

**Çoklu regresyon için ödev sorusu**

Montgomery, Peck, Vining, Doğrusal Regresyon Analizine Giriş kitabında **Tablo B1**'de verilen **"Milli futbol ligi"** verileri için aşağıdaki soruları cevaplayınız.

Bu verilere R programında **"library(MPV)"** ve **"data(table.b1)"** kullanarak ulaşabilirsiniz.

1. Kazanılan oyun sayısı ( $y$ ) olmak üzere, takımın pas mesafesi ( $x_2$ ), koşu oyunu yüzdesi ( $x_7$ ) ve rakplerin koşma mesafesi ( $x_8$ ) değişkenleriyle çoklu regresyon modelini oluşturunuz. (R sonuçlarını kullanınız.)
2. Oluşturduğunuz regresyon modeli için varyans analizi tablosunu oluşturunuz. (R sonuçlarını kullanarak oluşturunuz.)
3. Modelin anlamlılığı için hipotez testini oluşturarak, modelin anlamlılığını test ediniz.  
 $H_0 : \beta_2 = 0, H_1 : \beta_2 \neq 0$
4.  $H_0 : \beta_7 = 0, H_1 : \beta_7 \neq 0$  hipotezlerini ayrı ayrı  $t$  testini kullanarak test ediniz.  
 $H_0 : \beta_8 = 0, H_1 : \beta_8 \neq 0$   
Bu değişkenler model için önemli midir? Yorumlayınız.
5. Model için  $R^2$  ve  $R_{adj}^2$  değerlerini varyans analizi tablosundan hesaplayarak yorumlayınız.
6. Kısmi  $F$  testini kullanarak  $x_7$  ile gösterilen bağımsız değişkenin modele katkısını belirleyiniz. Bu kısmi  $F$  testinin değeri ile 4. soruda  $\beta_7$  için elde ettiğiniz  $t$  istatistiğinin değeri arasında nasıl bir ilişki vardır?
7.  $y_i$  değerleri ile  $\hat{y}_i$  tahmin değerlerinin korelasyon katsayısını bulunuz. Bu korelasyon katsayısının karesi hangi değere eşittir?
8.  $\beta_2, \beta_7$  ve  $\beta_8$  için %95 güven aralıklarını bulunuz.
9.  $x_2 = 2300, x_7 = 56$  ve  $x_8 = 2100$  olduğunda ortalama kazanılan oyun sayısı için %95 güven aralığını bulunuz.
10. Yukarıda üç bağımsız değişkenle oluşturduğunuz regresyon modelini sadece  $x_7$  ve  $x_8$  değişkenlerini kullanarak yeniden oluşturunuz.  
Bu yeni model için varyans analizi tablosunu oluşturarak, modelin anlamlılığını test ediniz.  
 $R^2$  ve  $R_{adj}^2$  değerlerini hesaplayarak 5. soruda bulduğunuz değerler ile karşılaştırınız.
11.  $\beta_7$  ve  $\beta_8$  için %95 güven aralıklarını bulunuz.  
Ayrıca,  $x_7 = 56, x_8 = 2100$  olduğunda ortalama kazanılan oyun sayısı için %95 güven aralığını bulunuz.  
Bulduğunuz bu güven aralıklarını 8. ve 9. soruda bulduklarınız ile karşılaştırınız.
12. Bu iki modeli karşılaştırdığınızda, modelden önemli bir bağımsız değişkenin çıkartılması ile oluşacak durum ile ilgili ne söyleyebilirsiniz.