



Visualizer介绍

议程

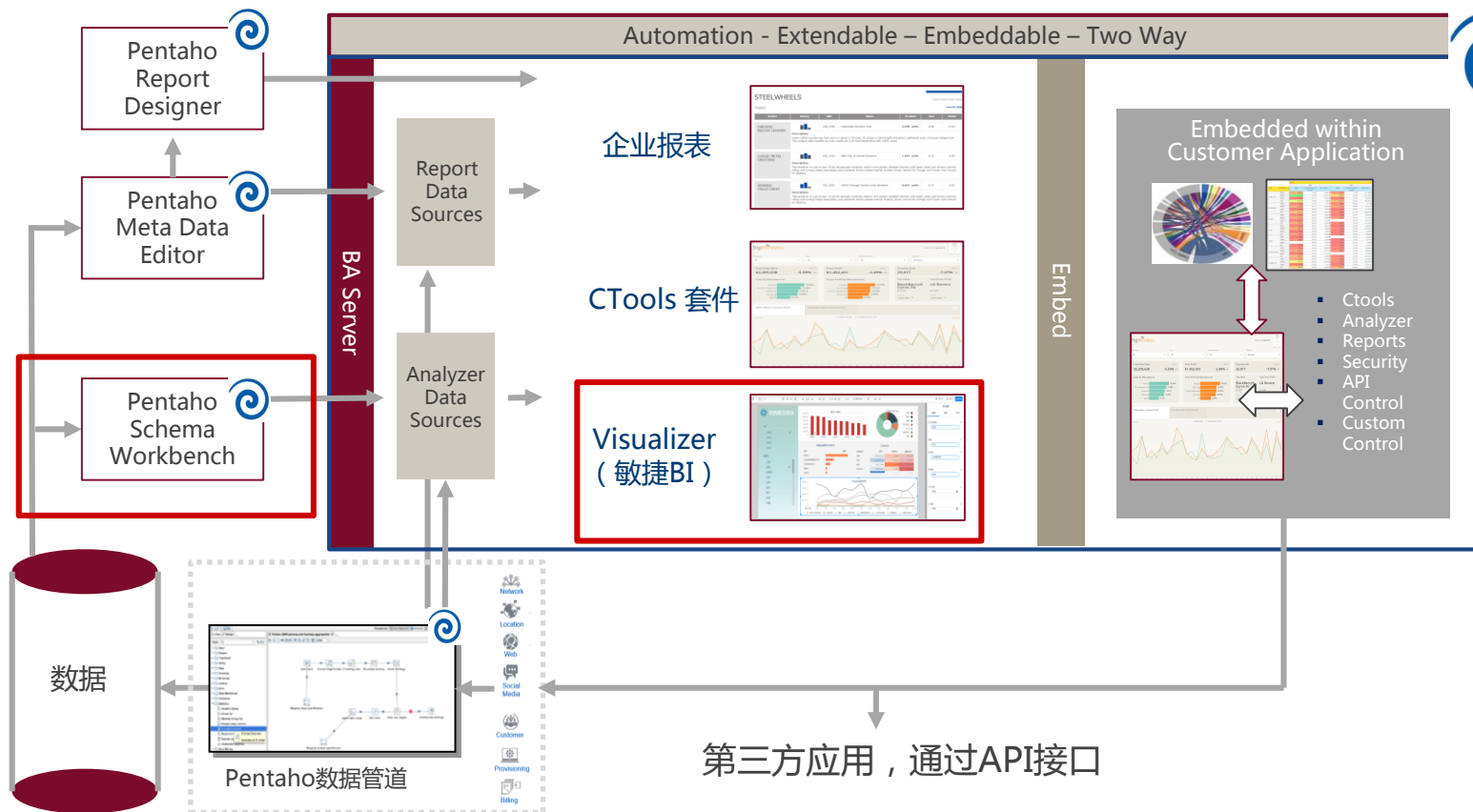
✓ 概述

- 多维建模介绍
- 建模练习
- Visualizer的介绍
- Visualizer练习
- Visualizer技巧
- 问与答

数据可视化与分析的趋势

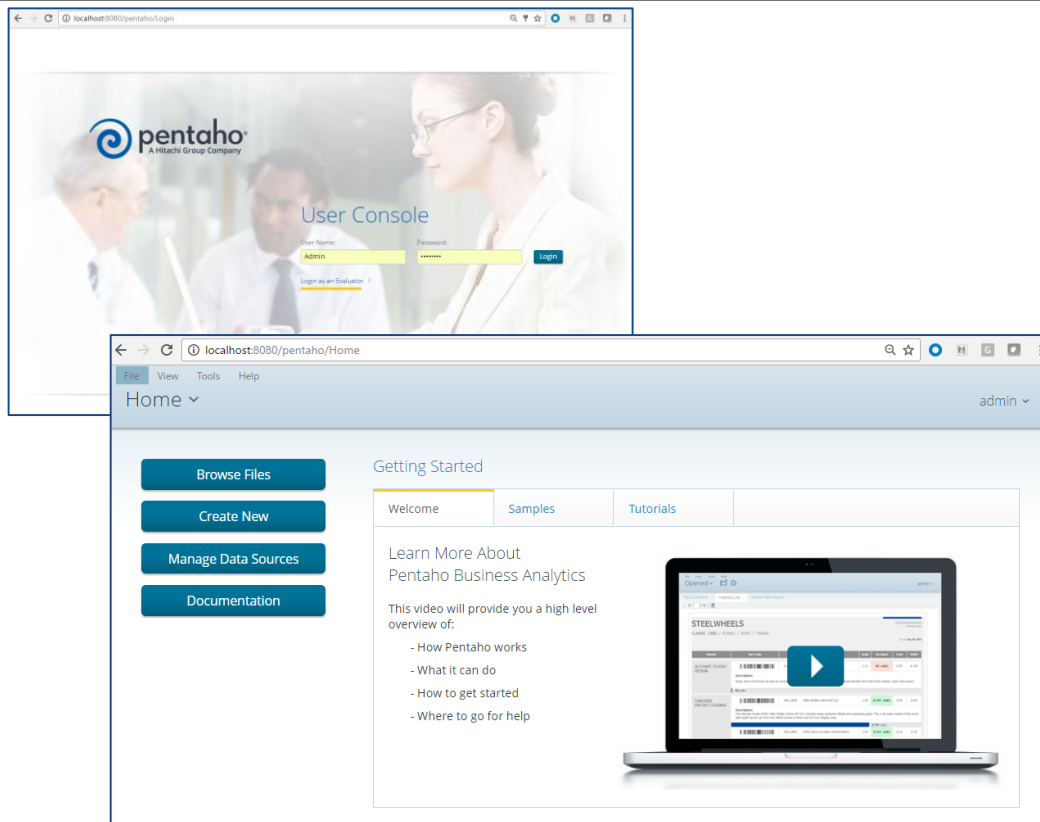


Pentaho Business Analytics(PBA)



Pentaho控制台

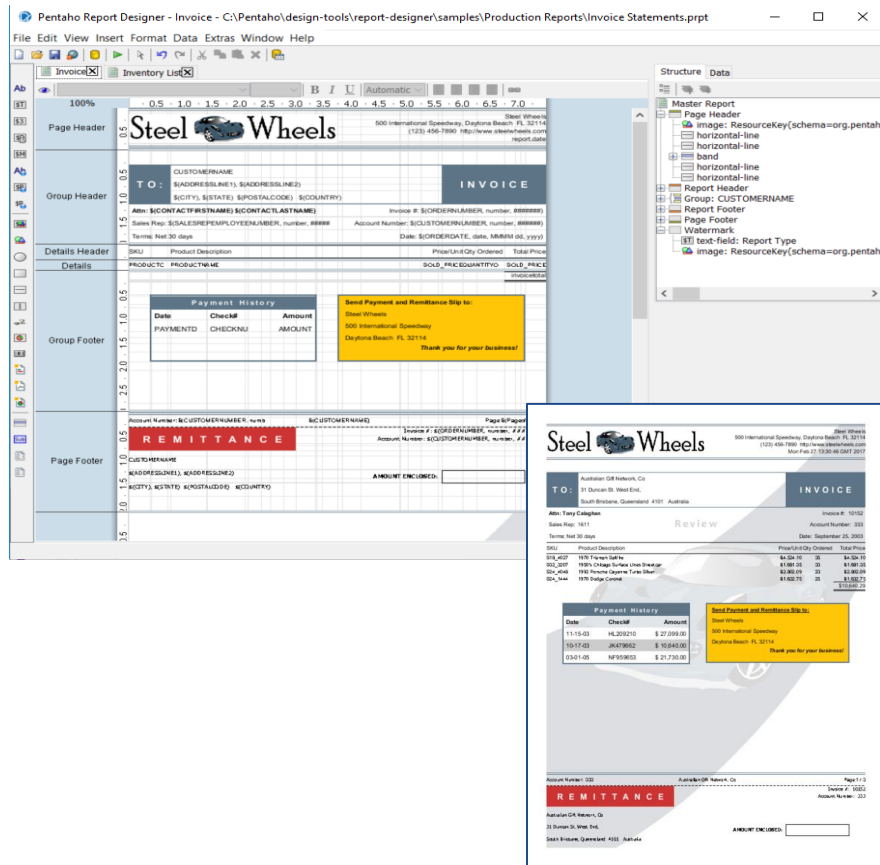
- 用户和角色管理
- 文件管理
- 数据源管理
- 工具入口
 - Ctools套件
 - Visualizer
- 计划报表



Pentaho Report Designer – 企业报表开发工具

复杂的企业报表开发

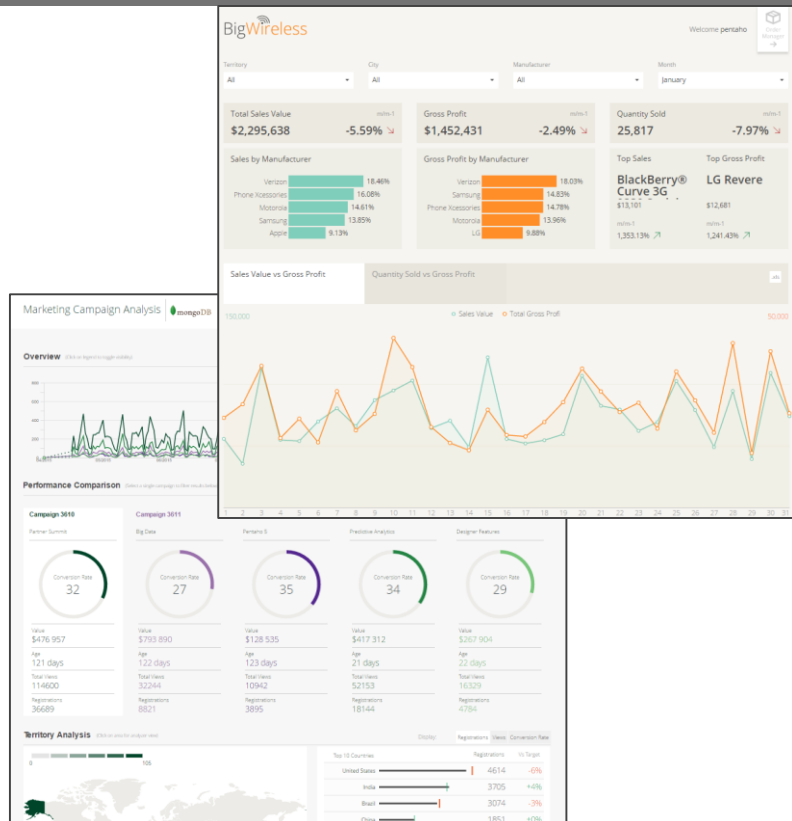
- 复杂格式
- 嵌套报表
- 可嵌入可视化图表
- 导出HTML, Excel, PDF, Text 或打印
- 直接连接PDI/Kettle任务



Ctools开发套件 — 适合IT的分析应用开发工具

CTools可以创建**复杂的、高度定制化的**仪表板或分析应用，特点如下：

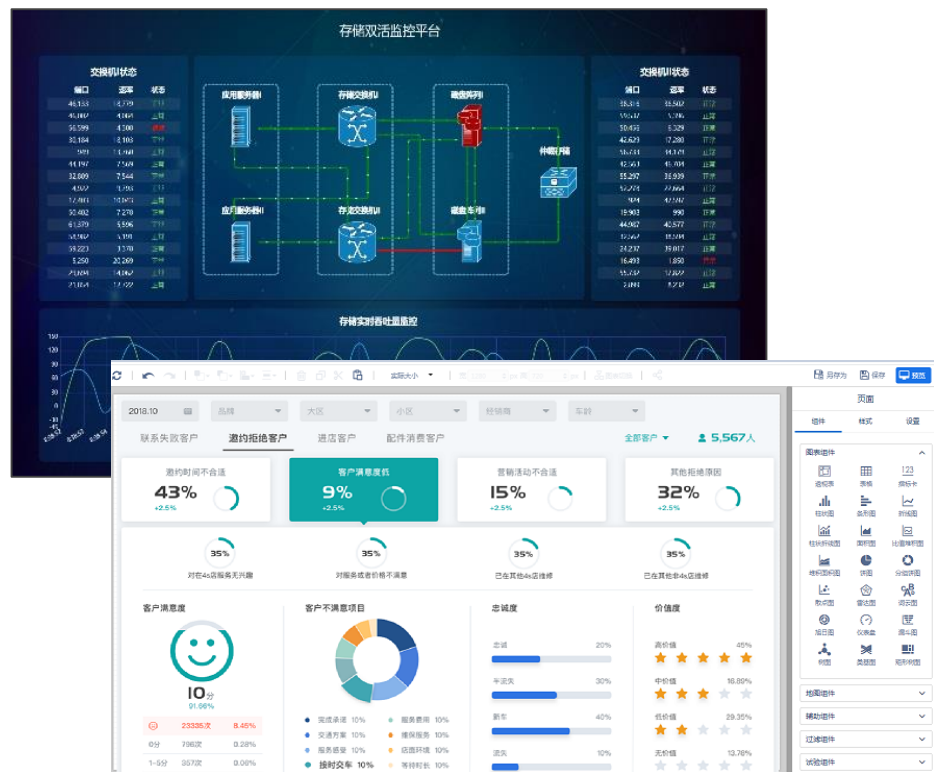
- 功能丰富、包括CDE、CDF、CDA等功能模块
- 参数传递框架，监听机制，简化了开发难度
- 组件丰富，可扩展
- 支持的数据源非常丰富
- 提供双向通信，通过PDI返回数据源，或通过API提供给其他应用程序。
- 通过Endpoint调用，无需后台编码，就能开发应用插件



Visualizer – 适合业务人员使用的BI工具

自助式敏捷BI工具

- 拖拽式设计
- 灵活的布局
- 丰富的数据可视化组件
- 个性化样式设置
- 多维分析和高级计算
- 嵌入和分享



议程

□ 概述

✓ 多维建模介绍

□ 建模练习

□ Visualizer的介绍

□ Visualizer练习

□ Visualizer技巧

□ 问与答

多维立方体

多维立方体 (Cube):用于支持联机分析应用 (OLAP),支持多角度多层次进行数据查询和分析。

- **维度**

维度提供了分类描述,表示一类分析角度,用户通过维度分析度量数据。比如上图中的三个维度:时间、产品、地域

- **度量 (Measures)**:

度量表示用来聚合分析的数值信息,度量的集合组合成了一个特殊的维度。如数量、销售额等。

- **层次 (Hierarchy)**

- **级别 (Level)**:

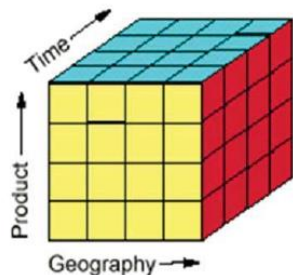
一个维度上可以包含的层次结构,表示特定的分类。如上图地域维度可以包含的级别层次级:国家、省、市;时间维度包含的级别层次包含:年、季度、月、日等。

- **成员 (Member)**:

成员是维度 (包括度量<Measures>) 上的项目值。如上图时间维度上”年”级别的成员就包含:2000,2001,2002,2003...月成员包含:1、2、3等。

- **计算成员 (Calculated member/measure)**:

表示式动态计算成员。通过各种数学表达式和各种函数定义,可以创建复杂的表达式。任何动态分析功能,都可以通过计算成员实现,比如实现占比,同期比等等。



名词解释:

- 共享维度
- 虚拟模型

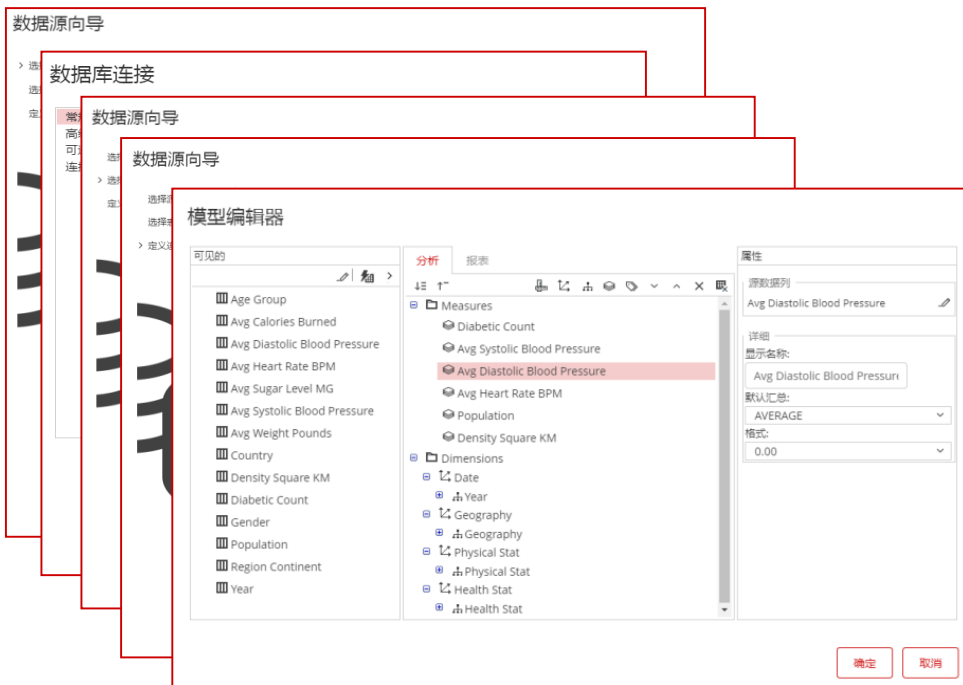
<https://mondrian.pentaho.com/documentation/schema.php>

建模工具

- Data Source Wizard
 - 向导式建模工具，Web版
 - 业务人员
- Schema workbench
 - 客户端工具，桌面版
 - IT人员
- Pentaho Data Refinery
 - 组件，集成在PDI/Kettle中
 - IT人员

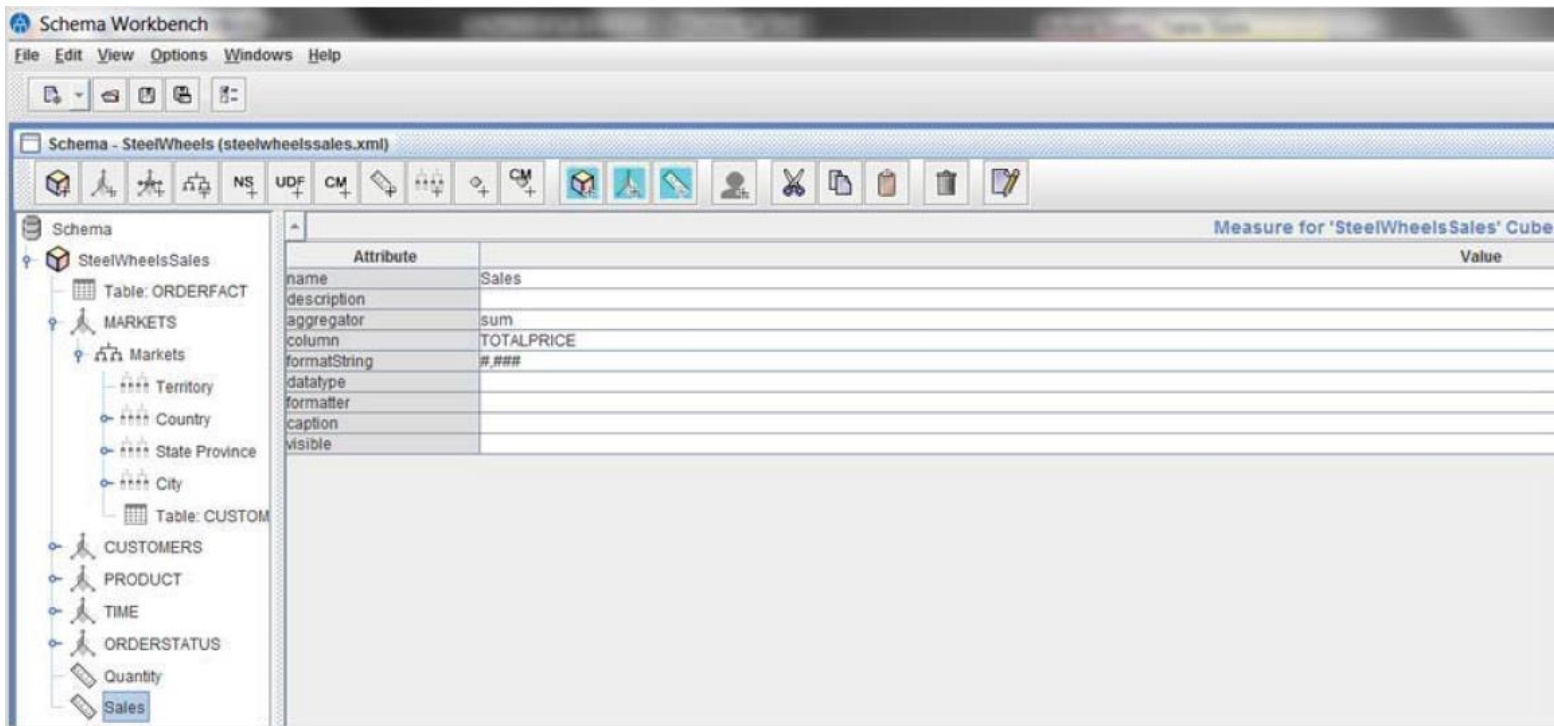
建模 : Data Source Wizard

1. 选择数据源类型
2. 选择 (新建) 连接
3. 选择数据表
4. 建立表关系
5. 创建模型



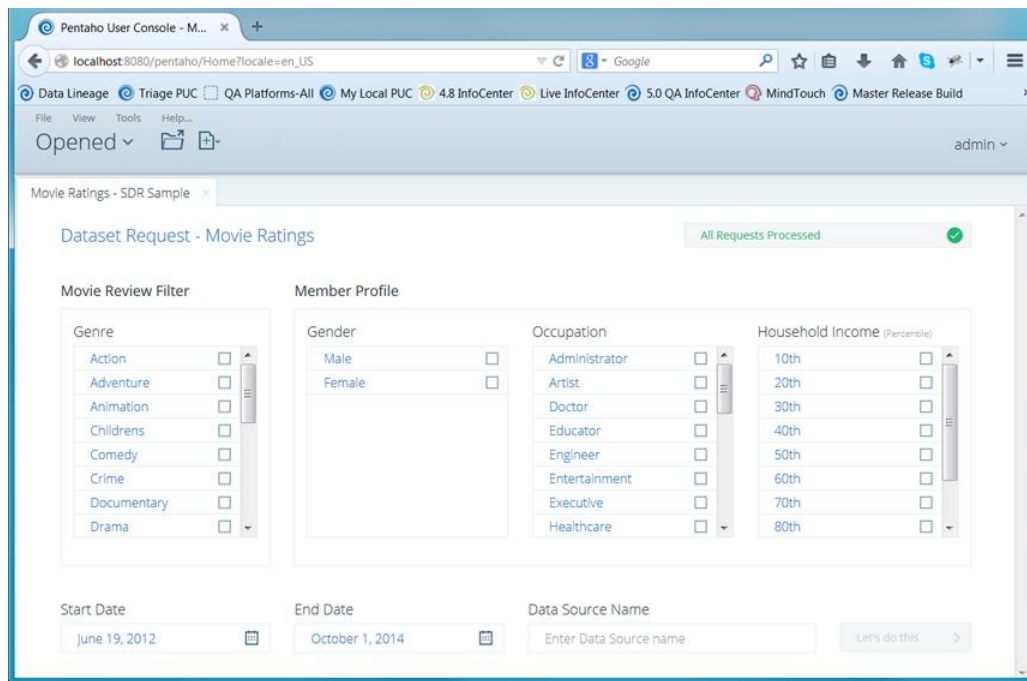
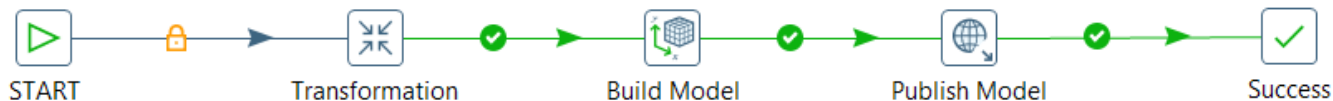
建模：Schema Workbench

- Schema Workbench is a visual application for building and testing Mondrian schemas.



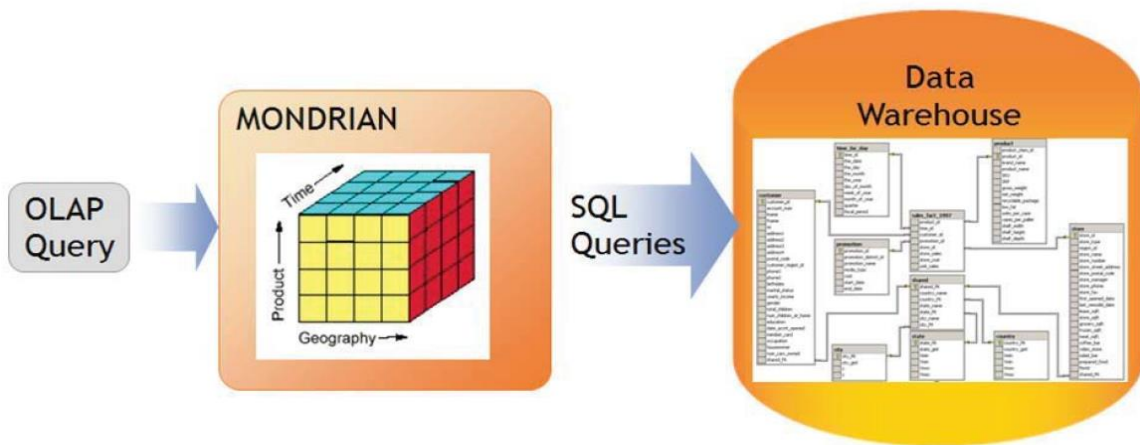
建模：Pentaho Data Refinery

https://help.pentaho.com/Documentation/9.0/Work_with_data/Use_the_Streamlined_Data_Refinery



Mondrian (OLAP引擎)

- Mondrian是一个在线分析处理（OLAP）引擎，使业务用户能够实时分析大量数据。
 - 支持对接大多数关系型数据库，和多个大数据平台诸如sap hana，kylin等。
 - 节省计算资源，保证高并发量压力下大数据量的可用性。
 - 超过250多个配置项，使引擎的运行状态更加贴合各种复杂的数据和应用场景。
 - 超过300多个内置分析计算函数和自定义分析函数扩展能力。
 - 支持多级权限管理，满足数据权限的精细控制。



Mondrian的性能

- 数据源自身的查询性能
 - 改变设计，避免创建高基数维度
 - 选择分析型数据库
 - 为大型事实表创建索引
 - 为频繁计算的视图创建聚合表
- 为大型数据仓库配置合适的Mondrian参数

例如：

- `mondrian.result.limit=5000000`
- `mondrian.rolap.iterationLimit=5000000`
- `mondrian.rolap.queryTimeout=300`
- `mondrian.native.crossjoin.enable=true`
- `mondrian.native.topcount.enable=true`
- `mondrian.native.filter.enable=true`

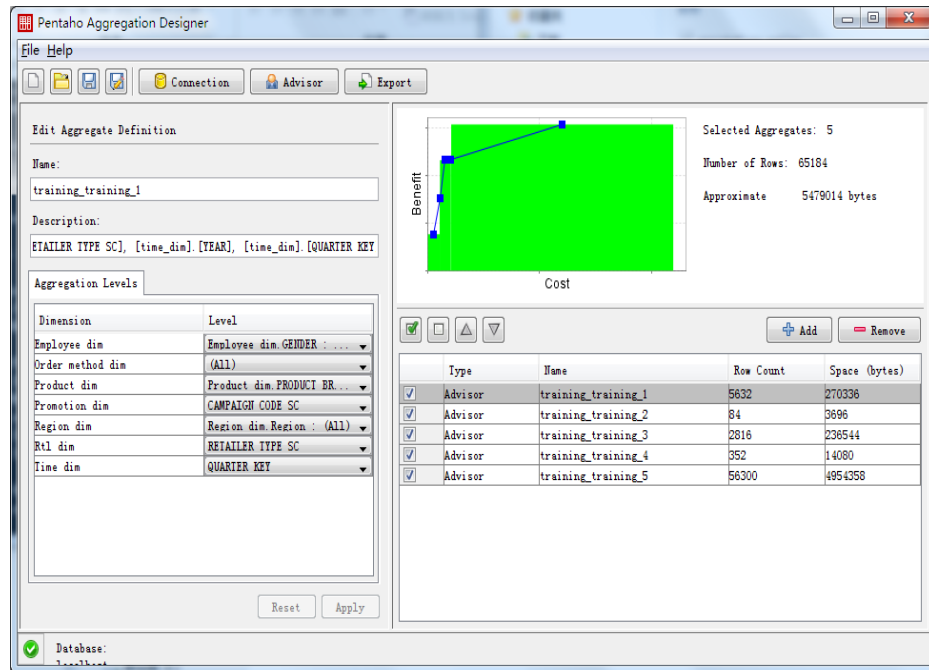
<https://mondrian.pentaho.com/documentation/configuration.php>

- Mondrian缓存控制
 - 例如：切换到Memcached（分布式高速缓存系统）
https://help.pentaho.com/Documentation/9.0/Work_with_data/Mondrian_cache_control

Pentaho Aggregation Designer – 空间换时间

简化了汇总表的创建和部署

- 创建聚合表
- 更新Mondrian Schema , 支持聚合表



数据权限

https://help.pentaho.com/Documentation/9.0/Work_with_data/Analysis_schema_security

```
01 <Role name="California manager">
02   <SchemaGrant access="none">
03     <CubeGrant cube="Sales" access="all">
04       <HierarchyGrant hierarchy="[Store]" access="custom" topLevel="[Store].[Store Country]">
05         <MemberGrant member="[Store].[USA].[CA]" access="all"/>
06         <MemberGrant member="[Store].[USA].[CA].[Los Angeles]" access="none"/>
07       </HierarchyGrant>
08       <HierarchyGrant hierarchy="[Customers]" access="custom" topLevel="[Customers].[State Province]" bottomLevel="[Customers].
[City]">
09         <MemberGrant member="[Customers].[USA].[CA]" access="all"/>
10         <MemberGrant member="[Customers].[USA].[CA].[Los Angeles]" access="none"/>
11       </HierarchyGrant>
12       <HierarchyGrant hierarchy="[Gender]" access="none"/>
13     </CubeGrant>
14   </SchemaGrant>
15 </Role>
```

数据权限（续）

配置文件：pentaho-solutions/system/pentahoObjects.spring.xml

- The Mondrian One-To-One UserRoleMapper

```
1 <bean id="Mondrian-UserRoleMapper"
2   name="Mondrian-One-To-One-UserRoleMapper"
3   class="org.pentaho.platform.plugin.action.mondrian.mapper.MondrianOneToOneUserRoleListMapper"
4   scope="singleton" />
```

- The Mondrian-SampleLookupMap-UserRoleMapper

```
01 <bean id="Mondrian-UserRoleMapper"
02   name="Mondrian-SampleLookupMap-UserRoleMapper"
03   class="org.pentaho.platform.plugin.action.mondrian.mapper.
04   MondrianLookupMapUserRoleListMapper"
05   scope="singleton">
06   <property name="lookupMap">
07     <map>
08       <entry key="ceo" value="California manager" />
09       <entry key="cto" value="M_CTO" />
10       <entry key="dev" value="M_DEV" />
11     </map>
12   </property>
13 </bean>
```

- The Mondrian-SampleUserSession-UserRoleMapper

```
1 <bean id="Mondrian-UserRoleMapper"
2   name="Mondrian-SampleUserSession-UserRoleMapper"
3   class="org.pentaho.platform.plugin.action.mondrian.mapper.
4   MondrianUserSessionUserRoleListMapper"
5   scope="singleton">
6   <property name="sessionProperty" value="MondrianUserRoles" />
7 </bean>
```

国际化

- 国际化

https://help.pentaho.com/Documentation/9.0/Products/Localization_and_internationalization_of_analysis_services

议程

- 概述

- 多维建模介绍

- ✓ 建模练习

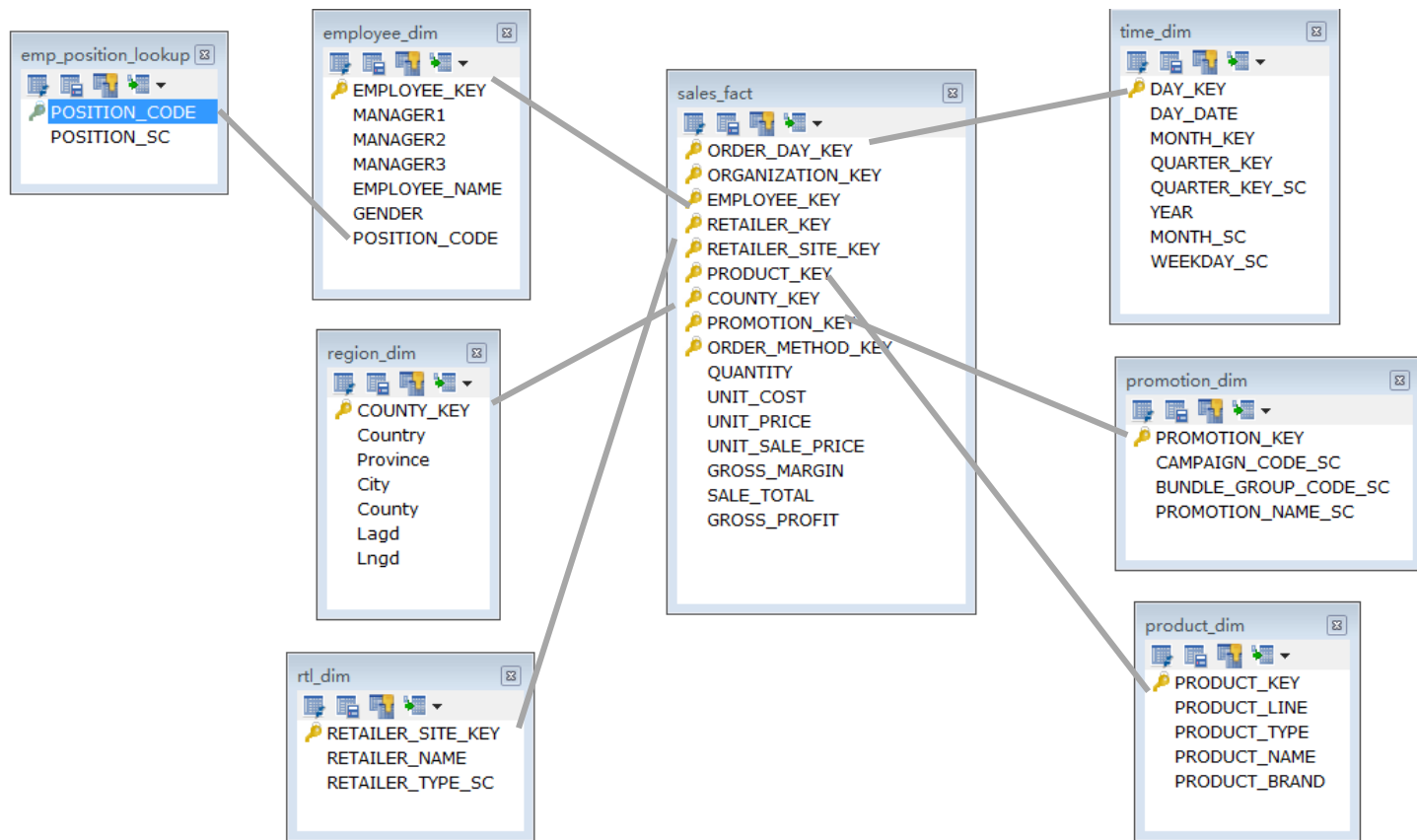
- Visualizer的介绍

- Visualizer练习

- Visualizer技巧

- 问与答

数据表



建模

议程

- 概述

- 多维建模介绍

- 建模练习

- ✓ Visualizer的介绍

- Visualizer练习

- Visualizer技巧

- 问与答

Visualizer组成

分享

Excel、CSV、PNG、PDF

Http/Https URL

Token URL

邮件

设计与交互

可视化元素

图表组件

筛选组件

参数
控制器

静态组件

功能组件

实时组件

数据

数据面板

过滤

排序

样式

颜色

尺寸

字体

格式

位置

形状

行为

定时
刷新

事件

交互

分析

筛选

跳转

联动

下钻
上卷

参考

计算

多维模型定义

图表扩展

国际化

审计

数据查询

BI服务 (PBA Services)

内容管理 | 安全 | Mondrian引擎 | Data Service | PDI引擎 | 单点登录 | 计划 | 邮件 | 审计

应用场景

- 自助式日常数据分析
- 报表、仪表盘、数字化报告
- 数据监控
- 数字大屏展示
- 嵌入式分析



敏捷

从数据到价值分享，只需4步

数据连接

- 文本
- 关系型数据库
- Data service
- SDR
- SparkSQL
- Kylin
- Impala
- Hadoop Hive
-



建立模型

- 多维模型
- 维度、层次、层级
- 雪花和星型
- 计算指标
- 权限控制



制作文档

- 拖拽设计
- 样式自定义
- 布局与层叠
- 页面自适应
- 数据联动
- 数据下钻
- 过滤筛选
- 排序与限制



嵌入与分享

- 分析应用嵌入
- 数字大屏
- 移动客户端
- 微信、钉钉等集成
- 数据导出

Visualizer



拖拽式设计数据文档

1. 选择可视化元素

图表组件



静态组件



功能组件

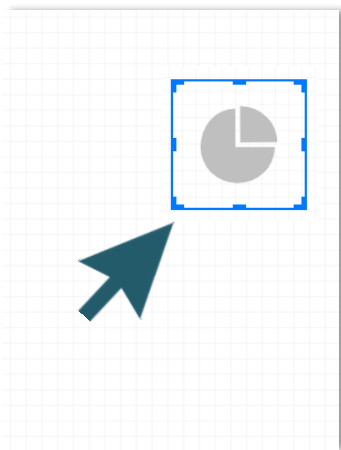


筛选数据

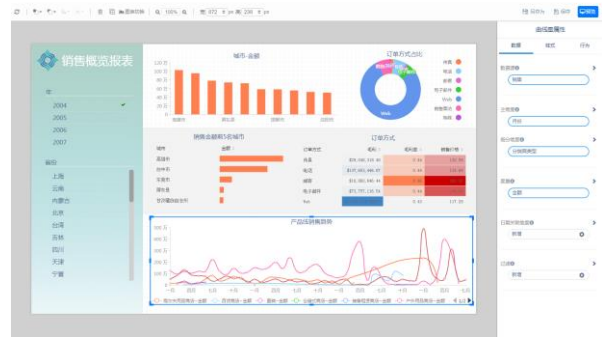


2. 在画布中拖拽元素、设计页面

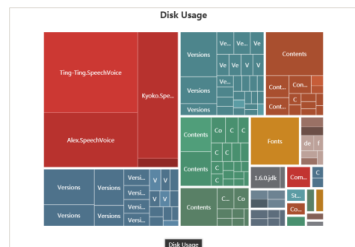
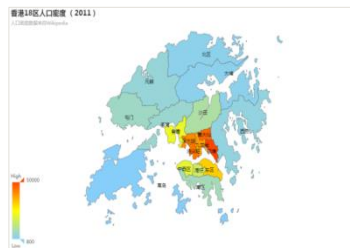
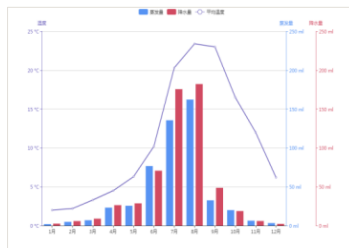
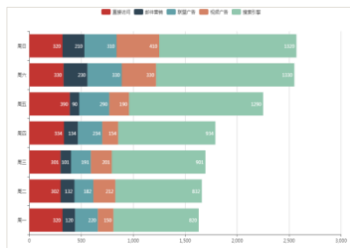
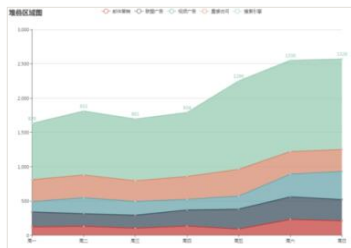
- 布局，位置、层叠等调整
- 绑定数据，添加筛选、排序等
- 调整样式，尺寸、大小、颜色等
- 设置组件行为



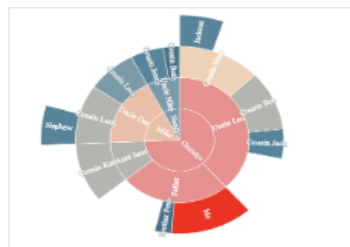
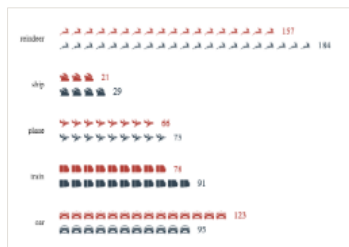
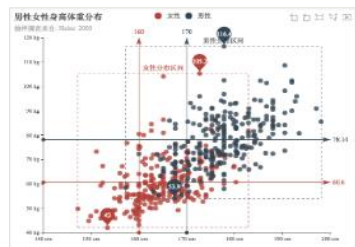
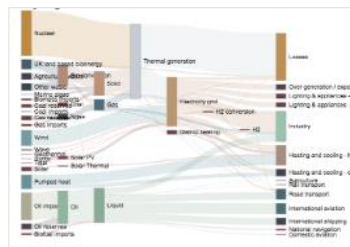
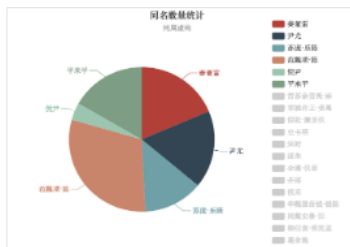
3. 设计出精美的“数据文档”



炫酷的图形组件（样例）



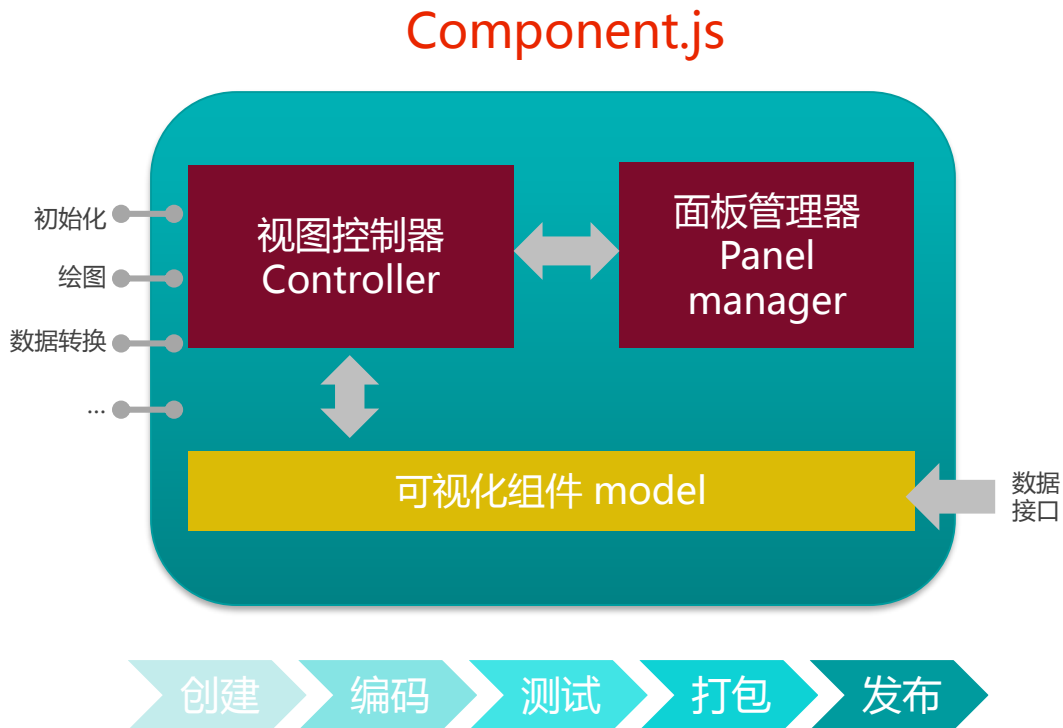
省份	金额	数量	毛利
上海	492,072.7	14,795.0	
云南	2,588,955.7	87,313.0	
内蒙古	2,384,636.5	59,078.0	
北京	187,157.6	4,365.0	
台湾	7,661,235.3	236,116.0	
吉林	1,351,799.7	23,459.0	
四川	4,061,146.1	73,367.0	
天津	351,532.5	8,954.0	
宁夏	384,285.3	12,783.0	
安徽	2,228,425.3	39,580.0	



创建自己的可视化元素

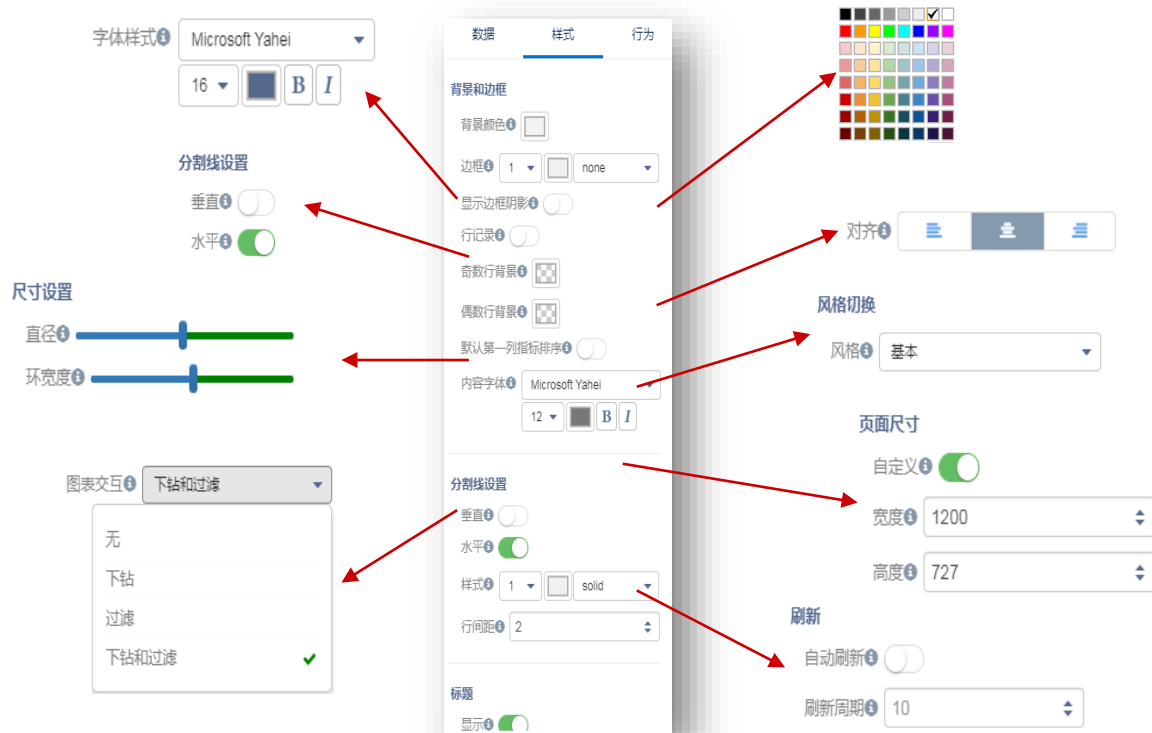
开发人员可使用自定义图表扩展 SDK 创建自定义可视化元素

- 图表
 - D3、Highchart、Fusioncharts 等
 - 3d效果three.js等
- 地图组件
 - Google地图、矢量地图、商圈地图等
- 动态矢量图
 - IOT设备、线路等

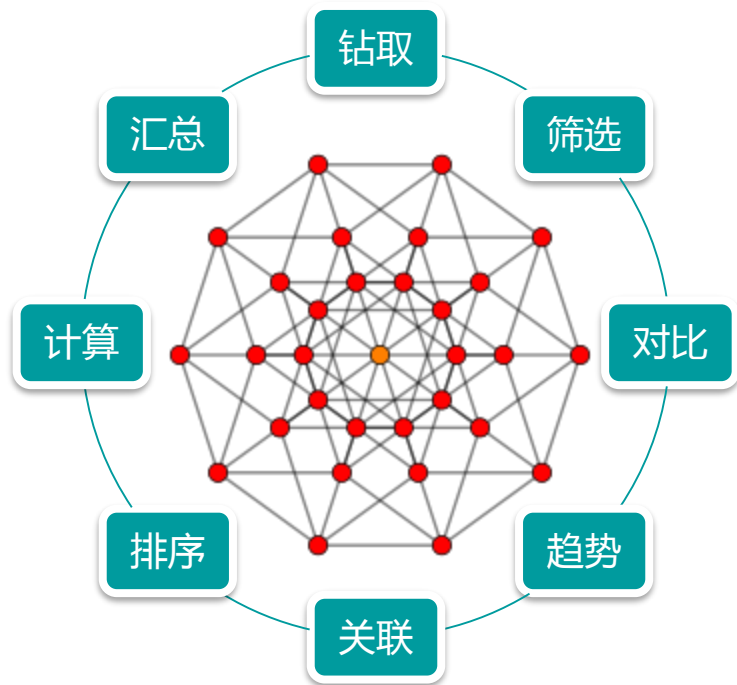


个性化样式

- 样式模板
- 全局样式、页面样式、组件样式
- 根据各个图表的不同特性设计的样式调整方案
- 颜色、大小、位置、形状、格式、字体等个性化元素



探索式分析方法



钻取：改变维的层次，变换分析的粒度。

筛选：组件级、页面级，筛选方式包括：列表选择、范围、通配比较、相对比较等。

对比：长度、大小、面积的对比。

趋势：通过趋势分析，了解变化。

关联：发现存在于大量数据集中的关联性或相关性。

排序：将“无序”的调整为“有序”。

计算：创建新的计算指标和计算成员。

汇总：包括累加、平均、计数、最小、最大等统计。

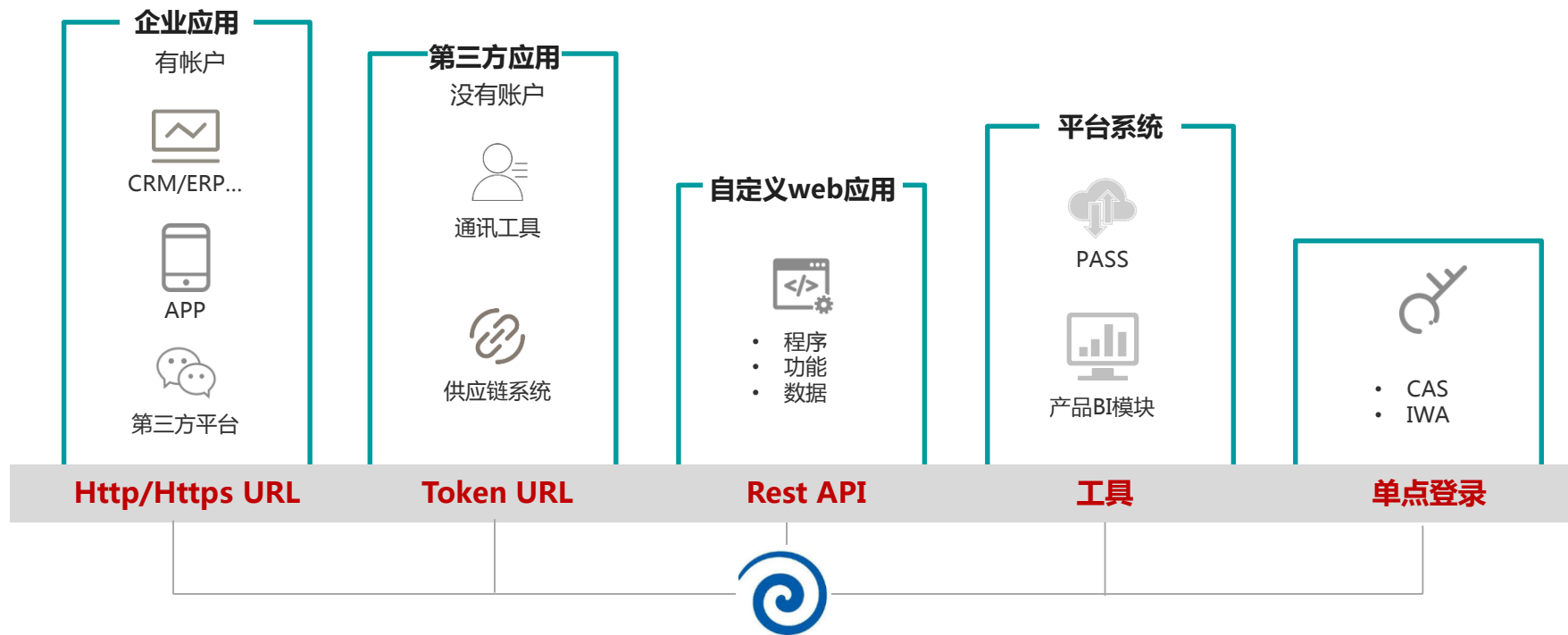
多屏应用

- 支持自适应布局、流式布局
- 桌面、平板、手机大屏，完美适应各种屏幕；



嵌入式 - 让分析融入所有业务环节

嵌入式分析通过赋予使用者按需获取数据、自主设计报表、自主可视化数据分析的能力。企业内部也可以使用嵌入式分析来提高内部运营效率。



优势

1 数据集成与业务分析紧密结合。
集成Pentaho分析数据管道。

2 炫酷的可视化，可扩展。强大的
分析能力

3 使用范围广。满足数据湖分析、
实时监控、嵌入式分析、日常数
据分析报告制作。

4 简单易用、适合业务人员的数据
分析工具

议程

- 概述
- 多维建模介绍
- 建模练习
- Visualizer介绍
- ✓ Visualizer练习
- Visualizer技巧
- 问与答

Visualizer安装

- ① 将datafor.zip解压到“pentaho-solutions\system”目录
- ② 修改“pentaho-solutions\system”目录下的 ImportHandlerMimeTypeDefinitions.xml 文件，添加如下代码：

```
<MimeTypeDefinition mimeType="application/octet-stream" hidden="false">  
  <extension>datafor</extension>  
</MimeTypeDefinition>
```

- ③ 修改“pentaho-solutions\system”目录下的applicationContext-spring-security.xml文件

- 在<bean id="filterInvocationInterceptor"下面添加如下代码

```
<sec:intercept-url pattern="\A/content/datafor/*.*\Z" access="Anonymous,Authenticated" />
```

- 在<bean id="filterInvocationInterceptorForWS"下添加如下代码

```
<sec:intercept-url pattern="\A/plugin/datafor/api/*.*\Z" access="Anonymous,Authenticated" />
```

- ④ 重启 PBA

页面制作

销售概览

月份

200501

金额

¥ 95.0M

↑ 12.8% (环比)

↑ 30.6% (年同比)

毛利

¥ 38.6M

↑ 14.4% (环比)

↑ 31.3% (年同比)

成本

¥ 136.4M

↑ 4.7% (环比)

↑ 20.8% (年同比)

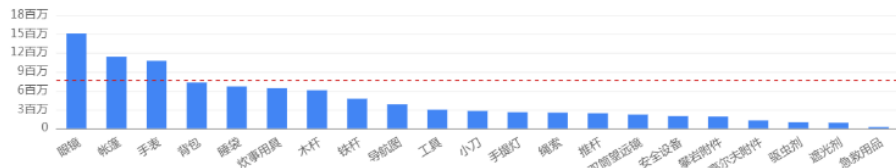
毛利率

43.0%

↑ 1.5% (环比)

↑ 0.4% (年同比)

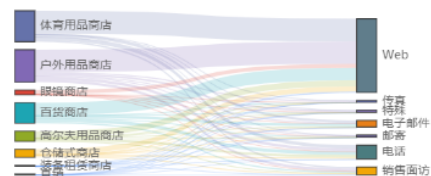
产品类型-金额



活动-金额

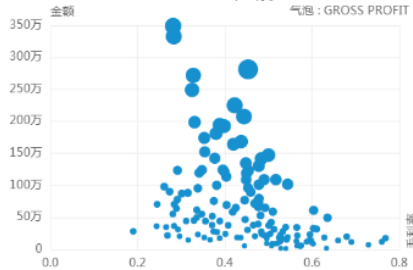


零售商类型-订单方式-金额

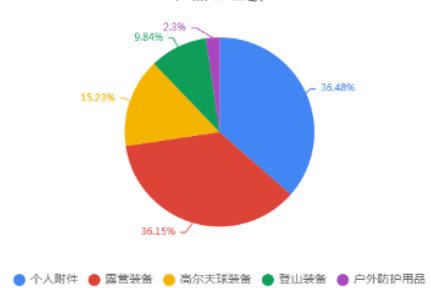


汇总 明细

产品名称-毛利率,金额,毛利



产品线-金额



议程

- 概述
- 多维建模介绍
- 建模练习
- Visualizer介绍
- Visualizer练习
- ✓ Visualizer技巧
- 问与答

Visualizer的技巧

- 技巧1：如何实现报表跳转、以及跳转传参
- 技巧2：如何实现高级分析（动态计算）
- 技巧3：如何扩展图表组件
- 技巧4：如何扩展矢量地图

如何实现报表跳转、以及跳转传参

- 跳转

- 从Training跳转到“全国销量分析页”

- 如何给页面跳转传参

- 结合全国销量分析页，调整链接参数，讲解参数传递

如何实现高级分析（动态计算）

■ 计算指标

- 由基础指标派生出新的指标。

例如，基于销售额和成本，派生出毛利

■ 计算成员

- 由基础维度成员派生新的成员。

例如，在省份维度上虚拟出新成员——其它，将指标值小于某个阈值的省份归入“其它”。

如何扩展图表组件

■ 步骤

- 步骤1. 按visualizer组件扩展规范，定义好组件描述文件
- 步骤2. 将描述文件放在插件extend components目录下
- 步骤3. 重启server

如何扩展矢量地图

■ 步骤

- 步骤1. 准备符合要求的geojson地图形状数据文件
- 步骤2. 将形状数据文件放在插件指定的路径
- 步骤3. 运行编译脚本，更新地图配置
- 步骤4. 重启server生效

谢谢