

Oracle让您用数据大“画”渠道

——基于营销渠道画像的新品投放

免费云试用及专家技术咨询：**400 699 8888**

ORACLE®

Copyright © 2019, Oracle and/or its affiliates. All rights reserved. |

目录

- 1 数据分析能够帮到我们吗？
- 2 数据探索与分析
- 3 敏捷的数据挖掘

数据分析能够帮到我们吗？

将数据分析挖掘技术与实践检验相结合，为消费品在新市场投放找到最具营销价值的零售户，利用他们稳定的客户群和市场影响力，从而提高市场推广效率和营销效果。

零售门店是消费品的重要销售和营销渠道。

零售户以其优势的地理位置和接触式销售模式使其具有独特的营销价值



便利店



连锁超市



食杂店



超市



零售业态

渠道精选开启“新品投放”的新征程

- 生产能力和工艺水平大幅度提高
- 研发能力大幅度提升



客户经理

用数字化的方法从城市的几万零售户中找到最具价值的营销对象，实现新品投放的精准营销



- 新品推出速度越来越快

• 谁会是新品投放的有力推动者

- 零售户的差异化市场细分正在形成；

在其他城市的销售经验，是目标城市投放的宝贵财富

- ① 通过数据分析探索的方法选出最具参考价值的城市；
- ② 通过数据加工处理构建零售户画像；
- ③ 通过机器学习的方法，基于参照城市建立模型，应用到目标市场，找出有能力、有意愿销售的零售户。



确立工程方法

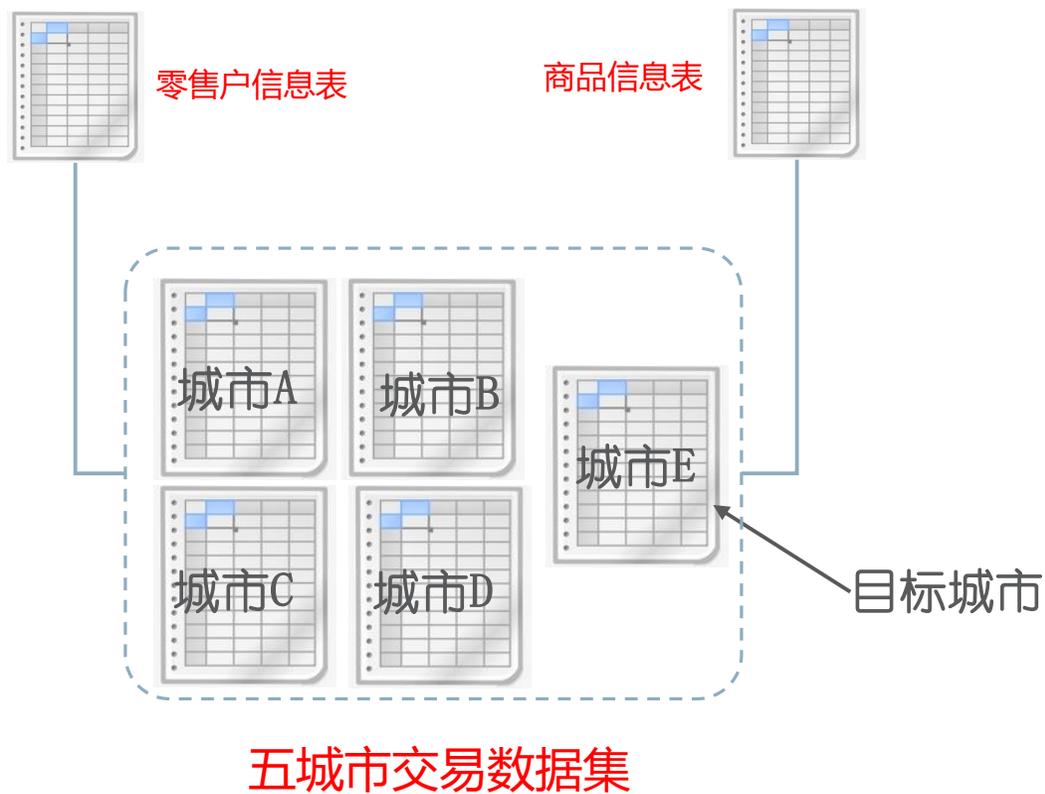


在实践中遵从了国际公认的数据挖掘标准化流程 CRISP-DM

目录

- 1 数据分析能够帮到我们吗？
- 2 数据探索与分析
- 3 敏捷的数据挖掘

客户数据



5个城市，10万零售户，
近亿条的年交易量

交易编号	零售户编码	交易日期	商品编号	交易量	交易金额
....
....
....
.....					
....

数据的探索与分析



客户数据



城市的探索与分析

零售户画像的探索与分析

城市的探索与分析

基于有限的数据和结合业务人员的经验，**快速**的确定出目标城市。

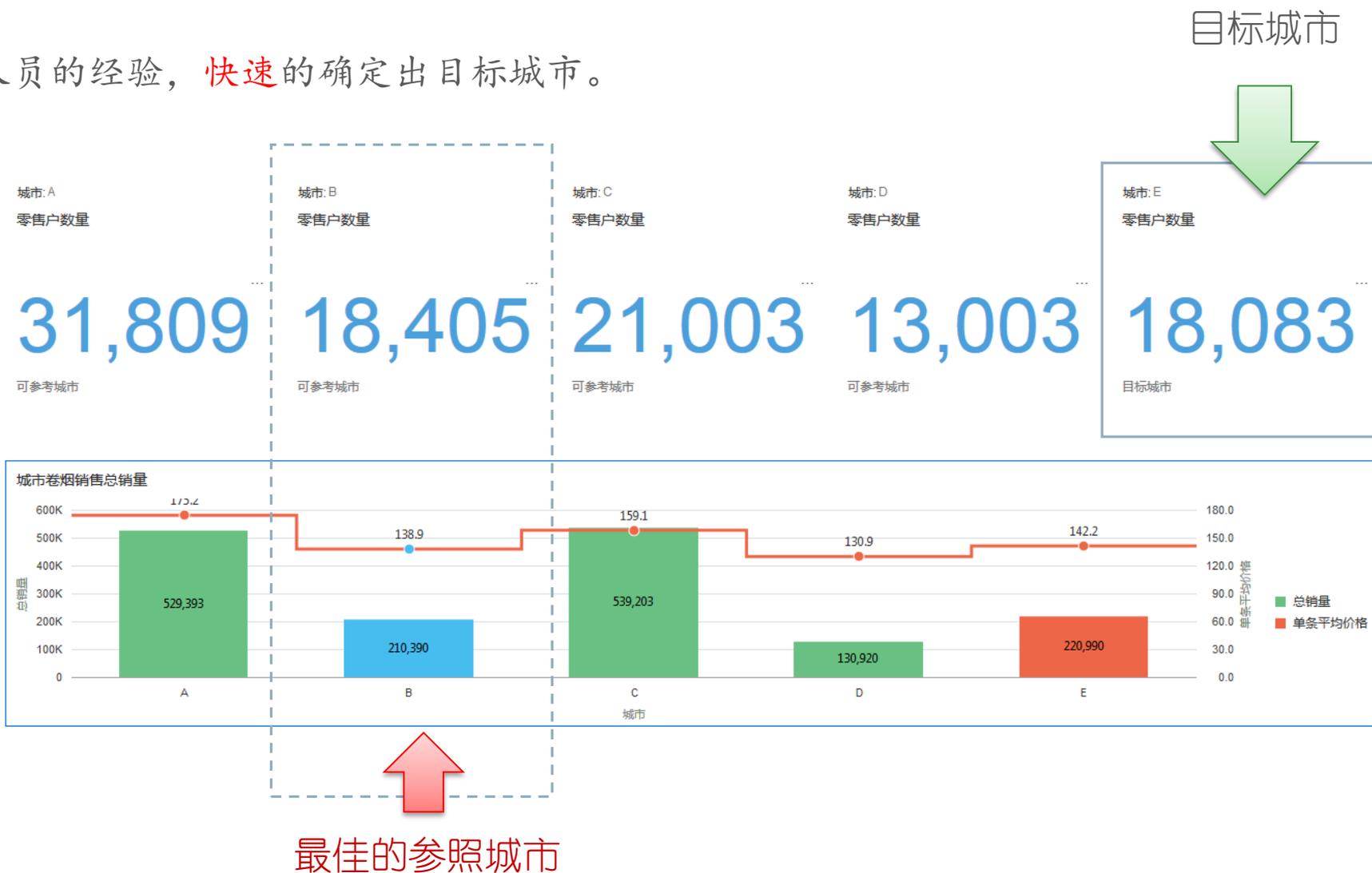
主要对比指标：

2018年零售户总量

2018年销售总量

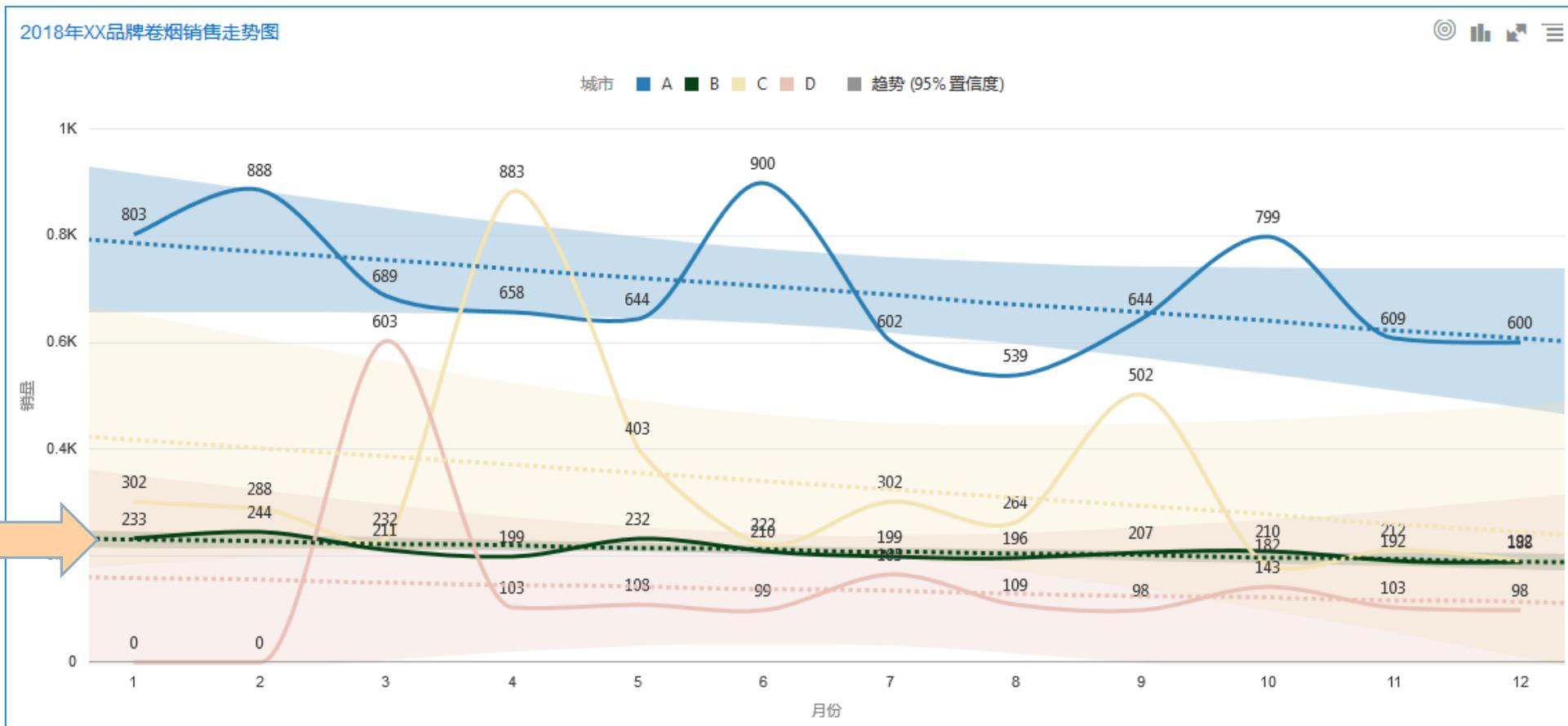
2018年品牌销售趋势

2018年零售户平均销售价格

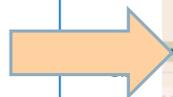


目标品牌在参照城市的市场表现

波动小代表市场供需稳定，市场成熟，更能反映出真实的市场原生需求



最佳参照城市B



平台强大的数据运算处理能力

5个城市，10万零售户，近亿条的年交易量，
在 Oracle 敏捷业务数据平台上进行实时的交互
式数据分析，**分析视图在秒级别呈现。**



零售户画像构建思路

基于新品投放的业务特点，从零售户新品销售意愿，能力出发，设计零售户画像。

有意愿销售 A	有能力销售 B
(1) 不抵触：各品牌都卖的都很好，品牌依赖度低。	(1) 同价格区间商品卖的好, 占比高
(2) 有意愿：爱卖新品。新品卖的多，占比高。	(2) 新品卖的好, 占比高。
(3) 有意愿：同品牌卖的好，销量占比高	(3) 同品牌卖的好，销量占比高
(4) 能接受。在卖同品牌。	(4) 关联品牌卖的好。
	(5) 总体卖的好，销售额高，销售量高

零售户画像的探索与分析



三大挑战

- ① 如何量化零售户的品牌均衡度？
- ② 如何找出具有关联性高的品牌？
- ③ 怎样建立一个机器学习模型，其中蕴含了XXX品牌在参照城市销售的零售户特征？



目录

- 1 数据分析能够帮到我们吗？
- 2 数据探索与分析
- 3 敏捷的数据挖掘

敏捷的数据挖掘

Oracle业务数据平台让数据挖掘不再是奢侈品，而是必需品!!!



数据准备

数据丰富

数据加工

数据过滤

人工补录

数据计算

机器学习

基于数仓快速开发能力



同价区销量占比

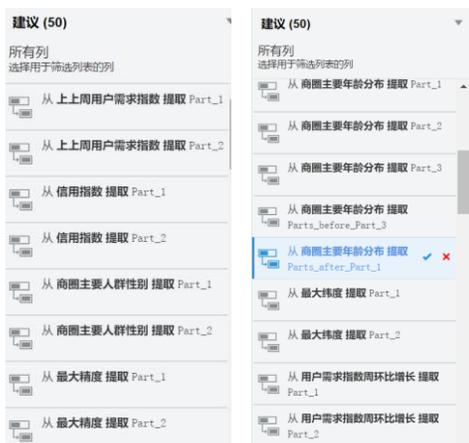
新品牌数量占比

新品销量占比

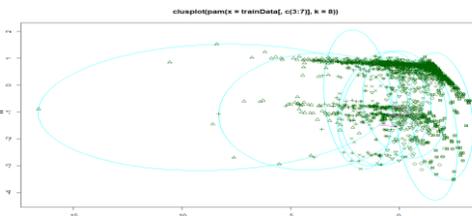
标准库丰富

智能判断

聚类补充



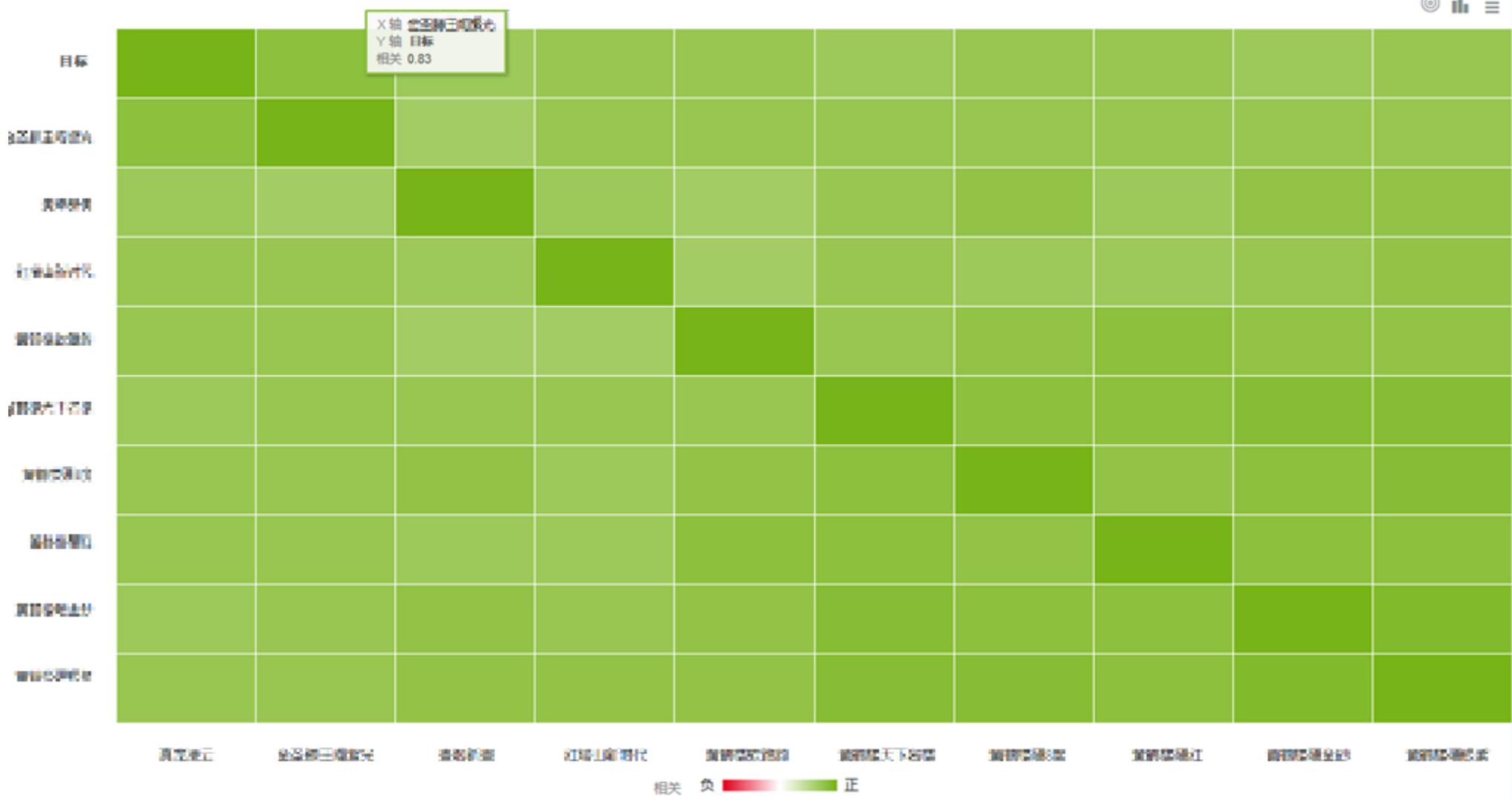
品牌均衡度是一个关键指标



数据准备

数据
丰富

相关
系数

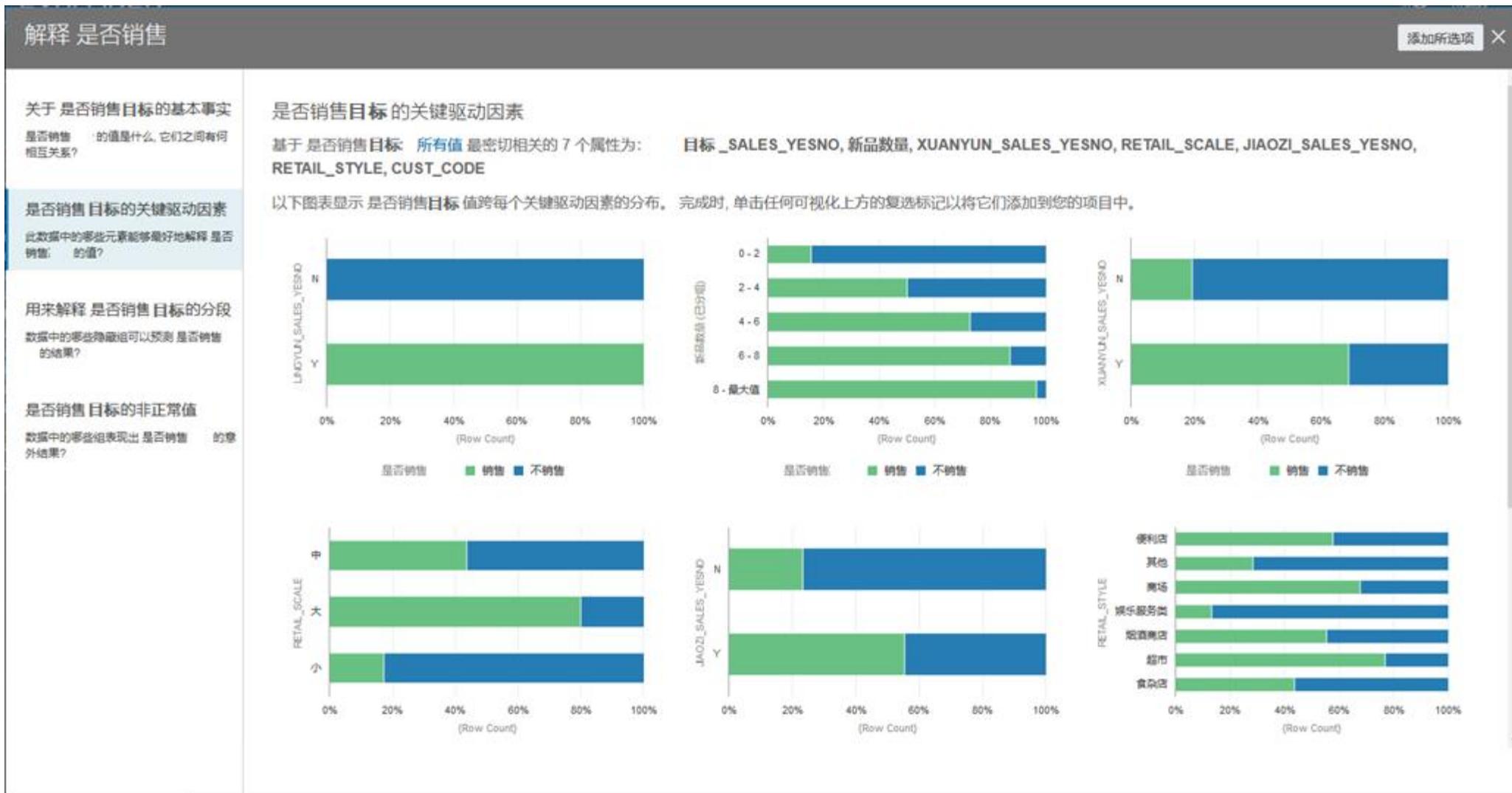


数据训练

数据
参数

模型
选择

模型
生成

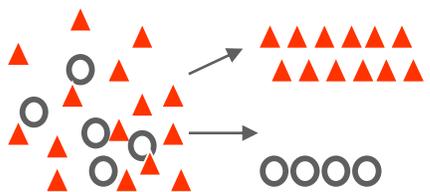


数据训练

数据
筛选

模型
选择

模型
生成



- 1) 我们以B市零售户是否购买“目标”作为预测目标，建立二分类数学模型，形成一套预测零售户是否售卖的规则。
- 2) 将该模型（规则）应用于B市零售户，以预测哪些零售户可能会购买“目标”；
- 3) 并以B市数据集训练形成的**置信度**作为依据，聚焦目标零售户。

天天天 A: 购买客户 vs. B 非购买客户 人

预测置信度

零售户购买目标的数学判断规则

零售户画像



训练二进制分类器

模型训练脚本 `CART for model training`

* Target `LINGYUN_SALES_YESNO`

target, the target(label) to learn/predict

Positive Class in Target

Positive class in the target value. Default is Yes.

Minimum Node Size

Minimum Node Size

Maximum Depth

Maximum Depth

Maximum Confidence

Maximum Confidence

Train Partition Percent

The percentage of original data used for training, default is 80%.

Balancing Method

The mode method for balancing unbalanced set.

选择训练二分类模型脚本

搜索



Logistic Regression for model training



SVM for Classification



Naive Bayes for Classification



CART for model training



Neural Network for Classification



Random Forest for model training

取消

确定

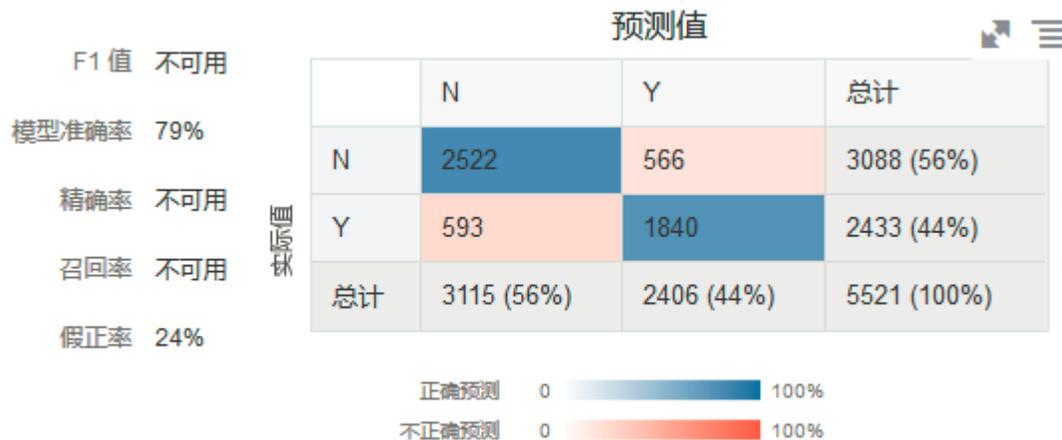
模型评估是对模型的价值和意义探索，通过“模型测试”和“业务经验”两个角度对模型质量进行评估，确保模型的可信性与可行性。如果模型的精确度不高，或置信度不高，模型都不具有实际价值；如果模型不能获得业务上的合理解释，同样会存在可行性风险。

数据筛选

模型选择

模型生成

样本数据集测试结果：精确度（命中率）：79%



基于B市全数据集的执行结果：

- 整体精确度（命中率）：80.3%
- 预测购买目标的准确度（命中率）：79.6%

	不销售	销售	总计
	零售户统计_B	零售户统计_B	
N	8,306	1,650	9,956
Y	1,972	6,477	8,449
总计	10,278	8,127	18,405

数据预测

PredictionConfidencePercentage
0.986321482958768

零售户数量统计_目标城市

484

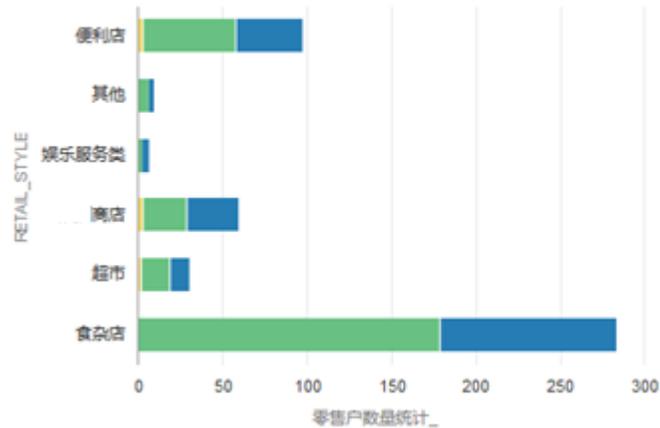
按 销区名称, MARTET_TYPE 列出 零售户数量统计_目标城市

数据
过滤

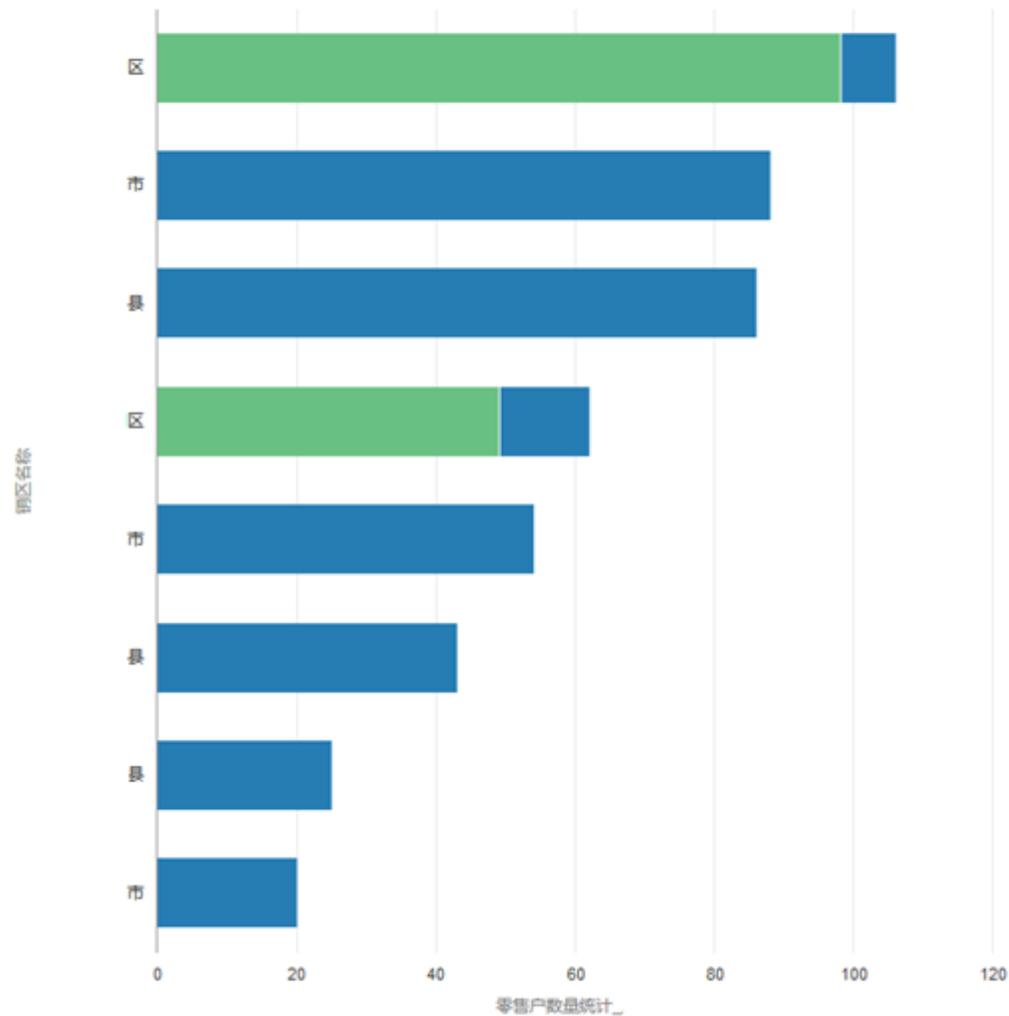
模型
应用

结果
输出

按 RETAIL_STYLE, RETAIL_SCALE 列出 零售户数量统计_目标城市



销区名称



RETAIL_SCALE 小 大 中

MARTET_TYPE 城镇 乡村

数据预测

数据
过滤

模型
应用

结果
输出

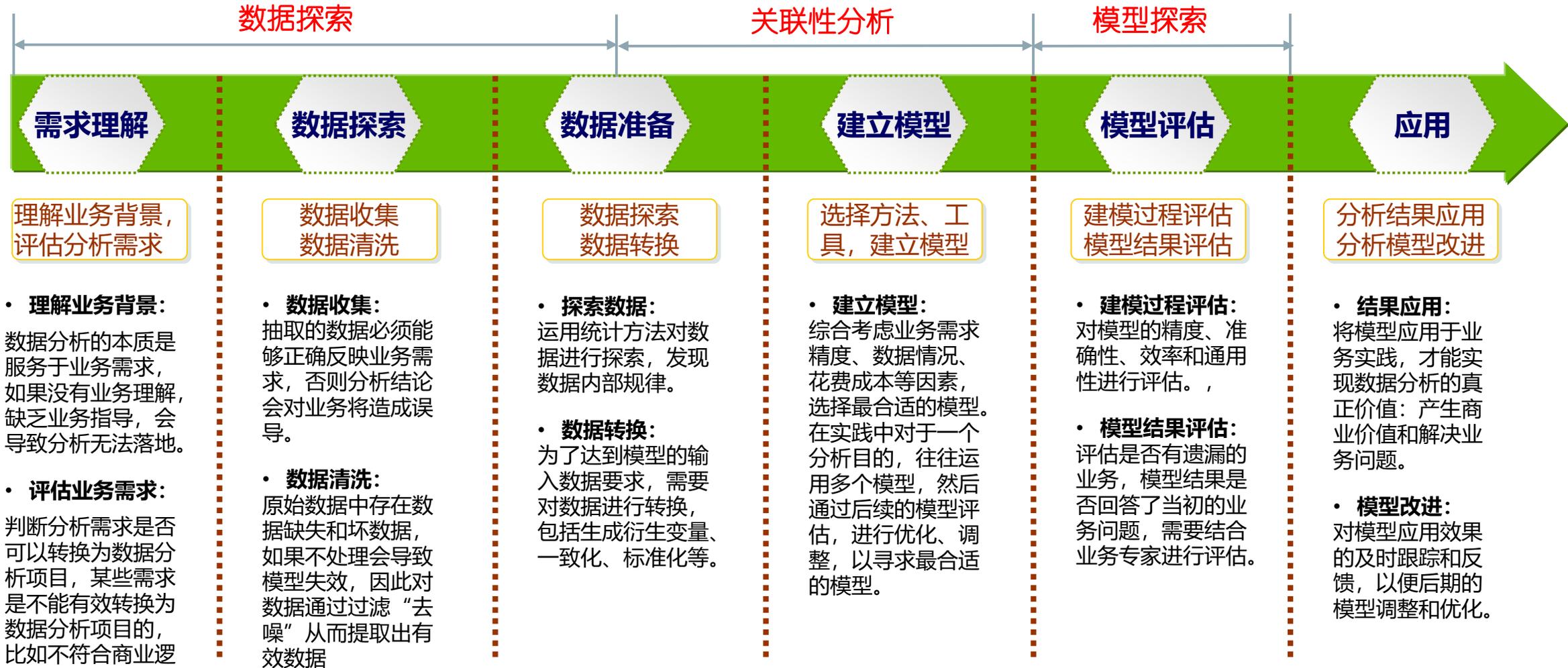
PredictedValue PredictionConfidencePercentage (7)
Y 0.83080137107243, 0.833333333333333, ...



零售户代码, PredictedValue, CUST_CODE, CUST_NAME, PredictionConfidencePercentage, 品牌均衡度, 新品占比

零售户代码	PredictedValue	CUST_CODE	CUST_NAME	PredictionConfidencePercentage	品牌均衡度	新品占比
421002100010	Y	421002100010	任立银	0.986321482958768	0.03	0.13
421002100011	Y	421002100011	裴德贵	0.954545454545455	0.02	0.18
421002100018	Y	421002100018	张保华	0.954545454545455	0.02	0.20
421002100024	Y	421002100024	张子兰	0.954545454545455	0.02	0.16
421002100025	Y	421002100025	赵忠武	0.954545454545455	0.03	0.13
421002100029	Y	421002100029	段云波	0.954545454545455	0.02	0.12
421002100031	Y	421002100031	刘才华	0.986321482958768	0.02	0.16
421002100036	Y	421002100036	杨群娥	0.954545454545455	0.03	0.06
421002100099	Y	421002100099	汪爱林	0.986321482958768	0.02	0.21
421002100100	Y	421002100100	任少冬	0.986321482958768	0.03	0.13
421002100103	Y	421002100103	邱昌旺	0.986321482958768	0.02	0.15
421002100118	Y	421002100118	马超男	0.954545454545455	0.03	0.06
421002100123	Y	421002100123	孙刚	0.954545454545455	0.03	0.19
421002100124	Y	421002100124	凡哲兵	0.954545454545455	0.03	0.11
421002100125	Y	421002100125	曾凡青	0.954545454545455	0.02	0.19
421002100126	Y	421002100126	郭翠兰	0.954545454545455	0.03	0.14
421002100129	Y	421002100129	郭昌梅	0.954545454545455	0.03	0.09
421002100135	Y	421002100135	瞿国珍	0.954545454545455	0.02	0.18
421002100141	Y	421002100141	蔡希金	0.954545454545455	0.04	0.16
421002100147	Y	421002100147	唐平	0.954545454545455	0.03	0.19
421002100148	Y	421002100148	朱学明	0.954545454545455	0.03	0.17
421002100154	Y	421002100154	蔡丽	0.954545454545455	0.03	0.15
421002100162	Y	421002100162	李林松	0.954545454545455	0.03	0.07
421002100167	Y	421002100167	李杰	0.954545454545455	0.02	0.18
421002100172	Y	421002100172	吴云	0.954545454545455	0.02	0.16
421002100193	Y	421002100193	王丕英	0.986321482958768	0.01	0.21
421002100194	Y	421002100194	熊建国	0.954545454545455	0.02	0.14
421002100219	Y	421002100219	李昌英	0.954545454545455	0.02	0.15

业务模型框架-基于数据探索的数据挖掘技术应用



在实践中遵从了国际公认的数据挖掘标准化流程CRISP-DM (cross-industry standard process for data mining)。

ADW 业务数据平台的创新方法

价值
导向

由小
及大

简化
程序

敏捷
高效

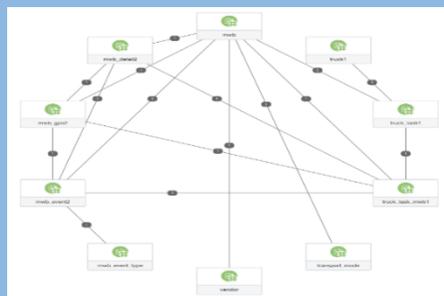
业务数据来自哪儿



ADW/业务系统/Excel...

智能管理

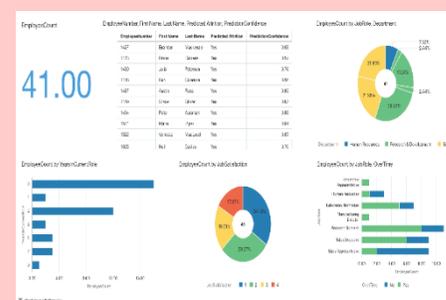
业务数据之间的关系



业务逻辑关联图
Data Diagram

迭代、验证、固化

业务数据怎么用



业务分析画布
Analysis Canvas

业务洞察

ORACLE®