

MSGSU İstatistik Bölümü**R Programlama Diline Giriş Final Sınavı****Ad-Soyad:****İmza:****Numara:**

Açıklama: 1. ve 2. soruları R script (.R uzantılı) dosyası biçiminde, 3. soruyu ise R Markdown (.Rmd uzantılı) dosyası biçiminde sisteme yükleyiniz. Başka dosya biçimlerinde gerçekleştirilen yüklemeler geçersiz sayılacaktır!!!

1) Bir tur firmasının aylık satış verileri ile ilgili tablosu şöyledir: (20 p.)

| Tur | Satış Fiyatı | Satış Adedi |
|---------------|--------------|-------------|
| Yunan Adaları | 2000 | 200 |
| İsviçre | 5600 | 100 |
| Japonya | 25000 | 0 |

- Tablodaki her bir gözede bulunan veriler için değişken tanımlayınız (Sayıları numeric olarak tanımlayınız.) Örneğin; `tur1<-“Yun”` (Toplamda 6 değişken tanımlanacak.) (5 p.)
- Tur firmasının aylık kazancını tanımladığınız değişkenleri kullanarak hesaplayınız. (5 p.)
- Tur firması veri girişi yapılırken satış adedine ondalıklı sayı girilmesini engellemek için satış adedi ile ilgili değişkenleri integer türüne dönüştürmek istemektedir. İlgili değişkenlerde gerekli integer dönüşümünü uygulayınız. (5 p.)
- Tur firması Japonya turunu hiç satamadığı için bu ürüne ait değişkenleri silmek istemektedir. Bu ürüne ait tanımladığınız değişkenleri siliniz. (5 p.)

2) Bir hava yolları şirketinin yaz dönemi İstanbul kalkışlı aylık uçuş listesi şöyledir: (30 p.)

| | Ankara | İzmir | Antalya |
|---------|--------|-------|---------|
| Haziran | 100 | 250 | 300 |
| Temmuz | 85 | 300 | 450 |
| Ağustos | 120 | 350 | 560 |

- Şirketin uçuş listesi bilgilerinden oluşan ucus isimli bir matris oluşturun. (8 p.)
- Tabloda verilen satır ve sütun isimlerini ucus matrisine satır ve sütun ismi olarak ekleyin. (6 p.)

- c) Her ay için toplam uçuş sayısı bilgilerini tutan aylık_ucus_toplami isimli bir vektör oluşturun. (8 p.)
- d) Oluşturmuş olduğunuz aylık_ucus_toplami vektörünü ucus matrisine sütun olarak ekleyin. (8 p.)

3) Bir kalp hastalığına ilişkin hasta bilgilerini içeren heart.csv dosyasını bilgisayarınıza indiriniz. R Markdown kullanarak veri setini R'a aktarıp veri setine kalp ismini atayınız. Tidyverse paketini ve pipe operatörünü kullanarak aşağıdaki soruları cevaplayınız. (.Rmd dosyasını yükleyiniz.) (50 p.)

a) Birinci kısım: (20 p.)

1. Age, ChestPaintType, MaxHR, HeartDisease değişkenlerinden oluşan verileri seçiniz ve oluşan veri setine SH ismini veriniz.
2. SH veri setindeki 40 yaş üstü tüm hastalardan ChestPainType değişken değeri ATA olan verileri seçiniz.

b) İkinci kısım: (20 p.)

1. ggplot fonksiyonunu kullanarak kalp veri setinin x ekseninde Cholesterol, y ekseninde MaxHR değişkenlerinin bulunduğu serpilme diyagramını çizdiriniz.
2. x eksenine Kolesterol, y eksenine Maksimum Kalp Atış Hızı isimlerini veriniz.
3. Serpilme diyagramındaki noktaları HeartDisease değişkenine göre renklendiriniz.

c) Üçüncü kısım: (10 p.)

kalp veri setinde Cholesterol değişkenine ait tanımlayıcı istatistikleri (min, max, kartiller, ortalama, medyan) ve basıklık-çarpıklık katsayılarını bulunuz.

Başarılar...

Dr. Bilge ÖZLÜER BAŞER