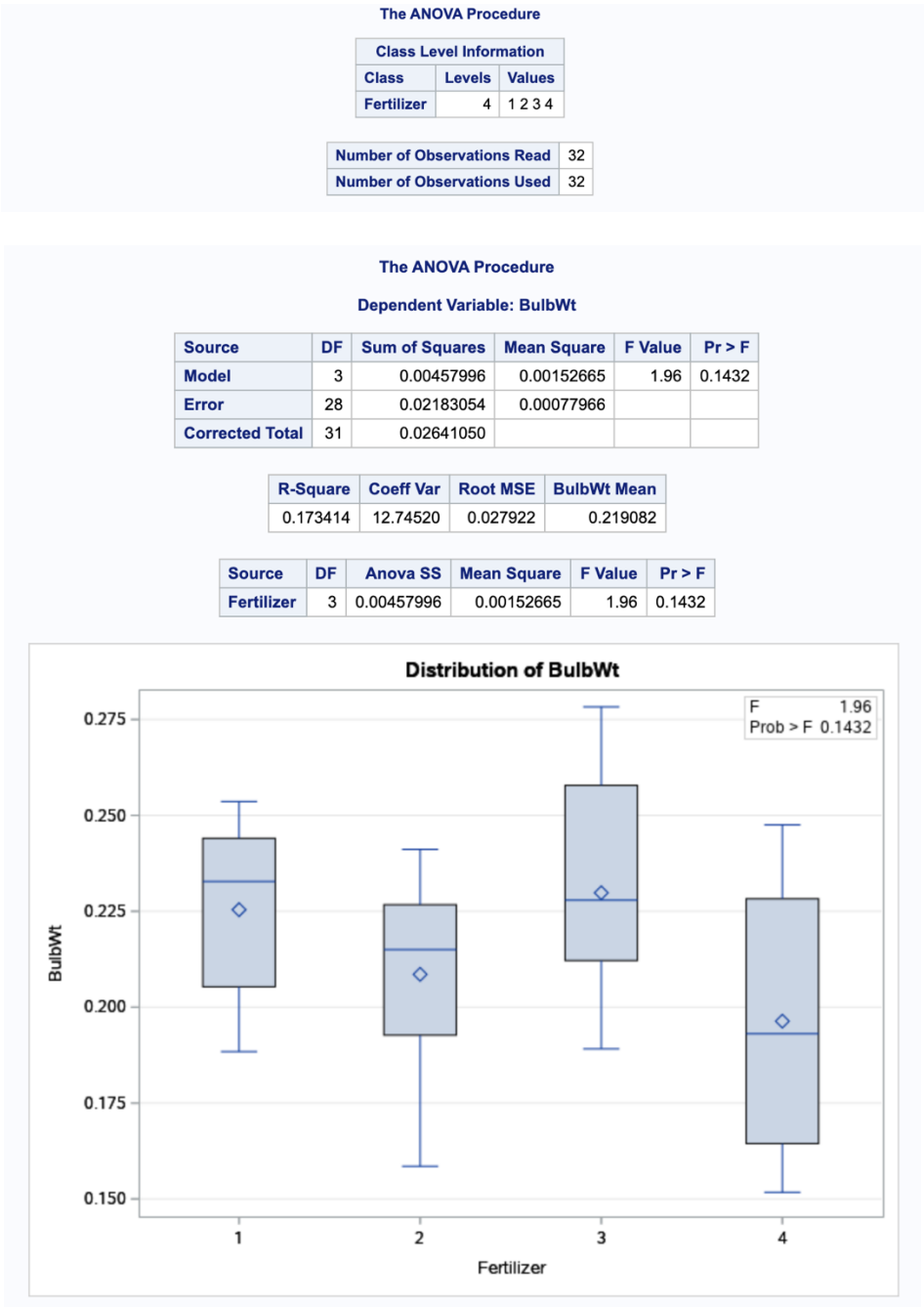




题目一

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：

The MEANS Procedure							
Analysis Variable : Strength							
Brand	Additive	N Obs	N	Mean	Std Dev	Minimum	Maximum
Consolidated	reinforced	5	5	25.8000000	2.3727621	22.7000000	29.3000000
	standard	5	5	22.6000000	1.5313393	20.4000000	24.2000000
EZ Mix	reinforced	5	5	27.2600000	1.9603571	25.5000000	30.2000000
	standard	5	5	24.4000000	3.7729299	19.8000000	28.0000000
Graystone	reinforced	5	5	30.6600000	1.3390295	29.5000000	32.6000000
	standard	5	5	25.2800000	3.1451550	21.2000000	29.8000000

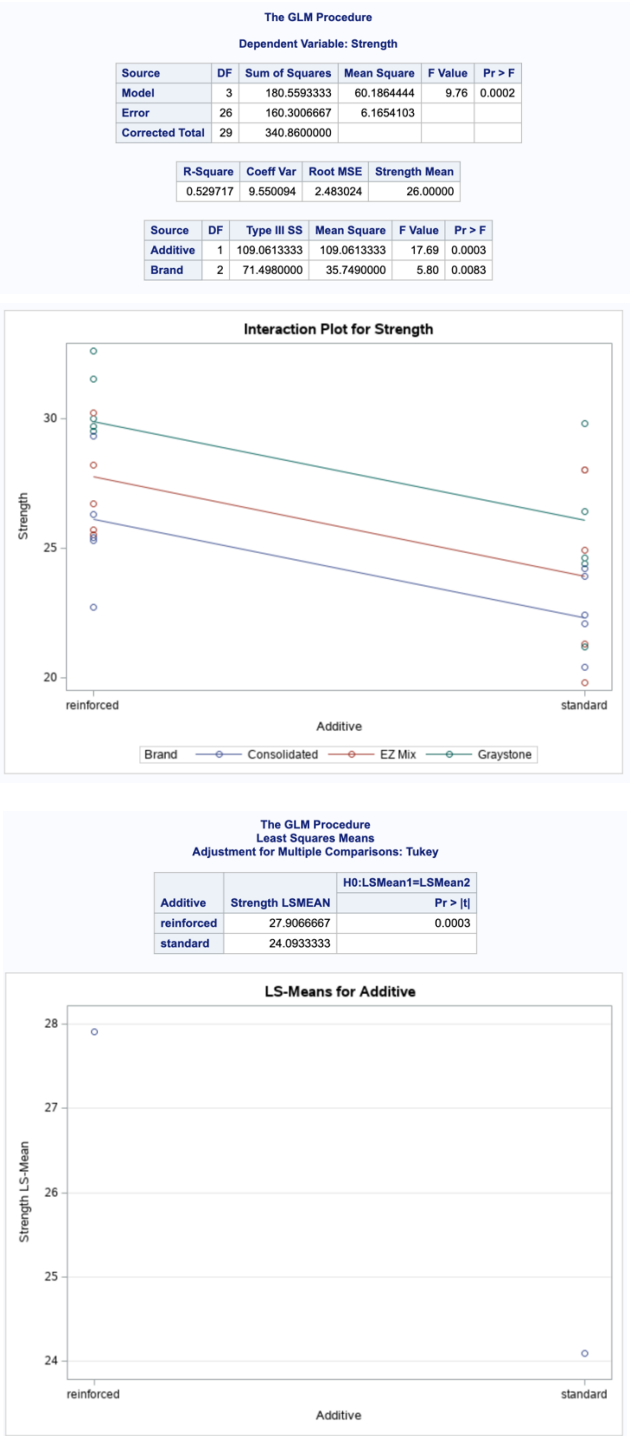
(2) ~ (3)

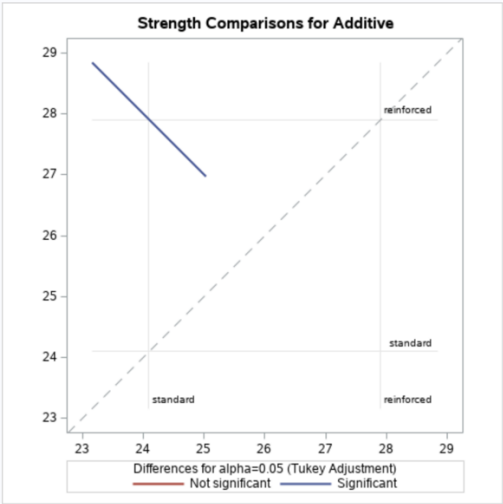
有必要做重复性实验，当在多个不同条件下测量随机样品的所有成员时使用重复测量分析。当样品存在多种不同条件时必须依次进行测量。在这种情况下使用标准 ANOVA 是不合适的，因为它不能建模重复测量不同条件之间的相关性。应该清楚重复测量多个不同条件之间的差异。对样本进行重复测量试验,每个试验测试了在不同条件下相同样品的特点。SAS 提供了 PROC GLM 用于进行重复测量分析。

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
- $R^2$  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



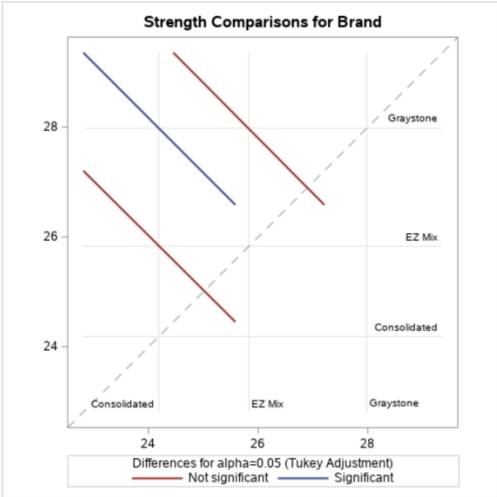
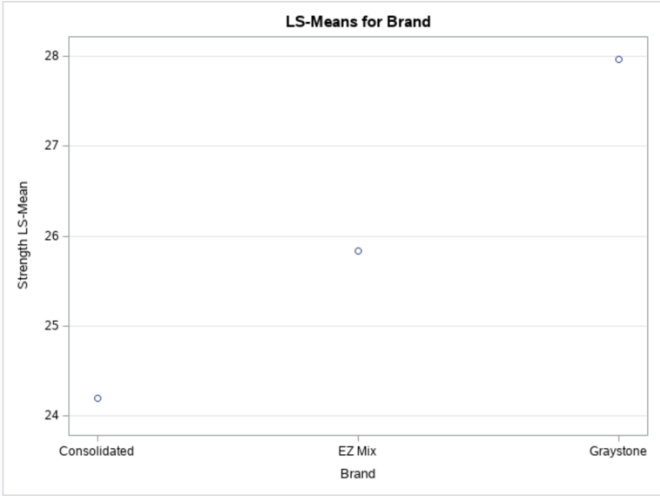


The GLM Procedure  
Least Squares Means  
Adjustment for Multiple Comparisons: Tukey

Brand	Strength LSMEAN	LSMEAN Number
Consolidated	24.2000000	1
EZ Mix	25.8300000	2
Graystone	27.9700000	3

Least Squares Means for effect Brand  
Pr > |t| for H0: LSMEAN(i)=LSMEAN(j)

Dependent Variable: Strength			
i/j	1	2	3
1		0.3224	0.0061
2	0.3224		0.1512
3	0.0061	0.1512	



## 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

