METOD

Veri toplama ve örneklem hakkında bilgi,

Amaç …. 6 grup vardır önsel bilginin araştırılması vb

Likert ölçeği ile kişilerin …. Değerlendirme sorularını kendilerine uygun şekilde doldurması ve anailizi ile yeni soruların boyut indirgenerek oluşturulması, ikili seçimlerde soru kombinasyonları belirlenirken birbirine en uzak 2 ifadenin tek soruda yer alması hedeflenmiştir.

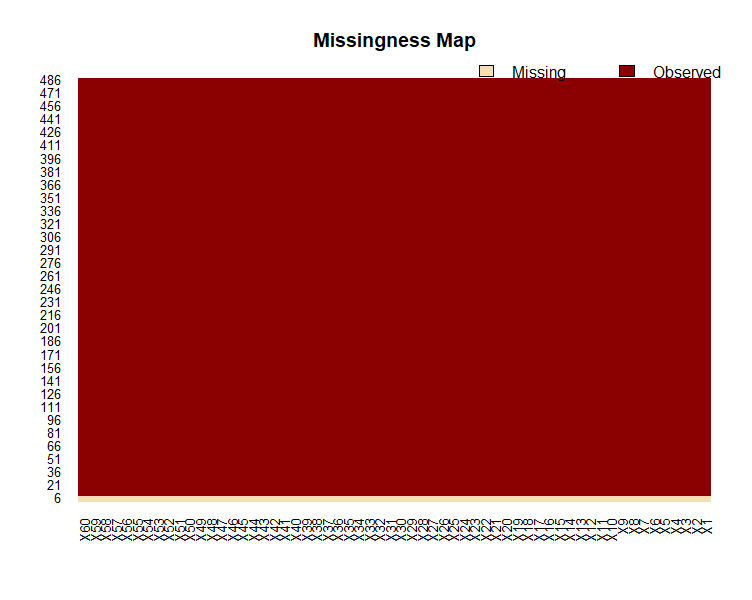
\*

İkili veri seti ve araştırma örneklem vb.

Yeni düzenlenen çalışmada her soru 2 ifade içermektedir. Katılımcılar soruları cevaplarken kendilerine en yakın olan ifadeyi seçmiştir. Toplamda 60 ifade yer almaktadır ve her bir boyutu temsilen 10 ifadeye yer verilmiştir.

Amaç: Birbirinden ayrı 6 homojen grubun varlığının sorgulanması, ifadelerin grupların varlığını temsil etme gücü ve ikili ifade kombinasyonlarının anlamlılığının araştırıldı.

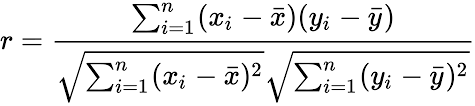
Veriseti: 486 gözlem içermektedir. Her bir ifade bir değişkeni temsil etmektedir. 60 değişken, 480 gözlemden oluşan veri setinde homojen grupları araştırmadan önce keşifçi veri analizi ve veri görselleştirme yöntemleri ile değişkenlerin gücünü ve birbirleri arasındaki ilişki incelendi.



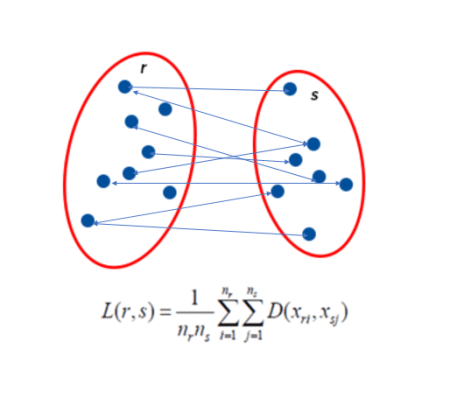
486 gözlemden 6’sı kayıp gözlem olmakla birlikte bu gözlemlerde hiçbir değişkene ilişkin bilgi yer almamaktadır. Bu nedenle kayıp gözlemler doğrudan veriden çıkartılmıştır.

İfadelerin grupları temsil etmede gücünü ölçmek için değişkenler arası hiyerarşik kümeleme analizi yapıldı. Analiz yapılırken veri seti binary (ikili) olduğu için değişkenler arası uzaklık matrisi oluşturmada korelasyon metodu kullanıldı.

Pearson uzaklık matrisi aşağıdaki formül ile hesaplanmıştır.

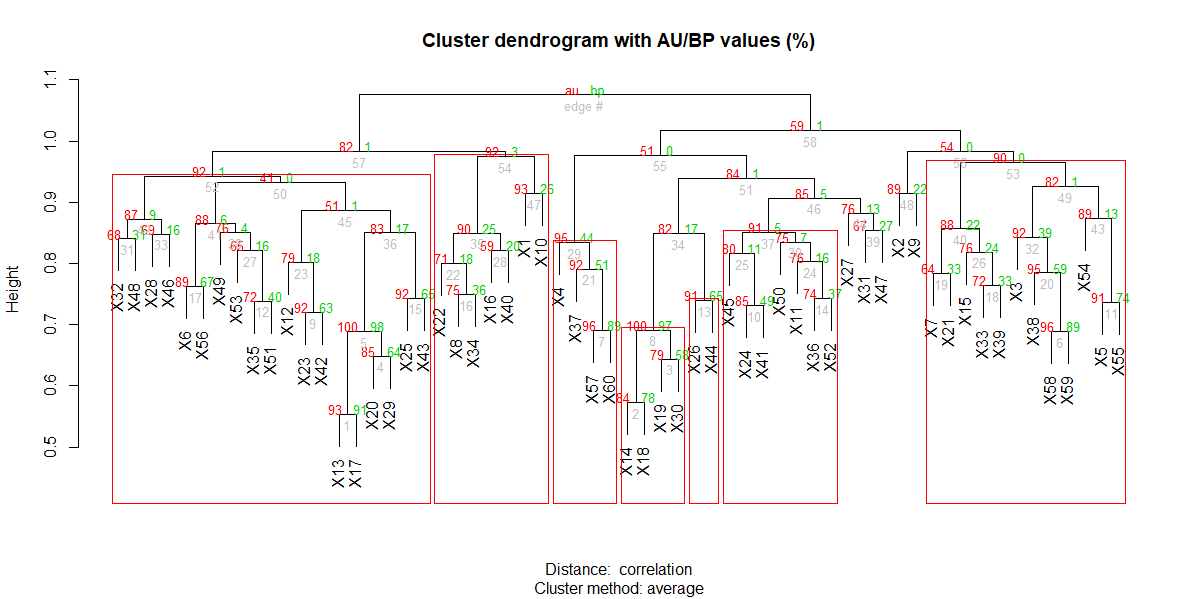


Kümeleme yöntemi ile homojen grupların belirlenmesinde çalışan ortalama bağlantı metodu (average linkage) seçilmiştir. Ortalama bağlantı kümeleme metodu, iki küme arasındaki uzaklığı; kümelerdeki nokta çiftleri arasındaki uzaklığın ortalaması olarak tanımlar ve küçük küme içi değişkenliğin fazla olduğu kümeleri birleştirme özelliğine sahiptir.



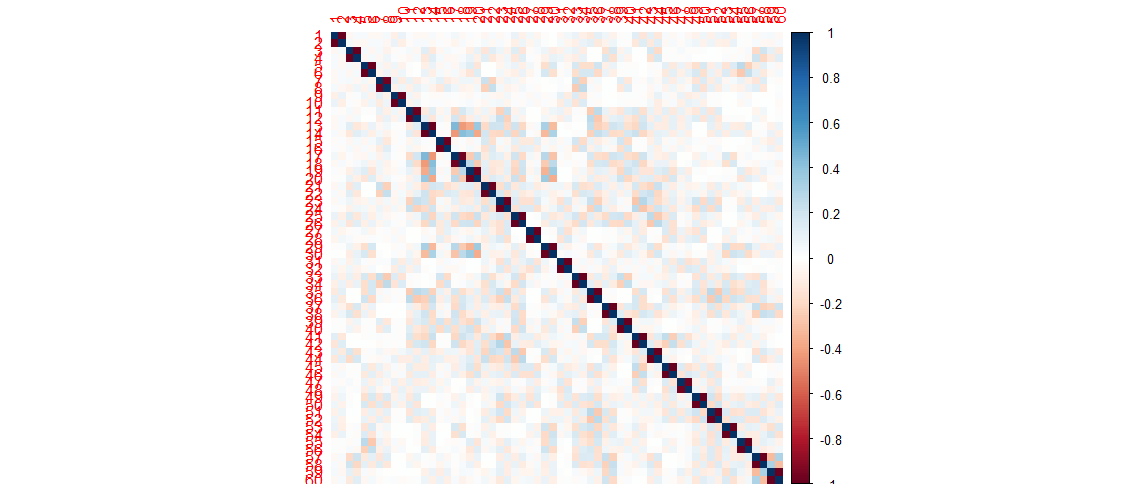
Bootstrap yeniden örnekleme metodu, örneklemin psikolojik ölçeklerde örneklemin küçük olduğu durumlarda kullanılmaktadır. Grupların genellenebilirliği ve istatistiksel anlamlılığını ölçmek için AU (Approximately Unbiased) ve BP (Bootstrap Probability) değerleri hesaplanarak aşağıdaki dendrogam elde edilmiştir. AU değerli multiscale bootstrap yöntemi olarak normal bootstrap’a göre bayes olmayan p değerleri elde etmede daha güvenilir sonuçlar vermektedir.

Kırmızı çerçeve ile işaretlenen kümeler p değeri 0.90 (alpha=0.10) olacak şekilde anlamlıdır.



Z3 ve Z5 boyutlarına ait değişkenler diğerlerinden kolaylıkla ayrışmaktadır. Fakat 1-4 ve 1-2-4 arasındaki ayrım anlamlı olmakla birlikte çok güçlü değildir. Bu kümeler birbirine yakın, gözlemlerin benzer davrandığı kümeler olabilir. Aynı zamanda kişilerin kendileri ile ilgili sorulara cevaplamada irrasyonel yaklaşımlarından kaynaklanan bayes de olabilir.

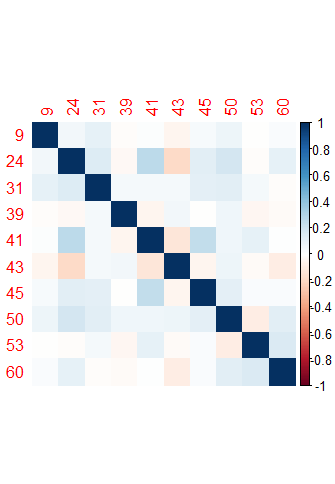
Değişkenlerin (ifadelerin) birbirleriyle olan ilişkilerini daha yakından incelemek için korelasyon matrisi kullandık.



Değişkenler arası korelasyonların sayısal değerlerini de içeren görsel… adresinde yakından incelenebilir.

İfadelerin boyutlara eğilimini ölçmedeki anlamlılığı için aynı boyuta ait ifadelerin pozitif yönlü güçlü ilişki içermesini, farklı boyutlardaki değişkenler arası ilişkinin de negatif yönlü güçlü olmasını bekliyoruz.

Z3, Z5 ve Z6 boyutlarına ait ifadelerin ayırt edici olduğunu %90 güvenle teyit ettik. Bu nedenle korelasyon matrisi kullanarak Z1, Z2 ve Z4 boyutlarına ilişkin tekli, ikili ve çoklu incelemeler yaptık.

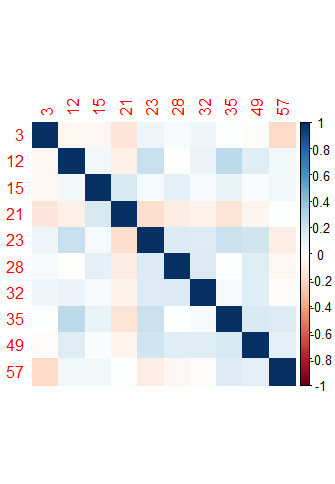


Z1 boyutundaki ifadeler arası korelasyon matrisi

Z1’e ait ifadeler arasında çoğunlukla pozitif yönlü bir ilişki olduğunu söyleyebiliriz. İlişki çok güçlü değil fakat bu seçeneklerden birine 1 yanıtı verenlerin diğerlerine de 1 yanıtı verme olasılığı var.

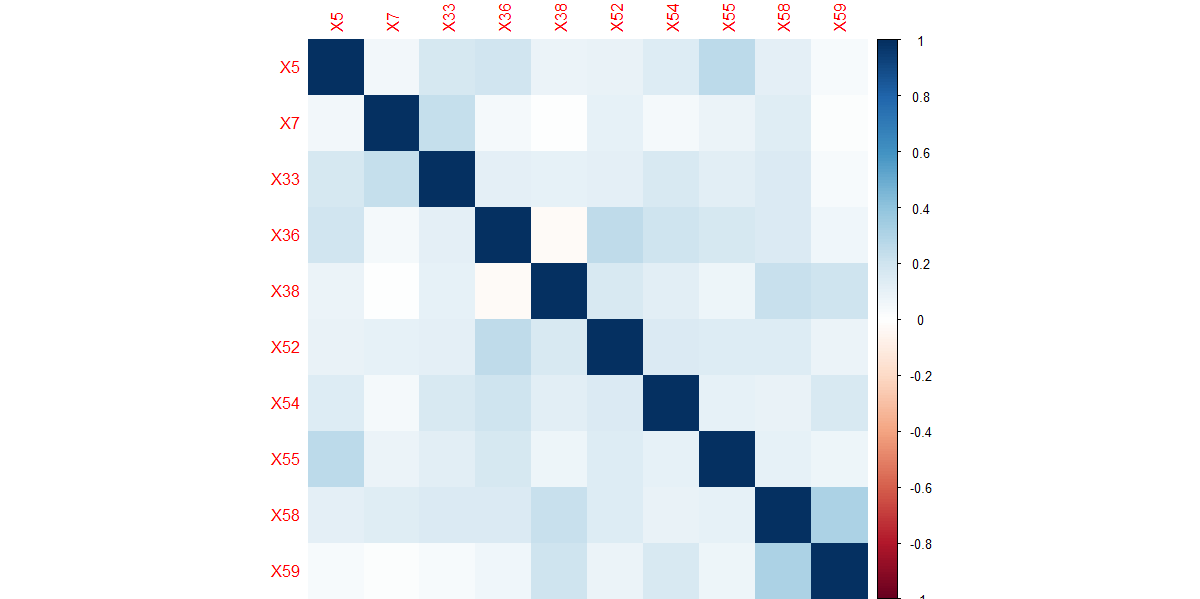
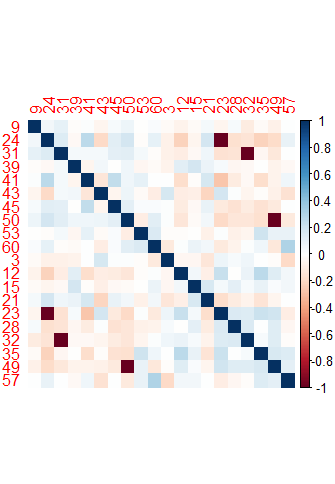
43. ifade ve 24. İfade arasında negatif yönlü bir ilişki var. İfadelerin her ikisi de Z1’de yer almasına rağmen birine evet yanıtı verenler diğerine hayır yanıtı vermeye meyillidir. 43. İfade boyutu ifade eden diğer değişkenlerle de negatif ilişki göstermektedir. Bu nedenle 43. İfadenin bir başka boyuta taşınması, ikili soruların kombinasyonunda yeni bir ifade ile değerlendirilmesi ya da kaldırılması önerilir.

İfadelerin bir başka boyuta taşınması ya da farklı ifade ile yer değiştirmesi ikili ve çoklu incelemelerde ele alınacaktır.



Z2 boyutundaki ifadeler arası korelasyon matrisi

Z2 boyutunda yer alan ifadeler birbiriyle çelişmektedir. Özellikle 3. 21. ve 57. İfade diğer gruplar ile negatif yönlü bir ilişkiye sahiptir. Bu ifadelerin boyutla ilişkisi tekrar değerlendirilmelidir.

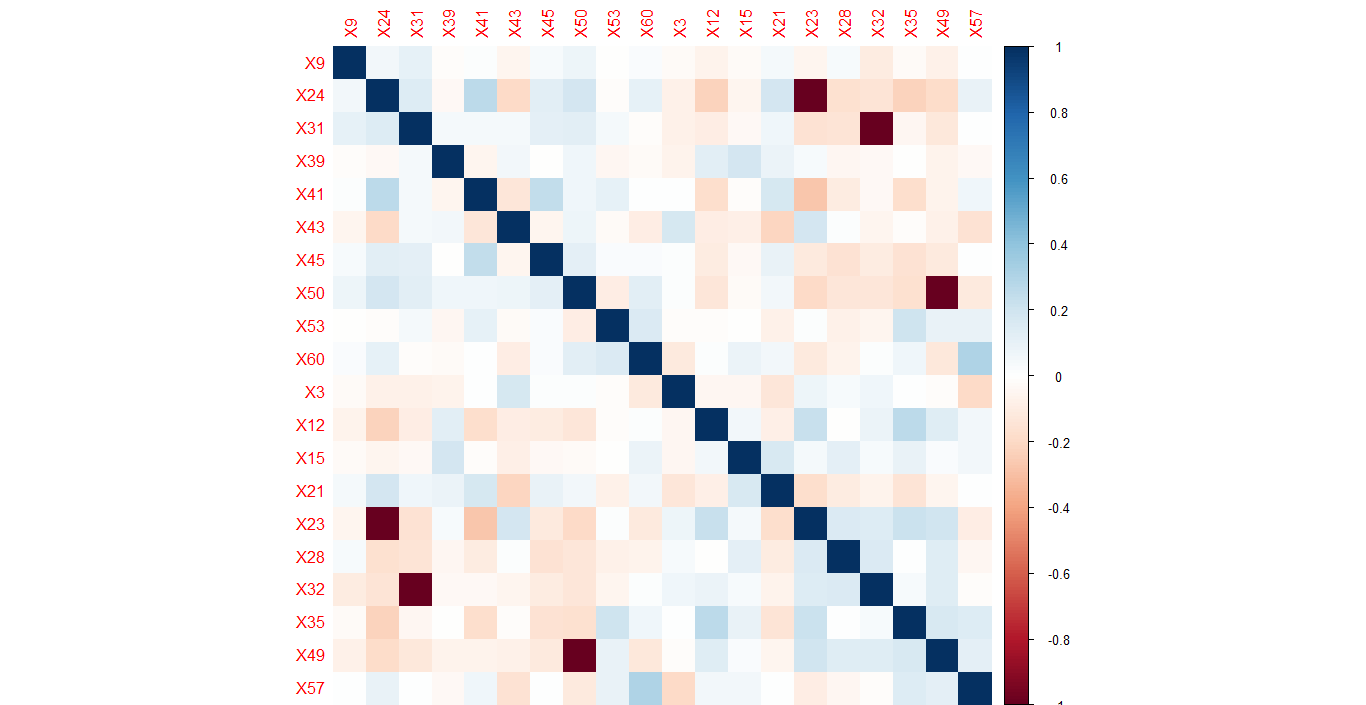


Z4 boyutunu temsil eden ifadeler arası negatif ilişki yoktur. Fakat pozitif ilişkiler de zayıftır. Bu nedenle ifadeler Z4 boyutuna düşecek gözlemi belirleme de zayıf kalabilir. Örneğin 38. İfade ile 36. İfade arasında ilişki yoktur. Bu ifadeler yeniden değerlendirilebilir. Nedenselliği sorgulanmalıdır.

Z4 boyutundaki ifadeler arası korelasyon matrisi

BOYUTLAR ARASI İFADELERİN AYRIŞTIRICI ETKİSİNİN İNCELENMESİ

Z1 ve Z2 boyutlarında yer alan ifadeler arası korelasyon matrisi



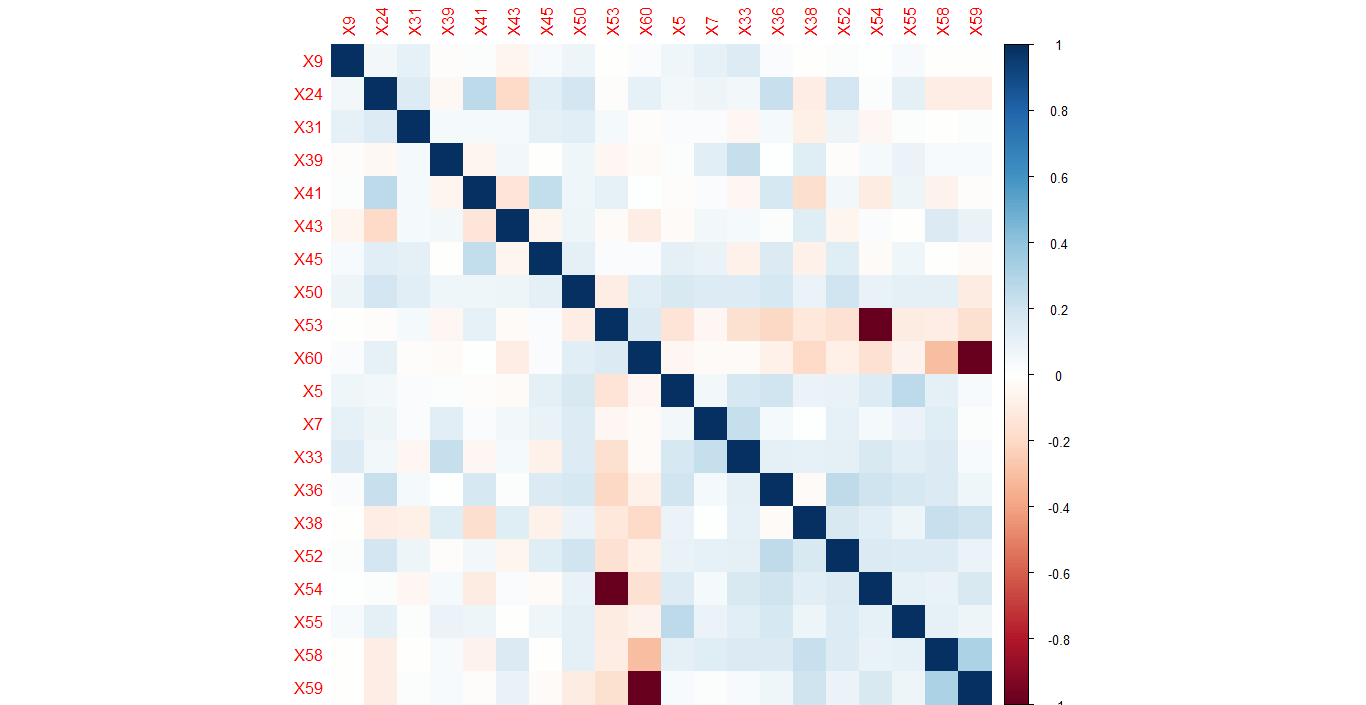
1-4

2-4

1-2-4

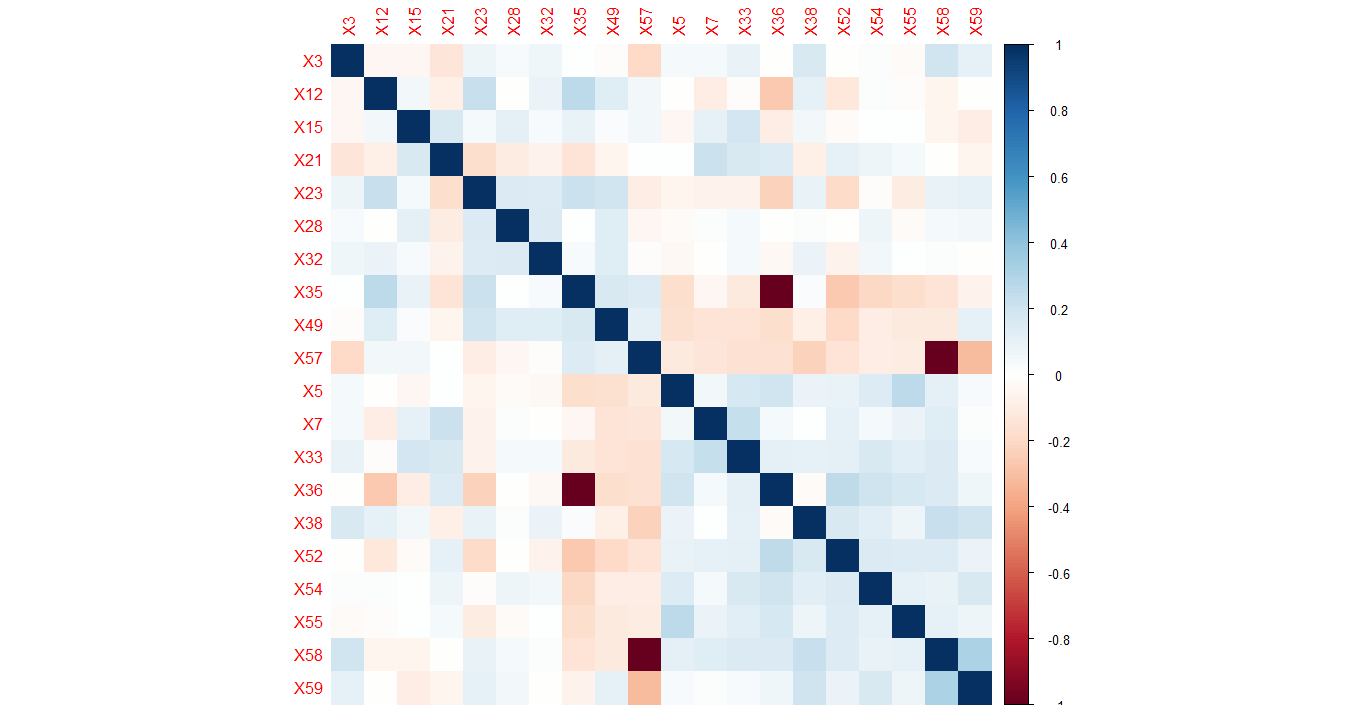
23-23, 31-31 ve 49-50 ifadeleri kişilerin Z1 ya da Z2 boyutuna eğilimi olduğunu anlamada etkin değişkenler fakat 39-15, 57-60 ifade çiftleri farklı boyutta yer alsa da iki boyutta da benzer davranış gözlemlenmektedir. Bu nedenle ayırt edici değildir.

Z1 ve Z4 boyutlarında yer alan ifadeler arası korelasyon matrisi

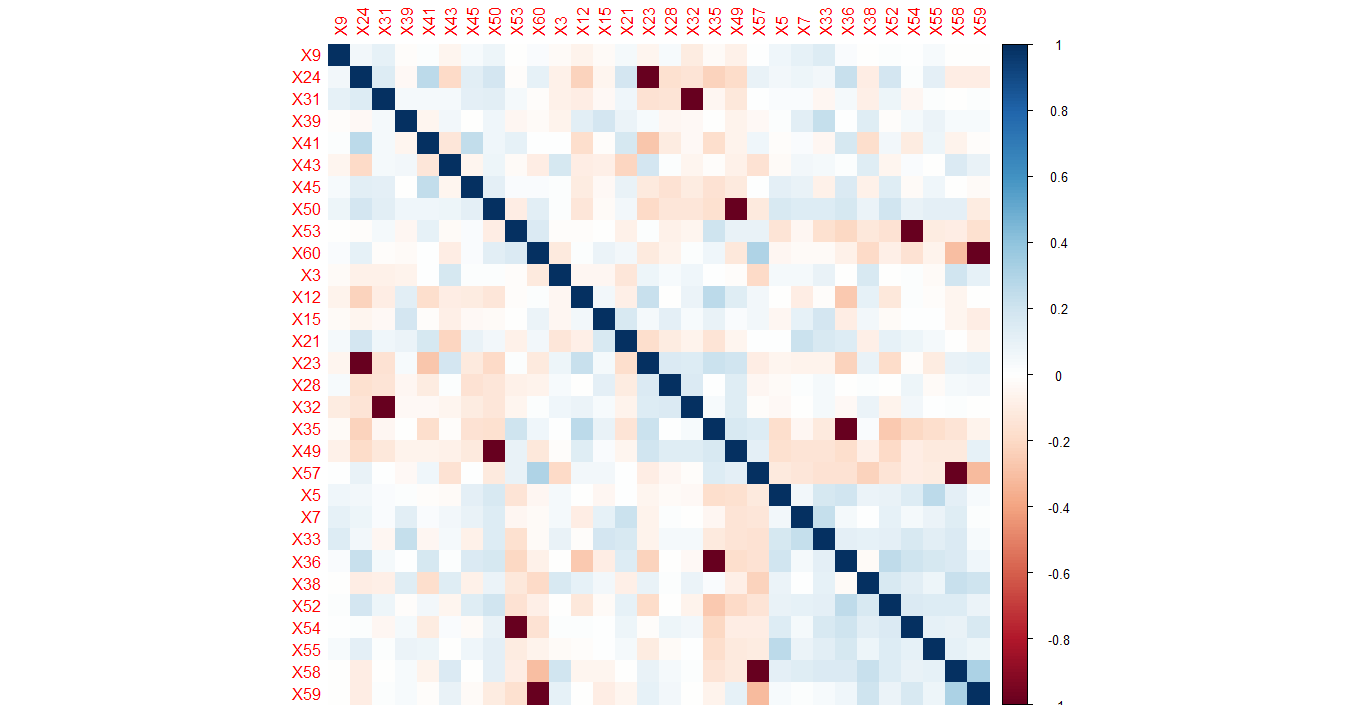


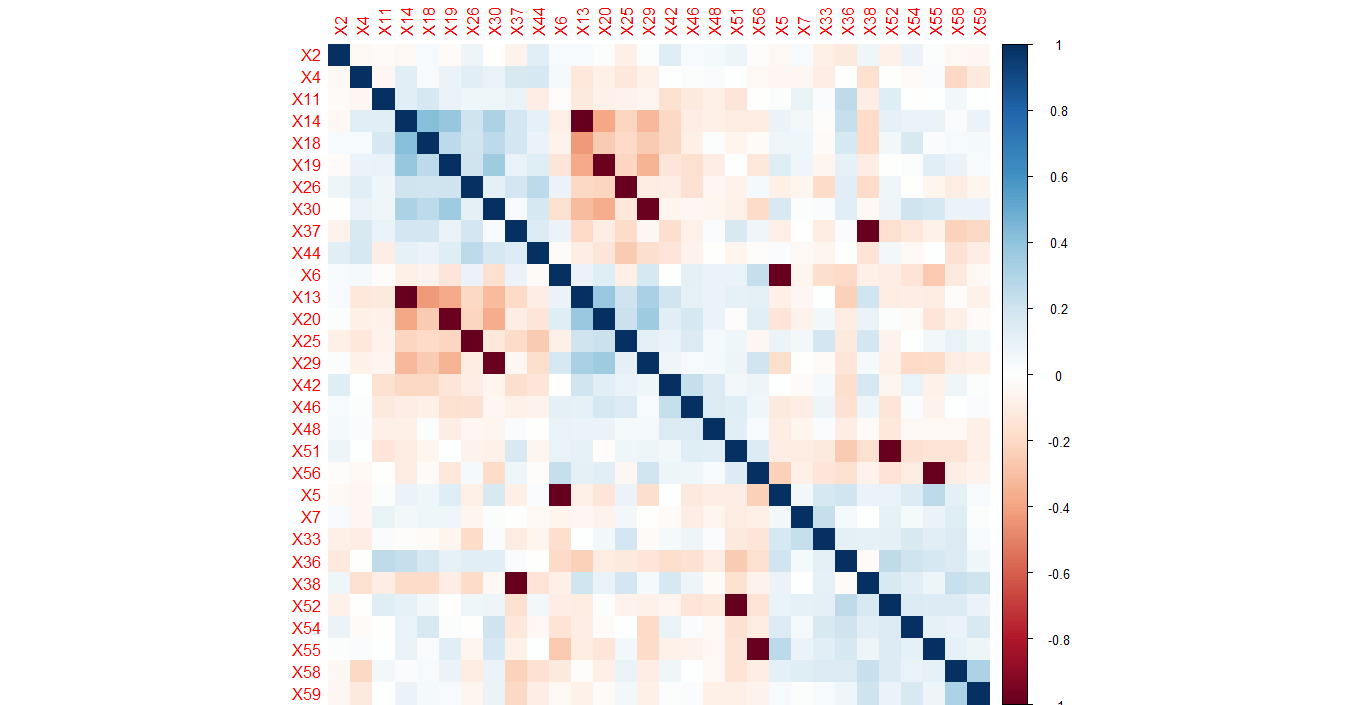
Z1 ve Z4 boyutlarına ilişkin ifadelerin ilişki matrisi verilmiştir. İfadeler arası korelasyona göre 54-55’e verilen yanıtlar ile 59-60 ifadelerine verilen yanıtlar gözlemin hangi gruba düştüğünü belirtmektedir. İfadelerin değerlendirme formu üzerinde kombinasyonu için ikili ve çoklu ilişkiler ele alınabilir. Örneğin, eğer xi. Gözlem kendi boyutunda negatif başka bir boyutta pozitif ilişki gösteriyorsa ve bu başka boyut içerisindeki değişken de xi.’nin yer aldığı boyutla pozitif, kendi boyutu ile negatif ilişki gösteriyorsa, ifadeler yer değiştirdiğinde, ifadelerin boyutlar arası ayırt edici gücü artacaktır.

Z2 ve Z4 boyutlarında yer alan ifadeler arası korelasyon matrisi



Boyut içinde ifadelerin pozitif, gruplar arası negatif ilişkiye sahip olmasını bekliyoruz. Yukarıda bahsedilen ifadelerin boyutlar arası yer değiştirmesi, nedenselliği, alan bilgisine aykırı değilse grupları ayırt etmede daha iyi sonuçlar verecektir. Z2-Z4 boyutlarının kıyaslanmasında bu çelişkiyi görüyoruz, Z2 boyutunda yer alan X3 değişkeni ile Z4 boyutunda yer alan X38 değişkeni yer değiştirebilir.

Z1, Z2, Z4 boyutlarında yer alan ifadeler arası korelasyon matrisi

Z5, Z6, Z4 boyutlarında yer alan ifadeler arası korelasyon matrisi

Z5, Z6 ve Z4 boyutları arası ifade ilişkilerini incelediğimizde, Z5 ve Z6 ifadelerinin kendi içinde pozitif yönlü ilişkiye ve gruplar arasında da negatif yönlü güçlü ilişkiye sahip olduğunu görüyoruz. Fakat Z4, Z1 ve Z2 ile olduğu gibi Z5 ve Z6 ile de karşılaştırıldığında ayırt edici gücü yüksek değil.

İkili soru grupları doğrudan grupları ayırmada doğrudan etkilidir. Bu nedenle eğer boyut x-boyut j ikili grubu belli gruplarda çoğunlukta ise bunun homojen dağılımı tekrar değerlendirilebilir.

HİYERARŞİK KÜMELEME ANALİZİ (HCA)

Hiyerarşik kümeleme analizi bir veri setinde yer alan kendi içinde homojen grupları tespit etmede kullanılır. Psikolojide kişilerin gruplara ayrılmasında, tıpta genlerin gruplarının belirlenmesinde ve hatta iş dünyasında müşterilerin segmentasyonunda kullanılmaktadır. Hiyerarşik kümeleme analizinde, gözlemlerin aldığı değerlere ve çeşitli index yöntemlere göre benzerlik matrisi oluşturup, grupları kümeleme yöntemleri ile tespit edilir. Kümelerin geçerliliği çeşitli istatistiksel yöntemlerle valide edilebilir.

Veri setimizi binary(ikili) değerler içerdiği için parametrelerimizi bunu dikkate alarak belirlememiz gerekmektedir.

BENZERLİK MATRİSİ (correlation based - Pearson)

Dissimilarity(uzaklık) matrisi oluştururken binary veri için Jaccard, Gower uzaklıkları ile Pearson korelasyon uzaklığı mevcuttur. Biz 3 uzaklık türünü de denedik ve Pearson diğerlerine göre daha iyi sonuç verdi.

(Jaccard/Gower nedir, açıkla)

KÜMELEME METODU (ward.D2 – min varyans gruplar içi)

Olan metodlar, bizim denediklerimiz ve neden ward.D2, çünkü data binary?

OPTİMUM KÜME SAYISI (wss/elbow, silhouette, dunn, gap statistic…)

Önerilen küme sayıları, 2-6-10 2 çünkü 5-6-3 vs 1-2-4, bundan sonraki minimum önerilen küme sayısı optimum 6’dır. İfadeleri de 6 boyutu temsil edecek şekilde girdiğimiz için, başarılı

DENDROGAM I – siyah beyaz

DENDROGAM II - renkli

DENDROGAM CIRCLIZED – renkli

AREAS & OVERLAPPING MAP (2 dim.)

CLUSTERING STATS(COUNT TABLE, SUMMARY STATS TABLE)

VALIDATION

CROSS VALIDATION TO DECIDE WARD.D2 with Rand İndex

INTERNAL VALİDATION, DUNN INDEX 0.35