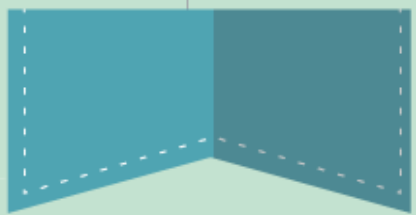
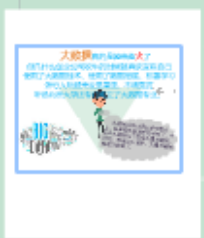
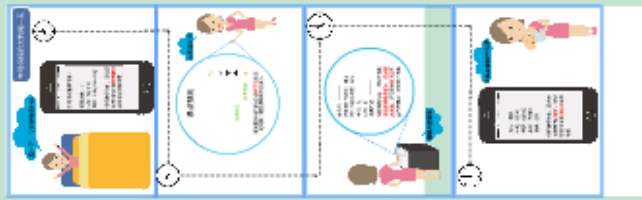
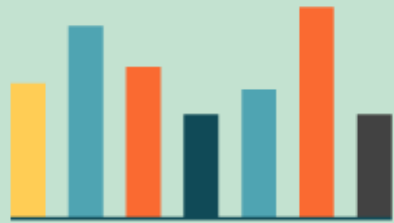


什么是大数据



2015年
大数据的定义

大数据的定义



大数据发展现状



谢谢观看

什么？你说太敷衍？
那下面来点干货好了！

大数据的定义



大数据 (big data)



是指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合，是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力来适应海量、高增长率和多样化的信息资产。



6

值日
周一
周二
周三
周四
王强
周五



Variety (多样)

Value (价值)

Velocity (高速)

Veracity (真实性)

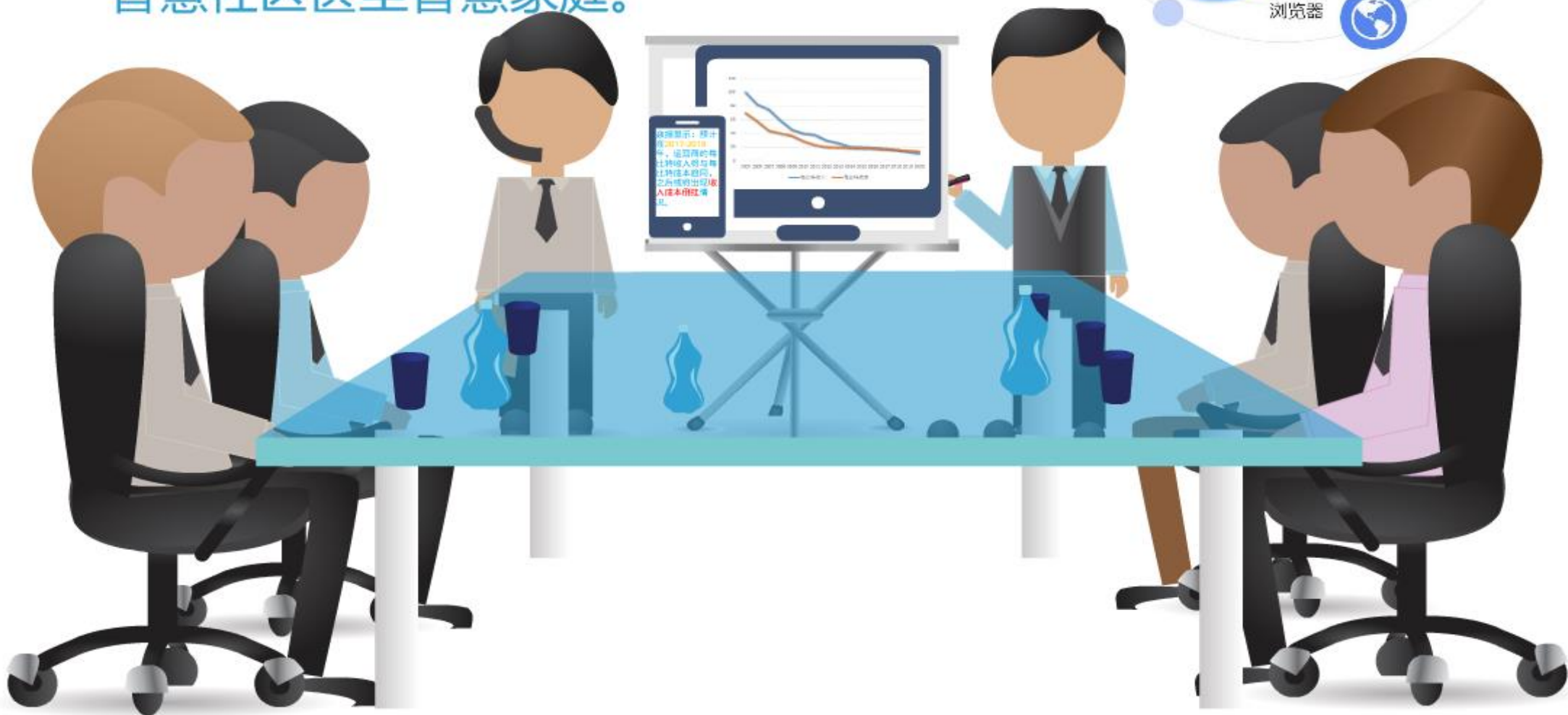
Volume (大量)

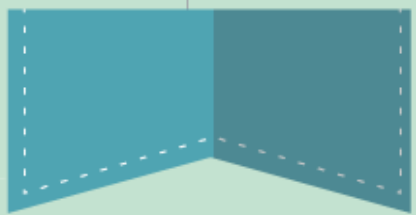
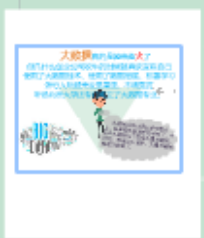
Variability (可变性)



人机交互

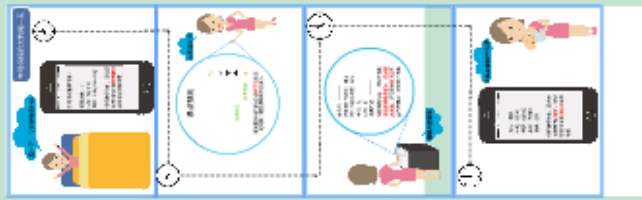
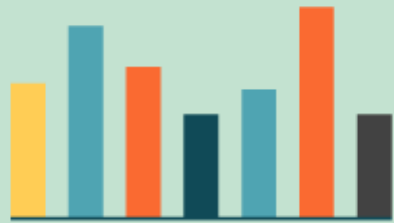
宽带和移动互联网在创造了一个用户可以
获得几乎任何所需的交互世界后，**大数据**
也能够充分应用于智慧政务、智慧医疗、智
慧交通、智慧城管、智慧旅游、智慧教育、
智慧社区甚至智慧家庭。





大数据的定义

大数据的定义



大数据发展现状

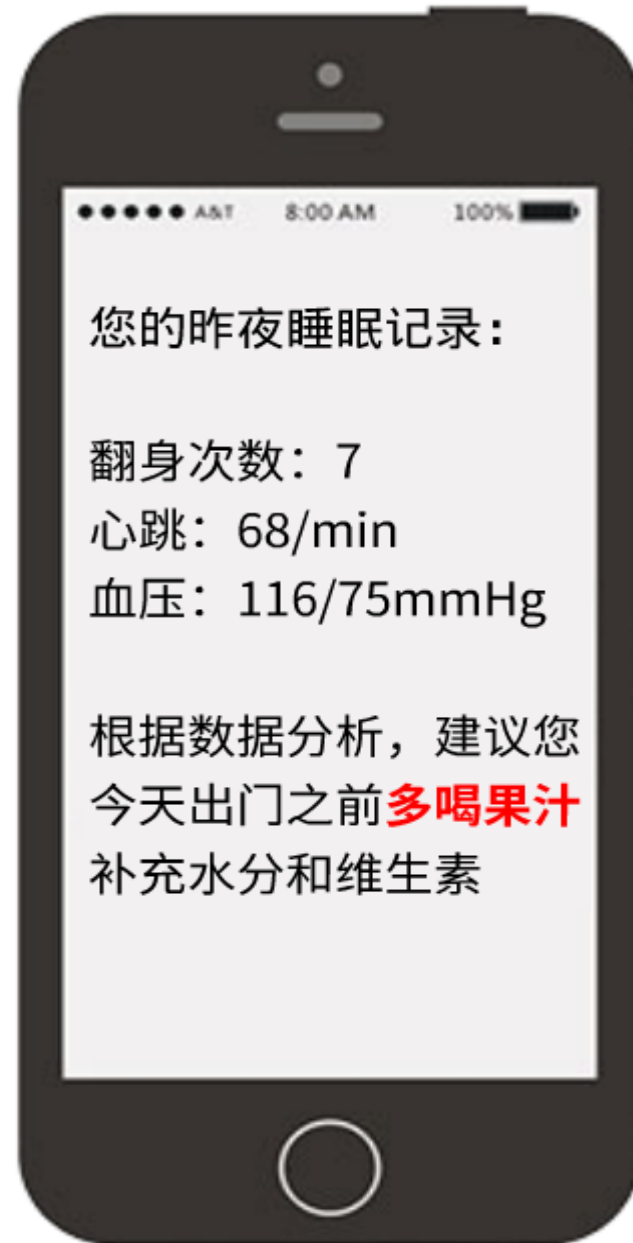


谢谢观看

通过实例来认识大数据

年轻妈妈的大数据一天

看一下，今天早餐吃什么





景点预测

景点	人数	交通	天气
栖霞寺	一般	畅通	
夫子庙	非常拥挤	严重堵塞	
玄武湖	舒适	堵塞	
梅花山	非常舒适	一般	
中山陵	拥挤	非常畅通	

根据数据分析建议您去栖霞寺游玩
避免路上遇到拥堵或不良天气

去哪里玩呢





请输入会员号

会员号：*****

您目前所在城市：南京

您所在城市湿度：42%

性别：女

年龄：27

消费历史：*****

根据数据分析，建议您购买
**烫后护卷型洗发水、滋润型
沐浴露**，您家里的**尿不湿**可能已经用完，建议及时购买





大晚上宝宝为什么哭





数据库（池塘捕鱼）

VS

大数据（大海捕鱼）



小（以MB为处理单位）

数据规模

大（以GB、TB、PB为处理单位）

单一（结构化为主）

数据类型

繁多（结构化、半结构化、非结构化）

先有模式后有数据
（先有池塘后有鱼）

模式和
数据关系

先有数据后有模式
模式随着数据增多不断演变

数据（池塘中的鱼）

处理对象

“鱼”，通过某些“鱼”判
断其他种类的鱼是否存在

One size fits all

处理工具

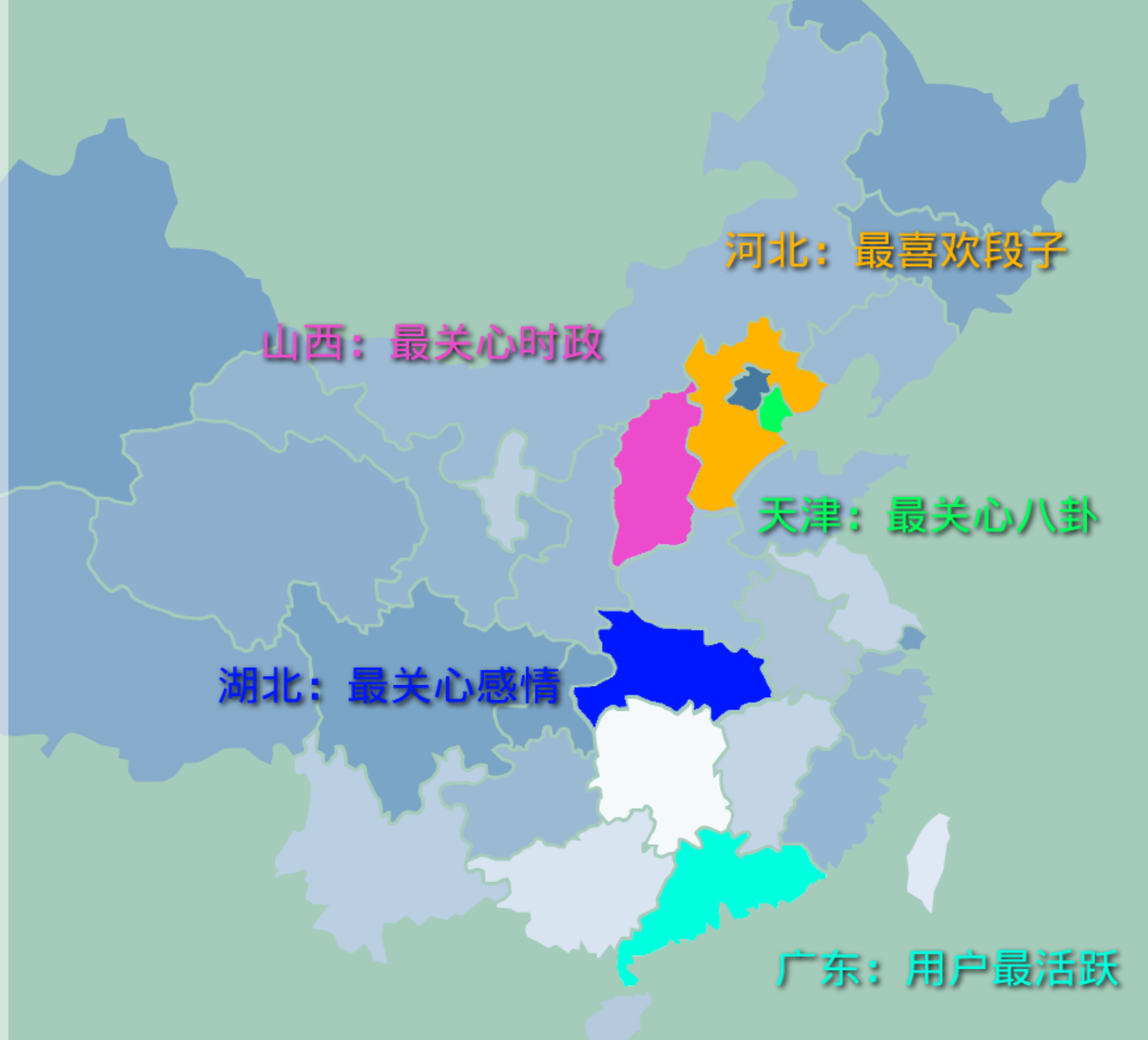
No size fits all



今日头条

目前超过**2.2亿**用户，每天有**2000万**用户使用今日头条阅读自己感兴趣的文章。

基于数据挖掘，从宏观的角度得到大众群体的阅读趋势，基于用户喜好进行文章推送。



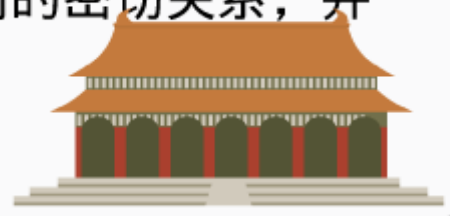


北京市旅游委 百度预测



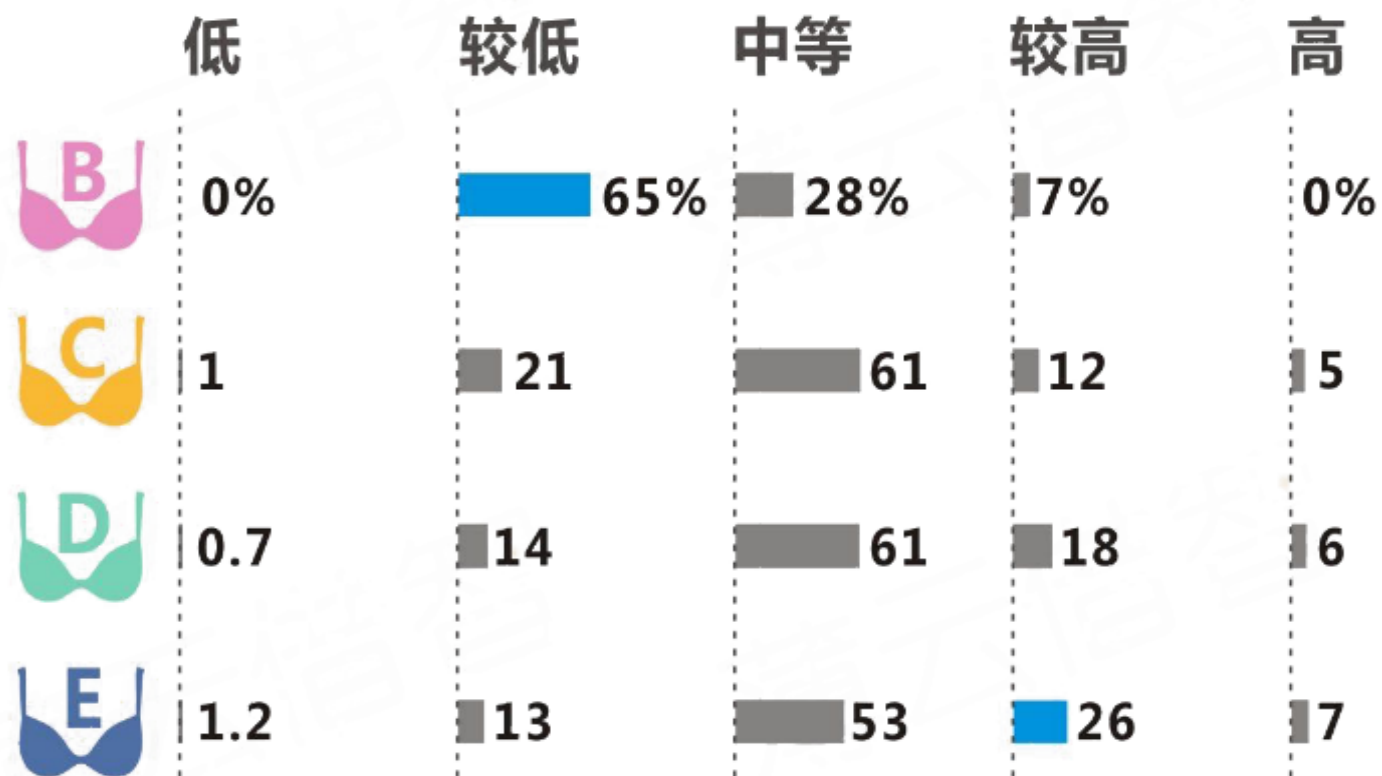
百度大数据部故宫旅游人数预测值和北京市旅游委提供的故宫当日实际人数的对比

百度从每天100亿搜索请求中，抽取旅游相关数据进行智能分析，同时利用百度后台每天数十亿次LBS数据进行计算，发现旅游相关词搜索数量和世界旅游人数之间的密切关系，并以此建立了旅游预测模型。



准确度达90%

罩杯大小与在线购买能力



coffee break

阿里巴巴分析师将内衣消费者按其购买能力分成了四类，发现女生胸越大越“败家”，65%的B罩杯女性都属于“购买能力低”这一类，而C及以上罩杯的女性大部分购买能力都是“中等”或以上！



Baidu 百度

阿里巴巴·中国
alibaba.com.cn

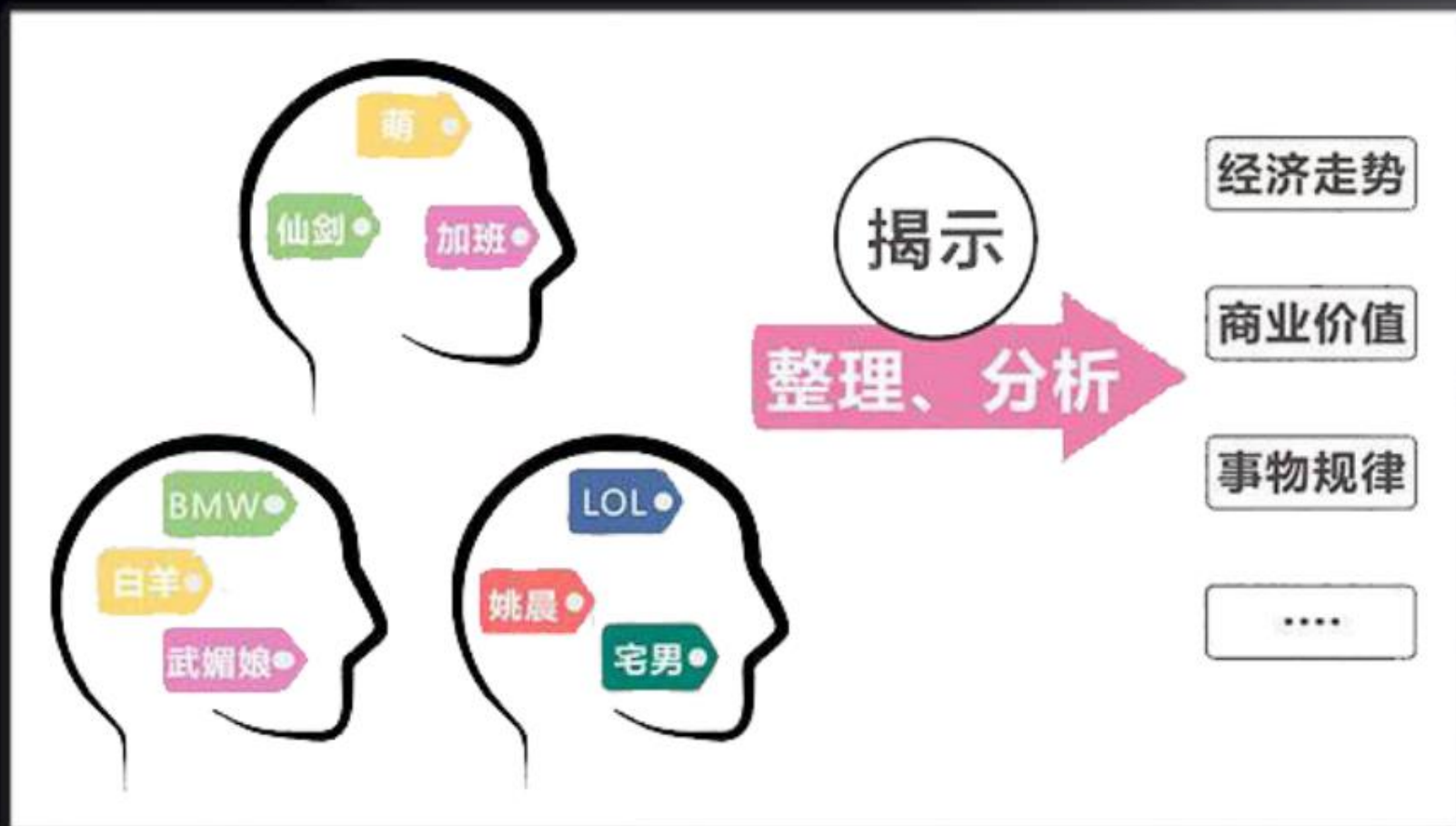
Tencent 腾讯

搜索引擎连接了信息与读者

购物平台连接了商品与消费者

社交网络连接了人与人

大数据则连接一切。拥有了数据就拥有了和万物、时间、空间的连接



大数据将具有价值的标签经过整理和分析，再结合其他的数据，将会揭示事物内部和外部的相关性，为个人、商家、社会带来巨大价值



好了，今天的课程就到这里，你对大数据了解了么！