## 6. 第一次作业-stata部分

**黑人的故事**

**背景：**

假设在样本数据Training.dta中，445个观测值既有黑人(black==1)，也有非黑人(black==0)，他们中的一部分人在1975年被随机抽取并接受长期培训(train==1)。所有样本在1974年(培训前)和1978年(培训后3年)的收入分别记为向量re74、re78.

**问题：**

1.生成变量edul表示教育水平，当受教育年限（educ）为1~6年时，edul取0；当受教育年限为07~12年时，edul取1；当受教育年限为12年以上时，edul取 2。做**两张**饼图，分别展示出**黑人、非黑人（black==0）**的教育水平分布情况。

2.利用 ttest 检验：**对于全体样本**，在1974年(培训前)，黑人和非黑人的收入均值有差异吗？1978年呢(培训后3年)？利用estpost和esttab将问题2的t检验结果输出为txt格式的表格。

3.利用 ttest 检验: 在1978年(培训后3年)，**对于接受培训的样本(train==1)**,黑人和非黑人的收入均值有差异吗？对于**没有接受培训的样本(train==0)**呢？

4. 问题2和3的对比说明了什么问题？

5. 自行help label并阅读帮助文档。在变量窗口，我们看到变量re78对应的label是“real earns., 1978, $1000s”，请将这个变量标签修改为中文“1978年的实际收入（千美元）”。

**要求【重要，影响分数】：**

1.10月8日前，将do文件和log文件发送到邮箱：jl2017f@126.com (计量2017fall)

2.do文件和log文件均以“学号\_姓名\_专业”命名，如 b151090001\_户纳东\_经济.do

3.此次作业两个文件【不】要打包发送

**4.【重要】邮件“主题”请务必填写为“ 10月8日第一次作业 ”**

10月8日0:00之后邮件不会正确归档，请大家不要延误提交，错误的文件命名、主题或打包会妨碍批改，也会酌情扣分。