

如何打造智能分析产品？

- 供应链需求预测产品的实践启示

陈玉君

华为 产品数据科学家

SPEAKER INTRODUCE

陈玉君 华为 产品数据科学家

- 华为质量与流程IT部门；
- 法国国家信息与自动化研究院实习并留任工程师；直读博士毕业于清华大学计算机系；
- VectorCG 时尚推荐产品、有利网金融产品、风险建模；
- 道法自然，热爱生活，迎接挑战



SPEAKER
ArchSummit 2017 ShenZhen

TABLE OF
CONTENTS 大纲

- **供应链领域的智能化应用**
- 需求预测产品和技术架构
- 智能算法在产品中应用
- 分析可视化在产品中呈现
- 总结和展望

供应链，从一个身边故事开始

物流、信息、资金流、周转率、商业利益最大化

实例：ZARA、AMAZON、GE、华为等

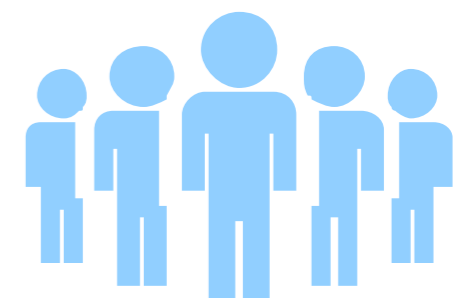


客户

消费者洞察



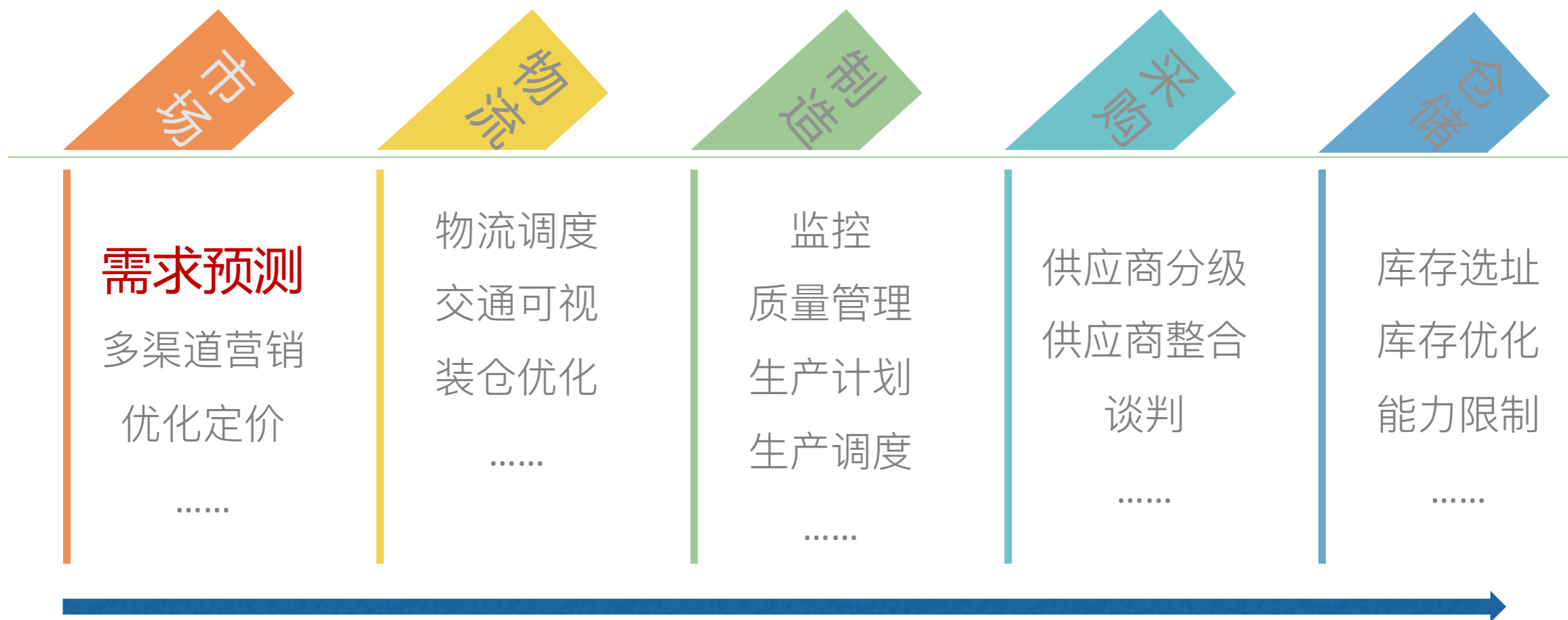
智能研发、生产决策和响应



供应商

供应商协同

数据智能让供应链更简单、更及时
 辅助/自动决策，业务增长黑客、寻找共性
 涉及预测、分类、最优化等多种算法模型



预测客户需求

供应、需求匹配



需求感知预测产品设计的前提 – 数据分析应用

相关性，找模式



提供给用户的预测分析模块的框架



依照这些方法，与业务共同设计辅助预测模块，数据支撑业务，提供洞察

TABLE OF
CONTENTS 大纲

- 供应链领域的智能化应用
- **需求预测产品和技术架构**
- 智能算法在产品中应用
- 分析可视化在产品中呈现
- 总结和展望

需求感知预测产品提供“一站式”需求预测
让预测更简单，更智能，更准确

行业现状痛点

预测业务

业务逻辑不透明
经验和方法传承性弱



目标

预测逻辑&建模

中长期&年度&短期预测逻辑
搭建预测模型加以验证
10 + 种预测模型库辅助

数据平台

统计工具效率低
缺少定制化数据分析
统计要素维护复杂



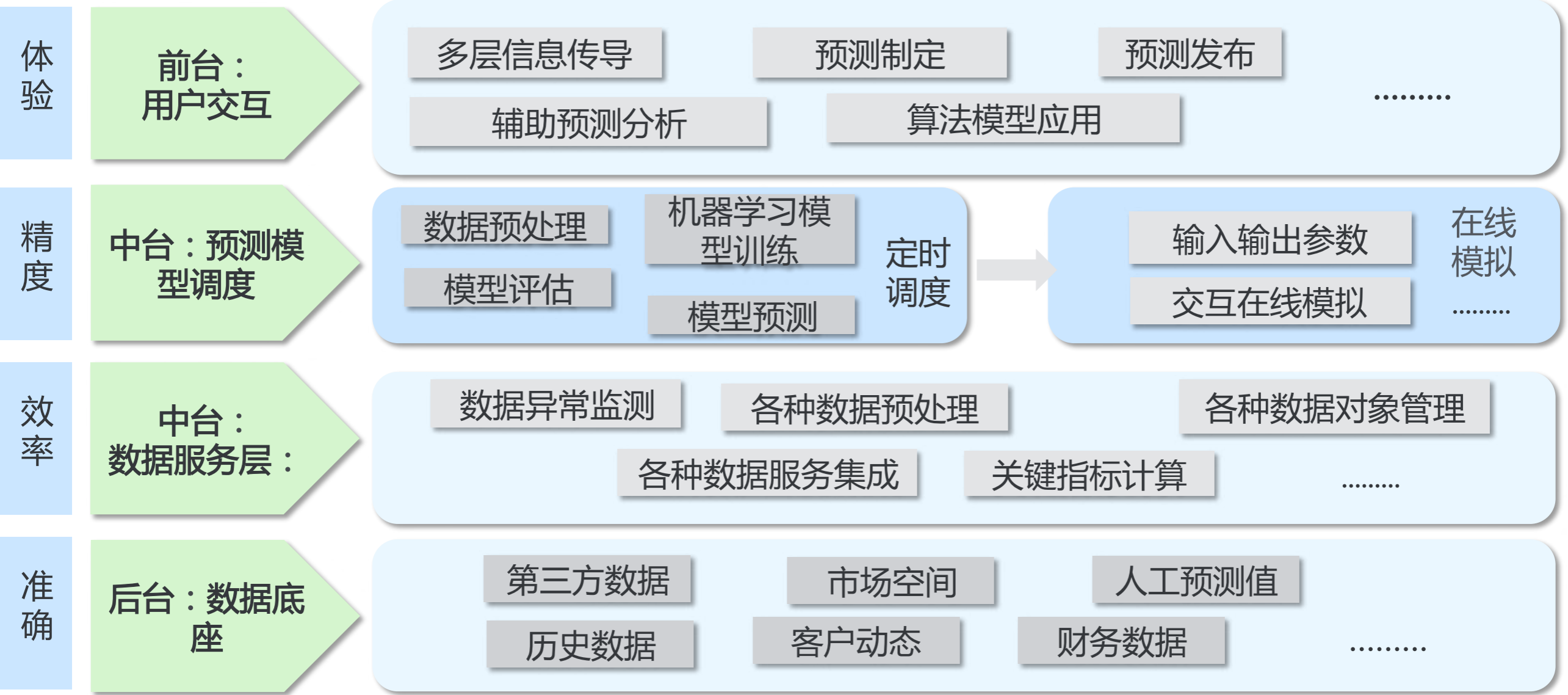
需求感知预测

数据底座承载数据
根据不同产品特征，智能定制数据
多维度、跨产品需求关联分析
KPI和智能预测模型结果分析

价值

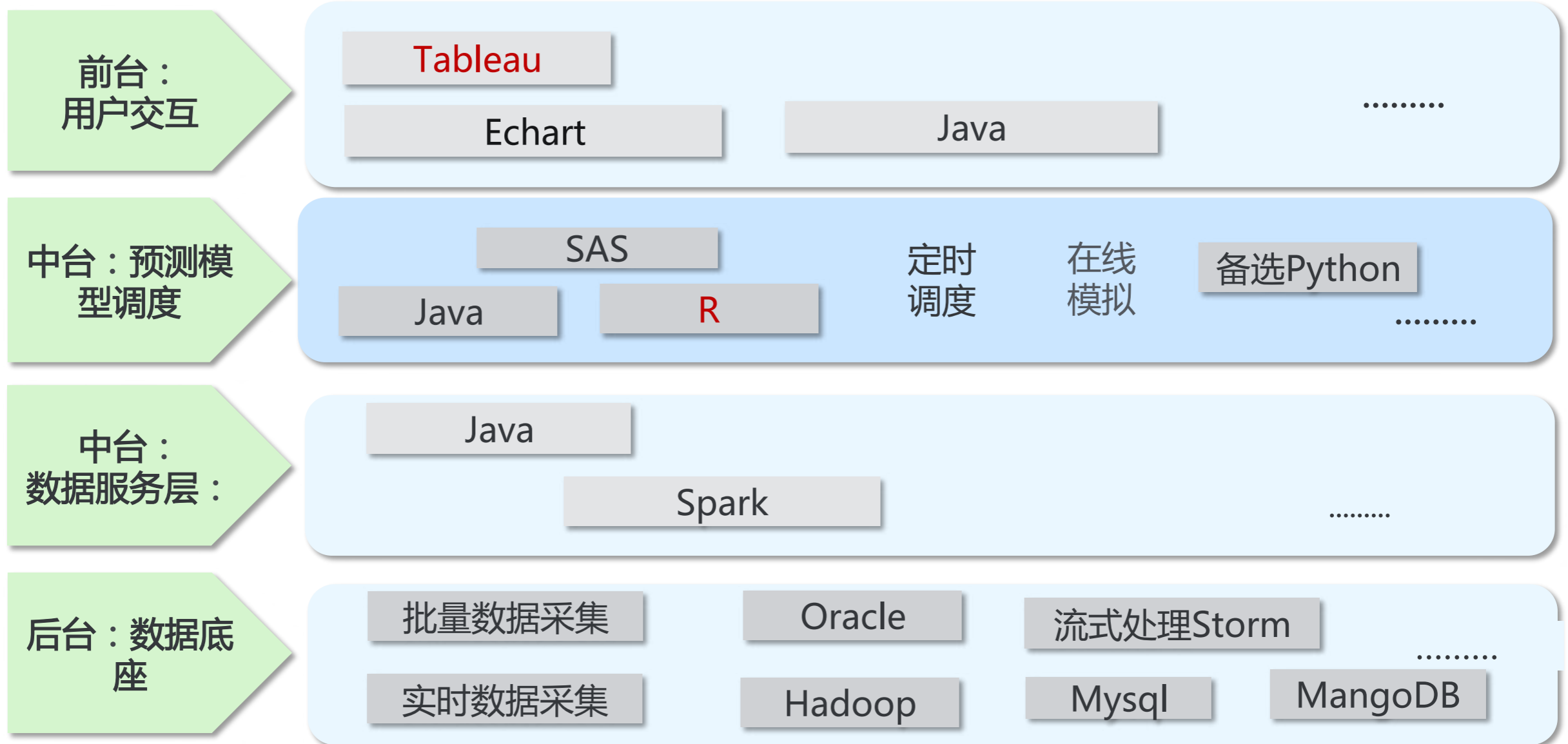
预测准确率提升，增加业务效率

产品架构设计：前、中、后台服务化设计



最优化特征转换

产品技术路线选型和优缺点分析



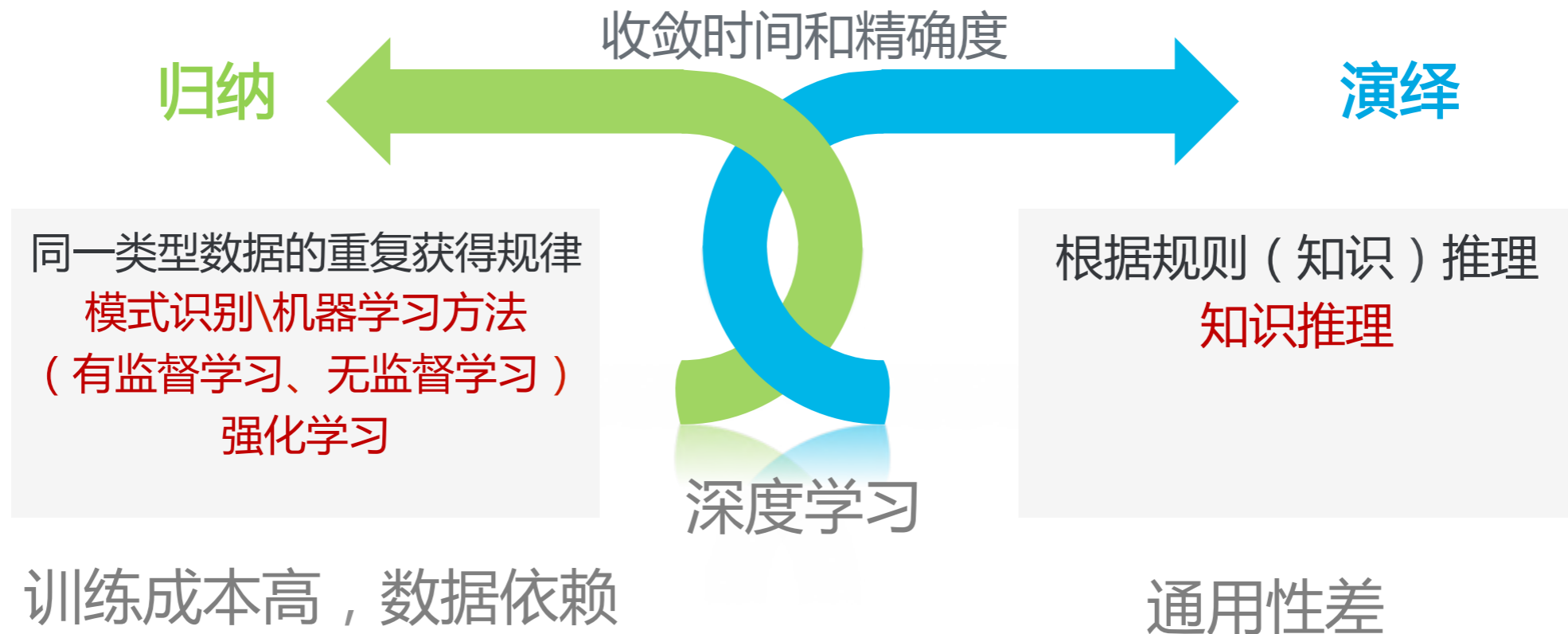
宗旨：敏捷，可扩展，各取所长

TABLE OF
CONTENTS 大纲

- 供应链领域的智能化应用
- 需求预测产品和技术架构
- **智能算法在产品中应用**
- 分析可视化在产品中呈现
- 总结和展望

机器学习的前世今生

用数据的广度，辅助思考的深度，用数据重构对世界的认识；用重复的计算，完成机器对于数据的理解



机器学习的基本原理和术语

$$y = f(x)$$

关系映射

参数估计

泛化

代价函数

收敛

过拟合



监督学习：
有标注的数据让机器学习
规律

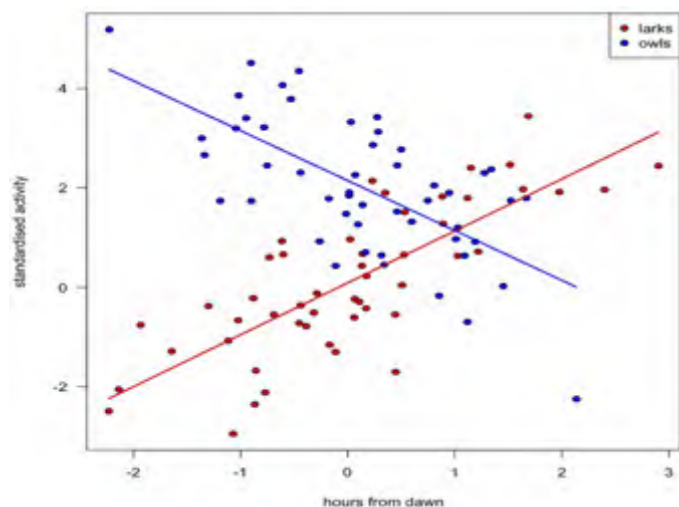


非监督学习：
大量数据让机器自己去学
习规律

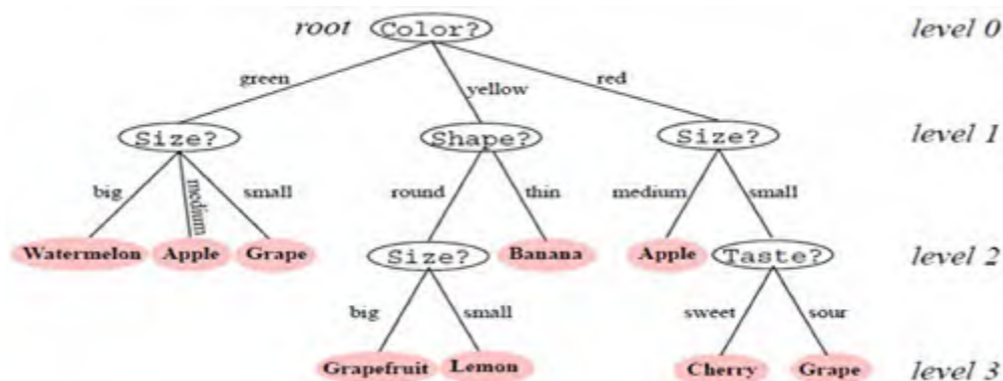


分类回归算法的三类模型

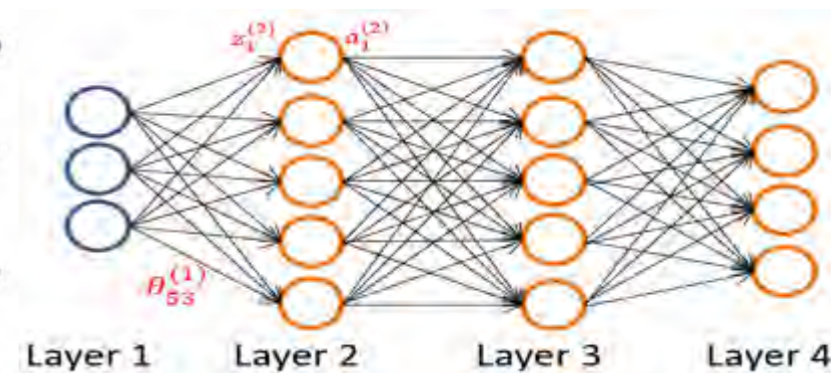
线状模型



树状模型



网状模型



ARIMA、LR、
Linear SVM

高维度
稀疏性
速度快
特征难

RF(Random Forest
DT(Decision Tree)
GBDT(Gradient Boosting
Decision Tree).....

连续性
稠密性
易解释
有限制

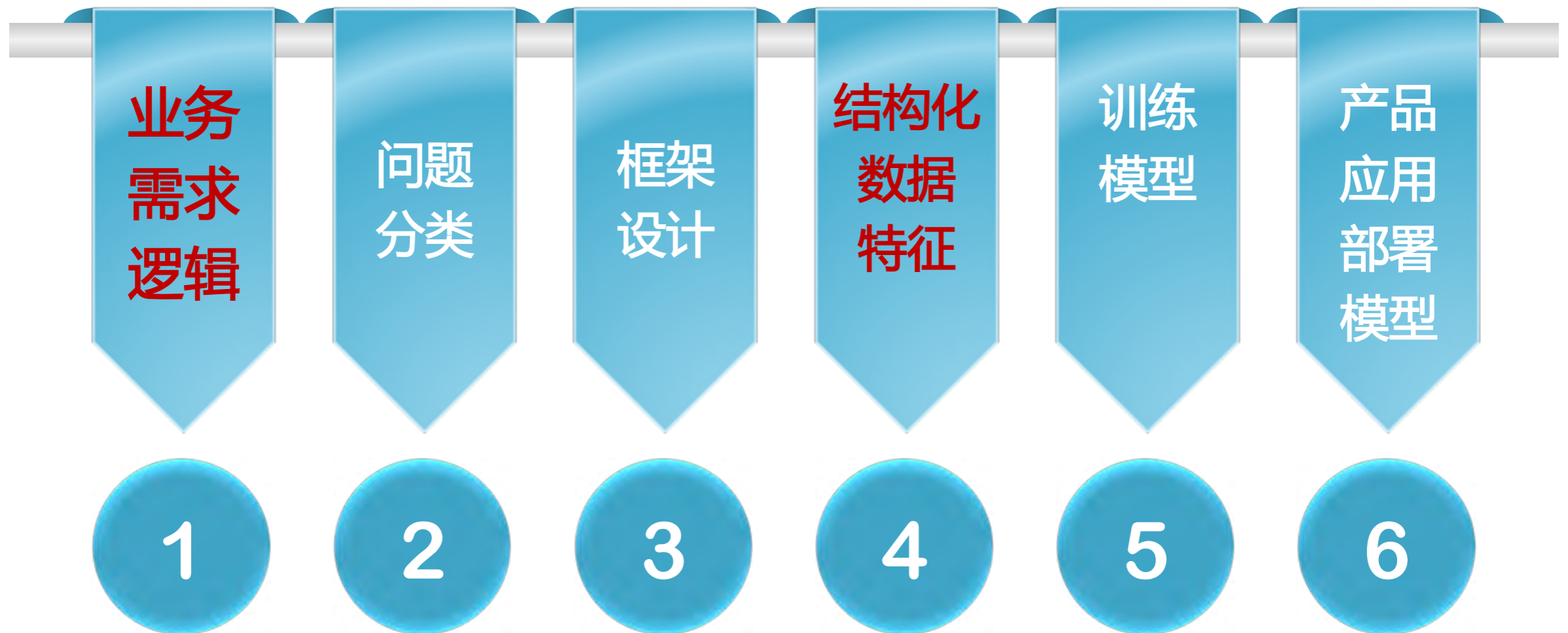
NN(Neural Network)
RNN/CNN

非线性
大规模
调参难

机器学习系统：多种方法集成并用

业务问题如何通过数据转换算法模型？

明确目标、了解特性流程、确定用户对象、设计算法和特征



重中之重，预测对象确定，约束转为特征

需求预测产品中机器学习集成算法架构

数据集成
初次分类

SAS
时间序列

R
机器学习

在线
应用

结果：
准确率提升8%
效率提升

特点：历史时间序列和当前情况，人工经验的结合

设置目标和约束 时间序列转换成特征



针对不同维度对象
定时调度

在线加入人工反馈

实践关键点：综合特征，多种模型，充分利用主流框架

What
How
Why

模型

多快
好省

组合方案
多试多用

业务中
特征提取

GPU提高效率
Tensorflow
Caffe等等

特征

平台
框架

关键会议期刊，跟踪前沿

学术界关注模型

产业界关注应用

借鉴应用

IJCAI 2017



ICML 2017
August 6-11 Sydney, Australia
官网 历年论文精选
重要日期
Paper Submission Deadline: Fri Feb 24, 2017
Author Feedback: Apr 14 - 19, 2017
Volunteer Application Deadline: May 22, 2017
Student Travel Application Deadline: May 29, 2017

ICCV 2017



中国 R 会议
The China-R Conference

CVPR 2017

JMLR

ICLR 2017

PAMI

KDD 2017

NIPS 2017

WWW 2017

wsdm 2017
Tenth ACM International Conference on Web Search and Data Mining

SIGMOD/PODS 2017
Chicago, Illinois, USA, May 14-19

TABLE OF
CONTENTS 大纲

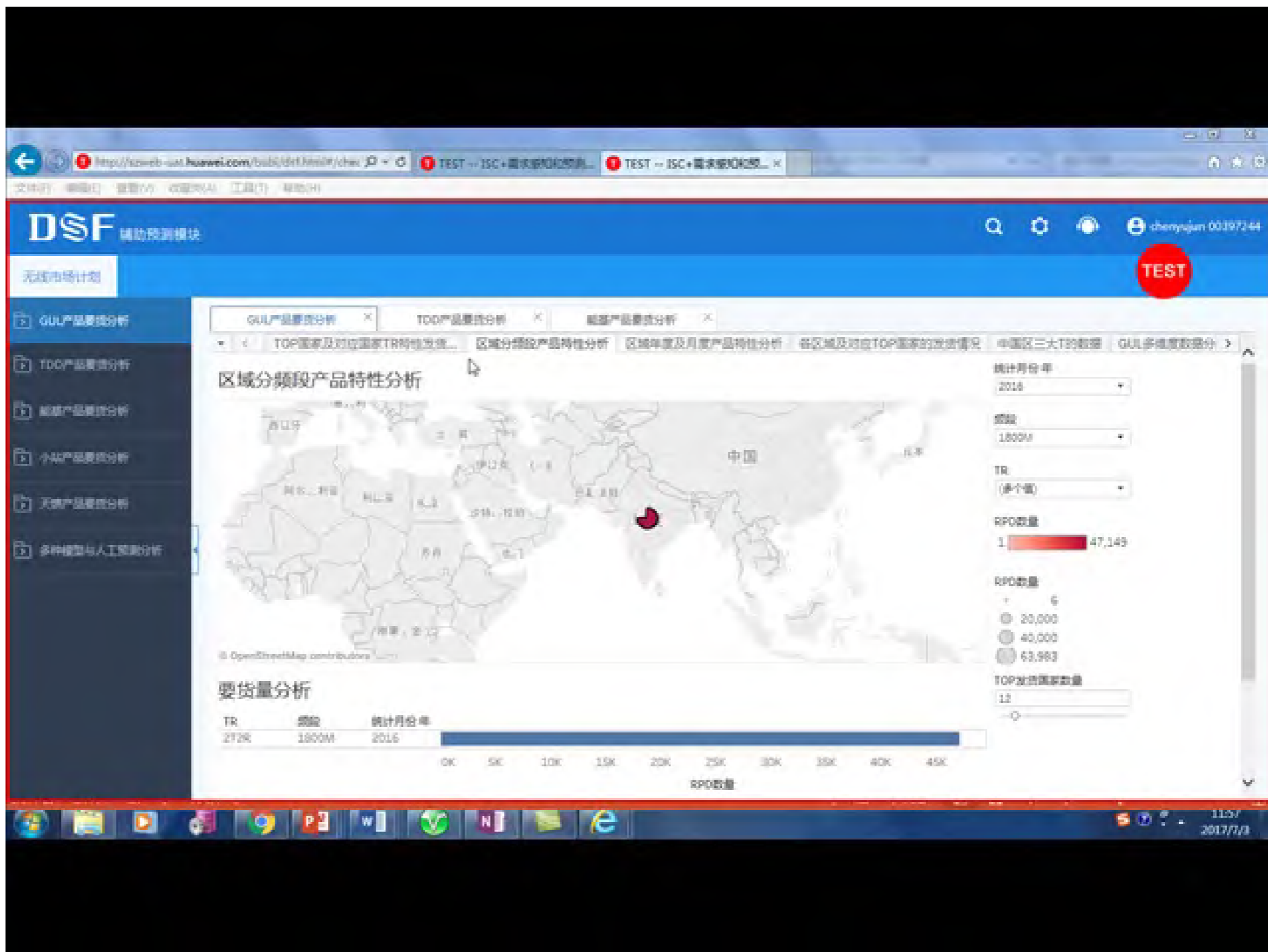
- 供应链领域的智能化应用
- 需求预测产品和技术架构
- 智能算法在产品中应用
- **分析可视化在产品中呈现**
- 总结和展望

需求感知预测数据分析现场演示

用户体验

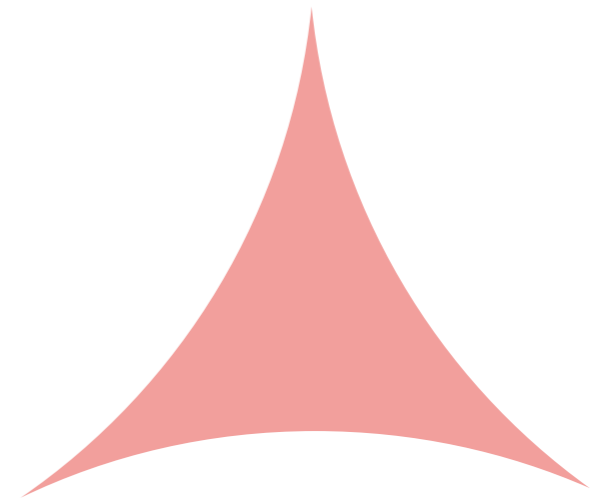
可视化

快速开发



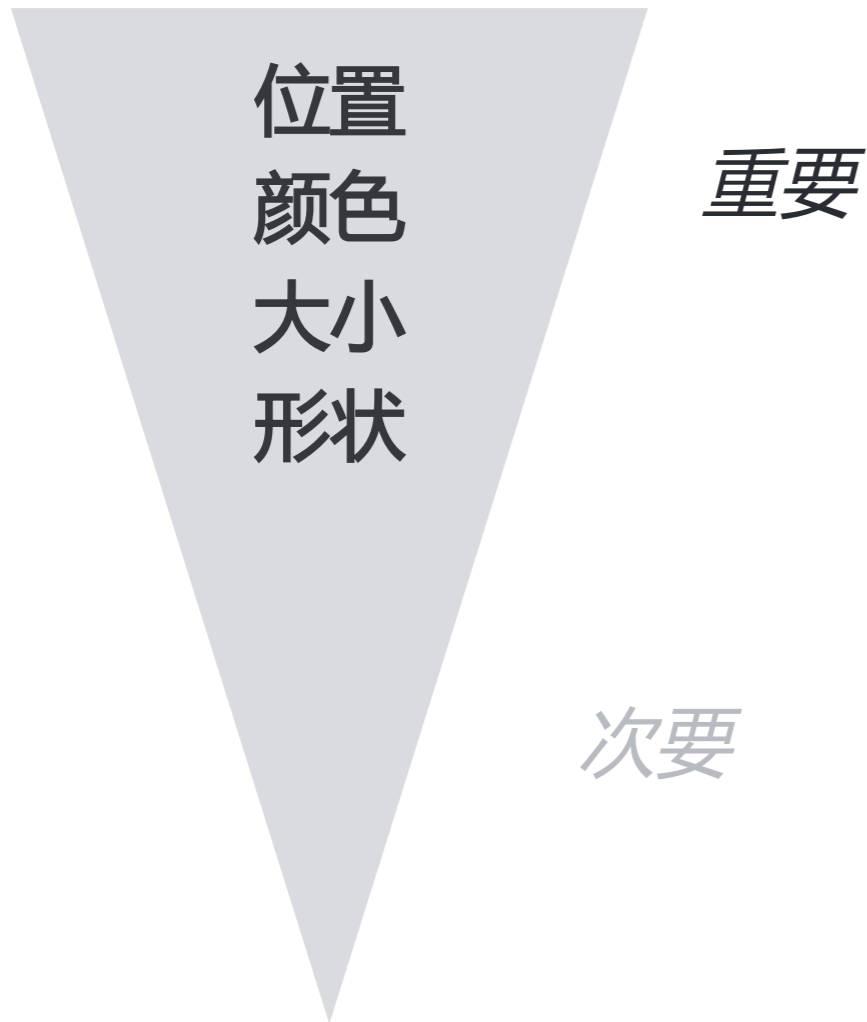
数据类产品中可视化是关键诉求之一

思维和设计原则



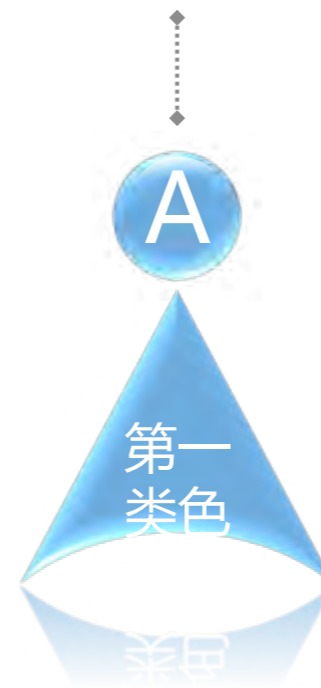
合理设计的数据可视化，可以甄别异常，启发思考，
辅助挖掘数据价值，准确明确有效传递信息。

总体原则- 人们如何观察可视化视图?



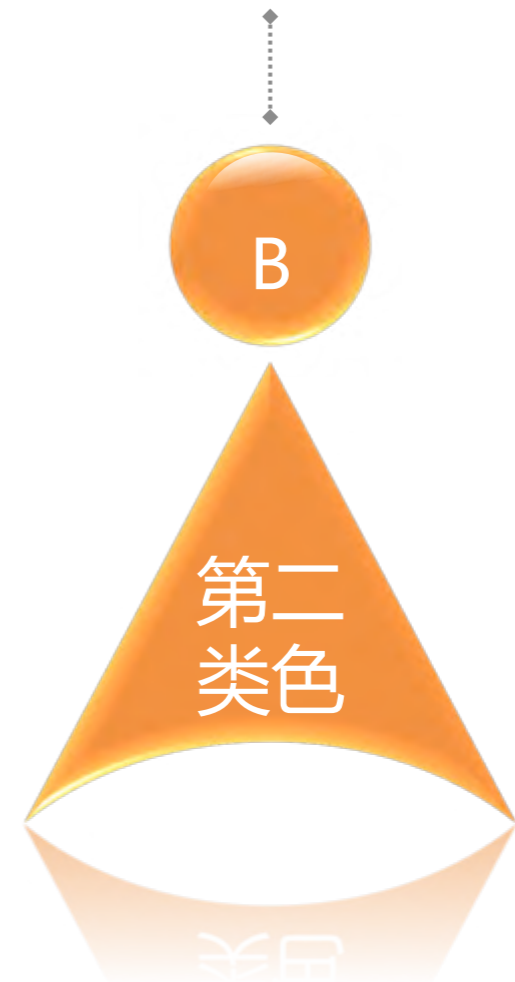
颜色原则

同一个模块中
颜色三类之内

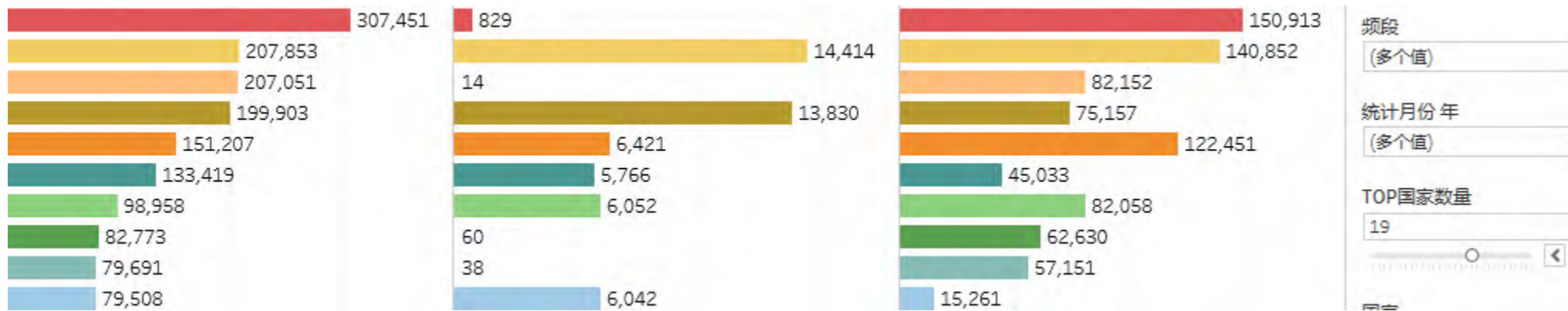


颜色搭配

主色 + 辅色 + 点睛色



视觉设计的基本原则 C.R.A.P- Robin

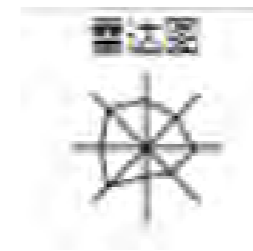


图表分类应用基本分类

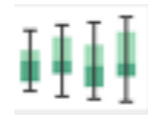
数据关系

适合图表样式

比较



分布



组成



相关



趋势

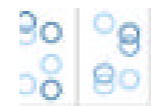


TABLE OF
CONTENTS 大纲

- 供应链领域的智能化应用
- 需求预测产品和技术架构
- 智能算法在产品中应用
- 分析可视化在产品中呈现
- **总结和展望**

总结需求感知预测产品落地之旅

核心：场景设计和业务行业经验

壁垒：模型能力



亮点：分析可视化能力

瓶颈：计算能力

Gartner预测的发展趋势

数据分析大众化，解耦模型算法，云化、产品化、模板化的综合应用

数据归一化

统一数据语言

算法集成化

算法库，集成
预测服务

数据标准化

特征标准化

算法标准化

场景标准化

特征标签化

提供特征工程工具
标签化，自动化的特征识别

场景模板化

分类、预测、决策、
分析服务模板化

总结和展望

发散到收敛
创新



THANKS!