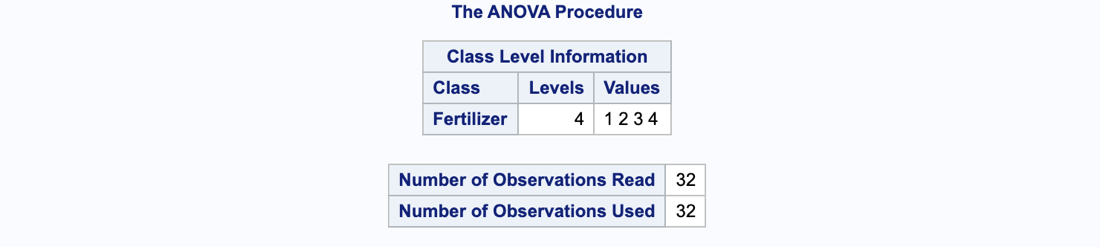
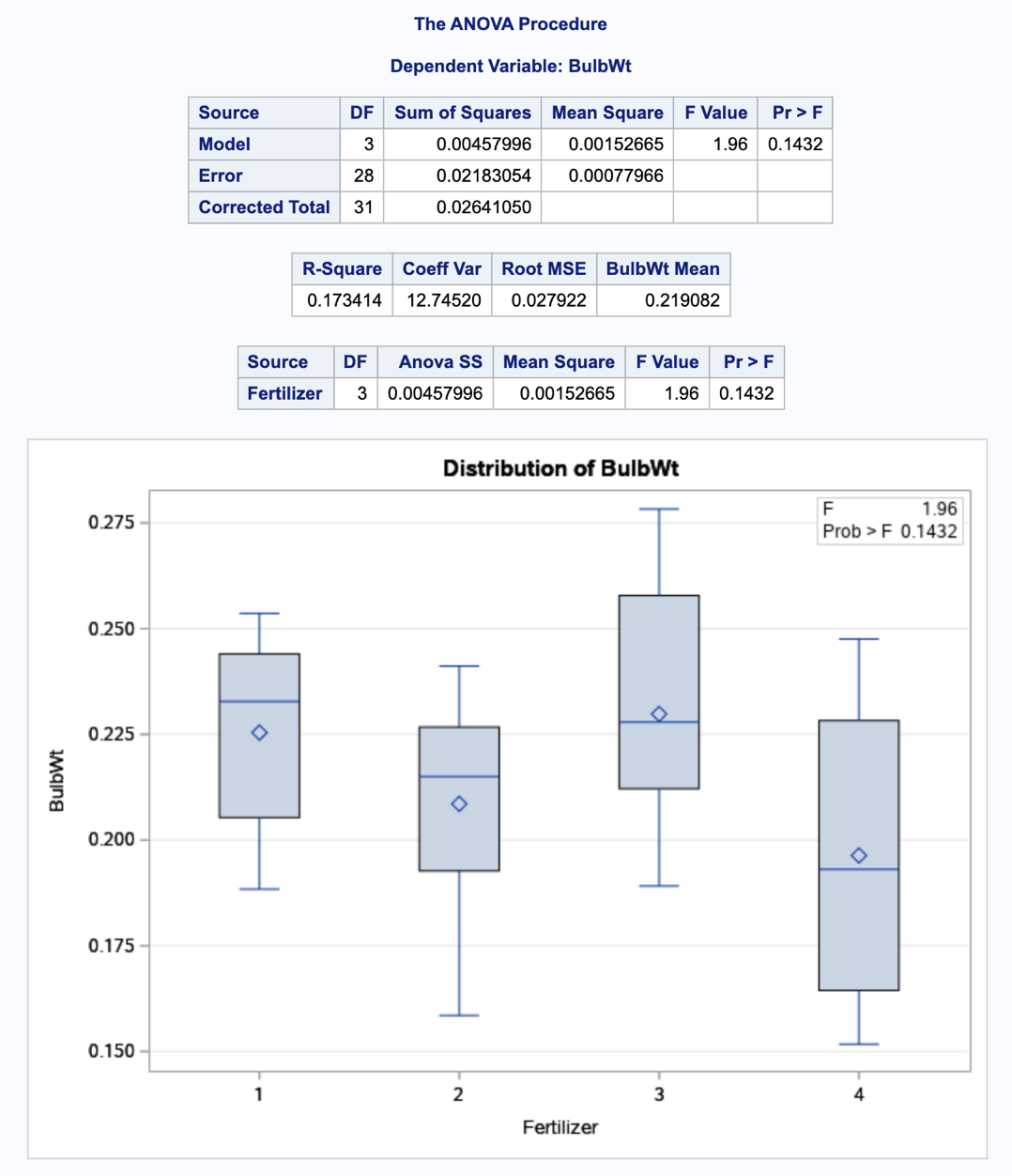
题目一

对MGGarlic数据集，以BulbWt作为响应变量（因变量），Fertilizer为预测变量（自变量），使用PROC ANOVA做单因素方差分析，并描述得到的结果。

【答】使用单因素方差分析源代码见附件./solution1.sas，分析结果如下：





分析结果说明：

* 过程输出的汇总信息部分：
  + 自变量因素Fertilizer水平为4，总观测数为32
* 方差分析部分：
  + 模型自由度为3，校正合计自由度为31，误差自由度为28
  + F值为1.96不具有显著性
  + p值为0.1432接受原假设
  + 我们可以认为，Fertilizer作用在BulbWt上不具有显著性差异

题目二

现有一个试验：测试三种不同品牌的混凝土，看添加剂(Additive)是否能使混凝土中的水泥更坚固。试验人员浇注了30个试验区，并在Concrete数据集中记录了以下数据：

Strength 混凝土强度

Additive 试验区是否使用了添加剂

Brand 接受测试的混凝土品牌

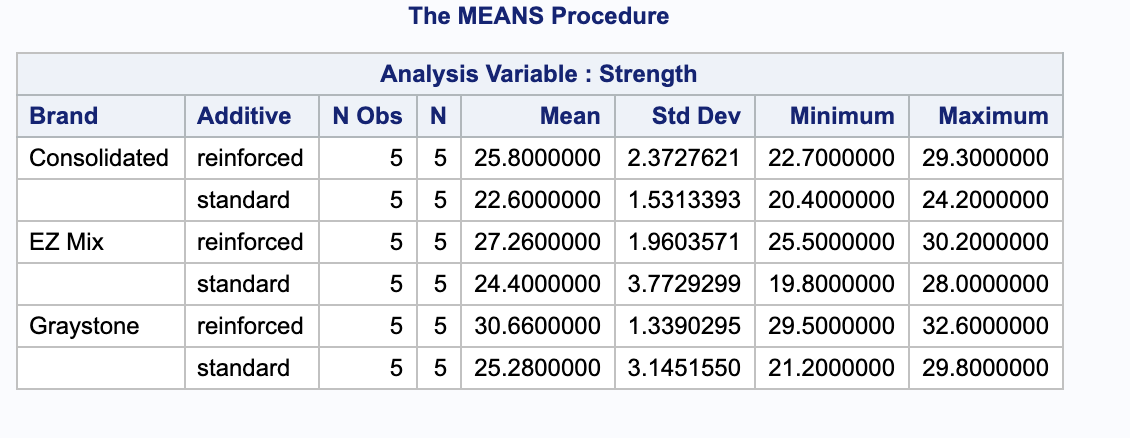
* 使用Strength作响应变量， Additive和Brand作预测变量检查数据。 你可以从中获得哪些信息？（提示：使用PROC MEANS）
* 使用方差分析的方法检验均值是否相等。到目前为止，你能得出什么样的结论？（提示：使用PROC GLM）
* 是否有必要做多重比较？如必要，请做多重比较；如非必要，请给出理由。

【答】

（1）源代码见附件./solution2.sas，运行程序可以得出如下结论：

* 由分析结果截图可以得出按照Additive和Brand分类分别得出的Strength的均值、方差、最大值和最小值
* Graystone 和reinforced的组合具有最大的Strength均值且表现效果稳定（方差小）赢推荐使用这种组合
* 添加剂的加入（从standered变为reinforced）可以使Strength增强，具有使水泥更坚固的效果

程序运行结果如下：

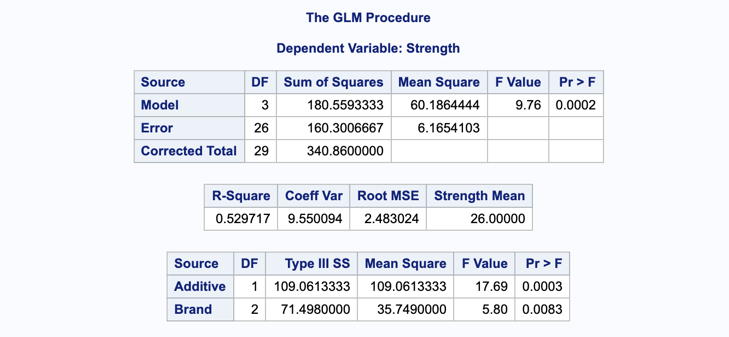


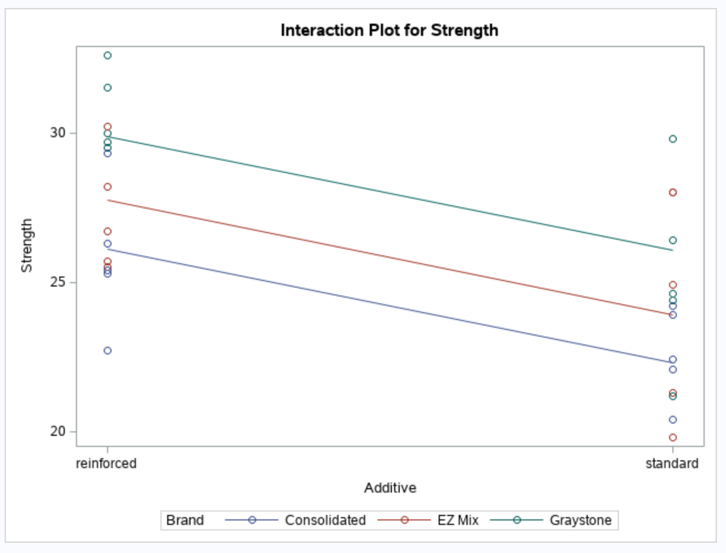
（2）～（3）

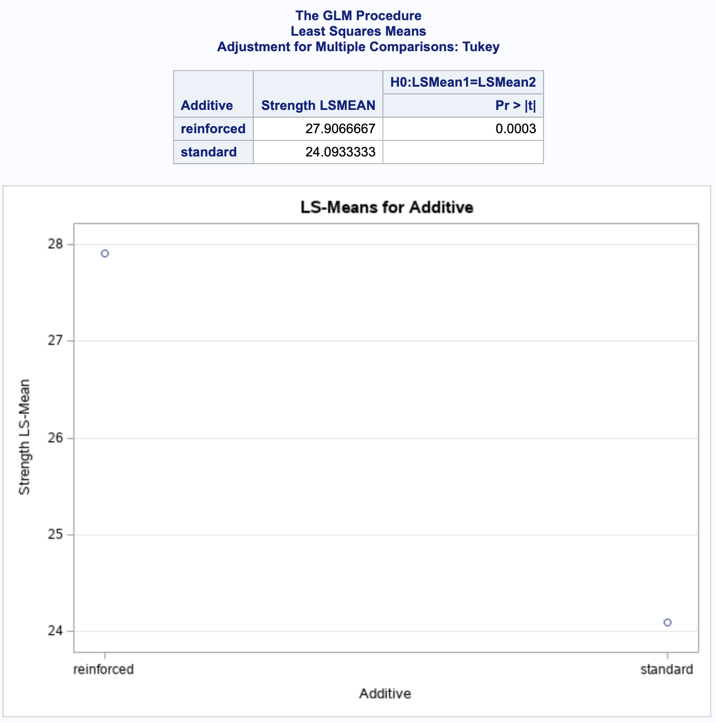
有必要做重复性实验，当在多个不同条件下测量随机样品的所有成员时使用重复测量分析。 当样品存在多种不同条件时必须依次进行测量。 在这种情况下使用标准ANOVA是不合适的，因为它不能建模重复测量不同条件之间的相关性。 应该清楚重复测量多个不同条件之间的差异。 对样本进行重复测量试验,每个试验测试了在不同条件下相同样品的特点。SAS提供了 PROC GLM用于进行重复测量分析。

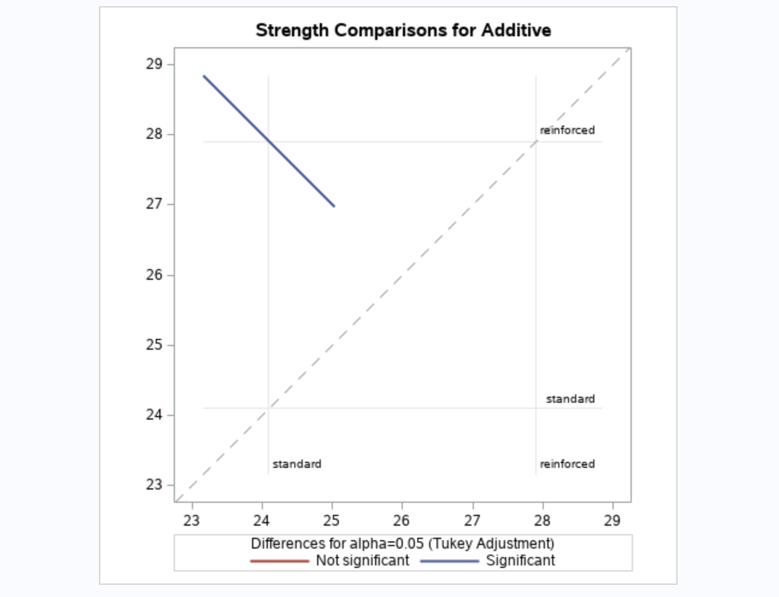
源代码见附件./solution2.sas，运行程序可以得出如下结论：

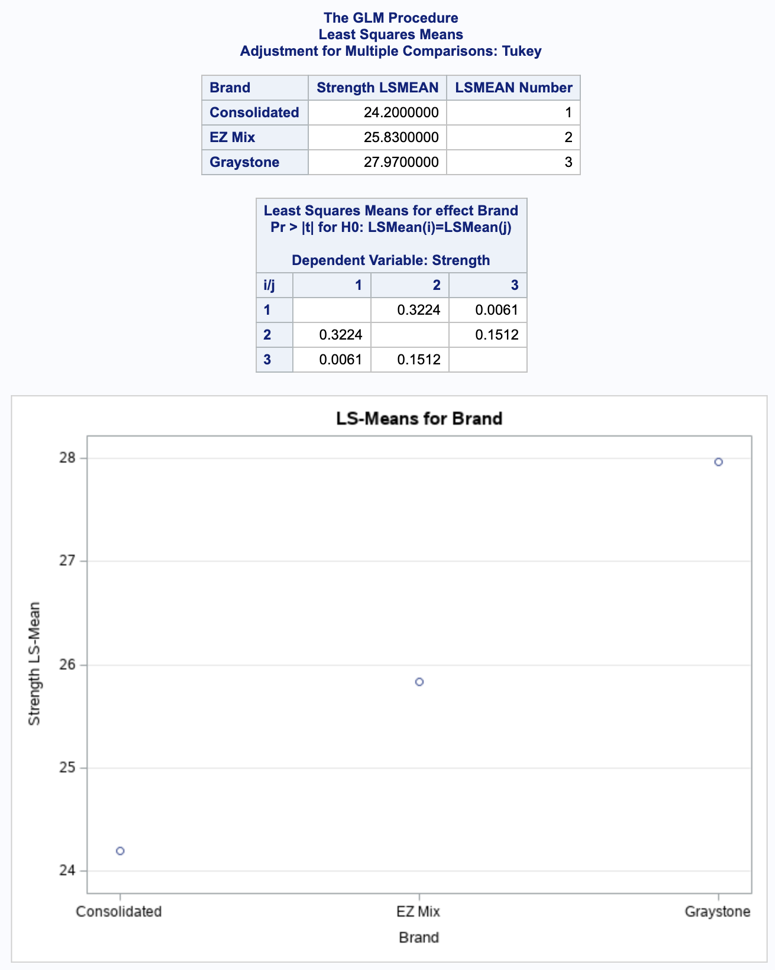
* 方差分析表中可以看到F统计量及相应的p值。F值9.76，p值0.0002小于0.05，拒绝原假设，即各组均值间有显著差异
* 统计量用于度量因变量变异中可由自变量解释部分所占的比例，以此判断统计模型的解释力。我们可以说在此模型中Addition和Brand解释52.9%的Strength值
* 由最小二乘均值表克制p值均小于0.05，拒绝原假设，不同的Brand和Addition均值具有显著差异
* 考虑本数据集中三种不同Brand，LS-Means图展现的是三组Brand对应的Strength的均值。参照各组Strength均值的数值，1组均值为24.2，2组均值为25.83，3组均值为27.97，对应到图中的3个点。显然，LS-Means图中均值点分布比较分散，我们有理由认为各组之间有显著差异
* 实验结果如下所示

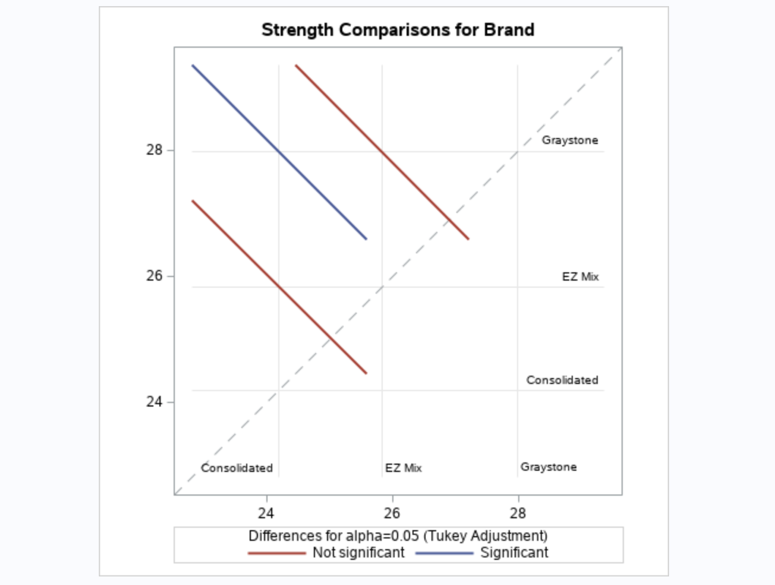












郑重声明

**本作业由作者独立完成。抄袭行为在任何情况下都是不能容忍的(COPY is strictly prohibited under any circumstances)！由抄袭所产生的一切后果由抄袭者承担，勿谓言之不预也。**

**陈麒先**