法学大数据分析平台

价值与应用

北京大学实证法务研究所
北京法意科技有限公司

2016年10月

目录

实证研究难点

平台价值特色

功能模块及示例

实证研究难点



01课题探索

- 缺少辅助工具
- 探索与筛选不连贯



02研究样本

- 少
- 不权威
- 更新不及时



03变量设计

- 变量梳理费时费力
- 不能快速扩展



04统计方法

- 缺少统计工具
- 缺少方法指引



05课题管理

- 维护成本高
- 不能持续比较

实证研究难点

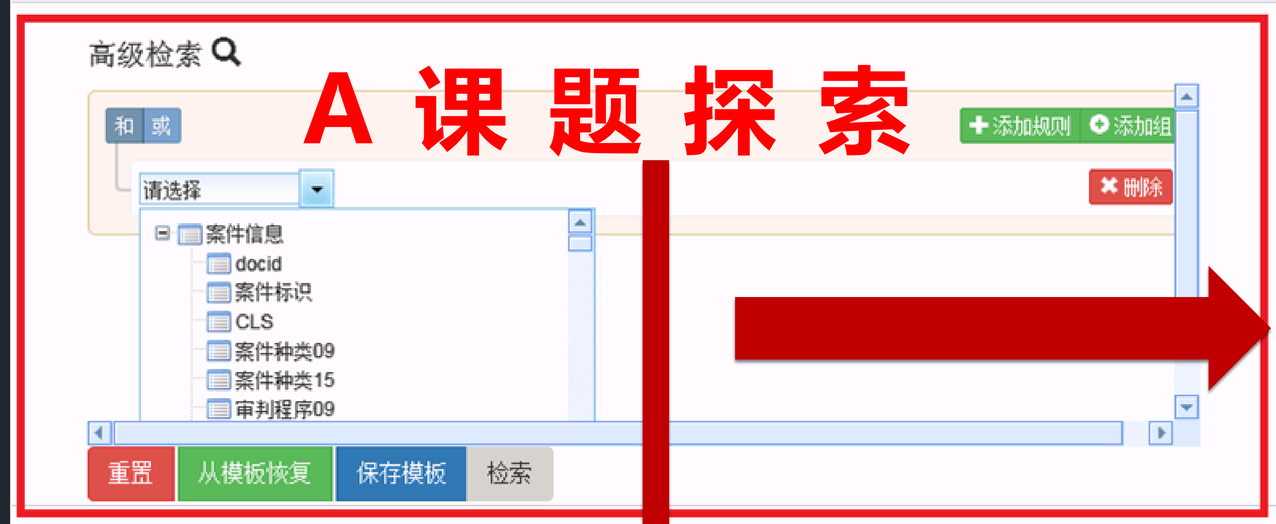
平台价值特色

功能模块及示例

01
探索

探索与研究
紧 密 结 合

探索与研究


高级检索
A 课 题 探 索
和
或
添加规则
添加组
请选择
删除
案件信息
docid
案件标识
CLS
案件种类09
案件种类15
审判程序09
重置
从模板恢复
保存模板
检索


法院层级
裁判日期
审判程序
B 列 表 展 现
全部选入
操作
案由
裁判日期
结案方式
1 查看 赤峰市元宝山区人民法院 （2015）元民初字第2212 民间借贷纠纷 2015/6/23 判决
2 查看 德阳市旌阳区人民法院 （2015）旌执字第979号 金融借款合同纠纷 2015/9/23 终结
3 查看 河南省中牟县人民法院 （2015）牟民初字第4557 民间借贷纠纷 2015/12/29 裁定撤诉
4 查看 广东省东莞市第二人民法 （2015）东二法刑初字第 过失致人死亡 2015/11/25 判决
5 查看 广东省东莞市第二人民法 （2015）东二法刑初字第 过失致人死亡 2015/11/25 判决
6 查看 安徽省合肥市中级人民法 （2015）合刑执字第 2015/12/29 0
7 查看 上海市第二中级人民法院 （2015）沪二中民一 民间借贷纠纷 2015/12/10 裁定按撤诉处理
8 查看 安徽省宿州市中级人民法 （2015）宿中刑终字第 骗取贷款、票据承兑、金 2015/12/12 维持
10
1 /12515
显示从1到10，总 125141 条。每页显示：10


C 分 布 状 况
单变量频次分析
按法院区域
安徽 2
北京 2
福建 1
甘肃 1
广东省 556
广西壮族自治区 82
安徽省 530
北京市 598
福建省 1684
甘肃省 226
广西 1
贵州省 264
按裁判年份
2001 1
2003 1
2005 1
2007 8
2009 12
2002 1
2004 2
2006 1
2008 6
2010 32
按审判程序
催告 10
督促 1
非诉财产保全审查 1
复核 1
行政赔偿 7
第三人撤销之诉 2
二审 2046
非诉行为保全审查 108
管辖 16
破产 1
按结案事由
被告人积极退赃 3
被告人有重大立功表现 1
不满足自诉案件范围 1
不属于民事诉讼范围 18
出现了新证据 7
当事人之间订有有效仲裁协议 1
被告人取得被害人谅解 2
被告死亡，没有遗产，也没有应当承… 2
不属于本院管辖 1
超过起诉期限且无正当理由 10
错列被告且拒绝变更 1
法律规定行政复议为必经程序未经过… 1

02
样本

样 本 完 整

来 源 权 威

同 步 更 新

样本完整

视为全样本

2200万+

来源权威

最高人民法院正式授权使用

裁判文书网资源

同步更新

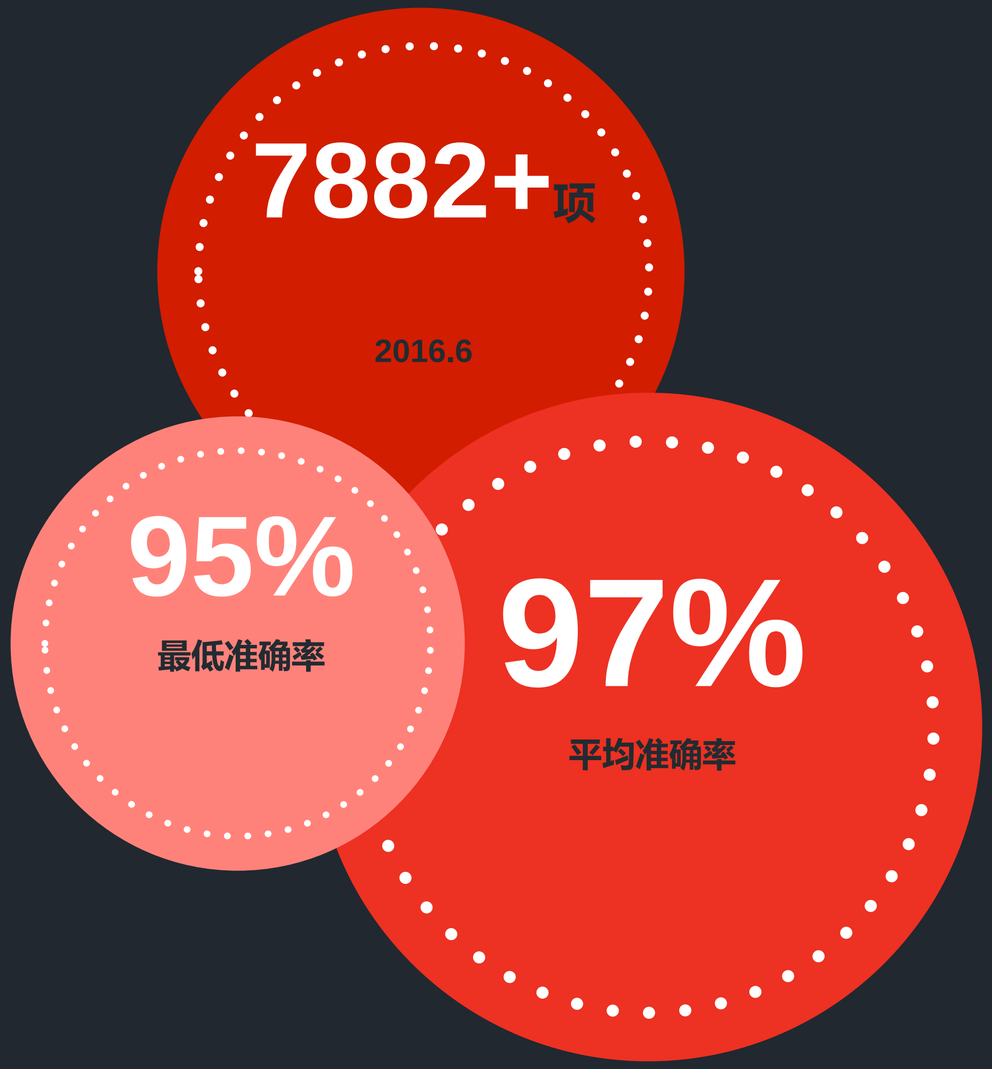
《中国裁判文书网》开发运营商，案例数据与中国裁判文书网保持同步

03
变量

量 多 又 准

自定义扩展

量多又准



刑事，4374+

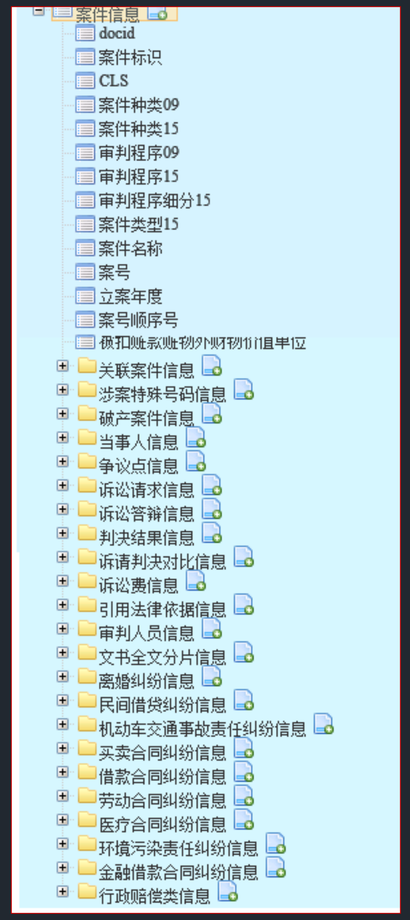
民事，1929+

行政，1205+

执行，196+

赔偿，178+

自定义扩展



基础变量



设置变量信息



		专利侵权抗辩事由											
		其他	抵触申请抗辩	非生产经营目的抗辩	合法来源抗辩	科研及实验目的抗辩	滥用权利抗辩	权利懈怠抗辩	权利用尽抗辩	无效称事由	先用权抗辩	现有技术或设计抗辩	医药行政审批抗辩	合计
	其他	2444	0	112	135	6	7	49	0	6363	38	9	1	9164
	败诉	628	1	110	84	0	26	10	5	18	14	16	7	919
	胜诉	2486	0	623	446	1	60	43	2	630	29	15	6	4341
合计		5558	1	845	665	7	93	102	7	7011	81	40	14	14424

结果预览

20+ 函数关系选择

"等于"
"不等于"
"在...之内"
"不在...之内"
"小于"
"小于或等于"
"大于"
"大于或等于"
"在...之间"
"不在...之间"
"以...开始"
"不以...开始"
"包含以下内容"
"不包含以下内容"
"以...结束"
"不以...结束"
"为空"
"不为空"
"为 null"
"不为 null"

04
统计方法

线上统计分析工具支撑

实证研究领域专家支持

统计分析工具

单变量频次分析

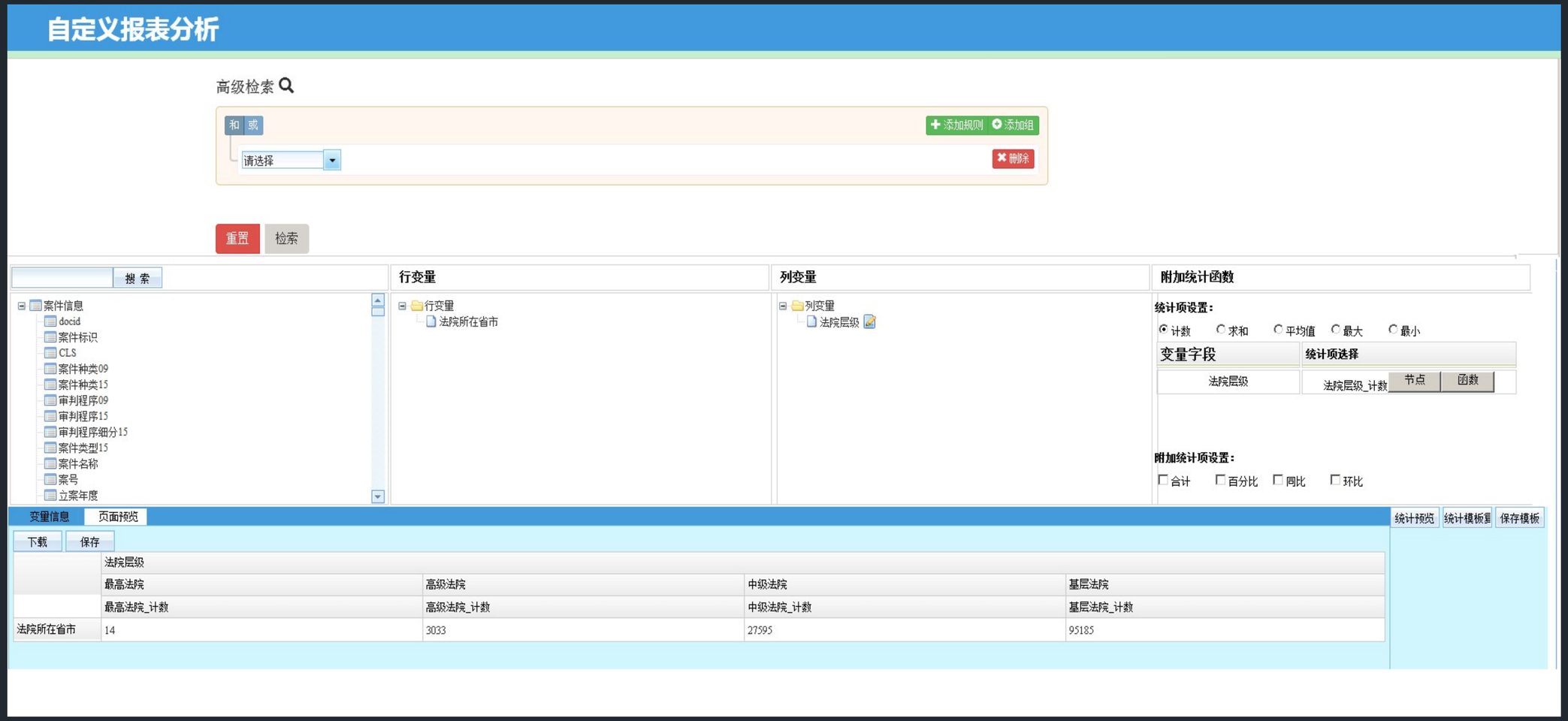
单变量频次分析是针对单个变量按照变量取值类型来统计每个取值类型出现的次数。比如针对犯罪人性别的频次分析，就是统计不同性别的犯罪人人数各为多少，即男性犯罪人人数为多少，女性犯罪人人数为多少

统计项设置： 结案主案由原文　　图形切换： 请选择　　是否显示统计值

下载　保存

	频次	百分比	有效百分比	累计百分比
滥伐林木	11	0.01%	0.01%	0.01%
销售假冒注册商标的商品	2	0%	0%	0.01%
电信服务合同纠纷	27	0.02%	0.02%	0.03%
侵害作品放映权纠纷	22	0.02%	0.02%	0.04%
对非国家工作人员行贿	2	0%	0%	0.05%
股东出资纠纷	7	0.01%	0.01%	0.05%
运输合同纠纷	26	0.02%	0.02%	0.07%
与企业有关的纠纷	5	0%	0%	0.07%
意外伤害保险合同纠纷	4	0%	0%	0.08%
网络侵权责任纠纷	1	0%	0%	0.08%
财产保险合同纠纷	32	0.02%	0.02%	0.1%
挂靠经营合同纠纷	11	0.01%	0.01%	0.11%
引诱、容留、介绍卖淫	4	0%	0%	0.11%
申请保全案件	11	0.01%	0.01%	0.12%
普通破产债权确认纠纷	1	0%	0%	0.12%
过失致人重伤	2	0%	0%	0.12%
侵占	3	0%	0%	0.12%
土地承包经营权纠纷	19	0.01%	0.01%	0.14%
票据纠纷	1	0%	0%	0.14%
共有权确认纠纷	2	0%	0%	0.14%

统计分析工具


自定义报表分析
高级检索
和 或
添加规则 添加组
请选择
删除
重置
检索
搜 索
行变量
列变量
附加统计函数
案件信息
docid
案件标识
CLS
案件种类09
案件种类15
审判程序09
审判程序15
审判程序细分15
案件类型15
案件名称
案号
立案年度
行变量
法院所在省市
列变量
法院层级
统计项设置：
计数 求和 平均值 最大 最小
变量字段
统计项选择
法院层级
法院层级_计数
节点 函数
附加统计项设置：
合计 百分比 同比 环比
变量信息 页面预览
统计预览 统计模板 保存模板
下载 保存
法院层级
最高法院 高级法院 中级法院 基层法院
最高法院_计数 高级法院_计数 中级法院_计数 基层法院_计数
法院所在省市 14 3033 27595 95185

北京大学实证法务研究所

领域专家支持

Image
Place Holder

your required image
just drug into me

领域专家支持

“同案同判的宪政意义及其实证研究”

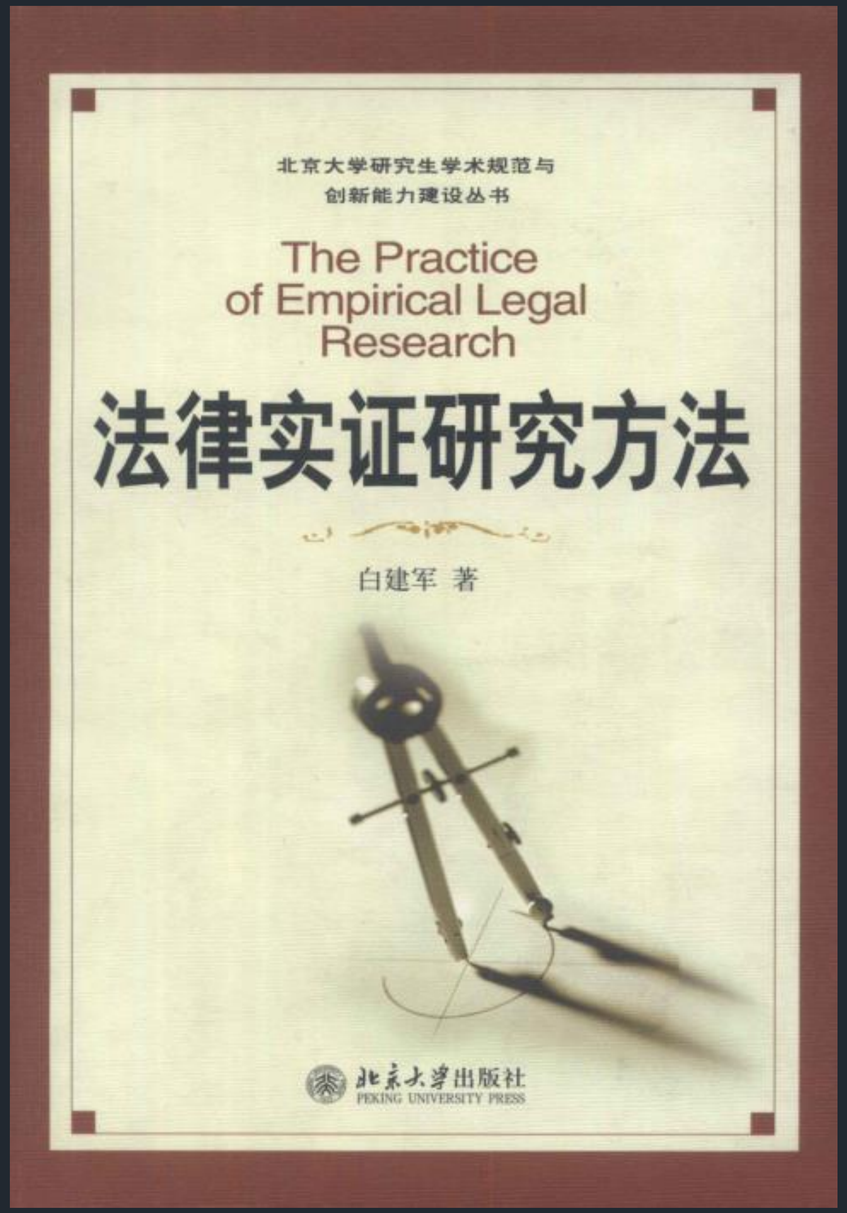
“刑罚轻重的量化分析”

“量刑基准实证研究”

“金融骗局的被害现象和要因分析”

“证监会60个处罚决定的实证分析”

…….



领域专家支持

浙江大学——“以互联网领域的侵犯知识产权类犯罪（假货犯罪）相关案件的电子证据应用情况”；“假货犯罪相关案例电子数据应用分析”

中山大学——“香港立法会质询权行使情况数据统计需求”

复旦大学——“环境公益诉讼案件整理”

北京师范大学——“农民工讨薪案例实证分析”

山西大学——“非法证据排除案例及数据分析”

江南大学——“互联网知识产权交易平台制度创新”

.....

05

报告可视化展现

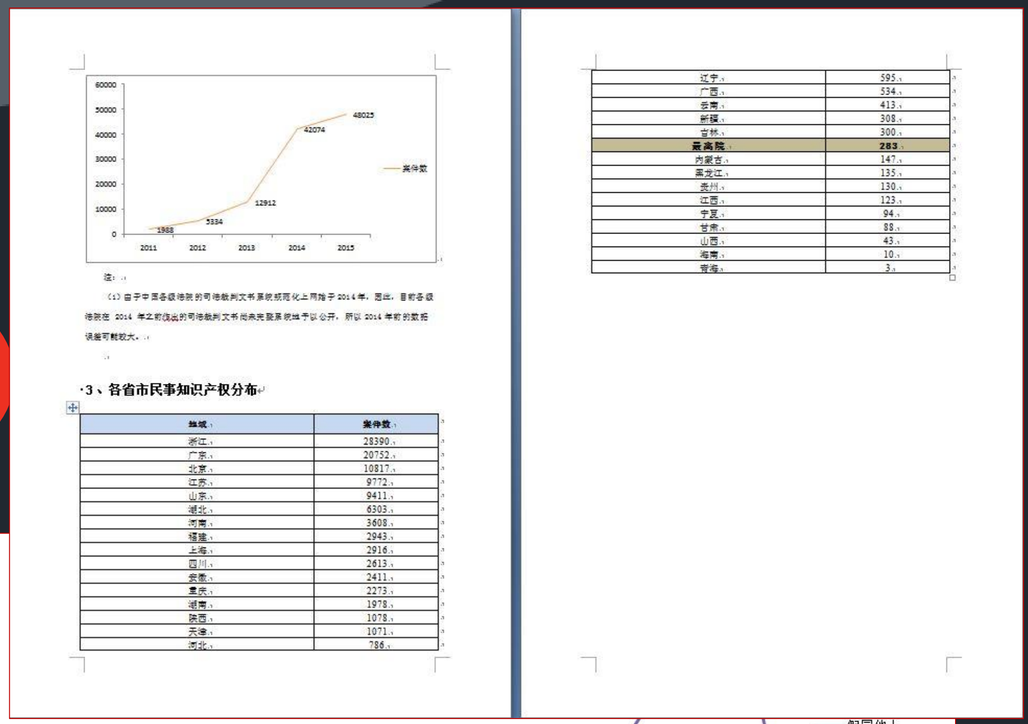
自由选择可视化效果

表格、饼图、柱状图、折线图、条形图、报告...

		专利侵权结案案由					
		假冒他人专利纠纷	侵害发明专利权纠纷	侵害实用新型专利权纠纷	侵害外观设计专利权纠纷	确认不侵害专利权纠纷	合计
裁判年份	缺失样本	1	76	198	477	1	753
	1970	0	0	0	1	0	1
	1991	0	0	1	1	0	2
	1992	0	2	0	0	0	2
	1993	0	1	2	0	0	3
	1994	0	3	2	2	0	7
	1995	0	1	6	1	0	8
	1996	0	3	3	1	0	7
	1997	0	4	2	4	0	10
	1998	0	4	2	2	0	8
	1999	0	13	25	21	0	59
	2000	0	7	31	17	0	55
	2001	0	7	32	21	0	60
	2002	0	8	38	28	0	74
	2003	0	25	110	74	0	209
	2004	0	43	86	145	1	275
	2005	0	113	203	328	1	645
	2006	1	136	252	388	6	783
	2007	1	112	183	402	17	715
	2008	0	155	283	518	13	969
	2009	1	206	428	887	9	1531
	2010	1	224	478	950	5	1658
	2011	0	293	510	1146	12	1961
	2012	7	236	418	1176	5	1842
	2013	0	341	622	1417	3	2383
	2014	2	183	315	804	3	1307
合计		14	2196	4230	8811	76	15327

01

02

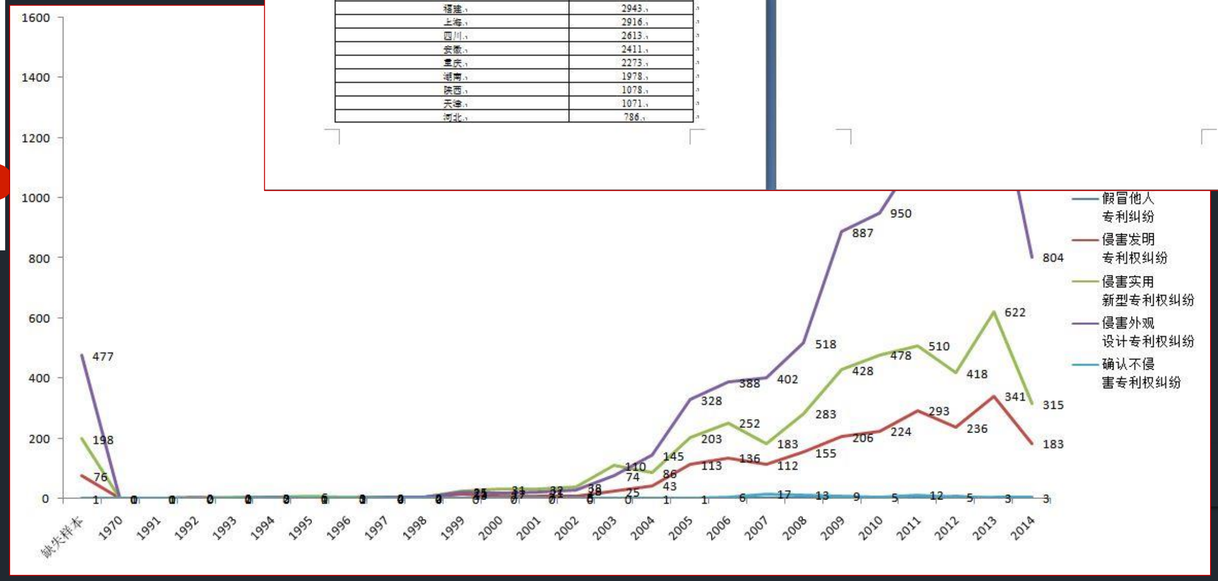


注：

（1）由于中国各级法院的司法裁判文书系统规范化上网始于2014年，因此，目前各级法院在2014年之前上网的司法裁判文书尚未完整系统地予以公开，所以2014年前的数据误差可能较大。

·3、各省市民事知识产权分布

地域	案件数
浙江	28390
广东	20752
北京	10817
江苏	9772
山东	9411
湖北	6303
河南	3608
福建	2943
上海	2916
四川	2613
安徽	2411
重庆	2273
湖南	1978
陕西	1078
天津	1071
河北	786
辽宁	595
广西	534
云南	413
新疆	308
吉林	300
最高院	283
内蒙古	147
黑龙江	135
贵州	130
江西	123
宁夏	94
甘肃	88
山西	43
海南	10
青海	3



06
课题自助管理

课题样本自由增删

分析成果实时跟进比对

07

课题增值服务

确立研究主题

选题

需求沟通确认

确认

主题分析实施

实施

① 研究框架设计
② 样本资源获取
③ 变量规则制作
④ 变量统计分析
⑤ 分析结果说明

分析结果发布

发布



应用示例

法学大数据分析平台功能模块&示例

功能
模块



01. 样本筛选



02. 选入课题



03. 统计分析



04. 研究报告



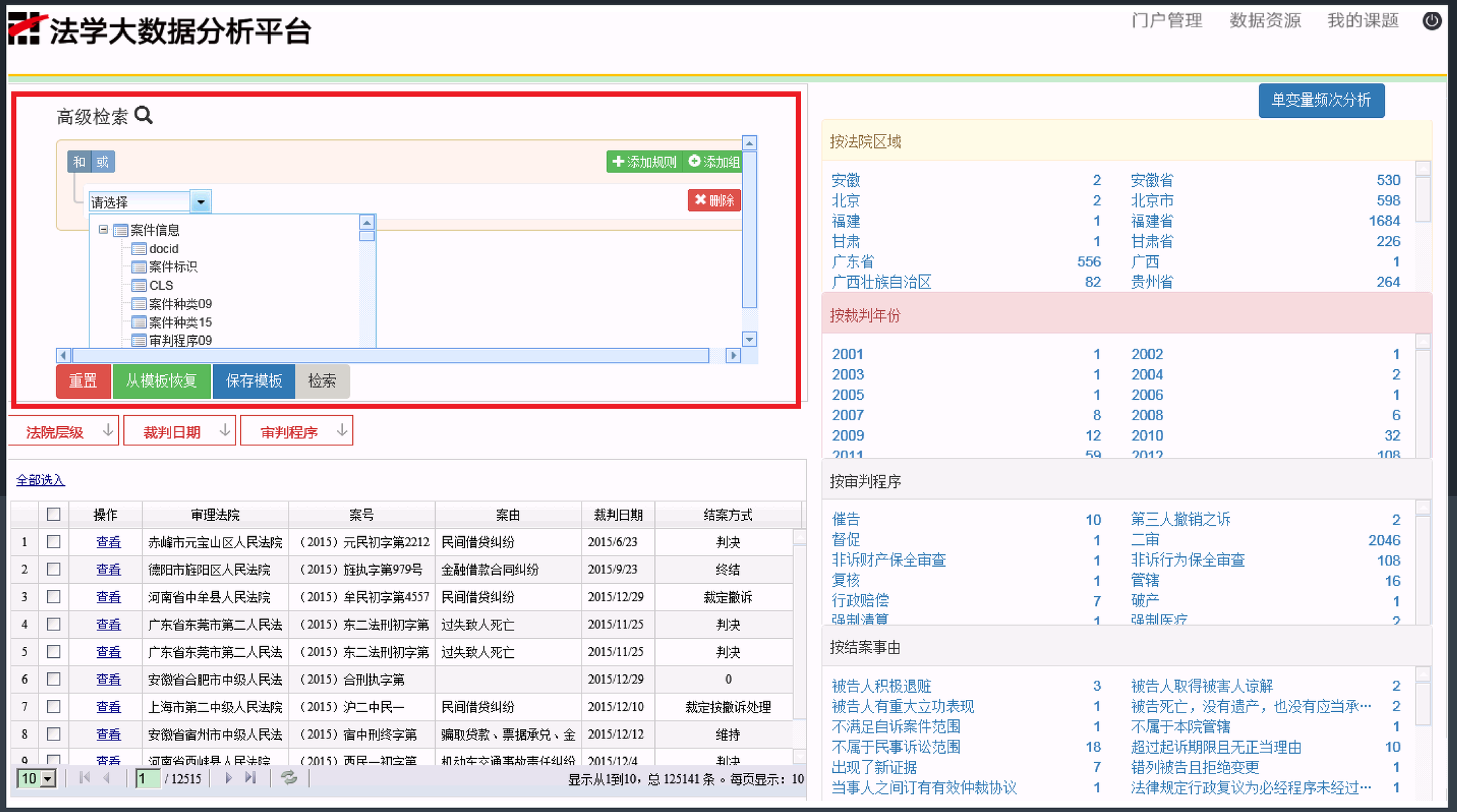
05. 自定义变量

功能
模块

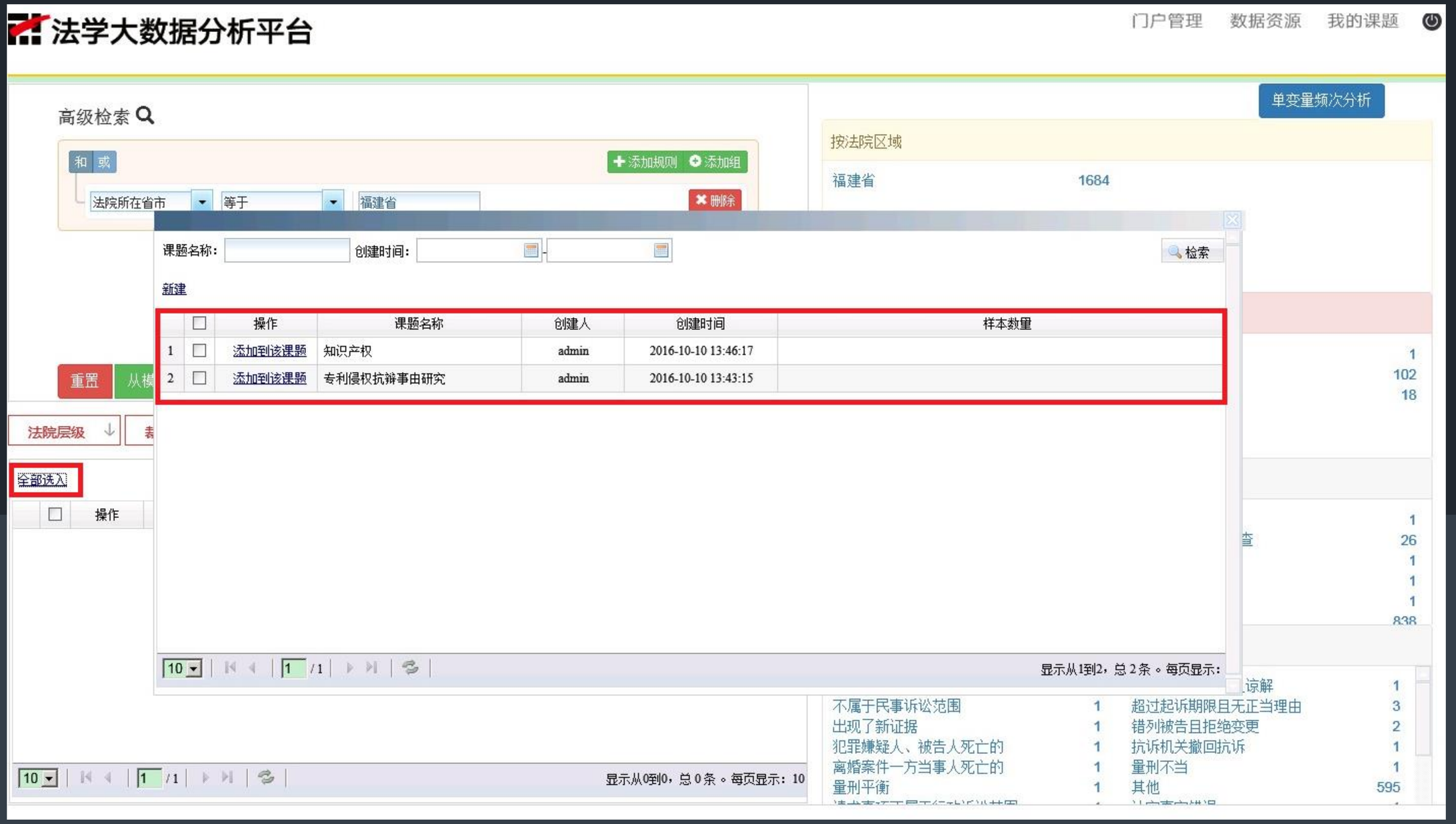
06.课题管理

07. 样本维护

01. 课题探索

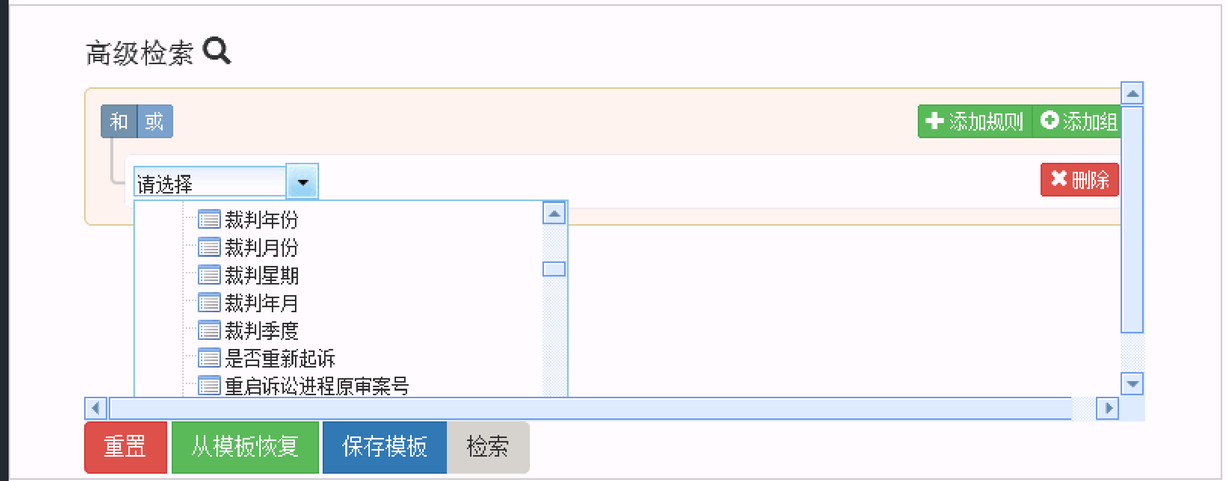

法学大数据分析平台
门户管理
数据资源
我的课题
高级检索
和
或
添加规则
添加组
删除
请选择
案件信息
docid
案件标识
CLS
案件种类09
案件种类15
审判程序09
重置
从模板恢复
保存模板
检索
单变量频次分析
法院层级
裁判日期
审判程序
全部选入
操作
审理法院
案号
案由
裁判日期
结案方式
1 查看 赤峰市元宝山区人民法院 （2015）元民初字第2212 民间借贷纠纷 2015/6/23 判决
2 查看 德阳市旌阳区人民法院 （2015）旌执字第979号 金融借款合同纠纷 2015/9/23 终结
3 查看 河南省中牟县人民法院 （2015）牟民初字第4557 民间借贷纠纷 2015/12/29 裁定撤诉
4 查看 广东省东莞市第二人民法 （2015）东二法刑初字第 过失致人死亡 2015/11/25 判决
5 查看 广东省东莞市第二人民法 （2015）东二法刑初字第 过失致人死亡 2015/11/25 判决
6 查看 安徽省合肥市中级人民法 （2015）合刑执字第 2015/12/29 0
7 查看 上海市第二中级人民法院 （2015）沪二中民一 民间借贷纠纷 2015/12/10 裁定按撤诉处理
8 查看 安徽省宿州市中级人民法 （2015）宿中刑终字第 骗取贷款、票据承兑、金 2015/12/12 维持
10 / 1 / 12515
显示从1到10，总 125141 条。每页显示：10
按法院区域
安徽 2 安徽省 530
北京 2 北京市 598
福建 1 福建省 1684
甘肃 1 甘肃省 226
广东省 556 广西 1
广西壮族自治区 82 贵州省 264
按裁判年份
2001 1 2002 1
2003 1 2004 2
2005 1 2006 1
2007 8 2008 6
2009 12 2010 32
按审判程序
催告 10 第三人撤销之诉 2
督促 1 二审 2046
非诉财产保全审查 1 非诉行为保全审查 108
复核 1 管辖 16
行政赔偿 7 破产 1
按结案事由
被告人积极退赃 3 被告人取得被害人谅解 2
被告人有重大立功表现 1 被告死亡，没有遗产，也没有应当承… 2
不满足自诉案件范围 1 不属于本院管辖 1
不属于民事诉讼范围 18 超过起诉期限且无正当理由 10
出现了新证据 7 错列被告且拒绝变更 1
当事人之间订有有效仲裁协议 1 法律规定行政复议为必经程序未经过… 1

02.选入课题


法学大数据分析平台
门户管理
数据资源
我的课题
高级检索
和
或
添加规则
添加组
法院所在省市
等于
福建省
删除
单变量频次分析
按法院区域
福建省
1684
课题名称：
创建时间：
检索
新建
操作
课题名称
创建人
创建时间
样本数量
1
添加到该课题
知识产权
admin
2016-10-10 13:46:17
2
添加到该课题
专利侵权抗辩事由研究
admin
2016-10-10 13:43:15
重置
法院层级
全部选入
显示从1到2，总 2 条 。每页显示：
显示从0到0，总 0 条 。每页显示：10
不属于民事诉讼范围 1
出现了新证据 1
犯罪嫌疑人、被告人死亡的 1
离婚案件一方当事人死亡的 1
量刑平衡 1
超过起诉期限且无正当理由 3
错列被告且拒绝变更 2
抗诉机关撤回抗诉 1
量刑不当 1
其他 595

03. 样本筛选

样本删除新增



法院层级 ↓ 裁判日期 ↓ 审判程序 ↓

删除 新增样本 全部删除

		操作	审理法院	案号	案由	裁判日期	结案方式
1	☐	查看	赤峰市元宝山区人民法院	（2015）元民初字	民间借贷纠纷	2015/6/23	判决
2	☐	查看	德阳市旌阳区人民法院	（2015）旌执字第	金融借款合同纠纷	2015/9/23	终结
3	☐	查看	河南省中牟县人民法院	（2015）牟民初字	民间借贷纠纷	2015/12/29	裁定撤诉
4	☐	查看	广东省东莞市第二人民法	（2015）东二法刑	过失致人死亡	2015/11/25	判决
5	☐	查看	广东省东莞市第二人民法	（2015）东二法刑	过失致人死亡	2015/11/25	判决
6	☐	查看	安徽省合肥市中级人民法	（2015）合刑执字		2015/12/29	0
7	☐	查看	上海市第二中级人民法院	（2015）沪二中民	民间借贷纠纷	2015/12/10	裁定按撤诉处理
8	☐	查看	安徽省宿州市中级人民法	（2015）宿中刑终	骗取贷款、票据承兑、金融票证	2015/12/12	维持
9	☐	查看	河南省西峡县人民法院	（2015）西民一初	机动车交通事故责任纠纷	2015/12/4	判决
10	☐	查看	北京市第一中级人民法院	（2015）一中民终	合同纠纷	2015/12/23	判决维持

10 | 1 / 12515 | 显示从1到10，总 125141 条。每页显示：10

按法院区域

安徽	2	安徽省	530
北京	2	北京市	598
福建	1	福建省	1684
甘肃	1	甘肃省	226
广东省	556	广西	1
广西壮族自治区	82	贵州省	264
海南省	34	河北	1

按裁判年份

2001	1	2002	1
2003	1	2004	2
2005	1	2006	1
2007	8	2008	6
2009	12	2010	32
2011	59	2012	108
2013	460	2014	1248

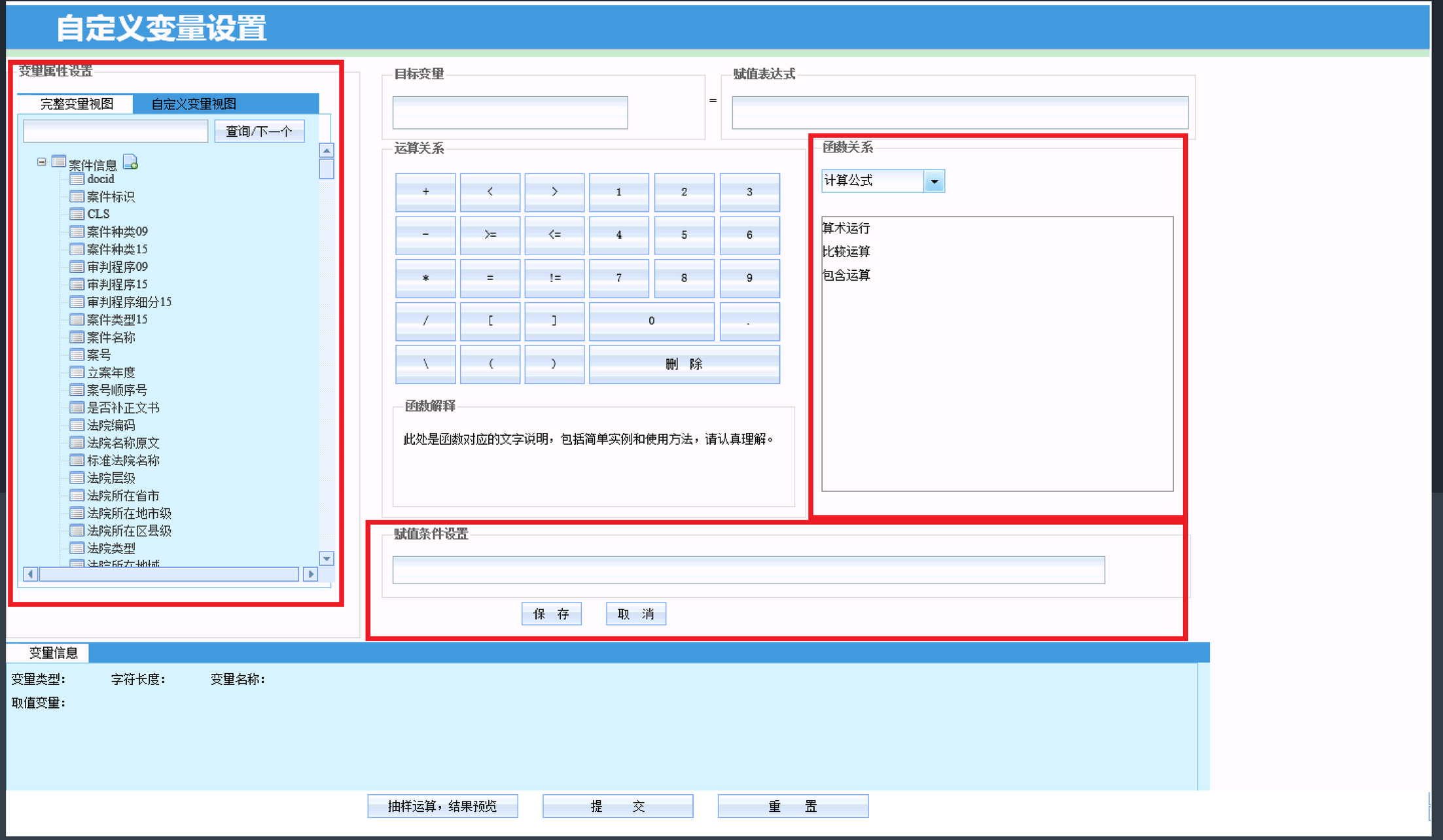
按审判程序

催告	10	第三人撤销之诉	2
督促	1	二审	2046
非诉财产保全审查	1	非诉行为保全审查	108
复核	1	管辖	16
行政赔偿	7	破产	1
强制清算	1	强制医疗	2
审判监督	223	司法赔偿	10

按结案事由

被告人积极退赃	3	被告人取得被害人谅解	2
被告人有重大立功表现	1	被告死亡，没有遗产，也没有应当承…	2
不满足自诉案件范围	1	不属于本院管辖	1
不属于民事诉讼范围	18	超过起诉期限且无正当理由	10
出现了新证据	7	错列被告且拒绝变更	1
当事人之间订有有效仲裁协议	1	法律规定行政复议为必经程序未经过…	1
犯罪嫌疑人、被告人死亡的	2	公诉机关撤诉	1

04.变量设计


自定义变量设置
变量属性设置
完整变量视图
自定义变量视图
查询/下一个
案件信息
docid
案件标识
CLS
案件种类09
案件种类15
审判程序09
审判程序15
审判程序细分15
案件类型15
案件名称
案号
立案年度
案号顺序号
是否补正文书
法院编码
法院名称原文
标准法院名称
法院层级
法院所在省市
法院所在地市级
法院所在区县级
法院类型
目标变量
=
赋值表达式
运算关系
+ < > 1 2 3
- >= <= 4 5 6
* = != 7 8 9
/ [ ] 0 .
\ ( ) 删 除
函数解释
此处是函数对应的文字说明，包括简单实例和使用方法，请认真理解。
函数关系
计算公式
算术运行
比较运算
包含运算
赋值条件设置
保 存
取 消
变量信息
变量类型:
字符长度:
变量名称:
取值变量:
抽样运算，结果预览
提 交
重 置

05统计分析

单变量频次分析

单变量频次分析是针对单个变量按照变量取值类型来统计每个取值类型出现的次数。比如针对犯罪人性别的频次分析，就是统计不同性别的犯罪人人数各为多少，即男性犯罪人人数为多少，女性犯罪人人数为多少

统计项设置： 结案主案由原文　　图形切换： 请选择　　是否显示统计值

下载　保存

	频次	百分比	有效百分比	累计百分比
滥伐林木	11	0.01%	0.01%	0.01%
销售假冒注册商标的商品	2	0%	0%	0.01%
电信服务合同纠纷	27	0.02%	0.02%	0.03%
侵害作品放映权纠纷	22	0.02%	0.02%	0.04%
对非国家工作人员行贿	2	0%	0%	0.05%
股东出资纠纷	7	0.01%	0.01%	0.05%
运输合同纠纷	26	0.02%	0.02%	0.07%
与企业有关的纠纷	5	0%	0%	0.07%
意外伤害保险合同纠纷	4	0%	0%	0.08%
网络侵权责任纠纷	1	0%	0%	0.08%
财产保险合同纠纷	32	0.02%	0.02%	0.1%
挂靠经营合同纠纷	11	0.01%	0.01%	0.11%
引诱、容留、介绍卖淫	4	0%	0%	0.11%
申请保全案件	11	0.01%	0.01%	0.12%
普通破产债权确认纠纷	1	0%	0%	0.12%
过失致人重伤	2	0%	0%	0.12%
侵占	3	0%	0%	0.12%
土地承包经营权纠纷	19	0.01%	0.01%	0.14%
票据纠纷	1	0%	0%	0.14%
共有权确认纠纷	2	0%	0%	0.14%

06研究报告

二、上网裁判文书的时间分布情况

截至 2016 年 6 月 30 日已上网的 18833561 篇裁判文书中，有 6267779 篇是 2014 年作出的裁判文书，占全部上网裁判文书的 33.28%，有 8340258 篇是 2015 年作出的裁判文书，占全部上网裁判文书的 44.28%，有 2233418 篇是 2016 年作出的裁判文书，占全部上网裁判文书的 11.86%，有 1992106 篇是其他年份作出的裁判文书，占全部上网裁判文书的 10.58%。（参见图表 3）

图表 3 已上网裁判文书作出年份分布



全国三十一省（区、市）及新疆生产建设兵团法院上传**2016 年作出**的裁判文书情况如下：





Thanks for Your Time