

可视化与可视计算概论第一次作业报告

1700012764 陈智斌

2019 年 10 月 28 日

目录

1	作业要求	2
2	数据描述与分析	2
2.1	按申请与资助情况统计	2
2.2	按单位隶属关系统计	3
2.3	按单位性质统计	3
3	设计宗旨和设计过程	4
3.1	表单 1	4
3.2	表单 2、3	4
3.3	可视化页面	4
4	结果描述	4
4.1	表单 1	4
4.2	表单 2	6
4.3	表单 3	6
5	总结与感受	6

1 作业要求

本次作业要求从 NSFC 资助项目统计资料中选取不少于三张有联系的表单，分别使用 D3.js 对各表单进行可视化，组成一个完整的可视化页面。在该份作业中，选择了以下三张表单：

- 1.2018 年面上项目按申请与资助情况统计；
- 2.2018 年面上项目资助情况（按单位隶属关系统计）；
- 3.2018 年面上项目资助情况（按单位性质统计）；

2 数据描述与分析

2.1 按申请与资助情况统计

2018年面上项目按申请与资助情况统计										
科学部	受理申请		批准资助					资助率		
	项数	金额	项数	金额	资助金额比例		单项平均 资助金额	项数	金额	
					占全委	占学部				
合计	92622	6320107.96	18947	1115289.00	100%	100%	58.86	20.46%	17.65%	
数理科学部	6544	442589.39	1743	104640.00	9.38%	100.00%	60.03	26.64%	23.64%	
数学	1835	100169.22	489	25428.00	2.28%	24.30%	52.00	26.65%	25.39%	
力学	1444	106578.25	383	24247.00	2.17%	23.17%	63.31	26.52%	22.75%	
天文学	384	28124.58	103	6504.00	0.58%	6.22%	63.15	26.82%	23.13%	
物理学I	1645	116055.08	438	27638.00	2.48%	26.41%	63.10	26.63%	23.81%	
物理学II	1236	91662.27	330	20823.00	1.87%	19.90%	63.10	26.70%	22.72%	
化学科学部	7811	553452.43	1737	113180.00	10.15%	100.00%	65.16	22.24%	20.45%	
合成化学	1287	92979.22	301	19584.00	1.76%	17.30%	65.06	23.39%	21.06%	
催化与表面化学	767	54139.50	176	11470.00	1.03%	10.13%	65.17	22.95%	21.19%	
化学理论与机制	522	36587.15	116	7561.00	0.68%	6.68%	65.18	22.22%	20.67%	
化学测量学	717	51306.31	159	10363.00	0.93%	9.16%	65.18	22.18%	20.20%	
材料化学与能源化学	1369	97150.88	293	19099.00	1.71%	16.87%	65.18	21.40%	19.66%	
环境化学	972	68960.08	214	13948.00	1.25%	12.32%	65.18	22.02%	20.23%	
化学生物学	608	43377.84	134	8734.00	0.78%	7.72%	65.18	22.04%	20.13%	
化学工程与工业化学	1569	108951.44	344	22421.00	2.01%	19.81%	65.18	21.92%	20.58%	
生命科学部	12664	890474.10	3048	177470.00	15.91%	100.00%	58.23	24.07%	19.93%	
微生物学、植物学	1575	111489.43	415	24240.00	2.17%	13.66%	58.41	26.35%	21.74%	
生态学、林学	1758	120487.69	399	23283.00	2.09%	13.12%	58.35	22.70%	19.32%	
生物物理、生化与生物分子学、生物力学与组织工程、免疫学	1293	93774.17	349	20315.00	1.82%	11.45%	58.21	26.99%	21.66%	

图 1: 表单 1 总览

这份数据按照科学部与进一步细分的一级学科为分类，记录了受理项目数、受理金额、批准资助项目数、批准金额及其占比、单项平均资助金额、资助率，其中从受理项目数、受理金额、批准资助项目数、批准金额出发，经过计算可以得到其他所有数据特征。如果以各一级学科为基础分类，该表单共计有学科、受理申请项数、受理金额、批准资助项数、批准金额、资助金额占全委比例、资助金额占学部比例、单项平均资助金额、资助项数率、资助金额率共计十个维度的数据，其中后五个维度数据可以由前五个维度推导而来，共计有 8 个科学部、55 个一级学科。考虑期单项平均资助金额、资助率大多相近，可视化应主要考虑体现前五个维度的数据，同时适当体现资助金额比例之间的关系，单项平均资助金额、资助率可以通过一些显性的数值计算、几何比例体现。

2018年面上项目按单位隶属关系统计										
科学部	合计		教育部		中国科学院		工、交、农、医、国防等部门		各省、自治区、市(直)	
	项数	金额	项数	金额	项数	金额	项数	金额	项数	金额
	项数	金额	项数	金额	项数	金额	项数	金额	项数	金额
合计	18947	1115289.00	8736	514641.70	2100	128261.50	2286	134998.50	5825	337387.30
%	100.00%	100.00%	46.11%	46.14%	11.08%	11.50%	12.07%	12.10%	30.74%	30.25%
数理科学部	1743	104640.00	740	43980.00	311	19660.00	254	15626.00	438	25474.00
%	100%	100%	42%	42%	18%	19%	15%	15%	25%	24%
化学科学部	1737	113180.00	871	56784.20	308	20114.00	112	7271.00	446	29010.80
%	100%	100%	50%	50%	18%	18%	6%	6%	26%	26%
生命科学部	3048	177470.00	1236	72574.00	444	26177.00	299	17393.00	1069	61326.00
%	100%	100%	41%	41%	15%	15%	10%	10%	35%	35%
地球科学部	1895	117210.00	735	45561.00	517	31983.00	276	17207.00	367	22459.00
%	100%	100%	39%	39%	27%	27%	15%	15%	19%	19%
工程与材料科学部	3199	192088.00	1668	100238.00	195	11732.00	418	25105.00	918	55013.00
%	100%	100%	52%	52%	6%	6%	13%	13%	29%	29%
信息科学部	2007	120037.00	910	55086.00	201	11873.00	297	17782.00	599	35296.00
%	100%	100%	43%	46%	10%	10%	15%	15%	30%	29%
管理科学部	803	38544.00	534	25667.30	28	1355.00	67	3211.00	174	8310.50
%	100%	100%	67%	67%	3%	4%	8%	8%	22%	22%
医学科学部	4515	252120.00	2042	114751.00	96	5467.50	563	31403.50	1814	100498.00
%	100%	100%	45%	46%	2%	2%	12%	12%	40%	40%

图 2: 表单 2 总览

2.2 按单位隶属关系统计

这份数据按照科学部为分类，记录了批准的总项数、总金额，教育部、中国科学院、公交农医国防部门、省市自治区四种单位隶属关系各自的批准项数、批准金额及其所占比例。该表单共计有科学部、总项数、总金额、四种隶属关系每种的项数及比例、金额及比例，共计 19 个维度的数据，其中各项比例可以推导得来。该表单中各科学部、各部门比例差距较大，可以比较好地展现出各科学部发展与不同隶属单位之间的关系，因此各维度数据都较为重要，并集中分析各比例之间的关系。

2.3 按单位性质统计

2018年面上项目资助情况（按单位性质统计）								
科学部	合计		高等院校		科研单位		其他	
	项数	金额	项数	金额	项数	金额	项数	金额
	项数	金额	项数	金额	项数	金额	项数	金额
合计	18947	1115289.00	15800	926009.00	2930	177112.50	217	12167.50
%	100.00%	100.00%	83.39%	83.03%	15.46%	15.88%	1.15%	1.09%
数理科学部	1743	104640.00	1398	82939.00	343	21574.00	2	127.00
%	100.00%	100.00%	80.21%	79.26%	19.68%	20.62%	0.11%	0.12%
化学科学部	1737	113180.00	1415	92183.50	318	20738.50	4	258.00
%	100.00%	100.00%	81.46%	81.45%	18.31%	18.32%	0.23%	0.23%
生命科学部	3048	177470.00	2293	133221.00	750	43953.00	5	296.00
%	100.00%	100.00%	75.23%	75.07%	24.61%	24.77%	0.16%	0.17%
地球科学部	1895	117210.00	1159	71601.50	706	43774.00	30	1834.50
%	100.00%	100.00%	61.16%	61.09%	37.26%	37.35%	1.58%	1.57%
工程与材料科学部	3199	192088.00	2928	175787.00	260	15644.00	11	657.00
%	100.00%	100.00%	91.53%	91.51%	8.13%	8.14%	0.34%	0.34%
信息科学部	2007	120037.00	1790	107157.00	212	12561.00	5	319.00
%	100.00%	100.00%	89.19%	89.27%	10.56%	10.46%	0.25%	0.27%
管理科学部	803	38544.00	778	37342.00	24	1153.00	1	49.00
%	100.00%	100.00%	96.89%	96.88%	2.99%	2.99%	0.12%	0.13%
医学科学部	4515	252120.00	4039	225778.00	317	17715.00	159	8627.00
%	100.00%	100.00%	89.46%	89.55%	7.02%	7.03%	3.52%	3.42%

图 3: 表单 3 总览

这份数据按照科学部为分类，记录了批准的总项数、总金额，高等院校、科研单位、其他单位三种单位性质的各自的批准项数、批准金额及其所占比例。该表单共计有科学

部、总项数、总金额、三种单位性质每种项数及比例、金额及比例，共计 15 个维度的数据，其中各项比例可以推导得来。与第二张表单类似，由于各种比例差距较大，应当集中分析各比例之间的关系。

3 设计宗旨和设计过程

3.1 表单 1

对这个表单可视化的关注点应该着重体现资助金额、资助项数之间的数量关系，以直观地表达出各科学部、各一级学科之间发展的差异、资助投入力度的大小。考虑使用柱状图来进行数据的可视化，以金额、项数为纵轴，以科学部、一级学科为横轴，为了直观、方便地比较科学部、一级学科之间的数量关系，分别以两种分类标准进行两张图的可视化绘制。

柱状图方法的好处是可以大体直观地体现数量关系，但难点在于无法体现维数过高的数据，进而无法将所有精确比例直接展现出来，因此考虑辅以标签、文字来进行可视化。

3.2 表单 2、3

这两张表单情况比较类似，都应该着重体现各个比例之间的数量关系，以直观地表现出不同科学部在不同隶属关系、不同单位性质下的资助差异、发展程度。由于需要体现的维数较高。考虑使用平行坐标轴图来进行数据的可视化，每个维度各自使用一个坐标轴。

平行坐标轴图可以方便地容纳高维数的数据、体现数据之间的大小关系，但难点在于各数据之间的绝对倍率关系体现不直观，且实现难度较大。

3.3 可视化页面

从上面的设计过程可以知道，整个可视化页面将由四张可视化图像组成，故考虑在页面内部使用 2*2 排布方式，以方便在各个可视化图像之间进行比较、关联，并且易于与可视化图像进行交互，但不足之处在于页面整体本身没有装饰、结构较为简单。

4 结果描述

4.1 表单 1

以科学部为分类标准，总共有 8 类。每一类都有受理申请项数、受理申请金额、批准项数、批准金额，分别用两种灰色及橘色、蓝色表示，在横轴上标出每个类别的名字，在柱状图上标出各类的精确数值。从图上可以直观地看出，医学科学部不论是申请项数还是申请金额都超过了其他所有的科学部，说明该资助项目在医学方向投入的资助量是最大的，而管理科学部的申请项数和申请金额都是最少的。此外，从该图中可以看除，项

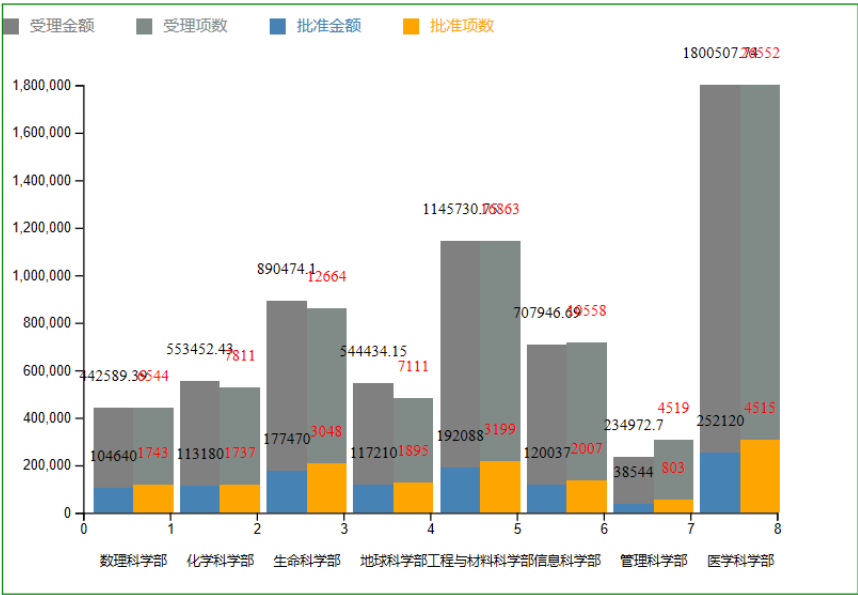


图 4: 表单 1 以科学部为分类标准的可视化

数和资助金额基本上是成正相关的，批准金额和受理申请金额也是正相关的，这说明资助项目的供需关系基本是比较稳定的。

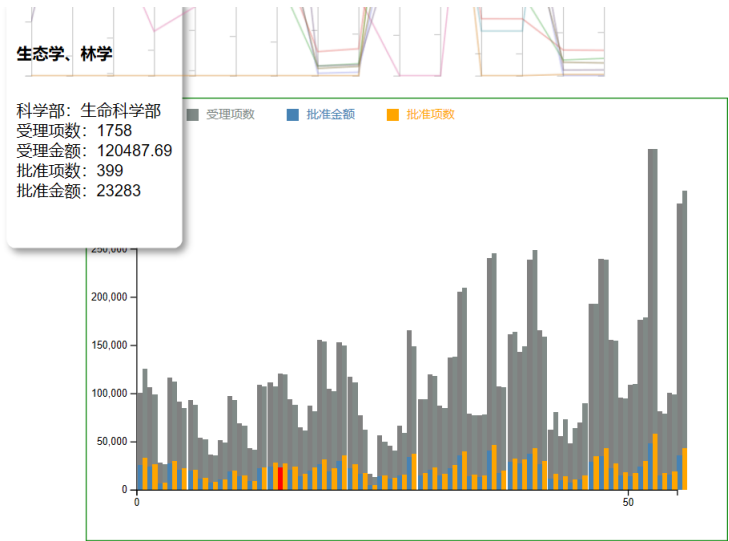


图 5: 表单 1 以一级学科为分类标准的可视化

以一级学科为分类标准，总共有 55 类，分属于不同的科学部的一级学科之间用白色细线作为间隔划分开来。由于类别数目太多，无法在横轴上标出每个一级学科的名称，也无法在柱状图上直接标出精确数值，因此引入交互标签。将鼠标移到对应的矩形上，会在标签中展示出该类别的具体信息，包括一级学科名、所属科学部、项数、金额。从该图中可以看出，在同一个学部内，不同学科的申请、批准情况的差异也是十分巨大的。

4.2 表单 2

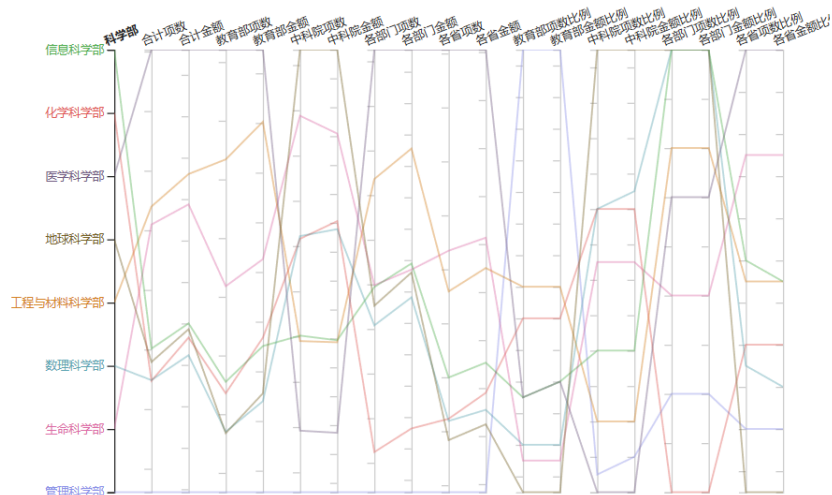


图 6: 表单 2 的可视化

将鼠标移到对应坐标轴上，可以看到该坐标轴对应的刻度，并且支持在坐标轴上拖动选区，以提取部分科学部的数据。从该图中，我们可以看出管理科学部的资助主要来自教育部，数理科学部、信息科学部、地球科学部的资助中，来自工、交、农、医、国防等部门的资助比例要比其他的学部高，医学科学部的资助来自各省市自治区的比例比较高。事实上，数理、信息、地球科学等领域在各生产部门之中的应用要比其他学部更为直接、广泛，而医学发展则受到各地政府较大力度的扶持。

4.3 表单 3

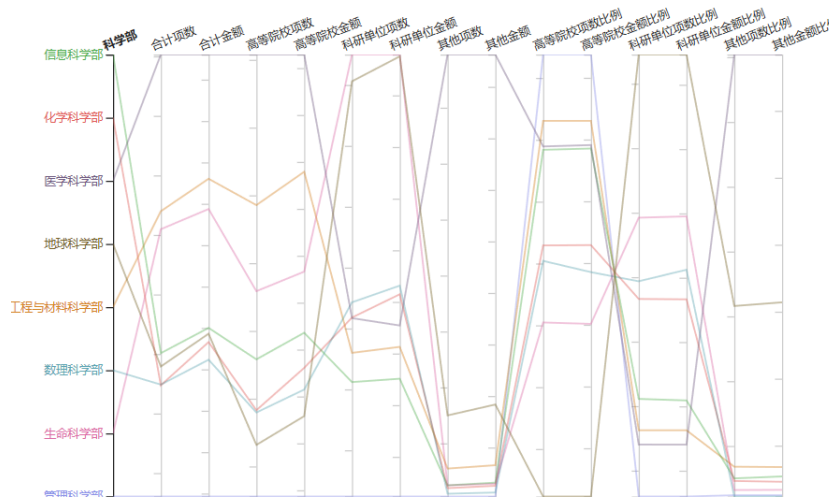


图 7: 表单 3 的可视化

与表单 2 类似，将鼠标移到对应坐标轴上，可以看到该坐标轴对应的刻度，并且支

持在坐标轴上拖动选区，以提取部分科学部的数据。从该图中，我们可以看出管理、信息、工程、医学科学部的资助中，高等院校项目所占的比例尤其高，地球科学、生命科学部的资助中科研单位项目所占比例高于其他科学部，其他性质单位的项目资助除了医学科学部外，其他科学部的资助几乎可以忽略不计。高等院校仍然是各学部资助项目申请、研究的主阵地，除了医学可以由医院的研究部门推进科研外，其他学部的科研活动主要还是集中在高等院校和科研单位，这也是符合常识的。

5 总结与感受

目前为止，如果我们粗略地使用资助项目的申请、批准情况来推测各学部的发展状况的话，各个学部之间的发展还是相当不平衡的，当然这与学科本身科研经费的消耗是有关的，并受市场调节。此外，各种项目主要是由隶属于教育部、中科院的各个高等院校、科研单位推进的，这充分说明了学术研究项目与工业界之间是有很大区别的。

从可视化实现的过程上看，选用合适的方法对数据进行可视化实现是一件需要技巧、难度较大的工作，需要有较强的理论、实践基础才能比较好地完成。