

数据库产品手册



PRODUCT MANUAL

CNRDS 中国研究数据服务平台
Chinese Research Data Services Platform

Open platform for high standard Chinese business research data

目 录

一、中国研究数据服务平台（CNRDS）简介.....	1
二、CNRDS 公司特色数据库.....	3
上市公司经营研究	4
（一）中国创新专利研究数据库-CIRD.....	4
（二）中国审计与风险数据库-CARD.....	7
（三）中国券商与分析师数据库-CBAS.....	9
（四）企业社会责任数据库-CCSR.....	12
（五）中国家族企业研究数据库-CFFD.....	14
（六）公司并购重组数据库-CMAD.....	16
（七）监管问询数据库-CRID.....	18
（八）中国上市公司内部控制研究数据库-ICRD.....	19
（九）上市公司参控股公司数据库-CPCD.....	21
（十）关键审计事项数据库-KAMD.....	23
（十一）中国上市公司供应链研究数据库-SCRD.....	24
（十二）中国上市公司异常收支数据库-ARED.....	25
（十三）投资者关系管理数据库-IRMD.....	26
（十四）海外上市研究数据库-COLD.....	28
（十五）中国法学研究数据库-CNLAW.....	30
上市公司新闻舆情	32
（一）中国上市公司财经新闻数据库-CFND.....	32
（二）中国上市公司股吧评论数据库-GUBA.....	35
（三）网络搜索指数数据库-WSVI.....	36
（四）中国上市公司社交媒体数据库-CSMD.....	37
（五）上市公司并购新闻数据库-MAND.....	38
（六）上市公司财务舞弊新闻数据库-FFND.....	39
（七）上市公司高管新闻数据库-CEND.....	40
上市公司文本信息	41
（一）年报文本语气数据库-ARTD.....	41
（二）管理层讨论与分析数据库-CMDA.....	42

(三) 上市公司业绩说明会数据库-ECCD.....	43
(四) 中国上市公司网上路演数据库-CNRD.....	44
(五) 中国上市公司澄清公告数据库-CCAD.....	45
资本市场人物特征.....	46
(一) 中国地方政府官员数据库-CGOD.....	46
(二) 董事长与总经理研究数据库-CCEO.....	48
(三) 上市公司其他高管数据库- OECD.....	50
(四) 独立董事研究数据库-IDRD.....	51
(五) 董秘信息数据库-CSBD.....	52
(六) 基金经理研究数据库-FMRD.....	53
(七) 金融机构高管数据库-FIED.....	54
银行及金融研究.....	55
(一) 银行研究数据库-CBRD.....	55
(二) 中国商业银行分支机构数据库-CCBD.....	57
(三) 中国风险投资和私募股权数据库-VCPE.....	58
(四) 中国金融理财研究数据库-CFSD.....	61
(五) 基金公司研究数据库-FCRD.....	63
(六) 券商投行数据库-CIBD.....	65
(七) 保险机构研究数据库-IIRD.....	67
(八) 信托机构研究数据库-CTID.....	68
社会经济组织研究.....	70
(一) 政府审计数据库-CGAD.....	70
(二) 中国高校科技统计数据库- USTS.....	71
(三) 中国非营利组织数据库-CNGO.....	73
(四) 中国商会研究数据库-CCCD.....	75
三、CNRDS 经济特色数据库.....	76
宏观经济研究系列.....	77
(一) 中国经济普查数据库-CECD.....	77
(二) 中国行政审批数据库-CAAD.....	79
(三) 全球夜间灯光数据库-GNLD.....	80
(四) 中国高铁航线数据库-CRAD.....	81

(五) 中国文化研究数据库-CCRD.....	82
(六) 中国环境数据库-CEDS.....	85
(七) 中国税收研究数据库-CTRD.....	88
(八) 中国商品交易市场统计数据库-CEMD.....	90
(九) 中国人口和就业统计数据库-PESD.....	92
区域经济研究系列.....	94
(一) 城市统计数据库-CCSD.....	94
产业经济研究系列.....	96
(一) 产业政策数据库-IPRD.....	96
(二) 中国房地产统计数据库-RESD.....	97
(三) 中国能源统计数据库-CESD.....	99
(四) 中国工业统计数据库-CISD.....	100
对外经济研究系列.....	101
(一) 中国对外直接投资数据库- OFDI.....	101
(二) 一带一路研究数据库-BRRD.....	103
(三) 海外并购数据库-COMA.....	105
三、CNRDS 基础数据库.....	106
上市公司股票基础数据.....	107
(一) 中国上市公司股价数据库-CNSP.....	107
(二) 中国股票异常交易数据库-CAST.....	109
(三) 中国特殊处理股票数据库-CSTS.....	110
(四) 中国融资融券数据库-CMTD.....	111
(五) 中国公司 IPO 综合数据库-CIPO.....	113
(六) 中国上市公司增发与配股数据库-CSEO.....	116
上市公司财务基础数据.....	117
(一) 上市公司业绩预告数据库-CEPD.....	117
(二) 财务报告披露时间数据库-FRDT.....	118
(三) 中国公司财务报表数据库-CNFS.....	119
(四) 上市公司财务报表附注数据库-NFSD.....	121
(五) 上市公司财务指标数据库-CNFI.....	124
(六) 中国上市公司盈利预测数据库-CEFD.....	126

上市公司治理基础数据	127
(一) 上市公司基本信息数据库-CBID.....	127
(二) 中国公司治理数据库-CCGD.....	128
(三) 审计意见与费用数据库-AUDIT.....	130
(四) 中国上市公司管理层变更数据库-MTDB.....	131
(五) 公司与高管违规处罚数据库-VPCE.....	132
(六) 高管薪酬与激励数据库-ECEL.....	133
(七) 中国上市公司关联交易数据库-CRTD.....	134
(八) 中国公司股权研究数据库-CERD.....	135
(九) 机构投资者持股研究数据库-IORD.....	137
(十) 中国公司股利分红数据库-CCDD.....	139
(十一) 中国上市公司诉讼仲裁数据库-CLAD.....	140
(十二) 中国上市公司内部人交易数据库-CITD.....	141
经济研究基础数据	142
(一) 中国宏观经济研究数据库-MACRO.....	142
(二) 中国区域经济研究数据库-CRED.....	143
(三) 中国债券研究数据库-BOND.....	144
四、中国财经大数据研究系统	146
五、最新学术动态	148
六、数据定制	149
七、数据合作	151

一、中国研究数据服务平台（CNRDS）简介

1. 平台概况

中国研究数据服务平台（*Chinese Research Data Services Platform, CNRDS*）是一个高质量、开放式、平台化的中国经济、金融与商学研究的综合数据平台。**CNRDS** 的宗旨在于借鉴 **WRDS** 等国外一流商学院打造的数据平台，构建中国特色的研究数据资源，促进中国研究数据的规范化、平台化、国际化，继而推动中国研究向纵深化、精细化方向发展。

CNRDS 平台提供的数据库主要来源于三个途径：1) 平台自有数据，从采集、整理到呈现均有该领域国内外知名专家和学者深度参与；2) 由相关领域的专家和学者提供，并由 **CNRDS** 提供技术支持和形式审核以及后续维护；3) 各专业领域的数据库公司提供原始数据，**CNRDS** 再邀请相关领域的专家和学者对数据质量进行严格审核。

我们相信每一位学者在支持和推动中国研究数据和资源构建上迈出的一小步，都将成为中国学术研究发展的一大步。

2. 平台特色

CNRDS 特色库紧跟学术热点和学术前沿，提供市场尚无或者获取难度较大的特色研究数据；基础库参阅各领域重点文献，整合了学术研究过程中可能用到的大部分基础数据。

CNRDS 数据库的特点可归纳为：

- 国内外知名高校教授合作开发
- 特色数据库满足各类研究需求
- 界面与模块设计契合学术研究

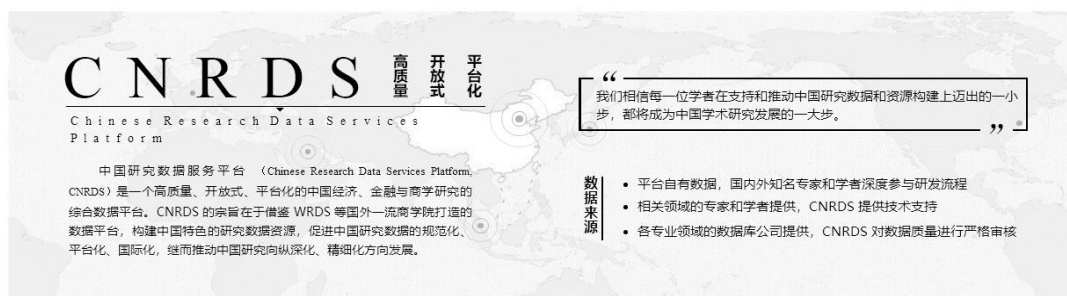
3. 平台网址

www.cnrds.com

4. 适用对象

中国经济、金融与商学研究方向的教授、学者；利用资本市场数据构建投资决策的金融投资机构；撰写报告的企业。

5. 平台界面



经济与金融		会计	
期刊简称	论文标题	发表日期	相关库
JBF	Political influence and financial flexibility:Evidence fro...	2019年02月	CGOD
JBF	Corporate pyramids,geographical distance,and investmen...	2019年02月	CERD
JBF	Politicians' promotion incentives and bank risk exposure ...	2019年02月	CBRD,CGOD
JCF	Price inversion and post lock-up period returns on privat...	2019年02月	MACRO
JEF	Stock liquidity and corporate diversification:Evidence fr...	2018年12月	CERD
RFS	Decentralized Privatization and Change of Control Right...	2018年10月	
JCF	Do firms care about investment opportunities?Evidence f...	2018年10月	CNFS

第1页, 共33页 < >



二、CNRDS 公司特色数据库

为了满足不同学科的研究，方便不同领域学者能够迅速查找、浏览和下载所需数据，**CNRDS** 参照国际标准并结合中国资本市场数据结构，基于 **CNRDS** 平台已有数据库及强大的研发能力，设立了公司特色库，主要提供中国资本市场参与机构（包括上市公司、金融机构以及其他社会经济组织等）的特色研究数据。**CNRDS** 公司特色数据库涉及的主题丰富多元，紧跟国内外中国资本市场问题研究的热点和前沿，目前已设立六个研究系列，具体如下：

上市公司经营研究 主要包括上市公司经营过程中涉及的有关数据，例如内部控制、创新专利、投资者关系管理、企业社会责任等数据。

上市公司新闻舆情 主要包括与上市公司相关的新闻报道、股吧评论和网络搜索等信息，并采用计算机技术对这类数据进行了深度清洗、加工和判断。

上市公司文本信息 主要采用已有文献的主流文本分析方法，对与上市公司相关的公告、年报等各类文本信息进行全面地分析与展示。

资本市场人物特征 主要囊括了资本市场发展中涉及到的诸多参与人员，例如上市公司高管、基金经理等多维度的个人特征数据。

银行及金融研究 主要包括资本市场的重要金融机构及金融行为数据，例如银行机构、基金公司、券商投行、风险投资等数据。

社会经济组织研究 主要列示了资本市场主流参与者之外的另一类重要组织机构的数据，例如，慈善基金会、商会、高校等机构的数据。

为保证数据质量及契合学术研究需要，**CNRDS** 公司特色数据库从数据采集、数据整理和数据呈现均有该领域知名的国内外教授和学者深度参与，并结合 **CNRDS** 平台严格的数据管理和建库流程，可最大程度为研究人员提供高水准、高质量、专业性的学术特色数据。

CNRDS 公司特色数据库具有以下三个特点：

- 国内外知名教授深度参与，保证研究数据质量
- 涉及不同研究领域数据，满足多种研究需求
- 各特色库自成体系，充分契合学术研究习惯

上市公司经营研究

(一) 中国创新专利研究数据库-CIRD

1. 数据库简介

中国创新专利研究数据库 (*Chinese Innovation Research Database, CIRD*) 主要基于对不同上市公司、非上市公司和不同地区的专利申请和授权情况而开发的专业数据库。*CIRD* 不仅考虑了不同主体申请或授权的专利数据, 其中区分了发明、实用新型和外观设计三种类别的专利, 还包括了诸如独立申请还是联合申请、专利的分类号汇总和统计等重要信息。

CIRD 包括六个模块的数据:

上市公司专利申请与获得模块 提供中国所有上市公司专利申请与授权情况, 该模块包括自 1991 年以来每年申请和获得授权的专利数量; 区分为每年独立和联合申请和获得授权的专利数量; 区分上市公司本身、子公司、联营公司和合营公司的专利申请和获得授权的专利数量。

上市公司专利分类号信息模块 提供中国所有上市公司专利分类号的汇总和统计情况, 该模块区分为申请和获得授权专利的分类号情况; 包括了发明和实用新型两类重要专利的分类号; 区分上市公司本身, 子公司、联营公司和合营公司等不同类型公司的专利分类号; 包括分类号汇总和分类号数量两个重要字段。

上市公司研发支出模块 提供中国所有上市公司自 2006 年以来每年的研发支出情况。

非上市公司专利申请与获得模块 提供所有中国非上市公司的专利信息, 该模块包括自 1985 年以来非上市公司每年申请和获得授权的专利数量; 区分为每年独立和联合申请和获得授权的专利数量。

科研机构专利申请与获得模块 提供中国所有科研机构 (包括高校) 的专利信息, 该模块包括自 1985 年以来科研机构每年申请和获得授权的专利数量; 区分为每年独立和联合申请和获得授权的专利数量。

各省市创新专利情况 提供中国各省、地级市的专利信息, 该模块包括了中国各省份、地级市每年的专利申请数量, 并区分为发明、实用新型和外观设计的专利数量。

2. 数据库特色

- 全面的公司专利信息

CIRD 对上市公司、及其子公司、联营企业和合营企业各自的专利数量进行统计, 精确反映公司集团的专利情况, 是目前唯一能够做到如此详尽的数据库。

- 唯一提供公司专利分类号信息

针对发明专利和实用新型专利, *CIRD* 包含了其对应的专利分类号及其统计结果, 可为专利研究的深度扩展提供数据支持。

- 提供地域创新力信息

CIRD 提供了不同省市的专利申请与授权情况, 可反映一个地区的创新力情况。

3. 典型数据举例

(1) 上市公司专利申请数量

股票代码	会计年度	公司类型	当年独立申请的发明数量	当年独立申请的实用新型数量	当年独立申请的外观设计数量	当年联合申请的发明数量	当年联合申请的实用新型数量	当年联合申请的外观设计数量
000030	2010	上市公司本身	2	8	6	1	13	5
000030	2011	上市公司本身	2	6	2	1	3	13
000030	2012	上市公司本身	1	2	6	1	13	6
000030	2013	上市公司本身	6	13	6	2	5	1
000032	2012	子公司	3	6	7	1	2	1
000063	2005	上市公司本身	1006	137	26	4	0	2

(2) 各省市专利申请情况

省份名称	地市名称	会计年度	当年申请的发明数量	当年申请的实用新型数量	当年申请的外观设计数量
安徽省	安庆市	2010	98	292	108
安徽省	安庆市	2011	165	421	115
安徽省	安庆市	2012	239	799	221
安徽省	安庆市	2013	1105	1471	234
安徽省	安庆市	2014	3294	2476	305

4. 数据库内容

数据模块	数据表	数据字段
上市公司专利申请与获得	上市公司专利申请情况	股票代码、会计年度、公司类型、申请时间、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	上市公司专利获得情况	股票代码、会计年度、公司类型、获得时间、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
上市公司专利分类号信息	上市公司发明申请专利分类号	股票代码、会计年度、公司类型、申请时间、发明申请 A 类、发明申请 A 类数量、发明申请 B 类、发明申请 B 类数量、发明申请 C 类、发明申请 C 类数量、发明申请 D 类、发明申请 D 类数量、发明申请 E 类、发明申请 E 类数量、发明申请 F 类、发明申请 F 类数量、发明申请 G 类、发明申请 G 类数量、发明申请 H 类、发明申请 H 类数量
	上市公司实用新型申请专利分类号	股票代码、会计年度、公司类型、申请时间、实用新型申请 A 类、实用新型申请 A 类数量、实用新型申请 B 类、实用新型申请 B 类数量、实用新型申请 C 类、实用新型申请 C 类数量、实用新型申请 D 类、实用新型申请 D 类数量、实用新型申请 E 类、实用新型申请 E 类数量、实用新型申请 F 类、实用新型申请 F 类数量、实用新型申请 G 类、实用新型申请 G 类数量、实用新型申请 H 类、实用新型申请 H 类数量

数据模块	数据表	数据字段
上市公司专利分类号信息	上市公司发明获得专利分类号	股票代码、会计年度、公司类型、获得时间、发明获得 A 类、发明获得 A 类数量、发明获得 B 类、发明获得 B 类数量、发明获得 C 类、发明获得 C 类数量、发明获得 D 类、发明获得 D 类数量、发明获得 E 类、发明获得 E 类数量、发明获得 F 类、发明获得 F 类数量、发明获得 G 类、发明获得 G 类数量、发明获得 H 类、发明获得 H 类数量
	上市公司实用新型获得专利分类号	股票代码、会计年度、公司类型、获得时间、实用新型获得 A 类、实用新型获得 A 类数量、实用新型获得 B 类、实用新型获得 B 类数量、实用新型获得 C 类、实用新型获得 C 类数量、实用新型获得 D 类、实用新型获得 D 类数量、实用新型获得 E 类、实用新型获得 E 类数量、实用新型获得 F 类、实用新型获得 F 类数量、实用新型获得 G 类、实用新型获得 G 类数量、实用新型获得 H 类、实用新型获得 H 类数量
上市公司研发支出	上市公司研发支出	股票代码、会计年度、研发投入总额
非上市公司专利申请与获得	非上市公司专利申请情况	公司名称、会计年度、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	非上市公司专利获得情况	公司名称、会计年度、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
科研机构专利申请与获得	科研机构专利申请情况	单位名称、会计年度、当年独立申请的发明数量、当年独立申请的实用新型数量、当年独立申请的外观设计数量、当年联合申请的发明数量、当年联合申请的实用新型数量、当年联合申请的外观设计数量
	科研机构专利获得情况	单位名称、会计年度、当年独立获得的发明数量、当年独立获得的实用新型数量、当年独立获得的外观设计数量、当年联合获得的发明数量、当年联合获得的实用新型数量、当年联合获得的外观设计数量
各省市创新专利情况	各省市专利申请情况	省份名称、地市名称、会计年度、当年申请的发明数量、当年申请的实用新型数量、当年申请的外观设计数量
	各省市专利获得情况	省份名称、地市名称、会计年度、当年获得的发明数量、当年获得的实用新型数量、当年获得的外观设计数量

5.数据范围

数据模块	数据起始时间
上市公司专利申请与获得	1991 年
上市公司专利分类号信息	1990 年
上市公司研发支出	2006 年
非上市公司专利申请与获得	1985 年
科研机构专利申请与获得	1985 年
各省市创新专利情况	1985 年

（二）中国审计与风险数据库-CARD

1.数据库简介

中国审计与风险数据库 (*Chinese Audit and Risk Database, CARD*) 是基于对中国公司审计与风险的深入考察和了解,建立的详细、准确和全面的公司审计与风险研究专业数据库。该库不仅包括中国注册会计师的个人特征和会计师事务所的基本信息,涵盖了这些事务所对上市公司出具的审计意见与收取的审计费用等信息,还包括与上市公司经营风险和财务风险相关的盈余重述情况等。

2.数据库特色

- 完整

CARD 包括研究中国上市公司审计情况的大部分数据,这些数据内容丰富、覆盖面广,为学者研究该领域问题提供了完整、全面的数据来源。

- 详细

CARD 详细列示了与研究相关的大量字段,例如,对注册会计师的个人特征从基本信息、教育背景、所学专业等进行了详细的区分,这些字段可以全方位体现相关数据的多维度信息。

3.典型数据举例

(1) 会计师事务所总所信息

事务所代码	会计师事务所名称	组织形式	批准设立时间	合伙人人数	注师人数	从业人员人数	本科人数
115359	湖南国信会计师事务所	有限责任	1999-11-02	6	21	21	6
113447	上海建正联合会计师事务所	普通合伙	2007-07-02	2	2	3	1
115455	湖南安信联合会计师事务所	普通合伙	2001-05-31	2	39	0	14
115458	安徽九通会计师事务所	有限责任	1999-12-15	5	29	16	10
115710	深圳业信会计师事务所	普通合伙	1997-12-23	2	7	17	3

(2) 注册会计师基本信息

审计师姓名	性别	所在事务所	学历	专业	资格取得方式	职位	是否合伙人
黄起华	男	深圳正声会计师事务所	本科	经济管理	考试	审计经理	是
胡建国	男	佛山中健会计师事务所	本科	会计学	考试	所长	是
胡柏青	男	深圳华起会计师事务所	本科	会计学	考试	所长	否

4.数据库内容

数据模块	数据表	数据字段
会计师事务所信息	会计师事务所总所信息	事务所代码、会计师事务所名称、组织形式、注册资本、批准设立时间、法定代表人、主任会计师、合伙人人数、注册会计师人数(含分所)、注册会计师人数、从业人员人数(含分所)、从业人员人数、小于等于40岁人数、小于等于60岁且大于40岁人数、小于等于70岁且大于60岁人数、大于70岁人数、博士研究生人数、硕士研究生人数、本科人数、大专及以下人数、分所数量

数据模块	数据表	数据字段
会计师事务所信息	会计师事务所分所信息	事务所分所代码、分所名称、分所负责人、注册会计师人数、从业人员人数、批准设立时间
	会计师事务所执业资格信息	事务所代码、资格取得时间、取得的执业资格名称、批准机关
	会计师事务所处罚信息	事务所代码、惩戒及处罚种类、实施惩戒及处罚的部门、惩戒及处罚事由、惩戒及处罚时间
	会计师事务所公益活动	事务所代码、主办单位、活动名称、参加时间、活动内容
注册会计师信息	注册会计师基本信息	审计师代码、审计师姓名、性别、所在事务所、所在分所、事务所代码、学历、专业、资格取得方式、批准注册文件号、职位、是否合伙人、批准注册时间、全科合格年份
	注册会计师奖励信息	审计师代码、奖励种类、奖励时间、奖励部门、奖励原因
	注册会计师其他执业资格	审计师代码、取得国内其他执业资格名称、取得时间
	注册会计师处罚信息	审计师代码、惩戒及处罚种类、实施惩戒及处罚的部门、惩戒及处罚事由、惩戒及处罚时间
审计意见与费用	合伙人信息	事务所代码、审计师代码、是否注师、合伙人姓名
	审计单位与审计意见	股票代码、会计年度、会计师事务所、审计师、审计意见、审计意见说明
	审计单位与审计费用	股票代码、会计年度、会计师事务所、审计费用
盈余重述	盈余重述	股票代码、会计年度、会计差错发生年度 1、会计差错发生年度 2、盈余变更方向、因税调整

5.数据范围

数据模块	数据起始时间
会计师事务所信息	1991 年
注册会计师信息	1991 年
审计意见与费用	1990 年
盈余重述	1999 年

（三）中国券商与分析师数据库-CBAS

1.数据库简介

中国券商与分析师数据库 (*Chinese Broker and Financial Analysts System, CBAS*) 是上海经禾信息技术有限公司与海外知名高校的金融与会计学者合作, 参照 *IBES* 等国际分析师数据库联合开发的有关中国券商与分析师研究的专业学术数据库, 不同于其他数据库集中在分析师报告的盈利预测和荐股评级上, 该数据库还包括了很多中国特色数据。数据库包括盈利预测、荐股评级、分析师背景信息、调研数据、券商数据、分析师报告文本分析和原始报告 7 大模块。

盈利预测模块和荐股评级模块 此两大模块从 30 多万份分析师报告中提取了盈利预测、荐股评级、目标股价等信息, 是目前为止较全面详实的分析师预测和荐股数据库。

分析师背景信息模块 此模块提供了近 3000 名分析师的个人详细信息, 包括基本信息(性别、工作开始时间、大学入学时间等)、教育背景(起止日期、学校、学历等)、以及工作经历(起止日期、工作单位、部门和职位)。

调研数据模块 此模块提供调研日期、参与人员、参与单位、接待人员等基本信息。

券商数据模块 此模块从券商年报中提取了券商财务信息、大股东及其持股比例信息、大股东性质信息、董事、监事、高管个人信息等独特数据。

分析师报告文本分析模块 此模块是大数据技术在学术数据库上的应用, 我们对每份分析师报告提供了用 LDA 主题模型提取的概率最高的 5 个主题, 以及文本句子数量和财务信息句子数量, 是目前为止唯一用大数据技术对分析师报告信息进行提炼的数据库。

原始报告模块 此模块提供近 30 万份的分析师报告、券商年报以及分析师调研报告等原始数据。研究者可以随时翻阅最原始的数据, 寻找灵感, 突破研究瓶颈。

2.数据库特色

- 分析师及券商唯一编码

CBAS 给数据库中 100 多家券商以及 6000 多名分析师分配了独特编码, 是目前为止唯一给券商和分析师独特编码的数据库, 有利于研究者进行数据匹配和深度挖掘。

- 独有的分析师背景信息数据

CBAS 包含分析师的个人特征数据, 包括基本信息、教育背景和工作经历等。

- 覆盖全面的调研信息

CBAS 包括 2007 年以来的分析师调研信息, 既包括来自年报的分析师调研情况, 也包括来自调研报告的调研情况。

- 独有的券商相关信息

CBAS 包括了主要券商的大股东、高管、财务数据信息等。

3. 典型数据举例

(1) 分析师基本信息

分析师姓名	分析师代码	券商名称	最高学历	性别
蔡雯娟	A00070	安信证券股份有限公司	硕士研究生	女
陈国栋	A00203	德邦证券股份有限公司	博士研究生	男
陈李	A00266	瑞银证券有限责任公司	博士研究生	男
陈雯	A00348	平安证券有限责任公司	硕士研究生	女
苏铨	A03001	安信证券股份有限公司	硕士研究生	男
严轩	A03520	长江证券股份有限公司	本科	男

(2) 券商股东构成

券商名称	券商代码	会计报告日期	大股东名称	大股东持股比例	大股东性质	最终控制人
爱建证券有限责任公司	B00001	2014-12-31	上海陆家嘴金融发展有限公司	0.51136	地方国有	地方国资局
安信证券股份有限公司	B00002	2014-12-31	国家开发投资公司	0.5801	中央国有	国资委
安信证券股份有限公司	B00002	2014-12-31	中国证券投资者保护基金有限责任公司	0.2588	中央国有	国资委
渤海证券股份有限公司	B00006	2014-12-31	天津市泰达国际控股(集团)有限公司	0.536695	地方国有	地方国资局

4. 数据库内容

数据模块	数据表	数据字段
盈利预测	预测情况	股票代码、报告日期、盈利预测结束日期、分析师姓名、分析师代码、是否明星分析师、券商名称、券商代码、EPS 预测、P/E 比率预测、净利润预测、主营业务收入预测、实际盈余公告日
	实际情况	股票代码、盈利预测结束日期、真实 EPS、真实净利润、真实营业利润
	结果分析	股票代码、盈利预测结束日期、CBAS 一致性预测日期、EPS 一致性预测、净利润一致性预测、营业利润一致性预测、一致性预测包括的报告数目
荐股评级	评级	股票代码、报告日期、分析师姓名、分析师代码、券商名称、券商代码、原始报告中的荐股评级、标准化荐股意见、标准化荐股意见代码、是否是明星分析师
分析师背景信息	基本信息	分析师姓名、分析师代码、券商名称、券商代码、大学入学时间、最高学历/学位、性别、工作开始日期
	教育背景	分析师姓名、分析师代码、教育背景序号、学位、教育经历起止日期、毕业学校、毕业学校代码
	工作经历	分析师姓名、分析师代码、工作经历序号、工作经历起止日期、工作单位、工作部门、工作职位
调研数据	访客	股票代码、公司简称、公开时间、调研时间、参与单位、参与单位类别、参与人员
	接待人员	股票代码、公司简称、调研时间、公开时间、调研接待人员 n 姓名、调研接待人员 n 职务
	调研内容	股票代码、公司简称、调研时间、调研报告内容

数据模块	数据表	数据字段
券商数据	股东构成	券商名称、券商代码、会计报告日期、大股东名称、大股东占股比例、大股东性质、最终控制人
	董监高	券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、工作职位、是否全职、会计报告日期
	IPO 承销排名	券商名称、券商代码、会计年度、当年承销 IPO 总募集资金金额、当年承销 IPO 总募集资金排名、当年承销 IPO 总募集资金占当年 IPO 市场总募集资金比例、当年 IPO 承销费用、当年 IPO 承销费用排名、当年 IPO 承销数量、当年联席保荐 IPO 数量、当年 IPO 承销数量排名
	财务报表	券商名称、券商代码、合并会计报表或母公司会计报表、会计报告日期、营业收入、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益 EPS、稀释每股收益 EPS、货币资金、客户资金存款、交易性金融资产、衍生金融资产、买入返售金融资产、存出保证金、可供出售金融资产、持有至到期投资、长期股权投资、固定资产、无形资产、资产总计、短期借款、质押借款、拆入资金、交易性金融负债、衍生金融负债、卖出回购金融资产款、代理买卖证券款、代理承销证券款、应付职工薪酬、应交税费、长期借款、应付债券、负债合计、股本、资本公积金、盈余公积金、未分配利润、一般风险准备、归属于母公司所有者权益合计、所有者权益合计、负债及股东权益总计、经营活动产生的现金流量净额、投资活动产生的现金流量净额、筹资活动产生的现金流量净额
	财务比率	券商名称、券商代码、合并会计报表或母公司会计报表、会计报告日期、证券公司净资本、营业收入、营业收入增长率、手续费及佣金净收入、手续费及佣金净收入在营业收入中的占比、利息净收入、利息净收入在营业收入中的占比、投资净收益、投资净收益在营业收入中的占比、公允价值变动净收益、公允价值变动净收益在营业收入中的占比、其他收入、其他收入在营业收入中的占比、代理买卖证券业务净收入、证券承销业务净收入、受托客户资产管理业务净收入
分析师报告文本分析	分析师报告文本分析	股票代码、报告日期、分析师姓名、分析师代码、券商名称、券商代码、分析师报告主题、句子数目、财务信息句子数目

5.数据范围

数据模块	数据起始时间
盈利预测	2001 年
荐股评级	2001 年
分析师背景信息	2006 年
调研数据	2011 年
券商数据	2006 年
分析师报告文本分析	2006 年

（四）企业社会责任数据库-CCSR

1.数据库简介

上市公司企业社会责任近年来受到各方面的关注，企业社会责任是指企业在创造利润、对股东承担责任的同时还要承担对员工、社区、产品和环境的责任。企业社会责任要求企业必须要超越把利润作为企业唯一目标的传统理念，强调在生产过程中对环境、对社会的贡献。企业落实社会责任，实现企业经济责任和社会责任的平衡，会提高企业的竞争力，为企业树立良好的声誉和形象，增强投资者信心。企业社会责任数据库（*Chinese Corporate Social Responsibilities Database, CCSR*）旨在为科研人员提供中国上市公司企业社会责任数据。该数据库的设计主要以 KLD STATS 的数据作为设计的主要模板，并根据我国企业社会责任所涉及的具体内容，从六个方面，用 58 个细分指标对企业社会责任进行衡量。*CCSR* 数据库包括 2006 年至今沪深两市所有披露过企业社会责任报告的上市公司的 CSR 数据。

2.数据库特色

- 设计科学

CCSR 参考了国外知名的企业社会责任数据库的设计标准，并且结合了我国企业的实际情况，设计了企业社会责任的六个方面：“慈善、志愿者活动以及社会争议”、“公司治理”、“多样化”、“雇员关系”、“环境”以及“产品”，共有 58 个字段。每个方面都从“优势”和“关注”两个角度设计字段。尤其是“关注”相关的字段，*CCSR* 通过大数据技术搜索所有信息来源进行构建，克服了别的企业社会责任数据库只能结合企业披露信息来构建字段的缺陷。

- 直观

CCSR 的大部分字段为哑变量，非常直观地表达了企业社会责任不同维度的特征，使用者可以很快地构建全面或者某一方面的企业社会责任指数。

3.典型数据举例

(1) 慈善、志愿者活动以及社会争议优势

股票代码	公司简称	会计年度	捐赠总额	支持教育	支持慈善	志愿者活动	带动就业	促进当地经济	其他优势
000001	平安银行	2012	92 万元	1	1	0	0	0	0
000002	万科 A	2012	1720 万元	1	1	1	1	1	1
000009	中国宝安	2012	110.256 万元	1	1	0	0	0	1

(2) 雇员关系关注

股票代码	公司简称	会计年度	雇员安全纠纷	裁员	其他关注
000527	美的电器	2012	0	0	1
000528	柳工	2012	0	0	1
000800	一汽轿车	2012	1	0	0
002557	洽洽食品	2012	1	0	0
600690	青岛海尔	2012	0	1	0

4.数据库内容

数据模块	数据表	数据字段
慈善、志愿者活动以及社会争议	优势	股票代码、公司简称、会计年度、捐赠总额、支持教育、支持慈善、志愿者活动、国际援助、带动就业、促进当地经济、其他优势
	关注	股票代码、公司简称、会计年度、融资纠纷
公司治理	优势	股票代码、公司简称、会计年度、CSR 报告全面性、CSR 报告页数、CSR 专栏、CSR 领导机构、CSR 愿景、CSR 培训、每股社会贡献值、可靠性保证、其他优势
	关注	股票代码、公司简称、会计年度、会计违规
多样化	优势	股票代码、公司简称、会计年度、党员、女性高管、女性董事席位、创新人力资源项目、其他优势
	关注	股票代码、公司简称、会计年度、没有女性高管
雇员关系	优势	股票代码、公司简称、会计年度、员工参股、员工福利、安全管理体系、安全生产培训、职业安全认证、职业培训、员工沟通渠道、其他优势
	关注	股票代码、公司简称、会计年度、雇员安全纠纷、裁员
环境	优势	股票代码、公司简称、会计年度、环境有益的产品、减少三废的措施、循环经济、节约能源、绿色办公、环境认证、环境表彰、其他优势
	关注	股票代码、公司简称、会计年度、环境处罚、污染物排放
产品	优势	股票代码、公司简称、会计年度、质量体系、售后服务、客户满意度调查、质量荣誉、专利数目、研发支出、研发人员比例、技术人员比例、反腐败措施、战略共享、诚信经营理念、其他优势
	关注	股票代码、公司简称、会计年度、产品纠纷

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

（五）中国家族企业研究数据库-CFFD

1.数据库简介

中国家族企业研究数据库 (*Chinese Family Firm Research Database, CFFD*)，是对中国上市公司家族企业进行调查和研究而建立的专业数据库。不同于以往将实际控制人可以追溯到个人的民营企业确定为家族企业，*CFFD* 所确定的家族企业包括三个标准：（1）实际控制人是一个自然人或某个以血缘、姻缘为连结的家族；（2）该自然人或家族直接或间接是上市公司第一大股东；（3）至少两个或以上家族成员在该上市公司或者关联公司持股或任职。*CFFD* 包括四张数据表：家族企业基本信息、IPO 时家族企业持股情况、家族企业控制链以及家族企业继承人信息。其中，家族企业基本信息包括实际控制人、成立和上市时间在内的一些企业基本特征信息；IPO 时家族企业持股情况包括企业在上市前后实际控制人的持股变化，家族其他相关人员的名称、持股和职务信息等；家族企业控制链则包括一个完整时间序列内家族企业实际控制人和相关人员的持股层级和持股比例情况；最后，家族企业继承人信息包括家族企业的创始人和继承人信息。

2.数据库特色

- 精确的家族企业定义

相比以往的家族企业定义，*CFFD* 对家族企业的定义更为准确，真正可以追溯到哪个家族对企业形成实际控制。

- 详细的家族企业成员持股信息

对 IPO 前后家族企业实际控制人、家族企业成员的持股和亲缘关系信息进行了详细列示，有助于发现家族企业的内部关系和持股关系。

- 时间序列上的控制链信息

根据年报和公告信息，构建了家族企业实际控制人在时间序列上的持股变化，动态跟踪家族企业所有权的变化情况。

- 家族继承人信息

丰富的家族企业继承人个人特征信息，为研究家族企业传承问题提供重要数据支持。

3.典型数据举例

IPO 时家族企业持股情况

股票代码	股东姓名	招股前持股比例 (%)	招股后持股比例 (%)	在上市公司的职位	关系说明
002010	徐冠巨	32.78	24.58	董事长	徐传化系徐冠巨、徐观宝之父
002010	徐观宝	18.89	14.17	董事	徐传化系徐冠巨、徐观宝之父
002010	徐传化	18.33	13.75	创始人	徐传化系徐冠巨、徐观宝之父
002194	孟庆南	39.83	29.8	董事长	孟庆南与王丽丽系夫妻关系
002194	王丽丽	39.83	29.8	副董事长、总经理	孟庆南与王丽丽系夫妻关系
002194	孟凡博	12	8.98	董事	孟凡博系孟庆南和王丽丽之子
002194	王凯	4.88	3.65	董事	王凯系王丽丽之弟
002194	王凌	0.06	0.04	副总经理	王凌系王丽丽之妹
002194	王莉萍	0.05	0.04	中层管理人员	王莉萍系王丽丽之妹

4.数据库内容

数据模块	数据表	数据字段
家族企业基本信息	家族企业基本信息	股票代码、公司名称、实际控制人、上市日期、成立日期、注册资本、法定代表人、经营范围、所属行业名称、所属行业代码、省份、城市、注册地址、办公地址、公司简介
IPO 时家族企业持股情况	IPO 时家族企业持股情况	股票代码、公司名称、序号、实际控制人、股东姓名、招股前持股比例(%)、招股后持股比例(%)、法人股东的大股东、法人股东的大股东持股比例(%)、法人股东的大股东职位、在上市公司的职位、关系说明
家族企业控制链	家族企业控制链	股票代码、公司名称、统计日期、持股层级数、一级股东、一级股东持股比例(%)、二级股东、二级股东持股比例(%)、三级股东、三级股东持股比例(%)、四级股东、四级级股东持股比例(%)、五级股东、五级股东持股比例(%)、六级股东、六级股东持股比例(%)、七级股东、七级级股东持股比例(%)、八级股东、八级股东持股比例(%)、九级股东、九级股东持股比例(%)
家族企业继承人信息	家族企业继承人信息	股票代码、公司名称、控股年份、创始人姓名、继承人姓名、与创始人关系、职位、任职开始日期、任职结束日期、性别、出生日期、国籍、籍贯/出生地、学历、职称与资格证书、简历

5.数据范围

数据模块	数据起始时间
家族企业基本信息	截面数据
IPO 时家族企业持股情况	截面数据
家族企业控制链	1999 年
家族企业继承人信息	截面数据

（六）公司并购重组数据库-CMAD

1.数据库简介

公司并购重组数据库 (*Mergers and Acquisitions Database, CMAD*) 是基于中国上市公司及其子公司等关联方的兼并、收购和重组事件整理而成的专业数据库。*CMAD* 收录了公司并购事件的基本信息, 并购相关方 (包括买方、买方和标的方) 的基本信息、公司股东和管理层等, 详细和完整的展示了并购过程中的重要信息。该数据库总共分为两个大模块, 并购事件信息模块主要包括并购时间、并购方式、并购是否成功、并购金额等重要字段; 并购涉及方信息主要包括涉及方基本信息、涉及方股东信息和涉及方高管等重要信息。

2.数据库特色

- 数据量巨大、字段丰富
CMAD 囊括了近几十年来的十多万例并购事件, 采用简约风格设计了大量重要字段。
- 极为详细的并购涉及方信息
CMAD 对并购涉及方的信息进行了深度挖掘, 为学者进行此类研究提供了重要数据支持。

3.典型数据举例

并购事件基本信息

股票代码	上市公司交易地位	买方名称	标的方名称	卖方名称	首次公告日	重组类型
000001	标的方	美国新桥投资集团公司	平安银行股份有限公司	深圳国际信托投资有限公司	2004-05-31	股权转让
000002	买方	上海万科城市花园发展有限公司	昆山嘉华投资有限公司		2004-02-09	资产收购
000002	买方	上海万科浦东置业有限公司	上海南都置地有限公司	上海中桥基建(集团)股份有限公司	2005-03-05	资产收购
000004	标的方	深圳长汇投资企业(有限合伙)	深圳中国农大科技股份有限公司		2011-07-21	股权转让
000005	卖方	深圳市恒裕(实业)集团有限公司	深圳市中环星苑房地产开发有限公司	深圳世纪星源物业发展有限公司	2017-12-28	资产剥离

4.数据库内容

数据模块	数据表	数据字段
并购事件信息	并购事件基本信息	事件 ID、股票代码、公司名称、上市公司交易地位、买方名称、标的方名称、卖方名称、首次公告日、最新公告日、交易是否成功、重组类型、是否关联交易、支付方式、交易金额、币种、标的类型、股权转让比例 (%)、控制权是否变更、交易简介

数据模块	数据表	数据字段
并购涉及方信息	涉及方基本信息	事件 ID、股票代码、上市公司交易地位、相关方、相关方身份、注册资本、经营状态、成立日期、公司类型、所属行业、曾用名、注册地址、所在省份、所在市、经营范围
	涉及方股东信息	事件 ID、股票代码、上市公司交易地位、相关方、相关方身份、序号、股东、持股比例 (%)、股东类型
	涉及方高管信息	事件 ID、股票代码、上市公司交易地位、相关方、相关方身份、序号、姓名、职务

5.数据范围

数据模块	数据起始时间
全部模块	1995 年

(七) 监管问询数据库-CRID

1.数据库简介

监管问询数据库 (*Regulatory Inquiries Database, CRID*) 是基于交易所等监管机构向上市公司问询有关其经营或治理相关问题的文件整理而成的专业数据库。*CRID* 的数据主要来源于上交所和深交所, 并对问询文件内容和公司回复内容进行梳理和展示。此外, 根据已有研究, 本数据库还提取了与监管问询以及公司回复相关的日期时间等重要字段。

2.数据库特色

- 字段丰富

CRID 除了包括直接来源于交易所的字段数据, 还对 *PDF* 文本内容以及相关日期等进行了提取和展示。

3.典型数据举例

监管问询

股票代码	函件类别	问询标题	发函日期	限期回复日期	公司回复日期
300043	非许可类重组问询函	关于对浙江互动娱乐股份有限公司的重组问询函	2014-12-9	2014-12-11	2014-12-11
002324	非许可类重组问询函	关于对浙江普利特股份有限公司的重组问询函	2015-1-9	2015-1-13	2015-1-13
002354	非许可类重组问询函	关于对浙江科冕木业股份有限公司的重组问询函	2015-1-12	2015-1-15	2015-1-19
002072	非许可类重组问询函	关于对浙江凯瑞德股份有限公司的重组问询函	2015-2-5	2015-2-10	2015-2-10
000035	非许可类重组问询函	关于对浙江中国天楹股份有限公司的重组问询函	2015-2-3	2015-2-5	2015-2-10
002164	年报问询函	关于对浙江宁波东力股份有限公司的重组问询函	2015-3-9	2015-3-11	2015-3-11
002160	年报问询函	关于对浙江常铝股份股份有限公司的重组问询函	2015-3-13	2015-3-20	2015-3-19

4.数据库内容

数据模块	数据表	数据字段
监管问询	监管问询	股票代码、公司简称、函件类别、问询标题、问询内容、公司回复内容、发函日期、限期回复日期、公司回复日期

5.数据范围

数据模块	数据起始时间
全部模块	2014 年

（八）中国上市公司内部控制研究数据库-ICRD

1.数据库简介

中国上市公司内部控制研究数据库（*Internal Control Research Database of Chinese Companies, ICRD*）是以研究为目的而设计的专业学术研究数据库，旨在为国内外从事上市公司内部控制研究的学者和机构从业人员提供精确、完整的学术研究数据。本数据库涵盖中国上市公司有关内部控制方面的重要数据信息，根据 *COSO* 内部控制框架的基本要素，不仅包括上市公司内部控制报告和年报中有关内部控制的内容，而且综合了其他与内部控制相关的内容，完整地体现了企业内部控制的总体状况。

2.数据库特色

- 基于 *COSO* 内控

ICRD 的基础模块设计基于 *COSO* 内控对公司内部控制的整体要求和评价，较为准确地反映了企业内部控制各个方面，有助于从不同角度研究企业内部控制情况。

- 全面完善

ICRD 还考虑了当前我国企业内部控制的各方面，全面、完善地考察了企业内部控制机制。

3.重要数据举例

内部风险披露

股票代码	会计年度	是否披露经营管理风险	是否披露技术创新风险	是否披露财务风险	是否披露人力资源风险	是否披露安全和环境风险
300301	2015	1	0	0	1	0
300351	2015	1	0	1	0	0
002572	2015	1	1	0	0	0
002168	2015	1	0	0	1	0
300298	2015	1	1	0	1	0
002349	2015	1	0	0	1	1
002522	2015	1	0	1	0	0

4.数据库内容

数据模块	数据表	数据字段
内部控制环境	董监高持股	股票代码、会计年度、管理层持股比例、董事会持股比例、监事会持股比例
	独立董事比例	股票代码、会计年度、董事总人数、独立董事人数、独立董事比例
	监事执行情况	股票代码、会计年度、监事会人数、职工监事人数
	两职兼任情况	股票代码、会计年度、董事长与总经理兼任情况
	企业社会责任	股票代码、会计年度、是否发布社会责任报告、是否有慈善捐赠、公司是否关注环境、安全和产品质量
	负面事件	股票代码、会计年度、是否卷入诉讼纠纷、公司是否受到监管部门惩罚、高管是否受到监管部门处罚、公司是否被媒体披露重大负面消息

数据模块	数据表	数据字段
风险管理	内部风险披露	股票代码、会计年度、是否披露经营管理风险、是否披露技术创新风险、是否披露财务风险、是否披露人力资源风险、是否披露安全和环境风险、是否披露其他内部风险、前三年是否有显著的损失
	外部风险披露	股票代码、会计年度、是否披露市场竞争风险、是否披露产品价格风险、是否披露原材料风险、是否披露行业风险、是否披露任何经济环境风险、是否披露政策风险、是否披露其他外部风险
	风险应对	股票代码、会计年度、风险数量、披露的风险类型、披露的风险内容、风险应对策略
控制活动	经营与业绩控制	股票代码、会计年度、存货是否有损失、资产减值损失是否增加、是否有业绩分析报告
	事故安全控制	股票代码、会计年度、是否有严重的生产事故、是否有严重的污染事故、是否有严重的资产安全事故
信息与沟通	信息交流	股票代码、会计年度、董事会秘书是否“金牌秘书”、是否有机制管理与投资者的关系、公司网站上是否有联系投资者的链接、是否有其他与投资者交流的平台
	信息及时性	股票代码、会计年度、定期报告是否公布于预定日期
	反舞弊情况	股票代码、会计年度、公司是否有重述、修正或补充的财务报表、公司是否有任何信息泄漏
监督活动	内部控制披露	股票代码、会计年度、是否有内部控制评价报告、内部控制是否有效、是否有提高内部控制的计划
	内部控制缺陷	股票代码、会计年度、是否有缺陷识别的标准、内部控制是否存在缺陷、是否有改正内部缺陷的计划
	内部审计情况	股票代码、会计年度、是否披露内控审计报告、审计意见类型、审计意见、非标准意见说明、会计师事务所、注册会计师名称

5.数据范围

数据模块	数据起始时间
全部模块	2010 年

（九）上市公司参控股公司数据库-CPCD

1.数据库简介

上市公司参控股公司数据库（*Controlled & Participated Company Database of Listed Company, CPCD*）基于上市公司年报、公告等来源，采集和整理上市公司的子公司、合营企业、联营企业等数据而开发建立的专业数据库。目前，上市公司通常在其财务报告中披露参控股公司信息，但是所披露的信息较为简单，仅可以获得列表信息，为了丰富该领域的研究数据，我司还从各个渠道对参控股公司的其他重要信息进行补充和完善，例如从国家企业信用信息公示系统获取了相关工商信息等。该数据库主要分为六个模块：参控股公司列表、参控股公司基本信息、参控股公司股东信息、参控股公司管理人员、参控股公司法律诉讼、参控股公司对外投资。其中参控股公司列表包含了时间序列上所有上市公司的参控股公司名称、持股比例等重要字段。需要说明的是，由于有关参控股公司的其他信息来自于其他渠道，因此诸如参控股公司基本信息等表中的参控股公司可能无法完全覆盖参控股公司列表中的全部公司。

2.数据库特色

- 唯一提供上市公司子公司、合营公司和联营公司的数据库

CPCD 主要基于年报数据，提供了时间序列上完整和丰富的上市公司参控股公司信息，数据量近 50 万条，是目前市场上唯一提供此类数据批量下载的数据库。

- 提供了上市公司参控股公司的多维度信息

CPCD 除了提供上市公司参控股列表信息以及持股比例等基本信息，还提供了各个参控股公司的重要信息，例如参控股公司的工商信息、股东、高管和法律裁决信息等。

3.典型数据举例

参控股公司列表

股票代码	会计年度	参控股公司名称	参控关系	参控比例	被参控股公司注册资本
000001	1996	深圳市元盛实业有限公司	子公司	100.00%	2101 万元
000001	1997	深圳市元盛实业有限公司	子公司	100.00%	2010 万元
000001	1998	上海银涛高尔夫俱乐部有限公司	子公司	75.00%	2000 万元
000001	1998	深圳市元盛实业有限公司	子公司	100.00%	2010 万元
000001	1998	上海元盛房地产有限公司	子公司	100.00%	1500 万元
000001	1998	大连元盛房屋开发有限公司	子公司	100.00%	1000 万元
000001	1998	上海银涛高尔夫别墅开发有限公司	子公司	100.00%	500 万元

4.数据库内容

数据模块	数据表	数据字段
参控股公司列表	参控股公司列表	股票代码、会计年度、序号、参控股公司名称、参控关系、参控比例、被参控股公司注册资本、货币单位
参控股公司基本信息	参控股公司基本信息	股票代码、参控股公司名称、法定代表人、注册资本、货币单位、经营状态、成立日期、登记机关、所属地区、所属行业、企业地址、经营范围

数据模块	数据表	数据字段
参控股公司 股东信息	参控股公司 股东信息	股票代码、参控股公司名称、参控股公司股东、持股比例、股东类型
参控股公司 管理人员	参控股公司 管理人员	股票代码、参控股公司名称、高管姓名、高管职务
参控股公司 法律诉讼	参控股公司 法律诉讼	股票代码、参控股公司名称、序号、案件名称、发布时间、案件编号、案件身份、执行法院
参控股公司 对外投资	参控股公司 对外投资	股票代码、参控股公司名称、被投资企业名称、被投资企业注册资本、货币单位、出资比例、被投资企业成立日期、被投资企业状态

5.数据范围

数据模块	数据起始时间
参控股公司列表	1990 年
参控股公司基本信息	截面数据
参控股公司股东信息	截面数据
参控股公司管理人员	截面数据
参控股公司法律诉讼	2007 年
参控股公司对外投资	截面数据

(十) 关键审计事项数据库-KAMD

1.数据库简介

关键审计事项数据库 (Key Audit Matters Database, **KAMD**) 是基于上市公司审计报告中披露的关键审计事项段整理而成的专业数据库。关键审计事项,是指注册会计师根据职业判断认为对本期财务报表审计最为重要的事项,从注册会计师与治理层沟通过的事项中选取。为了扩充报告使用者对公司价值判断的信息,同时增强报告使用者的决策相关性和有用性,2016年12月23日中国财政部发布了《中国注册会计师审计准则第1504号——在审计报告中沟通关键审计事项》,要求审计师披露在审计过程中遇到重要的、可能存在重大错误、需要特别关注的事项。**KAMD**根据上市公司年报披露关键审计事项的情况,及时搜集和整理了相关数据,便于学者获得该领域的详细研究数据。

2.数据库特色

- 数据的前沿性

目前,国际上已有多篇论文研究关键审计事项,国内对于该领域尚处于起步阶段,**KAMD**的推出紧跟了研究前沿和热点。

3.典型数据举例

关键审计事项

股票代码	会计年度	关键审计事项类型	关键审计事项内容	应对方法
000002	2016	存货的可变现净值的评估	2016年12月31日,贵公司及其子公司(以下合称“贵集团”)已完工开发产品、在建开发产品及拟开发产品(以下统称“存货”)的账面价值合计金额重大。该等存货按照成本与可变……	与评价存货的可变现净值相关的审计程序中包括以下程序:……
000002	2016	土地增值税的计提	贵集团应缴纳的主要税项之一为土地增值税。贵集团销售开发的房地产需要就土地增值额按照超率累进税率30%-60%缴纳土地增值税。……	与评价土地增值税的计提相关的审计程序中包括以下程序:评价管理层与计量预计的土地增值税相关的内部控制的运行有效性;……

4.数据库内容

数据模块	数据表	数据字段
关键审计事项	关键审计事项	股票代码、会计年度、关键审计事项类型、关键审计事项内容、应对方法

5.数据范围

数据模块	数据起始时间
全部模块	2016年

(十一) 中国上市公司供应链研究数据库-SCRD

1.数据库简介

中国上市公司供应链研究数据库 (*Supply Chain Research Database of Chinese Companies, SCRD*) 是基于对上市公司财务报告的分析而建立的研究上市公司供应商和客户的专业数据库。

2.数据库特色

- 数据完整、全面

相比于其他类似数据库, 无论从数据质量还是数量来看, *SCRD* 所包括的上市公司客户与供应商信息更为准确和完整。

3.典型数据举例

供应商名称及采购额

股票代码	公司简称	序号	会计年度	供应商名称	供应商采购额 (元)	供应商采购额占比
000038	深大通	1	2012	山东蒯建建筑集团有限公司	10520967.18	16.4%
000038	深大通	2	2012	兖州市建筑安装工程总公司	9883378.17	15.4%
000038	深大通	3	2012	青岛德泰建设工程有限公司	6400911	9.98%
000038	深大通	4	2012	山东港基建设集团有限公司	5378335.7	8.38%
000038	深大通	5	2012	天元建设集团有限公司	4023947.87	6.27%

4.数据库内容

数据模块	数据表	数据字段
主要供应商信息	供应商名称及采购额	股票代码、公司简称、序号、会计年度、供应商名称、供应商采购额、供应商采购额占比 (%)
	供应商总采购信息	股票代码、公司简称、会计年度、前五大供应商采购总额、前五大供应商采购额占比 (%)
主要客户信息	客户名称及收入	股票代码、公司简称、序号、会计年度、客户名称、客户收入、客户收入占比 (%)
	客户总收入信息	股票代码、公司简称、会计年度、前五大客户收入、前五大客户收入占比 (%)

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

(十二) 中国上市公司异常收支数据库-ARED

1. 数据库简介

中国上市公司异常收支数据库 (*Abnormal Revenue and Expenditure Database, ARED*) 涵盖我国上市公司高管在职消费、营业外收支明细等非正常收支数据。*ARED* 包括高管在职消费和营业外收支 2 个模块, 共 4 张表。高管在职消费模块依据前人研究经验, 区分为直接法下和间接法下的高管在职消费数据, 这些数据主要包括业务招待费、差旅费、办公费、会议费、交际应酬费等多类高管在职消费支出, 其中, 根据这些科目的对象又将其归类为管理费用和销售费用下的数据; 营业外收支明细模块包括公司的营业外收入明细和营业外支出明细模块。

2. 数据库特色

- 紧跟学术研究

ARED 紧跟该领域最新研究成果, 归纳整理了研究高管在职消费和营业外收支用到的研究数据。

3. 典型数据举例

高管在职消费_直接法

股票代码	会计年度	差旅费-销售费用 (元)	差旅费-管理费用 (元)	办公费-销售费用 (元)	办公费-管理费用 (元)	董事会会费 (元)
000713	2013	10305939.67	1116423.6	1820774.03	2972753.02	487224
000713	2014	10093385.54	1006174.5	1537123.12	2324831.78	652417.37
000713	2015	9606328.04	887482.41	1708404.77	3024887.01	560960

4. 数据库内容

数据模块	数据表	数据字段
高管在职消费	直接法	股票代码、会计年度、业务招待费-管理费用、差旅费-销售费用、差旅费-管理费用、办公费-销售费用、办公费-管理费用、交际应酬费-销售费用、交际应酬费-管理费用、通讯费-销售费用、通讯费-管理费用、出国培训费、董事会会费、小车费-销售费用、小车费-管理费用、会议费、高管总人数、主营业务收入
	间接法	股票代码、会计年度、管理费用、董事/监事/高管成员薪酬、坏账准备、存货跌价准备、无形资产摊销、总资产、存货、固定资产、营业收入
营业外收支明细	营业外收入明细	股票代码、公司简称、会计年度、非流动资产处置利得、债务重组利得、非货币性资产交换利得、接受捐赠、政府补助、罚没收入、赔偿款收入、特殊项目、其他
	营业外支出明细	股票代码、公司简称、会计年度、非流动资产处置损失、债务重组损失、非货币性资产交换损失、罚没支出、对外捐赠、特殊项目、其他

5. 数据范围

数据模块	数据起始时间
高管在职消费	1991 年
营业外收支明细	2006 年

(十三) 投资者关系管理数据库-IRMD

1. 数据库简介

投资者关系管理是上市公司通过与投资者进行有效沟通，促进投资者或潜在投资者对公司的了解和认同，让投资者发现公司价值，并保持对公司长期认可，提升投资者忠诚度，进而实现公司价值最大化的有效管理手段。

投资者关系管理数据库 (*Investor Relationship Management Database, IRMD*) 包含公司基本信息、投资者网络互动、投资者实地调研、投资者关系等四个模块，共 9 张表。

公司基本信息模块 主要是公司地址信息，如公司所属省市县、注册地址、办公地址、邮编等信息。

投资者网络互动模块 包含用户与公司问答（如投资者所提的具体问题和上市公司回答内容、以及问答时间和是否回复等）和公司互动情况统计（如每一年该公司被提问数及回复数量）。

投资者实地调研模块 包含访客、接待人员、调研内容、调研内容标签等四方面信息。

投资者关系模块 包括是否有其他与投资者交流的平台、公司董秘是否金牌董秘。

2. 数据库特色

- 多维和详尽

从公司基本信息、投资者网络互动、投资者实地调研和投资者关系等四个维度全面整理了投资者关系管理的相关信息。此外，本数据库包含的字段信息极其完整，如投资者网络互动模块，包含多个互动平台上的互动信息；投资者实地调研模块，包含私募投资者、基金公司、分析师、个人投资者等访客对公司进行的实地调研，本数据库还对调研内容进行分类编码。

3. 典型数据举例

公司与投资者互动情况统计

股票代码	公司简称	年度	被提问数	回复数
600004	白云机场	2012	23	17
600010	包钢股份	2012	92	80
600017	日照港	2012	99	37
600019	宝钢股份	2012	805	578

4. 数据库内容

数据模块	数据表	数据字段
公司基本信息	公司地址信息	股票代码、公司简称、公司中文名称、公司简介、所属省份、所属地级市、所属区县、注册地址、办公地址、邮编
投资者网络互动	用户与公司问答	股票代码、公司简称、浏览用户/注册用户、用户名、提问时间、提问内容、上市公司是否回复、回复时间、回复内容
	公司互动情况统计	股票代码、公司简称、年度、被提问数、回复数
投资者实地调研	访客	股票代码、公司简称、公开时间、调研时间、参与单位、参与单位类别、参与人员

数据模块	数据表	数据字段
投资者实地调研	接待人员	股票代码、公司简称、公开时间、调研时间、调研接待人员 n 姓名、调研接待人员 n 职务
	调研内容	股票代码、公司简称、调研时间、调研报告内容
投资者关系	是否有其他与投资者交流的平台	股票代码、公司简称、公司主页是否有投资者关系管理、公司主页是否有投资者咨询、公司主页是否有微信微博
	公司董秘是否金牌董秘	股票代码、公司简称、统计日期、董秘姓名、届数

5.数据范围

数据模块	数据起始时间
公司基本信息	1991 年
投资者网络互动	2009 年
投资者实地调研	2011 年
投资者关系	2006 年

(十四) 海外上市研究数据库-COLD

1. 数据库简介

海外上市研究数据库 (*Overseas Listing Research Database, COLD*) 是基于公司公告、招股说明书、年度报告等数据整理而成的研究中国公司海外上市的专业学术数据库。*COLD* 涵盖了中国公司在海外股票市场, 例如纽约证券交易所、纳斯达克证券交易所、伦敦证券交易所的上市情况。*COLD* 包括 *IPO* 概况、公司股东、公司高管、财务报表状况和审计单位等共 7 张表。*IPO* 概况表总结了与公司海外上市相关的上市日期、所在股票交易所、承销商等信息; 公司股东表包括股东名称、持股数和持股比例等字段; 公司高管包括高管姓名、职位、出生年份等字段; 审计单位表包括会计师事务所名称、审计意见等字段。此外, 该模块还包含上市后的三大财务报表信息。

2. 数据库特色

- 全面的数据和字段

COLD 不仅提供了市场上已有的海外上市公司的财务报表数据等信息, 还提供较难获得但对于研究又不可或缺的公司高管、股东和审计单位等重要数据。

3. 典型数据举例

(1) 海外 IPO 概况

股票代码	公司中文名称	交易所代码	上市日期	上市方式	发行股票数 (股)	募集资金 (USD)
ATAI	ATA 公司	NASDAQ	2008-01-29	境外直接上市	4874012	46300000
BCDS	BCD 半导体	NASDAQ	2011-01-28	境外直接上市	6000000	63000000
BIDU	百度	NASDAQ	2005-08-05	境外直接上市	4040402	109090000
BONA	保利博纳公司	NASDAQ	2010-12-09	境外直接上市	11740000	99790000
CTFO	北京北大千方科技有限公司	NASDAQ	2008-07-31	转板上市	1564945	10000000

4. 数据库内容

数据模块	数据表	数据字段
海外上市研究	IPO 概况	股票代码、公司中文名称、公司英文名称、交易所代码、交易所名称、上市日期、退市日期、上市方式、发行股票数目、募集资金、货币类型、行业名称、承销商
	公司股东	股票代码、公司名称、统计日期、股东排名、股东名称、持有股份、持股比例 (%)
	公司高管	股票代码、公司名称、高管姓名、职位、任职日期、离职日期、出生年份、性别、国籍、简历
	资产负债表	股票代码、截止日期、货币类型、现金及现金等价物、交易性金融资产、衍生金融资产、短期投资、应收账款、可收回税项、存货、其他流动资产、流动资

数据模块	数据表	数据字段
海外上市研究	资产负债表	产合计、可供出售金融资产、持有至到期投资、物业、厂房以及设备、投资性房地产、长期股权投资、长期投资、预付款项及其他应收款项、无形资产、商誉、递延所得税资产、其他非流动资产、非流动资产合计、资产总计、短期借款、交易性金融负债、衍生金融负债、应付账款、应交税费、应付股息及利息、其他应付款、流动负债-递延收益、一年内到期的非流动负债、其他流动负债、流动负债合计、长期借款、长期应付款、递延所得税负债、非流动负债-递延收益、其他非流动负债、非流动负债合计、负债合计、股本、专项储备、股本溢价、留存收益、其他储备、拟派股息、少数股东权益、归属于母公司所有者权益合计、所有者权益合计、负债及股东权益总计
	利润表	股票代码、起始日期、截止日期、货币类型、营业收入、其他业务收入、营业成本、销售费用、管理费用、折旧与摊销、员工薪酬、研发费用、资产减值损失、营业外收入、营业外支出、其他业务成本、投资收益、公允价值变动损益、对联营企业和合营企业的投资收益、财务费用、税前利润其他项目、利润总额、所得税费用、净利润、归属于母公司所有者的净利润、归属于少数股东的净利润、基本每股收益、稀释每股收益、其他综合收益、综合收益总额、归属于母公司的综合收益
	现金流量表	股票代码、起始日期、截止日期、货币类型、税前利润、资产减值准备、出售物业、厂房及设备的亏损(收益)、利息支出、利息收入、证券投资收益/亏损、基于股票的补偿费、减值及拨备、递延所得税、资产处置损益、投资损益、权益性投资损益、重估盈余、汇兑损益、经营业务调整其他项目、存货的减少(增加)、存款及限制性现金减少(增加)、待摊费用及其他资产、应收账款减少(增加)、预付款项、按金及其他应收款项减少(增加)、应付账款增加(减少)、预收账款、按金及其他应付款增加(减少)、递延收入、应付税项、应计费用及其他负债、应收利息减少(增加)、应付利息增加(减少)、经营资金变动其他项目、经营活动产生的现金、已收利息、已付利息、已付税项、经营业务其他项目、经营活动产生的现金流量净额、购建固定资产、处置固定资产、购买固定资产产生的现金流量净额、购买无形资产及其他资产、处置无形资产及其他资产、购买/出售无形资产收到的净额、购买或收购业务产生的现金、出售业务产生的现金流量、购买或出售业务产生的现金流量净额、已收利息及股息(投资)购买金融投资所支付的现金、出售金融投资所收到的现金、购买/出售投资产生的现金流量净额、存款增加(减少)、贷款偿还……
	审计单位	股票代码、会计年度、会计师事务所名称、审计意见、审计费用、货币类型

5.数据范围

数据模块	数据起始时间
海外上市研究	1999年

(十五) 中国法学研究数据库-CNLA

1. 数据库简介

中国法学研究数据库 (*Chinese Law Research Database, CNLAW*) 是面向广大高校和学者以及实务界人士, 旨在提供完整、精确的法学研究数据。此数据库全面涵盖法学研究所需要的各种信息, 不仅收录了海量的公司法律裁决信息, 还对上市公司诉讼仲裁事件进行汇总, 同时将律所、律师等法律相关机构和人员的信息进行全面整合。*CNLAW* 包含上市公司法律裁决、知识产权法律裁决、上市公司诉讼仲裁、律师事务所及律师等 5 个数据模块。

2. 数据库特色

- 数据维度多样

CNLAW 包括海量上市公司法律裁判信息、上市公司诉讼仲裁信息、法律相关机构及人员的多维度特征信息。

- 内容详细完善

CNLAW 对各个维度的信息进行了详细地描述, 比如, 在上市公司诉讼仲裁模块中, 不仅对一审进行详细描述, 更针对是否上诉, 对二审也进行了细致描述, 甚至还追踪了后续的执行情况。

3. 典型数据举例

(1) 上市公司诉讼仲裁

公司简称	公告日期	案件名称	原告方	被告方	诉讼类型	涉案金额	起诉日期
当代东方	2004-12-29	转让款纠纷	大同水泥股份有限公司	海南金牛膜料开发有限公司	民事诉讼	2436 万元	2003-08-22
益佰制药	2004-12-24	借款合同纠纷	交通银行上海浦东分行	上海源通医药有限公司	民事诉讼	1000 万元	2004-09-20
中原环保	2004-12-15	赔偿其担保损失纠纷	国泰君安证券股份有限公司	白鸽(集团)股份有限公司	民事诉讼	3998 万元	2001-12-31

(2) 律师情况统计

律师名称	是否合伙人	资格证取得日期	首次执业日期	人员类型	执业状态	执业机构
白云飞	是	1996-07-10	2001.4.30	专职律师	执业	北京市德渊律师事务所
薄智全	是	1997-07-28	1998.5.28	专职律师	派驻	北京市同一源律师事务所
蔡子明	是	2000-09-15	2001.1.1	专职律师	执业	北京市公衡律师事务所
曹三明	是	1985-05-20	1985.5.20	专职律师	执业	北京市华德莱律师事务所

4. 数据库内容

数据模块	数据表	数据字段
上市公司法律裁决	民事一审(原被告)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、原告、被告、第三人、法律依据

数据模块	数据表	数据字段
上市公司法律裁决	民事一审 (申请与被申请)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、申请人、被申请人、法律依据
	民事二审 (上诉与被上诉)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、上诉人、被上诉人、法律依据
	民事再审 (再审申请与被申请)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、再审申请人、被申请人、法律依据
	民事再审 (申诉与被申诉)	股票代码、案件年份、案件名称、案号、案件类型、案由、审理法院、法院层级、审理程序、裁判日期、申诉人、被申诉人、法律依据
	执行裁决	股票代码、案件年份、案件名称、案号、案件类型、审理法院、法院层级、裁判日期、申请执行人、被执行人、法律依据
	行政一审	股票代码、案件年份、案件名称、案号、案件类型、审理法院、法院层级、审理程序、裁判日期、原告、被告、第三人、法律依据
	行政二审	股票代码、案件年份、案件名称、案号、案件类型、审理法院、法院层级、审理程序、裁判日期、上诉人、被上诉人、原审第三人、法律依据
	行政审查	股票代码、案件年份、案件名称、案号、案件类型、审理法院、法院层级、审理程序、裁判日期、申请执行人、被执行人、法律依据
知识产权法律裁决	一审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、原告、被告、委托代理人、律师、律所、审判长、审判员、书记员
	二审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、上诉人、被上诉人、委托代理人、律师、律所、审判长、审判员、书记员
	再审程序	股票代码、案件年份、案件名称、案号、审理法院、法院层级、裁判日期、案由、审理程序、再审申请人、被申请人、委托代理人、律师、律所、审判长、审判员、书记员
上市公司诉讼仲裁	上市公司诉讼仲裁统计	股票代码、公司简称、当事人类型、公告日期、案件名称、案件简介、原告方、被告方、诉讼类型、涉案金额、币种、起诉日期、一审受理法院、一审判决日期、一审判决内容、是否上诉、上诉方、二审受理法院、二审判决日期、二审判决内容、执行情况
律师事务所	律所情况统计	律所名称、律所代码、律所地址、主管司法局、律所负责人、总所/分所、执业状态、组织形式、注册资金、律所简介
律师	律师情况统计	律师姓名、性别、民族、学历/学位、专业、政治面貌、是否合伙人、资格证取得日期、首次执业日期、人员类型、执业状态、执业机构、考核结果

5.数据范围

数据模块	数据起始时间
全部模块	1995 年

上市公司新闻舆情

（一）中国上市公司财经新闻数据库-CFND

1. 数据库简介

中国上市公司财经新闻数据库 (*Financial News Database of Chinese Listed Companies, CFND*) 是国内首个采用人工智能算法来采集、整理和分析上市公司财经新闻的大数据数据库。本数据的数据来源涵盖 400 多家网络媒体和 500 多家报纸刊物。海量的上市公司财经新闻数据成为本数据库的核心竞争力，以海量数据为基础，通过数据清洗、专家判断等手段，结合人工智能算法为学者提供上市公司财经新闻基本信息、财经新闻量化统计信息以及新闻相似度等。

CFND 包括网络财经新闻和报刊财经新闻 2 个模块，7 张表。网络财经新闻包括了来自 400 多家重要网络媒体的新闻报道数据，其中，最主要的是包括二十家主流网络财经媒体的新闻报道，如：和讯网、新浪财经、东方财富网、腾讯财经、网易财经、凤凰财经、中国经济网、搜狐财经、金融界、华讯财经、FT 中文网、全景网、中金在线、中国证券网、证券之星、财新网、澎湃新闻、第一财经、21CN 财经频道、财经网。这二十家网络财经媒体不仅在财经新闻报道数据和质量上属于国内前列，而且也是投资者经常浏览和关注的财经网站，因此其报道的新闻具有重要价值。此外，为了丰富数据来源，提供更多有用的新闻信息，该模块还包括了其他 400 余家大型重要网站、行业网站或地方性网站。该模块的新闻数据量为从 2002 年开始至今全部财经新闻，囊括了个股新闻、宏观经济报道、行业报道等，截止至 2018 年底，累计达 4000 多万条原始新闻。

报刊财经新闻包括了来自 500 多家重要报纸媒体的新闻数据，其中，最主要的是包含了研究常用到的八大主流财经报纸：中国证券报、上海证券报、第一财经日报、21 世纪经济报道、中国经营报、经济观察报、证券日报和证券时报。这八大财经报纸不仅具有财经新闻报道及时、质量高、影响力大等特点，还是国内经济、管理和商学研究媒体问题时经常用到的数据来源。同时，为了体现新闻报道的全面性，该模块还包括其他 500 余家重要报纸刊物，主要包括重要性中央报纸，地方性晨报、日报、晚报以及其他财经报刊。该模块的原始新闻数量为从 1994 年开始至今约 4000 多万条新闻报道。

无论是网络财经新闻还是报刊财经新闻，为了匹配得到上市公司的相关新闻，我们在几千万条原始新闻的基础上，经过数据清洗与匹配、新闻情感判断、相似新闻计算等方法，得到了最终较为科学的结果。

2. 数据库特色

- 新闻报道数量巨大

CFND 不仅涵盖网络上的主流财经网站、国家级网站、地方网站、行业网站，还包括主流财经报刊、国家级报刊等重要报刊新闻，总共 400 多家网络媒体和 500 多家报纸刊物，几千万条新闻数据，可以充分满足研究需要。

- 新闻情感识别准确

CFND 使用计算机领域成熟的文本情感判断方法，文章正、中、负面语气识别率较高，样本内的准确率高达 85%，可以较好地捕捉上市公司好、坏消息。

3. 典型数据举例

(1) 报刊财经新闻基本信息

股票代码	公司简称	新闻标题	报道时间	报道媒体	新闻情感	是否原创
000002	万科 A	城东新城将建设 杭州未来的“蓬皮社”中心	2015-12-07	今日早报	1	是
000063	中兴通讯	助产业转型升级 促外贸稳定增长	2015-12-04	中国国门时报	1	否
002464	金利科技	金利科技终止收购第七大道 标的估值两年后大幅缩水	2015-11-19	每日经济新闻报	-1	是
600519	贵州茅台	贵州茅台股价再创历史新高	2017-01-05	深圳商报	1	是
601107	四川成渝	遂西高速 明日 9 时开通试运营	2015-12-09	成都商报	0	是

注：新闻情感字段的数字含义如下：1 表示正面，-1 表示负面，0 表示中性。

(2) 网络新闻量化统计（按自然日）

股票代码	统计日期	全部正面新闻数	全部中性新闻数	全部负面新闻数	原创正面新闻数	原创中性新闻数	原创负面新闻数
000002	2016-05-25	17	28	6	3	3	3
000002	2016-05-26	14	33	16	2	4	6
000002	2016-05-27	15	34	32	4	3	2
000002	2016-05-28	2	6	10	1	0	3
000002	2016-05-29	9	1	9	0	0	1

4. 数据库内容

数据模块	数据表	数据字段
网络财经新闻	网络财经新闻基本信息	股票代码、公司简称、行业名称、行业代码、新闻 ID、新闻标题、报道时间、报道媒体、报道媒体所在地区、来源媒体、新闻情感、是否原创、采集网址、正文句子数、标题是否提及、正文提及句子数、正文提及次数、提及全部公司数、提及全部公司句子数、全部公司提及次数
	网络财经新闻量化统计（按自然日）	股票代码、公司简称、统计日期、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	网络财经新闻量化统计（按交易时间）	股票代码、公司简称、统计日期、时间段、星期时间、是否交易时间段、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	网络财经新闻相似度	新闻 1_ID、新闻 1_标题、新闻 1_所属公司代码、新闻 2_ID、新闻 2_标题、新闻 2_所属公司代码、新闻相似度
报刊财经新闻	报刊财经新闻基本信息	股票代码、公司简称、行业名称、行业代码、新闻 ID、新闻标题、报道时间、报道媒体、报道媒体所在地区、新闻情感、是否原创、正文

数据模块	数据表	数据字段
报刊财经新闻	报刊财经新闻基本信息	句子数、标题是否提及、正文提及句子数、正文提及次数、提及全部公司数、提及全部公司句子数、全部公司提及次数
	报刊财经新闻量化统计	股票代码、公司简称、统计日期、标题出现该公司的新闻总数、内容出现该公司的新闻总数、正面新闻数（全部）、中性新闻数（全部）、负面新闻数（全部）、正面新闻数（原创）、中性新闻数（原创）、负面新闻数（原创）
	报刊财经新闻相似度	新闻 1_ID、新闻 1_标题、新闻 1_所属公司代码、新闻 2_ID、新闻 2_标题、新闻 2_所属公司代码、新闻相似度

5.数据范围

数据模块	数据起始时间
网络财经新闻	2002 年
报刊财经新闻	1994 年

(二) 中国上市公司股吧评论数据库-GUBA

1.数据库简介

中国上市公司股吧评论数据库 (*Stocks Comments of Chinese Listed Companies, GUBA*) 是基于我国网络股吧论坛中上市公司的帖子评论进行统计研究而建立的专业性财经文本数据库。该数据库主要根据中国最大的股吧论坛网民对上市公司帖子进行文本分析、数量统计等, 从中得到维度众多、数量丰富的研究数据。**GUBA** 包括股吧文章统计 (自然日) 和股吧文章统计 (交易时间) 2 张表。股吧文章统计 (自然日) 根据自然日时间 (0:00-24:00) 对一天中的发帖总量、阅读总量和评论总量进行统计; 股吧文章统计 (交易时间) 则根据交易时间统计了不同时间段内的发帖情况, 这种计算方法可以直接反映交易时间内和非交易时间内投资者情绪对股市的影响。**GUBA** 采用机器学习方法对每个帖子的正面、负面和中性情感进行判断, 并统计当天该上市公司所在股吧的正面、负面和中性帖子总量, 这大大丰富了学者的研究维度。

2.数据库特色

- 数据量庞大, 计算方法准确

覆盖了从 2008 年以来, 所有上市公司的股吧股民发帖、评论和阅读情况, 原始数据达到几亿条, 可用于大数据挖掘和研究。从自然日和交易时间, 帖子正、中、负面情感等多个维度计算了每个上市公司的股吧情况, 这可为拓展研究广度和深度提供数据支持。

3.典型数据举例

股吧文章统计 (自然日)

股票代码	统计日期	帖子总量	正面帖子量	负面帖子量	中性帖子量	阅读数	评论数
000665	2016-07-14	44	12	12	20	24117	42
000665	2016-07-15	26	5	7	14	18997	23
000665	2016-07-16	14	6	2	6	7341	1

4.数据库内容

数据模块	数据表	数据字段
股吧文章统计	股吧文章统计 (自然日)	股票代码、统计日期、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数
	股吧文章统计 (交易时间)	股票代码、公司简称、获取时间、时间段、星期时间、是否交易时段、帖子总量、正面帖子量、负面帖子量、中性帖子量、阅读数、评论数

5.数据范围

数据模块	数据起始时间
股吧文章统计	2008 年

（三）网络搜索指数数据库-WSVI

1.数据库简介

网络搜索指数数据库 (*Web Search Volume Index of Chinese Listed Companies, WSVI*) 以各种网络搜索指数为基础, 涵盖各种主流网络搜索引擎对上市公司的搜索频率情况, 综合新闻舆情等信息, 可以反映网民情绪、公司搜索热度等, 是衡量上市公司关注度及其变化情况的关键指标。

2.数据库特色

- 搜索关键词全面、准确, 涵盖全国和地方性搜索指数

WSVI 不仅包括上市公司全称、简称、股票代码的搜索指数, 同时还包括全国性的网络搜索指数, 并对每个省份搜索指数进行统计, 反映不同地区对上市公司的搜索热度。

3.典型数据举例

上市公司网络搜索指数 (全国)

股票代码	行业代码	指数值 (股票代码)	指数值 (全部)	关键字个数	统计日期
000001	J66	6702	18273	2	2011-01-04
000001	J66	6371	17234	2	2011-01-05
000001	J66	6236	17135	2	2011-01-06
000001	J66	6658	16447	2	2011-01-07

4.数据库内容

数据模块	数据表	数据字段
上市公司网络搜索指数	上市公司网络搜索指数 (全国)	股票代码、行业代码、指数值 (股票代码)、指数值 (全部)、关键字个数、统计日期
	上市公司网络搜索指数 (地区)	股票代码、行业代码、地区代码、指数值 (股票代码)、指数值 (全部)、关键字个数、统计日期

5.数据范围

数据模块	数据起始时间
上市公司网络搜索指数	2011 年

（四）中国上市公司社交媒体数据库-CSMD

1.数据库简介

中国上市公司社交媒体数据库 (*Social Media Database of Chinese Listed Companies, CSMD*) 是基于对上市公司的社交媒体信息调查研究而建立的专业数据库。该库涵盖了券商微信公众号文章信息和财经类微信公众号文章信息。前者主要是指中国主要券商在其公众号上发布的对上市公司的推荐文章信息，后者主要指民间财经类公众号对上市公司的报道和分析等。

2.数据库特色

● 前沿

微信信息近期成为会计金融学术界研究热点，但目前还没有相应的数据库。*CSMD* 紧跟学术研究前沿，采用数据挖掘方法对这类信息进行文本分析、数量统计和分类规整等，可以为学术界提供全面、准确的财经文本数据。

3.典型数据举例

券商微信公众号

文章标题	发布时间	获取时间	来源公众号名称	文章字数	标题包含公司	正文包含公司
【重点推荐】恒逸石化 (000703) 化纤估值洼地，“供应链金融”增色	2015-06-01	2016-08-24	国泰君安证券研究	675	000703	000703
【重点推荐】中兴通讯 (000063) 中兴 2014: 从困顿反转到再次腾飞	2014-04-14	2016-08-24	国泰君安证券研究	698	000063	000063
【地产行业】韩思怡: 万科 A (000002) B 转 H 获证监会核准, 比价效应对 A 股利好明显	2014-03-04	2016-08-29	申万宏源研究	985	000002	000002 600340

4.数据库内容

数据模块	数据表	数据字段
微信公众号 财经新闻	券商微信公众号	文章标题、文章内容、发布时间、获取时间、来源公众号名称、文章字数、标题包含公司、正文包含公司、文章包含的所有公司
	财经微信公众号	文章标题、文章内容、发布时间、获取时间、来源公众号名称、文章字数、标题包含公司、正文包含公司、文章包含的所有公司

5.数据范围

数据模块	数据起始时间
全部模块	2012 年

(五) 上市公司并购新闻数据库-MAND

1.数据库简介

上市公司并购新闻数据库 (*Mergers and Acquisition News Database, MAND*) 是基于上市公司新闻库,对新闻事件类型进行判断识别后提取的数据库。市场上出现的并购重组新闻或者类似的小道消息、传闻等,都会影响上市公司股票的走势,因此对上市公司并购重组新闻的识别以及真实性判断具有重大意义。本数据库针对国内上市公司的新闻数据,获取并购重组事件发生前后的新闻,对并购新闻数据进度深度挖掘,抽取出并购各方名称、并购金额和转让股权比例等多个重要字段,同时判断该新闻是否属于小道消息或相关传闻。

2.数据库特色

- 精确的并购新闻提取

MAND 从海量上市公司新闻中准确提取出与并购重组相关的新闻,同时使用计算机处理和人工交叉核对流程,识别精度较高。

- 丰富的并购新闻字段

为了满足学者对并购领域的相关研究需求, *MAND* 展示了较为丰富的并购相关字段。

3.典型数据举例

上市公司并购新闻

股票代码	新闻日期	新闻标题	新闻来源	转让比例	并购价格
600533	2015/12/2	栖霞建设拟 13.84 亿收购河北银行 8.78%股份	中国财经时报网	8.78%	13.84 亿元
002464	2016/9/21	转型再升级 金利科技拟 18 亿收购微屏软件完善游戏产业链	证券日报网	93%	18.41 亿元
300282	2014/1/15	汇冠股份拟 10.7 亿元收购旺鑫精密	中国经济网	100%	10.7 亿元
600869	2014/6/25	远东电缆拟 1.8 亿并购艾能电力	中国经济网	51%	1.8 亿元
600378	2014/7/23	增强天科股份控制权 中化资产要约收购 1742 万股	中国经济网	5.86%	2.24 亿元

4.数据库内容

数据模块	数据表	数据字段
上市公司 并购新闻	上市公司 并购新闻	股票代码、公司简称、新闻日期、新闻标题、新闻内容、新闻来源、买方名称、卖方名称、标的方名称、转让比例、并购价格、货币单位、是否为传闻、说明

5.数据范围

数据模块	数据起始时间
全部模块	2007 年

（六）上市公司财务舞弊新闻数据库-FFND

1.数据库简介

上市公司财务舞弊新闻数据库 (*Financial Fraud News Database of Listed Companies, FFND*) 是基于上市公司财务舞弊的相关媒体报道, 从中分离出三类重要媒体报道, 包括财务舞弊相关正式公告发布之前的媒体报道、因财务舞弊被立案调查期间的媒体报道以及财务舞弊相关正式公告发布后的媒体报道。具体来说, 这三类财务舞弊新闻分别是: (1) 媒体质疑类: 该新闻是媒体对上市公司存在财务舞弊的质疑, 上市公司并没有发布公告也没有被监管机构立案调查, 只是媒体等民间机构对上市公司的一种纯粹质疑。(2) 立案调查类: 上市公司发生财务丑闻, 监管机构介入调查或者做出处罚。(3) 发布公告类: 上市公司发布公告承认存在财务舞弊, 或因收到处罚通知而发布公告, 或者对财务舞弊的传言进行澄清。

2.数据库特色

- 精确的财务舞弊新闻识别和分类

FFND 采用机器学习和人工识别的方法, 从海量新闻数据中准备识别出了上市公司的三类财务舞弊新闻, 数据质量较高。

3.典型数据举例

上市公司财务舞弊新闻

股票代码	新闻类别	文章标题	发布时间	新闻来源
000040	立案调查类	宝安地产深陷前任“虚假陈述”索赔案	2014-09-04	网易
000048	立案调查类	康达尔涉嫌虚假陈述 遭证监会立案调查	2012-12-03	网易
002131	发布公告类	业绩变脸解释站不住脚 利欧股份或涉嫌虚假陈述	2012-07-18	和讯网
002200	发布公告类	绿大地独董质疑公司财务造假	2011-10-28	金融界
601311	媒体质疑类	骆驼集团涉嫌财务造假 可能虚构6个亿以上收入	2011-04-17	网易
601388	媒体质疑类	怡球资源财务数据存蹊跷 涉嫌虚增利润数百万	2012-11-09	网易
601566	媒体质疑类	九牧王存货跌价计提过低 涉嫌虚增利润	2013-07-05	财经网

4.数据库内容

数据模块	数据表	数据字段
上市公司财务舞弊新闻	上市公司财务舞弊新闻	股票代码、公司简称、新闻类别、文章标题、文章内容、发布时间、新闻来源

5.数据范围

数据模块	数据起始时间
全部模块	2003年

(七) 上市公司高管新闻数据库-CEND

1.数据库简介

上市公司高管新闻数据库 (*Executives News Database of Listed Company, CEND*) 是基于中国财经大数据研究系统 (*CFBD*) 记录的来自 400 多家网络新闻媒体和 500 多家报纸新闻媒体所获得的 9000 多万条原始新闻数据, 通过确定公司高管姓名、上市公司关键词和相关规则等, 使用机器学习和人工检查分析新闻标题、新闻内容, 与相关领域专家讨论, 从而识别出的与上市公司高管相关的新闻报道。*CEND* 包括高管网络新闻和高管报纸新闻两张表, 前者主要提取自上市公司网络新闻报道, 后者提取自上市公司报纸新闻报道。这些来源网站与报纸均包括了财经类最重要的来源。

2.数据库特色

- 精确的高管新闻匹配

CEND 除从几千万条新闻数据中较精确的匹配出了与上市公司高管相关的新闻, 并从多个维度列出该高管新闻的相关字段, 例如新闻情感、相似新闻数量、新闻内容等。

3.典型数据举例

上市公司高管报刊新闻

股票代码	高管姓名	性别	新闻来源	新闻日期	新闻标题	相似新闻数量	新闻情感
000670	史浩樑	男	证券时报	2014-05-22	深耕智能产品芯片设计 掘金互联网时代大潮	4	1
002405	孙玉国	男	证券时报	2014-05-23	腾讯携四维图新布局智能 地图	2	0
002405	吴劲风	男	证券时报	2014-05-23	腾讯携四维图新布局智能 地图	2	0
600645	夏亮	男	证券时报	2014-05-23	中源协和重组方案确定 8 亿收购执诚生物	0	1

4.数据库内容

数据模块	数据字段
上市公司高管网络新闻	股票代码、公司简称、高管编码、高管姓名、职务、性别、报道媒体、报道时间、新闻标题、新闻内容、新闻情感、相似新闻数量
上市公司高管报刊新闻	股票代码、公司简称、高管编码、高管姓名、职务、性别、报道媒体、报道时间、新闻标题、新闻内容、新闻情感、相似新闻数量

5.数据范围

数据模块	数据起始时间
上市公司高管网络新闻	2006 年
上市公司高管报刊新闻	2001 年

上市公司文本信息

(一) 年报文本语气数据库-ARTD

1. 数据库简介

年报文本语气数据库 (*Annual Report's Tone Database, ARTD*) 基于上市公司 2007 年以来的年报文本信息, 根据已有中文语调研究成果, 采用多种方法计算上市公司年报文本的语气, 并建立消极词汇数、积极词汇数、年报语调等多个字段, *ARTD* 采用包括 LM 词典和台湾大学情感词典在内的多种前沿词典和方法计算年报文本语气, 为该领域的深度研究提供了数据支撑。

2. 数据库特色

- 准确和权威的情感判断词典

ARTD 制作过程中参考了发表在权威期刊上的多篇文献, 并借鉴这些已有研究成果使用的词典, 对年报语调进行了多字段展示。

- 数据唯一、全面

随着财经类文本语气研究成为热点, 为了减少学者搜集和整理数据的成本, 提供了多年度完整的年报文本信息, 为该领域研究提供了数据支持。

3. 典型数据举例

年报文本语气

股票代码	会计年度	LM 词典积极词汇数	LM 词典消极词汇数	台大词典积极词汇数	台大词典消极词汇数	基于 LM 词典的 TONE1	基于台大词典的 TONE
000001	2008	3828	3529	4021	1543	0.00479	0.44536
000001	2009	4111	3897	4467	1607	0.00321	0.47086
000001	2010	4535	4064	4896	1584	0.00683	0.51111
000001	2011	5772	4817	6334	1944	0.01108	0.53032
000001	2012	5216	4963	6116	1999	0.00307	0.50733
000001	2013	4069	3365	4752	1311	0.01230	0.56754

4. 数据库内容

数据模块	数据表	数据字段
年报文本语气	年报文本语气	股票代码、会计年度、年报字数、年报词汇数、年报句子数、LM 词典积极词汇数、LM 词典消极词汇数、台大词典积极词汇数、台大词典消极词汇数、基于 LM 词典的 TONE1、基于 LM 词典的 TONE2、基于台大词典的 TONE

5. 数据范围

数据模块	数据起始时间
年报文本语气	2007 年

(二) 管理层讨论与分析数据库-CMDA

1.数据库简介

管理层讨论与分析数据库(*Management's Discussion and Analysis Database, CMDA*)是针对上市公司财务报告中管理层讨论与分析模块而建立的专业数据库, 数据库涵盖了从 1999 年至今所有上市公司的年报、半年报和季报中与经营讨论与分析相关的数据资料, 并采用人工智能算法对文本内容涉及的正面、负面词汇进行判断识别。其中, 正面词汇和负面词汇的判断基础主要是根据 *Tim Loughran and Bill McDonald (2011)* 的英文正负面词典翻译为中文词典, 再结合中文语境进行扩充和完善得到。*CMDA* 可以有效帮助使用者分析上市公司所披露和预期的经营文本数据。

2.数据库特色

- 全面的管理层讨论与分析内容

CMDA 包括上市公司发布的季报、半年报和年报中的管理层讨论与分析内容, 可以充分满足研究的需求。

- 精确的正负面词汇统计

CMDA 基于 *Tim Loughran and Bill McDonald* 词典, 结合中文语境设定的正负面词汇字典, 既保证了词性判断的权威性, 又符合了中文的特殊语法结构。

3.典型数据举例

管理层经营讨论和分析

股票代码	会计年度	经营分析时间	经营讨论与分析内容	正面词汇数量	负面词汇数量	句子数量
002160	2016	2016-12-31	公司近 3 年(包括本报告期)的普通股股利分配方案(预案)……	18	3	6
300401	2016	2016-12-31	2016 年度,对公司来说是很不平凡的一年……	27	9	6
300399	2016	2016-03-31	报告期内,行业移动信息服务及保险产品……	38	3	7
603718	2015	2015-06-30	报告期内总体经营情况公司主营业务为动物疫苗的研发……	98	31	29

4.数据库内容

数据模块	数据表	数据字段
管理层经营讨论和分析	管理层经营讨论和分析	股票代码、公司简称、会计年度、经营分析时间、经营讨论与分析内容、正面词汇数量、负面词汇数量、句子数量、文字数量

5.数据范围

数据模块	数据起始时间
管理层经营讨论和分析	1999 年

（三）上市公司业绩说明会数据库-ECCD

1.数据库简介

上市公司业绩说明会数据库（*Earnings Communication Conferences Database of Chinese Listed Companies, ECCD*）是基于上市公司每年在交易所发布的业绩说明会上的提问与回答内容的文本信息进行整理和分析生成的专业文本分析数据库。**ECCD** 涵盖了从 2008 年至今的业绩说明会情况，采用计算机技术对提问回复内容涉及的正负面词汇以及语调进行识别、判断；其中，文本内容的正面词汇和负面词汇的判断基础主要是根据 *Tim Loughran and Bill McDonald* (2011) 的英文正负面词典翻译的中文词典结合中文语境进行扩充和完善得到。

2.数据库特色

- 精确的正负面词汇统计

ECCD 的词性判断基于 *Tim Loughran and Bill McDonald* 词典，并结合中文语境设定的正负面词汇字典，既保证了词性判断的权威性，又符合了中文的特殊语法结构。

3.典型数据举例

业绩说明会问答文本分析

股票代码	会计年度	提问内容	提问时间	回答正面词汇数比例	回答负面词汇数比例	回答净正面语调
000002	2013	公司 13 年乃至今后几年最大的挑战是什么？	2013-08-09	0.0417	0.2917	-0.7500
000002	2013	贵公司相对其它同行业的公司优越在那里？	2013-08-09	0.1271	0.0508	0.4286
000002	2013	在管理经营中，董事长最关注的问题是什么？	2013-08-09	0.0345	0.1724	-0.6667
000002	2013	请问 5 月份和半年业绩预报可能在 6 月 13 日前出吗？	2014-05-29	0.1765	0.0588	0.5000

4.数据库内容

数据模块	数据表	数据字段
业绩说明会概况	业绩说明会概况	股票代码、公司简称、会计年度、标题、举办时间、嘉宾姓名、嘉宾职务、嘉宾简历
业绩说明会问答文本分析	业绩说明会问答文本分析	股票代码、会计年度、问题序号、提问内容、提问时间、回答人、回答时间、回答内容、正面词汇数量、负面词汇数量、正面词汇数比例、负面词汇数比例、净正面语调

5.数据范围

数据模块	数据起始时间
业绩说明会概况	2016 年
业绩说明会问答文本分析	2008 年

（四）中国上市公司网上路演数据库-CNRD

1.数据库简介

中国上市公司网上路演数据库（*Net Roadshow Database of Chinese Listed Companies, CNRD*）提供了我国上市公司在 *IPO* 过程中的网上路演信息。这些信息主要来自于在主板、中小板和创业板上市的公司，在证券交易所指定的网站披露的路演信息。**CNRD** 包含以下三个数据模块，分别是公司基本信息、参与嘉宾信息以及路演问答信息。其中，公司基本信息包括股票代码、公司简称、所在板块、路演时间、公司简介、投资要点和风险提示等字段；嘉宾信息包括嘉宾姓名、嘉宾所在单位及任职以及其个人简介等字段；问答信息则包括投资者对上市公司的提问、回答的嘉宾以及回答内容等字段。

2.数据库特色

- 数据全面

涵盖上市公司网上路演过程中详尽的嘉宾信息和问答信息，数据字段全面、来源权威。

3.典型数据举例

参与嘉宾信息

股票代码	公司简称	所在板块	嘉宾姓名	嘉宾所在单位名称	嘉宾所在单位任职
002001	新和成	中小板	胡柏藩	浙江新和成股份有限公司	董事长
002001	新和成	中小板	张方治	浙江新和成股份有限公司	总经理
300100	双林股份	创业板	邬建斌	双林股份	董事长
300100	双林股份	创业板	曹靖	双林股份	财务总监
600908	无锡银行	沪市主板	任晓平	无锡农村商业银行股份有限公司	董事长
600908	无锡银行	沪市主板	邵辉	无锡农村商业银行股份有限公司	行长

4.数据库内容

数据模块	数据表	数据字段
公司基本信息	公司基本信息	股票代码、公司简称、所在板块、路演时间、公司简介、投资要点、风险提示
参与嘉宾信息	参与嘉宾信息	股票代码、公司简称、所在板块、嘉宾姓名、嘉宾所在单位、嘉宾所在单位任职、嘉宾个人简介
路演问答信息	路演问答信息	股票代码、公司简称、所在板块、问题、回答、回答人

5.数据范围

数据模块	数据起始时间
全部模块	2004 年

（五）中国上市公司澄清公告数据库-CCAD

1.数据库简介

中国上市公司澄清公告数据库（*Clarification Announcement Database of Chinese Listed Company, CCAD*）考察了上市公司针对市场传闻等信息发布的澄清公告，并将其转化为文本格式，提取其中的关键字段，从而建立的专业数据库。*CCAD* 包含了诸多重要字段，其中，传言日期除了提取自公告内容外，我们还根据传闻内容，通过网络搜索等找到相应的传言开始日期；传言事项则根据传言内容所述的事项，对这些事项的一个科学分类；传言类型则根据是否负面传言进行分类；传言媒体和传言标题均提取自公告内容中。

2.数据库特色

- 准确的文本信息，丰富的字段内容

CCAD 精确的将 *PDF* 格式的公告内容转为文本信息，学者可从中挖掘研究信息。此外，对澄清公告的重要字段进行了提取和整合，尤其是传言日期均由手工收集整理，可以体现传言的开始日期。

3.典型数据举例

上市公司澄清公告

股票代码	公司简称	发布日期	传言日期	传言类型	传播媒体
600535	天士力	2017/3/8	2017/3/17	1	中国证券报
600810	神马股份	2017/2/17	2017/2/17	-1	大众证券报、证券时报
600129	太极集团	2017/2/8	2016/12/17	-1	中国经营报
600275	武昌鱼	2017/1/25	2017/1/24	-1	东方财富网股吧
002560	通达股份	2017/1/25	2017/1/23	1	中国证券网

4.数据库内容

数据模块	数据表	数据字段
上市公司澄清公告	上市公司澄清公告	股票代码、公司简称、发布日期、公告内容、传言日期、传言类型、传播媒体、传言标题

5.数据范围

数据模块	数据起始时间
上市公司澄清公告	2001年

资本市场人物特征

(一) 中国地方政府官员数据库-CGOD

1. 数据库简介

中国地方政府官员数据库 (*Chinese Local Government Officials Database, CGOD*) 是有关中国地方政府官员信息的专业学术数据库, 其中包括地方政府官员的基本信息、教育背景与工作经历等信息。*CGOD* 包含省级政府官员信息、地市级政府官员信息与县级政府官员信息三个子模块, 该模块有以下几个特点: 省级政府官员信息包含自建国以来所有历任的省委书记与省长信息; 地市级政府官员信息包含自建市以来几乎所有历任市委书记与市长的信息; 县级政府官员信息包含建国以来几乎所有历任县委书记与县长的信息。该数据库的所有数据均来源于公开渠道。

2. 数据库特色

- 丰富的省市级官员信息

包括详尽的省级和市级官员教育背景、工作经历等信息, 数据量和字段大大高于市场上的其他数据库。

- 唯一提供县级官员信息数据库

包括详尽的县级官员信息, 是目前唯一提供县级官员信息的数据库。

3. 典型数据举例

(1) 省级政府官员基本信息

省级名称	职位	官员姓名	任职期间	性别	籍贯	出生日期	入党时间
安徽省	省长	李斌	2011.12-至今	女	辽宁抚顺	1954.1	1981.06
安徽省	省长	王金山	2003.01-2007.12	男	吉林公主岭	1945.02	1971.01
安徽省	省长	王金山	2002.10-2003.01	男	吉林公主岭	1945.02	1971.01
安徽省	省长	王三运	2008.01-2011.12	男	山东单县	1952.12	1979.07
安徽省	省委书记	郭金龙	2004.12-2007.12	男	江苏南京	1947.07	1979.04
安徽省	省长	许仲林	2000.01-2002.10	男	江苏常州市	1943.12	1964.11

(2) 地市级政府官员教育背景

官员姓名	官员代码	学历/学位	毕业学校	专业	教育起始时间	教育终止时间
艾文礼	23	本科	中央党校	经济管理	1993.08	1995.12
艾文礼	23	研究生	省委党校	经济管理	1999.09	2001.12
艾学峰	24	本科	东北工学院	计算机科学	1983.09	1987.07
艾学峰	24	研究生	中国人民银行	货币银行学	1989.09	1992.02

(3) 县级政府官员工作经历

官员姓名	官员代码	任期起始时间	任期结束时间	职位名称
陈贵	331	1987.06	1991.07	共青团河北省张家口市地委书记、党组书记
陈贵	331	1991.07	1992.08	河北怀安县委副书记 (正县级)
陈贵	331	1992.08	1993.02	中共河北省怀安县委副书记、代县长

官员姓名	官员代码	任期起始时间	任期结束时间	职位名称
陈贵	331	1993.02	1996.02	中共河北省怀安县委副书记、县长
陈贵	331	1996.02	1997.12	中共河北省怀安县委书记
陈贵	331	1997.12	1998.11	河北省张家口市副市长

4.数据库内容

数据模块	数据表	数据字段
省级政府 官员信息	基本信息	区域代码、省级名称、职位、职位代码、职位备注、官员姓名、任职期间、官员代码、性别、民族、籍贯、出生省、出生市、出生县、出生日期、入党时间、最高学历
	教育背景	官员姓名、官员代码、学历/学位、毕业学校、专业、教育起始时间、教育终止时间
	工作经历	官员姓名、官员代码、任期起始时间、任期结束时间、职位名称
地市级政府 官员信息	基本信息	区域代码、省级名称、地市级名称、职位、职位代码、职位备注、官员姓名、任期期间、官员代码、性别、民族、籍贯、出生省、出生市、出生县、出生日期、入党时间、最高学历
	教育背景	官员姓名、官员代码、学历/学位、毕业学校、专业、教育起始时间、教育终止时间
	工作经历	官员姓名、官员代码、任期起始时间、任期结束时间、职位名称
县级政府 官员信息	基本信息	区域代码、省级名称、地市级名称、县级名称、职位、职位代码、职位备注、官员姓名、任期期间、官员代码、性别、民族、籍贯、出生省、出生市、出生县、出生日期、入党时间、最高学历
	教育背景	官员姓名、官员代码、学历/学位、毕业学校、专业、教育起始时间、教育终止时间
	工作经历	官员姓名、官员代码、任期起始时间、任期结束时间、职位名称

5.数据范围

数据模块	数据起始时间
省级政府官员信息	1949 年
地市级政府官员信息	1949 年
县级政府官员信息	1949 年

（二）董事长与总经理研究数据库-CCEO

1.数据库简介

董事长与总经理研究数据库 (*Chairman and CEO Research Database, CCEO*) 是基于对中国上市公司年报、招股说明书、公告等多方来源数据的考察，并通过网络搜索、计算机提取和人工核对等多种方式补充完善个人特征相关信息而建立的专业数据库。*CCEO* 主要包括董事长与总经理基本信息、教育背景、任职信息、政府任职和从军背景五张表，详细列示了上市公司董事长与总经理的姓名、年龄、性别、简历、就读高校、所学专业、离任职信息、薪酬情况、工作经验、政府任职、从军经历等信息。

2.数据库特色

- 多维

CCEO 包括上市公司高管的多维度特征数据，包括年龄、性别、学历、专业、政府任职、海外经历、当前任职情况等多维度信息。

- 详细

CCEO 对高管的毕业院校的分类详细到其所就读的大专院校、本科院校、硕士院校、博士院校、MBA 院校、博士后院校等；对高管专业背景的描述则包括其大学所学专业、是否拥有海外教育背景等。这些字段极其详细地归纳了各个方面的高管特征。

- 数据丰富

CCEO 包含的籍贯/出生地、职称等字段的非空值远远多于市场上的已有数据库，可为学者研究提供便利。

3.典型数据举例

(1) 上市公司高管基本信息

股票代码	姓名	职务	性别	出生日期	国籍	籍贯或出生地
000997	林学杰	总经理	男	1960	中国	福建省福州市
002028	杨小强	总经理	男	1963	中国	北京市海淀区
002123	左强	董事长	男	1970	中国	北京市海淀区
002124	吴天星	董事长	男	1963	中国	北京市海淀区
002151	赵耀升	总经理	男	1959	中国	北京市海淀区
002208	张金生	总经理	男	1963	中国	北京市宣武区

4.数据库内容

数据模块	数据表	数据字段
上市公司高管特征	基本信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、性别、出生日期、国籍、籍贯、籍贯所在地区编码、出生地、出生地所在地区编码、籍贯或出生地、籍贯或出生地所在地区编码、学历、职称与资格证书、高管简历、是否兼任董事长和 CEO、是否兼任董事长和总经理、海外背景、学术背景、是否在其他上市公司兼职、兼职的公司、兼职的职务

数据模块	数据表	数据字段
上市公司高管特征	教育背景	股票代码、公司简称、统计日期、高管编码、高管姓名、所学专业、是否有海外教育背景、最高学历/学位、专科院校、专科院校代码、本科院校、本科院校代码、硕士院校、硕士院校代码、博士院校、博士院校代码、博士后院校、博士后院校代码、MBA 院校名称、MBA 院校代码、访学院校名称、访学院校代码
	任职信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、是否在职、任职开始日期、任职结束日期、报告期薪酬、年末持股数量、所属部门、是否政协委员或人大代表
	政府任职	股票代码、公司简称、统计日期、高管编码、高管姓名、任职政府单位、政治职务
	从军背景	股票代码、公司简称、统计日期、董事长与总经理编码、董事长与总经理姓名、任职军事单位、军事职务

5.数据范围

数据模块	数据起始时间
全部模块	1992 年

(三) 上市公司其他高管数据库- OECD

1.数据库简介

上市公司其他高管数据库 (*Other Executives Database of Listed Companies, OECD*) 是基于对中国上市公司除董事长、总经理、CEO 之外的其他高管简历、招股说明书、公告等来源披露的其他高管信息, 从多个维度进行其特征信息的整合和展示的专业数据库。*OECD* 主要包括上市公司其他高管基本信息、教育背景、任职信息和政府任职四张表。

2.数据库特色

- 全面

OECD 不仅包括上市公司其他高管个人基本信息, 还包括教育背景、政府任职等重要信息, 涵盖多个重要字段, 例如籍贯、出生地、所学专业、毕业院校、报告期薪酬、年末持股数量等。

3.典型数据举例

上市公司其他高管教育背景

股票代码	统计日期	姓名	本科院校	院校代码	硕士院校	院校代码	博士院校	院校代码
000608	2012/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008
000608	2013/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008
000608	2014/12/31	李鸣	北京大学	01-008	耶鲁大学	FA-032	北京大学	01-008

4.数据库内容

数据模块	数据表	数据字段
上市公司 高管特征	基本信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、性别、出生日期、国籍、籍贯、籍贯所在地区编码、出生地、出生地所在地区编码、籍贯或出生地、籍贯或出生地所在地区编码、学历、职称与资格证书、高管简历、海外背景、学术背景、是否在其他上市公司兼职、兼职的公司、兼职的职务
	教育背景	股票代码、公司简称、统计日期、高管编码、高管姓名、所学专业、是否有海外教育背景、最高学历/学位、专科院校、专科院校代码、本科院校、本科院校代码、硕士院校、硕士院校代码、博士院校、博士院校代码、博士后院校、博士后院校代码、MBA 院校名称、MBA 院校代码、访学院校名称、访学院校代码
	任职信息	股票代码、公司简称、统计日期、高管编码、高管姓名、高管职务、高管职务编码、是否在职、任职开始日期、任职结束日期、报告期薪酬、年末持股数量、所属部门、是否政协委员或人大代表
	政府任职	股票代码、公司简称、统计日期、高管编码、高管姓名、任职政府单位、政治职务

5.数据范围

数据模块	数据起始时间
全部模块	2010 年

（四）独立董事研究数据库-IDRD

1.数据库简介

独立董事研究数据库 (*Independent Directors Research Database, IDRD*) 基于对独立董事的简历、新闻报道、公司公告、监管机构公告等多方数据来源的考察,从多个维度对独立董事信息进行展示的专业数据库。独立董事研究从本世纪初就成为学术研究前沿与热点,并逐渐成为公司治理领域极为重要的一个研究方向,然而,目前市场上罕有单独成库的独立董事研究相关的数据,并且所涉及的数据深度和广度也远远不够。因此, *IDRD* 对该领域的研究数据进行了补充和扩展,包括独董工作经历、独董新闻报道、独董意见信息、独董处罚情况、独董辞任职、独董比例及独董交叉任职七张表格。

2.数据库特色

- 多维度独立董事动态信息

IDRD 从多个权威渠道获得了多维度独董动态信息,例如,新闻报道、处罚、辞任职信息等,这为独立董事研究的扩展提供了重要数据支持。

3.典型数据举例

独董任辞职

股票代码	会计年度	独董编码	独董姓名	任职开始日期	任职结束日期	离职原因
000505	2016	1313272	许来正	2016-09-13	2016-12-24	个人工作调动
000546	2016	1323162	赵燕	2016-04-30	2017-06-29	换届
000625	2016	1214869	帅天龙	2016-03-08	2016-06-24	去世
000633	2016	1313549	姜杰	2016-04-26	2017-06-09	工作变动原因

4.数据库内容

数据模块	数据表	数据字段
独立董事动态	独董工作经历	股票代码、公司简称、统计日期、独董编码、独董姓名、职务、工作单位、任职开始时间、任职结束时间、工作单位类型
	独董新闻报道	股票代码、公司简称、独董姓名、新闻类型、新闻来源、新闻日期、新闻标题、新闻内容、相似新闻数量、正负面评价
	独董意见信息	股票代码、公司简称、公告日期、独董姓名、发表内容、意见类别
	独董处罚情况	股票代码、公司简称、处罚日期、独董姓名、处分措施、违规行为、违规类型、实施处罚单位、处分类型
	独董任辞职	股票代码、公司简称、会计年度、独董编码、独董姓名、任职开始日期、任职结束日期、离职原因、是否正常离职、离职公告内容
	独董比例	股票代码、会计年度、董事总人数、独立董事人数、独立董事比例 (%)
	独董交叉任职	股票代码、公司简称、统计日期、独董编码、独董姓名、职务、任职开始日期、任职结束日期、是否交叉任职

5.数据范围

数据模块	数据起始时间
全部模块	1992 年

(五) 董秘信息数据库-CSBD

1.数据库简介

董秘信息数据库 (*Secretary of the Board Database, CSBD*) 是基于公开披露的上市公司董秘个人简历数据以及新财富年度金牌董秘评选整理而成的专业数据库。*CSBD* 除了包含上市公司董秘的个人基本信息外, 还包含了兼任情况、海外背景、学术背景、是否为金牌董秘等多维度信息。

2.数据库特色

- 多维

CSBD 包括上市公司董秘的个人多维度特征数据, 覆盖面广, 涵盖年龄、性别、学历、职业背景、职业资格证书、任职情况等丰富信息。

3.典型数据举例

上市公司董秘信息

股票代码	统计日期	高管编码	董秘姓名	是否为金牌董秘	性别	出生日期	学历
000002	2006-12-31	1200129	肖莉	1	女	1963	硕士
000002	2015-12-31	1200152	谭华杰	0	男	1973	本科
000002	2016-12-31	1209171	朱旭	0	女	1975	硕士
000003	2000-12-31	1200178	彭卫	0	男	1969	硕士
000004	2010-12-31	1200218	蒋伟诚	0	男	1970	硕士
000004	2010-12-31	1200256	杨斌	0	男	1972	硕士

4.数据库内容

数据模块	数据表	数据字段
上市公司董秘信息	上市公司董秘信息	股票代码、公司简称、统计日期、高管编码、董秘姓名、董秘的其他职务、是否为金牌董秘、性别、出生日期、国籍、籍贯、籍贯所在地区编码、出生地、出生地所在地区编码、籍贯或出生地、籍贯或出生地所在地区编码、学历、职称与资格证书、董秘简历、海外背景、学术背景

5.数据范围

数据模块	数据起始时间
全部模块	1999

（六）基金经理研究数据库-FMRD

1.数据库简介

基金经理研究数据库（*Fund Manager Research Database, FMRD*）是基于对基金经理的简历、负责的基金等多方数据来源的考察，对基金经理个人信息进行整合和展示的专业数据库。主要包括基金经理基本信息、工作经历、管理基金和教育背景四张表，包含基金经理所在任职机构名称、职位、所持基金情况、学历/学位、毕业院校等重要信息。

2.数据库特色

- 完善、全面的基金经理信息

FMRD 不仅包括了基金经理的个人基本信息，还包括了基金经理的教育背景、工作经历、管理基金等重要信息。

3.典型数据举例

基金经理工作经历

基金经理姓名	基金代码	基金简称	任职机构名称	职位	入职日期	离职日期
费逸	270021	广发聚瑞混合	广发基金管理有限公司	研究员	2010-07-06	2017-07-18
林高榜	110012	易方达科汇灵活配置混合	易方达基金管理有限公司	基金经理助理	2016-12-14	2017-05-10
刘晓龙	270025	广发行业领先混合 A	广发基金管理有限公司	基金经理	2010-11-23	2017-03-29
谈洁颖	519956	长信睿进混合 C	长信基金管理有限责任公司	基金经理	2012-07-23	2017-03-28
徐冶丹	004585	鹏扬汇利债券 A	北京鹏扬投资管理有限公司	基金经理	2017-06-02	2017-07-14

4.数据库内容

数据模块	数据表	数据字段
基金经理	基金经理基本信息	基金经理姓名、性别、出生年份、学历/学位、国籍、简介
	基金经理工作经历	基金经理姓名、基金代码、基金简称、任职机构名称、职位、入职日期、离职日期
	基金经理管理基金	基金经理姓名、基金代码、基金名称、基金公司名称、投资类型、开始日期、结束日期
	基金经理教育背景	基金经理姓名、基金代码、基金简称、学历/学位、专业、毕业学校

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

(七) 金融机构高管数据库-FIED

1.数据库简介

金融机构高管数据库 (*Financial Institution Executives Database, FIED*) 是从银行、券商和信托公司的年报以及官方网站等多方数据来源进行整理和分析后而建立的专业数据库。主要包括券商高管、银行高管、信托公司高管 3 个模块, 共 3 张表, 包含券商、银行和信托公司高管的职务、兼职单位以及简历信息等重要字段。*FIED* 包括几乎所有从 2000 年至今的券商高管、银行高管、信托公司高管个人特征数据, 大大减少了学者搜集和整理数据的时间。

2.数据库特色

- 提供深度的金融机构高管特征数据

FIED 从银行、券商和信托公司的年报以及官方网站等获得了这些金融机构的高管简历信息, 并在此基础上进行深度的整理和分析, 提供研究中常用的重要字段。

3.典型数据举例

银行高管

银行高管姓名	银行代码	银行名称	统计日期	职位	性别	年龄
陈光浩	100029	抚顺银行股份有限公司	2015-12-31	非执行董事	男	48
陈鹏飞	100029	抚顺银行股份有限公司	2015-12-31	监事、副行长	男	33
付蕾	100029	抚顺银行股份有限公司	2015-12-31	副行长	女	46
李向文	100029	抚顺银行股份有限公司	2015-12-31	非执行董事	男	41

4.数据库内容

数据模块	数据表	数据字段
金融机构高管	券商高管	统计日期、券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、职位、是否全职
	银行高管	银行高管姓名、银行代码、银行名称、统计日期、职位、性别、年龄、出生日期、任职期间、持有股份、是否领取薪酬、兼职单位的名称、兼职单位的职务、任职的股东单位名称、在股东单位的职务、简历
	信托公司高管	信托公司代码、信托公司名称、统计日期、信托公司高管姓名、性别、年龄、学历/学位、专业、职位、选任日期、兼职单位名称、兼职单位职务、任职的股东单位名称、任职的股东单位职务、金融从业年限、所推举的股东名称、推举股东的持股比例 (%)、简历

5.数据范围

数据模块	数据起始时间
券商高管	2007 年
银行高管	2000 年
信托公司高管	2004 年

银行及金融研究

(一) 银行研究数据库-CBRD

1.数据库简介

银行研究数据库 (*Chinese Bank Research Database, CBRD*) 是面向广大高校和学者以及实务界人士,对中国商业银行进行深入考察与了解建立的详细、准确和全面的研究数据库。该数据库涵盖了我国国有银行,股份制银行、城市商业银行、农村商业银行等银行。*CBRD* 包含银行基本信息、银行财务报表、银行信贷资产分析、银行重要业务指标等 4 个模块。其中,银行基本信息包括银行概况、股东构成、高管特征、审计信息、员工与股东信息等;银行财务报表模块包括了三张常用报表;银行信贷资产分析对信贷资产结构,十大贷款客户,贷款行业分布以及贷款担保方式进行了统计;银行重要业务指标对一些重要指标进行了罗列。

2.数据库特色

- 全面、准确

CBRD 不仅包括中国目前已经上市的银行机构,还包括大量未上市的银行机构数据,并进行多方信息来源的综合。为确保非上市银行数据的准确性,*CBRD* 进行了逐项、严格的校对工作。

3.典型数据举例

银行股东构成

银行代码	银行名称	统计日期	序号	股东名称	期末股份数 (股)	期末持股比例 (%)
100136	上海银行	2012-12-31	1	上海联和投资有限公司	814803808	19.24
100136	上海银行	2012-12-31	2	香港上海汇丰银行有限公司	338720000	8
100136	上海银行	2012-12-31	3	中国建银投资有限责任公司	296380000	7
100136	上海银行	2012-12-31	4	上海市黄浦区国有资产总公司	130409244	3.08
100136	上海银行	2012-12-31	5	上海商业银行有限公司	127020000	3

4.数据库内容

数据模块	数据表	数据字段
银行基本信息	银行概况	银行代码、银行名称、成立日期、银行性质、注册资本、是否上市、省级名称、注册地址、经营范围、法定代表人、总经理
	股东构成	银行代码、银行名称、统计日期、序号、股东名称、持股金额、期末股份数、期末持股比例 (%)、期初股份数、期初持股比例 (%)、股份质押数、股份冻结数
	高管特征	银行高管姓名、银行代码、银行名称、统计日期、职位、性别、年龄、出生日期、任职期间、持有股份、是否领取薪酬、兼职单位的名称、兼职单位的职务、任职的股东单位名称、在股东单位的职务、简历
	审计信息	银行代码、银行名称、统计日期、会计师事务所名称、审计日期、审计师姓名、审计费用

数据模块	数据表	数据字段
银行基本信息	员工与股东信息	银行代码、银行名称、统计日期、股东户数、机构股东数、自然人股东数、员工总数、高级职称员工数、高级职称员工占比(%)、中级职称员工数、中级职称员工占比(%)、初级职称员工数、初级职称员工占比(%)、中层以上管理人员数量、银行业务人员数量、博士研究生学历人数、硕士研究生学历人数、研究生学历占比(%)、大学本科学历人数、大学本科学历占比(%)、大专及以上学历人数、大专及以上学历占比(%)、大专及以下学历人数、大专及以下学历占比(%)、退养员工人数
银行财务报表	资产负债表	银行代码、银行名称、统计日期、现金及存放中央银行款项、存放同业款项、拆出资金、交易性金融资产、衍生金融资产、应收利息、发放贷款及垫款、可供出售金融资产、固定资产、资产总计、同业和其他金融机构存放款项、向中央银行借款、拆入资金、吸收存款、应交税费、应付利息、负债合计、实收资本、资本公积金、盈余公积金、未分配利润、一般风险准备、归属于母公司所有者权益合计、所有者权益合计、负债与所有者权益总计
	利润表	银行代码、银行名称、统计日期、营业收入、利息净收入、利息收入、利息支出、手续费及佣金净收入、手续费及佣金收入、手续费及佣金支出、营业支出、营业税金及附加、管理费用、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、归属于母公司所有者的净利润、少数股东损益、基本每股收益、稀释每股收益
	现金流量表	银行代码、银行名称、统计日期、客户存款和同业存放款项净增加额、向中央银行借款净增加额、向其他金融机构拆入资金净增加额、收取利息和手续费净增加额、收到其他与经营活动有关的现金、经营活动现金流入小计、客户贷款及垫款净增加额、经营活动产生的现金流量净额、投资活动产生的现金流量净额、筹资活动产生的现金流量净额
银行信贷资产分析	信贷资产结构	银行代码、银行名称、统计日期、正常类贷款、关注类贷款、不良类贷款
	十大贷款客户	银行代码、银行名称、统计日期、序号、客户名称、贷款余额、占贷款总额比例(%)、占资本净额比例(%)
	贷款行业分布	银行代码、银行名称、统计日期、序号、行业种类、贷款余额、占贷款总额比例(%)
	贷款担保方式	银行代码、银行名称、统计日期、质押贷款余额、抵押贷款余额、保证贷款余额、信用贷款余额
银行重要业务指标	银行重要业务指标	银行代码、银行名称、统计日期、存款总额、贷款总额、资本充足率(%)、核心资本充足率(%)、拆入资金比例(%)、拆出资金比例(%)、单一最大客户贷款比例(%)、最大十家客户贷款比例(%)、利息回收率(%)、逾期贷款、贷款呆帐准备金、不良贷款余额、不良贷款比例、不良贷款拨备覆盖率、资本净额、核心资本净额、成本收入比(%)

5.数据范围

数据模块	数据起始时间
银行基本信息	1992年
银行财务报表	2004年
银行信贷资产分析	2000年
银行重要业务指标	2003年

（二）中国商业银行分支机构数据库-CCBD

1.数据库简介

中国商业银行分支机构数据库（*Branch Database of Chinese Commercial Bank, CCBD*）分为两个数据模块，银行分支机构信息模块的数据主要来自中国银行业监督管理委员会对银行机构颁（换）发许可证的相关信息，包括机构设立情况、机构失控情况和机构退出情况三张表格。银行网点分布信息模块的数据主要来自高德地图以计算机技术获得的银行网点分布情况。由于银行网点分布信息除了包括银行分支机构信息，还包括诸如银行营业厅、ATM取款机等多种类型的银行服务机构，因此其数量要远多于银行分支机构信息，可为学术研究提供更多支持。

2.数据库特色

- 全面的银行分支机构和网点信息

CCBD 既包括官方批准的银行分支机构信息，又包括来自各个省市县的银行网点分布情况，可以充分反映一个地区的银行竞争、银行服务等情况。

3.典型数据举例

银行分支机构信息_机构设立情况

机构名称	机构地址	机构所在地	许可证发证日期	批准成立日期
中国农业发展银行 山西省分行	山西省太原市康乐街 38 号	山西省-太原市	2017-04-06	1995-04-25
中国农业发展银行 北京市分行	北京市丰台区南四环西路 186 号一区 1 号楼 5 层	北京市	2016-11-22	1995-05-19
中国农业发展银行 黑龙江省分行	黑龙江省哈尔滨市道里区 尚志大街 141 号	黑龙江省-哈尔 滨市	2007-05-29	1995-06-29

4.数据库内容

数据模块	数据表	数据字段
银行分支机 构信息	机构设立情况	机构编码、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期
	机构失控情况	机构编码、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期、失控原因、失控日期
	机构退出情况	机构编码、机构名称、机构地址、机构所在地、金融许可证发证日期、批准成立日期、退出日期
银行网点分 布信息	银行网点分布	省份名称、市县名称、银行网点名称、银行网点地址、获取时间

5.数据范围

数据模块	数据起始时间
银行分支机构信息	2007 年
银行网点分布信息	截面数据

(三) 中国风险投资和私募股权数据库-VCPE

1. 数据库简介

中国风险投资和私募股权数据库 (*Chinese Venture Capital and Private Equity Database, VCPE*) 是以研究为目的而设计的, 面向国内外从事风险和私募股权投资研究的学者, 旨在提供完整、精确的专业学术研究数据。该数据库全面覆盖股权投资研究所需要的各种信息, 不仅收录了基金信息、有限合伙人信息、被投企业信息、投资机构信息, 还将这些机构及企业的管理人员的信息进行汇总单独成库, 同时还涵盖了投资事件及退出事件的各个层面信息。本数据库包含 7 个数据模块, 即私募风险投资、私募风险基金 LP、私募风险基金 GP、被投资企业、私募风险基金管理层、投资事件、退出事件。

2. 数据库特色

- 内容详细完善

VCPE 对各个维度的信息进行了详细描述, 提供了中国股权投资市场情报、投资机构及企业、股权交易、财务数据等各个层面信息。例如, 被投资企业数据库中不仅包括被投资企业股权交易基本信息, 还包括选自财务报表中的重要财务指标数据。

- 学术性强

考虑到本数据库的学术导向, 我们在设计数据库时综合考虑了学者对风险投资和私募股权研究的需要, 并有针对性地对相关模块的字段设计进行了调整。

3. 典型数据举例

(1) 基金基本信息

基金名称	基金类型	资本类型	组织形式	是否开放投资	募集完成时间	募集规模	开始投资年度
IDG 技术创投基金	Venture	外资	有限合伙制	是	2000-06-27	2.00 亿美元	1997
腾讯产业共赢基金	Growth	中资	有限合伙制	是	2011-06-01	100.00 亿元	2011
新浪微博基金	Venture	合资	有限合伙制	是	2010-11-16	2.00 亿元	2011
启明三期美元基金	Venture	外资	有限合伙制	是	2011-05-19	4.50 亿美元	2011
纪源三期	Growth	外资	有限合伙制	否	2006-01-01	6.10 亿美元	2006

(2) 退出基本信息

退出事件名称	退出机构	退出时间	退出方式	总投资额
尚融成长退出超讯通信	尚融成长	2016/7/28	IPO	1200 万元
久德长盛退出超讯通信	久德长盛	2016/7/28	IPO	1500 万元
粤科金融退出华锋股份	粤科金融	2016/7/26	IPO	2121 万元

4. 数据库内容

数据模块	数据表	数据字段
私募风险投资	基金基本信息	基金名称、管理公司、基金类型、资本类型、组织形式、是否开放投资、基金所在地、募集状态、募集完成时间、募集规模、募集规模单位、投资币种、开始投资年度、基金描述

数据模块	数据表	数据字段
私募风险投资	基金募集信息	基金名称、开始募集时间、开始募集目标规模、开始募集目标规模单位、开始募集描述、募集完成时间、募集完成规模、募集完成规模单位、有限合伙人简称、LP 类别、承诺金额、承诺金额单位、投资金额、投资金额单位
	基金投资策略	基金名称、拟投资地域、拟投资中国比例(%)、投资上限、投资下限、拟投资阶段、拟投资行业、拟投资国标行业、策略描述
	基金投资分析	基金名称、基金募集规模、基金募集规模单位、开始投资年度、投资数量、累计投资金额、上市数量、并购数量、停业数量、企业简称、行业分类、地区、投资人、投资时间、投资性质、投资金额、投资金额单位
	基金退出分析	基金名称、退出企业、退出时间、退出方式、账面退出回报、账面退出回报单位、账面回报率(倍数)、累计投资金额、首次投资时间
私募风险基金 LP	LP 基本信息	LP 名称、总部所在地、成立时间、机构类型、资本类型、管理资金、管理资金单位、是否开放投资、机构描述
	LP 管理团队	LP 名称、管理人员姓名、管理人员职位
	LP 投资基金	LP 名称、基金名称、基金类型、资本类型、募集状态、募集完成时间、募集完成规模、募集完成规模单位、承诺金额、承诺金额单位
私募风险基金 GP	GP 基本信息	GP 名称、成立日期、资本类型、组织形式、所属行业、投资机构类型、管理资本量、管理资本量单位、注册地区、公司总部、是否备案、描述
	GP 管理团队	GP 名称、管理人员姓名、管理人员职位、管理人员状态
	GP 投资基金	GP 名称、基金名称、基金类型、募集状态、资本类型、募集完成时间、募集金额、募集金额单位、投资数量
被投资企业	被投资企业基本信息	融资企业名称、成立时间、商业状态、注册登记类型、员工人数、企业注册地、法定代表人、组织机构代码、相关产品、证监会行业分类、企业描述
	被投资企业融资情况	融资企业名称、融资时间、融资类型、融资金额、融资金额单位、股权(%)、投资机构、投资人、投资金额、投资金额单位
	被投资企业退出情况	融资企业名称、退出机构、退出基金、退出时间、退出方式、账面退出回报、账面退出回报单位、账面回报率(倍数)、累计投资金额、首次投资时间
	资产负债表	融资企业名称、统计日期、单位、应收账款、存货、流动资产合计、资产总计、应付账款、应交税费、流动负债合计、非流动负债合计、负债合计、所有者权益(股东权益)合计、负债和所有者权益(或股东权益)总计
	利润表	融资企业名称、统计日期、营业总收入、营业收入、利息收入、营业总成本、营业成本、利息支出、营业税金及附加、销售费用、管理费用、财务费用、投资收益、补贴收入、营业利润、营业外收入、利润总额、所得税费用
	现金流量表	融资企业名称、统计日期、单位、经营活动现金流入小计、经营活动现金流出小计、经营活动产生的现金流量净额、投资活动现金流入小计、投资活动现金流出小计、投资活动产生的现金流量净额、筹资活动现金流入小计、筹资活动现金流出小计、筹资活动产生的现金流量净额
私募风险基金管理层	管理层基本信息	管理人员姓名、管理人员性别、人物类别、管理人员描述
	管理层职业生涯	管理人员姓名、开始时间、结束时间、公司名称、职位
投资事件	投资方情况	投资事件名称、投资机构、基金名称、投资人、投资时间、投资金额、投资

数据模块	数据表	数据字段
投资事件	投资方情况	金额单位、股权(%)、投资类型
	融资方情况	投资事件名称、被投资企业、所在地区、所属行业、投资时间、融资金额、融资金额单位、投资阶段、轮次、事件描述
退出事件	退出基本信息	退出事件名称、退出企业、股票代码、交易所、企业所属行业、企业所在地区、退出机构、退出时间、退出方式、总投资额、总投资额单位、事件描述
	退出回报	退出事件名称、退出企业、退出机构、账面退出金额、账面退出金额单位、账面回报率(倍数)、账面 IRR

5.数据范围

数据模块	数据起始时间
全部模块	1990 年

（四）中国金融理财研究数据库-CFSD

1.数据库简介

中国金融理财研究数据库（*Chinese Financial Services Research Database, CFSD*）是以研究为目的而设计的专业学术研究数据库，旨在为国内外从事金融理财产品研究的学者和机构从业人员提供最精确、最完整的研究数据。本数据库覆盖中国投资者主要关注的理财产品类型，如银行理财产品、保险理财产品、阳光私募基金产品和 P2P 网贷。本数据库不仅包含相关产品的基本信息，还综合汇总了产品的具体信息。具体地，本数据库包含：

低风险理财产品 分为 2 个子模块：银行理财、保险理财。该数据库汇总了低风险理财产品的基本信息。

阳光私募基金产品 分为 3 个子模块：阳光私募基金产品信息、阳光私募基金经理信息、阳光私募基金收益率信息。该数据库包含所有私募基金的评级以及私募基金经理擅长的投资策略，内容非常完整且全面。

P2P 理财产品 分为 4 个子模块：P2P 网贷平台基本信息、P2P 网贷平台所属公司信息、P2P 问题平台信息、P2P 网贷平台高管基本信息。该数据库为投资者和研究人员提供详尽的 P2P 网贷平台信息。

2.数据库特色

- 数据完整性

中国金融理财研究数据库囊括了国内投资者与研究者关注的几乎所有的理财产品与项目，不仅包括产品本身的各方面详细信息，还包括产品的投资人、委托管理人、借贷人等多方位的综合信息。比如，P2P 网贷平台模块包含投资人与借款人的信用等级、各基金经理擅长的基金管理策略等信息，这为投资者和研究者全面掌握和了解各理财产品状况提供了参考。

- 数据唯一性

中国金融理财研究数据库是国内唯一专注于金融理财产品研究的学术数据库，考虑到学者对金融理财产品研究的需要，该数据库对不同的理财产品进行分类整理，将各类理财产品单独成库，各模块条理清晰。

3.典型数据举例

（1）P2P 网贷平台基本信息

平台名称	平均收益	投资期限	注册资金	债权转让	投标保障	业务类型
懒虫网	18%	1 月标 (100.0%)	1000 万元	不可转让	本金保障	车贷 (99%)
天天赢	16%	1 月标 (60.0%)、 天标 (40.0%)	3000 万元	1 个月	100%保 本保息	车贷 (30%)、银 行过桥 (70%)
宏量财富	16%	1 月标 (100.0%)	2000 万元	不可转让	VIP 保本 保息	房贷 (100%)
龙商 e 贷	10%	2 月标 (35.0%)、1 月标 (25.0%)	5000 万元	随时	风险保证 金制度	车贷 (35%)、房 贷 (35%)、供应 链 (30%)

(2) 阳光私募基金经理

姓名	学历/学位	从业年限	任职私募年限	任职日期	履历背景	基金公司	管理基金数	擅长策略
何翰宏	学士	6年	1年251天	2014-12-23	券商	金色木棉投资	42只	其他
范赞车	大专	20年	1年103天	2015-08-21	民间	象学资产	2只	股票型
侯玉成	学士	3年	4年84天	2012-06-08	媒体	向量多维资本	4只	市场中性
云峰	硕士	8年	126天	2016-04-26	期货	泰通资管	1只	管理期货
何淼	硕士	10年	266天	2015-12-08	公募	相聚资本	1只	其他
海乐	硕士	21年	2年11天	2014-08-20	券商	银领资产	19只	定向增发

4. 数据内容

数据模块	数据表	数据字段
低风险理财产品	银行理财产品基本信息	产品名称、银行性质、银行名称、收益类型、委托币种、起始金额、委托单位递增、预期收益率、与同期储蓄比、销售状态、风险评级、流动性评级、发行起始日、发行截止日、收益计起日、到期日、委托管理期限、是否能被质押、是否有权提前赎回、本息返还方式、产品管理费、投资标的、收益率说明、收益计算基础
	保险理财	产品名称、保险类别、投保年龄、发行机构、保障期限、主附险、缴费期限、保障范围、保险责任、产品特点
阳光私募基金产品	阳光私募基金产品信息	基金简称、基金类型、综合评级、起购资金、封闭期、浮动管理利率、基金公司、成立以来收益率、组织形式、单位净值、成立日期、结构形式、基金状态、托管银行、开放日期、备案号
	阳光私募基金经理	阳光私募基金经理姓名、学历/学位、从业年限、任职私募年限、任职日期、履历背景、基金公司、管理基金数、擅长策略、基金简称、经理介绍、公司介绍
	阳光私募基金收益率	基金名称、基金类型、成立以来收益率、数据更新时间、近一年收益率、近二年收益率、近三年收益率、近1月收益率、近3月收益率、近6月收益率、今年以来收益率
P2P理财产品	P2P网贷平台基本信息	平台名称、平台所属企业名称、平均收益、投资期限、注册资金、自动投标、债权转让、资金托管、投标保障、保障模式、担保机构、银行存管、风险准备金存管、业务类型、平台背景
	P2P网贷平台所属公司信息	平台名称、公司名称、公司法人、公司类型、股东结构、注册资本、实缴资本、注册地址、开业日期、核准日期、登记机关、营业执照注册号、组织机构代码、税务登记号
	P2P问题平台	平台名称、公司名称、所在地、发生问题时间、法定代表人、注册资本、利率范围、期限范围、上线时间、公司规模、问题原因
	P2P网贷平台高管基本信息	平台名称、P2P网贷平台高管姓名、职位、毕业学校、专业、高管简介

5. 数据范围

数据模块	数据起始时间
低风险理财产品	2004年
阳光私募基金产品	2003年
P2P理财产品	2000年

（五）基金公司研究数据库-FCRD

1.数据库简介

基金公司研究数据库（*Fund Company Research Database, FCRD*）基于相关基金公司的年报信息和披露信息等而建立的专业数据库。*FCRD* 包含基金公司基本信息、基金管理信息、基金经理信息 3 个模块，内容涵盖基金公司概况、股东情况、持股公司、旗下基金、管理费用等多方面数据，此外，还对基金经理的信息进行了详细展示。

2.数据库特色

- 较为完善的基金公司和基金经理数据

FCRD 包含了基金公司股东构成、持股公司、旗下基金等多类重要数据。*FCRD* 不仅介绍基本个人信息，还包括了基金经理的教育背景、工作经历、管理基金等重要信息。

3.典型数据举例

基金持股公司

基金公司代码	基金公司名称	统计日期	持股公司代码	持有基金数	持股总量 (股)	占该公司流通股比例 (%)
500038	华安基金管理有限公司	2003-12-31	600019	6	57994372	3.08973744
500038	华安基金管理有限公司	2003-12-31	600036	6	43984734	2.93231561
500038	华安基金管理有限公司	2003-12-31	600009	5	39830897	5.80345732

4.数据库内容

数据模块	数据表	数据字段
基金公司基本信息	基金公司概况	基金公司代码、基金公司名称、成立日期、公司性质、注册资本、法定代表人、总经理、注册地址、办公地址
	基金公司股东构成	基金公司代码、基金公司名称、股东名称、统计日期、持股数量、持股比例、股东性质
基金管理信息	基金持股公司	基金公司代码、基金公司名称、统计日期、持股公司代码、持股公司名称、持有基金数、持股总量、持股占该公司流通股比例、持股总市值、持股市值占基金净值比例、持股市值占基金股票投资市值比例
	基金公司旗下基金	基金公司代码、基金公司名称、基金代码、基金名称、基金类型、基金经理、基金成立日期
	基金管理费用	基金代码、基金名称、统计日期、基金公司代码、基金管理公司、基金份额、资产净值、管理费用
基金经理信息	基金经理基本信息	基金经理姓名、性别、出生年份、学历/学位、国籍、简介
	基金经理工作经历	基金经理姓名、基金代码、基金简称、任职机构名称、职位、入职日期、离职日期
	基金经理管理基金	基金经理姓名、基金代码、基金名称、基金公司名称、投资类型、开始日期、结束日期
	基金经理教育背景	基金经理姓名、基金代码、基金简称、学历/学位、专业、毕业学校

5.数据范围

数据模块	数据起始时间
基金公司基本信息	1998 年
基金管理信息	2001 年
基金经理信息	截面数据

（六）券商投行数据库-CIBD

1.数据库简介

券商投行数据库 (Chinese Investment Bank Database, **CIBD**) 从券商年报和披露信息中提取了券商财务信息、大股东及其持股比例信息、大股东性质信息, 董监高个人信息等独特数据, 提供了丰富和准确的券商投行相关数据。

2.数据库特色

- 独有的券商相关信息

CIBD 包括了在中国资本市场主要券商的大股东、高管、财务数据信息等, 是目前唯一提供此类数据的数据库。

3.典型数据举例

券商股东构成

券商名称	券商代码	会计报告日期	大股东名称	大股东持股比例	大股东性质	最终控制人
爱建证券有限责任公司	B00001	2014-12-31	上海陆家嘴金融发展有限公司	0.51136	地方国有	地方国资局
安信证券股份有限公司	B00002	2014-12-31	国家开发投资公司	0.5801	中央国有	国资委
安信证券股份有限公司	B00002	2014-12-31	中国证券投资者保护基金有限责任公司	0.2588	中央国有	国资委
渤海证券股份有限公司	B00006	2014-12-31	天津市泰达国际控股(集团)有限公司	0.536695	地方国有	地方国资局

4.数据库内容

数据模块	数据表	数据字段
券商基本信息	股东构成	券商名称、券商代码、统计日期、大股东名称、大股东占股比例 (%)、大股东性质、最终控制人
	高管特征	统计日期、券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、职位、是否全职
	承销排名	券商名称、券商代码、会计年度、当年承销 IPO 总募集资金金额、当年承销 IPO 总募集资金排名、当年承销 IPO 总募集资金占当年 IPO 市场总募集资金比例 (%)、当年 IPO 承销费用、平均承销费率 (%)、当年 IPO 承销费用排名、当年 IPO 承销数量、当年联席保荐 IPO 数量、当年 IPO 承销数量排名
	财务比率	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、券商净资本、营业收入、营业收入增长率 (%)、手续费及佣金净收入、手续费及佣金净收入在营业收入中的占比 (%)、利息净收入、利息净收入在营业收入中的占比 (%)、投资净收益、投资净收益在营业收入中的占比 (%)、公允价值变动净收益、公允价值变动净收益在营业收入中的占比 (%)、其他收入、

数据模块	数据表	数据字段
券商基本信息	财务比率	其他收入在营业收入中的占比(%)、代理买卖证券业务净收入、证券承销业务净收入、受托客户资产管理业务净收入
券商财务报表	资产负债表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、货币资金、客户资金存款、交易性金融资产、衍生金融资产、买入返售金融资产、存出保证金、可供出售金融资产、持有至到期投资、长期股权投资、固定资产、无形资产、资产总计、短期借款、质押借款、拆入资金、交易性金融负债、衍生金融负债、卖出回购金融资产款、代理买卖证券款、代理承销证券款、应付职工薪酬、应交税费、长期借款、应付债券、负债合计、股本、资本公积金、盈余公积金、未分配利润、一般风险准备、归属于母公司所有者权益合计、所有者权益合计、负债及股东权益总计
	利润表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、营业收入、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益、稀释每股收益
	现金流量表	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、经营活动产生的现金流量净额、投资活动产生的现金流量净额、筹资活动产生的现金流量净额

5.数据范围

数据模块	数据起始时间
券商基本信息	2003 年
券商财务报表	2006 年

(七) 保险机构研究数据库-IIRD

1.数据库简介

保险机构研究数据库 (*Insurance Institution Research Database, IIRD*) 基于中国财险、寿险等公司的年报数据以及中国保险年鉴的相关数据而建立的专业数据库。*IIRD* 囊括了 2005 年至今相关保险公司的经营和财务信息, 数据内容丰富完整, 结构清晰, 极大满足了学术研究的需求。

2.数据库特色

- 丰富的财报字段, 涵盖大多数保险公司

IIRD 数据来源于各大保险公司年报及中国保险年鉴, 因此展示了较为完整的财报字段和全面的财务数据。*IIRD* 通过多方数据来源, 收集整理了大部分保险公司的基本信息和财务信息。

3.典型数据举例

保险公司基本信息

保险公司代码	保险公司名称	保险公司类型	成立日期	所在城市
800005	安邦财产保险股份有限公司	保险公司(财险)	2011-12-31	深圳市
800006	安邦养老保险股份有限公司	保险公司(养老险)	2013-12-31	北京市
800007	安诚财产保险股份有限公司	保险公司(财险)	2006-12-31	重庆市
800008	安达保险有限公司	保险公司(财险)	2008-02-01	上海市

4.数据库内容

数据模块	数据表	数据字段
保险公司基本信息	保险公司基本信息	保险公司代码、保险公司名称、保险公司类型、成立日期、注册资本、所在城市、注册地址、经营范围、法定代表人、总经理
保险公司财务报表	资产负债表	保险公司代码、保险公司名称、统计日期、货币资金、交易性金融资产、衍生金融资产、买入返售金融资产、应收利息、应收代位追偿款、应收分保未到期责任准备金、应收分保未决赔款准备金、应收分保寿险责任准备金、应收分保长期健康险责任准备金、保户质押贷款、短期借款……
	利润表	保险公司代码、保险公司名称、统计日期、营业收入、已赚保费、保费总收入、分保费收入、分出保费、提取未到期责任准备金、营业支出、赔付支出、摊回赔付支出、提取保险合同准备金净额、摊回保险责任准备金、保单红利支出、分保费用、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、少数股东损益、归属母公司股东的净利润、每股收益、稀释每股收益
	现金流量表	保险公司代码、保险公司名称、统计日期、经营活动产生的现金流量、投资活动产生的现金流量、筹资活动产生的现金流量

5.数据范围

数据模块	数据起始时间
全部模块	2005 年

（八）信托机构研究数据库-CTID

1.数据库简介

信托机构研究数据库（*Chinese Trust Institution Database, CTID*）主要根据中国资本市场上信托机构披露的年报数据整理而成的数据库。涵盖了中国信托公司基本信息、股东构成、高管特征、审计信息、员工与股东信息以及财务信息等多类重要信息，为学者研究信托行业提供了完善而丰富的数据。

2.数据库特色

- 市场独有的信托公司数据

CTID 囊括了企业工商信息、股东信息、人员信息及财务信息。这些数据内容目前为市场独有，字段丰富而完善。

3.典型数据举例

信托公司基本信息

信托公司代码	信托公司名称	所在城市	成立日期	注册资本
200001	陕西省国际信托股份有限公司	西安市	1985-01-05	3964012846
200002	云南国际信托有限公司	昆明市	1997-09-03	1200000000
200003	西部信托有限公司	西安市	2002-07-18	1500000000
200004	华润深国投信托有限公司	深圳市	1982-08-24	6000000000
200005	五矿国际信托有限公司	西宁市	1997-09-23	6000000000

4.数据库内容

数据模块	数据表	数据字段
信托公司 基本信息	公司概况	信托公司代码、信托公司名称、所在城市、成立日期、注册地址、注册资本、经营范围、法定代表人、总经理
	股东构成	信托公司代码、信托公司名称、统计日期、序号、股东名称、持股金额、期末股份数、期末持股比例（%）、出资比例（%）、法定代表人、注册资本、注册地址、主要经营业务
	高管特征	信托公司代码、信托公司名称、统计日期、高管姓名、性别、年龄、学历/学位、专业、职位、选任日期、兼职单位名称、兼职单位职务、任职的股东单位名称、任职的股东单位职务、金融从业年限、所推举的股东名称、推举股东的持股比例（%）、简历
	审计信息	信托公司代码、信托公司简称、统计日期、审计日期、审计意见类型、审计师、会计师事务所名称、审计报告
	员工与股东信息	信托公司代码、信托公司名称、统计日期、股东总数、员工总数、董事/监事及高管人员数量、自营业务人员数量、信托业务人员数量、其他人员数量、博士研究生学历人数、硕士研究生学历人数、研究生学历占比（%）、大学本科学历人数、大学本科学历占比（%）、大专及以上学历人数、大专及以上学历占比（%）、员工平均年龄
信托公司 财务报表	资产负债表	信托公司代码、信托公司名称、统计日期、资产总计、流动资产、长期投资、流动负债、长期负债、负债合计、实收资本、所有者权益

数据模块	数据表	数据字段
信托公司 财务报表	利润表	信托公司代码、信托公司名称、统计日期、营业收入、营业支出、投资收益、利润总额、净利润
	其他指标	信托公司代码、信托公司名称、统计日期、信托资产运用合计、信托收入合计、信托费用合计、信托损益、资本收益率(%)、信托报酬率(%)、员工人数

5.数据范围

数据模块	数据起始时间
全部模块	2003 年

社会经济组织研究

(一) 政府审计数据库-CGAD

1. 数据库简介

政府审计数据库 (*Government Audit Database, CGAD*) 是基于中国审计年鉴、国家审计署等官方数据整理而成的专业数据库。该数据库全面记录了各省、市、县在各年度的政府审计情况、违规处理信息、各省份的审计厅领导信息等。*CGAD* 共包括三张表, 分别是地方政府审计信息、审计工作综合信息和地方审计部门领导。

2. 数据库特色

- 字段丰富

CGAD 除了包括各地市每年的审计结果信息, 还包括省级审计厅的领导信息, 以及审计署及其分派机构、地方审计单位的详细审计情况。

3. 典型数据举例

地方政府审计信息

省份名称	地市名称	会计年度	审计机关	完成审计项目 (个)	应上缴财政 (万元)	出具审计报告数量 (篇)	提出审计建议 (条)
安徽省	宿州市	2013	宿州市本级	26	68	37	626
安徽省	宿州市	2013	萧县审计局	17	48	21	120
安徽省	宿州市	2013	灵璧县审计局	36	50	62	93
安徽省	宿州市	2013	泗县审计局	13	1209	18	95
安徽省	淮北市	2013	淮北市本级	52	3847	90	170

4. 数据库内容

数据模块	数据表	数据字段
地方政府审计信息	地方政府审计信息	省份名称、地市名称、会计年度、审计机关、完成审计项目、审计查出主要问题金额、应上缴财政、应减少财政拨款或补贴、应归还原渠道资金、应调账处理金额、移送处理事项、出具审计报告数量、提出审计建议、提交审计信息
审计工作综合情况	审计工作综合情况	指标名称、计量单位、会计年度、合计、审计署本级汇总、审计署派出局汇总、审计署特派办汇总、地方审计机关汇总、北京市、天津市、河北省、山西省、内蒙古自治区、辽宁省、吉林省、黑龙江省、上海市、江苏省、浙江省、安徽省、福建省……
地方审计部门领导	地方审计部门领导	人员姓名、会计年度、省份名称、人员职位

5. 数据范围

数据模块	数据起始时间
全部模块	1999 年

(二) 中国高校科技统计数据库- USTS

1.数据库简介

中国高校科技统计数据库 (*Science and Technology Statistics Database of Chinese University, USTS*) 主要摘录和整理了来自中华人民共和国教育部科学技术司《高等学校科技统计资料汇编》的重要数据。该数据库记录了按学校和按地区统计的中国各大高校的科研情况,详细反映了高校的科技活动、研发情况和科技成果等情况。

2.数据库特色

- 数据权威可靠

USTS 的主要数据来自教育部,因此统计数字翔实可靠,字段丰富,可直接为研究者使用。

3.典型数据举例

高校科技活动分学校统计情况

学校名称	会计年度	教研人员总数	科技经费总额	科技课题总数	专著数量	成果授奖
北京大学	2012	12601	2342999	3815	31	51
清华大学	2012	5293	3892514	5135	45	115
北京交通大学	2012	1418	718431	3311	17	29
北京工业大学	2012	1796	640682	2275	14	31
北京邮电大学	2012	1393	407970	1388	22	7

4.数据库内容

数据模块	数据表	数据字段
高校科技活动	高校科技活动分学校统计情况	学校名称、学校类型、会计年度、教学与科研人员总数、其中:科学家与工程师人数 a、其中:具有高级职称的科学家与工程师人数 a、研究与发展人员总数、其中:科学家与工程师人数 b、其中:具有高级职称的科学家与工程师人数 b、全时当量研究与发展人员、科技经费总额、科技经费当年拨入金额、其中:政府资金拨入、其中:当年企事业单位委托拨入(进入学校财务)、其中:当年其他拨入、当年科技经费内部支出金额、科技课题总数、科技课题当年投入人数、科技课题当年拨入经费、科技课题当年支出经费、专著数量、专著字数、学术论文……
	高校科技活动分地区统计情况	地区名称、会计年度、教学与科研人员总数、其中:科学家与工程师人数 a、其中:具有高级职称的科学家与工程师人数 a、研究与发展人员总数、其中:科学家与工程师人数 b、其中:具有高级职称的科学家与工程师人数 b、全时当量研究与发展人员、科技经费总额、科技经费当年拨入金额、其中:政府资金拨入、其中:当年企事业单位委托拨入(进入学校财务)、其中:当年其他拨入、当年科技经费内部支出金额、科技课题总数、科技课题当年投入人数、科技课题当年拨入经费、科技课题当年支出经费、专著数量、专著字数、学术论文、其中:国外及全国性刊物发表……

数据模块	数据表	数据字段
高等学校研发情况	高校研发机构分地区统计情况	地区名称、会计年度、机构数、从事研究与发展人员合计数、其中：具有高级职称人数、其中：具有中级职称人数、其中：具有初级职称人数、其中：具有其他职称人数
	高校研发项目分地区统计情况	地区名称、会计年度、研发项目合计数、研发项目合计当年投入人员、研发项目合计参与研究生、研发项目合计当年拨入经费、研发项目合计当年支出经费、基础研究项目数、基础研究当年投入人员、基础研究参与研究生、基础研究当年拨入经费、基础研究当年支出经费、应用基础项目数、应用基础当年投入人员、应用基础参与研究生、应用基础当年拨入经费、应用基础当年支出经费、试验发展项目数、试验发展当年投入人员、试验发展参与研究生、试验发展当年拨入经费、试验发展当年支出经费
高等学校科技成果	高校科技成果分地区统计情况	地区名称、会计年度、出版科技著作数量、出版科技著作字数、发表学术论文合计、其中：发表在国外学术刊物的学术论文数、国家级项目验收合计数、其中：与其他单位合作的国家级项目验收、验收项目来源于 973 计划的数量、验收项目来源于科技攻关计划的数量、验收项目来源于 863 计划的数量……
	高校技术转让分地区统计情况	地区名称、会计年度、签订技术转让合同合计数、与国有企业签订的技术转让合同数、与外资企业签订的技术转让合同数、与民营企业签订的技术转让合同数、与其他机构签订的技术转让合同数、签订技术转让合同合计金额、与国有企业签订的技术转让合同金额……
	高校科技获奖分地区统计情况	地区名称、会计年度、国家最高科学技术奖特等奖、国家自然科学奖二等奖、国家发明奖一等奖、国家发明奖二等奖、国家科技进步奖特等奖、国家科技进步奖一等奖、国家科技进步奖二等奖、国务院各部门科技进步奖、省、自治区、直辖市科技进步奖
其他	高校科技人力分地区统计情况	地区名称、会计年度、学校数、教学与科研人员合计、其中：科学家和工程师 a、研究与发展人员合计、其中：科学家和工程师 b、研究与发展全时人员合计、……
	高校科技经费分地区统计情况	地区名称、会计年度、学校数、拨入经费合计、其中：拨入科研事业费、其中：拨入主管部门专项费、其中：拨入其他政府部门专项费、其中：拨入企事业单位委托经费、其中：各种收入中转为科技经费、其中：拨入其他经费、支出经费合计、其中：支出科研人员费、其中：支出业务费、其中：支出固定资产购置费、其中：支出上缴税金、其中：支出其他经费、其中：支出转拨给外单位经费
	高校国际科技交流分地区统计情况	地区名称、会计年度、国际合作研究派遣、国际合作研究接受、国际学术会议出席人员、国际学术会议交流论文、国际学术会议特邀报告、国际学术会议主办

5.数据范围

数据模块	数据起始时间
全部模块	2003 年

（三）中国非营利组织数据库-CNGO

1.数据库简介

中国非营利组织数据库 (*Chinese Non-Government Organization Database, CNGO*) 是面向广大高校教授和学者以及实务界人士, 研究中国非营利组织机构特征、财务状况、人员概况以及捐赠情况的数据库。目前该数据库主要包括我国所有公募和非公募慈善基金会的独特数据, 是反映中国慈善基金会发展状况较为全面的学术数据库。

2.数据库特色

- 全面

在时间跨度上, *CNGO* 包括所有中国慈善基金会 2005 年至今的所有基本信息, 学术研究提供较长的时间序列数据; 在空间跨度上, 除了个别未公开披露、数据缺失的基金会, *CNGO* 收录了中国大部分的基金会信息, 这为学术研究提供了大量的横截面数据。

- 唯一

CNGO 的高管模块收集了基金会大部分理事和监事在其他单位的任职背景, 这为考察慈善基金会的各种行为提供了不同视角。同时, 该库提供的审计信息、公益活动信息等也是当前学术市场所没有的。

3.典型数据举例

(1) 基金会信息

基金会代码	基金会名称	关注领域	基金会类型	基金会范围	登记部门	成立时间	原始基金数额
10001	爱佑慈善基金会	儿童、医疗救助	非公募	全国	民政部	2008/5	2000 万
10002	北京大学教育基金会	教育、科学研究	非公募	全国	民政部	1995/7	2000 万
10293	北京当代艺术基金会	教育、文化	公募	地方	省民政部	2008/12	400 万

(2) 基金会财务概况

基金会名称	会计年度	年度总收入 (元)	捐赠收入 (元)	年度总支出 (元)	公益事业支出 (元)
北京大学教育基金会	2014	786623183	567967683	309490517	302509385
北京大学教育基金会	2013	1489656769	1395841805	175105873	209366023
北京大学教育基金会	2012	428954733	383523157	267340159	259211412
北京大学教育基金会	2011	622791928	618194642	290777664	282008922

(3) 基金会理事背景

基金会名称	会计年度	姓名	职务	性别	出生日期	工作单位及职务
北京市温暖基金会	2012	惠远霖	理事	男	1954-01-05	北京市海淀区总工会主席
北京市温暖基金会	2012	刘欢	理事	女	1967-12-03	北京市教育工会常务副主席
北京市温暖基金会	2012	刘丽臣	理事	女	1952-06-07	北京市建工集团工会主席
北京市温暖基金会	2012	司建	副理事长	男	1963-01-14	北京市总工会权益部部长

4.数据库内容

数据模块	数据表	数据字段
基本信息	基金会信息	基金会代码、基金会名称、基金会地址、所属省份、宗旨、负责人中国工作人员数、关注领域、负责人中担任过省部级职务人员数、基金会类型、全职员工数量、基金会范围、志愿者数量、登记部门、评估等级、业务主管单位、成立时间、原始基金数额
	基金会关联方	基金会代码、基金会名称、会计年度、关联方、与基金会的关系
	审计信息	基金会代码、基金会名称、会计年度、审计意见、审计机构、注册会计师_1、注册会计师_2
财务信息	财务概况	基金会代码、基金会名称、会计年度、净资产、年度总收入、捐赠收入、投资收入、服务收入、政府补助收入、其他收入、年度总支出、用于公益事业的支出、工作人员工资福利支出、行政办公支出、其他支出
	业务活动表	基金会代码、基金会名称、会计年度、捐赠收入_非限定性、会费收入_非限定性、提供服务收入_非限定性、商品销售收入_非限定性、政府补助收入_非限定性、投资收益_非限定性、其他收入_非限定性、收入合计_非限定性、业务活动成本_非限定性、管理费用_非限定性、筹资费用_非限定性、其他费用_非限定性、费用合计_非限定性……
	资产负债表	基金会代码、基金会名称、会计年度、货币资金、短期投资、应收款项、预付账款、待摊费用、一年内到期的长期债权投资、其他流动资产、流动资产合计、长期股权投资、长期债权投资、长期投资合计、固定资产原价、累计折旧、固定资产净值、在建工程、文物文化资产、固定资产清理、固定资产合计、无形资产……
捐赠情况	公益支出情况	基金会代码、基金会名称、会计年度、基金会类型、上年度基金余额（非公募）/上年度实际收入合计(公募)、本年度总支出、本年度用于公益事业的支出、工作人员工资福利支出、行政办公支出
	接受捐赠情况	基金会代码、基金会名称、会计年度、本年度捐赠收入、来自境内的捐赠、来自境内自然人的捐赠、来自境内法人或者其他组织的捐赠、来自境外的捐赠、来自境外自然人的捐赠、来自境外法人或者其他组织的捐赠
	大额捐赠情况	基金会代码、基金会名称、会计年度、捐赠方、捐赠金额、捐赠用途
	重大公益项目	基金会代码、基金会名称、会计年度、项目名称、大额支付对象、支付金额、占年度公益总支出比例(%)、用途
	业务活动情况	基金会代码、基金会名称、会计年度、项目名称、项目本年度收入、项目本年度支出、项目介绍
高管信息	理事背景	基金会代码、会计年度、基金会名称、理事姓名、理事会职务、性别、出生日期、年度出席会议次数、工作单位及职务
	监事背景	基金会代码、会计年度、基金会名称、监事姓名、性别、出生日期、年度出席会议次数、工作单位及职务

5.数据范围

数据模块	数据起始时间
全部模块	2005年

（四）中国商会研究数据库-CCCD

1.数据库简介

中国商会研究数据库 (*Chamber of Commerce Research Database, CCCD*) 中所涉及的商会主要指异地商会,是指由异地同一籍自然人或法人在某级行政辖区内投资兴办并自愿发起组成,以原籍地行政区域名称命名为基本特征的非营利性社会团体。具体来说, *CCCD*数据库包含商会基本信息、理事会成员、主要会员企业 3 个模块,共 4 张表。商会基本信息模块主要包括商会自身的一些基本信息,例如,名称、原籍地、活动地域、成立时间等;理事会成员包括理事的个人信息,如姓名、性别、最高学历、毕业学校、籍贯等信息;主要会员企业信息包括会员企业基本信息以及股东出资信息情况。

2.数据库特色

- 唯一

CCCD 是目前国内唯一一个关注商会研究的数据库,弥补了该领域研究的数据缺陷。

3.典型数据举例

商会会员基本信息

商会代码	商会名称	会员企业名称	公司类型	法定代表人	注册资本	成立日期
A00001	上海市浙江商会	上海复星高科技(集团)有限公司	有限责任公司	郭广昌	480000 元	2005-03-08
A00001	上海市浙江商会	万丰奥特控股集团	有限责任公司	陈爱莲	12000 元	1998-03-04
A00001	上海市浙江商会	上海均瑶(集团)有限公司	有限责任公司	王均金	80000 元	2001-02-14

4.数据库内容

数据模块	数据表	数据字段
商会基本信息	基本信息	商会名称、原籍地、活动地域、简介、章程、成立日期
理事会成员	理事会成员	商会名称、理事职位、理事姓名、性别、最高学历/学位、毕业学校、籍贯、政治面貌、理事所在单位及职位、理事其他社会职务、理事简历
主要会员企业	会员基本信息	商会名称、会员企业名称、经营状态、公司类型、成立日期、法定代表人、注册资本、注册资本单位、营业期限、登记机关、发照日期、所属行业、企业地址、经营范围、公司简介
	股东与出资	商会名称、公司名称、股东、认缴出资金额、认缴出资金额单位、认缴出资时间、实缴出资金额、实缴出资金额单位、实缴出资时间

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

三、CNRDS 经济特色数据库

随着中国成为国际上举足轻重的经济体，中国的经济发展模式和经济奇迹受到世界各国的关注，因此对于中国经济问题的研究逐渐成为国内外学术界的研究热点和前沿，而这些研究过程中的数据成为学者们的关键需求。**CNRDS**正是基于国内外学者在研究中国经济问题过程中的数据需求和数据质量要求，推出了一系列重要的经济特色数据库。其特色之处在于这些数据库非常符合中国经济问题研究过程中学者所关注的热点问题，虽获取难度大，但数据价值高且各有特色，可满足不同研究领域的独特数据需求。目前，**CNRDS**根据已有研究，结合经济热点问题，推出了以下四个系列数据：

宏观经济研究系列 主要包括中国经济发展过程中具有一定特色的宏观经济相关数据，例如中国环境问题数据、经济普查数据、高铁航线数据等。

区域经济研究系列 主要包括中国区域经济发展的数据，例如中国城市统计数据、中国农村统计数据等。

产业经济研究系列 主要分不同产业建立了相应数据库，例如，中国能源数据、房地产数据、科技统计数据等。

对外经济研究系列 囊括中国经济对外开放过程中的相关数据，包括对外直接投资、海外并购、一带一路等重要数据。

总之，经济特色库的推出不仅对具体数据库进行了清晰、正确的归类，更方便了经济研究领域学者的使用。这些经济特色数据库都自成体系，几乎囊括了经济研究该领域可能用到的关键数据，这大大缩短了繁琐的数据寻找、合并过程，真正减少了学者处理数据的时间成本。

CNRDS 经济特色数据库具有以下三个特点：

- 数据多数来自官方渠道，权威可靠
- 深度整合源数据，数据结构简约
- 多维度经济数据，满足不同领域研究

宏观经济研究系列

（一）中国经济普查数据库-CECD

1.数据库简介

中国经济普查数据库 (*Chinese Economic Census Database, CECD*) 包含了我国目前已完成的 2004、2008、2013 三次大规模全国经济普查数据,涵盖了我国境内从事第二产业和第三产业的全部法人单位、产业活动单位和个体经营户。*CECD* 分为中央数据和地区数据,数据内容包括产业单位、法人单位以及 31 个省份的经济普查情况,数据内容全面而丰富,可为学术研究提供更多支持。

2.数据库特色

- 数据来源权威

CECD 来源于国家统计局、地方统计机构以及统计年鉴等多个渠道,包含了三次全国经济普查数据,在保证每次普查数据特色的基础上,按照具体行业进行分类归纳,将零散的表格进行了合并整理。

3.典型数据举例

按地区、单位类别分组的法人单位数及从业人数

地区	单位类别	单位数 (个)	从业人数 (人)
福州市	法人单位合计	48267	2144823
鼓楼区	法人单位合计	13091	551817
台江区	法人单位合计	5347	164388

4.数据库内容

数据模块	数据表
综合	按机构类型、行业(门类)分组的法人单位数、产业活动单位数及从业人数;按行业(小类)分组的法人单位数、产业活动单位数及从业人数;按地区、开业(成立)时间分组的法人单位数、产业活动单位数及从业人数;按登记注册类型分组的法人单位数、产业活动单位数及从业人数;按地区、单位类别分组的法人单位数及从业人数;按行业(大类)、地区分组的法人单位数及从业人数;按行业(大类)、开业(成立)时间分组的法人单位数及从业人数;按地区、开业(成立)时间分组的法人单位数及从业人数;按地区、学历、技术分组的法人单位数及从业人数;按行业(小类)、登记注册类型、学历、技术分组的法人单位数及从业人数;按登记注册类型、行业(中类)分组的单位信息化情况;按地区分组的单位信息化情况;按行业(大类)分组的个体经营户情况;按地区分组个体经营户情况;按登记注册类型、地区、行业(大类)分组的单位数及从业人数……
餐饮和住宿业	住宿和餐饮企业分登记注册类型经营情况;住宿和餐饮企业分行业经营情况;住宿和餐饮企业分地区经营情况;住宿和餐饮企业分登记注册类型主要财务指标;住宿和餐饮企业分行业主要财务指标;住宿和餐饮企业分地区主要财务指标;限额以上住宿和餐饮企业分登记注册类型主要财务指标;限额以上住宿和餐饮企业分行业主要财务指标;限额以上住宿和餐饮企业分地区主要财务指标;限额以下住宿和餐饮企业分登记注册类型主要财务指标;限额以下住宿和餐饮企业分行业主要财务指标;限额以下住宿和餐饮企业分地区主要财务指标;按规模分住宿和餐饮企业分登记注册类型设施状况;住宿和餐饮企业分登记注

数据模块	数据表
餐饮和住宿业	册类型设施状况；按规模分住宿和餐饮企业分行业设施状况；按规模分住宿和餐饮企业分地区设施状况；各市、县住宿和餐饮企业年末从业人员及收入情况；住宿和餐饮企业主要指标
房地产业	房地产业企业个数、年末从业人数、固定资产原价、营业收入；房地产开发企业个数和从业人数、投资情况、新增固定资产；房地产开发按构成、工程用途分的投资；房地产开发土地购置、土地开发；房地产开发资金来源；房地产开发按用途分的商品房面积和销售额、出租房屋面积；房地产开发资产、资本、经营情况；房地产物业管理企业人员、资本、经营情况；房地产中介服务人员、资本、经营情况；其他房地产企业人员、资本、经营情况；房地产开发企业主要指标
工业	工业企业单位数、总产值、人数；工业企业主要经济效益指标；工业企业主要产品产量；工业企业按年份分单位、总产值、人数；工业企业产值指标；规模以上工业产品生产、销售和库存；主要工业产品生产能力；规模以上工业企业经济效益综合指数；规模以上工业企业主要财务指标；工业企业主要财务指标；规模以上工业企业劳动生产率；各种工业企业主要经济效益指标；规模以上“三资”工业企业主要经济效益指标；各市规模以上工业企业主要产值、效益指标；各市、县规模以上工业总产值；各市、县规模以上工业企业从业人员平均人数；各市、县规模以上工业企业劳动生产率；各市、县规模以上工业企业主要财务指标……
建筑业	分地区分行业建筑业企业主要经济指标；总承包和专业承包建筑业企业合同；总承包和专业承包建筑业企业承包工程完成情况；总承包和专业承包建筑业企业总产值和竣工产值；总承包和专业承包建筑业企业房屋建筑面积和竣工面积；总承包和专业承包建筑业企业按主要用途分的房屋建筑竣工面积；总承包和专业承包建筑业企业按主要用途分的房屋竣工价值；总承包和专业承包建筑业企业主要生产效益；总承包和专业承包建筑业企业机械设备装备；总承包和专业承包建筑业企业从业人员；总承包和专业承包建筑业企业建筑材料消耗量；总承包和专业承包建筑业企业资产负债；总承包和专业承包建筑业企业所有者权益、实收资本……
批发和零售业	批发和零售企业分登记注册类型商品销售情况；批发和零售业企业分行业商品销售情况；批发和零售业企业分行业、分规模商品销售情况；设区市批发和零售企业分登记注册类型单位、人员、销售情况；设区市批发企业分登记注册类型单位、人员、销售情况；设区市零售企业分登记注册类型单位、人员、销售情况；设区市批发企业分行业单位、人员、销售情况；设区市零售企业分行业单位、人员、销售情况；设区市限额以上批发和零售企业分登记注册类型单位、人员、销售情况；设区市限额以上批发企业分登记注册类型单位、人员、销售情况；设区市限额以上零售企业分登记注册类型单位、人员、销售情况……
运输、仓储、邮政业	按登记注册类型分组的交通运输、仓储、邮政业单位数、从业人员数；按地区分组的交通运输、仓储、邮政业单位数、人员数及财务情况；交通运输、仓储、邮政业信息化情况；按行业分组的交通运输、仓储、邮政业主要财务指标；按登记注册类型、地区分组的交通运输、仓储、邮政业主要财务指标；各设区市交通运输、仓储、邮政业其他业务利润；按地区分组的交通运输个体户主要经营情况；按行业(大类)分组的交通运输、仓储、邮政业主要指标

5.数据范围

数据模块	数据起始时间
全部模块	2003年

（二）中国行政审批数据库-CAAD

1.数据库简介

中国行政审批数据库（*Chinese Administrative Approval Database, CAAD*）是基于全国 300 多个地级市行政审批中心数据而建立的数据库。随着中国经济改革的深入，行政审批中心的设立可以实现审批项目“一站式”办理，从而反作用于经济发展。因此，为了支持学者对这一领域的深入研究，*CAAD* 从多个维度提供了行政审批中心数据，重要字段包括是否设立行政审批中心、审批中心成立时间、进驻部门或单位、进驻事项和进驻窗口等。由于中国行政审批制度改革并未停止，近年来在行政审批中心基础上，多个地方纷纷设立行政审批局，启用行政审批专用章，从而实现了“一颗印章管审批”。因此，本文基于此还设立了是否组建行政审批局、行政审批局成立时间等重要字段，扩展了学者的研究思路。

2.数据库特色

- 数据唯一

近年来，行政审批效率研究已然成为学术研究的重要议题，但是这类研究数据的搜集需要学者耗费大量时间和精力，尚无公开数据库提供该类数据，*CAAD* 弥补了这类公开数据缺失的不足。

- 字段丰富

CAAD 基于全国 300 多个地级市审批中心数据，提供成立时间、进驻事项、进驻窗口等近 20 个重要字段。

3.典型数据举例

行政审批中心

省级名称	地级名称	行政审批单位名称	设立时间	进驻部门或单位	进驻事项	进驻窗口数量	进驻工作人员
河北省	唐山市	唐山市行政服务中心	2002-11-20	41	278	263	400
山西省	太原市	太原市政务服务中心	2003-01-00	53	450	70	160
内蒙古	包头市	包头市政务服务中心	2002-00-00	40	600	150	400
吉林省	长春市	长春市政务中心	2002-10-00	29	277	183	345
吉林省	松原市	松原市政务服务中心	2002-09-00	45	302	94	127

4.数据库内容

数据模块	数据表	数据字段
行政审批中心	行政审批中心	省级名称、地级名称、是否设立行政审批单位、行政审批单位名称、设立时间、单位介绍、是否组建行政审批局、行政审批局成立时间、进驻部门或单位、审批部门、服务部门、职能科室、公共服务单位、公共资源部门、内设机构、进驻事项、行政许可事项、进驻窗口数量、进驻工作人员

5.数据范围

数据模块	数据起始时间
全部模块	2000 年

（三）全球夜间灯光数据库-GNLD

1.数据库简介

全球夜间灯光数据库（Global Night-time Light Database, GNLD）是基于美国国家海洋与大气管理局(NOAA)的 DMSP/OLS 影像数据和 VIIRS/DNB 影像数据开发而成的，反映中国各省、地级市夜间灯光数据情况。DMSP/OLS 影像灯光数据目前已经应用于经济类研究，并且已有较多文献发表在诸多高水平杂志上。VIIRS/DNB 影像灯光数据在国内文献中还未被广泛应用，但基于其具有许多优点因此逐渐受到学者关注。由于学者通常只能获得 NOAA 上面的原始图片资料，将这些图像资料转化为可用数据需要用到较为复杂的计算机和编程技术，因此，为了方便学者使用该数据，促进中国问题研究，我司联合相关领域专家开发了全球夜间灯光数据库。具体的开发步骤请见数据库说明书。

2.数据库特色

- 丰富多样的数据字段

GNLD 既提供了连续时间序列的校正后的 DN 均值和灯光指数，还提供不同时期不同卫星测定的影像数据，方便学者进行深入研究。

- 提供地级市灯光数据

GNLD 不仅包括省份灯光数据，还包括所有地级市的灯光数据，为学者拓宽研究思路提供了支持。

3.典型数据举例

DMSP 中国各地级市灯光数据（校正后）

省份名称	城市名称	年度	DN 均值
江苏省	常州市	1998	10.3544
江苏省	常州市	1999	10.6286
江苏省	常州市	2000	10.7672

4.数据库内容

数据模块	数据表	数据字段
DMSP/OLS 全球灯光数据	DMSP 中国各省份灯光数据（校正后）	省份名称、年度、DN 均值
	DMSP 中国各省份灯光数据（原始）	省份名称、年度、所属卫星、DN 均值
	DMSP 中国各地级市灯光数据（校正后）	省份名称、城市名称、年度、DN 均值
	DMSP 中国各地级市灯光数据（原始）	省份名称、城市名称、年度、所属卫星、DN 均值
VIIRS/DNB 全球灯光数据	VIIRS 中国各省份灯光数据	省份名称、年度、月份、DN 均值
	VIIRS 中国各地级市灯光数据	省份、城市、年度、月份、DN 均值

5.数据范围

数据模块	数据起始时间
DMSP/OLS 全球灯光数据	1992 年
VIIRS/DNB 全球灯光数据	2013 年

（四）中国高铁航线数据库-CRAD

1.数据库简介

中国高铁航线数据库 (*Chinese High-speed Rail and Airline Database, CRAD*) 主要包含国内各省市所开通的高铁线路、高铁火车站点、飞机航线和航站的情况。其中, 高铁列车信息包括车站所在省市、列车车次、出发站、到达站、运行时间、里程等信息; 高铁站开通时间数据主要来自国家铁路总公司公布的自 2003 年以来高铁线路建设和开通情况, 该表提供各高铁站名称、开通时间、所属地区和所属高铁线路等信息; 高铁线路信息包括近百条已开通高铁的开通时间、站点和运营时速等信息; 飞机航线信息包括机场所在城市、机场名称、航班信息等。

2.数据库特色

CRAD 详细记录了中国各省、市、县所开通的列车站和飞机场的情况, 尤其是, 是否开通高铁站信息以及高铁站开通时间, 可以在一定程度上反映一个城市对外沟通和交流的便利程度, 可为学者研究提供更多的数据支持。

3.典型数据举例

高铁列车信息

列车车次	出发站	出发站类型	到达站	车型	运行时间	里程
G9464	南京南	过路车	徐州东	高速铁路	01 小时 38 分	331
D5641	南京南	始发车	芜湖	动车组	00 小时 43 分	20
G1113	南京南	始发车	广州南	高速动车	07 小时 35 分	1583
G7041	南京南	始发车	上海	高速动车	02 小时 09 分	311

4.数据库内容

数据模块	数据表	数据字段
高铁列车信息	高铁列车信息	列车车次、出发站、出发站所属地级市、出发站所属省份、出发站类型、到达站、到达站所属地级市、到达站所属省份、到达站类型、车型、开车时间、到站时间、运行时间、里程
	高铁站开通时间	高铁站名称、开通时间、所在省份、所在城市、所属线路名称、备注
	高铁线路信息	高铁线路名称、起点名、终点名、开通时间、线路长度(km)、设计速度(km/h)、沿途主要车站
飞机航线信息	飞机航线信息	起点城市、起点城市所属地级市、起点城市所属省份、起点机场、终点城市、终点城市所属地级市、终点城市所属省份、终点机场、航空公司、航班、机型、出发时间、到达时间、准点率、平均延误、班次_周一、班次_周二、班次_周三、班次_周四、班次_周五、班次_周六、班次_周日、有效期_起始、有效期_结束

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

（五）中国文化研究数据库-CCRD

1.数据库简介

中国语言类别众多，官话、方言、少数民族方言差异巨大，语言的差异体现着文化的差异；宗教的种类、分布，庙宇的数量也在一定程度上体现出各地区人们的宗教信仰程度，以及各宗教在中国的传播状况；中国不同地区因历史、文化和风俗等差异，对于购买彩票的态度也存在差异，博彩文化的地区差异在一定程度上反应了各地冒险精神或赌博偏好。中国文化研究数据库（*Chinese Culture Research Database, CCRD*）通过对中国语言地图信息、中国宗教庙宇信息、中国博彩文化信息的全面考察和整理，为学者进行文化研究提供了详实的数据来源。

中国语言地图信息 包含各省、市、县对应的语言片区，包括官话、汉语方言、少数民族方言的大类、小类与细分信息。

中国宗教庙宇信息 包含中国所有寺庙、道观、清真寺和教堂的多维度信息，如名称、建立日期、负责人姓名、所属地区、是否为国家文物保护单位、具体地址、邮编以及简介等。

中国博彩文化信息 包含各省、直辖市自 2008 年以来的地区不同类型彩票销售额、增长率信息，数据主要来自财政部官网相关内容披露。

2.数据库特色

- 独特性

CCRD 对中国各地区方言分布、中国宗教庙宇分布进行了综合梳理和总结，对中国博彩文化信息数据进行了整理，对研究中国文化传统和文化遗产具有重要参考价值。

- 完整性

CCRD 中国语言地图数据涵盖了中国现存的所有方言分布信息，中国宗教庙宇信息包含五个宗教所有的庙宇信息，中国博彩文化信息包括 2008 年至今的彩票销售情况。

- 多维性

CCRD 的各个模块都包含详细的字段内容。其中，中国语言地图信息模块将官话、汉语方言、少数民族方言全部进行片区与大小类的细分；中国宗教庙宇信息对基督教与天主教各个教堂的礼拜与弥撒时间都详细标出，对佛教、道教、伊斯兰教、基督教、天主教，各教的建筑状况均做出标识与介绍。

3.典型数据举例

(1) 汉语方言片区

区域代码	省级名称	市级名称	县级名称	官话区/汉语方言区	汉语方言 (大类)	汉语方言 (小类)	汉语方言 (细分)
350182	福建省	福州市	长乐市	汉语方言区	闽语	闽东区	侯官片
350205	福建省	厦门市	海沧区	汉语方言区	闽语	闽南区	泉漳片
350421	福建省	三明市	明溪县	汉语方言区	客家话	客家话	汀州片
350423	福建省	三明市	清流县	汉语方言区	客家话	客家话	汀州片

(2) 佛寺信息

寺庙名称	支派类别	建立日期	所属省份	所属地级市	所属区县	是否为国家文物保护单位	景区级别
白居易	藏语系	公元 1427 年	西藏自治区	日喀则市	江孜县	全国重点文物保护单位	国家 AAAA 级旅游景区
白马寺	汉语系	公元 68 年	河南省	洛阳市	洛龙区	全国重点文物保护单位	国家 AAAA 级旅游景区
风穴寺	汉语系	公元 190 年	河南省	平顶山市	汝州市	全国重点文物保护单位	国家 AAA 级旅游景区

4.数据库内容

数据模块	数据表	数据字段
中国语言地图信息	官话片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、官话（大类）、官话（片区）、官话（小片区）
	汉语方言片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、汉语方言（大类）、汉语方言（小类）、汉语方言（细分）
	少数民族方言片区	省级名称、地市级名称、县级名称、区域代码、官话区/汉语方言区/少数民族方言区、少数民族方言（语族）、少数民族方言（语族大类）、少数民族方言（语族小类）、少数民族方言（语族细分 1）、少数民族方言（语族细分 2）
中国宗教庙宇信息	佛寺	寺庙名称、支派类别、负责人姓名、建立日期、所属省份、所属地级市、所属区县、具体地址、佛寺简介、是否为国内名寺、是否为国家文物保护单位、景区级别
	道观	道观名称、支派类别、负责人姓名、建立日期、所属省份、所属地级市、所属区县、具体地址、道观简介、是否为国内名观、是否为国家文物保护单位、景区级别
	清真寺	清真寺名称、建立日期、所属省份、所属地级市、所属区县、具体地址、清真寺简介、是否为国内名寺、是否为国家文物保护单位、景区级别
	基督教堂	教堂名称、建立日期、所属省份、所属地级市、所属区县、具体地址、礼拜时间、基督教堂简介
中国博彩文化信息	全国各地区彩票销售情况表	地区、年份、月份、福利彩票本月销售额、福利彩票本月销售额比上年同期增长（%）、福利彩票本年累计销售额、福利彩票本年累计销售额比上年同期增长（%）、体育彩票本月销售额、体育彩票本月销售额比上年同期增长（%）、体育彩票本年累计销售额、体育彩票本年累计销售额比上年同期增长（%）、本月合计销售额、本月合计销售额比上年同期增长（%）、本年累计销售额、本年累计销售额比上年同期增长（%）、销量累计排序
	全国各类型彩票销售情况表	类型、年份、月份、本月销售额、上年同期销售额、本月销售额比上年同期增长（%）、该月环比增长（%）、本年累计销售额、上年累计销售额、本年累计销售额比上年同期增长（%）

5.数据范围

数据模块	数据起始时间
中国语言地图信息	截面数据
中国宗教庙宇信息	截面数据
中国博彩文化信息	2008 年

（六）中国环境数据库-CEDS

1. 数据库简介

环境状况对于经济发展与资源配置起着重要作用。鉴于环境对经济发展和企业行为可能产生的重要影响，本数据库为学者提供研究中国气候与环境状况的详细、全面、准确的字段信息。

中国环境数据库 (*Chinese Environment Database, CEDS*) 对中国环境状况进行全方面的考察与总结，共包含气象、环境、自然灾害、城市地理信息、自然生态、环境治理、环境影响评价、环评工作人员全国各省、地级市 1998 年以来各年年均 PM2.5 等 9 个模块。

(关于 1998 年以来各年年均 PM2.5 模块详细数据处理流程可见说明书附录)

气象数据模块 包含历史天气数据，如城市当天的天气状况、最高气温、最低气温、风向和风力等信息。

环境数据模块 包括历史空气质量，如城市各项空气质量指标，二氧化硫、二氧化氮、一氧化碳、PM2.5、PM10、臭氧等；污染物排放数据，如不同省份不同年份不同污染物排放种类的排放数据；水体环境数据，如水质状况、溶解氧等化学物质含量、河流名称、PH 值等。

自然灾害模块 包含台风灾害与地震灾害两个部分，如灾害发生时间、地点、力度等级与类型等信息。

城市地理信息 不仅包含省市县名称及区号等常见信息，还包括城市区县的经纬度信息与海拔数据。

自然生态模块 包含全国自然保护区名录（如保护区具体位置、面积大小、保护区主要保护对象、级别）、国家级生态乡镇名单、国家生态工业示范园区、环境标志获证企业（如企业所在地、证书有效期限、认证产品种类等）。

环境治理模块 包含污染物排放国家重点监控企业名单，具体包含被监控企业所属行政区、企业法人代码、监控类别等具体内容。

环境影响评价机构 包含环境影响评价机构详细信息（如机构资质等级，出具报告书的评价范围，相关联系人）、环境影响评价机构基本信息、环境影响评价机构诚信信息（如机构受到奖惩的时间、奖惩类型以及奖惩的具体情况）。

环评工作人员 包含环境影响评价工程师和岗位证书持有人员。

1998 年以来各年年均 PM2.5 包含全国省级、地市级自 1998 年以来各年年均 PM2.5 数据。

2. 数据库特点

● 独特性

中国环境数据库是国内唯一专注于中国环境研究的学术数据库，该数据库对中国环境各方面的数据都进行了梳理总结，不仅包含气象、环境和自然灾害数据，还汇总了各保护区、自然生态、环评机构与环评工作人员的信息。

● 多维性

中国环境数据库不仅包括基础的大气环境与自然灾害方面的信息，还包含有城市地理位置、环评机构与人员、国家污染排放重点监控企业名单等环境相关的内容。本数据库还将大气环境的内容细分为污染物排放、空气质量、大气环境等多个方面。

● 完整性

中国环境数据库包括大气与灾害、自然生态、环境影响评价等信息，数据库内容丰富，对环境相关的内容进行延伸，做到字段的完整以及整个环境系列数据的全面覆盖。

3.典型数据举例

空气质量信息

城市名称	日期	质量等级	AQI指数	当天AQI排名	二氧化硫	二氧化氮	一氧化碳	PM2.5	PM10
上海	2015/12/01	良	90	200	16	75	1.03	66	52
上海	2015/12/02	良	97	259	20	54	0.98	71	75
上海	2015/12/03	良	97	332	29	54	0.92	54	142
上海	2015/12/04	良	71	256	36	54	1.00	43	87
上海	2015/12/05	轻度污染	103	268	31	87	1.43	77	115
上海	2015/12/06	轻度污染	143	321	33	80	1.61	107	137
上海	2015/12/07	良	96	206	21	69	1.12	69	90
上海	2015/12/08	优	49	75	15	63	0.74	32	48

4.数据库内容

数据模块	数据表	数据字段
气象数据	天气	城市编号、所属省份、所属地级市、城市名称、日期、星期时间、最高气温、最低气温、天气情况、风向、风力
环境数据	空气质量	城市编号、城市名称、所属省份、日期、AQI指数、质量等级、当天AQI排名、PM2.5 24小时均值、PM10 24小时均值、二氧化硫 24小时均值、二氧化氮 24小时均值、一氧化碳 24小时均值、臭氧指标 24小时均值
	污染物排放	省级名称、年度、指标种类、排放量指标数值、排放量指标单位、子指标、具体污染指标
	水体环境	水系名称、点位名称、断面情况、溶解氧(mg/l)、氨氮含量(mg/l)、高锰酸钾化学需氧量(mg/l)、PH值、河流或湖泊名称、上周水质、本周水质、主要污染物指标、报告年份、报告周数
自然灾害	台风灾害	台风编号、台风中文名称、台风英文名称、年度、台风数据来源、台风中心风速(m/s)、时间、台风中心风力、台风中心气压、7级风力半径(km)、类型、10级风力半径(km)、台风中心经度、台风中心纬度
	地震灾害	发震日期、发震时间、震源纬度、震源经度、震源深度、震级、震源位置
城市地理信息	城市经纬度	城市编码、城市级别、省级名称、地市级名称、县级名称、区号、邮政编码、经度、纬度
	城市海拔	城市编号、城市级别、省级名称、地市级名称、县级名称、区号、邮政编码、海拔
自然生态	全国自然保护区名录	保护区编号、保护区名称、所属省份、所属地级市、所属区县、总面积、主要保护对象、类型、级别、始建时间、主管部门
	国家级生态乡镇名单	年度、省级名称、地市级名称、县级名称、乡镇名称

数据模块	数据表	数据字段
自然生态	国家生态工业示范园区	示范园区编号、示范园区名称、类别、批准文号、批准时间
	环境标志获证企业	企业名称、企业编号、所属省份、产品种类、批准时间、有效期限、证书编号、证书状态
环境治理	污染物排放国家重点监控企业名单	行政区划代码、法人代码、公司名称、监控类别、所属省份、所属年份
环境影响评价	环境影响评价机构资质信息	所属省份、所属地级市、机构名称、机构资质等级、机构资质证书号、甲级报告书评价范围、乙级报告书评价范围、报告表评价范围、有效期
	环境影响评价机构基本信息	所属省份、所属地级市、机构名称、成立时间、首次发证时间、经济类型、通讯地址、邮编
	环境影响评价机构诚信信息	所属省份、所属地级市、机构名称、奖惩时间、奖惩部门、文号、奖惩类型、具体情况
环评工作人员	环境影响评价工程师	环评工程师姓名、登记单位、登记证号、登记类别、登记有效期起始日期、登记有效期终止日期、职业资格证书号
	岗位证书持有人员	岗位持证人姓名、岗位证书号、所在单位、所属省份
1998 年以来各年年均 PM2.5	省级 1998 年以来各年年均 PM2.5	年度、省级名称、地区编码、PM2.5 年均浓度 (微克/立方)
	地市级 1998 年以来各年年均 PM2.5	年度、省级名称、地市级名称、地区编码、PM2.5 年均浓度 (微克/立方)

5.数据范围

数据模块	数据起始时间
气象数据	2011 年
环境数据	2009 年
自然灾害	2002 年
城市地理信息	截面数据
自然生态	2007 年
环境治理	截面数据
环境影响评价	2005 年
环评工作人员	2013 年
1998 年以来各年年均 PM2.5	1998 年

(七) 中国税收研究数据库-CTRD

1. 数据库简介

中国税收研究数据库 (*Chinese Tax Research Database, CTRD*) 基于对中国国税、地税以及上市公司税收的深入考察, 建立的详细、准确和全面的中国税收研究专业数据库。该库不仅包括中国相关税务部门分地区、分税种、分企业类型、分产业等的税收收入情况, 详细的分市县各税种税收收入, 上市公司税收信息, 还包括国家税务局与各地区地税部门领导信息、相关从业人员基本情况。

2. 数据库特色

- 唯一性与多维度

CTRD 是目前市场上唯一提供中国税收收入详细信息的数据库, 弥补了我国税收研究数据的空缺。该数据库从中国税收、地税收入、上市公司税收和税务官员信息四个方面, 分地区、分税种、分企业类型、分产业等多个维度详细记录了我国税收的发展情况。

- 丰富的税务官员信息

CTRD 囊括中国税务部门的领导信息以及从业人员基本情况, 包括了这些官员的姓名、所在单位、职务、任期等信息, 这为研究提供了丰富的人物信息。

3. 典型数据举例

(1) 各省地税部门分税种分市、县收入情况

所在省级名称	所在市级名称	所在县级名称	会计年度	地税部门组织收入总计 (万元)	一、税收收入 (万元)	二、基金费收入合计 (万元)	三、税务部门其他罚没收入 (万元)
安徽省	合肥市	合肥市合计	2009	1722736	1192646	529463	627
安徽省	芜湖市	繁昌县	2010	64391	43258	21131	2
安徽省	宣城市	宣州区	2011	121581	88805	32771	5
安徽省	滁州市	来安县	2012	97850	72550	25299	1
安徽省	合肥市	市区	2013	3673069	2497914	1174956	199
安徽省	亳州市	亳州市合计	2014	725812	551119	174669	24

(2) 各省地税部门领导信息

地区名称	会计年度	部门	职位	姓名
安徽省地税局	2011	局领导	党组书记、局长	汪建国
合肥市地税局	2012	合肥市地税局	党组书记、局长	邢孝鸿
天长市地税局	2013	天长市地税局	副局长	姜培忠、高立新
阜阳市地税局	2014	阜阳市地税局	政治处主任	陶小杰

4.数据库内容

数据模块	数据表
中国税收收入信息-全国税收收入	税收收入分税种分产业情况、税收收入分税种分地区情况、税收收入分税种分级次情况、税收收入分税种分企业类型情况、税收收入分地区分产业情况、税收收入分地区分企业类型情况
中国税收收入信息-具体税收收入	各类税收分地区分企业收入情况、增值税收入分项目分企业类型收入情况、资源税分项目分企业类型收入情况、营业税分行业分企业类型收入情况、营业税分税目分企业类型收入情况、涉外税收分地区分税种收入情况、涉外税收分地区分项目收入情况、涉外税收分项目分税种收入情况、企业所得税分地区分行业收入情况、企业所得税分项目分企业类型收入情况、个人所得税分地区分项目收入情况
国税收入-全部税收收入	国家税务总局税收收入分税种分产业情况、国家税务总局税收收入分税种分地区情况、国家税务总局税收收入分税种分级次情况、国家税务总局税收收入分税种分企业类型情况、国家税务总局税收收入分地区分产业情况、国家税务总局税收收入分地区分企业类型情况
国税收入-具体税收收入	国家税务总局营业税分地区分企业类型收入情况、国家税务总局资源税分地区分企业类型收入情况、国家税务总局涉外税收分地区分税种收入情况、国家税务总局企业所得税分地区分企业类型收入情况
地税收入-全部税收收入	地方税务局税收收入分税种分产业情况、地方税务局税收收入分税种分地区情况、地方税务局税收收入分税种分级次情况、地方税务局税收收入分税种分企业类型情况、地方税务局税收收入分地区分产业情况、地方税务局税收收入分地区分企业类型情况
地税收入-具体税收收入	地方税务局营业税分地区分企业类型收入情况、地方税务局资源税分地区分企业类型收入情况、地方税务局涉外税收分地区分税种收入情况、地方税务局企业所得税分地区分企业类型收入情况
地税收入详细信息	安徽省、北京市、福建省、广东省、广西壮族自治区、黑龙江省、湖北省、山东省、云南省、浙江省、河南省、湖南省、四川省、陕西省、云南省
上市公司税收信息	上市公司税收信息
税务官员信息-国家税务总局总局官员信息	国家税务总局总局领导信息、国家税务总局部门领导信息、国家税务总局从业人员基本情况
税务官员信息-地方税务局官员信息	各省地税部门领导信息、地方税务局从业人员基本情况

5.数据范围

数据模块	数据起始时间
中国税收收入信息	2001 年
地税收入详细信息	1994 年
上市公司税收信息	2001 年
税务官员信息	2004 年

(八) 中国商品交易市场统计数据库-CEMD

1. 数据库简介

中国商品交易市场统计数据库 (*Statistical Database of China Commodity Exchange Market, CEMD*) 是基于国家统计局数据和商品交易市场统计年鉴数据整理开发的专业数据库。*CEMD* 收集整理了中国近十几年各地区交易市场的成交额、摊位数、市场数量等数据内容, 为学术界进行相关研究提供了重要数据支持。*CEMD* 包含五个数据模块, 分别为: 市场综合; 省、自治区、直辖市; 东中西部及东北地区; 36 城市; 三大地带。各部分都按照经营状态、经营环境、经营方式、市场类别等详细展现了全国不同区域的市场交易详细情况。

2. 数据库特色

- 数据来源权威。

CEMD 数据来源于国家统计局以及历年商品交易市场统计年鉴, 指标覆盖完整, 数据权威。

- 结构的一致性。

CEMD 在数据整理后的模块结构与表格形式上与最新几年的年鉴格式保持了一致, 便于进行数据比对。

3. 典型数据举例

市场综合_历年商品交易市场情况

会计年度	市场数(个)	年末出租摊位数(个)	营业面积(平方米)	成交额(万元)
2011	5075	3334787	262345039	820172688
2012	5194	3494122	278993712	930237663
2013	5089	3488170	288683344	983651254
2014	5023	3534757	295679201	1003099019
2015	4952	3468638	300657328	1001337865
2016	4861	3457899	300233736	1021396989

4. 数据库内容

数据模块	数据表
市场综合	历年商品交易市场情况; 历年商品交易市场变动; 市场综合商品交易市场总体情况; 市场综合商品交易市场情况 (按营业状态分); 市场综合商品交易市场情况 (按经营方式分); 市场综合商品交易市场情况 (按经营环境分); 市场综合商品交易市场成交情况 (按摊位分); 市场综合商品交易市场成交情况 (按摊位与经营状态分); 市场综合商品交易市场成交情况 (按摊位与经营方式分); 市场综合商品交易市场成交情况 (按摊位与经营环境分)
省、自治区、直辖市	省、自治区、直辖市商品交易市场总体情况; 省、自治区、直辖市商品交易市场情况 (按市场类别分); 省、自治区、直辖市商品交易市场情况 (按营业状态分); 省、自治区、直辖市商品交易市场情况 (按经营方式分); 省、自治区、直辖市商品交易市场情况 (按经营环境分); 省、自治区、直辖市商品交易市场成交情况 (按摊位分)

数据模块	数据表
东中西部及东北地区	东中西部及东北地区商品交易市场总体情况；东中西部及东北地区商品交易市场情况(按市场类别分)；东中西部及东北地区商品交易市场情况（按营业状态分）；东中西部及东北地区商品交易市场情况（按经营方式分）；东中西部及东北地区商品交易市场情况（按经营环境分）；东中西部及东北地区商品交易市场成交情况（按摊位分）
36 城市	36 城市商品交易市场总体情况；36 城市商品交易市场情况(按市场类别分)；36 城市商品交易市场情况（按营业状态分）；36 城市商品交易市场情况（按经营方式分）；36 城市商品交易市场情况（按经营环境分）；36 城市商品交易市场成交情况（按摊位分）
三大地带	三大地带商品交易市场总体情况；三大地带商品交易市场情况(按市场类别分)；三大地带商品交易市场情况（按营业状态分）；三大地带商品交易市场情况（按经营方式分）；三大地带商品交易市场情况（按经营环境分）；三大地带商品交易市场成交情况（按摊位分）

5.数据范围

数据模块	数据起始时间
全部模块	2000 年

(九) 中国人口和就业统计数据库-PESD

1.数据库简介

中国人口和就业统计数据库 (*Chinese Population and Employment Statistics Database, PESD*) 收集了全国和各省、自治区、直辖市人口就业统计的主要数据,同时收录了世界部分国家和地区的相关数据,全面反映了我国人口和就业状况。**PESD** 数据主要来源于国家统计局发布的《中国人口和就业统计年鉴》,根据年鉴中的数据内容,将其归类为六大模块:综合数据、劳动力抽样调查、城镇单位就业人员、全国户籍统计人口、计划生育统计人口、世界人口与就业。

2.数据库特色

- 准确的数据归类

PESD 根据《中国人口和就业统计年鉴》各表数据内容,准确地将其归类为六大模块,直观反映了我国人口和就业状况。

- 简洁的数据结构

PESD 整合了多个年度的相同数据到一张表格中,数据结构简洁,下载方便,节约了数据搜寻和整理时间。

3.典型数据举例

综合数据_人口_按性别分人口数

会计年度	总人口 (年末)	男性人口数	男性人口比重	女性人口数	女性人口 比重
2002	128453	66115	51.47	62338	48.53
2003	129227	66556	51.50	62671	48.50
2004	129988	66976	51.52	63012	48.48
2005	130756	67375	51.53	63381	48.47
2006	131448	67728	51.52	63720	48.48

4.数据库内容

数据模块	数据表
综合数据	按性别分人口数;人口年龄结构和抚养比;各地区人口出生率、死亡率和自然增长率;流动人口数;各地区人口平均预期寿命;全国历年人口密度;就业基本情况;分城乡就业人员年末人数;分产业就业人员年末人数;城镇登记失业人数及失业率(年末数);分地区城镇登记失业情况(年末数);分行业城镇非私营单位就业人员年末人数;分登记注册类型城镇非私营单位就业人员年末人数;分地区按行业分私营企业和个体就业人数;分地区按行业分城镇私营企业和个体就业人数;分地区私营企业就业人数;分地区个体就业人数;分行业城镇非私营单位女性就业人员年末人数;分登记注册类型城镇非私营单位就业人员平均工资;分登记注册类型城镇非私营单位就业人员平均工资指数;分行业城镇非私营单位就业人员平均工资;分地区城镇非私营单位就业人员平均工资;分地区按行业分城镇单位类别就业人员平均工资

数据模块	数据表
劳动力抽样调查	全国按就业身份、性别分的就业人员年龄构成；城镇按就业身份、性别分的就业人员年龄构成；城镇按行业、性别分的就业人员年龄构成；城镇按职业、性别分的就业人员的年龄构成；城镇按未工作原因、性别分的失业人员年龄构成；全国分地区就业人员受教育程度构成；全国分地区男性就业人员受教育程度构成；全国分地区女性就业人员受教育程度构成；全国按年龄、性别分的就业人员受教育程度构成；全国接受教育程度、性别分的就业人员年龄构成；全国按行业、性别分的就业人员受教育程度构成；全国按职业、性别分的就业人员受教育程度构成；全国接受教育程度、性别分的就业人员职业构成；全国接受教育程度、性别分的就业人员就业身份构成；全国按就业身份、性别分的就业人员受教育程度构成；城镇按受教育程度、性别分的就业人员就业身份构成；城镇按就业身份、性别分的就业人员受教育程度构成；城镇按受教育程度、性别分的就业人员行业构成；城镇按行业、性别分的就业人员受教育程度构成……
城镇单位就业人员	各地区分行业国有单位就业人员数；各地区分行业城镇集体单位就业人员数；各地区分行业其他单位就业人员数；各地区国有单位就业人员数；各地区城镇集体单位就业人员数；各地区其他行业就业人员数
全国户籍统计人口	各地区总户数、总人口；各地区市总户数、总人口；各地区镇总户数、总人口；各地区县总户数、总人口；各地区非农业、农业人口；各地区市非农业、农业人口；各地区镇非农业、农业人口；各地区县非农业、农业人口；各地区区县人口数；按总人口排序的市及人口数；按非农业人口排序的市及人口数
计划生育统计人口	各地区分孩次出生政策符合率与上年同期比较；各地区已婚育龄妇女领证情况、避孕率与上年同期比较；各地区采用各种节育措施人数；各地区采用各种节育措施人数与上年同期比较；各地区采取各种避孕措施分布；各地区采取各种避孕措施分布与上年同期对比；各地区实施各种节育手术的例数；各地区实施各种节育手术例数与上年同期对比
世界人口与就业	各国人口出生率、死亡率、自然增长率；各国人口年龄构成；各国人口指标；人口主要指标(粗)；人口主要指标；面积和人口主要指标；各国就业与失业情况；劳动力和就业状况；劳动人口及失业状况；经济活动人口及失业状况；各国失业人数；失业人数和失业率；各国按三次产业分就业人员构成；按行业划分的就业人数；按每月就业收入划分的就业人数；按行业划分的就业人口和月工作收入中位数；按行业划分督导级(不包括经理级与专业雇员)及以下雇员的工资指数；各国消费价格指数；分类别的消费价格指数；居民消费价格分类指数

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

区域经济研究系列

(一) 城市统计数据库-CCSD

1. 数据库简介

中国城市统计数据库 (*Chinese City Statistics Database, CCSD*) 涵盖了我国 600 多个城市 (含地级以上城市和县级城市) 社会经济发展方面的主要统计数据。*CCSD* 数据主要来源于国家统计局发布的《中国城市统计年鉴》，根据年鉴中的数据内容，将其归类为 11 个大模块，分别是行政区划、人口就业、公共设施、土地农业、教育教学、工业生产、财政投资、环境环保、经济金融、邮政电信、其他。这些模块既包括了地级以上城市的统计资料，又囊括县级城市统计资料。根据学术研究需要，*CCSD* 在横向上从数据类别对数据进行归纳总结，在纵向上从时间序列对数据进行多年度整合，极大满足了学术研究的需求。

2. 数据库特色

- 准确的数据归类

CCSD 根据《中国城市统计年鉴》各个表格数据内容，准确地将其归类为 11 个大模块，直观体现了城市社会经济发展方面的主要统计数据。

- 简洁的数据结构

CCSD 不仅将地级市和县级市数据分别列示，并整合了多个年度的相同数据到一张表格中，数据结构简洁，下载方便，节约了数据搜寻和整理时间。

3. 典型数据举例

在岗职工人数及工资状况_地级市

省份名称	城市名称	会计年度	在岗职工平均人数	在岗职工工资总额	职工平均工资	统计范围
河北省	石家庄市	1999	96	758877	7938	全市
河北省	唐山市	1999	88	640579	7288	全市
河北省	秦皇岛市	1999	33	272117	8314	全市
河北省	邯郸市	1999	69	425401	6169	全市
河北省	邢台市	1999	41	255322	6255	全市

4. 数据库内容

数据模块	数据表
行政区划	行政区划和区域分布; 行政区划土地面积_县级市
人口就业	人口状况_地级市; 人口状况_县级市; 劳动力与就业_地级市; 劳动力与就业_县级市; 在岗职工人数及工资状况_地级市; 工资水平_县级市; 按产业划分的年末单位从业人员_地级市; 按产业划分的年末单位从业人员构成(比重)_地级市; 按行业分组的单位从业人员_地级市; 社会保障主要指标_地级市
公共设施	体育场馆、剧场、影剧院数及公共图书馆图书藏量_县级市; 剧场、影剧院数及公共图书馆图书藏量_地级市; 医院、卫生院数、床位数和医生数_地级市; 医院、卫

数据模块	数据表
公共设施	生院床位数及医护人员数_县级市; 道路面积及公共汽车、出租车拥有情况_地级市
土地农业	主要农产品产量_地级市; 主要农产品产量_县级市; 人均农产品占有量_地级市; 城市建设用地状况_市辖区
教育教学	学校数量_地级市; 学校数量_县级市; 专任教师数_地级市; 专任教师数_县级市; 在校学生数_地级市; 在校学生数_县级市
工业生产	主要财务指标_地级市; 工业企业数_地级市; 规模以上工业总产值_地级市; 规模以上工业企业情况_县级市
财政投资	公共财政收支_地级市; 公共财政收支_县级市; 利用外资情况_地级市; 固定资产投资情况_地级市; 固定资产投资情况_县级市
环境环保	工业固体废物综合利用率和污水及生活垃圾处理率_全市; 工业废水排放和二氧化硫产生及排放量_全市; 工业烟(粉)尘产生及排放量_全市; 绿地面积及建成区绿化覆盖面积_市辖区; 二氧化硫、氮氧化物及烟(粉)尘排放量和污水处理情况_县级市
经济金融	地区生产总值_地级市; 地区生产总值_县级市; 地区生产总值构成_地级市; 年末金融机构存贷款余额_地级市; 年末金融机构存贷款余额_县级市; 社会消费品零售总额及批发零售贸易业情况_地级市; 社会消费品零售总额及批发零售贸易业情况_县级市
邮政电信	电话及互联网用户数_地级市; 电话及互联网用户数_县级市; 邮政局数量及收入_地级市
其他	按运输方式分类的客运量_全市; 按运输方式分类的货运量_全市; 供水、用水及用电情况_地级市; 煤气及液化石油气供应及利用情况_地级市; 全社会用电情况_县级市

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

产业经济研究系列

(一) 产业政策数据库-IPRD

1. 数据库简介

产业政策研究数据库 (*Industrial Policy Research Database, IPRD*) 是基于我国中央政府和各省级政府每五年发布一次的国民经济和社会发展五年规划纲要, 并结合已有产业政策方面的学术研究成果, 整理归纳“五年规划”中提及的相关产业制作而成的数据库。*IPRD* 不仅提取了五年规划纲要中提及的相关产业及规划内容, 还依据文中对提及产业的描述语气, 归纳了国家及各省对提及的相关产业的政策态度, 分为鼓励、中性和抑制三类。此外, 还根据证监会发布的“上市公司行业分类指引 2001”对五年规划大纲中提及的产业进行标准化分类。

2. 数据库特色

- 丰富的产业政策相关字段

IPRD 不仅提取了五年规划中提及的相关产业及规划内容, 还对这些产业进行了标准化分类, 并对产业政策态度进行了区分, 为该领域研究提供更多直观可用的数据。

- 囊括中央和地方产业政策

除了主流研究使用的国家产业政策, *IPRD* 还提供各省级政府产业政策, 进一步拓展了研究范围。

3. 典型数据举例

中央产业政策

五年规划	批准日期	规划提及产业	行业分类 2001	行业代码 2001	政策态度	是否重点支持产业
十五计划	2001-03-15	旅游业	旅游业	K34	鼓励	是
十五计划	2001-03-15	零售业和餐饮业	零售业; 餐饮业	H11; K30	鼓励	否
十五计划	2001-03-15	职业培训产业	专业、科研服务业	K20	鼓励	否
十一五规划	2006-03-14	建材业	非金属矿物制品业	C61	抑制	否

4. 数据库内容

数据模块	数据表	数据字段
产业政策	中央产业政策	五年规划、批准日期、规划提及产业、规划内容、所属行业大类、行业分类 2001、行业代码 2001、政策态度、是否重点支持产业
	省级产业政策	所属省份、五年规划、批准日期、规划提及产业、规划内容、所属行业大类、行业分类 2001、行业代码 2001、政策态度、是否重点支持产业、是否国家重点支持产业

5. 数据范围

数据模块	数据起始时间
中央产业政策	2001 年
省级产业政策	2006 年

(二) 中国房地产统计数据库-RES D

1.数据库简介

中国房地产统计数据库 (*Real Estate Statistics Database, RES D*) 是基于国家统计局数据和中國房地產统计年鉴数据整理开发的专业数据库。**RES D** 收集整理了中国近十几年来的房地产开发企业的基本信息、房地产企业经营财务数据以及全国 35 个大中城市的房地产经济统计数据等, 提供了全面的数据和丰富的字段, 为学术界对房地产领域的研究提供了重要数据支持。**RES D** 包含五个数据模块, 分别为: 房地产开发企业基本概况, 房地产开发投资与资金状况, 房屋开竣工面积、商品房销售及土地情况, 房地产开发企业主要财务状况, 35 个大中城市房地产综合情况。其中, 为了重点反映全国大、中城市房地产开发企业运行状况, 在模块“35 个大中城市房地产综合情况”中包含了直辖市、省会城市及计划单列市的房地产完成投资、房屋销售、建设、待售、企业主要财务指标等重要指标。

2.数据库特色

- 数据来源权威

RES D 数据来源于国家统计局以及历年房地产统计年鉴, 指标覆盖完整, 数据权威。

- 结构的一致性

RES D 在数据整理后的模块结构与表格形式上与最新几年的年鉴格式保持了一致, 便于进行数据比对。

3.典型数据举例

35 个大中城市按用途分房地产开发企业商品房平均销售价格

地区名称	会计年度	商品房销售平均价格 (元/平方米)	住宅 (元/平方米)	别墅、高档公寓 (元/平方米)	办公楼 (元/平方米)	商业营业用房 (元/平方米)
北京	2006	8280	7375	12181	13555	14965
北京	2007	11553	10661	15740	15152	17585
北京	2008	12418	11648	19541	16554	17148
北京	2009	13799	13224	20385	16857	19091
北京	2010	17782	17151	28680	23413	22452
北京	2011	16852	15518	28543	23702	24920

4.数据库内容

数据模块	数据表
房地产开发企业基本概况	房地产开发企业主要指标完成情况; 各地区按登记注册类型分房地产开发企业个数; 各地区按资质等级分房地产开发企业个数; 各地区按隶属关系分房地产开发企业个数; 各地区按登记注册类型分房地产开发企业从业人数; 各地区按资质等级分房地产开发企业从业人数; 各地区按隶属关系分房地产开发企业从业人员
房地产开发投资与资金状况	各地区房地产开发企业投资规模与完成情况; 各地区按登记注册类型分房地产开发企业计划总投资; 各地区按资质等级分房地产开发企业计划总投资; 各地区按登记注册类型分房地产开发企业完成投资; 各地区按资质等级分房地产开发企业完成投资; 各地区按隶属关系分房地产开发企业完成投资; 各地区按构成分房地产开发企业完成投资; 各地区按用途分的房地产开发企业完成投资; 各地区按控股情况分房

数据模块	数据表
房地产开发投资与资金状况	地产开发企业完成投资；各地区按投资规模分房地产开发企业完成投资；各地区按登记注册类型分房地产开发企业住宅完成投资；各地区按资质等级分房地产开发企业住宅完成投资；各地区按资质等级分房地产开发企业 90 平方米以下住宅完成投资；各地区按资质等级分房地产开发企业 144 平方米以上住宅完成投资；各地区按资质等级分房地产开发企业别墅、高档公寓完成投资；各地区按登记注册类型分房地产开发企业办公楼完成投资；各地区按资质等级分房地产开发企业办公楼完成投资；各地区按登记注册类型分的房地产开发企业商业营业用房完成投资……
房屋开竣工面积、商品房销售及土地情况	各地区按用途分的房地产开发企业房屋施工面积；各地区按资质等级分房地产开发企业房屋施工面积；各地区按资质等级分的房地产开发企业住宅施工面积；各地区按用途分房地产开发企业房屋新开工面积；各地区按资质等级分房地产开发企业房屋新开工面积；各地区按资质等级分房地产开发企业住宅新开工面积；各地区按用途分房地产开发企业房屋竣工面积；各地区按资质等级分房地产开发企业房屋竣工面积；各地区按资质等级分房地产开发企业住宅竣工面积；各地区按资质等级分房地产开发企业 90 平方米及以下住宅竣工面积；各地区按资质等级分房地产开发企业别墅、高档公寓竣工面积；各地区按用途分房地产开发企业不可销售面积；各地区按资质等级分的房地产开发企业不可销售面积；各地区按用途分房地产开发企业住宅竣工套数；各地区按资质等级分房地产开发企业住宅竣工套数；各地区按资质等级分房地产开发企业 90 平方米及以下住宅竣工套数……
房地产开发企业主要财务状况	各地区房地产开发企业主要财务指标；各地区按登记注册类型分房地产开发企业资产总计；各地区按资质等级分房地产开发企业资产总计；各地区按登记注册类型分房地产开发企业负债合计；各地区按资质等级分房地产开发企业负债合计；各地区按登记注册类型分房地产开发企业所有者权益；各地区按资质等级分房地产开发企业所有者权益；各地区按登记注册类型分房地产开发企业营业利润；各地区按资质等级分房地产开发企业营业利润；各地区按登记注册类型分的房地产开发企业利润总额；各地区按资质等级分房地产开发企业利润总额；各地区按登记注册类型分房地产开发企业主营业务收入；各地区按资质等级分房地产开发企业主营业务收入
35 个大中城市房地产综合情况	35 个大中城市按登记注册类型分房地产开发企业个数；35 个大中城市按资质等级分房地产开发企业个数；35 个大中城市按登记注册类型分房地产开发企业从业人数；35 个大中城市按资质等级分房地产开发企业从业人数；35 个大中城市房地产开发企业投资规模与完成情况；35 个大中城市按用途分房地产开发企业完成投资；35 个大中城市房地产开发企业实际到位资金情况；35 个大中城市按用途分房地产开发企业房屋施工面积；35 个大中城市按用途分房地产开发企业房屋新开工面积；35 个大中城市按用途分房地产开发企业房屋竣工面积；35 个大中城市按用途分房地产开发企业房屋竣工价值；35 个大中城市按用途分房地产开发企业住宅竣工套数；35 个大中城市按资质等级分房地产开发企业住宅竣工套数；35 个大中城市按资质等级分房地产开发企业别墅、高档公寓套数；35 个大中城市按用途分房地产开发企业商品房销售面积；35 个大中城市按用途分市房地产开发企业商品房销售额；35 个大中城市按用途分房地产开发企业商品房平均销售价格；35 个大中城市按用途分房地产开发企业房屋出租面积；35 个大中城市按用途分房地产开发企业商品房待售面积；35 个大中城市房地产开发企业主要财务指标

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

（三）中国能源统计数据库-CESD

1.数据库简介

中国能源统计数据库（Chinese Energy Statistics Database, CESD）涵盖了中国多个能源品种、多个行业和城市的能源生产和消费情况数据。CESD 数据主要来源于国家统计局发布的《中国能源统计年鉴》，根据年鉴中的数据内容，将其归类为 5 个大模块，分别是综合、能源建设、能源生产、能源消费、能源平衡表。这些模块既包括能源总量和分能源品种的数据，还包括分行业分地区、分地区分能源品种等数据，在横向上从数据类别对数据进行归纳总结，在纵向上从时间序列对数据进行多年度整合，极大满足了学术研究的需求。

2.数据库特色

- 准确的数据归类和简洁的数据结构。

CESD 根据《中国能源统计年鉴》各个表格数据内容，在整合归纳的基础上，将其准确地归类为 5 个大模块，直观体现了中国能源发展的主要统计数据。

3.典型数据举例

能源生产与消费增长速度

会计年度	能源生产增长速度	能源消费增长速度	能源生产弹性系数	能源消费弹性系数
2011	9.0%	7.3%	0.95	0.77
2012	3.2%	3.9%	0.40	0.49
2013	2.2%	3.7%	0.28	0.47
2014	0.9%	2.1%	0.12	0.29

4.数据库内容

数据模块	数据表
综合	能源生产与消费增长速度；能源经济主要指标；平均每万元国内生产总值能源消费量；能源加工转换效率；人均能源生产和消费量；人均生活用能量；交通运输设备拥有量；主要能源品种进、出口量；主要高耗能产品进、出口量；分地区主要污染物排放情况 1994~2010；分地区主要污染物排放情况 2011~2016
能源建设	国有经济能源工业分地区行业固定资产投资；能源工业分地区分行业投资；能源工业分行业投资
能源生产	一次能源生产量和构成；分地区分品种能源生产供应
能源消费	能源消费总量和构成；工业分行业终端能源消费量(实物量)；工业分行业终端能源消费量(标准量)；分行业分品种能源消费量；分地区分品种能源消费量
能源平衡表	中国能源平衡表(实物量)；中国能源平衡表(标准量)；分品种能源平衡表；分地区能源平衡表

5.数据范围

数据模块	数据起始时间
全部模块	1980 年

（四）中国工业统计数据库-CISD

1.数据库简介

中国工业统计数据库 (*Chinese Industrial Statistics Database, CISD*) 是基于历年中国工业统计年鉴整理而成的数据库, 系统地收录了全国各经济类型、各工业行业和各省、自治区、直辖市等工业经济各方面的统计数据。全库分为三个模块: 综合数据、分行业数据及分地区数据, 主要反映中国工业经济的发展、中国工业经济的行业结构、中国工业经济的地区布局, 可供学者了解和研究我国工业经济发展情况。

2.数据库特色

- 数据整合和归类

CISD 根据《中国工业统计年鉴》各表数据内容, 准确地将其归类为三个大模块, 并对属于类似范围的表格进行了合并, 数据结构更为简洁和直观。

3.典型数据举例

工业统计_工业企业单位数

年度	全国总计(万个)	国有控股(万个)	私营(万个)	外商投资和港澳台商投资(万个)
2011	32.56	1.71	18.06	5.72
2012	34.38	1.79	18.93	5.69
2013	36.98	1.86	20.84	5.74
2014	37.79	1.88	21.38	5.52

4.数据库内容

数据模块	数据表
综合数据	工业企业单位数; 工业企业主要经济指标; 分登记注册类型规模以上工业企业主要经济指标; 分行业规模以上工业企业主要效益指标; 分地区规模以上工业企业主要效益指标; 历年主要工业产品产量
分行业数据	分行业规模以上工业企业主要经济指标; 分行业分企业性质工业企业主要经济指标; 分行业大中型工业企业主要经济指标
分地区数据	分地区规模以上工业企业主要经济指标; 分地区分企业形式工业企业主要经济指标; 分地区分企业性质工业企业主要经济指标; 分地区分企业规模工业企业主要经济指标; 分地区采矿业-合计数; 分地区采矿业主要经济指标; 分地区制造业-合计数; 分地区制造业主要经济指标; 电力、热力、燃气及水生产和供应业; 分地区主要工业产品产量

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

对外经济研究系列

(一) 中国对外直接投资数据库- OFDI

1.数据库简介

中国对外直接投资数据库 (*Outward Foreign Direct Investment Database, OFDI*) 是基于商务部统计公报以及国内对外投资相关新闻报道等信息整理而成的专业数据库。*OFDI* 数据库包括中国境外投资企业、对外直接投资事件、对外直接投资情况 (分国家地区)、对外直接投资情况 (分省市) 四种类型的表格, 不仅分年度展示了中国对世界各地和国家的投资情况, 还列明了重要投资事件的具体信息, 包括投资年份、投资方、交易方、投资事件、投资金额、持股比例、投资方所属行业、交易方所属行业、投资国家或地区等重要字段。

2.数据库特色

- 数据来源权威

OFDI 基于政府部门或公益性组织等统计数据, 来源权威可靠。

- 丰富的统计数据 and 事件信息

OFDI 不仅包括一定时期内中国对各国家和地区直接投资的总体数据, 还列明了该时期内的重要事件具体信息, 为学术研究提供了重要数据支持。

3.典型数据举例

对外直接投资事件

年份	月份	投资方	交易方	投资事件	投资金额	持股比例	投资国家
2016	1	海信集团有限公司	欧洲杯	海信将成为 2016 年欧洲杯的顶级赞助商			法国
2016	1	航天科技控股集团股份有限公司	Hiwinglux S.A.	向益圣国际有限公司发行股份购买 Hiwinglux S.A.100%股权	28.78	100%	美国
2016	1	航天科技控股集团股份有限公司	Naviligh	以现金购买国新国际投资有限公司持有的 Navilight S.a.r.l.股权	21.1	100%	美国
2016	1	海尔集团	General Electric	拟以 54 亿美元购买通用电气及其子公司所持有家电业务资产	5400		美国
2016	1	两家中资公司	阿尔及利亚政府	签约成立合资公司, 共同兴建阿尔及利亚境内最大港口	3300		阿尔及利亚

4.数据库内容

数据模块	数据表	数据字段
中国境外投资企业	中国境外投资企业	境外投资企业 (机构) 名称、境内投资者名称、统计日期、投资国家 (地区)
对外直接投资事件	对外直接投资事件	年份、月份、投资方、交易方、投资事件、投资金额、持股比例、投资方所属行业、交易方所属行业、投资国家 (地区)

数据模块	数据表	数据字段
对外直接投资情况 (分国家地区)	对外直接投资流量情况 (分国家地区)	国家(地区)、序号、年份、投资金额
	对外直接投资存量情况 (分国家地区)	国家(地区)、序号、年份、投资金额
对外直接投资情况 (分省市)	对外直接投资流量情况 (分省市)	地区、序号、年份、投资金额
	对外直接投资存量情况 (分省市)	地区、序号、年份、投资金额

5.数据范围

数据模块	数据起始时间
中国境外投资企业	截面数据
对外直接投资事件	2000年
对外直接投资情况(分国家地区)	2003年
对外直接投资情况(分省市)	2003年

（二）一带一路研究数据库-BRRD

1.数据库简介

一带一路研究数据库 (*The Belt and Road Research Database, BRRD*) 是根据我国提出的“一带一路”发展战略, 充分调查和研究了一带一路沿线 65 个国家的政治、经济、人口、环境、社会等信息开发而成的专业数据库。为了给学者提供准确、详细、权威的数据字段, *BRRD* 借鉴了多方数据, 例如, 世界银行、国际货币基金组织、联合国、国际劳工组织等多个机构的独有数据, 涵盖了财政税收、对外贸易、发展援助、国际收支、基础设施、教育教学、经济金融、企业经营调查、人工就业、土地农业、外部债务、医疗卫生、资源环境等 13 个重要方面。*BRRD* 根据 65 个国家所在区域分为七个子库, 分别为: 东南亚 11 国、东亚-蒙古、南亚 8 国、中亚 5 国、西亚 17 国、中东欧 16 国、独联体 7 国。

东南亚 11 国包括: 越南、老挝、柬埔寨、泰国、缅甸、马来西亚、新加坡、印度尼西亚、文莱、菲律宾、东帝汶; 南亚 8 国包括: 印度、巴基斯坦、阿富汗、孟加拉国、斯里兰卡、尼泊尔、不丹和马尔代夫; 中亚 5 国包括: 土库曼斯坦、吉尔吉斯斯坦、乌兹别克斯坦、塔吉克斯坦、哈萨克斯坦; 西亚 17 国包括: 伊朗、伊拉克、土耳其、叙利亚、约旦、黎巴嫩、以色列、沙特阿拉伯、也门、阿曼、阿联酋、卡塔尔、科威特、巴林、希腊、塞浦路斯和埃及的西奈半岛; 中东欧 16 国: 波兰、捷克共和国、斯洛伐克、匈牙利、斯洛文尼亚、克罗地亚、罗马尼亚、保加利亚、塞尔维亚、黑山、马其顿、波黑、阿尔巴尼亚、爱沙尼亚、立陶宛和拉脱维亚。独联体 7 国包括: 俄罗斯、乌克兰、白俄罗斯、格鲁吉亚、阿塞拜疆、亚美尼亚和摩尔多瓦。

2.数据库特色

- 1000 多个特色指标

BRRD 囊括了 1000 多个特色指标, 从多个维度反映了一带一路各国的基本情况。

- 数据来源权威

BRRD 整合了来自多个渠道的各国数据, 包括世界银行、国际货币基金组织、联合国、国际劳工组织等多个机构的独有数据, 具有较强的权威性。

3.典型数据举例

菲律宾_财政税收

国家名称	会计年度	指标名称	数据值
菲律宾	2015	一般政府最终消费支出占 GDP 百分比	10.96
菲律宾	2015	一般政府最终消费支出 (年增长率)	7.766
菲律宾	2015	企业与税务官员见面的平均次数	1
菲律宾	2015	企业其他应缴税种 (占商业利润的百分比)	13.9
菲律宾	2015	利润税占商业利润的百分比	20.3
菲律宾	2015	劳动税和缴费占商业利润的百分比	8.7
菲律宾	2015	筹纳税所需时间 (小时)	191
菲律宾	2015	纳税项 (数量)	36

4.数据库内容

数据模块
财政税收
对外贸易
发展援助
国际收支
基础设施
教育教学
经济金融
企业经营调查
人口就业
土地农业
外部债务
医疗卫生
资源环境

5.数据范围

数据模块	数据起始时间
全部模块	1960 年

（三）海外并购数据库-COMA

1.数据库简介

海外并购数据库 (*Overseas M&A Database, COMA*) 是基于公司公告、新闻报道、商务部官方网站等数据建立的专业数据库。*COMA* 包括 1988 年以来中国公司 (包括上市公司和非上市公司) 进行海外并购的事件数据, 主要包括四大表格: 并购事件、并购方、被并购方以及被并购方财务信息。

2.数据库特色

- 整合海外并购事件, 提供丰富字段

COMA 提供了较为全面的海外并购事件信息, 并囊括研究所用到的常用字段, 弥补了该领域研究数据匮乏的缺陷。

3.典型数据举例

海外并购事件

并购 ID	首次公告日	并购生效日	转让股权比例 (%)	转让后并购方持股比例 (%)	转让股权估值 (美元)	现金支付比例 (%)
331144	1992/12/3	1992/12/3	30.1	30.1	21071000	100
331283	1998/3/4	1998/7/16	85	85	20000000	100
331422	2001/9/12	2001/10/1	29	29	17372000	46.05
333349	2013/5/29	2013/9/26	100	100	4752384000	100

4.数据库内容

数据模块	数据表	数据字段
海外并购	并购事件	并购 ID、并购方名称、被并购方名称、首次公告日、并购生效日、转让股权比例 (%)、转让后并购方持股比例 (%)、转让股权估值、货币类型、现金支付比例 (%)、竞买方数量、完成状态
	并购方	并购 ID、并购方名称、并购方英文名称、并购方企业性质、并购方所在交易所、并购方股票代码、并购方上市状态、并购方所在地区、并购方所属行业
	被并购方	并购 ID、被并购方名称、被并购方所在国、被并购方所属行业、被并购方企业性质
	被并购方财务信息	并购 ID、被并购方名称、首次公告日、并购生效日、货币类型、被并购方资产负债率、被并购方净销售额、被并购方 EBIT、被并购方净资产、被并购方总资产、被并购方 EPS、被并购方并购公告后一周股价、被并购方并购公告后 180 天股价、被并购方并购公告后四周股价、被并购方并购公告后 90 天股价、被并购方并购公告当天股价、被并购方并购公告前一天股价、被并购方并购公告前一周股价、被并购方并购公告前四周股价

5.数据范围

数据模块	数据起始时间
全部模块	截面数据

三、CNRDS 基础数据库

基础数据是指在研究过程中必须用到的一系列基本数据，包括上市公司股价数据、财务数据、股权结构、公司治理数据等。这些数据为研究提供了各种原始资料，其质量高低可能会影响到研究结果的可靠性。因此，在有关中国问题的研究过程中，基础数据的涵盖范围、数据质量是学者关注的一个重点。**CNRDS** 参照国际标准并结合中国资本市场数据结构，基于 **CNRDS** 基础数据库整合了经济管理领域学术研究过程中可能用到的大部分研究基础数据。参阅各领域重点文献，依托已有的学术研究成果，**CNRDS** 平台根据研究需求对各类数据进行整合和重构，归纳了以下四个系列的基础数据：

上市公司股票基础数据 主要包括中国股票市场相关数据，例如股价数据、融资融券数据、股权融资数据。

上市公司财务基础数据 主要包括中国上市公司相关财务数据，例如财务报表数据、盈利预测数据、财务指标数据等。

上市公司治理基础数据 包括与公司治理相关的上市公司数据，例如关联交易、内部人交易、高管薪酬等重要数据。

经济研究基础数据 主要涵盖了经济研究所需要的基本研究数据，例如来自国家统计局的基本宏观经济数据和区域经济数据等。

CNRDS 基础数据库从数据生产、数据整理和数据呈现均有一套严格的标准和程序。首先，基础数据库的字段构成必须由相关领域的专家学者讨论通过；其次，数据库体系必须符合学术研究需求和学者浏览、下载数据的习惯；最后，基础数据的数据质量必须经过多次核查方可最终上线。

CNRDS 基础数据库具有以下三个特点：

- 数据库设计和布局合理，契合使用习惯
- 数据严格审核，数据质量较高
- 多维度基础数据，满足不同研究需求

上市公司股票基础数据

(一) 中国上市公司股价数据库-CNSP

1. 数据库简介

中国上市公司股价数据库 (*Securities Prices of Chinese Public Firms, CNSP*) 是以研究为目的而设计的, 对来自上海证券交易所和深圳证券交易所中国上市公司全部交易数据进行全面深入研究, 旨在提供完整精确的股价信息。本数据库提供准确、真实、全面的中国上市公司股价信息, 并结合中国股票市场自身特点进行校正。目前, 本数据库主要包含了两个数据模块, 即个股回报率和市场回报率。

个股回报率 分为三个子模块: 个股日回报率、个股周回报率、个股月回报率。该模块主要提供中国上市公司来源于权威信息源的股票信息, 严格按照 *CRSP* 国际标准数据库的调整技术计算个股回报率, 以日、周、月进行详细划分, 有利于对任一股票的行情做出合理判断。

市场回报率 分为三个子模块: 市场日回报率、市场周回报率、市场月回报率。该模块主要提供中国上市公司来源于权威信息源的股票信息, 严格按照 *CRSP* 国际标准数据库的调整技术计算市场回报率, 以日、周、月进行详细划分, 有利于对整个市场股票的行情做出合理判断。

2. 数据库特色

- 字段设计严格参照权威数据库

CNSP 严格按照 *CRSP* 国际标准数据库的调整技术计算个股回报率和市场回报率, 以日、周、月进行详细划分, 有利于对任一股票的行情做出合理判断。

3. 典型数据举例

个股日回报率

股票代码	公司简称	交易时间	日开盘价 (元)	日最高价 (元)	日最低价 (元)	日收盘价 (元)	日涨跌幅 (%)
000001	平安银行	2012-12-18	15.15	15.47	15.04	15.25	0.1971
000001	平安银行	2012-12-19	15.2	15.34	15.14	15.18	-0.459
000001	平安银行	2012-12-20	15.1	15.15	14.78	15.11	-0.4611
000001	平安银行	2012-12-21	15.08	15.14	14.85	14.87	-1.5884
000001	平安银行	2012-12-24	14.84	15.21	14.82	14.92	0.3362
000002	万科 A	2012-12-25	9.55	10.26	9.5	10.12	5.9686
000002	万科 A	2012-12-31	10.12	10.12	10.12	10.12	0

4. 数据库内容

数据模块	数据表	数据字段
个股回报率	个股日回报率	股票代码、公司简称、交易时间、日开盘价、日最高价、日最低价、日收盘价、前日收盘价、日涨跌幅 (%)、复权日收盘价、前复权日收盘价、复权日涨跌幅 (%)、日成交量、日成交金额、日总股本、日流通股本、日总市值、日流通市值、日换手率 (%)

数据模块	数据表	数据字段
个股回报率	个股周回报率	股票代码、公司简称、交易时间、周开盘价、周最高价、周最低价、周收盘价、前周收盘价、周涨跌幅(%)、复权周收盘价、前复权周收盘价、复权周涨跌幅(%)、周成交量、周成交金额、周总股本、周流通股本、周总市值、周流通市值、周换手率(%)
	个股月回报率	股票代码、公司简称、交易时间、月开盘价、月最高价、月最低价、月收盘价、前月收盘价、月涨跌幅(%)、复权月收盘价、前复权月收盘价、复权月涨跌幅(%)、月成交量、月成交金额、月总股本、月流通股本、月总市值、月流通市值、月换手率(%)
市场回报率	市场日回报率	市场类型编码、市场类型名称、交易时间、日开盘价、日最高价、日最低价、日收盘价、前日收盘价、日市场涨跌幅(%)、日成交量、日成交金额
	市场周回报率	市场类型编码、市场类型名称、交易时间、周开盘价、周最高价、周最低价、周收盘价、前周收盘价、周市场涨跌幅(%)、周成交量、周成交金额
	市场月回报率	市场类型编码、市场类型名称、交易时间、月开盘价、月最高价、月最低价、月收盘价、前月收盘价、月市场涨跌幅(%)、月成交量、月成交金额

5.数据范围

数据模块	数据起始时间
全部模块	1990年

(二) 中国股票异常交易数据库-CAST

1.数据库简介

中国股票异常交易数据库 (*Chinese Abnormal Stock Trading Database, CAST*) 是以研究为目的而设计的专业学术研究数据库,旨在为国内外进行中国股票市场交易情况研究的学者与研究人员提供专业、准确的学术研究数据,本数据库收集整理了2006年以来的中国股票市场异常交易的数据情况。

本数据包含两个数据模块,分别是:交易公开信息列表与交易公开信息明细共两张表格,22个字段内容,其中包括进行买卖行为的交易方以及交易成交的金额与数量情况。

2.数据库特色

- 唯一提供该类数据

CAST 根据交易所数据整合而成,是目前市场上唯一提供此类数据的数据库。

3.典型数据举例

上市公司异常交易公开信息列表

股票代码	公司简称	公告日期	成交金额(元)	成交量(股/份)	披露原因
002716	金贵银业	2015/11/20	83603794	1132146171	日换手率达到 29.29%
002722	金轮股份	2015/11/20	13167915	499729705	日换手率达到 28.79%
002741	光华科技	2015/11/20	34627238	1490742790	日换手率达到 38.47%
002751	易尚展示	2015/11/20	8902984	1065388112	日价格振幅达到 16.22%

4.数据库内容

数据模块	数据表	数据字段
交易公开信息列表	交易公开信息列表	股票代码、公司简称、公告日期、偏离值%、成交金额(元)、成交量(股/份)、披露原因、振幅%、换手率%、累计偏离值%、累计成交量、累计成交金额(万元)、异常期间、累计换手率%、成交占比%
交易公开信息明细	交易公开信息明细	股票代码、公司简称、公告日期、买/卖、会员营业部名称、买入金额(元)、卖出金额(元)

5.数据范围

数据模块	数据起始时间
全部模块	2006年

(三) 中国特殊处理股票数据库-CSTS

1.数据库简介

中国特殊处理股票数据库 (*Chinese Special Treatment Stock Database, CSTS*) 收集整理了 1998 年以来中国沪深两市受到特殊处理的股票数据。数据库包括股票代码、公司简称、特殊处理前的公司简称、特殊处理后的公司简称、处理类别、处理类别编号、信息发布日期、变动原因、以及处理实施日期在内一共 9 个字段。

2.数据库特色

- 字段丰富内容完整

CSTS 收录了非常完整的特殊处理股票样本以及全面的字段数据。包括了详细的变动类别和明确的变动原因以及完整的日期信息。

3.典型数据举例

上市公司特殊处理内容

股票代码	公司简称	处理前的公司简称	处理后的公司简称	处理类别	信息发布日	变动原因	处理实施日
000005	世纪星源	世纪星源	ST 星源	被 ST	2003-04-15	两年亏损	2003-04-16
000005	世纪星源	ST 星源	*ST 星源	ST 变为 *ST	2003-04-30	两年亏损	2003-05-12
000005	世纪星源	*ST 星源	ST 星源	*ST 变为 ST	2004-05-17	两年亏损	2004-05-18
000005	世纪星源	ST 星源	世纪星源	撤销 ST	2008-06-24	两年亏损	2008-06-25
000007	全新好	深达声 A	ST 达声	被 ST	2001-04-18	两年亏损	2001-04-19
000007	全新好	ST 达声	深达声 A	撤销 ST	2002-04-11	恢复盈利	2002-04-12

4.数据库内容

数据模块	数据表	数据字段
特殊处理内容	特殊处理内容	股票代码、公司简称、特殊处理前的公司简称、特殊处理后的公司简称、处理类别、处理类别编号、信息发布日期、变动原因、处理实施日期

5.数据范围

数据模块	数据起始时间
全部模块	1998 年

（四）中国融资融券数据库-CMTD

1.数据库简介

中国融资融券数据库 (*Chinese Margin Trading Database, CMTD*) 是基于对我国证券市场融资融券行为的分析而建立的专业数据库。该库不仅包括个股融资融券的详细信息, 还包括市场统计、券商统计、标的物 and 担保物等多方统计信息。

CMTD 包含个券交易明细、个券交易市场汇总、券商交易信息、标的物 and 担保物、市场交易统计 5 个模块, 7 张表, 共 50 多个字段。其中, 个券交易明细包括每天各个证券的融资融券具体情况; 个券交易市场汇总是每天不同交易所融资融券情况的汇总; 券商交易信息统计了每个月每个券商的融资融券情况; 标的物 and 担保物列示了融资融券过程中涉及的标的物 and 担保物信息; 最后, 该数据库还统计了沪深两市的整体融资情况和融券情况。*CMTD* 数据库包括 2010 年至今沪深交易所的所有融资融券信息。

2.数据库特色

- 全面

该库不仅包括个股融资融券的详细信息, 还包括市场统计、券商统计、标的物 and 担保物等多方统计信息。

3.典型数据举例

融资融券个券交易明细

股票代码	统计日期	融资余额 (元)	融资买入额 (元)	融资偿还额 (元)	融券余量 (股)	融券卖出量 (股)	融券偿还量 (股)
000060	2011-11-11	266180162	4615081	5438343	148349	76200	72700
000060	2011-11-14	268433670	7944975	5691465	524163	581956	206142
000060	2011-11-15	266133661	5768534	8068537	126445	87200	484918
000060	2011-11-16	262340073	8282910	12076501	121629	79700	84516
000060	2011-11-17	263635801	5726912	4431191	132200	96866	86295
000060	2011-11-18	255632669	1990704	9993843	77513	35400	90087
000060	2011-11-21	246946076	2385813	11072409	152766	286204	210951
000060	2011-11-22	235041264	1462217	13367033	90439	143999	206326

4.数据库内容

数据模块	数据表	数据字段
个券交易明细	个券交易明细	股票代码、公司简称、统计日期、融资余额、融资买入额、融资偿还额、融券余量、融券卖出量、融券偿还量
个券交易市场汇总	个券交易市场汇总	统计日期、交易所、融资余额、融资买入额、融资偿还额、融券余量、融券卖出量、融券偿还量
券商交易信息	券商交易信息	统计日期、机构简称、期末融资融券余额、期末融资余额、期间累计融资买入额、期间累计卖券还款额、期末融券余额、期间累计融券卖出额、期间累计买券还券额

数据模块	数据表	数据字段
标的物 and 担保物	标的物信息	股票代码、公司简称、标的种类、生效起始日期
	担保物信息	股票代码、公司简称、生效日期、折算率(%)、交易所
市场交易统计	融资规模分析	统计日期、融资余额、融资余额占流通市值(%)、融资余量、融资交易买入额、融资交易买入额占 A 股成交额(%)、融资交易偿还额、融资买入个股数
	融券规模分析	统计日期、融券余额、融券余额占 A 股流通市值(%)、融券余量、融券交易卖出额、融券交易卖出额占 A 股成交额(%)、融券交易偿还额、融券买入个股数

5. 数据范围

数据模块	数据起始时间
全部模块	2010 年

（五）中国公司 IPO 综合数据库-CIPO

1.数据库简介

IPO 综合数据库 (*IPO Integrated Database of Chinese Companies, CIPO*) 是基于对公司 IPO 上市全过程的深入考察和了解建立的详细、准确和全面的公司 IPO 研究专业数据库。该库几乎涵盖了中国公司上市前、上市时和上市后的全部活动和参与者信息，不仅包括已经成功上市的 IPO 公司情况，还涉及多次报审以及审核未通过的公司。

2.数据库特色

- 全面

CIPO 不仅包括已经 IPO 成功的公司数据，还包括审核未通过的数据，与 IPO 密切相关的证券承销商以及募集资金的使用和变更情况等，全面覆盖了 IPO 过程中的所有重要事项和重要参与者，

3.典型数据举例

IPO 多次报审公司

股票代码	公司全称	审核会议日期	审核结果	主承销商	预计发行股数 (万股)	预计募集资金 (万元)
002279	北京久其软件股份有限公司	2008-07-30	通过	申万宏源	1530.00	21848.10
002281	武汉光迅科技股份有限公司	2008-08-04	通过	广发证券	4000.00	32792.00
002288	广东超华科技股份有限公司	2008-08-13	通过	南京证券	2200.00	14322.00
002296	河南辉煌科技股份有限公司	2008-08-22	通过	中德证券	1550.00	13827.00

4.数据库内容

数据模块	数据表	数据字段
IPO 报审情况	IPO 多次报审公司	股票代码、公司全称、省份、审核会议日期、上市板、审核结果、审核委员、主承销商、预计发行股数、预计募集资金、交易所
	IPO 审核通过公司	公司全称、上市板、审核结果、预计发行股数、预计募集资金、会议年度、会议届次、会议日期、审核委员、主承销商、省份、注册资本、主要产品及业务、交易所
	IPO 审核未通过公司	股票代码、公司全称、省份、审核会议日期、上市板、审核结果、审核结果说明、审核委员、主承销商、预计发行股数、预计募集资金、交易所
IPO 公司信息	公司基本信息	股票代码、公司简称、行业代码、行业名称、公司全称、成立日期、招股开始日期、上市日期、注册地、注册地所在地、招股时注册资本
	IPO 前资产负债表	股票代码、招股开始日期、统计日期、货币资金、交易性金融资产、应收票据、应收账款、其他应收款、预付账款、应收股利、应收利息、存货、消耗性生物资产、待摊费用、划分为持有待售的资产、一年内到期的非流动资产、结算备付金、拆出资金、应收保费、应收分保账款、应收分保合同准备金、买入返售金融资产、其他流动资产、流动资产合计、可供出售金融资产、持有至到期投资、投资性房地产、长期股权投资、长期应收款、固定资产、工程物资、在建工程、固定资产清理、生产性生物资产、油气资产、无形资产……

数据模块	数据表	数据字段
IPO 公司信息	IPO 前利润表	股票代码、招股开始日期、统计日期、营业总收入、主营业务收入、利息收入、已赚保费、手续费及佣金收入、营业总成本、主营业务成本、利息支出、手续费及佣金支出、营业税金及附加、销售费用、管理费用、财务费用、资产减值损失、退保金、赔付支出净额、提取保险责任准备金、保单红利支出、分保费用、公允价值变动损益、投资收益、对联营企业和合营企业的投资收益、汇兑净收益、营业利润、营业外收入、营业外支出、非流动资产处置净损失、利润总额、所得税费用、未确认的投资损失、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益、稀释每股收益、其他综合收益、综合收益总额、归属于少数股东的综合收益总额、归属于母公司普通股股东综合收益总额
	IPO 前现金流量表	股票代码、招股开始日期、统计日期、销售商品、提供劳务收到的现金、收到的税费返还、收到其他与经营活动有关的现金、保户储金净增加额、客户存款和同业存放款项净增加额、向中央银行借款净增加额、向其他金融机构拆入资金净增加额、收取利息和手续费净增加额、收到的原保险合同保费取得的现金、收到的再保业务现金净额、处置交易性金融资产净增加额……
IPO 信息	招股基本情况	股票代码、公司简称、招股开始日期、首发上市日期、首发价格、发行数量合计、募集资金总额、首发募集资金、首发募集资金净额、网上发行数量、网下发行数量、向战略投资者配售数量、首发后总股本、首发前总股本、首发市盈率(摊薄)、首发市盈率(加权)、首发超募资金、首发预计募集资金、承销商认购余额、回拨比例(%)、预计发行股数、预计发行后总股本、预计募投项目投资总额、发行时每股面值、发行方式、承销方式、主承销商、审计机构、资产评估机构、律师事务所
	重要日期及费用	股票代码、公司简称、招股意向书公告日期、发行公告日、招股公告日、招股开始日期、招股截至日期、初步询价公告日、初步询价起始日、初步询价截止日、上市公告日、发行结果公告日、网上申购开始日期、网上申购截止日期、网下申购开始日期、网下申购截止日期、承销保荐费、审计费用、资产评估费用、律师费用、信息披露费用、其他费用、费用合计
	IPO 申购结果	股票代码、公司简称、招股开始日期、网上发行中签率(%)、网上申购数量上限、网上申购资金上限、网上发行有效申购数量、网上发行有效申购户数、网上超额认购倍数、网上冻结资金、网下申购配售比例(%)、网下申购总量、网下冻结资金、网下有效报价下限、网下有效报价上限、网下有效报价申购量、网下超额认购倍数、网下超额认购倍数(回拨前)、网下有效报价申购量比例(%)、网下高于有效报价上限的申购量、网下高于有效报价上限的申购量比例(%)、网下申购数量上限、网下申购数量下限、网下最低累进申购数量、网下申购报价数量、网下申购配售对象家数、网下申购询价对象家数、网下公募基金获配数量、网下社保基金获配数量、网下企业年金获配数量、网下保险资金获配数量、初步询价配售对象家数、初步询价询价对象家数、初步询价下限、初步询价上限、初步询价申购总量、初步询价申购倍数(回拨前)
	承销商情况	股票代码、公司简称、招股开始日期、首发保荐人、首发保荐人代表、首发上市推荐人、首发发行方式、首发承销方式、首发主承销商、首发副主承销商、首发分销商
	首发上市表现	股票代码、公司简称、招股开始日期、首日上市数量、上市首日开盘价、上市首日最高价、上市首日最低价、上市首日收盘价、上市首日成交均价、上

数据模块	数据表	数据字段
IPO 信息	首发上市表现	市首日成交量、上市首日成交额、上市首日涨跌幅 (%)、上市首日换手率 (%)
证券承销商信息	股东构成	券商名称、券商代码、统计日期、大股东名称、大股东占股比例、股东排名、大股东性质、最终控制人
	高管特征	统计日期、券商名称、券商代码、董监高姓名、职位类型、性别、出生年份、国籍、学历/学位、现任或者曾任、工作经历起止日期、工作单位、职位、是否全职
	报表信息	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、营业收入、营业利润、营业外收入、营业外支出、利润总额、所得税、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益 EPS、稀释每股收益 EPS、货币资金、客户资金存款、交易性金融资产、衍生金融资产、买入返售金融资产、存出保证金、可供出售金融资产、持有至到期投资、长期股权投资、固定资产、无形资产、资产总计、短期借款、质押借款、拆入资金、交易性金融负债、衍生金融负债、卖出回购金融资产款、代理买卖证券款、代理承销证券款、应付职工薪酬、应交税费、长期借款、应付债券、负债合计、股本、资本公积金、盈余公积金、未分配利润、一般风险准备、归属于母公司所有者权益合计、所有者权益合计、负债及股东权益总计、经营活动产生的现金流量净额、投资活动产生的现金流量净额、筹资活动产生的现金流量净额
	财务比率	券商名称、券商代码、合并会计报表或母公司会计报表、统计日期、证券公司净资本、营业收入、营业收入增长率(%)、手续费及佣金净收入、手续费及佣金净收入在营业收入中的占比(%)、利息净收入、利息净收入在营业收入中的占比(%)、投资净收益、投资净收益在营业收入中的占比(%)、公允价值变动净收益、公允价值变动净收益在营业收入中的占比(%)、其他收入、其他收入在营业收入中的占比(%)、代理买卖证券业务净收入、证券承销业务净收入、受托客户资产管理业务净收入
	承销排名	券商名称、券商编码、会计年度、当年承销 IPO 总募集资金金额、当年承销 IPO 总募集资金排名、当年承销 IPO 总募集资金占当年 IPO 市场总募集资金比例、当年 IPO 承销费用、当年 IPO 承销费用排名、当年 IPO 承销数量、当年联席保荐 IPO 数量、当年 IPO 承销数量排名
募集资金使用及变更	募集资金使用及变更	股票代码、公司简称、首发募集资金、首发募集资金净额、首发上市日期、T 年度、T 累计变更用途的募集资金总额、T 项目名称、T 项目承诺总投资额、T 调整后项目投资总额、T 项目实际投入金额、T 是否变更项目、T 截至期末累计投资金额、T 实现的收益金额、T 项目可行性是否发生重大变化、T+1 年度、T+1 累计变更用途的募集资金总额、T+1 本年项目名称、T+1 项目承诺总投资额、T+1 调整后投资总额、T+1 本年度实际投入金额……

5.数据范围

数据模块	数据起始时间
IPO 报审情况	2003 年
IPO 公司信息	1990 年
IPO 信息	1984 年
证券承销商信息	2006 年
募集资金使用及变更	1999 年

(六) 中国上市公司增发与配股数据库-CSEO

1.数据库简介

中国上市公司增发与配股数据库 (*SEOs and Share Allotment Database of Chinese Companies, CSEO*) 收集整理了 1991 年以来中国上市公司增发股票与配股的数据, 建立的详细、准确和全面的上市公司权益再融资研究学术数据库。该库涵盖了上市公司增发股票所涉及的几乎所有重要指标。

2.数据库特色

- 覆盖全面

本数据库中增发模块和配股模块包含多个字段, 内容详细。

3.典型数据举例

上市公司增发资料

股票代码	会计年度	增发公告日	增发数量 (股)	增发募集资金 (元)	增发实际募集资金 (元)	主承销商
000402	2004	2004-12-02	76000000	707740000	669210000	国泰君安
000407	2014	2014-12-04	31138790	175000000	164300000	东方花旗
000407	2015	2016-04-19	106035888	650000000	637010000	广发证券

4.数据库内容

数据模块	数据表	数据字段
增发资料	增发资料	股票代码、公司简称、会计年度、增发进度、增发发行方式、发行对象、定价方式、认购方式、预案公告日期、股东大会公告日、增发公告日、增发上市日、发审委通过公告日、证监会审核公告日、预计增发价格下限、增发目的、增发价格、增发数量、增发募集资金、增发实际募集资金、发行总费用、增发承销方式、增发主承销商、财务顾问
配股资料	配股资料	股票代码、公司简称、会计年度、配股进度、配股价格、配股募集资金、配股费用、配股实际募集资金、基准股本、每股配股数、计划配股数、实际配股数、预案公告日、股东大会公告日、股权登记日、配股除权日、配股上市日、缴款起始日、缴款终止日、配股获准公告日、配股公告日、配股主承销商、配股方式、配股承销方式、配股分销商、配股预案价上限、配股预案价下限、承销商认购余股、国有股理论配股数、法人股理论配股数、职工股理论配股数、转配股理论配股数、已流通股理论配股数、国有股实际配股数、法人股实际配股数、职工股实际配股数、转配股实际配股数、已流通股实际配股数、持股 5%以上的大股东持股数、持股 5%以上的大股东理论认购股数、持股 5%以上大股东认购股数、国家股认购方式、法人股认购方式、国有法人股认购方式、募集资金用途、发行人律师事务所、发行人律师负责人、资产评估机构、资产评估经办人、会计师事务所、会计师

5.数据范围

数据模块	数据起始时间
全部模块	1991 年

上市公司财务基础数据

(一) 上市公司业绩预告数据库-CEPD

1.数据库简介

上市公司业绩预告数据库 (*Estimated Performance Database of Listed Companies, CEPD*) 是基于上市公司业绩预告和业绩快报公告整理而成的专业数据库。CEPD 包含上市公司业绩预告、上市公司业绩快报两张表, 包含了上市公司正式披露季报、半年报、年报之前, 交易所对上市公司提前预告经营业绩和披露业绩快报情况的披露, 反映了上市公司基本面的最新重大变化, 为该领域研究提供了全面、高质量的数据支持。

2.数据库特色

- 字段全面, 覆盖时间长

预计预告与快报表均涵盖了相关的重要字段, 数据可追溯至 2006 年, 提供了长时间的序列数据。

3.典型数据举例

业绩预告

股票代码	统计日期	发布时间	预告类型	预告净利润下限	预告净利润上限
000005	2016-06-30	2016-07-15	扭亏	11000.00	13000.00
000005	2016-12-31	2017-01-26	扭亏	10500.00	13000.00
000006	2016-06-30	2016-07-12	预减	7742.49	12904.14
000006	2016-12-31	2017-01-11	预增	75015.50	87518.09

4.数据库内容

数据模块	数据表	数据字段
业绩快报	业绩快报	股票代码、公司简称、统计日期、发布时间、本期营业收入、上年同期营业收入、营业收入增长率(%)、本期营业利润、上年同期营业利润、营业利润增长率(%)、本期利润总额、上年同期利润总额、利润总额增长率(%)、本期归属母公司股东的净利润、上年同期归属母公司股东的净利润、归属母公司股东的净利润增长率(%)、本期基本每股收益、上年同期基本每股收益、基本每股收益增长率(%)、本期加权平均净资产收益率(%)、上年同期加权平均净资产收益率(%)、本期总资产、上年同期总资产、总资产增长率(%)、本期净资产、上年同期净资产、净资产增长率(%)、业绩简要说明
业绩预告	业绩预告	股票代码、公司简称、统计日期、发布时间、预告类型、预告摘要、预告净利润下限、预告净利润上限、预告净利润同比增长下限(%)、预告净利润同比增长上限(%)、预告内容、业绩变动原因

5.数据范围

数据模块	数据起始时间
全部模块	2006 年

(二) 财务报告披露时间数据库-FRDT

1.数据库简介

财务报告披露时间数据库 (*Financial Report Disclosure Time Database*, **FRDT**) 包含了 1999 年以来中国 A 股和 B 股上市公司的年报、中报和季报的预披露时间、更正时间和实际披露时间等多个重要时间字段。

2.数据库特色

- 连续的时间序列数值

FRDT 包含了 1999 年以来上市公司在各个披露节点上的财务报告披露时间及其变更信息。

3.典型数据举例

财务报告披露时间

股票代码	统计日期	首次预约时间	第一次变更日期	第二次变更日期	第三次变更日期	实际披露时间
000005	2007-12-31	2008-03-29	2008-04-18	2008-04-24	2008-04-26	2008-04-26
000011	2006-12-31	2007-04-27	2007-04-27	2007-04-30	2007-04-28	2007-04-28
000017	2007-03-31	2007-04-20	2007-04-24	2007-04-30	2007-04-28	2007-04-28
000040	2007-12-31	2008-04-16	2008-04-19	2008-04-25	2008-04-22	2008-04-22
000055	2004-12-31	2005-02-04	2005-04-08	2005-04-16	2005-04-21	2005-04-21
000150	2009-12-31	2010-04-24	2010-02-12	2010-03-02	2010-02-27	2010-02-27

4.数据库内容

数据模块	数据表	数据字段
财务报告披露时间	财务报告披露时间	股票代码、公司简称、统计日期、首次预约时间、第一次变更日期、第二次变更日期、第三次变更日期、实际披露时间

5.数据范围

数据模块	数据起始时间
全部模块	1999 年

(三) 中国公司财务报表数据库-CNFS

1.数据库简介

中国公司财务报表数据库 (*Financial Statements of Chinese Firms, CNFS*) 是以研究为目的而设计的, 对 186 个重要会计科目进行全面深入研究, 旨在提供完整精确的财务信息。*CNFS* 以 2006 年财政部出台的企业会计准则为基本框架, 包含了中国自 1990 年以来公布的各公司全部财务信息, 数据库的设计基本遵循企业会计准则规定的合并报表列报格式, 适用于所有行业, 并能往前兼容历史数据。目前, 本数据库包含三个数据模块, 即资产负债表、利润表和现金流量表。

2.数据库特色

- 字段设计科学

CNFS 严格按照企业会计准则所规定会计科目进行字段设置, 最大程度保证字段的有用性。对 2007 年之前的公司财务报表数据进行了调整, 这使得时间序列上的数据具有一致性。

3.典型数据举例

上市公司利润表

股票代码	统计日期	销售费用 (元)	管理费用 (元)	财务费用 (元)	净利润 (元)
000009	2012-12-31	537867968	491278560	255133456	307301120
000010	2012-12-31	21200886	37293092	2028002	3593386
000011	2012-12-31	43951176	109680544	-1504589	374822144

4.数据库内容

数据模块	数据表	数据字段
资产负债表	资产负债表	股票代码、统计日期、货币资金、交易性金融资产、应收票据、应收账款、其他应收款、预付账款、应收股利、应收利息、存货、消耗性生物资产、待摊费用、划分为持有待售的资产、一年内到期的非流动资产、结算备付金、拆出资金、应收保费、应收分保账款、应收分保合同准备金、买入返售金融资产、其他流动资产、流动资产合计、可供出售金融资产、持有至到期投资、投资性房地产、长期股权投资、长期应收款、固定资产、工程物资、在建工程、固定资产清理、生产性生物资产、油气资产、无形资产、开发支出、商誉、长期待摊费用、递延所得税资产、发放贷款及垫款、其他非流动资产、非流动资产合计、资产总计、短期借款、交易性金融负债、应付票据、应付账款、预收账款、应付职工薪酬、应交税费、应付利息、应付股利、其他应付款、预提费用、递延收益-流动负债、划分为持有待售的负债、一年内到期的非流动负债、短期应付债券、向中央银行借款、吸收存款及同业存放、拆入资金、卖出回购金融资产款、应付手续费及佣金、应付分保账款、保险合同准备金、代理买卖证券款、代理承销证券款、其他流动负债、流动负债合计、长期借款、应付债券、长期应付款、长期应付职工薪酬、专项应付款、预计负债、递延所得税负债、递延收益-非流动负债、其他非流动负债、非流动负债合计、负债合计、实收资本(股本)、其他权益工具、其他权益工具:优先股、资本公积、盈余公积、未分配利润、库存股、其他综合收益、专项储备、一般风险准备、外币报表折算差额、未确认的投资损失、少数股

数据模块	数据表	数据字段
资产负债表	资产负债表	东权益、归属于母公司所有者权益合计、所有者权益合计、负债及股东权益总计
利润表	利润表	股票代码、统计日期、营业总收入、主营业务收入、利息收入、已赚保费、手续费及佣金收入、营业总成本、主营业务成本、利息支出、手续费及佣金支出、营业税金及附加、销售费用、管理费用、财务费用、资产减值损失、退保金、赔付支出净额、提取保险责任准备金、保单红利支出、分保费用、公允价值变动损益、投资收益、对联营企业和合营企业的投资收益、汇兑净收益、营业利润、营业外收入、营业外支出、非流动资产处置净损失、利润总额、所得税费用、未确认的投资损失、净利润、少数股东损益、归属于母公司所有者的净利润、基本每股收益、稀释每股收益、其他综合收益、综合收益总额、归属于少数股东的综合收益总额、归属于母公司普通股股东综合收益总额
现金流量表	现金流量表	股票代码、统计日期、销售商品、提供劳务收到的现金、收到的税费返还、收到其他与经营活动有关的现金、保户储金净增加额、客户存款和同业存放款项净增加额、向中央银行借款净增加额、向其他金融机构拆入资金净增加额、收取利息和手续费净增加额、收到的原保险合同保费取得的现金、收到的再保业务现金净额、处置交易性金融资产净增加额、拆入资金净增加额、回购业务资金净增加额、经营活动现金流入小计、购买商品、接受劳务支付的现金、支付给职工以及为职工支付的现金、支付的各项税费、支付其他与经营活动有关的现金、客户贷款及垫款净增加额、存放央行和同业款项净增加额、支付原保险合同赔付款项的现金、支付手续费的现金、支付保单红利的现金、经营活动现金流出小计、经营活动产生的现金流量净额、收回投资收到的现金、取得投资收益收到的现金、处置固定资产、无形资产和其他长期资产收回的现金、处置子公司及其他营业单位收到的现金、收到其他与投资活动有关的现金、投资活动现金流入小计、购建固定资产、无形资产和其他长期资产支付的现金、投资支付的现金、取得子公司及其他营业单位支付的现金、支付其他与投资活动有关的现金、投资活动现金流出小计、投资活动产生的现金流量净额、吸收投资收到的现金、子公司吸收少数股东投资收到的现金、取得借款收到的现金、收到其他与筹资活动有关的现金、发行债券收到的现金、筹资活动现金流入小计、偿还债务支付的现金、分配股利、利润或偿付利息支付的现金、子公司支付给少数股东的股利、利润、支付其他与筹资活动有关的现金、筹资活动现金流出小计、筹资活动产生的现金流量净额、汇率变动对现金的影响、现金及现金等价物净增加额、期初现金及现金等价物余额、期末现金及现金等价物余额、净利润、资产减值准备、固定资产折旧、油气资产折耗、生产性生物资产折旧、无形资产摊销、长期待摊费用摊销、待摊费用减少、预提费用增加、处置固定资产、无形资产和其他长期资产的损失、固定资产报废损失、公允价值变动损失、财务费用、投资损失、递延所得税资产减少、递延所得税负债增加、存货的减少、经营性应收项目的减少、经营性应付项目的增加、未确认的投资损失、其他、间接法-经营活动产生的现金流量净额、债务转为资本、一年内到期的可转换公司债券、融资租入固定资产、现金的期末余额、现金的期初余额、现金等价物的期末余额、现金等价物的期初余额、间接法-现金及现金等价物净增加额

5.数据范围

数据模块	数据起始时间
全部模块	1990年

(四) 上市公司财务报表附注数据库-NFSD

1.数据库简介

上市公司财务报表附注数据库 (*Notes to Financial Statements Database, NFSD*) 是基于上市公司年报数据, 对报表附注数据进行加工整理而成的专业数据库。根据报表附注中的披露项目, *NFSD* 设置了资产项目、应收项目、债务项目、费用明细、损益项目、其他项目等 6 个方面的财务数据, 可为财务报表使用者提供完整、准确的报表附注数据支持。

2.数据库特色

- 丰富的报表附注字段

NFSD 包含多种类型报表附注字段, 从资产、债务、债权和损益等多个方面体现了附注披露情况。

3.典型数据举例

(1) 存货

股票代码	统计日期	存货项目	账面余额	跌价准备	账面价值
000009	2016-12-31	原材料	241215375.31	2926636.82	238288738.49
000011	2016-12-31	原材料	1230918.63	484743.69	746174.94
000009	2016-12-31	包装物	19169821.78	207385.90	18962435.88
000019	2016-12-31	包装物	6497436.43	649825.76	5847610.67
000926	2016-12-31	产成品	58009374.04	881008.21	57128365.83
000959	2016-12-31	产成品	1837318375.00	83182519.78	1754135855.22

(2) 政府补助

股票代码	统计日期	项目	金额	说明
000004	2015-12-31	化学原料科技奖	8000.00	与收益相关
000004	2016-12-31	烟台市专利补助资金	4000.00	与收益相关
000009	2015-12-31	惠州贝特瑞厂区土地出让补助金	121355.16	与资产相关
000009	2015-12-31	山西贝特瑞厂区土地出让补助金	84717.96	与资产相关

4.数据库内容

数据模块	数据表	数据字段
资产项目	存货	股票代码、统计日期、存货项目、账面余额、跌价准备、账面价值
	固定资产	股票代码、统计日期、固定资产科目、数据类型、总计、房屋及建筑物、机器设备、机电设备、运输设备、办公电子设备、专用设备、通用设备、其他设备
	无形资产	股票代码、统计日期、无形资产科目、数据类型、总计、专利技术、非专利技术、商标权、特许经营权、土地使用权、房屋使用权、软件、其他
	资产减值损失	股票代码、统计日期、数据类型、坏账准备合计、坏账准备-应收账款、坏账准备-其他应收款、短期投资跌价准备合计、存货跌价准备合计、存货跌价准备-库存商品、存货跌价准备-原材料、存货跌价准备-低值易

数据模块	数据表	数据字段
资产项目	资产减值损失	耗品、存货跌价准备-包装物、存货跌价准备-在产品、存货跌价准备-自制半成品、长期投资减值准备合计、固定资产减值准备合计、固定资产减值准备-房屋、建筑物、固定资产减值准备-机器设备、固定资产减值准备-运输工具、固定资产减值准备-电子设备、固定资产减值准备-其它设备、无形资产减值准备合计、无形资产减值准备-专利权、无形资产减值准备-商标权、在建工程减值准备、委托贷款减值准备、资产减值准备合计、抵债资产减值准备、融资租赁风险准备、存放同业及其他金融机构款项、拆出资金、应收利息、发放贷款和垫款、长期待摊费用减值准备、长期应收款减值准备、应收款项类投资减值准备、融出资金坏账准备、融出证券坏账准备
应收项目	应收账款明细	股票代码、统计日期、债务人名称、欠款金额、占应收账款总额比例(%)、拖欠时间、拖欠原因
	应收账款大股东欠款	股票代码、统计日期、债务人名称、欠款金额、占总应收款比例(%)
	应收账款账龄结构	股票代码、统计日期、账龄、应收账款金额、占总应收款比例(%)、提取的坏账准备、坏账准备计提比例(%)
	其他应收款明细	股票代码、统计日期、债务人名称、欠款金额、占其他应收款总额比例(%)、拖欠时间、拖欠原因
	其他应收款大股东欠款	股票代码、统计日期、债务人名称、欠款金额、占总其他应收款比例(%)
	其他应收款账龄结构	股票代码、统计日期、账龄、其他应收账款金额、占总其他应收款比例(%)、提取的坏账准备、坏账计提比例(%)
债务项目	公司借款	股票代码、统计时间、借款类型、借款项目、期末数、期初数、说明
	应付职工薪酬	股票代码、统计日期、数据类型、工资、奖金、津贴和补贴、职工福利费、社会保险费、其中：医疗保险费、其中：基本养老保险费、其中：年金缴费、其中：失业保险费、其中：工伤保险费、其中：生育保险费、住房公积金、工会经费和职工教育经费、非货币性福利、辞退福利、因解除劳动关系给予的补偿、以现金结算的股份支付、离职后福利、其中_基本养老保险费、其中_年金缴费、其中_失业保险费、其他薪酬、合计
	应交税费	股票代码、统计日期、增值税、营业税、企业所得税、城市维护建设税、房产税、个人所得税、消费税、土地使用税、印花税、车船使用税、土地增值税、资源税、教育费附加、车辆购置税、地方教育费附加、其他
费用明细	管理费用	股票代码、统计日期、项目名称、本期发生额、上期发生额
	销售费用	股票代码、统计日期、项目名称、本期发生额、上期发生额
	财务费用	股票代码、统计日期、项目名称、本期发生额、上期发生额
损益项目	投资收益	股票代码、统计日期、成本法核算的长期股权投资收益、权益法核算的长期股权投资收益、处置长期股权投资产生的投资收益、持有交易性金融资产期间取得的投资收益、持有至到期投资取得的投资收益期间取得的投资收益、持有可供出售金融资产等期间取得的投资收益、处置交易性金融资产取得的投资收益、持有至到期投资取得的投资收益、可供出售金融资产等取得的投资收益、其他、合计
	非经常性损益	股票代码、统计日期、非流动性资产处置损益、偶发性损益、政府补助、资金占用费、享有被投资单位可辨认净资产公允价值产生的收益、非货

数据模块	数据表	数据字段
损益项目	非经常性损益	币性资产交换损益、委托他人投资或管理资产的损益、偶发性资产减值准备、债务重组损益、企业重组费用、企业合并净损益、或有事项损益、公允价值变动损益、应收款项减值准备转回、对外委托贷款取得的损益、托管费收入、其他营业外收支、其他非经常性损益、对所得税影响额、归属于母公司所有者的非经常性损益净额、归属于少数股东的非经常性损益净额
	营业外收入	股票代码、统计日期、非流动资产处置利得合计、其中：固定资产处置利得、其中：无形资产处置利得、其中：其他非流动资产处置利得、债务重组利得、接受捐赠、政府补助、罚款收入、赔偿款收入、营业外收入特殊项目、其他营业外收入、营业外收入合计
	营业外支出	股票代码、统计日期、非流动资产处置损失合计、其中：固定资产处置损失、其中：无形资产处置损失、其中：其他非流动资产处置损失、债务重组损失、罚没支出、对外捐赠、营业外支出特殊项目、其他营业外支出、营业外支出合计
	政府补助	股票代码、统计日期、项目、金额、说明
其他项目	财务数据变动说明	股票代码、统计日期、项目、本期金额、上期金额、相比上年同期变动情况、变动原因
	关联方债权债务	股票代码、统计日期、关联方名称、上市公司向关联方提供资金发生额、上市公司向关联方提供资金余额、关联方向上市公司提供资金发生额、关联方向上市公司提供资金余额
	上市银行存贷款	股票代码、统计日期、数据类型、企业存款、企业活期存款、企业定期存款、储蓄存款、储蓄活期存款、储蓄定期存款、短期贷款、中长期贷款

5.数据范围

数据模块	数据起始时间
全部模块	1992 年

(五) 上市公司财务指标数据库-CNFI

1.数据库简介

上市公司财务指标数据库 (*Financial Indicator Database of Listed Companies, CNFI*) 是为研究中国上市公司财务状况和经营成果而建立的专业数据库。本数据库的数据建立在 *CNDRS* 中国公司财务报表数据库的基础上, 包含了中国自 1991 年以来公布的各上市公司财务指标信息, 数据库的设计基本遵循企业会计准则规定的合并报表列报格式, 根据相关的会计制度、会计准则做必要的调整, 适用于所有行业, 并能往前兼容历史数据。*CNFI* 提供了偿债能力、盈利能力、成长能力、每股指标、杜邦分析、现金流分析、经营能力等 7 个方面的财务指标。

2.数据库特色

- 计算结果精确性

对数据进行了细致地校对及检验工作, 避免了可能在运算过程中发生逻辑错误的情况, 保证了数据的精确性。

- 指标多样性

在盈利指标中, 对同一指标提供多种算法, 方便使用者自行选择。

- 更新及时性

每年对数据库进行定期更新, 保证数据的及时性和持续性。

3.典型数据举例

偿债能力

股票代码	统计日期	流动比率	现金比率	资产负债率	产权比率	权益乘数	息税前利润	现金流利息保障倍数
000002	2016-12-31	1.2436	0.1501	0.8054	4.1379	5.1379	40845679693	24.8520
000004	2016-12-31	3.6575	2.7555	0.2035	0.2554	1.2554	48792564	56.7457
000017	2016-12-31	1.2580	0.7128	0.6805	2.1298	3.1298	4220942	-1.0991
000018	2016-12-31	1.4062	0.4533	0.7708	3.3629	4.3629	688774984	-13.2448

4.数据库内容

数据模块	数据表	数据字段
偿债能力	偿债能力	股票代码、统计日期、流动比率、速动比率、现金比率、营运资金、营运资金与借款比、带息负债比率、资产负债率、产权比率、权益乘数、息税前利润、已获利息倍数、现金债务总额比、现金带息债务比、现金流动负债比、经营活动产生的现金流量净额/净债务、现金非流动负债比、现金流利息保障倍数、现金流到期债务保障倍数、非筹资性现金净流量与流动负债的比率、非筹资性现金净流量与负债总额的比率、有形资产负债率、有形资产带息债务比、有形净值债务率、权益对负债比率、长期借款与总资产比、长期资本负债率、长期负债权益比率、长期债务与营运资金比率

数据模块	数据表	数据字段
盈利能力	盈利能力	股票代码、统计日期、资产报酬率 A (%)、资产报酬率 B (%)、总资产净利润率 A (%)、总资产净利润率 B (%)、流动资产净利润率 A (%)、流动资产净利润率 B (%)、固定资产净利润率 A (%)、固定资产净利润率 B (%)、净资产收益率 A (%)、净资产收益率 B (%)、息税前利润、息前税后利润、净利润与利润总额比 (%)、利润总额与息税前利润相比 (%)、营业净利率 (%)、营业毛利率 (%)、营业成本率 (%)、营业利润率 (%)、总营业成本率 (%)、销售费用率 (%)、管理费用率 (%)、财务费用率 (%)、销售期间费用率 (%)、成本费用利润率 (%)、主营业务比率 (%)、资产减值损失/营业总收入 (%)、资产减值损失/营业利润 (%)、息税前营业利润率 (%)、现金与利润总额比 (%)、归属于母公司净资产收益率 A (%)、归属于母公司净资产收益率 B (%)
成长能力	成长能力	股票代码、统计日期、基本每股收益增长率 (%)、稀释每股收益增长率 (%)、每股经营活动产生的现金流量净额增长率 (%)、营业总收入增长率 (%)、营业收入增长率 (%)、营业利润增长率 (%)、利润总额增长率 (%)、净利润 (同比增长率)、归属母公司股东的净利润增长率 (%)、经营活动产生的现金流量净额增长率 (%)、总负债增长率 (%)、总资产增长率 (%)、货币资金增长率 (%)、固定资产增长率 (%)
每股指标	每股指标	股票代码、统计日期、每股收益 EPS (基本)、每股收益 EPS (稀释)、每股收益 EPS (期末股本摊薄)、每股净资产 BPS、每股经营活动产生的现金流量净额、每股营业总收入、每股营业收入、每股资本公积、每股盈余公积、每股未分配利润、每股留存收益、每股现金流量净额、每股息税前利润
杜邦分析	杜邦分析	股票代码、统计日期、净资产收益率、权益乘数、销售净利率、总资产周转率、净利润/利润总额、利润总额/息税前利润、息税前利润/营业总收入
现金流分析	现金流分析	股票代码、统计日期、销售收现比、销售现金比率、经营活动产生的现金流量净额/经营活动净收益、资本支出/折旧和摊销、经营活动产生的现金流量净额占比 (%)、投资活动产生的现金流量净额占比 (%)、筹资活动产生的现金流量净额占比 (%)、经营性现金净流量/营业总收入、现金满足投资比率、现金营运指数、全部资产现金回收率 (%)、每股股利 (税后)、总股本、现金股利保障倍数
经营能力	经营能力	股票代码、统计日期、存货周转率、应收账款周转率、流动资产周转率、营运资本周转率、固定资产周转率、非流动资产周转率、总资产周转率、应付账款周转率、存货周转天数、应收账款周转天数、应付账款周转天数、营业周期、净营业周期

5.数据范围

数据模块	数据起始时间
全部模块	1991 年

（六）中国上市公司盈利预测数据库-CEFD

1.数据库简介

中国上市公司盈利预测数据库 (*Earnings Forecast Database of Chinese Companies, CEFD*) 是以研究为目的设计开发的学术数据库。

本数据库共由两个模块组成：分析师荐股评级与分析师盈利预测。两个模块的具体字段内容均是提取自 30 多万份分析师报告，包括有具体的分析师标准化荐股意见以及各相关主要指标与比率的预测情况，字段内容丰富、数据质量高、信息全面，便于研究者与学者的分析使用。

2.数据库特色

- 分析师及券商唯一编码

CEFD 给数据库中 100 多家券商以及 6000 多名分析师分配了独特编码，是目前为止唯一给券商和分析师独特编码的数据库，有利于研究者进行数据匹配和深度挖掘。

- 数据覆盖全面

相比市场已有数据库，*CEFD* 所覆盖的分析师预测对象和预测数据更为全面和准确。

3.典型数据举例

分析师盈利预测

股票代码	报告日期	盈利预测结束日期	分析师姓名	分析师编码	每股净利润预测	主营业务收入预测	实际盈余公告日
000001	2009-04-24	2009-12-31	余晓宜	A05010	1.4 元/股	15371830 元	2010-03-12
000001	2009-04-24	2009-12-31	朱琰	A05846	1.47 元/股	15037000 元	2010-03-12
000001	2009-04-24	2010-12-31	陈水祥	A00326	1.68 元/股	17679000 元	2011-02-25
000001	2009-04-24	2010-12-31	范坤祥	A00669	1.76 元/股	19781000 元	2011-02-25

4.数据库内容

数据模块	数据表	数据字段
分析师荐股评级	分析师荐股评级	股票代码、报告日期、分析师姓名、分析师编码、券商名称、券商编码、原始报告中的荐股评级、标准化荐股意见、标准化荐股意见编码、IPO 询价区间预测、IPO 发行价预测、目标股价预测、是否是明星分析师
分析师盈利预测	分析师盈利预测	股票代码、报告日期、盈利预测结束日期、分析师姓名、分析师编码、是否是明星分析师、券商名称、券商编码、每股净利润预测、P/E 比率预测、净利润预测、主营业务收入预测、实际盈余公告日

5.数据范围

数据模块	数据起始时间
全部模块	2001 年

上市公司治理基础数据

(一) 上市公司基本信息数据库-CBID

1. 数据库简介

上市公司基本信息数据库 (*Basic Information Database of Listed Companies, CBID*) 包含了 1990 年以来在中国 A 股和 B 股市场上市或者曾经上市的公司信息。*CBID* 提供了这些上市公司的股票代码、经营范围、注册资本、所属行业、曾用名称等多个维度的重要基本信息。

2. 数据库特色

- 包含各类重要基本信息

CBID 包含了上市公司多方面的基本信息，尤其是列示了时间序列上的公司曾用名变化情况。

3. 典型数据举例

公司曾用名

股票代码	名称变更日期	变更前名称	变更后名称
000001	2012-07-27	深圳发展银行股份有限公司	平安银行股份有限公司
000002	1993-12-31	深圳万科企业股份有限公司	万科企业股份有限公司
002004	2015-09-01	华邦颖泰股份有限公司	华邦生命健康股份有限公司
600398	2014-04-03	凯诺科技股份有限公司	海澜之家股份有限公司

4. 数据库内容

数据模块	数据表	数据字段
公司基本信息	公司基本信息	股票代码、公司简称、上市板块、股票种类、公司中文名称、公司英文名称、成立日期、首发上市日期、摘牌日期、同公司 B 股代码、同公司 H 股代码、注册资本、注册地址、办公地址、所在省份、所在地级市、所在县级市、主营产品名称、经营范围、行业名称 A、行业代码 A、行业名称 B、行业代码 B
公司曾用名	公司曾用名	股票代码、名称变更日期、变更前名称、变更后名称

5. 数据范围

数据模块	数据起始时间
全部模块	1990 年

(二) 中国公司治理数据库-CCGD

1. 数据库简介

中国公司治理数据库 (*Chinese Corporate Governance Database, CCGD*) 是基于对中国公司治理研究的深入考察而建立的专业数据库, 包括了研究中国公司治理问题几乎所有的指标数据, 能够为学者提供全面、准确的研究数据。**CCGD** 包括 3 个模块, 分别为股东治理、管理层治理、外部治理。

2. 数据库特色

- 全面

CCGD 对研究公司治理问题可能涉及到的内外部治理指标进行了详细而全面的归类, 有助于减少学者的数据搜集时间和成本, 大大便利了其研究。

3. 典型数据举例

上市公司股权性质

股票代码	会计年度	第一大股东名称	第一大股东持股数量 (股)	第一大股东持股比例 (%)	实际控制人	实际控制人类别
000034	2012	中国希格玛有限公司	78306968	21.88	王晓岩	个人
000034	2013	中国希格玛有限公司	78306968	21.88	王晓岩	个人
000034	2014	中国希格玛有限公司	78306968	21.88	王晓岩	个人
000034	2015	中国希格玛有限公司	78306968	21.88	王晓岩	个人

4. 数据库内容

数据模块	数据表	数据字段
股东治理	主要股东持股	股票代码、统计日期、序号、股东名称、持股数量、持股比例 (%)
	公司股权性质	股票代码、公司简称、会计年度、实际控制人、实际控制人类别、第一大股东名称、第一大股东持股数量、第一大股东持股比例 (%)
	国有股持股情况	股票代码、会计年度、国有股比例 (%)、国有股股数、总股数
	机构持股情况	股票代码、公司简称、统计日期、公司总股本、公司流通股股本、基金持股家数、基金持股合计、基金持股占流通 A 股比例 (%)、基金持仓市值、券商持股家数、券商持股合计、券商持股占流通 A 股比例 (%)、券商持仓市值、保险公司持股家数、保险公司持股合计、保险公司持股占流通 A 股比例 (%)、保险公司持仓市值、社保基金持股家数、社保基金持股合计、社保基金持股占流通 A 股比例 (%)、社保基金持仓市值、QFII 持股家数、QFII 持股合计、QFII 持股占流通 A 股比例 (%)、QFII 持仓市值、其他机构持股家数、其他机构持股合计、其他机构持股占流通 A 股比例 (%)、其他机构持仓市值
	股权集中度	股票代码、统计日期、第 1 大股东持股比例、前 3 大股东持股比例之和、前 5 大股东持股比例之和、前 10 大股东持股比例之和、Herfindahl 指数 1、Herfindahl 指数 3、Herfindahl 指数 5、Herfindahl 指数 10、Z 指数、S 指数

数据模块	数据表	数据字段
管理层治理	董监高持股	股票代码、会计年度、高管持股比例（%）、管理层持股比例（%）、董事会持股比例（%）、监事会持股比例（%）、高管持股数量、管理层持股数量、董事会持股数量、监事会持股数量、总股数、董事总人数、监事总人数、独立董事人数、独立董事比例（%）
	高管及员工薪酬	股票代码、公司简称、会计年度、董监高年薪总额、前三名高管报酬总额、前三名董事报酬总额、董事会秘书薪酬、财务总监薪酬、总经理薪酬、董事长薪酬、员工人均薪酬
	两职兼任情况	股票代码、会计年度、董事长与总经理兼任情况、董事长名称、总经理名称
外部治理	发行 B&H 股	股票代码、会计年度、是否发行 B 股或 H 股
	会计师事务所排名	会计师事务所、会计年度、事务所排名、总收入、当年收入占行业收入比重（%）、注册会计师总人数
	审计单位及审计意见	股票代码、会计年度、会计师事务所、审计师、审计意见、审计费用

5.数据范围

数据模块	数据起始时间
全部模块	1990 年

(三) 审计意见与费用数据库-AUDIT

1.数据库简介

审计意见与费用数据库 (*Audit Opinion & Expenses Database of Chinese Companies, AUDIT*) 包括审计师事务所对上市公司出具的审计意见与收取的审计费用等信息。具体包括上市公司聘请的注册会计师事务所及审计师姓名信息、出具的审计意见和支付的审计费用等; *AUDIT* 收录了 1990 年至今所有中国上市公司的审计意见与审计收费信息。

2.数据库特色

- 覆盖时间长

AUDIT 覆盖了 1990 年以来所有可获得上市公司审计单位、审计意见与审计费用等信息, 可充分满足研究需要。

- 数据准确

我们对审计费用和审计意见等数据进行了精确的核对和校正, 保证数据的准确性。

3.典型数据举例

上市公司审计单位与审计费用

股票代码	会计年度	会计师事务所	审计费用
002543	2012	立信会计师事务所 (特殊普通合伙)	780000 元
002543	2013	立信会计师事务所 (特殊普通合伙)	730000 元
002543	2014	广东正中珠江会计师事务所 (特殊普通合伙)	760000 元
002543	2015	广东正中珠江会计师事务所 (特殊普通合伙)	790000 元
002544	2011	信永中和会计师事务所	350000 元
002544	2012	信永中和会计师事务所	370000 元
002544	2013	众华会计师事务所	392200 元

3.数据库内容

数据模块	数据表	数据字段
审计单位与审计意见	审计单位与审计意见	股票代码、会计年度、会计师事务所、审计师、审计意见、审计意见说明
审计单位与审计费用	审计单位与审计费用	股票代码、会计年度、会计师事务所、审计费用

4.数据范围

数据模块	数据起始时间
全部模块	1990 年

（四）中国上市公司管理层变更数据库-MTDB

1.数据库简介

中国上市公司管理层变更数据库 (*Management Turnover Database of Chinese Companies, MTDB*) 系统全面地统计了与上市公司管理层岗位变更和持股变更相关的重要数据信息。数据库涵盖了高管离职时间、离职日期、离职原因, 高管持股变动日期、关联人、变动股份等。

2.数据库特色

- 兼具相关性与完整性

MTDB 定位面向学术研究, 因此在数据库字段设计上充分贴合学术研究的需求, 同时在数据量上力求做到完整。

3.典型数据举例

上市公司管理层岗位变动

股票代码	会计年度	高管姓名	出生日期	担任职务	任职日期	离职日期	变动原因
000063	2004	倪勤	1960-01-01	副总裁	2004-01-16	2013-01-15	解聘
000063	2004	谢大雄	1964-01-01	副总裁	2004-01-16	2013-01-15	解聘
000539	2014	祝德俊	1960-02-01	董事	2014-04-24	2015-07-04	辞职
000548	2006	皮钊	1966-01-01	董事	2006-11-23	2015-07-08	辞职

4.数据库内容

数据模块	数据表	数据字段
管理层岗位变动	管理层岗位变动	股票代码、会计年度、高管姓名、出生日期、担任职务、任职日期、离职日期、变动原因、是否在任
管理层持股变动	管理层持股变动	股票代码、公司简称、会计年度、公告日期、变动开始日期、变动截止日期、变动人姓名、关联董监高姓名、变动人与关联董监高关系、变动方向、变动股数、变动后持股数量、变动后持股比例 (%)、交易均价

5.数据范围

数据模块	数据起始时间
管理层岗位变动	1991 年
管理层持股变动	2001 年

（五）公司与高管违规处罚数据库-VPCE

1.数据库简介

公司与高管违规处罚数据库（*Violation and Punishment of Chinese Companies & Executives, VPCE*）包括上市公司违规处罚与上市公司高管违规处罚数据，涵盖违规主体、违规类型、处分措施等字段。这些处罚主要来自沪深证券交易所、证监会、财政部等监管部门。

2.数据库特色

- 全面收录高管与公司处罚信息

VPCE 根据各方来源，全面收录了高管与公司的处罚信息，并对这些信息进行了适当的整理，所设计的字段更符合学术需求。

3.典型数据举例

上市公司高管违规与处罚

股票代码	公告日期	违规类型	处罚对象	与上市公司关系	处罚金额
000007	2015/12/23	信息披露违规; 涉嫌违反相关法律法规	戴光	公司高管	50000 元
000035	2004/6/17	涉嫌违反相关法律法规	董志刚	公司高管	30000 元
000035	2006/10/12	信息披露违规; 涉嫌违反相关法律法规	曹传德	公司高管	50000 元

4.数据库内容

数据模块	数据表	数据字段
上市公司违规与处罚	上市公司违规与处罚	事件 ID、股票代码、公司简称、公告日期、违规类型、违规类型分类、处罚对象、与上市公司关系、违规行为、处分类型、处分措施、处理人、处罚金额、相关法规、行业名称
上市公司高管违规与处罚	上市公司高管违规与处罚	事件 ID、股票代码、公司简称、公告日期、违规类型、违规类型分类、处罚对象、与上市公司关系、违规行为、处分类型、处分措施、处理人、处罚金额、相关法规、行业名称

5.数据范围

数据模块	数据起始时间
上市公司违规与处罚	1994 年
上市公司高管违规与处罚	2001 年

（六）高管薪酬与激励数据库-ECEI

1.数据库简介

高管薪酬与激励数据库 (*Executive Compensation & Equity Incentive Database, ECEI*) 系统全面地统计了与上市公司管理层薪酬相关的重要数据信息。数据库不仅统计了管理层薪酬信息与管理层激励信息, 同时还将公司股权激励方案与股权激励实施明细模块纳入数据库内。

2.数据库特色

- 学术化

ECEI 不仅统计了高管薪酬数据、高管股权激励数据, 还将公司股权激励方案与股权激励实施明细作为两个独立模块纳入数据库内, 充分考虑到了学术研究的需要。

3.典型数据举例

上市公司管理层薪酬

股票代码	公司简称	会计年度	董监高年薪总额 (元)	前三名高管薪酬总额 (元)	董事会秘书薪酬 (元)	财务总监薪酬 (元)	总经理薪酬 (元)	董事长薪酬 (元)
600881	亚泰集团	2009	13932100	2939500	442100	900600	1138300	1138300
600881	亚泰集团	2010	21697000	4474500	776000	1363600	1747300	1747300
600881	亚泰集团	2011	27402200	4922300	860800	1490800	1940700	1940700
600881	亚泰集团	2012	18501700	3422300	860800	1040800	2381500	1340700

4.数据库内容

数据模块	数据表	数据字段
管理层薪酬	管理层薪酬	股票代码、公司简称、会计年度、董监高年薪总额、前三名高管薪酬总额、前三名董事薪酬总额、董事会秘书薪酬、财务总监薪酬、总经理薪酬、董事长薪酬
	管理层股权激励	事件 ID、股票代码、公告日期、高管姓名、担任职务、激励标的物、授予的激励数量、占授予激励数量的比例 (%)、行使的激励数量
公司股权激励	股权激励方案	股票代码、公司简称、最新公告日期、预案公告日、股东大会公告日、首次实施公告日、方案进度、激励标的物、激励方式、激励总数、激励总数占当时总股本比例 (%)、期权初始行权价格、有效期、方案说明、激励股票出售说明、激励授予条件、行权特别条件
	股权激励实施明细	股票代码、公司简称、最新公告日期、期权激励授权日、方案进度、激励标的物、激励方式、激励数量、每股转让价格、期权初始行权价格、提取激励基金数量、股权激励是否全部完成

5.数据范围

数据模块	数据起始时间
管理层薪酬	2001 年
公司股权激励	2005 年

(七) 中国上市公司关联交易数据库-CRTD

1.数据库简介

中国上市公司关联交易数据库 (*Related Transactions Database of Chinese Companies, CRTD*) 搜集整理了 1997 年以来的上市公司关联方交易数据。该数据库由 13 个字段组成, 包括了关联交易的公司简称、会计年度、公告日期、关联交易的具体内容以及关联方的相关信息等。

2.数据库特色

- 数据量大且字段全面

CRTD 囊括了数十万条上市公司关联交易数据, 同时字段全面, 仅与关联交易内容相关的字段就超过十个, 可以充分满足不同方向研究的需求。

3.典型数据举例

上市公司关联交易

股票代码	会计年度	公告日期	关联交易方式	关联方	关联关系	关联交易金额(元)
600209	2010	2011/4/23	提供商品或劳务	海南大宇实业有限公司	参股股东	350604.78
600209	2010	2011/4/23	提供商品或劳务	海南黄金海岸集团有限公司	控股股东	1246895.74
600209	2011	2012/4/28	提供商品或劳务	北京罗顿沙河公司	合营企业	274166.00

4.数据库内容

数据模块	数据表	数据字段
关联交易	关联交易	股票代码、公司简称、会计年度、公告日期、关联交易方式类型编码、关联交易方式、关联方、关联关系、关联交易金额、关联交易定价依据、关联交易简介、关联方与公司是否存在控制关系、支付方式

5.数据范围

数据模块	数据起始时间
全部模块	1997 年

（八）中国公司股权研究数据库-CERD

1.数据库简介

中国公司股权研究数据库 (*Equity Research Database of Chinese Company, CERD*) 是基于对中国上市公司股权结构和股东信息的深入考察,从而建立的详细、准确和全面的公司股权研究学术数据库。*CERD* 包含股权结构、股权变动、股份质押与冻结、股东持股、股份回购 5 个模块,共 9 张表。其中,股权结构包括股本结构、实际控制人;股权变动包括股本变动、重要股权交易;股份质押与冻结包括相应股份的质押与冻结情况;股东持股包括大股东持股、机构持股;最后,股份回购模块展示了公司的股份回购情况。

2.数据库特色

- 覆盖全面

CERD 基本覆盖股权与股东研究的全部方面,例如,该数据库涵盖了机构持股和股份回购等数据,为全面研究上市公司股权信息提供了保障。

3.典型数据举例

上市公司大股东持股

股票代码	统计日期	序号	股东名称	持股数量 (股)	持股比例 (%)
000010	2013/12/31	1	深圳五岳乾坤投资有限公司	176360000	29.99
000010	2013/12/31	2	重庆瑞达投资有限公司	62586842	10.64
000010	2013/12/31	3	信达投资有限公司	49772478	8.46
000010	2013/12/31	4	四川省创源投资管理有限公司	10357315	1.76
000010	2013/12/31	5	北京永安商业有限公司	5178658	0.88
000010	2013/12/31	6	中国纺织机械和技术进出口有限公司	4142926	0.7
000010	2013/12/31	7	戚观家	2826380	0.48
000010	2013/12/31	8	黄贤清	2146560	0.37
000010	2013/12/31	9	杨喜宽	1993600	0.34
000010	2013/12/31	10	南通纺织控股集团纺织染有限公司	1835635	0.31

4.数据库内容

数据模块	数据表	数据字段
股权结构	股本结构	股票代码、公司简称、统计日期、截止日期、总股本、流通 A 股、流通 B 股、流通 H 股、境外流通股、流通股合计、限售 A 股、限售 B 股、限售/国有股、限售/其他内资持股合计、限售/境内法人持股、限售/高管持股、限售/境内自然人持股、限售/外资持股合计、限售/境外法人持股、流通 A 股比例(%)、流通 B 股比例(%)、流通 H 股比例(%)、流通股合计比例(%)、限售 A 股比例(%)
	实际控制人	股票代码、公司简称、会计年度、实际控制人、实际控制人类别、第一大股东名称、第一大股东持股数量、第一大股东持股比例 (%)
股权变动	股本变动	股票代码、公司简称、变动日期、变动原因、变动前总股本、变动前流通 A 股、新增总股本、新增流通 A 股、变动后总股本、变动后流通 A 股

数据模块	数据表	数据字段
股权变动	重要股权交易	股票代码、公司简称、会计年度、公告日期、变动开始日期、变动截止日期、股东名称、股东类型、关联董监高姓名、股东与关联董监高关系、变动方向、变动股数、通过大宗交易系统的变动股数、变动后持股数量、变动后股东持股比例(%)、交易均价、变动期间股票均价、是否为减持解禁、说明
股份质押与冻结	股份质押	股票代码、统计日期、股东名称、质押方、质押股数、质押起始日期、质押截止日期、解押日期、相关说明
	股份冻结	股票代码、统计日期、股东名称、执行冻结机构、冻结股数、冻结起始日期、冻结截止日期、解冻日期、相关说明
股东持股	大股东持股	股票代码、统计日期、序号、股东名称、持股数量、持股比例(%)
	机构持股	股票代码、公司简称、统计日期、公司总股本、公司流通股股本、基金持股家数、基金持股合计、基金持股占流通 A 股比例(%)、基金持仓市值、券商持股家数、券商持股合计、券商持股占流通 A 股比例(%)、券商持仓市值、保险公司持股家数、保险公司持股合计、保险公司持股占流通 A 股比例(%)、保险公司持仓市值、社保基金持股家数、社保基金持股合计、社保基金持股占流通 A 股比例(%)、社保基金持仓市值、QFII 持股家数、QFII 持股合计、QFII 持股占流通 A 股比例(%)、QFII 持仓市值、其他机构持股家数、其他机构持股合计、其他机构持股占流通 A 股比例(%)、其他机构持仓市值
股份回购	股份回购	股票代码、公司简称、公告日期、截止日期、回购金额上限、已回购股数、已回购金额、最高价、最低价、回购进度

5.数据范围

数据模块	数据起始时间
股权结构	1990 年
股权变动	1991 年
股份质押与冻结	1998 年
股东持股	2003 年
股份回购	2012 年

（九）机构投资者持股研究数据库-IORD

1.数据库简介

机构投资者持股研究数据库 (*Institutional Ownership Research Database, IORD*) 包括机构持股与机构重仓股两大模块。机构持股模块主要包括公司基本信息、各大机构持股数量、各大机构持股比例以及持股机构的数量等字段。

2.数据库特色

- 完整

IORD 包含机构重仓股与机构持股两个主模块。重仓股亦包含数个子模块，可以满足不同研究方向学者的需求。

- 详细

IORD 不仅数据量巨大，并且对数据进行了详细的手工处理，方便了数据使用者对数据进行进一步筛选和处理。

3.典型数据举例

QFII 重仓流通股

股票代码	公司简称	机构名称	统计日期	持仓类型	本期持仓数量 (股)	本期占流通股比例 (%)	变动占流通股比例 (%)
600125	铁龙物流	富通银行	2005-09-30	减持	4559485	1.86	-0.49
600125	铁龙物流	富通银行	2005-12-31	增持	8222155	3.35	1.49
600125	铁龙物流	法国巴黎银行	2008-12-31	增持	11745823	1.67	0.99
600125	铁龙物流	法国巴黎银行	2009-06-30	减持	7884623	1.12	-0.55

4.数据库内容

数据模块	数据表	数据字段
机构持股	机构持股	股票代码、公司简称、统计日期、公司总股本、公司流通股股本、基金持股家数、基金持股合计、基金持股占流通 A 股比例 (%)、基金持仓市值、券商持股家数、券商持股合计、券商持股占流通 A 股比例 (%)、券商持仓市值、保险公司持股家数、保险公司持股合计、保险公司持股占流通 A 股比例 (%)、保险公司持仓市值、社保基金持股家数、社保基金持股合计、社保基金持股占流通 A 股比例 (%)、社保基金持仓市值、QFII 持股家数、QFII 持股合计、QFII 持股占流通 A 股比例 (%)、QFII 持仓市值、其他机构持股家数、其他机构持股合计、其他机构持股占流通 A 股比例 (%)、其他机构持仓市值
机构重仓股	基金重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例 (%)、变动占流通股比例 (%)、上期占流通股比例 (%)
	券商重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例 (%)、变动占流通股比例 (%)、上期占流通股比例 (%)

数据模块	数据表	数据字段
机构重仓股	保险重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)
	QFII 重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)
	社保基金重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)
	信托公司重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)
	年金重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)
	财务公司重仓流通股	股票代码、公司简称、机构名称、统计日期、持仓类型、本期持仓数量、变动持仓数量、上期持仓数量、本期占流通股比例(%)、变动占流通股比例(%)、上期占流通股比例(%)

5.数据范围

数据模块	数据起始时间
机构持股	2004 年
机构重仓股	2003 年

(十) 中国公司股利分红数据库-CCDD

1.数据库简介

中国公司股利分红数据库 (*Dividend Database of Chinese Companies, CCDD*) 基于对上市公司股利公告的全面收录, 结合学术研究领域的需求建立了完整和详细的专业数据库。该库搜集整理了1991年至今中国上市公司全部的利润分配数据, 涵盖了会计年度、预案内容、实施进度以及股本基准日、股权登记日、除权除息日、派息日、红股上市日等多维度数据。

2.数据库特色

- 内容完整

数据库囊括了自1991年以来全部A股上市公司的全部分红数据, 可供数据使用者依据自己的需求进行任意选择。

- 覆盖全面

数据库一共两个模块, 基本覆盖了股利分红全部方面, 并且涵盖了预案的详细内容与实施的详细内容, 为全面研究公司股利政策实施提供了保障。

3.典型数据举例

上市公司分红实施内容

股票代码	会计年度	派息比 (税前)	基准股本 (万股)	股本基准日	实施方案公告日	股权登记日	除权除息日	派息日
600246	2000	0.08	9200	2000-12-31	2001/6/1	2001/6/7	2001/6/8	2001/6/15
600246	2001	0.08	9200	2001-12-31	2002/6/25	2002/6/28	2002/7/1	2002/7/8
600246	2002	0.1	9200	2002-12-31	2003/7/8	2003/7/11	2003/7/14	2003/7/18

4.数据库内容

数据模块	数据表	数据字段
分红预案内容	分红预案内容	股票代码、公司简称、会计年度、报告期、预案内容、预案配股比、预案配股价格、预案派息比 (含税)、预案每股派息金额单位、预案送股比、预案转增比
分红实施内容	分红实施内容	股票代码、公司简称、会计年度、报告期、方案实施进度、派息比 (税前)、派息比 (税后)、送股比例、转增比例、基准股本、股本基准日、实施方案公告日、股权登记日、除权除息日、派息日、红股上市日、分红对象

5.数据范围

数据模块	数据起始时间
分红预案内容	1996年
分红实施内容	1991年

(十一) 中国上市公司诉讼仲裁数据库-CLAD

1.数据库简介

中国上市公司诉讼仲裁数据库 (*Litigation & Arbitration Database of Chinese Companies, CLAD*) 是以研究为目的而设计的,旨在提供完整、精确的中国上市公司诉讼仲裁数据。该数据库提供上市公司的诉讼仲裁信息,不仅详细罗列了一审、二审的诉讼情况,还对后续执行情况进行了跟踪,保障了信息的完整性和有效性。

2.数据库特色

- 全面、准确的公司诉讼信息

CLAD 根据上市公司公告、财务报告等信息,全面且准确记录了公司的诉讼与仲裁情况。

3.典型数据举例

上市公司诉讼仲裁统计

股票代码	公告日期	案件名称	原告方	被告方	诉讼类型	涉案金额	起诉日期
000673	2004-12-29	转让款纠纷	大同水泥股份有限公司	海南金牛膜料开发有限公司	民事诉讼	2436万元	2003-08-22
600168	2004-12-29	托管纠纷	武汉三镇实业控股股份有限公司	天元控股有限公司	民事诉讼	2768万元	2004-08-30

4.数据库内容

数据模块	数据表	数据字段
上市公司诉讼仲裁统计	上市公司诉讼仲裁统计	股票代码、公司简称、当事人类型、公告日期、案件名称、案件简介、原告方、被告方、诉讼类型、涉案金额、币种、起诉日期、一审受理法院、一审判决日期、一审判决内容、是否上诉、上诉方、二审受理法院、二审判决日期、二审判决内容、执行情况

5.数据范围

数据模块	数据起始时间
上市公司诉讼仲裁统计	1995年

（十二）中国上市公司内部人交易数据库-CITD

1.数据库简介

中国上市公司内部人交易数据库 (*Insider Trading Database of Chinese Companies, CITD*) 搜集整理了 1996 年以来的上市公司内部人股权交易数据。该数据库由 19 个字段组成, 包括内部人交易的公司简称、公告日期、内部人交易涉及的股东和高管信息以及内部人股权交易的具体信息。

2.数据库特色

- 数据量大且字段全面

CITD 包括十万条以上的内部人交易数据, 与内部人交易具体内容相关的字段达到 15 个以上, 可以充分满足不同方向研究的需求。

3.典型数据举例

上市公司内部人交易

股票代码	公告日期	内部人名称	关联董监高姓名	与关联董监高关系	变动方向	变动股数 (股)	交易均价 (元)
000001	2007/7/17	王魁芝	王魁芝	本人	增持	1600	27.90
000001	2013/4/11	白国华	孙先朗	配偶	增持	4000	20.05
000002	2007/5/30	杨新宇	莫军	配偶	减持	20000	21.90
000002	2013/1/26	丁长峰	丁长峰	本人	增持	850000	8.66

4.数据库内容

数据模块	数据表	数据字段
内部人交易	内部人交易	股票代码、公司简称、会计年度、公告日期、变动开始日期、变动截止日期、内部人名称、内部人类型、关联董监高姓名、内部人与关联董监高关系、变动方向、变动股数、通过大宗交易系统的变动股数、变动后持股数量、变动后内部人持股比例 (%)、交易均价、变动期间股票均价、是否为减持解禁、说明

5.数据范围

数据模块	数据起始时间
全部模块	1996 年

经济研究基础数据

(一) 中国宏观经济研究数据库-MACRO

1.数据库简介

中国宏观经济研究数据库 (*Chinese Macro-economic Research Database, MACRO*) 是以研究为目的设计的专业学术研究数据库,旨在为国内外从事宏观经济研究的学者和研究人员提供完整、可靠的学术研究数据。

本数据库包含以下 28 个方面的经济数据指标:综合;国民经济核算;人口;就业人员和工资;固定资产投资和房地产;对外经济贸易;能源;财政;价格指数;人民生活;城市概况;资源和环境;农业;工业;建筑业;运输和邮电;社会消费品零售总额;住宿和餐饮业;旅游业;金融业;教育;科技;卫生;社会服务;文化;体育;公共管理、社会保障及其他。

2.数据库特色

- 数据来源权威

MACRO 主要根据国家统计局、地方统计机构以及统计年鉴等多个渠道数据整合而成。

- 数据覆盖全面

MACRO 从 28 个方面反映了我国自 1995 年以来 20 多年的整体、宏观的经济变化过程与社会发展情况,囊括了所有行业与方面的经济数据指标。

3.数据库内容

数据模块	
综合	建筑业
国民经济核算	运输和邮电
人口	社会消费品零售总额
就业人员和工资	批发和零售业
固定资产投资和房地产	住宿和餐饮业
对外经济贸易	旅游业
能源	金融业
财政	教育
价格指数	科技
人民生活	卫生
城市概况	社会服务
资源和环境	文化
农业	体育
工业	公共管理、社会保障及其他

4.数据范围

数据模块	数据起始时间
所有模块	1995 年

（二）中国区域经济研究数据库-CRED

1.数据库简介

中国区域经济研究数据库 (*Chinese Regional Economy Database, CRED*) 从 28 个方面反映了我国 31 个省市的区域经济与社会发展状况, 囊括了各种类型的区域经济数据指标。*CRED* 包含: 综合; 国民经济核算; 人口; 就业人员和工资; 固定资产投资和房地产; 对外经济贸易; 能源; 财政; 价格指数; 人民生活; 城市概况; 资源和环境; 农业; 工业; 建筑业; 运输和邮电; 社会消费品零售总额; 批发和零售业; 住宿和餐饮业; 旅游业; 金融业; 教育; 科技; 卫生; 社会服务; 文化; 体育; 公共管理、社会保障及其他等 28 个方面的经济数据指标, 涉及 31 个省市。

2.数据库特色

- 数据来源权威

CRED 主要根据国家统计局、地方统计机构以及统计年鉴等多个渠道数据整合而成。

- 数据覆盖全面

CRED 数据库囊括了 1995 年以来近 20 多年的主要社会经济统计指标, 可为学者提供完善的研究数据。

3.数据库内容

数据模块	
综合	建筑业
国民经济核算	运输和邮电
人口	社会消费品零售总额
就业人员和工资	批发和零售业
固定资产投资和房地产	住宿和餐饮业
对外经济贸易	旅游业
能源	金融业
财政	教育
价格指数 (上年=100)	科技
人民生活	卫生
城市概况	社会服务
资源和环境	文化
农业	体育
工业	公共管理、社会保障及其他

4.数据范围

数据模块	数据起始时间
所有模块	1995 年

(三) 中国债券研究数据库-BOND

1.数据库简介

中国债券研究数据库 (*Chinese Bond Research Database, BOND*) 涵盖债券市场各方面信息, 不仅收录债券及债券发行主体的信息, 还对债券发行时的重要信息进行汇总, 同时也整理了债券的信用评级及债券发行主体的财务报表信息。**BOND** 包含 5 个数据模块, 即债券基本信息、债券发行主体基本信息、债券发行信息、债券信用评级、债券发行主体财务报表。

2.数据库特色

- 内容详细完善

中国债券研究数据库对与债券相关的各个维度信息进行了详细描述, 全面汇总了债券市场几乎所有交易品种的全方位数据。该数据库不仅包含债券本身基本信息和债券发行主体基本信息, 还汇总了债券及债券发行主体的信用评级信息。

- 数据准确性高

为确保本数据库的数据准确性, 本数据库在设计字段和汇总数据信息时采用三级数据复核制度, 以确保数据的准确性和完整性。较高的数据准确性为学者进行债券市场研究提供了基础条件。

3.典型数据举例

债券基本属性

债券代码	债券简称	发行面值	发行总额	债券期限	票面利率	计息基准
101551098.IB	15 龙源电力 MTN001	100	3000000000	5	4.44%	ACT/ACT
101558030.IB	15 龙岩汇金 MTN001	100	500000000	5	5.38%	ACT/ACT
041555045.IB	15 龙岩工贸 CP001	100	200000000	1	4.07%	A/366

4.数据库内容

数据模块	数据表	数据字段
债券基本信息	债券基本属性	债券代码、债券简称、债券全称、债券发行面值、发行总额、起息日期、计息截止日、到期日期、债券期限、债券期限备注、利率类型、债券发行票面利率(%)、债券票面利率备注、计息方式、计息基准、每年付息次数、付息日说明、上市日期、上市地点、发行币种、摘牌日期
	债券分类	债券代码、债券简称、中债债券一级分类、中债债券二级分类、NAFMII 债券一级分类、NAFMII 债券二级分类、上清所债券分类、是否城投债、城投行政级别、是否跨市场交易、是否次级债、是否混合资本债券
	债券担保信息	债券代码、债券简称、担保人、担保人公司简介、担保人公司属性、担保方式、担保条款文字
债券发行主体基本信息	发行主体基本信息	债券代码、债券简称、债务主体、发行人中文简称、成立日期、注册资本、法定代表人、董事长、总经理、经营范围、主要产品及业务、员工人数、省份、城市、注册地址、办公地址、是否上市公司、公司属性
	发行主体行业分类	债券代码、债券简称、所属证监会大类行业名称、所属证监会大类行业代码

数据模块	数据表	数据字段
债券发行主体基本信息	发行主体股东信息	债券代码、债券简称、统计日期、大股东名称、大股东持股比例(%)、大股东类型
债券发行信息	债券发行重要日期	债券代码、债券简称、发行起始日期、发行截止日期、发行公告日、上市公告日
	债券发行中介机构	债券代码、债券简称、主承销商、会计师事务所、律师事务所
债券信用评级	债券发行时信用评级	债券代码、债券简称、发行时债项评级、发行时主体评级、发行时主体评级展望、发行人委托评级机构
	债券历史信用评级	债券代码、债券简称、公司名称、发布日期、评级类型、信用评级、评级机构
	债券发行主体历史信用评级	债券代码、债券简称、公司名称、发布日期、评级类型、信用评级、评级机构
债券发行主体财务报表	资产负债表	债券代码、债券简称、统计日期、货币资金、交易性金融资产、应收票据、应收账款、预付账款、其他应收款、存货、流动资产合计、长期应收款、长期股权投资、投资性房地产、固定资产、在建工程、无形资产、开发支出、商誉、递延所得税资产、非流动资产合计、资产总计、短期借款、交易性金融负债、应付票据、应付账款、预收账款、应付职工薪酬、应交税费、应付利息、应付股利、短期应付债券、一年内到期的非流动负债、流动负债合计、长期借款、应付债券、预计负债、递延所得税负债、其他非流动负债、非流动负债合计、负债合计、实收资本(股本)、资本公积、盈余公积、未分配利润、归属于母公司所有者权益合计、所有者权益合计
	利润表	债券代码、债券简称、统计日期、营业总收入、主营业务收入、主营业务成本、营业税金及附加、销售费用、管理费用、财务费用、资产减值损失、公允价值变动损益、投资收益、营业利润、营业外收入、营业外支出、利润总额、所得税费用、净利润、归属于母公司所有者的净利润
	现金流量表	债券代码、债券简称、统计日期、销售商品、提供劳务收到的现金、收到的税费返还、购买商品、接受劳务支付的现金、支付给职工以及为职工支付的现金、支付的各项税费、经营活动产生的现金流量净额、收回投资收到的现金、取得投资收益收到的现金、处置固定资产、无形资产和其他长期资产收回的现金、处置子公司及其他营业单位收到的现金、购建固定资产、无形资产和其他长期资产支付的现金、投资支付的现金、取得子公司及其他营业单位支付的现金、投资活动产生的现金流量净额、吸收投资收到的现金、发行债券收到的现金、取得借款收到的现金、偿还债务支付的现金、分配股利、利润或偿付利息支付的现金、筹资活动产生的现金流量净额、现金及现金等价物净增加额

5.数据范围

数据模块	数据起始时间
全部模块	1990年

四、中国财经大数据研究系统

1. 数据库简介

目前大数据分析技术的方法已经开始与资本市场研究相结合，使用的方法各不相同，但主要对象都是对资本市场文本信息的处理，具体包括分析师报告、年报、IPO 申报稿、公司公告和新闻报道等。迄今国内与财经大数据研究相关的文章还较少，主要原因在于数据获取的困难和跨学科结合的障碍。为解决这一难题，我司结合资本市场资深专家、文本挖掘和财经领域知名教授的专业知识，嵌入较前沿的大数据分析技术，开发了中国财经大数据研究系统，简称 **CFBD** (*Chinese Financial Bigdata Systems*)。

2. CFBD 包含以下六大模块

- 财经新闻库

本数据库包括来自主流财经网站、大型重要网站、行业网站等近 400 家网络媒体的财经新闻报道。这些财经新闻囊括个股新闻、宏观经济报道、行业报道等，由于该库提供最原始的财经新闻信息，因此并不是每个新闻都能匹配到对应个股，但是这些全面、完整的财经新闻信息为学者研究多维度财经信息提供了很好的文本来源。

- 网络新闻库与网络量化舆情库

本数据库包括来自近 400 家网络媒体报道能够匹配到个股的新闻信息。提供包括所属公司、股票代码、新闻内容、发布时间、新闻正负面、报道媒体、报道媒体所属地区、来源媒体等多个字段。量化舆情提供新闻属性表、舆情统计表（按自然日）、舆情统计表（按交易时间）、新闻相似度四张表。新闻属性表提供了股票代码、新闻情感、报道媒体、发布时间、原创性等多个字段；网络财经新闻量化统计（按自然日）根据自然日时间（0:00-24:00）对一天中的新闻数量进行统计；网络财经新闻量化统计（按交易时间）则根据交易时间统计了不同时间段内的新闻文章情况，这种计算方法可以直接反映交易时间内和非交易时间内新闻报道对市场的影响；网络财经新闻相似度则使用文本分析方法，计算分析了两篇新闻的相似度，本数据库只保留相似度大于 0.5 的结果。

- 报纸新闻库和报纸量化舆情库

除了网络新闻，**CFBD** 还包括学术研究比较关注的纸质报纸和报刊中的财经信息。本数据库不仅包括媒体研究常用到的六大报纸，还包括总共 500 多家主流报纸的财经新闻报道，提供的字段内容也比较丰富。与网络量化舆情库类似，报纸量化舆情库包括新闻属性表、舆情统计表、新闻相似度三张表。

- 股吧评论库

CFBD 主要根据中国最大股吧论坛的帖子信息，对上市公司发帖评论标题、时间、内容、阅读量等信息进行文本分析、数量统计等，从中得到维度众多，数量丰富的研究数据。其中，日帖子统计（按自然日）根据自然日时间（0:00-24:00）对一天中的发帖总量、点击总量和评论总量进行统计；日帖子统计（按交易时间）则根据交易时间统计了不同时间段内的发帖情况，这种计算方法可以直接反映交易时间内和非交易时间内投资者情绪对股市的影响。

- 公司关注指数

CFBD 整合了主流搜索引擎对上市公司的搜索指数，提供上市公司每天的搜索趋势，跟踪投资者对上市公司每日的关注情况。该数据库不仅包括全部公司关键字的网络搜索指数加总值，为了研究需要，还包括以股票代码为关键字的搜索值。同时，该库还分地区（地级市）展示了不同区域的搜索指数。

- 财经文本分析

基于大数据技术开发的财经文本自动分析系统，学者可以自己上传文章，通过该系统进行自动分词、词频统计、关键词统计、正负面分析以及文本分类。为了提高分析的准确率，该模块不仅提供不同类型的财经文本分析，如年报、分析师报告、招股说明书、公司公告、新闻报道等，还提供不同的机器学习方法，如 SVM、NB、Rocchio 算法、KNN 和神经网络等方法，以满足不同的研究需要。

3.数据库特色

- 全面覆盖

丰富多元的底层数据来源，包括财经新闻、搜索指数、主流贴吧等市场关注信息；近千家媒体机构，网络媒体与纸质报刊全覆盖；2.5 亿条历史新闻和评论记录。

- 算法精确

采用分布式抓取、清洗、分类和识别，数十道数据清洗程序，拒绝无关信息；国内外计算机和财经专家教授深度参与，确保产品质量；最新机器学习技术，语气识别正确率高达 85%。

- 学术市场唯一的财经大数据研究系统

CFBD 旨在为高校老师和研究者研究中国资本市场大数据提供数据来源和分析方法，弥补了我国会计界大数据研究的空白。

五、最新学术动态

(一) 模块简介

为了使广大学者便于跟踪国内外权威学术期刊最新论文等信息，**CNRDS** 开辟了“最新学术动态”一栏。该栏目将定期及时更新国内外顶级财经期刊的文章、国内外学术会议信息。

(二) 模块功能

1.及时更新国内外各类权威期刊学术动态

CNRDS 可为您提供便于浏览的学术文章题目、摘要、链接地址等重要信息。该栏目将及时更新来自 20 本会计金融类国际顶级期刊上的最新论文列表，以及来自 13 本中国财经类权威期刊的最新论文列表。

2.总结近期国际期刊上的中国问题研究文章

来自 20 本会计金融类国际顶级期刊上的中国问题研究文章，可为学者浏览、查阅最前沿的学术论文，扩展研究思路提供有益参考。

六、数据定制

（一）定制项目

1.数据咨询

该服务主要为海内外高校学者免费提供中国学术研究所需数据的来源以及数据质量的咨询和建议。根据您提出的问题，我们将比较中国已有数据库之间的差异，找出最适合您需求的数据库；针对某些您不确定在中国是否存在的数据或者数据来源不明确的数据，我们将为您提供免费的数据咨询服务。我们的咨询团队都是由国内一流的数据和财经领域的专家组成，他们不仅长期关注中国资本市场的发展，还与证监会、银监会、财政部等中国资本市场监管部门保持良好的交流与联系。他们所提供的咨询服务能为您节约研究成本，提高研究效率。

2.RA 服务

我们可为国内外学者提供收集与整理数据服务，我司专业、高效的 RA 团队能为您及时提供准确的研究数据。您既可以将您需要搜集的数据项目告诉我们，也可以将您已有的数据项目提供给我们，我们将秉承准确性、全面性、及时性、保密性原则，在第一时间为您收集和整理出电子化的标准格式数据。

3.学术数据定制

该服务将根据您在研究和实务中对独特数据的需要，为您提供专业的财经类数据查找、搜集、整理等定制服务。我们拥有长期、稳定的数据采集渠道，高效、可靠的数据整理团队，采用网页爬虫、数据抓取、大数据标注、人工搜集和整理等手段和方法，能够在第一时间获取您所需的定制数据。

（二）定制流程



1.提交数据定制需求

您可通过电话、邮件、微信或在线留言等方式提交您的数据定制需求。联系邮箱：lihui@cnrds.com；联系固话：021-66181082；手机：15000386334（李慧，数据服务部经理）。

2.评估与反馈

我司对您的数据定制需求进行评估，并会在 3 个工作日内把数据可获得性、数据样本、报价以及完成所需时间等信息反馈给您。

3.签订合同

我们双方之间就数据、价格、交付时间等方面达成一致后，我司拟定数据定制电子版合同电邮给您。最终确定好合同各项条款后，双方通过邮寄或扫描等方式签订正式的数据定制合同。

4.付款

我司可以配合您的工作单位科研经费的报销流程，出具合同、发票等财务报销相关资料。

5.交付数据

我司按合同约定的时间、交付方式交付数据。

6.售后跟踪

交付数据后，为了了解您对定制数据的使用情况，我司会通过邮件、电话或实地拜访等对您进行回访，以便及时、完整地解答您使用数据过程中所遇到的问题。

七、数据合作

（一）数据合作

我们诚邀各专业领域的数据库公司、相关领域的专家和学者与 **CNRDS** 平台进行数据合作，我们相信每一位学者在支持和推动中国研究基础数据和资源构建上迈出的小小一步，都将成为中国学术研究工作的一大步。

（二）合作流程



1.合作咨询

若您有需要商业化的特色学术数据，可与我们联系，并咨询我司数据研发部工作人员，联系邮箱：liuchaoran@cnrds.com；联系电话：13816318829（刘超然，数据研发部）。

2.提交数据

我们双方确定好合作意向后，您可通过邮件等方式发送合作数据的样本给我司。

3.专家团队评审

我方邀请相关领域的专家和学者对您的数据合作项目进行匿名审核。我方会在 10 个工作日内把专家团队的匿名评审结果反馈给您。

4.签订合同

我们双方之间就数据合作各项细节达成一致后，我方拟定数据合作合同电邮给您，最终确定好合同各项条款后，双方通过邮寄或扫描等方式签订正式的数据合作合同。

5.制作数据库并上线

您有以下两种数据呈现形式：第一，数据内容由贵研究团队独立制作，我方提供技术支持和形式审核以及后续维护；第二，贵研究团队进行顶层设计，我方进行数据库内容搜集和整理，并提供制作和后续维护。最后，我司将制作好的数据库发布在 **CNRDS** 平台上。

合作高校

截止目前，CNRDS为海内外诸多高校和科研机构提供了数据服务，其中不乏国内顶尖高校和海外知名学府。我司所提供的学术研究数据库、财经大数据、数据定制、数据采集等产品和服务得到了一致好评和肯定。

国内合作高校



海外合作高校



◆ 国内外知名高校教授合作开发

各领域专家深度参与，严格的数据质量控制体系，为学术研究提供高质、准确、权威的研究数据

◆ 特色数据库满足各类研究需求

结合中国问题研究，紧跟学术热点和前沿，提供各类特色研究数据，为高质量学术研究提供坚实的数据保障

◆ 界面与模块设计契合学术研究

参阅各领域重点学术文献，整合学术研究各类基础数据，提供科学、便捷的数据浏览、查询和下载服务

CNRDS 中国研究数据服务平台
Chinese Research Data Services Platform

Website: <http://www.cnrds.com>

数据服务部

联系人：李慧

电话：(+86) 021-66181082

手机：(+86) 15000386334

邮箱：lihui@cnrds.com

地址：上海市宝山区长江南路180号长江软件园



CNRDS微信公众号