

Уровень значимости для всех гипотез задается 0.01. Зеленым обозначено условие для первого варианта, желтым для второго. Переслать SAS код делающий все последовательности шагов.

- 1) Проверить влияет ли Страна производитель (Origin) (или тип Кузова (Type)) на мощность (HorsePower) (или размер двигателя EngineSize) с заданным уровнем значимости.
- 2) Преобразовать категориальные переменные так, чтобы не было «неразличимых» групп (используя график diffogram, попарный t-test и шаг данных для объединения «неразличимых» групп). Написать соответствующий код.
- 3) Добавить предиктор тип Кузова (Type) (или Страна производитель (Origin) соответственно), чтобы понять улучшается ли модель (по RMSE). Объединения «неразличимых» групп не проводить.
- 4) Проверить нужен ли предиктор Origin\*Type и если не нужен, то исключить из модели. Построить финальную модель.
- 5) Сформировать pdf отчет, содержащий ТОЛЬКО таблицы: «Least Squares Means for effect ...» шага 2, ANOVA таблицу и таблицу статистик (fit statistics) последнего шага.

Бонус задание: Перед вторым шагом написать шаблон с помощью процедуры template, чтобы в таблице попарных сравнений «Least Squares Means for effect ...» (имя объекта stat.GLM.Pdiff) значение P-values больше заданного уровня значимости (0.01) подсвечивались красным или отмечалось «\*».