

福建农林大学 经济与管理学院
旅游学

实验报告

课程名称: 统 计 学

专业班级: 2010 金融

学 号: _____

学生姓名: _____

指导教师: 张 玉 容

成 绩: _____

2012 年 6 月 6 日

实验一：EXCEL 的数据整理与显示

一、实验目的及要求：

(一)目的

1. 了解 EXCEL 的基本命令与操作、熟悉 EXCEL 数据输入、输出与编辑方法；
2. 熟悉 EXCEL 用于预处理的基本菜单操作及命令；
3. 熟悉 EXCEL 用于整理与显示的基本菜单操作及命令。

(二)内容及要求

根据学生实验数据 2011-2012,

1. 用 Excel 制作一张学生生源的频数分布表，并绘制一张柱状图，反映不同生源的学生人数分布情况。
2. 对学生的体重进行等距分组，整理成频数分布表，并绘制直方图。
3. 编制生源与专业交叉分布的列联表，并画出雷达图，比较两个专业的学生生源的分布是否相似。

二、仪器用具

硬件：计算机（安装 Windows98 、Windows2000 或 Windows XP 或以上）

软件：EXCEL

三、实验原理

统计中数据整理与显示的相关理论。

四、实验方法与步骤

1、方法：应用数据透视表和数据透视图编制学生生源的频数分布表。

步骤：“数据”——“数据透视表和数据透视图”——“数据源类型”——“下一步”——选定数据源区域——“下一步”——“布局”——将“生源”拖至左边的行区域和数据区域——“确定”——“完成”——复制数据透视表——“选择性粘贴”——选“数值”——“确定”

2、方法：应用“frequency”函数编制学生体重的频数分布表。

步骤：（1）按等距分组分成 5 组，组距为 10——选定数据接收区域——“插入”——“函数”——“frequency”——“确定”——在 Data_array 中选定体重列的数据，在 Bins_array

选中组距上限列的数据——按“Ctrl+Shift”，同时按下“Enter”——计算出频率把表格补充完整

(2) 选定如下数据（上下各多选中一行）：

40-50	58
50-60	70
60-70	40
70-80	11
80-90	1

——“插入”——“图表”——选择柱形图——“下一步”——“确定”——输入图表标题和 Y 轴、X 轴的名称——“下一步”——“完成”——把光标放在图表的柱上右击选择“数据系列格式”——“选项”——把“分类间距”改为 0

3、方法：应用数据透视表和数据透视图编制生源与专业交叉分布的列联表。

步骤：(1) “数据”——“数据透视表和数据透视图”——“数据源类型”——“下一步”——选定数据源区域——“下一步”——“布局”——将“生源”拖至左边的列区域，“专业”拖至行区域，“生源”拖至数据区域——“确定”——“完成”——复制数据透视表——“选择性粘贴”——选“数值”——“确定”

(2) 选中表格——“插入”——“图表”——“雷达图”

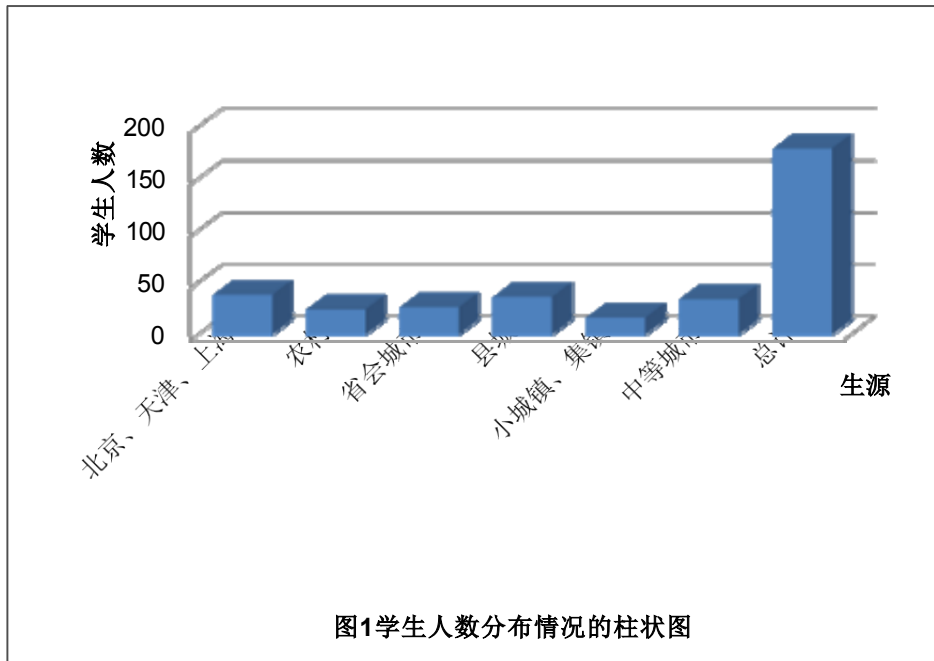
五、实验结果与数据处理

1、

表 1 学生生源的频数分布表

生源	学生人数
北京、天津、上海	39
农村	25
省会城市	27
县城	37

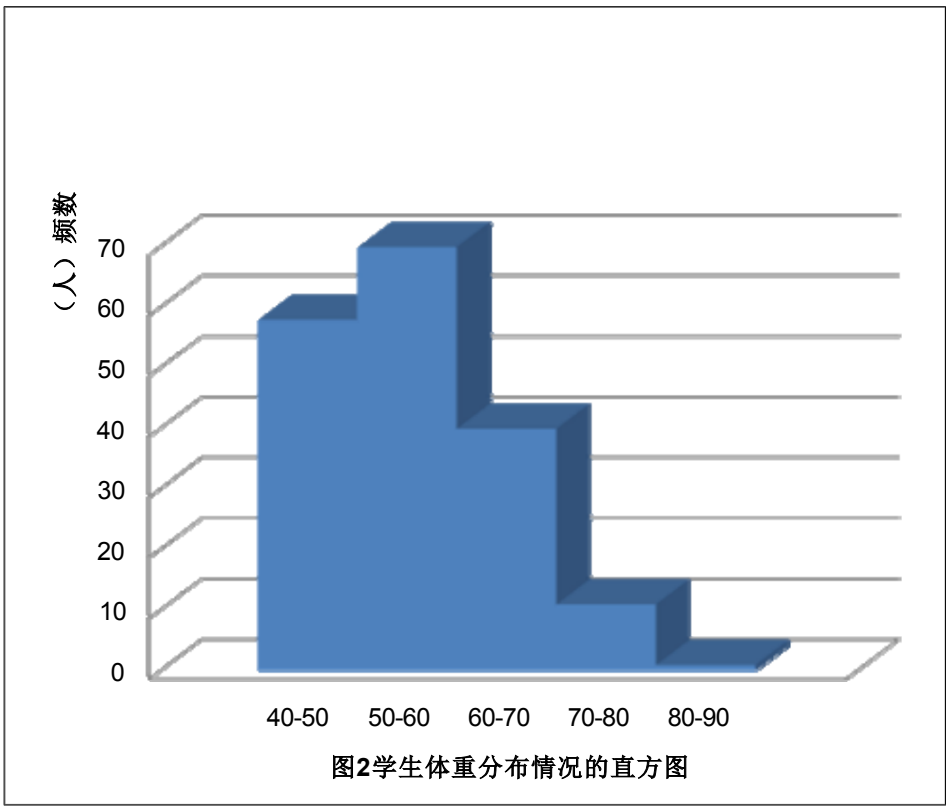
小城镇、集镇	17
中等城市	35
<hr/>	
总计	180
<hr/>	



2、

表 2 学生体重频数分布表

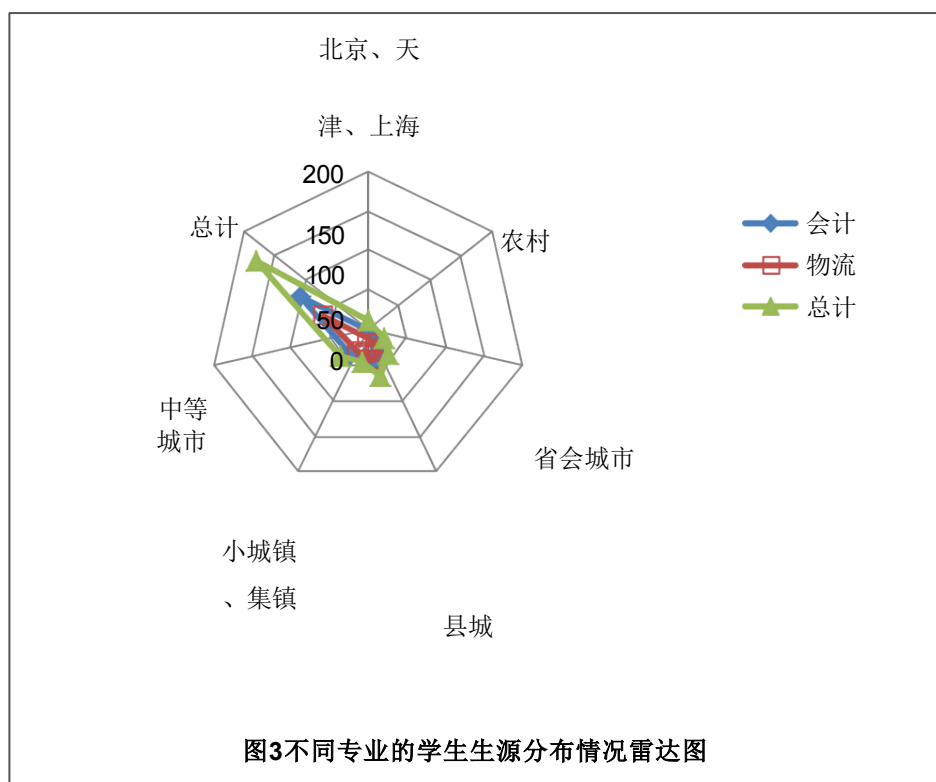
按体重分组	频数	频率 (%)
40-50	58	32.22
50-60	70	38.89
60-70	40	22.22
70-80	11	6.11
80-90	1	0.56
合计	180	100



3、

表 3 不同专业的学生生源的分布表

生源专业	北京、天津、上海	农村	省会城市	县城	小城镇、集镇	中等城市	总计
会计	25	13	14	23	12	23	110
物流	14	12	13	14	5	12	70
总计	39	25	27	37	17	35	180



由上图可知，会计和物流的学生生源的分布相似。

六、讨论与结论

- 1、在制作直方图的时候，遇到了比较多的问题，后来请教同学后就解决了。
- 2、对图表的制作比较生疏。

实验报告评分表

学生姓名	林雪冰	学号	102263004030	专业 年级	2010 金融		
实验项目 名称	实验一：EXCEL 的数据整理与显示			实验 学时	3 学时		
评价项目	权重	评价内容	评价结果				得分
			A	B	C	D	
实验态度	20%	实验态度端正，遵守实验室守则，严格按照实验要求进行操作。	20	16	14	12	
实验过程	30%	实验项目符合大纲，实验方法科学；步骤操作合理，逻辑条理清晰，符合指导书要求。	30	24	21	18	
实验结论 与讨论	30%	实验结论正确，分析、讨论深入。	30	24	21	18	
实验报告 描述	20%	语言精炼、流畅、准确、灵活，逻辑性强；结构严谨规范，条理清晰，布局合理，系统严密。	20	16	14	12	
总分							
教师签名							

实验二：EXCEL 的数据特征描述、列联分析、多元回归分析

一、实验目的及要求：

(一)目的

熟悉 EXCEL 用于数据描述统计、列联分析、多元回归的基本菜单操作及命令。

(二)内容及要求

根据学生实验数据 2011-2012，

1. 对学生的身高进行描述统计，说明学生身高的一些基本特征。
2. 对不同户口中贫困生的分布情况编制列联表，并分析贫困生的数量的多少是否与户口的不同是有关。
3. 根据学生的体育成绩(y)与学生的年龄(x1)、体重(x2)、身高(x3) 和性别 (x4) 建立一个多元线性回归模型，并判断此模型对于解释学生体育成绩差异有无意义。

二、仪器用具

硬件：计算机（安装 Windows98 、 Windows2000 或 Windows XP 或以上）

软件：EXCEL

三、实验原理

统计中数据整理与显示的相关理论。

四、实验方法与步骤

- 1、方法：应用数据分析中的描述统计法对学生的身高进行描述统计。

步骤：“工具”——“数据分析”——“描述统计”——输入区域：“身高”列——选中“标志位于第一行”——“确定”

学生身高的一些基本特征：平均身高为 164.77cm，身高为 155cm 的学生人数最多，学生身高分布不均匀，全体学生的身高趋势呈稍微扁平 and 稍微右偏。

- 2、方法：应用数据透视表和数据透视图对不同户口中贫困生的分布情况编制列联表。

步骤：“数据”——“数据透视表和数据透视图”——“下一步”——选定数据源区域：全选——“下一步”——“布局”——将“户口”拖至行区域，“是否贫困生”拖至列区域，“是否贫困生”拖至数据区域——“完成”；复制列联表的数据求出观察值和 P 值。

3、方法：应用回归分析根据学生的体育成绩(y)与学生的年龄(x1)、体重(x2)、身高(x3) 和性别 (x4) 建立一个多元线性回归模型。

步骤：全选——“数据”——“筛选”——“自动筛选”——将女生的体育成绩全部除以 0.8，男生的体育成绩全部除以 1——复制因变量：体育成绩，自变量：年龄(x1)、体重(x2)、身高(x3) 和性别 (x4) ——用自动筛选将自变量：“性别”改为虚拟变量，男生=0，女生=1——“数据分析”——“回归”——Y 值区域：“体育成绩”列，X 值区域：“年龄”、“体重”、“身高”和“性别”4 列——“确定”

五、实验结果与数据处理

1、表 1 学生身高的描述统计

身高	
平均	164.772222
标准误差	0.59512976
中位数	163.5
众数	155
标准差	7.98450354
方差	63.7522967
峰度	-0.6969825
偏度	0.24812533
区域	36
最小值	150
最大值	186
求和	29659
观测数	180
最大(1)	186
最小(1)	150
置信度(95.0%)	1.17437276

由上表可知学生身高的一些基本特征：全体学生的身高趋势呈稍微扁平 and 稍微右偏分布，众数 $<$ 中位数 $<$ 平均数；学生的平均身高为 164.77cm，身高为 155cm 的学生人数最多，

中位数为 163.5cm；另外，学生身高的方差和标准差都比较大，说明学生中身高差距大，分布比较不平衡。

2、

表 2 不同户口中贫困生的分布情况表

是否贫困生 户口	是否贫困生		总计
	否	是	
城市户口	122	14	136
农村户口	36	8	44
总计	158	22	180

表 3 期望值表

是否贫困生 户口	是否贫困生	
	否	是
城市户口	119.38	16.62
农村户口	38.62	5.38

表 4 求 P 值

观察值	期望值	求 P 值
122.00	119.38	0.59
36.00	38.62	
14.00	16.62	
8.00	5.38	

结论：因为求出的相关性为：0.59，即 $P=0.59 > \alpha=0.05$ ，不能拒绝原假设，所以可以判断出贫困生的数量的多少与户口的不同是无关的。

3、表 5 多元线性回归模型

SUMMARY OUTPUT

Multiple R	0.7969335
R Square	0.6351031
Adjusted R Square	0.6267626
标准误差	25.019962

以上内容仅为本文档的试下载部分，为可阅读
页数的一半内容。如要下载或阅读全文，请访
问：

<https://d.book118.com/198101052126007004>