6.7.2: Using SAS

In SAS we would set up the ANOVA as:

```
proc mixed data=school covtest method=type3;
    class Region SchoolType Teacher Class;
    model sr_score = Region SchoolType Region*SchoolType;
    random Teacher(Region*SchoolType);
    store out_school;
run;
```

In SAS `proc mixed` , we see that the fixed effects appear in the model statement, and the nested random effect appears in the random statement.

We get the following partial output:

Type 3 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
Region	1	564.062500	564.062500	Var(Residual) + 2 Var(Teach(Region*School)) + Q(Region,Region*SchoolType)	MS(Teach(Region*School))	4	24.07	0.0080
SchoolType	1	76.562500	76.562500	Var(Residual) + 2 Var(Teach(Region*School)) + Q(SchoolType,Region*SchoolType)	MS(Teach(Region*School))	4	3.27	0.1450
Region*SchoolType	1	264.062500	264.062500	Var(Residual) + 2 Var(Teach(Region*School)) + Q(Region*SchoolType)	MS(Teach(Region*School))	4	11.27	0.0284
Teach(Region*School)	4	93.750000	23.437500	Var(Residual) + 2 Var(Teach(Region*School))	MS(Residual)	8	5.00	0.0257
Residual	8	37.500000	4.687500	Var(Residual)	.	.	.	.

The results for hypothesis tests for the fixed effects appear as:

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Region	1	4	24.07	0.0080
SchoolType	1	4	3.27	0.1450
Region*SchoolType	1	4	11.27	0.0284

Given that the *Region*SchoolType* interaction is significant, the `PLM` procedure along with the `lsmeans` statement can be used to generate the Tukey mean comparisons and produce the groupings chart and the plots to identify what means differ significantly.

```
ods graphics on;
proc plm restore=out_school;
lsmeans Region*SchoolType / adjust=tukey plot=meanplot cl lines;
run;
```

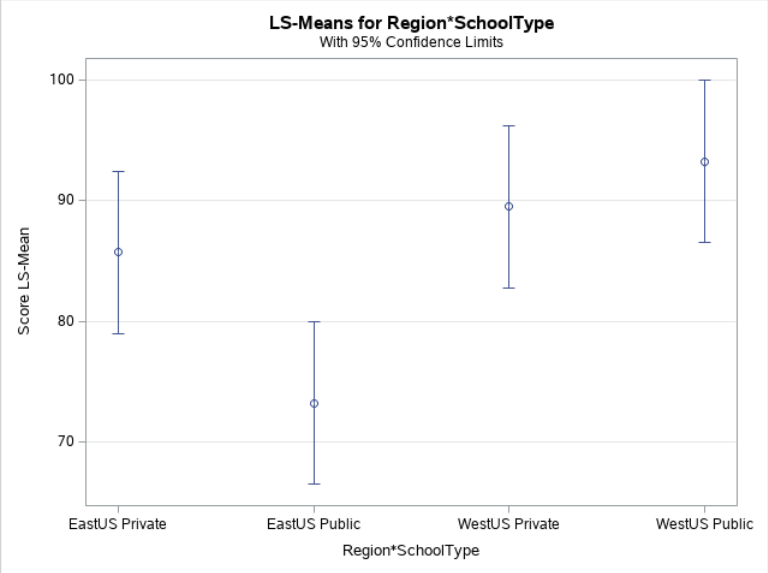

Figure 6.7.2.1: Plot of score LS-means for Region*SchoolType, with 95% confidence limits.

Differences of Region*SchoolType Least Squares Means Adjustment for Multiple Comparisons: Tukey														
Region	SchoolType	_Region	_SchoolType	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
EastUS	Private	EastUS	Public	12.5000	3.4233	4	3.65	0.0217	0.0703	0.05	2.9955	22.0045	-1.4356	26.4356
EastUS	Private	WestUS	Private	-3.7500	3.4233	4	-1.10	0.3349	0.7109	0.05	-13.2545	5.7545	-17.6856	10.1856

Differences of Region*SchoolType Least Squares Means Adjustment for Multiple Comparisons: Tukey													
Region	SchoolType	_Region	_SchoolType	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
EastUS	Private	WestUS	Public	-7.5000	3.4233	4	-2.19	0.0936	0.2677	0.05	-17.0045	2.0045	-21.4356	6.4356
EastUS	Public	WestUS	Private	-16.2500	3.4233	4	-4.75	0.0090	0.0301	0.05	-25.7545	-6.7455	-30.1856	-2.3144
EastUS	Public	WestUS	Public	-20.0000	3.4233	4	-5.84	0.0043	0.0146	0.05	-29.5045	-10.4955	-33.9356	-6.0644
WestUS	Private	WestUS	Public	-3.7500	3.4233	4	-1.10	0.3349	0.7109	0.05	-13.2545	5.7545	-17.6856	10.1856

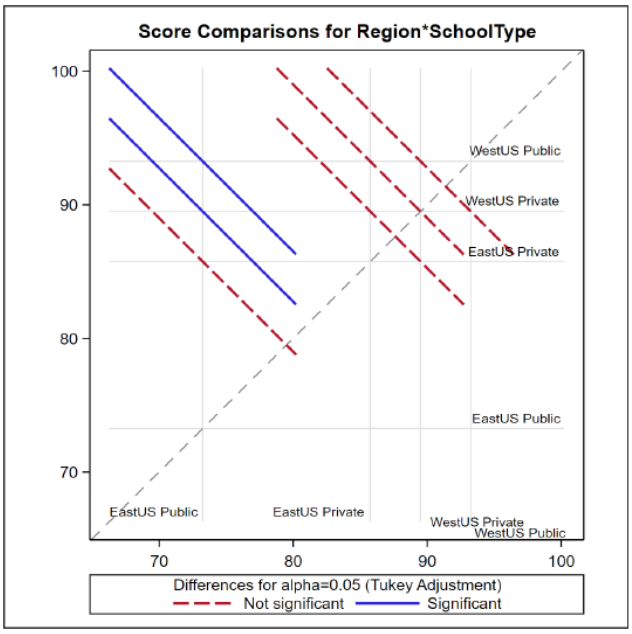

Figure 6.7.2.2: Diffogram of score comparisons for Region*SchoolType.

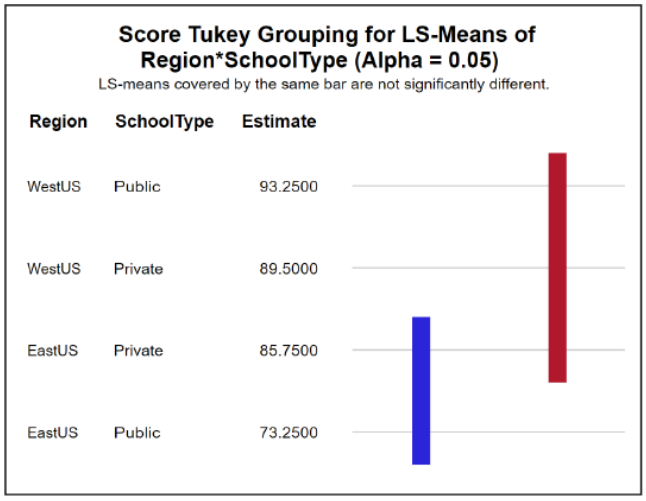

Figure 6.7.2.3: Score Tukey grouping for LS-means of Region*SchoolType.

From the results, it is clear that the mean self-rating scores are highest for the public school in the west region. The difference mean scores for public schools in the west region is significantly different from the mean scores for public schools in the east region as well as the mean scores for private schools in the east region.

The `covtest` option produces the results needed to test the significance of the random effect, *Teach(Region*SchoolType)* in terms of the following null and alternative hypothesis:

$$H_0 : \sigma^2_{teacher} = 0 \text{ vs. } H_a : \sigma^2_{teacher} > 0$$

However, as the following display shows, `covtest` option uses the Wald Z test, which is based on the z-score of the sample statistic and hence is appropriate only for large samples—specifically, when the number of random effect levels is sufficiently large. Otherwise, this test may not be reliable.

Covariance Parameter Estimates

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z
Teach(Region*School)	9.3750	8.3689	1.12	0.2626
Residual	4.6875	2.3438	2.00	0.0228

Therefore, in this case, as the number of teachers employed is few, Wald's test may not be valid. It is more appropriate to use the ANOVA F-test for *Teacher(Region*SchoolType)*. Note that the results from the ANOVA table suggest that the effects of the teacher within the region and school type are significant (Pr > F = 0.0257), whereas the results based on Wald's test suggest otherwise (since the p-value is 0.2626).