**EPS数据平台使用指南**

EPS数据平台是集丰富的数值型数据资源和强大的数据分析处理系统为一体的综合性信息服务平台。EPS数据平台大规模集成整合了各类数据资源，形成了96个专业的数据库，并将数据分析处理软件与数据库置于同一系统平台中，为用户提供了数据资源获取与数据分析处理的一站式服务。

希望《EPS数据平台使用指南》能够帮助您更好地了解和使用EPS数据平台的数据和功能，提高学习研究效率。

**一、用户登录**

**1.EPS数据平台入口**

在浏览器中输入网址https://www.epsnet.com.cn/或直接点击本单位连接地址，即可进入EPS DATA官网，

**2.登录数据库**

若当前IP在授权范围内时，点击“EPS数据平台”或页面主体部分的“具体数据库名称”则直接进入EPS数据平台。



若当前IP不在授权范围内时，则弹出登录窗口，需要通过手机号验证码或机构账号进入平台。



校外可以使用学校的VPN链接登录，具体链接可以查看学校的图书馆官网中的校外访问链接或者询问图书馆老师。已经开通CARSI的学校也可以使用CARSI在校外进行访问登录。

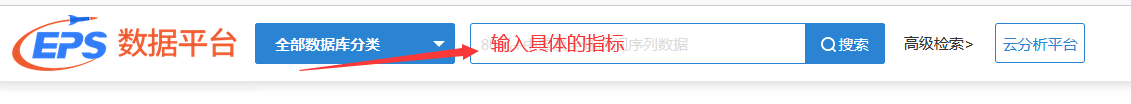


**二、数据检索**

EPS数据平台为用户提供两种数据检索模式。分别为：跨库检索和库内检索。

**1.跨库检索**

跨库检索可以进行语义检索，登录EPS数据平台并进入数据库后，用户可以输入需要查询的时间+地区/国家/行业+指标，系统会在所有数据库中搜索包含相关输入内容的指标并显示查询结果。



例如：我们需要查询2020年全国的国内生产总值增加值的情况，我们可以在搜索栏中输入“国民总收入”这一指标，系统显示查询到几十条包含“国民总收入”的指标及他们的相关指标。我们在检索结果页提供了频度筛选、数据时间、区域筛选、来源筛选等选项，方面用户快速定位检索指标。点击“国民总收入”指标即可进入到中国宏观经济数据库年度(全国)中显示数据。



高级检索可以根据需要选择要检索的字段，在全部数据库或者部分数据库中检索指标。



如您点击具体指标页面没有跳转到具体的数据页面，您可以检查一下您的浏览器设置，是否有设置了阻止弹窗弹出或者是广告拦截等设置，将这些设置关闭即可跳转到具体的数据页面。

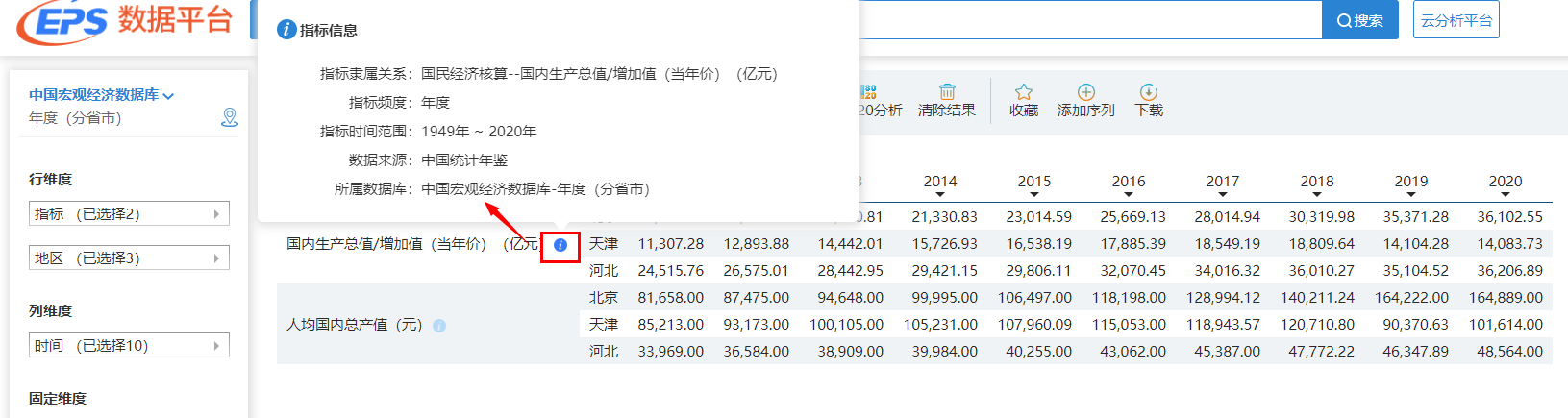
**2.库内检索**

用户登录平台后，可以在界面的左端进入相应的数据库进行指标检索。例如：我们仍然查询全国各省国内生产总值的情况，首先在头部“全部数据库分类”找到“中国宏观经济数据库”，由于我们查找的是省级数据，所以我们进入“中国宏观经济数据库-年度数据（分省市）”子库，然后在“指标”维度下拉框内进行库内模糊搜索，在搜索框内输入“国内生产总值”关键词，点击放大镜搜索图标或按回车键，“指标”维度下拉框内即会显示含有该关键词的所有指标高亮显示，选中“国内生产总值/增加值（当年价）（亿元）”指标，点击“查询”，界面右边会出现该指标的数据表格。



|  |  |
| --- | --- |
|  |  |

另外，在维度下拉框指标处右键，会出现指标选择的便捷功能，方便客户对指标进行筛选。如：选择同级、选择子项等。

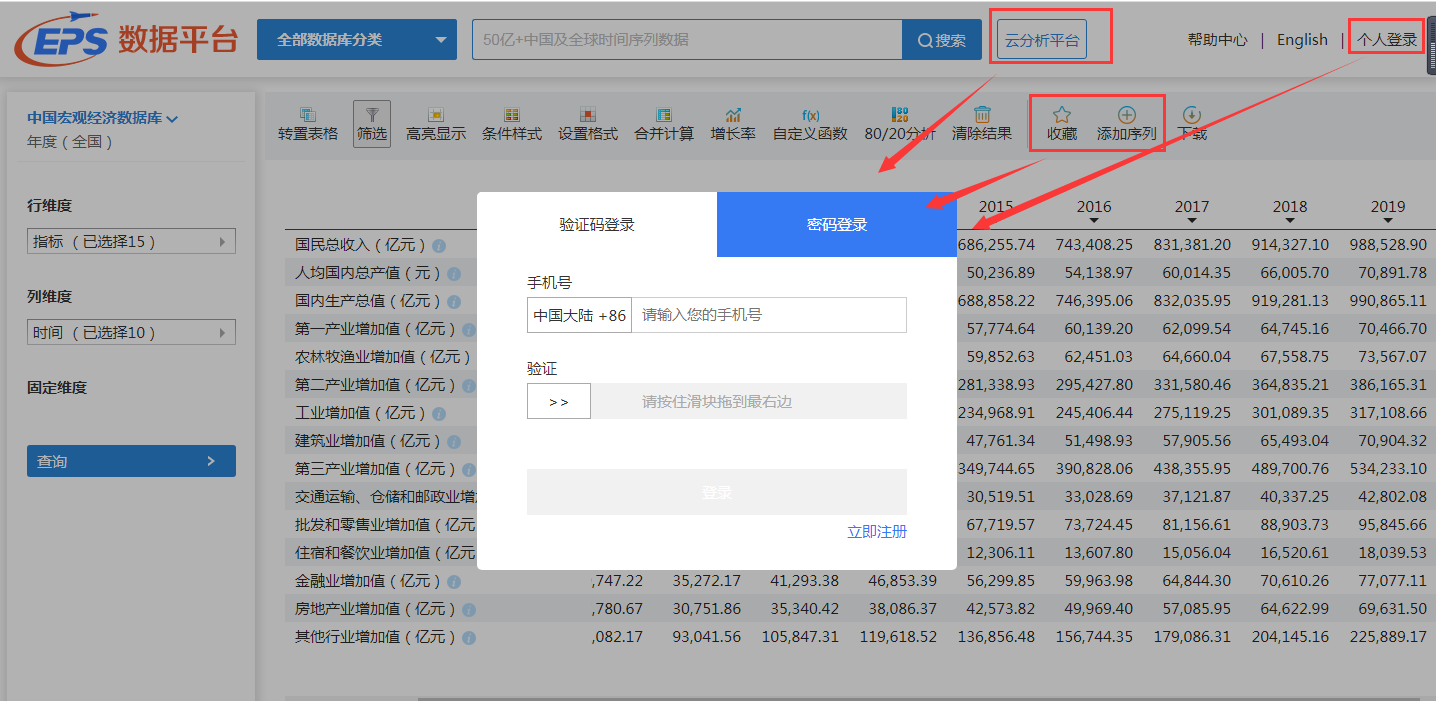
同时，在数据区域展示的指标中，每条指标后都有一个icon图标,当鼠标放在该图标上时，便会出现该指标所对应的基本信息，包含指标隶属关系、指标频度、指标时间范围、数据来源以及所属数据库。部分指标的指标信息中还包括指标批注信息。

**3.个人中心**

为了方便用户在线保存和管理数据，EPS数据平台提供**个人中心**功能。

对于IP用户：

1. 当IP登录或机构账号密码登录用户点击“收藏”、“添加序列”以及进入“个人中心”和“云分析平台”时，系统会自动弹出我的数据中心登录窗口。



对于没有个人中心账号的用户，在相应页面都提供了“注册”功能，注册只需手机号即可完成注册，注册流程如下：





这里需要注意的是，“个人中心”账号仅限在注册时所在的机构IP授权范围内或者机构账号登录范围内再次登录时有效。

**手机验证码用户**在进入EPS数据平台数据库后，相应功能自动开通。

**4.身份认证**

所谓身份认证服务，是为了保障机构用户在IP范围外仍可正常访问平台所提供的一种增值服务，是EPS DATA创新产品服务体系的重要组成部分。具体流程如下：

第一步：填写信息。

信息包含学校/机构、真实姓名、性别、出生日期、邮箱、职业身份、院系/部门、专业、入学时间、毕业时间、证件编号、证件照片，所有选项为必填项。同一个机构下证件编号不能有重复；



第二步：提交信息，等待审核。审核结果将在1个工作日内短信通知，同时审核结果可在个人中心查看。



**三、数据保存和下载**

**1.数据的保存**

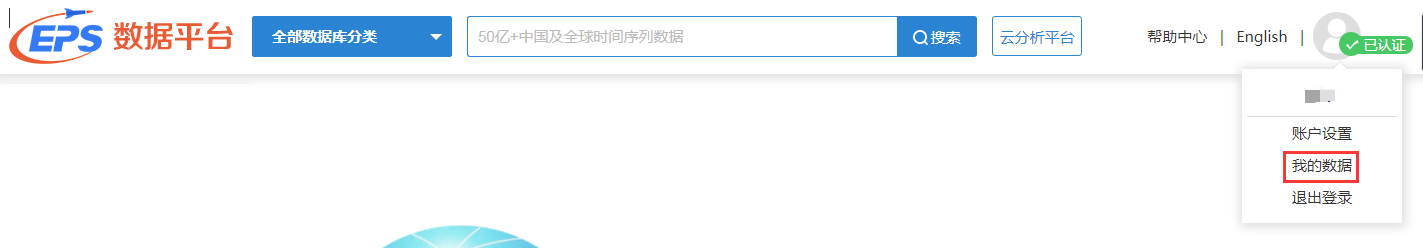
EPS数据平台的保存分为两种，一种保存的是查询条件，一种保存的是查询结果。这里的数据保存是指前者，而后者定义为"下载"。用户在进行数据查询检索的时候，是通过设定一定的检索条件才能得到数据结果的。为了下次登录时方便调用上次的检索结果，我们可以对查询条件进行保存。

点击 “收藏”图标，弹出“保存到收藏夹”窗口，用户自定义需要保存数据的名称，选择保存位置，并点击“收藏”，即可保存当前数据到收藏夹。



**2.打开已保存数据**

如果要调出上次的检索结果可以通过点击导航栏“我的数据”，在“我的收藏”中查找。



**3.下载**

EPS数据平台为了方便用户的使用，提供了多种下载方式。在数据查询页面下提供Excel、csv、txt三种下载格式。图表和地图功能提供了图片、Excel、PDF三种下载格式。

**四、数据处理分析功能**

**1、表格转置功能**

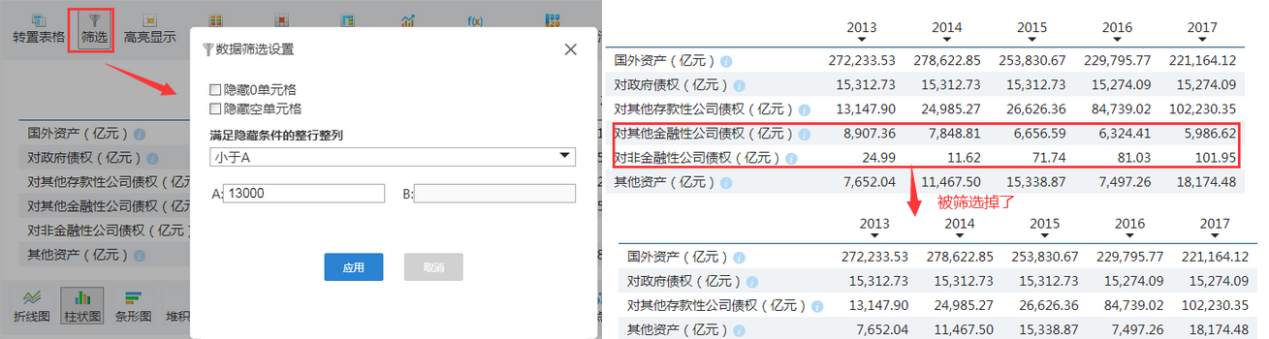
用户可以通过点击“转置表格”图标，轻松实现数据的行列转换。另外，通过拖拽改变维度栏的位置，也可以实现行列的转换，还可以实现其他不同样式的表格形式。



注：当某一维度被拖拽到“固定”栏时，只能选择该维度中的一项指标，不能实现多选。同时，被固定的指标信息会显示在表格上方。

**2. 数据筛选功能**

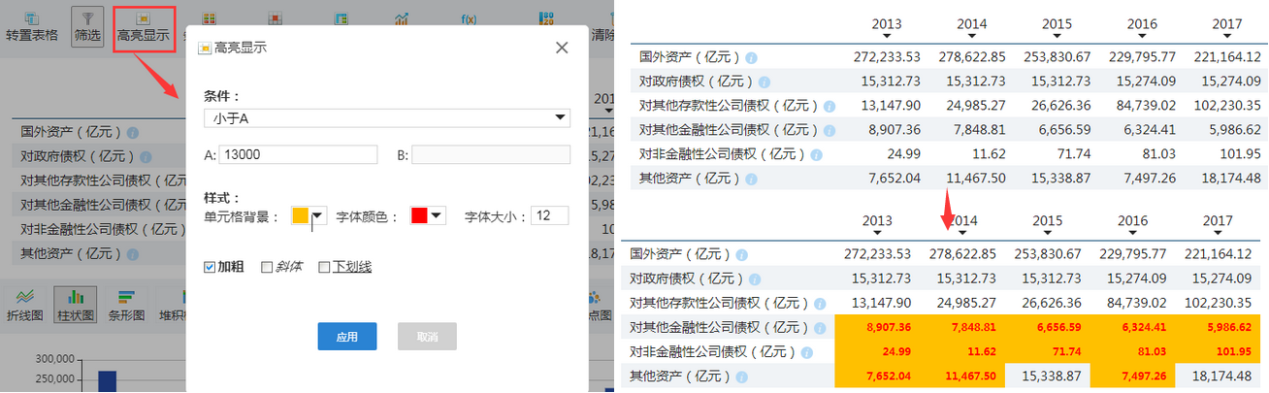
用户可以通过点击“筛选”图标，实现对被检索数据的筛选。点击“筛选”图标，弹出“数据筛选”窗口，窗口上方是两个常用的筛选条件：隐藏0单元格和隐藏空单元格，直接勾选，点击应用即可。



用户也可以在窗口下方自定义筛选条件。需要说明的是，当某行或某列中只要有一个数不满足筛选条件时，该行或该列将不会被隐藏。

**3、高亮显示功能**

当用户需要在检索出的大量数据中观察某些特定数据时，可以通过“高亮显示”功能，用不同的颜色、字体、背景色彩使数据凸显出来，可更加直观地展现数据情况。在高亮显示设置对话框中输入相应的参数，设置样式，点击应该即可。



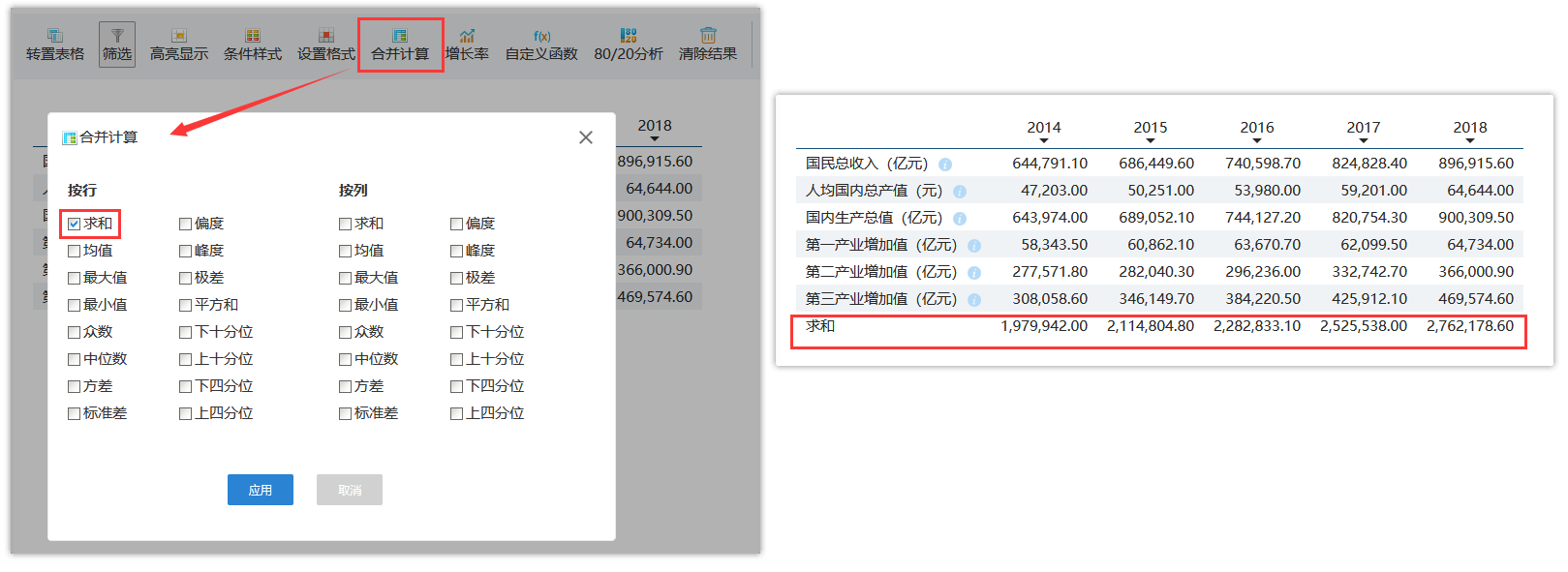
**4、条件样式功能**

当被检索出的数据过多时，用户可以通过“条件样式”功能，通过表格色调或图标的样式变化来区分数据数值的大小，有助于用户更直观地观察全部的检索数据，发现数据的特征。系统提供4种表格的色调样式以及4种图标样式，用户可根据自己的喜好进行选择。



**5、合并计算功能**

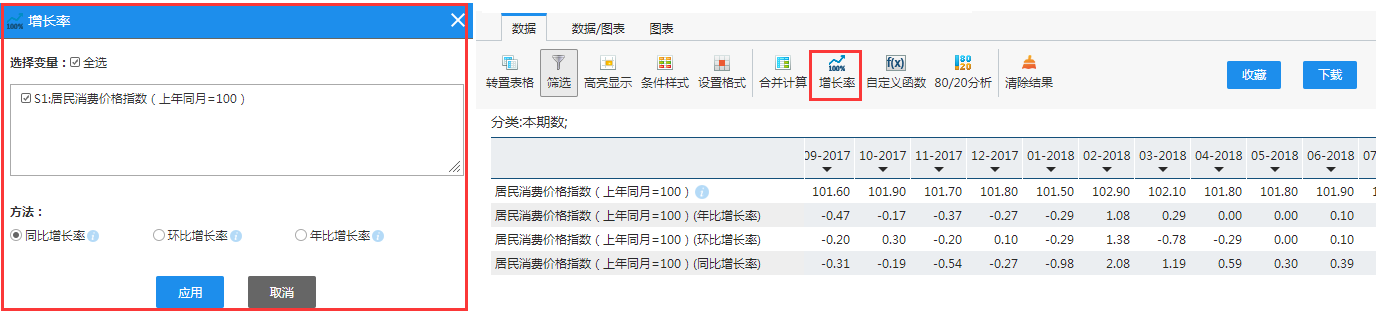
EPS数据平台还为用户提供了常用的数据计算功能，可以在线直接对检索出的数据进行十多种相关运算，大大节约了用户进行数据计算和处理的时间。



**6. 增长率功能**

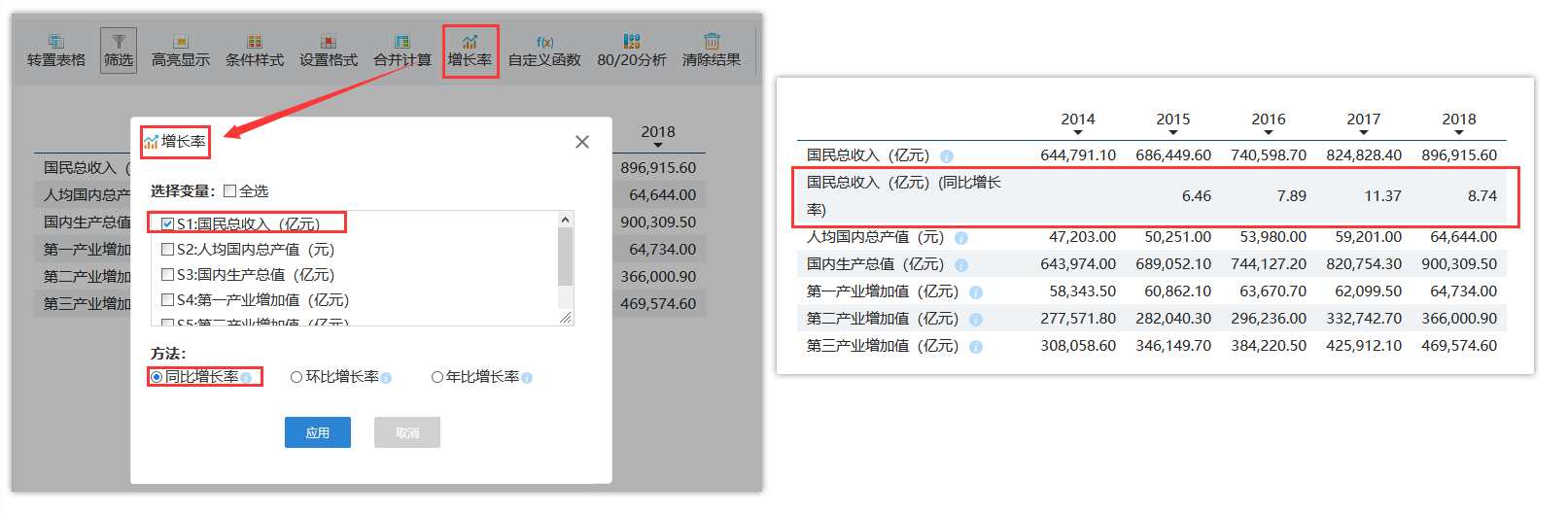
计算数据查询结果的某一行或多行时间序列数据的增长率，EPS数据平台提供同比增长率、环比增长率和年比增长率三种。若数据是季、月度数据，则年比增长率为（当前数值-去年12月或去年第四季度数值）/去年第四季度数值\*100。

点击“增长率”功能按钮，在弹出框中选择变量和方法，点击应用即可。



**7.自定义函数**

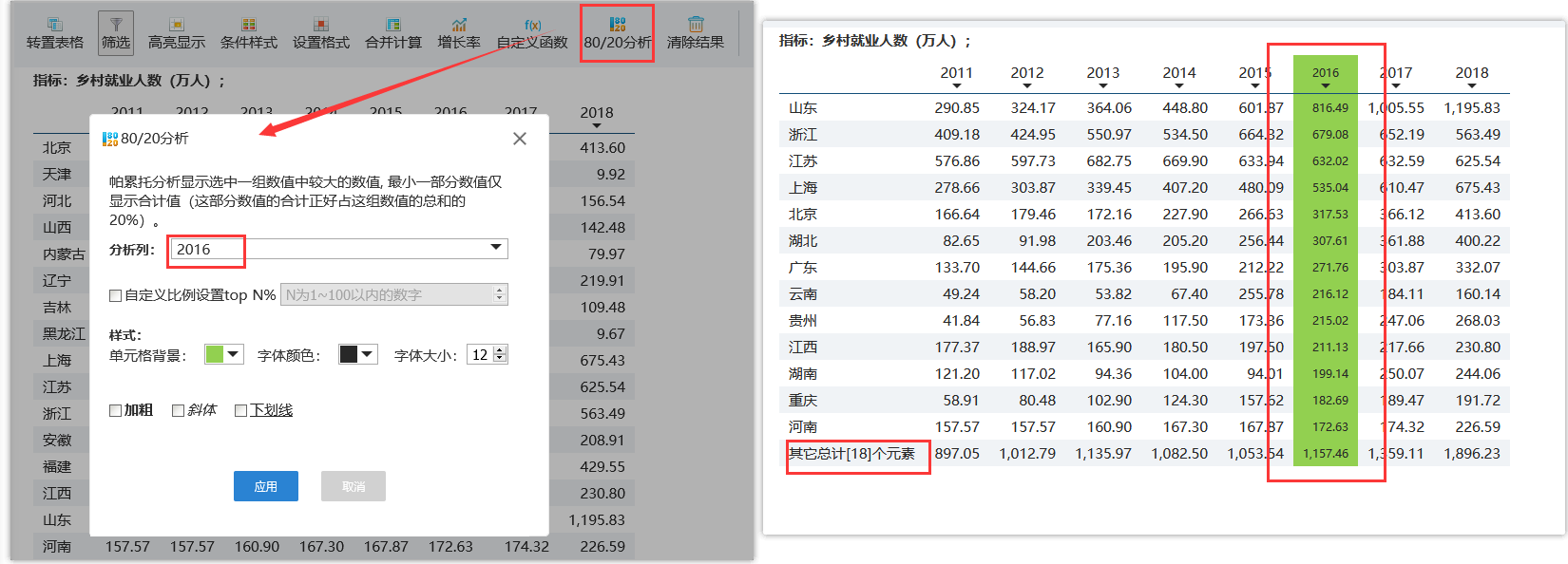
若要对序列做较复杂的运算时，那么“自定义函数”无疑是一个有力的工具。“自定义函数”可以对数据查询结果进行各种算数计算。包括加减乘除、倒数、乘幂、绝对值、对数、开方、余数等。比如：我们要求山东省粮食单位产量，即用粮食总产量除以粮食播种面积得到，打开“自定义函数”，通过鼠标或键盘选择或输入变量、计算符号的方式添加到“计算表达式”中，点击“应用”，即可得到结果。



**8、80/20分析功能**

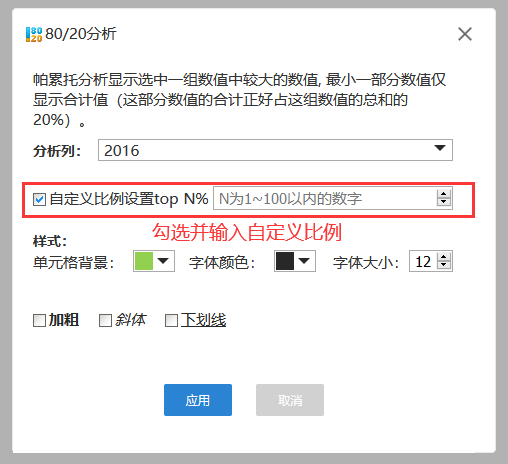
EPS数据平台的80/20分析是根据意大利经济学家帕雷托的80/20法则进行数据分析的一种方法，通过这种方法的分析，可以将已经查询到的数据分为重要的并且占比例达到80％的数据和不重要的只占比例20％的数据。 这样，用户就可以在经济研究、能源研究或者管理研究等过程中轻松的判断出重要的占80％的因素是来源于哪方面或者哪个地区。从而，为相关决策提供数据支持。

例如：我们对查找到的数据（2011-2015年31个省市私营企业乡村就业人数）进行80/20分析。用户可以通过点击“80/20分析”图标，实现对数据进行80/20分析。我们可以在参数设置对话框中选择我们需要分析的某一列，然后可以对分析结果的字体和其背景颜色加以设置。



可以看到，根据选择系统对2016年这一列进行了80/20分析：依降序排列此列数据，表格正常显示前80％的结果，而后20％的结果则以总和的形式显示。 在分析结果中，可以看出，私营企业中乡村就业人数最多的是浙江，按照乡村就业人数对省市进行降序排列，山东、浙江、江苏等13个省市的私营企业中乡村就业人数占中国31个省市总的私营企业中乡村就业人数的80%，剩余18个省市占20%。这样，我们一眼就可以判断全国私营企业中乡村就业人数中占比较大的省市，而不用从众多数据中逐一查找。可以得出结论：私营企业吸纳的乡村就业人员中相当大比例集中于东部私营经济发达的省（市）。进一步可对比研究国企和集体企业的数据，来说明若要带动农村就业，需大力发展私营经济。

同时，EPS数据平台V2.0对80/20分析增加了自定义比例，可以满足用户多样性需求。

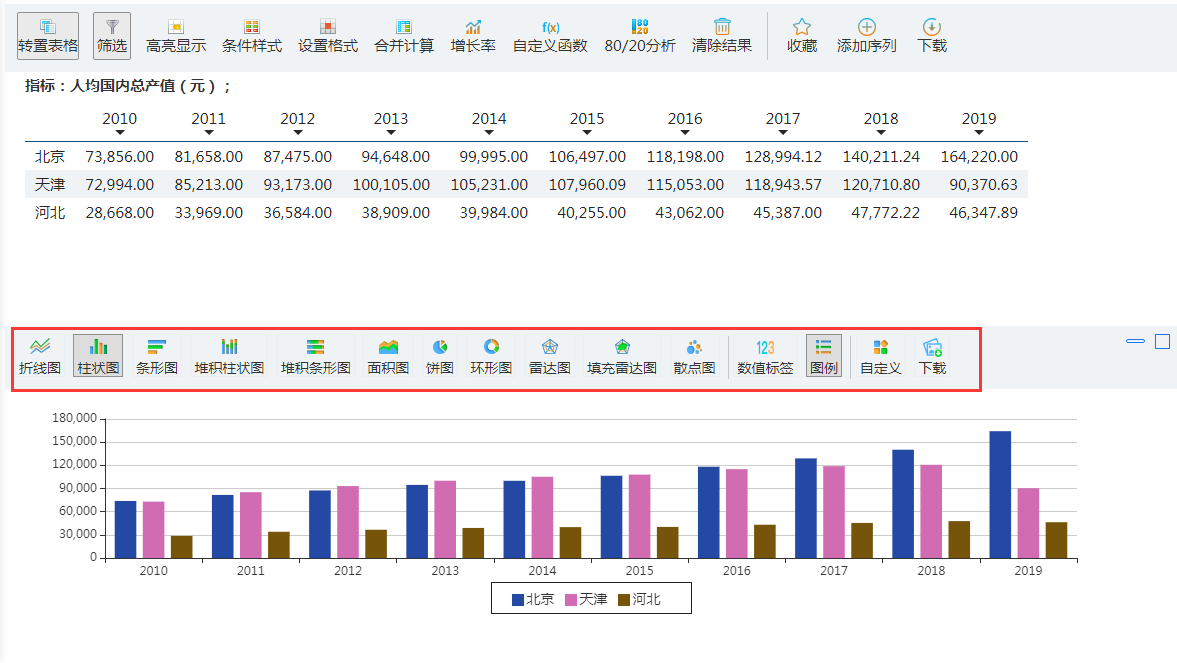


**五、数据可视化功能**

**1、图表可视化功能**

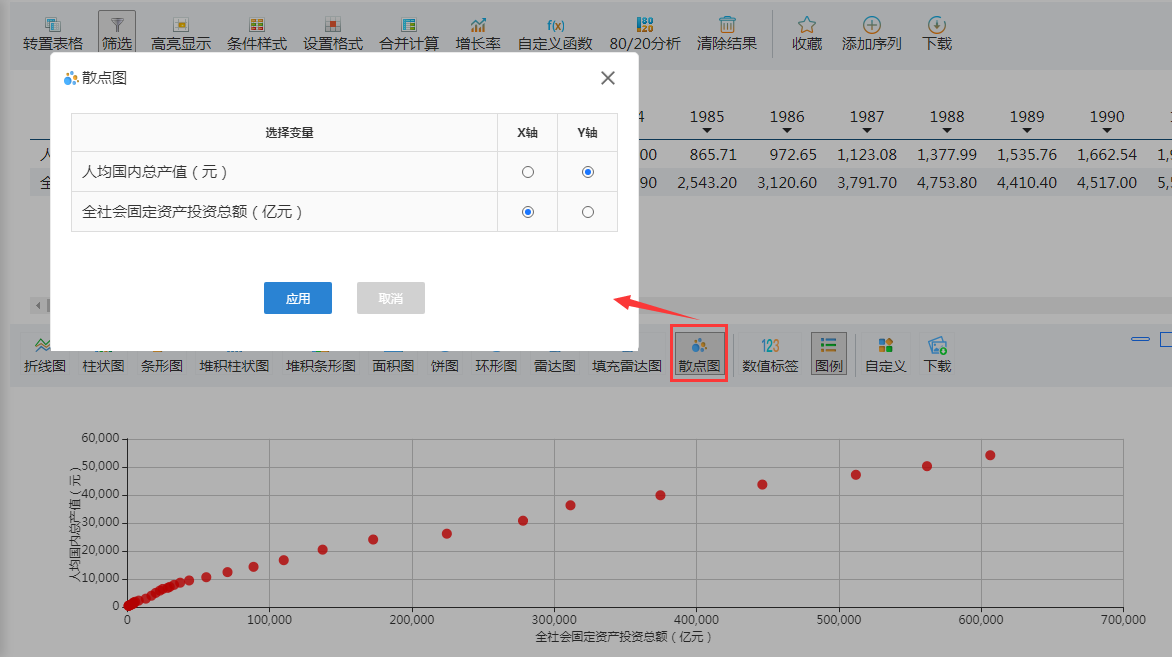
EPS数据平台提供了11种图表展示，用户可以根据自己的需求选取合适的图表。选择图表显示功能时，系统默认以柱形图显示表格所有的区域。用户也可以点击表格中的年份或者地区，单独展示某年或某地区的数据，甚至可以只选取某个数据点，展示该点的数据。

例如：我们要比较北京、天津、河北三个地区的人均国内总产值，在查询数据并选择“表格/图表”栏后，图表中显示了2010-2019年三个地区人均国内总产值的柱形图，由图表可知，北京和天津的人均国内总产值远高于河北的人均国内总产值。

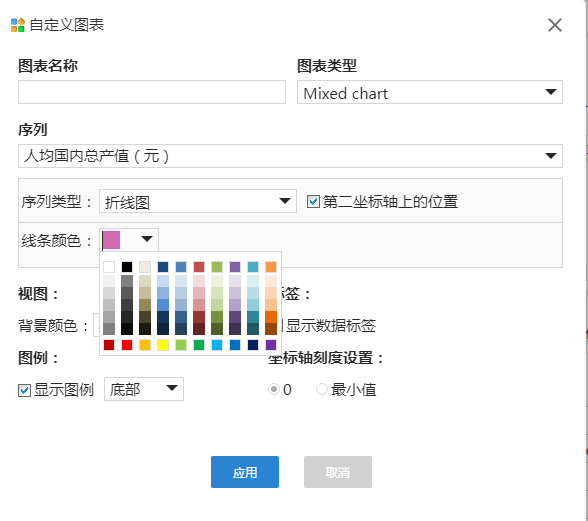


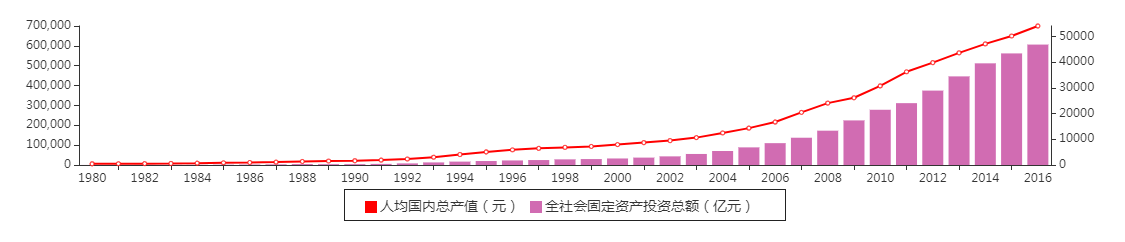
用户还可以自行选择是否添加数据值标签和图例。

如果用户需要研究两组指标的关系，可以选用散点图。散点图需要自行选择X轴和Y轴，然后绘制图表。



同时，用户可根据具体需求，通过点击“自定义图表”按钮，实现图表属性的个性化设定。另外，自定义图表功能的图表类型中新添了“Mixed Chart” 类型，可以在图表中同时显示折线图和柱形图，满足了用户多样化的需求。同时，还可以设定第二坐标轴，是否显示网格线，选择背景颜色，图例显示位置（底部和右边）等。

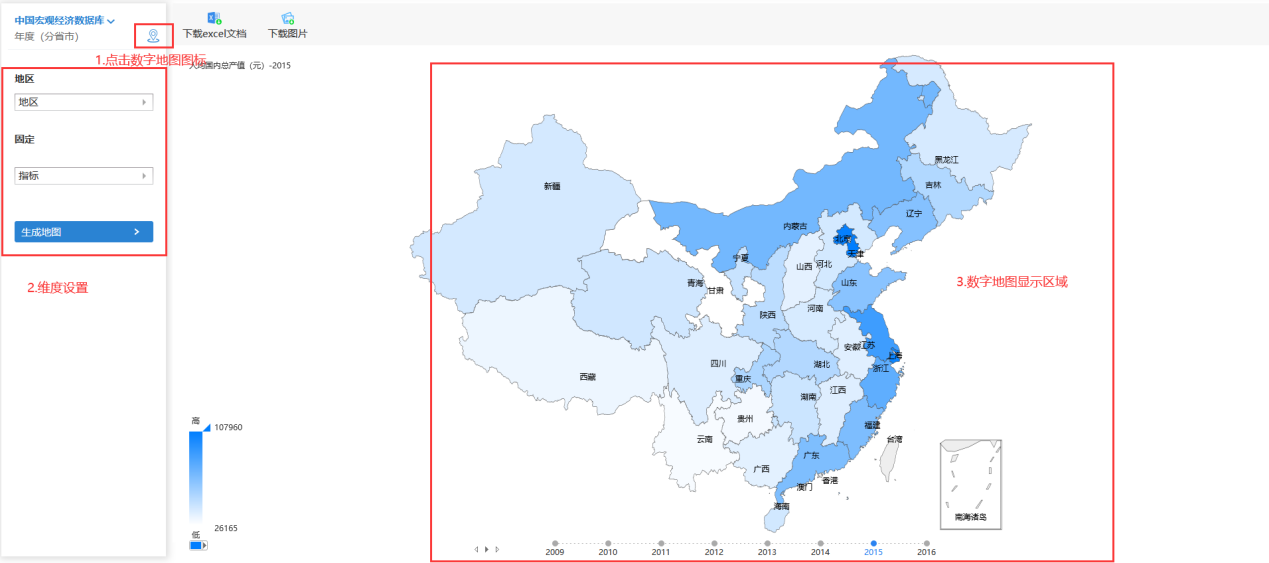




如果用户需要下载图表，只需点击“下载图片”即可。用户还可以点击“图表”栏，只显示图表不显示表格数据，让用户更清晰地观察图表的细节。

**2. 数字地图功能**

数字地图功能是通过对地图上不同区域颜色的深浅变化，以直观的形式展现中国各省或者世界各国的数据变化情况。选择地区和指标后，点击“地图”，便可按地图模式查看数据，地图上颜色的深浅代表数值的大小。用户可以在下方时间序列选择时间，获得相应时间的地图，也可以按“播放/停止”键，依次播放所有时间的地图。另外，用户可以自行设置地图的颜色，也可以直接下载地图（excel、jpg两种格式）。



**六、分析预测**

EPS数据平台采用SaaS(软件即服务)模式，通过云分析为用户提供高质量、高效率、低成本的数据处理、可视化展现、分析预测等软件服务，为科学研究或论文撰写提供专业强大的工具支持。

比如，根据理论和经验分析，研究影响粮食产量的主要因素有化肥施用量、粮食播种面积、成灾面积、农业机械总动力以及第一产业就业人数。

首先用户通过跨库检索或库内检索找到相应指标，通过“添加序列”功能将来自不同数据库的指标时间序列添加到云分析平台。



注意：“添加序列”功能仅适用于时间序列，并且对于IP登录用户需要登录“个人中心” 。

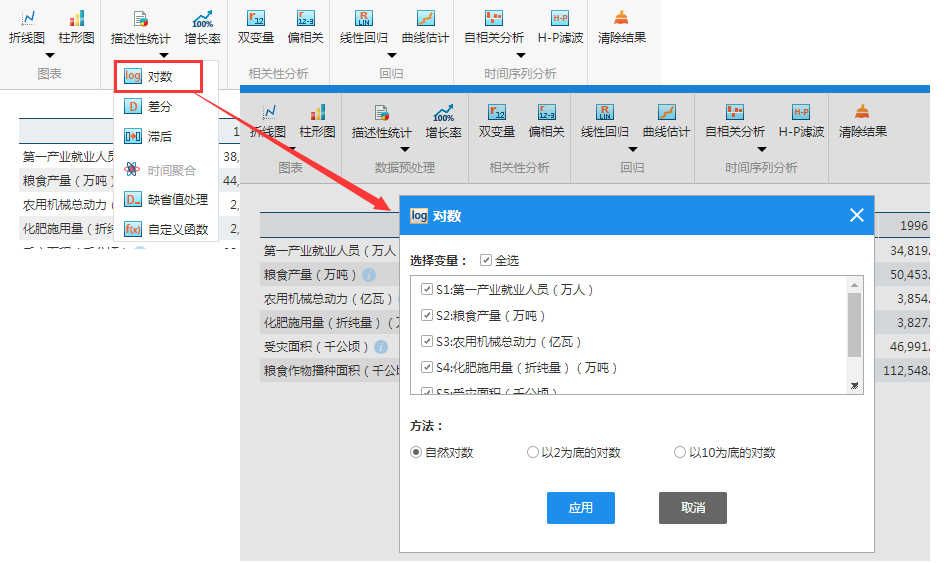
在云分析平台中，提供图表、数据预处理、相关性分析、回归以及时间序列分析等功能。

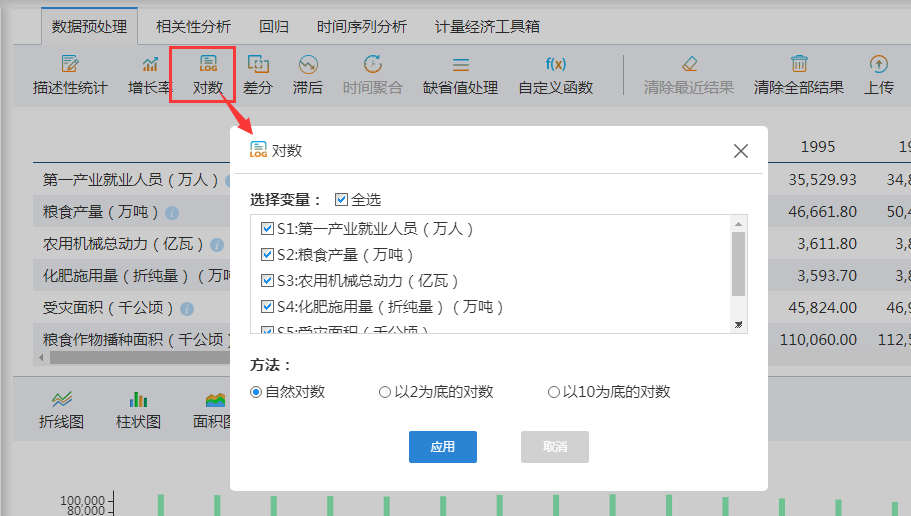
点击 “云分析平台”进入，在页面左侧找到添加的序列集“粮食生产”，时间跨度为1990年到2016年。



1. **数据预处理**

先对所有序列取自然对数，点击“对数”功能，在弹出框中选择变量、方法，点击应用即可得到结果。

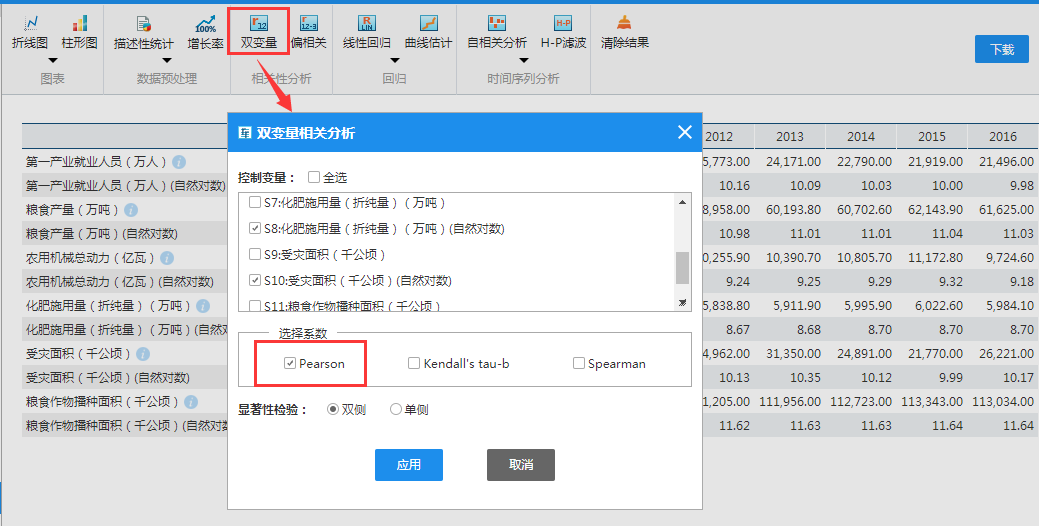


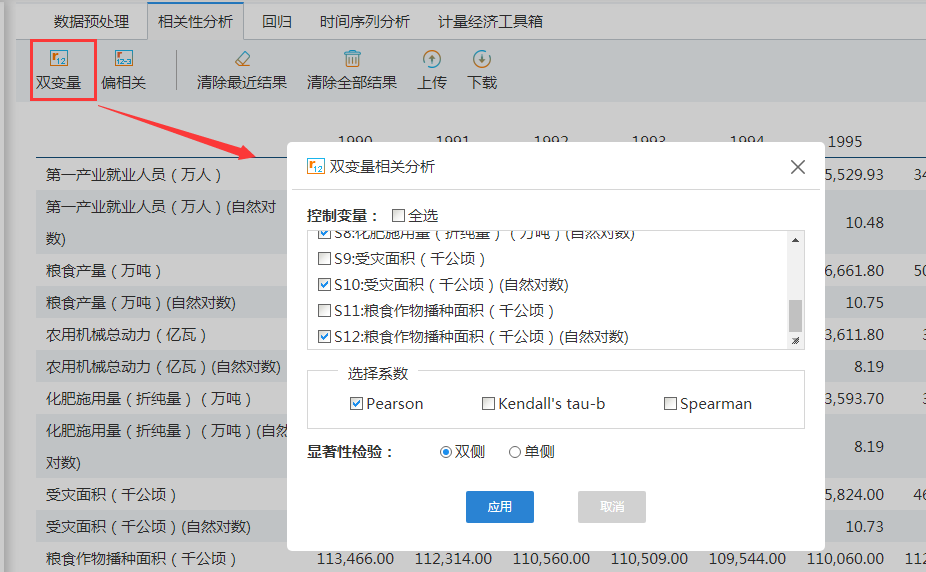




1. **相关性分析**

在对变量间进行双变量相关性分析，点击“双变量”，选择影响粮食产量的取自然对数后的序列，选择方法，点击应用，得到双变量相关分析结果。



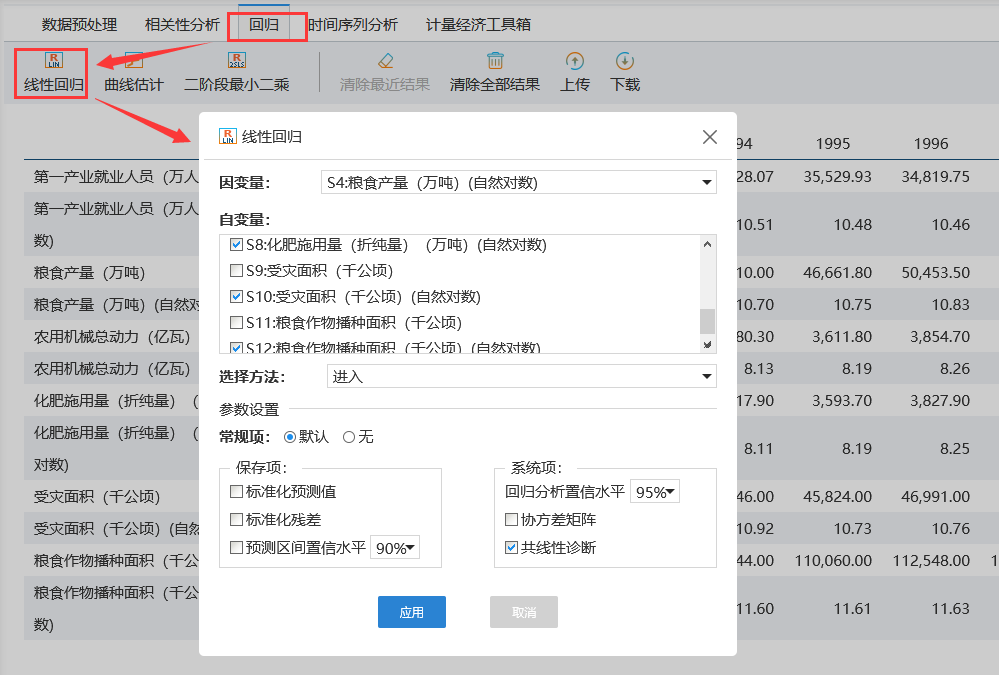




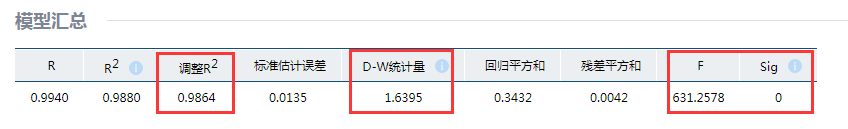
有分析结果可以看出：S2和S4、S2和S6，S2和S8之间有较强的相关性,可以初步判断 S2、S6、S8之间可能存在多重共线性，可以考虑从模型中剔除S2、S6.

1. **线性回归**

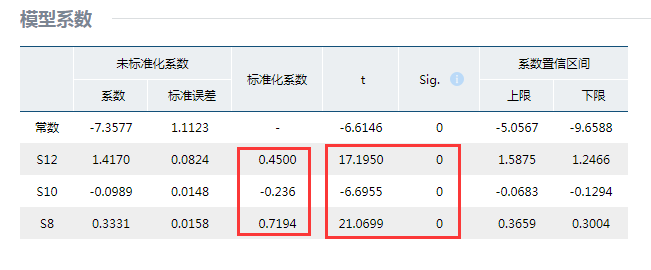
最后进行回归分析，根据上述结果，点击“线性回归”功能，在弹出框中变量选择S4,自变量选择S8、S10、S12，选择方法、保存项、设置置信水平，点击应用。



输出结果：

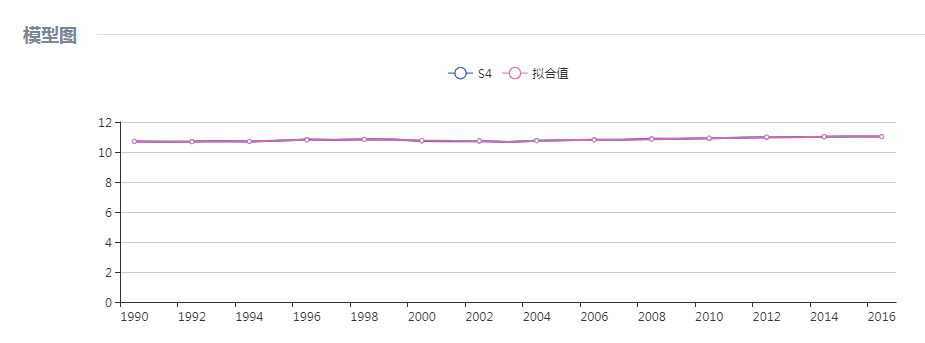


由F=631.2578，p<0.05统计学显著，并且，调整R^2=0.9864,可以看出粮食产量与自变量总体线性关系显著。

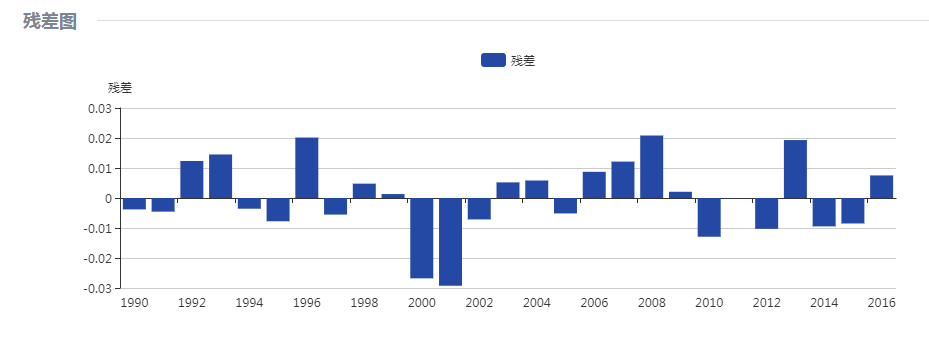


此外， S8、S10、S12通过t检验。

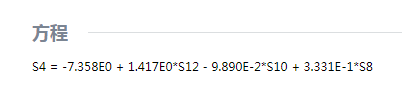
模型调整R^2=0.9864，表明模型拟合程度比较理想。通过模型图也可以看出，拟合值和实际值之间误差较小。



另外，由于模型使用的是时间数据建模，因此残差可能存在自相关问题。查看残差图，认为残差序列可能不存在显著的自相关问题。



模型表达式为





结合标准化回归系数，不难看出， 化肥施用量与粮食播种面积对粮食产量影响为正，即在其他条件不变的情况下，化肥用量值在一定范围内越大，粮食播种面积越大，粮食产量越高。并且，化肥施用量的影响相对更大些。 成灾面积与粮食产量成负相关。在其他条件不变的情况下，成灾面积越大，粮食产量越低。