

ZJU·IDG
Interactive Data Group

VIS投稿经验分享

谢潇
浙江大学

A Visual Analytics Approach for Exploratory Causal Analysis: Exploration, Validation, and Applications



Xiao Xie¹



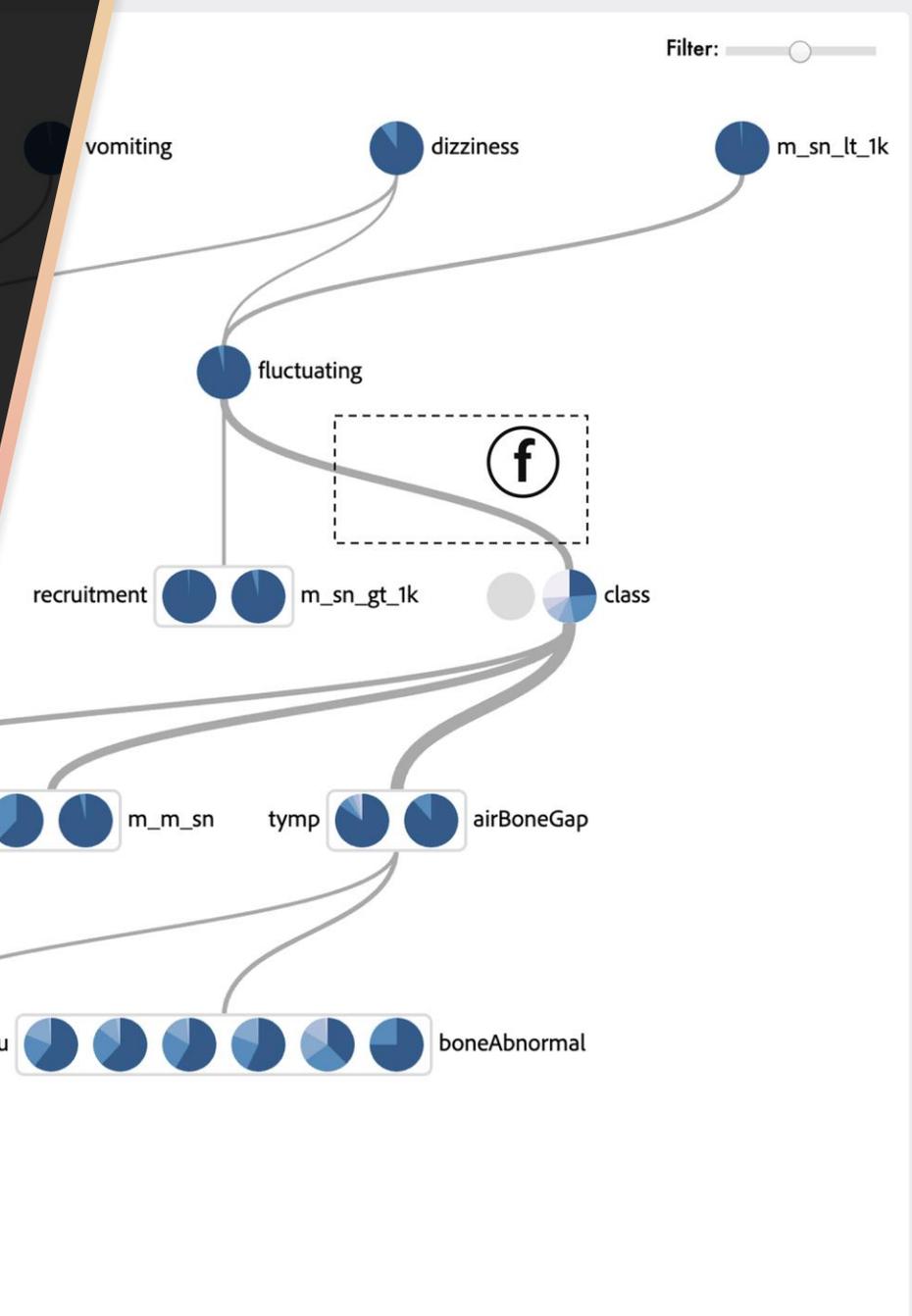
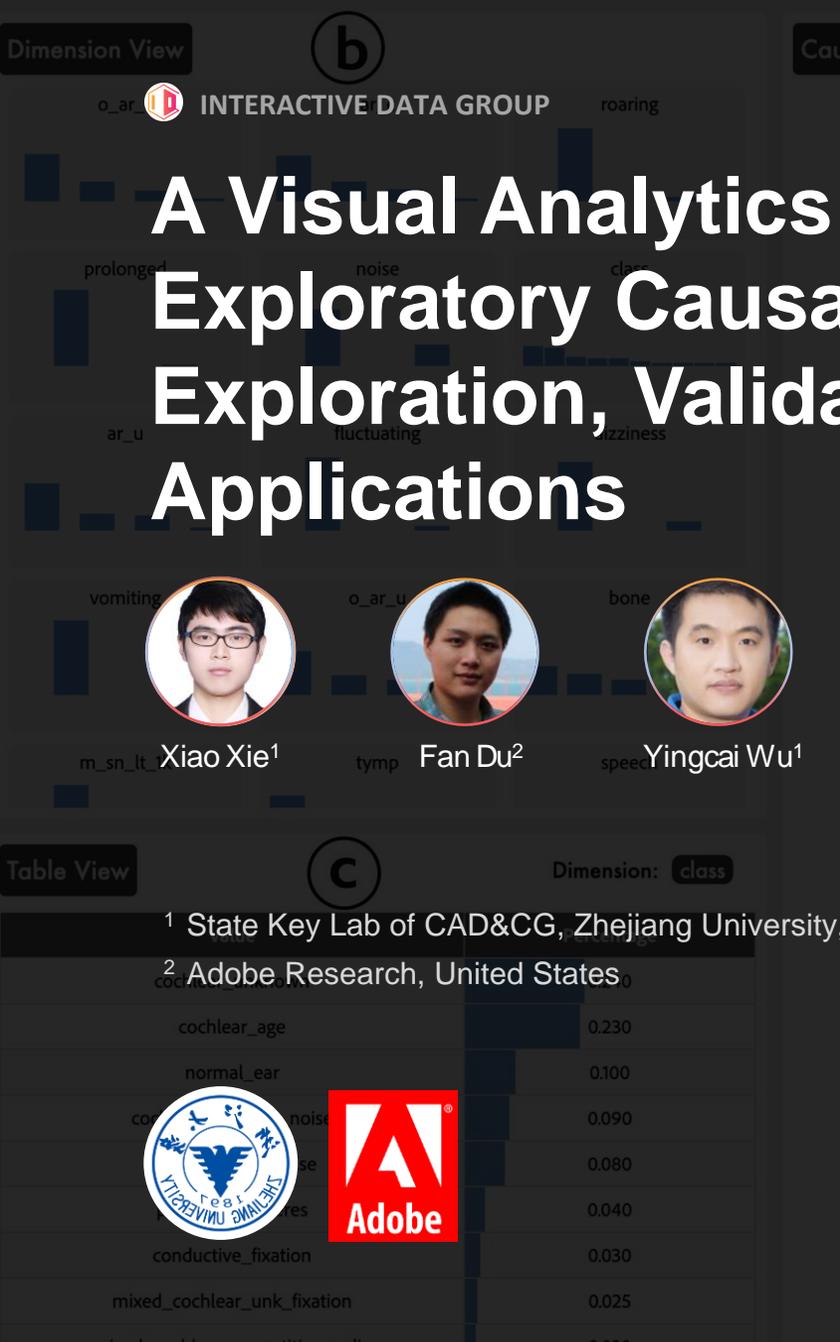
Fan Du²



Yingcai Wu¹

¹ State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China

² Adobe Research, United States



因果关系

因果分析是数据分析中的常见任务



医疗数据分析



市场营销数据分析



教育数据分析

从应用场景出发构思论文

从与用户的讨论中发现领域面临的痛点

缺少因果关系，难以判断每类平台的“准确”的影响

高度相关 同时受到影响



分析师

提高 →



订阅

← 促进

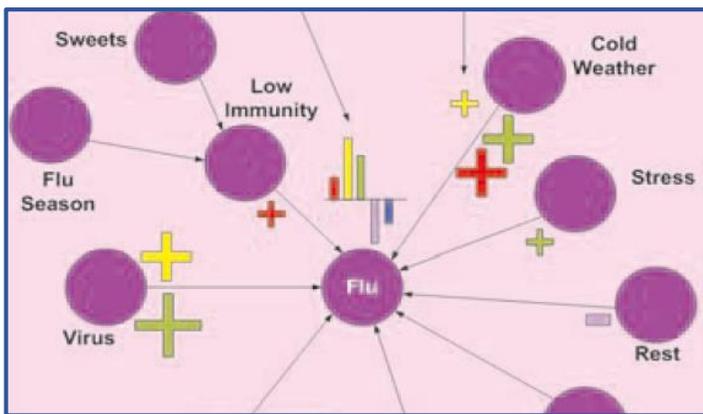


社交平台

因果可视化 (1/2)

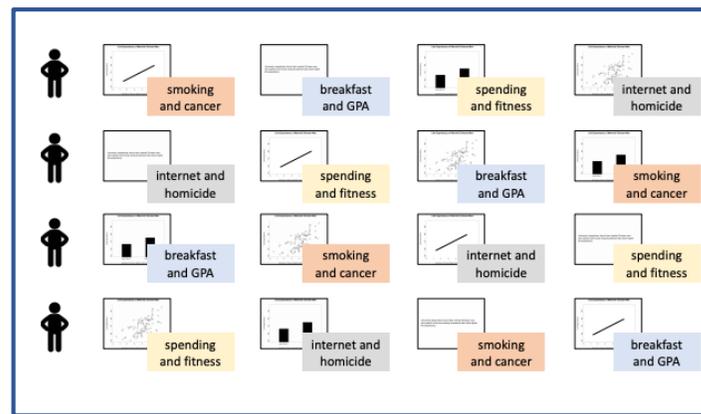
关于因果可视化的用户实验研究

提供了丰富的设计准则



展示因果的静态和动态的视觉通道

N. R. Kadaba et.al., 2007



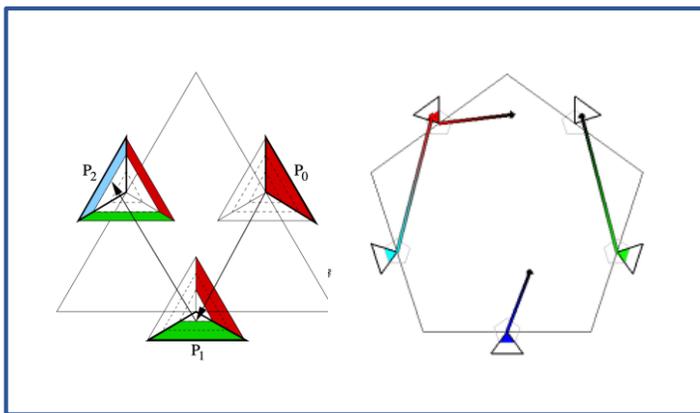
不同可视化表达的不同程度的因果

C. Xiong et.al., 2020

因果可视化 (2/2)

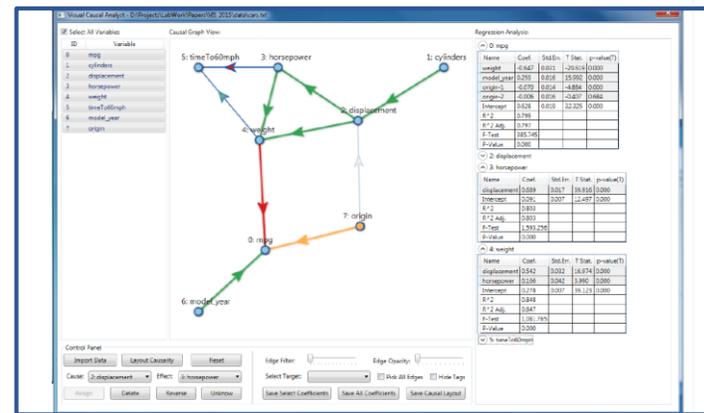
面向因果的可视化分析系统
根据调研结果将论文想法“落地”

专注于小数据集的探索



Animated growing polygons

N. Elmquist and P. Tsigas, 2003



The Visual Causality Analyst

J. Wang and K. Mueller, 2015

Causality Explorer

Dimension View

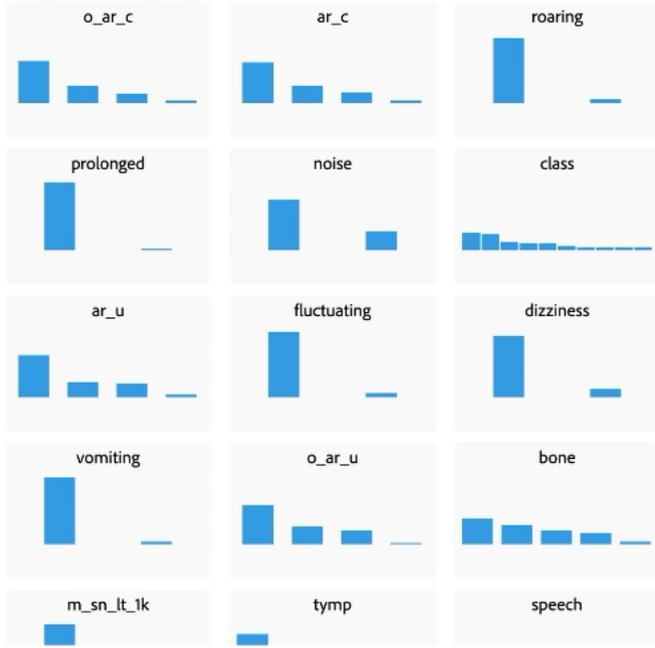
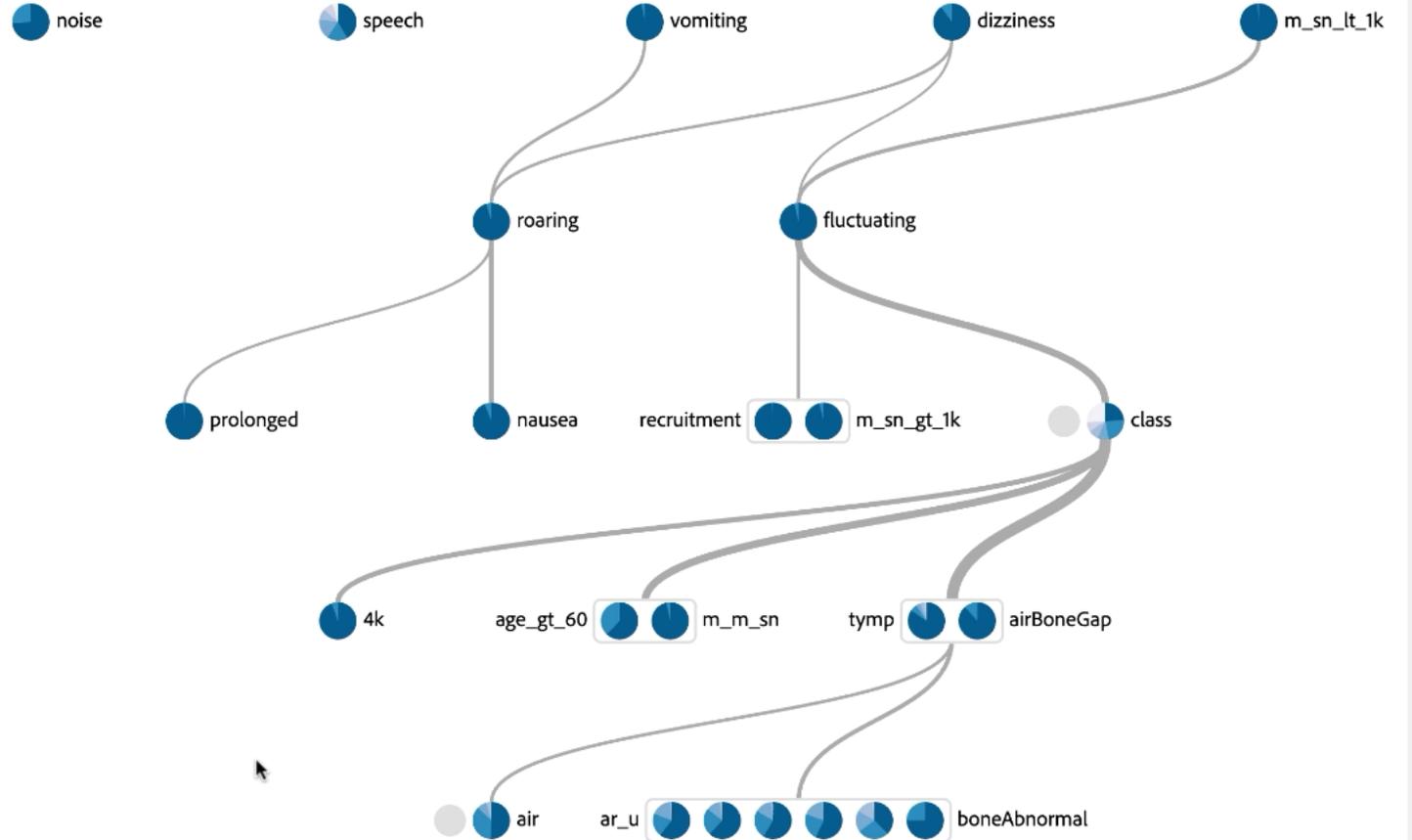


Table View

Dimension: **speech**

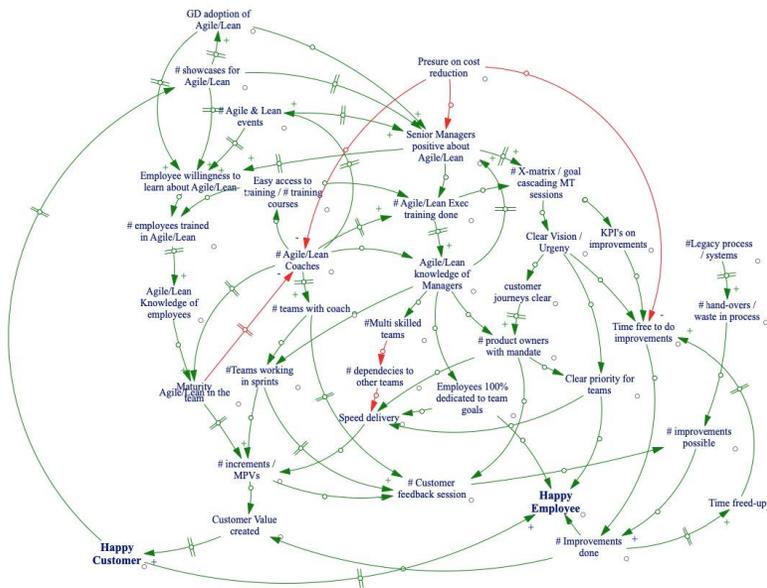
Causal Graph View

Filter:

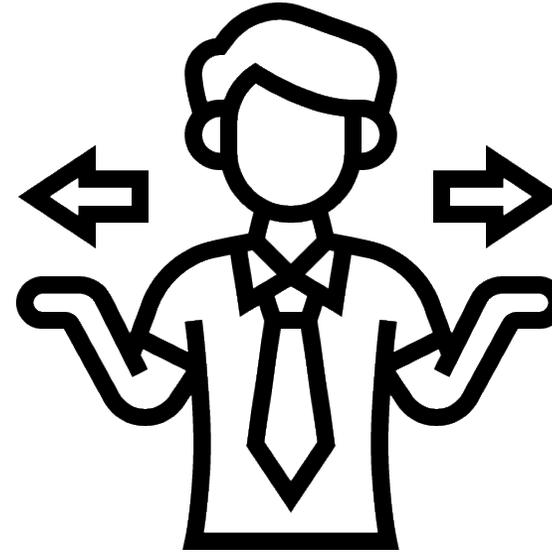


挑战

- 一个包含大量不确定因果关系的因果图的可视化表达



大量的因果关系



因果关系的不确定性

挑战

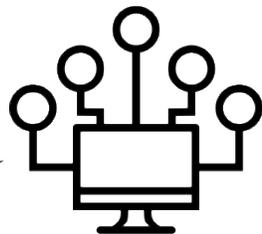
- 一个包含大量不确定因果关系的因果图的可视化表达
- 一个支持因果模拟推演的交互式可视分析系统的实现

从实践出发，根据迭代过程中遇到的问题进行总结



What's the result of Y if ...

The value of Y will reduce to ...



缺乏交互式工具



探索分析

数据格式

3种类别型数据

观测值



个人记录



动作

目标



治愈率

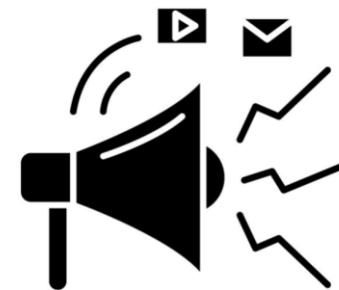


消费

干预



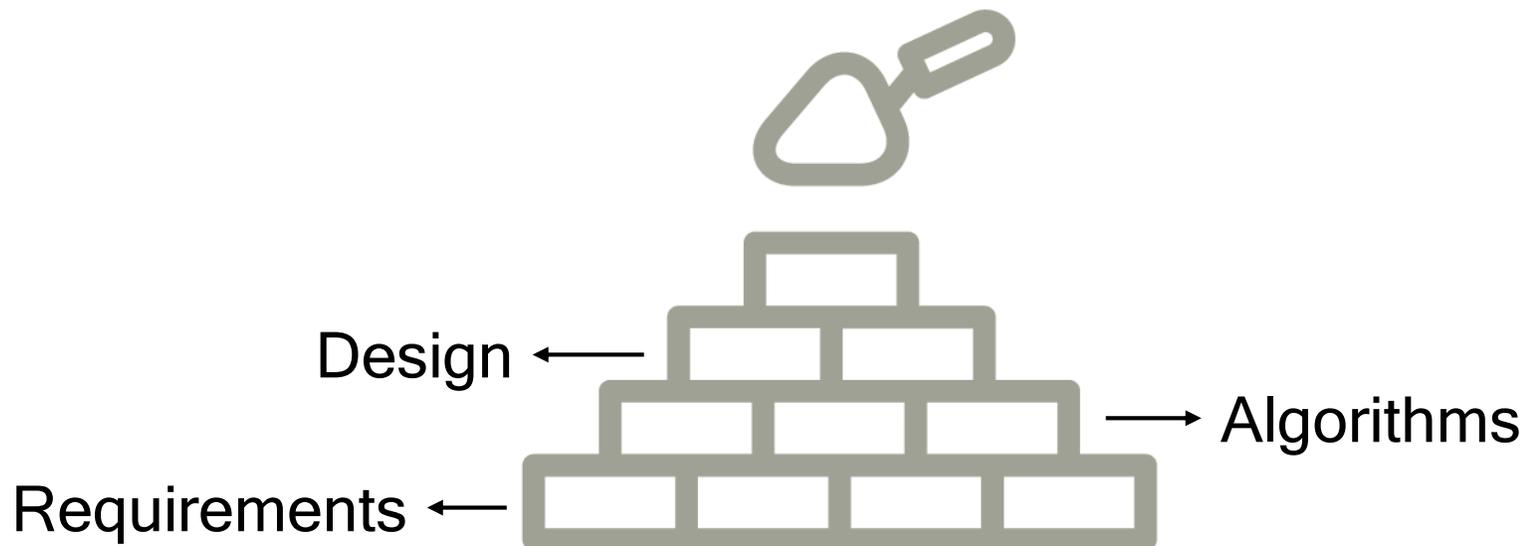
药物



促销

从专家访谈中发现需求

- 邀请了5位专家
 - 3位市场营销分析师, 2位因果分析专家
- 开展了2轮半结构化的访谈
- 总结了3大类的需求和7项详细的任务



投入的精力



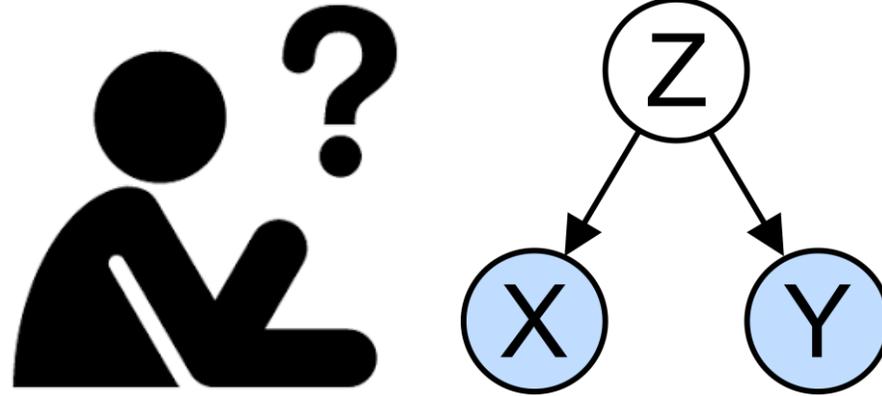
需求分析

- 支持因果关系的感知和检查(R1)

展示大因果图

显示因果关系的
不确定性

提高可靠性



需求分析

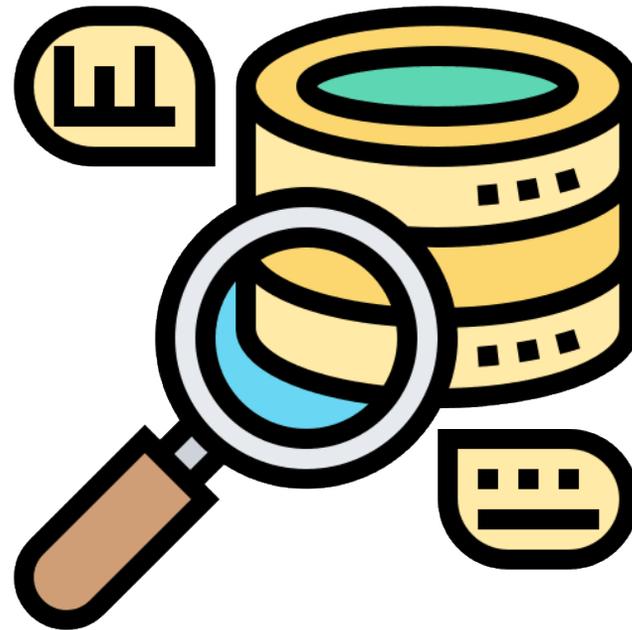
- 支持因果关系的感知和检查(R1)

展示大因果图

显示因果关系的
不确定性

检查数据质量

校验数据假设



需求分析

- 支持因果关系的感知和检查(R1)
- 支持发现具有影响力的因素(R2)

估计不同因素的
影响力

发现重要因素



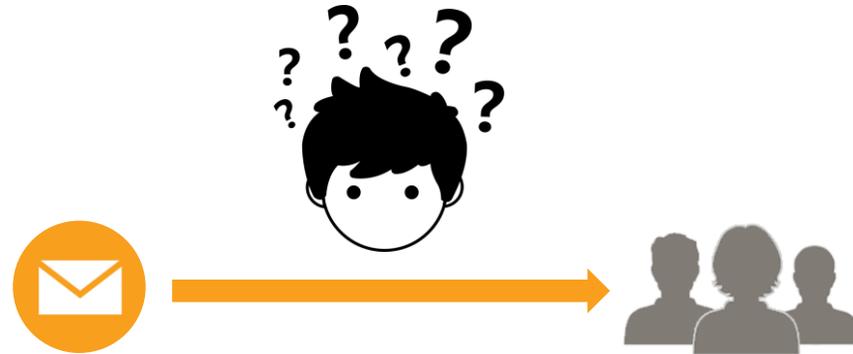
需求分析

- 支持因果关系的感知和检查(R1)
- 支持发现具有影响力的因素(R2)

估计不同因素的
影响力

可视化影响因素
和因果关系

理解影响的机制



需求分析

- 支持因果关系的感知和检查(R1)
- 支持发现具有影响力的因素(R2)
- 支持完成基于因果的模拟推演(R3)

根据不同方案
完成模拟

预测未来态势



需求分析

- 支持因果关系的感知和检查(R1)
- 支持发现具有影响力的因素(R2)
- 支持完成基于因果的模拟推演(R3)

根据不同方案
完成模拟

展示不同方案造
成的影响

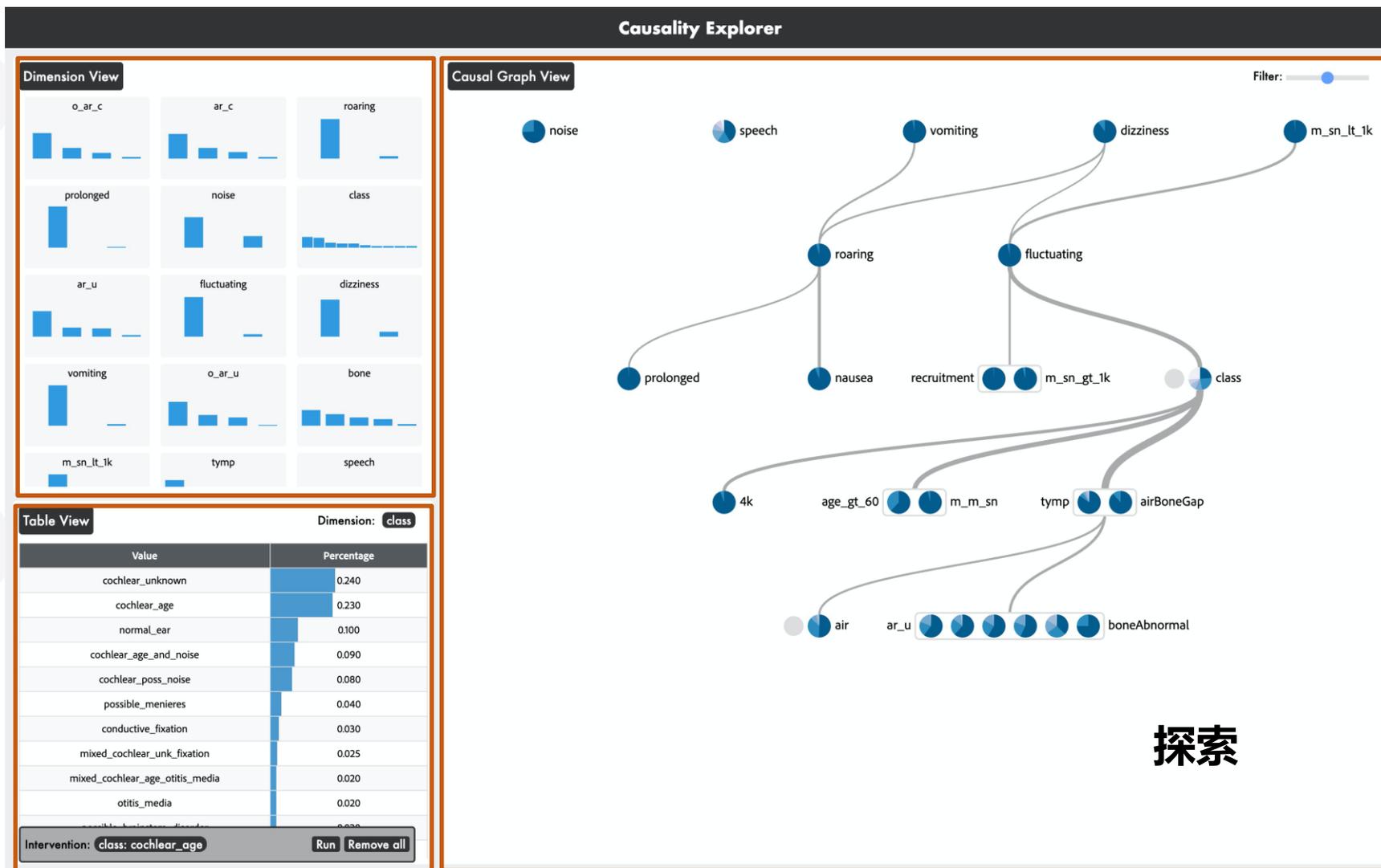
避免不必要的副作用



系统设计

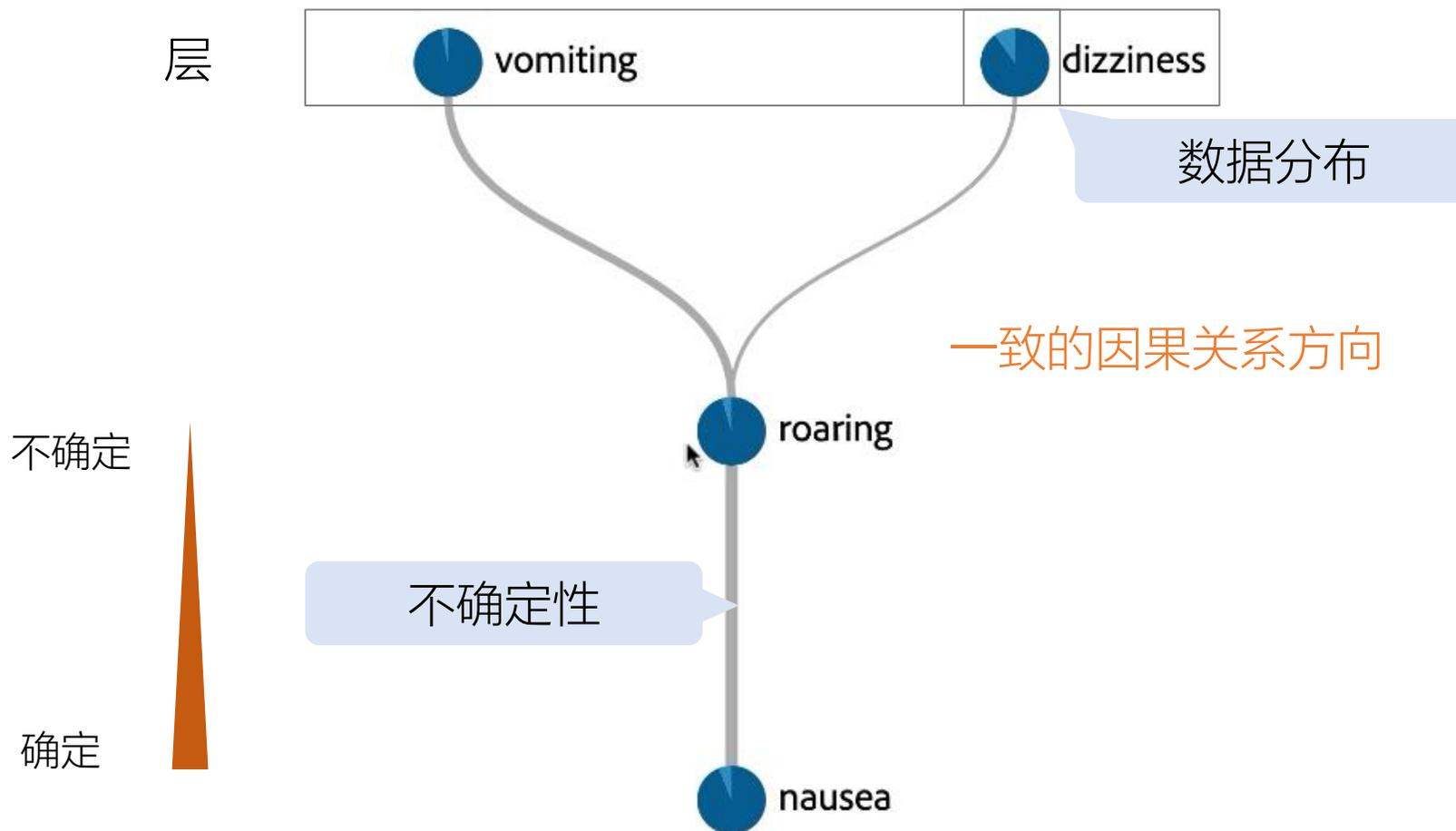
验证

应用



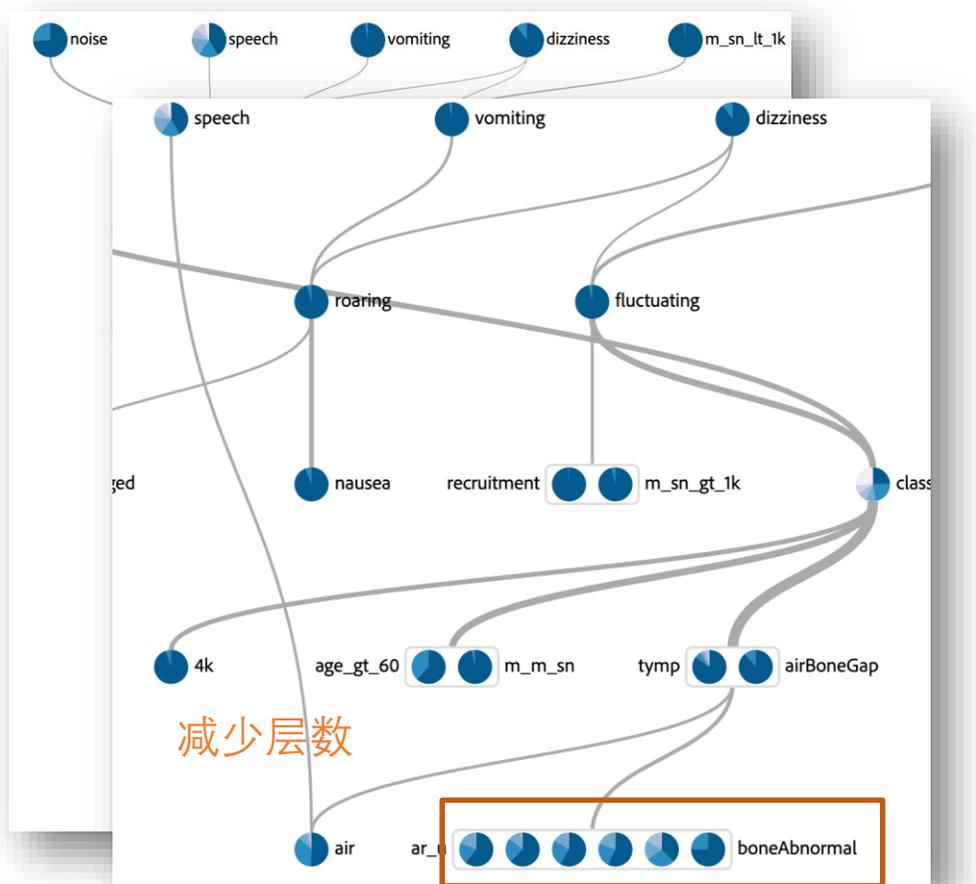
因果图的设计

层次化的布局展示因果关系



布局策略

加强可扩展性以展示更多的因果关系

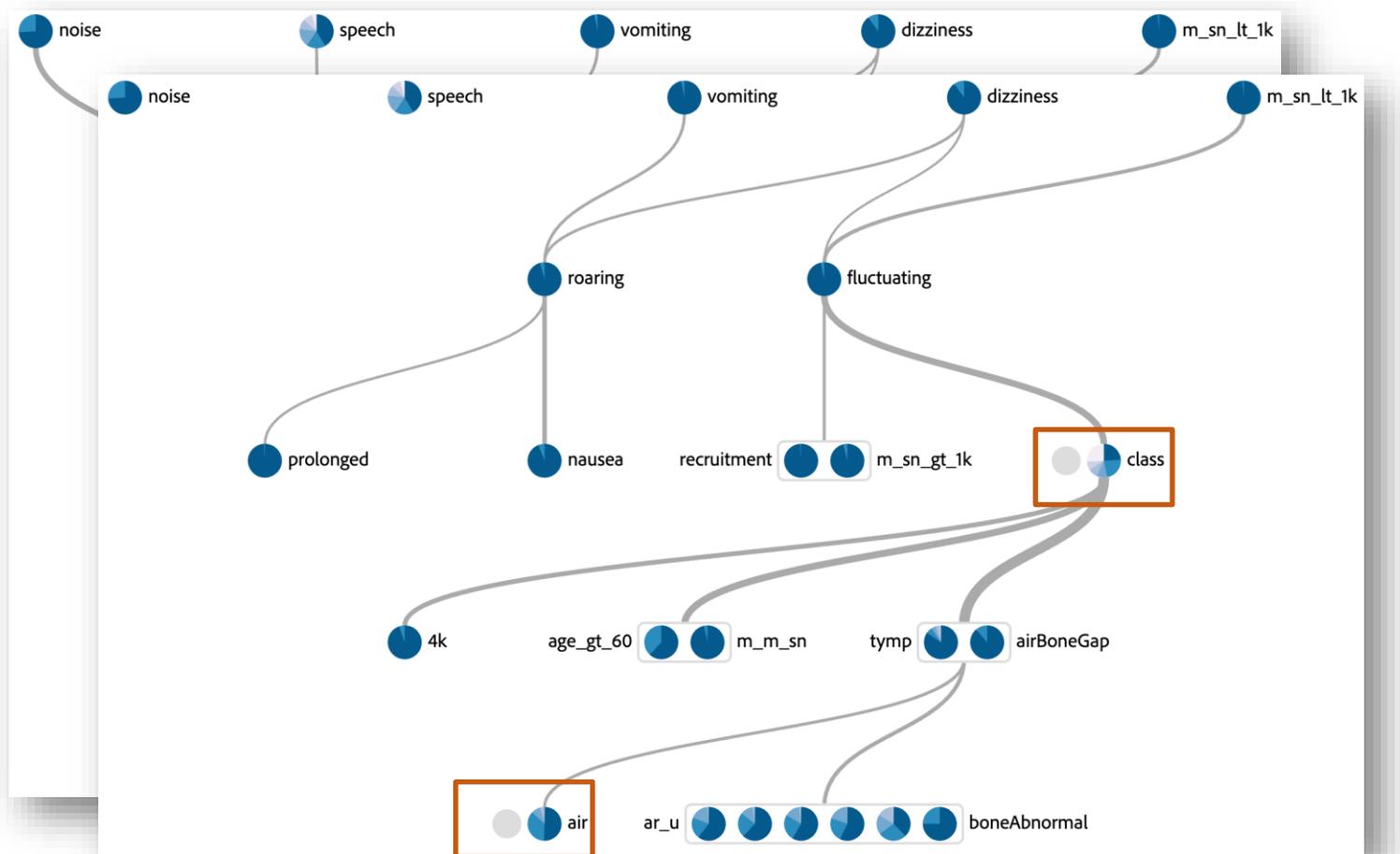


点的聚合

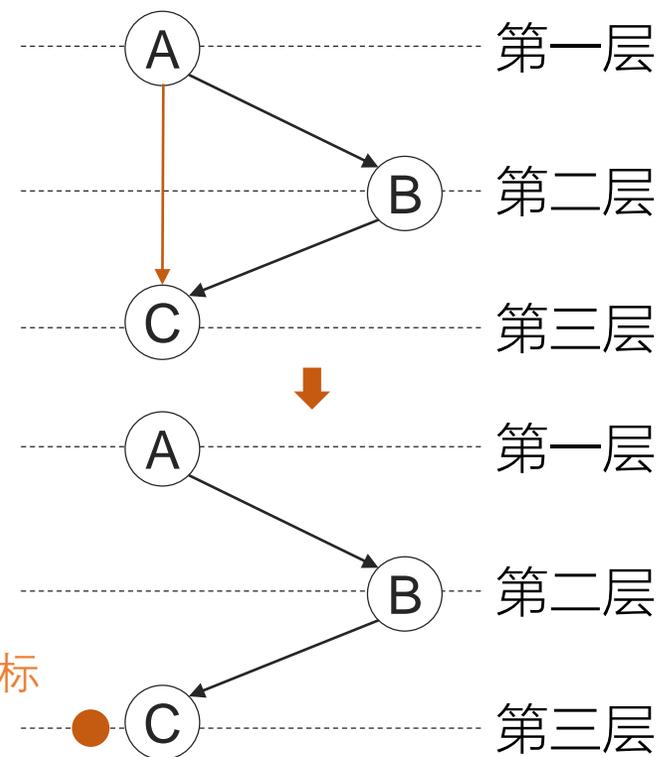


布局策略

加强可扩展性以展示更多的因果关系



边的消减



模拟推演的交互

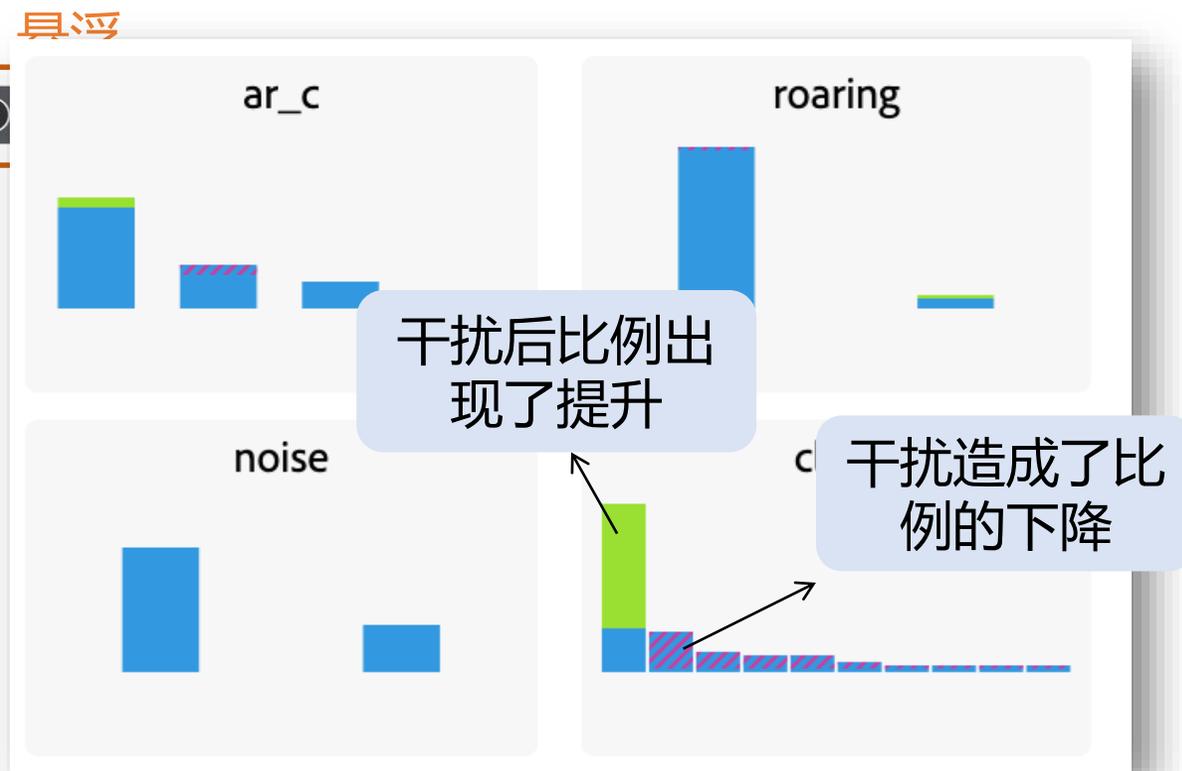
Table View

Dimension: class

Value	Percentage
cochlear_unknown	0.240
cochlear_age	0.230
normal_ear	0.100
cochlear_age_and_noise	0.090
cochlear_poss_noise	0.080
possible_menieres	0.040
conductive_fixation	0.030
mixed_cochlear_unk_fixation	0.025
mixed_cochlear_age_otitis_media	0.020
otitis_media	0.020
possible_humintan_disorder	0.020

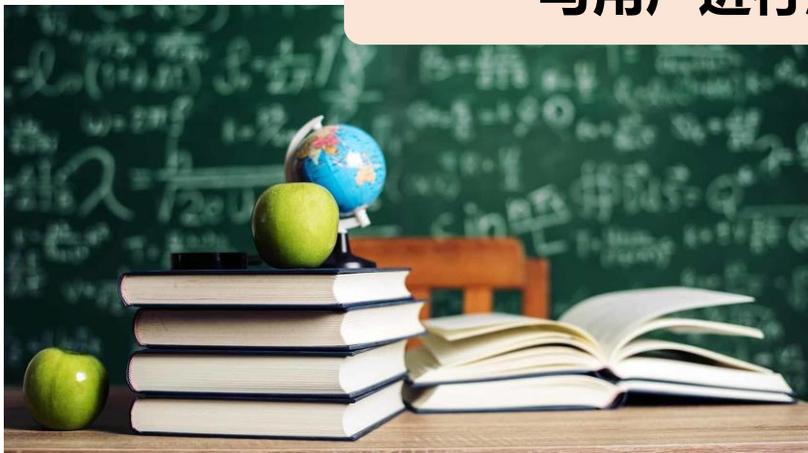
Intervention: class: cochlear_unknown

Run Remove all



案例分析

与用户进行深入的讨论，收集丰富的反馈



Education

3500 students from a collage

8 dimensions

Reduce the dropout rate



Marketing

10000 visit samples of an online retail store

32 dimensions

Increase the purchase

用户反馈

优势

- 用户通过因果图的布局可以清晰地感知因果方向
- 不确定性促进了用户对因果关系的探索和验证
- 系统提供了灵活的交互功能

建议

- 系统中可以加入更多对数据属性的聚合、分离等交互
- 系统可以提供模拟推演的历史记录促进方案的选择

A Visual Analytics Approach for Exploratory Causal Analysis: Exploration, Validation, and Applications



Xiao Xie¹



Fan Du²



Yingcai Wu¹

¹ State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China

² Adobe Research, United States



Q&A

Xiao Xie

xxie@zju.edu.cn