

CSMAR Solution

用户手册

CSMAR

深圳希施玛数据科技有限公司

版权所有 侵权必究

1. 引言

1.1. 编写目的

为使用 CSMAR Solution 的用户提供操作指引。

1.2. 背景

CSMAR Solution 是深圳希施玛数据科技有限公司开发的 CSMAR 数据库网页版查询工具，您可以通过登录个人注册账号或者学校机构账号查询下载已有权限的 CSMAR 数据，同时，您还能直接通过微信支付或者支付宝支付在线购买 CSMAR 数据，方便快捷。

1.3. 定义

无

1.4. 参考资料

《CSMAR Solution V4.1 用户手册》、《CSMAR Solution V4.2 用户手册》、《CSMAR Solution V4.3 用户手册》、《CSMAR Solution V4.4 用户手册》、《CSMAR Solution V4.5 用户手册》、《CSMAR Solution V4.6 用户手册》、《CSMAR Solution V4.7 用户手册》、《CSMAR Solution V4.8 用户手册》。

2. 用途

2.1. 功能特点

1、数据查询

直接对 CSMAR 数据进行查询下载，操作方便简单。同时，提供数据下载到邮箱功能及方案保存功能，方便您下次调用。

2、数据超市

提供 CSMAR 系列数据库自主便捷选购服务，可以对数据库资源进行自由组合下单购买，购买流程简单易操作。在线支付成功后，可立即进行已购买数据表的数据查询下载操作。

2.2. 性能特点

无

3. 运行环境

3.1. 硬件环境

本系统为 B/S 架构。对用户的硬件环境没有特殊要求。用户只需将所使用电脑连接到网络，正常打开网页，即可使用本系统。

3.2. 软件环境

页面推荐分辨率：1280*960 以上

浏览器：Chrome(谷歌)浏览器、搜狗高速浏览器、IE11、火狐浏览器、Safari 浏览器、

Edge 浏览器，推荐使用 Chrome(谷歌)浏览器。

4. 使用过程

4.1. 安装与初始化

登录网址：cn.gtadata.com

4.2. 使用说明

4.2.1 登录及注册

- 已经拥有账号——登录

输入网址，进入 CSMAR Solution 首页，点击导航栏右侧【登录】，在登录弹框中输入账号、密码登录，验证通过即可进入。



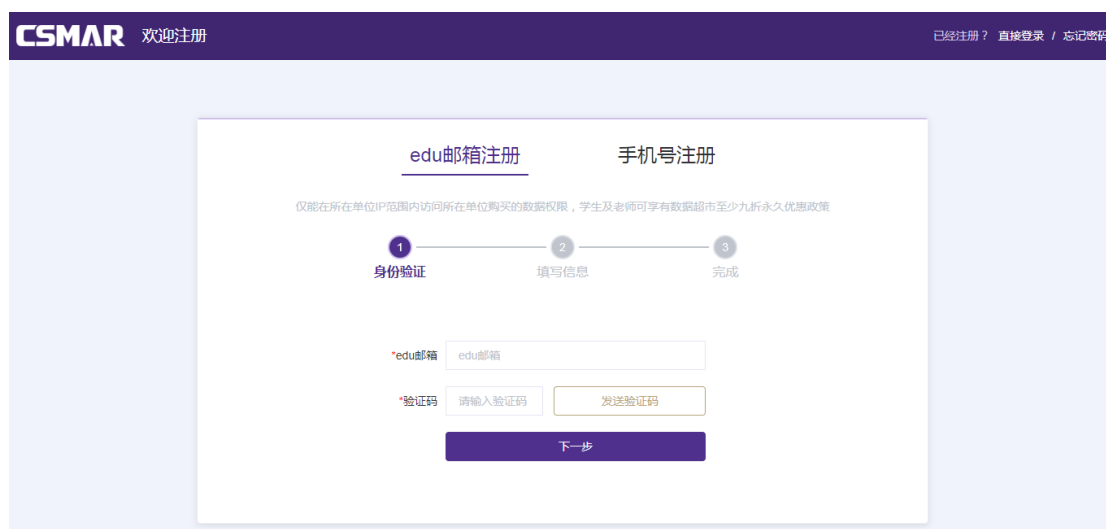
登录账号分学校机构账号和个人注册账号，其中：

个人注册账号：使用注册时的电话、邮箱或用户名均可登录；

学校机构账号：允许多人同时在线，最大同时在线人数一般为 50-200 人，具体根据学校购买情况而定。同时，学校机构账号需要验证 IP 范围段，IP 范围内才能正常登录使用，支持使用 IP 自动登录。

● 没有账号——注册

您可以通过注册成为网站成员，点击首页导航栏右侧【注册】，注册可选择 Edu 邮箱注册或者手机号注册（手机号注册需要上传证件），在注册页面填写相关信息即可。注册成功后该个人注册账号即可拥有所在单位的 CSMAR 数据权限，但是数据权限仅在所在单位提供的 IP 范围内使用。注意，“所在单位名称”必须如实填写，后续如需更改（Edu 邮箱注册用户无法修改所在单位，需用新的 Edu 邮箱进行注册），手机号注册可在个人中心-基本信息上传相应的证件-勾选“修改单位名称”，提交审核通过即可。



The screenshot shows the CSMAR registration interface. At the top, there is a navigation bar with the CSMAR logo and '欢迎注册' (Welcome to Register) on the left, and links for '已经注册?' (Already Registered?), '直接登录' (Direct Login), and '忘记密码' (Forgot Password) on the right. The main content area features two tabs: 'edu邮箱注册' (Edu Email Registration) and '手机号注册' (Mobile Number Registration). Below the tabs, a progress indicator shows three steps: 1. 身份验证 (Identity Verification), 2. 填写信息 (Fill in Information), and 3. 完成 (Complete). The 'edu邮箱注册' form includes a text input for 'edu邮箱' (Edu Email) with the placeholder 'edu邮箱', a text input for '验证码' (Verification Code) with the placeholder '请输入验证码' (Please enter verification code), and a '发送验证码' (Send Verification Code) button. At the bottom of the form is a '下一步' (Next Step) button.

● 个人注册账号密码丢失——忘记密码

您可以通过“忘记密码”功能找回您丢失的密码。具体操作为：点击登录弹框【忘记密码】或者注册页面的【忘记密码】，在重置密码页面选择重置密码方式，Edu 邮箱或者手机号，身份验证完成后，输入新密码，系统验证通过即可重新登录。

找回密码方式 edu邮箱 手机号码

1 身份验证 2 重置密码 3 完成

*edu邮箱 邮箱

*验证码 请输入验证码 发送验证码

下一步

4.2.2 首页

登录成功，进入网站首页，首页主要包括 CSMAR 最新滚动资讯、功能模块介绍、产品动态、和现有客户 4 个模块，其中：功能模块包括 Python 数据接口、跨表查询、单表查询三个模块的功能介绍和快捷入口。

4.2.3 数据中心_单表查询

数据中心单表查询模块提供 CSMAR 系列数据的查询下载服务，您可以进行字段搜索、数据下载、保存方案等操作；在此模块中，CSMAR 数据库结合实证研究专题，按研究方向进行数据分类，满足不同研究者的需求。CSMAR 数据库目前包含 **18 个系列**和 **160+个子数据库**，同时我们的数据库也在不断更新。其中 18 个系列分别为：

因子研究系列/绿色经济系列/人物特征系列/银行研究系列/股票市场系列/公司研究系列/基金市场系列/债券市场系列/衍生市场系列/经济研究系列/行业研究系列/货币市场系列/海外研究系列/板块研究系列/市场资讯系列/专题研究系列/科技金融系列/商品市场研究系列。

1. 关键字搜索（支持模糊搜索）

在搜索框中输入关键字，搜索框下显示包含关键字的十个相关字段、数据表、数据库信息，点击搜索，包含此关键字的字段/表/数据库结果将查找出来。比如输入“股票”，点击“股票价格”进行搜索，在弹出来的对话框中显示搜索的结果。如图所示：



以下为搜索到包含“股票价格”的信息，您可以点击名称快速进入：

字段结果	表结果	库结果
字段名 - 来源表 - 来源库		
股票价格 - 股票支付表 - 并购重组		
股票价格 - 股票支付表 - 新三板并购重组		
股票价格 - 股票风格表 - 股票市场衍生指标		
股票价格 - 前十名股东持股变动情况表（季） - 行为金融		
股票价格 - 股票支付表 - 海外直接投资		
股票价格 - 股票支付表 - 家族企业		
股票价格指数 - 股票价格指数（年） - 一带一路		
标的股票价格 - 权证定价重要参数表 - 权证市场		
合并股票价格 - 吸收合并表 - 并购重组		
合并股票价格 - 吸收合并表 - 新三板并购重组		

您可以通过点击字段结果、表结果或者库结果的信息，进入相应页面。如点击字段结果的“股票价格”字段，树节点展开，打开字段所在表，字段高亮显示。如下图所示：

海外并购重组 — 支付情况 — 股票支付表

本表数据总记录数：181 条

字段设置

请输入关键字进行字段搜索

搜索

选择： 全选 5/6

已选：1/6 全部删除

事件ID

支付金额

股票价格

发行数量

买方相对卖方换股比例(%)

股票来源编码

2. 数据查询

如果您希望查看某个数据库的数据，只需点击【数据中心】- “单表查询” 指定系列界面的某个数据库名称，将会进入当前数据库的数据查询页面。有权限的数据库为黑色字体显示，无权限的数据库为灰色字体显示。

首页 / 数据中心 / 单表查询

字段/表名/库名

单表查询 跨表查询

显示设置 全部

热门推荐 热门数据库

银行研究 有权限数据库

分析师预测

股票市场系列

债券市场

因子研究系列

宏观经济学

公司治理

资本结构

公司研究系列

基本分析

人物特征系列

财务报表附注

基金市场系列

股东

债券市场系列

首次公开发行 (A股)

衍生市场系列

增发配股

经济研究系列

股票市场交易

股票市场衍生指标

绿色经济系列

内部控制

并购重组

市场指数 无权限数据库

违规处理

财务报告审计意见

财务报表

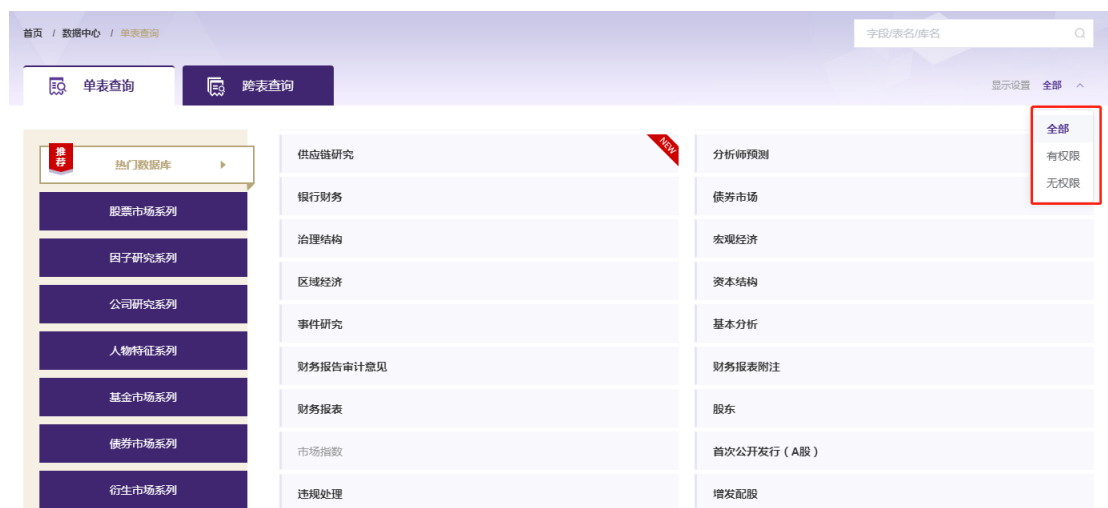
事件研究

区域经济学

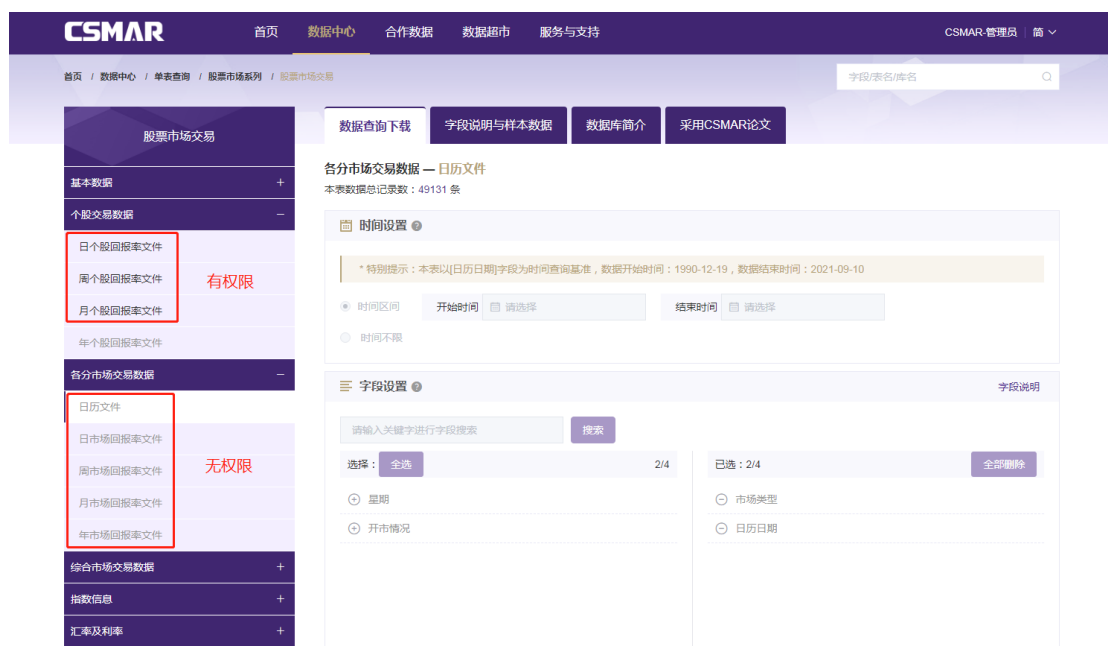
治理结构

银行财务

如果您只需查看有权限或无权限的数据库，可进行显示设置为“有权限”或“无权限”，默认是全部显示。



数据库的数据查询页面展示当前数据库的数据表信息，其中：黑色字体显示有权限的数据表，灰色字体显示无权限的数据表。



3. 时间设置

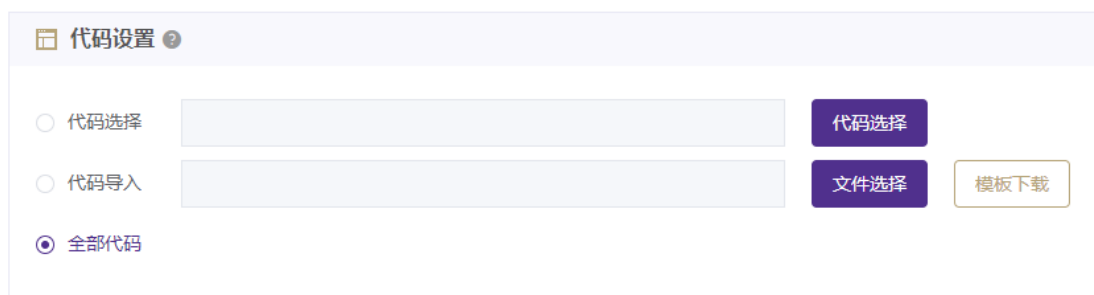


设置所需查询的数据时间区间，无权限的数据表不可进行时间设置；有权限的数据表时

间设置范围是权限内的时间区间。针对一些数据量较大的数据表，限制每次查询数据的年份。

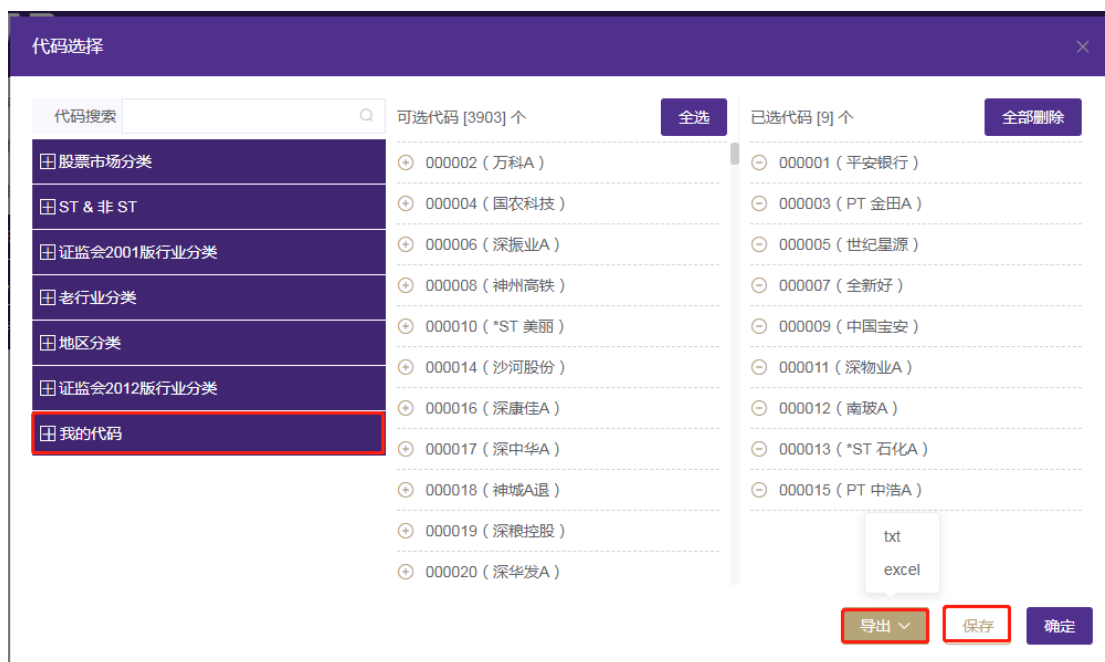
4.代码设置：

您可以通过以下三种方式进行代码设置：全部代码、代码选择、代码导入（请严格按照代码模板进行导入）：



代码选择：点击【代码选择】，选择代码分类（可以组合分类），树节点有【我的代码】分类。

- **可选代码区域：**点击“全选”，默认将可选代码区域的代码全部选中移至已选代码区域；点击可选区域的某个代码，将自动移至已选代码区域。
- **已选代码区域：**点击“全部删除”，默认将已选代码区域的代码全部选中移至可选代码区域；当点击已选代码区域的某个代码，将自动移至可选代码区域。
- **代码导出和保存：**可以对已选代码区域的代码进行导出，导出格式有 TXT 和 EXCEL 两种，同时，还能将已选代码区域的代码保存，输入代码方案名称即可保存。在左侧【我的代码】模块进行查看历史保留的代码方案。**注：**【我的代码】只保存的 10 条记录，如果需要新增代码方案，需删除历史记录后才能进行保存。



代码导入：点击【文件选择】按钮，在弹出对话框中选中本地代码导入文件。请严格按照【模板下载】下载的模板进行导入。

全部代码：在所有代码范围内进行检索。

5. 字段选择:

选择需要查询下载的字。您可以通过以下操作进行字段的选择和删除：

- 单击字段项或者点击【全选】按钮进行字段选择。
- 单击字段项或者点击【全部删除】按钮进行取消选择。

如果当前数据表包含字段太多，您可以通过在输入框内输入您想查找的字段关键字进行实时搜索。如需了解当前数据表的字段说明，可点击【字段说明】进行查看。



6.条件筛选:

如果您想对查询数据设定筛选条件,可以通过设置条件方式进行条件限定,以满足符合一个或者多个条件组合的数据结果。 **例如**:财务报表-资产负债表中,我们希望查出在2018-2019年年报的财务数据。操作步骤:

(1) 选择字段: 会计期间

(2) 选择条件运算符: 包含

(说明:针对不同类型的数据,系统自动判别加载运算符符合相应的取值输入控件,比如日期,将弹出日期选择窗体,避免您误输入错误格式的日期,减少操作错误的几率。)

(3) 条件取值: 设置为____12-31。

(4) 点击添加按钮,则此条件添加到条件列表中。

条件筛选
?

字段

运算符

条件值

单位: [没有单位]

字段类型: [时间]

添加

序号	字段	运算符	条件值	单位	条件关系	操作
1	会计期间	包含	____-12-31		AND	删除

注:您在作多个条件组合筛选时,选择的条件间组合默认是 AND 的关系。您可以自行设置,查询数据的计算逻辑是:(((序号 1 序号 2 的条件关系 序号 2)序号 3 的条件关系 序号 3)序号 4 的条件关系 序号 4) ...

7. 下载数据

设置好查询条件后,点击【下载数据】。新页面打开数据下载概要,如下图:

🔗
下载数据概要

下载表名	资产负债表
数据区间	1990-12-31 至 2020-06-30
选择代码	全部代码
输出类型	Excel2007格式 (*.xlsx)
选择字段	证券代码[Stkcd] 会计期间[Accper] 报表类型[Typrep] 货币资金[A001101000] 结算备付金[A0d1102000] 现金及存放中央银行款项[A0b1103000] 贵金属[A0b1105000] 拆出资金净额[A0f1106000] 衍生金融资产[A0f1108000]
筛选条件	[Accper] LIKE '%____-12-31%'

下载提示

- 此次下载数据仅在站内保存48小时
- 点击鼠标右键“链接另存为...”即可保存
- 数据输出完成点击下方链接进行保存

资产负债表134931432
.zip (2.9MB, 共 81854 条)

您将看到具体的条件设置信息,同时,下方显示查询数据的下载压缩包。点击压缩包名称,可直接下载数据到本机。**注:**下载数据仅保存 48 小时,您可以在【个人中心-下载记录】模块找到历史下载信息。

8. 保存方案和下载到邮箱

如果您想保存查询条件设置,可以在数据查询页面选中“保存方案”,同时输入方案名称,即可在【个人中心-我的方案】模块进行查询调用。

您还能直接选中“邮箱”，输入邮箱地址，系统将自动发送此次查询下载的数据包链接到您输入的邮箱，您可以通过登录邮箱，点击链接自动获取压缩包，邮箱链接只保存 48 小时。

邮箱

不支持gmail邮箱和hotmail邮箱

保存方案

保存方案可在【个人中心-我的方案】页面查询

[下载数据](#) [预览数据](#) [重置](#)

9. 字段说明与样本数据、数据库说明书、采用 CSMAR 论文

如果您想查看某个数据表的样本数据或字段说明，请点击【[字段说明和样本数据](#)】。或者，您可以通过点击【[数据库说明书](#)】查看当前数据库的简介和下载数据库说明书。当有此数据库的表权限时，才能下载数据库说明书。

CSMAR 首页 数据中心 合作数据 数据超市 服务与支持 CSMAR-管理员 | 简

首页 / 数据中心 / 单表查询 / 股票市场系列 / 股票市场交易

数据查询下载 **字段说明与样本数据** 数据库简介 采用CSMAR论文

个股交易数据 — 日个股回报率文件

样本数据

证券代码	交易日期	日开盘价	日最高价	日最低价	日收盘价	日个股交
603897	2019-01-02	23.01	23.28	22.74	22.81	2176373
603898	2019-01-02	15.33	15.7	15.11	15.33	852951
603899	2019-01-02	30.15	30.24	28.04	28.86	3435768
603900	2019-01-02	11.88	11.99	11.71	11.82	629531
603901	2019-01-02	6.59	6.69	6.59	6.62	968100
603903	2019-01-02	20.84	21.5	20.12	20.15	1865700
603906	2019-01-02	8.65	8.72	8.55	8.63	1106160
603908	2019-01-02	24.38	24.68	23.88	24.51	963600
603909	2019-01-02	25.52	26.12	25.41	25.7	1089096
603912	2019-01-02	13.32	13.95	13.32	13.7	4097856
603916	2019-01-02	11.67	11.9	11.46	11.47	5060825
603917	2019-01-02	12.08	12.17	11.96	12.01	1209320
603918	2019-01-02	11.88	12.03	11.65	11.99	592202

字段说明

序号	字段	数据类型	字段标题	字段说明
1	Stkcd	Nvarchar	证券代码	以上交所、深交所公布的证券代码为准
2	Trdtd	Datetime	交易日期	以YYYY-MM-DD表示
3	Opnprc	decimal	日开盘价	A股以人民币元计，上海B以美元计，深圳B以港币计

如果您想了解采用数据库发表的文献情况，可点击【采用 CSMAR 论文】查看，目前主要展示《会计研究》、《经济研究》、《金融研究》等期刊中采用 CSMAR 数据的文献列表。



期刊名称	文献名称	发表年份	相关数据库
Journal of Corporate Finance	Cross-border acquisitions and employment policies	2020	财务报表 / 股票市场交易
Journal of Corporate Finance	Corporate board reforms around the world and stock price crash risk	2020	股票市场交易
Journal of Corporate Finance	Do technology spillovers affect the corporate information environment?	2020	股票市场交易 / 业绩预告
Journal of Corporate Finance	Is language an economic institution? Evidence from R&D investment	2020	财务报表 / 股票市场交易 / 上市公司研发创新
Journal of Corporate Finance	The impact of monetary policy on M&A outcomes	2020	股票市场交易 / 财务指标分析
Journal of Corporate Finance	Protection of proprietary information and financial reporting opacity: Evidence from a natural experiment	2020	股票市场交易
Journal of Corporate Finance	It's who you know that counts: Board connectedness and CSR performance	2020	治理结构 / 财务报表 / 股票市场交易 / 财务指标分析
Journal of Corporate Finance	So far away from me: Firm location and the	2020	股票市场交易

4.2.4 数据中心_跨表查询



指标列表

可选指标

请输入关键字

- 股票市场系列
- 因子研究系列
- 公司研究系列

已选指标[0/50]

请从左侧“可选指标”区域选择需要的指标

提示:

- 单击“可选指标”区域指标名称,即成功选择指标,“已选指标”区域同步显示此指标。
- “已选指标”区域只能通过(删除)减少指标。

数据中心_跨表查询模块实现了 CSMAR 不同数据表中指标的自由组合，内容主要涉及股票市场、公司研究、基金研究、经济研究等研究系列，包含 80+个数据库（如股票市场交易、财务报表、财务指标分析、治理结构、分析师预测、家族企业等）、700+张表、10000+个字段。

如果只需查看有权限或无权限的数据库，可进行显示设置为“有权限”或“无权限”，

默认是全部显示。

1. 指标选择

选择需要查询下载的字。您可以通过以下操作进行字段的选择和删除：

- 单击左侧可选指标区域的字段项进行字段选择。
- 单击左侧可选指标区域的字段项或者点击右侧已选指标区域的【删除】按钮进行已选指标删除。

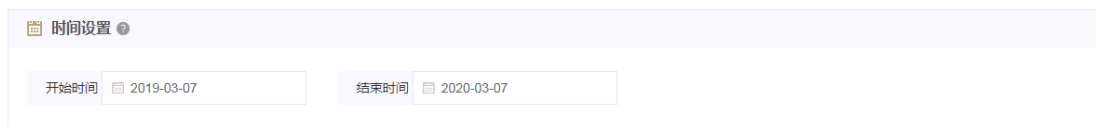
股票和基金相关指标不能同时选择。如果某个数据表有多种数据类别，比如财务报表有“合并报表”和“母公司报表”两种类别，可以点击右侧数据表后面灰色字段进行设置。



如果希望快速查询指标，可在搜索框中输入关键字，包含此关键字的字段/表/数据库结果将查找出来。比如输入“资产负债”，如图所示：



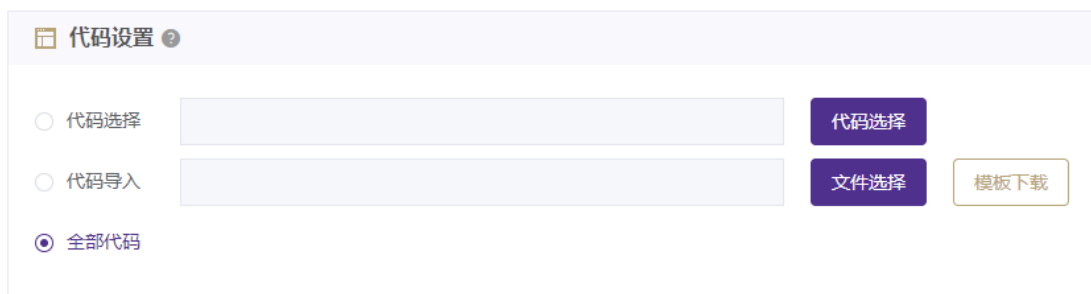
2. 时间设置



设置所需查询的数据时间区间，如果选择的指标来源数据表中包含数据量较大的数据表，限制每次查询 3 年的数据。

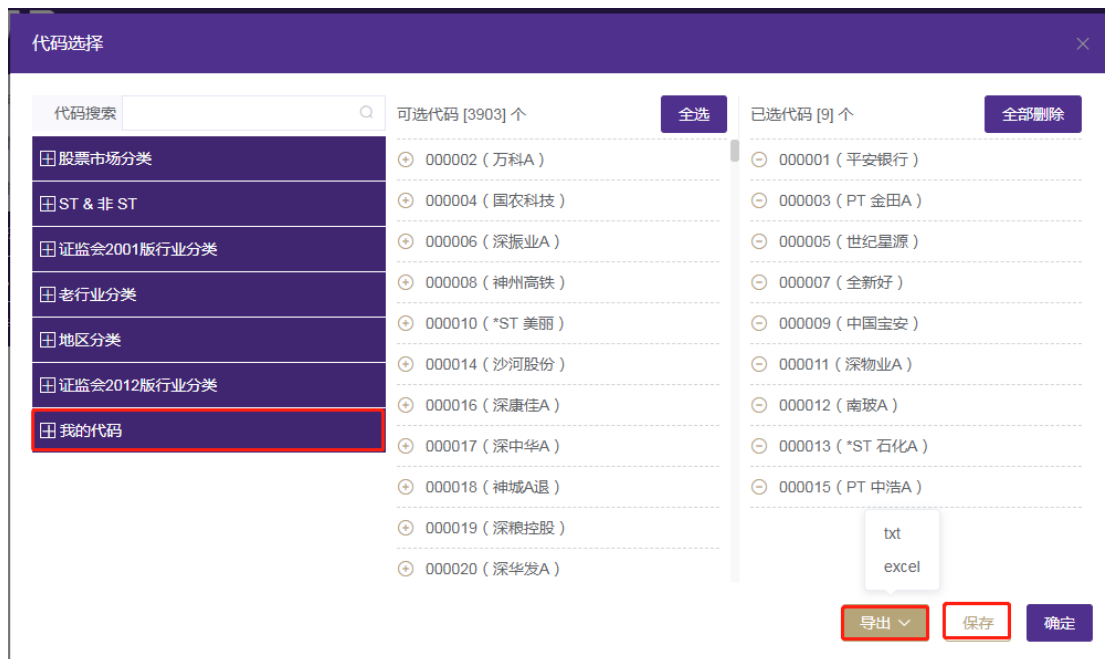
3. 代码设置：

您可以通过以下三种方式进行代码设置：全部代码、代码选择、代码导入（请严格按照代码模板进行导入）：



代码选择：点击【代码选择】，选择代码分类（可以组合分类），树节点有【我的代码】分类。

- **可选代码区域：**点击“全选”，默认将可选代码区域的代码全部选中移至已选代码区域；点击可选区域的某个代码，将自动移至已选代码区域。
- **已选代码区域：**点击“全部删除”，默认将已选代码区域的代码全部选中移至可选代码区域；当点击已选代码区域的某个代码，将自动移至可选代码区域。
- **代码导出和保存：**可以对已选代码区域的代码进行导出，导出格式有 TXT 和 EXCEL 两种，同时，还能将已选代码区域的代码保存，输入代码方案名称即可保存。在左侧【我的代码】模块进行查看历史保留的代码方案。**注：**【我的代码】只保存的 10 条记录，如果需要新增代码方案，需删除历史记录后才能进行保存。

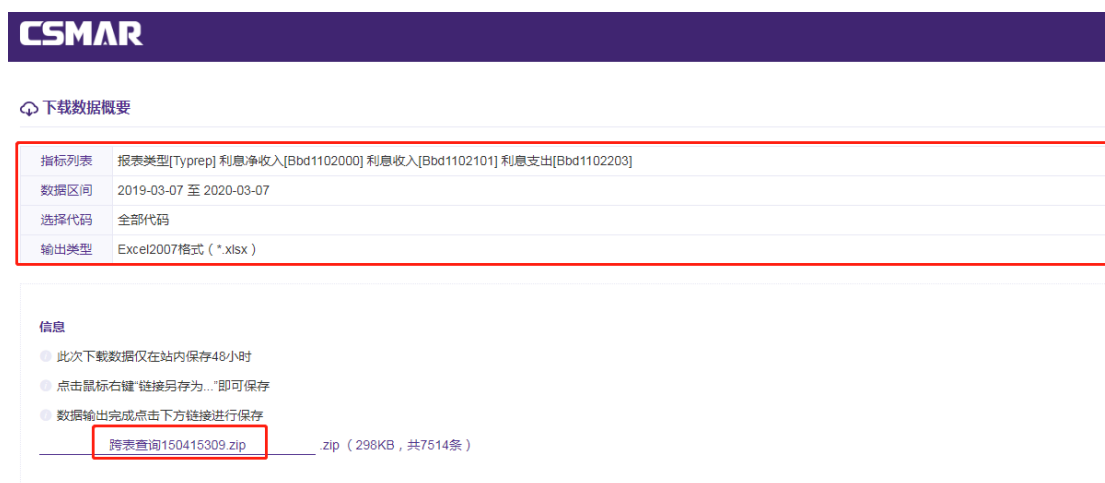


代码导入：点击【文件选择】按钮，在弹出对话框中选中本地代码导入文件。请严格按照【模板下载】下载的模板进行导入。

全部代码：在所有代码范围内进行检索。

4. 下载数据

设置好查询条件后，点击【下载数据】。新页面打开数据下载概要，如下图：



您将看到具体的条件设置信息，同时，下方显示查询数据的下载压缩包。点击压缩包名称，可直接下载数据到本机。**注**：下载数据仅保存 48 小时，您可以在【个人中心-下载记录】模块找到历史下载信息。

5.下载到邮箱

您还能直接选中“邮箱”，输入邮箱地址，系统将自动发送此次查询下载的数据包链接到您输入的邮箱，您可以通过登录邮箱，点击链接自动获取压缩包，邮箱链接只保存 48 小时。

邮箱

不支持gmail邮箱和hotmail邮箱

[下载数据](#) [预览数据](#)

4.2.5 合作数据



合作数据模块展示我司与专家学者共同开发的数据库信息，如果您希望与我们合作，可点击“我有数据，我要合作”，填写您的数据内容，提交成功后我司将有专人与您进行合作洽谈。合作数据的数据查询下载相关模块设置同 4.2.3 数据中心-单表查询。

WinGo 财经文本数据介绍

WinGo 财经文本分析数据平台是一整个大的文本分析的模块，里面有几十种不同的文本分析的应用和基于文本分析获得的数据库。WinGo 财经文本数据平台是中国首家基于中美上市公司披露文本的人工智能财经数据平台，包含中美上市公司的文本分析，中国政府文件的文本分析，新冠疫情以及专利的数据库。平台从学术研究需求出发，针对中美两国截然不同的文本披露规则和财经文本特点，应用自然语言处理、深度学习和人工智能技术对财

经文本进行深度加工，给用户 提供财经文本的词频、相似词、文本特征等全新深度处理的数据，从而为学术研究、投资决策应用等提供多方位支持。

点击导航栏右侧【登录】，在登录弹框中输入个人在线注册账号/学校公共账号、密码登录。



1) 权限说明

WinGo 财经文本数据平台仅限正式用户及试用用户使用。其中正式用户指用户所在单位/个人购买了该平台；试用用户指用户所在单位/个人申请试用了该平台。

2) 平台功能操作说明

词频搜索

单词词频指某个单词在文本中出现的频率。单词词频是词频数据库实现的最基础功能。

单词词频数据库除向用户提供单词的词频之外，还向用户提供了语料的基本信息表。以财务报告为例：

第一步：选择词频或基本信息表，其中更常用的是选择词频。

第二步：输入关键词，即需要统计的在文本中出现的词汇，按回车确定。可输入 200 个以内。

第三步：选择报告类型，可以多选。财务报告分一季度财报、中期财报、三季度财报和年度财报四种。

第四步：选择代码，对应需要研究的上市公司。

第五步：选择时间范围。

第六步：选择字段内容，分精确词频，扩展词频，精确句频，扩展句频四个。具体含义为：

精确词频：该词汇的精确词频，“精确”指关键词本身在报告中出现的次数，如报告中“资产”这一词汇出现的次数。扩展词频：该词汇的扩展词频，“扩展”指包含关键词的所

有词汇出现的总次数，如报告中“资产”的所有扩展词汇（如“资产”、“资产负债表”、“资产收益率”等）出现的次数。精确句频：该词汇的精确句频，“精确”指包含关键词本身的句子在报告中出现的次数，如报告中包含“资产”这一词汇的句子出现次数。扩展句频：该词汇的扩展句频，“扩展”指包含关键词扩展词汇的句子出现的总次数，如报告中包含“资产”扩展词汇（如“资产”、“资产负债表”、“资产收益率”等）的句子出现次数。

完成以上 6 步后点击查询即可得到相应的结构化数据。

邻近搜索

“邻近搜索”指的是用户将词语间的间隔字符数限定为 n 后（ n 的范围为 0-15），数据库将输出词语间间隔字符数小于等于 n 个字符数的词组的频率。

第一步：选择功能，选择“邻近搜索”进行搜索，选择“基本信息表”显示文本基本信息。

第二步：输入词组。

输入词组对，以空格隔开，按回车确认。

第三步：制定词组间相隔字符数（0-15）。

第四步：选择报告类型（财务报告分第一季度，中报，第三季度和年报）。

第五步：选择代码，即相应的上市公司。

第六步：选择时间范围。

最后点击保存会弹出对话框，供用户选择下载格式和收取数据的邮箱。

高级搜索

“高级搜索”指的是用户可对语料中的指定词汇设置搜索条件，包括 And, Not, Or, Near 等，从而获取出现符合搜索条件的指定词汇的文件列表和语句信息。对于包含英文字母的关键词，本系统默认为小写；关键词只针对两个及两个以上字符的词汇（纯数字除外）。逻辑

运算优先级顺序为：括号内为第一优先级，括号外 NOT>AND>OR。

第一步：输入关键词。输入需要搜索的条件，样例如下：

第一步：输入关键词 

AND	(价格	AND	价值)	⊕ ⊖
AND	(股票	AND	债券)	⊕ ⊖
AND	(经济	AND	金融)	⊕ ⊖

第二步：选择逻辑运算的范围，可以是全文或者单个句子。

第三步：选择财务报告类型。

第四步：选择上市公司代码。

第五步：选择时间区间。

最后点击保存，选择格式和输入邮箱即可获得查询数据。

深度学习相似词

深度学习相似词是利用神经网络构建出的词语相似度计算模型，其可计算出每个词语与其他词语之间的相似度大小。相似度越大，则这两个词的语义越相近、句法越相似；反之，则它们之间的语义、句法差异越大。通过对每个词的相似词词集的检测并根据相似度大小排序，我们取每个词相似度最高的前 30 个词作为该词的相似词词集结果。

第一步：输入关键词，按回车确定，可以输入 200 个以内关键词。

第二步：选择是否保留相似词推荐结果中的关键词扩展词。以“风险”关键词为例，假定“风险”的相似词为“经营风险”、“潜在风险”、“可能性”、“损失”，如果选择不保留相似词推荐结果中的关键词扩展词，那么相似词推荐结果就是“可能性”、“损失”；反之，则相似词推荐结果为“经营风险”、“潜在风险”、“可能性”、“损失”。因为“经营风险”、“潜在风险”都是“风险”的扩展词汇。以此类推。搜索“价值”，保留关键词的结果如下：

查询结果预览

关键词	相似词	相似度	词频
价值	市场价值	0.6812	2,679
价值	商业价值	0.5467	1,029
价值	公司价值	0.5403	4,957
价值	资产价值	0.5382	2,536
价值	企业价值	0.5288	2,272
价值	实际价值	0.5055	565
价值	经济价值	0.4930	794
价值	评估价值	0.4930	15,275
价值	投资价值	0.4898	3,176
价值	内在价值	0.4818	1,295

不保留关键词的结果如下：

查询结果预览

关键词	相似词	相似度	词频
价值	市值	0.4785	10,083
价值	净值	0.4547	23,153
价值	增值	0.4445	12,583
价值	作价	0.4316	18,297
价值	账面值	0.4264	4,652
价值	账面净值	0.4152	8,541
价值	评估价格	0.4033	1,555
价值	估值	0.4019	8,462
价值	评估价	0.3833	2,460
价值	售价	0.3694	15,654

同义词词林

同义词词林是基于传统主流同义词词林结构体系实现的同义词推荐。同义词指的是词义完全相同或者相近的词语。输入一个关键词，返回该关键词的一个或多个同义词结果。输入关键词即可进行查询，例如输入价格和价值，查询结果如下：

查询结果预览

关键词	相似词
价格	价钱
价格	价
价格	标价
价格	价位
价值	价
价值	值

语义相似词

语义相似词结合传统知识库，充分利用词语在不同知识库中的语义信息，根据它们的分布情况，按照一定的策略综合利用传统知识库中的语义信息，从而分别计算出词语的两个相似度，然后按照一定的算法计算出综合的词语语义相似度。在查询界面，分别输入两行关键词（——对应），返回这两行关键词各自对应的相似度大小值。同时输入多组词组，需要计

算的词组对一一对应。例如关键词 1 输入“价格”和“回报”，关键词 2 输入“价值”和“收益”，点击计算相似度，结果为：

查询结果预览

关键词	相似词	相似度
价格	价值	0.5747
回报	收益	0.0429

自定义特征

自定义特征数据库集成了 WinGo 词频数据库与 WinGo 深度学习相似词数据库中的两大基础功能，旨在为用户提供一站式的文本指标的构建工具。具体来讲，自定义特征指标的构建流程分为三步：（1）用户根据已有的研究理论，定义所构建指标的原始词集（又称种子词集）；（2）用户使用系统集成的 WinGo 深度学习相似词推荐功能对种子词集进行相似词扩充；（3）系统自动计算自定义指标词集中每个词的词频，加总归一化得到最终文本指标。

根据以上自定义指标构建流程，自定义数据库提供“特征词典定制”和“特征计算”两大模块。其中，“特征词典定制”模块用于实现自定义指标构建流程的（1）和（2），通过集成“种子词集输入+深度学习相似词扩充”功能，完成用户自定义文本指标的词典构建；“特征计算”模块用于快捷实现自定义指标流程的（3），提供用户已有自定义文本指标的自动计算和下载功能。可供计算下载的文本指标对应的数据源选择范围涵盖包括财务报告在内的系统所有内置数据源。

第一步：特征词典定制

指标名称	操作		
指标1	修改名称	下载词集	删除
指标2	修改名称	下载词集	删除
指标3	修改名称	下载词集	删除

[+ 添加指标](#)

可以删除已有指标或者添加新的指标。举例如下：

新建自定义指标 ×

指标4

输入种子词（回车键确认输入种子词）

价格 ×
价值 ×
价钱 ×
×

回车键确认输入关键词，总词数不超过1000个
关键词批量导入

取消
确定

定义好以后可以在相似词列表中看到，同时可以针对每一个关键词查找相似词：

是否保留相似词推荐结果中的关键词扩展词

是 否

*说明：以“风险”关键词为例，假定“风险”的相似词为“经营风险”、“潜在风险”、“可能性”、“损失”，如果选择不保留相似词推荐结果中的关键词扩展词，那么相似词推荐结果就是“可能性”、“损失”；反之，则相似词推荐结果为“经营风险”、“潜在风险”、“可能性”、“损失”，因为“经营风险”、“潜在风险”都是“风险”的扩展词汇，以此类推。

关键词列表

关键词名称	来源	操作	
价格	种子词	查找相似词	删除
价值	种子词	查找相似词	删除
价钱	种子词	查找相似词	删除

第二步：特征计算

定义好关键词后就可以使用该关键词进行特征计算。选择指标，选择相应的文本，点击

确定就进入数据查询页面：

中国上市公司文本数据库
自定义特征
特征计算

自定义特征

指标1
 指标2
 指标3
 指标4

财务报告

财务报告是指企业对外提供的反映企业某一特定日期财务状况和某一会计期间经营成果、现金流量的文件。目前国内外文针对财报文本的研究涵盖会计、金融、战略管理、管理科学与工程等多个领域，研究的文本特征包括可读性、语确、语义、文本相似性、风险、创新、诚信、前瞻性、业绩自利性归因、自我指涉度等。

董事会报告章节

董事会报告是对财务报告中的数据和其他必要的统计数据以及报告期内发生和未发生的重大事项进行讨论和分析的报告文件，受到投资者的高度关注和重视。根据中国证监会颁布的上市公司信息披露内容与格式准则，从2015年开始，董事会报告章节更改为经营情况讨论与分析。

中国上市公司文本数据库 > 自定义特征 > 指标4 > 财务报告

第一步：关键词列表（最多显示前100条）

价格 价值 价钱

第二步：财务报告类型选择

一季度财务报告 中期财务报告 三季度财务报告 年度财务报告

第三步：代码选择

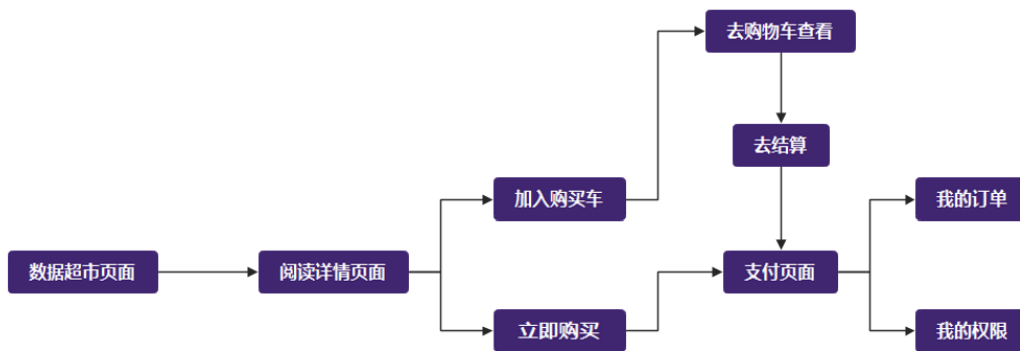
请选择股票代码

格式要求：按照单列每行一个股票代码的格式进行排列，文件支持txt格式

再按照之前的介绍进行相应的设置和查询，即可得到结果。

4.2.7 数据超市

数据超市模块提供 CSMAR 数据表的在线购买，目前仅支持个人在线注册用户购买，针对使用学校机构账号的用户，需注册新用户（登录页面 cn.gtadata.com 点击【注册】按钮进行注册）才能购买。购买数据库或者数据表的流程如下：



1. 数据超市页面

主菜单点击【数据超市】进入该模板，页面布局展示如下：



页面左侧为 CSMAR 数据库 19 个系列分类（数据库来源于 18 个数据库系列+合作数据），当有优惠活动时会展多一个特惠活动专区。如，点击股票市场系列，页面右侧区域则会显示股票市场系列下包含的所有数据库。在每个数据库模块中，展示有数据库特色、使用期限（30 天）对应单表最低价格、阅读详情按钮。点击【阅读详情】按钮，或者点击数据库名称，即可在当前页面跳转至数据库详情展示页面。

您可在数据库模块列表右上方搜索框内输入需求关键词，即可快速定位关键词相关的数据库/表。例：在右上方的搜索框内输入“分析师”，点击搜索，页面右侧区域将显示数据库名称或表名称中包含“分析师”关键词的数据库模块列表。



2. 阅读详情页面

在数据超市页面点击列表下数据库名称或者【阅读详情】按钮即可从当前页面跳转到数据库详情展示页面，具体页面内容展示如下：

CSMAR 首页 数据中心 合作数据 数据超市 服务与支持 CSMAR-管理员 简

首页 / 数据超市 字段/表名/库名

环境研究 数据库说明

产品介绍

主要收录了上市公司的环境绩效（绩效明细、排放明细、资源消耗明细）；环境披露信息（环境管理、披露载体、环境负债披露、环境业绩与治理）；环境投资信息（投资明细、投资统计）。 more

购买前须知

数据超市购买只提供发票不提供合同（如需要合同请联系在线客服进行线下购买），购买前请确认数据内容，一旦购买成功，概不退款。购买成功后30天内可在数据中心使用已购买数据表的全部数据。

选择购买表单（可勾选多项，点击表名称查看样本数据和字段说明）

全选：¥40000

<input type="checkbox"/> 上市公司环境绩效明细表	¥4255	<input type="checkbox"/> 上市公司环境排放明细表	¥4255	<input type="checkbox"/> 上市公司资源消耗明细表	¥4255
<input type="checkbox"/> 上市公司环境管理披露情况表	¥4134	<input type="checkbox"/> 上市公司环境监管与认证披露情况表	¥4166	<input type="checkbox"/> 上市公司环境信息披露载体信息表	¥1415
<input type="checkbox"/> 上市公司环境负债披露情况表	¥4158	<input type="checkbox"/> 上市公司环境业绩与治理披露情况表	¥4116	<input type="checkbox"/> 上市公司环境投资明细表	¥5605
<input type="checkbox"/> 上市公司环境投资统计表	¥3639				

合计：¥0

立即购买 加入购物车 返回上一级

产品介绍：展示当前数据的信息介绍；

购买前须知：由于数据商品的特殊性，一旦购买，概不退款。购买成功后提供 30 天数据使用权限。例如：2016 年 5 月 20 日购买某些表 30 天的使用权限，2016 年 6 月 20 日之后，数据权限失效；

选择购买表单：该模块显示当前数据库包含的所有上架数据超市的数据表，在此模块用户可自主勾选所需要的数据表进行购买，同时页面下方将实时更新展示当前所勾选数据表的合计金额；

选择购买表单（可勾选多项，点击表名称查看样本数据和字段说明）

全选：¥11287

<input type="checkbox"/> 资产负债表	¥4508	<input type="checkbox"/> 利润表	¥2202	<input type="checkbox"/> 现金流量表(直接法)	¥1852
<input type="checkbox"/> 现金流量表(间接法)	¥1013	<input type="checkbox"/> 所有者权益变动表	¥1712		

合计：¥0

立即购买 加入购物车 返回上一级

立即购买：点击【立即购买】按钮，当前页面将跳转至填写订单页面，进入支付环节；

加入购物车：点击【加入购物车】按钮，则当前已勾选的数据表商品将成功加入购物车，此时可在弹出框中选择“去购物车查看”或“继续查看商品”；

返回上一级：点击【返回上一级】按钮，则当前页面下跳转回上一级页面，即数据超市系列模块展示页面。



3. 填写订单页面

用户信息填写

* 联系电话 18025343790

* 电子邮箱 请输入电子邮箱

留言

支付方式  支付宝 ALIPAY  微信支付 WeChat Pay

使用优惠 否 是

发票需求 否 是

合计： ¥11290

应付金额： ¥11290

去支付

合计：显示当前订单金额合计。

支付方式：目前支持微信支付和支付宝支付。

使用优惠：可以使用由 CSMAR 团队发送的数据超市优惠码支付，或者使用符合条件的活动期间领取的优惠券。通过证件审核的老师或者学生及 edu 邮箱注册的老师或者学生将享受在此优惠基础上至少 9 折优惠。

发票需求：默认是不开发票，如果需要开发票，则选择“是”，提供选择“增值税普通发票（纸质版）”、“增值税普通发票（电子版）”和“增值税专用发票（纸质版）”，当去除优惠金额小于 500 元时，如需开具**纸质**发票需支付 22 元邮寄费用。请按照要求输入发票及收票人相关信息。

发票需求 否 是 (开具发票金额为实际支付金额, 低于500元时, 需支付22元快递费, 默认顺丰速运)

增值税普通发票 (纸质版) 增值税普通发票 (电子版) 增值税专用发票 (纸质版)

* 发票抬头

* 纳税人识别号

* 收货人姓名

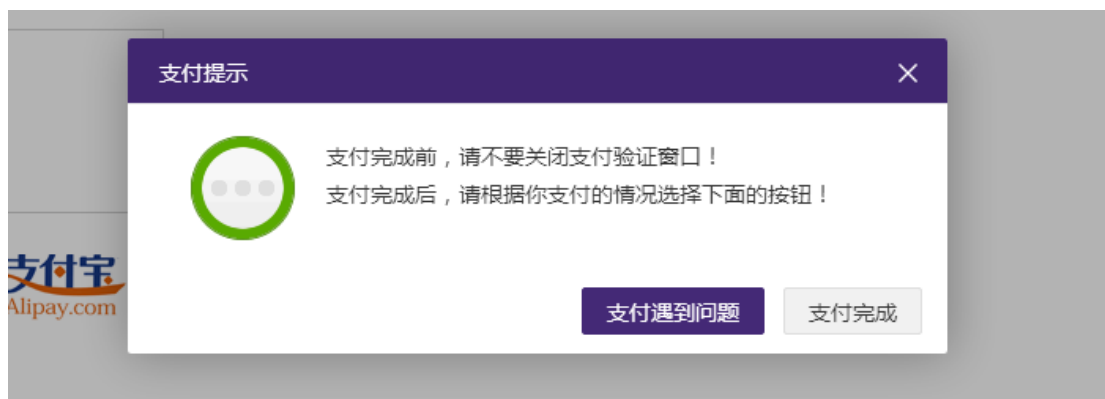
* 收货人联系电话

* 发票邮寄地址

温馨提示: 本公司开具发票内容为“信息技术服务*CSMAR数据库信息服务费”

去支付: 点击“去支付”, 跳转到第三方支付平台 (支付宝和微信在线支付) 页面。

成功支付后, 页面将进行如下文字提示: 恭喜您, 您的订单*****已经支付成功, 具体订单信息可在“我的订单”中查询!



4.2.8 服务与支持

为了让您享受更方便、快捷、周到的服务, 我们提供 Python/R/stata/matlab 数据接口文档、数据定制、公开数据、操作演示、联系我们和在线客服功能。

如果您是第一次使用我们的产品, 可点击右侧悬浮“操作演示”查看主要功能的使用说明。

如果您对我们的产品有任何疑问或者建议, 都可以第一时间拨打 400-639-8883 客服热线或者点击右侧悬浮“在线客服”与我司工作人员联系。

如果您是 Python/R/stata/matlab 使用者,我们也提供 Python/R/stata/matlab 数据接口可以直接调用有权限的 CSMAR 数据,提高您的研究效率。

The screenshot shows the CSMAR website's '数据接口文档' (Data Interface Documentation) page for Python. The page is structured as follows:

- Navigation:** 首页, 数据中心, 合作数据, 数据超市, 服务与支持. User: CSMAR-管理员.
- Left Sidebar:** 操作演示, 数据定制, 数据接口文档 (selected), 公开数据, 联系我们.
- Search:** 字段/表名/库名
- Language Tabs:** Python, R, Matlab, Stata.
- Main Content:**
 - 一. 功能介绍:** CSMAR-PYTHON 是基于Python语言的数据接口 (目前仅支持Windows系统), 以函数调用的形式实现CSMAR经济金融研究数据库数据查询、下载等功能, 方便用户在进行分析、实证研究时高效获取CSMAR数据。
 - 二. 初始配置:**
 - 安装python3.6.X工具【压缩包】, 并安装类库urllib3, websocket, websocket_client.


```
pip install urllib3
pip install websocket
pip install websocket_client
```
 - CSMAR-PYTHON安装方式
 - 下载CSMAR-PYTHON【压缩包】
 - 解压至python安装目录下的Lib/site-packages. 如:\Programs\Python36\Lib\site-packages

系统中所有的函数返回数据格式默认都是json, 可以通过列表打印类库(prettytable)来格式化数据以便查看

```
pip install prettytable
```
 - 三. 使用CSMAR-PYTHON:**
 - 使用CSMAR-PYTHON


```
from csmarapi.CsmarService import CsmarService
csmar = CsmarService()
```

如果需要以表格形式展示数据, 则需添加

```
from csmarapi.ReportUtil import ReportUtil
```
- Right Sidebar:** 一 功能介绍, 二 初始配置, 三 使用CSMAR-PYTHON, 四 示例.

公开数据

The screenshot shows the '公开数据' (Public Data) page on the CSMAR website. The page is structured as follows:

- Navigation:** 首页, 数据中心, 合作数据, 数据超市, 服务与支持. User: CSMAR-管理员.
- Left Sidebar:** 操作演示, 数据定制, 数据接口文档, 公开数据 (selected), 联系我们.
- Search:** 字段/表名/库名
- Official Statistical Data (官方统计数据):**
 - 联合国商品贸易统计数据
 - 全球可持续发展指标数据库
 - 服务贸易数据库
 - 世界发展指标数据库
 - 全球金融发展数据库
 - 环境、社会和公司治理数据库
- Academic Institution Data (科研机构数据):**
 - 中国老年健康影响因素跟踪调查数据
 - 中国企业创新创业调查数据
 - 中国家庭追踪调查数据
 - 中国健康与养老追踪调查
 - 中国家庭收入调查数据
 - 网民社会意识调查数据
 - 中国管理会计研究基础数据库
 - 中国家庭金融调查
 - 中国小微企业调查
 - 中国老年社会追踪调查数据
 - 中国综合社会调查数据
 - 中国教育追踪调查数据
 - 中国余寿调查数据
 - 中国社会状况综合调查数据
 - 中国劳动力动态调查数据
 - 北京大学数字普惠金融指数

展示经济金融相关的公开免费数据,包含官方统计数据 and 科研机构数据,点击数据库名称可查看数据库简介及来源链接。

数据定制

定制服务

我们针对具体的需求，提供专属于您的结构化数据，包括基础数据进行加工、分析（仅限简单的分析结果，不提供论文撰写服务）、衍生指标计算、模型定制等服务，节省您数据处理时间，有效提高研究效率。

付费定制流程

- 原始需求**
客户根据研究方向在线提交需求
- 需求评估&报价**
与客户详细沟通确认需求，进行报价及完成周期的反馈
- 样本提供&付款**
根据最终确认需求提供样本，客户确认无误即可进行付款（可提供发票）
- 数据交付**
确认收到款项后，客户可以在个人中心-我的定制 进行数据下载

定制信息填写

* 邮箱 * 频率

* 定制类型 * 数据区间 —

* 定制内容

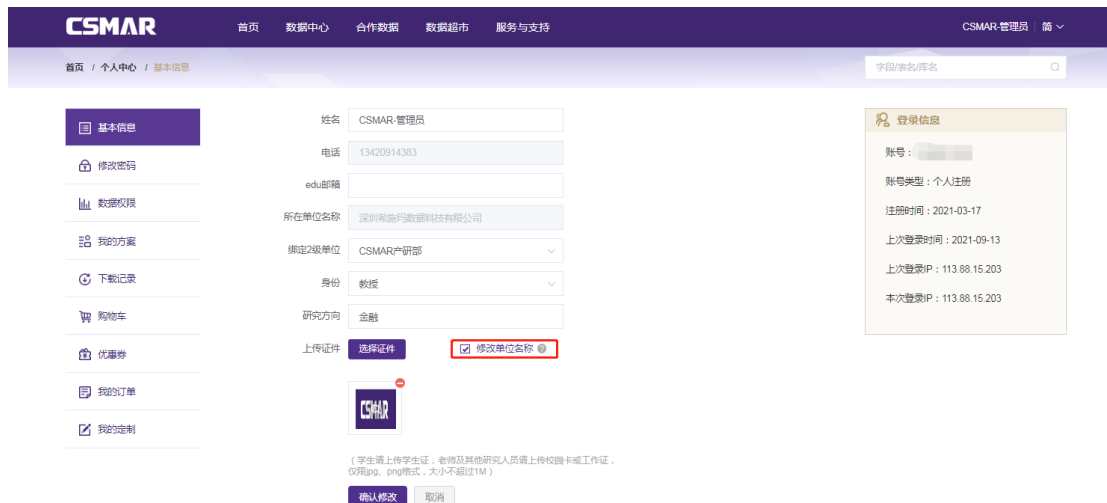
附件 仅限选择jpg、png、jpeg、bmp、gif、Word、Excel格式文件

若您提交申请则代表您已阅读并同意我们的服务协议
我们会在1-3个工作日内进行评估，并主动联系您，请及时查看邮箱&手机信息
如需求紧急，请直接联系在线客服或拨打服务热线400-639-8883

如果您需要购买数据库或者进行数据/模型定制，可以通过数据定制模块在线提交数据定制需求，填写您的邮箱、数据频率、定制类型、定制数据时间范围、定制内容等信息，点击【提交定制】，后期我司工作人员会以邮件形式和您沟通。

4.2.9 个人中心

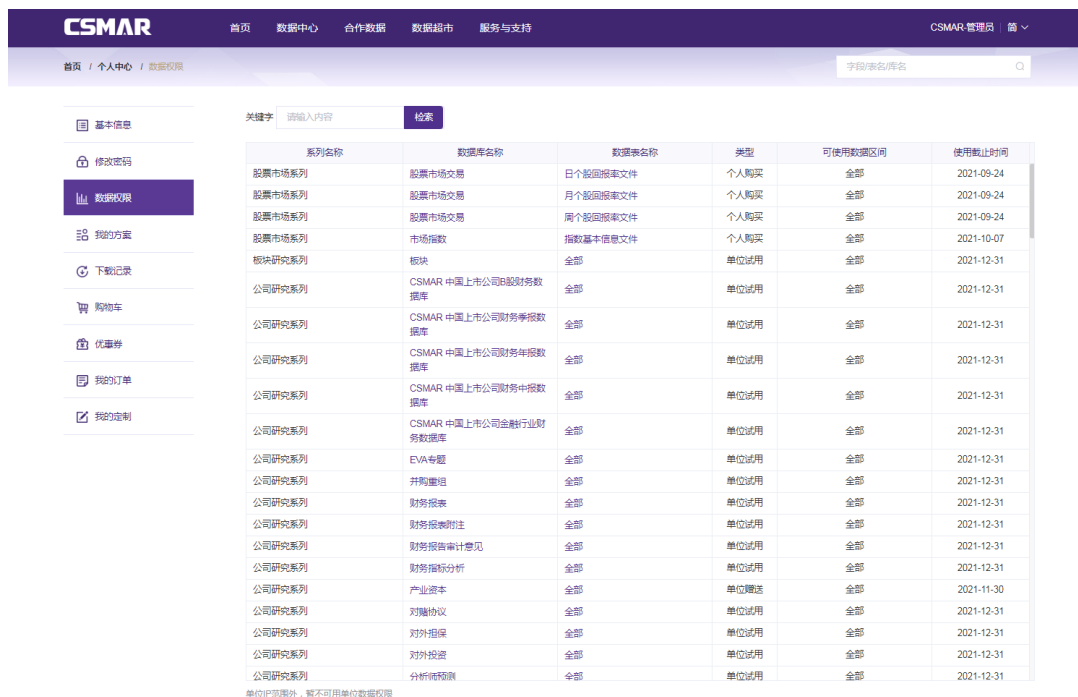
点击页面顶端【姓名】，进入个人中心页面，您可以在个人中心页面查看基本信息，修改登录密码，同时查看我的数据权限、我的下载、我的方案、我的定制、我的优惠券、我的购物车和我的订单。



1. 基本信息

个人注册账号可修改基本信息，电话邮箱一旦确定，不可修改；所在单位不可修改，edu 邮箱注册用户，如需修改，请重新使用最新学校 edu 邮箱注册；手机号注册用户请上传相应的证件，并勾选“修改单位名称”，提交成功后，我司工作人员会进行证件审核，审核通过后后台将自动帮您修改所在单位。如果一个单位有 2 个（及以上）2 级单位购买/试用 CSMAR 数据库，可根据需要进行修改绑定。

2. 数据权限



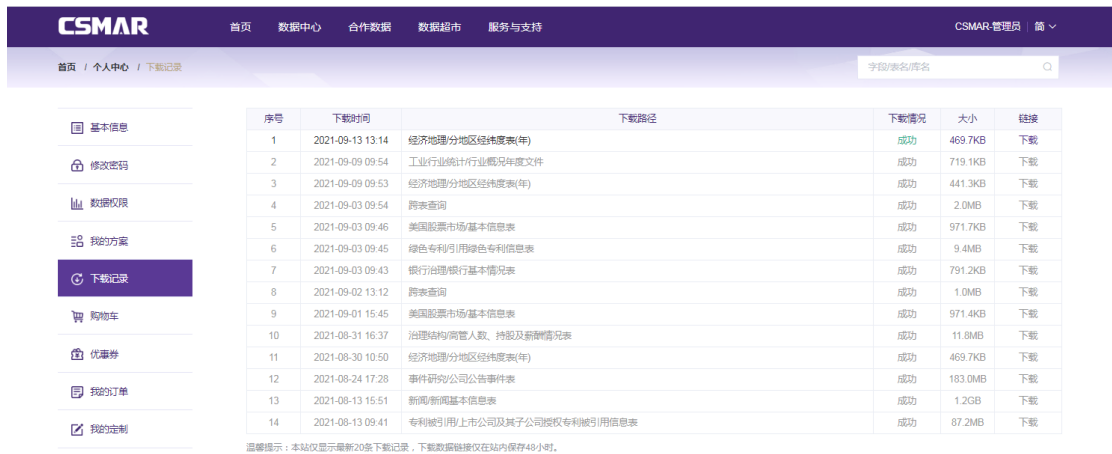
展示账号对应的数据权限,其中个人注册账号拥有所绑定单位的数据权限以及个人在数据超市购买的数据表权限,单位数据权限必须在单位 IP 范围内才能查看使用,IP 范围外不能使用。

3. 我的方案



我的方案包括对代码设置、时间设置、字段设置、条件筛选设置的保存,需要在数据中心_单表查询模块的数据查询过程中设置,即可展示在我的方案模块。

4. 下载记录



下载记录显示针对 CSMAR 数据表的下载记录,下载数据链接仅在站内保存 48 小时,最多显示 20 条下载记录。

5. 购物车

CSMAR 首页 数据中心 合作数据 数据超市 服务与支持 CSMAR-管理员 简

首页 / 个人中心 / 购物车

购物车

全选 批量删除 已选产品: 2 项 合计: ¥13213 去结算

可以单击“修改”修改当前数据库下勾选的数据表；同时也能删除购物车的数据库；点击“去结算”进入填写订单页面。

6. 我的订单

CSMAR 首页 数据中心 合作数据 数据超市 服务与支持 CSMAR-管理员 简

首页 / 个人中心 / 我的订单

订单状态: 请选择 下单时间: 开始日期 至 结束日期 查询

已支付 订单号: 20210907113154800757872 下单时间: 2021-09-07 11:31:55 支付时间: 2021-09-07 11:32:18

市场指数 使用期限30天 金额: ¥504

指数基本信息文件 ¥504

实际支付金额: ¥504

已支付 订单号: 20210825101029800757872 下单时间: 2021-08-25 10:10:29 支付时间: 2021-08-25 10:10:29

股票市场交易 使用期限30天 金额: ¥2867

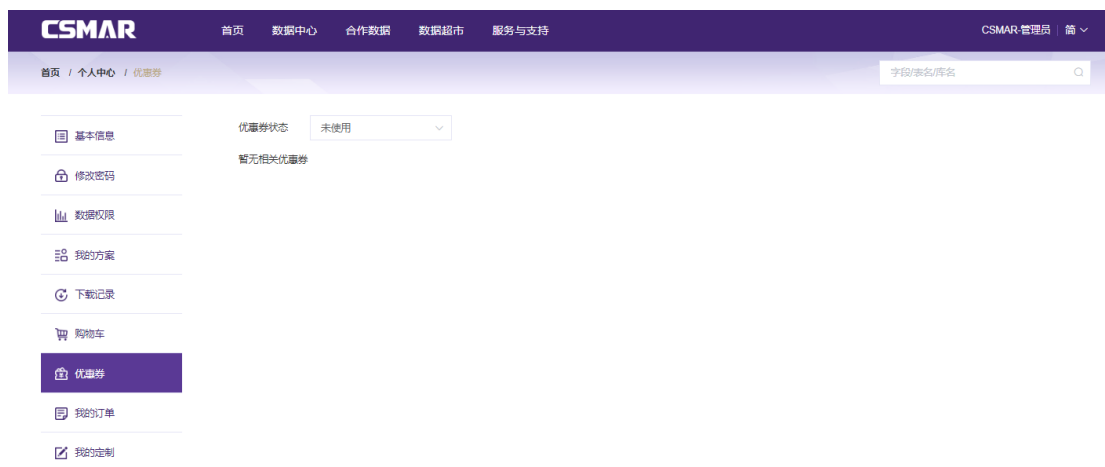
日个股回报率文件 ¥1141 周个股回报率文件 ¥853 月个股回报率文件 ¥873

实际支付金额: ¥2867

我的订单包含“待支付”订单、“已支付”订单和“已失效”订单，每个订单包含订单号、下单时间、使用期限、数据表列表以及订单金额合计等信息。您可以通过设置订单状态和下单时间，点击“查询”来搜索定位需要的订单。针对待支付订单，可点击“继续支付”进行支付或点击“删除订单”删除此订单。待支付订单 12 小时内如果未完成支付，将自动

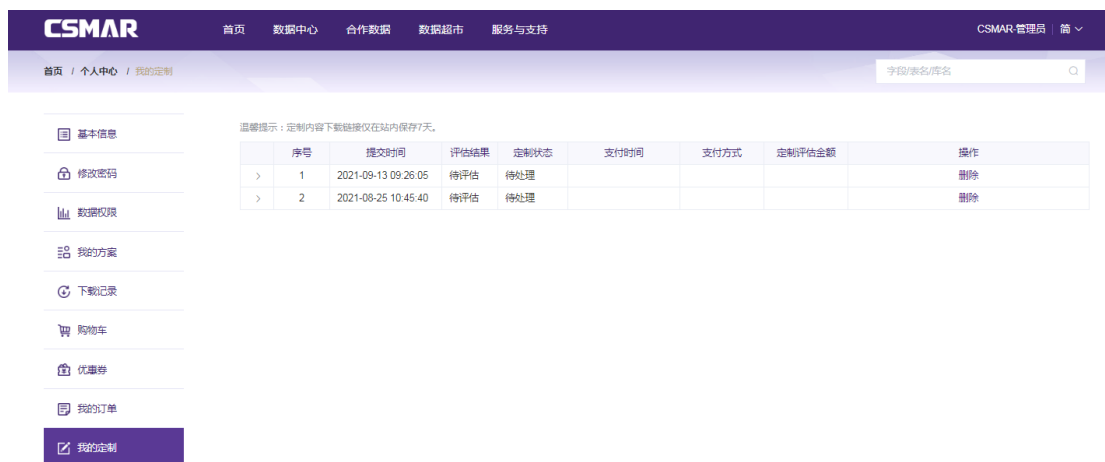
变为“已失效”订单。

7. 优惠券



展示活动期间领取的优惠券，包含已到期、已使用和可使用三种。

8. 我的定制



展示通过【数据定制】模块在线提交的所有定制需求，可查看定制评估进度，如果您同意定制评估金额，可在线付款并在线下载定制数据，方便快捷。