



T.C.
AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ
ANA BİLİM DALI

YÜKSEK
LİSANS
TEZİ

KAYIP VERİLERİN BENZERLİK TEMELLİ
ATAMA YÖNTEMİ İLE BELİRLENMESİ VE
BASİT ATAMA YÖNTEMLERİ İLE
KARŞILAŞTIRILMASI

Ahmet ŞENGÜL

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
BİLİM DALI

Antalya, 2022

**AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI EĞİTİMDE
ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK
LİSANS PROGRAMI**

**KAYIP VERİLERİN BENZERLİK TEMELLİ ATAMA YÖNTEMİ İLE
BELİRLENMESİ VE BASİT ATAMA YÖNTEMLERİ İLE KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

AHMET ŞENGÜL

**Danışman:
DOÇ.DR. ALPER SİNAN**

Antalya, 2022

DOĞRULUK BEYANI

Yüksek lisans tezi olarak sunduğum bu çalışmayı, bilimsel ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçalardan gösterilenlerden oluştuğunu ve bu eserleri her kullanımında alıntı yaparak yararlandığımı belirtir; bunu onurumla doğrularım. Enstitü tarafından belli bir zamana bağlı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

22/ 07 / 2022

AHMET ŞENGÜL

AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Ahmet ŞENGÜL 'un bu çalışması 28 / 06 / 2022 tarihinde jürimiz tarafından Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme yüksek lisans tezi olarak **oy birliği/oy çokluğu** ile kabul edilmiştir.

İmza

Başkan : Prof. Dr. Bayram BIÇAK
(Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri ABD)

Üye: : Dr. Öğ. Üyesi Aydın KARAKOCA
(Necmettin Erbakan Üniversitesi, Fen Fakültesi, İstatistik ABD)

Üye (Danışman) : : Doç. Dr. Alper SİNAN
(Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri ABD)

YÜKSEK LİSANS TEZİNİN ADI: KAYIP VERİLERİN BENZERLİK TEMELLİ ATAMA YÖNTEMİ İLE BELİRLENMESİ VE BASİT ATAMA YÖNTEMLERİ İLE KARŞILAŞTIRILMASI

ONAY: Bu tez, Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulunun tarihli ve sayılı kararıyla kabul edilmiştir.

Enstitü Müdürü

TEŐEKKÜR

Ders ve tez yazım sürecinde hayatımı kolaylařtırmak için olađanüstü bir özveride bulunan eřim Sema ŐENGÜL'e, arařtırma alıřmalarımı yürütme sürecinde yařadığım her türlü durumda desteklerini esirgemeyen tez danışmanım Do. Dr. Alper SİNAN'a, yürekten kardeşlerim Ođuz ÖZYAZI'ya ve Mehmet ERSOY'a sonsuz teşekkür ederim.

ÖZET

KAYIP VERİLERİN BENZERLİK TEMELLİ ATAMA YÖNTEMİ İLE BELİRLENMESİ VE BASİT ATAMA YÖNTEMLERİ İLE KARŞILAŞTIRILMASI

ŞENGÜL Ahmet

Yüksek Lisans Tezi

Eğitim Bilimleri (Eğitimde Ölçme ve Değerlendirme) Ana Bilim Dalı

Tez Danışmanı: Doç. Dr. Alper SİNAN

Temmuz 2022, 51 sayfa

Araştırmada kayıp gözlem tamamlamada yeni yöntem olarak benzerlik temelli atama yönteminin geliştirilmesi ve bu yeni yöntem ile tanıtıcı istatistik yöntemlerinin güvenilirlik ve doğru atama oranları açısından karşılaştırılması amaçlanmaktadır. Araştırma, temel araştırma modeli bağlamında kurgulanmıştır. Akdeniz Üniversitesi eğitim bilimleri alanında yapılan ilgi ölçüğünü eksiksiz cevaplayan 3. Sınıf öğrenciler arasından rastgele sayılar tablosu ile seçilen 450 öğrenciye ait veriler araştırmanın çalışma grubunu oluşturmaktadır. Araştırmada elde edilen veriler üzerinde Cronbach alfa güvenilirlik testi ve doğru atama oranlarının karşılaştırma analizleri ve betimleyici analizler spss paket program ile gerçekleştirilmiştir. Yapılan analizler ile benzerlik temelli atama yönteminin güvenilirlik katsayısının tanıtıcı istatistik yöntemlerine göre daha yüksek olduğu görülmüştür. Ayrıca benzerlik temelli atama yöntemi ile kayıp gözlem tamamlamada doğru atama oranının tanıtıcı istatistik yöntemlerine göre daha büyük olduğu belirlenmiştir. Elde edilen bulgulara göre en çok doğru atama gerçekleştiren yöntemin benzerlik temelli atama yöntemi (% 5) olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Kayıp gözlem, Benzerlik temelli atama, Basit atama yöntemleri, Cramer's V katsayısı

ABSTRACT

DETERMINING MISSING DATA WITH SIMILARITY BASED IMPITATION METHOD AND COMPARISON WITH SIMPLE IMPUTATIONS METHOD

ŞENGÜL Ahmet

Master's Degree Thesis

Department of Educational Sciences (Assessment and Evaluation in Education)

Thesis Consultant: Doç.Dr. Alper SİNAN

July 2022, 51 page

In the research, it is aimed to develop similarity-based assignment method as a new method in completing missing observations and to compare this new method and descriptive statistics methods in terms of reliability and correct assignment rates. The research was designed in the context of the basic research model. The study group of the research consists of the data of 450 students selected by random numbers table among the 3rd grade students who answered the interest scale in the field of educational sciences at Akdeniz University. Cronbach's alpha reliability test and comparison analyzes of correct assignment rates and descriptive analyzes on the data obtained in the study were performed with the spss package program. With the analyzes made, it was seen that the reliability coefficient of the similarity-based assignment method was higher than the descriptive statistics methods. In addition, it has been determined that the correct assignment rate in completing the missing observation with the similarity-based assignment method is higher than the descriptive statistics methods. According to the findings, it was concluded that the most accurate assignment method was the similarity-based assignment method (5%).

Keywords: Missing observation, Similarity-based assignment, Simple assignment methods, Cramer's V coefficient

İÇİNDEKİLER

TEŞEKKÜR	IV
ÖZET	V
ABSTRACT	VI
TABLolar LİSTESİ	IX
ŞEKİLLER LİSTESİ	X

BÖLÜM I

GİRİŞ

1.1. Problem Durumu.....	3
1.2. Araştırmanın Amacı ve Alt Problemler	4
1.3. Araştırmanın Önemi	5
1.4. Araştırmanın Sayıltıları	5
1.5. Araştırmanın Sınırlılıkları	6
1.6. Tanımlar.....	6

BÖLÜM II

KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

2.1. KURAMSAL BİLGİLER	7
2.1.1. Kayıp Gözlem.....	7
2.1.2. Kayıp Gözlem Nedenleri	9
2.1.3. Kayıp Gözlem Türleri.....	9
2.1.3.1. Tamamıyla Rassal Olan Kayıp (TROC).....	10
2.1.3.2. Rassal Olan Kayıp (ROC)	10
2.1.3.2. İhmal Edilemez Kayıp Gözlemler (IEK)	11
2.1.4. Kayıp Gözlem Tamamlama Yöntemleri.....	12
2.1.4.1. Eğilim Skorları Eşleştirmesi.....	12
2.1.4.2. Bayesci Veri Tamamlama	12
2.1.4.3. Beklenti Maksimizasyonu Algoritması	12
2.1.4.4. Regrasyon Ataması Yöntemi.....	13
2.2. İLGİLİ ARAŞTIRMALAR	13
2.2.1. Kayıp Gözlem Atama Yöntemleri İle İlgili Yurtiçi Çalışmalar	13
2.2.2. Kayıp Gözlem Atama Yöntemleri İle İlgili Yurtdışı Çalışmalar	15

BÖLÜM III

YÖNTEM

3.1. Araştırmanın Modeli	17
3.2. Çalışma Grupları.....	17
3.3. Veri Toplama Araçları	18
3.4. Veri Analizi.....	18
3.4.1. Kayıp Gözlem Oluşturulması	19
3.4.2. Tanıtıcı İstatistiksel Yöntemler İle Kayıp Gözlem Atama.....	21
3.4.3. Cramer's V Katsayısı İle Kayıp Gözlem Atama.....	24
3.4.4. Benzerlik Temelli Kayıp Gözlem Atama Yöntemi	27
3.4.4.1. Kayıp Gözlem Oluşturma.....	28
3.4.4.2. Kayıp Gözlemlerin Algoritmaya Yüklenmesi.....	28
3.4.4.3. En Yüksek Benzerliklerin Belirlenmesi	30
3.4.4.4. Benzerlik Temelli Atama İle Kayıp Gözlem Atama	31

BÖLÜM IV

BULGULAR

4.1. Kayıp Gözlem Atama Yöntemlerinde Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular	32
4.1.1. %5 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular.....	32
4.1.2. %10 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular.....	33
4.1.3. %15 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular.....	34
4.2. Kayıp Gözlem Atama Yöntemlerinde Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular	34
4.2.1. %5 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular.....	35
4.2.2. %10 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular.....	35
4.2.3. %15 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular.....	36

BÖLÜM V
SONUÇ TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma:	37
5.2. Öneriler.....	39
KAYNAKÇA.....	40
EKLER.....	44
Ek 1.....	44
ÖZGEÇMİŞ	50
YAYIN İNTİHAL RAPORU	51

TABLolar

Tablo 3.1: Örneklem Demografik Bilgileri	17
Tablo 3.2.: Araştırma Grubunun Merkezi Eğilim Ölçülerinin Basıklık/Çarpıklığın İncelenmesi.....	18
Tablo 3.3.: Kayıp Gözlem Little's MCAR Testi Bulguları.....	19
Tablo 3.4.: %5 Kayıp Gözlem Setinin Madde Dağılımı.....	20
Tablo 3.5.: %10 Kayıp Gözlem Setinin Madde Dağılımı.....	20
Tablo 3.6.: %15 Kayıp Gözlem Setinin Madde Dağılımı.....	21
Tablo 3.7.: %5 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı.....	22
Tablo 3.8.: %10 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı.....	23
Tablo 3.9.: %15 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı.....	24
Tablo 3.10.: %5 Kayıp Gözlem Setinin Cramer's V katsayısına Göre Atama Dağılımı.....	25
Tablo 4.1.: %5 Kayıp Gözlem Doğru Atama Oranı Bulguları.....	33
Tablo 4.2.: %10 Kayıp Gözlem Doğru Atama Oranı Bulguları.....	33
Tablo 4.3.: %15 Kayıp Gözlem Doğru Atama Oranı Bulguları.....	34
Tablo 4.4.: %5 Kayıp Gözlem Güvenirlilik Analizi Bulguları.....	35
Tablo 4.5.: %10 Kayıp Gözlem Güvenirlilik Analizi Bulguları.....	35
Tablo 4.6.: %15 Kayıp Gözlem Güvenirlilik Analizi Bulguları.....	36

ŞEKİLLER

Şekil 3.1. Kayıp Gözlem Oluşturma Örneği.....	28
Şekil 3.2. B.T.A. Yöntemi Algoritmasının İşleyişi.....	29
Şekil 3.3. Benzerlik sıralaması.....	30

BÖLÜM I

GİRİŞ

Bilimsel arařtırmalarda hipotezler toplanan verilerin analiz edilmesiyle test edilir. Toplanan verilerden yüksek geçerlik ve güvenilirliğe sahip sonuçlar elde etmek için verilerin eksiksiz olması bilinen bir gerçektir. Kayıp gözlem, amaçlanan ölçmeye yönelik bir gözlem değerinin bulunmaması durumudur (Hohensinn & Kubinger, 2011). Çoğu araştırma her ne kadar özenle yapılırsa yapılsın birçok farklı nedenden dolayı verilerin tam alınmaması sorunuyla karşı karşıya kalabilmektedir. Kayıp gözlem, en sade hâliyle elde edilmesi planlanan veriler ile elde edilen veriler arasındaki oluşan fark olarak tanımlanmıştır (Longford 2005).

Analizlerde kullanılan istatistiksel yöntemler ve analiz programları verilerin tam olduğu varsayımıyla çalışır. Bütün analiz yöntemleri veri setinin tam olduğu varsayımı altında geliştirilmiştir. Bu nedenle kayıp gözlemler; analizler için kullanılacak olan istatistiksel yöntemler için önemli bir sorun oluşturur (Allison, 2003; Osborne, 2013; Pigott, 2001). Kayıp gözlemler göz ardı edilerek yapılan analizler doğru sonuçlar vermeyebilir. Ayrıca bu sonuçların araştırma örnekleminin evreni temsil etme gücünü azalttığı söylenebilir. Analizler için hazırlanan programlar tam veri setlerine göre oluşturulduğundan, verideki kayıp gözlemler analizlerde taraflı sonuçların ortaya çıkması, örneklemin evreni temsil etme gücünü, standart hatanın artması gibi sorunları ortaya çıkaracaktır. (Acock, 2005; Bal, 2003; Demir ve Parlak, 2012). Bu nedenle yapılacak analizlerden tarafsız ve güvenilir sonuçlar elde edebilmek için toplanan verilerin eksiksiz olması gerektiği bilinmektedir. Maalesef, çoğu veri nadiren eksiksizdir (Cheema, 2012)

Arařtırmalarda kayıp gözlemlerin oluşmasının birden çok sebebi olabilir. Kayıp gözlemler arařtırmacı, yanıtlayıcı veya ölçek kaynaklı oluşabilir. Ölçme araçları tarafından elde edilen veriler, yanıtlayıcıdan, ölçme aracından veya uygulayıcıdan kaynaklanan sebeplerden dolayı her zaman eksiksiz şekilde elde edilememektedir (Işıkođlu, 2017). Brown ve Kros (2003) kayıp gözlemlerin oluşma sebeplerini üç grupta sınıflandırmıştır. Bunlardan ilki veri giriři sırasında karşılaşılan problemleri ele alan süreçle ilgili nedenlerdir. İkinci neden, cevaplamayı reddetme ile ilgili nedenlerdir. Üçüncü neden ise herhangi bir durum, grup veya konuyla ilgisi olmayan soruların sorulmasıdır. Cheema (2012)'ya göre gözlemin dikkatsizliğinden, veri giriři yapılırken yapılan hatalardan ve yanlış işaretlemelerden dolayı kayıp gözlemler oluşabilir. Alan

yazıda belirtilen etkilerden dolayı bir araştırmada kayıp gözlemlerin varlığı azımsanmayacak düzeydedir.

Kayıp gözlemlerin oluşturduğu sorunları ortadan kaldırmak için birden fazla yöntem geliştirilmiştir. Kayıp gözlemler çalışma sonuçlarını farklı şekillerde etkilediği için, araştırmacıların örneklem büyüklüğü, kayıp gözlem oranı vb. faktörleri dikkate alarak uygun yöntem seçimi yapması gerekmektedir (Cheema, 2012). Kayıp gözlemlerle çalışmaya devam edildiği veya uygun yöntem seçilmediği takdirde çalışma sonuçları bu durumdan olumsuz yönde etkilenecektir (Ginkel ve diğerleri, 2007). Bu nedenle toplanan veriler üzerinde herhangi bir analiz yapmadan önce eksik verilerin miktarı ve yapısıyla ilgili gerekli hazırlıklar yapılmalı ve oluşacak olumsuz etki ortadan kaldırılmalıdır.

Araştırmalarda kayıp gözlemlerin varlığı durumunda: (1) yeni gözlemciler eklenmesi, (2) kayıp gözlemleri bulunan gözlemcinin tamamen veri setinden silinmesi veya (3) kayıp gözlemler yerine yaklaşık değer ataması yapılarak, kayıp gözlem sorunuyla baş edilebilir. Günümüzde yapılan çalışmalarda araştırmacılar analizler öncesi kayıp gözlemler yerine yaklaşık değer atamak için birçok farklı yöntem kullanmaktadır. Kayıp gözlem yerine yaklaşık değer atama yöntemleri veri kümesindeki kayıp gözlemlerin yapısı ve miktarına göre belirlenmektedir. Araştırmacıların sıklıkla kullandıkları yaklaşık değer atama yöntemleri başlıca şunlardır;

- (1) Ortalama Atama (Mean Substitution),
- (2) Yakın Noktalar Medyan Ataması (Median of Nearby Points),
- (3) Doğrusal Değerleme (Linear Interpolation),
- (4) Beklenti Maksimizasyonu Algoritması (Expectation Maximization Algorithm),
- (5) Eğilim Skorları Eşleştirmesi (Propensity Score Matching),
- (6) Markov Zincirleri Monte Carlo (Markov Chain Monte Carlo),
- (7) Bayesci Veri Atama (Bayesian Imputation),
- (8) Stokastik Regresyon Ataması (Stochastic Regression Imputation),
- (9) K-Ortalama Kümeleme Ataması (K-Means Clustering Imputation) vb. şeklindedir.

Araştırmacıların birçoğu yapılan çalışmanın niteliğine ve veri setinin büyüklüğüne göre bu yöntemlerden herhangi birini veya birkaçını kullanmaktadır. Günümüzde analiz

programlarının gelişmesi sonucunda geleneksel olarak kullanılan yöntemlerden ziyade çalışmanın beklentilerine cevap verebilecek gelişmiş yöntemler tercih edilmektedir.

Yapılan çalışmaların geçerlik ve güvenilirliğinin kayıp gözlemlerden etkilenmemesi için geleneksel yöntemler yerine yaklaşık değer atama yöntemlerinin kullanılmasının daha etkili olacağı öngörülmektedir. Bu nedenle yaklaşık değer atama yöntemlerinin geliştirilmesi ve yaygınlaştırılmasına önem verilmelidir.

Birçok araştırmada yaklaşık değer atama yöntemlerinin performanslarının karşılaştırılmıştır. (Allison, 2003; Engels ve Diehr, 2003; Köse, 2014; Köse ve Öztemur, 2014; Saunders, Morrow-Howell, Spitznagel, Proctor ve Pescarino, 2006; Çokluk ve Kayrı, 2011; Altaş ve Kaspar, 2012; Demir, 2013; Akbaş ve Tavşancıl, 2015; Koçak ve Çokluk Bökeoğlu, 2017;). Bahsi geçen araştırmaların çoğunda basit atama yöntemleri ve gelişmiş atama yöntemleri hazır veriler üzerinden karşılaştırılmıştır. Analizi en az etkileyecek şekilde birçok kayıp gözlem tamamlama yöntemi geliştirilmiştir. Ancak bu yöntemler hiçbir zaman tam anlamıyla gerçek bir çözüm değildir. Gerçek çözüm, araştırmalarda kayıp gözlem olmaması veya araştırmayı etkilemeyecek düzeyde olması gerekmektedir. Bu nedenle çalışmaların yeterli özveri ve dikkatle yapılması gerekmektedir. Özensiz ve dikkatsiz bir şekilde yürütülen bir araştırmada herhangi bir istatistiksel düzeltmenin bir anlamı da yoktur (Allison, 2002)

1.1. Problem Durumu

Araştırmacılar evreni temsil eden örneklemelerinde her ne kadar eksiksiz veri elde etmeye çalışsalar da bazı nedenlerden dolayı bu mümkün olmamaktadır. Özellikle büyük gruplar ile çalışma yürütüldüğünde eksik veriler ile sıkça karşılaşmaktadır. Başarı testi, tutum ölçeği, anket gibi ölçme aracı uygulamalarında katılımcılar zamanla ilgili veya kişisel nedenlerden dolayı maddelerin tamamına cevap vermeyebilirler. Bazen de araştırmacılar dikkatsizlik nedeniyle verileri eksik kaydedebilirler. Bu nedenlerden dolayı üzerinde analizlerin yapılacağı veri setlerinde eksik bilgi ile karşılaşılması araştırmacıların birçoğunun karşılaştığı durumdur.

Kayıp gözlemler araştırma çalışmasında yapılması planlanan veri çözümleme analizlerinin sonuçları üzerinde etki etmektedir. Çünkü veri analiz tekniklerinin tamamına yakını eksiksiz veri üzerine modellenmiştir. Kayıp gözlemler dikkate alınmadan yapılacak olan analizler ise yanıltıcı sonuçlar verebilir. Ayrıca bu yanıltıcı sonuçlar araştırmadan elde edilen bulguların araştırma evrenine genellemesi noktasında sorunlar oluşturur. Kayıp gözlem

miktarının artması sonuçların güvenilirliğini ve yapılacak istatistiksel analizlerin sonuçlarını büyük ölçüde etkileyecektir. Ayrıca kayıp gözlemlerin varlığı çalışmanın geçerliğini de olumsuz yönde etkileyecektir. Alan yazında kayıp gözlemler bulunan veri setlerine herhangi bir düzenleme olmadan yapılan analizlerin, kayıp gözlemin niteliğine ve miktarına bağlı olarak analiz sonuçlarını etkileyeceği belirtilmektedir. Bu nedenle, elde edilen veriler üzerinden analizler yapılmadan önce kayıp gözlemlerin yapısı ve miktarı ile ilgili araştırmaların yapılması ve oluşabilecek sorunlar için önlemlerin alınması gerekmektedir.

Dolayısıyla araştırmalarda elde edilen veriler üzerinde yapılacak her türlü analizin doğru neticeler ile sonuçlandırılması için eksik veri tamamlama yöntemlerinin geliştirilmesi ve mevcut yöntemlere yeni yöntemlerin eklenmesi alanında çalışmaların yapılması gerekmektedir. Bu amaçla mevcut yöntemlerin karşılaştırılması, yeni yöntem geliştirilmesi ve sağlamlasının yapılması bu araştırmada hedeflenmektedir.

1.2. Araştırmanın Amacı ve Alt Problemler

Çoğu araştırmada çeşitli nedenlerden dolayı kayıp gözlem sorunu oluşmaktadır. Bu sorun yapılan araştırmanın sonuçlarını ve güvenilirliğini etkilemekte, araştırmanın geçerliliğini düşürmektedir. Bundan dolayı kayıp gözlemlerin araştırmada yapılan analizlerin güvenilirliğini artıracak şekilde tamamlanması büyük önem taşımaktadır. Bu nedenle araştırmada kayıp gözlem tamamlamada yeni yöntem geliştirilmesi ve bu yeni yöntem ile yaygın olarak kullanılan kayıp gözlem tamamlama yöntemlerinin (Mod'a göre atama, Medyan'a göre atama, Cramer's V katsayısı) karşılaştırılması amaçlanmaktadır.

Araştırma problemi; Benzerlik temelli atama yöntemi ile basit atama yöntemleri arasında doğru atama yüzdesi ve güvenirlilik katsayısına etkisi açısından fark var mıdır? şeklindedir.

Bu çalışma ile aşağıda yer verilen alt problemler incelenmiştir. Bu alt problemler:

(1) Benzerlik temelli atama yöntemi ile basit atama yöntemlerinin güvenirlilik katsayılarının tam veri setinin güvenirlilik katsayısına yakınlığı arasında fark var mıdır?

(2) Benzerlik temelli atama yöntemi ile basit atama yöntemlerinin kayıp gözlem tamamlamada doğru atama yüzdesi arasında fark var mıdır?
şeklindedir.

1.3. Araştırmanın Önemi

Yapılan her türlü akademik araştırmada kayıp gözlemlerin yoğunluğu ters orantılı olarak çalışmanın güvenilirliğini etkilemektedir. Bundan dolayı kayıp gözlem tamamlama yöntemleri büyük önem taşımaktadır. Birçok değişik alanda yapılan her bir çalışmanın veri türleri de çeşitlilik göstermektedir. Bu çeşitliliğe bağlı olarak veri tamamlama yöntemlerinde sadece birkaç yöntemin kullanışlı olması düşünülemez. Bundan dolayı araştırmacıların karşı karşıya kaldıkları kayıp gözlem türlerine göre yöntemler kullanılması çalışmanın analiz bulgularının geçerliliğini artıracaktır. Kayıp gözlemi belirlemek ve kayıp gözlem bulunduran veri kümelerini analiz etmek için zaman içerisinde çeşitli yöntemler geliştirilmiş olup her geçen gün araştırmacılar tarafından daha kullanışlı yeni yöntemler oluşturulmaktadır. Yeni yaklaşımlar ile kayıp gözlemlerin tamamlanmasında güvenilirlik ve geçerliliği artıran yöntemler ortaya konulmaktadır.

Araştırmacıların çalışmalarında elde ettikleri veri kümelerinin analiz bulgularını yorumlanabilir ve gerçekçi bir şekilde ortaya koyabilmeleri için veri türüne uygun yöntem ve tekniği seçmeleri gerekmektedir. Uygun yöntem ve teknik ile kayıp gözlemler yerine gerçekçi atamalar yapılabilecek ve dolayısıyla analizler daha sağlıklı uygulanabilecektir.

Çalışmada günümüzde yapılan araştırmalarda oluşan kayıp gözlem sorununun çözülebilmesi için yeni bir yöntem geliştirilmesi amaçlanmıştır. Çalışmada yeni yöntemin ağırlıklı olarak kullanılan yaygın kayıp gözlem atama yöntemleri ile karşılaştırılması yapılmıştır. Kayıp gözlem tamamlama yöntemlerinin yapılacak analizlerin güvenilirliğine etkileri ortaya konulmuştur. Güvenirliğin artırılması ile analiz bulgularının daha gerçekçi bir şekilde yorumlanabilmesi için önemle gerekmektedir.

1.4. Araştırmanın Sayıltıları

Bu araştırmada üzerinde çalışma yapılacak veri setinin oluşumunda herhangi bir hatalı veri girişinin yapılmadığı varsayımında bulunulmuştur.

Çalışmanın kapsamıyla ilgili alanyazının genelini tarandığı varsayılmıştır.

1.5. Araştırmanın Sınırlılıkları

Bu araştırma eğitim bilimleri alanında yapılmış ilgi ölçeği çalışmasından elde edilen 3725 denekten alınan SPSS verileri kapsamında yapılacaktır. Yapılacak analizler Cronbach Alpha güvenilirlik analizi ve doğru atama belirleme çalışması ile sınırlıdır.

1.6. Tanımlar

Veri seti: Değişkenle ilgili denek ya da objenin değerlerinin oluşturduğu kümeye denir.

Güvenirlilik: bir niteliğe ait ölçmelerin aynı kişilere uygulandığında benzer şartlarda benzer sonuçları vermesidir.

Geçerlilik: bir ölçme aracının ölçmeyi amaçladığı özelliği, başka herhangi bir özellikle karıştırmadan, doğru ölçebilme derecesidir.

Mod (Tepe Değer): Bir sayısal veri serisi içinde en çok tekrar eden sayıdır.

Medyan (Ortanca). Bir sayısal veri serisi sıralandığında ortada kalan sayıdır.

Cramer's V Katsayısı (Cramer's V): Satır ve sütun sayısından bağımsız, iki değişken arasındaki ilişkinin gücünü ölçer. 0 ile 1 arasında değerler alır. 0 ilişkinin olmadığını gösterir.

Kayıp gözlem: Veri setlerinde bir veya daha fazla değer elde edilememesi ya da eksik olması durumudur.

BÖLÜM II

KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

Araştırmanın bu kısmında akademik çalışmalarda yapılan veri analizlerinde karşılaşılan kayıp gözlem, kayıp gözlem silme yöntemleri, kayıp gözlem atama yöntemleri, farklı kayıp gözlem atama yöntemlerinin güvenilirliğe etkileri ile ilgili kuramsal bilgiler sunulmuş ve bu kuramsal çerçevedeki konu başlıkları ile ilgili yurt içi ve yurt dışı yapılan akademik çalışmaların sonuçlarına yer verilmiştir.

2.1. KURAMSAL BİLGİLER

Bu bölümde kayıp gözlem, kayıp gözlem silme yöntemleri, kayıp gözlem atama yöntemleri, farklı kayıp gözlem atama yöntemlerinin güvenilirliğe etkileri ile bu değişkenler arasındaki ilişkilere yönelik kuramsal bilgiler sunulmuştur.

2.1.1. Kayıp Gözlem

Eğitim alanında yapılan çalışmalarda çeşitli veri toplama araçları ile nicel veriler toplanmakta ve bu veriler üzerinde değişik analizler uygulanmaktadır. Bu analizlerde en önemli göstergelerden biri güvenilirlik bulgularıdır. Elde edilen güvenilirlik katsayısının büyüklüğü ile doğru orantılı bir şekilde çalışma bulgularının geçerliliği artmaktadır. Ancak analizlerin güvenilirliğini olumsuz etkileyen etmenler araştırmacılar için sorun teşkil etmektedir. Bu etmenlerin başında; araştırma örnekleminde toplanan veriler içerisinde kayıp gözlem bulunmasıdır. Matris halinde oluşturulan veri setlerinde hücre birimlerine değişik nedenlerle girilemeyen sayısal değerler kayıp gözlem anlamına gelmektedir (Little & Rubin, 1987).

Akademik çalışmalarda yapılacak olan istatistiksel analizler için kayıp gözlem yoğunluğuna göre sorun yaratır. Çünkü araştırmacı tarafından kayıp gözlem bulunan veri gruplarının analiz dışına çıkarmak istemesidir. Kayıp gözlem gruplarının analiz dışına çıkarılması, veri setinin büyüklüğünü olumsuz etkileyip, olası homojen cevapları yok edebilmektedir. Bu konunun araştırmalar için bir sorun olduğu 1930 lu yıllarda kabul edilmiştir. İstatistik programlarında bu sorunu giderici yöntem ve teknikler geliştirilmeye başlanılmıştır. 1980 li yıllarda önemli gelişmeler sağlanmıştır. İstatistiksel programlarda kayıp gözlem ile ilgili analizler 1990'lı yıllarda oldukça yaygınlaşmıştır (Kürşad ve Nartgün, 2015; Bayhan, 2018). Kayıp gözlemlerin oluşması örneklemin yanıtlamama, boş bırakma davranışından ya da araştırmacının analizler

enasında yaptığı hatalardan kaynaklanmaktadır. Alanyazında yanıtlamama, boş bırakma ve kayıp gözlem işleme aynı anlamda, birbirinin yerine kullanılabilir. Alanyazın incelendiğinde kayıp gözlemlerin, sonuçların güvenilirliğini, geçerliliğini, genellenebilirliği, ortalama ve standart sapma vb. gibi birçok istatistiksel değerlerini olumsuz yönde etkilediği görülmüştür (Little ve Rubin, 1987; Bal, 2003; Hohensinn & Kubinger, 2011). Ölçme araçları tarafından elde edilen veriler, yanıtlayıcıdan, ölçme aracından veya uygulayıcıdan kaynaklanan sebeplerden dolayı her zaman eksiksiz şekilde elde edilememektedir (Işıkoğlu, 2017). Cheema (2012)'ya göre gözlemin dikkatsizliğinden, veri girişi yapılırken yapılan hatalardan ve yanlış işaretlemelerden dolayı kayıp gözlemler oluşabilir.

Brown ve Kros (2003) kayıp gözlemlerin oluşma sebeplerini üç grupta sınıflandırmıştır. Bunlardan ilki veri girişi sırasında karşılaşılan problemleri ele alan süreçle ilgili nedenlerdir. İkinci neden, cevaplamaı reddetme ile ilgili nedenlerdir. Üçüncü neden ise herhangi bir durum, grup veya konuyla ilgisi olmayan soruların sorulmasıdır. Maksimum düzeyde yansız ve güvenilir sonuçlar elde edebilmek için veri setlerinin eksiksiz ve yeterli büyüklükte olması gerektiği bilinmektedir. Maalesef, çoğu veri nadiren eksiksizdir (Cheema, 2012). Kayıp gözlemler analizler için kullanılacak olan istatistiksel yöntemlerin hemen hemen hepsi için önemli bir sorun oluşturur çünkü tüm yöntemler veri setinin eksiksiz olduğu varsayımı altında geliştirilmiştir (Allison, 2003; Osborne, 2013; Pigott, 2001). Kayıp gözlemler dikkate alınmadan yapılacak olan analizler yanıltıcı sonuçlar verebilir. Ayrıca bu sonuçların araştırma evrenine genellenebilirliği noktasında hata miktarının artacağı da söylenebilir (Byrne, 2000).

Çüm. S. (2018) Bu bakımdan, toplanan veriler üzerinden yürütülecek olan ileri düzey analizlere geçilmeden önce kayıp gözlemlerin miktarı ve yapısıyla ilgili incelemelerin yapılması ve karşılaşılabilecek sorunların bertaraf edilmesi için gerekli önlemlerin alınması araştırmacılar için bir zorunluluk olarak ortaya çıkmaktadır.

Araştırma sürecinde meydana gelen kayıp gözlemler analizlerin etkililiği azalmaktadır. Peng, Harwell, Liou ve Ehman (2007)'a göre, kayıp gözlemlerin oluşması nedeniyle araştırmalarda ortaya çıkacak problemler şu şekildedir:

1. Yapılacak olan istatistiksel analizlerden elde edilen bulgularda yanlışlık oluşabilmektedir.
2. Veri seti üzerinde yapılacak olan analizlerin çeşitliliğini azaltabilmektedir.
3. İstatistiksel yöntem çeşitliliğini azaltabilmektedir.

4. Araştırma sürecinde değerlendirilebilir kaynakların kullanılmamasına neden olabilmektedir.

2.1.2. Kayıp Gözlem Nedenleri

Birçok araştırma alanında kayıp gözlem problemleri ile karşılaşmaktadır. Bunun yanı sıra kayıp gözlem sorunu en çok sosyal bilimlerde alanlarında karşılaşmaktadır (Vansteelandt, Carpenter ve Kenward, 2010). Bunun en önemli nedenlerin başında kişisel veri toplama tekniklerinin yoğunlukla kullanılmasıdır.

Bir araştırmada soru adedinin fazla olması nedeni ile cevaplayanların bazı soruları boş bırakması, cevaplayanların dikkatsizliği veya bilinçli olarak soruların boş bırakılması kayıp gözleme neden olan başlıca sebeplerdendir. Brown ve Kros (2003)'e göre, kayıp gözlem oluşma nedenleri 3 gruba ayrılabilir. Birinci olarak araştırma süreci ile ilgili durumları içerir. Bu süreç araştırmacı tarafından veri giriş işlemleri sırasında yapılan hataları kapsamaktadır. İkinci olarak veri toplama süreci ile ilgili durumları içerir. Bu süreç cevaplayanların bilinçli veya rastgele olarak soruları cevapsız bırakmasını kapsamaktadır. Üçüncü süreç ise, veri toplama aracı ile ilgili durumları içerir. Bu süreç uygulanan anket veya ölçeğin bazı sorularının konu ile ilgili olmamasını, ölçeğin uygulandığı örneklemin özellikleri ile ölçek sorularının uygun olmamasını içerir.

McKnight, Sidani ve Figueredo (2007)'e göre, kayıp gözlem oluşma nedenleri 3 gruba ayrılır. Bu nedenler;

- 1- Katılımcılar ile ilgili durumlar
- 2- Çalışma sürecinin tasarımı ile ilgili durumlar
- 3- Katılımcılar ile çalışma süreci arasındaki etkileşimle ilgili durumlar şeklindedir.

Örneğin; ölçek sorularından bazılarının katılımcılar tarafından benimsenmeyip cevaplanmaması, ölçeğin uygulanma süresinin fazla zaman alıyor olması veya katılımcının soruları özel yaşamıyla ilgili görüp cevaplamaması gibi durumlar kayıp gözleme neden olmaktadır.

2.1.3. Kayıp Gözlem Türleri

Kayıp gözlemin oluşma şekli bu kayıp gözlemin tamamlama yönteminin belirlenmesinde ve uygulanmasında önemli bir etkidir. Bu nedenle kayıp gözlemlerin meydana gelme mekanizmaları sınıflandırılmıştır. Kayıp gözlem göstergeleri ilk olarak 1976 yılında formüle edilmeye başlanmıştır. Kayıp gözlem oluşma mekanizmaları 3 temel kategoriye

ayrılmıştır. (1) Tamamıyla rassal olan kayıp gözlemler (TROK), (2) Rassal olan kayıp gözlemler (ROK) ve (3) İhmal edilemez kayıp gözlemler (IEK) şeklindedir.

2.1.3.1. Tamamıyla Rassal Olan Kayıp (TROK)

Bir veri seti içerisinde herhangi bir verinin boş kalması yani kayıp gözlem olma olasılığının, bu verinin yer aldığı değişkenden ve diğer verilerin yer aldığı farklı değişkenlerden bağımsız bir şekilde oluşmasıdır (Little ve Rubin, 1987). Tüm değişkenleri için bu önermeyi doğrulayan bir veri setinin bulunduğu örneklem, çalışmanın yapılacağı tam veri setinin rastgele seçilmiş örnekleme olarak kabul edilebilir (Allison, 2001 & Enders, 2010). Trok mekanizmasının formüle edilmiş hali şu şekildedir;

A= Gözlenen değişken

B= Kayıp gözlem içeren değişken olmak üzere

TROK mekanizması $P(B \text{ kayıp} | A, B) = P(B \text{ kayıp})$

Bu tür kayıp gözlemler, yer aldığı değişkenin düşük veya yüksek olmasıyla bağlantısı bulunmayan ve veri setinin diğer değişkenlerinin etkisi ile oluşmayan gözlemlerdir. Tamamıyla rastlantısal olarak oluşan kayıp gözlemlerdir. Örneğin aylık gider miktarlarından oluşan bir değişkenden yüksek veya düşük giderlerde aynı düşüklük ya da yükseklikte kayıp gözlemlerin oluşmamasıdır. Bunun yanı sıra oluşan bu kayıp gözlemin çalışmanın yaş veya cinsiyet gibi değişkenleri ile ilişkisinin olmaması da bu kayıp gözlemin tamamıyla rastsal olarak (TROK) dağıldığını gösterir.

2.1.3.2. Rassal Olan Kayıp Gözlemler (ROK)

Bir veri seti içerisinde herhangi bir verinin boş kalması yani kayıp gözlem olma olasılığının, bu verinin yer aldığı değişkenden bağımsız olması ancak veri setinin diğer değişkenlerine bağımlı olması durumuna rassal kayıp gözlem denir (Little & Rubin, 1987). TROK mekanizmasında olduğu gibi rassal kayıp gözlem oluşumu da formüle edilecek olursa;

A= Gözlenen değişken

B= Kayıp gözlem içeren değişken olmak üzere

$P(B \text{ kayıp} | A, B) = P(B \text{ kayıp} | A)$ şeklindedir.

Bir veri setindeki herhangi bir değişkendeki kayıp gözlemlerin diğer değişkenlerle ilgili olup değişkenin kendisinden bağımsız olması ile oluşan kayıp gözlemlerdir. Örneğin aylık giderin cinsiyetin harcama türlerine bağlı olarak kayıp gözlemlerin oluşması ancak gider miktarının

yüksek ya da düşük olmasıyla bir ilişkisinin olmaması durumunda bu kayıp gözlemlerin rassal olan kayıp gözlemler olduğu söylenebilir.

Kayıp gözlemlerin tamamıyla rassal ya da rassal olarak oluşumlarının belirlenmesinde üç farklı yöntem kullanılır. Bunlar; (1) Kayıp gözlem bulunan ve bulunmayan veri grupları arasındaki ortalama farkına bakıldığı t-testidir. T-testi sonucunda anlamlı bir fark çıkmaması kayıp gözlemlerin tamamen rassal olarak (TROK) dağıldığını göstermektedir. (2) Kayıp gözlem bulunan ve bulunmayan veri grupları arasındaki Pearson korelasyon katsayısını karşılaştırmaktadır. Düşük korelasyonların var olması rassallığı göstermektedir. (3) Little'ın TROK testidir. TROK testinde, aynı kayıp gözlem dizilimine sahip gözlemlerin alt gruplarının ortalama farklılıklarını karşılaştırmaktadır. Bir başka deyişle; Alt grupların ortalamaları ve genel ortalama arasındaki farkların ağırlıklı toplamıdır. Bu testte anlamlı fark çıkmaması kayıp gözlemlerin tamamen rassal olarak dağıldığını göstermektedir (Baygül, 2007; Enders, 2010).

2.1.3.2. İhmal Edilemez Kayıp Gözlemler (IEK)

Bir veri setinde yer alan kayıp gözlemin bulunduğu değişkene bağlı olarak oluşmasına ihmal edilemez kayıp gözlem denilir (Little & Rubin, 1987). Çalışmadaki diğer değişkenlerin etkileri sıfırlandığında, bir değişkende yer alan kayıp gözlemler aynı değişkende yer alan tam veri durumlarından önemli ölçüde farklılık gösteriyorsa oluşan kayıp gözlemler ihmal edilemez şeklindedir (Cheema, 2012). TROK mekanizmasında olduğu gibi ihmal edilemez kayıp gözlem oluşumu da formüle edilecek olursa;

A= Gözlenen değişken

B= Kayıp gözlem içeren değişken olmak üzere

$P(B \text{ kayıp} | A, B) = P(B \text{ kayıp} | B)$ şeklindedir.

Kayıp gözlemlerin olduğu çalışmalarda araştırmacılar, (1) yeni tekrar gözlemlerin yapılması, (2) kayıp gözlem bulunan veri gruplarının çalışmanın veri setinden çıkartılması, (3) kayıp gözlemlere ilişkin yaklaşık değerlerin belirlenip kayıp gözlemler yerine kullanılması yöntemlerinden birisi ile kayıp gözlemlerin çalışma analizlerini olumsuz etkilemesine önlem alabilirler. Bu yöntemlerin arasında yeni gözlem yapılıp çalışmaya eklenmesi hem zaman hem de kaynak bakımından yeni masraflar ortaya çıkaracağından dolayı elverişli bir yöntem değildir. Bunun yanı sıra kayıp gözlemlerin bulunduğu veri gruplarının çalışma veri setinden çıkartılması yöntemi veri miktarının azalmasına ve çalışmanın güvenilirliğine olumsuz etkisi olacaktır. Bu istatistiksel analizlerin gücünün azalmasına yol açacaktır (Roth, 1994; Alpar, 2011). Oluşacak olan bu sorunlardan dolayı kayıp gözlemler yerine yaklaşık değerlerin belirlenmesi ve bu değerlerin

kayıp gözlemler yerine kullanılması her yönden daha verimli ve istatistiksel analizlerin gücünü arttırıcı yöntem olacağı ortaya çıkmaktadır.

2.1.4. Kayıp Gözlem Tamamlama Yöntemleri

Araştırmacılar tarafından çalışmalarında oluşan kayıp gözlemlere yönelik yaklaşık değerlerin belirlenmesi ve kayıp gözlemlerin bu değerlerle tamamlanması çalışma sürecinde zaman ve kaynak tasarrufu sağlamaktadır. Bu nedenle araştırmacılar tarafından elde ettikleri veri çeşitliliğine göre farklı yöntemler kullanılmaktadır. Bu yöntemler arasında araştırmacılar tarafından ağırlıklı olarak kullanılan yöntemler; Ortalama atama, mod atama, medyan atama, doğrusal değerlendirme gibi basit atama yöntemleridir. Bu yöntemlerin yanı sıra daha ileri düzey olarak nitelendirilen beklenti maksimizasyonu algoritması, çoklu atamaya dayalı yöntemler olan eğilim skorları eşleştirmesi gibi yöntemler kullanılmaktadır. Kayıp gözlem miktarının toplam veri seti içerisinde yoğunluğunun az olması ve TROK olması durumunda basit atamaya dayalı yöntemler etkili olabilir. Ancak farklı kayıp gözlem oluşma durumlarında yeni geliştirilen ve kompleks işleyen yöntemlerin kullanılması daha güvenilir sonuçlar verecektir (Osborne, 2013; Schafer, 1999). Söz konusu kayıp gözlem atama yöntemlerinden bazıları şu şekildedir.

2.1.4.1. Eğilim Skorları Eşleştirmesi

Regresyon yardımı ile tahmini eğilim skorları eşleştirilmiş veri setlerinin oluşturulması için bir işlemdir. Eşit veya birbirine yakın eğilim skoruna sahip birimlerin, ortak değişkenlerine bağlı olarak, aynı (veya benzer) dağılıma sahip olma eğiliminde olacağı varsayılır (Kaspar, 2011).

2.1.4.2. Bayesci Veri Tamamlama

Bayesci veri tamamlama yönteminde, rasgele bir sürece bağlı olarak meydana gelen rastgele bir A olayı ile diğer bir rastgele B olayı için koşullu olasılıklar ve sıra dışı olasılıklar arasındaki ilişkiyi açıklamaktadır (Çüm S. & ark.,2018).

2.1.4.3. Beklenti Maksimizasyonu Algoritması

Beklenti maksimizasyonu süreci beklenti ve maksimizasyon adımıyla oluşan iki aşamalı yenilemeli bir yöntemdir. Beklenti adımıyla kayıp gözlemler yerine regresyon

yöntemiyle yaklaşık değerler atanır. Maksimizasyon sürecinde ise tamamlanmış olan veri üzerinden öngörüler yenilenir. (Hedderley ve Wakeling, 1995).

2.1.4.4. Regresyon Ataması Yöntemi

Bu yöntemde veri seti, kayıp gözlem içeren verilerin yordanan; kayıp gözlem içermeyen verilerin yordayıcı konumda olduğu regresyon hesaplamalarıyla tamamlanmaktadır (Enders, 2010). Little ve Rubin (1987) tarafından yapılan çalışmaya göre kayıp gözlemler yerine bilinçsizce atanan değerler sorunları ortadan kaldırmadığı, bunun yanı sıra ortaya çözümü daha güç olan yeni sorunlar meydana getirmektedir. Örneğin, kayıp gözlemlerin TROK mekanizmasına uygun olmadığı durumlarda basit atama yöntemleri ile yapılacak bir tamamlama sonrası kayıp gözlemlerden dolayı meydana gelen yanlılık sorunu ortadan kaldırılmamış olabilir (Çüm S. & ark., 2018).

2.2. İLGİLİ ARAŞTIRMALAR

2.2.1. Kayıp Gözlem Atama Yöntemleri İle İlgili Yurtiçi Çalışmalar

Bayhan A. (2018) tarafından yapılan çalışmada farklı oranlarda oluşan kayıp gözlem içeren veri setlerinin güvenilirlik değerindeki değişiklik, test uzunluğu, örneklem büyüklüğü ve puanlama türü değişkenleri açısından incelenmesi amaçlanmıştır. Çalışmanın örneklemini wigen 3 programı koşulları altında 20'li tekrarlar halinde üretilmiş ve rasgele şekilde %5, %10, %20 oranlarında silinerek elde edilmiş kayıp gözlem seti oluşturmaktadır. Çalışma ilişkisel bir çalışma olup Cronbach Alpha güvenilirlik analizleri yapılmıştır. Elde edilen bulgulara göre, örneklem büyüklüğünün artması ile güvenilirlikteki değişiminin azaldığı, benzer şekilde madde sayısı arttıkça da güvenilirlik değişiminin azaldığı görülmüştür. Bunun yanı sıra örneklem büyüklüğü, testin uzunluğu ve dağılım biçimi farklı olsa bile kayıp gözlem oranının artmasıyla güvenilirlikteki değişimin de arttığı sonucuna varılmıştır.

Akbaş U. (2014) tarafından yapılan çalışmada ölçeklerin psikometrik özelliklerinin değişik kayıp gözlem atama yöntemleriyle incelenmesi amaçlanmıştır. Çalışmanın örneklemini 250, 500 ve 1000 çalışma grubu büyüklüğüne sahip ve %10 ila %15 oranlarında kayıp gözlem meydana getirilen veri setleri oluşturmaktadır. Tam veri setleri üzerinde liste bazında silme, Öklid uzaklığına dayalı benzer tepki örüntüsüne göre atama, skolastik regresyonla atama ve çoklu atama teknikleri kullanılarak kayıp gözlem setleri elde edilmiştir. Örneklem üzerinde güvenilirlik analizleri, geçerlilik analizleri, toplam varyans ve doğrulayıcı faktör analizleri gerçekleştirilmiştir. Elde edilen bulgulara göre, liste bazlı veri silmenin önemli problemlere yol

açabileceği, beklenti-maksimizasyon algoritması ve çoklu değer atama yöntemlerinin genel olarak yüksek oranda iyi performans sergilediği görülmüştür. Bunun yanı sıra tüm kayıp gözlem durumlarında kullanılabilir ve kesin olarak en iyi sonuçları veren tek bir yöntemin bulunmadığı sonucuna varılmıştır.

Ak D.(2020) tarafından yapılan çalışmada, kayıp gözlem setlerinde G kuramında varyans bileşenlerinin ve güvenilirlik katsayılarının belirlenmesi için kayıp gözlem atama yöntemlerinin karşılaştırılması olarak incelenmesi amaçlanmıştır. Çalışmanın örneklemini 689 kişilik bir çalışma grubu ve 4 kişilik bir puanlayıcı oluşturmaktadır. Tüm kişiler test maddelerinin tümünü yanıtlamış ve puanlayıcılar tüm kişileri maddelerin tamamında puanlamıştır. Veri matrisi random bir desen olarak oluşturulmuştur. R programı aracılığı ile veri setinde %5, %10, %20 oranlarında seçkisiz şekilde kayıp gözlem elde edilmiştir. Örneklemden elde edilen veriler üzerinde varyans analizi, Phi katsayısı hesaplama, uyum RMSE analizleri gerçekleştirilmiştir. Elde edilen bulgulara göre, LBS yönteminin kayıp gözlem oranlarında referans değere en yakın sonuçları verdiği, OA yönteminin tüm kayıp gözlem oranlarında evren puanı varyansı ve mutlak hata varyansında farklılaşmanın en fazla olduğu görülmüştür. Ayrıca tüm istatistiklerde ve kayıp gözlem atama yöntemlerinde kayıp gözlem oranı arttıkça az da olsa hata değerlerinin arttığı sonucuna varılmıştır.

Arıkan Ç.A. & Soysal S.(2018) tarafından yapılan çalışmada, güvenilirlik katsayısı kestirimlerinin kayıp gözlem tamamlama yöntemlerine göre çalışma grubu büyüklüğü, kategori sayısı ve kayıp gözlem oranları açısından incelenmesi amaçlanmaktadır. Çalışmanın örneklemini 500 ve 5000 büyüklüğünde 20 maddeden oluşan tek faktörlü iki veri seti oluşturmaktadır. Bu örneklem üzerinde %5, %10, %20 ve %30 oranlarında veri silme işlemi yapılarak seçkisiz biçimde kayıp gözlem setleri elde edilmiştir. Kayıp gözlemlere regresyon, ortalama ve çoklu atama yöntemleri ile değerler atanmıştır. Güvenirlik analizleri yapılarak farklı yöntemler arasında karşılaştırma gerçekleştirilmiştir. Elde edilen bulgulara göre, kayıp gözlem atama yöntemine bağlı olarak kayıp gözleme daha dayanıklı olan tek bir katsayının olmadığı belirlenmiştir. Kayıp gözlem tamamlama yöntemleri arasında çoklu atama ve regresyon atama yöntemlerinin kullanımının güvenilirlik açısından sorun yaratmadığı görülmüştür. Bunun yanı sıra sıfır atama yönteminin kullanılmaması sonucuna varılmıştır.

Bayram İ. (2020) tarafından yapılan çalışmada, değişik test uzunluklarında, değişik örneklem büyüklüğüne sahip veri setlerinde uygulanan kayıp gözlem tamamlama yöntemlerinin güvenilirlik bulguları üzerine etkisinin incelenmesi amaçlanmıştır. Araştırmanın çalışma grubunu farklı büyüklükteki yapay veri setleri oluşturmaktadır. Çalışma verileri üzerinde dizin silme, regresyon atama, çoklu atama ve beklenti maksimizasyonu kayıp gözlem

atama yöntemleri uygulanmıştır. Bu yöntemlerle tamamlanan tam veri setlerinde Cronbach Alpha güvenilirlik analizleri gerçekleştirilmiştir. Elde edilen bulgulara göre, çalışmada uygulanan farklı kayıp gözlem atama yöntemlerinin veri setlerinin güvenilirlikleri arasında anlamlı bir fark olmadığı, referans değerlere en yakın sonuçları regresyon ve çoklu atama yöntemlerinin verdiği sonuçlarına varılmıştır.

Kürşad Ş.M. & Nartgün Z. (2015) tarafından yapılan çalışmada, kayıp gözlem probleminin çözümünde araştırmacılar tarafından kullanılan farklı veri atama yöntemlerinin geçerlilik ve güvenilirlik üzerindeki etkileri açısından karşılaştırılması amaçlanmıştır. Çalışmada veri toplama aracı olarak PISA 2012 “Matematik ve çalışma Etiği” ölçeği kullanılmıştır. Bu ölçek ile 200 kişilik Türkiye örneklemini araştırmanın çalışma grubunu oluşturmaktadır. Kayıp gözlem atamalarında seri ortalaması, yakın noktaların ortalaması, yakın noktaların medyanı, doğrusal değer kestirimi, noktanın doğrusal eğimi, liste bazında silme, beklenti maksimizasyonu, regresyon ataması ve çoklu atama kayıp yöntemleri uygulanmıştır. Elde edilen tam veri setlerinde geçerlik ve güvenilirlik analizleri yapılmıştır. Yapılan analizlere göre, liste bazında silme yöntemi için elde edilen değerler genel olarak tam veri setinden elde edilen değerlere en az benzerlik gösteren değerler olduğu, en yakın değer veren yöntemlerin çoklu atama ve regresyon atama yöntemlerinin olduğu sonucuna varılmıştır.

2.2.2. Kayıp Gözlem Atama Yöntemleri İle İlgili Yurtdışı Çalışmalar

Enders (2004) tarafından yapılan çalışmada, tamamıyla seçkisiz kayıp gözlem altında liste bazında silme, ortalama değer atama ve beklenti maksimizasyon algoritması yöntemlerinin Cronbach Alpha güvenilirlik katsayısı üzerindeki etkilerinin karşılaştırılması amaçlanmıştır. Araştırmanın çalışma grubunu 1000 veri bulunduran yapay veri seti oluşturmaktadır. Örneklemdaki kayıp gözlemlerin tamamlama yöntemlerinin yanlılık, hataların ortalama karekökü ve güven aralığı açısından incelenmiştir. Kayıp gözlemler beklenti maksimizasyonu algoritması, liste bazında silme ve çift bazında silme yöntemleri kullanılmıştır. Yapılan analizler ile liste bazında silme ve çift bazında silme tekniklerinin seçkisiz kayıp koşulu için isabetli sonuçlar verdiği ancak seçkisiz kayıp koşulu altında kabul edilebilir sınırın altında olduğu sonucuna varılmıştır. Ayrıca ortalamanın atanması yöntemiyle elde edilen sonuçların her iki durumda büyük ölçüde isabetsiz olduğu belirlenmiştir.

Bernaards ve Sijtsma (1999) tarafından yapılan çalışmada, çok boyutlu birden fazla kategorili gizil değişken modeline uygun olarak oluşturulmuş yapay veri seti çalışma grubudur. Kayıp gözlem oranı %5, %10 ve %20 olan veri setleri üzerinde liste bazında silme, beklenti

maksimizasyon algoritması, rastgele deęer atama, genel ortalamannın atanması, ortak deęişkenlere göre ortalama deęer atama, faktörleştirme yöntemi, madde ortalaması atama ve eksen döndürme yöntemleri kullanılmıştır. Yapılan analizler ile bütün durumlar için beklenti maksimizasyon algoritması deęerlerinin ortalamasının ve standart sapmasının en az olduęu sonucuna varılmıştır.

Chen, Wang & Chen (2011) tarafından yapılan çalışmada, tamamıyla seçkisiz kayıp gözlem koşulu altında %10, %20, %30 ve %50 oranlarında kayıp gözlem içeren veri setleri çalışma gruplarıdır. Kayıp gözlem atama yöntemlerinden liste bazında silme, seri ortalaması atama, doğrusal eğim, doğrusal interpolasyon, regresyon ve beklenti maksimizasyon algoritması yöntemi, açımlayıcı ve doğrulayıcı faktör analizleri açısından karşılaştırılmalı incelenmesi amaçlanmıştır. Yapılan analizler ile faktör analizi varyans oranının %54 ile %99 arasında deęiştiięi, uygulanan farklı kayıp gözlem atama yöntemlerinin benzer sonuçlar ürettięi sonucuna varılmıştır.

Enders & Bandalos (2001) tarafından yapılan çalışmada, kayıp gözlem veri atama yöntemlerinden liste bazında silme, benzer tepki örüntüsüne dayalı atama ve çift bazında silme yöntemlerinin doğrulayıcı faktör analizleri açısından karşılaştırılmalı incelenmesi amaçlanmıştır. Araştırmada çalışma grubu olarak %2, %5, %10, %15 ve %25 oranlarında kayıp gözleme sahip 100, 250, 500 ve 750 büyüklüğünde veri seti kullanılmıştır. Yapılan analizler ile çift bazında silme ve tam bilgi maksimum olabilirlik yöntemlerinin en yüksek yakınsama oranını gösterdikleri; liste bazında silme yönteminin en düşük yakınsama oranını gösterdikleri sonucuna varılmıştır.

BÖLÜM III

YÖNTEM

Yöntem bölümünde araştırma modeli, araştırma çalışma grupları, veri toplama araçları, veri toplama süreçleri ile toplanan araştırma verileri ile verilerin çözümlene ve analizlerde kullanılan istatistiksel teknikler ile ilgili açıklamalar yer almaktadır.

3.1. Araştırmanın Modeli

Bu araştırma kayıp gözlemler yerine yaklaşık değer atama yöntemlerinin etkililiğinin karşılaştırıldığı bir temel araştırmadır. Temel araştırmalar olaylar arasındaki ilişkileri keşfetmek ve teori geliştirmek amacıyla yapılan deneysel nitelikte çalışmalardır.

3.2. Çalışma Grupları

Araştırmada kullanılan tam veri setini, Akdeniz Üniversitesi Eğitim Bilimleri alanında yapılan ilgi ölçeğini eksiksiz cevaplayan 3. Sınıf öğrenciler (n=3725) arasından rastgele sayılar tablosu ile seçilen 450 öğrenciye ait veriler oluşturmaktadır. Güvenirlik katsayıları için ideal örneklem büyüklüğü incelendiğinden alan yazıda farklı değerlerde örneklem büyüklüklerinin ideal olarak belirtildiği görülmüştür. Bu tür çalışmalarda 300-400 örneklem büyüklüğünün yeterli olduğu, ancak 500 örneklem büyüklüğünün en kesin sonuçları vermektedir (Charter,1999; Yurdugül, 2008). Bu nedenle yapılacak araştırma için yeterli örneklem büyüklüğü 450 olarak alınmıştır.

Araştırma örneğine ait demografik bilgiler Tablo- 3.1.'de verilmiştir.

Tablo 3.1. Örneklem Demografik Bilgileri Analizi Bulguları

Yaş	Sınıf	Cinsiyet	Toplam Öğrenci Sayısı (f)	Yüzde (%)
20-21	3. sınıf	Erkek	171	38
20-21	3. sınıf	Kadın	279	62

* Tablo 3.1.'de görüldüğü gibi, örnekte bulunan 450 öğrencinin 171 (%38)'i erkek, 279 (%62)'u kadın olup, öğrencilerin tümü 3.sınıftır.

Araştırma örnekleminde elde edilen veriler üzerinde gerçekleştirilecek analizlere uygunluğunu belirlemek amacıyla merkezi eğilim, basıklık ve çarpıklık analizleri yapılmış, elde edilen bulgular Tablo 3.2.'de verilmiştir.

Tablo 3.2. Araştırma Grubunun Merkezi Eğilim Ölçülerinin Basıklık/Çarpıklığın İncelenmesi

Ölçekler	Kolmogorov-Smirnov	Skewness Katsayısı	Kurtosis Katsayısı
İlgi Ölçeği	,080	-,617	-,261

* Araştırma grubunda bulunan 450 öğrenciden elde edilen verilere normallik testi uygulanmıştır. Yapılan normallik testi sonucunda çarpıklık katsayıları Skewness= -,617 ve Kurtosis= -,261 ($-1 < \text{Ç.K.} > 1$) olduğu, Kolmogorov-Smirnov p-değerinin $x=,080$ ($>,05$) olduğu görülmüştür. Çarpıklık katsayılarının -1 ile $+1$ aralığında olması, puanların normal dağılımdan önemli bir sapma göstermediğini ayrıca grup büyüklüğünün 50'den büyük olması durumunda bakılan Kolmogorov-Smirnov p-değerinin, 05'ten büyük çıkması ($p=,080 >,05$) bu anlamlılık düzeyinde verilerin normal dağılımdan anlamlı sapma göstermediği, uygun olduğunu göstermektedir. (Büyüköztürk, 2018)

3.3. Veri Toplama Araçları

Araştırmada kullanılan veri toplama aracı Akdeniz Üniversitesi Eğitim Fakültesi öğrencilerinin okul içi yaşantılarını ve eğilimlerini ölçen 18 soruluk tutum ölçeğidir. Tek faktörlü bir yapıya sahip olan bu ölçek, “Kesinlikle katılıyorum (5)”, “Katılıyorum (4)”, “Kararsızım (3)”, “Katılmıyorum (2)”, “Kesinlikle katılmıyorum (1)” şeklinde 5’li sınıflama ölçeği formatındadır.

3.4. Veri Analizi

450 öğrenciden oluşan tam veri setinden, tamamen rassal olarak kayıp (TROC) mekanizmasına uygun, belirli oranlarda (%5, %10, %15) veriler silinerek eksik veri setleri oluşturulmuş ve daha sonra bu setler kayıp gözlem sorununun çözümünde kullanılan farklı yöntemler ile yeni tam veri setlerine dönüştürülmüştür. Tam veri setine dönüştürülürken eksik veri bulunan maddelerin mod ve medyanlarına göre, Cramer’s V katsayısı ile hesaplanan en yakın kişinin verdiği cevaba göre ve eksik veri bulunan kişiye en yakın cevapları veren kişilerin tespit edilip onların verdiği cevaplara göre eksik veriler tamamlanmıştır.

Çalışmada tam veri seti üzerinde TROK mekanizması ile rastgele bazı veriler silinmiş ve kayıp gözlemler oluşturulmuştur. Öncelikle Excel programında tüm veri setinin %5'i sıfır(0) geri kalan %95'i bir (1) olacak şekilde rasgele sayılar tablosu oluşturulmuştur. Tam veri seti ile oluşturulan rasgele sayılar tablosu çarpılarak veri setinde değeri sıfır olan bölümler oluşturulmuştur. Daha sonrasında sıfır olan değerler silinerek %5'i eksik verilerle yeni veri seti elde edilmiştir. Aynı işlemler %10 ve %15 kayıp gözlem olacak şekilde tekrarlanmıştır. %5'i eksik gözlem verili veri seti tamamen rastlantısal olarak kayıp (TROK) veri olduğunu tespit etmek için ise Little'in TROK testi uygulandı. Bu test sonucunda anlamlı fark çıkmazsa veri yapısı TROK olarak kabul edilir (Baygöl, 2007; Kaspar, 2011; Şahin Kürşad, 2014) Çalışmada TROK mekanizması ile oluşturulan kayıp gözlemlerin uygunluğu ile ilgili Little's MCAR testi yapılmış ve elde edilen analiz sonuçları Tablo-3.3 de verilmiştir.

Tablo-3.3 Kayıp Gözlem Little's MCAR Testi Bulguları

Kayıp gözlem Yüzdesi	Test	X ²	df	Sig.
%5	Little's MCAR	1630,785	1721	,940
%10	Little's MCAR	3533,592	3469	,218
%15	Little's MCAR	4834,763	4830	,478

*Tablo-3.3 'de elde edilen bulgulara göre tam veri seti üzerinde; %5 kayıp gözlem oluşturulmasının TROK mekanizmasına uygun olduğu ($x= ,940 > .005$) görülmüştür. %10 kayıp gözlem oluşturulmasının TROK mekanizmasına uygun olduğu ($x= ,218 > .005$) görülmüştür. %15 kayıp gözlem oluşturulmasının TROK mekanizmasına uygun olduğu ($x= ,478 > .005$) görülmüştür.

TROK varsayımını test etmenin bir diğer yolu Little's MCAR testidir.

Bu test her değişkendeki ortalama farkları eş zamanlı olarak değerlendiren çok değişkenli bir t testi yaklaşımıdır. Tek değişkenli t testlerinin aksine bu tüm veri kümesine uygulanan bir TROK testidir. Little's MCAR testi $p<.005$ olması durumunda verinin TROK olmadığını gösterir (Enders, 2010).

3.4.1. Kayıp Gözlem Oluşturulması

Tam veri seti üzerinde farklı büyüklüklerde (%5, %10 ve %15) kayıp gözlem oranına sahip veri setleri TROK mekanizmasına uygun olarak oluşturulmuştur. Veri setinde oluşan % 5 oranında kayıp gözlem madde yüzdeleri Tablo-3.4 'de verilmiştir.

Tablo-3.4. %5 Kayıp Gözlem Setinin Madde Dağılımı

Madde	Cevaplanan veri	Kayıp gözlem	Kayıp gözlem yüzdesi
Madde 1	426	24	5,3
Madde 2	426	24	5,3
Madde 3	431	19	4,2
Madde 4	431	19	4,2
Madde 5	436	24	5,3
Madde 6	429	21	4,7
Madde 7	418	32	7,1
Madde 8	441	9	2,0
Madde 9	420	30	6,7
Madde 10	421	29	6,4
Madde 11	426	24	5,3
Madde 12	435	15	3,3
Madde 13	423	27	6,0
Madde 14	430	20	4,4
Madde 15	426	24	5,3
Madde 16	428	22	4,9
Madde 17	430	20	4,4
Madde 18	424	26	5,8
Toplam	7701	409	5

*Tablo-3.4 de görüldüğü gibi, Madde-1 ile Madde-18'e kadar en az %2,0 en çok % 6,7 olmak üzere farklı oranlarda kayıp gözlem bulunmaktadır.

Veri setinde oluşan % 10 oranında kayıp gözlem madde yüzdeleri Tablo-3.5 'de verilmiştir.

Tablo-3.5 %10 Kayıp Gözlem Setinin Madde Dağılımı

Madde	Cevaplanan veri	Kayıp gözlem	Kayıp gözlem yüzdesi
Madde 1	404	46	10,2
Madde 2	401	49	10,9
Madde 3	401	49	10,9
Madde 4	408	42	9,3
Madde 5	407	43	9,6
Madde 6	408	42	9,3
Madde 7	403	47	10,4
Madde 8	402	48	10,7
Madde 9	412	38	8,4
Madde 10	399	51	11,3
Madde 11	412	38	8,4
Madde 12	400	50	11,1
Madde 13	408	42	9,3
Madde 14	399	51	11,3
Madde 15	415	35	7,8
Madde 16	413	37	8,2
Madde 17	405	45	10,0
Madde 18	400	50	11,1
Toplam	7297	803	10

*Tablo-3.5' de görüldüğü gibi, Madde-1 ile Madde-18'e kadar en az %8,4 en çok % 11,3 olmak üzere farklı oranlarda kayıp gözlem bulunmaktadır.

Veri setinde oluşan % 15 oranında kayıp gözlem madde yüzdeleri Tablo-3.6 'da verilmiştir.

Tablo-3.6 %15 Kayıp Gözlem Setinin Madde Dağılımı

Madde	Cevaplanan veri	Kayıp gözlem	Kayıp gözlem yüzdesi
Madde 1	368	82	18,2
Madde 2	377	73	16,2
Madde 3	378	72	16,0
Madde 4	386	64	14,2
Madde 5	381	69	15,3
Madde 6	386	64	14,2
Madde 7	382	68	15,1
Madde 8	377	73	16,2
Madde 9	399	51	11,3
Madde 10	381	69	15,3
Madde 11	370	80	17,8
Madde 12	371	79	17,6
Madde 13	380	70	15,6
Madde 14	384	66	14,7
Madde 15	384	66	14,7
Madde 16	396	54	12,0
Madde 17	386	64	14,2
Madde 18	384	66	14,7
Toplam	6870	1230	15

*Tablo-3.6' da görüldüğü gibi, Madde-1 ile Madde-18'e kadar en az %12,0 en çok % 18,2 olmak üzere farklı oranlarda kayıp gözlem bulunmaktadır.

Çalışmanın yapıldığı veri setinde bulunan 8100 toplam veriden %5, %10 ve %15 oranlarında kayıp gözlem oluşturulmuştur. Elde edilen kayıp gözlemlerin Trok mekanizmasına uygun olduğu belirlenmiş olup, bu kayıp gözlemlere tanıtıcı istatistiksel yöntemler, Cramer's V katsayısı ile tamamlama ve geliştirilen benzerlik temelli atama yöntemi ile tamamlama yapılmıştır.

3.4.2. Tanıtıcı İstatistiksel Yöntemler ile Kayıp Gözlem Atama

Çalışmada farklı oranlarda TROK mekanizmasına uygun olarak oluşturulan kayıp gözlemlerin tamamlanmasında tanıtıcı istatistiksel hesaplamalara dayalı olan mod'a göre atama, medyan'a göre atama ve Cramer's V katsayısı ile hesaplanan en yakın kişinin verdiği cevaba göre atama yapılmıştır. Veri setinde bulunan kayıp gözlemlerin tamamlanmasında diğer araştırma verilerinin bütünüyle dikkate alınıp, verilerin dağılımına göre atama yapılması oldukça önemlidir. Bunu sağlayacak farklı yöntemler araştırmacılar tarafından sıklıkla kullanılmaktadır (Çiğdem, 2011). Kullanılan bu yöntemlere her geçen gün yeni geliştirilen yöntemler eklenmektedir. Çalışmada benzerlik temelli atama yöntemi geliştirilerek, diğer

tanıtıcı istatistiksel yöntemlere karşılaştırılması yapılmıştır. Bu karşılaştırmada, kullanılan yöntemlerin sağladıkları doğruluk yüzdeleri ve güvenilirlik katsayısında tam veri güvenilirlik katsayısına olan yakınlıkları incelenmiştir.

%5,%10 ve %15 kayıp gözlem setinde maddelerin mod ve medyanlarına göre eksik veriler tamamlanmıştır. Ayrıca sadece %5 kayıp gözlem seti Cramer's V katsayısı ile tamamlama yapılmış ve tam veri setine dönüştürülmüştür.

%5 oranında kayıp gözlem yer alan veri setinde tanıtıcı istatistiksel yöntemlere (mod ve medyan) göre yapılan atama dağılımı Tablo-3.7 'de verilmiştir.

Tablo-3.7 %5 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı

Madde	Cevaplanan Madde (N)	Kayıp gözlem	Medyan	Mod
Madde 1	426	24	3,0	3,0
Madde 2	426	24	3,0	3,0
Madde 3	431	19	3,0	3,0
Madde 4	431	19	3,0	3,0
Madde 5	426	24	4,0	4,0
Madde 6	429	21	4,0	5,0
Madde 7	418	32	4,0	4,0
Madde 8	441	9	4,0	4,0
Madde 9	420	30	4,0	5,0
Madde 10	421	29	4,0	5,0
Madde 11	426	24	4,0	5,0
Madde 12	435	15	3,0	3,0
Madde 13	423	27	4,0	4,0
Madde 14	430	20	4,0	4,0
Madde 15	426	24	4,0	4,0
Madde 16	428	22	4,0	4,0
Madde 17	430	20	3,0	4,0
Madde 18	424	26	4,0	5,0

*Tablo 3.7'de görüldüğü gibi, Madde-1 ile Madde-18 aralığındaki tüm maddelerde yer alan %5 oranında kayıp gözlemlere mod ve medyana göre atama yapılmıştır. Mod'a göre yapılan atamada 3,00 ile 5,00 aralığında değerler atanmış, medyana göre yapılan atamada 3,00 ile 4,00 aralığında değerler atanmıştır. Sınıflama ölçeğinde 3,00= Kararsızım, 4,00= Katılıyorum, 5,00= Kesinlikle Katılıyorum şeklindedir.

%10 oranında kayıp gözlem yer alan veri setinde tanıtıcı istatistiksel yöntemlere (mod ve medyan) göre yapılan atama dağılımı Tablo-3.8 'de verilmiştir.

Tablo-3.8 %10 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı

Madde	Cevaplanan Madde (N)	Kayıp gözlem	Medyan	Mod
Madde 1	404	46	3,0	3,0
Madde 2	401	49	3,0	3,0
Madde 3	401	49	3,0	3,0
Madde 4	408	42	3,0	3,0
Madde 5	407	43	4,0	4,0
Madde 6	408	42	4,0	5,0
Madde 7	403	47	4,0	4,0
Madde 8	402	48	4,0	4,0
Madde 9	412	38	4,0	4,0
Madde 10	399	51	4,0	5,0
Madde 11	412	38	4,0	5,0
Madde 12	400	50	3,0	3,0
Madde 13	408	42	4,0	4,0
Madde 14	399	51	4,0	4,0
Madde 15	416	35	4,0	4,0
Madde 16	413	37	4,0	4,0
Madde 17	405	45	3,0	4,0
Madde 18	400	50	4,0	5,0

*Tablo 3.8' de görüldüğü gibi, Madde-1 ile Madde-18 aralığındaki tüm maddelerde yer alan %10 oranında kayıp gözlemlere mod ve medyana göre atama yapılmıştır. Mod'a göre yapılan atamada 3,00 ile 5,00 aralığında değerler atanmış, medyana göre yapılan atamada 3,00 ile 4,00 aralığında değerler atanmıştır. Sınıflama ölçeğinde 3,00= Kararsızım, 4,00= Katılıyorum, 5,00= Kesinlikle Katılıyorum şeklindedir.

%15 oranında kayıp gözlem yer alan veri setinde tanıtıcı istatistiksel yöntemlere (mod ve medyan) göre yapılan atama dağılımı Tablo-3.9 'da verilmiştir.

Tablo-3.9 %15 Kayıp Gözlem Setinin Mod ve Medyan'a Göre Atama Dağılımı

Madde	Cevaplanan Madde (N)	Kayıp gözlem	Medyan	Mod
Madde 1	368	82	3,0	3,0
Madde 2	377	73	3,0	3,0
Madde 3	378	72	3,0	3,0
Madde 4	386	64	3,0	3,0
Madde 5	381	69	4,0	4,0
Madde 6	386	64	4,0	5,0
Madde 7	382	68	4,0	4,0
Madde 8	377	73	4,0	4,0
Madde 9	399	51	4,0	5,0
Madde 10	381	69	4,0	5,0
Madde 11	370	80	4,0	5,0
Madde 12	371	79	3,0	3,0
Madde 13	380	70	4,0	4,0
Madde 14	384	66	4,0	4,0
Madde 15	384	66	4,0	4,0
Madde 16	396	54	4,0	4,0
Madde 17	386	64	3,0	4,0
Madde 18	384	66	4,0	5,0

*Tablo 3.9' da görüldüğü gibi, Madde-1 ile Madde-18 aralığındaki tüm maddelerde yer alan %15 oranında kayıp gözlemlere mod ve medyana göre atama yapılmıştır. Mod'a göre yapılan atamada 3,00 ile 5,00 aralığında değerler atanmış, medyana göre yapılan atamada 3,00 ile 4,00 aralığında değerler atanmıştır. Sınıflama ölçeğinde 3,00 = Kararsızım, 4,00= Katılıyorum, 5,00= Kesinlikle Katılıyorum şeklindedir.

3.4.3. Cramer's V Katsayısı İle Kayıp Gözlem Atama

Cramer's V katsayısı; iki değişkenin iki veya daha fazla boyutlu tablolarında birlikte değişimi ölçmek amacıyla geliştirilmiş katsayıdır. Kayıp gözlemin yer aldığı veri grubu ile ilişki katsayısı en yüksek olan veri grubunun Cramer's V katsayısı belirlenmektedir. Belirlenen bu veri grubunun eşleştirilmesi yapılarak kayıp gözlem ataması gerçekleştirilmektedir. Dolayısıyla veri gruplarında kayıp gözlem oranının artması, ilişkinin yüksek belirlendiği veri gruplarında bu eşleştirmeyi zayıflatmaktadır. Kayıp gözlem bulunan madde ile Cramer's V katsayısı en yüksek bulunan veri grubundaki maddede veri bulunması olasılığını arttırmak amacıyla kayıp gözlem oranı düşük tutulmaktadır.

%5 oranında kayıp gözlem yer alan veri setinde Cramer's V katsayısına göre yapılan atama dağılımı Talo-3.10 'da verilmiştir.

Tablo-3.10 %5 Kayıp Gözlem Setinin Cramer's V katsayısına Göre Atama Dağılımı

Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.
2	148	0,811	0,003	138	45	0,723	0,02	295	68	0,898	0
3	1	0,81	0	140	258	0,81	0,003	296	52	0,857	0
4	113	1	0,003	141	406	0,8	0,009	299	131	1	0
5	235	0,863	0,001	143	113	1	0	300	307	1	0
6	305	0,865	0	146	251	1	0,001	301	350	0,819	0
9	132	0,884	0	147	265	0,851	0,001	303	402	0,82	0
10	22	0,811	0,003	149	155	0,911	0	305	6	0,865	0
11	450	0,782	0,005	150	387	0,874	0,001	306	15	0,856	0,001
12	113	1	0	151	235	0,405	0,757	308	444	0,9	0
15	74	0,865	0,003	156	113	1	0,003	309	186	1	0,001
16	316	0,873	0,001	158	186	1	0,001	312	65	0,769	0,001
17	4	0,81	0,002	159	147	0,751	0,002	313	71	0,82	0
19	305	0,884	0	160	336	0,874	0,001	314	113	1	0,001
21	8	1	0	161	131	0,894	0	315	232	0,787	0,004
22	396	0,87	0,002	162	191	1	0,003	317	210	0,824	0
23	288	0,937	0	164	450	0,78	0,001	318	1,00	0,919	0
25	398	0,877	0,003	165	151	0,531	0,437	319	259	1	0
26	200	0,818	0,002	168	73	0,893	0,001	320	259	0,886	0,001
28	113	1	0,001	171	29	0,721	0,005	321	397	0,877	0,003
30	152	0,921	0	172	179	0,766	0,032	323	237	0,818	0,002
31	91	0,687	0,009	173	186	0,791	0,01	327	142	0,771	0,006
32	4	0,846	0,001	174	356	0,843	0,001	329	357	0,8	0,003
33	403	0,759	0,002	175	288	0,837	0,001	331	217	0,902	0
34	302	0,901	0	176	87	0,803	0,009	332	181	0,856	0,004
35	85	0,873	0	178	256	0,793	0	335	156	0,774	0,006
36	98	0,768	0,002	182	398	1	0	337	148	1	0
38	74	0,72	0,022	184	168	0,833	0,08	338	97	0,839	0
39	113	1	0,001	188	132	0,854	0	340	131	0,882	0
40	89	0,881	0	190	307	1	0	342	357	0,861	0,003
43	187	0,817	0	193	415	0,701	0,007	343	61	0,791	0,013
45	155	0,894	0,001	194	414	0,816	0,008	344	190	0,83	0
46	260	0,795	0,001	195	113	1	0	351	416	0,779	0,005
48	153	0,818	0	198	259	0,871	0,008	352	442	0,892	0
50	152	0,788	0,001	199	185	0,837	0,006	354	186	1	0
51	436	1	0	204	221	0,886	0	355	397	1	0,001

Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.
53	113	1	0,001	205	283	0,804	0,003	356	186	1	0,001
55	398	1	0	206	39	0,754	0,009	357	113	1	0,003
56	186	1	0	207	443	0,803	0,001	358	345	0,932	0,001
58	406	1	0,001	208	157	0,803	0,009	359	19	0,789	0,001
59	66	0,787	0,004	209	283	1	0	361	7	0,847	0
60	258	0,736	0,003	211	404	0,793	0,001	363	37	0,804	0,001
61	436	1	0	214	223	0,862	0	365	277	0,767	0,001
63	90	0,877	0,003	215	210	0,808	0,001	366	261	0,818	0,002
66	170	0,93	0	219	21	0,926	0	367	51	0,849	0,004
67	163	0,926	0	220	265	0,846	0	368	47	0,764	0,007
69	113	1	0	222	302	0,859	0,001	369	258	0,854	0,001
70	406	0,825	0,016	226	288	0,667	0,1	371	224	0,835	0
71	113	1	0	227	358	0,786	0,014	372	420	0,52	0,49
72	163	0,845	0,001	228	318	0,803	0,003	373	186	1	0,001
74	233	0,904	0,001	231	52	0,857	0,003	374	258	0,916	0
77	132	0,853	0,001	232	253	0,811	0,001	375	87	1	0,001
78	374	0,679	0,035	233	74	0,904	0,001	377	113	1	0,003
80	87	0,856	0,01	235	5	0,863	0,001	379	177	0,835	0,002
81	421	0,911	0	238	125	0,884	0	380	289	0,839	0,001
82	181	0,837	0,013	239	251	1	0,001	381	157	0,813	0,018
83	113	1	0,001	240	185	0,856	0,01	382	90	0,87	0,009
85	113	1	0	241	376	0,772	0,006	388	87	0,856	0,01
88	167	0,9	0	242	292	0,88	0	390	382	0,887	0
93	383	0,883	0	243	87	0,811	0,008	394	414	0,825	0,006
94	113	1	0,001	245	111	0,806	0,001	396	85	1	0
95	323	0,748	0,003	248	4	0,782	0,005	399	260	0,856	0,001
96	168	0,812	0,008	249	156	0,91	0	401	259	0,856	0,004
98	383	0,782	0,005	250	191	1	0	403	90	0,882	0,007
100	261	0,901	0	253	131	0,945	0	404	397	1	0,001
101	185	0,825	0,007	256	407	0,391	0,481	405	52	1	0
102	265	0,854	0	257	157	0,907	0,002	409	8	0,856	0,01
104	376	0,764	0,007	260	242	0,869	0,002	412	330	1	0,001
105	251	1	0,001	261	153	0,909	0	414	52	0,857	0,003
106	90	0,882	0,003	262	318	0,815	0,002	417	420	0,825	0,002
108	181	1	0,001	263	35	0,728	0,018	420	417	0,825	0,002
109	251	0,791	0,024	264	373	0,71	0,02	422	189	0,95	0

Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.	Madde	Kişi	Cramer's V	Sig.
111	330	1	0,003	265	102	0,854	0	424	259	0,84	0,002
114	139	0,833	0,001	269	200	0,828	0	425	52	0,857	0,001
115	201	1	0	270	445	0,787	0,011	426	15	0,774	0,001
116	137	0,765	0,032	271	237	0,874	0	427	406	0,837	0,006
117	113	1	0	272	350	0,854	0,001	428	125	0,825	0,002
119	376	1	0	273	8	0,894	0,002	429	397	0,755	0,016
121	119	0,82	0,002	274	376	0,867	0,001	430	357	0,753	0,01
122	217	0,811	0,001	275	302	0,867	0,003	433	330	1	0,001
123	398	0,903	0,002	278	398	0,877	0,008	438	181	1	0,003
127	52	0,738	0,12	280	398	0,877	0,003	439	200	0,857	0
128	445	0,791	0,01	282	113	1	0	441	113	1	0
129	8	1	0	283	209	1	0	446	280	0,772	0
130	186	0,791	0,01	286	113	1	0	448	113	1	0,001
131	299	1	0	288	23	0,937	0	449	406	0,825	0,007
134	316	0,82	0,002	293	251	1	0,001				
135	397	0,87	0,003	294	259	0,871	0,003				

* Cramer's V katsayısı iki değişkenin iki veya daha fazla boyutlu tablolarında birlikte değişimi ölçmek amacıyla geliştirilmiş katsayıdır. Cramer's V katsayısı 0 ile 1 arasında değişim gösterir. 0 değeri iki değişken arasında birliktelik olmadığını, 1 ise tam birliktelik olduğunu belirtir (Özdamar, 2004;206).

3.4.4. Benzerlik Temelli Kayıp Gözlem Atama Yöntemi

Çalışmada yer alan Benzerlik Temelli Atama(B.T.A) yönteminin çalışma prensibi maddelerin vermiş olduğu cevapların benzerliğine dayanır. Maddelerin vermiş olduğu cevapları karşılaştırarak benzerliği en yüksek maddelerin verileri belirlenir ve kayıp gözlem bulunan kısma elle atama yapılır. Geliştirilen yöntemin işlem basamakları şu şekildedir;

- (1) Tam veri seti üzerinde %5, %10 ve %15 oranında kayıp gözlem oluşturulur.
- (2) Excel ve java programları aracılığıyla işlem algoritmasına kayıp gözlemler yüklenir.
- (3) Kayıp gözlem bulunan her bir kişiye ait veriler ile diğer tüm kişiler karşılaştırılarak aralarındaki benzerlik belirlenir. Belirlenen benzerliklerden toplam kişi sayısının %5 i kadar en yüksek yakınlığa sahip kişilerin verileri algoritma tarafından sıralanır.

(4) Algoritma tarafından sıralanan kişilerin verilerine göre 2 ayrı yaklaşım ile atama gerçekleştirilir. Atama manuel olarak yapılır. B.T.A. (en yakın) yaklaşımında; benzer olan kişinin verileri kayıp gözleme elle atanır. B.T.A. (%5) yaklaşımında; %5 yakınlıktaki benzer kişilerin verilerinin modu kayıp gözleme elle atanır.

3.4.4.1. Kayıp Gözlem Oluşturma

Örnekleme yer alan veri seti üzerinde Excel programı aracılığıyla seçkisiz olarak rastgele sayılar tablosu ile kayıp gözlemler oluşturulmuştur. Yapılan işleme ilişkin örnek şekil-3.1 'de verilmiştir.

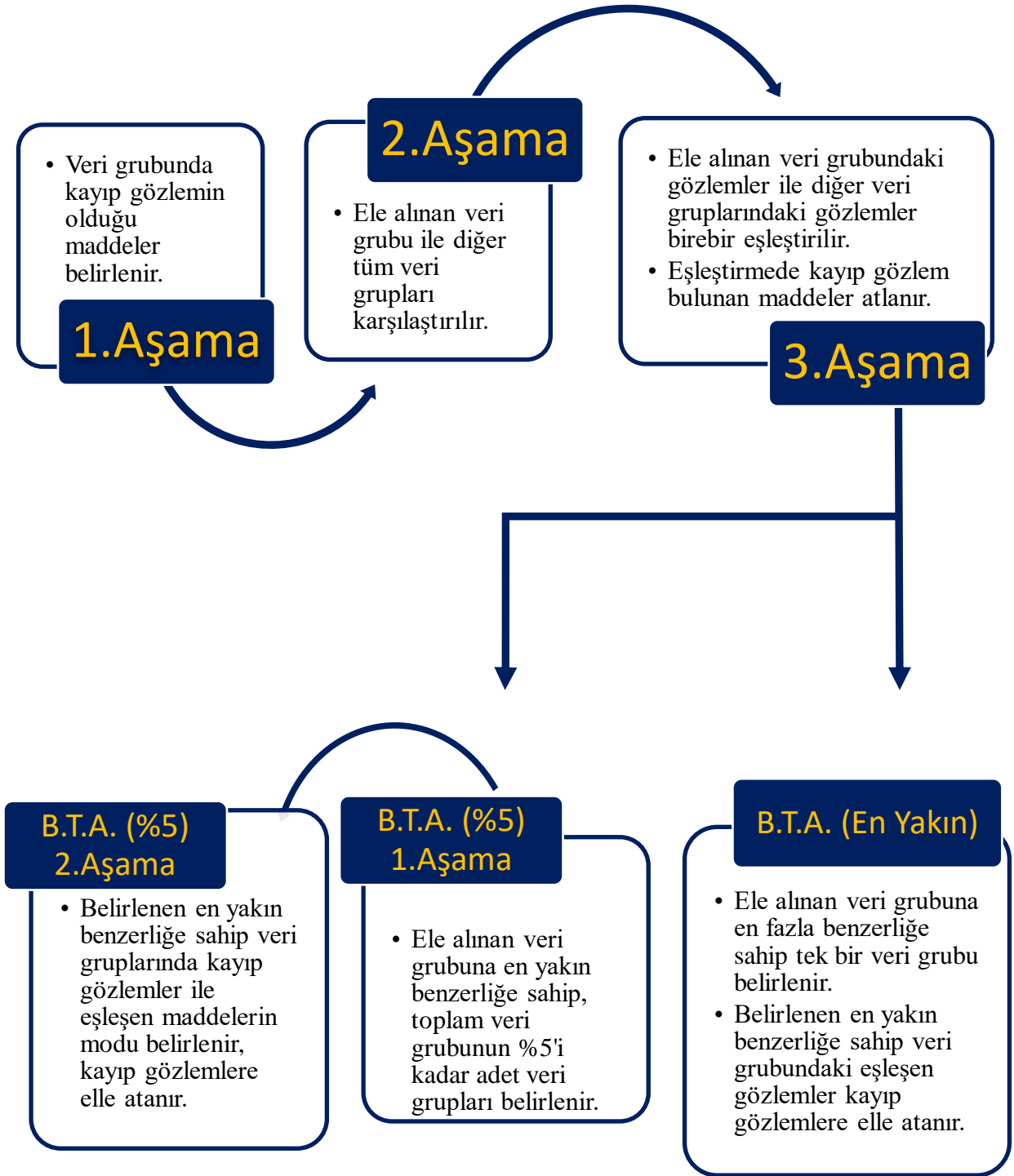
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	kişi	SORU 1	SORU 2	SORU 3	SORU 4	SORU 5	SORU 6	SORU 7	SORU 8	SORU 9	SORU 10	SORU 11	SORU 12	SORU 13	SORU 14	SORU 15	SORU 16	SORU 17	SORU 18
2	1	5	5	1	5	4	4	5	4	4	4	5	5	4	4	5	3	5	4
3	2	3	3	4	4	4	4	3	3	3	3	4	4	3	4	4	3	3	4
4	3	3	3	2	4	2	5	4	5	4	4	4	4	5	5	4	4	3	4
5	4	3	3	2	3	4	3	5	4	5	5	4	3	5	4	1	5	4	4
6	5	4	4	4	3	3	4	3	3	4	4	4	4	4	3	4	3	4	3
7	6	4	4	4	4	5	5	3	4	4	4	4	5	4	4	5	4	5	5
8	7	3	2	3	3	4	5	3	4	5	5	5	3	4	4	4	4	4	4
9	8	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
10	9	3	3	3	3	4	4	4	2	4	4	4	4	4	4	3	3	4	4
11	10	2	2	2	2	2	2	2	2	2	2	4	1	2	2	3	2	2	2
12	11	3	3	3	4	3	5	3	4	5	4	2	4	4	3	3	3	3	4
13	12	3	4	3	3	4	4	4	4	4	4	4	4	4	3	3	4	4	4
14	13	3	2	3	3	4	3	4	4	3	4	4	2	3	3	2	4	3	3
15	14	5	3	5	3	3	4	3	2	4	4	3	3	2	2	2	5	1	3
16	15	4	4	4	5	4	4	4	4	4	4	4	4	4	5	4	4	4	4
17	16	3	3	4	3	2	5	4	2	5	4	5	3	5	5	4	4	4	5
18	17	4	4	4	2	3	4	4	2	3	4	2	4	3	2	4	4	4	2
19	18	3	3	3	3	4	5	4	5	5	5	5	5	5	4	4	4	4	5
20	19	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	3	3	3
21	20	5	4	3	2	3	4	4	2	4	3	2	5	4	3	2	3	4	2
22	21	4	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
23	22	3	3	3	3	3	4	3	4	3	4	3	3	3	3	3	3	3	3
24	23	4	4	3	4	4	4	4	4	3	4	4	4	4	4	4	4	3	5
25	24	4	3	5	5	4	4	4	5	5	5	4	4	4	4	4	4	3	5
26	25	2	2	2	2	2	4	4	3	4	4	4	3	2	3	3	3	3	4
27	26	3	2	4	3	4	4	3	5	5	5	5	4	5	5	4	4	3	4

Şekil-3.1. Kayıp Gözlem Oluşturma Örneği

Şekil-3.1'de görüldüğü gibi, kayıp gözlem elde etmek amacıyla tam veri seti üzerinde Excel programında rastgele sayılar tablosu kullanılmıştır. Seçkisiz olarak veriler belirlenmiş ve bu veriler silinerek kayıp gözlemler elde edilmiştir. Yapılan işleme örnek olarak; kişi-10'nun soru-1 ve soru-11 cevapları silinerek kayıp gözlem oluşturulmuştur.

3.4.4.2. Kayıp Gözlemlerin Algoritmaya Yüklenmesi

Şekil-3.1'de görüldüğü gibi, %5 kayıp gözlem bulunan veri setinde kişilerin sorulara vermiş olduğu cevaplar tablo üzerinde Excel programı ile doldurulmuştur. Benzerlik Temelli Atama yönteminde Java üzerinde yazılan yazılım algoritması eksik gözlem bulunan maddeyi tespit ederek maddenin diğer sorulara verdiği cevapları kaydeder. Benzerlik temelli atama algoritmasının işleyiş şeması Şekil-3.2.'de verilmiştir.



Şekil-3.2. B.T.A. Yöntemi Algoritmasının İşleyiş Şeması

*Şekil 3.2.'de görüldüğü gibi, benzerlik temelli atama yönteminin başlangıç aşamalarında; 1. Aşamada kayıp gözlemler belirlenir, 2. Aşamada veri grupları birbiri ile karşılaştırılır, 3. Aşamada kayıp gözlem bulunan veri grubu, diğer veri grupları ile birebir eşleştirilir.

Başlangıçtaki ilk üç aşama B.T.A. (En Yakın) ve B.T.A.(%5) yaklaşımları için ortak gerçekleşir. B.T.A(En Yakın) yaklaşımında kayıp gözlem bulunan veri grubu ile en yakın benzerliğe sahip veri grubu belirlenip, kayıp gözlemler yerine eşleşen maddedeki veriler elle atanır. B.T.A(%5) yaklaşımında kayıp gözlem bulunan veri grubu ile en yakın benzerliğe sahip veri setinin %5'i kadar veri grubu belirlenip, kayıp gözlemler yerine eşleşen maddelerin modu elle atanır.

3.4.4.3. En Yüksek Benzerliklerin Belirlenmesi

Verilen cevaplar ile diğer maddelerin verdiği cevapları karşılaştırarak aynı cevabın verilmesi durumunda benzerliğin yakınlığını belirlenir. Benzerliği en yüksek olan maddeden başlayarak benzerlik sıralaması yapılır. Program toplam veri setinin %5 yakınlık oranı olan kısmını kaydeder. Yapılan işleme ilişkin örnek şekil-3.3'de verilmiştir.

	BENZERLİK	BİREY		
2			BİREY	2
4	EKSİK VERİ BULUNAN MADDELER		3	4
6				
7	12	179	3	4
8	11	287	4	4
9	10	127	3	5
10	10	150	4	4
11	10	411	4	4
12	9	8	3	4
13	9	11	3	
14	9	50		3
15	9	98	3	3
16	9	131	4	
17	9	137	4	3
18	9	159	4	
19	9	167	2	4
20	9	236	4	3
21	9	247	3	3
22	9	259	3	3
23	9	266	2	3
24	9	269	3	3
25	9	330	3	4
26	9	409	2	4
27	9	444	4	2
28	8	9	3	3
31	MOD		3	3

Şekil-3.3. Benzerlik sıralaması

3.4.4.4. Benzerlik Temelli Atama İle Kayıp Gözlem Atama

Benzerlik Temelli Atama yönteminde kayıp gözlemlerin tamamlanmasında iki yaklaşım ele alınmıştır.

- *Benzerlik Temelli Atama (En Yakın) Yönteminde;* eksik gözlem bulunan madde ile cevapları en benzer olan maddeler tespit edilir. En yakın maddenin eksik gözlem olan veriye verdiği cevaplar kaydedilir ve eksik gözlem yerine elle atama yapılır. Benzerlik Temelli Atama yöntemi algoritması öncelikle kayıp gözlem bulunan maddeyi tespit eder. İşlem örneği şekil-3.1’de verilmiştir. Örnek olarak;

Şekil-3.1’de görüldüğü gibi, Madde-2’ye ait eksik gözlem bulunan 3 ve 4. sorulara diğer maddelerin verdiği cevaplar Excel tablosunda kaydedilir.

Şekil-3.3’de görüldüğü gibi, Madde-2’ye en yakın olan Madde-179’un verdiği cevaplar eksik gözlem bulunan sorulara elle kaydedilir.

- *Benzerlik Temelli Atama (%5) Yönteminde;* eksik gözlem bulunan madde ile cevapları en benzer olan maddeler tespit edilir. Maddeler en yakından – uzağa doğru sıralanır. Veri setindeki toplam madde sayısının %5’i olacak şekilde en yakın maddeler gruplanır. Grubun eksik gözlem olan veriye verdiği cevapların modu hesaplanır ve eksik gözlem yerine elle atama yapılır. Örnek olarak;

Şekil-3.1’de gösterilen Madde-2’ye ait eksik gözlem bulunan 3 ve 4. sorulara diğer maddelerin verdiği cevaplar Excel tablosunda kaydedilir.

Şekil-3.3’de görüldüğü gibi, Madde-2’ye en yakın cevabı veren veri setindeki toplam madde sayısının %5’ i olacak şekilde (22 madde) maddeler sıralanır. Maddelerin verdiği cevapların modu hesaplanarak eksik gözlem bulunan sorulara elle kaydedilir.

BÖLÜM IV

BULGULAR

Bu bölümde araştırmanın problemlerine yönelik elde edilen verilerin analiz sonuçları tablo, grafik ve şekillerle verilmiştir.

Araştırmalarda veri toplama sürecindeki en büyük sorunlar eksik/kayıp gözlem veya uç değer verilerinin olduğu durumlardır. Evren hakkında bir önerme yapmak adına evrenden çekilmiş olan örneklemin eksiksiz olması araştırmanın sonucunu belirlemek için son derece önemlidir. Araştırmada kayıp gözlemlerin tamamlanması adına kayıp gözlem oranlarına göre çeşitli yöntemler kullanılmıştır. Çalışma için belirlenen veri setleri %5, %10 ve %15 kayıp gözlemler içermekte olup, %5 kayıp gözlem bulunan set için verilerin atanmasında Cramer's V katsayısı, mod, medyan ve benzerlik temelli atama yöntemleri uygulanmıştır. %10 ve %15 kayıp gözlem içeren veri seti için verilerin atanmasında mod, medyan ve benzerlik temelli atama yöntemi uygulanmıştır. Kayıp gözlemlerin atanmasında kullanılan yöntemler doğrultusunda bulgular kayıp gözlemler miktarına göre sınıflandırılmıştır.

Kayıp gözlemlerin atamasında kullanılan benzerlik temelli atama yöntemi çalışma için geliştirilmiş olup çalışma birimlerinin maddelere verdiği cevaplara göre birbirine en benzer birimleri tespit ederek kayıp gözlemlerin atanması temeline dayanmaktadır.

4.1. Kayıp Gözlem Atama Yöntemlerinde Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada incelenen kayıp gözlem atama yöntemlerinin doğru atama oranları incelenmiştir. Araştırmada incelenen tanıtıcı istatistiksel yöntemler (mod, medyan), Cramer's V katsayısı ve B.T.A. yöntemi ile tamamlanan verilere ait doğru atama oranı bulguları aşağıda verilmiştir.

4.1.1. %5 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %5 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen doğru atama oranlarının karşılaştırılması Tablo- 4.1 'de verilmiştir.

Tablo 4.1. % 5 Kayıp Gözlem Doğru Atama Oranı Bulguları

Atama Yöntemi	n(k)	Doğru Atama Sayısı	Doğru Atama Yüzdesi
Cramer's V katsayısı ile atama yöntemi		118	%28.85
Moda göre atama yöntemi		144	%35.20
Medyana göre atama yöntemi	409	142	%34.72
Benzerlik temelli atama yöntemi (%5)		232	%56.72
Benzerlik temelli atama yöntemi		208	%50.85

*Kayıp gözlem atama yöntemlerinin %5 kayıp gözlem veri setindeki atamaları incelendiğinde, Cramer's V katsayısı ile atama yöntemi 409 kayıp gözlemden 118'ini (%28.85) doğru atamıştır. Moda göre atama yöntemi 409 kayıp gözlemden 144'ünü(%35.20) doğru atamıştır. Medyana göre atama yöntemi 409 kayıp gözlemden 142'sini(%34.72) doğru atamıştır. Benzerlik temelli atama yöntemi(%5) 409 kayıp gözlemden 232'sini(56.72) doğru atamıştır. Benzerlik temelli atama yöntemi 409 kayıp gözlemden 208'ini (50.85) doğru atamıştır. Ulaşılan sonuçlara göre en çok doğru atama yapan yöntem benzerlik temelli atama yöntemi(%5) olarak bulunmuştur.

4.1.2. %10 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %10 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen doğru atama oranlarının karşılaştırılması Tablo- 4.2. 'de verilmiştir.

Tablo 4.2. % 10 Kayıp Gözlem Doğru Atama Oranı Bulguları

Atama Yöntemi	n(k)	Doğru Atama Sayısı	Doğru Atama Yüzdesi
Moda göre atama yöntemi		286	%35.60
Medyana göre atama yöntemi		294	%36.60
Benzerlik temelli atama yöntemi (%5)	803	462	%57.53
Benzerlik temelli atama yöntemi		428	%53.30

*Kayıp gözlem atama yöntemlerinin %10 kayıp gözlem veri setindeki atamaları incelendiğinde, moda göre atama yöntemi 803 kayıp gözlemden 286'sını(%35.60) doğru atamıştır. Medyana göre atama yöntemi 803 kayıp gözlemden 294'ünü(%36.60) doğru atamıştır. Benzerlik temelli

atama yöntemi(%5) 803 kayıp gözlemden 462'sini(57.53) doğru atamıştır. Benzerlik temelli atama yöntemi 803 kayıp gözlemden 428'ini (53.30) doğru atamıştır. Ulaşılan sonuçlara göre en çok doğru atama yapan yöntem benzerlik temelli atama yöntemi(%5) olarak bulunmuştur.

4.1.3. %15 Kayıp Gözlem Durumunda Atama Yöntemlerinin Doğru Atama Oranlarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %15 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen doğru atama oranlarının karşılaştırılması Tablo- 4.3. 'da verilmiştir.

Tablo 4.3. % 15 Kayıp Gözlem Doğru Atama Oranı Bulguları

Atama Yöntemi	n(k)	Doğru Atama Sayısı	Doğru Atama Yüzdesi
Moda göre atama yöntemi		460	%37.39
Medyana göre atama yöntemi		409	%33.25
Benzerlik temelli atama yöntemi (%5)	1230	667	%54.27
Benzerlik temelli atama yöntemi		615	%50.00

*Kayıp gözlem atama yöntemlerinin %15 kayıp gözlem veri setindeki atamaları incelendiğinde, moda göre atama yöntemi 1230 kayıp gözlemden 460'mı(%37.39) doğru atamıştır. Medyana göre atama yöntemi 1230 kayıp gözlemden 409'unu(%33.25) doğru atamıştır. Benzerlik temelli atama yöntemi(%5) 1230 kayıp gözlemden 667'sini(54.27) doğru atamıştır. Benzerlik temelli atama yöntemi 1230 kayıp gözlemden 615'ini (%50.00) doğru atamıştır. Ulaşılan sonuçlara göre en çok doğru atama yapan yöntem benzerlik temelli atama yöntemi(%5) olarak bulunmuştur.

4.2. Kayıp Gözlem Atama Yöntemlerinde Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada incelenen kayıp gözlem atama yöntemlerinin veri setinin güvenirlilik katsayısı üzerindeki etkileri incelenmiştir. Araştırmada incelenen tanıtıcı istatistiksel yöntemler (mod, medyan), Cramer's V katsayısı yöntemi ve B.T.A. yöntemi ile tamamlanan veri setlerinin güvenirlilik katsayılarına ait bulgular aşağıda verilmiştir.

4.2.1. %5 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %5 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen Cronbach Alpha güvenirlik katsayılarının karşılaştırılması Tablo 4.4. 'de verilmiştir.

Tablo 4.4. % 5 Kayıp Gözlem Güvenirlik Analizi Bulguları

Atama Yöntemi	n(k)	Cronbach Alpha (α)
Tam Veri		,941
Cramer's V katsayısı ile atama yöntemi		,934
Moda göre atama yöntemi	450	,935
Medyana göre atama yöntemi		,937
Benzerlik temelli atama yöntemi (%5)		,944
Benzerlik temelli atama en yakın		,942

*Tablo 4.4.'de elde edilen bulgulara göre, tam verinin güvenirlik katsayısı $\alpha=,941$ ($\alpha> ,40$), Cramer's V katsayısı yöntemi ile güvenirlik katsayısı $\alpha=,934$ ($\alpha> ,40$), moda göre atama yöntemi ile güvenirlik katsayısı $\alpha=,935$ ($\alpha> ,40$), medyana göre atama yöntemi ile güvenirlik katsayısı $\alpha=,937$ ($\alpha> ,40$), B.T.A. (%5) yöntemi ile güvenirlik katsayısı $\alpha=,944$ ($\alpha> ,40$), B.T.A. yöntemi ile güvenirlik katsayısı $\alpha=,942$ ($\alpha> ,40$) olduğu görülmüştür(Özdamar, 1999). Buna göre B.T.A. yöntemi ile yapılan kayıp gözlem atama yönteminin denenen diğer atama yöntemlerine kıyasla daha yüksek güvenirlik katsayısına sahip olduğu belirlenmiştir.

4.2.2. %10 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlik Katsayılarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %10 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen Cronbach Alpha güvenirlik katsayılarının karşılaştırılması Tablo 4.5. 'de verilmiştir.

Tablo 4.5. % 10 Kayıp Gözlem Güvenirlik Analizi Bulguları

Atama Yöntemi	n(k)	Cronbach Alpha(α)
Tam Veri		,941
Moda göre atama yöntemi		,927
Medyana göre atama yöntemi	450	,929
Benzerlik temelli atama yöntemi (%5)		,946
Benzerlik temelli atama en yakın		,942

*Tablo 4.5.' de elde edilen bulgulara göre, tam verinin güvenirlilik katsayısı $\alpha=,941$ ($\alpha> ,40$), moda göre atama yöntemi ile güvenirlilik katsayısı $\alpha=,927$ ($\alpha> ,40$), medyana göre atama yöntemi ile güvenirlilik katsayısı $\alpha=,929$ ($\alpha> ,40$), B.T.A. (%5) yöntemi ile güvenirlilik katsayısı $\alpha=,946$ ($\alpha> ,40$), B.T.A. yöntemi ile güvenirlilik katsayısı $\alpha=,942$ ($\alpha> ,40$) olduğu görülmüştür(Özdamar, 1999). Buna göre B.T.A. yöntemi ile yapılan kayıp gözlem atama yönteminin denenen diğer atama yöntemlerine kıyasla daha yüksek güvenirlilik katsayına sahip olduğu belirlenmiştir.

4.2.3. %15 Kayıp Gözlem Durumunda Atama Yöntemlerinin Güvenirlilik Katsayılarının Karşılaştırılmasına İlişkin Bulgular

Çalışmada yer alan veri setinde oluşturulan %15 oranında kayıp gözlemlere ilişkin farklı yöntemler ile veri atama yapılmıştır. Elde edilen Cronbach Alpha güvenirlilik katsayılarının karşılaştırılması Tablo 4.6. 'de verilmiştir.

Tablo 4.6. % 15 Kayıp Gözlem Güvenirlilik Analizi Bulguları

Atama Yöntemi	n(k)	Cronbach Alpha(α)
Tam Veri		,941
Moda göre atama yöntemi		,917
Medyana göre atama yöntemi	450	,921
Benzerlik temelli atama yöntemi (%5)		,948
Benzerlik temelli atama en yakın		,942

*Tablo 4.6.' de elde edilen bulgulara göre, tam verinin güvenirlilik katsayısı $\alpha=,941$ ($\alpha> ,40$), moda göre atama yöntemi ile güvenirlilik katsayısı $\alpha=,917$ ($\alpha> ,40$), medyana göre atama yöntemi ile güvenirlilik katsayısı $\alpha=,921$ ($\alpha> ,40$), B.T.A. (%5) yöntemi ile güvenirlilik katsayısı $\alpha=,948$ ($\alpha> ,40$), B.T.A. yöntemi ile güvenirlilik katsayısı $\alpha=,942$ ($\alpha> ,40$) olduğu görülmüştür(Özdamar, 1999). Buna göre B.T.A. yöntemi ile yapılan kayıp gözlem atama yönteminin denenen diğer atama yöntemlerine kıyasla daha yüksek güvenirlilik katsayına sahip olduğu belirlenmiştir.

BÖLÜM V

SONUÇ TARTIŞMA VE ÖNERİLER

Bu bölümde çalışmada elde edilen bulgulara göre ulaşılan sonuçlara yer verilmiştir. Araştırma sorularına yönelik elde edilen sonuçlar ile daha önce yapılan çalışmalarda sonuçlar karşılaştırılmıştır. Buna bağlı olarak öneriler oluşturulmuş ve sunulmuştur.

5.1. Sonuç ve Tartışma:

Bu çalışmada sınıflama ölçeğinde kayıp gözlemlerin tamamlanmasında benzerlik temelli atama yöntemi geliştirilmiştir. Geliştirilen benzerlik temelli atama yöntemi, TROK mekanizmasına uygun farklı büyüklüklerde ki (%5 ,%10 ve %15) kayıp gözlem oranlarına sahip veri setini, tanıtıcı istatistikler(mod, medyan) ile kayıp gözlemi tamamlama ve Cramer's V katsayısı ile tespit edilen en yakın kişinin vermiş olduğu cevaplar ile tamamlama yapılarak karşılaştırılmıştır. Ayrıca tamamlama yöntemlerinin ölçeğin güvenilirliğine etkisi karşılaştırılıp etkinliği incelenmiştir.

Yapılan araştırma bulgularına göre %5 kayıp gözlem oranına sahip veri setinin belirtilen yöntemler ile tamamlanması sonucunda en düşük doğru atama oranı(%28.85) Cramer's V katsayısı ile yapılan tamamlama yönteminde görülmüştür. Geliştirilen Benzerlik Temelli Atama yönteminde iki yöntem denenmiş olup benzerlik temelli atama(%5)'nın daha iyi sonuç verdiği (%56.72) görülmüştür.

%10 kayıp gözlem oranına sahip veri setinin belirtilen yöntemler ile tamamlanması sonucunda en düşük doğru atama oranı(%35,60) mod ile yapılan tamamlama yönteminde görülmüştür. Geliştirilen benzerlik temelli atama yönteminden Benzerlik Temelli Atama(%5)'nın doğru atama oranı(%57.53) en yüksek olduğu, diğer Benzerlik Temelli Atama En Yakın'ın yarıdan fazla doğru atama (%53.30) yaptığı görülmüştür.

%15 kayıp gözlem oranına sahip veri setinin belirtilen yöntemler ile tamamlanması sonucunda en düşük doğru atama oranı(%33,25) medyan ile yapılan tamamlama yönteminde görülmüştür. Geliştirilen benzerlik temelli atama yönteminden Benzerlik Temelli Atama(%5)'nın doğru atama oranı(%54.27) en yüksek olduğu, diğer Benzerlik Temelli Atama En Yakın'ın yarıdan fazla doğru atama (%50.00) yaptığı görülmüştür.

Farklı büyüklüklerdeki kayıp gözlem oranına sahip veri setlerinin tamamlanması yöntemleri incelendiğinde Benzerlik Temelli Atama(%5) yönteminin diğer yöntemlerden daha iyi doğru atama yaptığı tespit edilmiştir. Benzerlik Temelli Atama(%5) yöntemi kayıp gözlem bulunan verinin verdiği cevaplara en benzer cevap veren %5 yakınlıktaki verilerin (bizim çalışmamızın örneklemini 450 kişi olduğundan %5'i yani en benzer cevap veren 22 kişinin verileri) eksik gözlem için verilen cevapların modu şeklinde belirlenmiştir.

%5 kayıp gözlem oranına sahip veri setinde yapılan tamamlama işlemleri sonucunda Cronbach's Alpha güvenirlik katsayıları karşılaştırıldığında tam veri setinden elde edilen katsayı($\alpha = .941$) ile genel olarak birbirlerinden farklılaştığı ancak bu farkın çok küçük olduğu gözlenmiştir. Cronbach's Alpha güvenirlik katsayısı en düşük moda göre tamamlama yönteminde ($\alpha = .937$) görülmüştür. Benzerlik Temelli Atama(%5) yöntemi ile yapılan tamamlama sonrasında Cronbach's Alpha güvenirlik katsayısı ($\alpha = .944$) en yüksek değerde olduğu hesaplanmıştır.

%10 kayıp gözlem oranına sahip veri setinde yapılan tamamlama işlemleri sonucunda Cronbach's Alpha güvenirlik katsayıları karşılaştırıldığında tam veri setinden elde edilen katsayı($\alpha = .941$) ile genel olarak birbirlerinden farklılaştığı ancak bu farkın çok küçük olduğu gözlenmiştir. Cronbach's Alpha güvenirlik katsayısı en düşük moda göre tamamlama yönteminde ($\alpha = .927$) görülmüştür. Benzerlik Temelli Atama(%5) yöntemi ile yapılan tamamlama sonrasında Cronbach's Alpha güvenirlik katsayısı ($\alpha = .946$) en yüksek değerde olduğu hesaplanmıştır.

%15 kayıp gözlem oranına sahip veri setinde yapılan tamamlama işlemleri sonucunda Cronbach's Alpha güvenirlik katsayıları karşılaştırıldığında tam veri setinden elde edilen katsayı($\alpha = .941$) ile genel olarak birbirlerinden farklılaştığı ancak bu farkın çok küçük olduğu gözlenmiştir. Cronbach's Alpha güvenirlik katsayısı en düşük moda göre tamamlama yönteminde ($\alpha = .917$) görülmüştür. Benzerlik Temelli Atama(%5) yöntemi ile yapılan tamamlama sonrasında Cronbach's Alpha güvenirlik katsayısı ($\alpha = .948$) en yüksek değerde olduğu hesaplanmıştır.

Benzerlik Temelli Atama yöntemleri ile farklı büyüklüklerde ki (%5 ,%10 ve %15) kayıp gözlem oranlarına sahip veri setleri tamamlandığında, Tam veri seti güvenirlik katsayısından daha yüksek değerlerin görülmesinin sebebi tamamlama yönteminin iç tutarlılığı arttırmasından kaynaklıdır. Kayıp gözlem oranı arttıkça Benzerlik Temelli Atama ile

tamamlama yöntemi iç tutarlılığı daha da arttırmakta bu durumun sonucunda Cronbach's Alpha güvenirlik katsayısı da artmaktadır.

5.2. Öneriler

Araştırmada yer alan çalışmalardan elde edilen bulgular ve sonuçlar ile aşağıda belirtilen öneriler sunulmaktadır;

1- Yapılan çalışmalarda örneklem büyüklüğü, kayıp gözlem oranı, test uzunluğu gibi değişkenler göz önüne alındığında kayıp gözlem tamamlamada tüm şartları sağlayan tek bir yöntemin olmadığı, bundan dolayı çalışma örnekleminin özelliklerine en uygun kayıp gözlem atama yönteminin araştırmacılar tarafından seçilmesi önerilmektedir.

2- Yapılan analizlerin güvenirliğinin yüksek değerlerden sağlanması için, araştırmanın sonuçlarına dayalı olarak tanıtıcı istatistiksel yöntemler yerine B.T.A. yönteminin de kullanılması önerilmektedir. Benzerlik Temelli Atama yöntemlerinin doğru atama yüzdesinin araştırmada denenen tüm şartlarda (%5,%10 ve %15 kayıp gözlem durumu) %50 oranının üzerinde doğru atama yapması dikkate alınması gereken bir durumdur. Özellikle B.T.A(%5)'in diğer atama yöntemleri ile karşılaştırılması ve doğru atama yüzdelerinin hesaplandığı çalışmaların yapılması önerilmektedir.

3- B.T.A. yönteminin güvenirlik üzerindeki etkisinin belirlenmesi amacıyla 3 yaygın yöntem ile karşılaştırmalı incelenmesi yapılmıştır. Geliştirilen yeni yöntemin etkinliğinin daha kapsamlı incelenmesi için farklı kayıp gözlem tamamlama yöntemleri ile karşılaştırıldığı çalışmaların yapılması önerilmektedir.

4- Yapılan araştırmalarda elde edilen veriler birçok değişken açısından farklılık gösterdiğinden, kayıp gözlem tamamlama yöntemlerinde paralel çeşitliliğin sağlanması gerekmektedir. Bu nedenle kayıp gözlem tamamlamada yeni yöntem geliştirme çalışmalarının artarak devam etmesi önerilmektedir.

KAYNAKÇA

- Acock, Alan C.(2005), “Working with missing values”, *Journal of Marriage and Family*, Vol.67, s.1012-1028. <https://doi.org/10.1111/j.1741-3737.2005.00191.x>
- Akbaş, U. ve Tavşancıl, E. (2015). Farklı örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp gözlem baş etme teknikleri ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 38-57. <https://doi.org/10.21031/epod.26476>
- Allison, PD. (2002). *Missing data*. Thousands Oaks: Sage Publication. <https://doi.org/10.4135/9781412985079>
- Allison. P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*. 4(1), 545-557. <https://doi.org/10.1037/0021-843X.112.4.545>
- Altaş, D. ve Kaspar, E.Ç. (2012). Propensity skor ve hot-deck veri atama yöntemlerinin karşılaştırılması. *Trakya Üniversitesi İktisadi ve İdari Bilimler Fakültesi E-Dergi*, 1(1), 26-41.
- Bal C. (2003). *Çok gruplu veri setlerinde kayıp veri sorununun çözümlenmesi ve sağlık alanında bir uygulama*, Doktora Tezi, Eskişehir Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü, Eskişehir.
- Bayhan, A. (2018). *Farklı koşullardaki kayıp veri oranının iç tutarlığa etkisi*. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Büyüköztürk Ş. (2018), *Sosyal Bilimler İçin Veri Analizi El Kitabı*, Baskı:24, Pegem akademi, Ankara. <https://doi.org/10.14527/9789756802748>
- Brown, M. L., & Kros, J. F. (2003). Data mining and the impact of missing data. *Industrial Management & Data Systems*. <https://doi.org/10.1108/02635570310497657>
- Byrne, B. (2000). *Structural equation modelingwith AMOS: Basic concepts, applications, andprogramming*. Mahwah, NJ: Lawrence Erlbaum.
- Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology*, 21, 559-566. <https://doi.org/10.1076/jcen.21.4.559.889>

- Cheema, J. (2012). *Handling Missing Data In Educational Research Using Spss*. Unpublished Doctoral Dissertation. George Mason University, USA.
- Çokluk, Ö. ve Kayrı, M. (2011). Kayıp değerlere yaklaşık değer atama yöntemlerinin ölçme araçlarının geçerlik ve güvenilirliği üzerindeki etkisi. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(1), 289-309.
- Çüm, S., Demir, EK., Gelbal, S., & Kışla, T. (2018). Kayıp veriler yerine yaklaşık değer atamak için kullanılan gelişmiş yöntemlerin farklı koşullar altında karşılaştırılması. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, (45), 230-249. <https://doi.org/10.21764/maeuefd.332605>
- Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS örneği. *Eğitim Bilimleri Araştırmaları Dergisi*, 3(2), 47-68. <https://doi.org/10.12973/jesr.2013.324a>
- Demir, E. ve Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde Ve Psikolojide Ölçme Ve Değerlendirme Dergisi*. 3(1), 230-241.
- Enders, C. K. (2010). *Applied missing data analysis*. NY: The Guilford Press.
- Engels, J. M. ve Diehr, P. (2003). Imputation of missing longitudinal data: a comparison of methods. *Journal of Clinical Epidemiology*, 56(1), 968-976. [https://doi.org/10.1016/S0895-4356\(03\)00170-7](https://doi.org/10.1016/S0895-4356(03)00170-7)
- Ginkel, J.R.V., Van der Ark, L.A., Sijtsma, K. & Vermunt, J.K. (2007). Two-Way imputation: a bayesian method for estimating missing scores in tests and questionnaires, and an accurate approximation. *Computational Statistics & Data Analysis*, 51, 4013 – 4027. <https://doi.org/10.1016/j.csda.2006.12.022>
- Hohensinn, C. & Kubinger K. D. (2011). On the impact of missing values on item fit and the model validness of the rasch model. *Psychological Test and Assessment Modeling*, 53, 380-393.
- Işıkoğlu, M. A. (2017). *Kayıp veri ile baş etme yöntemlerinin ölçme değişmezliğine etkisi açısından karşılaştırılması*. Yayınlanmamış Yüksek Lisans Tezi. Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

- Koçak, D. ve Çokluk Bökeoğlu, Ö. (2017). Kayıp veriyle baş etme yöntemlerinin model veri uyumu ve madde model uyumuna etkisi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(2), 200-223. <https://doi.org/10.21031/epod.303753>
- Köse, A. (2014). The effect of missing data handling methods on goodness of fit indices in confirmatory factor analysis. *Educational Research and Reviews*, 9(8), 208-215. <https://doi.org/10.5897/ERR2014.1709>
- Köse, İ. A. ve Öztemur, B. (2014). Kayıp veri ele alma yöntemlerinin t-testi ve anova parametreleri üzerine etkisinin incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 14(1), 400-412. <https://doi.org/10.17240/aibuefd.2014.14.1-5000091519>
- Little, R. ve Rubin, D. (1987). *Statistical Analysis With Missing Data*. New York: Wiley.
- McKnight, P., McKnight, K., Sidani, S, S., & Figueredo, A. (2007). *Missing data: A gentle introduction*. Guilford Press.
- Osborne, J. W. (2013). *Best practices in data cleaning*. California: Sage Publication. Inc.
- Peng, C., Harwell, m., Liou, S., & Ehman, L. (2006). Advances in missing data methods and implications for educational research. *Real data analysis*, 3178.
- Pigott, T. D. (2001). A Review of methods for missing data. *Educational Resarch and Evaluation*. 7(1), 353-383. <https://doi.org/10.1076/edre.7.4.353.8937>
- Saunders, J. A., Morrow-Howell, N., Spitznagel, P. D., Proctor, E.K. ve Pescarino, R. (2006). Imputing missing data: a comparison of methods for social work researchers. *Social Work Research*, 30(1). <https://doi.org/10.1093/swr/30.1.19>
- Şahin Kürşad, M. (2014). *Sıklıkla kullanılan kayıp veri yöntemlerinin betimsel istatistik, güvenirlik ve geçerlik açısından karşılaştırması*. Yüksek Lisans Tezi, Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü, Bolu. <https://doi.org/10.21031/epod.95917>
- Şahin Kürşad, M. ve Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin ölçüklerin geçerlik ve güvenirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 254-267.
- Vansteelandt, S., Carpenter, J. & Kenward, M.G. (2010). Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology*, 6(1), 37-48.

Yurdugül, H. (2008). Minimum sample size for cronbach's coefficient alpha: A Monte- Carlo Study. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35, 397-405.
<https://doi.org/10.1027/1614-2241/a000005>

EKLER

Ek 1. :

```
public class MainActivity extends AppCompatActivity {

    private static final String TAG = "MainActivity";

    private String[] FilePathStrings;
    private String[] FileNameStrings;
    private File[] ListFile;
    File file;

    Button btnUpDirectory, btnSDCard;

    ArrayList<String> pathHistory;
    String lastDirectory;
    int count = 0;
    int dataCount=0;

    ArrayList<XYValue> uploadData;

    ListView lvInternalStorage;

    @RequiresApi(api = Build.VERSION_CODES.M)
    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate(savedInstanceState);

        setContentView(R.layout.activity_main);
        lvInternalStorage = (ListView) findViewById(R.id.lvInternalStorage);
        btnUpDirectory = (Button) findViewById(R.id.btnUpDirectory);
        btnSDCard = (Button) findViewById(R.id.btnViewSDCard);
        uploadData = new ArrayList<>();

        checkFilePermissions();

        lvInternalStorage.setOnItemClickListener(new AdapterView.OnItemClickListener() {
            @Override
            public void onItemClick(AdapterView<?> adapterView, View view, int i, long l) {
                lastDirectory = pathHistory.get(count);
                if(lastDirectory.equals(adapterView.getItemAtPosition(i))){
                    Log.d(TAG, msg: "lvInternalStorage: Selected a file for upload: " + lastDirectory);

                    readExcelData(lastDirectory);

                }else
                {
                    count++;
                    pathHistory.add(count, (String) adapterView.getItemAtPosition(i));
                    checkInternalStorage();
                    Log.d(TAG, msg: "lvInternalStorage: " + pathHistory.get(count));
                }
            }
        });
    }
}
```

```

btnUpDirectory.setOnClickListener(new View.OnClickListener() {
    @Override
    public void onClick(View view) {
        if(count == 0){
            Log.d(TAG, msg: "btnUpDirectory: You have reached the highest level directory.");
        }else{
            pathHistory.remove(count);
            count--;
            checkInternalStorage();
            Log.d(TAG, msg: "btnUpDirectory: " + pathHistory.get(count));
        }
    }
});

btnSDCard.setOnClickListener(new View.OnClickListener() {
    @Override
    public void onClick(View view) {
        count = 0;
        pathHistory = new ArrayList<String>();
        pathHistory.add(count, System.getenv( name: "EXTERNAL_STORAGE"));
        Log.d(TAG, msg: "btnSDCard: " + pathHistory.get(count));
        checkInternalStorage();
    }
});
}

```

```

/**
 *reads the excel file columns then rows. Stores data as ExcelUploadData object
 * @return
 */
private void readExcelData(String filePath) {
    Log.d(TAG, msg: "readExcelData: Reading Excel File.");

    File inputFile = new File(filePath);

    try {
        InputStream inputStream = new FileInputStream(inputFile);
        XSSFWorkbook workbook = new XSSFWorkbook(inputStream);
        XSSFSheet sheet = workbook.getSheetAt( index: 0);
        int rowCount = sheet.getPhysicalNumberOfRows();
        FormulaEvaluator formulaEvaluator = workbook.getCreationHelper().createFormulaEvaluator();
        StringBuilder sb = new StringBuilder();

        for (int r = 1; r < rowCount; r++) {
            Row row = sheet.getRow(r);
            int cellsCount = row.getPhysicalNumberOfCells();
            for (int c = 0; c < cellsCount; c++) {
                if(c>cellsCount){
                    Log.e(TAG, msg: "readExcelData: ERROR. Excel File Format is incorrect! ");
                    toastMessage("ERROR: Excel File Format is incorrect!");
                    break;
                }else{
                    String value = getCellAsString(row, c, formulaEvaluator);
                    String cellInfo = "r:" + r + " ; c:" + c + " ; v:" + value;
                    sb.append(value + " , ");
                }
            }
        }
    }
}

```



```

    }
    Log.d(TAG, msg: "readExcelData: STRINGBUILDER: " + sb.toString());

    parseStringBuilder(sb);

} catch (FileNotFoundException e) {
    Log.e(TAG, msg: "readExcelData: FileNotFoundException. " + e.getMessage() );
} catch (IOException e) {
    Log.e(TAG, msg: "readExcelData: Error reading inputstream. " + e.getMessage() );
}
}

```

```

/**
 * Method for parsing imported data and storing in ArrayList<XYValue>
 */

```

```

public void parseStringBuilder(StringBuilder mStringBuilder) {
    Log.d(TAG, msg: "parseStringBuilder: Started parsing.");
    String[] rows = mStringBuilder.toString().split( regex: ":" );
    for (int i = 0; i < rows.length; i++) {
        String[] columns = rows[i].split( regex: "," );
        try {
            String ilqi1 = (columns[0]);
            String ilqi2 = (columns[1]);
            String ilqi3 = (columns[2]);
            String ilqi4 = (columns[3]);
            String ilqi5 = (columns[4]);
            String ilqi6 = (columns[5]);
            String ilqi7 = (columns[6]);
            String ilqi8 = (columns[7]);
            String ilqi9 = (columns[8]);
            String ilqi10 = (columns[9]);

```

```

            String ilqi11 = (columns[10]);
            String ilqi12 = (columns[11]);
            String ilqi13 = (columns[12]);
            String ilqi14 = (columns[13]);
            String ilqi15 = (columns[14]);
            String ilqi16 = (columns[15]);
            String ilqi17 = (columns[16]);
            String ilqi18 = (columns[17]);

```

```

            String cellInfo = "(ilqi1,ilqi2,ilqi3,ilqi4,ilqi5,ilqi6,ilqi7,ilqi8,ilqi9,ilqi10,ilqi11,ilqi12,ilqi13,ilqi14,ilqi15,ilqi16,ilqi17,ilqi18)";
            uploadData.add(new XYValue(ilqi1, ilqi2, ilqi3, ilqi4, ilqi5, ilqi6, ilqi7, ilqi8,

```

```

            System.out.println(uploadData.get(i).toString());

```

```

        } catch (NumberFormatException e) {

```

```

            Log.e(TAG, msg: "parseStringBuilder: NumberFormatException: " + e.getMessage());

```

```

        }

```

```

    }

```

```

    System.out.println(uploadData.get(0).ilqi1);

```

```

    System.out.println(uploadData.get(1).ilqi1);

```

```

    for (int a = 0; a < uploadData.size(); a++) {

```

```

        for (int w = 0; w < uploadData.size(); w++) {

```

```

    if (uploadData.get(a).ilgi1.equals(uploadData.get(w).ilgi1)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi2.equals(uploadData.get(w).ilgi2)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi3.equals(uploadData.get(w).ilgi3)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi4.equals(uploadData.get(w).ilgi4)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi5.equals(uploadData.get(w).ilgi5)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi6.equals(uploadData.get(w).ilgi6)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi7.equals(uploadData.get(w).ilgi7)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi8.equals(uploadData.get(w).ilgi8)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi9.equals(uploadData.get(w).ilgi9)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi10.equals(uploadData.get(w).ilgi10)) {
        dataCount++;
    }

```

```

    if (uploadData.get(a).ilgi11.equals(uploadData.get(w).ilgi11)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi12.equals(uploadData.get(w).ilgi12)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi13.equals(uploadData.get(w).ilgi13)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi14.equals(uploadData.get(w).ilgi14)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi15.equals(uploadData.get(w).ilgi15)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi16.equals(uploadData.get(w).ilgi16)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi17.equals(uploadData.get(w).ilgi17)) {
        dataCount++;
    }
    if (uploadData.get(a).ilgi18.equals(uploadData.get(w).ilgi18)) {
        dataCount++;
    }
    System.out.println("öğrenci" + a + "karsılastırma" + w + "=" + dataCount);
    dataCount = 0;
}

```

```

private void printDataToLog() {
    Log.d(TAG, "msg: \"printDataToLog: Printing data to log...\");

    for(int i = 0; i< uploadData.size(); i++){
        String ilgi1 = String.valueOf(uploadData.get(i).getilgi1());
        String ilgi2 = String.valueOf(uploadData.get(i).getilgi2());
        String ilgi3 = String.valueOf(uploadData.get(i).getilgi3());
        String ilgi4 = String.valueOf(uploadData.get(i).getilgi4());
        String ilgi5 = String.valueOf(uploadData.get(i).getilgi5());
        String ilgi6 = String.valueOf(uploadData.get(i).getilgi6());
        String ilgi7 = String.valueOf(uploadData.get(i).getilgi7());
        String ilgi8 = String.valueOf(uploadData.get(i).getilgi8());
        String ilgi9 = String.valueOf(uploadData.get(i).getilgi9());
        String ilgi10 = String.valueOf(uploadData.get(i).getilgi10());
        String ilgi11 = String.valueOf(uploadData.get(i).getilgi11());
        String ilgi12 = String.valueOf(uploadData.get(i).getilgi12());
        String ilgi13 = String.valueOf(uploadData.get(i).getilgi13());
        String ilgi14 = String.valueOf(uploadData.get(i).getilgi14());
        String ilgi15 = String.valueOf(uploadData.get(i).getilgi15());
        String ilgi16 = String.valueOf(uploadData.get(i).getilgi16());
        String ilgi17 = String.valueOf(uploadData.get(i).getilgi17());
        String ilgi18 = String.valueOf(uploadData.get(i).getilgi18());
    }
}

```

```

/**
 * Returns the cell as a string from the excel file
 * @param row
 * @param c
 * @param formulaEvaluator
 * @return
 */
private String getCellAsString(Row row, int c, FormulaEvaluator formulaEvaluator) {
    String value = "";
    try {
        Cell cell = row.getCell(c);
        CellValue cellValue = formulaEvaluator.evaluate(cell);
        switch (cellValue.getCellType()) {
            case Cell.CELL_TYPE_BOOLEAN:
                value = ""+cellValue.getBooleanValue();
                break;
            case Cell.CELL_TYPE_NUMERIC:
                double numericValue = cellValue.getNumberValue();
                if(HSSFDateUtil.isCellDateFormatted(cell)) {
                    double date = cellValue.getNumberValue();
                    SimpleDateFormat formatter =
                        new SimpleDateFormat( pattern: "MM/dd/yy");
                    value = formatter.format(HSSFDateUtil.getJavaDate(date));
                } else {
                    value = ""+numericValue;
                }
                break;
            case Cell.CELL_TYPE_STRING:
                value = ""+cellValue.getStringValue();
                break;
        }
    }
}

```

```

        default:
    }
} catch (NullPointerException e) {

    Log.e(TAG, msg: "getCellAsString: NullPointerException: " + e.getMessage() );
}
return value;
}

private void checkInternalStorage() {
    Log.d(TAG, msg: "checkInternalStorage: Started.");
    try{
        if (!Environment.getExternalStorageState().equals(
            Environment.MEDIA_MOUNTED)) {
            toastMessage("No SD card found.");
        }
        else{
            file = new File(pathHistory.get(count));
            Log.d(TAG, msg: "checkInternalStorage: directory path: " + pathHistory.get(count));
        }

        listFile = file.listFiles();

        FilePathStrings = new String[listFile.length];

        FileNameStrings = new String[listFile.length];

        for (int i = 0; i < listFile.length; i++) {
            FilePathStrings[i] = listFile[i].getAbsolutePath();
            FileNameStrings[i] = listFile[i].getName();
        }
    }
}

```

```

        for (int i = 0; i < listFile.length; i++)
        {
            Log.d(tag: "Files", msg: "FileName:" + listFile[i].getName());
        }

        ArrayAdapter<String> adapter = new ArrayAdapter<String>( context: this, android.R.layout.simple_list_item_1, FilePathStrings);
        lvInternalStorage.setAdapter(adapter);

    }catch(NullPointerException e){
        Log.e(TAG, msg: "checkInternalStorage: NULLPOINTEREXCEPTION " + e.getMessage() );
    }
}

@RequiresApi(api = Build.VERSION_CODES.N)
private void checkFilePermissions() {
    if(Build.VERSION.SDK_INT > Build.VERSION_CODES.LOLLIPOP){
        int permissionCheck = this.checkSelfPermission("Manifest.permission.READ_EXTERNAL_STORAGE");
        permissionCheck += this.checkSelfPermission("Manifest.permission.WRITE_EXTERNAL_STORAGE");
        if (permissionCheck != 0) {

            this.requestPermissions(new String[]{Manifest.permission.WRITE_EXTERNAL_STORAGE, Manifest.permission.READ_EXTERNAL_STORAGE}, requestCode: 1001); //Any number
        }
    }else{
        Log.d(TAG, msg: "checkBTPermissions: No need to check permissions. SDK version < LOLLIPOP.");
    }
}

```

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı: AHMET ŞENGÜL

Doğum Yeri ve Tarihi:

Eğitim Durumu

Lisans Öğrenimi:

Yüksek Lisans Öğrenimi:

Bildiği Yabancı Diller:

Bilimsel Faaliyetleri:

İş Deneyimi

Stajlar :

Projeler :

Çalıştığı Kurumlar:

İletişim:

Tarih : 22/ 07/ 2022

YAYIN İNTİHAL RAPORU

EĞİTİM BİLİMLERİNDE EKSİK GÖZLEM VERİLERİ TAMAMLAMA YÖNTEMLERİNİN KARŞILAŞTIRMASI VE YENİ YÖNTEM ÇALIŞMASI

ORJİNALLİK RAPORU

% **19**
BENZERLİK ENDEKSİ

% **19**
İNTERNET KAYNAKLARI

% **3**
YAYINLAR

% **5**
ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	www.openaccess.hacettepe.edu.tr:8080 İnternet Kaynağı	% 4
2	acikbilim.yok.gov.tr İnternet Kaynağı	% 3
3	dergipark.org.tr İnternet Kaynağı	% 3
4	dergipark.ulakbim.gov.tr İnternet Kaynağı	% 2
5	Submitted to Akdeniz University Öğrenci Ödevi	% 1
6	dspace.akdeniz.edu.tr:8080 İnternet Kaynağı	% 1
7	www.efdergi.hacettepe.edu.tr İnternet Kaynağı	% 1
8	docplayer.biz.tr İnternet Kaynağı	% 1

www.fed.sakarya.edu.tr

EK-9. BİLDİRİM SAYFASI

BİLDİRİM

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kâğıt ve elektronik kopyalarının Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- Tezim/Raporum sadece Akdeniz Üniversitesi yerleşkelerinden erişime açılabilir. Tezimin/Raporumun yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

22.07.2022

Ahmet ŞENGÜL