# 实验一 了解各类商务智能工具平台

# 一、实验目的

1. 了解IBM Cognos；
2. 了解IBM spss statistics；

# 二、实验步骤

## （1）了解IBM Cognos

Cognos是在BI核心平台之上，以服务为导向进行架构的一种数据模型，是唯一可以通过单一产品和在单一可靠架构上提供完整业务智能功能的解决方案。它可以提供无缝密合的报表、分析、记分卡、仪表盘等解决方案，通过提供所有的系统和资料资源，以简化公司各员工处理资讯的方法。作为一个全面、灵活的产品，Cognos业务智能解决方案可以容易地整合到现有的多系统和数据源架构中。

### 简介

Cognos展现的报表基于统一的元数据模型。统一的元数据模型为应用提供了统一、一致的视图。用户可以在浏览器中自定义报表，格式灵活，元素丰富，而且可以通过Query Studio进行及时的开放式查询。Cognos还具有独特的穿透钻取（Drill Through）、切片（slice）和切块（dice）、以及旋转（pivot）等功能，使分析人员、管理人员或执行人员能够从多角度对信息进行快速、一致、交互地存取，从而获得对数据的更深入了解，有效地将各种相关的信息关联起来，使用户在分析汇总数据的同时能够深入到自己感兴趣的细节数据中，以便更全面地了解情况，做出正确决策。

Cognos强大的报表制作和展示功能能够制作/展示任何形式的报表，其纯粹的Web界面使用方式又使得部署成本和管理成本降到最低。同时Cognos还可以同数据挖掘工具、统计分析工具配合使用，增强决策分析功能。

2007年11月美国波士顿当地时间本周一，IBM宣布将以50亿美元现金收购Cognos。故又称IBM Cognos

### 组件功能

从大模块来看， Cognos产品组件只有三个：

Cognos Powerplay Transformation Server：负责将数据源变成数据立方体；

Cognos Powerplay Enterprise Server：负责将数据立方体以OLAP分析、OLAP报表等方式展现出来；

Cognos ReportNet Server：负责实现基于数据库的数据查询、报表制作、仪表盘制作、报表/仪表盘展示等；

每个Cognos组件的功能：

**Cognos Powerplay Transformation Server部分：**

Cognos Impromptu：主要用来连接数据库，形成数据源定义（IQD文件），Transformer会根据数据源定义到源数据库中抽取数据；

Cognos Transformer：在Windows界面中提供一个图形化的模型设计界面，供开发人员设计模型和调试模型；在UNIX版本中，这个模块名称为 Cognos Powerplay Transformer UNIX Client，增加了设计界面对服务器的控制菜单；

Cognos Transformation Server：后台的OLAP数据抽取转换引擎，用来把源数据抽取出来形成数据立方体。在Windows版本中，它与前端设计界面Cognos Transformer是结合在一起的；在UNIX版本中，它与前端的模型设计界面是分离的，安装在UNIX环境中，接受命令行或来自于客户端的调度。

**Cognos Powerplay Enterprise Server部分：**

PowerConnect：Cognos Powerplay Enterprise Server与第三方数据立方体的连接接口，通过PowerConnect，可以将MSOLAP、Essbase、IBM DB2 OLAP、SAP R/3 BW等第三方数据立方体通过Cognos展现出来；

Powerplay Enterprise Server：负责通过Web方式或Client方式展现数据立方体、提供OLAP界面和报表界面的服务器产品，是整个Cognos OLAP应用的核心；

Cognos Upfront：Cognos Powerplay Enterprise Server提供的门户界面，可以定制外观、功能。

**Cognos ReportNet Server部分：**

Cognos ReportNet Server：Cognos 负责查询、报表、仪表板的设计、制作以及展现的核心服务器模块；

Cognos Framework Manager：负责定义Cognos ReportNet Server使用的元数据，包括定义表结构、表连接、全局过滤条件、虚拟视图、业务映射等等；

Cognos Connection：Cognos ReportNet Server附送的门户界面，功能丰富，允许用户自定义内容、布局、外观、功能等等。

**Cognos Access Manager 部分：**

Access Manager：Cognos全部产品的安全性管理模块，是Cognos安全性贯穿始终的基础，是Cognos安全性的核心模块；

LDAP Server：一般使用Cognos随产品附送的 Netscape/iPlanet Directory Server，或由第三方提供，用来保存Cognos安全性信息。

**Cognos Powerplay Client 部分：**

Cognos Powerplay Client：Cognos Powerplay Enterprise Server的Windows客户端，是强大的OLAP报表制作工具，用来制作报表，并将报表发布到 Cognos Powerplay Enterprise Server上。

## （2）了解IBM spss statistics

### 简介

SPSS（Statistical Product and Service Solutions）是世界上最早的统计分析软件，由美国斯坦福大学的三位研究生Norman H. Nie、C. Hadlai (Tex) Hull 和 Dale H. Bent于1968年研究开发成功。中文名为：“统计产品与服务解决方案”。SPSS为IBM公司推出的一系列用于统计学分析运算、数据挖掘、预测分析和决策支持任务的软件产品及相关服务的总称，有Windows和Mac OS X等版本。

2009年7月28日，IBM公司宣布将用12亿美元现金收购统计分析软件提供商SPSS公司。如今SPSS的最新版本为26，并且更名为IBM SPSS Statistics。迄今，SPSS公司已有40余年的成长历史。

使用SPSS Statistics，可以：

通过用户友好的界面分析并更好地理解您的数据，并解决复杂的业务和研究问题。

借助高级统计程序，可以更快地了解大型和复杂的数据集，从而有助于确保高精度和高质量的决策制定。

使用扩展，Python和R编程语言代码与开源软件集成。

借助灵活的部署选项，可以更轻松地选择和管理软件。

### 功能介绍

**1.数据管理**

在10版以后，SPSS的每个新增版本都会对数据管理功能作一些改进，以使用户的使用更为方便。13版中的改进可能主要有以下几个方面：

1）超长变量名：在12版中，变量名已经最多可以为64个字符长度，13版中可能还要大大放宽这一限制，以达到对当今各种复杂数据仓库更好的兼容性。

2）改进的Autorecode过程：该过程将可以使用自动编码模版，从而用户可以按自定义的顺序，而不是默认的ASCII码顺序进行变量值的重编码。另外，Autorecode过程将可以同时对多个变量进行重编码，以提高分析效率。

3）改进的日期/时间函数：本次的改进将集中在使得两个日期/时间差值的计算，以及对日期变量值的增减更为容易上。

**2.结果报告**

从10版起，对数据和结果的图表呈现功能一直是SPSS改进的重点。在16版中，SPSS推出了全新的常规图功能，报表功能也达到了比较完善的地步。13版将针对使用中出现的一些问题，以及用户的需求对图表功能作进一步的改善。

**1）统计图：**在经过一年的使用后，新的常规图操作界面已基本完善，本次的改进除使得操作更为便捷外，还突出了两个重点。首先在常规图中引入更多的交互图功能，如图组（Paneled charts），带误差线的分类图形如误差线条图和线图，三维效果的简单、堆积和分段饼图等。其次是引入几种新的图形，已知的有人口金字塔和点密度图两种。

**2）统计表：**几乎全部过程的输出都将会弃用文本，改为更美观的枢轴表。而且枢轴表的表现和易用性会得到进一步的提高，并加入了一些新的功能，如可以对统计量进行排序、在表格中合并/省略若干小类的输出等。此外，枢轴表将可以被直接导出到PowerPoint中，这些无疑都方便了用户的使用。

**3.统计建模**

Complex Samples是12版中新增的模块，用于实现复杂抽样的设计方案，以及对相应的数据进行描述。但当时并未提供统计建模功能。在13版中，这将会有很大的改观。一般线形模型将会被完整地引入复杂抽样模块中，以实现对复杂抽样研究中各种连续性变量的建模预测功能，例如对市场调研中的客户满意度数据进行建模。对于分类数据，Logistic回归则将会被系统的引入。这样，对于一个任意复杂的抽样研究，如多阶段分层整群抽样，或者更复杂的PPS抽样，研究者都可以在该模块中轻松的实现从抽样设计、统计描述到复杂统计建模以发现影响因素的整个分析过程，方差分析模型、线形回归模型、Logistic回归模型等复杂的统计模型都可以加以使用，而操作方式将会和完全随机抽样数据的分析操作没有什么差别。可以预见，该模块的推出将会大大促进国内对复杂抽样时统计推断模型的正确应用。

**4.模块**

这个模块实际上就是将以前单独发行的SPSS AnswerTree软件整合进了SPSS平台。笔者几年前在自己的网站上介绍SPSS 11的新功能时，曾经很尖锐地指出SPSS的产品线过于分散，应当把各种功能较单一的小软件，如AnswerTree、Sample Power等整合到SPSS等几个平台上去。看来SPSS公司也意识到了这一点，而AnswerTree就是在此背景下第一个被彻底整合的产品。

Classification Tree模块基于数据挖掘中发展起来的树结构模型对分类变量或连续变量进行预测，可以方便、快速的对样本进行细分，而不需要用户有太多的统计专业知识。在市场细分和数据挖掘中有较广泛的应用。已知该模块提供了CHAID、Exhaustive CHAID和C&RT三种算法，在AnswerTree中提供的QUEST算法尚不能肯定是否会被纳入。

为了方便新老用户的使用，Tree模块在操作方式上不再使用AnswerTree中的向导方式，而是SPSS近两年开始采用的交互式选项卡对话框。但是，整个选项卡界面的内容实际上是和原先的向导基本一致的，另外，模型的结果输出仍然是AnswerTree中标准的树形图，这使得AnswerTree的老用户基本上不需要专门的学习就能够懂得如何使用该模块。

由于树结构模型的方法体系和传统的统计方法完全不同，贸然引入可能会引起读者统计方法体系的混乱。为此，本次编写的高级教程并未介绍该模块，而将在高级教程的下一个版本，以及关于市场细分问题的教材中对其加以详细介绍。

**5.兼容性**

随着自身产品线的不断完善，SPSS公司的产品体系已经日益完整，而不同产品间的互补和兼容性也在不断加以改进。在13版中，SPSS软件已经可以和其他一些最新的产品很好的整合在一起，形成更为完整的解决方案。例如，SPSS、SPSS Data Entry和新发布的SPSS Text Analysis for Surveys一起就形成了对调查研究的完整解决方案。而新增的SPSS Classification Trees模块将使得SPSS软件本身就能够针对市场细分工作提供更为完整的方法体系。

# 三、实验总结

本次实验后，大致了解了IBM Cognos和IBM spss statistics的发展与功能，对这两个商务智能工具平台有了进一步理解。