

COVID Foundation data issues

以下所有结果都是基于香港科技大学数据 2022-12-12.xlsx 这个文件得到的。

子表格: 香港大学数据查询 (就是基本组织信息)

机构类型、公募资格上没有数据问题。

Column: 机构类型

Missing value count: 0

Number of organizations with '0' in this column: 0

Unique value count: 5

Unique values: ['基金会' '红十字会' '慈善会' '社会团体' '社会服务机构']

Column: 公募资格

Missing value count: 0

Number of organizations with '0' in this column: 0

Unique value count: 2

Unique values: ['是' '否']

员工人数上, 我们有两个疑问。第一, 为什么会有机构的员工人数是 0? 第二, 对于缺失值, 员工人数是“未知”还是“0”?

Column: 员工人数

Missing value count: 51

Number of organizations with '0' in this column: 12

Unique value count: 57

Unique values: [97. 58. 5. 3. 7. 2. 4. 11. 10. 8. 6. !没有填写! 0. 26.
41. 9. 23. 16. 18. 25. 1. 20. 15. 100. 50. 13. 21. 12.
39. 14. 19. 42. 22. 27. 17. 325. 30. 111. 31. 35. 45. 28.
61. 87. 36. 181. 49. 40. 84. 120. 83. 60. 78. 70. 93. 208.
43. 29.]

企业/家族基金会所属企业或个人上，对于缺失值，隶属的企业是“未知”还是“没有隶属”？

Column: 企业/家族基金会所属企业或个人

Missing value count: 550

Number of organizations with '0' in this column: 0

Unique value count: 18

Unique values: [! 没有填写! '百度（中国）有限公司' '德天下国际健康管理（北京）有限公司' '正荣
'三七互娱网络科技集团股份有限公司' '山东京博控股集团有限公司' '上海国乾投资管理有限公司' '上
'上海相宜本草化妆品股份有限公司' '上海云开教育培训有限公司' 'TCL科技集团股份有限公司' '深圳
'花样年集团（中国）有限公司' '陈灵梅' '深圳市松禾创业投资有限公司' '杭州锦江集团有限公司']

评估等级上，“! 没有填写!” “无” “未知” 有什么区别？更具体的说，“! 没有填写!” 应该被看作是” 未知” 还是看作” 无”？

Column: 评估等级

Missing value count: 226

Number of organizations with '0' in this column: 0

Unique value count: 8

Unique values: ['5A' '4A' ! 没有填写! '3A' '2A' '未参加' '1A' '无' '未知']

子表格: 财务信息

问题是一致的: 第一，为什么有基金会的钱是'0.0' (除了政府补助收入可以理解以外)? 第二，“! 没有填写!” 代表的是” 未知” 吗? 还是” 0.0”? 第三，政府补助收入下， '! 没有填写!' 应该被视作没有政府收入，还是看作” 未知”？

Column: 总资产合计

Missing value count: 1

Number of organizations with '0' in this column: 0

Unique value count: 485

Unique values: ['! 没有填写!']

Column: 净资产合计

Missing value count: 10

Number of organizations with '0' in this column: 3

Unique value count: 474

Unique values: ['0.0', '! 没有填写!']

Column: 收入合计

Missing value count: 2

Number of organizations with '0' in this column: 1

Unique value count: 484
Unique values: ['!没有填写!', '0.0']

Column: 捐赠收入
Missing value count: 2
Number of organizations with '0' in this column: 12
Unique value count: 473
Unique values: ['!没有填写!', '0.0']

Column: 政府补助收入
Missing value count: 40
Number of organizations with '0' in this column: 312
Unique value count: 120
Unique values: ['0.0', '!没有填写!'] <-- 这里很重要