**广州商学院**

**课程论文**

**题目：动漫受众可视化分析**

1. **需求分析**

动漫是世界上最受欢迎的娱乐类型之一，它是以电视视频、电影、同人小说等形式制作的动画。动漫的例子有进击的巨人、死亡笔记、钢之炼金术师兄弟会、一拳超人、刀剑神域等。本次论文主要研究电视类型的动漫受欢迎的类型背后的人群和市场分析。本次调查研究使用的数据集来自于https://www.kaggle.com/datasets/harits/anime-database-2022网址，该数据集主要是来自MyAnimeList.net的网络抓取获得，网络抓取恰好在2022年9月20日实施，获得21460条动漫数据，28种词条。本次需求分析主要针对电视类型，查看动漫根据人群分布情况，得出动漫的主要市场；同时查看最受欢迎的前十部电视类型动漫跟最不受欢迎的前十电视类型动漫，查看它们的评分跟主题，查看最受欢迎的电视类型动漫的主题和最不受欢迎的电视类型动漫是否有相同点，得出什么主题吸引人群，和该主题吸引人群的主要原因。

1. **总体设计**







1. **详细设计**
2. 读取数据集，使用pandas库中的read\_csv方法读取，代码如下：

Anime = pd.read\_csv("C:/Users/jie/Desktop/202006050040+吕佩洁.csv")

1. 观察Type类型情况，使用matplotlib库绘制条形图1-1得出Type类型数目和每一种Type的数目。
2. 清洗数据
   1. 使用drop方法将Status类型等于Not yet aired的动漫进行去除
   2. 使用drop方法将Title类型等于Unknown的动漫进行去除
   3. 使用drop方法将Score类型等于Unknown的动漫进行去除
   4. 使用drop方法将Type类型等于Unknown的动漫进行去除
   5. 使用fillna方法将Score为空的值等于0，再使用drop方法把Score为0的动漫进行去除，代码如下：

new\_anime = Anime.drop(index= Anime.loc[( Anime['Status']=='Not yet aired')].index)

new\_anime = new\_anime.drop(index=new\_anime.loc[(new\_anime['Title']=='Unknown')].index)

new\_anime = new\_anime.drop(index=new\_anime.loc[(new\_anime['Score']=='Unknown')].index)

new\_anime = new\_anime.drop(index=new\_anime.loc[(new\_anime['Type']=='Unknown')].index)

new\_anime = new\_anime.fillna(0,inplace=False)

new\_score = new\_anime.drop(index=new\_anime.loc[(new\_anime['Score']==0)].index)

1. 使用loc方法将动漫按照Type类型进行分类（除去Unknown类型），一共分为六类，代码如下：
2. 观察这六个不同类型的的评分分布，使用matplotlib库中的boxplot方法绘制箱型图观察六个类型中评分最集中的类型（评分最集中的类型为电视类型），使用seaborn库和pyplot将箱型图1-2中的中文显示，代码如下：

TV = new\_score.loc[new\_score['Type'] == 'TV']

Movie = new\_score.loc[new\_score['Type'] == 'Movie']

OVA = new\_score.loc[new\_score['Type'] == 'OVA']

ONA = new\_score.loc[new\_score['Type'] == 'ONA']

Special = new\_score.loc[new\_score['Type'] == 'Special']

Music = new\_score.loc[new\_score['Type'] == 'Music']

plt.rcParams['font.sans-serif'] = ['SimHei']

sns.set\_style({'font.sans-serif':['SimHei','Arial']})

TV\_score = TV['Score']

Movie\_score = Movie['Score']

OVA\_score = OVA['Score']

ONA\_score = ONA['Score']

Special\_score = Special['Score']

Music\_score = Music['Score']

plt.boxplot((TV\_score,Movie\_score,OVA\_score,ONA\_score,Special\_score,Music\_score),\

labels=('电视类型','电影类型','OVA类型','ONA类型','特殊类型','音乐类型'))

plt.show()

1. 观察电视类型动漫Demographics种类和数目分布情况，使用 Counter函数统计Demographics种类和数目，利用pyecharts库中Bar函数绘制条形图1-3，代码如下：

from collections import Counter

Counter(TV['Demographics'])

from pyecharts.charts import Bar

from pyecharts import options as opts

attr = ['少年','青年','未知','少女','儿童','女性漫画','男性儿童','女性儿童']

data = [819,350,2417,298,300,52,17,10]

bar = (

Bar()

.add\_xaxis(attr)

.add\_yaxis('人群分类数目',data)

.set\_global\_opts(title\_opts=opts.TitleOpts(title='电视类型动漫按照人群分布分类情况'))

)

bar.render\_notebook()

.add\_xaxis(attr)

.add\_yaxis('人群分类数目',data)

.set\_global\_opts(title\_opts=opts.TitleOpts(title='电视类型动漫按照人群分布分类情况'))

)

bar.render\_notebook()

1. 观察电视类型动漫根据性别定位数目，使用pyecharts库中Pie函数和JsCode函数绘制环形图1-4，内圆需使用自定义函数绘制，代码如下：

c = (

Pie()

.add(

"",

[list(z) for z in zip(["男性", "女性"], [1186, 360])],

center=["30%", "30%"],

radius=[45, 65],

label\_opts=new\_label\_opts(),

)

.set\_global\_opts(

title\_opts=opts.TitleOpts(title="电视类型动漫性别占比"

),

legend\_opts=opts.LegendOpts(

type\_="scroll", pos\_top="20%", pos\_left="80%", orient="vertical"

),

)

)

c.render\_notebook()

1. 找出Popularity类型前十动漫跟倒数前十动漫，使用sort\_values函数按照ascending值进行排列，ascending值为True则按从大到小排列（若ascending值为False则按从小到大排列），再截取前十，代码如下：

TV\_Popularity = TV.sort\_values(by="Popularity",axis=0,ascending=True)[:10]

1. 查看前十和倒数前十的评分，利用折线图观察，收藏人数利用条形图观察，使用pyecharts库中Grid函数将Bar跟Line结合绘制为Grid图2-1与Grid图2-2，由于x轴字体过长需要旋转角度，代码如下：

from pyecharts.charts import Line

import pyecharts.options as options

from pyecharts.charts import Bar

from pyecharts.charts import Grid

from pyecharts.globals import ThemeType

from pyecharts.charts import \*

attr = np.array(TV\_Popularity['Title']).tolist()

line\_data = np.array(TV\_Popularity['Score']).tolist()

bar\_data = np.array(TV\_Popularity['Favorites']).tolist()

from pyecharts import options as opts

from pyecharts.charts import Bar, Grid, Line

bar = (

Bar()

.add\_xaxis(attr)

.add\_yaxis(

"收藏人数",

bar\_data,

yaxis\_index=0,

color="#5793f3", #稍微修改下颜色

)

.extend\_axis(

yaxis=opts.AxisOpts(

type\_="value",

name="评分",

min\_=0,

max\_=10,

position="left",

axisline\_opts=opts.AxisLineOpts(

linestyle\_opts=opts.LineStyleOpts(color="#d14a61")

),

axislabel\_opts=opts.LabelOpts(formatter="{value}"),

splitline\_opts=opts.SplitLineOpts(

is\_show=True, linestyle\_opts=opts.LineStyleOpts(opacity=1)

),

)

)

.set\_global\_opts(

yaxis\_opts=opts.AxisOpts(

name="收藏数",

min\_=30000,

max\_=210000,

position="right",

offset=0, #这里是Y轴间距，由80改为0即两个Y轴重合

axisline\_opts=opts.AxisLineOpts(

linestyle\_opts=opts.LineStyleOpts(color="#5793f3")

),

axislabel\_opts=opts.LabelOpts(formatter="{value}"),

),

title\_opts=opts.TitleOpts(title="电视类型前十受欢迎评分和收藏数"),

tooltip\_opts=opts.TooltipOpts(trigger="axis", axis\_pointer\_type="cross"),

xaxis\_opts=opts.AxisOpts(name\_rotate=60, axislabel\_opts={"rotate": 15})

)

)

.extend\_axis(

yaxis=opts.AxisOpts(

type\_="value",

name="评分",

min\_=0,

max\_=10,

position="left",

axisline\_opts=opts.AxisLineOpts(

linestyle\_opts=opts.LineStyleOpts(color="#d14a61")

),

axislabel\_opts=opts.LabelOpts(formatter="{value}"),

splitline\_opts=opts.SplitLineOpts(

is\_show=True, linestyle\_opts=opts.LineStyleOpts(opacity=1)

),

)

)

.set\_global\_opts(

yaxis\_opts=opts.AxisOpts(

name="收藏数",

min\_=30000,

max\_=210000,

position="right",

offset=0, #这里是Y轴间距，由80改为0即两个Y轴重合

axisline\_opts=opts.AxisLineOpts(

linestyle\_opts=opts.LineStyleOpts(color="#5793f3")

),

axislabel\_opts=opts.LabelOpts(formatter="{value}"),

),

title\_opts=opts.TitleOpts(title="电视类型前十受欢迎评分和收藏数"),

tooltip\_opts=opts.TooltipOpts(trigger="axis", axis\_pointer\_type="cross"),

xaxis\_opts=opts.AxisOpts(name\_rotate=60, axislabel\_opts={"rotate": 15})

)

)

title\_opts=opts.TitleOpts(title="电视类型前十受欢迎评分和收藏数"),

tooltip\_opts=opts.TooltipOpts(trigger="axis", axis\_pointer\_type="cross"),

xaxis\_opts=opts.AxisOpts(name\_rotate=60, axislabel\_opts={"rotate": 15})

)

)

1. 查看前十和倒数前十的Themes，绘制词云图3-1和词云图3-2代码如下：

TV\_Popularity['Themes']

from pyecharts.charts import WordCloud

s\_data = TV\_Popularity['Themes'].value\_counts()

s\_data = pd.DataFrame(s\_data)

words = list(s\_data.index)

num = list(s\_data['Themes'].values)

data = [k for k in zip(words, num)]

data = [(i,str(j)) for i, j in data]

workcloud = (WordCloud()

.add(series\_name=' ', data\_pair=data, word\_size\_range=[0,40])

.set\_global\_opts(title\_opts=opts.TitleOpts(

title='动漫主题词云图',title\_textstyle\_opts=

opts.TextStyleOpts(font\_size=23)),

tooltip\_opts=opts.TooltipOpts(is\_show=True))

)

workcloud.render\_notebook()

1. 查看前十和倒数前十的Rating，绘制扇形图4-1和扇形图4-2，代码如下：

from pyecharts import options as opts

from pyecharts.charts import Pie

from pyecharts.commons.utils import JsCode

Counter(TV\_Popularity['Rating'])

c = (

Pie()

.add(

"",

[list(z) for z in zip(['R - 17+ (violence & profanity)','PG-13 - Teens 13 or older'],[6,4])],

)

.set\_global\_opts(

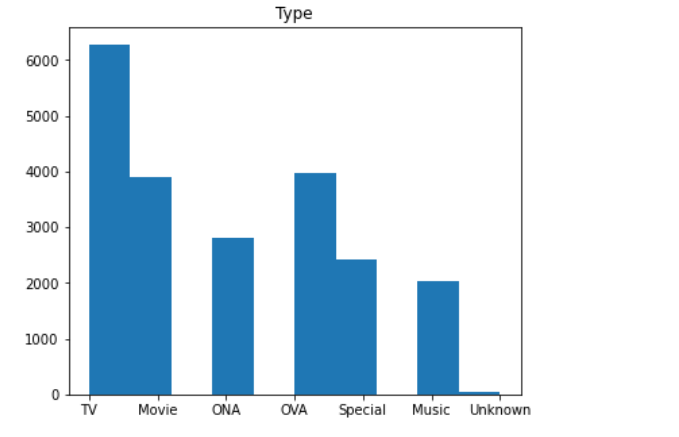
title\_opts=opts.TitleOpts(title="电视类型动漫前十受欢迎定位"),

)

)

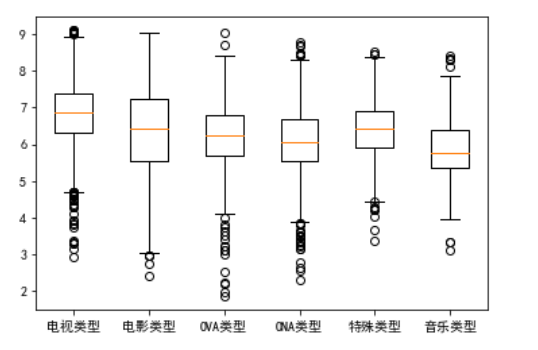
c.render\_notebook()

1. **程序运行结果测试与分析**
2. 由条形图1-1可得，若按Type类型对动漫进行分类，共有七类，其TV种类的动漫占市场大额



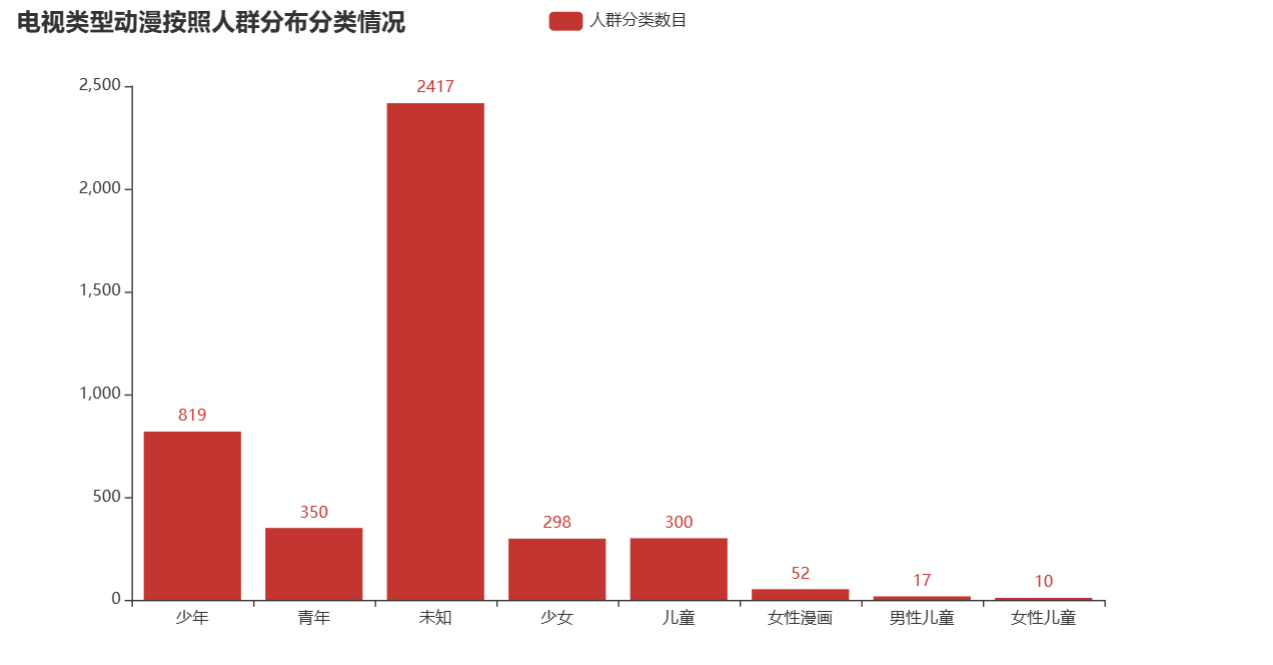
条形图1-1

1. 由箱型图1-2可得已知Type类型的评分分布，TV类型的评分相比其他较为集中且平均评分较高，因此本次分析主要选取TV类型进行分析



箱型图1-2

1. 由条形图1-3可得，有定位迎合人群的动漫主要迎合人群为少年，是10到17岁的男生，而有2417部动漫是没有定位迎合人群



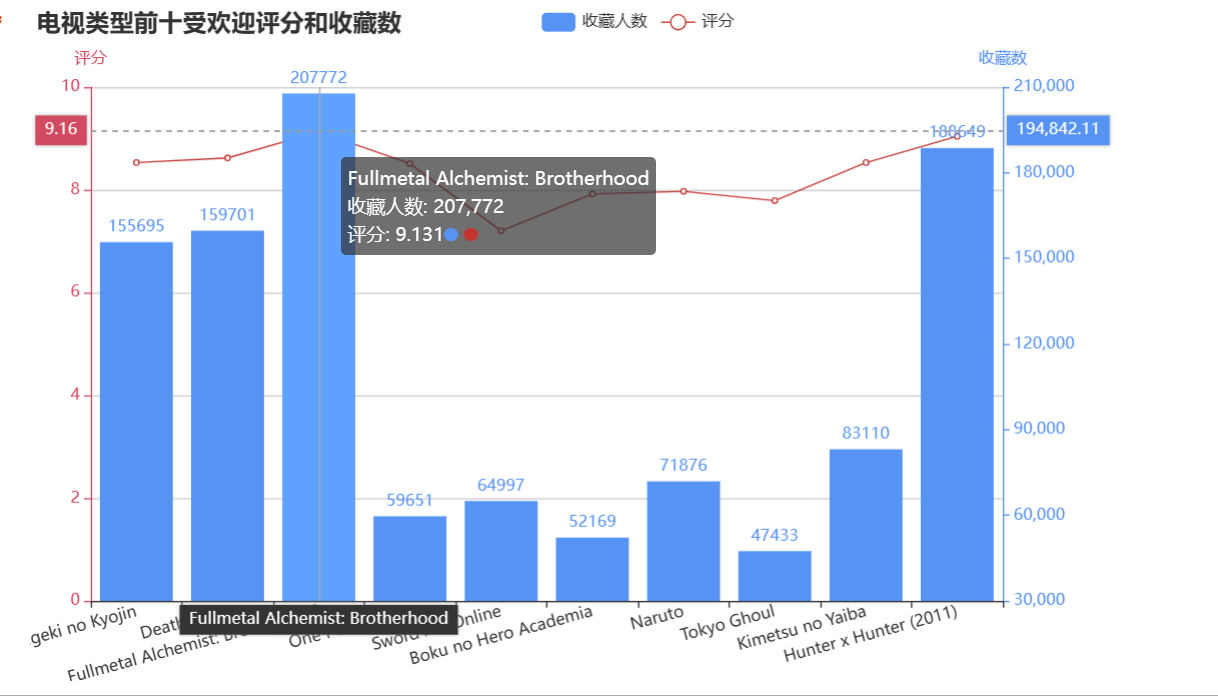
条形图1-3

若按性别划分，由扇形图1-4可得，以男性为主体目标的动漫有1186部，占统计动漫数目的四分之三。

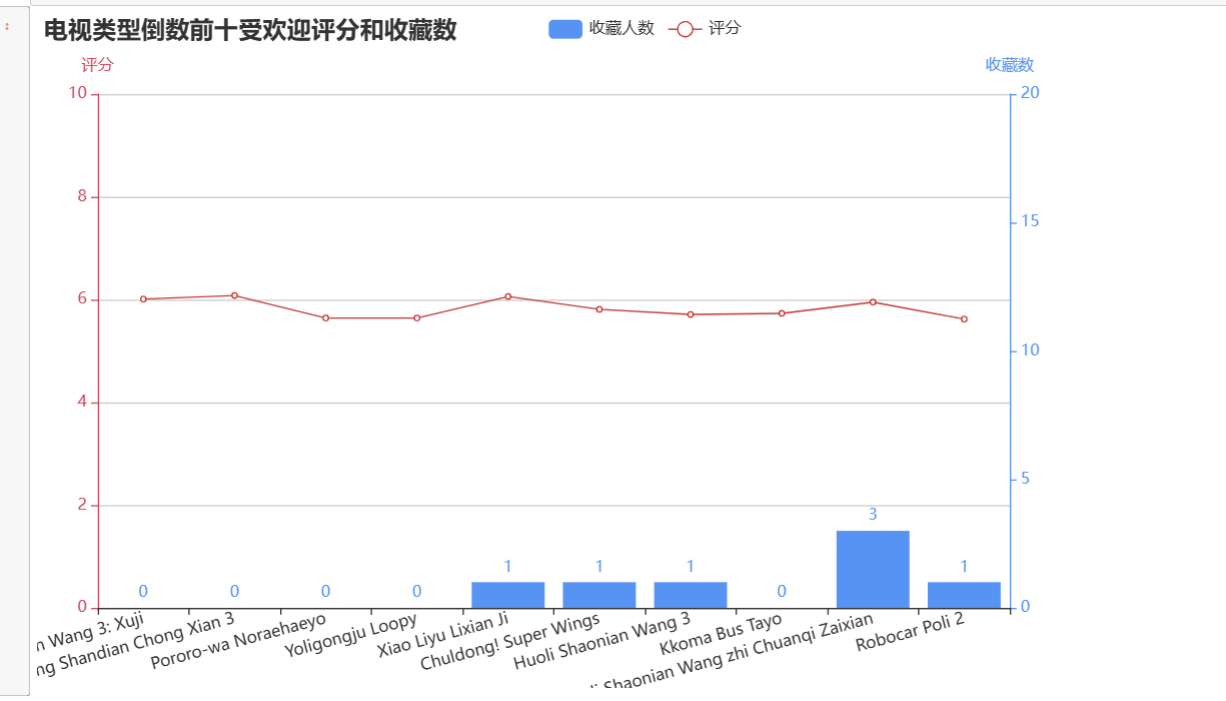


扇形图1-4

1. 由Grid图2-1可得最受欢迎前十动漫的评分在7到9分之间，收藏人数在4万以上，其中排在第三的钢之炼金术师：兄弟会评分最高，并且收藏人数也最多，而由Grid图2-2可得最不受欢迎前十动漫中有一半没有收藏，评分主要集中在6分左右

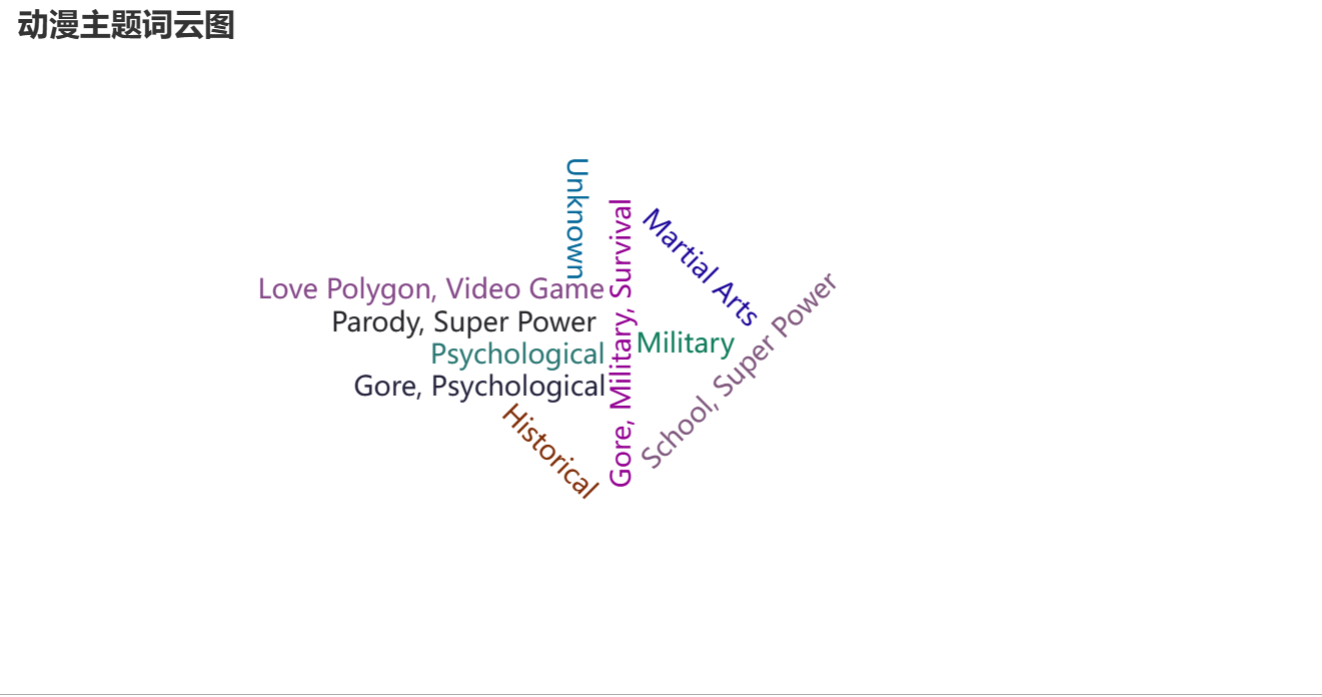


Grid图2-1



Grid图2-2

1. 由词云图3-1可得最受欢迎前十动漫的Themes，分析这十个关键词可得主题偏奇幻与暴力，而最不受欢迎前十动漫的Themes由词云图3-2可得，策略游戏跟赛车主题制作的动漫最容易不受人们欢迎。

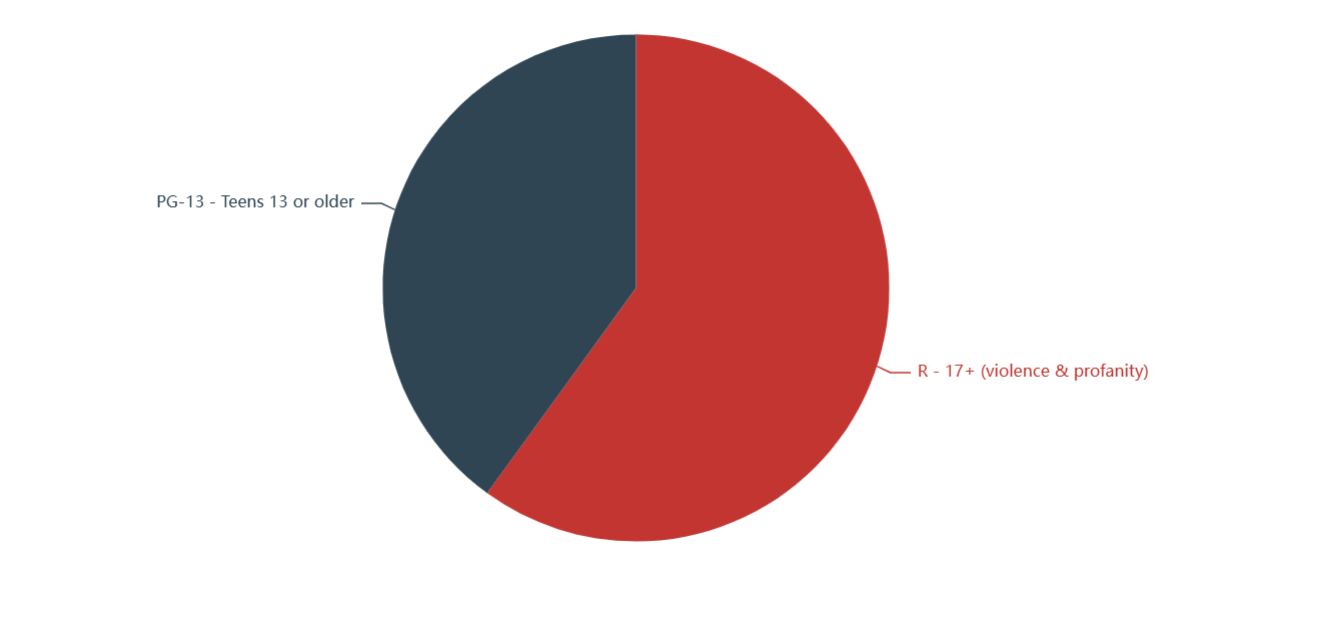


词云图3-1

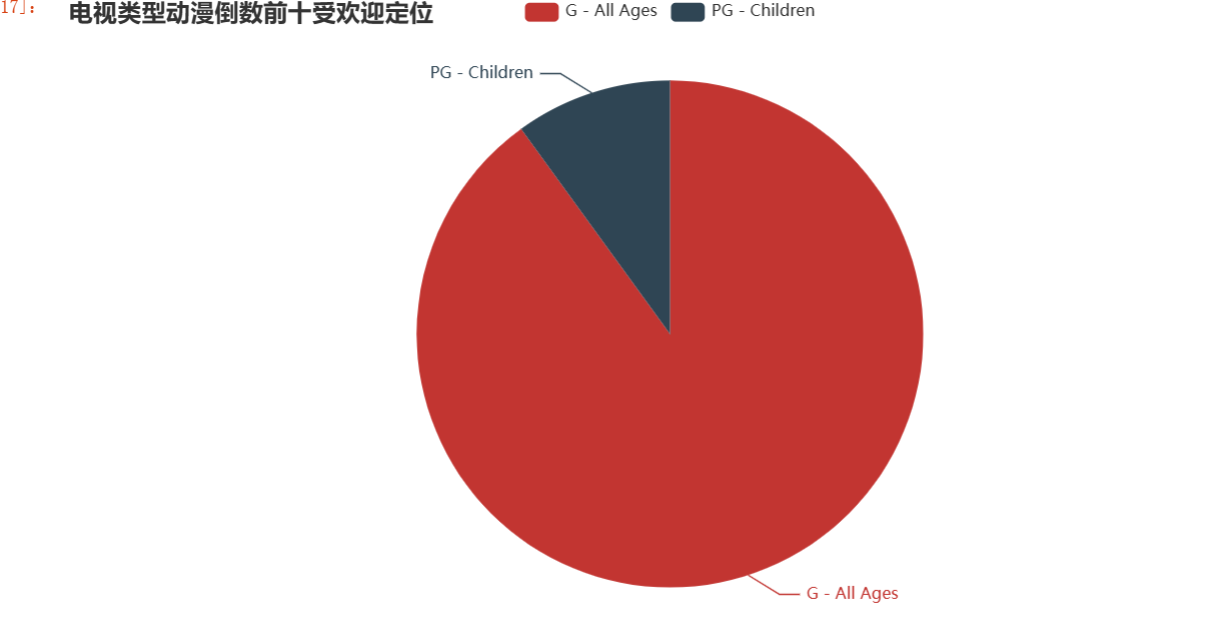


词云图3-2

1. 由扇形图4-1看出电视类型最受欢迎前十动漫年龄分段超过一半是17岁以上，偏暴力题材，剩下的也是13岁以上；而由扇形图4-2看出电视类型最受欢迎倒数前十动漫年龄分段中有九个都是所有年龄段，剩下一个是年龄段为儿童；由上述可得动漫类型主体在13岁以上



扇形图4-1



扇形图4-2

1. **结论与心得**

首先，动漫作品要针对受众心理特征。年龄不同，心理特点不一样，对作品的选择、理解和评价明显不同。

幼儿心底单纯，好奇心强，爱幻想，同时思维能力和理解能力不足，看的动画片需要卡通一些，形象可爱，情节直接，充满童趣，瑞士著名的心理学家皮亚杰认为幼儿期的孩子普遍具有一种特殊的心理现象———泛灵心理，所以他们喜欢跟动物说话，认为树会怕疼还会流眼泪。因此，幼儿喜欢童话和神话故事，喜欢拟人类的作品，像《猫和老鼠》《葫芦兄弟》等。

青少年延续着内心的好奇心、求知欲，同时也滋生着逆反心理，对自身意义的探索，因此喜欢唯美的人物形象，反映理想主义、探索类、少年维特式烦恼、叛逆、刺激、暴力、情感类的故事，最讨厌说教性强的作品。同时他们也需要轻松、快乐、新奇和幻想，新奇制胜的故事情节更能赢得青少年的喜爱。同时，虚拟的形象带给青少年的不仅仅是欣赏，更有一种人格代替的诱惑，精神分析学家认为，虚幻故事中的人物和动物被阐释为人们内心经历的投影，就像是映照人类内心的神奇镜子。青少年正处于寻求自我，渴望友情，展开人生理想的阶段，能够吸引他们的艺术作品不仅可以满足感官愉悦和刺激、尽情驰骋的想象空间，还可以体验戏剧化的生活和情感世界，作品中蕴含的思维理念和生活方式也将成为青年人伦理价值观的重要参考，积极的人生观念，如乐观精神、挑战困难的决心、团队意识、友情与亲情、环保观念等，消极的人生观，如暴力、同性恋、颓废、个人主义等，从而改变他们对自己、对他人、对社会、对世界的看法和行为。

成年人则因其工作与生活的压力、思想的成熟更加喜欢消遣、幽默、讽刺、寓意丰富、与自身生活的酸甜苦辣密切相关的作品等。老年人特有的怀旧心理则更喜欢看怀旧类的或富含童心的作品。中国动画片《魔比斯环》耗资1．3亿，制作周期长达5年，虽然动用了众多全球3D动画界的顶尖高手，体现“正义与邪恶的较量”这一永恒的主题，却从一进入市场就陷入了“大人看了嫌幼稚，小孩看了嫌深奥”的尴尬之中，角色形象不符合青少年的审美眼光，情节直露无法引人入胜。这部作品的失败与没有深入了解受众的接受心理有很大关系。

其次，接受主体的图式也是受众欣赏和接受艺术作品的重要影响因素。

皮亚杰认为一切刺激只有被主体的图式“同化”或“顺应”才能够发挥作用，否则人对外界刺激会视而不见、充耳不闻、感而不觉。格式塔心理学家阿恩海姆提出“异质同构”说，认为人们对某种东西之所以会喜欢和感兴趣，是因为这个东西在结构图式上与人的心理结构图式上有异质同构性。受众的接受图式影响他对接受作品的选择和感受的强弱，如果受众所选的作品内容或情绪表达与其阅历与期望相似，他就会产生强烈的共鸣，对作品的感受也会更强烈，但如果作品的内涵彻底超出主体的接受图式，他也会无法理解和接受这样的作品。同时受众的接受图式也会随欣赏和接受艺术作品的熏陶而得到拓展与提升。在接受图式中，情感是一项非常重要的变量。受众在欣赏艺术作品时，情感贯穿始终，只有积极情感的唤起才能充分的激活大脑中已经积累下来的丰富的表象因素，情感越是丰富，表象活动也越频繁，多样的表象是受众理解和接受艺术作品的重要前提。从这个角度来看，能够唤起受众的情感也成为艺术作品的重要追求，从打动情感入手调动受众心理上对作品的认同。行为主义心理学家贝里尼揭示出情绪与受众欣赏艺术过程的关系，他认为艺术品最典型的特征就是当其出现在受众面前时，会引起某种兴奋。兴奋是由于艺术品中包含某种诱发受众产生联想或解释及冲突的刺激物，受众通过对艺术品中刺激物的全面把握形成一种图式，冲突得到解决，兴奋度逐渐下降，愉悦感随之产生。“以情动人”使受众接受艺术作品所传达的内涵并与之共鸣，然后到“理性思考”，并通过受众自己的思考，得到心灵的净化。

**参考文献：**[1]李莉.动漫受众心理分析[J].中国校外教育,2013(12):34-35.

**课程论文（设计）评分表**

|  |  |
| --- | --- |
| 评分项目 | 得分 |
| **创新性、合理性（10分）**  9-10：表明设计具有创新性、合理性；  6-8：表明设计具有新意、合理性；  5分及以下：表明设计合理性较差。 |  |
| **选题难度、实现复杂性（30分）**  26-30：难度系数高，设计实现功能复杂；  21-25：难度系数较高，设计实现功能较复杂；  16-20：难度系数一般，设计实验功能一般  15分以下：难度系数很低，设计实现功能一般。 |  |
| **完成情况包括功能完整性、工作量、界面美观友好（30分）**  20-30：按要求完成课题的全部功能，有完整的符合标准的文档，文档有条理、文笔通顺，格式正确，其中有总体设计思想的论述，有正确的流程图，程序完全实现设计方案，设计方案先进，软件可靠性好；  10-19：完成课题规定的功能，有完整的符合标准的文档，文档有条理、文笔通顺，格式正确；有基本实现设计方案的软件，设计方案正确，但有少数失误；  9以下：完成课题规定的大部分功能，有完整的符合标准的文档，有基本实现设计方案的软件，设计方案基本正确，个别功能没有实现。 |  |
| **课程设计报告（20分）**  18-20：优秀，表明报告清楚详细地说明了设计的过程，格式规范、内容完整，阐述清晰，层次分明；  15-17：良好，表明报告说明了设计的过程，格式较规范、内容比较完整，阐述清晰，有层次；  10-14：中，表明报告说明了设计的过程说明不太清晰，格式不规范、内容比较简单；  9以下：差，表明报告过于简单。 |  |
| **答辩表现（10分）**  9-10：自己的设计阐述清晰，回答问题准确，基本概念非常清楚；  6-8：自己的设计阐述比较清晰，回答问题较准确，基本概念比较清楚；  5分以下：自己的设计阐述不清晰，回答问题不准确，基本概念不清楚。 |  |
| **总得分**： |  |