

○○○○

应用统计学 分析报告

71086032 曾诗仪

○○○○

目录 / 流程

- 背景分析
- 确定研究问题
- 选着统计分析方法
- 收集样本数据
- 数据分析
- 结论及建议
- 调查方案优缺点分析



确定研究问题：

北航大学生娱乐消费情况调查

确定研究问题

■ 背景分析：

随着经济的发展，大学生消费观念逐渐多元化，娱乐消费市场逐渐成为各类企业争夺的焦点。这也对位于北京的企业的服务和营销提出了更高的要求。

■ 确定研究问题：

为此，开展大学生娱乐消费情况调查有助于了解大学生娱乐消费的趋势、特点和需求变化，为相关企业和部门制定更加精准、高效的营销策略和服务，提供决策参考。

选择统计方法

■ 问卷设计：

- 本次调查问卷的设计，首选由要分析的数据出发，并结合实际，设计出一系列与该课题有关的问题。

■ 选择处理软件：


- 利用问卷星收集相关问卷，再利用EXCEL统计软件进行综合统计分析。




选着统计方法

- 问卷内容：可被分为3个部分：
 - 第一部分为**基础的个人信息**，涉及的信息包括被调查者的性别、年级、生活费水平以及每月花费在娱乐方面消费，这一部分的数据主要是为了与后两段的数据一起做相关性的分析。
 - 第二部分为**大学生娱乐消费情况**，包括大学生选择某种娱乐消费选择支出、频率等等。
 - 第三部分为**相关原因调查**，主要涉及第二部分的问题，对其进行说明（原因），便是被调查者选择某种娱乐消费的对于娱乐消费的观点以及评分。

收集样本数据

北航大学生消费情况调查调查 

ID:216770908 • 未发布 答卷:70 5月09日 12:54

 设计问卷 ▾  发送问卷 ▾  分析&下载 ▾

 发布  复制  删除  文件夹  提醒

- 本次问卷题数为最多27题，其中包含逻辑选填题。
- 本次问卷发放采用的是线上分享的形式进行收集。
- 为了增加问卷分析有效率，期望问卷填写人数为60人，实际收集到的问卷数为70份。

数据分析

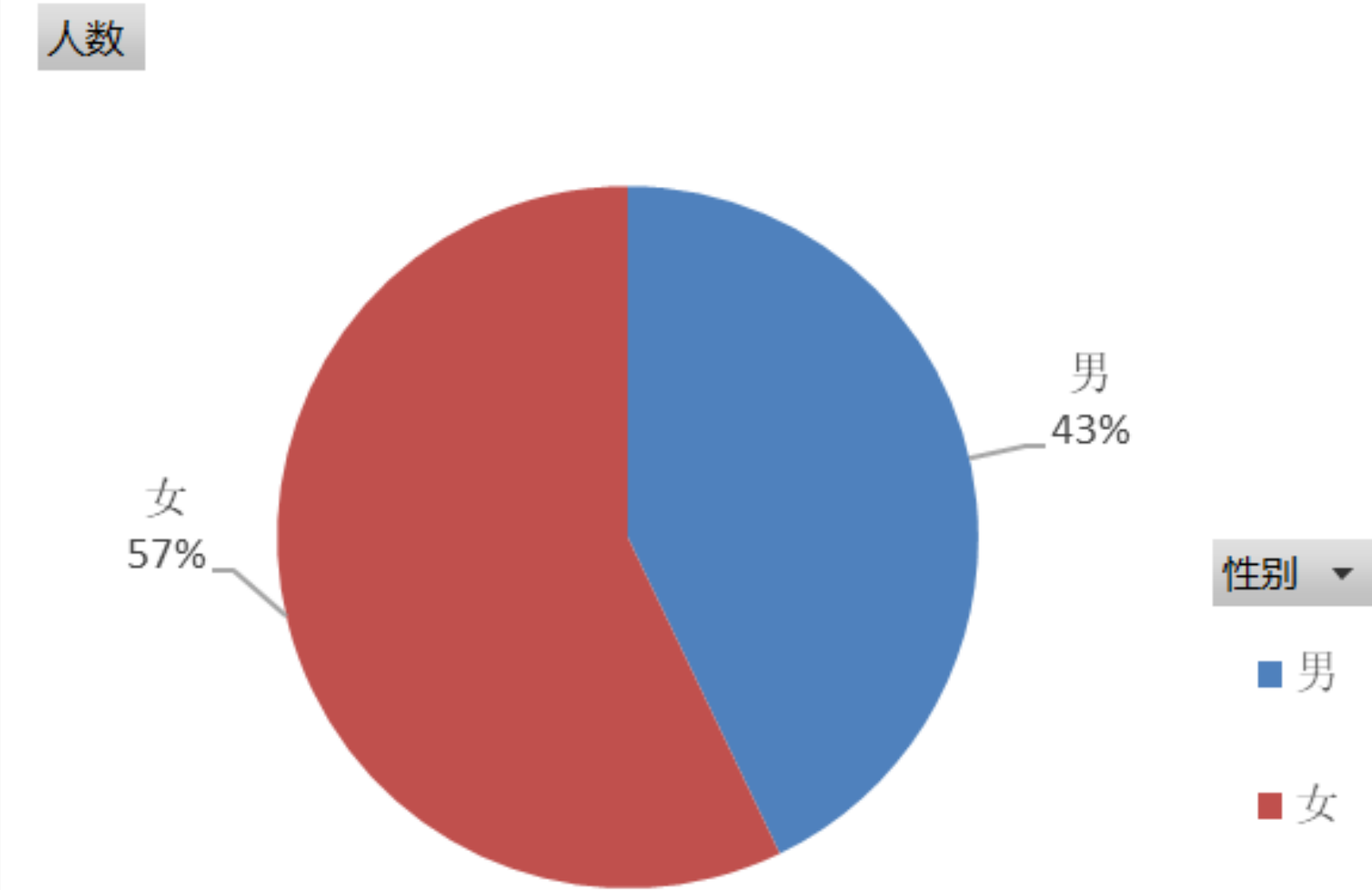
■ 初步分析：

- **描述性统计分析：** 使用平均数、中位数、众数、标准差、四分位数等方法对娱乐消费情况进行描述和总结，以便了解大学生的娱乐消费水平和消费倾向。
- **频率分析：** 使用频率分布表、直方图或数据分析对大学生的娱乐消费情况进行分析，以便了解大学生在不同的娱乐项目上的消费频率和消费金额。
- **单因素方差分析：** 将大学生按照性别、年级等因素进行分类，以便了解不同因素对大学生的娱乐消费有无显著差异。
- **假设性分析：** 根据大学生的娱乐消费水平和频率等因素，进行相关的假设性分析，如：男女对于直播打赏频率相等的假设性分析

数据分析-单项统计

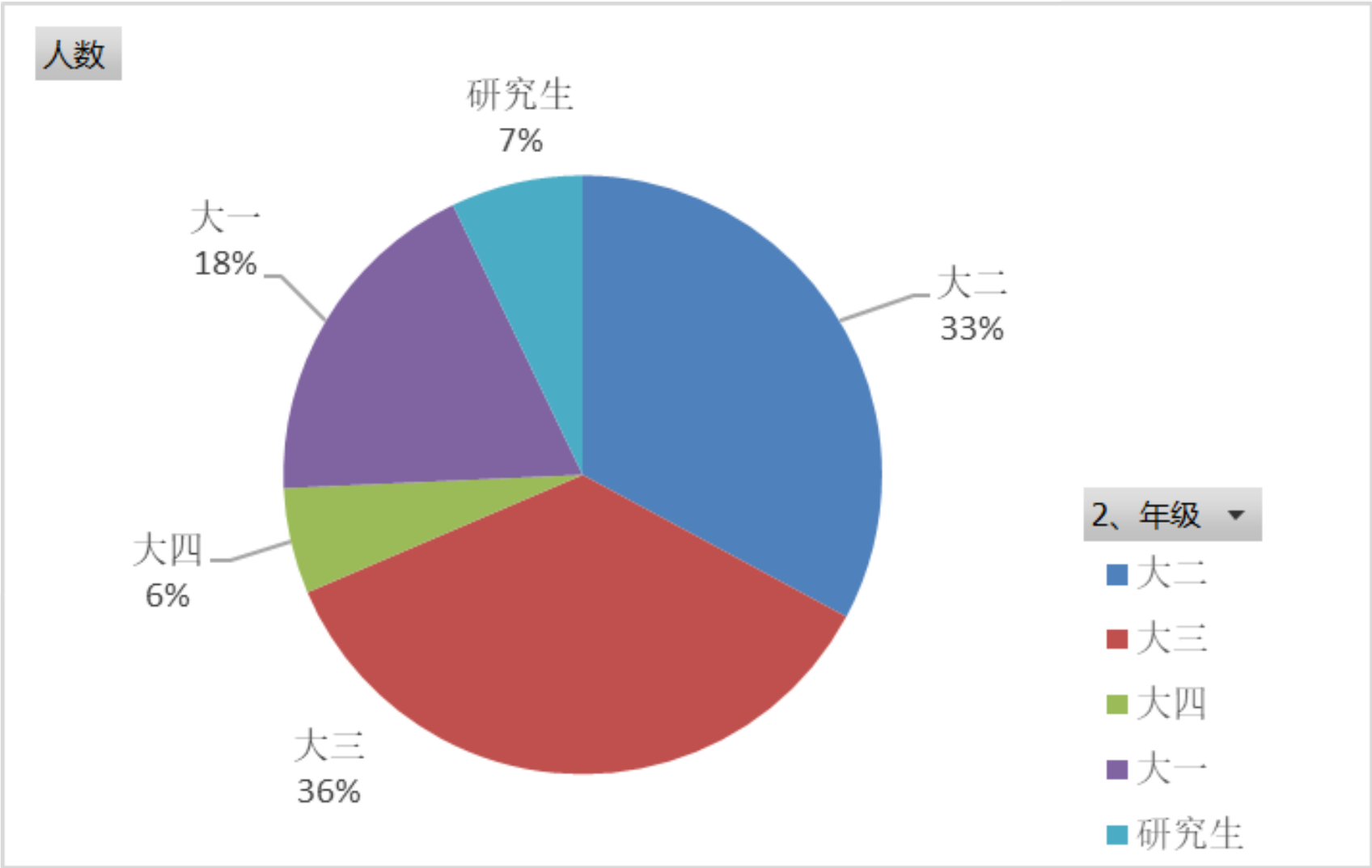
性别	人数
男	30
女	40

■ 性别分布：
在调查获得的有效数据中，男生共有30人，女生共有40人。



数据分析-单项统计

年级分布	人数
大二	23
大三	25
大四	4
大一	13
研究生	5
总计	70



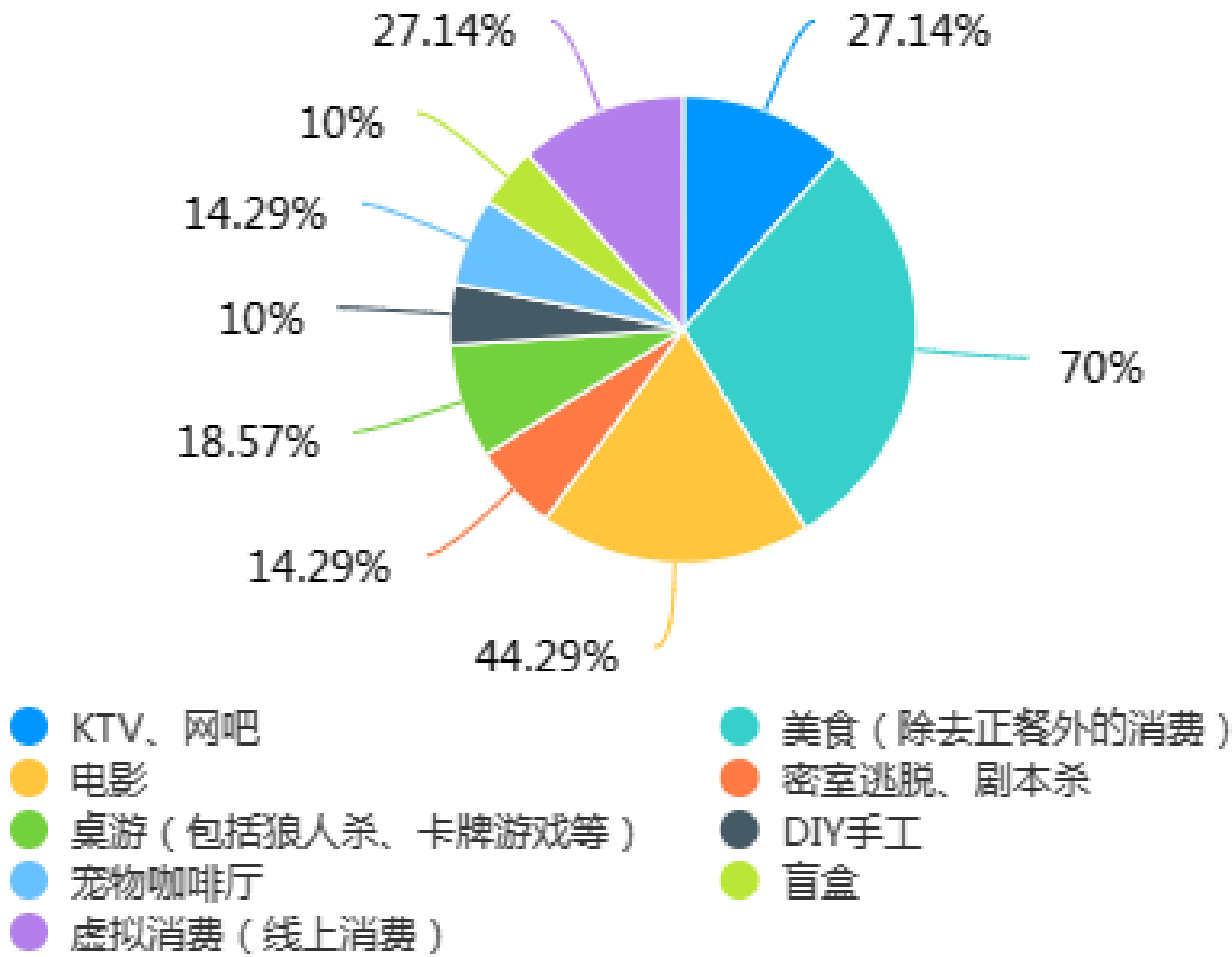
■ 被调查人群分布（年级）：

在身份调查中，我们对所有人群进行调查。其中包括本校各年级的本科生以及研究生，被调查者人群占比最大的是北航大三的学生，其次是大二的。

数据分析-单项统计

第5题：通常在会选择什么娱乐方式（最多3项） [多选题]

选项	小计	比例
美食（除去正餐外的消费）	49	70%
电影	31	44.29%
KTV、网吧	19	27.14%
虚拟消费（线上消费）	19	27.14%
桌游（包括狼人杀、卡牌游戏等）	13	18.57%
密室逃脱、剧本杀	10	14.29%
宠物咖啡厅	10	14.29%
DIY手工	7	10%
盲盒	7	10%
本题有效填写人次	70	

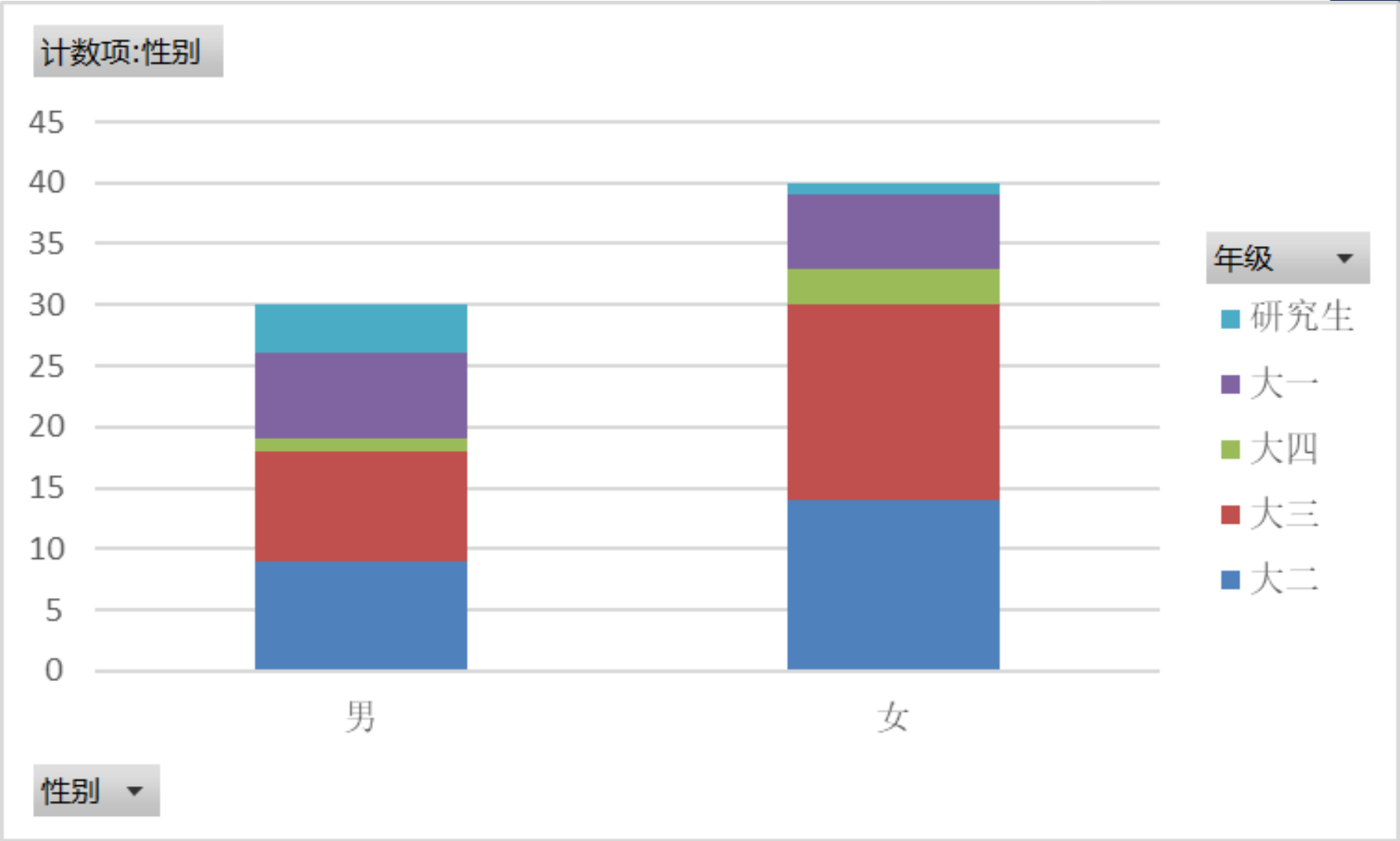


■ 消费项目：

在此问卷调查中，北航大学生选择美食作为娱乐消费的人群是非常庞大的，共占了**70%**，其次是电影以及KTV或网吧等；占比最少的是盲盒以及DIY手工。

数据分析-双项统计

计数项:性别		年级					总计
性别		大二	大三	大四	大一	研究生	
男		9	9	1	7	4	30
女		14	16	3	6	1	40
总计		23	25	4	13	5	70

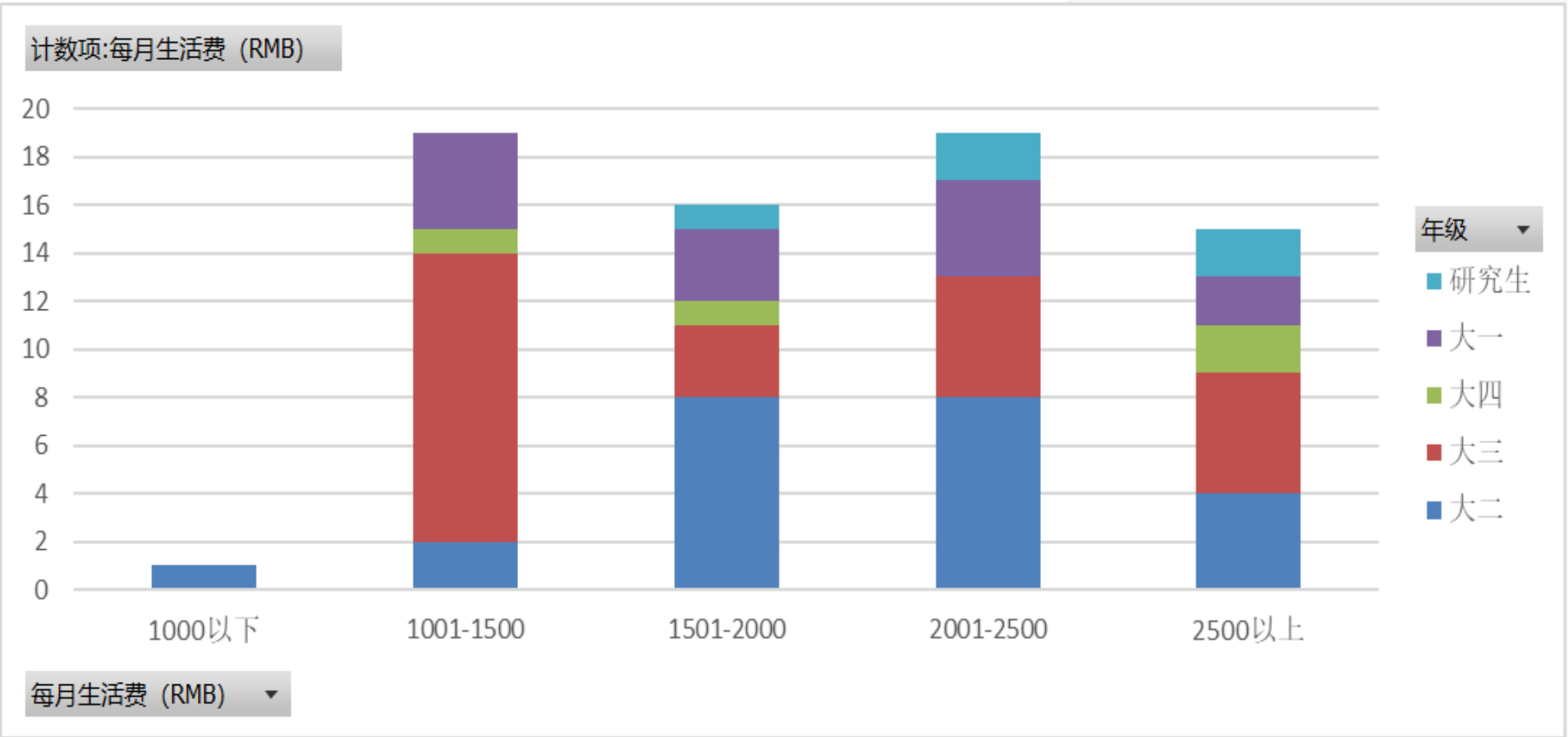


■ 被调查人群分布（年级与性别）：

从数据显示可以看出男性被调查者大二大三的占比是较多的，分别有**9**人；而女性的则是大三的居多，共有**16**人。

数据分析-双项统计

计数项:每月生活费 (RMB)	年级						
每月生活费	大二	大三	大四	大一	研究生	总计	
1000以下	1					1	
1001-1500	2	12	1	4		19	
1501-2000	8	3	1	3	1	16	
2001-2500	8	5		4	2	19	
2500以上	4	5	2	2	2	15	
总计	23	25	4	13	5	70	

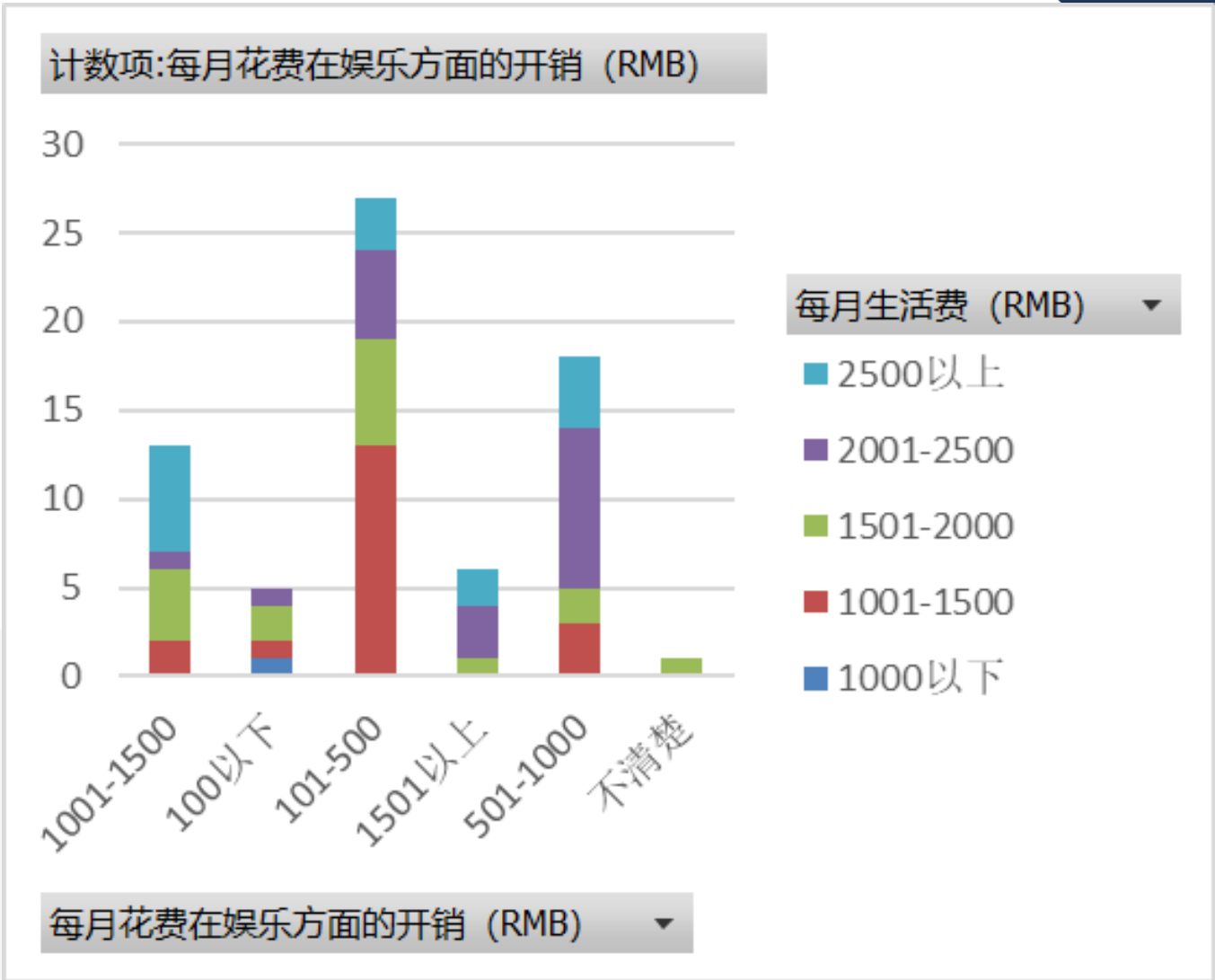


■ 被调查人群分布（年级与每月的生活费）：

从数据显示可以看出大多数的北航大三的学生（12名），每月的生活费大约为1001-1500以内，而只有一位大一的学生每月的生活费为1000以内。

数据分析-双项统计

计数项:每月花费在娱乐方面的开销 (RMB)	每月生活费	1000以下	1001-1500	1501-2000	2001-2500	2500以上	总计
每月花费在娱乐的金额							
1001-1500			2	4	1	6	13
100以下		1	1	2	1		5
101-500			13	6	5	3	27
1501以上				1	3	2	6
501-1000			3	2	9	4	18
不清楚				1			1
总计		1	19	16	19	15	70



■ 调查每月生活费（因）以及花费在娱乐上的金额）：

从数据显示可以看出每月生活费处于1001-1500的北航大学生，每月花费101-500在娱乐上的占比是最多的。

数据分析-假设检验

■ 由于 $t < \text{左临界值}$ ，落入拒绝域，因此拒绝原假设，接受备选假设。

备选假设H1：男性每月花费在美食次数<女性每月花费在美食次数

即：男女性别每月花费在美食的次数是有显著性差别的。

值	列标签	男	女	总计
样本数		20	29	49
平均值		3.95	4.310344828	4.163265306
标准差		1.050062655	0.967450563	1.007201957
原假设H0： 男性每月花费在美食次数大于等于女性每月花费在美食次数				
备选假设H1： 男性每月花费在美食次数<女性每月花费在美食次数				
不同性别每月每月平均花费在美食的次数两样本均值之差的假设检验				
• 两总体方差未知，不相等，统计量服从t分布				
• 输入的数据来自样本统计指标				
假设检验表				
项目	数值或结论			
检验常数	0			
原 假 设H0 均值A-均值B≥	0			
备选假设H1 均值A-均值B<	0			
显著性水平α	5%			
均值之差xA-xB	-0.3603			
自由度df	38.7465			
统计量t	-1.2188			
拒绝域左临界值t(α)	-1.6860			
统计量是否落入拒绝域	是			
是否拒绝原假设	是			
是否接受备选假设	是			

数据分析-假设检验

■ 由于 $t <$ 右临界值，未落入拒绝域，因此不能拒绝原假设，不能接受备选假设。

即：无法判断男性每月的娱乐支出是否大于女性。

值	列标签		总计
	男	女	
样本数	30	40	70
均值	960	867.5	907.1428571
标准差	576.3140906	523.0029171	544.3574669
原假设H0：男性每月娱乐支出小于等于女性每月娱乐支出			
备选假设H1：男性每月娱乐支出>女性每月娱乐支出			
不同性别每月每月平均花费在美食的次数两样本均值之差的假设检验			
• 两总体方差未知，不相等，统计量服从t分布			
• 输入的数据来自样本统计指标			
• 两样本均值右边尾t检验：			
假设检验表			
项目	数值或结论		
检验常数	0		
原 假 设H0 均值A-均值B≤	0		
备选假设H1 均值A-均值B>	0		
显著性水平α	5%		
均值之差xA-xB	92.5000		
自由度df	59.1174		
统计量t	0.6912		
拒绝域右临界值t(2*α,df)	1.6711		
统计量是否落入拒绝域	否		
是否拒绝原假设	否		
是否接受备选假设	否		

数据分析-显著性差异

1月的娱乐支出是否有显著性差异												
	每月生活费（千）数据取最大值					方差分析：单因素方差分析						
样本数	1千以下	1001-1500	1501-2000	2001-2500	2500以上	SUMMARY						
1	100	100	2000	1000	500	组	观测数	求和	平均	方差		
2		500	1500	1000	1500	1千以下	1	100	100	#DIV/0!		
3		500	500	500	1500	1001-1500	19	12600	663.1579	133567.3		
4		1000	100	1000	1000	1501-2000	16	13200	825	375333.3		
5		1000	500	500	1000	2001-2500	19	19100	1005.263	294970.8		
6		500	100	1000	1500	2500以上	15	18500	1233.333	245238.1		
7		1500	0	500	2000							
8		500	1500	1000	500							
9		500	500	1500	500							
10		500	500	1000	1000	方差分析						
11		500	1500	2000	2000	差异源	SS	df	MS	F	P-value	F crit
12		1000	1000	500	1500	组间	3669411	4	917352.8	3.554144	0.011034	2.51304
13		1500	500	500	1500	组内	16777018	65	258108			
14		500	1500	1000	1000	总计	20446429	69				
15		500	1000	2000	1500							
16		500	500	2000								
17		500		1000								
18		500		1000								
19		500		100								

- 方差分析1：不同月生活费下每月的娱乐支出是否有显著性差异；
- P值小于0.05，因此不同月生活费下每月的娱乐支出有显著性差异。
- F值大于F-crit（临界值），因此不同月生活费下每月的娱乐支出有显著性差异。

数据分析-显著性差异

2.数据筛选： 每月平均花费在美食的次数以及附件美食店评分数据

	每月平均花费在美食的消费次数（次）				
样本数	1	2	3	4	5
1		4	3	3	4
2		4	4	4	5
3		4	4	3	3
4		3	3	2	3
5			3	4	4
6			4	4	4
7			5	4	1
8			3	5	4
9			5	5	4
10				3	5
11				5	2
12					5
13					3
14					5
15					4
16					4
17					5
18					4
19					5
20					5
21					5
22					5
23					5
24					5
25					5

方差分析：单因素方差分析

SUMMARY				
组	观测数	求和	平均	方差
1	0	0	#DIV/0!	#DIV/0!
2	4	15	3.75	0.25
3	9	34	3.777778	0.694444
4	11	42	3.818182	0.963636
5	25	104	4.16	1.14

方差分析

差异源	SS	df	MS	F	P-value	F crit
组间	1.677673	4	0.419418	0.42618	0.788902	2.583667
组内	43.30192	44	0.984135			
总计	44.97959	48				

不能肯定各组均值是否有显著性差异

不能肯定各组均值是否有显著性差异

- 方差分析2： 每月平均花费在美食次数以及美食附近美食店评分数据是否有显著性差异；
- P值大于0.05， 因此不能肯定两者之间是否有显著性差异。
- F值小于F-crit（临界值）， 因此不能肯定两者之间是否有显著性差异。



THANK YOU

