

# 《传播统计学》

## 统计分析与Excel操作

教师：林志良

邮箱：[linzhl@nfu.edu.cn](mailto:linzhl@nfu.edu.cn)

个人网站：[www.zhilianglin.com](http://www.zhilianglin.com)

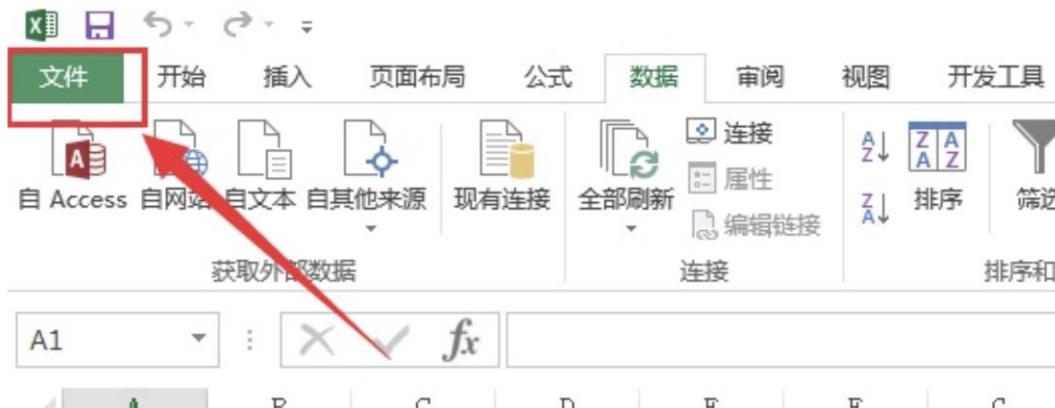
# 目录

- Excel统计分析说明
- 描述性统计：
  - 类别型变量
  - 数值型变量
- 推断性统计：
  - 双变量分析
  - 回归分析

# Excel统计分析说明

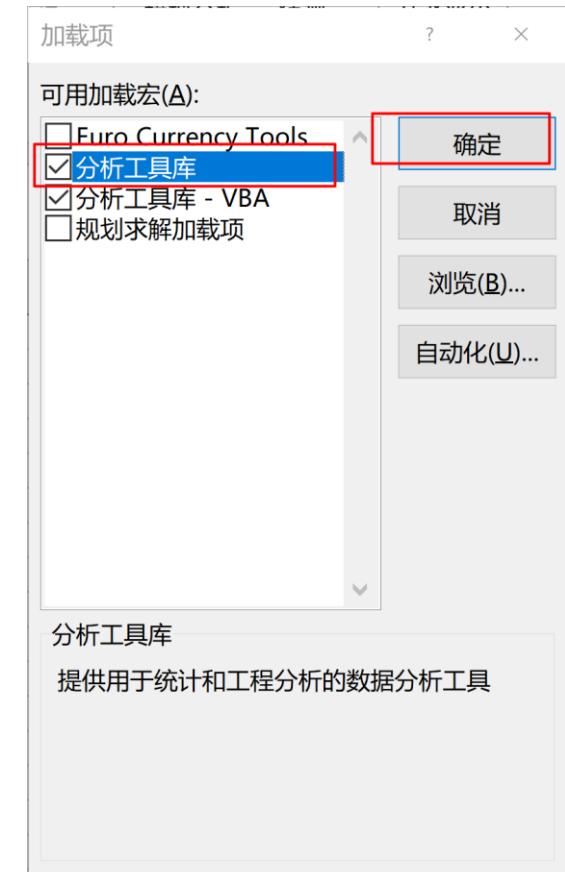
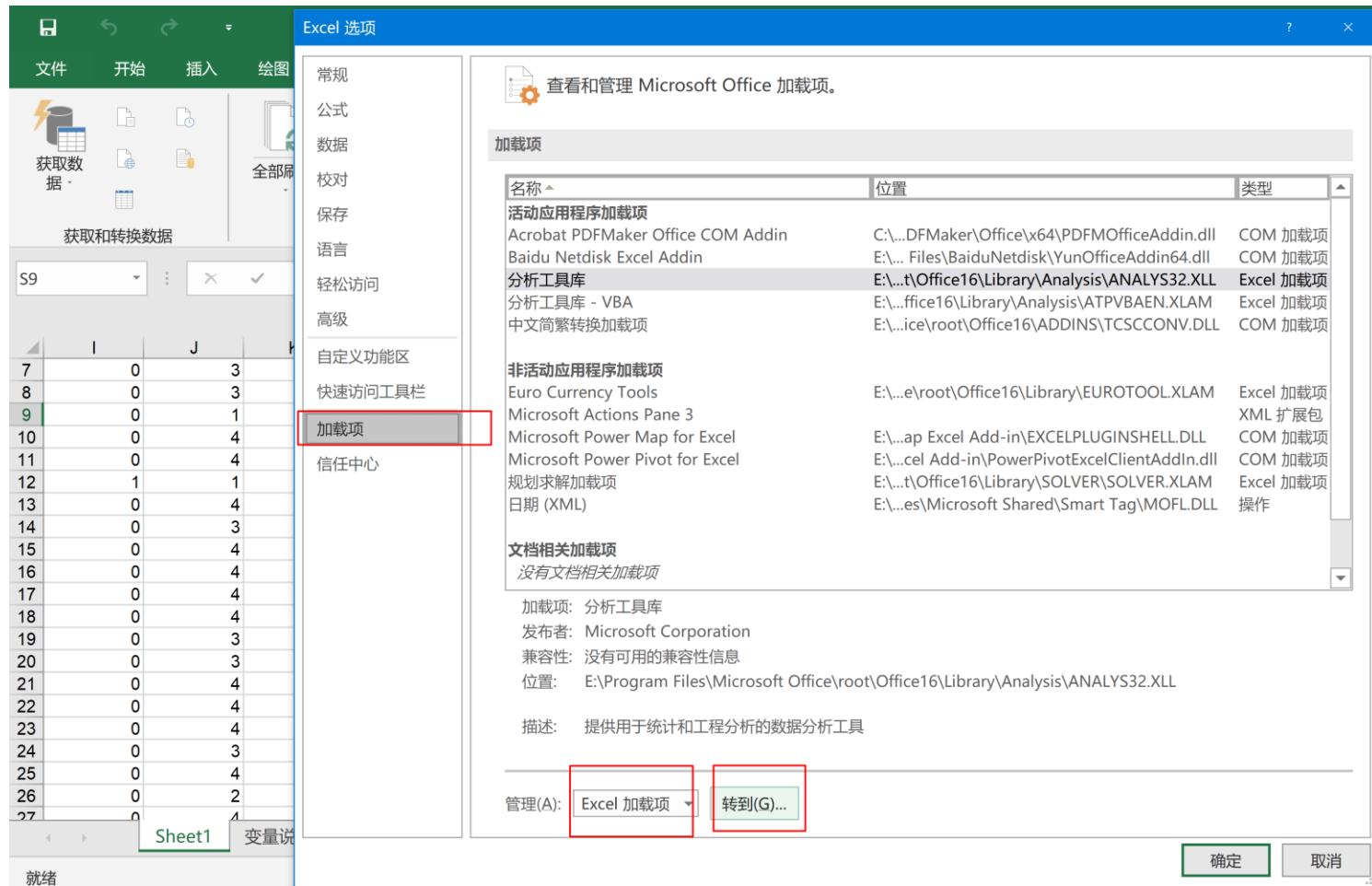
## Excel数据工具包准备

【文件】>【选项】>【加载项】>【Excel 加载项】>【转到】>勾选【分析工具库】>确定



# Excel统计分析说明

## Excel数据工具包准备



# Excel统计分析说明



# 描述性统计

- 要可视化**类别型变量**, 可以选择**条形图** (bar chart) 或者**饼状图** (pie chart)
- 要可视化**数值型变量**, 可以选择**直方图** (histogram)
- 注意: **条形图**和**直方图**虽然长得比较像, 但是是适用的变量类型是不同的

# 描述性统计

数据透视表字段

选择要添加到报表的字段:

搜索

gender

更多表格...

在以下区域间拖动字段:

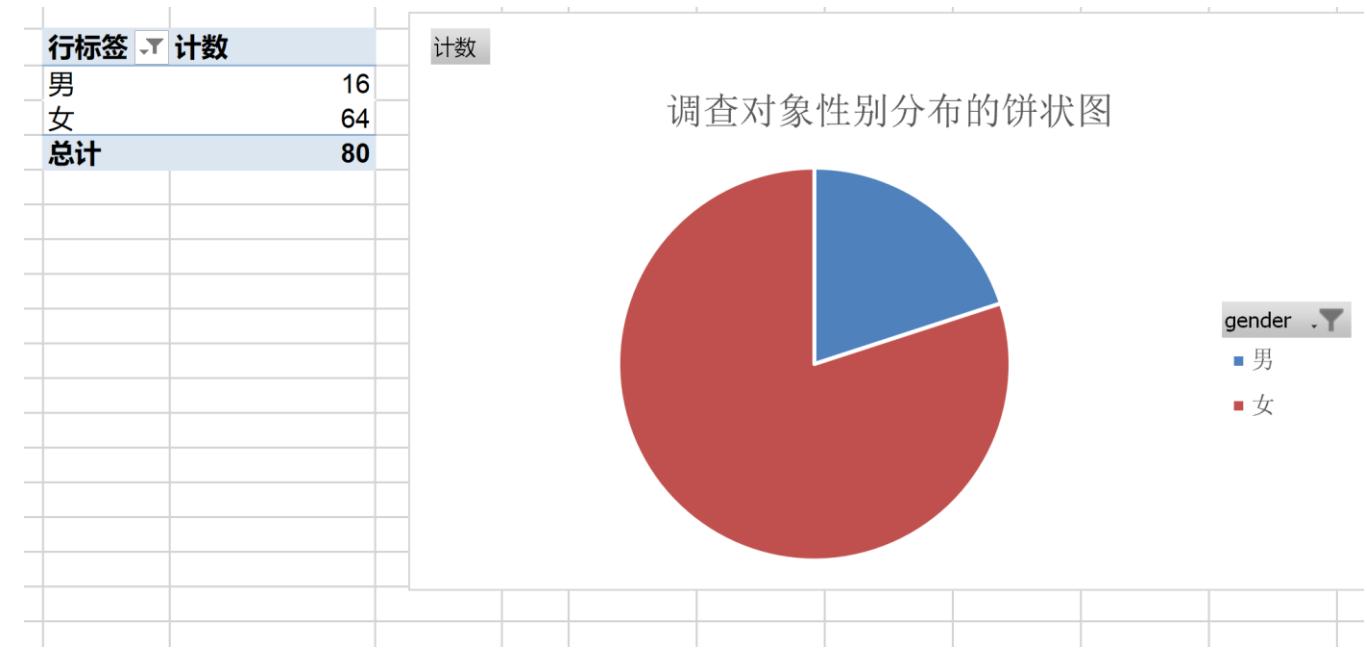
筛选	列

行	Σ 值
gender	计数项:gender

延迟布局更新

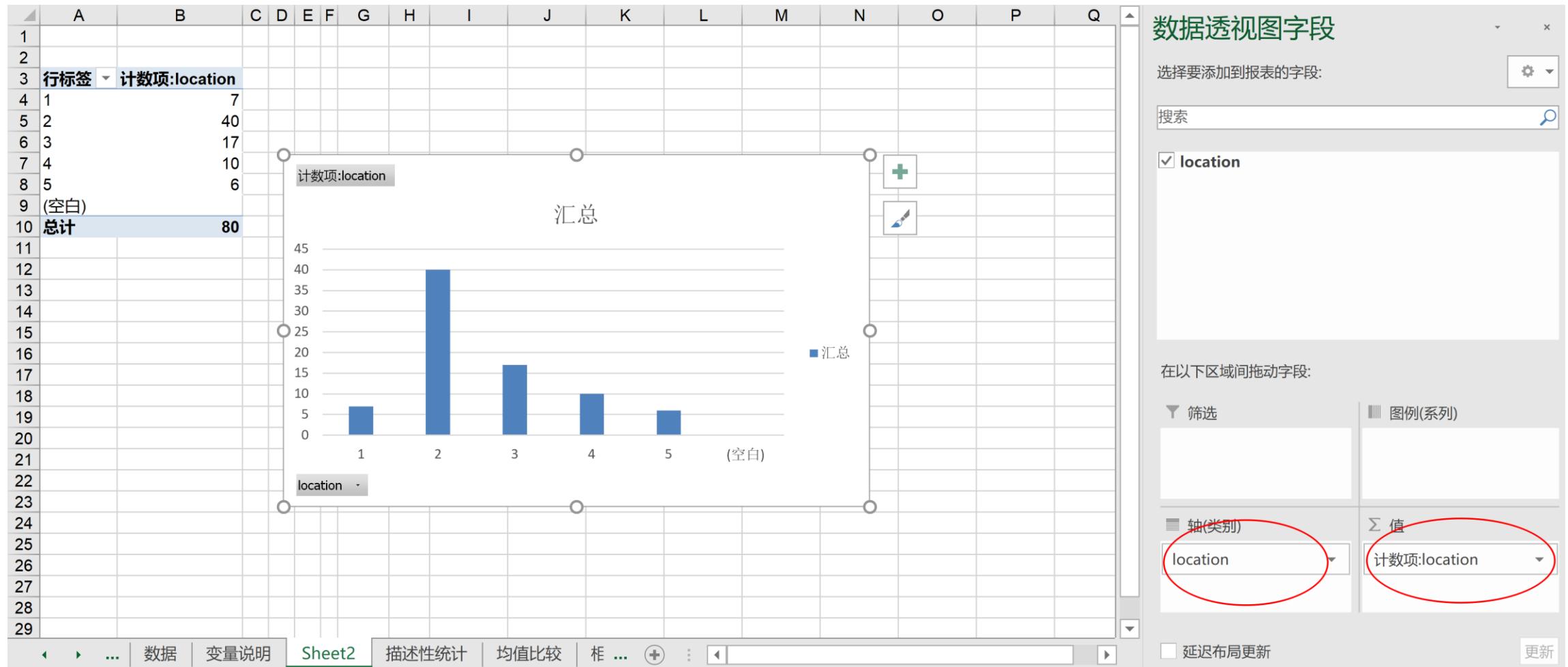
## 类别型变量

先用数据透视表功能整理完数据后  
再绘图



# 描述性统计

## 类别型变量



# 描述性统计

## 数值型变量

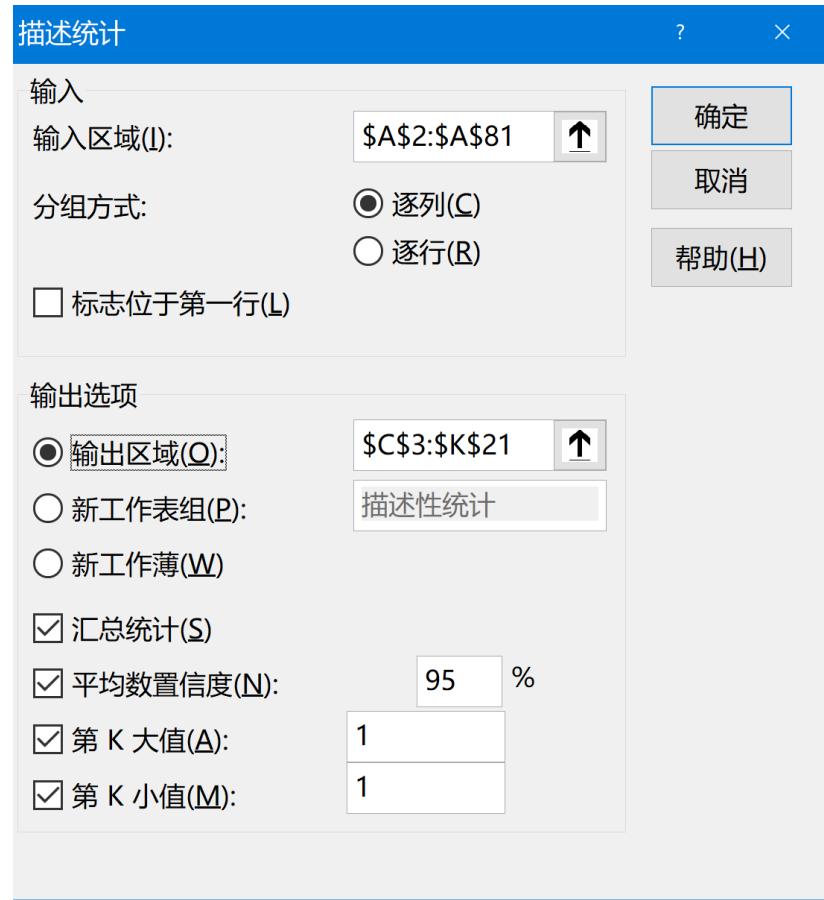
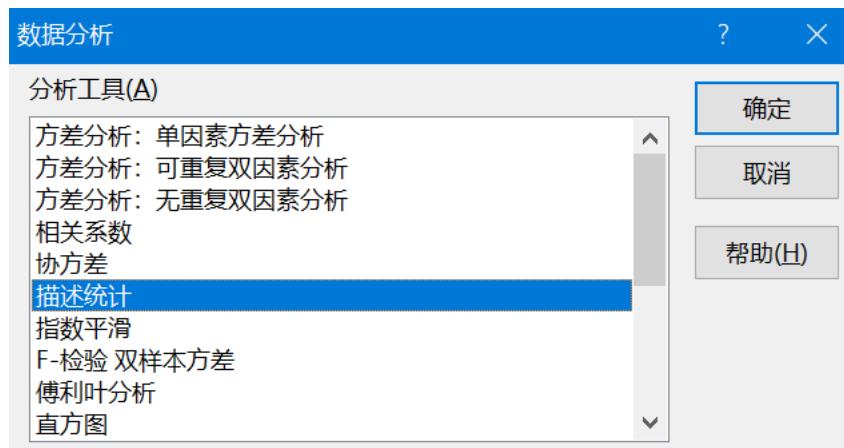
### Excel函数

数据类型	Excel函数	
平均数	average	
中位数	median	
样本方差(variance)	var.s	总体方差： var.p
样本标准差(standard deviation)	stdev.s	总体标准差： stdev.p

# 描述性统计

## 数值型变量

### 数据分析工具包的【描述统计】

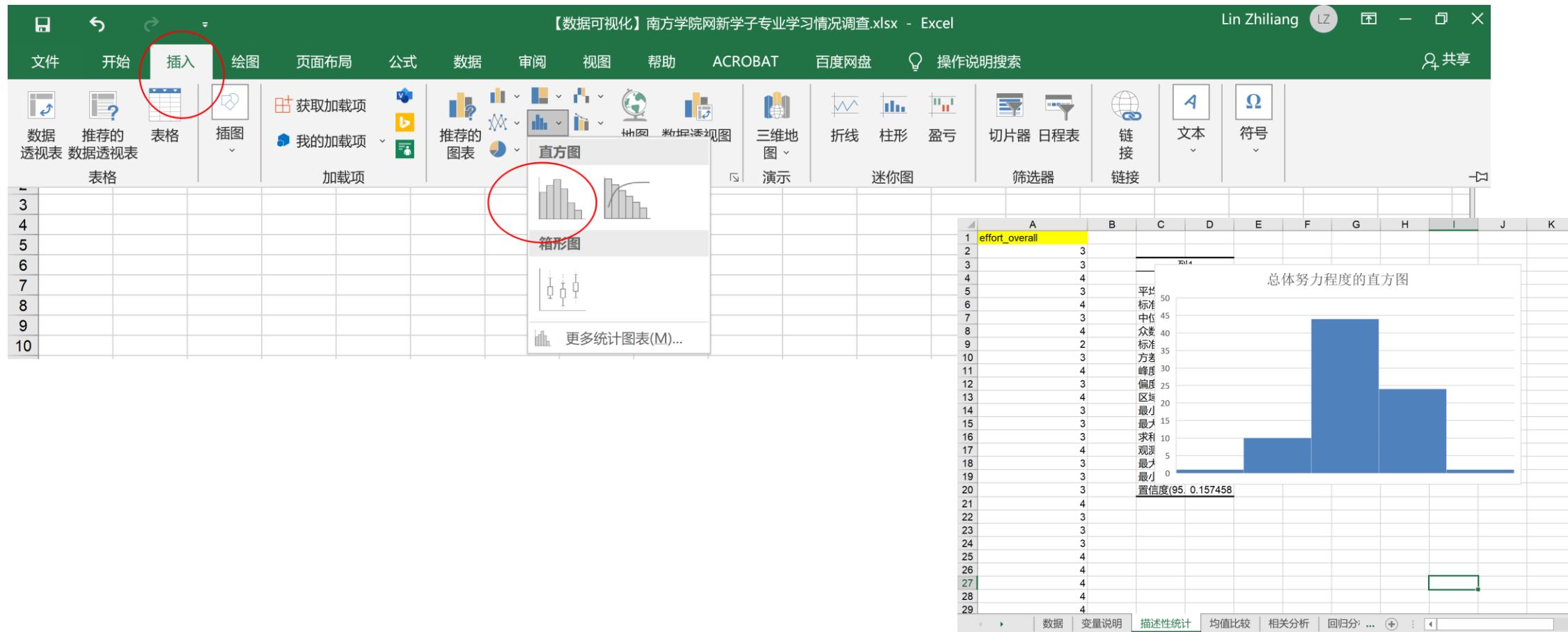


A	B	C	D
effort_overall			
3			
3			列1
4			
3		平均	3.175
4		标准误差	0.079107
3		中位数	3
4		众数	3
2		标准差	0.707554
3		方差	0.500633
4		峰度	0.353272
3		偏度	-0.26361
4		区域	4
3		最小值	1
3		最大值	5
3		求和	254
4		观测数	80
3		最大(1)	5
3		最小(1)	1
3		置信度(95%)	0.157458
4			
3			
3			
4			
4			

# 描述性统计

## 数值型变量

- 选中数据，点击【插入】，选择【直方图】



# 推断性统计

## 双变量分析梗概

- 分类型变量 + 数值型变量

- 二分类型变量 + 数值型变量



独立样本t 检验

- 多分类型变量 + 数值型变量



单因素方差分析

差异

---

- 分类型变量 + 分类型变量



卡方检验

关联

- 数值型变量 + 数值型变量



相关分析

# 推断性统计

## 双变量分析梗概

- 分类型变量 + 数值型变量

- 二分类型变量 + 数值型变量



独立样本t 检验

- 多分类型变量 + 数值型变量



单因素方差分析

箱型图

---

- 分类型变量 + 分类型变量



卡方检验

簇状条形图/堆叠条形图

- 数值型变量 + 数值型变量



相关分析

散点图

## 双变量分析梗概

- 独立样本t检验、单因素方差分析和卡方检验我们分别汇报检验统计量 ( $t$ 值/ $F$ 值/  $\chi^2$ 值) 和p值，根据p是否小于0.05判断显著性；相关分析我们汇报相关系数和p值，根据p是否小于0.05判断显著性。

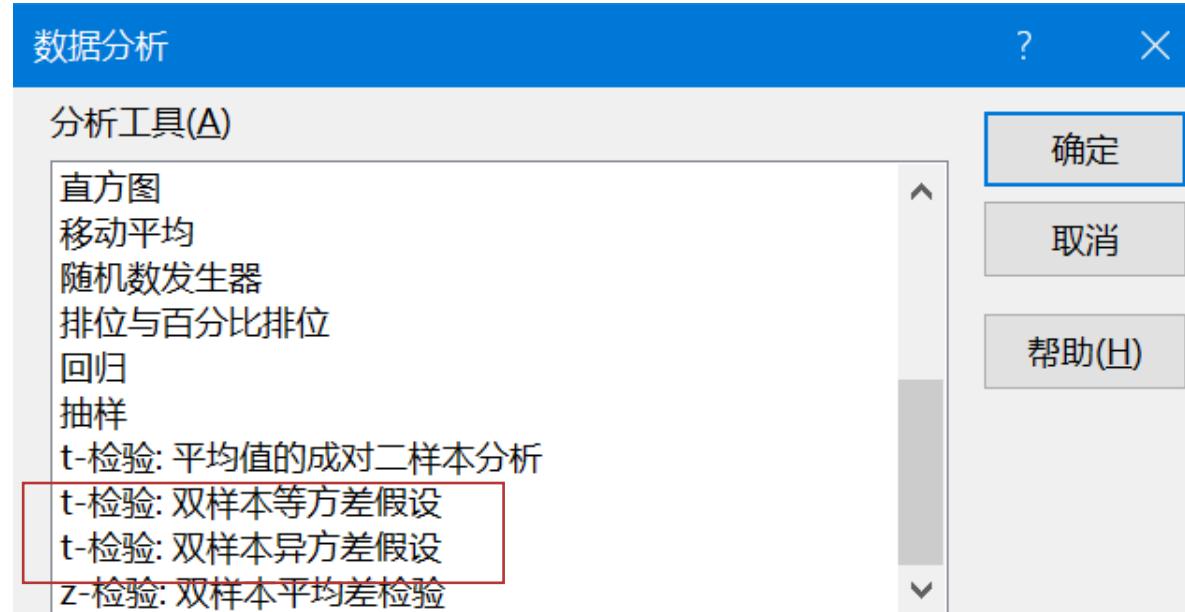
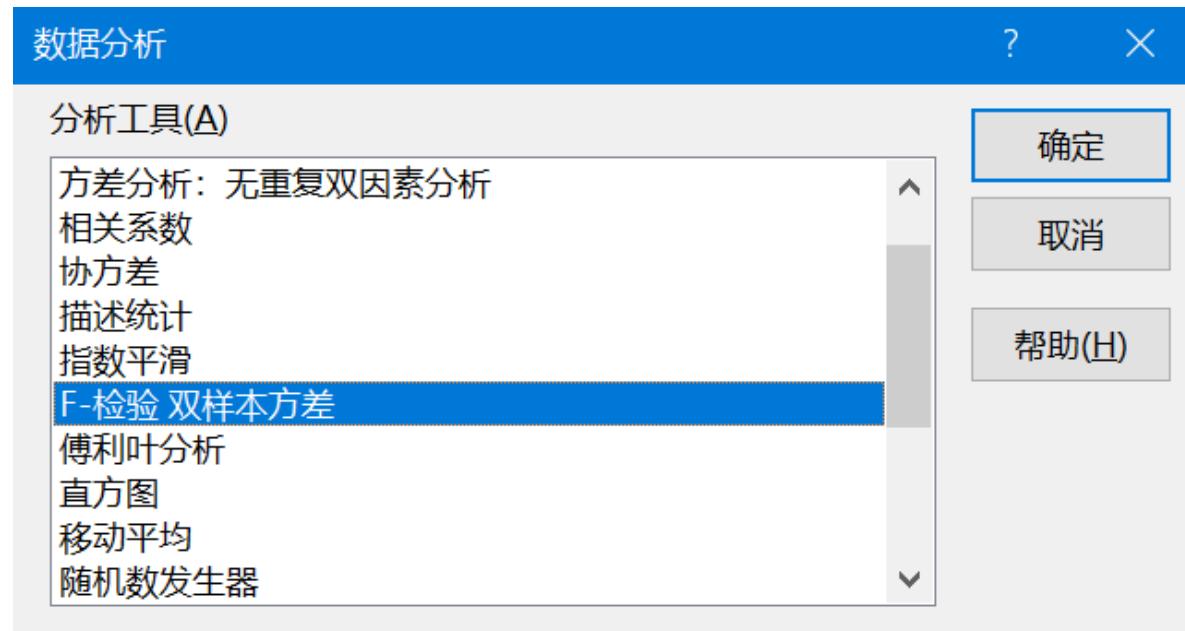
# 独立样本t检验

【工具】>【数据分析】>【F-检验 双样本方差分析】>选择数据范围



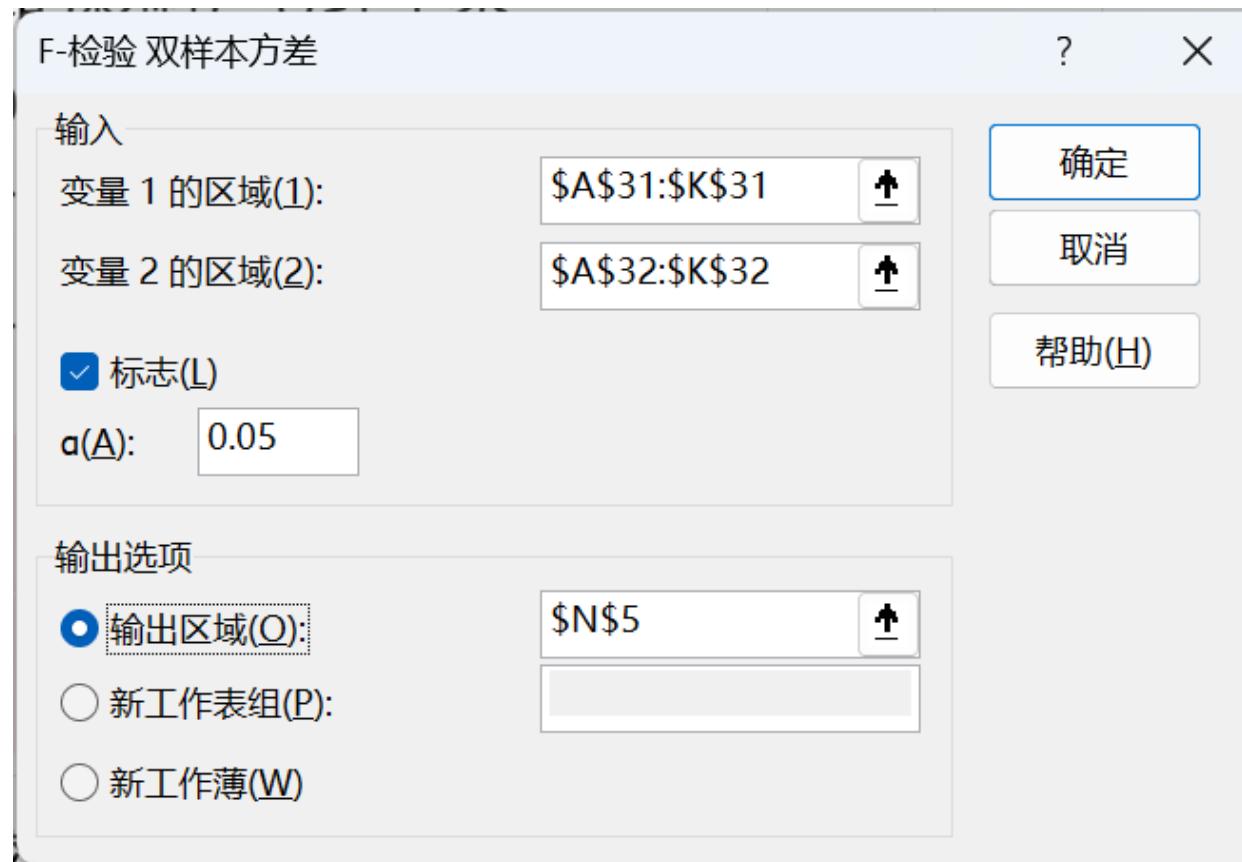
如果F检验的结果是 $p < 0.05$ , 则选择异方差; 如果 $p > 0.05$ , 选择同方差

【工具】>【数据分析】>【t检验：双样本等（异）方差假设】>选择数据范围



# 独立样本t检验

## F检验：双样本方差检验



# 独立样本t检验

## F检验

F检验

F-检验 双样本方差分析		
	变量 1	变量 2
平均	26	18
方差	22.22222	17.77778
观测值	10	10
df	9	9
F	1.25	
P(F<=f) 单尾	0.372501	
F 单尾临界	3.178893	

方差齐性检验，如果  
 $p < 0.05$ ，使用【t检验：  
双样本异方差假设】；  
如果 $p > 0.05$ ，则使用  
【t检验：双样本等方差  
假设】

# 独立样本t检验

## t检验：双样本等方差假设

t-检验: 双样本等方差假设

	变量 1	变量 2
平均	26	18
方差	22.22222	17.77778
观测值	10	10
合并方差	20	
假设平均差	0	
df	18	
t Stat	4	
P(T<=t) 单尾	0.00042	
t 单尾临界	1.734064	
P(T<=t) 双尾	0.00084	
t 双尾临界	2.100922	

我们一般看双尾的p值

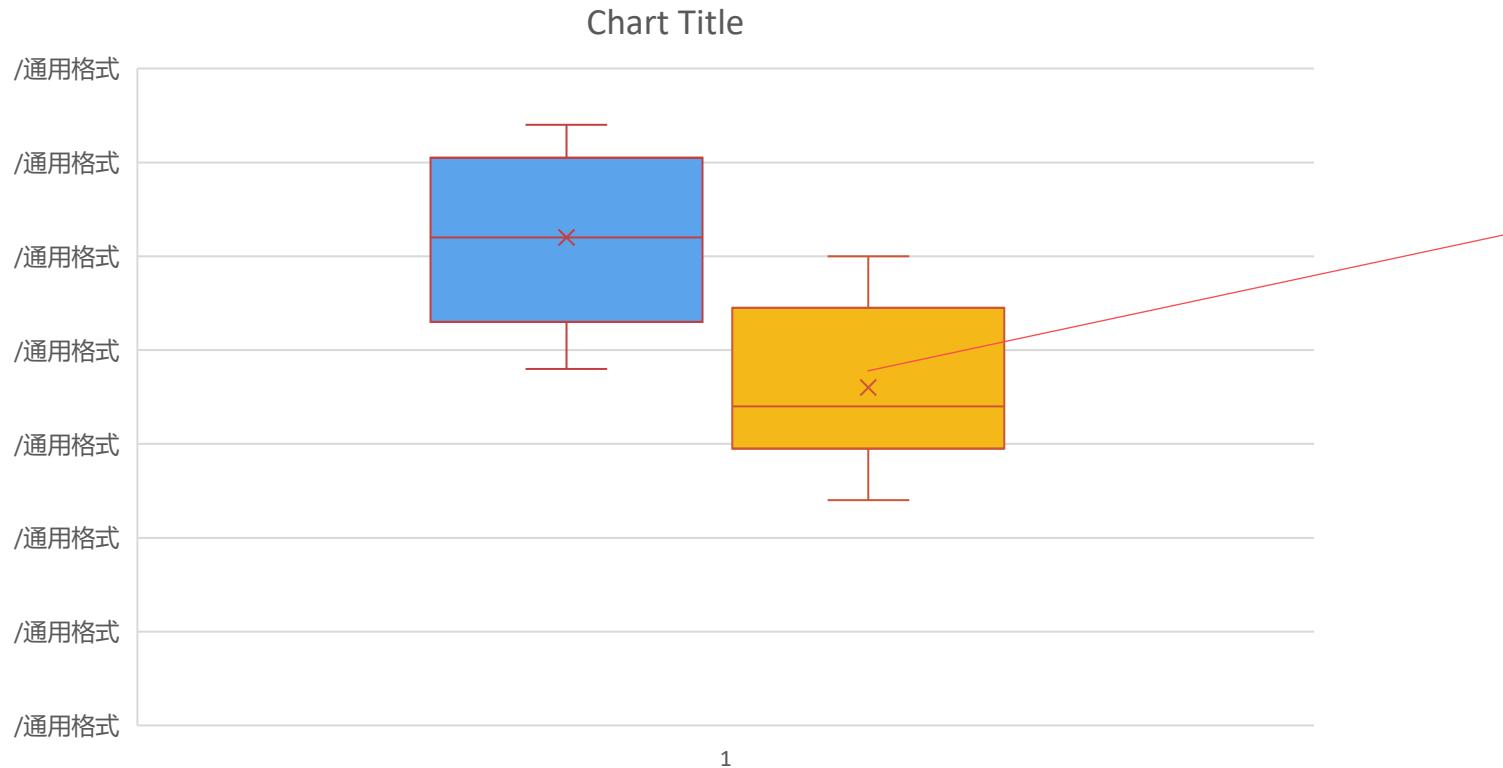
# 独立样本t检验

## 箱型图的解读



# 独立样本t检验

## 箱型图



注：这个符号代表的是平均值  
(其它软件画箱线图一般没有)

# 单因素方差分析



【工具】>【数据分析】>【方差分析：单因素方差分析】>选择数据范围



# 单因素方差分析

A	B	C	D	E	F	G	H	I	J	K	L	M
254	234	200			0.0000498523504603869							
263	218	222										
241	235	197										
237	227	206										
251	216	204										

方差分析: 单因素方差分析

SUMMARY

	组	观测数	求和	平均	方差
列 1		5	1246	249.2	108.2
列 2		5	1130	226	77.5
列 3		5	1029	205.8	94.2

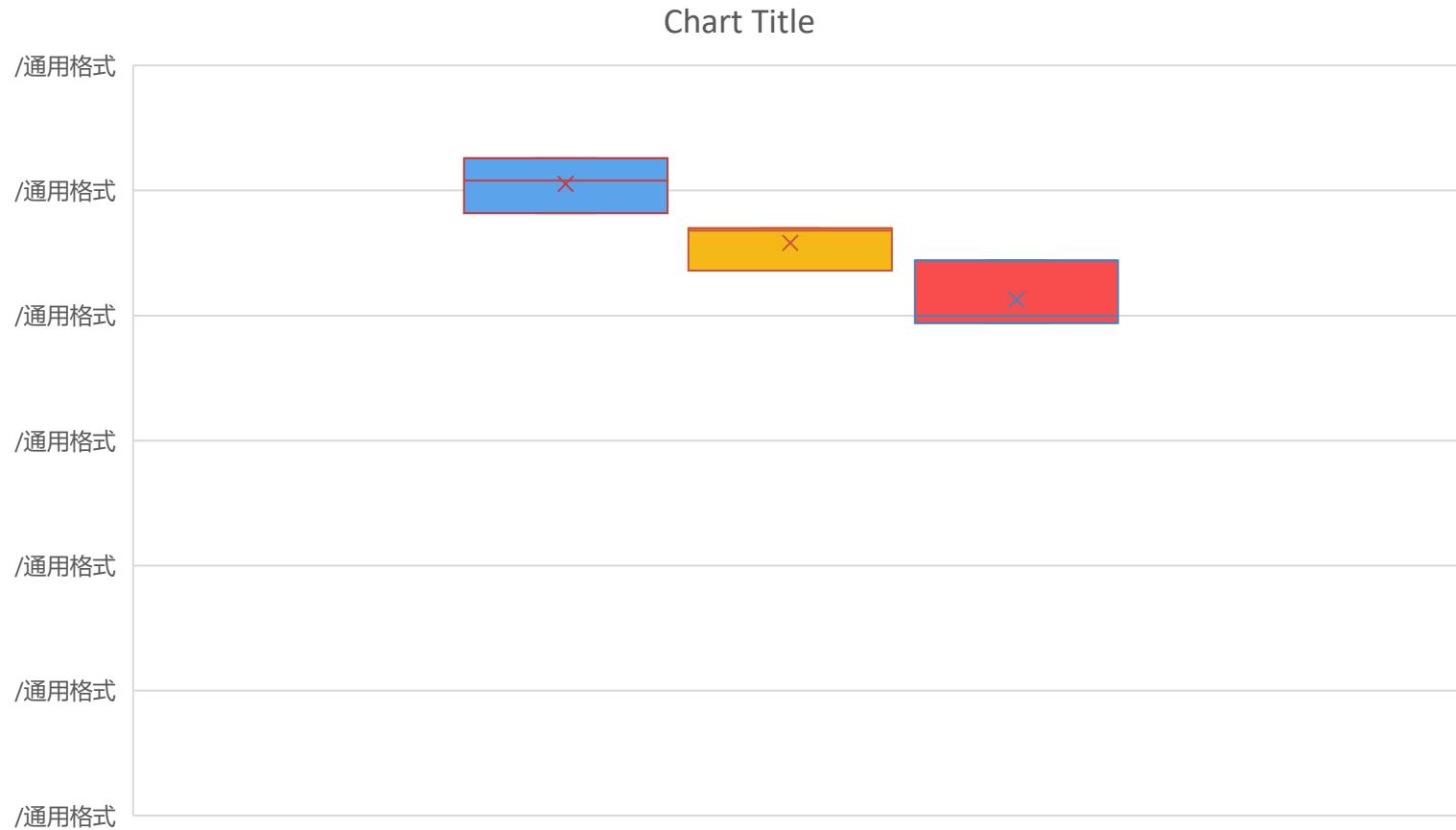
这是科学记数法

方差分析

差异源	SS	df	MS	F	P-value	F crit
组间	4716.4	2	2358.2	25.27546	4.98524E-05	3.885293835
组内	1119.6	12	93.3			
总计	5836	14				

# 单因素方差分析

## 箱型图



# 卡方检验

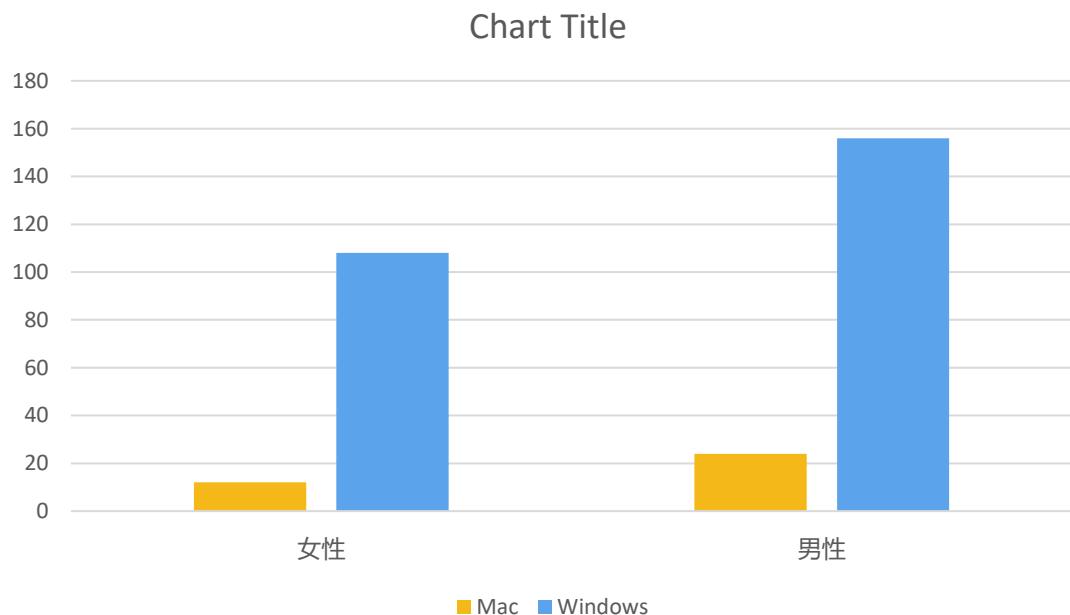
如果使用Excel做卡方检验，需要手动整理观察频数表和期望频数表，然后使用函数**CHISQ.TEST**计算卡方检验的p值。

The screenshot shows a Microsoft Excel spreadsheet with two data tables. The first table, labeled '观察频数表' (Observation Frequency Table), contains observed counts for Mac and Windows users across gender categories. The second table, labeled '期望频数表' (Expected Frequency Table), contains expected counts for Mac and Windows users. The formula bar shows the function =CHISQ.TEST(B2:C3,F2:G3) entered into the formula bar.

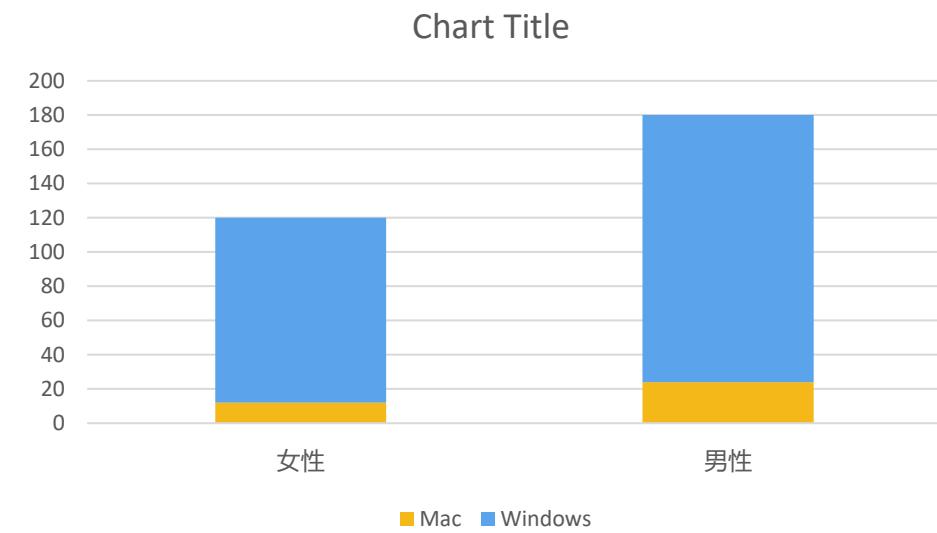
	A	B	C	D	E	F	G	H	I
	Mac	Windows			Mac	Windows			
女性		12	108		女性	14.4	105.6		0.384088
男性		24	156		男性	21.6	158.4		
观察频数表					期望频数表				

# 卡方检验

## 簇状条形图



## 堆叠条形图



# 相关分析

## 计算相关系数

注意：Excel没有便捷的方法对相关分析进行显著性检验

- 方法一：使用函数**CORREL**(每次只能分析两个变量之间的相关关系)

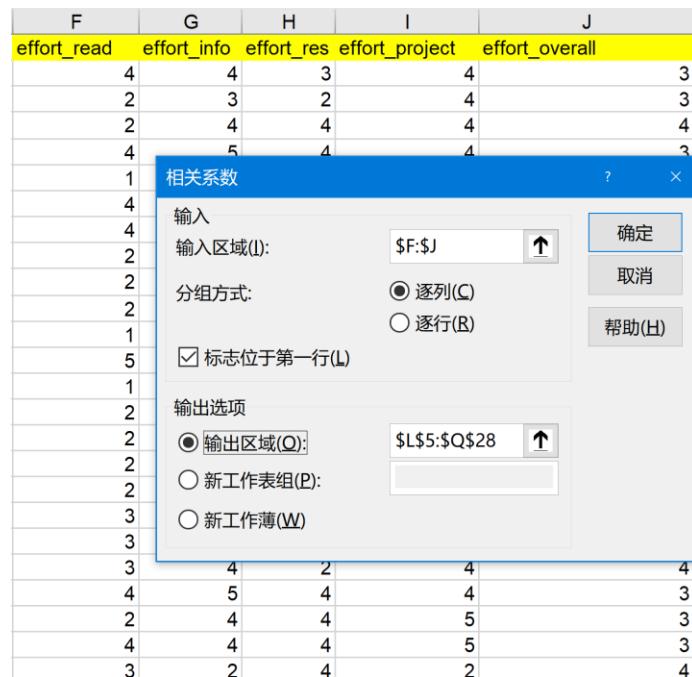
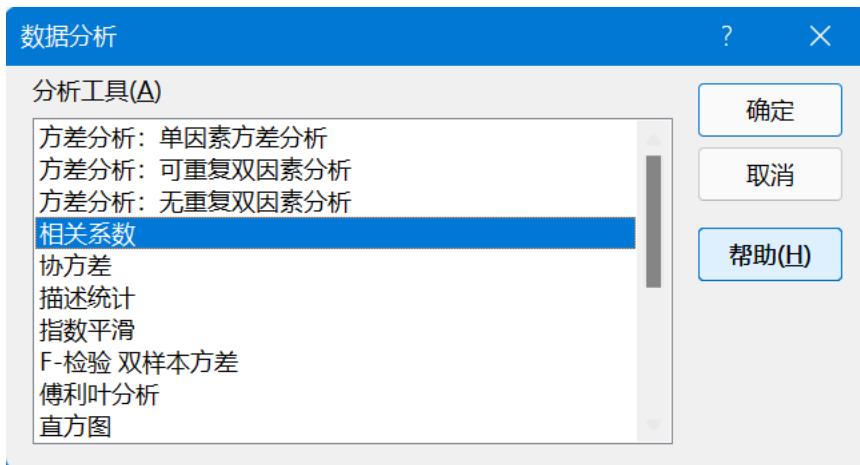
	A	B	C	D	E	F	G
1	ranking	effort_overall					
2	3	3		0.28433			
3	2	3					
4	4	4					
5	3	3					
6	4	4					
7	3	3					
8	3	4					
9	3	2					
10	4	3					
11	5	4					
12	3	3					
13	4	4					
14	3	3					
15	3	3					
16	3	3					
17	2	4					
18	3	3					
19	3	3					
20	3	3					
21	3	4					
22	4	3					
23	3	3					
24	4	3					
25	2	4					
26	3	4					
27	5	4					

# 相关分析

## 计算相关系数

- 方法二：使用【分析工具】中的【相关系数】(可同时分析多个变量两两之间的相关系数)

这个表格叫：**相关系数表**



	effort_read	effort_info	effort_res	effort_project	effort_overall
effort_read	1				
effort_info	0.467343532	1			
effort_res	0.418507689	0.642064177	1		
effort_project	0.247580107	0.46872072	0.372938049	1	
effort_overall	0.239919349	0.252594288	0.324248546	0.32795651	1

# 相关分析

## 散点图

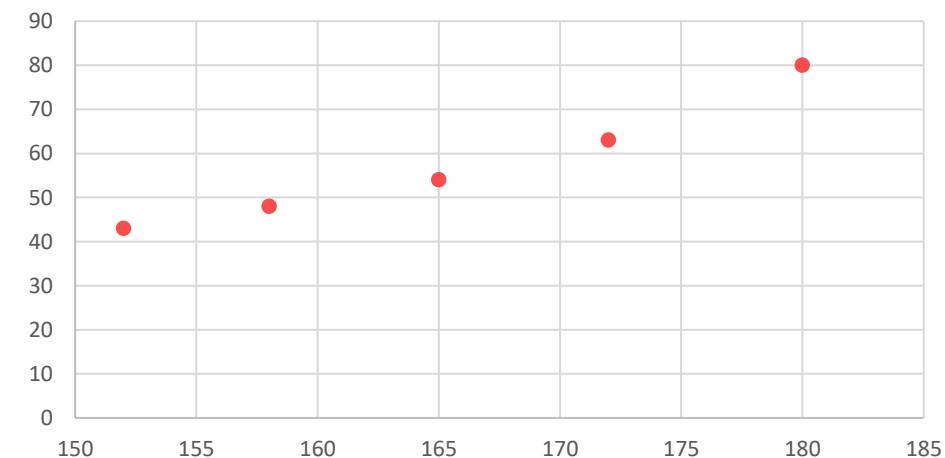
选中数据，点击【插入】，选择散点图的图形

文件 开始 插入 绘图 页面布局 公式 数据 审阅 视图 帮助 操作说明

数据透视表 数据透视表 表格 插图  
推荐的图表  
我的加载项  
获取加载项  
加载项  
推荐的图表  
散点图  
气泡图  
更多散点图(M)...

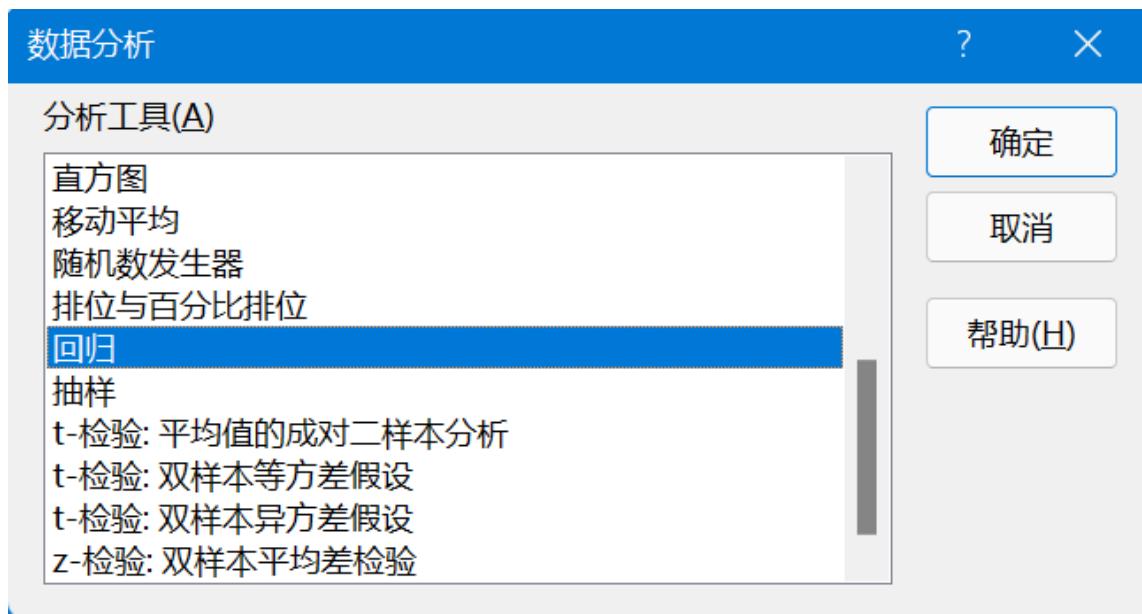
	152	158	165	172	180
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14	152	158	165	172	180
15	43	48	54	63	80
16					
17					
18					

身高与体重的散点图



# 回归分析

## Excel操作



可勾选此项



# 回归分析

## Excel结果

回归方程：

$$\widehat{\text{体重}} = -154.81 + 1.28 \times (\text{身高})$$

SUMMARY OUTPUT								
回归统计								
Multiple R	0.976691735							
R Square	0.953926746							
Adjusted R Square	0.938568995							
标准误差	3.611343571							
观测值	5							
方差分析								
	df	SS	MS	F	Significance F			
回归分析	1	810.0746	810.0746	62.1137	0.00425671			
残差	3	39.12541	13.0418					
总计	4	849.2						
	Coefficients	标准误差	t Stat	P-value	Lower 95%	Upper 95%	下限 95.0%	上限 95.0%
Intercept	-154.8070033	26.09936	-5.73373	0.010533	-240.73103	-68.882981	-240.73103	-68.882981
X Variable 1	1.284201954	0.162944	7.881224	0.004257	0.76563992	1.80276398	0.76563992	1.80276398
RESIDUAL OUTPUT								
观测值	预测 Y	残差						
1	40.39169381	2.608306						
2	48.09690554	-0.09691						
3	57.08631922	-3.08632						
4	66.0757329	-3.07573						
5	76.34934853	3.650651						

## 参考资料

- Lizongzhang的个人空间-合集 · Excel 数据分析实战

谢谢！