

社会科学中的政治学

马啸

北京大学政府管理学院

上周内容回顾

- 什么是政治学
- 几种对政治学的“误解”
- 政治学要解决什么问题
- 政治学的现实意义
- 课程要求

课程要求补缺

- 课程主要内容：课件+课堂讲授
 - 课程录像上线一周（选课完成前特殊安排）
- 缺席需事先请假，事后不补
 - 哪些是请假的合理事由？
- 读书报告与观察报告
 - 合理、规范的引证（参考文献不算字数，但解释性的脚注算字数）
 - 不能自我抄袭
 - 之后会提供一些往年的优秀范例

本周内容

- 实证与规范社会科学
- 什么是实证社会科学
- 概率性与决定性理论
- 因果关系
- 相关性与因果关系
- 如何识别因果关系
- 对政治学科学性的反思

政治学是科学吗？

- 关于科学性的争论
- 从如何翻译“政治学系”窥见一斑：politics, government, political science
- 整体趋势：由政治哲学向政治科学，由定性向定量转变
- 政治学者更多的称呼自己为政治科学家（political scientists）
- 转变出现的原因：围绕特定问题的论战的深入，对确定答案的追求
- 通过科学的方法解决学术争议

1. 两种类型的社会科学

- 实证社会科学：Positive Social Science
 - 以证据为导向，结论由证据得出（“世界为什么是这样？”）
 - 实证社会科学（应用经济学、社会学），比较政治学、政治经济学、公共政策等
- 规范社会科学：Normative Social Science
 - 以观点为导向，论证围绕观点展开（“世界应该是什么样”）
 - 政治哲学、部分以政策目标为导向的社会科学

判断下列问题（实证or规范？）

- 为什么工业革命在英国发生而不是在同时期的中国？（实证）
- 为什么必须先发展好经济才能建设政治文明？（规范）
- 什么是人类社会最理想的政治模式？（规范）
- 依法治国与以德治国哪种手段更有效？（实证）
- 为什么民主制度在北美获得成功却在南美经历曲折？（实证）
- 怎么样增强人民的幸福感和获得感？（规范）
- 经济援助对发展中国家是福还是祸？（实证）
- 民主和专制哪个更好战？（实证）

两种类型的社会科学

- 学术对话中的逻辑一致性
- 实证社会科学研究中并非没有规范性元素
 - 核心概念的界定同样是一个建构的过程
 - $1+1=?$ 中的规范元素

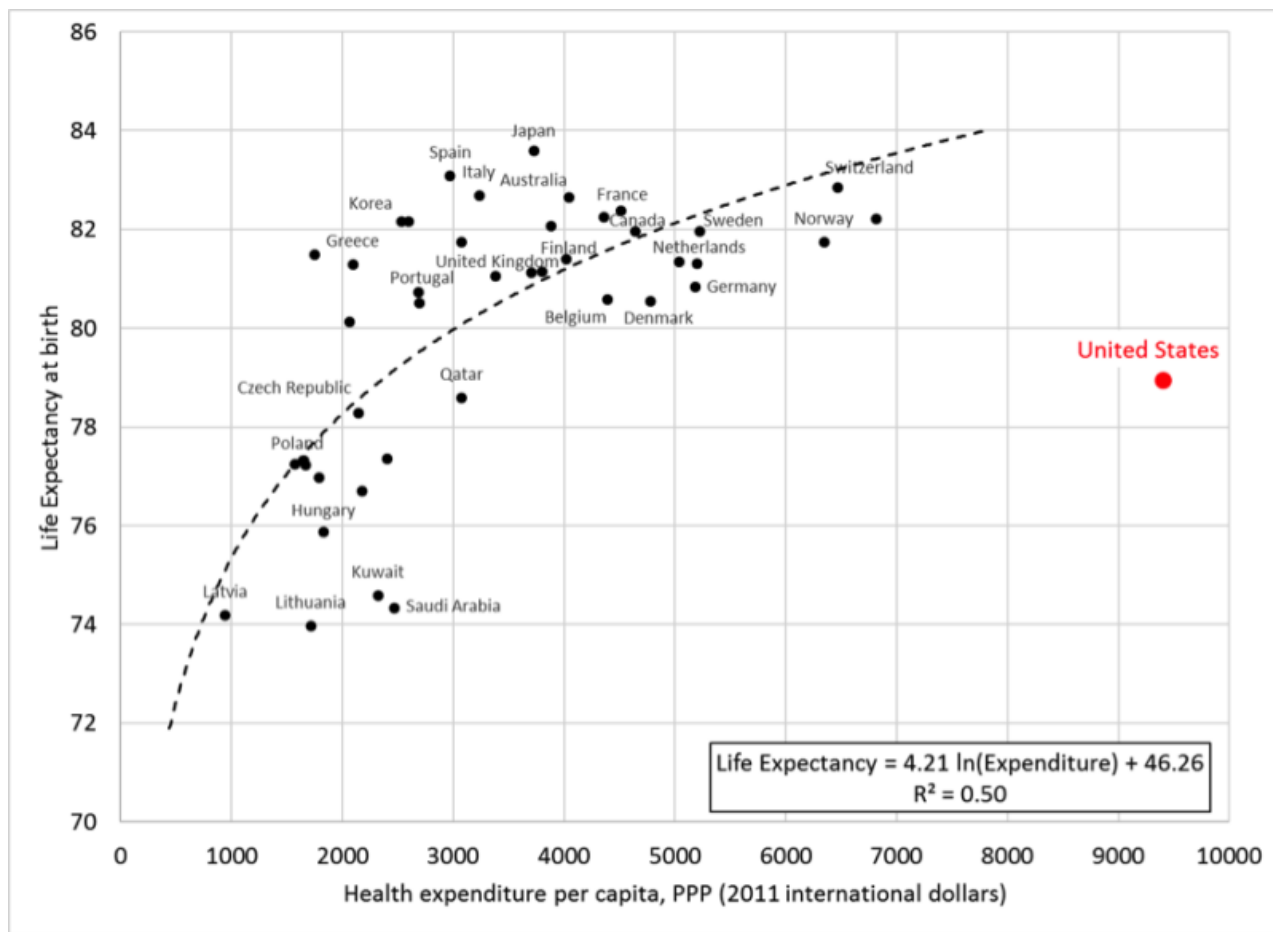
2. 什么是实证科学

- 科学的 (scientific) 假设、命题、理论的标准：可观察 (observability)、可检验 (testability)、可重复 (repeatability)、可证伪 (falsifiability)
- 一个反例：“车库里喷火的龙” (卡尔·萨根)
- 注意：科学命题与“对错”无关，“可证伪”只是一种逻辑条件
- 科学本身就是不断证伪已有命题和理论，寻求“真理”的过程
- 两个例子，哪个是科学的命题？：
 - 所有的天鹅都是白的
 - 上帝是万能的
- 思考：日常生活中的哪些说法不是科学的命题/假设

3. 概率性与决定性理论

- 理论：对事实的一般规律的总结
- 概率性理论（probabilistic theory）：注重个体间的相似性，但承认个体间差异的存在
 - 更高的教育水平将带来更高的收入
 - 吸烟有害健康
 - 左翼选民倾向支持民主党候选人
- 决定性理论（deterministic theory）：强调预测结果的唯一性
 - 万有引力定律
 - 所有的天鹅都是白的
- 能用个案否定一般规律吗？

能因为个案否定一般规律吗？



概率性理论：很难

例：

人均医疗支出与人均寿命

决定性理论：一个反例即可证伪

例：

万有引力定律

在日常讨论中，我们需要注意区分两种不同的理论及其内在逻辑。避免用决定性理论的逻辑去讨论概率性理论。

作为概率性理论的社会科学

- 人类行为的复杂性
- 几乎所有的社会科学理论都属于概率性理论
- 与个体经验、直觉相反的观察未必一定能证伪概率性理论
- 需要更为严谨的研究设计去检验

4. 不同类型的实证社会科学研究

- 情形描述：发生了什么
 - 介绍一个未知的事实，或者对存在争议的事实进行辨析
 - 情形描述不代表研究可以脱离理论（社会科学和历史学的区别）
- 寻求因果关系：什么导致了什么
- 因果机制解释：什么如何导致什么
 - 因果机制的解释多与情形描述相结合

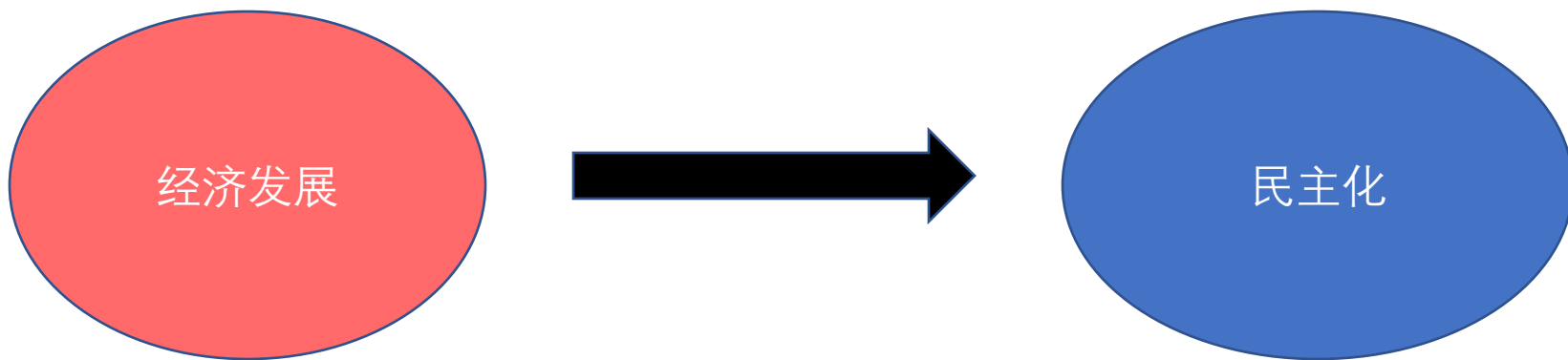
5. 寻求因果关系

- 什么导致了什么：社会科学最常见的研究类型
- 尝试用某种原因（因）解释一种现象（果）
- A导致B：A的变化导致了B的变化
- 例如如下假设：
- 个人受教育（A）水平越高，则长期收入（B）越高
- 一个国家的经济越依赖自然资源出口（A），则越有可能成为专制国家（B）
- 少数族裔选民（A）容易遭到政治家的忽视（B）

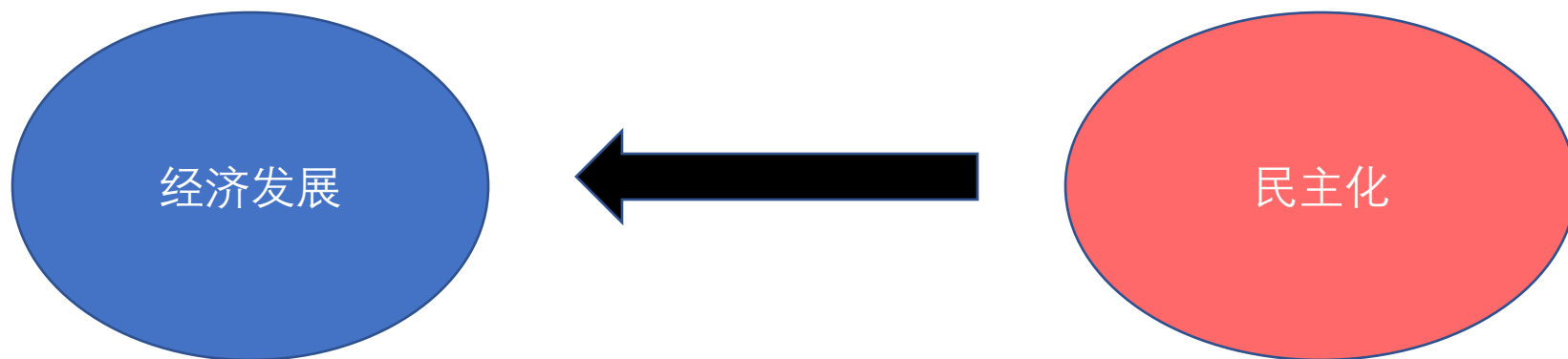
相关性与因果关系

- 我们可以在日常生活中观察到很多相关性
- 相关性：两个现象之间存在显著的线性关系（A变化B也同时变化）
- A, B可以是连续变量（例如：收入、身高、经济发展水平、体温），也可以是离散变量（例如：性别、死亡、是否是民主政体、是否吃了某种药）
- 例如：受教育水平越高，收入越高；一国的经济发展水平越高，越有可能是民主国家；服用名贵中草药的人寿命较长
- 相关性不等于因果关系：A与B呈现相关性的情况下，A也不一定导致B

反向因果关系



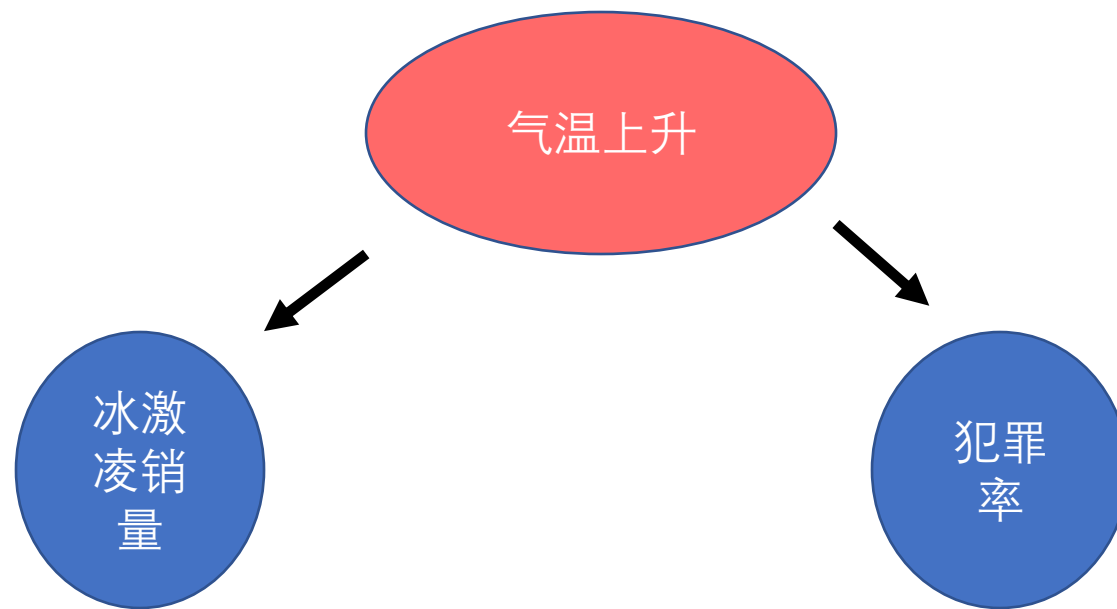
反向因果关系



遗漏变量偏差



遗漏变量偏差

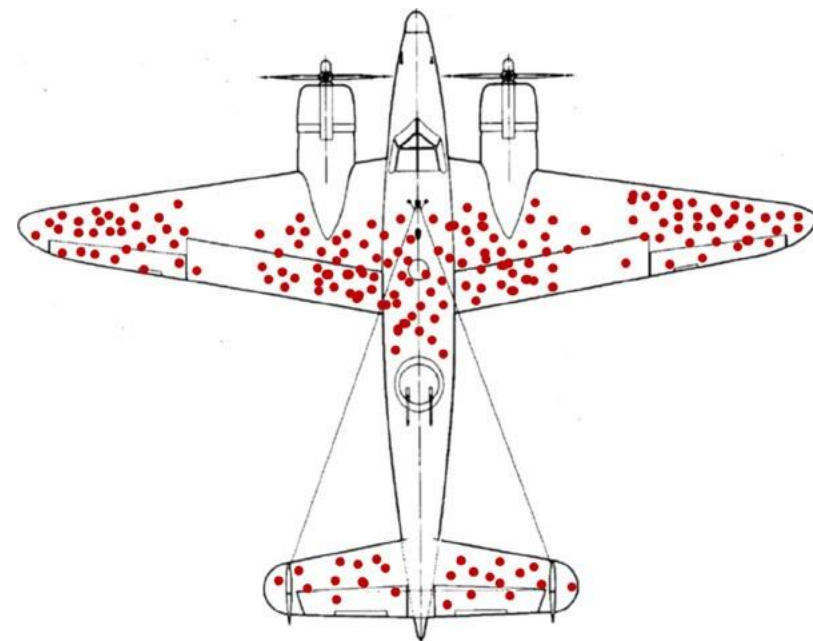


以下相关性一定是因果关系吗？

- 无家可归人数较多的地区犯罪率较高
- 黑人犯罪率更高
- “女司机更容易发生交通事故”
- 经常服用灵芝、燕窝等名贵药材，寿命更长
- 尝试戒烟的人死亡率（比正常吸烟的人）更高
- 对烟草征重税的国家肺癌人数较少
- 经济高度依赖资源出口（石油、天然气）的国家一般都是专制国家

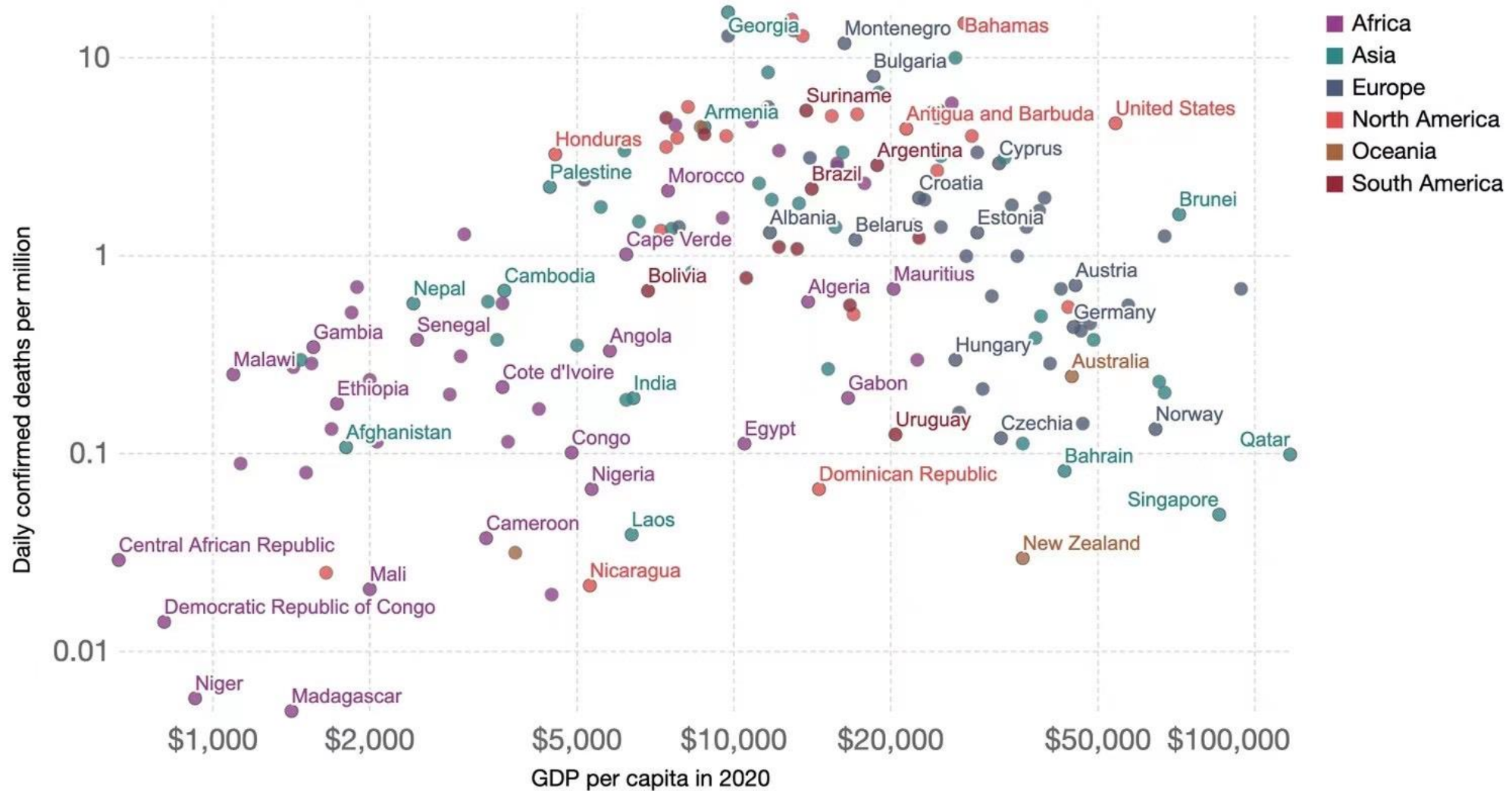
其他一些认知世界的偏差

- 样本选择偏差：观察对象的群体选择差异，会影响结论
- 例如：尝试戒烟的人死亡率（比不试图戒烟的人）更高
- 收据收集时要考虑数据的代表性，同时注意数据收集方法中可能引入的偏误（bias）
- 例如：电话访谈问人们是否使用电话
- “中国人不诚信？”的实验
- 2016年美国总统选举结果的预测错误
- 经典案例：战斗机上的弹孔



Confirmed COVID-19 deaths per million vs GDP per capita, Sep 9, 2021

Shown is the 7-day rolling average. Limited testing and challenges in the attribution of the cause of death means that the number of confirmed deaths may not be an accurate count of the true number of deaths from COVID-19. GDP per capita is adjusted for price differences between countries (it is expressed in international dollars).

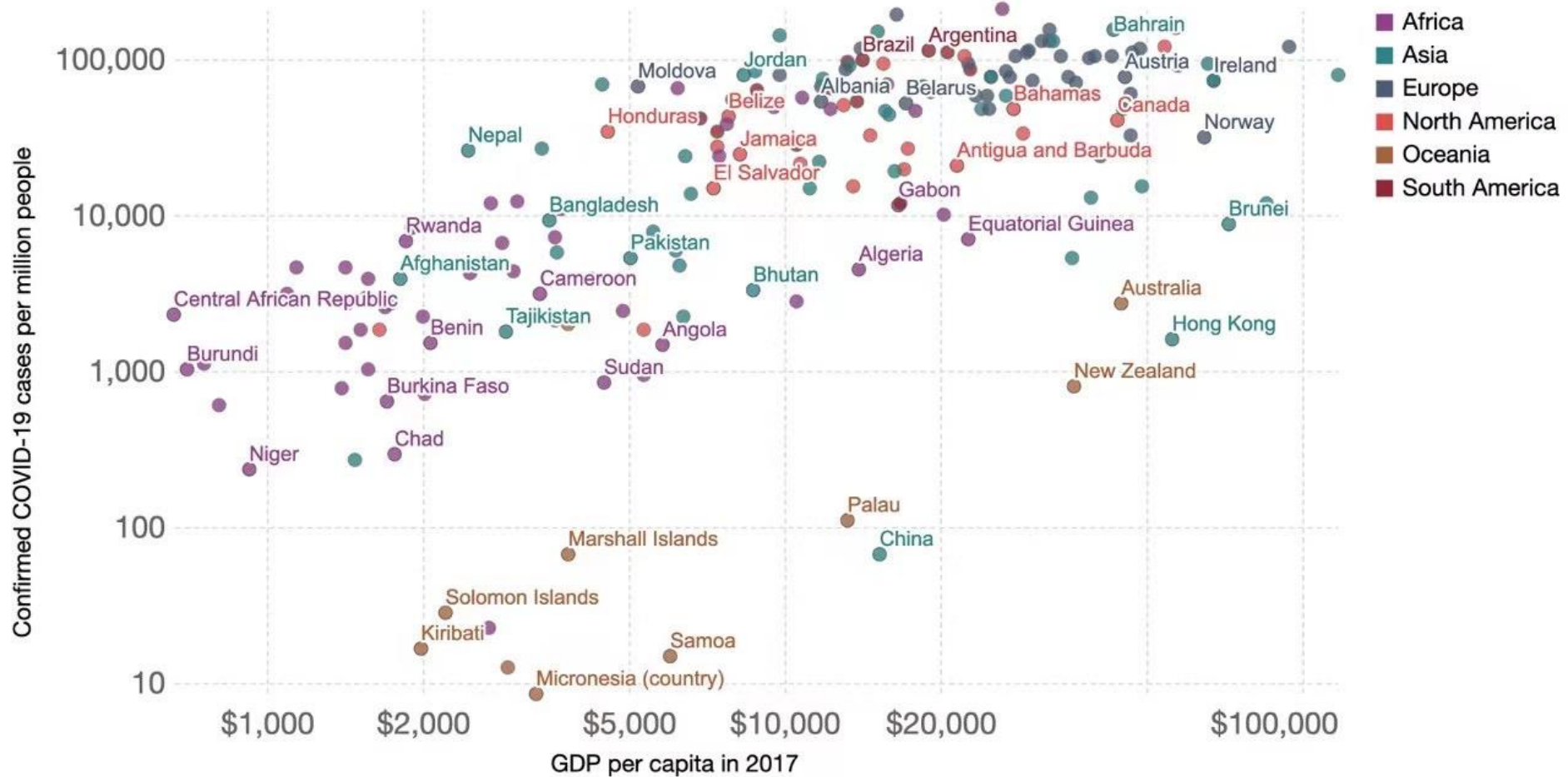


Cumulative confirmed COVID-19 cases per million vs. GDP per capita, Sep 9, 2021

Our World
in Data

The number of confirmed cases of COVID-19 is lower than the number of total cases. The main reason for this is limited testing.

GDP per capita is adjusted for price differences between countries (it is expressed in international dollars).



样本选择偏差

样本选择偏差本质上也是一种“遗漏变量偏差”，即某一未观察到的变量影响了研究对象是否更容易被研究者观测到

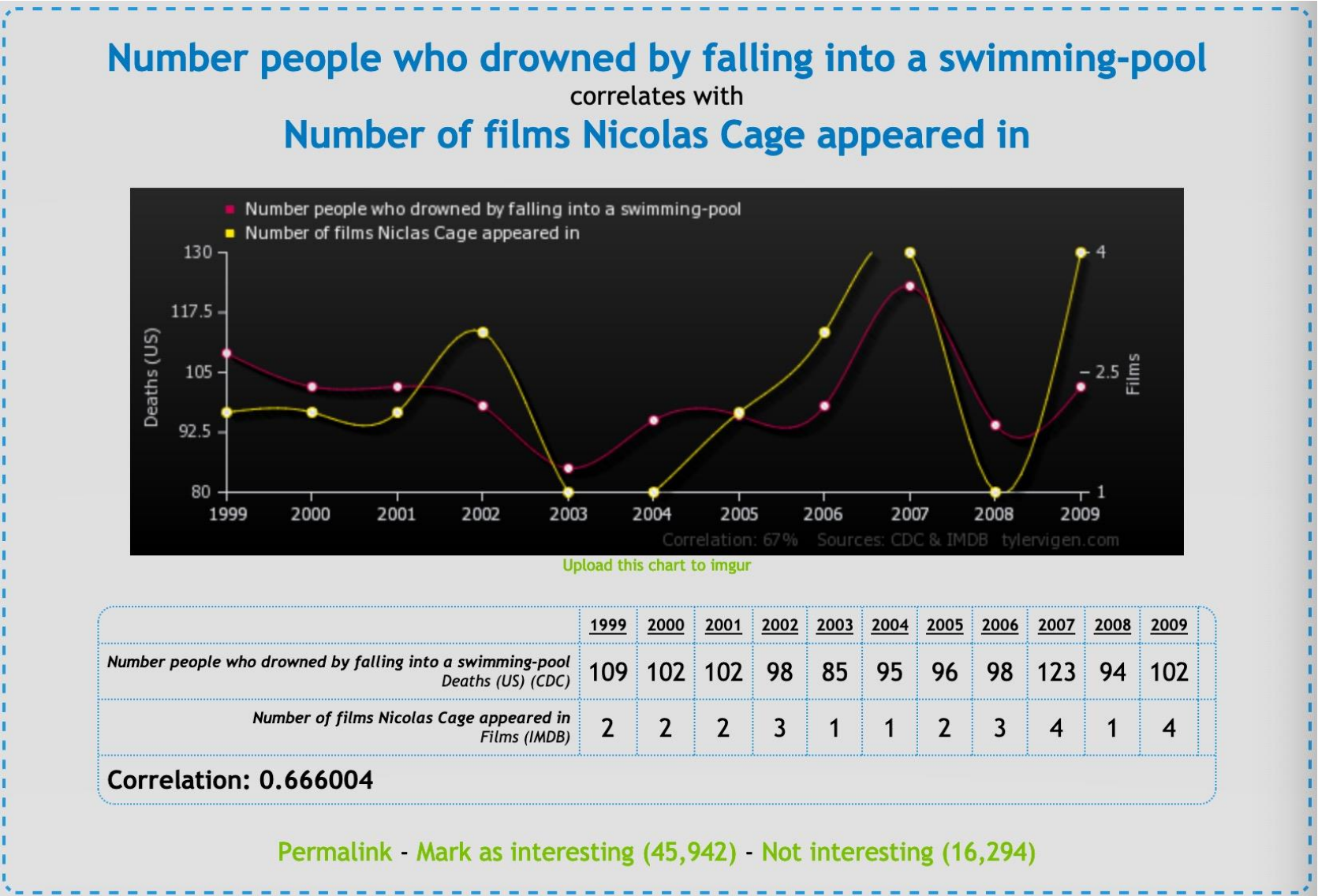
例如：尝试戒烟的人死亡率（比不试图戒烟的人）更高

遗漏变量：个体的健康状况

在生活中我们还能想到哪些样本选择偏差的例子？

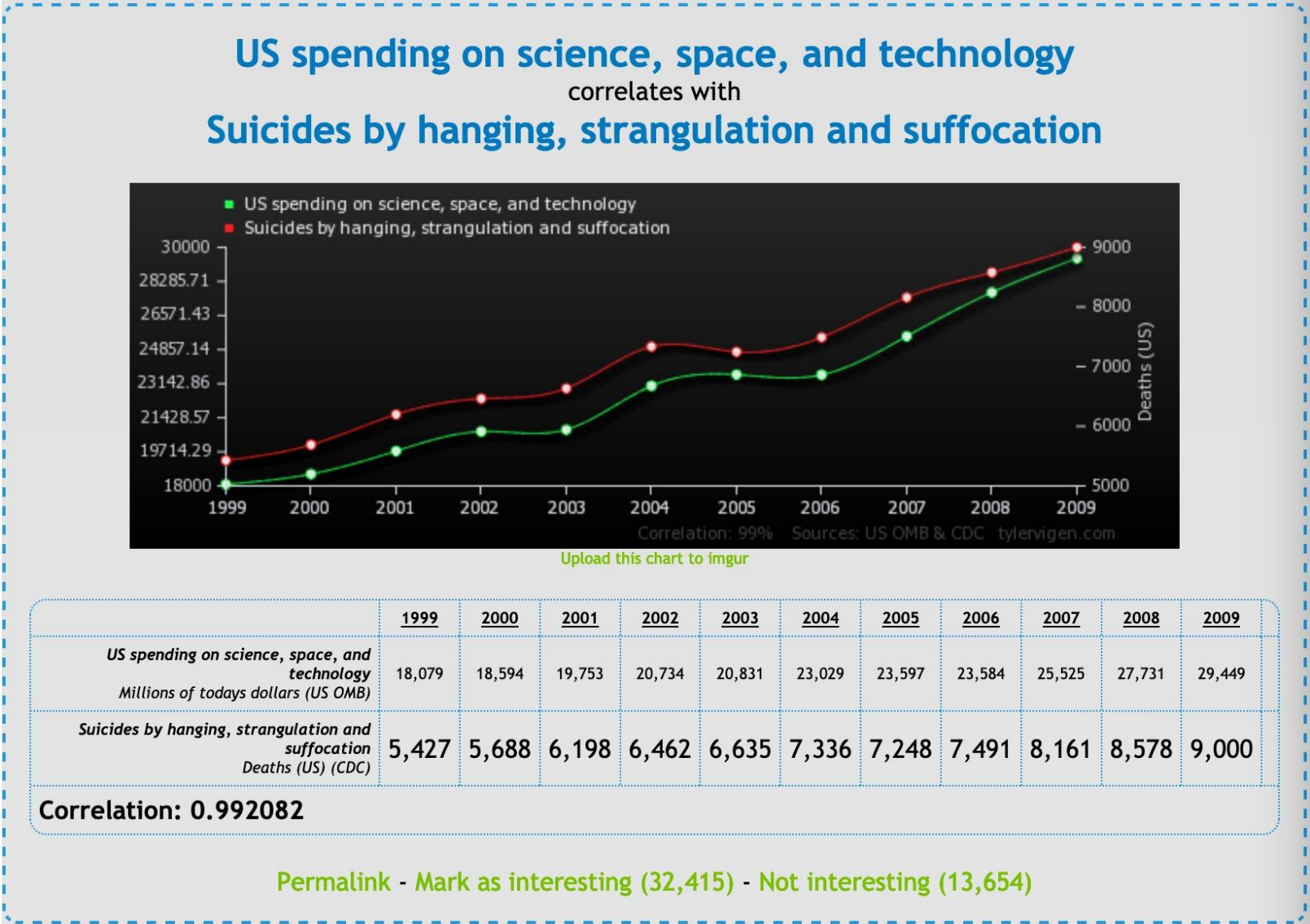
- 公共政策制定者需要特别注意样本选择偏差：
- 公共政策的制定需要广泛吸纳公众的意见
- 弱势群体发声能力最弱，而这些人却最恰恰需要政策倾斜

此外，相关性的存在有可能仅仅是巧合！



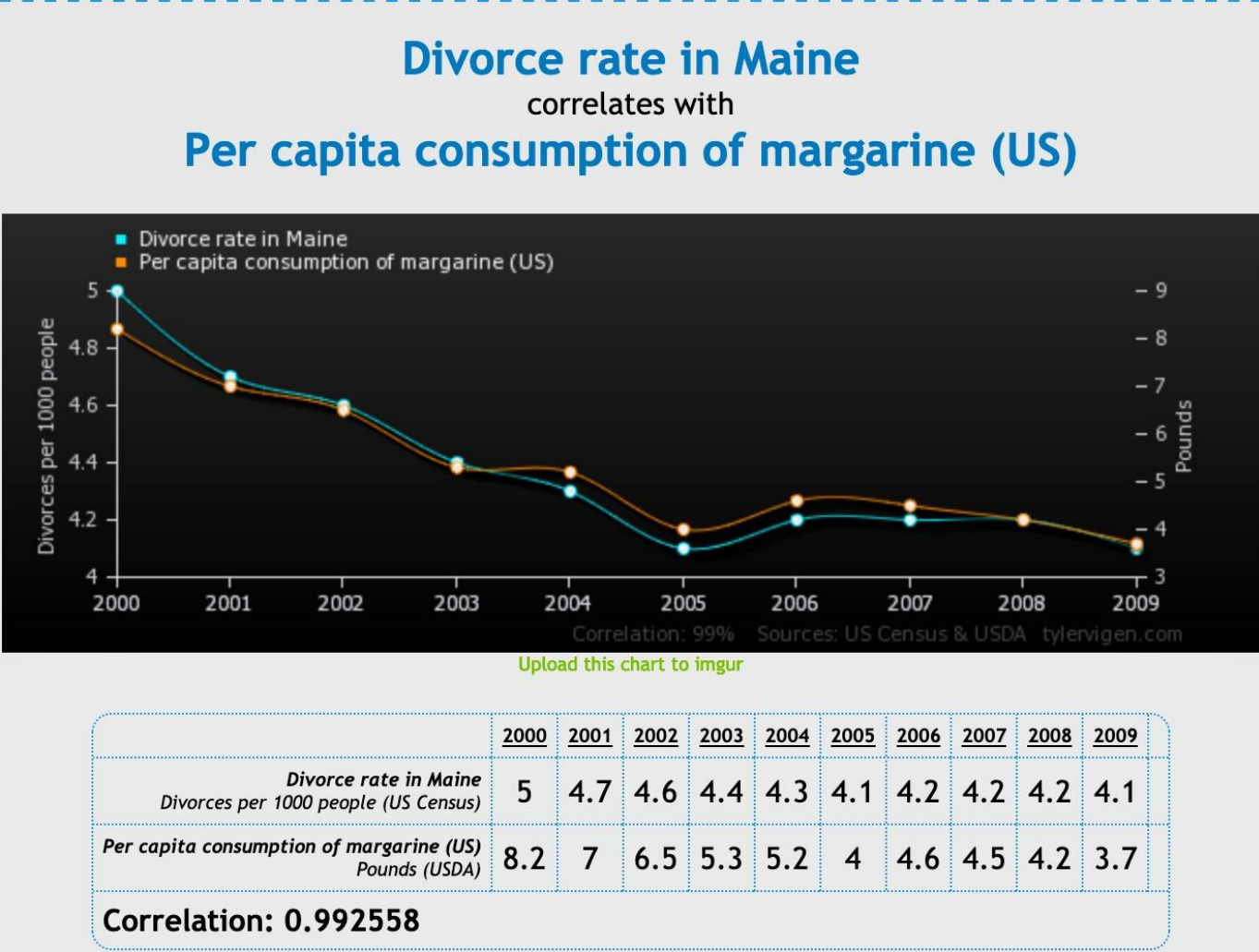
截图来源：
<https://tylervigen.com/page?page=1>

此外，相关性的存在有可能仅仅是巧合！



截图来源：
<https://tylervigen.com/page?page=1>

此外，相关性的存在有可能仅仅是巧合！

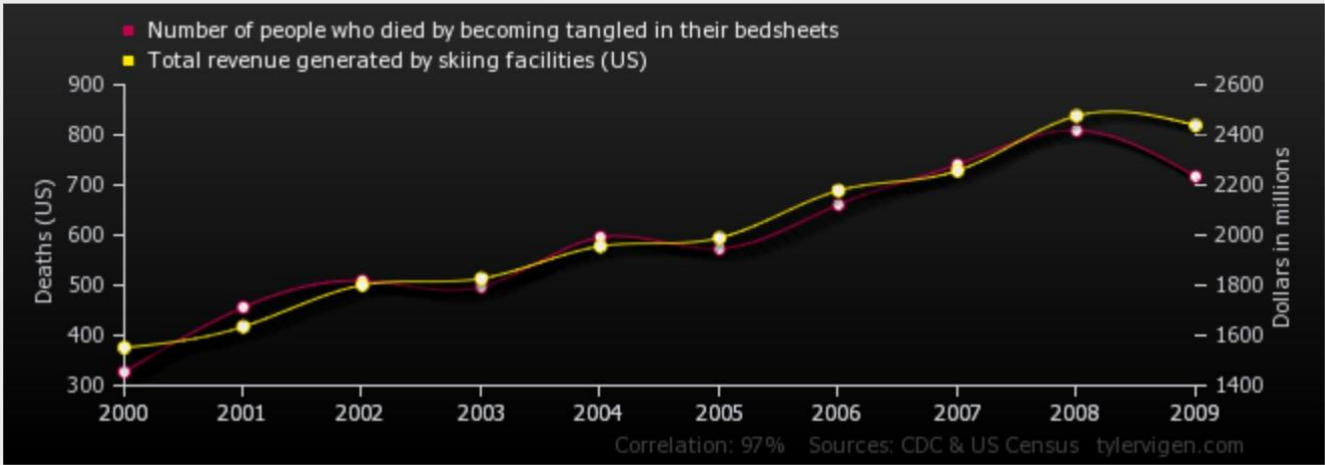


Permalink - Mark as interesting (23,284) - Not interesting (10,149)

截图来源：
<https://tylervigen.com/page?page=1>

此外，相关性的存在有可能仅仅是巧合！

Number of people who died by becoming tangled in their bedsheets
correlates with
Total revenue generated by skiing facilities (US)



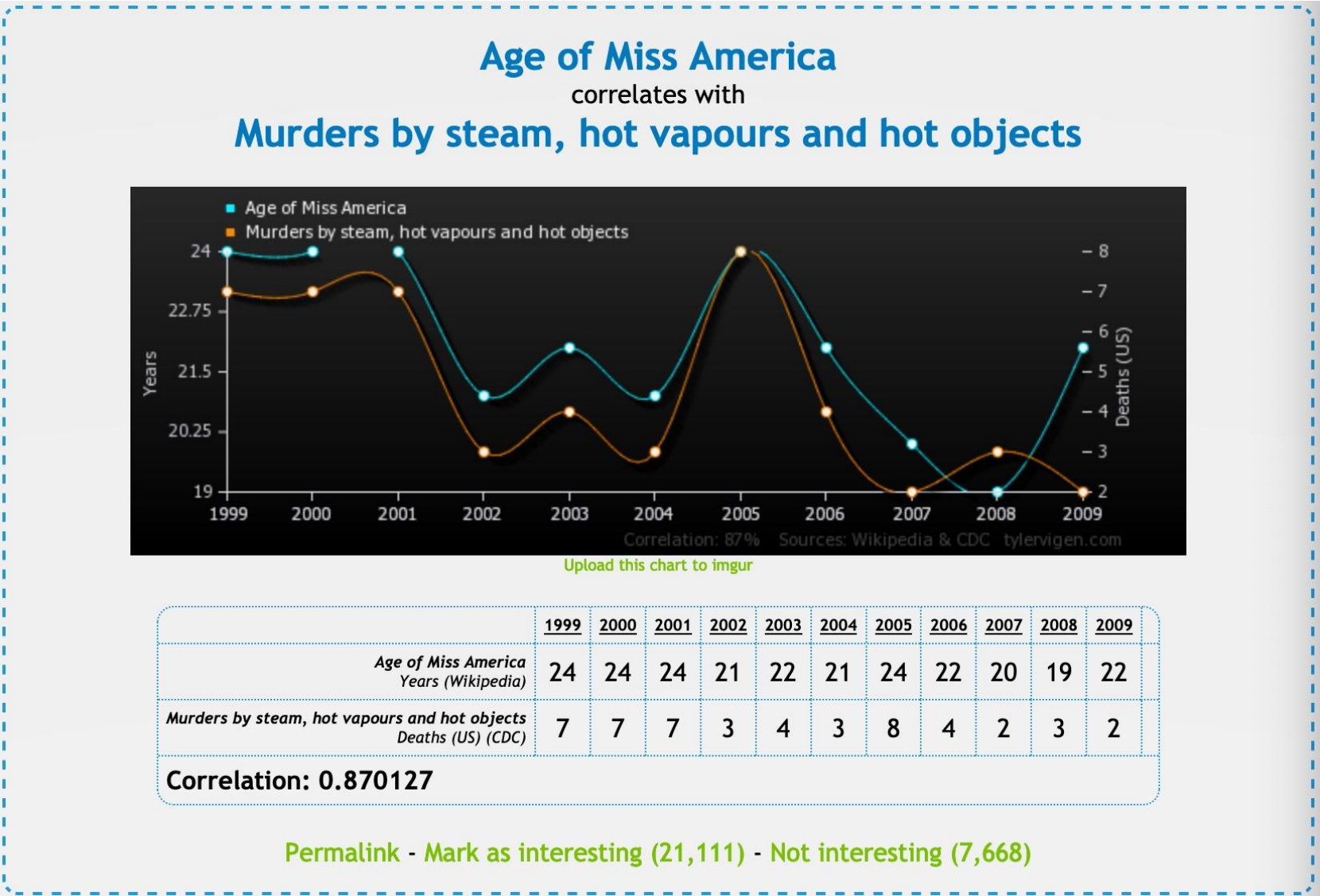
Upload this chart to imgur

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Number of people who died by becoming tangled in their bedsheets Deaths (US) (CDC)	327	456	509	497	596	573	661	741	809	717
Total revenue generated by skiing facilities (US) Dollars in millions (US Census)	1,551	1,635	1,801	1,827	1,956	1,989	2,178	2,257	2,476	2,438
Correlation: 0.969724										

Permalink - Mark as interesting (27,876) - Not interesting (7,977)

截图来源：
<https://tylervigen.com/page?page=1>

此外，相关性的存在有可能仅仅是巧合！



截图来源：
<https://tylervigen.com/page?page=1>

相关性与因果关系小结

- 寻求因果关系是实证社会科学的核心任务
- 相关性不等于因果关系（因果关系一定是相关性）：必要但不充分条件
- 哪些因素可能会影响相关性被解读为因果关系
 - 反向因果
 - 遗漏变量偏差
 - 样本选择偏差
 - 巧合！