

**T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İSTATİSTİK ANA BİLİM DALI**



**PARAMETRİK OLMAYAN ÇOK DEĞİŞKENLİ VARYANS
ANALİZİ VE BEYAZ EŞYA SEKTÖRÜNDE BİR UYGULAMA**

Yüksek Lisans Tezi

Serhat YILMAZ

Danışman

Prof. Dr. Yüksel TERZİ

SAMSUN
2021

TEZ KABUL VE ONAYI

Serhat YILMAZ tarafından, Prof. Dr. Yüksel TERZİ danışmanlığında hazırlanan “Parametrik Olmayan Çok Değişkenli Varyans Analizi ve Beyaz Eşya Sektöründe Bir Uygulama” başlıklı bu çalışma, jürimiz tarafından 7.7.2021 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Yüksek Lisans Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	İmza	Sonuç
Başkan	Prof. Dr. Muhammet BEKÇİ Sivas Cumhuriyet Üniversitesi İstatistik Anabilim Dalı		<input checked="" type="checkbox"/> Kabul <input type="checkbox"/> Ret
Üye (Danışman)	Prof. Dr. Yüksel TERZİ Ondokuz Mayıs Üniversitesi İstatistik Anabilim Dalı		<input checked="" type="checkbox"/> Kabul <input type="checkbox"/> Ret
Üye	Doç. Dr. Talat ŞENEL Ondokuz Mayıs Üniversitesi İstatistik Anabilim Dalı		<input checked="" type="checkbox"/> Kabul <input type="checkbox"/> Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

ONAY
... / ... / ...
Prof. Dr. Ali BOLAT
Enstitü Müdürü

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım Dönem Projesi tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

İmza
19/06/ 2021
Serhat YILMAZ

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı: Parametrik Olmayan Çok Değişkenli Varyans Analizi ve Beyaz Eşya Sektöründe Bir Uygulama

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 19.06.2021 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 19

Tek kaynak oranı : % 3 çıkmıştır.

İmza
19/06/2021
Prof. Dr. Yüksel TERZİ

ÖZET

PARAMETRİK OLMAYAN ÇOK DEĞİŞKENLİ VARYANS ANALİZİ VE BEYAZ EŞYA SEKTÖRÜNDE BİR UYGULAMA

Serhat YILMAZ

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

İstatistik Anabilim Dalı

Yüksek Lisans, Temmuz/2021

Danışman: Prof. Dr. Yüksel TERZİ

Perakende satış ve pazarlama alanında etkenlerin çoğunlukla birbirleriyle ilişkili birden fazla sonucu olabilir. Bu ilişkilerin ayrı ayrı analiz edilmesi I. Tip hatanın artmasına sebep olabilir. Ancak uygulamaların veri analizi kısmında bu hata sıklıkla yapılmakta, tek değişkenli analiz yöntemlerinden yararlanılmakta ve değişkenlerin önemli bir bölümü tek değişkenli model varsayımlarından olan normal dağılım ve varyansların homojenliği varsayımlarını sağlamamaktadır. Bu nedenle daha doğru ve güvenilir sonuç üretmesi açısından yeni metot ve yöntemlere ihtiyaç zorunlu hale gelmektedir. Bu çalışmada, değişkenler arasındaki ilişkileri dikkate alan ve varsayımları az olan parametrik olmayan PERMANOVA (Parametrik Olmayan Çok Değişkenli Varyans Analizi) modelinin teorik özellikleri anlatılmış, perakende satış ve pazarlama üzerine elde edilen bir veri setine uygulanarak sonuçlar tartışılmıştır. Çalışma sonucunda perakende satış mağazalarının bağlı oldukları satış grupları ile ürün grupları arasındaki ilişkileri, il-ilçe performansları, yıllık cirosal karşılaştırmaları incelenmiştir. İllerin iki yıllık kıyası yapıldığında cirosal farklılık çok olmamasına rağmen ürün grupları arasında yüksek oranda değişimler olduğu tespit edilmiştir. Birçok değişkenin birbirlerini etkileyen ilişkilerini birlikte dikkate alarak incelemek adına PERMANOVA modelinin kullanılması önerilmiştir.

Anahtar Sözcükler: PERMANOVA, Permütasyon, Yapay F değeri

ABSTRACT

NON-PARAMETRIC MULTIVARIATE ANALYSIS OF VARIANCE AND AN APPLICATION IN THE WHITE GOODS INDUSTRY

Serhat YILMAZ

Ondokuz Mayıs University
Institute of Graduate Studies
Department of Statistics

Master, July/2021

Supervisor: Prof. Dr. Yüksel TERZİ

There are more than one consequences of factors in retail sales and marketing. Analyzing these relationships separately might cause an increase in Type I error. Yet, in the data analysis of these applications, this error is frequently made. Univariate analysis methods are used, and a significant part of the variables do not provide the assumptions of normal distribution and homogeneity of variances being univariate model assumptions. Consequently, the need for new methods and methods becomes mandatory to produce more accurate and reliable results. This study tries to explain the theoretical aspects of the non-parametric PERMANOVA (Non-Parametric Multivariate Analysis of Variance) model by analyzing the relationships between the variables and few assumptions. As a result of this study, the relations between the sales groups and product groups of the retail stores, their provincial-district performances, annual turnover comparisons were investigated. When the two-year comparison of the provinces is made, this study examines that there are high changes between the product groups, despite the turnover difference is not much. As a result, this study suggests to use the PERMANOVA model in order to examine the relationships of many variables affecting each other together.

Keywords: PERMANOVA, Permutation, Artificial F value

ÖNSÖZ VE TEŞEKKÜR

Tezimin başlangıcını oluşturan, gerekli bütün yardım, tavsiye ve yönlendirmeleri yapan, bilgi ve deneyimlerinden yararlandığım, eğitimim boyunca olumlu katkı ve eleştirilerinden faydalandığım OMÜ Fen Edebiyat Fakültesi İstatistik Ana Bilim Dalı Bölüm Başkanı Sayın Hocam Prof. Dr. Yüksel TERZİ'ye şükranlarımı sunuyorum.

Çalışmalarım boyunca destekleriyle hayatımın hiçbir döneminde beni yalnız bırakmayan, kimi zaman sevgisinde boğulduğum, kimi zaman küçük bir tebessümü için gözlerinin içinde kaybolduğum çok kıymetli, sevgili eşim Elif YILMAZ'a, oğlu olmaktan her zaman gurur duyduğum anneme ve babama, sevgi dolu teşekkürlerimi sunuyorum.

Profesyonel iş hayatımda yolumu aydınlatan, maddî ve manevî desteklerini esirgemeyen, vizyonu ile ufkumu açan, satış tecrübesini yaşamama fırsat veren, her daim destekleyen, yönlendiren ve yolumu açan, bana kattığı değerleri ve tecrübeleri hiçbir zaman ödeyemeyeceğim, savaşçı ruhlu çok kıymetli büyüğüm, ağabeyim, Sayın Cevdet GÜRMAN'a tez çalışmamı ithaf ediyorum.

Serhat YILMAZ

İÇİNDEKİLER

TEZ KABUL VE ONAYI	i
ETİK BEYAN	ii
ÖZET	iii
ABSTRACT	iv
TEŞEKKÜR	v
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR	viii
ŞEKİLLER DİZİNİ	ix
TABLolar DİZİNİ	x
1. GİRİŞ	1
2. KURAMSAL TEMELLER VE KAYNAK ÖZETLERİ	3
3. MATERYAL VE YÖNTEM	6
4. ÇOK DEĞİŞKENLİ VARYANS ANALİZİ (MANOVA)	7
4.1. Model Kurulması.....	7
4.1.1. Wilks'in Lambda istatistiği.....	9
4.1.2. Pillai İz istatistiği	10
4.1.3. Hotelling-Lawley İz istatistiği	10
4.1.4. Roy'un en büyük özdeğere dayalı test istatistiği	11
5. PERMÜTASYONEL ÇOK DEĞİŞKENLİ VARYANS ANALİZİ	12
5.1. Tek Yönlü Analiz	12
5.2. İki Yönlü Analiz	15
5.3. Yöntemde Kullanılan Uzaklık Katsayıları	16
5.3.1. Öklid uzaklık ölçüsü	16
5.3.2. Manhattan uzaklık ölçüsü	17
5.3.3. Minkowski uzaklık ölçüsü	17
5.3.4. Ki-kare uzaklık ölçüsü	18
5.3.5. Bray Curtis uzaklık ölçüsü.....	18
5.3.6. Jaccard uzaklık ölçüsü	18
6. UYGULAMA	19
6.1. Normal Dağılım Testi.....	19
6.2. Tek Yönlü Modelin Oluşturulması.....	29
6.2.1. Öklid uzaklığı ile analiz çıktıları	29
6.2.2. Bray-Curtis uzaklığı ile analiz çıktıları.....	30

6.2.3. Manhattan uzaklığı ile analiz çıktıları	31
6.3. İki Yönlü Modelin Oluşturulması	35
6.3.1. İki yönlü analiz çıktıları (Bölge ve yıl).....	35
6.3.2. İki yönlü analiz çıktıları (Bölge ve çeyrekler)	36
6.3.3. İki yönlü analiz çıktıları (Bölge ve segmentler)	38
6.3.4. İki yönlü analiz çıktıları (Yıl ve çeyrekler)	40
6.3.5. İki yönlü analiz çıktıları (Yıl ve segmentler).....	41
6.3.6. İki yönlü analiz çıktıları (Çeyrek ve segmentler)	42
7. SONUÇ VE ÖNERİLER	44
KAYNAKLAR	46

SİMGELER VE KISALTMALAR

Λ	: Wilks'in Lambda istatistiđi
T	: Pillai İz istatistiđi
ϵ_{ij}	: Hata terimi
p	: F istatistiđi olasılık deđeri
F^π	: Yapay F istatistiđi
g	: Grup sayısı
W	: Çok boyutlu vektör
$D_{(i,j)}$: Uzaklık deđeri
d_{ij}	: Uzaklık ölçüsü
$S^{(j)}$: Jaccard uzaklık ölçüsü
x_{ik}	: i. gözlemin k. deđişken deđeri
d^*	: Çok boyutlu vektör
μ	: Grup ortalama deđeri
N	: Permütasyon sayısı
SS_T	: Genel kareler toplamı
SS_A	: Gruplar arası kareler toplamı
SS_W	: Grup içi karaler toplamı
GKT	: Genel kareler toplamı
GİKT	: Grup içi kareler toplamı
MANOVA	: Çok deđişkenli varyans analizi
PERMANOVA	: Permütasyonel çok deđişkenli varyans analizi

ŞEKİLLER DİZİNİ

Şekil 5.1. Ham veri ile mesafe matrisi grafiği	13
Şekil 5.2. Mesafe matrisinin gruplara ayrılmış grafiği	13
Şekil 6.1. Klima ürün grubu normal olasılık grafiği	20
Şekil 6.2. Dondurucu ürün grubu normal olasılık grafiği	21
Şekil 6.3. Televizyon ürün grubu normal olasılık grafiği	22
Şekil 6.4. Buzdolabı ürün grubu normal olasılık grafiği.....	24
Şekil 6.5. Çamaşır makinesi ürün grubu normal olasılık grafiği	25
Şekil 6.6. Bulaşık makinesi ürün grubu normal olasılık grafiği	26
Şekil 6.7. Fırın ürün grubu normal olasılık grafiği	27
Şekil 6.8. Bölgelerin ürün grubu üzerindeki dağılımı.....	33
Şekil 6.9. Bölgelerin segmentlere göre ürün grubu üzerindeki dağılımı	33
Şekil 6.10. Bölgelere göre ürün grubu üzerindeki yoğunluk dağılımı grafiği	34
Şekil 6.11. Segmentlere göre ürün grubu üzerindeki yoğunluk dağılımı grafiği.....	34
Şekil 6.12. Bölge ve yılların ürün grubu dağılım grafiği	36
Şekil 6.13. Bölge ve çeyreklerin ürün grubu dağılım grafiği.....	37
Şekil 6.14. Bölge ve segmentlerin ürün grubu dağılım grafiği	39
Şekil 6.15. Yıl ve çeyreklerin ürün grubu dağılım grafiği	40
Şekil 6.16. Yıl ve segmentlerin ürün grubu dağılım grafiği	42
Şekil 6.17. Çeyrek ve segmentlerin ürün grubu dağılım grafiği	43

TABLolar DİZİNİ

Tablo 5.1. İki yönlü analiz kareler toplamı	15
Tablo 6.1. Klima ürün grubu normallik test değerleri	20
Tablo 6.2. Derin dondurucu ürün grubu normallik test değerleri	22
Tablo 6.3. Televizyon ürün grubu normallik test değerleri	23
Tablo 6.4. Buzdolabı ürün grubu normallik test değerleri	24
Tablo 6.5. Çamaşır makinesi ürün grubu normallik test değerleri.....	26
Tablo 6.6. Bulaşık makinesi ürün grubu normallik test değerleri.....	27
Tablo 6.7. Fırın ürün grubu normallik test değerleri.....	28
Tablo 6.8. Ürün gruplarının çoklu normallik testi Mardia normal test değeri.....	28
Tablo 6.9. Tek yönlü model için Öklid uzaklığı çıktıları (N=10).....	29
Tablo 6.10. Tek yönlü model için Öklid uzaklığı çıktıları (N=50).....	29
Tablo 6.11. Tek yönlü model için Öklid uzaklığı çıktıları (N=500).....	30
Tablo 6.12. Tek yönlü model için Öklid uzaklığı çıktıları (N=1000).....	30
Tablo 6.13. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=10).....	30
Tablo 6.14. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=50).....	30
Tablo 6.15. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=500).....	31
Tablo 6.16. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=1000).....	31
Tablo 6.17. Tek yönlü model için Manhattan uzaklığı çıktıları (N=10).....	31
Tablo 6.18. Tek yönlü model için Manhattan uzaklığı çıktıları (N=50).....	31
Tablo 6.19. Tek yönlü model için Manhattan uzaklığı çıktıları (N=500).....	32
Tablo 6.20. Tek yönlü model için Manhattan uzaklığı çıktıları (N=1000).....	32
Tablo 6.21. Tek yönlü model için uzaklık ölçülerinin karşılaştırılması (N=1000)....	32
Tablo 6.22. Bölge ve yılların Öklid uzaklığı ile analiz çıktıları.....	35
Tablo 6.23. Bölge ve yılların Bray-Curtis uzaklığı ile analiz çıktıları.....	35
Tablo 6.24. Bölge ve yılların Manhattan uzaklığı ile analiz çıktıları.....	35
Tablo 6.25. Bölge ve çeyreklerin Öklid uzaklığı ile analiz çıktıları	36
Tablo 6.26. Bölge ve çeyreklerin Bray-Curtis uzaklığı ile analiz çıktıları	37
Tablo 6.27. Bölge ve çeyreklerin Manhattan uzaklığı ile analiz çıktıları	37
Tablo 6.28. Bölge ve segmentlerin Öklid uzaklığı ile analiz çıktıları	38
Tablo 6.29. Bölge ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları.....	38
Tablo 6.30. Bölge ve segmentlerin Manhattan uzaklığı ile analiz çıktıları	38
Tablo 6.31. Yıl ve çeyreklerin Öklid uzaklığı ile analiz çıktıları.....	40
Tablo 6.32. Yıl ve çeyreklerin Bray-Curtis uzaklığı ile analiz çıktıları.....	40
Tablo 6.33. Yıl ve çeyreklerin Manhattan uzaklığı ile analiz çıktıları.....	40
Tablo 6.34. Yıl ve segmentlerin Öklid uzaklığı ile analiz çıktıları.....	41
Tablo 6.35. Yıl ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları	41
Tablo 6.36. Yıl ve segmentlerin Manhattan uzaklığı ile analiz çıktıları.....	41
Tablo 6.37. Çeyrek ve segmentlerin Öklid uzaklığı ile analiz çıktıları	42
Tablo 6.38. Çeyrek ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları.....	42
Tablo 6.39. Çeyrek ve segmentlerin Manhattan uzaklığı ile analiz çıktıları	43

1. GİRİŞ

Türkiye, küresel beyaz eşya ihracatında, önde gelen ülkelerden birisi olup Avrupa'da lider konumdadır. Beyaz eşya sektörü her geçen yıl daha fazla ihracat sağlayan, büyüme rakamları en yüksek olan sektörlerden birisidir. Bir ülkenin gelişmesi ve büyümesinde üretim vazgeçilmez konumda yer alıyor. Beyaz eşya sektörü, iç pazarda tüketici ihtiyacını doğru yorumlama, kaliteli ve esnek üretim anlayışı, güçlü sanayi altyapısı ile Avrupa'nın ana tedarikçisi olma potansiyeli taşıyor. Gelişme sürecinin başarılı ilerlemesinde, Ar-Ge, garanti ve satış sonrası hizmetlerin büyük yeri olmasına rağmen ancak ve ancak pazarın iyi analiz edilip, tüketici ve pazarın ihtiyacına yönelik doğru ürünlere yönelik doğru stratejiler ile üretim yapılmasından geçer. Ürünlere yönelik talepler, geri bildirimler, satış adetleri ile birçok kaynak değerler birlikte analiz edilmesi için çok değişkenli istatistiksel yöntemlere ihtiyacın olduğunu göstermektedir.

Analizlerde genellikle deneysel faktör türlerinin, veriler üzerindeki etkilerinin bireysel etkinin yanı sıra aynı anda test edilmesine ihtiyaç duyulmaktadır. Böylelikle çok değişkenli varyans analizi son dönemlerde birçok alanda sıklıkla kullanılmaya başlanmıştır. Bunun en önemli sebebi bir değişkeni etkileyen birden fazla etken olmasıdır.

Çok değişkenli analizlerde ilgilenilen veri setlerinin birbirleriyle de mutlaka bir ilişkisi vardır. Bu ilişkiler göz ardı edilerek, tek değişkenli analizler ile değerlendirilmesi 1. tip hata yapma olasılığını arttırmaktadır. Bu gibi araştırmalarda değişkenlerin birden fazla ilişkili sonuçlarının olması durumunda, tutarlı sonuçlar veren tek değişkenli varyans analizinin genelleştirilmiş hali olan çok değişkenli varyans analizi yönteminin kullanılması gerekmektedir. Bu yöntem sonuç değişkenlerinin çok değişkenli normal dağılım ve kovaryans matrislerinin homojenliği gibi varsayımların sağlanması gerekmektedir. Sonuç değişkenlerinin çok değişkenli normal dağılım göstermesi, varyans ve kovaryans matrislerinin homojenliği, değişken sayısı fazla olan birçok alanda karşılaşılabilecek bir durum değildir. Değişkenlerin çok değişkenli normal dağılıma uymaması veya varyans kovaryans matrisinin homojen olmaması durumunda çok değişkenli varyans analizi tutarlı sonuçlar vermeyecektir. Değişken sayısının örneklem sayısından büyük olduğu çalışmalarda ise genellikle hesaplanamamaktadır. Parametrik olan bu yöntemin varsayımları olduğu için kullanılıp, kullanılmaması her çalışma için farklılık göstermektedir. Varsayımların

sağlanmaması durumunda parametrik olmayan yöntemler kullanılır.

Bu çalışmada, parametrik analizlerden olan MANOVA'nın teorik özellikleri, varsayımları anlatılmış ve bu varsayımların sağlanmadığı durumlar için parametrik olmayan analizlerden PERMANOVA'nın teorik özelliklerine değinilmiştir.

Uygulama bölümünde beyaz eşya sektöründe öncü firmalardan birinin, belirli bölgelerinden rastgele dört il seçilip, bu illerdeki mağazalarda gerçekleşen klima, derin dondurucu, buzdolabı, bulaşık makinesi, çamaşır makinesi ve televizyon satışlarının, bölgesel, yıllar ve üçer aylık çeyrekler ve ürün segmentlerine yönelik çok değişkenli analiz varsayımlarının sağlanıp, sağlanmadığı test edilmiştir.

Varsayımlar sağlanmadığı için parametrik olmayan analizlerden PERMANOVA analizi kullanılmış, bu analiz için tek yönlü modeller ile çok yönlü modeller kurulmuş olup, ilgili değişkenler için literatürde sıklıkla kullanılan uzaklık ölçüleri için test değerleri birbirleri ile karşılaştırılmıştır. Matematiksel işlemlerin çok fazla olması ve uzun sürmesinden dolayı paket programlarına ihtiyaç duyulmuştur. Anderson (2005), uzaklık ölçülerini kullanarak oluşturduğu matrisin özdeğerini kullanıp, PERMANOVA yöntemini kolaylaştıran bir paket programı önermiştir.

2. KURAMSAL TEMELLER VE KAYNAK ÖZETLERİ

Güvenilir sonuçlar vermesi sebebiyle, veri analizlerinde genellikle parametrik yöntemler kullanılır. Ancak varsayımların sağlanmaması durumunda parametrik yöntemlerin kullanılması doğru sonuçlar vermeyecektir.

Parametrik testlerin varsayımı sağlanmadığı durumda, varsayımları daha az olan parametrik olmayan testler kullanılmalıdır. Bu durumda parametrik olmayan yöntemlerde parametrik olan yöntemlerde olduğu gibi gruplar arası değişkenlere karşı grup içi değişkenlik karşılaştırılır. Parametrik yöntemlerde deney tasarımı kısıtı var iken parametrik olmayan yöntemlerde çoğunlukla deney tasarımı için herhangi bir sınırlama yoktur. Çalışma çaprazlama, iç içe geçmiş veya karışık bir deney tasarımı da içerebilir. Bu da yöntemin avantajları arasında yer almaktadır.

Çok değişkenli varyans analizinde, parametrik ve parametrik olmayan testlerin arasındaki farkları incelediğimizde; parametrik testlerin tutarlı sonuçlar vermesi ve yapılabilmesi için normallik varsayımı ve grup varyansların eşitliği istenmektedir. Bağımlı değişkenler arasındaki korelasyonlara hassas ve örneklemdaki birim sayısının değişken sayısından daha fazla olması şartını gerektirir, değişken sayısı örneklemdaki birim sayısından daha fazla olmamalıdır. Parametrik olmayan testlerde normallik sağlanması varsayımı olmayıp, gruplar arasındaki varyanslar homojen olmayabilir. Korelasyon hassaslığı gerektirmez. Kitle üzerindeki dağılım varsayımları için oldukça güçlüdür (Kıroğlu, 2001). Sonuç olarak, parametrik olan MANOVA testi gözlemlerin sıfır değer alması durumuna karşı aşırı derecede hassas iken parametrik olmayan PERMANOVA testi bu duruma karşı hassas ve duyarlı değildir (Pasin vd, 2016).

PERMANOVA testinin tek gerektirdiği varsayım, gözlemlerin bağımsız olduğunu ve dağılımlarının benzer olduğunu varsaymaktır. Bunun dışında testin herhangi bir varsayımı yoktur. Testin varsayımlarının az olması yöntemin kullanım alanlarını da artırmaktadır. Bilindiği üzere varsayımları çok fazla olan testlerin uygulanması ve gerçek hayatta varsayımların sağlanması oldukça güçtür. Bunun yanı sıra PERMANOVA testi çapraz, iç içe geçmiş veya karışık deneme düzenlerinde de kullanılabilir (Pasin vd, 2016).

Parametrik olmayan çok değişkenli veri analizleri incelendiğinde varsayım şartları olmamasına rağmen literatürde çok sayıda çalışma bulunmamaktadır. Bunun en büyük sebebi ise çok değişkenli verilerde varsayımların sağlanması durumunda verilere yönelik analizler ile varsayımların sağlanması durumuna getirilmesidir.

PERMANOVA analizi ile ilgili olarak literatürde bulunan çalışmalara bakıldığında parametrik analiz testleri varsayımları sağlanmadığı, özellikle sağlık, biyoloji ve pazarlama alanlarında kullanıldığı ve bunlardan bazıları aşağıdaki gibidir;

Anderson (2001a), çalışmasında varyans analizinin tek değişkenli veriler için güçlü olduğunu bununla birlikte çok değişkenli verilerde varsayımlarının çok katı olduğunu varsayımların sağlanmaması durumunda Permutasyon tekniklerine dayalı parametrik olmayan yöntemlerin kullanılması gerektiğinden bahsetmiştir. Yeni bir parametrik olmayan yöntem PERMANOVA yöntemini önermiştir.

Anderson ve Walsh (2013), yaptıkları çalışmada Ekolojide yaygın olarak kullanılan PERMANOVA, ANOSIM ve Mantel testinin benzerliklerinden bahsetmişler. Özellikle heterojenliğin etkilerini incelemek için tasarlanmış simülasyon çalışması yapmışlardır. Çalışma sonucunda PERMANOVA testi dipersiyonlardaki heterojenliğe karşı hassas olmadığı, ANOSIM ve Mantel testinin çok hassas olduğu tespit edilmiştir. Ayrıca PERMANOVA korelasyon yapısındaki farklılıklardan etkilenmediği gözlemlenmiştir.

Watson vd. (2007), balık türlerine yönelik bu çalışmasında, hedeflenen türde artışın olması hedeflenmeyen türlerin popülasyonundaki düşüşün sebebi araştırılıyor. Sığ ve derin bölgelerden alınan örneklerin parametrik test varsayımları sağlanmadığı için PERMANOVA analizi ile test edilip anlamlı farklılıklar tespit etmişler.

Tang vd. (2016), mikrobiyal toplulukların disbiyoz üzerindeki etkilerini incelemişler. Mesafeye dayalı yöntemlerin çoğu tek bir mesafe kullanır ve yanlış seçilir ise yanlış sonuçlar verdiğini söyleyerek, PERMANOVA analizi ile çoklu mesafeleri analize dahil ederek, yeni bir mesafe analizi yöntemi olan PERMANOVA-S'i önermişlerdir.

Pasin vd. (2016), Hashimoto hastalığı varlığı ve cinsiyet ile lipid profili değerleri arasındaki ilişkiler parametrik olmayan çok değişkenli analiz yöntemini kullanarak sonuçları değerlendirmiş ve PERMANOVA yöntemini kullanmışlardır.

Zhu vd. (2020), Mikrobiyom çalışmalarında eşleşen küme verileri üzerine PERMANOVA ile doğrusal ayrışma modelini (LDM) kullanılarak iki analizden hangisinin daha esnek olduğunu çıkarmışlar. İki analiz yönteminde eşleşmiş küme verilerinde esneklik sağladığı görülmüştür.

Peralta vd. (2012), bitki, mikrobiyal ve genel bakteri topluluklarına çevresel faktörlerin etkileri araştırılıyor. Elde edilen veri seti parametrik test varsayımlarını

sağlamadığı için PERMANOVA analizi ile incelenmiştir.

Kelly ve Gross (2015), Mikrobiyom çalışmalarında genellikle mikrobiyal gruplarının karşılaştırması üzerine çalışma yapıldığından, çevresel faktörlerinde etkileri olabileceği için bu etkilerinde çalışmaya eklenmesi ve çok değişkenli analizler yapılması gerektiğini vurgulamışlar. Yaptıkları çalışmada parametrik olmayan PERMANOVA analizini uygulamışlardır.

Gower ve Krzanowski (1999), birçok veri setinin çok değişkenli varyans analizi (MANOVA) yapısına uymasına rağmen varsayımları sağlanmadığı durumda kullanılamayacağını söylemişler. Bu nedenle veri setlerini bölünerek tek değişkenli varyans analizi ile incelemişler.

3. MATERYAL VE YÖNTEM

Tek yönlü MANOVA analizi birden fazla bağımlı değişkene tek bir bağımsız değişkenin etki ettiği durumlarda kullanılmaktadır. Yokluk hipotezinde bağımsız değişkenin ortalama vektörlerinin farklı olmadığı varsayılırken, alternatif hipotez de ise en az iki grubun ortalama vektörünün farklı olduğu şeklinde kurulur.

Bu hipotezler şu şekilde formüle edilir;

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_a$$

$$H_1 : \text{En az biri } \mu_i \neq \mu_j$$

Bağımlı değişken sayısının birden fazla ve bağımsız değişken sayısı ikiden fazla olduğunda ise iki yönlü MANOVA kullanılır. İki den fazla bağımsız değişken olduğunda ise Faktöriyel MANOVA testi kullanılır.

MANOVA testinin uygulanabilmesi için;

- Verinin çok değişkenli normalliğe uygun olması,
- Varyans-kovaryans matrisinin homojen olması,
- Çoklu doğrusal bağımlı ve teklik,
- Uç değerler ve doğrusallık,

Varsayımların sağlanması gerekmektedir. Ancak genellikle satış pazarlama alanında bu varsayımların sağlandığı veri seti çok kısıtlı sayıdadır. Dolayısıyla varsayımların sağlanmadığı durumda alternatif olarak parametrik olmayan PERMANOVA testi tercih edilmektedir. MANOVA gözlemlerin sıfır değeri almasına çok hassas iken, PERMANOVA bu duruma hassas ve duyarlı değildir. (Anderson, 2001b)

4. ÇOK DEĞİŞKENLİ VARYANS ANALİZİ (MANOVA)

Bir bağımlı değişkenin iki ve daha fazla bağımsız grubun ortalamalarının karşılaştırılması için tek yönlü varyans analizi kullanılır (ANOVA). Her bir grupta iki ve daha fazla bağımlı değişken var ise çok değişkenli varyans analizi (MANOVA) kullanılır. Varyans analizlerinde, bağımlı değişken sayısı bir tane olduğunda tek değişkenli varyans analizi, iki ve daha fazla olduğunda çok değişkenli varyans analiz teknikleri kullanılır (Terzi, 1995).

Varyans analizi bağımlı değişkenin bağımsız değişkenler tarafından açıklanmasıdır. Analizde F istatistiği muameleler arasında anlamlı bir farklılık olup olmadığı hakkında bilgi verebilmektedir. Sabit modelde, fark var ise farkın hangi gruptan geldiği çoklu karşılaştırmalar ile tespit edilebilmektedir. Bunlardan bazıları; Scheffe, Tukey, Newman-Keuls, Brown- Forsythe, Duncan testleridir (Koç, 2018).

Varyans analizi tek değişkenli ve çok değişkenli olmak üzere ikiye ayrılır. Çok değişkenli varyans analizi, tek yönlü, iki yönlü ve faktöriyel olmak üzere üçe ayrılır.

4.1. Model Kurulması

Çok değişkenli normal dağılım gösteren g popülasyon ile ortalama vektör hipotezlerinin test edilmesinde kullanılır. Tek yönlü MANOVA sıfır hipotezi g kitle ortalama vektörlerinin birbirine eşit olduğunu varsaymaktadır. Çok değişkenli varyans analizinde H_0 hipotezi ikiden fazla ortalama vektörün eşit olduğu şeklinde kurulur.

$$H_0: \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{p1} \end{bmatrix} = \begin{bmatrix} \mu_{1g} \\ \mu_{2g} \\ \vdots \\ \mu_{pg} \end{bmatrix}$$

$$H_1: \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{p1} \end{bmatrix} \neq \begin{bmatrix} \mu_{1g} \\ \mu_{2g} \\ \vdots \\ \mu_{pg} \end{bmatrix}$$

g: grup sayısı

p: her gruptaki deęişken sayısı

B : gruplar arası kareler ve çarpımlar toplamı matrisi

W : grup içi kareler ve çarpımlar toplamı matrisi

T : genel kareler ve çarpımlar toplamı matrisi

n_i : i. grubun gözlem sayısı

\bar{x} : genel ortalama

n: toplam gözlem sayısı

\bar{x}_i : i-inci gruba ait ortalama vektörü

S_i : i. grubun varyans-kovaryans matrisi

Şeklinde olmak üzere B ve W matrisleri ve test istatistiklerinin formülleri ile tek deęişkenli modelin genelleştirilmiş hali aşağıdaki gibidir.

$$x_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad j = 1, 2, \dots, n_i, \quad i = 1, 2, \dots, g$$

μ : tüm popülasyon ortalaması

τ_i : i-inci deneme etkisi

ε_{ij} : $N \sim (0, \sigma_e^2)$ dağılımına sahip bağımsız deęişkenlerdir.

$$B = \sum_{i=1}^g n_i (x_i - \bar{x})(x_i - \bar{x})' \quad (4.1)$$

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)' \quad (4.2)$$

$$T = B + W = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x})(x_{ij} - \bar{x})' \quad (4.3)$$

(4.1) numaralı denklem, gruplar arası kareler ve çarpımlar toplamı matrisi olup, grup ortalama vektörünün genel ortalama vektöründen farklı olup gruplar arasındaki etki farklılığını gösterir.

(4.2) numaralı denklem, hata kareler ve çarpımlar toplamı matrisi olup, her gruptaki gözlemler vektöründeki verilerin kendi grup ortalama vektöründen ayrılığı olup hataları belirtmektedir.

(4.3) numaralı denklem, genel kareler ve çarpımlar toplamı matrisidir.

Çok deęişkenli varyans analizinin matematiksel yönüne ait çalışmalar Anderson (1958), Timm (1975), Morrison (1976), Johnson ve Wichern (2007) gibi yabancı

kaynaklarda yer almaktadır. Bu kaynakların hemen hemen tamamında çok deęişkenli varyans analizinin modeli sabit etkili, rastgele etki modeli, karışık etki modeli olarak oluşturulmuştur (Terzi, 1995).

Çok deęişkenli varyans analizlerin varsayımları, varyans kovaryans matrisinin homojen olması ile her bir grubun normal dağılışı göstermesidir. Ortalama vektörleri arasında anlamlı bir farklılık olup olmadığını incelemekte kullanılan birçok test yöntemi vardır. Bu test yöntemlerine çoklu karşılaştırma testleri denir. Bu test istatistiklerinden dört tanesi aşağıda verilmiştir.

4.1.1. Wilks'in Lambda istatistięi

Wilks (1932) tarafından geliştirilmiş bu iki determinantın oranı genelleştirilmiş varyans oranı olan Λ 'dır. Box (1949) tarafından kullanılabilir hale getirilmiştir (Balibeyoęlu, 1989). Λ oranı varyans analizinde F testi istatistięi yerine kullanılmaktadır.

$$\Lambda = L = \frac{|W|}{|B+W|} \quad (4.4)$$

Bu oranın sifıra yaklaşması ortalama vektörlerde farklılık olduğunu göstermektedir. Eęer grup etkisi yoksa $B=0$, Λ deęeri 1'dir. İstatistik 1'e yakın deęer aldığıında H_0 hipotezi red edilemez. Eęer W matrisi B 'den küçük ise istatistik sifıra yaklaşır. Bu durumda H_0 hipotezi reddedilir. Λ ile T^2 arasında bir ilişki vardır.

$k=2$ olduğu durumda;

$$\Lambda = \frac{1}{1 + \frac{T^2}{N-k}} \quad (4.5)$$

Λ ile T^2 istatistikleri; p ve $N-p-1$ serbestlik dereceli F dağılımını göstermektedir.

$p=1$ olduğunda ise,

$$\Lambda = \frac{1}{1 + \frac{B}{W}} = \frac{B}{W} \frac{N-k}{k-1} = F \quad (4.6)$$

olduęu görülür. Wilks Λ istatistięinin, $\min(p, k-1) \leq 2$ olduğu özel durumlarda F dağılımları Tablo 4.1'de verilmiştir.

Tablo 4.1. Grup içi kareler toplamı tablosu

Değişken sayısı	Grup sayısı	F dağılımına yaklaşım	Serbestlik derecesi
$p = 1$	$g \geq 2$	$\left(\frac{n-g}{g-1}\right) \left(\frac{1-\Lambda}{\Lambda}\right)$	$g-1, n-g$
$p = 2$	$g \geq 2$	$\left(\frac{n-g-1}{g-1}\right) \left(\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\right)$	$2(g-1), 2(n-g-1)$
$p \geq 2$	$g = 3$	$\left(\frac{n-p-2}{p}\right) \left(\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\right)$	$2p, 2(n-p-2)$

Örnek genişlikleri büyük ise, Brattlet (1937) tarafından yapılan çalışmalar sonucunda Λ 'ın χ^2 dağılımı gösterdiği bilinmektedir (Johnson ve Wichern 2007).

4.1.2. Pillai İz istatistiği

$$T = \sum_{i=1}^s \frac{\lambda_i}{1+\lambda_i} \quad (4.7)$$

T istatistiği yardımıyla şu şekilde bulunur.

$$\frac{2\tilde{n}+s+1}{2m+s+1} \times \frac{T}{s-T} \quad (4.8)$$

(4.8) değeri $s(2m + s + 1)$ ve $s(2\tilde{n} + s + 1)$ serbestlik dereceleri ile F dağılımı gösterir (Alpar, 2013).

4.1.3. Hotelling-Lawley İz istatistiği

λ_i 'ler BW^{-1} matrisinin özdeğerleri olmak üzere, Hotelling iz istatistiği (T),

$$T = \sum_{i=1}^s \lambda_i \quad (4.9)$$

ile verilir. Denek sayısı yeterli iken,

$$nT = \chi_{[p(k-1); \alpha]}^2 \quad (4.10)$$

ise ortalama vektörleri arasında anlamlı bir farklılık olduğu söylenebilir. T istatistiğini test etmek için F dağılımından da yararlanılır. Test istatistiği şu şekilde bulunur (Alpar, 2013).

$$F = \frac{2(s\tilde{n}+1)}{s^2(2m+s+1)} T \quad (4.11)$$

Değeri $s(2m + s + 1)$ ve $s(2\tilde{n} + s + 1)$ serbestlik dereceleri ile F dağılımı gösterir.

4.1.4. Roy'un en büyük özdeğere dayalı test istatistiği

En büyük özdeğer λ_{max} ile gösterilirse, Roy'un en büyük özdeğere dayalı test istatistiği,

$$T = \frac{\lambda_{max}}{1+\lambda_{max}} \quad (4.12)$$

ile gösterilir. Bulunan test istatistiği s , m ve \tilde{n} parametrelili Heck grafik değeri ile karşılaştırılır. T istatistiğinin Heck grafik değerinden büyük olması durumunda ortalama vektörleri arasında anlamlı bir farklılık olduğu söylenebilir. Bazı kaynaklarda ve bilgisayar çıktılarında bu istatistik T yerine θ (teta) ile de gösterilir (Alpar, 2013).

5. PERMÜTASYONEL ÇOK DEĞİŞKENLİ VARYANS ANALİZİ

Parametrik testlerin varsayımlarının sağlanmadığı durumlarda, parametrik olmayan testlere ihtiyaç duyulmuştur. MANOVA testinin varsayımlarının sağlanmaması durumunda parametrik olmayan çok değişkenli varyans analizi olan PERMANOVA kullanılabilir (Anderson, 2001a).

Bu yöntem uzaklık veya benzerlik ölçülerinden yararlanmaktadır. Noktalar arasındaki uzaklıkların kareler toplamını, merkezlerin nokta sayısına bölünen ara noktaların uzaklıkların kareler toplamına eşit olmasıdır. F test istatistiği kullanılmakla beraber, p değeri hesaplanırken permütasyon teknikleri kullanılır. Bu sebeple bu yöntem dağılım varsayımı gerektirmemektedir (Anderson, 2001a).

Testin tek gerektirdiği varsayımı gözlemlerinin bağımsız olduğunu ve benzer dağılım gösterdiklerini varsaymaktadır. Bunun dışında herhangi bir varsayımı yoktur. Testin çok katı varsayımının olmaması kullanım alanlarını da arttırmaktadır.

5.1. Tek Yönlü Analiz

Modelde a tane grup ve her grupta n tane gözlem olduğunu düşünürsek toplam gözlem sayısı $N=an$ 'dir. Gözlemlerin i. ve j. bireylerinin gözlemler arasındaki uzaklık d_{ij} olmak üzere grup için kareler toplamı ve genel kareler toplamı aşağıdaki gibi olacaktır.

$$SS_T = \frac{1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij} \quad (5.1)$$

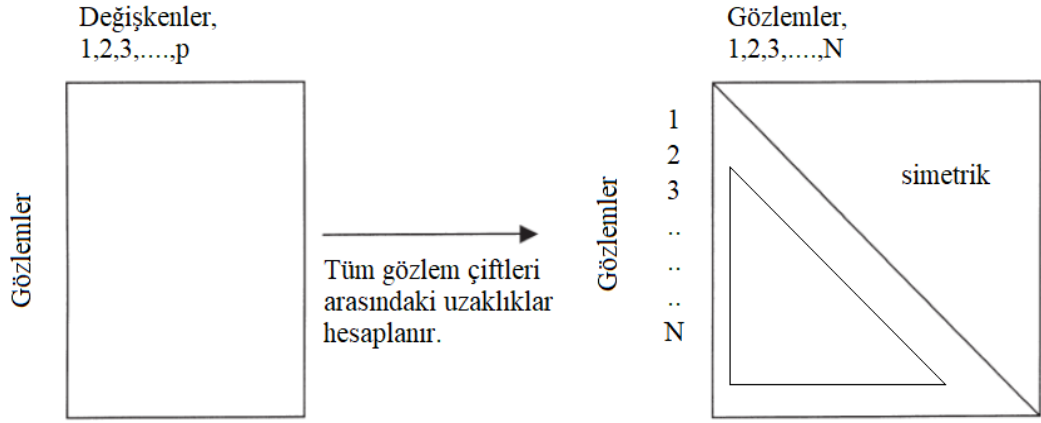
$$SS_W = \frac{1}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij}^2 \varepsilon_{ij} \quad (5.2)$$

$$SS_A = SS_T - SS_W \quad (5.3)$$

(5.1) numaralı denklem genel kareler toplamının formülüdür.

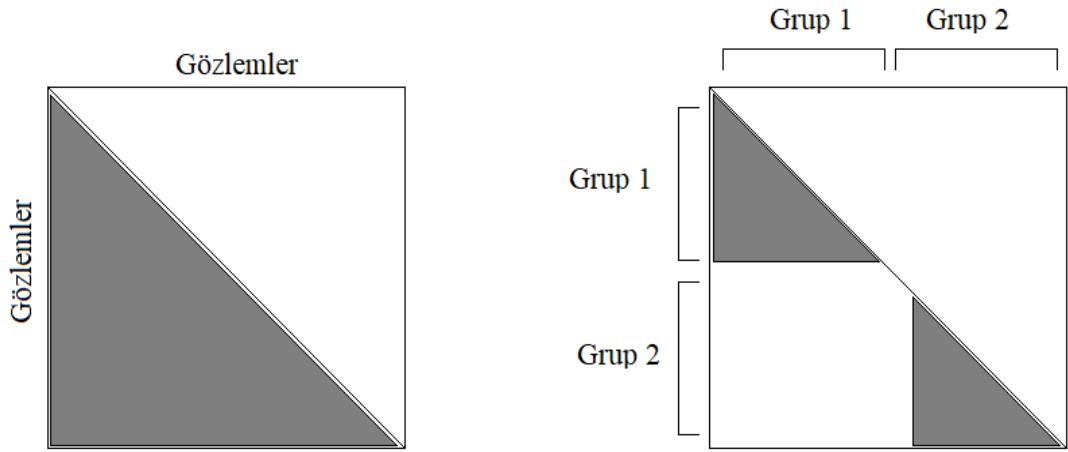
(5.2) numaralı denklem grup için kareler toplamının formülüdür.

(5.3) numaralı denklem gruplar arası kareler toplamının formülüdür.



Şekil 5.1. Ham veri ile mesafe matrisi grafiği

Genel kareler toplamı hesaplanırken uzaklık matrisinin alt köşegenine ait uzaklıkların kareleri toplanır ve N 'ye bölünür (Pasin, 2016).



Şekil 5.2. Mesafe matrisinin gruplara ayrılmış grafiği

Grup içi kareler toplamının hesaplanabilmesi için hata terimi dikkate alınmalıdır. Burada hata teriminin i . ve j . gözlemleri aynı grupta ise 1, farklı gruplarda ise 0 değerini alan bir katsayıdır. Bu katsayıya aynı grupta yer alan gözlemler arasındaki uzaklıkların kareleri eklenir Şekil 3.2.' deki gibi gösterilir (Pasin, 2016).

Gruplar arasındaki farklılıklar için kareler toplamı hesaplanması ise genel kareler toplamından grup için kareler toplamının çıkarılması ile bulunur.

F değeri hesaplanırken gruplar arası serbestlik derecesi $a-1$, grup için serbestlik derecesi ise $N-a$ olarak alınır. Bu bilgiler ile hesaplanan yapay F test istatistiği, Fisher'in F dağılımına uygunluk göstermediği için p değerinin elde edilebilmesi için bilinen F tablosu kullanılmaz. Bu PERMANOVA yöntemini diğer yöntemlerden ayıran en önemli özelliktir (Pasin, 2016).

Klasik F tablosu kullanılmayacağı için 1. Tip hata yani p değerinin hesaplanmasında permütasyon yöntemlerinden yararlanılmaktadır. Permütasyon yöntemi ile F değeri hesaplanacağı için kurulan sıfır hipotezinin doğru olduğu varsayılmaktadır. Ancak ve ancak gruplar arasında farklılıklar yok ise gözlem değerleri gruplara rasgele olarak dağılabilecektir. Gözlemler her adımda gruplara rasgele dağıtılarak belirli bir gruba tanımlanır. Permütasyonlar sonucu elde edilen F değeri F^π ile gösterilir. Bu rasgele elde edilen F^π değeri olası tüm sonuçlarla ilişkili satırların yeniden oluşturulmuş hali için hesaplanır. Grup sayısı a tane olmak üzere gruplarda n tekrar ile tek yönlü bir testte F^π istatistiğinin olası ayrı sonuçlarının sayısı $(an)!/(a!(n!)^a)$ kadardır. Sıfır hipotezinin doğruluğu varsayımı altında permütasyonlar sonucunda hesaplanan F değerinin dağılımı, orijinal satırlardan hesaplanan F değeri ile karşılaştırılır ve p değeri şu şekilde hesaplanır (Pasin, 2016).

$$p = \frac{\text{(Her bir } F^\pi \geq F \text{ koşulunun sağlanma sayısı)}}{\text{(Toplam } F^\pi \text{ lerin sayısı)}} \quad (5.4)$$

Denklemden F değeri permütasyonlar sonucunda elde edilen F^π dağılımının bir üyesi niteliğindedir. Tüm olası permütasyonların hesaplamak çok vakit alacağından pek tercih edilmemektedir. Bu nedenle p değeri olası permütasyonlardan rasgele seçilen bir alt seti kullanarak hesaplanabilir. Permütasyon sayısının fazla olması p değerlerinin duyarlılığını arttırdığı unutulmamalıdır. Çoğunlukla $\alpha = 0,05$ alınıp, permütasyon sayısı en az 1.000, $\alpha = 0,01$ için permütasyon sayısı 5.000 belirlenmelidir (Pasin, 2016).

PERMANOVA yönteminde üç farklı permütasyon yöntemi vardır. Bunlar satır veri, indirgenmiş model artıkları, tam model artıkları yöntemleridir. Örneklem hacmi, kullanılan veri çeşidi ve kullanım amacına göre dezavantaj veya avantajı vardır. Permütasyon yöntemlerinin açıklama ve avantajları şu şekildedir,

Satır veri yöntemi: Manly (1997) tarafından önerilen bu yöntem karmaşık ANOVA çalışmaları için tutarlı sonuçlar vermektedir. 1. Tip hata α değerine yakındır. Örneklem hacmi büyük ise daha tutarlı sonuç vermektedir. Lakin çok büyük örneklem hacmine gerek yoktur. (Pasin, 2016)

İndirgenmiş model artıkları yöntemi: Veri seti karmaşık bir yapıya sahip ise en iyi gücü ve en kesin 1. Tip hatayı vermektedir. Sonuçlar kavramsal kesin bir teste yakın olurlar. (Pasin, 2016)

Tam model artıkları yöntemi: Braak (1992) tarafından önerilen bu yöntemde tam modelin artık değerleri her bir gözlem değeri karşılık gelen satır ortalamasından çıkarılarak elde edilir. Hataların tahmin edilmesi ise her bir tekrar için ilişkili olan artıklardan elde edilir. Bu artıklar daha sonra permütasyonlanır ve test istatistiği sadece hatalar için hesaplanır (Pasin, 2016).

İndirgenmiş model ve tam model artıkları yöntemlerinde artıklara ihtiyaç olduğu için ortalamalardan gözlemler çıkarılarak yeni ortalamalar tahmin edilerek elde edilen artıklardır. Permütasyonlar sonucu elde edilen artıkların gerçek hatalara yakın olması için örneklem hacminin büyük olması gereklidir. Eğer örneklem hacmi küçük ise satır veri yöntemi önerilmektedir (Pasin, 2016).

5.2. İki Yönlü Analiz

İki faktörlü bir tasarımı ele alalım. a tane seviyesi olan birinci faktör A olarak adlandırılınsın. Aynı şekilde b tane seviyesi olan ikinci faktör B olarak adlandırılınsın. İki faktörün her ab kombinasyonundan n tekrar yapıldığı varsayılır ise gözlem sayısı $N=abn$ olacaktır (Pasin, 2016).

Tablo 5.1. İki yönlü analiz kareler toplamı

Değişim kaynağı	Kareler toplamı
A (B faktörü göz ardı edildiğinde)	$SS_{W(A)} = \frac{1}{bn} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij}^2 \varepsilon_{ij}^{(A)}$
B (A faktörü göz ardı edildiğinde)	$SS_{W(B)} = \frac{1}{an} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij}^2 \varepsilon_{ij}^{(B)}$
A	$SS_A = SS_T - SS_{W(A)}$
B	$SS_B = SS_T - SS_{W(B)}$
AB	$SS_{AB} = SS_T - SS_A - SS_B - SS_R$
Artık	$SS_R = \frac{1}{n} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij}^2 \varepsilon_{ij}^{(AB)}$

A faktörü göz ardı edilir ise B faktörleri için, B faktörü göz ardı edilir ise A faktörleri için hesaplama formülleri Tablo 5.1'deki gibidir

Tablo 5.1'deki formüllerde yer alan $\varepsilon_{ij}^{(A)}$ ve $\varepsilon_{ij}^{(B)}$ terimleri i ile j terimleri sırasıyla A ve B faktörlerinin aynı grubunda yer alıyorsa "1" yer almıyorsa "0" değerini almaktadır. $\varepsilon_{ij}^{(AB)}$ terimi ise i ve j gözlemleri A ve B faktörlerinin aynı kombinasyonunda yer alıyorsa "1" yer almıyorsa "0" değerini almaktadır. Analizde her bir ana etki için karışık gelen kareler toplamı A ve B faktörleri için SS_A ve SS_B

şeklinde gösterilmektedir. Artık kareler toplamı, A ve B faktörleri için her ab kombinasyonları içinde ara nokta uzaklıklar dikkate alınarak hesaplama yapılmaktadır ve SS_R şeklinde gösterilmektedir (Pasin, 2016).

Faktörlerin iç içe geçtiği modellerde etkileşim terimi olmadan aynı yaklaşım kuralları uygulanmaktadır. $SS_{B(A)} = SS_T - SS_A - SS_R$ formülü ile hesaplanır. Formülde yer alan $SS_{B(A)}$ terimi, B faktörünün A faktörü içine geçtiğini ifade etmektedir. Tek yönlü analiz yönteminde benzer şekilde hesaplanan F değerleri ile F tablo değerleri kullanılamaz ve p değerlerinin hesaplanmasında permütasyon yöntemleri kullanılmaktadır. Permütasyon yöntemlerinin çok iyi seçilmesi faktör sayısı ile doğrudan orantılıdır. Yanlış seçilen bir permütasyon yöntemi sonuçları da etkileyeceği için yorumlamanın hatalı olmasına yol açmaktadır (Pasin, 2016)

Permütasyon yöntemleri ile kesin bir p değeri elde etmek için sıra faktör sayısının az olması gerekmektedir. Sıra faktör sayısı arttıkça kesin bir p değeri elde etmek zor olacak ve çok fazla vakit almaktadır. Bu durumun avantaja dönüşmesi için kesin permütasyon testleri yerine analizde yer alan bütün terimlere karşılık gelen satır verilerin ve artıkların beraberce permütasyonunu içeren yaklaşımsal permütasyon yöntemi kullanılmalıdır (Pasin, 2016).

5.3. Yöntemde Kullanılan Uzaklık Katsayıları

PERMANOVA modelinde uzaklık ölçülerinden metrik ve yarı metrik uzaklık ölçüleri kullanılmaktadır. Bunlardan metrik uzaklık ölçüleri Öklid, Manhattan ve Minkowski iken yarı metrik olanlar ise ki-kare, Bray Curtis, Jaccard'tır. Bu ölçülerin benzer özellikleri aşağıdaki gibidir.

Metrik uzaklık özellikleri; i ve j herhangi iki nokta olmak üzere, $i=j$ ise i ve j arasındaki mesafeyi gösteren $D_{(i,j)} = 0$ 'dır. Her zaman pozitifdir (eğer $i \neq j$ ise $D_{(i,j)} > 0$ 'dır). Simetriktir ($D_{(i,j)}=D_{(j,i)}$). Üçgen eşitsizliği kuralı vardır.

Yarı metrik uzaklık özellikleri; i ve j herhangi iki nokta olmak üzere, $i=j$ ise i ve j arasındaki mesafeyi gösteren $D_{(i,j)} = 0$ 'dır. Her zaman pozitifdir (eğer $i \neq j$ ise $D_{(i,j)} > 0$ 'dır). Simetriktir ($D_{(i,j)}=D_{(j,i)}$).

5.3.1. Öklid uzaklık ölçüsü

Metrik uzaklık ölçülerinden olan Öklid uzaklık ölçüsü iki grup arasındaki benzerliklerin ölçüsüdür. Doğru standardizasyon sağlanmaz ise ölçü sağlıklı sonuçlar vermemektedir. Ölçümün bir üst limiti olmamakla beraber tanımlayıcı sayısı artar ise sonsuza kadar artmaktadır.

İki değişken olduğunda Öklid uzaklığı formülü,

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2} \quad (5.5)$$

Üç değişken olduğunda Öklid uzaklığı formülü,

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2} \quad (5.6)$$

Dolayısıyla p değişkenli bir yapıda i. ile j. gözlem arasındaki genelleştirilmiş Öklid uzaklığı aşağıdaki gibidir (Alpar, 2013).

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (5.7)$$

x_{ik} : i. gözlemin k. değişken değeri,

x_{jk} = j. gözlemin k. değişken değeri,

p: değişken sayısı

Eşitlikten anlaşılacağı üzere, Öklid uzaklığında etkili değişkenin diğer değişkene göre ölçüm biriminin yüksek olması gerekir. Her değişkenin uzaklığa katkısının aynı düzeyde olabilmesi için değişkenlerin yaklaşımlar ile standartlaştırılır (Alpar, 2013).

5.3.2. Manhattan uzaklık ölçüsü

Metrik uzaklık ölçülerinden olan Manhattan uzaklık ölçüsünün bir diğer adı City Block'tur. Doğru standardizasyon sağlanmaz ise ölçü önerilmemektedir. Öklid uzaklığına benzer özellikleri olup büyük değerli değişkenler tarafından yönlendirilebilmektedir.

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}| \quad (5.8)$$

Farkların mutlak değeri olarak tanımlanır. Kesikli sayısal veriler için önerilir. Öklid uzaklığı iki nokta arasındaki en kısa uzaklığı verirken, Manhattan uzaklık ölçüsü üçgen kenarlarının toplamı olarak düşünülebilir (Alpar, 2013).

5.3.3. Minkowski uzaklık ölçüsü

Metrik uzaklık ölçülerinden olan Minkowski uzaklık ölçüsü Öklid ve Manhattan uzaklık ölçülerinin bir farklı versiyonudur. Farklı sayıda veriler için değişkenlerin benzerliklerini karşılaştırmada etkili bir uzaklık ölçüsüdür.

$$d_{ij} = \left[\sum_{k=1}^p |x_{ik} - x_{jk}|^m \right]^{1/m} \quad (5.9)$$

Burada, m=1 değeri verilirse Manhattan uzaklık ölçüsü, m=2 değeri verilirse Öklid uzaklık ölçüsünü verir.

5.3.4. Ki-kare uzaklık ölçüsü

Yarı metrik uzaklık ölçülerinden olan ki-kare uzaklık ölçüsü bir benzerlik ölçüsüdür. Değişkenler sayılabilir olmalıdır. Öklid uzaklığındaki farkların $1/s_k^2$ ile standartlaştırılması ile elde edilir.

$$d_{ij} = \sqrt{\sum_{k=1}^p \frac{1}{s_k^2} (x_{ik} - x_{jk})^2} \quad (5.10)$$

5.3.5. Bray Curtis uzaklık ölçüsü

Yarı metrik uzaklık ölçülerinden olan Bray Curtis uzaklık ölçüsü Manhattan uzaklık ölçüsünün bir farklı versiyonudur. Bu ölçü değişkenlerdeki sıfır değerlerini göz ardı etmektedir (Adek ve Quran, 2015). Sıfır ile bir arasında sonuçlar vermektedir.

$$d_{ij} = \frac{\sum_{k=1}^p |x_{ik} - x_{jk}|}{\sum_{k=1}^p x_{jk} + \sum_{k=1}^p x_{ik}} \quad (5.11)$$

x_{ik} : i. gözlemin k. değişken değeri,

x_{jk} = j. gözlemin k. değişken değeri,

p: değişken sayısı

5.3.6. Jaccard uzaklık ölçüsü

Yarı metrik uzaklık ölçülerinden olan Jaccard uzaklık ölçüsü Bray Curtis uzaklığının modifiye edilmiş halidir. Formülünden de anlaşılacağı üzere bu uzaklık ölçüsü sıfır değerini alan değişken bir değerini alan değişkenden daha az bilgilendiricidir. Değişkenlerin yokluğu-varlığına dayalı ikili bir asimetric katsayıdır.

$$S^{(j)}(d, d^*) = \frac{d \cap d^*}{d \cup d^*} \quad (5.12)$$

d ve d^* birbirinden farklı çok boyutlu iki vektörü temsil eder. Genelleştirilmiş Jaccard ölçüsü ise şu şekilde gösterilebilir (Işık, 2008).

$$S^{(j)}(d, d^*) = \frac{d \cdot d^*}{|d|^2 + |d^*|^2 - d \cdot d^*} \quad (5.13)$$

6. UYGULAMA

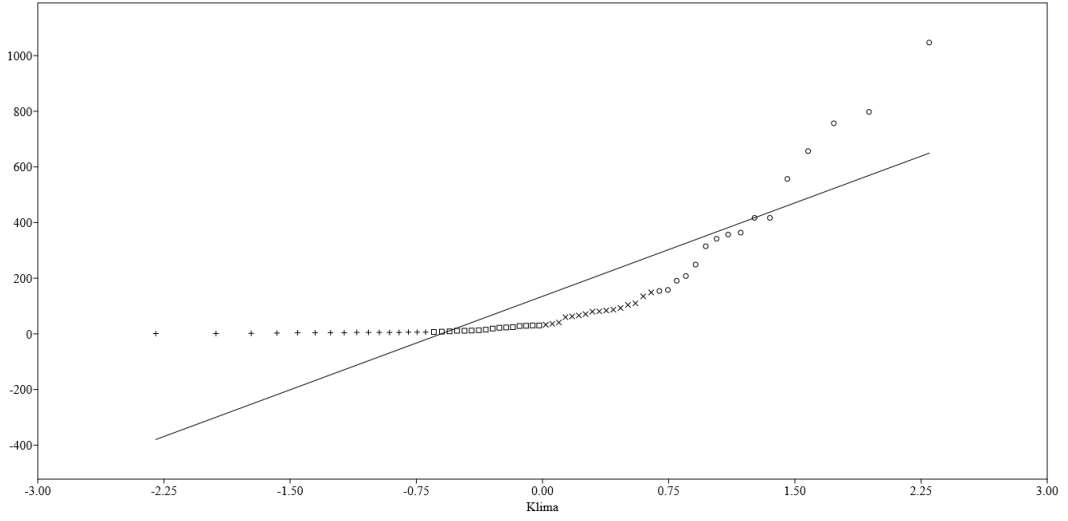
Beyaz eşya sektöründe ihracatın ve iç pazarda bayilik yapısının mutlak yeri çok önemli olmakla beraber her geçen gün daha da önemli hale gelmiştir. Oldukça rekabet olan bu pazarda bayilerin devamlılığı ile yeni pazarlama stratejileri ve birçok etkenle doğrudan ilişkili olup büyüme ve devamlılıklarını sağlamak için günbegün satışlarını arttırması gerekmektedir.

Türkiye pazarında öncü bir firmanın 2019 ve 2020 yıllarındaki perakende satış performansları incelenerek bölgesel çıkarımlarda bulunma, satış performansları ve bunun sunucunda pazarın ne ölçüde yanıt verdiği dönemsel ve ürün grupları bazında karşılaştırmalar yapılacaktır. Türkiye’de dört bölge müdürlüğü ve 1.200 bayi ile hizmet veren firmanın bölgelerine bağlı rasgele dört il seçilip, bu illerdeki bayilerin perakende satışları ürün grubu ve dönemsel olarak incelenmek istenmiştir. Seçilen bu değişkenlerin alt verileri için kodlama yapılmıştır. Bölge değişkeni dört farklı ili gösteriyor olup, 1 İç Anadolu, 2 Orta Anadolu, 3 Karadeniz, 4 Doğu Anadolu’dur. Yıl değişkeni iki yıl dönemini gösteriyor olup 1 2019, 2 2020’dir. Çeyrek değişkeni dört alt dönemi ifade ediyor olup, Q1 Ocak Şubat Mart, Q2 Nisan Mayıs Haziran, Q3 Temmuz Ağustos Eylül, Q4 Ekim Kasım Aralıktır. Segment değişkeni ise 1 orta segment 2 üst segment ürün satışlarını göstermektedir.

Elde edilen verilerin normal dağılım göstermediği ve varyans-kovaryans matrisinin homojen olmamasından dolayı parametrik olmayan ve literatürde oldukça çok az çalışması olan PERMANOVA yöntemi kullanılmıştır. İşlemlerin matematiksel işlemler ile çok uzun sürmesinden dolayı PAST istatistik paket programı kullanılmıştır.

6.1. Normal Dağılım Testi

İlgilenilen değişken sayısı arttıkça çoklu normalliğin sağlanması çok mümkün değildir. Normalliğin sağlandığı durumlarda parametrik olan analizler daha doğru sonuçlar verebilir. Normallik sağlanması ya da sağlanmaması durumunda parametrik olmayan testler yapılabilir. PERMANOVA analizinin normallik varsayımı olmayıp, yine de tekli ve çoklu normallik testleri incelenecek, sağlanmaz ise parametrik testlerin kullanılmaması gerektiği gösterilmiş olacaktır. Ürün gruplarının olasılık grafikleri ve test istatistikleri ile normallik durumları incelenmelidir.



Şekil 6.1. Klima ürün grubu normal olasılık grafiği

Normal dağılım sağlanması için değerlerin düz çizgiye yakın ve sürekli dağılması gerekmektedir. Klima ürün grubunda veriler çizgiye yakın ve rasgele dağılmadığı için normallik sağlanmamıştır.

Tablo 6.1. Klima ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,915	0,139	0,516	0,162
2	0,908	0,109	0,525	0,153
3	0,820	0,005	0,929	0,014
4	0,835	0,008	1,035	0,007

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,646	0,000	4,644	0,000
2	0,711	0,000	3,754	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,814	0,004	1,222	0,002
2	0,694	0,000	2,043	0,000
3	0,684	0,000	2,036	0,000
4	0,685	0,000	2,373	0,000

Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,616	0,000	4,857	0,000
2	0,697	0,000	3,477	0,000

Hipotezler,

H_0 : Veriler normal dağılışı gösterir

H_1 : Veriler normal dağılışı göstermez

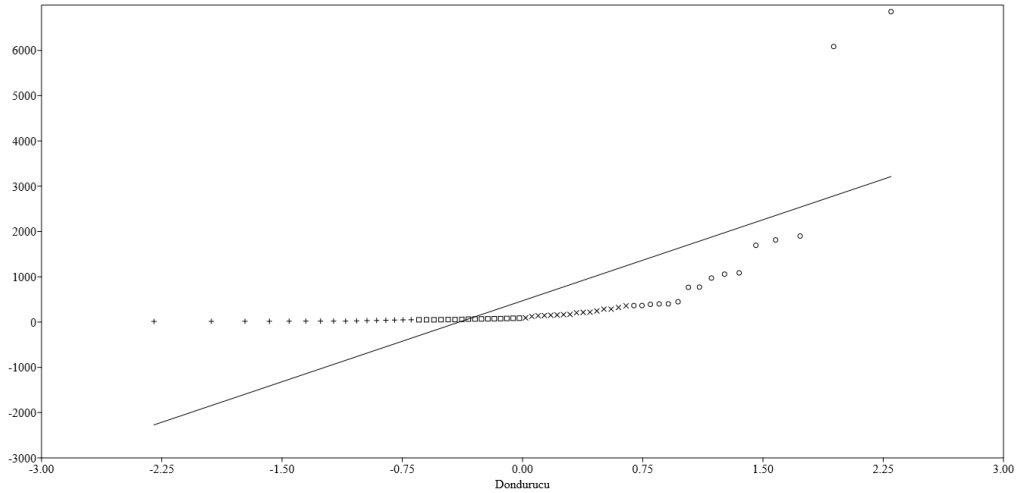
Bölgesel bakıldığında İç Anadolu ve Orta Anadolu bölgesi için $p > 0,05$ 'den büyük olduğu için H_0 hipotezi red edilemez. Veriler normal dağılım göstermektedir.

Doğu Anadolu ve Karadeniz bölgesi için $p < 0,05$ olduğundan H_0 hipotezi red edilir. Veriler normal dağılım göstermemektedir.

2019 ve 2020 yılları grubunda $p < 0,05$ olduğundan H_0 hipotezi red edilir. Veriler normal dağılım göstermemektedir.

2019 ve 2020 yılları üçer aylık dönemsel bakıldığında $p < 0,05$ H_0 hipotezi red edilir. Veriler normal dağılım göstermemektedir.

Segment grubu için $p < 0,05$ olduğundan H_0 hipotezi red edilir. Veriler normal dağılım göstermemektedir.



Şekil 6.2. Dondurucu ürün grubu normal olasılık grafiği

Normal dağılım sağlanması için değerlerin düz çizgiye yakın ve sürekli dağılması gerekmektedir. Dondurucu ürün grubunda veriler çizgiye yakın ve rasgele dağılmadığı için normallik sağlanmamış, bazı veriler çizginin çok uzağında olması normalliğin net sağlanmadığını göstermektedir.

Tablo 6.2. Derin dondurucu ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,611	0,000	2,646	0,000
2	0,615	0,000	2,580	0,000
3	0,719	0,000	1,871	0,000
4	0,738	0,000	1,633	0,000

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,373	0,000	7,091	0,000
2	0,423	0,000	6,488	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,628	0,000	2,441	0,000
2	0,547	0,000	3,087	0,000
3	0,592	0,000	2,842	0,000
4	0,621	0,000	2,590	0,000

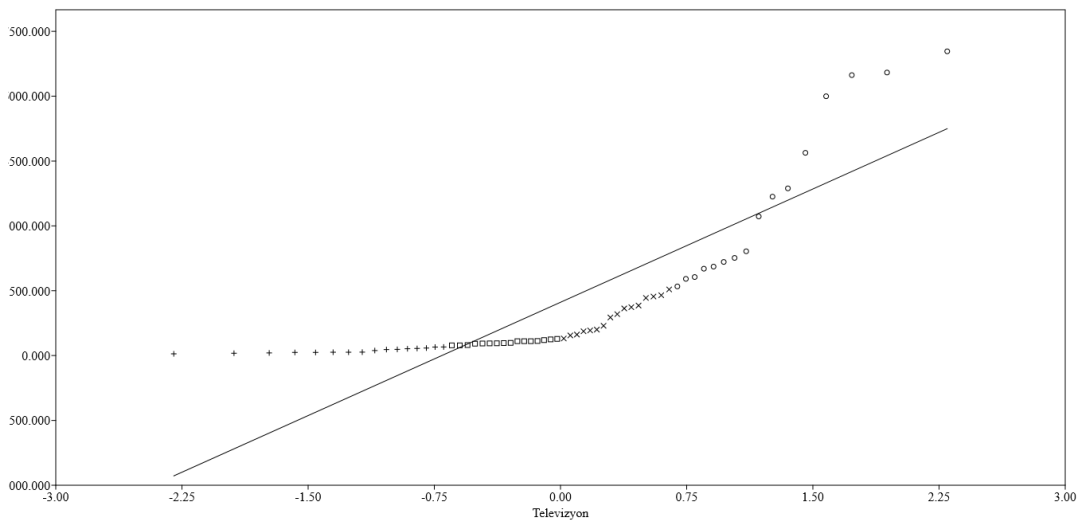
Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,607	0,000	4,320	0,000
2	0,509	0,000	5,904	0,000

Hipotezler,

H_0 : Veriler normal dağılış gösterir

H_1 : Veriler normal dağılış göstermez

Tüm dondurucu verileri için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Veriler normal dağılım göstermemektedir.



Şekil 6.3. Televizyon ürün grubu normal olasılık grafiği

Normal dağılım sağlanması için değerlerin düz çizgiye yakın ve sürekli dağılması gerekmektedir. Dondurucu ürün grubunda veriler çizgiye yakın ve rasgele dağılmadığı için normallik sağlanmamış, bazı veriler çizginin çok uzağında olması normalliğin net sağlanmadığını göstermektedir.

Tablo 6.3. Televizyon ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,881	0,041	0,695	0,056
2	0,852	0,014	0,947	0,012
3	0,867	0,024	0,731	0,045
4	0,770	0,001	1,051	0,007

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,687	0,000	3,820	0,000
2	0,680	0,000	3,763	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,649	0,000	2,123	0,000
2	0,707	0,000	1,689	0,000
3	0,730	0,000	1,666	0,000
4	0,689	0,000	2,036	0,000

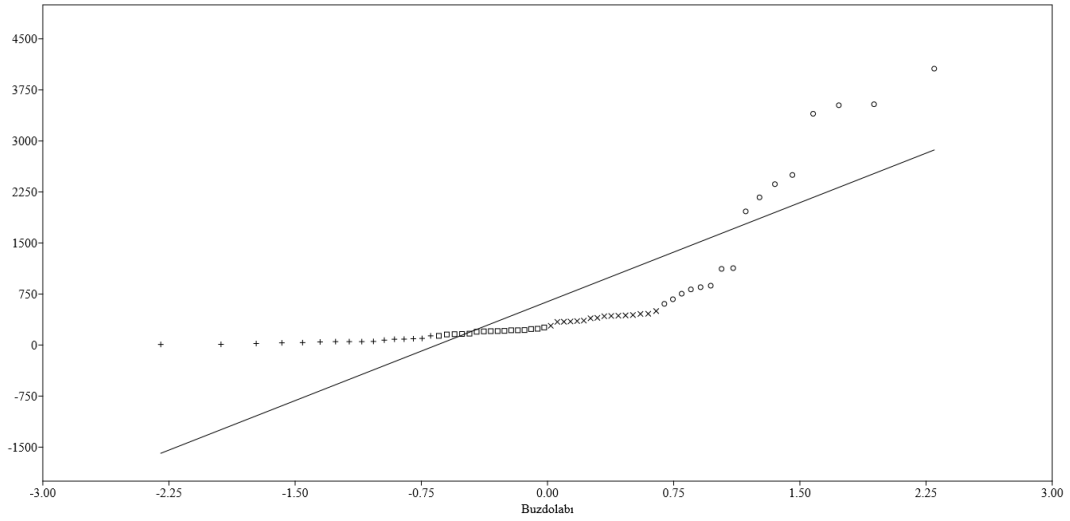
Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,782	0,000	2,665	0,000
2	0,712	0,000	3,917	0,000

Hipotezler,

H_0 : Veriler normal dağılım gösterir

H_1 : Veriler normal dağılım göstermez

Bölgesel bazlı bakıldığında İç Anadolu bölgesi için $p > 0,05$ olduğundan H_0 hipotezi red edilemez ve veriler normal dağılım gösterir. Diğer değişkenler için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Veriler normal dağılım göstermemektedir.



Şekil 6.4. Buzdolabı ürün grubu normal olasılık grafiği

Normal dağılım sağlanması için değerlerin düz çizgiye yakın ve sürekli dağılması gerekmektedir. Buzdolabı ürün grubunda veriler çizgiye yakın ve rasgele dağılmadığı için normallik sağlanmamış, bazı veriler çizginin çok uzağında olması normalliğin net sağlanmadığını göstermektedir.

Tablo 6.4. Buzdolabı ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,855	0,016	0,918	0,015
2	0,910	0,117	0,508	0,170
3	0,892	0,059	0,610	0,093
4	0,879	0,038	0,742	0,042

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,629	0,000	4,528	0,000
2	0,616	0,000	5,001	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,616	0,000	2,627	0,000
2	0,603	0,000	2,708	0,000
3	0,650	0,000	2,338	0,000
4	0,606	0,000	2,696	0,000

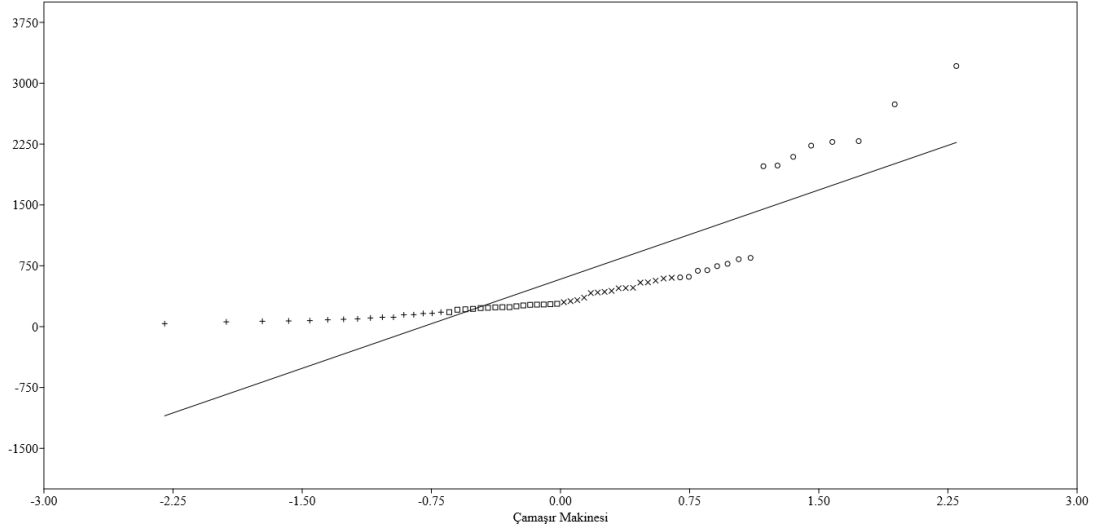
Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,819	0,000	1,713	0,000
2	0,746	0,000	3,393	0,000

Hipotezler,

H_0 : Veriler normal dağılış gösterir

H_1 : Veriler normal dağılış göstermez

Bölgesel bazlı bakıldığında Orta Anadolu ve Karadeniz bölgesi için $p > 0,05$ olduğundan H_0 hipotezi red edilemez ve veriler normal dağılım gösterir. Diğer deęişkenler için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Veriler normal dağılım göstermemektedir.



Şekil 6.5. Çamaşır makinesi ürün grubu normal olasılık grafięi

Normal dağılım sağlanması için deęerlerin düz çizgiye yakın ve sürekli dağılması gerekmektedir. Çamaşır makinesi ürün grubunda veriler çizgiye yakın ve rasgele dağılmadığı için normallik sağlanmamıştır.

Tablo 6.5. amařır makinesi rn grubu normallik test deęerleri

Blge	Shapiro-Wilk	p-deęeri	Anderson-Darling	p-deęeri
1	0,868	0,025	0,961	0,011
2	0,956	0,587	0,271	0,626
3	0,877	0,034	0,699	0,055
4	0,934	0,282	0,504	0,174

Yıl	Shapiro-Wilk	p-deęeri	Anderson-Darling	p-deęeri
1	0,631	0,000	4,806	0,000
2	0,700	0,000	3,722	0,000

eyrek	Shapiro-Wilk	p-deęeri	Anderson-Darling	p-deęeri
1	0,620	0,000	2,582	0,000
2	0,624	0,000	2,487	0,000
3	0,735	0,000	1,580	0,000
4	0,657	0,000	2,268	0,000

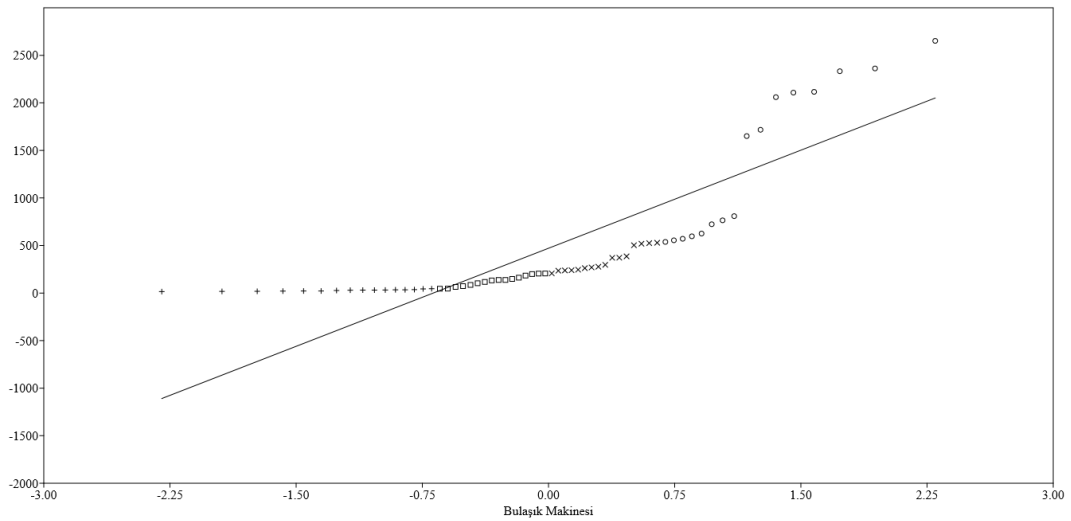
Segment	Shapiro-Wilk	p-deęeri	Anderson-Darling	p-deęeri
1	0,862	0,001	1,571	0,000
2	0,735	0,000	3,812	0,000

Hipotezler,

H_0 : Veriler normal daęılıř gsterir

H_1 : Veriler normal daęılıř gstermez

Blgesel bazda Orta Anadolu ve Doęu Anadolu $p > 0,05$ olduęundan H_0 hipotezi red edilemez. Veriler normal daęılım gstermektedir. Dięer veriler iin $p < 0,05$ olduęundan H_0 hipotezi reddedilir. Veriler normal daęılım gstermemektedir.



řekil 6.6. Bulařık makinesi rn grubu normal olasılık grafięi

Tablo 6.6. Bulaşık makinesi ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,836	0,008	1,133	0,004
2	0,825	0,006	1,228	0,002
3	0,843	0,011	0,972	0,011
4	0,819	0,005	1,233	0,002

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,659	0,000	4,248	0,000
2	0,669	0,000	4,125	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,640	0,000	2,393	0,000
2	0,652	0,000	2,262	0,000
3	0,729	0,000	1,629	0,000
4	0,622	0,000	2,500	0,000

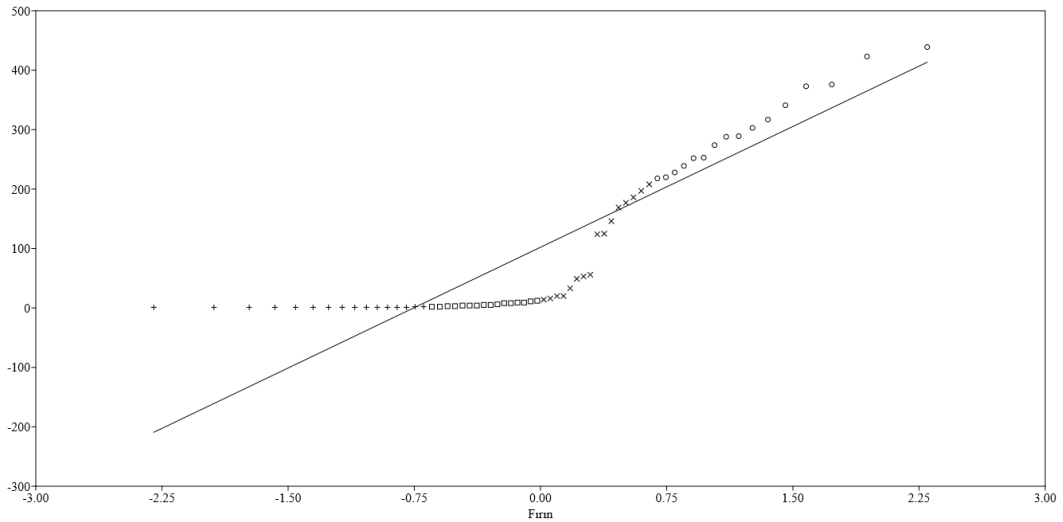
Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,762	0,000	3,254	0,000
2	0,702	0,000	3,954	0,000

Hipotezler,

H_0 : Veriler normal dağılış gösterir

H_1 : Veriler normal dağılış göstermez

Tüm veriler için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Veriler normal dağılım göstermemektedir.



Şekil 6.7. Fırın ürün grubu normal olasılık grafiği

Tablo 6.7. Fırın ürün grubu normallik test değerleri

Bölge	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,784	0,002	1,505	0,000
2	0,781	0,002	1,589	0,000
3	0,774	0,001	1,459	0,001
4	0,788	0,002	1,467	0,001

Yıl	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,768	0,000	3,046	0,000
2	0,740	0,000	3,758	0,000

Çeyrek	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,775	0,001	1,566	0,000
2	0,769	0,001	1,647	0,000
3	0,762	0,001	1,687	0,000
4	0,757	0,001	1,705	0,000

Segment	Shapiro-Wilk	p-değeri	Anderson-Darling	p-değeri
1	0,955	0,199	0,382	0,379
2	0,768	0,000	2,784	0,000

H_0 : Veriler normal dağılım gösterir

H_1 : Veriler normal dağılım göstermez

Segment grubunda üst segment ürünlerin $p > 0,05$ olduğu için H_0 hipotezi red edilemez, veriler normal dağılım göstermektedir. Diğer değişkenler için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Veriler normal dağılım göstermemektedir.

Tablo 6.8. Ürün gruplarının çoklu normallik testi Mardia normal test değeri

Parametre	Değer	İstatistik	s.d.	p-değeri
Çarpıklık katsayısı	76,56	816,70	84,00	0,00
Çarpıklık, düzeltilmiş	865,00	84,00	0,00	
Basıklık katsayısı	138,60	26,96	0,00	
Ep:	880,80			
p-değeri	0,00			

Hipotezler,

H_0 : Grup değişkenleri ve ürün grubu verileri birlikte normal dağılım gösterir

H_1 : Grup değişkenleri ve ürün grubu verileri birlikte normal dağılım göstermez

Bölgeler, 2019, 2020, çeyrekler ve segmentler ile ürün gruplarının birlikte çok değişkenli normalliği incelendiğinde $p < 0,05$ olduğu için H_0 hipotezi reddedilir ve çok değişkenli normal dağılımın sağlanmadığını söyleyebiliriz.

6.2. Tek Yönlü Modelin Oluşturulması

Veri grubumuzda bölgeler, yıllar, çeyrekler ve segmentler olmak üzere dört tane grup değişkenimiz mevcuttur. Bu değişkenlerin ürün grupları ile tek yönlü incelenmesi yapılarak farklılıkların olup olmadığına bakılacaktır.

Hipotezler şu şekilde kurulur,

H_0 : Grup faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur.

H_1 : Grup faktör seviyelerinde bağımlı değişkenden oluşan ortalama vektörlerinin en az ikisi arasında anlamlı bir farklılık vardır.

6.2.1. Öklid uzaklığı ile analiz çıktıları

Tablo 6.9. Tek yönlü model için Öklid uzaklığı çıktıları (N=10)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permüstasyon sayısı (N)	10	10	10	10
Genel kareler toplamı	391800000	391800000	391800000	391800000
Grup için kareler toplamı	216500000	388500000	371600000	326500000
F istatistiği	16,20	0,52	1,09	12,39
p değeri	0,09	0,73	0,64	0,09

Analizde metrik olan Öklid uzaklığı yöntemi kullanılmıştır. Permüstasyon sayısı 10 olarak belirlenmiştir. Tüm grup değişkenleri için $p > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur.

Tablo 6.10. Tek yönlü model için Öklid uzaklığı çıktıları (N=50)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permüstasyon sayısı (N)	50	50	50	50
Genel kareler toplamı	391800000	391800000	391800000	391800000
Grup için kareler toplamı	216500000	388500000	371600000	326500000
F istatistiği	16,20	0,52	1,09	12,39
p değeri	0,02	0,65	0,29	0,02

Tablo 6.11. Tek yönlü model için Öklid uzaklığı çıktıları (N=500)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	500	500	500	500
Genel kareler toplamı	391800000	391800000	391800000	391800000
Grup için kareler toplamı	216500000	388500000	371600000	326500000
F istatistiği	16,20	0,52	1,09	12,39
p değeri	0,00	0,66	0,37	0,00

Tablo 6.12. Tek yönlü model için Öklid uzaklığı çıktıları (N=1000)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	1000	1000	1000	1000
Genel kareler toplamı	391800000	391800000	391800000	391800000
Grup için kareler toplamı	216500000	388500000	371600000	326500000
F istatistiği	16,20	0,52	1,09	12,39
p değeri	0,00	0,61	0,39	0,00

Tablo 6.12’de analiz permütasyon sayısı 1000’dir. Yıl ve çeyrek grupları için $p > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur. Bölge ve Segment grubu için $p < 0,05$ olduğundan grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık vardır.

Analiz sonuçları incelendiğinde permütasyon sayısı arttırıldıkça p-değeri daha belirgin olup 1000 permütasyon sonrasında sabit kalmaktadır. Doğru sonuçlar almak için permütasyon sayısının yüksek olması gerekmektedir.

6.2.2. Bray-Curtis uzaklığı ile analiz çıktıları

Tablo 6.13. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=10)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	10	10	10	10
Genel kareler toplamı	13	13	13	13
Grup için kareler toplamı	7	12	12	10
F istatistiği	15,15	0,42	0,94	18,10
p değeri	0,09	0,91	0,73	0,09

Tablo 6.14. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=50)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	50	50	50	50
Genel kareler toplamı	13	13	13	13
Grup için kareler toplamı	7	12	12	10
F istatistiği	15,15	0,42	0,94	18,10
p değeri	0,02	0,75	0,65	0,02

Tablo 6.15. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=500)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	500	500	500	500
Genel kareler toplamı	13	13	13	13
Grup için kareler toplamı	7	12	12	10
F istatistiği	15,15	0,42	0,94	18,10
p değeri	0,00	0,78	0,50	0,00

Tablo 6.16. Tek yönlü model için Bray-Curtis uzaklığı çıktıları (N=1000)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	1000	1000	1000	1000
Genel kareler toplamı	13	13	13	13
Grup için kareler toplamı	7	12	12	10
F istatistiği	15,15	0,42	0,94	18,10
p değeri	0,00	0,82	0,46	0,00

Tablo 6.16.'da analiz permütasyon sayısı 1000'dir. Yıl ve çeyrek grupları için $p > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur. Bölge ve Segment grubu için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık vardır.

6.2.3. Manhattan uzaklığı ile analiz çıktıları

Tablo 6.17. Tek yönlü model için Manhattan uzaklığı çıktıları (N=10)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	10	10	10	10
Genel kareler toplamı	1212000000	1212000000	1212000000	1212000000
Grup için kareler toplamı	822100000	1202000000	1144000000	987300000
F istatistiği	9,48	0,52	1,18	14,10
p değeri	0,09	0,91	0,27	0,09

Tablo 6.18. Tek yönlü model için Manhattan uzaklığı çıktıları (N=50)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	50	50	50	50
Genel kareler toplamı	1212000000	1212000000	1212000000	1212000000
Grup için kareler toplamı	822100000	1202000000	1144000000	987300000
F istatistiği	9,48	0,52	1,18	14,10
p değeri	0,02	0,65	0,25	0,02

Tablo 6.19. Tek yönlü model için Manhattan uzaklığı çıktıları (N=500)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	500	500	500	500
Genel kareler toplamı	1212000000	1212000000	1212000000	1212000000
Grup için kareler toplamı	822100000	1202000000	1144000000	987300000
F istatistiği	9,48	0,52	1,18	14,10
p değeri	0,00	0,58	0,30	0,00

Tablo 6.20. Tek yönlü model için Manhattan uzaklığı çıktıları (N=1000)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Permütasyon sayısı (N)	1000	1000	1000	1000
Genel kareler toplamı	1212000000	1212000000	1212000000	1212000000
Grup için kareler toplamı	822100000	1202000000	1144000000	987300000
F istatistiği	9,48	0,52	1,18	14,10
p değeri	0,00	0,60	0,32	0,00

Tablo 6.20.'de analiz permütasyon sayısı 1000'dir. Yıl ve çeyrek grupları için $p > 0,05$ olduğundan dolayı H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur. Bölge ve Segment grubu için $p < 0,05$ olduğundan H_0 hipotezi reddedilir. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık vardır.

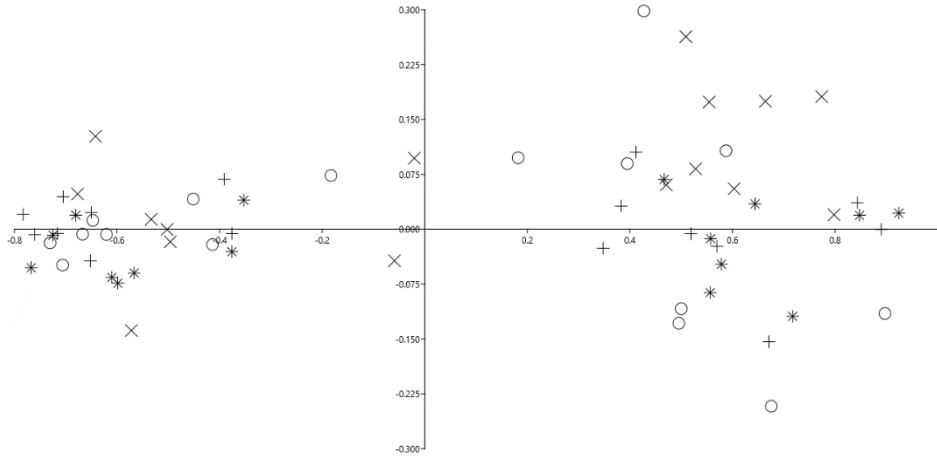
Tablo 6.21. Tek yönlü model için uzaklık ölçülerinin karşılaştırılması (N=1000)

Açıklama	Bölge	Yıl	Çeyrek	Segment
Öklid				
F istatistiği	9,48	0,52	1,18	14,1
p değeri	0,00	0,6	0,32	0,00
Bray-Curtis				
F istatistiği	15,15	0,42	0,94	18,1
p değeri	0,00	0,82	0,46	0,00
Manhattan				
F istatistiği	16,2	0,52	1,09	12,39
p değeri	0,00	0,61	0,39	0,00

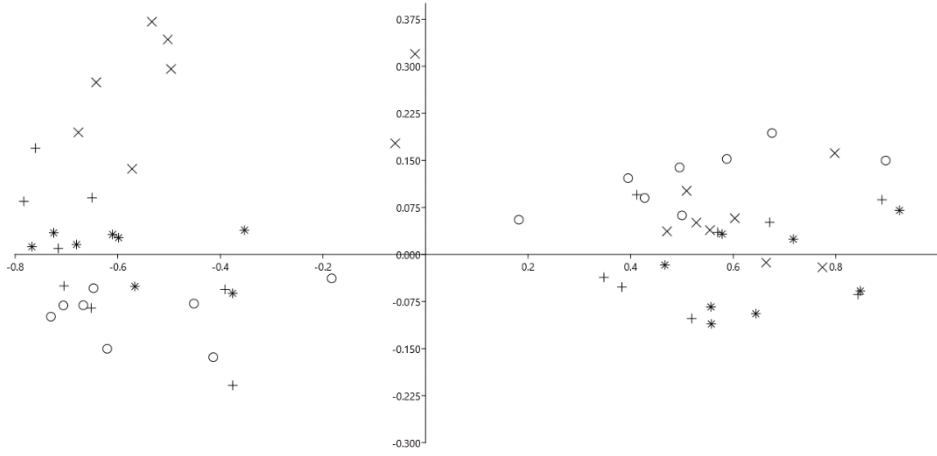
Tek yönlü model için Öklid, Bray-Kurtis ve Manhattan uzaklık ölçüleri kullanılarak analiz yapılmıştır. Tablo 6.21.'de uzaklık ölçülerinin p-değerleri karşılaştırılmış ve ölçüler arasında hipotezin red edilmesi ya da red edilememesi için bir fark olduğunu görünmemektedir.

Üç uzaklık ölçüsünde 1000 permütasyon yapıldığında $p < 0,05$ olduğundan bölge

ve ürün segmentleri arasında anlamlı bir farklılık olduğu, $p>0,05$ olduğundan yıl ve çeyrek grupları arasında anlamlı bir farklılık olmadığı söylenebilir.



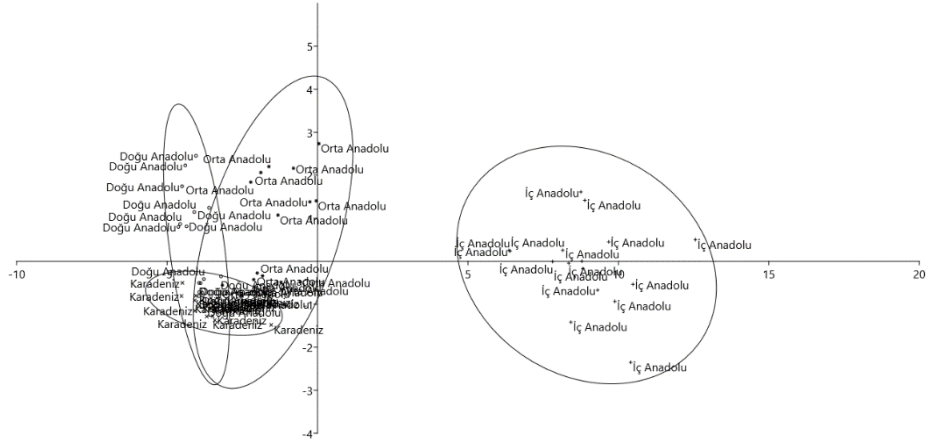
Şekil 6.8. Bölgelerin ürün grubu üzerindeki dağılımı



Şekil 6.9. Bölgelerin segmentlere göre ürün grubu üzerindeki dağılımı

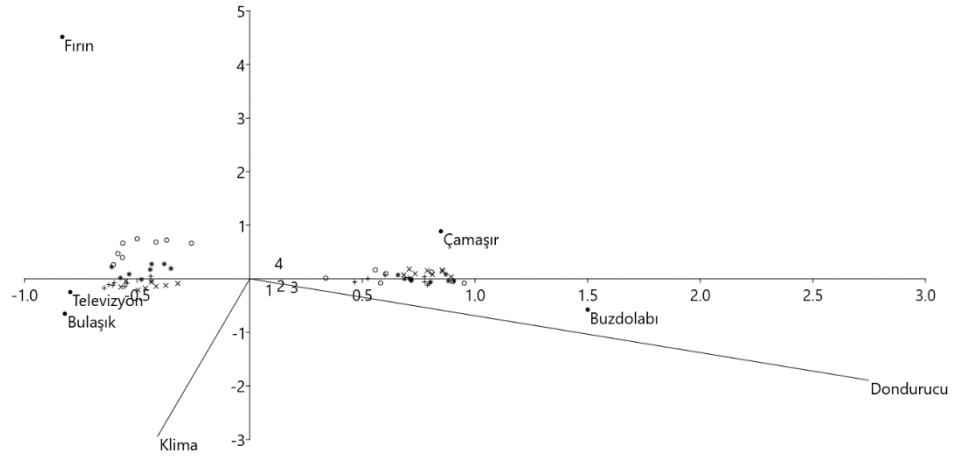
Şekil 6.8.'de bölgelerin ürün gruplarına göre doğrudan satış adetlerinin dağılımı göstermektedir. Bölgelerin dağılışı ortak bir noktada birleşmemesinden dolayı bir farklılık olduğu dikkat çekmektedir.

Şekil 6.9.'da bölgelerin segmentlere göre satış adetlerinin ürün grupları üzerindeki dağılımını göstermektedir. Segmentlere göre bölgelerin ürün grubu üzerindeki satışların farklılık gösterdiği söylenebilir.



Şekil 6.10. Bölgelere göre ürün grubu üzerindeki yoğunluk dağılımı grafiği

Şekil 6.10.'da ürün grubu bazında yoğunlaşan bölgeler gösterilmektedir. Grafikte görüleceği üzere Orta Anadolu, Karadeniz ve Doğu Anadolu bölgesi satışları ortak bölgede yoğunlaşmaktadır. İç Anadolu bölgesi diğer bölgelerden uzakta kümelenmiş durumdadır. Farklılık en çok İç Anadolu bölgesinden kaynaklanmaktadır.



Şekil 6.11. Segmentlere göre ürün grubu üzerindeki yoğunluk dağılımı grafiği

Bölgelerin ürün gruplarına göre gerçekleşen satışların segment dağılımı göstermektedir. Şu yorumlar yapılabilir; bulaşık makinesi, televizyon ve çamaşır makinesi ürün grubunda bölgeler arasında orta ve üst segment satışlarında bir farklılık yoktur. Buzdolabı ürün grubu benzerlik göstermektedir. Derin dondurucu ve klimada ise bölgeler arasında anlamlı farklılık çok yüksektir. Özetle, segmentlere göre anlamlı farklılık çıkmasındaki en büyük pay, buzdolabı, derin dondurucu ve klima ürün grubundan kaynaklanmaktadır.

6.3. İki Yönlü Modelin Oluşturulması

Veri grubumuzda bölgeler, yıllar, çeyrekler ve segmentler olmak üzere dört tane grup değişkenimiz mevcuttur. Bu değişkenlerin ikişerli olarak ürün grupları ile iki yönlü incelenmesi yapılarak farklılıkların olup olmadığına bakılacaktır. Tek yönlü analizde permütasyon sayısının artırılmasının tutarlı sonuçlar verdiği gösterilmiştir. İki yönlü analiz yapılırken permütasyon sayısı 1000 olarak alınacaktır.

6.3.1. İki yönlü analiz çıktıları (Bölge ve yıl)

Bölgeler ve yılların ürün gruplarına göre satış adetleri arasında anlamlı bir farklılık olup olmadığı test edilmiştir. Analizde üç farklı uzaklık ölçüsü kullanılmış ve karşılaştırılmasına yer verilmiştir.

Tablo 6.22. Bölge ve yılların Öklid uzaklığı ile analiz çıktıları

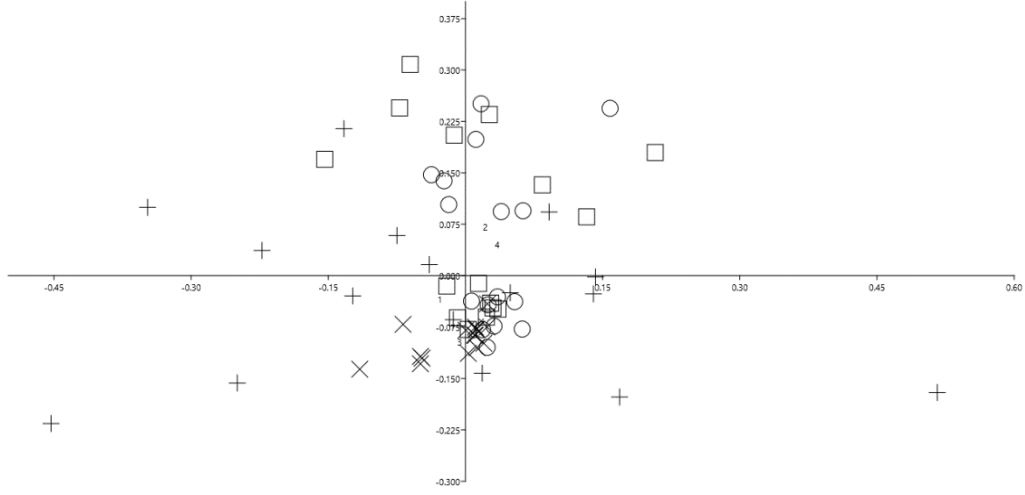
Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	175310000	3	58435000	15,54	0,00
Yıl	3257900	1	3257900	0,87	0,45
Etkileşim	2675700	3	891910	0,24	0,98
Artıklar	210550000	56	3759900		
Toplam	391790000	63			

Tablo 6.23. Bölge ve yılların Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	G.K.T.	S.D.	G.İ.K.T.	F istatistiği	p-değeri
Bölge	5	3	2	14,65	0,00
Yıl	0	1	0	0,68	0,59
Etkileşim	0	3	0	0,45	0,95
Artıklar	7	56	0		
Toplam	13	63			

Tablo 6.24. Bölge ve yılların Manhattan uzaklığı ile analiz çıktıları

Kaynak	G.K.T.	S.D.	G.İ.K.T.	F istatistiği	p-değeri
Bölge	389690000	3	129900000	9,02	0,00
Yıl	10129000	1	10129000	0,70	0,49
Etkileşim	5892800	3	1964300	0,14	0,99
Artıklar	806110000	56	14395000		
Toplam	1211800000	63			



Şekil 6.12. Bölge ve yılların ürün grubu dağılım grafiği

Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görülmektedir. Bölgelerin p-değeri 0,05'den küçük olduğu için bölgeler arasında anlamlı bir farklılık vardır. Yılların p-değeri 0,05'den büyük olduğu için yıllar arasında anlamlı bir farklılık olmadığı söylenebilir.

Etkileşim hipotezleri şu şekilde kurulabilir,

H_0 : Bölge ve yıl faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı farklılık yoktur.

H_1 : Bölge ve yıl faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörlerinin en az ikisi arasında anlamlı bir farklılık vardır.

Bölge ve yılların ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakılmalıdır, p-değeri $0,95 > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur.

6.3.2. İki yönlü analiz çıktıları (Bölge ve çeyrekler)

Tablo 6.25. Bölge ve çeyreklerin Öklid uzaklığı ile analiz çıktıları

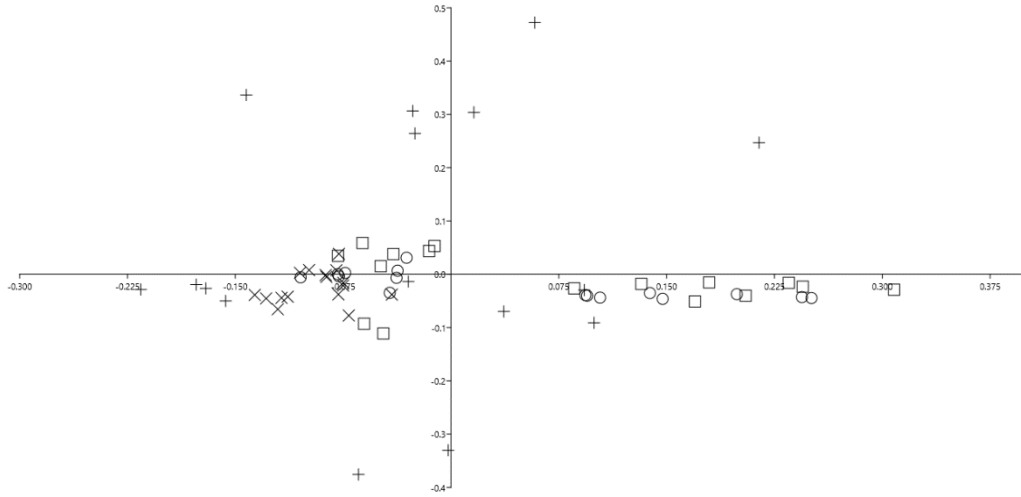
Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	175310000	3	58435000	16,29	0,00
Çeyrek	20180000	3	6726600	1,87	0,09
Etkileşim	24086000	9	2676200	0,75	0,79
Artıklar	172220000	48	3587900		
Toplam	391790000	63			

Tablo 6.26. Bölge ve çeyreklerin Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	5	3	2	13,88	0,00
Çeyrek	1	3	0	1,44	0,16
Etkileşim	0	9	0	0,29	0,99
Artıklar	6	48	0		
Toplam	13	63			

Tablo 6.27. Bölge ve çeyreklerin Manhattan uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	389690000	3	129900000	8,97	0,00
Çeyrek	67605000	3	22535000	1,56	0,15
Etkileşim	59339000	9	6593200	0,46	0,98
Artıklar	695180000	48	14483000		
Toplam	1211800000	63			



Şekil 6.13. Bölge ve çeyreklerin ürün grubu dağılım grafiği

Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görülmektedir. Bölgelerin p-değeri 0,05'den küçük olduğu için bölgeler arasında anlamlı bir farklılık vardır. Çeyreklerin p-değeri 0,05'den büyük olduğu için çeyrekler arasında anlamlı bir farklılık olmadığı söylenebilir.

Etkileşim hipotezleri şu şekilde kurulabilir,

H_0 : Bölge ve çeyrek faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı farklılık yoktur.

H_1 : Bölge ve çeyrek faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörlerinin en az ikisi arasında anlamlı bir farklılık vardır.

Bölgelerin ve çeyreklerin ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakılmalıdır, p-değeri $0,79 > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur. Bölgeler ve çeyreklerin birlikte etkisinin olmadığı söylenebilir.

6.3.3. İki yönlü analiz çıktıları (Bölge ve segmentler)

Tablo 6.28. Bölge ve segmentlerin Öklid uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	175310000	3	58435000	39,34	0,00
Segment	65260000	1	65260000	43,93	0,00
Etkileşim	68033000	3	22678000	15,27	0,00
Artıklar	83192000	56	1485600		
Toplam	391790000	63			

Tablo 6.29. Bölge ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Bölge	5	3	2	49,99	0,00
Segment	3	1	3	78,65	0,00
Etkileşim	2	3	1	21,13	0,00
Artıklar	2	56	0		
Toplam	13	63			

Tablo 6.30. Bölge ve segmentlerin Manhattan uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GKO	F istatistiği	p-değeri
Bölge	389690000	3	129900000	37,77	0,00
Segment	224510000	1	224510000	65,28	0,00
Etkileşim	405040000	3	135010000	39,26	0,00
Artıklar	192590000	56	3439000		
Toplam	1211800000	63			

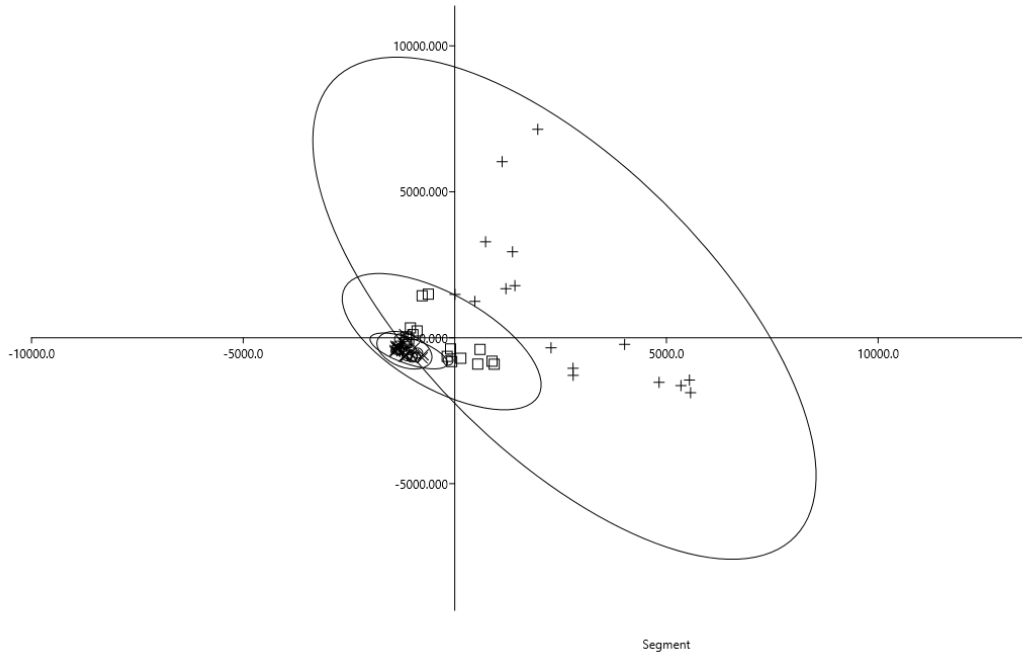
Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görünmektedir. Bölgeler ve segmentlerin p-değeri 0,05'den küçük olduğu için bölgeler arasında anlamlı bir farklılık vardır.

Etkileşim hipotezleri şu şekilde kurulabilir,

H_0 : Bölge ve segment faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı farklılık yoktur.

H_1 : Bölge ve segment faktör seviyelerinde 7 tane bağımlı değişkenden oluşan ortalama vektörlerinin en az ikisi arasında anlamlı bir farklılık vardır.

Bölgelerin ve segmentlerin ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakılmalıdır, p-değeri $0,00 < 0,05$ olduğundan H_0 hipotezi reddedilir. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık vardır.



Şekil 6.14. Bölge ve segmentlerin ürün grubu dağılım grafiği

Şekil 6.14.'de bölgeler ve segmentlerin ürün gruplarına göre dağılımları benzer olduğu ve etkileşimin önemli olduğu söylenebilir. Buradan, alt ve üst segment ürünler satış hızına göre bölgesel bazda değişkenlik göstermektedir.

6.3.4. İki yönlü analiz çıktıları (Yıl ve çeyrekler)

Tablo 6.31. Yıl ve çeyreklerin Öklid uzaklığı ile analiz çıktıları

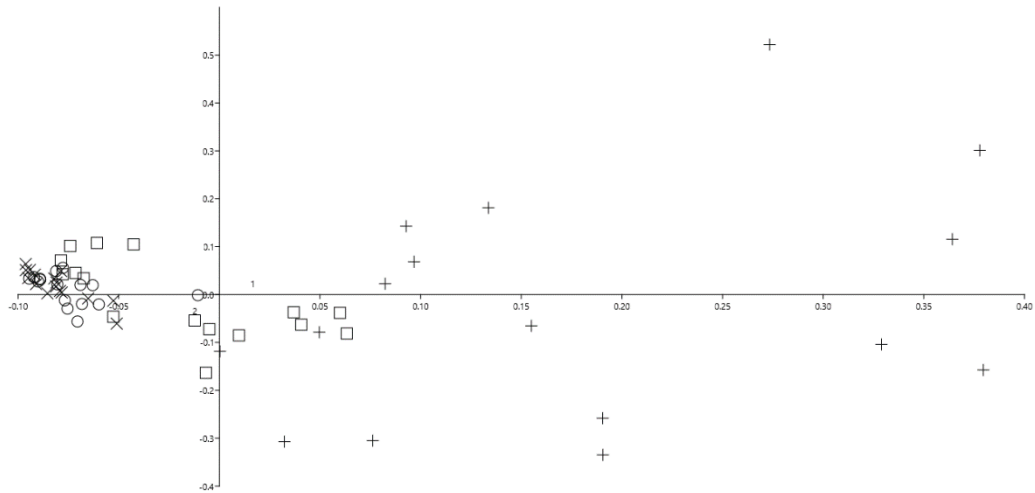
Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	3257900	1	3257900	0,50	0,65
Çeyrek	20180000	3	6726600	1,03	0,39
Etkileşim	2808600	3	936190	0,14	0,99
Artıklar	365540000	56	6527600		
Toplam	391790000	63			

Tablo 6.32. Yıl ve çeyreklerin Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	0	1	0	0,40	0,85
Çeyrek	1	3	0	0,89	0,54
Etkileşim	0	3	0	0,10	0,99
Artıklar	12	56	0		
Toplam	13	63			

Tablo 6.33. Yıl ve çeyreklerin Manhattan uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	10129000	1	10129000	0,50	0,65
Çeyrek	67605000	3	22535000	1,12	0,35
Etkileşim	7971200	3	2657100	0,13	0,99
Artıklar	1126100000	56	20109000		
Toplam	1211800000	63			



Şekil 6.15. Yıl ve çeyreklerin ürün grubu dağılım grafiği

Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görülmektedir. Yıl ve çeyrek bazda bakıldığında p-değeri 0,05'den büyük olduğu için anlamlı bir farklılık olmadığı söylenebilir.

Çeyrekler ve yılların ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakıldığında, p-değeri $0,99 > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur.

6.3.5. İki yönlü analiz çıktıları (Yıl ve segmentler)

Tablo 6.34. Yıl ve segmentlerin Öklid uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	3257900	1	3257900	0,61	0,59
Segment	65260000	1	65260000	12,16	0,00
Etkileşim	1210800	1	1210800	0,23	0,84
Artıklar	322060000	60	5367700		
Toplam	391790000	63			

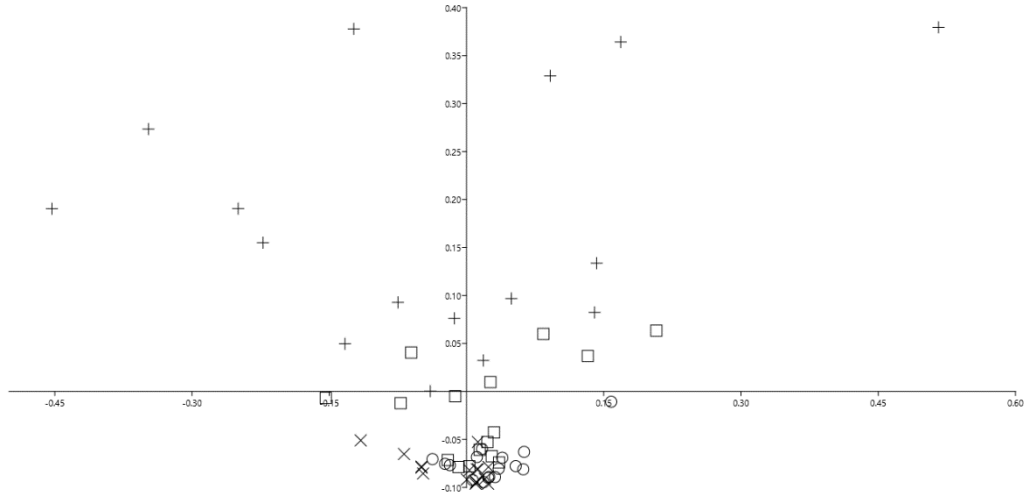
Tablo 6.35. Yıl ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	0	1	0	0,53	0,72
Segment	3	1	3	17,82	0,00
Etkileşim	0	1	0	0,52	0,73
Artıklar	10	60	0		
Toplam	13	63			

Tablo 6.36. Yıl ve segmentlerin Manhattan uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Yıl	10129000	1	10129000	0,62	0,58
Segment	224510000	1	224510000	13,82	0,00
Etkileşim	2298800	1	2298800	0,14	0,88
Artıklar	974880000	60	16248000		
Toplam	1211800000	63			

Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görülmektedir. Segment bazında bakıldığında p-değeri 0,05'den küçük olduğundan segmentler arası anlamlı bir farklılık olduğu söylenebilir.



Şekil 6.16. Yıl ve segmentlerin ürün grubu dağılım grafiği

Segmentler ve yılların ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakıldığında, p-değeri $0,73 > 0,05$ olduğundan H_0 hipotezi red edilemez. Etkileşimin anlamlı olmadığı söylenebilir.

6.3.6. İki yönlü analiz çıktıları (Çeyrek ve segmentler)

Tablo 6.37. Çeyrek ve segmentlerin Öklid uzaklığı ile analiz çıktıları

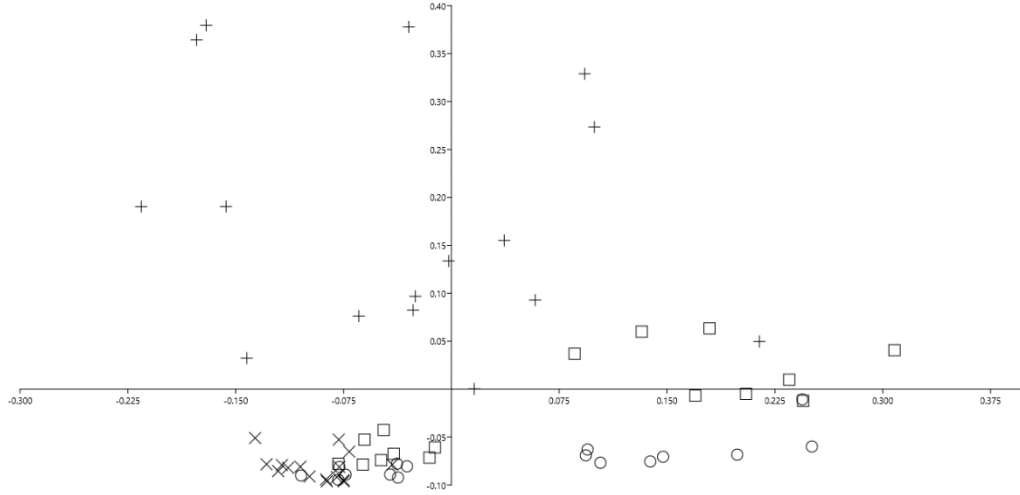
Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Çeyrek	20180000	3	6726600	1,27	0,28
Segment	65260000	1	65260000	12,28	0,00
Etkileşim	8773400	3	2924500	0,55	0,80
Artıklar	297580000	56	5313900		
Toplam	391790000	63			

Tablo 6.38. Çeyrek ve segmentlerin Bray-Curtis uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Çeyrek	1	3	0	1,16	0,29
Segment	3	1	3	17,50	0,00
Etkileşim	0	3	0	0,15	0,99
Artıklar	9	56	0		
Toplam	13	63			

Tablo 6.39. Çeyrek ve segmentlerin Manhattan uzaklığı ile analiz çıktıları

Kaynak	GKT	SD	GİKT	F istatistiği	p-değeri
Çeyrek	67605000	3	22535000	1,38	0,20
Segment	224510000	1	224510000	13,76	0,00
Etkileşim	6056300	3	2018800	0,12	0,99
Artıklar	913650000	56	16315000		
Toplam	1211800000	63			



Şekil 6.17. Çeyrek ve segmentlerin ürün grubu dağılım grafiği

Tablolardaki analiz sonuçları incelendiğinde uzaklık ölçüleri arasında sonucu değiştirecek bir fark olmadığı görülmektedir. Segment bazında bakıldığında p-değeri 0,05'den küçük olduğu için segmentler arası anlamlı bir farklılık olduğu söylenebilir.

Segmentler ve çeyreklerin ürün grupları üzerinde birlikte etkisi için etkileşim parametresine bakıldığında, p-değeri $0,80 > 0,05$ olduğundan H_0 hipotezi red edilemez. Grup faktör seviyelerinde, bağımlı değişkenden oluşan ortalama vektörleri arasında anlamlı bir farklılık yoktur.

7. SONUÇ VE ÖNERİLER

Birçok alanda olduğu gibi satış pazarlama alanında da bazı değişkenler normal dağılmamakta ve parametrik testlerin gerekli varsayımları sağlanamamaktadır. Ayrıca gözlem sayısı, değişken sayısından daha az olduğunda parametrik testlerin kullanımı uygun olmayacaktır.

Bu çalışmada, değişkenleri birbirleriyle ilişkili birden fazla sonucunu birlikte değerlendiren, normallik varsayımı gibi parametrik testlerin varsayımları sağlanmadığı için PERMANOVA yöntemi ile incelenmiştir. Hangi modelin uygulanması gerektiği anlatılmış ve sıklıkla kullanılan uzaklık ölçülerinin sonuçları kıyaslanmıştır. Parametrik olmayan çok değişkenli varyans analizi yöntemlerinden biri olan PERMANOVA analizi ile incelenen veri setinde permütasyon sayısı arttıkça uzaklık ölçülerinin birbirine yakın sonuçlar verdiği gözlenmiştir. Permütasyon sayısı 50'nin altında iken uzaklık ölçü sonuçları birbirinden uzak ve seçilmesi durumunda sonucu değiştirdiği, uygun olan uzaklık ölçüsünün doğru seçilmesi gerektiği ve permütasyon sayısı artırılarak, hem tek yönlü hem de çift yönlü analizler yaparak karşılaştırmalara yer verilmiştir.

Parametrik olmayan analizlerin tek varsayımı verilerin doğruluğunun olmasına rağmen veri setinde tekli ve çoklu normallik sağlanıp sağlanmadığı özellikle kontrol edilmiştir. Tüm veri grupları hem tekil hem de çok değişkenli normal dağılım göstermediği gözlemlenmiştir. Yarı metrik ve tam metrik olan uzaklık ölçülerinden Öklid, Bray Curtis ve Manhattan'ın ayrı ayrı çıktıklarına yer vererek analiz sonuçları tablolar halinde verilmiştir. Bu uzaklık ölçülerinde Bray Curtis ile elde edilen sonuçlarda etkileşim parametresi önemsizken, Öklid yönteminde önemli çıktığı görülmüştür. Bu sebeple uzaklık ölçüsünün doğru seçilmesinin önemli olduğu tespit edilmiştir.

Parametrik olmayan yöntemlerden PERMANOVA yöntemi literatürde çok karşılaşılmayan bir yöntem olmasına rağmen son dönemlerde özellikle sağlık alanında çok ilgi görmüştür. Perakende satış verilerinde de parametrik testlerin varsayımları sağlanmadığı için ve parametrik olmayan yöntemlerin daha tutarlı sonuçlar verdiği gözlemlenmiştir.

Çalışmada, beyaz eşya perakende satışlarına yönelik değişkenler için tek ve iki yönlü modeller oluşturulmuş olup, uzaklık ölçülerinden p değeri en düşük olan uzaklık ölçü çıktıları ile hipotezler değerlendirilmiştir.

Tek yönlü model sonuçları incelendiğinde, bölgeler ile ürün grupları arasında sonuçlar incelendiğinde, İç Anadolu, Orta Anadolu, Karadeniz ve Doğu Anadolu bölgeleri arasında anlamlı bir farklılık olduğu görülmüştür. Yıl dönemleri ile ürün grupları arasında sonuçlar incelendiğinde, 2019 ve 2020 yılları arasında anlamlı bir farklılık olmadığı görülmüştür. Çeyrekler ile ürün grupları arasında sonuçlar incelendiğinde, 2019 ve 2020 yılları arasında anlamlı bir farklılık olmadığı görülmüştür. Segmentler ile ürün grupları arasında sonuçlar incelendiğinde, orta ve üst segment arasında anlamlı bir farklılık olduğu görülmüştür.

İki yönlü model sonuçları incelendiğinde, bölgeler ve yıl dönemlerinin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olmadığı görülmüştür. Bölgeler ve çeyrek dönemlerinin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olmadığı görülmüştür. Bölgeler ve segmentlerin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olduğu görülmüştür. Bu farklılığın İç Anadolu bölgesinin üst segment satışlarından kaynaklandığı Şekil 6.14’de grafiksel olarak gösterilmiştir. Yıl ve çeyrek dönemlerinin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olmadığı görülmüştür. Yıl dönemleri ve segmentlerin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olmadığı görülmüştür. Çeyrek dönemleri ve segmentlerin ürün grupları üzerine etkileri birlikte incelendiğinde, anlamlı bir farklılık olmadığı görülmüştür.

Çalışma sonucunda, bölgeler ve segmentler arasında anlamlı farklılık olduğu görülmüştür. İç Anadolu bölgesinin ürün grubu satış adetleri incelendiğinde diğer bölgelere kıyasla farklılık olduğu tespit edilmiştir. Tüm bölgelerde ürün gruplarının orta ve üst segment satışları incelendiğinde, anlamlı farklılıklar görülmüş ve farklılığın en yüksek olduğu bölge İç Anadolu bölgesidir. İç Anadolu bölgesinde üst segment satışları, orta segmentlere göre anlamlı derecede farklılık olduğu görülmüştür.

Uygulamadan elde ettiğimiz sonuçlarda, üst segment satışlarının Orta Anadolu, Karadeniz ve Doğu Anadolu bölgelerinde gerçekleşen satışlar ile uyumlu olduğu yıl ve çeyrek dönemlerinde anlamlı bir farklılık olmadığı görülmüştür. Bu bölgelerde üst segment satışları için firma tarafında çalışmalar yapılabilir.

KAYNAKLAR

- Adek, T. R. and Quran, F. P. (2015). Detection system Tajwid Al Quran on image using Bray Curtis distance, *International Journal of Computing and Technology*, 2:8, 293-300.
- Alpar, R. (2013). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler* (Dördüncü Baskı). Detay Yayıncılık, 167-180, Türkiye.
- Anderson, M. J. (2001a). A new method for non-parametric multivariate analysis of variance, *Austral Ecology*, 26, 32-46. doi: 10.1111/j.1442-9993.2001.01070. pp.x
- Anderson, M. J. (2001b). Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of Fisheries and Aquatic Sciences* 58, 626-639. doi: 10.1139/cjfas-58-3-626
- Anderson, M. J. (2005). Permutational multivariate analysis of variance a computer program, *University of Auckland Department of Statistics*, 1-24.
- Anderson, M. J. and Walsh, D. C. I. (2013). PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing, *Ecological Monographs*, 83:4, 557-574. doi: 10.1890/12-2010.1
- Balibeyoğlu, J. (1989). Çok değişkenli varyans analizi ve yazlık mercimek kalitesi üzerine uygulaması. Doktora Tezi. Gazi Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, 112, Ankara.
- Gower, J. C. and Krzanowski, W. J. (1999). Analysis of distance for structured multivariate data and extensions to multivariate analysis of variance, *Journal of the Royal Statistical Society*, 48, 505-519. doi: 10.1111/1467-9876.00168
- Hammer, Ø., Harper, D. and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4-9
- Işık, M. and Çamurlu, A. Y. (2008). Web belgeleri kümelemede benzerlik ve uzaklık ölçütleri başarılarının karşılaştırılması, *Marmara Fen Bilimleri Dergisi*, 20:1, 35-49
- Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (Sixth Edition). Pearson Education, 771-773, United States.
- Kelly, B. J. and Gross, R. (2015). Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA, *Bioinformatics*, 31:15, 2461-2468. doi: 10.1093/bioinformatics/btv183
- Kıroğlu G. (2001). *Uygulamalı Parametrik Olmayan İstatistiksel Yöntemler* (Altıncı Baskı), 148-187, Türkiye.
- Koç, Ş. (2018). Çok değişkenli varyans analizinde kullanılan test istatistiklerinin Monte-Carlo simülasyonu ile karşılaştırılması, Doktora Tezi, Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, 28-33, Kahramanmaraş.
- Manly, BFJ., (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology* (2nd edition), Chapman and Hall, UK.

- Pasin, Ö., Ankaralı, H. ve Cangür Ş. (2006). Parametrik olmayan çok deęişkenli varyans analizi ve saęlık alanında bir uygulaması, *Bilişim Teknolojileri Dergisi*, 9:1, 13-20. doi: 10.17671/btd.47787
- Peralta, A. L., Matthews, J. W., Flanagan, D. N. and Kent, A. D. (2012). Environmental factors at dissimilar spatial scales influence plant and microbial communities in restored wetlands, *Society of Wetland Scientists*, 32, 1125-1134. doi: 10.1007/s13157-012-0343-3
- Tang, Z. Z., Chen, G. and Alekseyenko, A. V. (2016). PERMANOVA-S: Association test for microbial community composition that accommodates confounders and multiple distances, *Human Genetics*, 96:5, 797-807.
- Terzi, Y. (1995). Çok deęişkenli varyans analizi (Manova) ve bir uygulama, Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi İstatistik Anabilim Dalı, 43, Samsun.
- Watson, D. L. and Harvey E.S. (2007). Protection from fishing alters the species composition of fish assemblages in a temperate-tropical transition zone, *Australian Institute of Marine Science*, 152, 1197-1206. doi: 10.1007/s00227-007-0767-0
- Zhu, Z., Satten, G., Caroline, M. and Hu, Y.J. (2020). Analyzing matched sets of microbiome data using the LDM and PERMANOVA, *Research Square*, 1-16. doi: 10.21203/rs.3.rs-17148/v1

ÖZ GEÇMİŞ

Serhat YILMAZ, Nefise Andıçen Lisesini bitirdikten sonra Ondokuz Mayıs Üniversitesi Fen Edebiyat Fakültesi'nden 2015 yılında mezun oldu.

İletişim Bilgileri

ORCID ID: 0000-0002-6061-8529