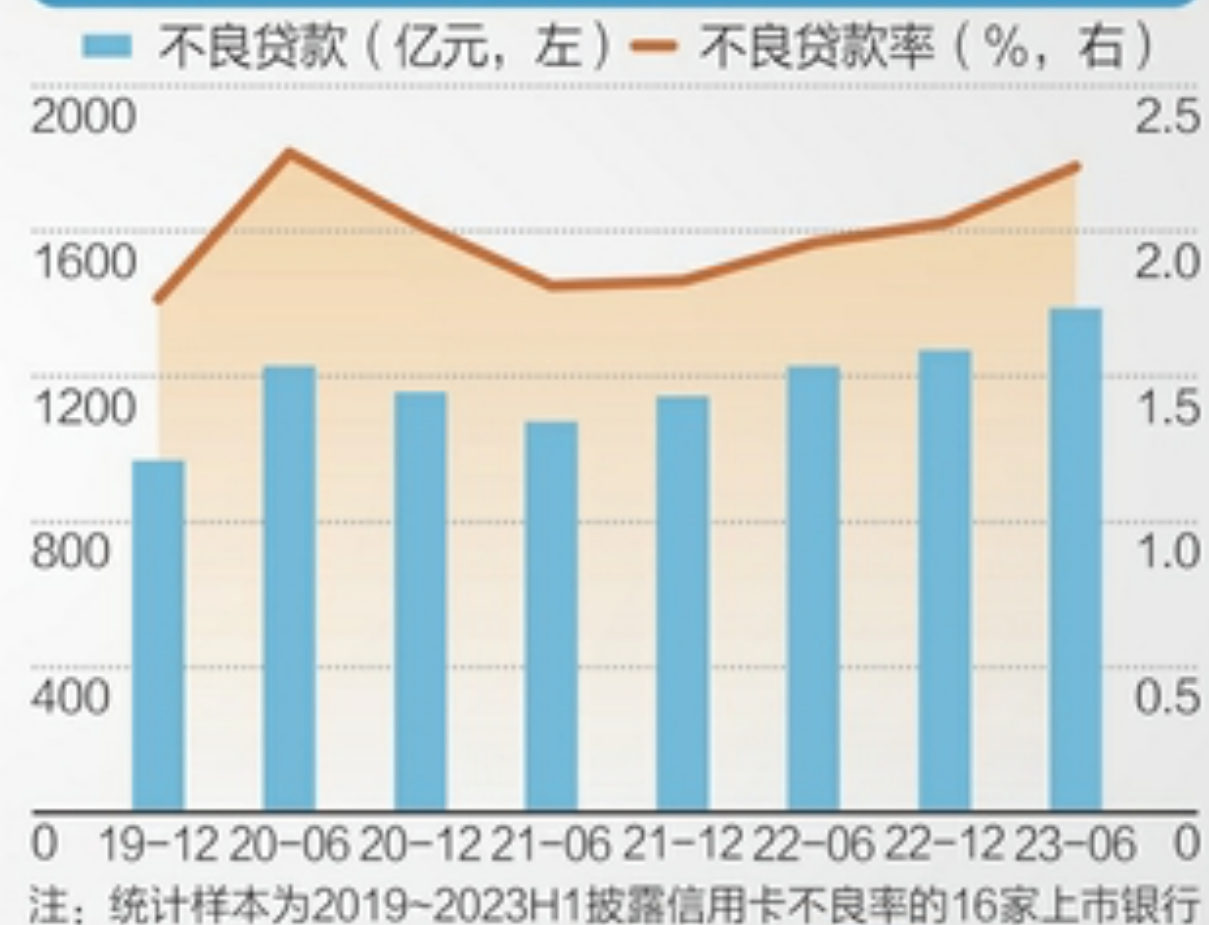
## 信用卡市场增长放缓、竞争加剧

目前我国信用卡市场增长放缓、竞争加剧。从卡量来看，央行今年发布的《2023年第二季度支付体系运行总体情况》显示，截至2023年第二季度末，信用卡和借贷合一卡在用发卡数量7.86亿张，较去年同期减少2100万张。专家指出，银行传统营销获客成本在上升，提升信用卡用户的体验与黏性，是未来信用卡业务的核心竞争力。

### 信用卡累计发卡量及10年复合增速（2022）



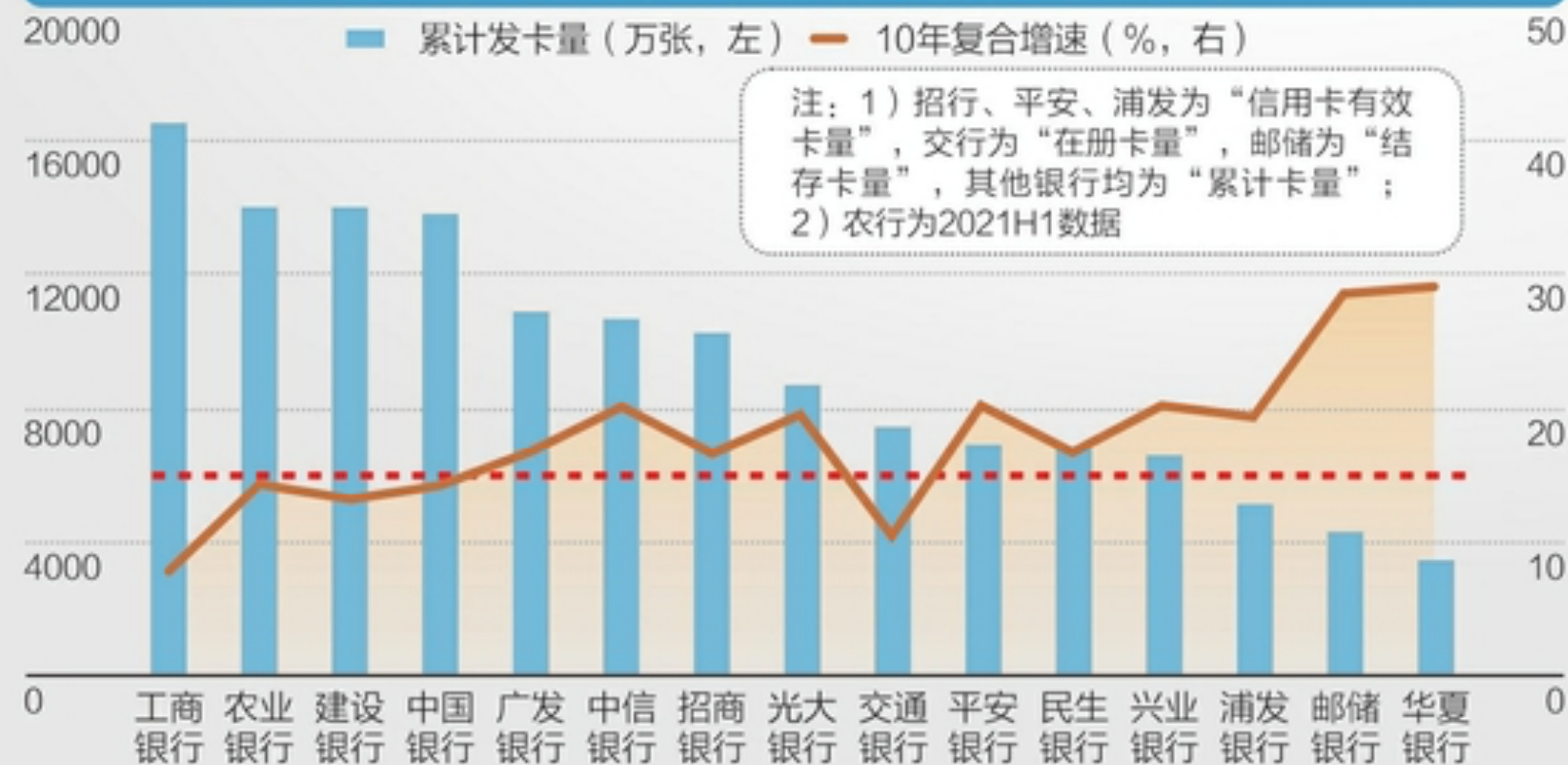
### 上市银行信用卡不良贷款余额及不良率



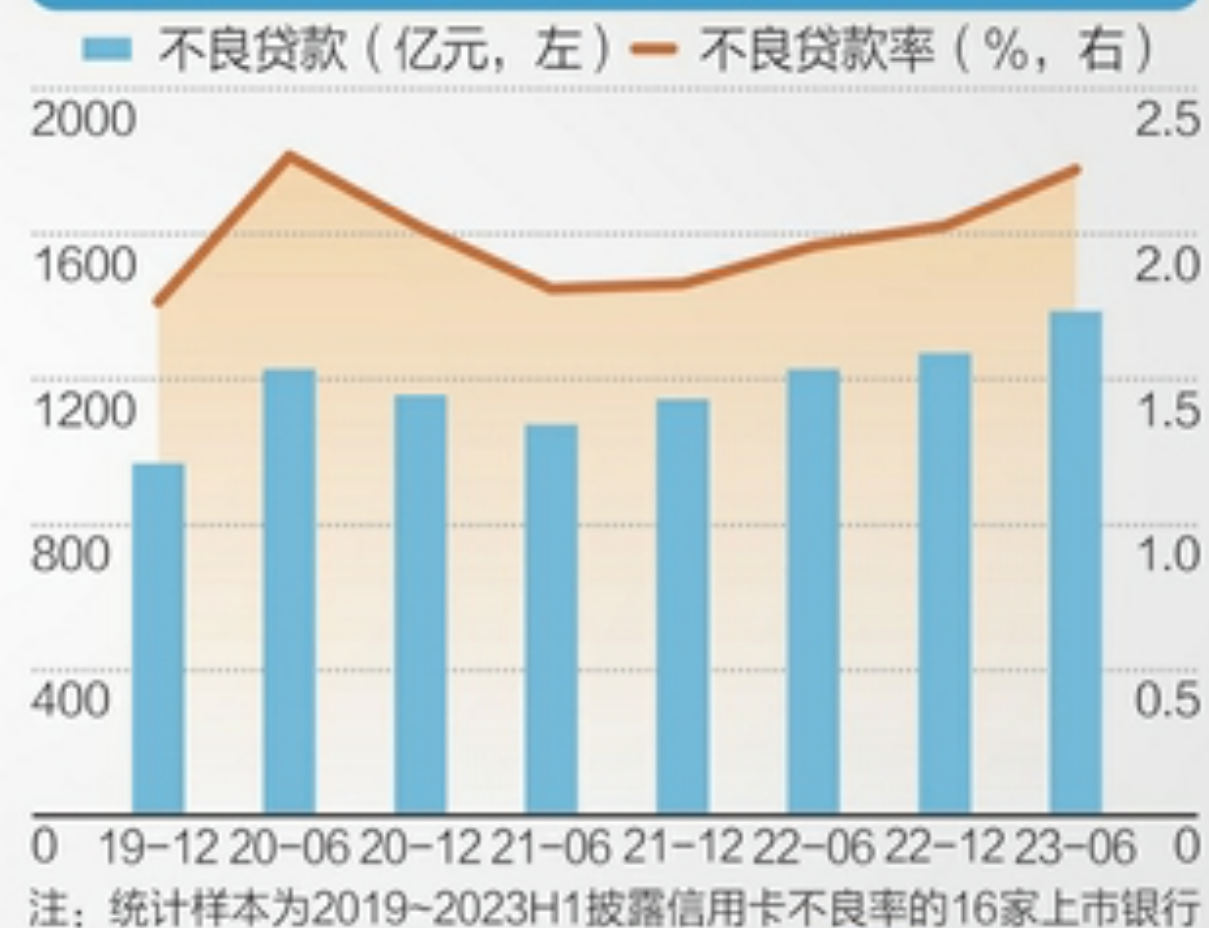
## 信用卡市场增长放缓、竞争加剧

目前我国信用卡市场增长放缓、竞争加剧。从卡量来看，央行今年发布的《2023年第二季度支付体系运行总体情况》显示，截至2023年第二季度末，信用卡和借贷合一卡在用发卡数量7.86亿张，较去年同期减少2100万张。专家指出，银行传统营销获客成本在上升，提升信用卡用户的体验与黏性，是未来信用卡业务的核心竞争力。

### 信用卡累计发卡量及10年复合增速（2022）



### 上市银行信用卡不良贷款余额及不良率



1. 左图为分类数据（不是时序数据），不宜用折线图
2. 没有标注水平虚线的含义（猜测为10年复合增速的平均值）

# 常用统计图表的起源

以及更好的数据展示方法

# 折线图、柱形图、饼图

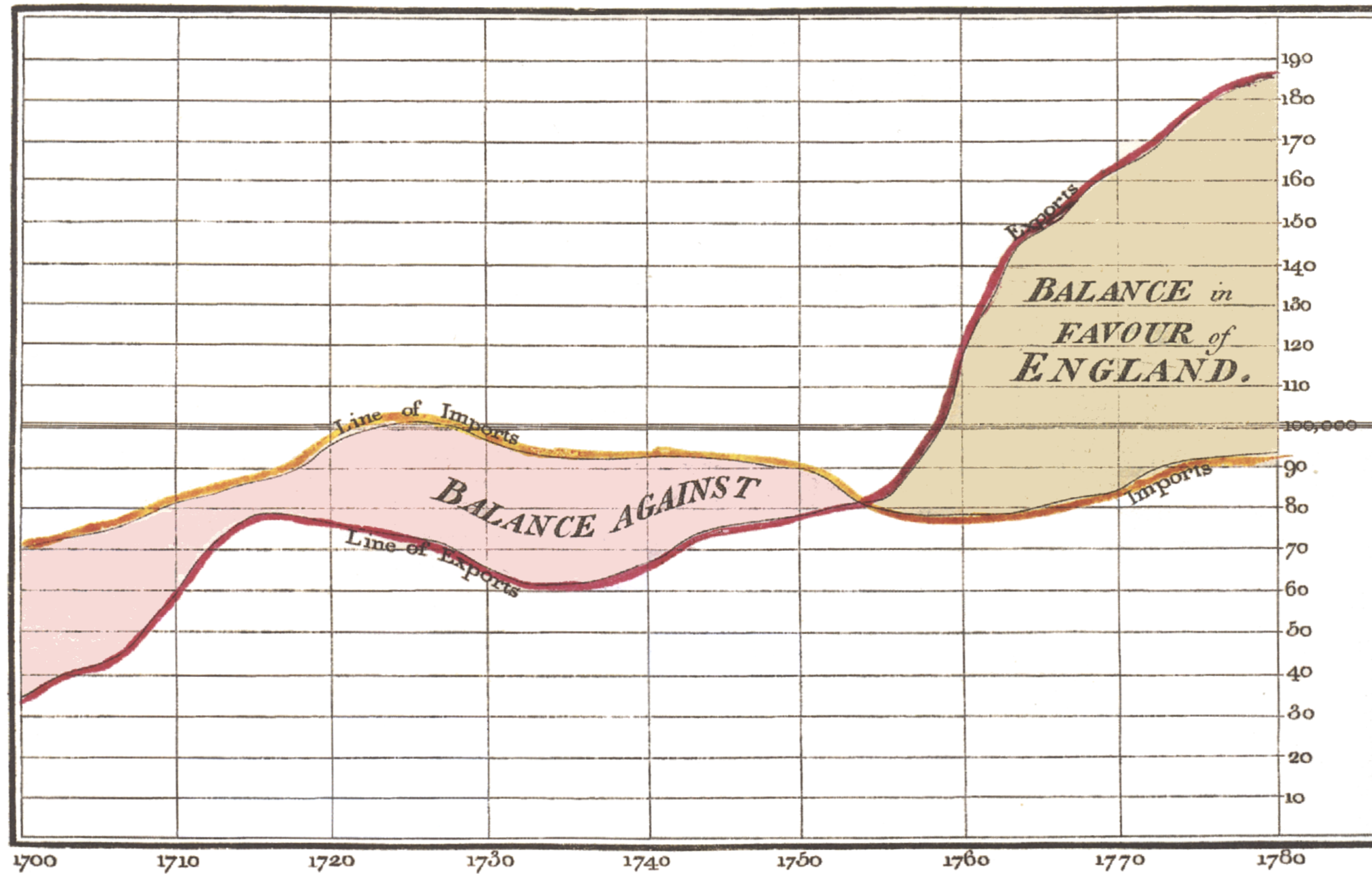
- 第一个将统计数据可视化付诸实践的是18世纪苏格兰工程师 William Playfair (1759 – 1823), 被誉为现代统计图表之父, 在他的著作中可以发现许多折线图、柱形图和饼图的应用。
- Playfair, W. (1786). *The Commercial and Political Atlas*.  
关于英格兰的国际贸易和其他经济数据, 共包含 44 幅图, 包括折线图和柱形图。在此之前, 学者们通常只使用表格展示数据。
- Playfair, W. (1801). *Statistical Breviary*.  
首次使用饼图。



William Playfair

来源: <https://www.playfairprize.com/william-playfair>

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.

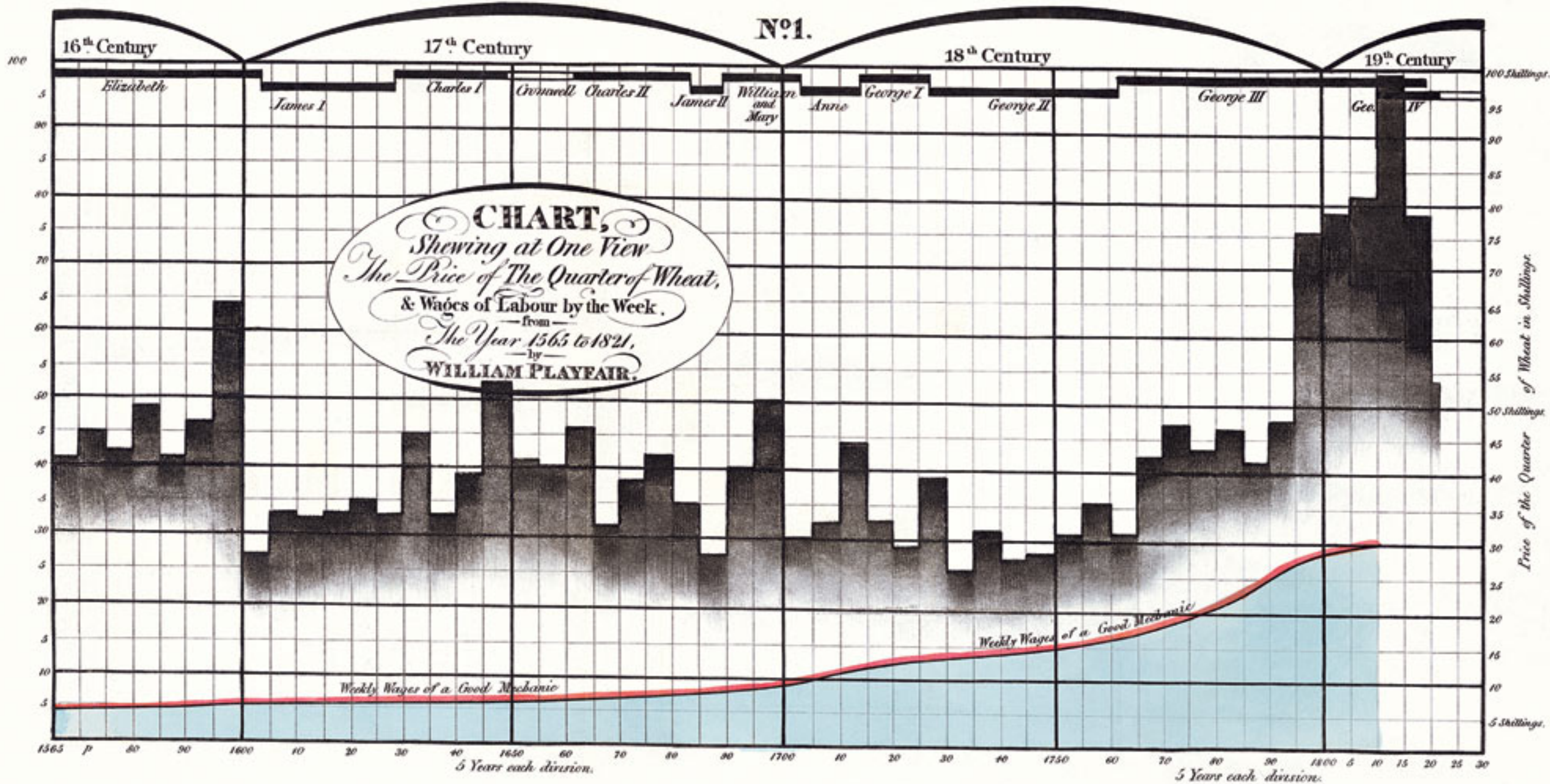


The Bottom line is divided into Years, the Right hand line into £10,000 each.

Published as the Act directs, 1<sup>st</sup> May 1786. by W<sup>m</sup> Playfair

Neele sculpt 352, Strand, London.

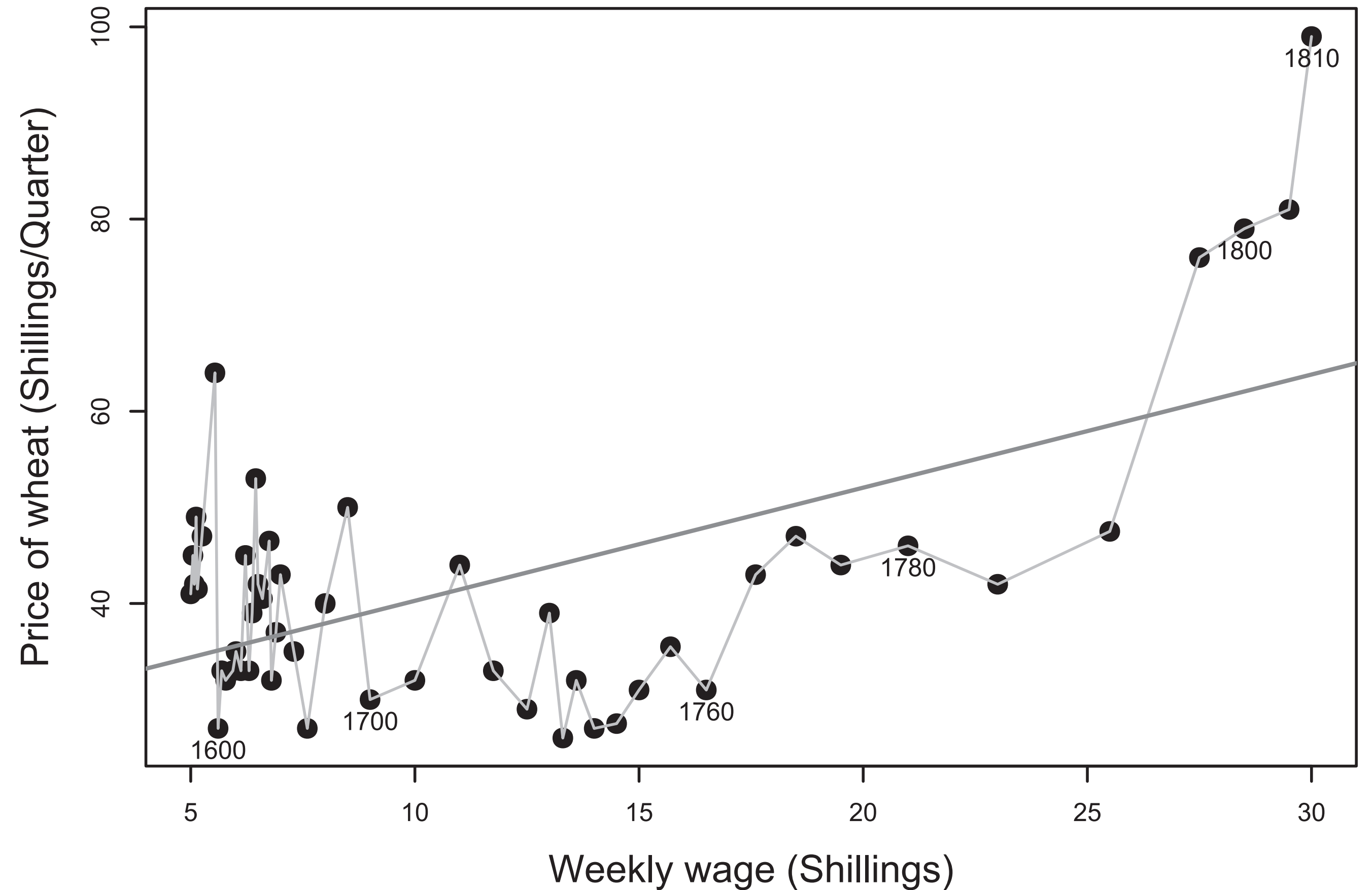
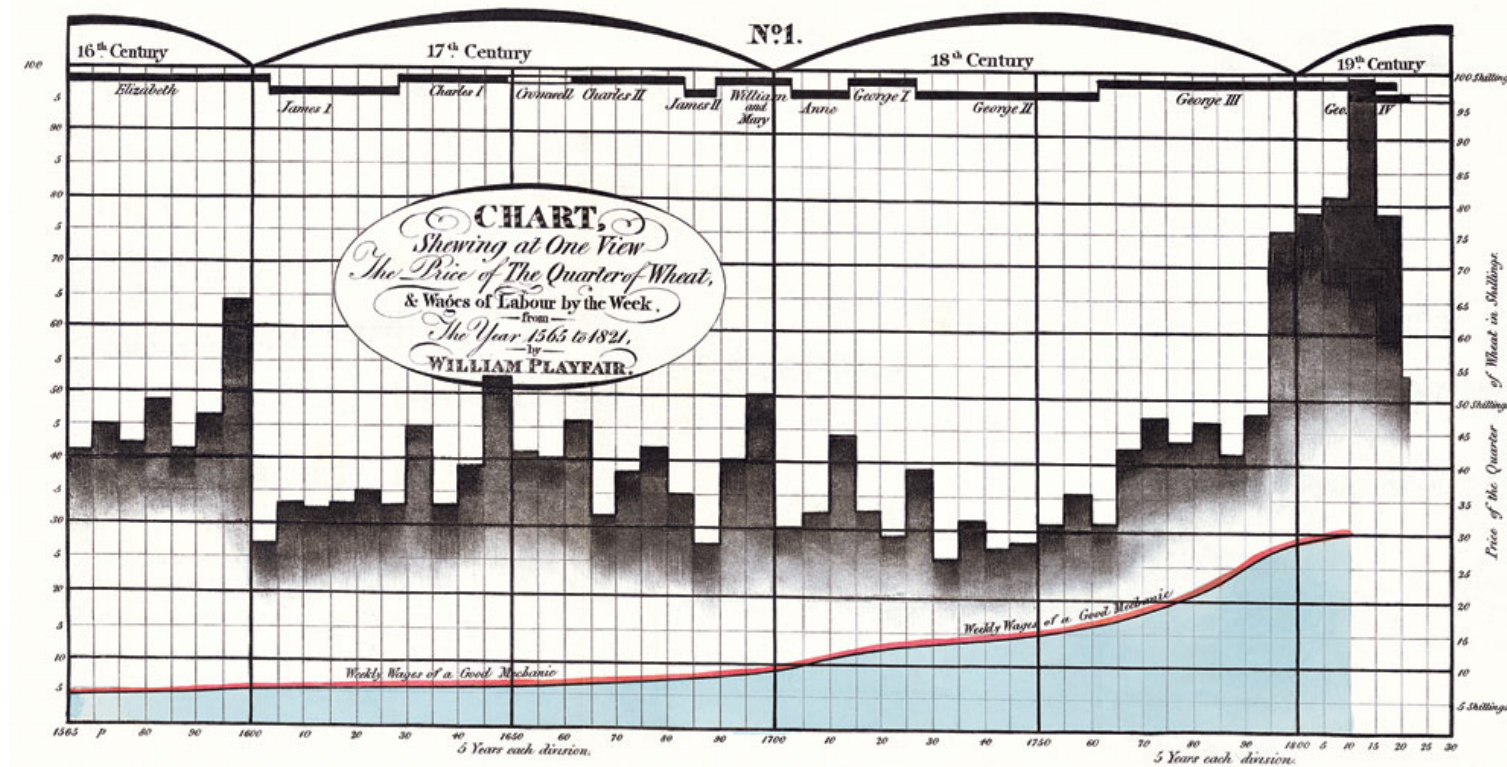
Playfair, W. (1786). *The Commercial and Political Atlas*. 英格兰在1700-1780年间对丹麦和挪威的进出口额。



Playfair, W. (1821). *A letter on our agricultural distresses, their causes and remedies*. 小麦（1/4磅）的货币价值（先令），一个优秀技工的周薪变化，以及英国的朝代更替。

# 散点图

- 散点图可以清晰地展示两个变量之间的相关关系。
- 右图是根据 Playfair 小麦价格和工资水平的时序图重新绘制的散点图，横轴为周薪，纵轴为小麦价格。



Friendly, M. & Wainer, H. (2021). *A History of Data Visualization and Graphic Communication*. Harvard University Press.

# 南丁格尔与玫瑰图

- Florence Nightingale (1820 – 1910) 是我们熟知的现代护理工作的奠基人，同时她也是一名应用统计学家和社会活动家。
- 南丁格尔最有影响力的统计可视化作品是她的玫瑰图，也称为极坐标区域图（polar area diagram）或鸡冠花图（coxcombs，南丁格尔自己的命名），用来描述在克里米亚战争期中后期（1854 – 1856）英军士兵死亡人数和死亡原因的分布及演化。
- 南丁格尔极具视觉冲击效果的玫瑰图帮助她说服了英国政府建设战地医院以提高士兵的存活率。

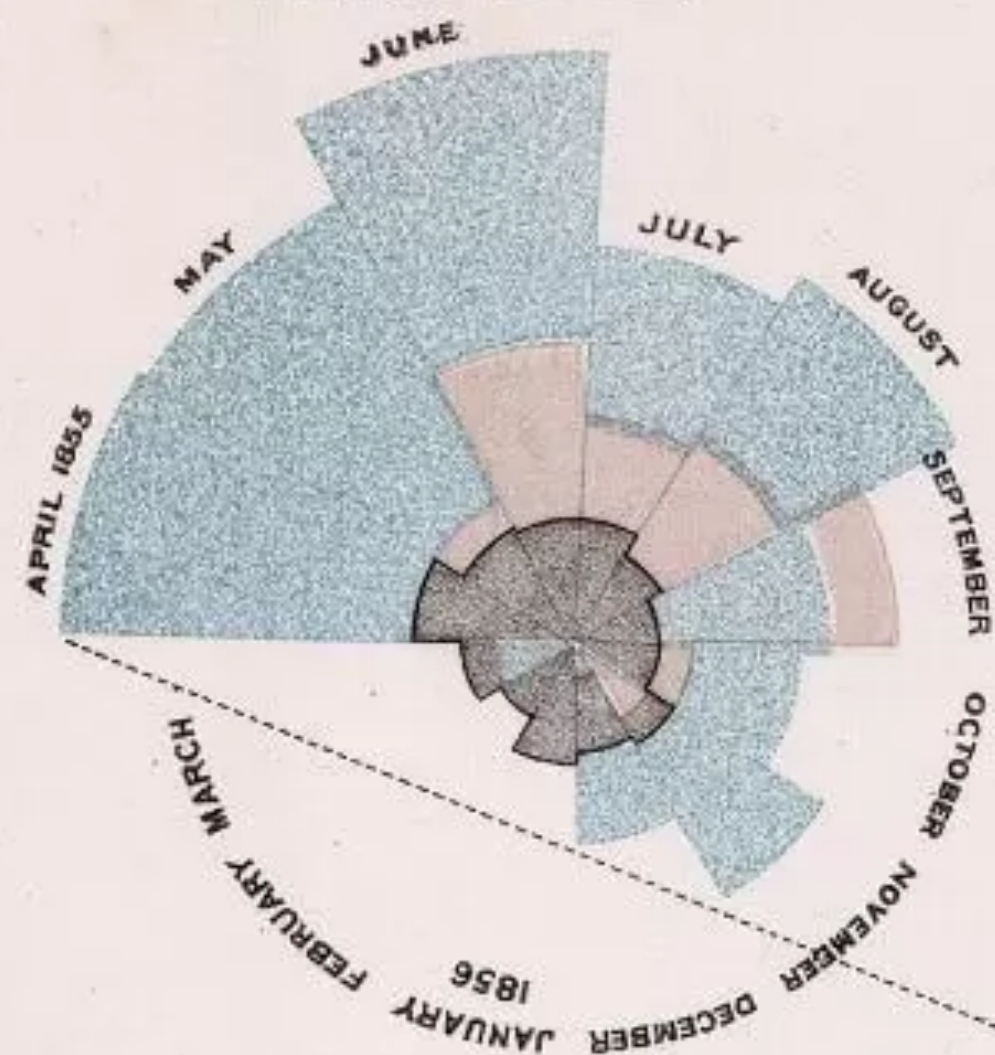


Florence Nightingale

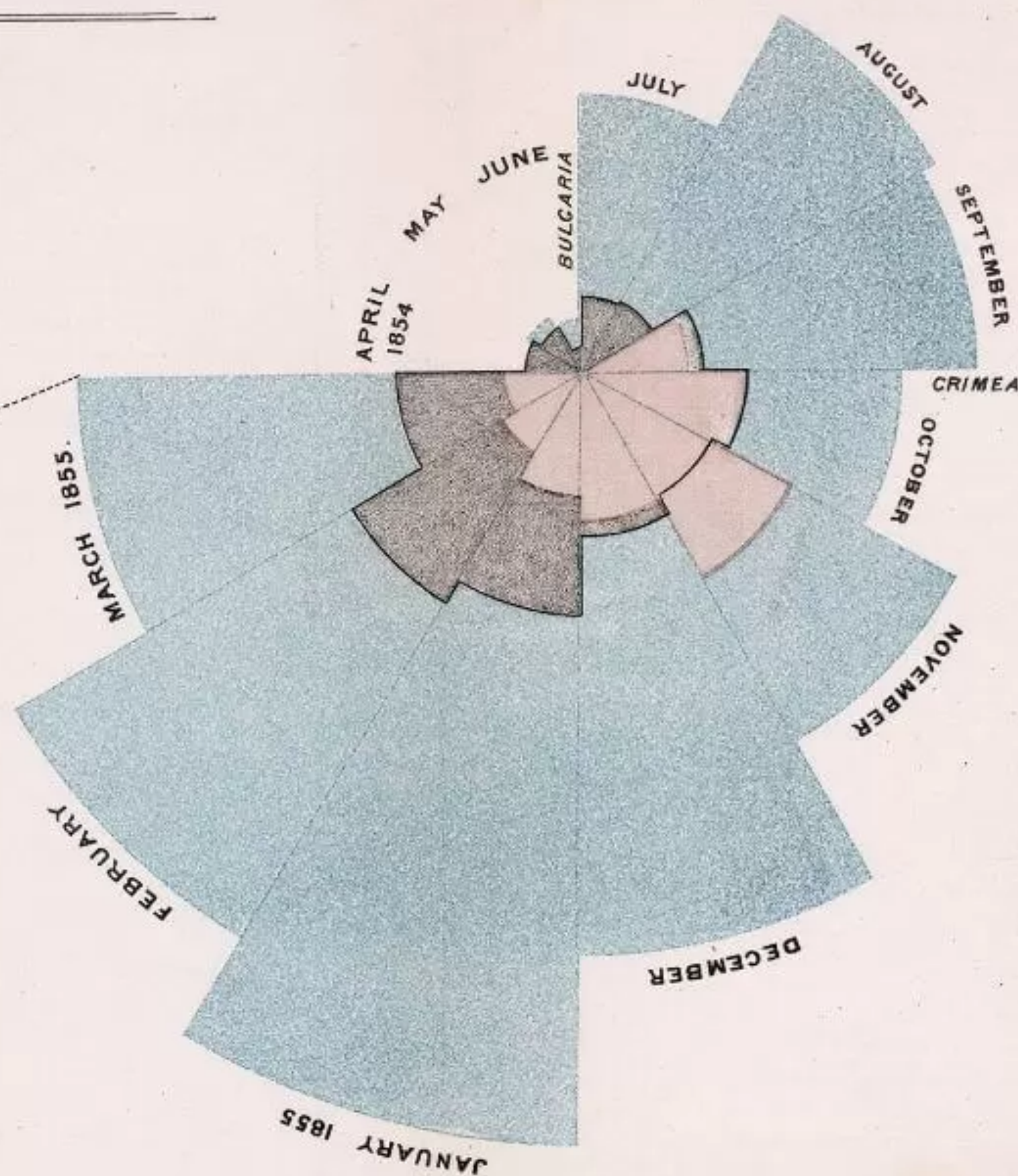
来源：<https://www.playfairprize.com/florence-nightingale>

# DIAGRAM OF THE CAUSES OF MORTALITY IN THE ARMY IN THE EAST.

2.  
APRIL 1855 TO MARCH 1856.



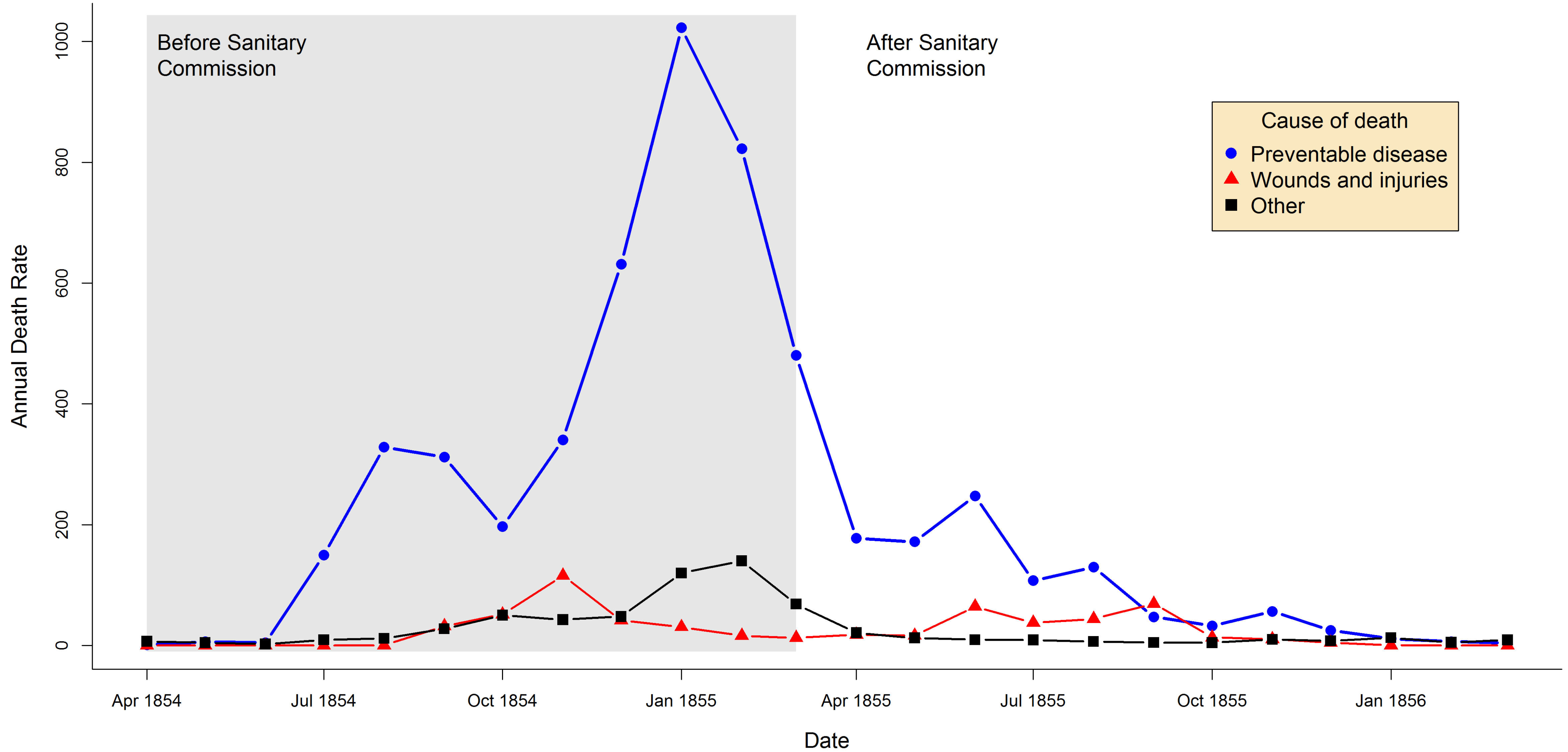
1.  
APRIL 1854 TO MARCH 1855.



*The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.*  
*The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.*  
*The black line across the red triangle in Nov<sup>r</sup> 1854 marks the boundary of the deaths from all other causes during the month.*  
*In October 1854, & April 1855, the black area coincides with the red, in January & February 1856, the blue coincides with the black.*  
*The entire areas may be compared by following the blue, the red & the black lines enclosing them.*

Nightingale, F. (1859). *A contribution to the sanitary history of the British army during the late war with Russia*. 南丁格尔玫瑰图。每一个扇形代表一个自然月，面积代表死亡人数，蓝红黑三色重叠。蓝色部分为传染病导致的可以避免的死亡，红色部分为受伤导致的死亡，黑色部分为其他原因导致的死亡。1855年4月（左图），英国政府成立卫生委员会改善战地卫生条件，死亡人数也因此下降。

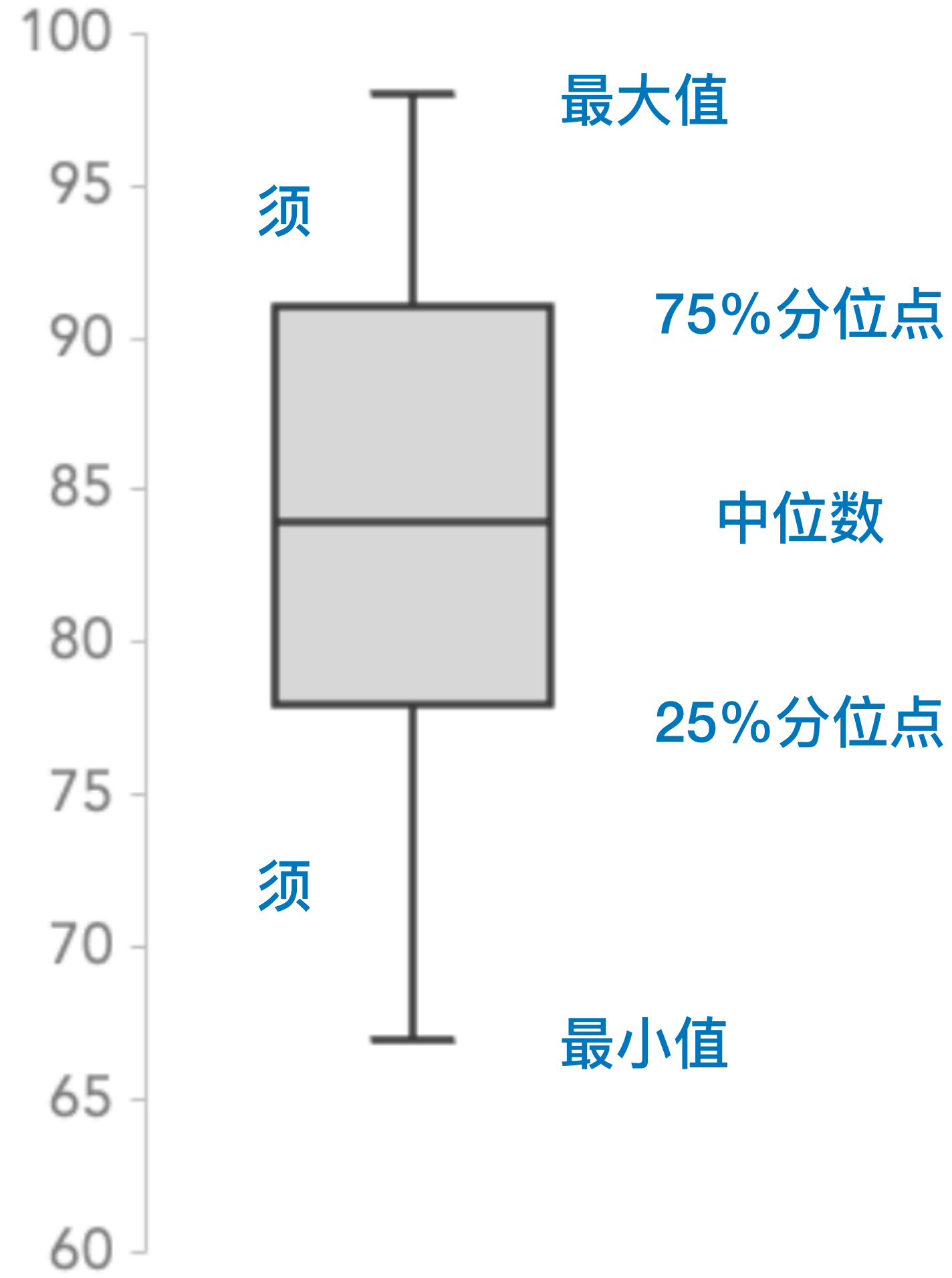
### Causes of Mortality of the British Army in the East



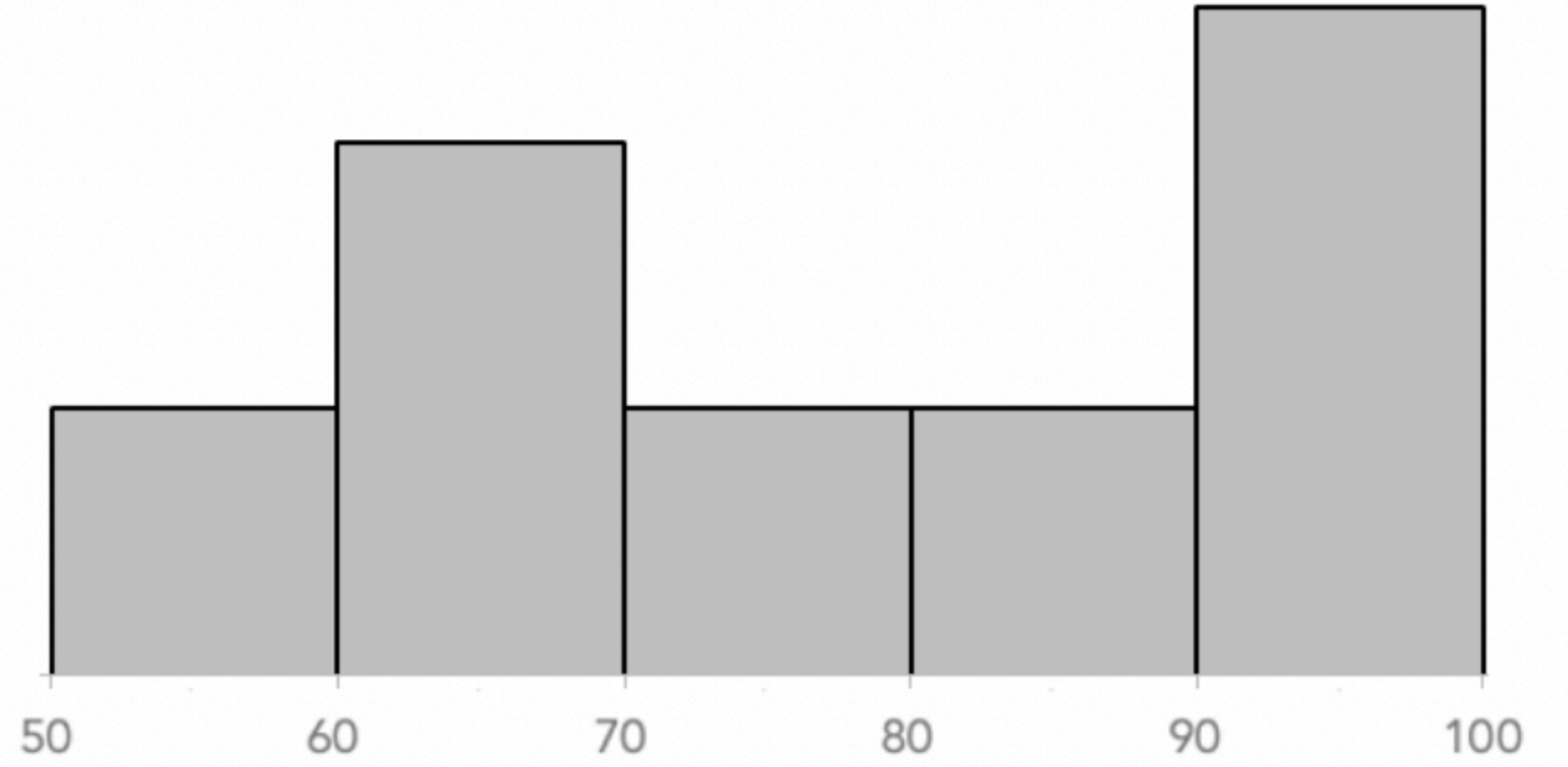
Friendly, M. & Wainer, H. (2021). *A History of Data Visualization and Graphic Communication*. Harvard University Press.  
作者将南丁格尔玫瑰图中的数据以折线图的形式重新描绘。

# 展示数据的分布

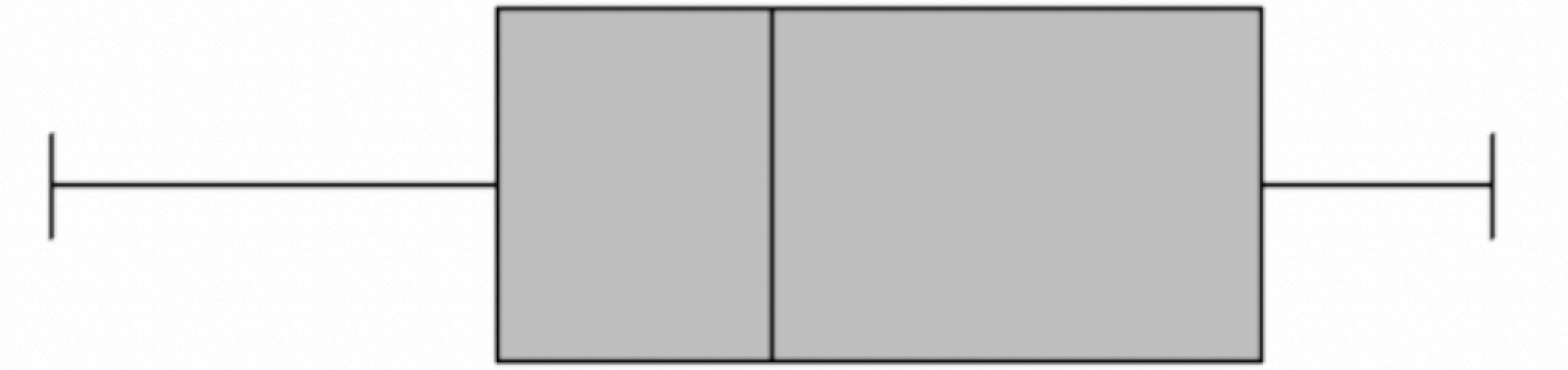
# 箱线图



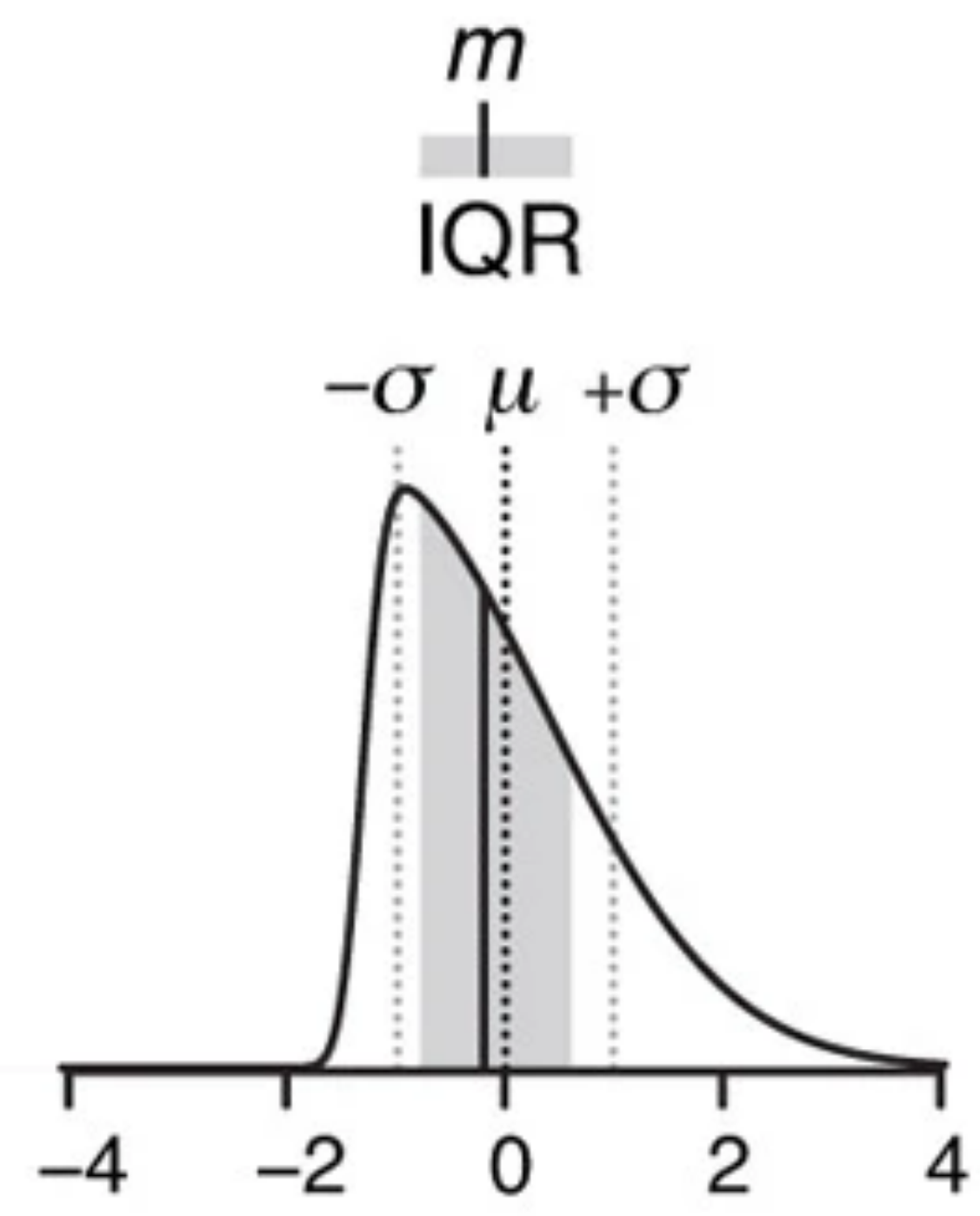
# 直方图



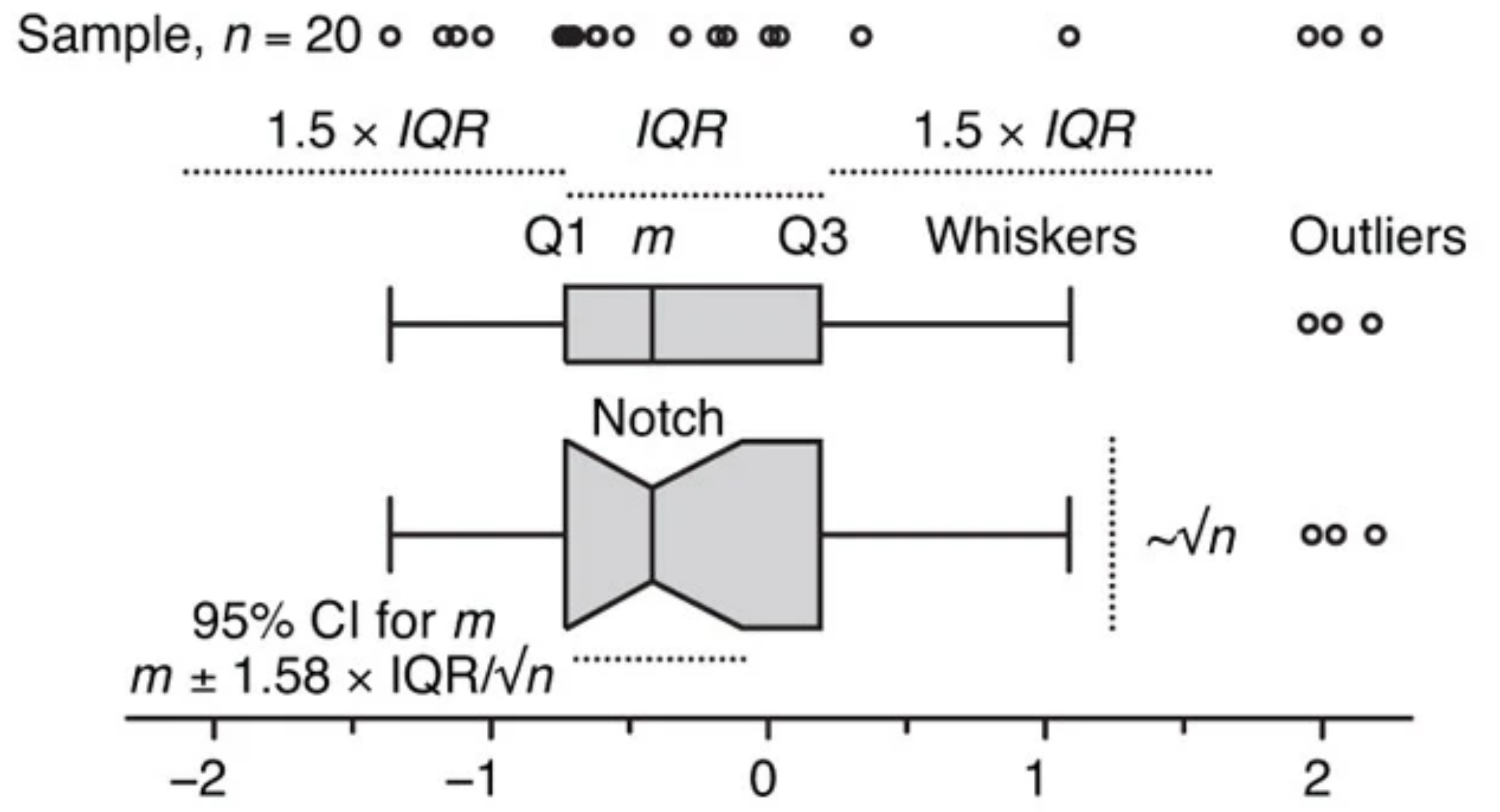
# 箱线图



**a** Population distribution



**b** Construction of a box plot

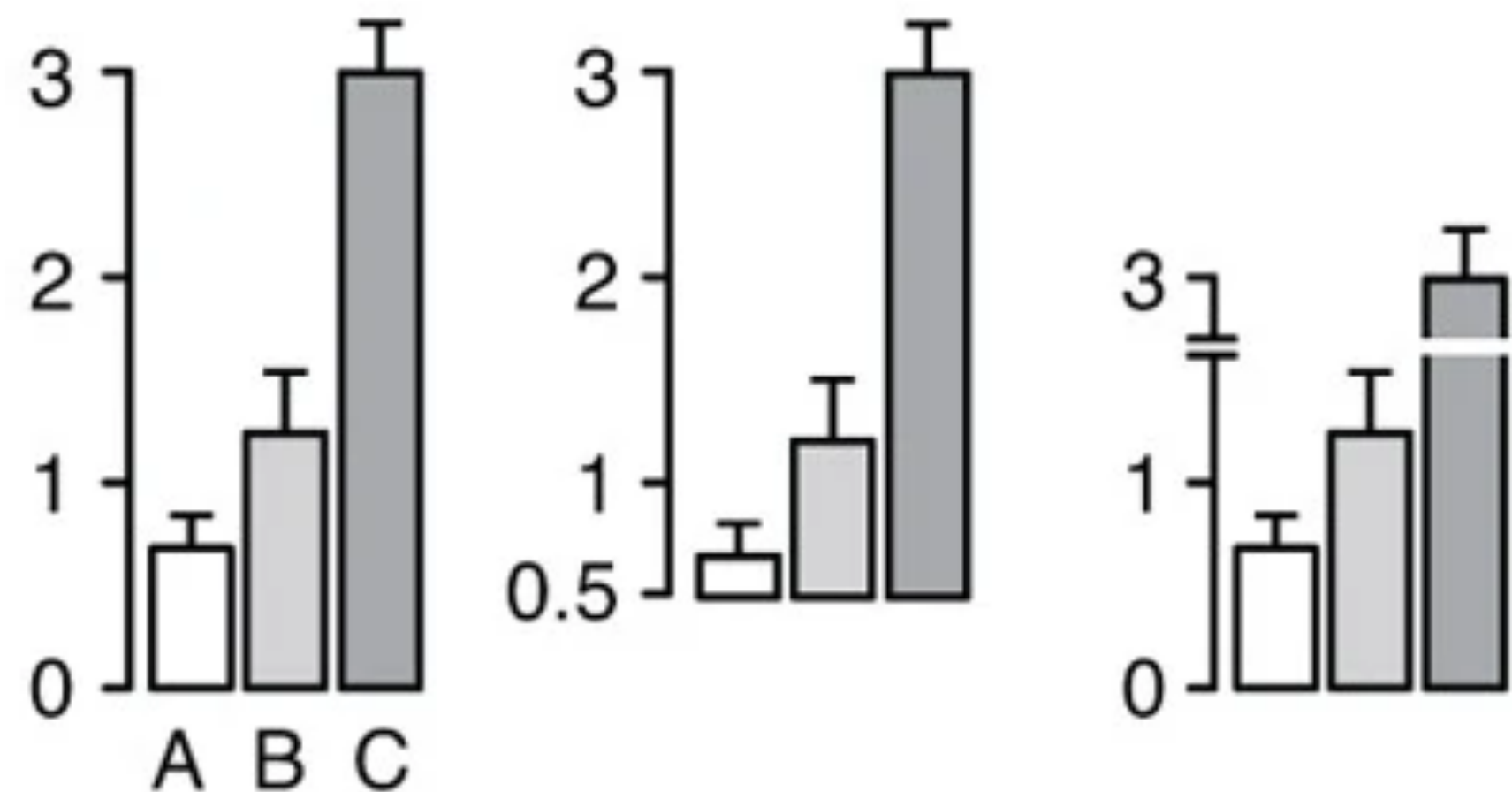


Krzywinski, M. & Altman, N. (2014). Visualizing samples with box plots. *Nature Methods*, 11:119-120.

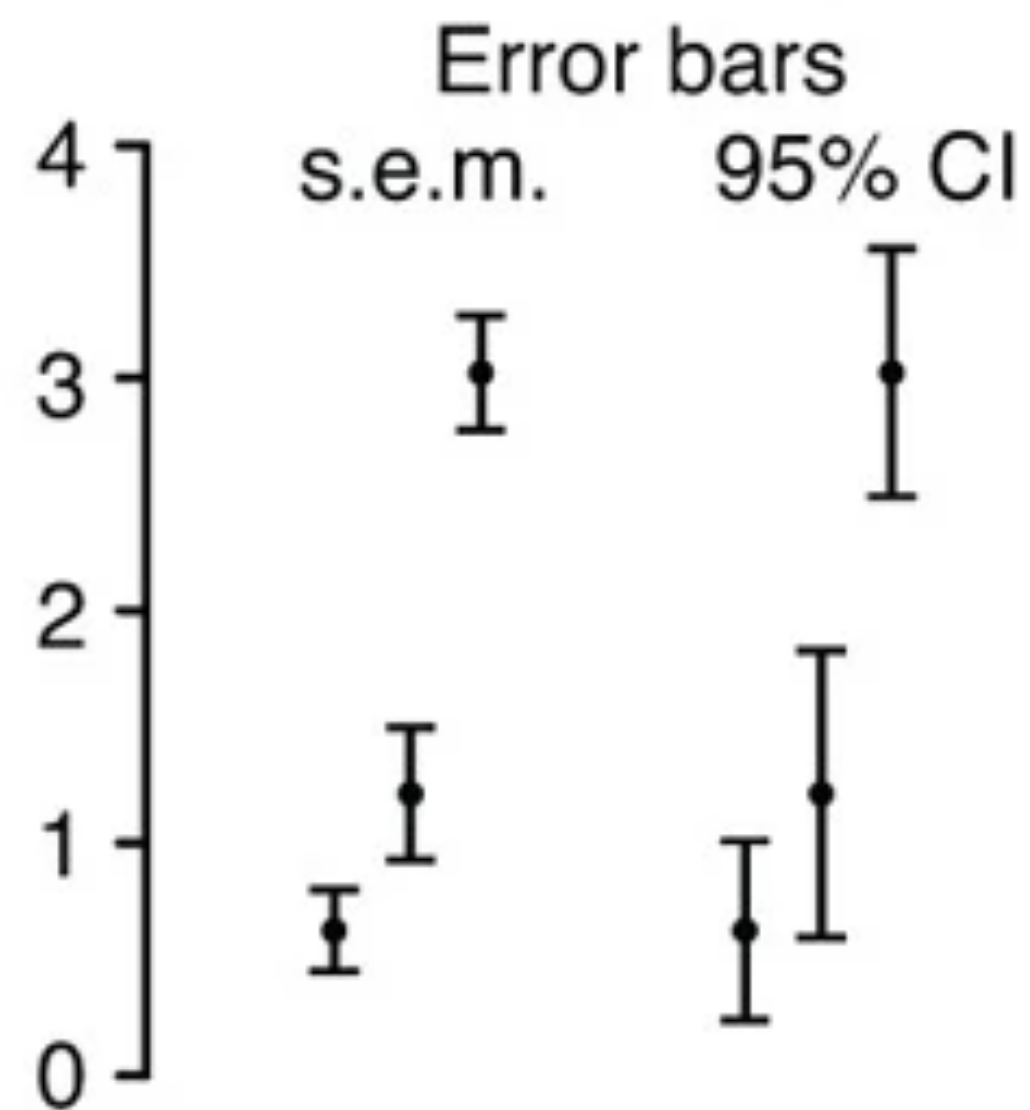
另一种常见的箱线图将须的两端绘制在箱的两端正负1.5箱长范围内的最大和最小观测值处，并将剩余的观测值作为极值绘制。如果数据服从正态分布，则须的两端之间约等于  $\mu \pm 2.7\sigma$ ，大约包含总数据的 99.3%。

**a** Means as bar plots

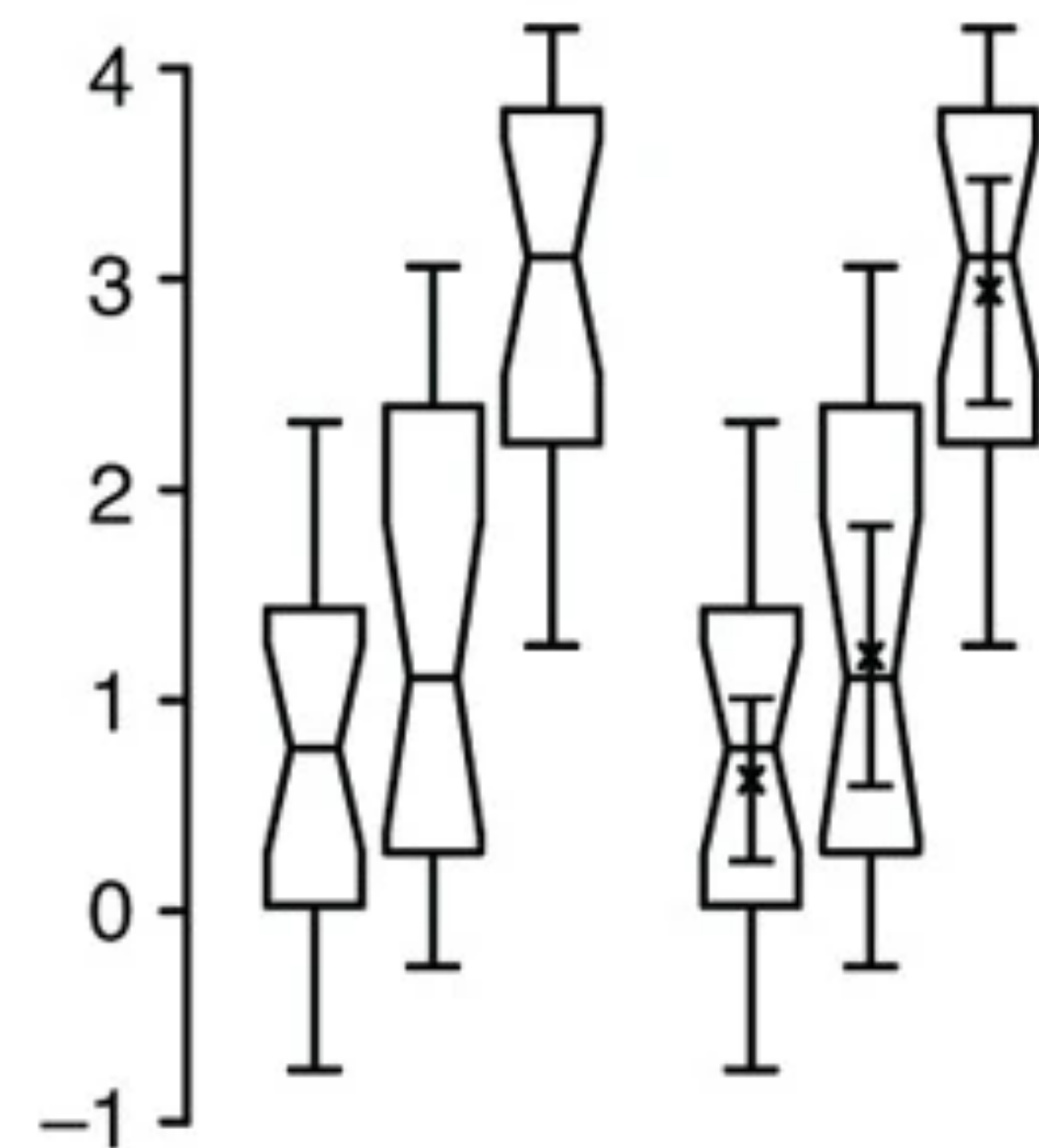
Not recommended



**b** Means as scatter plots

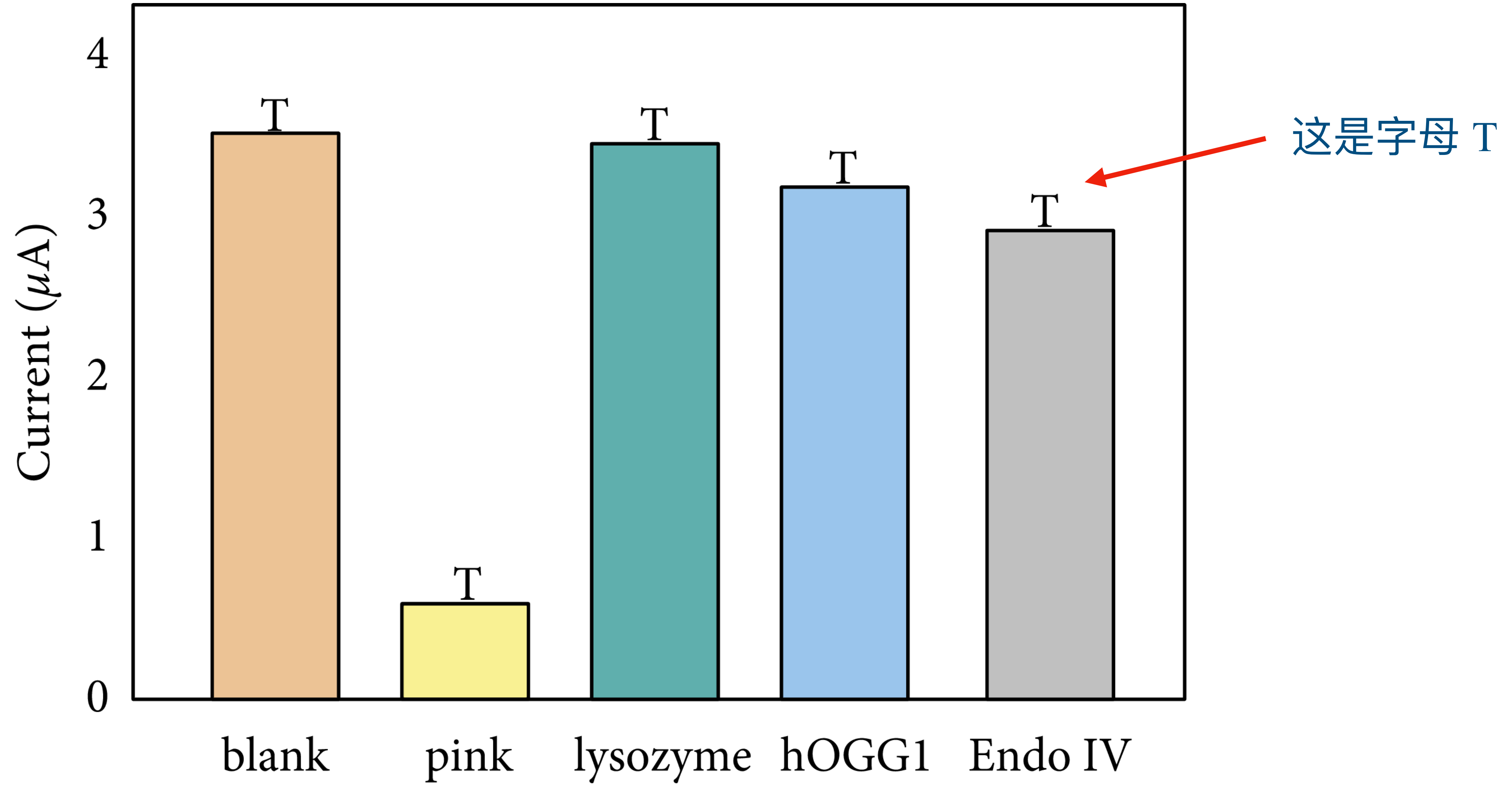


**c** Box plots with optional means and 95% CI



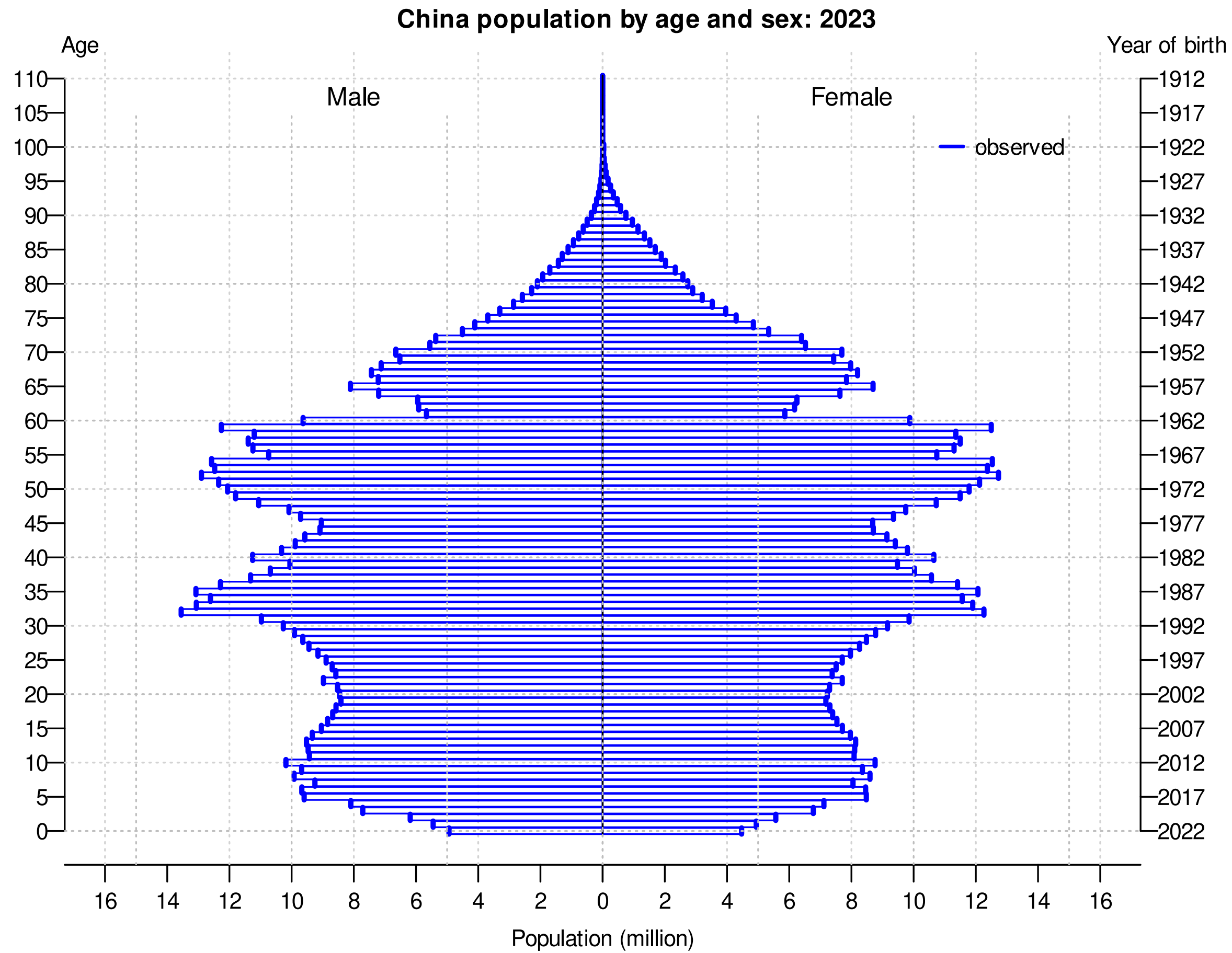
Krzywinski, M. & Altman, N. (2014). Visualizing samples with box plots. *Nature Methods*, 11:119-120.

(a) 常见的以柱形图的形式绘制样本均值和标准误。不推荐以这种形式描绘样本的分布特征。(b) 样本均值  $\pm$  标准误或样本均值的95%置信区间更加合适。(c) 箱线图与置信区间相结合。



Gong, R. & Liu, B. (2022). Monitoring of Sports Health Indicators Based on Wearable Nanobiosensors, *Advances in Materials Science and Engineering*, vol. 2022, Article ID 3802603, 11 pages, [Retracted]. Figure 9.

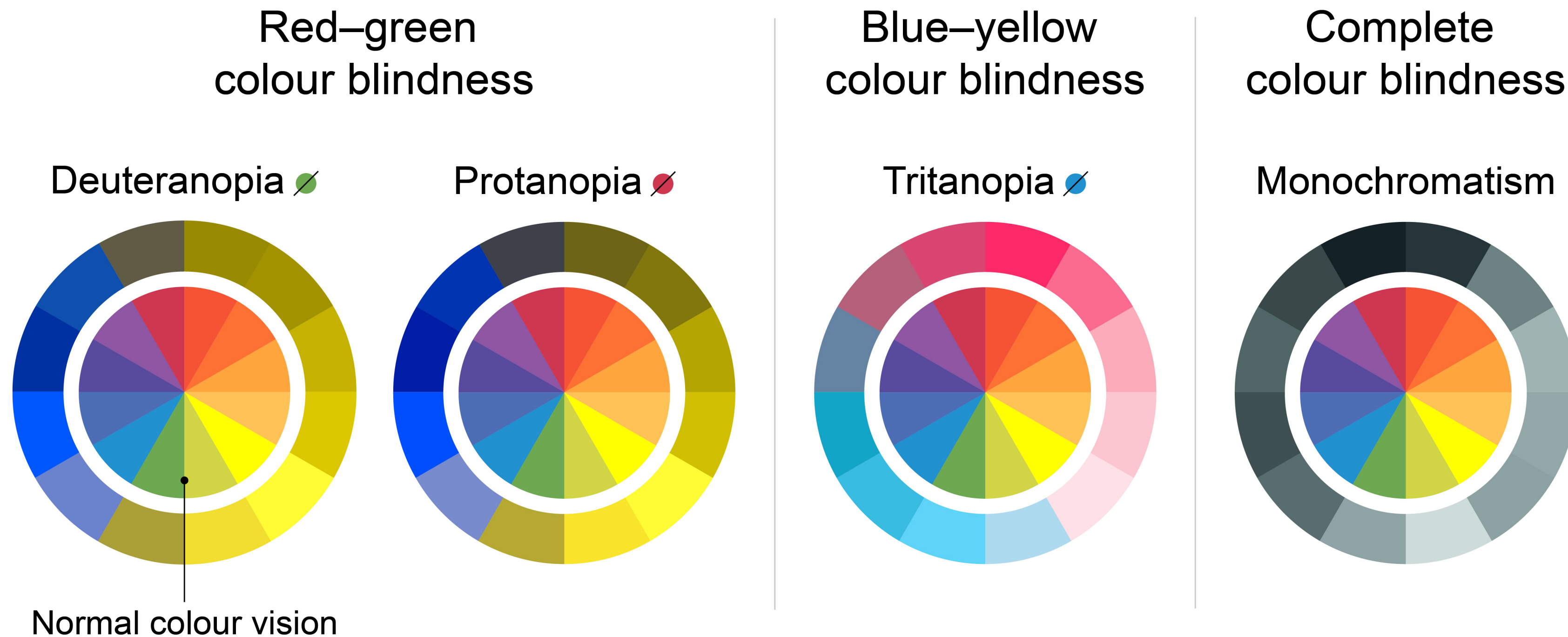
# 人口金字塔



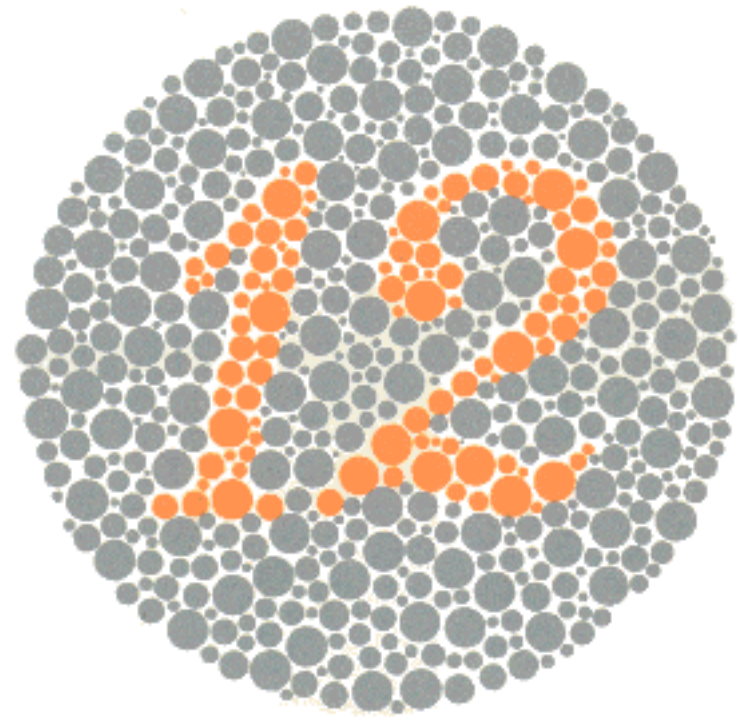
# 颜色感知偏差

# 并不是所有人都能正确感知颜色

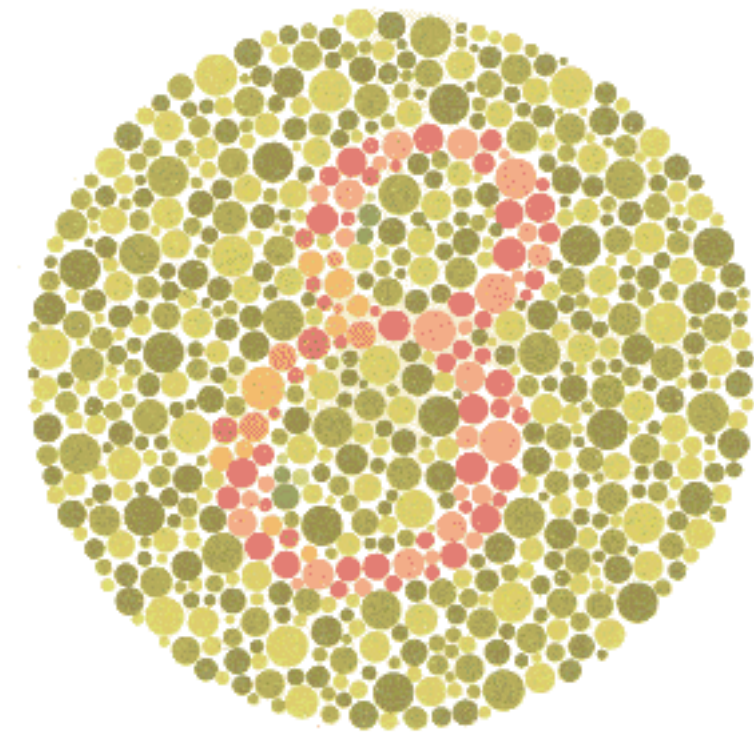
- 男性中有近 8% 为先天性色弱或色盲，女性中的比例约为 0.5%。
- 大多数色弱为红绿系色弱，极少数为蓝黄系色弱或色盲。



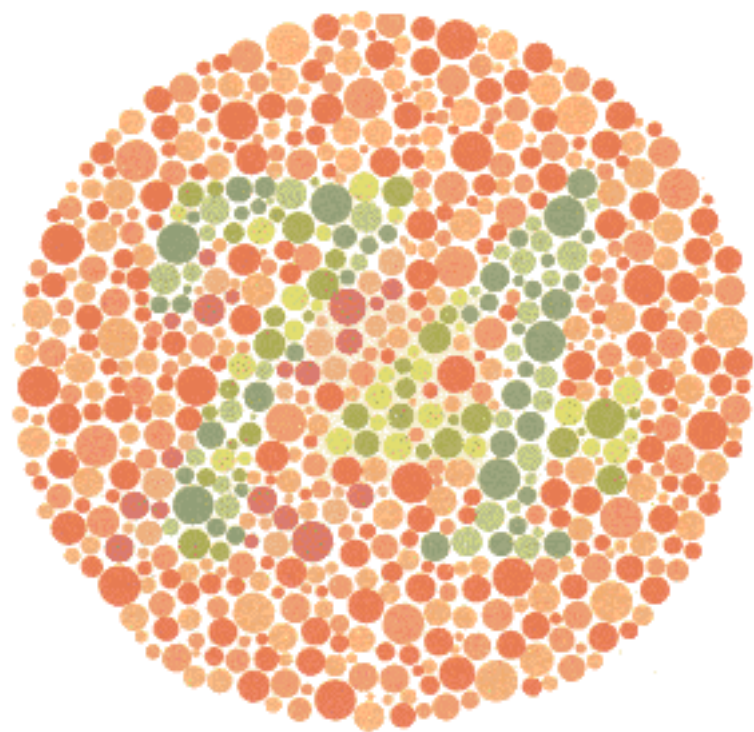
<https://www.healthdirect.gov.au/colour-blindness>



所有人都应该能够看到数字 12，包括色弱和色盲。



视觉正常者能看到数字 8，红绿系色弱者能看到 3，色盲者看不到数字。



视觉正常者能看到数字 74，红绿系色弱者能看到 21，色盲者看不到数字。

Colour-vision deficient (CVD)

Colour-blind

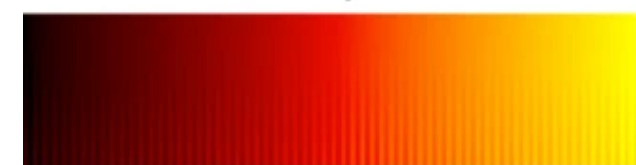
viridis



cividis



kry



thermal



batlow



jet



Deuteranopia



Protanopia



Tritanopia



Grey scale



✓

✓

✓

✓

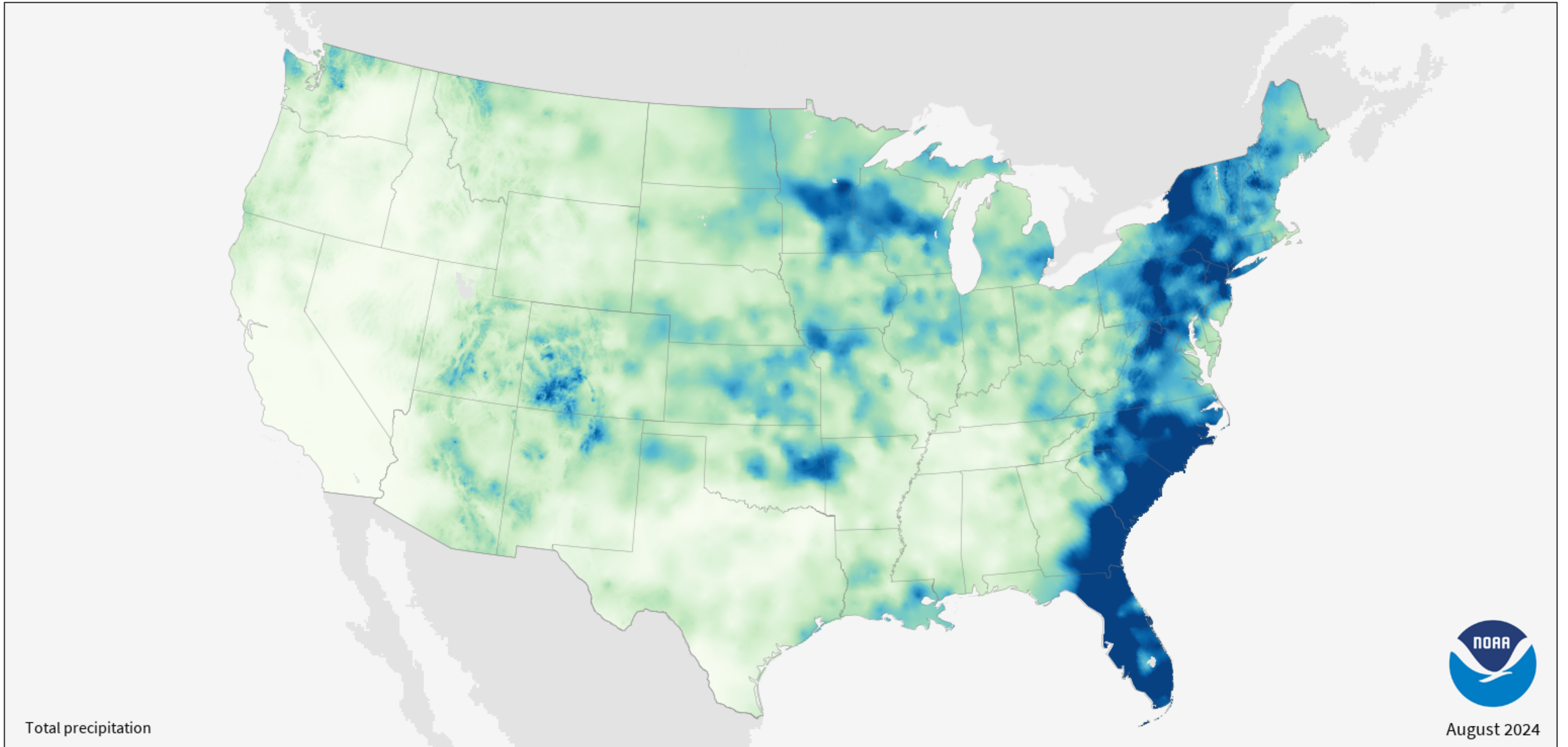
✓

✗

CVD-friendly



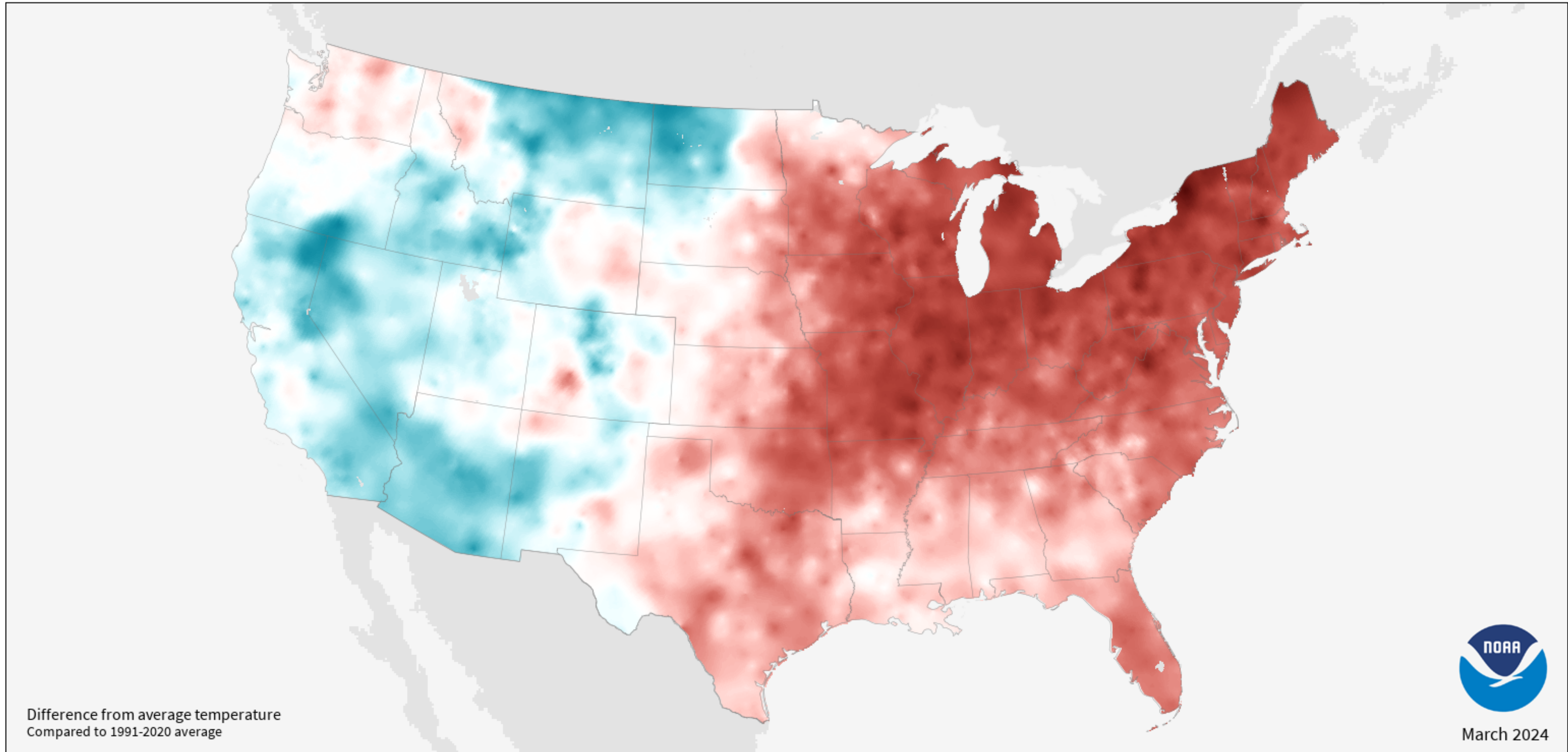
最常见的彩虹色系无法提供单向平滑的颜色变化



◀ less

more ▶

<https://www.climate.gov/>



◀ cooler

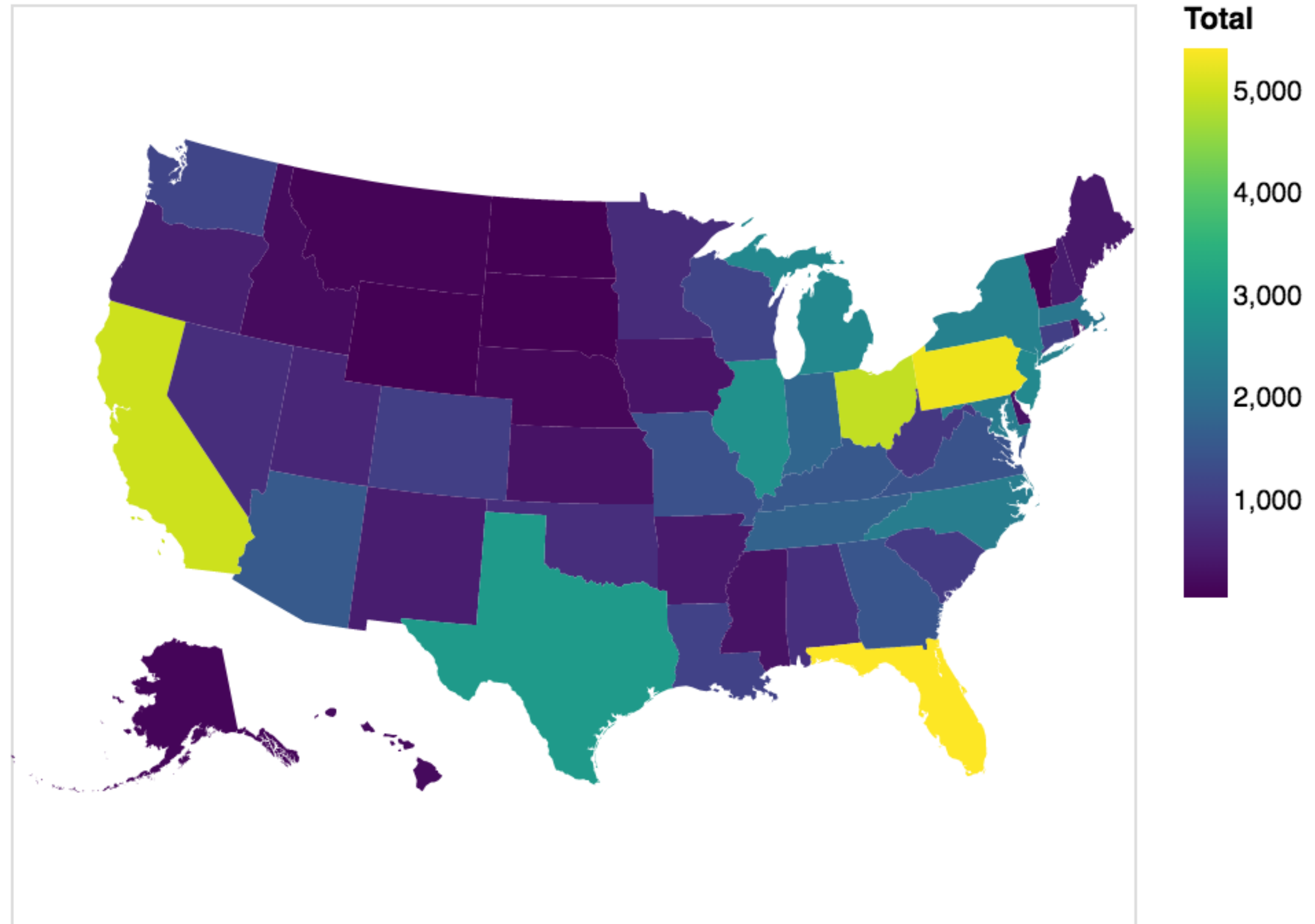
warmer ▶

<https://www.climate.gov/>

# A growing drug crisis

Drug overdose deaths by US state, Jan 2018

**Total drug overdose deaths by state, January 2018**



Source: CDC

<https://www.playfairprize.com/worked-examples>

# 关于颜色的选择：更多信息

- Picking A Colour Scale For Scientific Graphics  
<https://betterfigures.org/2015/06/23/picking-a-colour-scale-for-scientific-graphics/>
- Scientific colour maps  
<https://www.fabiocrameri.ch/colourmaps/>
- ColorCET  
<https://colorcet.com/>
- Beautiful colormaps for oceanography: cmocean  
<https://matplotlib.org/cmocean/#why-jet-is-a-bad-colormap-and-how-to-choose-better>
- mpl colormaps  
<https://bids.github.io/colormap/>
- ColorBrewer 2.0  
<https://colorbrewer2.org/>