

您的**信息** 您的**智慧**

**2011 IBM 信息管理与业务分析论坛**

# SPSS: 数据分析的新纪元

刘 海 亮

预测分析软件 大中华区 销售经理





# 目录

1

新纪元的特点和业务分析软件

2

SPSS简介

3

SPSS案例及应用说明





# 卓越企业善于在以下三个领域来推动企业发展

## 1. 重塑客户关系

- 在产品/服务生命周期的各个阶段增加客户互动，从而更加贴近客户；
- 通过多样化的渠道战略为客户提供更加丰富并持久的客户体验；
- 运用数据分析，更加智慧地理解、洞察客户需求，并为其提供更好的产品或服务。

## 2. 展现创新领导力

- 企业领导者不仅要以快速增长为企业目标，更要高瞻远瞩，方能安然应对复杂环境的挑战；
- 创新的CEO会带领企业通过差异化的价值定位，持续实现业务创新；
- 创建可持续发展的组织，以更灵活的运营模式满足客户日益提升的需求。

## 3. 构建灵活的运营

- 优化组织运营模式和产品，以更好处理复杂性；
- 运用数据分析工具和手段，提高科学决策水平和速度；
- 构建灵活运营模式以提高战略执行速度。





# IBM 业务分析软件的价值定位

Strategic Advantage

**Analysis and Insight**  
Enhance customer insight to enable profitable growth



**Strategic Alignment**  
Enterprise visibility and aligned execution



**Execution**  
*Transaction and Operational Systems*



Operational Decisions

**Information Access**  
Foster informed / collaborative decision making



**Business Agility**  
Drive real-time decision optimization



**Foundational Capabilities:**



**Proven Practices:** *Business strategy, industry and technical expertise*

**GBS**

Awareness

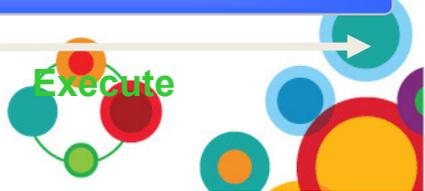
您的信息 您的智慧

Insight

2011 IBM 信息管理与业务分析论坛

In Context

Execute





# 目录

- 1 新纪元的特点和业务分析软件
- 2 SPSS简介
- 3 SPSS案例及应用说明

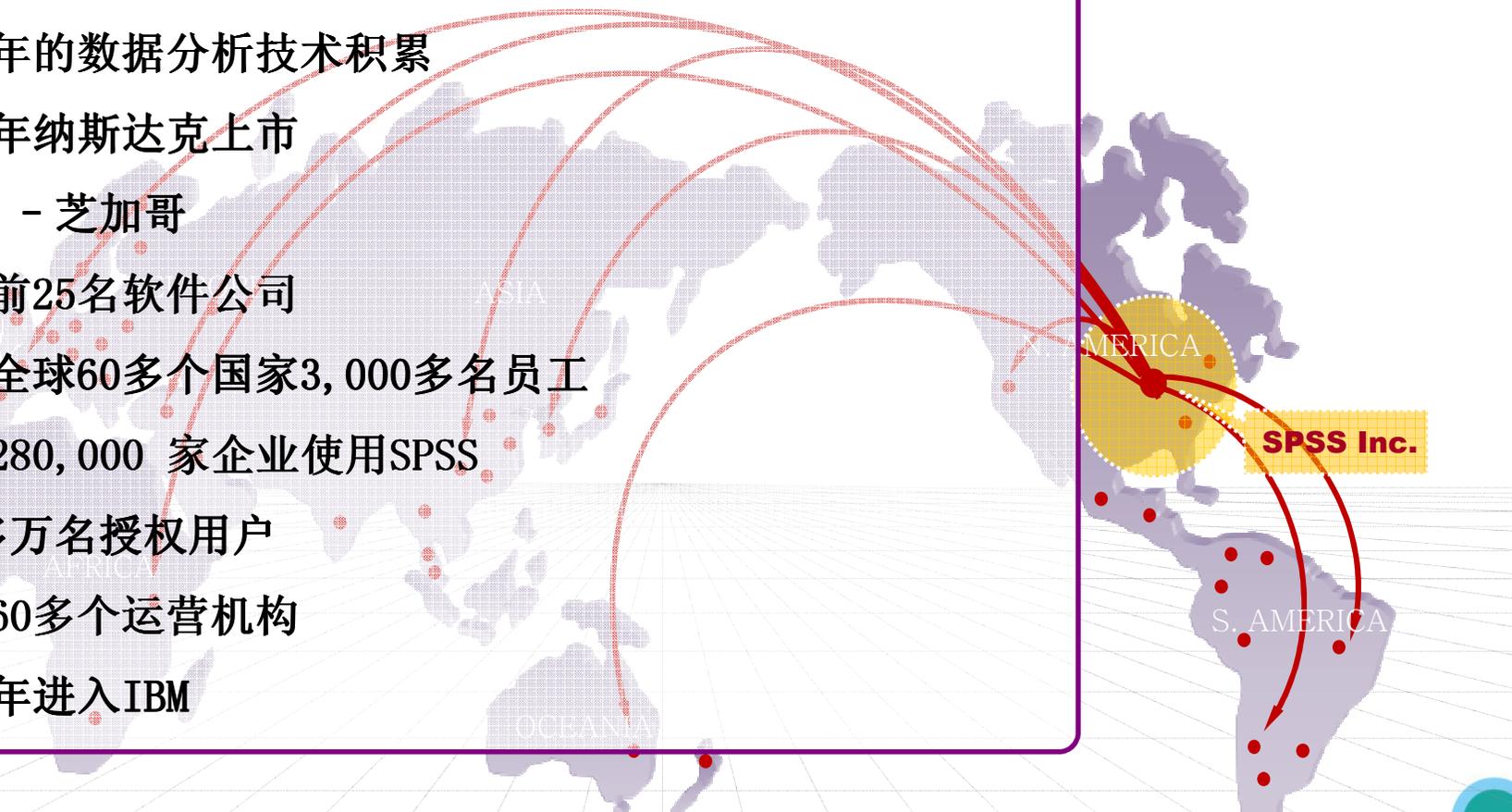


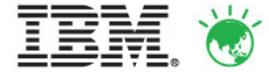


# SPSS介绍

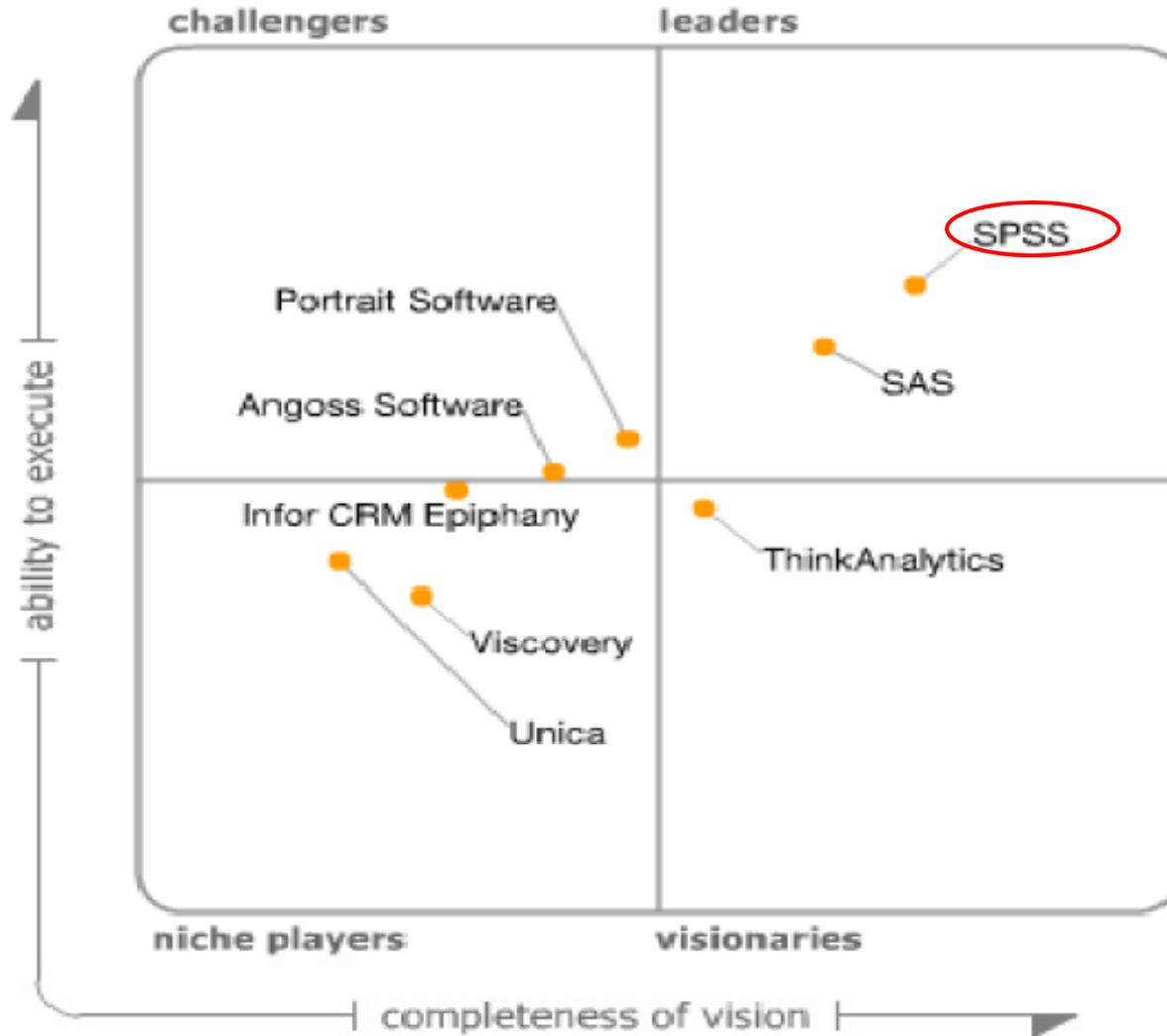


- 1968年成立, 是世界上最早的统计分析软件
- 40多年的数据分析技术积累
- 1993年纳斯达克上市
- 总部 - 芝加哥
- 全球前25名软件公司
- 遍布全球60多个国家3,000多名员工
- 超过280,000 家企业使用SPSS
- 2百多万名授权用户
- 全球60多个运营机构
- 2009年进入IBM





# SPSS是数据挖掘领域的领导者



您的信息 |

Source: Gartner (June 2008)

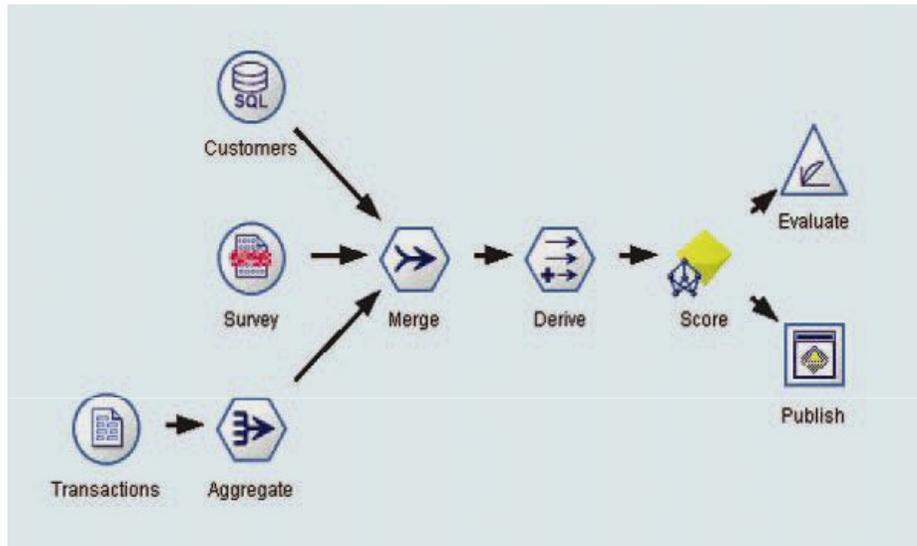
As of June 2008





# 业务分析软件的特点: 易用性

## ①简单直观的图形化数据挖掘



## ②丰富的数据挖掘算法

**预测 & 分类**  
 Neural net, C5.0, C&R Tree, QUEST, CHAID, Linear Regression, Logistic & Multinomial Logistic Regression

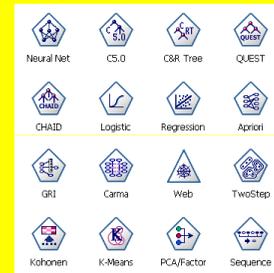
**关联发现**  
 Apriori, GRI, Carma, Web graph

**聚类 & 分群**  
 TwoStep, Kohonen, K-Means

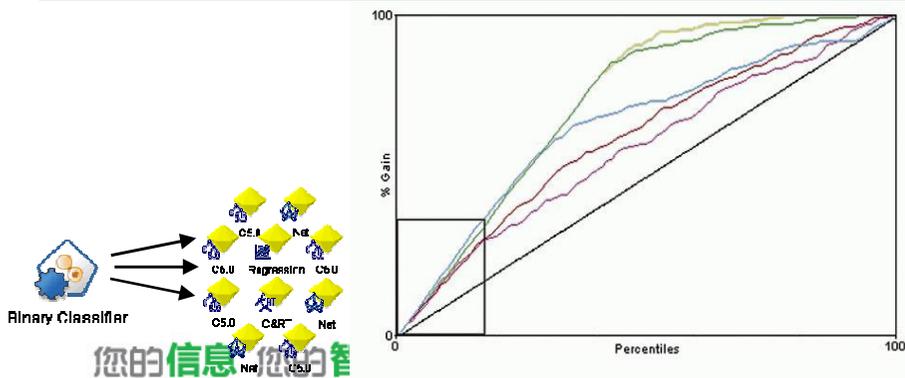
**数据降维**  
 Factor, Principle Components

**序列分析**  
 Capri, Neural Net, Regression

**时间序列**  
 Time Series Forecasting



## ③快速自动发现最优模型



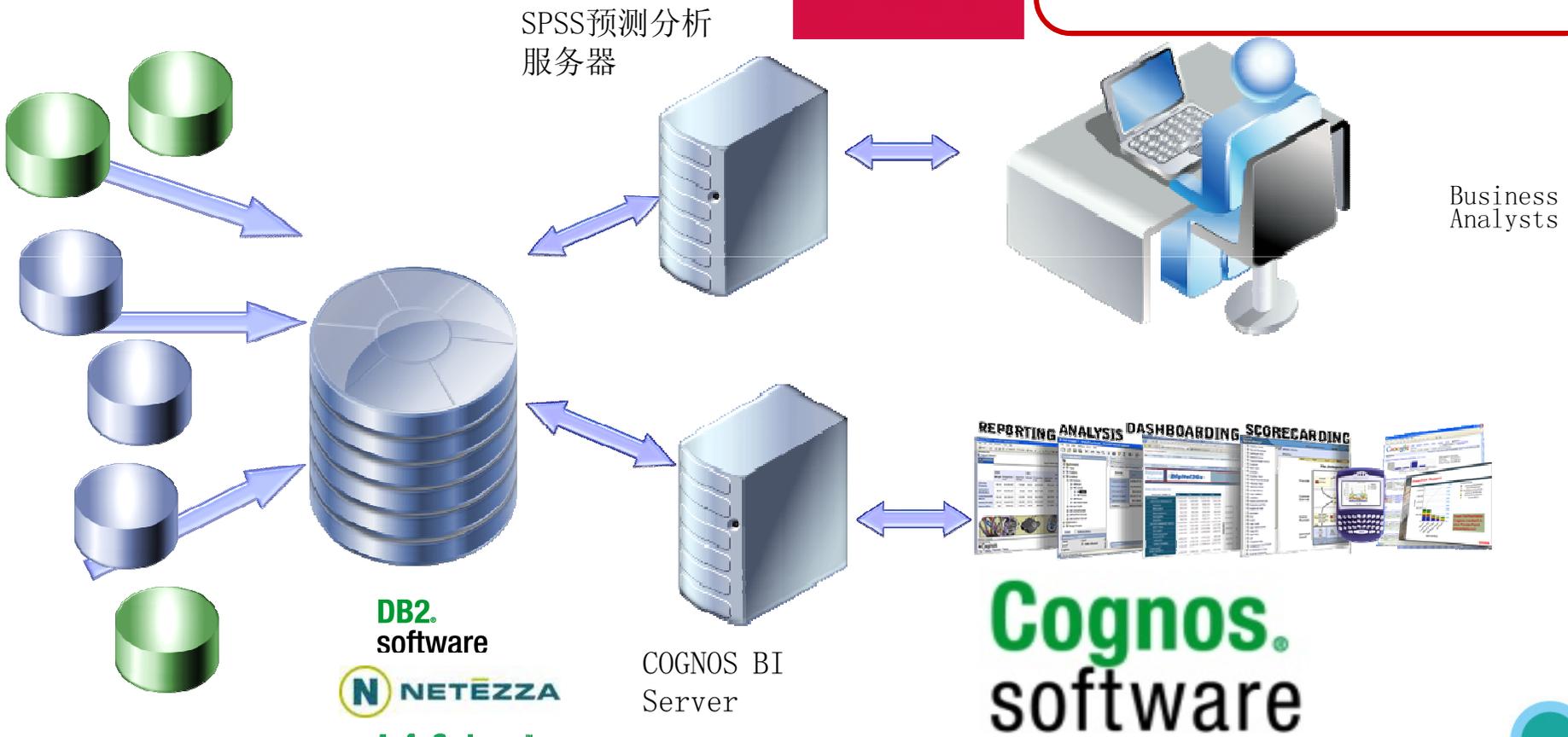
## ④标准流程、整合性好



# 业务分析软件的特点: 整合性



增加业务导向的预测性模型  
和模型管理功能





# SPSS 主要产品与服务

## Data Collection



- 数据采集类产品:  
SPSS Data Collection

## Statistics



- 统计分析类产品:  
SPSS Statistics

## Modeler

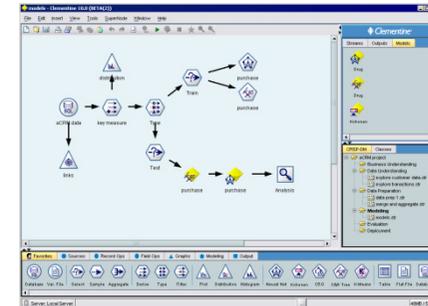
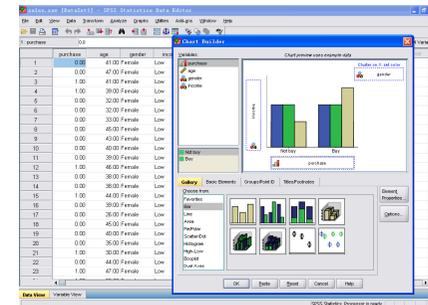


- 数据挖掘类产品:  
SPSS Modeler

## Deployment



- 数据发布类产品:  
SPSS Collaboration and Deployment Services





# 目录

- 1 新纪元的特点和业务分析软件
- 2 SPSS简介
- 3 SPSS案例及应用说明



# 主要解决方案

## 金融领域

**客户细分**：针对不同客户提供不同的服务  
**客户流失**：挽留高价值客户  
**交叉销售**：提高收入和客户的收益率  
**客户满意度分析**：提高客户忠诚度  
**欺诈监测**：分析欺诈模式，识别欺诈行为  
**信用风险管理**：计量和规避信用风险  
**降低索赔**：识别出具有较低索赔可能的客户.....

## 通讯领域

**客户细分、客户流失、营销响应、满意度分析**  
**销售提升**：预测特定消费群的行为特征，设定有效的销售战略  
**客户市场推广**：模拟计费 and 模拟出账  
**客户欠费分析和动态防欺诈**：欠费行为的事前预测  
**呼叫中心管理**：话务量预测，降低运营成本.....

**辅助税收管理**：发现未申报的潜在税务责任人、税种、；识别容易偷漏税的纳税人、针对不同的征收案例制订正确的征收策略；  
**犯罪行为分析**：分析案件关联、串并案；对高危人群及地区犯罪进行预警；  
.....

**门店选址**：找出最佳的新店铺位置  
**库存管理**：满足客户需求，降低库存成本  
**产品设计**：在尽可能少的试验次数下确定影响质量的因素并选择合适的生产水平；  
**质量控制**：实现品管七大手法.....

## 政府及公共事业

## 教育/零售/制造业



# SPSS银行业用户





# 业务分析软件在政府行业的客户



**中华人民共和国人力资源和社会保障部**  
Ministry of Human Resources and Social Security of the People's Republic of China



**国家税务总局**  
State Administration of Taxation

税收取之于民 用之于民 造福于民



**中华人民共和国国家发展和改革委员会**  
National Development and Reform Commission



**中华人民共和国民政部**  
Ministry of Civil Affairs of the People's Republic of China

全国民政门户



**中华人民共和国国家人口和计划生育委员会**  
National Population and Family Planning Commission of P.R.China

中国人口 CHINAPOP

**中国疾病预防控制中心**  
CHINESE CENTER FOR DISEASE CONTROL AND PREVENTION



**国家广播电影电视总局**  
SARFT THE STATE ADMINISTRATION OF RADIO FILM AND TELEVISION



**中华人民共和国环境保护部**  
Ministry of Environmental Protection of the People's Republic of China



**中华人民共和国交通运输部**  
Ministry of Transport of the People's Republic of China



**中华人民共和国农业部**  
Ministry of Agriculture of the People's Republic of China



**中华人民共和国公安部**  
The Ministry of Public Security of the People's Republic of China

简体 | 繁体 | English



**中华人民共和国商务部**  
Ministry of Commerce of the People's Republic of China



**国家级西安高新技术产业开发区**  
HTTP://WWW.XDZ.COM.CN



**中华人民共和国国家统计局**  
National Bureau of Statistics of China



**中华人民共和国国家工商行政管理总局**  
State Administration for Industry & Commerce of the People's Republic of China

**SFDA 国家食品药品监督管理局**  
State Food and Drug Administration



**中国人民银行**  
THE PEOPLE'S BANK OF CHINA



**卫生部卫生经济研究所**  
China Health Economics Institute



cablecom  
touch your worlds





# 英国国家司法部 通过对犯罪记录的分析改善公共安全



Ministry of  
JUSTICE



## 业务需求:

通过对于犯罪记录数据的分析，英国国家法务部希望发现哪些刑满释放人员容易再次进行犯罪活动，寻找其中的关联性。从而进行预防，并改善监狱中对于不同犯人的改造及教育。

## 解决方案:

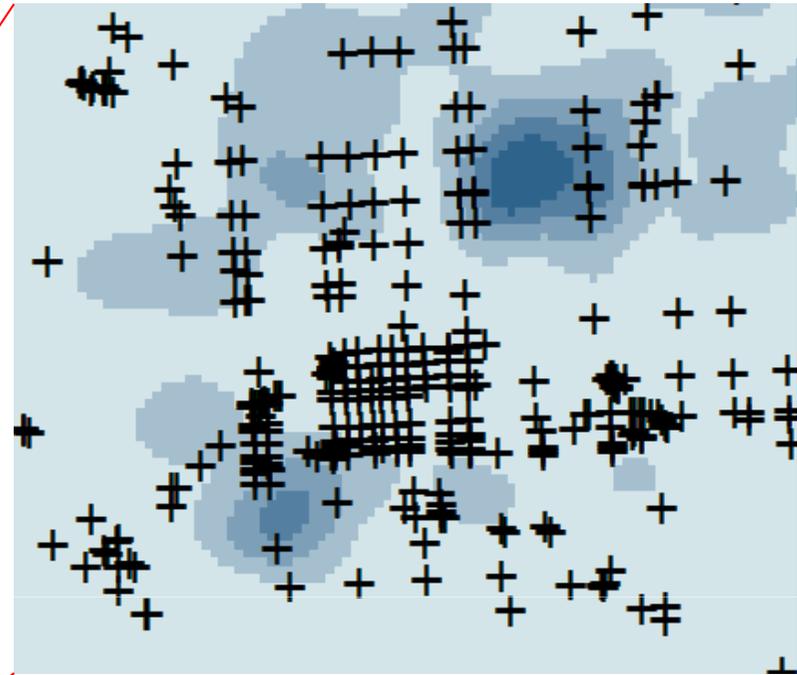
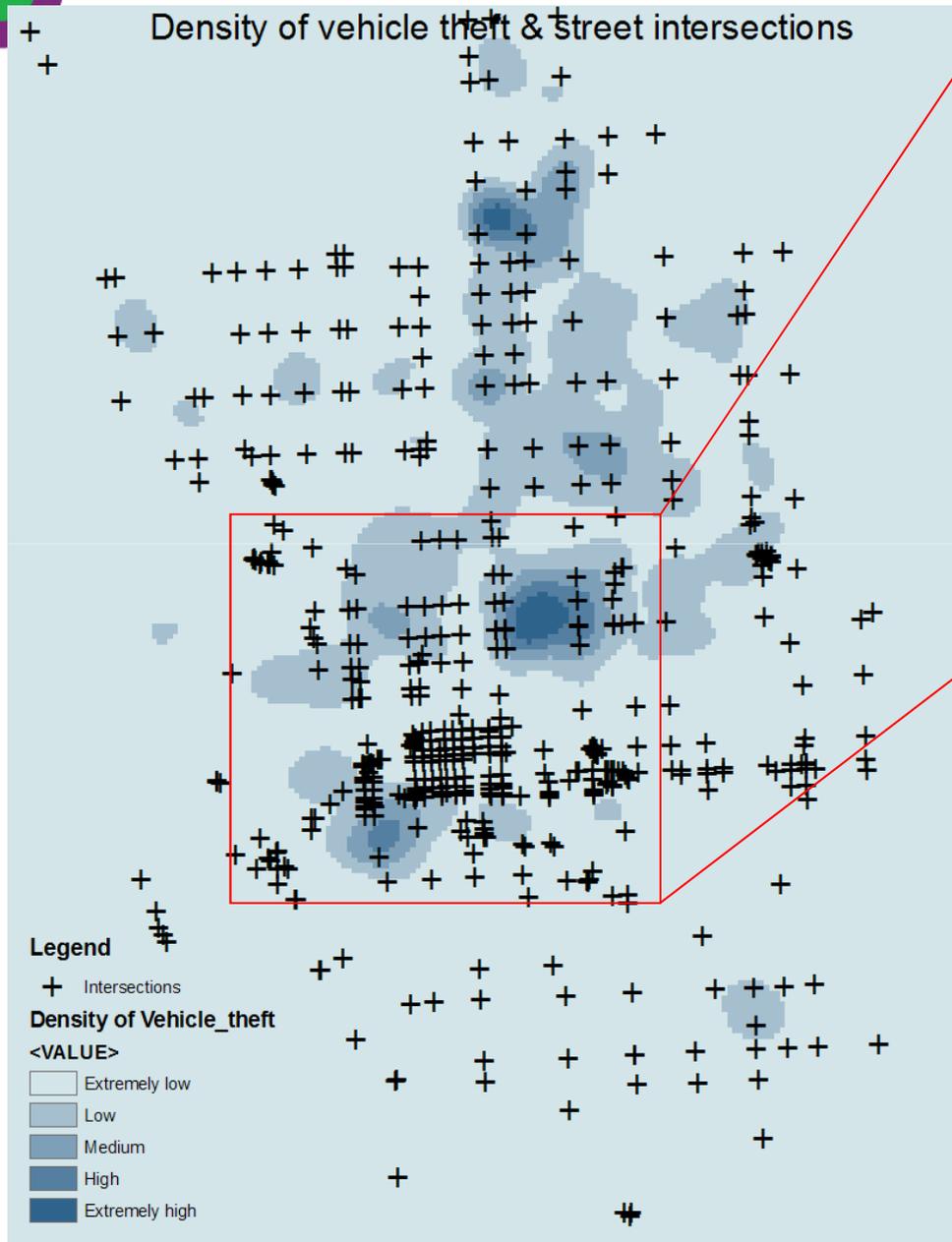
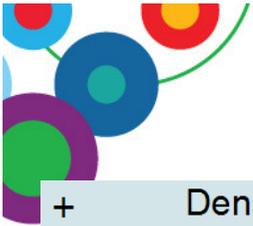
通过使用SPSS Modeler，英国国家法务部对于140所监狱的犯人的犯罪记录和罪犯的行为特征、个性偏好等因素进行综合考量并进行分析。从中发现隐含于其中的大量宝贵信息。

*“With almost 4 million records on file, it simply wouldn't be feasible to trawl through this manually...SPSS technology gives us valuable insight into offender data, helping us predict who may re-offend and enabling us to advise on preventive measures.”*

**Solution  
Component:**

- IBM® SPSS Modeler  
— Spokesperson,  
Ministry of Justice,  
United Kingdom



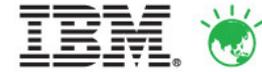


•左边的图显示了偷车事件的点密度.深蓝色是偷车的高密度区.

•小的十字是十字路口的位置

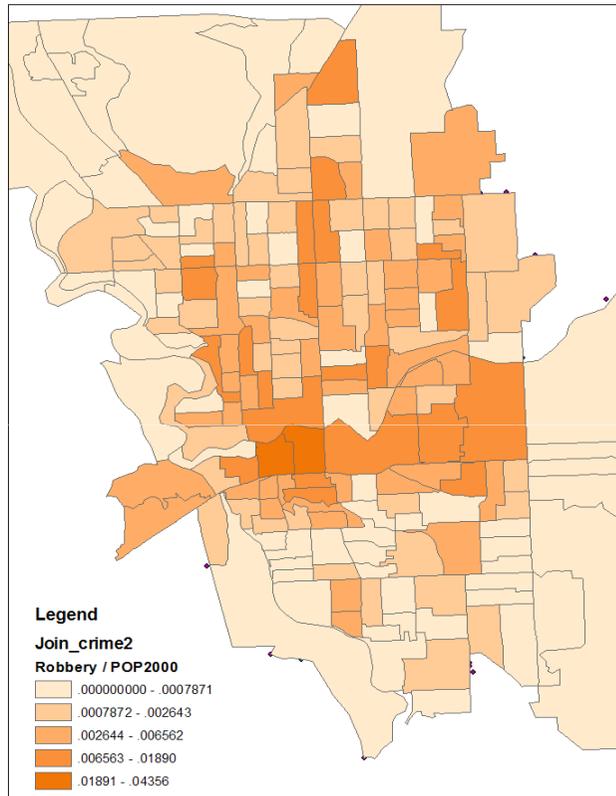
•通过将偷车分布图和十字路口的位置的叠加,可以看出偷车常常发生在十字路口较为分散的一些地点



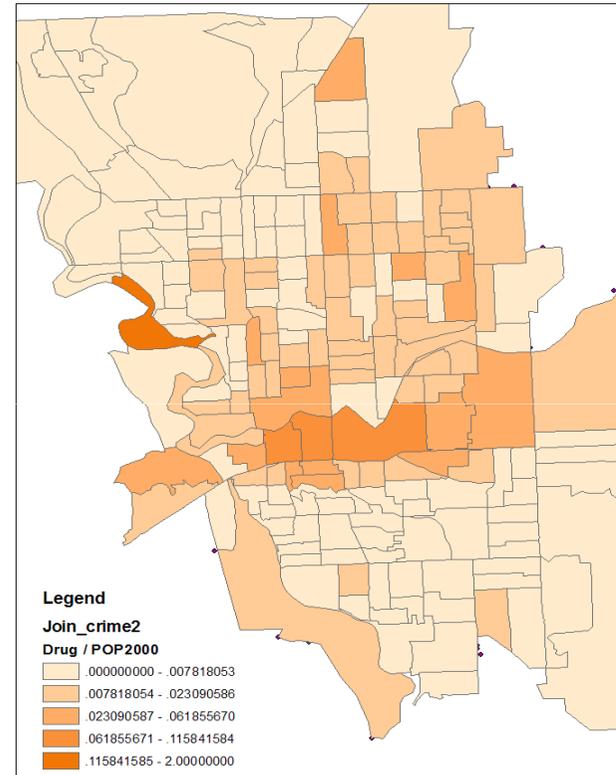


# 事件区间密度值制图 - 抢劫和毒品

Counts of robbery standardized by pop



Counts of drugs standardized by pop



- 人们常常把.抢劫和毒品联系在一起.左边的图显示了2009年一年抢劫在人口普查组区的发生人口百分比.右边的图显示了2009年一年毒品在人口普查组区的发生人口百分比.人口普查组区是比人口普查小区更小的一个人口普查单位。





中华人民共和国国家工商行政管理总局  
State Administration for Industry & Commerce of the People's Republic of China

国家工商行政管理总局数据中心  
State Administration For Industry & Commerce DataCenter

数据校验首页 平时检查 集中检查 问题数据追溯反馈 数据质量模型管理 数据质量分析 分环节报表对比

任务：请选择 区域：广西壮族白 确定 导出EXCEL

### 2.2.1.1.3 内资企业数据冗余检查

单位：广西壮族自治区

项目	分值	百分比	得分
企业基本信息市场主体身份代码冗余	10	25%	100.0000
企业名称冗余	10	25%	99.9971
注册地址冗余	10	25%	100.0000
经营范围冗余	10	25%	99.9971
其他企业	10	40%	99.9985
其他企业	10	33%	100.0000
其他企业	10	33%	100.0000
其他企业	10	33%	100.0000
其他企业	10	30%	100.0000
其他企业	10	33%	100.0000
其他企业	10	33%	100.0000
其他企业	10	33%	100.0000
其他企业	10	30%	100.0000
其他企业	100		99.9994

国家工商行政管理总局数据中心  
State Administration For Industry & Commerce DataCenter

数据校验首页 平时检查 集中检查 问题数据追溯反馈 数据质量模型管理 数据质量分析 分环节报表对比

### 平时检查得分 (2010年11月15日)

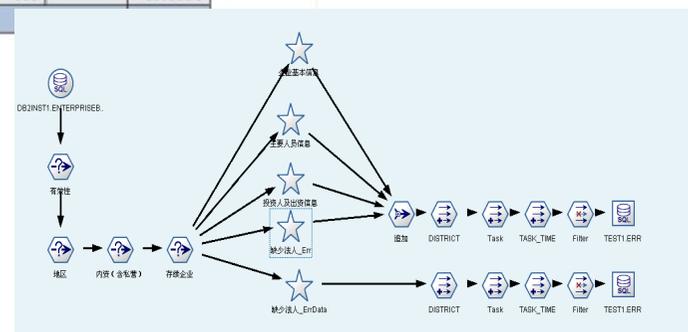
上海

平时检查得分 (2010年11月15日)

上海

时标条件

- 内资企业: 92.3
- 外资企业: 95.2
- 内资企业: 98.2
- 外资企业: 99.2
- 证照: 96.7
- 黑名单: 100.0
- 一人公司: 100.0

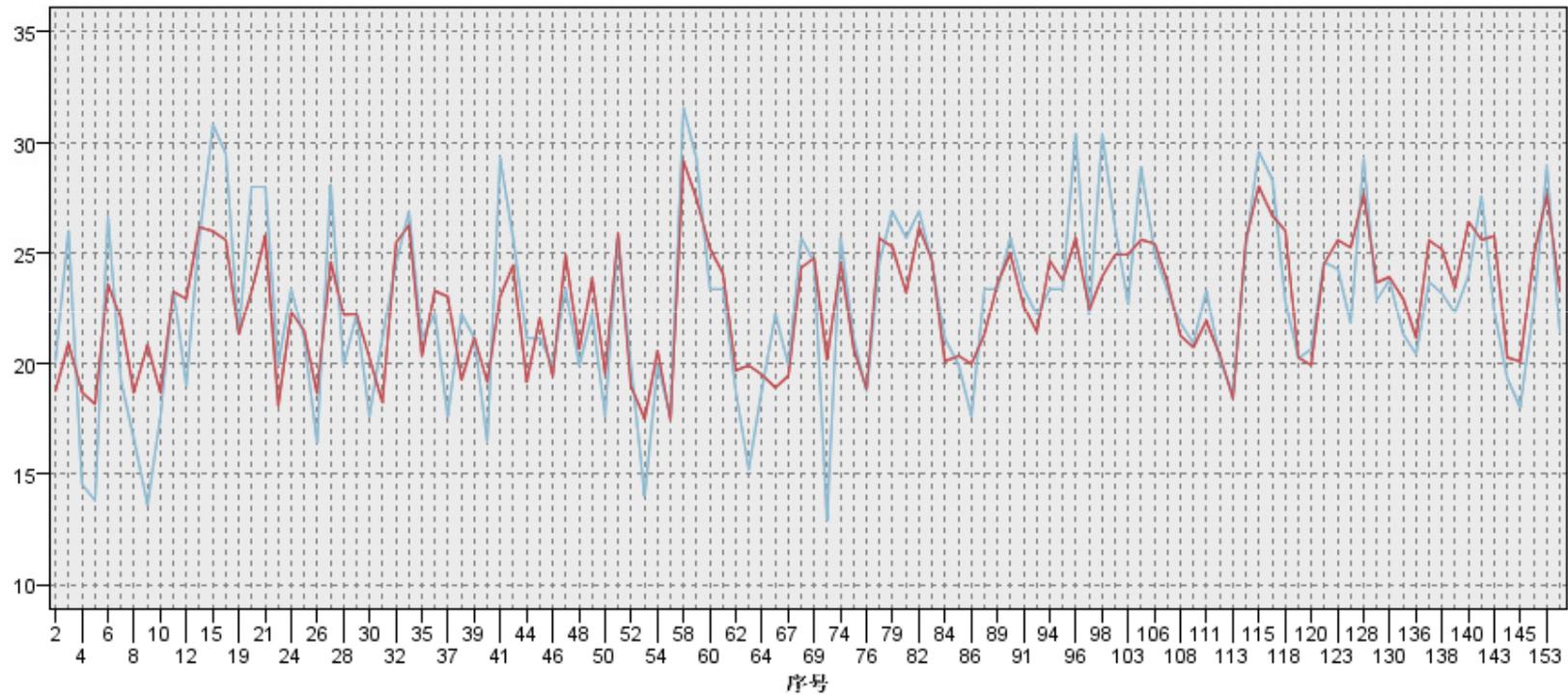




中华人民共和国水利部  
The Ministry of Water Resources of the People's Republic of China



# 土壤含水率预测趋势图



10cm土壤含水率: 预测平均误差百分比为8.2





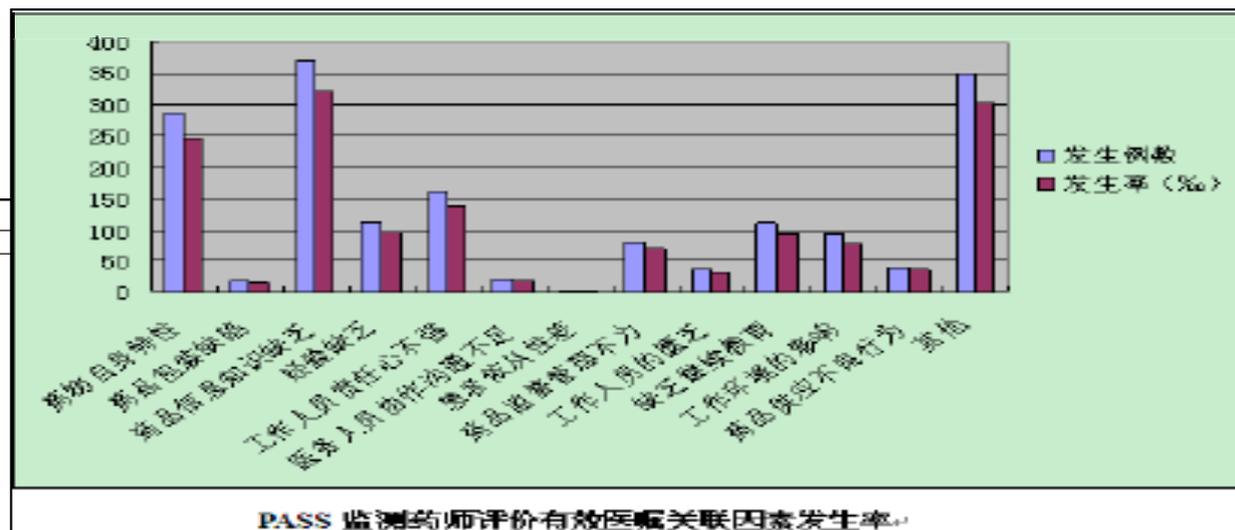
## 成分药评价分析

根据药品不良反应上报评价分析的数据，从药品成分药的角度对相关问题药品进行分析对比，

## OTC 转换评价分析

对药品在OTC 转换中进行不良反应历史比较和分析。通过与其他智能搜索工具相结合，及时发现、评价、转换和更新我国OTC 药品目录，并对相关药品做出准确、清晰评价。

各医院科室不合理用药发生率 <sup>1</sup>			
医院科室 <sup>2</sup>	总病例数 <sup>3</sup>	不合理用药例数 <sup>4</sup>	例次 <sup>5</sup>
<b>解放军总医院<sup>6</sup></b>			
骨科 <sup>7</sup>	544 <sup>8</sup>	137 <sup>9</sup>	200 <sup>10</sup>
呼吸科 <sup>11</sup>	537 <sup>12</sup>	182 <sup>13</sup>	298 <sup>14</sup>
普外科 <sup>15</sup>	423 <sup>16</sup>	126 <sup>17</sup>	148 <sup>18</sup>
心内科 <sup>19</sup>	378 <sup>20</sup>	112 <sup>21</sup>	255 <sup>22</sup>
合计 <sup>23</sup>	1882 <sup>24</sup>	557 <sup>25</sup>	901 <sup>26</sup>
<b>沈阳军区总医院<sup>27</sup></b>			
骨科 <sup>28</sup>	559 <sup>29</sup>	101 <sup>30</sup>	101 <sup>31</sup>
呼吸科 <sup>32</sup>	383 <sup>33</sup>	9 <sup>34</sup>	16 <sup>35</sup>
普外科 <sup>36</sup>	432 <sup>37</sup>	26 <sup>38</sup>	47 <sup>39</sup>
心内科 <sup>40</sup>	621 <sup>41</sup>	33 <sup>42</sup>	50 <sup>43</sup>
合计 <sup>44</sup>	1995 <sup>45</sup>	169 <sup>46</sup>	214 <sup>47</sup>
<b>成都军区总医院<sup>48</sup></b>			
骨科 <sup>49</sup>	180 <sup>50</sup>	61 <sup>51</sup>	108 <sup>52</sup>
呼吸科 <sup>53</sup>	203 <sup>54</sup>	126 <sup>55</sup>	336 <sup>56</sup>
普外科 <sup>57</sup>	358 <sup>58</sup>	180 <sup>59</sup>	416 <sup>60</sup>
心内科 <sup>61</sup>	292 <sup>62</sup>	214 <sup>63</sup>	619 <sup>64</sup>
合计 <sup>65</sup>	1033 <sup>66</sup>	581 <sup>67</sup>	1479 <sup>68</sup>





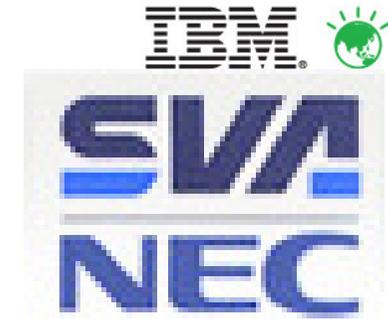
# 美国海军设备主动维护

- 故障数据被收集并被保存，包括：
  - 叙述性数据（由维护人员按流程描述）
  - 结构化数据（例如:舰船类型、故障日期、停工期和**维护成本**）
- 准确的预测哪些因素可能导致频繁的故障发生
- 识别哪些因素可能导致成本增加以及系统生命周期中过长的停工期
- 快速挖掘数以千计描述系统维护状况的日志并确定哪些是维护团队需要关注的重点





# 精细生产



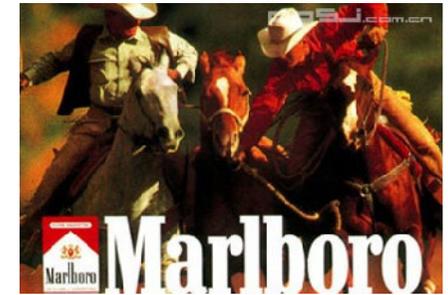
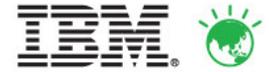
对于制造商来说，产品质量是企业生存的根本。由于生产过程的各个环节都可能对产品最终质量产生影响，而且生产过程中的各个环节性能参数总处在随机波动状态，因此质量管理人员很难使用常用的报表统计方式了解生产过程对产品质量的影响，从而优化生产过程，提高最终成品质量。

过去,SVA-NEC一直专注于工艺改良来提高产品合格率。随着工艺的不断改良，SVA-NEC生产部门发现使用该方法已无法明显改善产品合格率。因此，SVA-NEC希望通过数据分析来帮助生产部门优化生产过程参数配置，从而明显提高产品合格率。



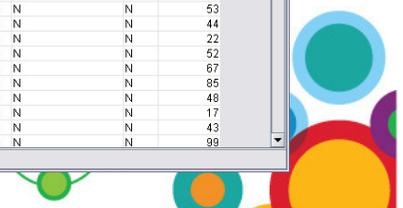


# 门店流量的预测



- 预测每个门店的客流量
- 预测结果实时发布给门店经理
- 门店经理根据预测值优化人员分配
- 建立自动化的系统，使人工操作将到最低
- 取代基于Excel宏编写的预测功能
- 解放业务分析人员，为企业提供更专业、更有价值的分析

	Week	Day	Hour	DayoftheWeek	Sales (\$)	Holidays	Season	Marketing Campaign	Event	Foot Traffic
459	4/3/2005	4/3/2005	16	Sunday	14560	Y	Spring	N	N	101
460	4/3/2005	4/3/2005	18	Sunday	12273	Y	Spring	N	N	86
461	4/3/2005	4/4/2005	10	Monday	7357	Y	Spring	N	N	53
462	4/3/2005	4/4/2005	12	Monday	4514	Y	Spring	N	N	34
463	4/3/2005	4/4/2005	14	Monday	5483	Y	Spring	N	N	41
464	4/3/2005	4/4/2005	16	Monday	12086	Y	Spring	N	N	85
465	4/3/2005	4/4/2005	18	Monday	14861	Y	Spring	N	N	103
466	4/3/2005	4/5/2005	10	Tuesday	11596	Y	Spring	N	N	81
467	4/3/2005	4/5/2005	12	Tuesday	7594	N	Spring	N	N	55
468	4/3/2005	4/5/2005	14	Tuesday	9202	N	Spring	N	N	65
469	4/3/2005	4/5/2005	16	Tuesday	12388	N	Spring	N	N	87
470	4/3/2005	4/5/2005	18	Tuesday	7366	N	Spring	N	N	53
471	4/3/2005	4/6/2005	10	Wednesday	6023	N	Spring	N	N	44
472	4/3/2005	4/6/2005	12	Wednesday	2754	N	Spring	N	N	22
473	4/3/2005	4/6/2005	14	Wednesday	7234	N	Spring	N	N	52
474	4/3/2005	4/6/2005	16	Wednesday	9375	N	Spring	N	N	67
475	4/3/2005	4/6/2005	18	Wednesday	12097	N	Spring	N	N	85
476	4/3/2005	4/7/2005	10	Thursday	6536	Y	Spring	N	N	48
477	4/3/2005	4/7/2005	12	Thursday	1915	Y	Spring	N	N	17
478	4/3/2005	4/7/2005	14	Thursday	5622	Y	Spring	N	N	43
479	4/3/2005	4/7/2005	16	Thursday	14220	Y	Spring	N	N	99





# 污水溢出和污染事故



- **面临挑战:**

减少“污水溢出及其他事故”（DG5定义的所有相关事故）的数量。

- **问题:**

DG5类事故类维修通常都是被动等待居民举报的。每个突发性事故都会带来大笔支出——包括工程师出勤时间、设备替换或维修、清理、法定/保险支出。DG5在OPA中也是一个基本的指标。

- **SPSS能力:**

通过一个预测分析应用，对历史和实时数据进行分析，预测最有可能出现污水溢出（DG5）的位置，从而使管理人员能在污水溢出事故发生前采取各类预防措施予以避免。

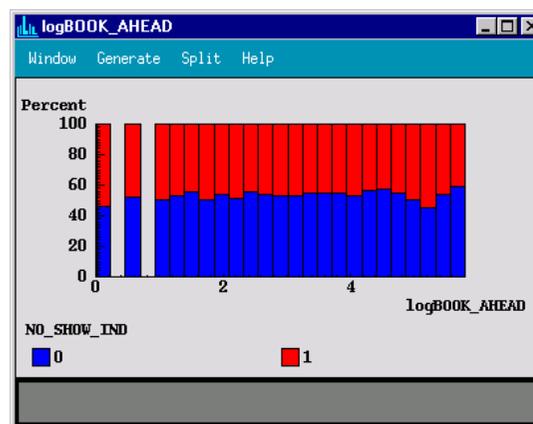
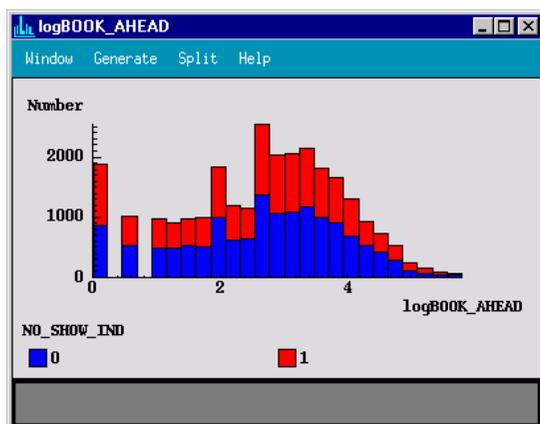
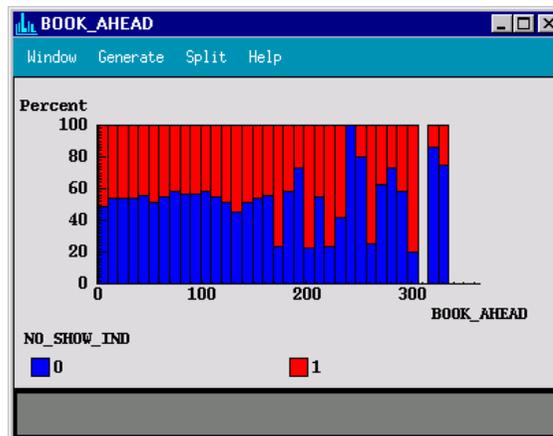
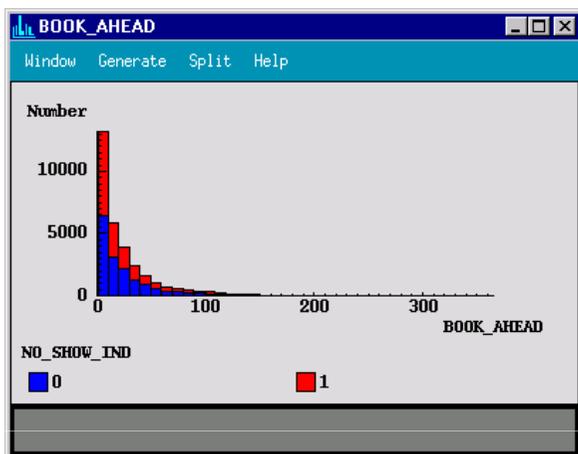
- **收益:**

通过预测DG5类溢流事故最可能发生的地点，管理人员可以在问题发生前安排工程师或检查员进行相应的检查和维修。通过这种主动管理的模式，管理人员可以更有效的利用工程师资源，使居民更少受到污水溢出等事件的困扰、了解设备更新评估的模式，最终减少DG5类事故的数量，提高OPA等级评定——对管理层和股东非常重要。





# 评估顾客放弃预订座位的概率





HSBC  汇丰 环球金融 地方智慧

- 背景描述  
美国汇丰银行是HSBC 集团成员之一，通过其380多个分支机构为140多万银行客户提供核算、投资、借贷和其它金融服务。
- 面临的问题  
同一地区可能有多家银行设有分支机构，在吸引和保持潜在客户方面存在着持续的竞争压力。为了能够保持高水平的客户获取、降低客户流失率，并维持服务的可赢利性，银行经常要实现以下几项目标：
  - 扩展与现有客户之间的关系
  - 控制营销费用、降低成本以维持利润水平
  - 用新的智能决策快速转移市场关注力
- 解决方案  
应用SPSS 对不断增长的客户数据进行挖掘分析，建立预测模型以发现交叉销售机会。定位每个产品的最有价值客户来使销售利润最大化、营销费用最小化。
- 结果
  - ① 了解特定客户的需求，销售增加50%
  - ② 营销费用减少30%
  - ③ 提高了建立和开展适时营销战略的能力





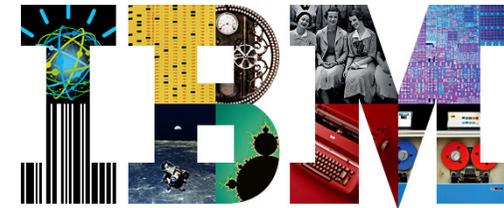
## 基金销售举例

等级分类	最低分数	最高分数	总客户数	购买基金 客户数	占总数比	占总数比 积累	购买基金 率	K-S
一级（最高）	898	995	8,288	70	45.8%	45.8%	0.84%	35.8%
二级	783	897	8,289	32	20.9%	66.7%	0.39%	46.8%
三级	698	782	8,289	14	9.2%	75.8%	0.17%	45.9%
四级	550	697	8,289	16	10.5%	86.3%	0.19%	46.4%
五级	486	549	8,289	5	3.3%	89.5%	0.06%	39.6%
六级	419	485	8,289	4	2.6%	92.2%	0.05%	32.2%
七级	329	418	8,289	3	2.0%	94.1%	0.04%	24.2%
八级	269	328	8,289	4	2.6%	96.7%	0.05%	16.8%
九级	150	268	8,289	4	2.6%	99.3%	0.05%	9.4%
十级（最高）	57	149	8,288	1	0.7%	100.0%	0.01%	0.0%
<b>合计</b>	<b>898</b>	<b>149</b>	<b>82,888</b>	<b>153</b>	<b>100.0%</b>		<b>0.18%</b>	





Thank  
YOU



© Copyright IBM Corporation 2008. All rights reserved. The information contained in these materials is provided for informational purposes only, and is provided AS IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, these materials. Nothing contained in these materials is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software. References in these materials to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. Product release dates and/or capabilities referenced in these materials may change at any time at IBM's sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way. IBM, the IBM logo, Cognos, the Cognos logo, and other IBM products and services are trademarks of the International Business Machines Corporation, in the United States, other countries or both. Other company, product, or service names may be trademarks or service marks of others.

您的信息 您的智慧 2011 IBM 信息管理与业务分析论坛

