

数据分析的基本概念

灵魂拷问：什么是数据分析

广州和深圳哪个发展潜力更大？

经验感觉：深圳更大

数据分析：广州和深圳的GDP增长分别是多少？
创业公司数量增长对比如何？
人才流动情况怎样？
消费水平对比如何？



产品销售的重要影响因素有哪些？

经验感觉：季节因素、产品质量、广告投放

数据分析：产品销售随时间流逝如何波动？（识别）
广告投放带来的预期收益如何？（预测）

简单来说，数据分析就是利用数据来理性思考和决策的过程。

数据挖掘 (Data Mining)：是一个跨学科的计算机科学分支，它是用人工智能、机器学习、统计学和数据库的交叉方法在**相对较大型的数据集中发现模式**的计算过程。



EDIT数字化模型：



- 探索 → 指标体系
- 诊断 → 性质分析法，数量分析法
- 指导 → 知识库，策略库，流程模板
- 工具 → 数据模型，算法模型，优化模型

1. 业务运行探索（E）

探索关注企业各项业务的**运行状态**、各项指标是否合规以及各项业务的**具体数据**情况等等。

2. 问题根原因诊断（D）

当业务指标偏离正常值时，采用定性和定量相结合的方式，在中、微观层面**定位和分析问题**。

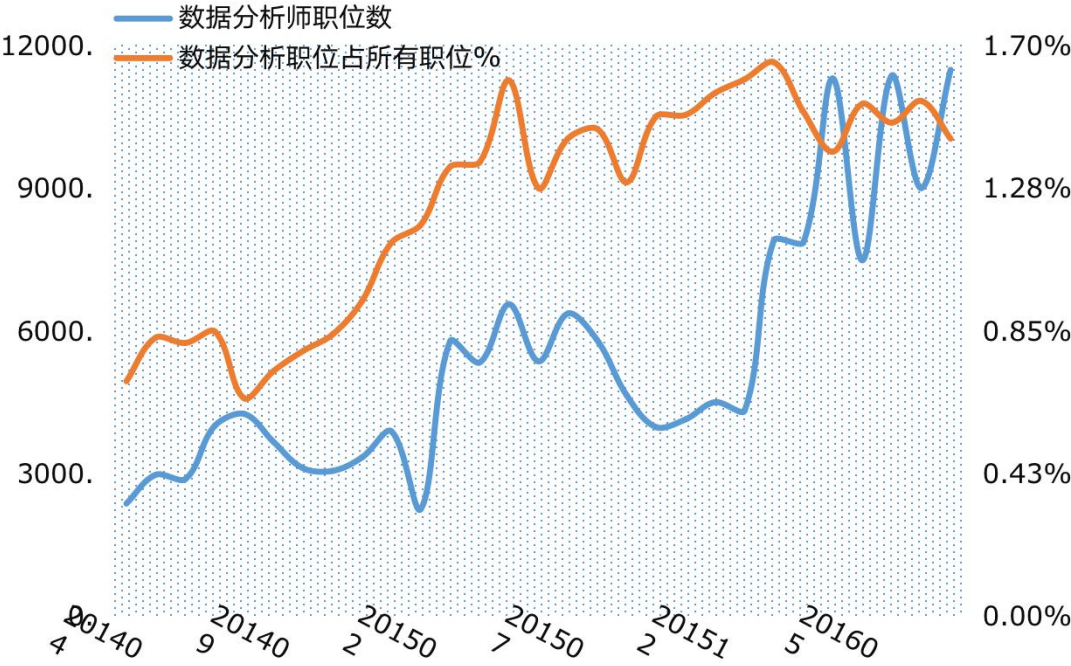
3. 业务策略指导（I）

在确定获客群体后，我们就需要考虑下如何制定业务策略指导。根据知识库、策略库、流程模版可以便捷地制定**优化策略**。

4. 智能算法工具（T）

算法架构是EDIT数字化模型的**发动机**，服务于价值经营、客群运营、全面质量管理、全面风险管理等各个方面。

这是数据分析



数据来源：猎聘网

这也是数据分析

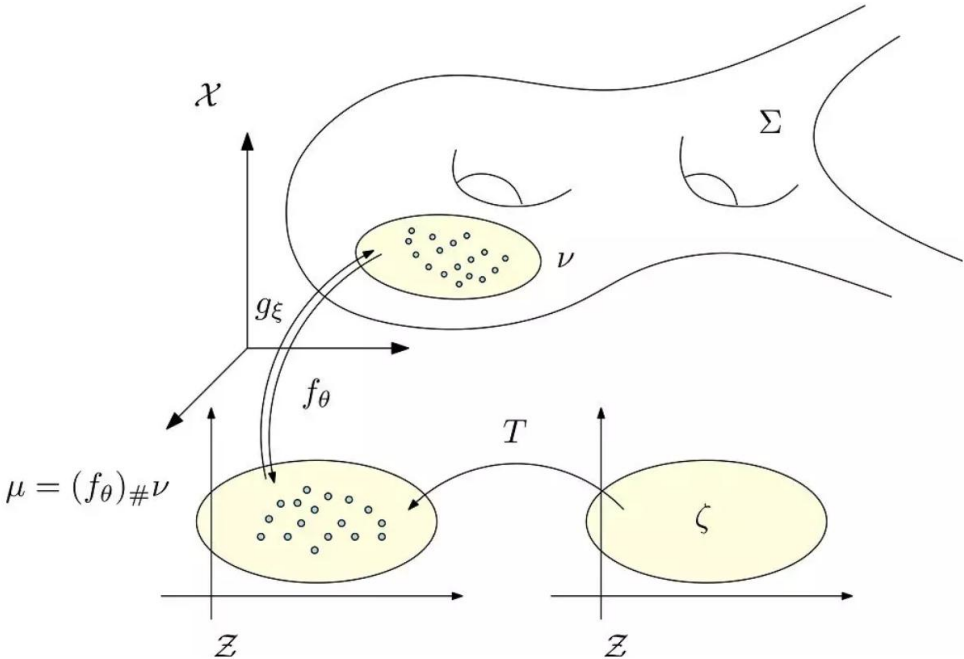


Fig 6. AE-OMT 半透明生成模型。

图例来源：《计算共形几何》

“谷歌流感趋势”把大数据推上风口浪尖

搜索流感信息的人数

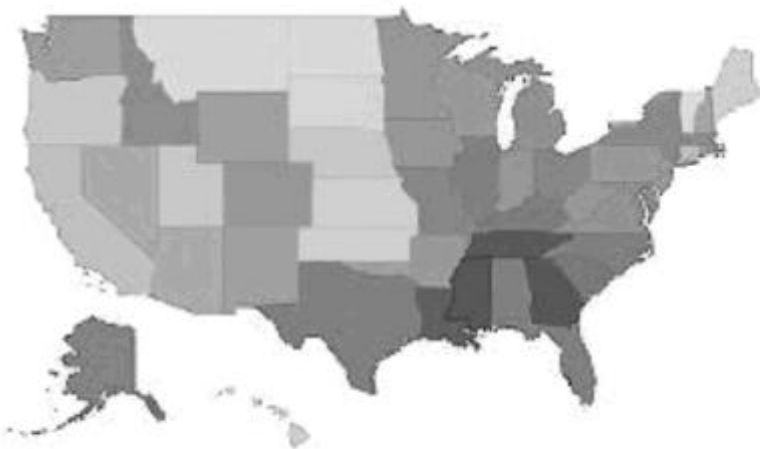
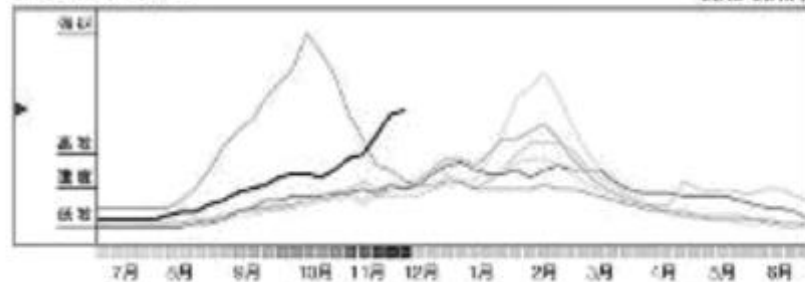
实际患病人数

流感爆发

- 美国疾病控制中心要在流感爆发**两周**后才知道
- 谷歌的大数据预测只需要**一天**

全美感冒趋势

2012-2013年



欢迎同学们来到数据分析的精彩世界!