

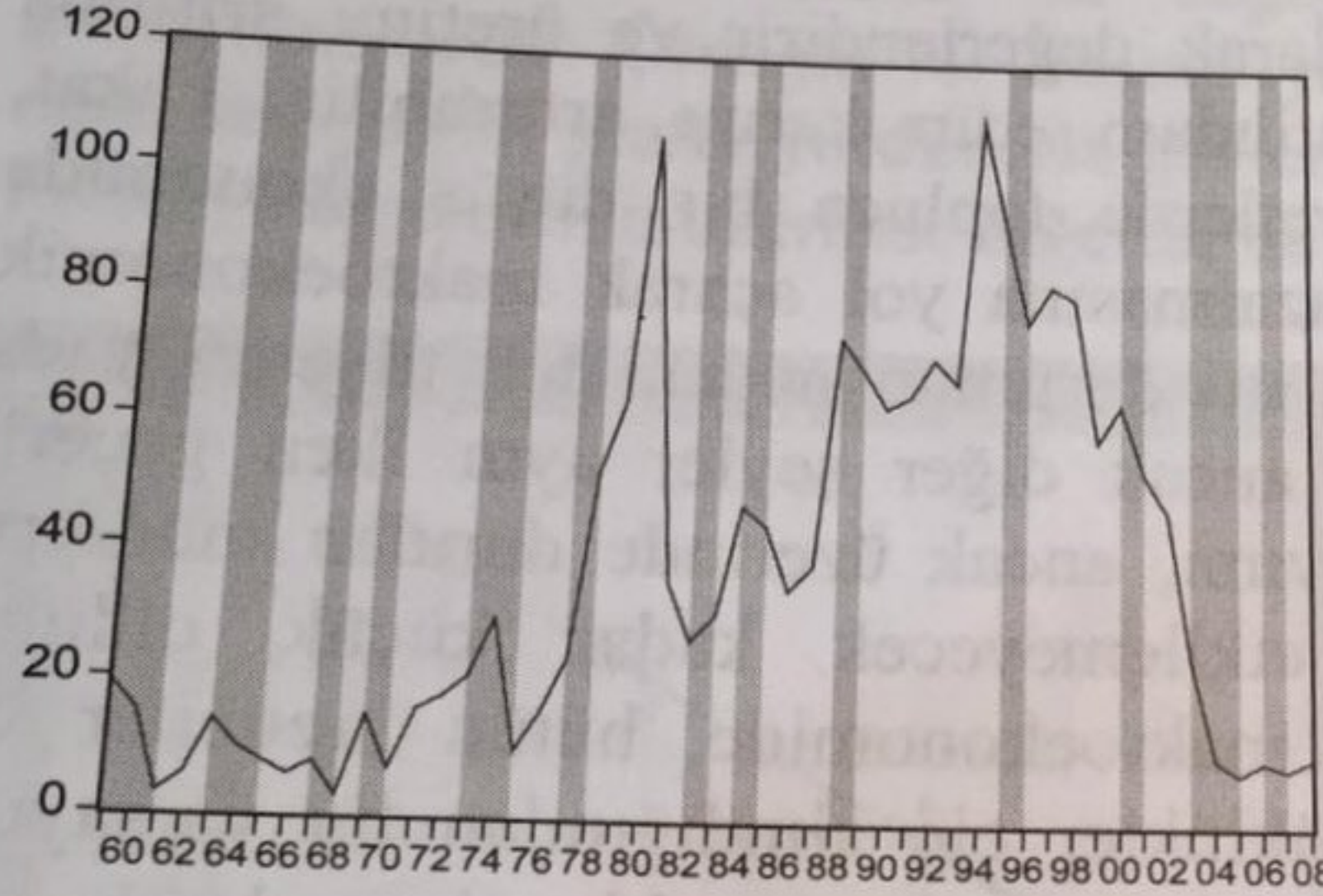
25.11.2022

Veri Görselleştirme

Hafta 6: İlişkilerin Görselleştirilmesi



İKTİSADİ BAKIŞ 4



Yandaki grafik bize Türkiye’de enflasyon oranının 1959-2010 döneminde izlediği seyri gösteriyor. Grafikte görülen taralı alanlar ise Fenerbahçe’nin şampiyon olduğu futbol



sezonlarını ifade ediyor. Grafiğe göre, Fenerbahçe’nin ligde iyi performans gösterdiği dönemlerde enflasyon oranı düşüyor, kötü performans gösterdiği dönemlerde ise, genellikle, enflasyon yükselme eğilimine giriyor. Sizce birkaç sezon üst üste Fenerbahçe’yi daha sezon başında şampiyon ilan ederek Türkiye ekonomisini enflasyon sorunundan kurtarmamız mümkün mü?

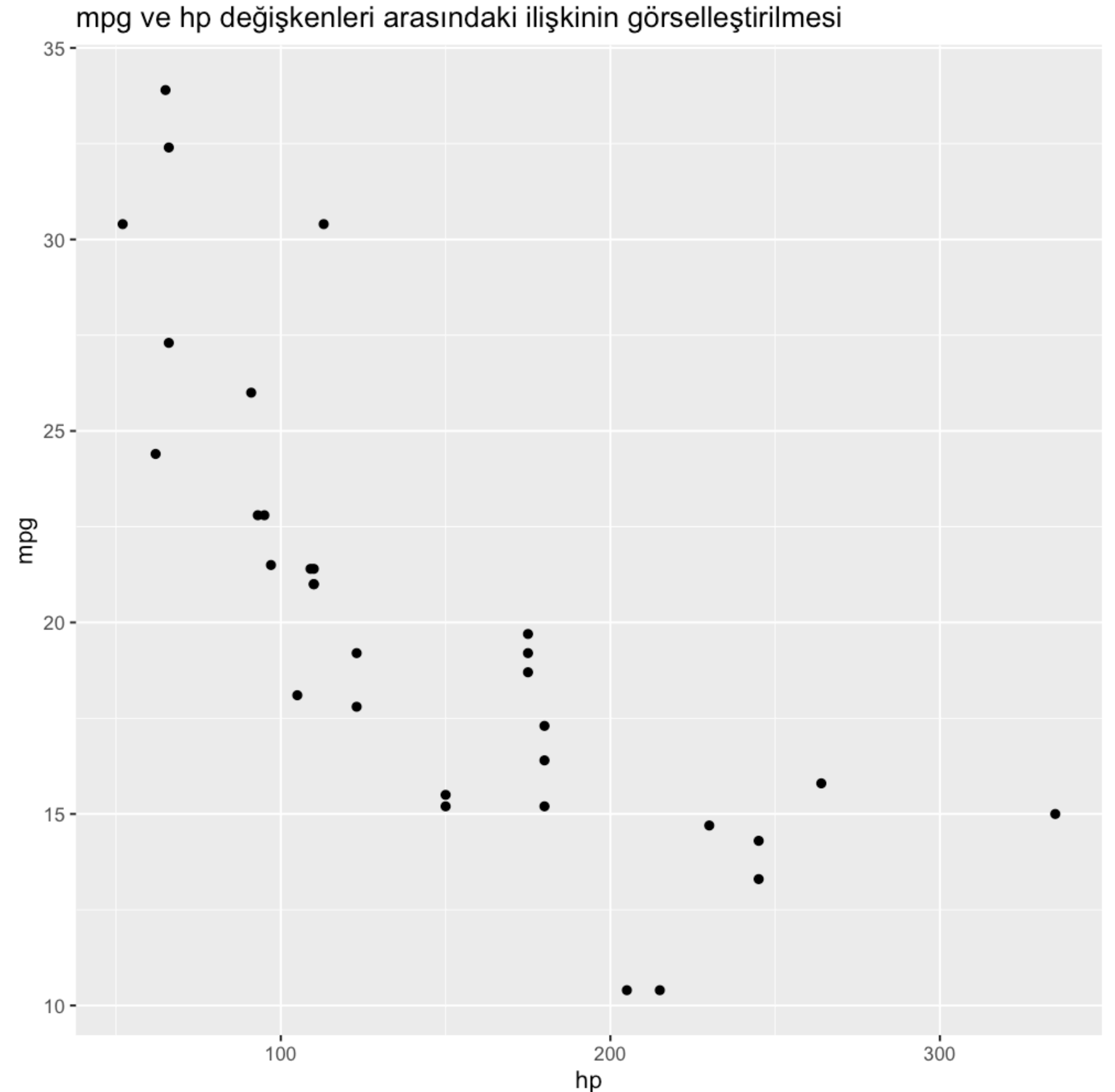
Korelasyon, nedensellik belirtmez;
ancak bir nedensellik aramak için iyi bir sebeptir!

Giriş

- İki değişken arasındaki ilişkinin belirlenmesi, bilimsel çalışmalar ve pratik araştırmalarda sıklıkla karşılaşılan bir problemdir.
- Bu ilişki iki boyut üzerinden ifade edilir:
 1. Yön (pozitif veya negatif)
 2. Şiddet (düşük-orta-yüksek gibi)
- Örneğin, **hava sıcaklıkları ve günlük ihtiyaç duyulan enerji miktarı arasındaki ilişki** veya **kişi başına düşen milli gelir ve ortalama yaşam süresi arasındaki ilişki** ölçülmek istenebilir.
- Bu gibi ilişkileri görselleştirmek için kullanılan en temel araç, saçılım grafiğidir (scatter-plot). Yanı sıra korelogram ve balon grafikleri (bubble-chart) de kullanılabilir.

Saçılım grafikleri (scatter-plot)

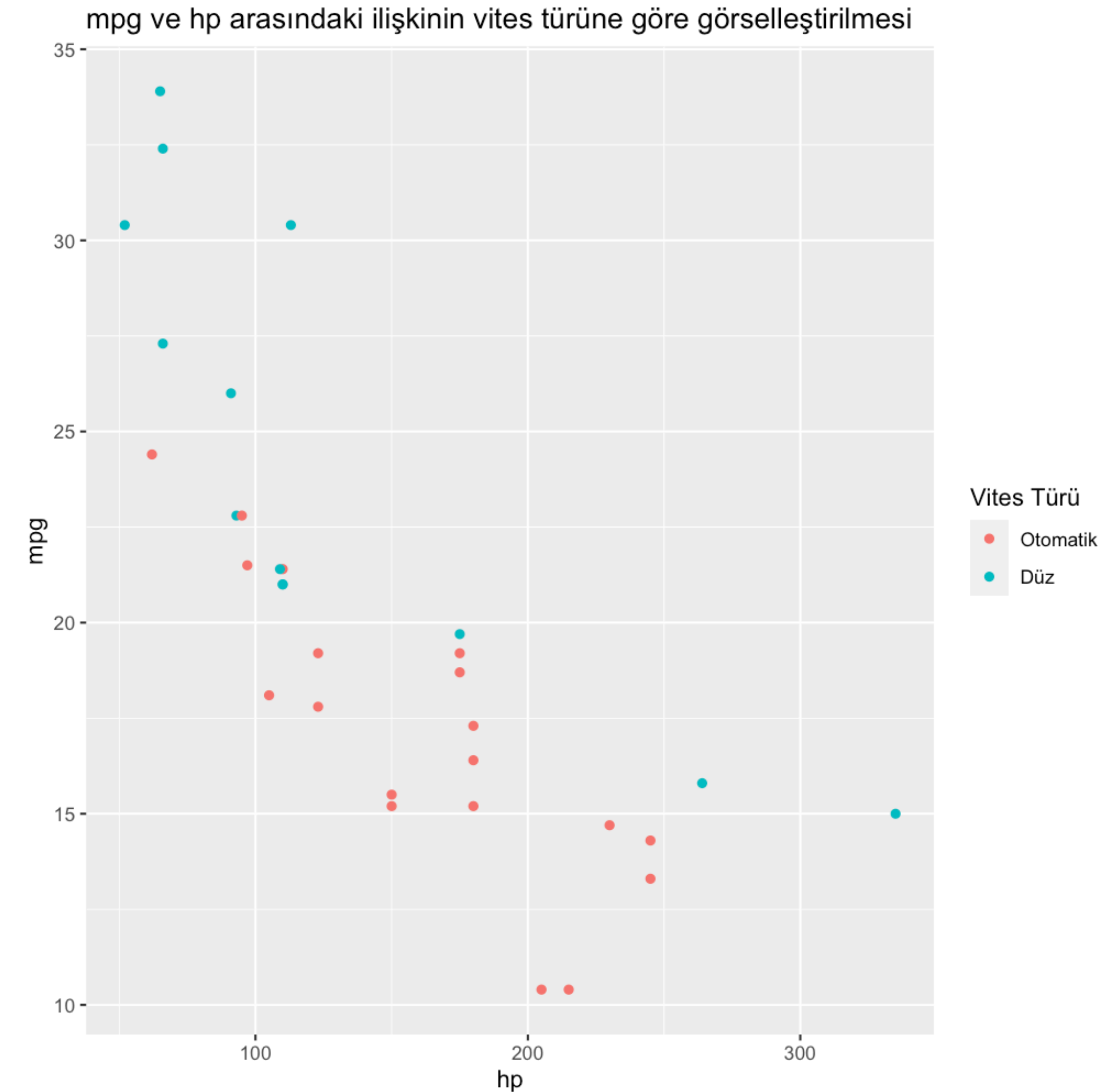
- İki değişken arasındaki ilişkinin araştırılması için kullanılan temel araçlardan biridir.
- İki boyutlu bir eksen üzerinde, her ekseninde bir değişkenin gözlem değerleri diğer eksenindeki değişkenin gözlem değerleriyle eşleştirilerek oluşturulur.



Saçılım grafikleri (scatter-plot)

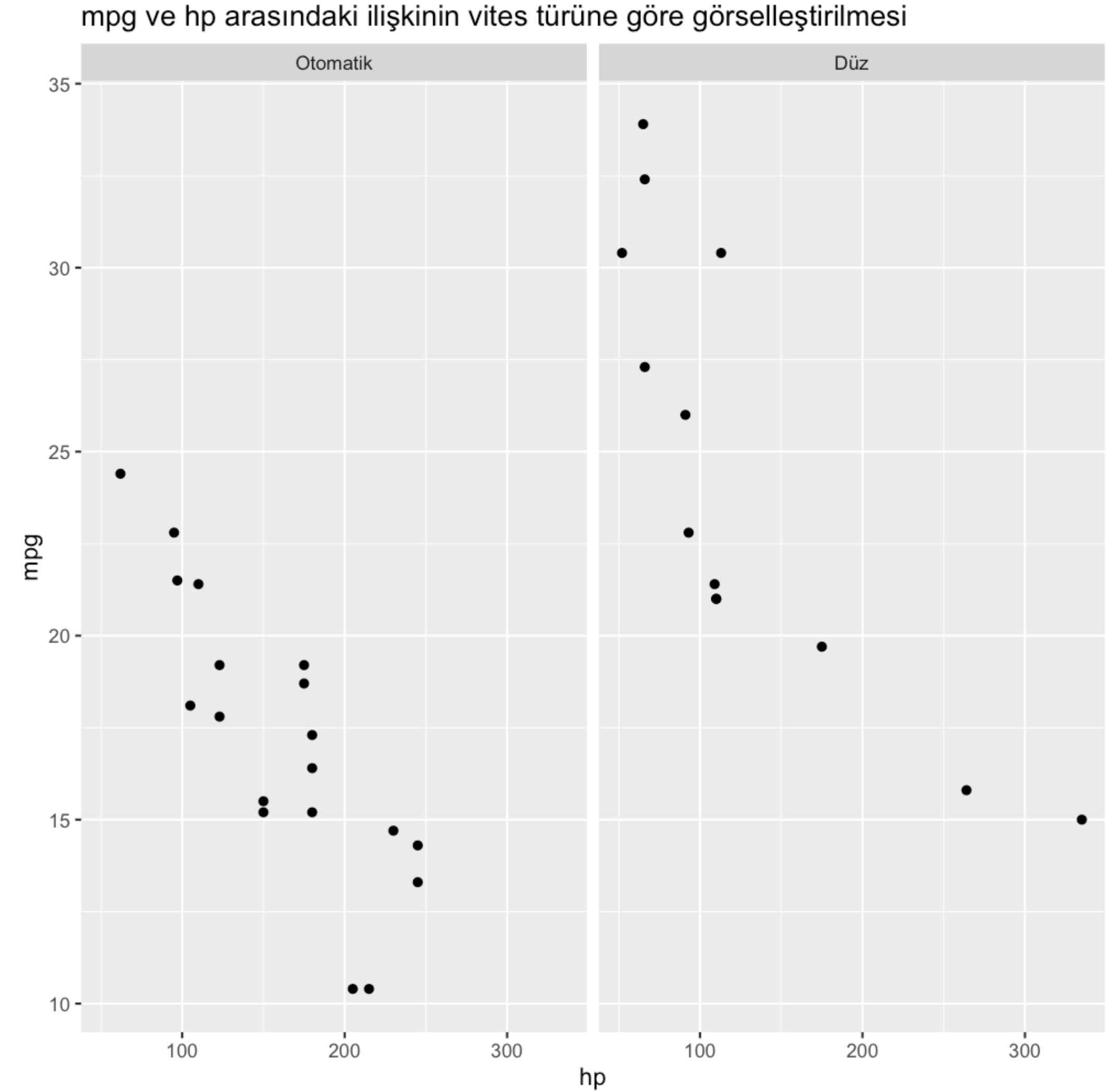
Saçılım grafiklerine renk estetiği kullanarak bir kategorik değişkeni de görselleştirme çalışmasına dahil edebiliriz.

Burada farklı bir görselleştirme stratejisi kullanabilir miydik?

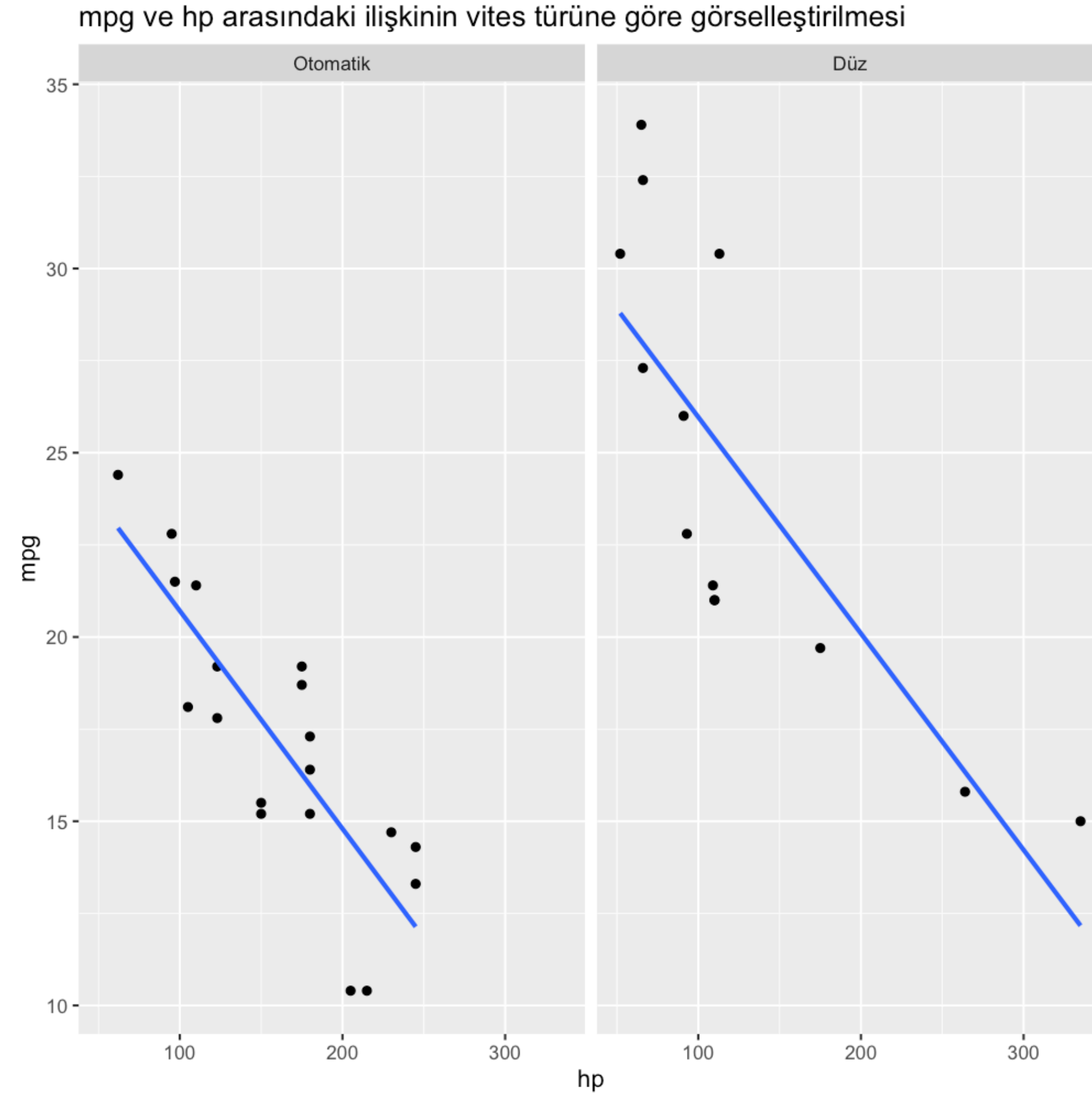


Saçılım grafikleri (scatter-plot)

Saçılım grafiklerine regresyon doğrusu ekleyerek, gözlem değerlerinin eğilimini ortaya koyabiliriz.

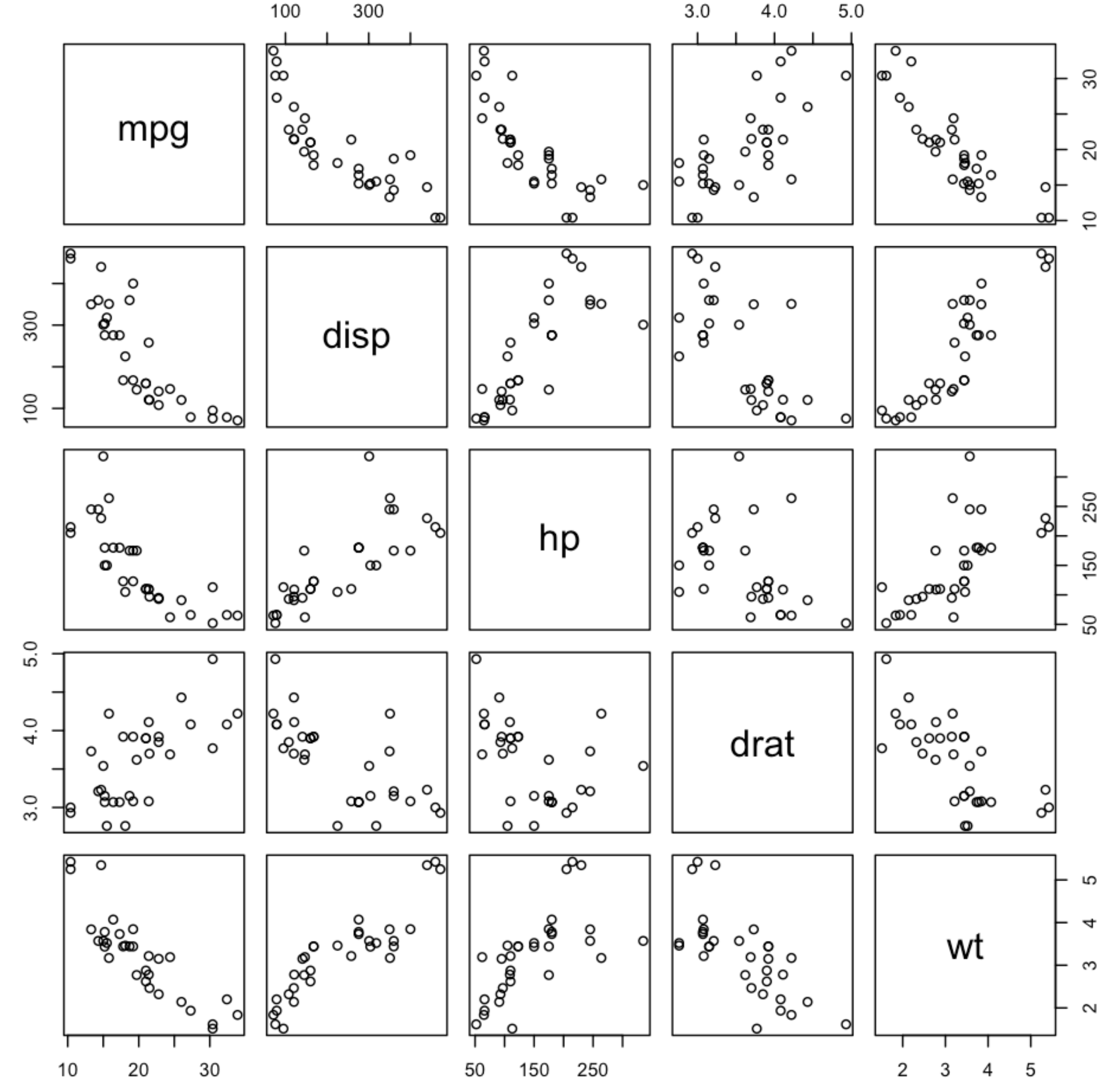


Saçılım grafikleri (scatter-plot)

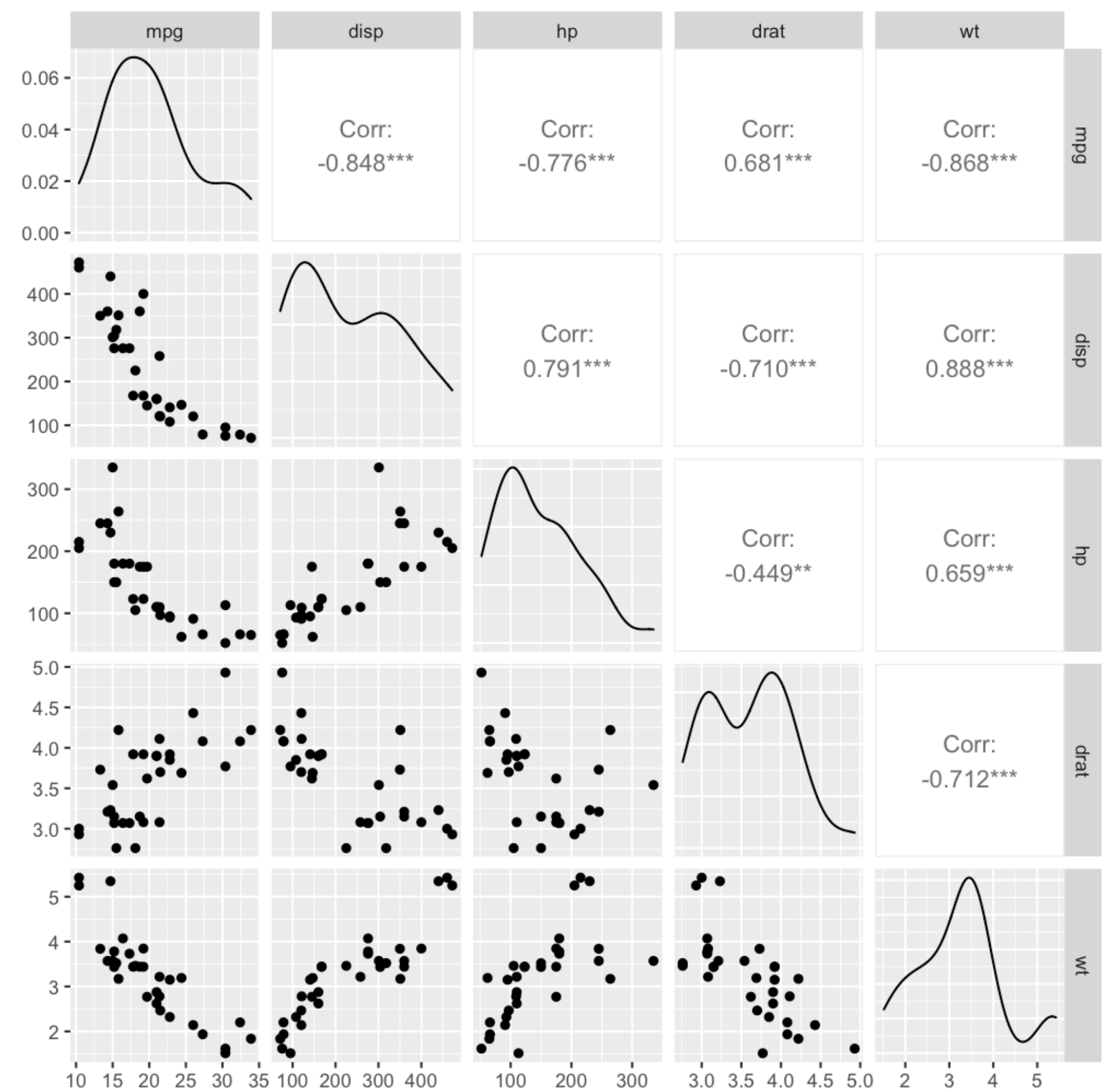


Saçılım matrisleri

Birden fazla değişken ikilisinin görselleştirilmesi için kullanılırlar.



Saçılım matrisleri

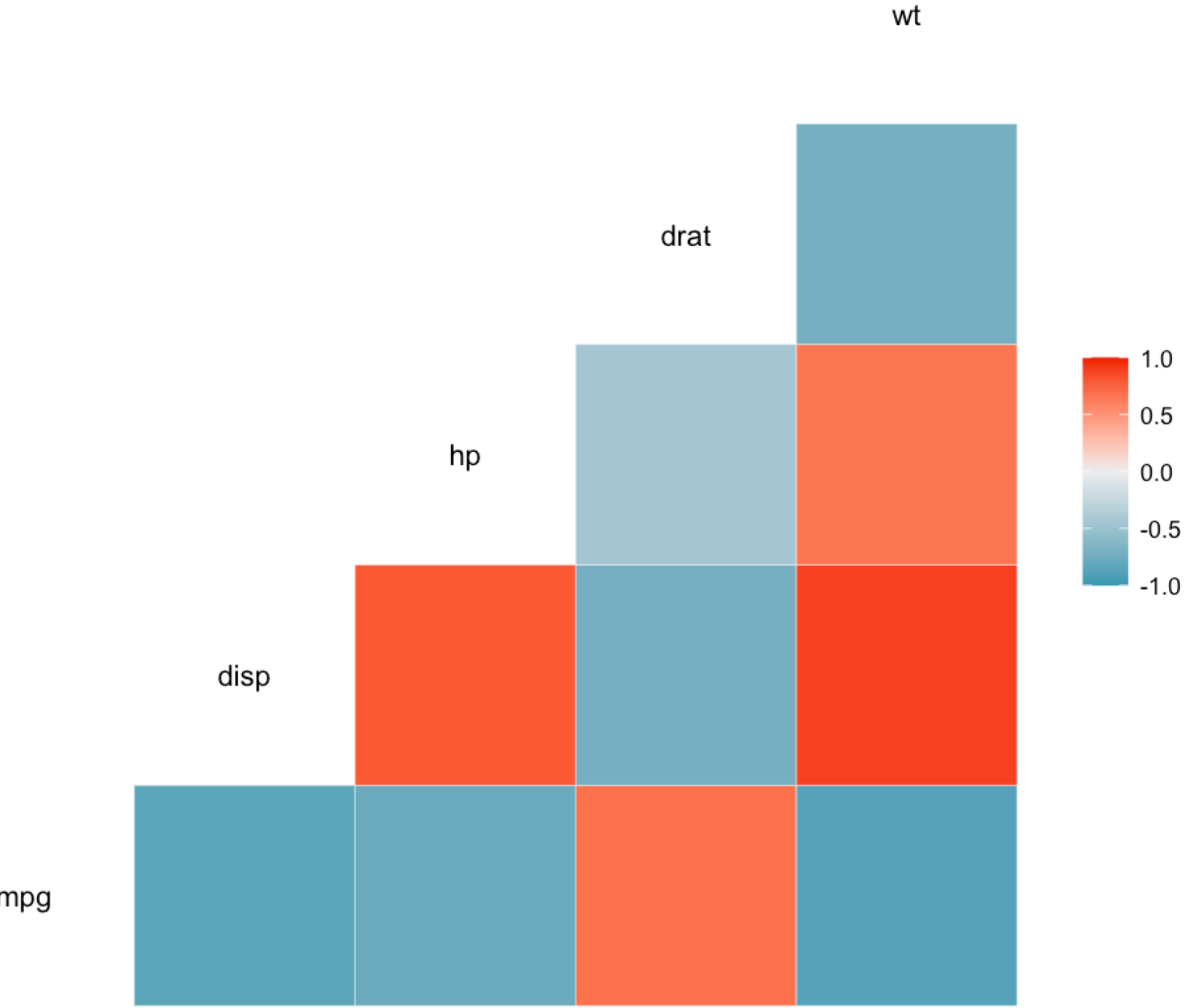


Korelogram (correlogram)

Değişken sayısı arttığında saçılım matrislerini okumak zorlaşmaktadır. Bu durumda ilişkiyi görselleştirmektense ilişki miktarını ölçmek ve bu ölçümleri görselleştirmek daha kullanışlıdır.

- Bunu yapabilmenin en kolay yolu, korelasyon katsayısını (r) hesaplamaktır. Korelasyon katsayısı -1 ve +1 arasında değerler alır.
- $r = 0$ değerini aldığı anda iki değişken arasında herhangi bir ilişki olmadığını, -1 ve ya +1 değerini aldığı anda ise yüksek düzeyde bir ilişki olduğunu gösterir.
- Korelasyon katsayısının işareti ise ilişkinin yönünü gösterir.
- Korelasyon katsayılarının görselleştirilmesi için korelogram kullanılır.

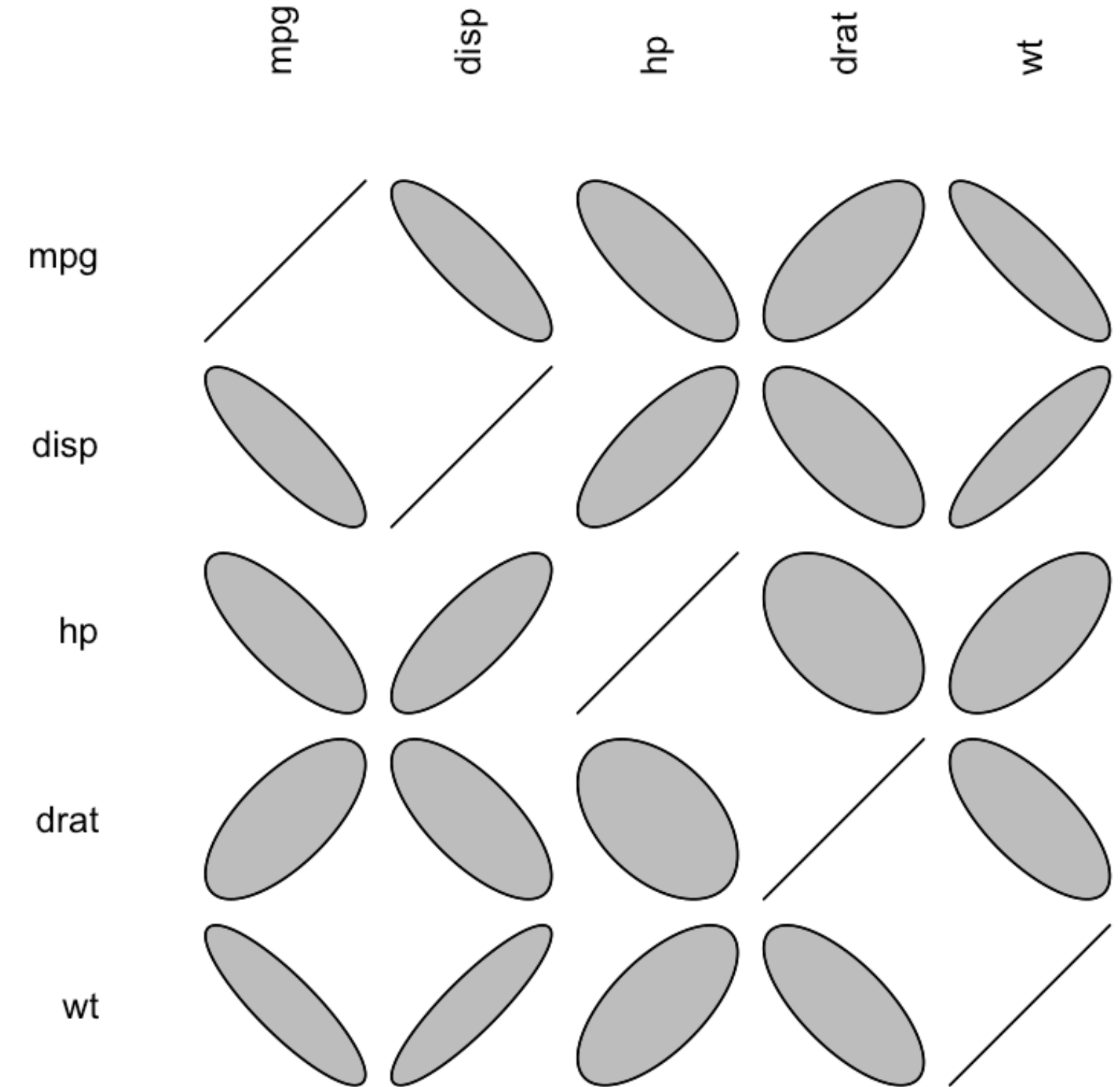
Korelogram (correlogram)



Korelogram (correlogram)

Burada oval nesneler korelasyonun yönüne göre konumlanır ve şiddetine göre şekil alırlar.

Korelasyon katsayısı arttıkça nesneler inceleşir, azaldıkça genişler ve daire formuna yaklaşır.



Kaynak

Bu derste yer alan not ve görseller, Claus O. Wilke'nin "Fundamentals of Data Visualization" isimli kitabından derlenmiştir.

