**指标编码细则**

|  |  |
| --- | --- |
| **研究框架** | **说明** |
| **文章类型** |  |
| 找不到论文信息 | 可能发表于医学或检察日报等期刊，无法查看相应文献 |
| 非研究 | 多指会议综述和报道等 |
| 非实证研究 | 虽然议题符合犯罪学的常见议题，但是摘要和题目的描述并未表明针对其使用了量化方法的文献 |
| 非犯罪学研究 | 虽然使用了量化方法，但议题不符合犯罪学 |
| 实证犯罪学 | 使用量化研究方法分析犯罪学 |
| **数据来源** |  |
| 个人调查问卷 | 通过问卷的方式收集到的数据，有别于大型调查数据库 |
| 业务数据 | 公检法等司法机关产生的业务数据；法律年鉴；企业、机构等产生的数据（如省级面板数据） |
| 大型调查数据库 | 国家或高校收集的调查数据库 |
| 案例数据库 | 来源于裁判文书网、无讼网、北大法宝等裁判文书数据库 |
| 其他 | 其他未具名的渠道 |
| **分析单位** |  |
| 个人 | 常见数据来源于问卷调查 |
| 案例 | 常见数据来源于案例数据库，同时也包含了通过调研等方式收集到的案例 |
| 集合数据 | 通常指由同属于某个类别的所有个体数据的某些统计量（常见均值或百分比）构成，例如某地区的犯罪率或百分比 |
| 其他 | 以从知网下载的文献进行数据分析或其他 |
| **研究设计** |  |
| 问卷调查 | 通过制定问卷，要求被调查者据此进行回答以收集资料的方法。 |
| 实验 | 使用实验设计进行研究的主要方法 |
| 二手数据分析 | 非介入性调查，本文主要指的是利用业务数据或者统计年鉴等进行二手资料分析 |
| **时间维度** |  |
| 截面数据 | 所使用的数据跨度没有超过一年 |
| 历时数据 | 所使用的数据跨度超过一年 |
| **抽样方法** |  |
| 随机抽样 | 文章指出用了随机的抽样方法 |
| 非随机抽样 | 抽样方法未知；如果对从裁判文书网获取的数据总体进行了筛选或者剔除，定义为选择样本且是非随机 |
| 总体 | 文章所使用的数据是从司法裁判文书等获取的全部数据 |
| **研究对象范围** |  |
| 全国 | 研究范围为全国 |
| 省、直辖市和自治区 | 研究范围为省、直辖市和自治区。此外，有些文献使用了多个省的研究数据，但达不到全国，本文纳入了省 |
| 地级市 | 研究范围为地级市 |
| 区县 | 研究范围为区县 |
| 其他 | 研究对象为国外或未知 |
| **获取的样本量** | 指最初获取的数据样本量 |
| **分析的样本量** | 剔除或筛选一些数据后，最后研究进行分析的样本量 |
| **分析工具** | 0=未提及/不清楚，1=Excel，2=SPSS,3=SAS,4=STATA,5=R，6=Python，7=GIS,8=Eviews，其他=MINITAB数学软件、SmartPLS2软件、GeoDa软件、Matlab R2012b软件、citespace、Depthmap等软件分析工具 |
| **量化方法使用** |  |
| **描述性统计：** |  |
| **频次/频率** | 1=描述了研究变量的频次，0=否 |
| **均值** | 1=描述了变量的集中趋势（均值），0=否 |
| **标准差、方差** | 1=描述了变量的离散趋势，0=否 |
| **列联表** | 1=文章中使用了列联表进行分析，0=否 |
| **均值比较** | 1=使用了均值比较，0=否 |
| **推论性统计：** |  |
| **列联相关系数** | 1=使用了卡方检验或斯皮尔曼（Spearman）等，0=否 |
| **T、ANOVA** | 1=使用了T检验/F检验，0=否 |
| **皮尔逊相关系数** | 1=使用了皮尔逊相关（Pearson），0=否 |
| **回归** | 1=使用了回归分析，0=否 |
| **进阶分析** | 1=使用了更为进阶的方法，0=否 |
| **论文撰写** |  |
| **数据获取过程** | 1=指出了具体的数据来源，0=否 |
| **研究假设** | 1=提出了研究假设；0=否 |
| **复杂概念的测量** | 1=对自变量、因变量等做出了具体的解释或测量，0=否 |
| **讨论研究局限** | 1=讨论了文章的研究局限，0=否 |