

恩师 Lizhongzhang up 主数据分析视频发布习惯及其赞赏值差异研究

陈瑞祺

关键词：数据分析 赞赏值 心得

（一）引言

统计学的入门以及相关统计软件的掌握，李宗璋老师给予了我很大的帮助。老师不仅仅从课堂上传授知识，而且在课后也会为同学耐心地答疑解惑，更值得敬佩的是，老师有坚持在 B 站发布教学视频，包含统计学知识以及相关统计软件的使用，像 R studio 的学习，我也因老师的视频受益匪浅。本文旨在研究 Lizhongzhang up 主数据分析视频发布习惯及其赞赏值差异，借此机会，对老师的恩情表示衷心的感谢。

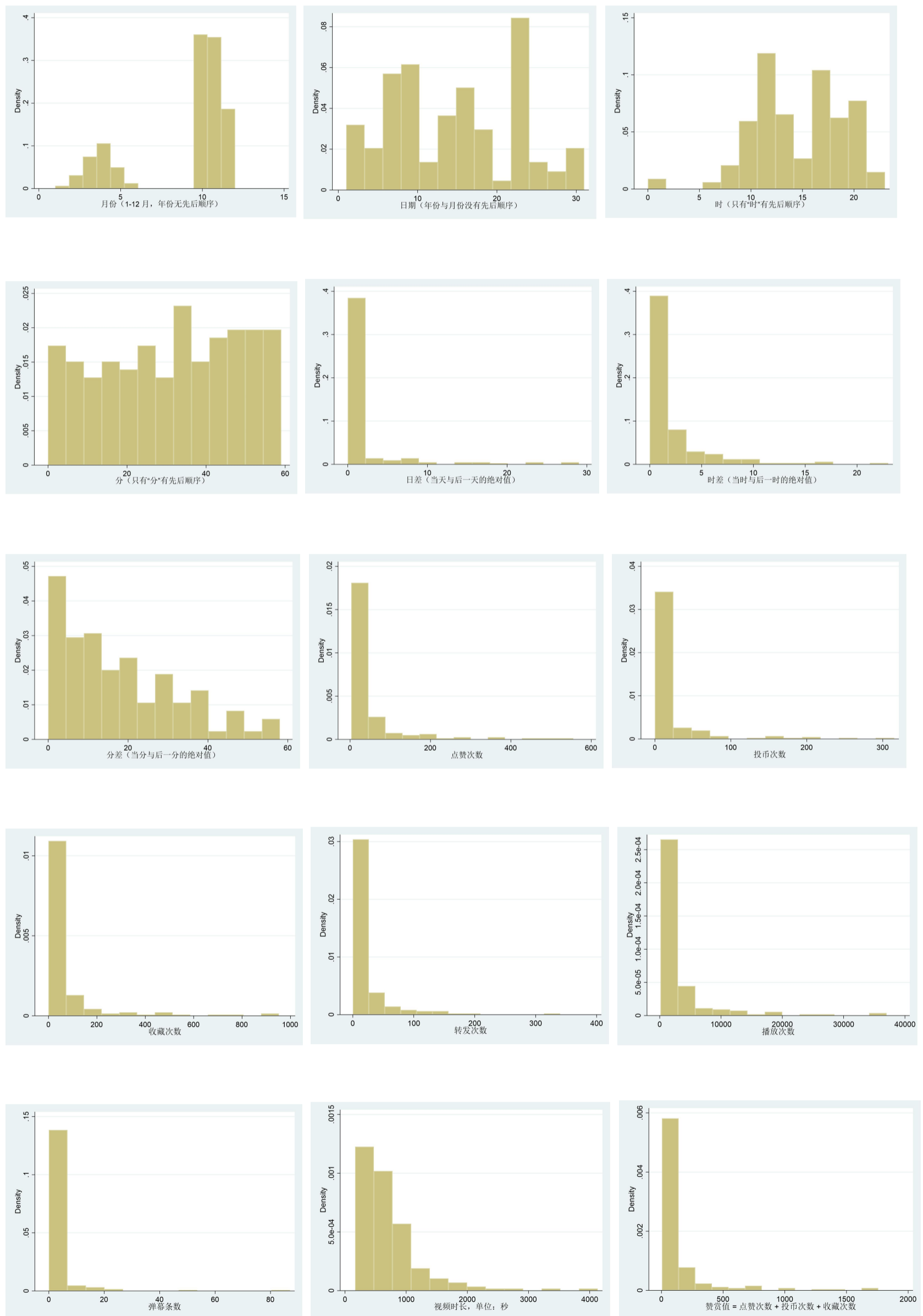
（二）数据

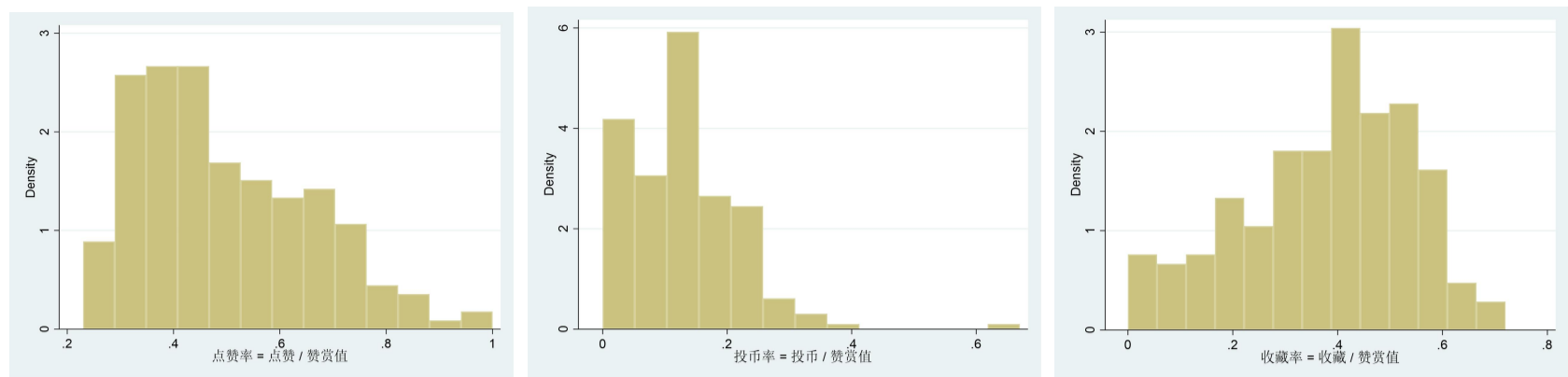
数据于 2022 年 6 月 3 日从 B 站收集而来，包含了 190 个 Lizhongzhang up 主至今有关数据分析的视频，时间从 2020 年 10 月 3 日到 2022 年 5 月 27 日，一些合集视频做了剔除处理。老师对 R 软件的精彩讲解，令我开始关注老师过往的教学视频，发现既有大量统计学以及计量经济学知识的讲解，又有其他统计软件的详细教程。为了对这些视频的特征进行尽可能全面的分析，由此设置了以下变量。

variable name	storage type	display format	value label	variable label
年份	int	%8.0g	2020 & 2021 & 2022	
月份	byte	%8.0g	1-12 月, 年份无先后顺序	
日期	byte	%8.0g	年份与月份没有先后顺序	
时	byte	%8.0g	只有“时”有先后顺序	
分	byte	%8.0g	只有“分”有先后顺序	
日差	byte	%8.0g	当天与后一天的绝对值	
时差	byte	%8.0g	当时与后一时的绝对值	
分差	byte	%8.0g	当分与后一分的绝对值	
点赞	int	%8.0g	点赞次数	
投币	int	%8.0g	投币次数	
收藏	int	%8.0g	收藏次数	
转发量	int	%8.0g	转发次数	
播放量	long	%8.0g	播放次数	
弹幕数	byte	%8.0g	弹幕条数	
时长	int	%8.0g	视频时长, 单位: 秒	
软件	str6	%9s	Eviews & Excel & SPSS & R & 其他	
类型	str9	%9s	数据 & 绘图 & 函数 & 检验 & 计量 & 小知识	
赞赏值	int	%8.0g	赞赏值 = 点赞次数 + 投币次数 + 收藏次数	
点赞率	float	%8.0g	点赞率 = 点赞 / 赞赏值	
投币率	float	%8.0g	投币率 = 投币 / 赞赏值	
收藏率	float	%8.0g	收藏率 = 收藏 / 赞赏值	

在软件的分类中，“其他”表示习题讲评、使用录屏软件 Sreen Brush 进行知识点的讲解以及数据抓取网页八爪鱼客户端的介绍等，而在类型的划分中，“数据”包含了描述性统计、各种分布类型以

及方差分析等，“小知识”包含了置信区间的构造、参数估计和 p 值的理解等。对其中的定量变量分别绘制出频率分布直方图来观察其分布特征，发现大部分分布都属于右偏分布。老师发布的视频在 3-5 月，尤其在 10-12 月最集中，而且日差和时差绝大部分都很小，说明习惯于连续更新。





Decriptive statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max	(6) Var	(7) skewness	(8) kurtosis	(9) sum	(10) p25	(11) p50	(12) p75
月份	190	9.126	3.141	1	12	9.868	-1.180	2.796	1,734	10	10	11
日期	190	14.50	7.997	1	31	63.96	0.210	1.954	2,755	7	14	22
时	190	14.47	4.376	0	23	19.15	-0.340	3.092	2,750	11	15	18
分	190	31.22	17.29	0	59	299.0	-0.144	1.845	5,932	16	33	47
赞赏值	190	138.9	268.8	4	1,757	72,228	3.831	19.14	26,390	23	49	106
点赞	190	49.05	81.16	3	555	6,587	3.793	19.14	9,319	13	22.50	43
投币	190	21.07	44.62	0	315	1,991	3.961	20.45	4,004	2	6	18
收藏	190	68.77	147.8	0	952	21,855	3.927	19.94	13,067	6	20	52
转发量	190	20.29	40.33	0	340	1,627	4.201	26.51	3,856	1	6	19
播放量	190	3,004	5,411	129	37,000	2.928e+07	3.794	19.98	570,820	395	1,105	2,961
弹幕数	190	2.089	8.112	0	87	65.81	7.649	72.20	397	0	0	1
时长	190	718.2	535.4	166	4,130	286,665	2.926	15.26	136,452	394	575	861
点赞率	190	0.493	0.164	0.230	1	0.0271	0.689	2.773	93.61	0.350	0.455	0.610
投币率	190	0.126	0.0908	0	0.670	0.00824	1.275	8.778	23.97	0.0700	0.120	0.180
收藏率	190	0.381	0.161	0	0.720	0.0260	-0.449	2.557	72.36	0.270	0.400	0.500

由描述性统计的结果表明，对于所统计的 190 个数据分析视频，老师已累计赞赏值 26390，中位数为 49，说明每条视频也具有一定的认可度。然而赞赏值并不是认可度的唯一体现，以目前 1.2 万的粉丝量作为参考，可以推测大部分粉丝应该尚未形成点赞投币的好习惯。在数据统计的过程中，出于对老师的敬意，大部分视频都有点赞，而投币也做到了力所能及。

老师坚持更新视频展现了浓厚的学习热情，也诠释了“学习使我快乐”，对知识与技能的探索难免会遭遇种种困难，在软件教学视频当中也发现老师有时会运行失败，但老师总能冷静分析并一步步地解决问题，当中也增强了我对软件的熟知度，减轻了我对系统报错的恐惧感。伴随着熟练度的提高，我也能够自行摸索并绘制出自己想要的图表，大多数就是成功的，也难免有失败的时候，但我依然能感受到这种学习的快乐，由此我非常感谢老师的引导。

(三) 实证分析

一键三连即点赞、投币和收藏能够支持 up 主们积极投稿，相信这也是被众多 up 主关注的指标。本次研究通过将点赞、投币和收藏的次数累加生成一个新变量——赞赏值，以此体现出一部分的认可度，再以日期、点赞次数、投币次数、收藏次数、点赞率、投币率、收藏率、转发量、播放量、弹幕数以及视频时长作为可选择的解释变量，对被解释变量赞赏值进行多元回归分析。

VARIABLES	赞赏值						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
日期	-2.480*** (-3.172)	-2.478*** (-3.177)	-2.479*** (-3.179)	-2.479*** (-3.179)	0.003 (0.012)	-0.080 (-0.978)	-0.021 (-0.236)
点赞					2.139*** (12.714)	1.745*** (43.880)	
投币					1.102*** (5.376)		1.884*** (39.925)
收藏						1.013*** (37.849)	1.177*** (45.110)
点赞率	-442.618 (-0.354)	60.949 (1.359)	-207.240*** (-2.928)				
投币率	-236.090 (-0.189)	267.662*** (3.697)		206.434*** (2.912)			
收藏率	-505.236 (-0.403)		-269.247*** (-3.715)	-61.726 (-1.373)			
转发量	1.666*** (4.286)	1.687*** (4.390)	1.675*** (4.361)	1.685*** (4.387)	0.448*** (3.940)	-0.104** (-2.462)	0.061 (1.320)
播放量	0.031*** (10.333)	0.031*** (10.376)	0.031*** (10.396)	0.031*** (10.378)	0.009*** (7.158)	-0.004*** (-8.904)	0.003*** (5.162)
弹幕数	5.235*** (5.337)	5.248*** (5.366)	5.239*** (5.357)	5.245*** (5.362)	-2.112*** (-6.537)	0.344*** (2.704)	0.378*** (2.727)
时长	0.028** (2.390)	0.028** (2.399)	0.028** (2.397)	0.028** (2.399)	0.007** (2.083)	0.001 (0.753)	-0.001 (-0.944)
Constant	457.635 (0.366)	-46.544 (-1.440)	221.953*** (3.589)	14.716 (0.558)	-25.186*** (-5.138)	-3.362* (-1.789)	9.581*** (4.941)
Observations	190	190	190	190	190	190	190
R-squared	0.908	0.908	0.908	0.908	0.993	0.999	0.999

t-statistics in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

结果表明，月初发布的视频和月末相比会得到更高的赞赏值，在模型 5-7 中，对于赞赏值的积极影响，贡献最大的是点赞次数，其次是投币次数，最后是收藏次数。在模型 2-4 中，只有投币率对赞赏值有极其显著的积极影响，而点赞率和收藏率都对赞赏值表现出了极其显著的消极影响。为了进一步探讨赞赏值与这些解释变量的相关性，再加入分类变量进行回归分析，构建模型为：

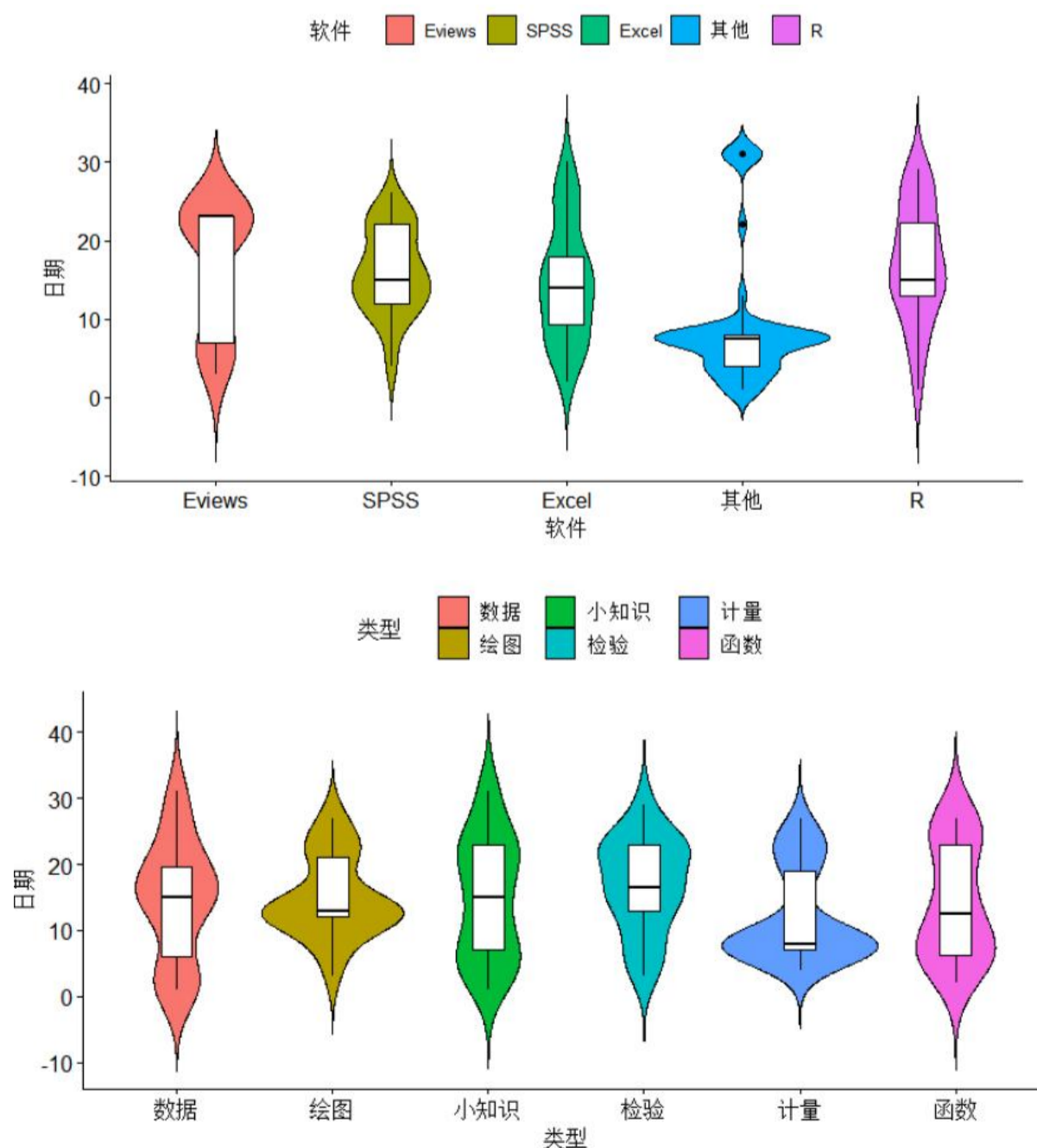
$$\text{赞赏值} = \alpha + \beta_1 \text{日期} + \beta_2 \text{转发量} + \beta_3 \text{播放量} + \beta_4 \text{弹幕数} + \beta_5 \text{时长} + \varepsilon$$

VARIABLES	(1) Eviews	(2) Excel	(3) SPSS	(4) R	(5) 其他
日期	0.070 (0.201)	-1.504 (-0.963)	0.449 (0.308)	-0.541* (-1.889)	-7.296*** (-3.275)
转发量	2.139*** (4.041)	6.875*** (6.438)	3.278*** (3.722)	5.719*** (4.088)	5.604*** (3.677)
播放量	0.013*** (4.677)	-0.016 (-1.566)	0.015*** (3.217)	-0.004 (-0.412)	0.037*** (6.200)
弹幕数	14.396** (2.238)	-2.923 (-0.553)	7.868** (2.184)	8.914** (2.449)	-5.109* (-1.846)
时长	0.002 (0.069)	-0.004 (-0.106)	0.058* (1.761)	-0.008 (-1.457)	0.021 (1.024)
Constant	8.705 (0.915)	23.603 (0.594)	-46.015 (-1.519)	35.572*** (5.080)	55.883 (1.534)
Observations	33	36	45	32	44
R-squared	0.946	0.969	0.945	0.684	0.928

VARIABLES	(1) 数据	(2) 绘图	(3) 函数	(4) 检验	(5) 计量	(6) 小知识
日期	0.665 (0.444)	0.039 (0.034)	-0.811** (-2.770)	-1.507 (-1.122)	-0.325 (-0.685)	-3.993* (-2.023)
转发量	0.853 (1.557)	2.254*** (3.146)	-1.356* (-2.160)	3.894*** (5.577)	6.045*** (12.780)	2.827** (2.406)
播放量	0.024*** (6.536)	0.013** (2.296)	0.035*** (6.922)	0.005 (0.783)	0.004 (1.074)	0.042*** (7.488)
弹幕数	20.124*** (8.372)	2.777 (0.724)	2.580 (0.647)	31.577*** (5.035)	-0.331 (-0.051)	-2.140 (-0.970)
时长	0.053 (1.277)	0.016 (0.736)	0.032** (2.971)	0.045* (2.001)	0.007 (1.506)	0.047 (1.533)
Constant	-32.902 (-0.969)	2.194 (0.115)	5.381 (0.766)	4.564 (0.138)	2.180 (0.232)	21.150 (0.493)
Observations	30	26	18	30	44	42
R-squared	0.983	0.936	0.983	0.940	0.991	0.930

t-statistics in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

结果表明，转发量越大，往往赞赏值也会越大，但对于函数视频会呈现一种负相关关系。对于 Eviews 和检验类型的视频，弹幕数对赞赏值的积极影响最显著。对于 SPSS 软件操作的讲解，增加时长有可能会最大赞赏值。还值得关注的是，对于月初发布的视频和月末相比会得到更大的赞赏值这一结论，应该是由于月初和月末视频的讲解内容不同而导致的，这也与老师视频发布习惯相联系。



结果表明，习题讲评、函数以及计量类的视频习惯在月初被发布，SPSS、Excel、R 和绘图方面的视频则在月中，Eviews 和检验等则在月末。

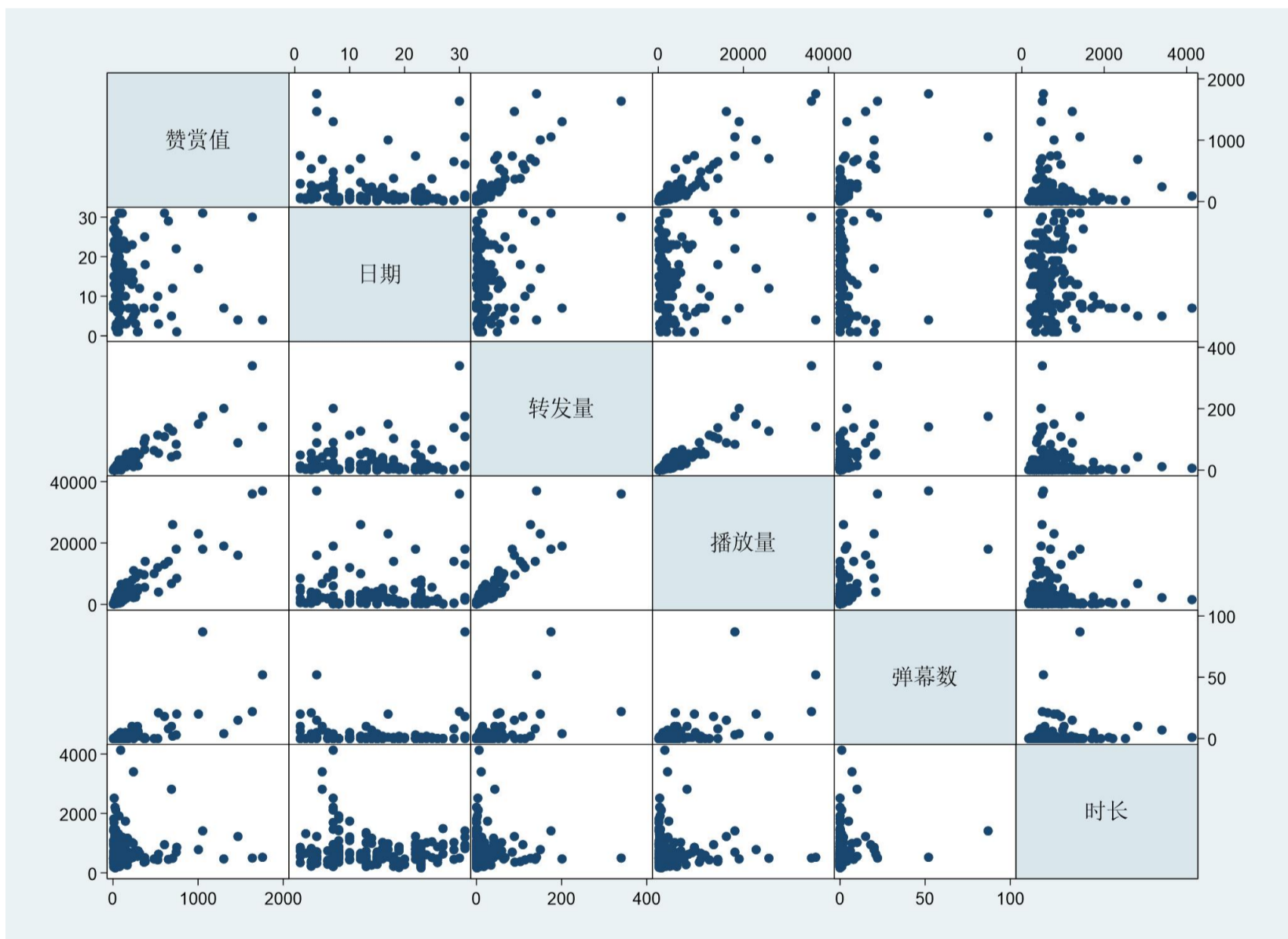
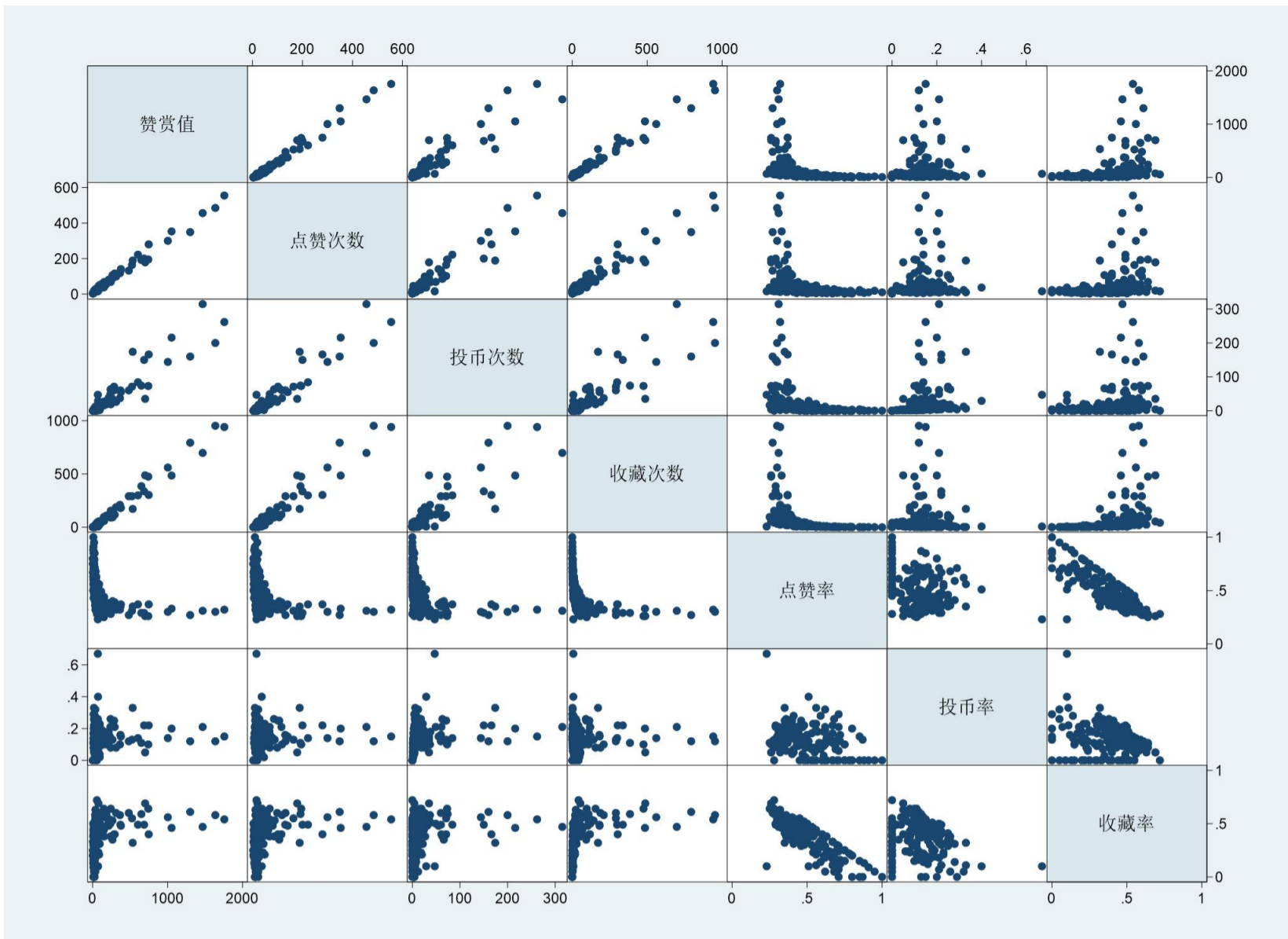
软件 * 类型 Crosstabulation

		类型							
		函数	绘图	计量	检验	数据	小知识	Total	
软件	其他	Count	0	2	22	3	5	12	44
		% within 软件	0.0%	4.5%	50.0%	6.8%	11.4%	27.3%	100.0%
Eviews		Count	2	5	6	5	2	13	33
		% within 软件	6.1%	15.2%	18.2%	15.2%	6.1%	39.4%	100.0%
Excel		Count	10	7	1	7	8	3	36
		% within 软件	27.8%	19.4%	2.8%	19.4%	22.2%	8.3%	100.0%
R		Count	6	2	1	6	8	9	32
		% within 软件	18.8%	6.3%	3.1%	18.8%	25.0%	28.1%	100.0%
SPSS		Count	0	10	14	9	7	5	45
		% within 软件	0.0%	22.2%	31.1%	20.0%	15.6%	11.1%	100.0%
Total		Count	18	26	44	30	30	42	190
		% within 软件	9.5%	13.7%	23.2%	15.8%	15.8%	22.1%	100.0%

软件 * 类型 Crosstabulation

		类型							
		函数	绘图	计量	检验	数据	小知识	Total	
软件	其他	Count	0	2	22	3	5	12	44
		% within 类型	0.0%	7.7%	50.0%	10.0%	16.7%	28.6%	23.2%
Eviews		Count	2	5	6	5	2	13	33
		% within 类型	11.1%	19.2%	13.6%	16.7%	6.7%	31.0%	17.4%
Excel		Count	10	7	1	7	8	3	36
		% within 类型	55.6%	26.9%	2.3%	23.3%	26.7%	7.1%	18.9%
R		Count	6	2	1	6	8	9	32
		% within 类型	33.3%	7.7%	2.3%	20.0%	26.7%	21.4%	16.8%
SPSS		Count	0	10	14	9	7	5	45
		% within 类型	0.0%	38.5%	31.8%	30.0%	23.3%	11.9%	23.7%
Total		Count	18	26	44	30	30	42	190
		% within 类型	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

此外，老师在习题讲解方面，主要是以计量为主，而计量又主要是通过 SPSS 来演示的。函数知识主要通过 Excel 来讲解，小知识在 Eviews 中演示得最多，SPSS 还包含了较多的绘图教学和检验知识。最后，为了更好地呈现定量变量之间的相关性，采用矩阵图做可视化处理。



图像呈现的是定量变量两两对应的相关关系，对于被解释变量赞赏值，转发量、播放量和弹幕数的作用与多元回归一样都是积极的，然而视频时长对赞赏值的影响偏向于消极的。此外，点赞率和收藏率呈现出强烈的负相关关系，转发量、播放量和弹幕数两两分别呈现正相关关系。

(四) 总结和展望

本文主要发现 Lihongzhang up 主的视频点赞值与其视频发布习惯有一定的联系，不同软件的教学以及不同知识点的讲解会导致赞赏值的较大差异，就研究结果来看，粉丝对经济计量学的讲解的外显认可度最高，但考虑到老师在 B 站讲计量的时间相对较早，而这一结论就不太严谨。

于 2022 年 2 月 19 日老师发布了第一个有关 R studio 的视频，至数据收集当天已累计有 32 个，不到 4 个月，已累计了 866 的赞赏值，从最初讲解 Eviews 的视频，到后来的 Excel 和 SPSS，赞赏值有较大的提升，就自身而言，视频的教学确实起到了非常不错的效果，因此，对于 R 的实验课，我更倾向于看视频教学来完成作业，相信这些的视频在未来将获得更高的关注度与认可度。

Decriptive statistics & 赞赏值

VARIABLES	(1) N	(2) mean	(3) sd	(4) p25	(5) p50	(6) p75	(7) min	(8) max	(9) Var	(10) skewness	(11) kurtosis	(12) sum
R	32	27.06	14.99	16.50	22	38	8	73	224.6	1.190	4.108	866
SPSS	45	143.1	206.8	45	72	124	17	1,004	42,749	2.752	10.12	6,439
Excel	36	213.9	342.0	50	88	200	23	1637	116,992	3.047	11.95	7,699
Eviews	33	56.55	60.01	20	31	70	9	243	3,601	1.917	5.871	1,866

再次感谢李宗璋老师的教导，老师辛苦了！