

动漫评分影响因素分析

摘要：本文主要以日本 TV 动漫为研究对象，通过统计分析探究导演水平、编剧水平、是否为原创作品、改编来源和动画题材等相关因素对动画评分的影响作用，通过线性回归模型来探究具有显著影响的变量。结论表明导演、编剧对动画评分有显著正向作用，原创与否影响不显著，热血、励志等题材有显著正向作用，后宫、异世界等题材具有显著负向作用。

一、背景介绍

伴随着日本动漫作品和生活方式的输入，中国的二次元人群正在逐步扩容，相关文化市场的热情也在逐步攀升。根据艾瑞咨询和东北证券测算数据，2020 年我国泛二次元用户规模已达 4.1 亿人，为以动漫为中心的文化及周边产业提供了庞大的消费潜力，Mob 研究院预测 2025 年时中国动画相关行业的总市值有望达到 2500 亿。整体上来看，我国经济和市场已经具备支持国内动漫产业高速发展的基础，国内动漫产业市场发展潜力巨大。然而，美好愿景之下是残酷的竞争，日本动漫依旧占据着国内相关市场的霸主地位，国内动漫也面临着高烈度的市场竞争，动漫产业已然成为高度内卷的红海。创作者如何才能把握观众喜好，打造爆点产品？我们将探索动画作品评分的影响因素，助力动画产业振兴。

二、数据来源与说明

动漫评分数据来源于动漫评分网站 bangumi，相较于另一更为大众的动漫网站 bilibili 来说，该网站使用者多具有较为丰富的阅番量，更有可能形成一套完整而系统的评价准则，其评分更具客观性和

代表性。通过网络爬虫，我们获取了共计 3230 条记录，其中 TV 动画共 2058 条。所获取动漫的播出时间涵盖 1963-2021 年，虽然大部分数据集中在 2000-2021 年，但由于在此之前的多数动漫直到今天仍然经久不衰广受好评，非常具有参考价值，故并未作特殊处理。

对于影响因素的选取，参考组内动漫专家的建议，我们主要关注三个方面：动画制作团队、动画属于原创作品还是改编作品和动画属于什么题材，这也是大多数观众在浏览新番时会关注的方面。对于动画制作团队，最受关注的是导演和编剧的水平，导演作为团队的总指挥决定着团队实力的发挥，而动画若想要脱离泯然众人的境地，离不开优秀的剧情，故而负责剧情的编剧决定着动画的上限，对此我们选取了经标准化处理的导演评分和编剧评分作为自变量。动画是否原创则是决定了制作团队的发挥空间，改编作品有着原作水平托底，即使改变水平泛善可陈，也容易在原作的基础上给人较好的观感，而原创番没有原作的束缚，其上限更高，相对应地下限也更低，容易出神作的同时也更容易暴雷，对此我们设定是否原创为分类变量，并收集了对应原作的评分。更进一步地，我们将改编动画按其原作类型分为小说改、漫画改、游戏改、GAL 改四个类别，由此可以在改编动画的分组下进行更精细的回归，研究原作类型对评分的影响。最后是动画题材方面，我们选取了三十个动画题材变量，包括奇幻、百合、热血、治愈、异世界等主流要素，每个动画都可以同时符合多个题材。表 1 提供了一个变量概览。

表 1 变量概览

变量类型	变量名	详细说明	取值范围
因变量	动画评分	来自 bangumi 网站评分	2.34-9.14
	导演标准分	标准化处理的导演评分	-4.42-3.52
	编剧标准分	标准化处理的编剧评分	-4.46-3.23
自变量	是否原作	原创为 1，改编为 0	0 或 1
	原作评分	标准化处理的原作评分	-4.63-3.29
	原作类型	4 分类变量	1-4 分别代表 小说改、漫画 改、游戏改、 GAL 改
	动画题材	30 个二分类变量	0 或 1

三、描述性统计

经计算，动画评分的均值为 6.62，标准差为 0.86，导演标准分、编剧标准分和原作标准分众数为 0，相当多的样本在这三个方面上处于平均水平。



图 1 动画题材 MEKKO 图

从 MEKKO 图（图 1）可以看出，搞笑、奇幻、战斗、校园、日常等题材最为大众，而音乐、魔法、运动等属于小众题材。

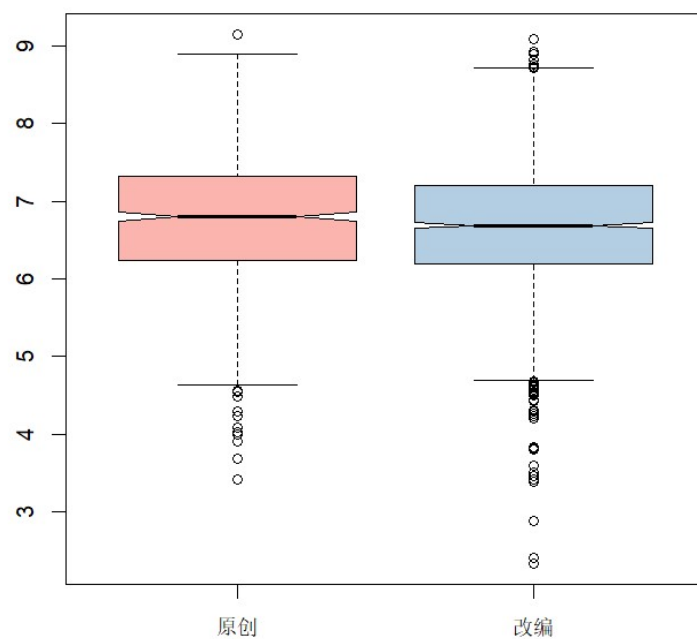


图 2

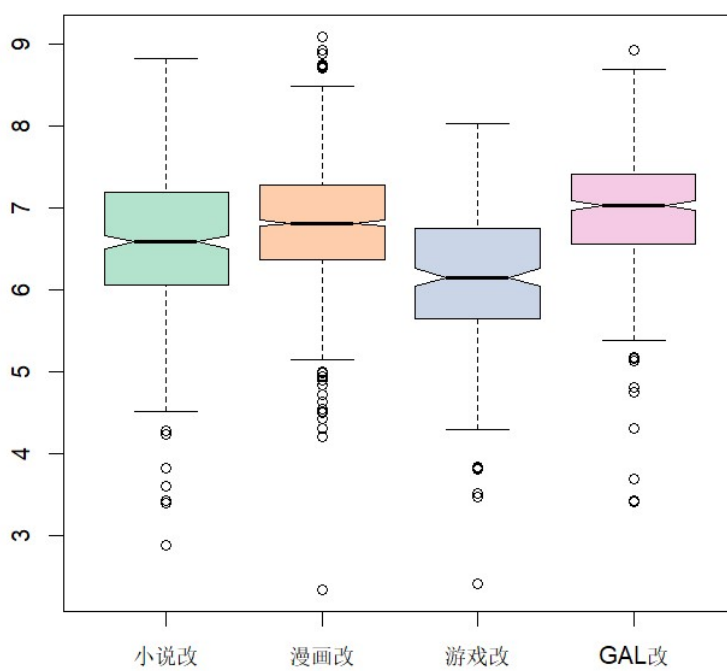


图 3

如图 2 和图 3 所示，原创作品和改编作品的分数平均来说没有显著差异，但值得注意的是改编作品有更多的离群点。从改编作品的四个改编来源来看，游戏改的作品评分平均来说明显要低于其他三个

来源，其他三个来源的平均水平没有明显差异。

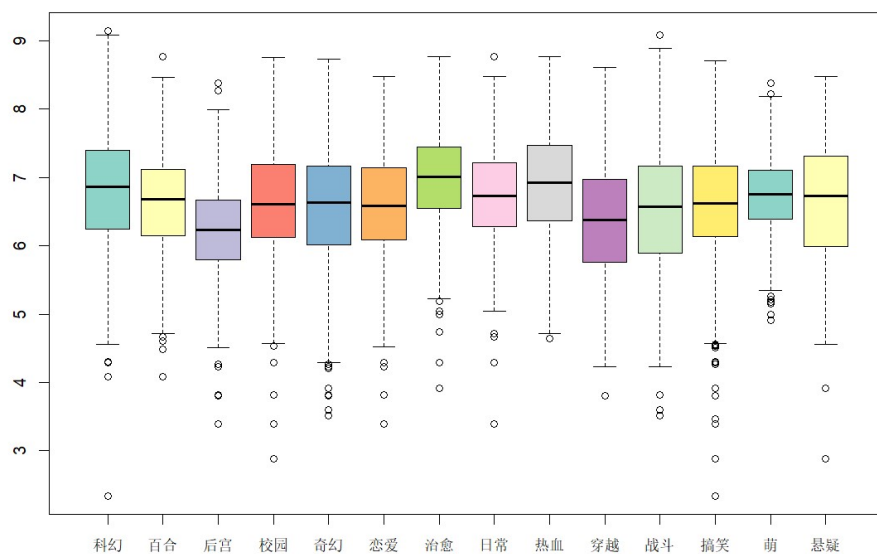


图 4

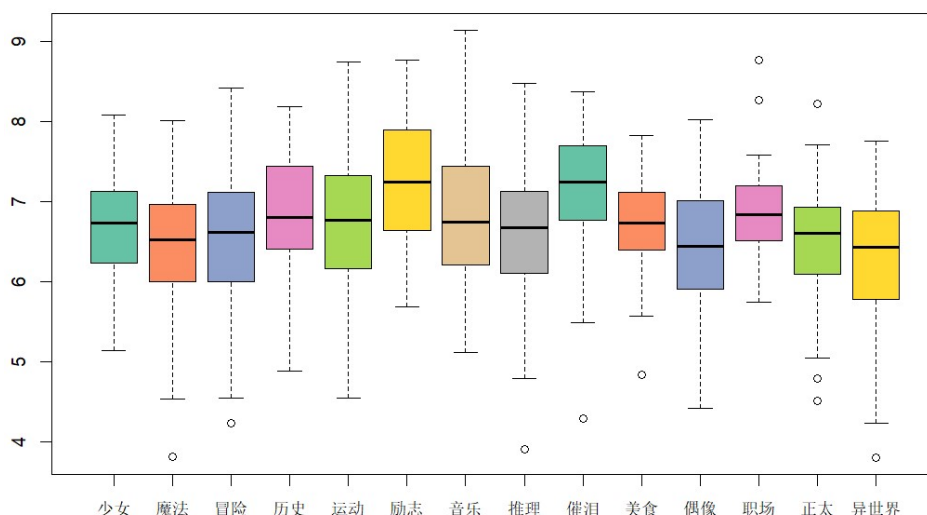


图 5

图 4 和图 5 展示了三十个动画题材对应的番剧评分状况，注意到含有治愈、励志、热血、催泪四个标签的动画评分平均来说要高于其他动画，而含有穿越、偶像、和异世界三个标签的动画评分平均来说要低于其他动画。

此外，我们还对坊间传言“自古四月/十月出神番”做了验证，

分别以排名前 25%和排名前 10%作为高分动画标准，结果如图 6 和图 7 示，四月和十月的高分动画相比其他两个季度来说确实更多，但我们进一步对每个季度的动画总数作出频数图（图 8），发现四月和十月的动画总数也较多，在基量大的情况下出现更多高分动画是很自然的情况，对四个季度动画评分所作的卡方检验结果 p 值为 0.31，也即可以认为四个季度并无显著差别，这说明不是四月和十月有什么特殊之处，二仅仅是因为其总量更多。

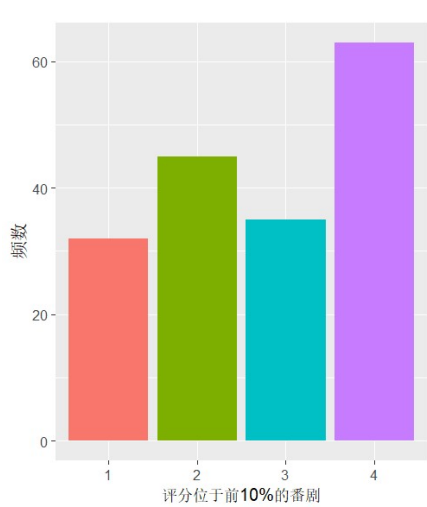


图 6

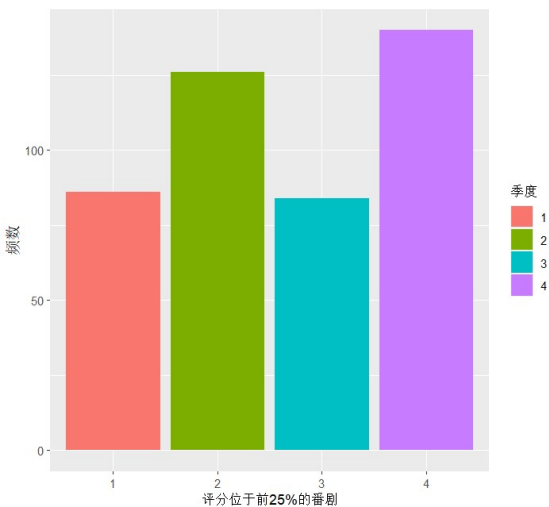


图 7

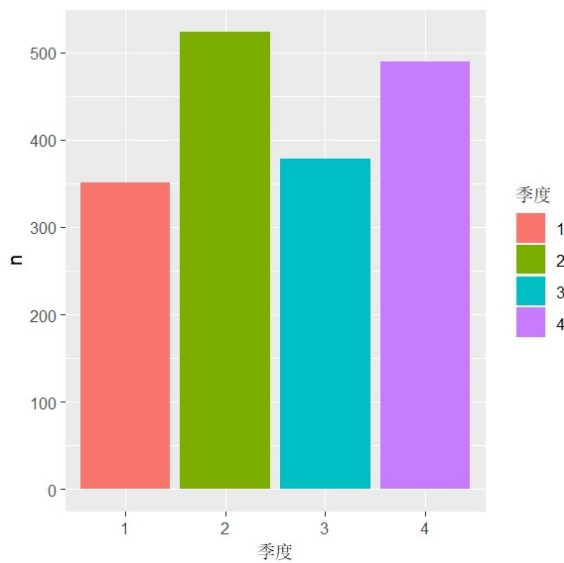


图 8

四、回归分析

首先对数据建立简单的线性回归模型，回归结果如表 2 所示

表 2 基本回归结果

变量/模型	所有 TV 动画	非原创 TV 动画
截距项	6.537***	6.548***
导演标准分	0.278***	0.364***
编剧标准分	0.300***	0.220***
是否原创	-0.033	
原作评分		0.235***
修正 R^2	0.163	0.327
F 值	135***	250.5***

注：***, **, *分别表示在 0.001, 0.01, 0.05 显著性水平下显著, 下文同此

导演标准分、编剧标准分和原作评分的系数均和理论结果一致，均对动画评分具有显著的正向影响，从所有 TV 动画的回归结果来看，编剧的影响略大于导演，而是否原创虽然系数为负，但并不显著，这说明是否原创对动画的评分影响不大，在后续回归中删去此变量；从非原创 TV 动画的回归结果来看，导演的作用在较大程度上高于编剧，原作评分的作用也略微高于编剧，这也和理论结果一致。

注意到两个简单模型的修正 R^2 都比较低，我们需要加入更多变量来解释动画评分的变动，下面我们引入非原创动画的四个改编来源作进一步回归。回归结果如表 3 所示。

表 3 引入改编来源后回归结果

变量/模型	非原创 TV 动画 (未处理)	非原创 TV 动画 (平方处理)
-------	--------------------	---------------------

截距项	6.548***	43.0956***
导演标准分	0.364***	5.123***
编剧标准分	0.220***	2.610***
原作评分	0.235***	2.676***
小说改（基准）		
漫画改	0.199***	2.518***
游戏改	-0.499***	-6.282***
GAL 改	-0.325***	-4.486***
修正R ²	0.400	0.418
F 值	148***	185.4***

注意到此时的残差 QQ 图（图 9）有所弯曲，K-S 检验结果为在 0.1 的显著性水平上拒绝原假设，这意味着正态性假定没有很好地满足，鉴于后续需要根据 AIC 进行变量筛选，而 AIC 需要正态性假定。为此考虑作 Box-Cox 变换，计算结果如图 10 所示，可将因变量平方处理，平方后的 QQ 图如图 11 所示，其几乎呈现为一条直线，K-S 检验结果显示在 0.1 的显著性水平上接受原假设。平方处理后的回归结果如表 4 所示。

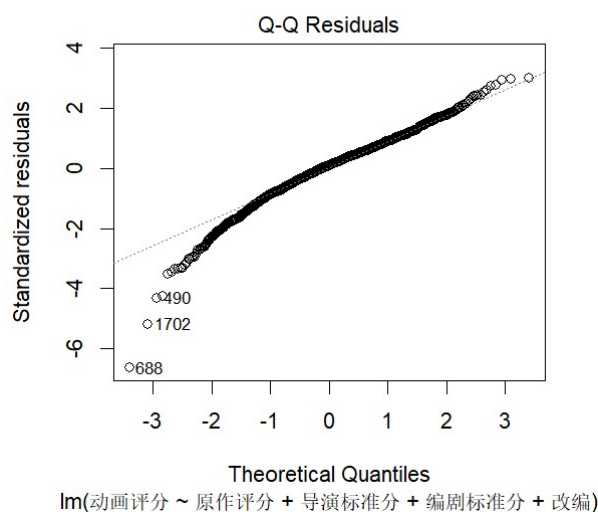


图 9 残差 QQ 图

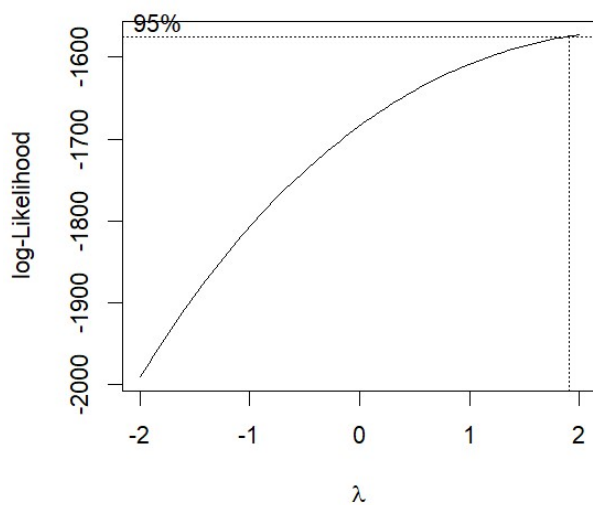


图 10 Box-Cox 变换参数空间 MLE 示意图

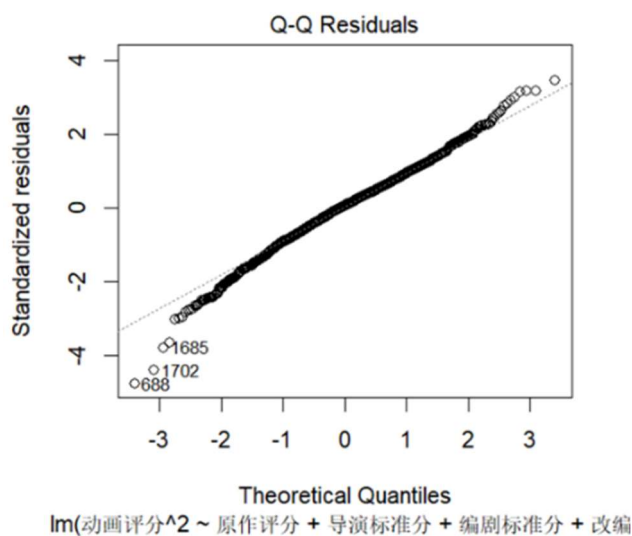


图 11 平方处理后的残差 QQ 图

回归结果显示所有系数均显著，其中漫画改编具有正向的贡献，而游戏改和 GAL 改均有着负向的贡献。游戏改具有较高负向作用是很自然的结果，因为大量这样的作品都是通过情怀赚取短期利益，并未经过仔细打磨。而 GAL 改的负向作用却与常规印象相反，GAL 改动画给人的印象是佳作频出，如缘之空、CLANNAD、悠久之翼等，但数据

否定了这一结论，图 12 展示了 GAL 改动画评分的分布，相当多的作品评分位于 TV 动画均值之下。我们推断这有两个可能因素，一是 GAL 改的原作党（玩过原作品进而对改编动画感兴趣的人）可能对改编质量有较高要求；二是 GAL 改作品有一定观看门槛，如缘之空为了符合原作的多分支剧情，各集之间有特定的观看顺序，不了解者可能会因此打低分。

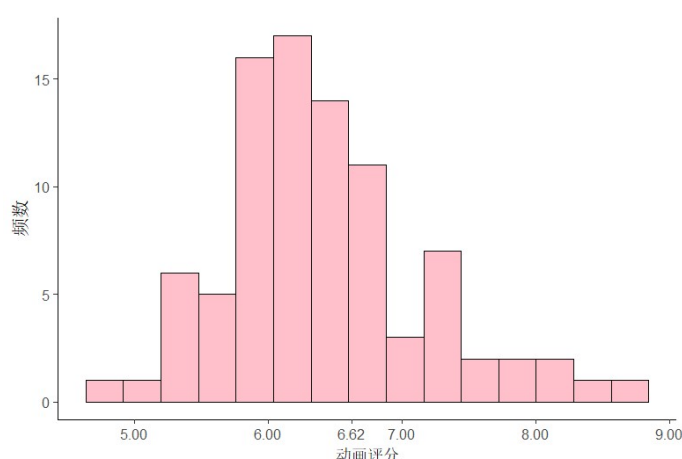


图 12

现基于上文回归结果，将 TV 动画分为原创和非原创两个类别考虑题材问题，将 30 个分类变量全部加入模型中，由于变量较多，下面仅展示根据 AIC 删减变量后的回归结果，列于表 5 中。

注意到热血、励志、催泪等风格具有较高的正向系数，这与主流审美比较契合，可以称之为经典叙事剧情，动漫圈内有三大民工漫之称，指经久不衰的死神、海贼王和火影忍者三部作品，其剧情正是属于这一风格。后宫标签具有较高的负向作用，推测这可能是因为大部分后宫作品剧情较为糟糕，在打分较为客观的 bangumi 成为了负面标签。在原创番中，音乐标签具有显著正向作用，《轻音少女》、《吹响吧！上低音号》、《四月是你的谎言》等以音乐为主要元素的番剧都是广受好评的佳作。治愈番无论是原创作品还是非原创作品都具有很高

的正向作用，这同样也属于主流审美的一个方面，人们总是倾向于从美好的事物中寻求心灵的慰藉。

表 4 引入动画题材变量后回归结果

变量/模型	原创 TV 动画 (平方处理)	非原创 TV 动画 (平方处理)	所有 TV 动画 (平方处理)
截距项		40.889***	41.051***
导演标准分	39.671***	2.480***	3.977***
编剧标准分	3.439***	2.975***	1.713***
原作评分		4.723***	4.720***
漫画改		1.146*	
游戏改		-6.353***	
GAL 改		-3.542***	
科幻	3.652***		1.713*
恋爱		1.098*	
治愈	6.477***	3.228***	4.720***
日常		2.275***	2.611***
百合	1.558		-1.206
热血	3.348*	4.847***	5.790***
战斗			-1.968***
搞笑		0.725	1.015
萌		1.229*	1.602*
悬疑			2.186*
历史		3.274*	3.590*
运动		-1.959*	
励志	3.945	4.264***	5.051***
音乐	5.204**		
推理	-4.691	1.938	

催泪		3.982*	5.895**
偶像	-3.344	4.084**	
后宫		-1.894**	-4.2354***
修正R ²	0.243	0.483	0.243

五、总结与反思

经过上述统计分析，我们得到如下结论：

- 是否为原创作品平均来说并不影响动画评分
- 导演和编剧的水平对动画评分具有至关重要的作用
- 漫画改作品评分较其他改编来源更高，事实上也只有足够优秀的漫画才能被选中改编；游戏改则多出于短期盈利目的，质量一般；GAL 改作品可能因为受众群体问题而获得较低评分
- 剧情和画质虽未经量化分析，但从异常值点作品来看，其暴雷原因无一例外是出于剧情或画质问题
- 符合主流审美的作品更容易获得高分，如带有热血、战斗、励志、治愈等标签的作品。
- 后宫标签已经成为负面标签

对此我们给出如下建议：

- 动画题材要贴近主流审美，选择热血、励志、治愈等标签可以提高动画观感。避开后宫、异世界等负面题材
- 剧情和画质要保持在平均水平，否则容易暴雷
- 导演和编剧对动画质量至关重要，必须选取可靠负责的主创团队

本研究还存在许多不足，如所选取的变量是较为简陋的，各模

型 R^2 普遍偏低，说明所选取因素远未能充分解释动画评分的变动，还有更多的因素影响着一部动漫的评分，譬如较好量化的声优的选取、是否得到视频平台的推广、导演和编剧的具体选择等分类变量，再譬如动画的剧情、画质等较难量化的变量，它们或许可以采取一些间接的方式来进行度量。

进一步的回归分析也是可行的，动画题材之间可能存在交互作用，如考虑百合向动画，注意到百合*少女的交互项显著地具有较高正向作用，这也极具分析价值。更深入地，构建词向量并用机器学习方法进行聚类分析，这或许可以找出动漫的主题，即一些涵盖多种标签的集合，如热血番总是与战斗、励志、友情等标签联系在一起。

最后，本研究的范围也仅局限于日本动漫，其对国产动漫的适用程度还需要进一步的探索。

表 5 百合*少女交互项回归

变量/模型	所有 TV 动画 (平方处理)
截距项	77.933***
导演标准分	4.714***
编剧标准分	5.688***
百合	0.150
少女	-1.894
恋爱	-1.591*
治愈	8.923***
百合*少女	9.517*
R^2	0.228