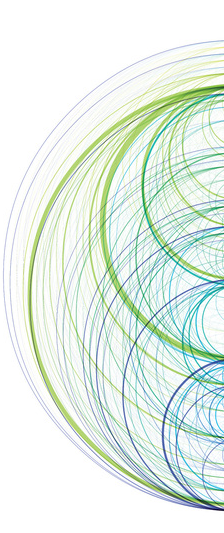


DataDiscovery

|  |
| --- |
| 大数据平台 可视化UE |
| **产品白皮书** |



成都四方伟业软件股份有限公司

Chengdu Sefon Software Co., LTD.

CHENGDU SEFON SOFTWARE CO.,LTD

**成都四方伟业软件股份有限公司**

# SDC UE产品总体介绍

## 产品架构

不同于传统的商业智能BI，SDC UE创新性地开辟了一条以数据可视为核心的行业方向，同时兼顾数据分析能力。

SDC可视化设计**User-interface Editor**（以下简称SDC UE）产品，是一款利用计算机图形学，图像处理技术以及大数据技术，通过简单拖拽、配置编排出可视化应用页面的企业级大数据可视化设计平台。

SDC UE平台整体架构是基于大数据技术架构进行数据处理，采用最新的客户端Web技术进行平台构建和数据可视化。

平台总体架构如图2-1所示，主要为四大模块。

SDC UE总体功能架构图



* 数据访问接入层

主要负责对各种数据源的接入和数据交换进行适配，包含RDBMS类、文本类、Hadoop/hive大数据平台以及其他JDBC数据源等。同时支持实时数据接口以及第三方Webservice、http、socket等接口数据。

* 数据分析处理层

采用大数据分布式实时查询引擎，并基于该引擎进行数据分析、数据处理的能力封装。通过数据集的定义，支持灵活的多维分析、即席查询，支持各种常见的数理统计函数和高级的表计算函数。并提供GIS Server引擎和各种基于geo标准的GIS应用服务接口。

* 数据可视化层

提供基于浏览器的图形化编排集成开发环境（IDE）和运行框架，支持可视化页面编排、数据配置、事件交互配置和样式效果配置。内置丰富的WEB交互组件、统计图表组件、关系图形组件、3D组件、GIS组件，可自由轻松设计出各种多样化的可视化场景，并支持在可视化运行框架的支持下，完成页面的解析、数据的加载和转换、事件的执行和图形的渲染等。

* 集成与管理层

负责整个可视化平台的所有基础管理、数据管理、作品管理、框架集成。

# 

# SDC UE产品能力特性

## 强大的数据接入和处理能力

### 多数据源支持

基于大数据多源异构数据源适配技术，SDC UE支持多种数据源的连接访问和数据交换，包括主流关系型 RDBMS、Excel/CSV文本数据源、基于Hadoop的大数据平台数据源以及其他多种JDBC数据源等，具体说明如表3-1所示。

数据源类型及说明表

|  |  |
| --- | --- |
| 类型 | 数据源描述 |
| 关系型数据源 | 主流关系型数据库支持：Oracle、MySQL、SQLServer、PostgraSQL、DB2等。  国产数据库：达梦、金仓、神通等。 |
| 大数据、分布式数据库 | 支持Hadoop/Hive、星环inceptor等，支持大数据量的分布式计算 |
| 文件数据源支持 | 可直接上传.xls .xlsx 或.csv 文件，并可读取其中的每个页签。 |
| 第三方数据接口 | 提供第三方API接口的接入能力，预留接口适配层。支持实时数据接口。 |

### 自助式数据建模

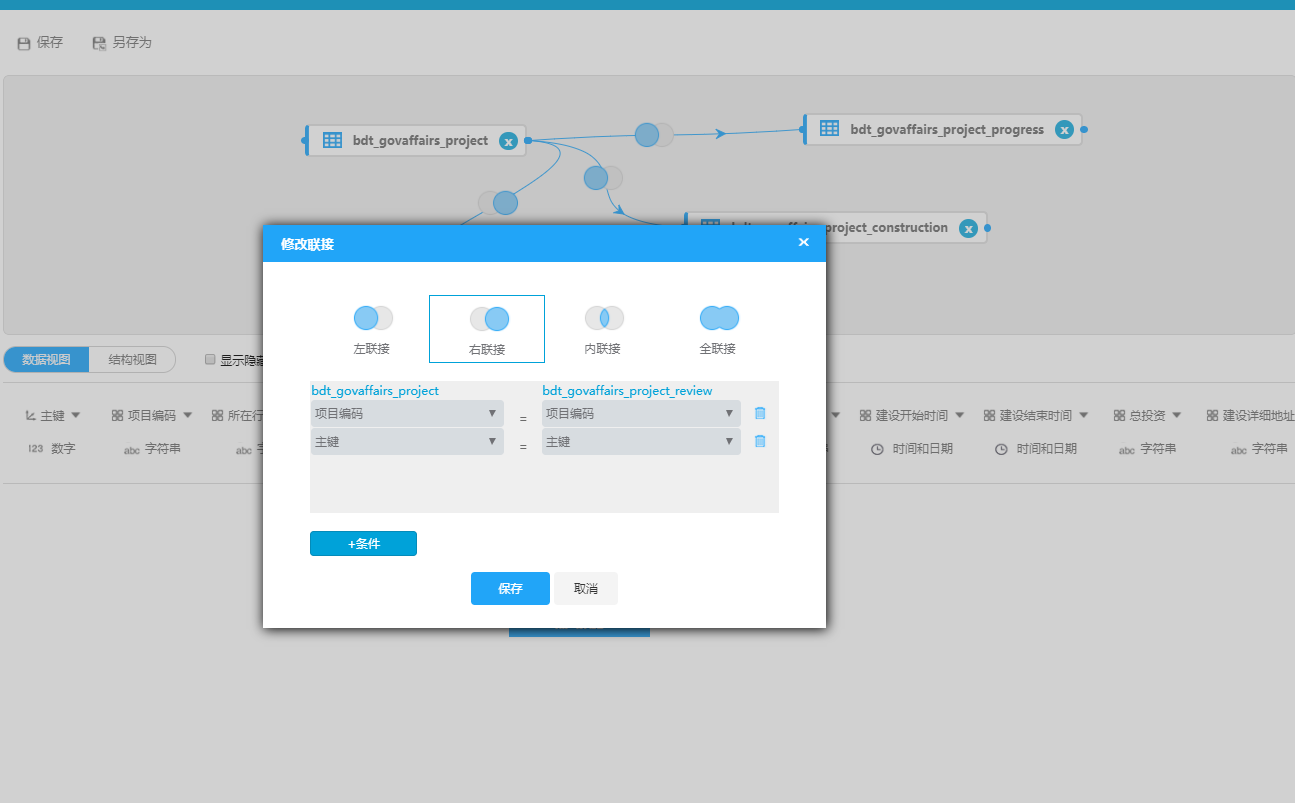
在可视化设计器的页面上，当数据源接入后，给组件绑定数据时，为方便用户快速简单地直接选择数据字段，需要将这些不同的数据源统一转换为单个独立的数据集。一个数据集中定义了多个有业务关联的数据字段，数据字段的类型分为维度（对数据进行分类的字段，比如时间、年份、地区、类别、名称等）、度量（含有定量数值的字段，表示某种数据的数值，比如销售量、利润值、商品库存数等）。

创建数据集的过程，定义为数据建模。SDC UE可以将一个简单的表做成单个的数据集，也支持多个表、多个视图、多个自定义SQL的逻辑视图进行相互关联。SDC UE不仅支持传统的基于数据仓库的星型模型、雪花模型，还可以根据客户现有的数据环境自定义数据模型。SDC UE支持两种数据集建模方式：

* 图形化建模

用户可以通过简单的可视化操作，将数据库中的数据表创建成待分析的数据模型，简单易用的操作界面，非技术人员也可以快速上手。表与表之间支持多个条件关联。

可视化建模



* 自定义SQL建模

用户可以通过自己编写SQL语句建立一张虚拟的表创建模型，支持SQL定义的虚拟表之间进行关联，也支持SQL定义表与数据表和数据视图关联建模。可视化建模如图3-2所示。

SQL建模



平台智能识别出维度和指标信息，并且支持对维度、指标的转换，如图3-3所示。

模型的维度、度量管理



这两种建模方式，并不会实际在用户的数据源环境中新建任何的表、视图，而是在本地建立了一个虚拟的数据集映射（数据模型），并且提供了数据安全机制，让用户只能访问自己受信任的数据模型，也可以通过授权方式共享给其他用户使用，既保证了数据的访问权限，也实现了开放共享。

### 聚合和表计算

在数据分析工作中，除了对普通数据的维度、度量进行直接分析和观察数据的整体效应外，分析平均数、最大值、最小值等需求也是非常常见的。如果汇总方式仅有单一的求和汇总，则很难满足这类需求。

在SDC UE中，可以针对所选的任意度量进行求和、求最值、求平均等，并且它们之间可以实时切换。

除此之外还可以基于当前所选的“数据视图”的计算结果进行行级数据运算，包括：差异计算、百分比计算、差异百分比计算、总额百分比计算、累计和计算、同比计算和环比计算。

表计算



### 自定义计算度量

度量（指标）可以来自于原始的数据字段，同时也可通过公式计算得来。

SDC UE支持创建自定义计算字段。这些新增字段是通过原有的一个或者多个数据字段，采用几十种普通的数学公式函数计算而来。比如，有一个销售总量字段、一个成本字段，那么在原本没有发送短信总字段的时候，现在就可以通过这两个字段相加而得到。

计算字段表达式



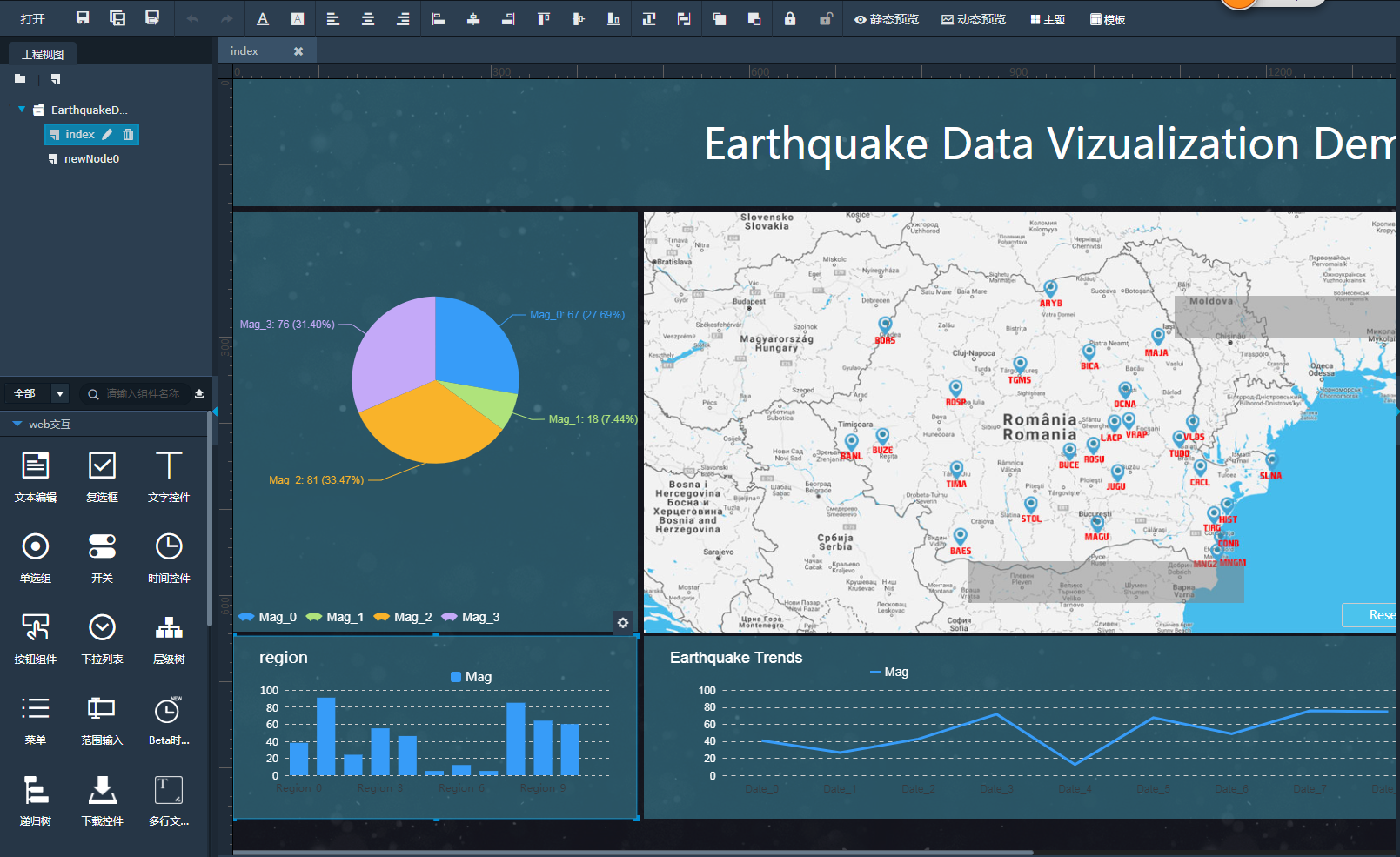
SDC UE内置多达30余种计算函数，包括方差、偏差等统计函数、截位、取整、大小写转换等字符操作函数、if-else、case-when、空值判断等逻辑函数、Max、Min等聚合函数、日期转换、四舍五入等格式转换函数。通过这些函数用户可以构建复杂的计算表达式，实现更加复杂的业务定义。

## 快速炫酷的可视化设计

SDC UE的设计器是基于纯Web前端代码开发的IDE，可以在各种浏览器下进行数据可视化设计，无需单独安装客户端。产品提供大量的可视化组件（多达100种以上），包括各种Web表单组件、图形组件，用户只需要通过简单的拖拽和参数配置，无需编程能力，在设计器上任意发挥创意，即可创造出专业的数据分析报表、dashboard仪表盘、数据可视化大屏以及可视化数据展现Web页面。

可视化编排页面





SDC UE的可视化设计器提供丰富的工具栏，灵活自如的画布区，画布区提供标尺及辅助对齐线，提供组件管理列表，简单方便地控制每个组件，支持拷贝、粘贴、上下左右移、回退、重做、保存、全部保存等快捷键。具有类PPT模式的编排体验，让页面设计轻而易举。

### 敏捷极致的可视化编排

设计器采用自由布局的模式，允许用户直接通过拖拽的方式，自定义组件的位置和大小，并且提供对齐辅助线和边线吸附功能，方便用户手动调整布局的过程中保持组件对齐。

SDC UE产品设计的页面，除了支持普通显示屏页面和个人电脑的分辨率以外，画布区的大小可以达到4w以上的超大分辨率，并支持页面的分辨率自适应能力，以确保编排的页面能够在一定范围内适应多种不同分辨率的显示终端。

### 丰富多彩的图表组件

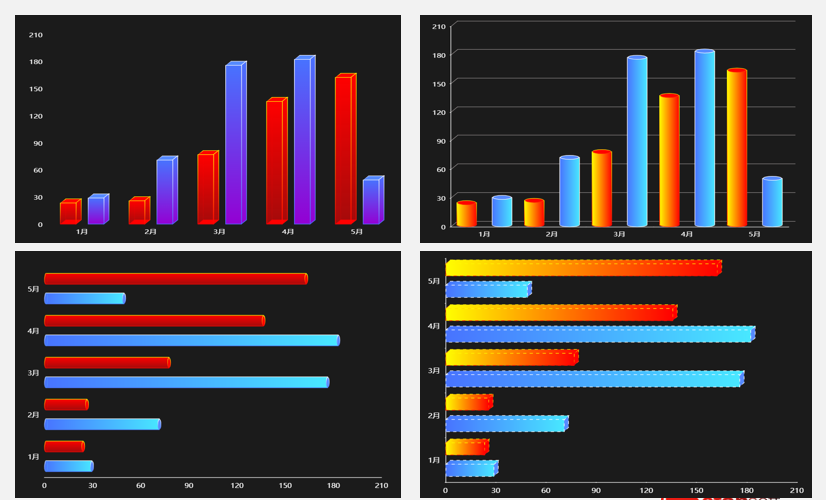
随着信息技术的不断发展，Web前端技术也取得了长足的进步，丰富炫酷的效果已经能在浏览器上实现。SDC UE产品使用了eCharts2/3/4、D3.js、Three.js等框架，以及SVG、Canvas、CSS3、WebGL等原生技术，构建了强大的图表组件库，具备多样化的数据表现能力，除具备展示统计分析数据的统计图形以外，还具备展示关系型数据、层级结构数据的多种图表组件。部分组件的截图如图3-3所示。

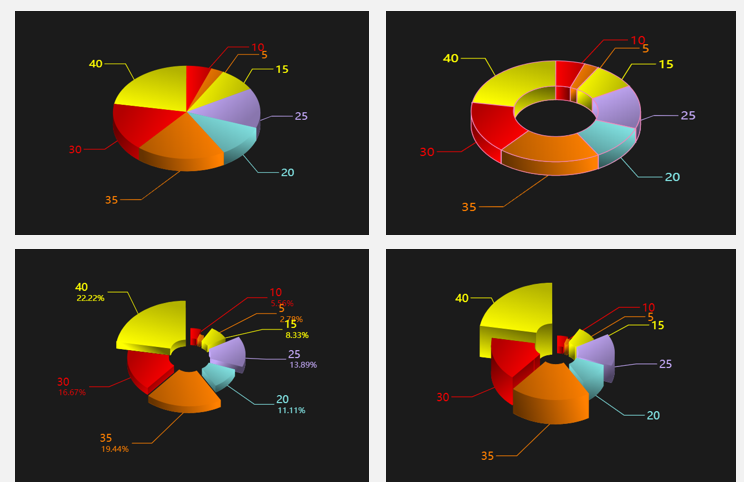
部分图表组件截图



图表组件具备多样化的配置，比如柱状图。

同时支持2.5D的图表及多样化效果配置





### 数据表格多样化

表格是最原始也是最常用的数据可视化方式，SDC UE平台除了提供图形化的组件外，也支持将数据以多样化的表格进行呈现。SDC UE提供了多个表格组件，比如：统计表、交叉表等。

表格组件支持多个表头合并以实现二级表头，支持排序、下钻、按照条件进行过滤、按条件渲染。同时还支持不同方式的导出——当前页导出、全量导出、原数据导出，当数据量大时，可生成多个Excel表，并压缩为一个完整的zip包。

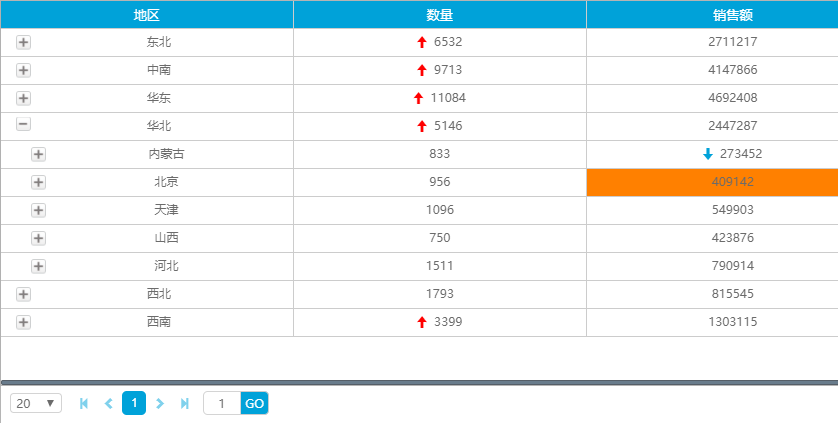
统计表、交叉表



表格支持按照不同的条件进行渲染，可选择针对整行或单个单元格进行渲染，渲染样式支持图标、框线、文字、背景色以及自定义方式。渲染方式灵活多变，可实现预警、对比、重点标注、分类标注等。



设置条件渲染和树形下钻后，支持将数据以交叉表的方式进行呈现，平台自动进行行列转换，效果图如下。





### 炫丽的动态视觉效果

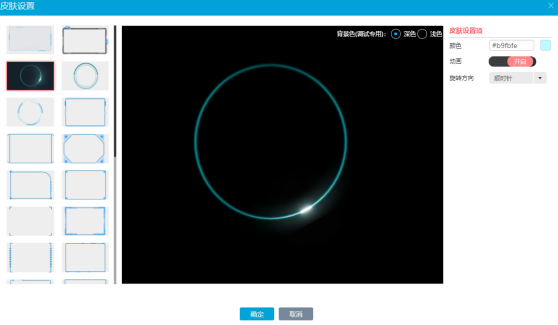
SDC UE平台除了炫酷丰富的图表外，为了增强可视化呈现能力，从多个方面进行炫酷和动态效果的设计。

* 内置多种静态/动态背景图片，加强页面的整体效果。
* 大量动态边框皮肤，让图表更美观、更动感。
* 丰富动感的3D装饰组件。

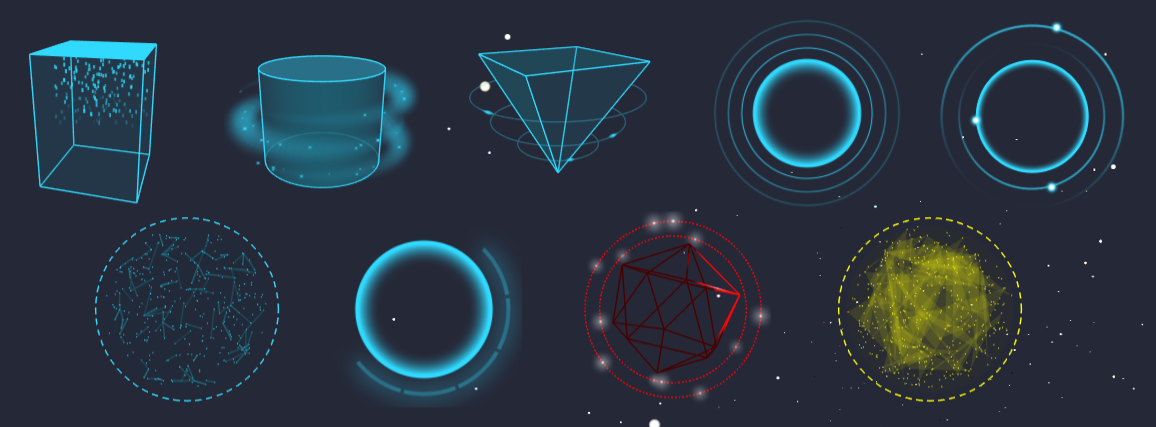
利用产品提供的大量视觉功能，用户轻易就能够设计出美观、动感的可视化页面。

部分功能截图如下：

边框皮肤、背景图示例

装饰类组件图示例



## 多形态空间数据可视化

### 普通矢量地图

产品提供基于geojson数据的矢量地图的轻量级空间数据可视化方案，通过分布地图、迁移地图、标注地图、热力地图四种方式，简单数据绑定即可实现空间数据可视。

支持全国、省（直辖市）、地市（州）三级分。布地图，如图3-5所示。

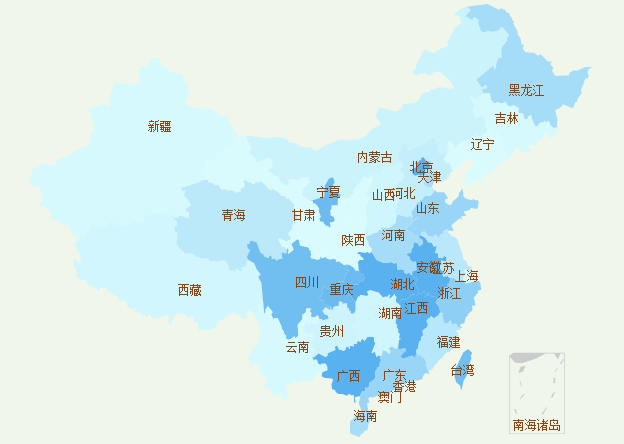
让用户非常直观地监控不同地区的业务发展和绩效情况。地图区域采用不同颜色进行区域显示，决策层可以清楚了解企业的战略执行，并通过在地图上实现下钻操作及联动图形，探索问题的根源，增强洞察力。

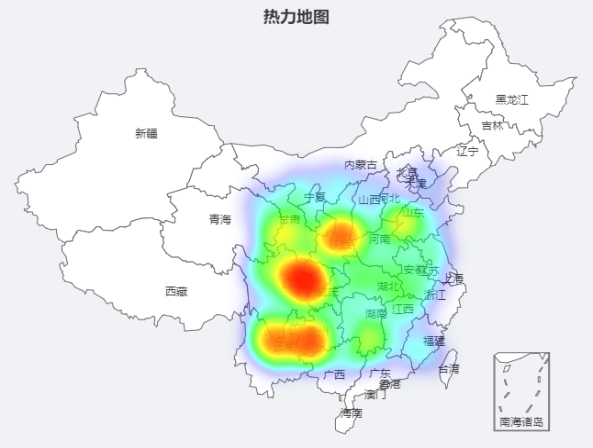
支持迁移、流向场景的矢量地图，用于直观地体现各种数据在省（直辖市）、市（州）的流动，如图3-5所示。

丰富的标注图形，丰富的动态炫光效果，支持大规模数据模式（超过200个标注对象），大量标注点也可以流畅显示。

矢量热力图，可以灵活的调配各项热力参数，方便美观的展示密度型数据，如图3-5所示

矢量地图

### SVG 3D地图

同时支持通过导入SVG图片和Geojson文件，自动生成自定义的3D矢量地图。该地图的能力支持：内置至少四种投影方式，8种及以上的光点样式；支持下钻配置；地图可配置旋转、悬浮、缩放、拖动漫游，并设置投影、发光等3D效果；支持类似GIS地图一样在地图底图上，叠加多个图层，图层的类型包含标注图层、热力图层、迁移图层。

提供丰富的参数配置能力：



支持4种图层配置，每个图层都具备独立的样式配置和数据配置：



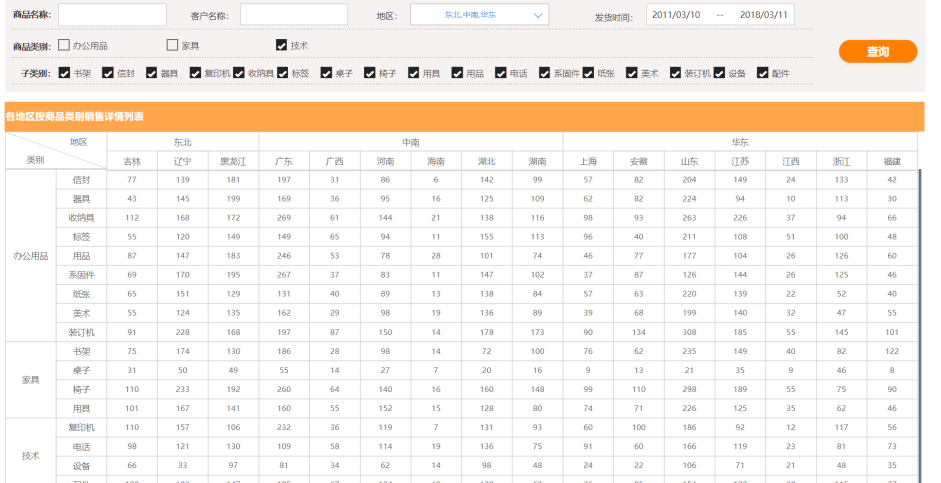
## 灵活的数据交互式设计

随着大数据量的越来越大，数据维度越来越复杂，数据之间的关联性也越来越复杂。在进行大数据可视化展现时，不能再像以往dashboard仪表板一样仅适用一个图表就表达出所有信息，而是需要让用户与数据交流，通过数据的传递、联动，从不同的角度、或者组合条件来交互式地分析和查看数据。

那么当我们在一个可视化设计平台制作好数据呈现后，如何让页面上的数据和图表具有交互性呢？UE产品在数据交互式分析可视化上，独辟蹊径，创新性地引入了两个秘密武器：Web表单组件和事件机制。

通过Web表单组件和事件机制，UE能编排出一个个的交互式查询数据报表页面，再结合事件机制的跳转、弹窗和页面之间参数的传递，就能将这些页面组装成一个完整的报表查询Web应用系统。





## 完备的基础系统管理

### 用户权限管理

* 权限粒度：提供用户、角色、权限三级管理机制。
* 提供地区管理、部门管理，实现分区式和集团式管理。
* 通过用户管理控制进行按照部门、地区、不同角色、不同用户进行精细化授权。

### 安全策略

* URL的访问安全：同时支持免登录模式和认证模式。对于免登录模式，可以直接打开URL页面。对于一些需要进行权限保护的URL页面，系统提供统一的权限管理机制，所有试图打开UE所发布的页面URL的第三方系统，都需要申请和登记，系统会为这些系统分配唯一标志，只有符合条件的系统才能够调用UE的报表页面。因此，就算外部系统或者其他人员拿到了这些URL页面地址，也无法访问。
* 数据安全：在报表配置页面，针对每一个维度、指标，都可以进行脱敏处理，包含对数据加密显示或者屏蔽显示。
* 敏感数据加密：系统中使用到的用户名和密码进行加密存储和传输。
* 具备防XSS注入\SQL注入\CRSF攻击等能力，系统的输入项前端和服务器端均进行了验证。

## 结果应用与集成

### 可视化结果应用

SDC UE平台发布的报表工程、页面，能通过URL或者API接口方式嵌入到第三方应用系统。根据安全要求以及第三方系统特性，同时支持免登录方式和权限认证方式。

### 平台部署与外部集成

SDC UE与其他应用的集成也比较简单，提供专门的SDK组件、丰富的二次开发调用接口，配以深入浅出的帮助和样例文档，让开发者能轻松上手。

在平台集成和客户化功能方面，SDC UE开发的报表是完全基于HTML5技术，可以灵活和第三方进行集成应用和跨平台部署。

平台支持基于CAS认证的标准SSO单点登录技术和接口，可以很轻松地集成其他业务系统，同时也能被集成到第三方业务系统（要求提供标准的SSO接口）中去。

## 部署环境依赖

### 部署方式

UE可以根据用户现场的环境情况、性能要求，提供单机部署、集群部署和云部署三种部署模式，各部署方式的说明如表3-3所示。

部署方式说明表

|  |  |
| --- | --- |
| 部署方式 | 部署说明 |
| 单机部署 | 建议客户单机的最低配置：CPU：单路/6核/1.7GHz；内存：16G，可以满足基本运行能力，如果需要提供较好的性能建议使用CPU：单路/8核/2.1GHz；内存：32G的服务器。 |
| 集群部署 | 用户分析数据量较大时的环境需要采用集群部署方式，以支持较重数据使用量和较高的用户负载。需要两台应用服务器，一台负载均衡服务器，n（n>2）台分析服务器和两台存储服务器。 |
| 虚拟机或基于云的部署 | 在虚拟机上或云部署中运行UE时，应该注意，每种虚拟化平台都提供了足够的管理基础结构，用以在可扩展的高性能拓扑中部VM。还请遵循虚拟基础结构提供商建议的最佳做法，确保UE可以获得相应的计算资源、内存资源和数据资源。 |

### 环境要求

服务端环境

服务端对硬件配置、操作系统、软件环境、应用服务器和数据库均有限定，具体配置要求请参见表3-4。

服务端环境配置说明表

|  |  |  |
| --- | --- | --- |
|  | 最低配置 | 推荐配置 |
| 硬件配置 | 普通服务器  CPU: Intel至强/单路/6核/1.7 GHz  内存：16G以上  磁盘：TB级 | 普通服务器  CPU: Intel至强/单路/8核/2.1GHz  内存：32G以上  磁盘：TB级 |
| 操作系统 | SUSE/CentOS 6.5及以上操作系统 | |
| 软件环境 | JDK 1.8 | |
| 应用服务器 | Tomcat | |
| 数据库 | Mysql，PostgreSQL | |

客户端环境

浏览器：推荐使用Chrome 42以上版本

硬件配置：4G内存，1G以上显存