

Nicel Arařtırmalarda Kayıp Veriler ve Uç Deęerler

Çözüm Önerileri ve SPSS Uygulamaları

Ufuk AKBAŞ • Hakan KOĖAR

3. Baskı





Doç. Dr. Ufuk AKBAŞ - Doç. Dr. Hakan KOĞAR

NİCEL ARAŞTIRMALARDA KAYIP VERİLER VE UÇ DEĞERLER Çözüm Önerileri ve SPSS Uygulamaları

ISBN 978-625-7052-22-1

Kitap içeriğinin tüm sorumluluğu yazarlarına aittir.

© 2023, PEGEM AKADEMİ

Bu kitabın basım, yayım ve satış hakları Pegem Akademi Yay. Eğt. Dan. Hizm. Tic. AŞ'ye aittir. Anılan kuruluşun izni alınmadan kitabın tümü ya da bölümleri, kapak tasarımı; mekanik, elektronik, fotokopi, manyetik kayıt ya da başka yöntemlerle çoğaltılamaz, basılamaz ve dağıtılamaz. Bu kitap, T.C. Kültür ve Turizm Bakanlığı bandrolü ile satılmaktadır. Okuyucularımızın bandrolü olmayan kitaplar hakkında yayineimize bilgi vermesini ve bandrolsüz yayınları satın almamasını diliyoruz.

Pegem Akademi Yayıncılık, 1998 yılından bugüne uluslararası düzeyde düzenli faaliyet yürüten **uluslararası akademik bir yayinevi**dir. Yayımladığı kitaplar; Yükseköğretim Kurulunca tanınan yükseköğretim kurumlarının kataloglarında yer almaktadır. Dünyadaki en büyük çevrimiçi kamu erişim kataloğu olan **WorldCat** ve ayrıca Türkiye'de kurulan **Turcademy.com** tarafından yayınları taranmaktadır, indekslenmektedir. Aynı alanda farklı yazarlara ait 1000'in üzerinde yayını bulunmaktadır. Pegem Akademi Yayınları ile ilgili detaylı bilgilere <http://pegem.net> adresinden ulaşılabilmektedir.

1. Baskı: Mart 2020, Ankara
3. Baskı: Mart 2023, Ankara

Yayın-Proje: Ferdi Akkaya
Dizgi-Grafik Tasarım: Tuğba Kaplan
Kapak Tasarım: Pegem Akademi

Baskı: Sonçağ Yayıncılık Matbaacılık Reklam San Tic. Ltd. Şti.
İstanbul Cad. İstanbul Çarşısı 48/48 İskitler/Ankara
Tel: (0312) 341 36 67

Yayıncı Sertifika No: 51818
Matbaa Sertifika No: 47865

İletişim

Macun Mah. 204. Cad. No: 141/A-33 Yenimahalle/ANKARA
Yayınevi: 0312 430 67 50
Dağıtım: 0312 434 54 24
Hazırlık Kursları: 0312 419 05 60
İnternet: www.pegem.net
E-ileti: pegem@pegem.net
WhatsApp Hattı: 0538 594 92 40

Doç. Dr. Ufuk AKBAŞ

1982 yılında Bucak'ta doğdu. İlk ve ortaöğrenimi Antalya'da tamamladı. Lisans eğitimini 2007 yılında Gazi Üniversitesi Gazi Eğitim Fakültesi İlköğretim Matematik Öğretmenliği; lisansüstü eğitimini ise 2014 yılında Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme anabilim dalında tamamladı. 2010 – 2015 yılları arasında Hoca Ahmet Yesevi Uluslararası Türk – Kazak Üniversitesi Türkiye Türkçesi ile Uzaktan Eğitim Programları (TÜRTEP) bünyesinde yer alan Ölçme, Değerlendirme ve Sınav Hizmetleri biriminde görev yaptı. 2015 yılından itibaren Hasan Kalyoncu Üniversitesi Eğitim Fakültesi'nde akademik ve idari personel olarak görev yapmaktadır.

Lisans, yüksek lisans ve doktora düzeyinde ölçme ve değerlendirme, bilimsel araştırma yöntemleri, ölçek geliştirme ve uyarlama, eğitim istatistiği, test kuramları, tamamlayıcı ölçme – değerlendirme uygulamaları ve çok değişkenli istatistik gibi dersleri yürütmektedir. İlgili alanları arasında çok değişkenli istatistiksel yöntemler, ölçme aracı geliştirme ve uyarlama, test kuramları ve bilimsel araştırma yöntemleri yer almaktadır.

Doç. Dr. Hakan KOĞAR

1986 yılı Eskişehir doğumludur. İlk ve orta öğretimini Eskişehir'de tamamladıktan sonra lisans eğitimini 2008 yılında Hacettepe Üniversitesi Eğitim Fakültesi Sınıf Öğretmenliği, yüksek lisansını eğitimini 2010 yılında Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme, doktora eğitimini ise 2014 yılında Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme anabilim dalında tamamladı. 2009-2014 yılları arasında Başkent Üniversitesi Eğitim Fakültesinde araştırma görevlisi, 2014-2015 yılları arasında ise aynı fakültede öğretim görevlisi olarak çalışmıştır. 2015 yılının ekim ayından bugüne Akdeniz Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Bölümü Eğitimde Ölçme ve Değerlendirme ABD'nde doktor öğretim üyesi olarak görev almaktadır. 2019 yılının haziran ayında doçent unvanını almıştır.

Lisans düzeyinde ölçme ve değerlendirme, istatistik, bilimsel araştırma yöntemleri; yüksek lisans düzeyinde test geliştirme teknikleri, maksimum performansın ölçülmesi, psikolojik ölçmenin temelleri, bilgisayarda ölçme uygulamaları, ölçme ve değerlendirmede güncel konular; doktora düzeyinde ise ileri bilimsel araştırma teknikleri ve etik gibi dersleri yürütmektedir. Madde tepki kuramı uygulamaları, mokken ölçekleme analizleri, parametrik olmayan madde tepki kuramı, boyutluluk analizleri, test geliştirme ve çok değişkenli istatistiksel teknikler akademik ilgi alanlarını oluşturmaktadır.

ÖN SÖZ

Hemen hemen tüm arařtırmalarda toplanan veriler, bir miktar kayıp veri ve uç deęer içermektedir. Kayıp veriler ve uç deęerler, geręekleřtirilen istatistiksel analizlerin sonuçlarını olumsuz etkileyebilmekte hatta tümüyle isabetsiz kılabilir. İstatistiksel analizleri konu edinen pek çok kaynakta, kayıp veriler ve uç deęerler için yapılabilecek incelemelere ve işlemlere yer verilmektedir. Fakat bu kaynakların odağında çeřitli istatistiksel analizler bulunduęu için verilen bilgiler ve örnekler de büyük ölçüde sınırlı kalmaktadır. Yapılan çalıřmalar da arařtırmacıların, verilerin analize hazırlığı sürecindeki işlemlere yeterince ağırlık vermediğini ya da veremediğini göstermektedir.

Lisansüstü eğitimimiz süresince ilgimizi çeken ve kendi arařtırmalarımıza konu ettiğimiz kayıp veriler ve uç deęerlerle ilgili okumalarımızı yaparken, yurtdışında bu konular üzerine yazılmış pek çok kaynağın bulunduğunu gördük. Yukarıda da bahsettiğimiz gibi, arařtırmacıların bu konulardaki incelemeleri ve işlemleri yapmakta zorlanmalarının, Türkiye’de bu konularla ilgili kılavuz niteliğinde bir kaynağın bulunmamasıyla ilgili olabileceğini düşündük. Bu noktadan hareketle arařtırmacılara, kayıp verileri ve uç deęerleri nasıl önleyebileceklerini, önleyemedikleri durumlarda ise hangi işlemleri nasıl yapabileceklerini gösteren bir kitap hazırlamaya karar verdik.

Şüphesiz bu kitap, konuyla ilgili tüm ayrıntıları ve olası durumlar içermektedir. Örneğin, sunduğumuz senaryolar ve kullandığımız veri setleri büyük ölçüde sürekli deęişkenleri içermekte ve genel olarak büyük örneklemeleri temsil etmektedir. Benzer şekilde, literatürde kayıp verilerle baş etmede ve uç deęerleri belirlemede kullanılabilen çok sayıda teknik arasından bir kısmına yer verdik. Bu sınırlamaları, kitaptan faydalanmasını umduğumuz arařtırmacıların alanlarının ve eğitim düzeylerinin çok farklı olabileceğini dikkate alarak gerekli ve uygun gördük. Konularla ilgili teknik açıklamalara yeri geldikçe girmeye; mümkün olduğunca sade ve yalın bir anlatım tarzı benimsemeye çalıştık. Bu kitabın hem ileri düzey arařtırmacılar için temel bir kaynak hem de akademik çalıřmalara yeni başlayan arařtırmacılar için doyurucu bir kılavuz olmasını umuyoruz; anlatım ve örneklerimizdeki hataların bizlere bildirilmesinden mutluluk duyacağımızı beyan ediyoruz.

Doç. Dr. Ufuk AKBAŞ

ORCID No: 0000-0002-6122-154X

Doç. Dr. Hakan KOĞAR

ORCID No: 0000-0001-5749-9824

İÇİNDEKİLER

Ön Söz.....v

1. BÖLÜM

KAYIP VERİLER

Kayıp Veri Nedir?	1
Kayıp Veriler Neden Sorun Oluşturur?	5
Kayıp Veri Mekanizması ve Kayıp Veri Örüntüsü	6
Kayıp veri mekanizmasının incelenmesi	10
a. Ön inceleme.....	11
b. Kayıp veri örüntüsünün incelenmesi	12
c. Kayıp veri içeren ve içermeyen satırlar üzerinden ortalama puanların karşılaştırılması	18
d. Kayıp verilerin kategorik değişkenlere göre dağılımının incelenmesi	20
e. Kayıp veri içeren değişkenler arasındaki ilişkilerin incelenmesi	21
f. Little'ın TSK testi (Little's MCAR test)	22
Mekanizmanın incelenmesinde dikkat edilmesi gereken hususlar	23
Kayıp Veri Teknikleri	24
a. Liste bazında silme (Listwise deletion)	26
b. Çift bazında silme (Pairwise deletion)	26
c. Ortalama değer atama (Mean imputation).....	28
d. Diğer basit atama teknikleri	28
e. Regresyonla değer atama (Regression imputation)	31
f. Stokastik regresyonla değer atama (Stochastic regression imputation)	35
g. Beklenti – Maksimizasyon (Expectation – Maximization / EM)	36
h. Çoklu değer atama (Multiple imputation / MI)	38
Kayıp Verilerin Azaltılmasına Yönelik Öneriler.....	43
Kayıp Veri Tekniğinin Belirlenmesi.....	47

2. BÖLÜM

UÇ DEĞERLER

Uç Değer Kavramı.....	51
Uç Değerlerin Tarihçesi.....	52
Uç değerlerin olası ortaya çıkma sebepleri	54
a. Veri hatalarından kaynaklanan uç değerler	54
b. Kasıtlı ya da farklı bir nedenden kaynaklı yanlış raporlamalar sonucunda ortaya çıkan uç değerler	54

c. Örneklem hatasından kaynaklanan uç deęerler	54
d. Uygulamadan kaynaklanan uç deęerler	55
e. Hatalı ya da yanlış ölçme yapan ölçme araçlarından kaynaklanan uç deęerler	55
f. Dağılıma ait varsayımlara ilişkin hatalardan kaynaklanan uç deęerler	55
g. Doğru evrenden alınan uygun örneklemeye rağmen gözlenen uç deęerler	56
h. Arařtırmanın odağından kaynaklanan uç deęerler	56
Uç Deęerlerin İstatistiksel Analiz Sonuçları Üzerindeki Olası Etkileri	56
Uç Deęer Belirleme Yöntemleri	59
a. Tek deęişkenli uç deęer belirleme yöntemleri	59
1. Z puanına dönüřtürme	59
2. SPSS ile grafiksel teknikler kullanılarak uç deęerlerin belirlenmesi	62
b. Çok Deęişkenli Uç Deęer Belirleme Yöntemleri	67
1. Mahalanobis Uzaklığı	67
Uç Deęerlerle Baş Etme Yöntemleri	71
Uç deęerlere veri dönüřümü (transformasyon) uygulanması	71
a. Karekök dönüřümü (square-root transformation)	72
b. Ters dönüřüm (inverse transformation)	72
c. Logaritmik dönüřüm (logarithmic transformation)	72
d. Arc sinüs dönüřümü (arc-sinus transformation)	72
SPSS İle Logaritmik Dönüřüm Kullanılarak Uç Deęerlerin Belirlenmesi	74
Uç Deęerlerin Uyumunu Saęlama	75
a. Kırpılmış ortalama	75
b. Merkezi eğilim ortalaması	76
c. Ucunu kesme	77
Uç Deęerlerin Silinmesi	77
KAYNAKÇA	79

1. BÖLÜM

KAYIP VERİLER

Kayıp Veri Nedir?

Bilimsel çalışmalar problemin fark edilmesi, araştırmanın tasarlanması, yürütülmesi ve bulgularının sunulması gibi kapsamlı süreçlerin izlenmesini içerir. Büyük özveri ve dikkat gerektiren bu süreçlere konu olan hipotezler, araştırma kapsamında toplanan veriler kullanılarak uygun analizlerle test edilir. Genellikle bilgisayar üzerinden uygulanan veya tüm maddelerin yanıtlanmasının zorunlu olduğu özel durumlar haricinde, elde edilen verilerde bazı kayıpların bulunması sık karşılaşılan bir sorundur. Kayıp veri terimi en genel şekliyle, bir araştırmada yer verilen değişkenlere ait ölçümlerden en az birinin elde edilemediğini ifade etmektedir. Kayıp veriler, ortaya çıkma şekline göre madde yanıtlanmama (item nonresponse), birim yanıtlanmama (unit nonresponse), dalgalı kayıp (wave nonresponse) ve dönüşsüz kayıp (dropout / attrition) şeklinde adlandırılmaktadır.

Madde yanıtlanmama şeklindeki kayıplar, araştırmaya katılan bir bireyin en az bir maddeye ait verisinin elde edilemediği durumları ifade etmektedir. Madde yanıtlanmamanın arkasında, testi alan bireyin dikkatsizliğine bağlı olarak maddeyi atlaması, maddeye nasıl cevap vereceğinden emin olamaması, daha sonra cevap vermek üzere cevapsız bıraktığı maddeye dönmeyi unutması gibi sebepler bulunabilir. Bireyin, cevabına bağlı olarak zarar görebileceğini düşünmesi veya okuma hızı düşük olduğu için testin sonundaki sorulara erişememesi madde yanıtlanmama şeklindeki kayıplara zemin oluşturabilir (Graham, 2012). Bunların dışında Field (2018) bireylerin, cinsellik gibi mahremiyet içeren konularla ilgili maddelere cevap vermek istemeyebileceklerini ve ölçme işleminde kullanılan mekanik araçlardaki arızaların da veri kaybına neden olabileceğini ifade etmektedir.

Gerçekte gözlenmiş olduğu halde, veri girişi esnasında kayıp verileri temsil etmek üzere seçilen değer (ya da aralığın) ölçüm noktalarına karşılık gelmesi de kullanılan istatistik paket programının ilgili değeri kayıp veri olarak algılamasına neden olabilir. Örneğin kayıp verilerin “9” değeri ile kodlandığı bir araştırmada dokuz yaşındaki bir bireyin yaşı, kayıp veri olarak işlem görecektir. Benzer şekilde,

veri giriři esnasında bazı hücreler boş bırakılmıř olabilir (*system missing*). İstatistik paket programları boş hücreleri kayıp veri olarak deęerlendirmektedir. Kayıp veriler için özel bir sayısal deęer belirlenmiřse, veri giriřinde tutarlı bir řekilde bu deęerin dikkate alınması avantaj saęlayacaktır (*user defined missing*). Bylece veri setindeki boş hücrelerin dikkatsizlikten kaynaklandıęı kolayca anlařılabilir. Veri analizinden nce bu boş hücreler, uygulanan formlar zerinden yapılacak kontrollerle tamamlanabilir.

SPSS'te "Variable View" arayznde yer alan "Missing" stunu zerinden c farklı kayıp veri tri için rnek bir tanımlama řekil 1'de verilmiřtir. "Range plus one optional discrete missing value" seeneęi ile belirlenen aralıktaki kalan deęerlerin kayıp verileri temsil etmesi de saęlanabilir.

řekil 1. Kayıp verilerin tanımlanması

Grldę zere SPSS, farklı anlamlara gelen kayıplar için birden ok deęerin eř zamanlı olarak kullanımına imkn tanımaktadır. rneęin, bir ęrenci 20 maddeden oluřan bir testte sadece cnc soruyu cevaplamamıř ise "9"; son drt soruyu cevaplamamıř ise "99" deęeri tercih edilebilir. Burada ilk durum, ęrencinin maddeyi herhangi bir sebeple cevapsız bıraktıęını (atladıęını); ikinci durum ise testin sonundaki maddelere eriřemedięini ifade etmektedir. Benzer řekilde, oktan semeli bir testteki bir soruda iki seeneęin birden iřaretlenmesi gibi bir durum için farklı bir kodlama (rneęin, 999) yapılabilir.

Bazı durumlarda da ilgilenilen zellikler bireyler yerine kurumlara, coęrafi blgelere veya herhangi bir nesneye ait olabilir. rneęin farklı okulların ęretmen, ęrenci ve sınıf sayılarına, bahelerinin yzlcmlerine ve okulların hangi yılda inřa edildiklerine iliřkin elde edilen veri setinde, bir okul için ęrenci sayısına

ulaşılamamış olabilir. Benzer şekilde, belli bir tarihte bir ildeki hava sıcaklığına ya da bir bilgisayarın kapasitesine ait veriye ulaşılamadığı durumlarda karşılaşılan veriler de madde yanıtlamama kategorisi altında ele alınabilir.

Öte yandan, trafikte seyir halindeyken sergilenen davranışları içeren bir araştırmada, ehliyeti olmadığını söyleyen bir kişinin; algılanan eş desteğini içeren bir araştırmada, romantik ilişkisi bulunmayan bir kişinin uygulanan ölçme araçlarındaki bazı maddeleri atlaması gerekebilir. Bu gibi durumlarda atlanan maddeler için herhangi bir veri elde edilmesi zaten anlamlı olmayacağından, veri setindeki eksiklerin “kayıp” olarak ele alınması da anlamlı değildir. Görüldüğü üzere bu gibi durumlar genellikle, ölçme aracının önceki kısımlarında yer alan bir maddeye verilen cevabın ilerleyen kısımlardaki bazı maddelerin atlanmasını gerektirdiğinde ortaya çıkar. Burada verilen örnekler benzer nitelikte kayıplarla karşılaşılması beklenen bir durum ise ortada bir kayıp veri sorunu olduğu söylenemez. Fakat verilerin eksiksiz olması gerekiyorsa, bu bireylerin araştırmanın örnekleminde gerçekten yer alıp almamaları gerektiğini kontrol etmekte fayda vardır.

Birim yanıtlamama şeklindeki kayıplarla ise bireye (kuruma, nesneye, vb.) ait ölçümlerin hiçbirine ulaşılamadığı durumlarda karşılaşılır. Örnekleme giren bir öğrencinin o gün okula gelmemesi, araştırmaya katılmak istememesi veya maddelere cevap verecek zamanının olmaması gibi durumlarda, veri setinde bu kişiye ait hiç veri bulunmayacaktır.

Aynı bireylerden tekrarlı ölçümlerin alındığı boylamsal çalışmalarda dalgalı ve dönüşsüz kayıplarla karşılaşılabilir. Dalgalı kayıpta, katılımcılara belli bir ölçüm zamanında ulaşılamamakta fakat sonraki zamanlarda ulaşılabilir. Verinin bir noktadan sonraki tüm ölçüm zamanları için kayıp olması ise dönüşsüz kayıp olarak adlandırılmaktadır (Graham, 2012). Dalgalı ve dönüşsüz kayıpların daha iyi anlaşılabilmesi için ön test, son test ve izleme testi olmak üzere toplam üç farklı zamanda ölçüm alınmasının planlandığı ve tüm bireylerin ön teste katıldığı bir araştırma düşünelim. Sadece son teste katılmamış olan bireyler için kayıp veriler dalgalı kayıp olarak adlandırılırken, hem son teste hem de izleme testine katılmamış bireyler için dönüşsüz kayıptan söz edilebilir. Anlaşılacağı üzere, aynı veri setinde hem dalgalı hem de dönüşsüz kayıplar bulunabilir.

Buraya kadar olan kısımda bahsedilen kayıp veriler, araştırmacının kontrolü veya isteği dışında ortaya çıkmaktadır. Bazı araştırmalar özel olarak belli bir örnekte sergileyen kayıplar bulunacak şekilde tasarlanabilir. Örneğin, PISA ve TIMSS gibi geniş ölçekli test uygulamalarında, planlı kayıp veri desenine sahip veri setleri elde edilir. Bu uygulamalarda öğrencilere, farklı alt testler (ve dolayısıyla farklı maddeler) içeren kitapçıklar belli bir düzen içinde dağıtılmaktadır (OECD, 2014).

Dolayısıyla öęrenciler havuzdaki maddelerin bir kısmı ile karřılařmakta, karřılařmadıkları maddelere de doęal olarak cevap verememektedirler. Bu gibi durumlar için ise kayıp veriler yerine “999” deęeri kullanılabilir. Kısaca ifade etmek gerekirse testteki bir maddenin atlanması, testin sonundaki maddelere eriřilememesi ya da maddelerin katılımcıların bir kısmına uygulanmamıř olması gibi durumlar farklı anlamlara geldięi için kayıp verilerin, veri giriři esnasında farklı řekillerde kodlanması arařtırmacılara daha fazla bilgi saęlayabilir. Kayıplar arasındaki farkları ayırt etmek isteyen arařtırmacılara bir kodlama stratejisi geliřtirmeleri önerilir.

Bir arařtırmada veriler geniř ölçekli uygulamalara benzer řekilde, planlı kayıp veri deseni (planned missing data design / missing by design) olarak adlandırılan veri toplama süreci ile toplanabilir. Kullanılan form sayısına göre tek, iki ve üç formlu desen olarak adlandırılan veri toplama süreçleri ile veri setindeki kayıpların önceden belirlenmiř bir plana göre toplanması mümkündür (Graham, Hoffer ve MacKinnon, 1996). Sırasıyla tek, iki ve üç formun kullanıldıęı planlı kayıp desenleri Tablo 1’de gösterilmiřtir.

Tablo 1. Tek, iki ve üç formun yer aldıęı planlı kayıp veri desenleri

Desen	Uygulama grupları	Kullanılan ölçme araçları			
		X	A	B	C
Tek form	Grup 1	✓	✓	✓	✓
İki form	Grup 1	✓	✓	-	
	Grup 2	✓	-	✓	
Üç form	Grup 1	✓	✓	✓	-
	Grup 2	✓	✓	-	✓
	Grup 3	✓	-	✓	✓

Tablo 1’de yer alan “-“ iřareti, ölçme aracının ilgili gruba uygulanmadıęı anlamına gelmektedir. Bir arařtırma kapsamında yer alan bütün ölçme araçlarının örnekleme de yer alan tüm bireylere uygulanması halinde tek form desenden yararlanılmaktadır. Arařtırmada birden fazla ölçme aracı yer alıyor olsa bile bunları tüm katılımcılar cevaplamaktadır. Esasen tek form deseni için veri setinde planlı kayıpların bulunması beklenmemektedir. Arařtırmacının elinde üç farklı ölçme aracının olduęu durumlarda iki form deseni ile veri toplanabilir. Örneęin, ölçme araçlarının X, A ve B řeklinde adlandırıldıęı bir arařtırmada örneklemin yarısına X testi ile birlikte A; dięer yarısına ise X testi ile birlikte B testinin uygulanması halinde, elde edilen veriler önceden bilinen (planlı) bir örüntüye uygun kayıplar içerecektir. Uygulanacak test sayısının X, A, B ve C gibi dört adet olduęu durumlarda ise örneklemin üç gruba ayrılıp birinci gruba X-A-B, ikinci gruba X-A-C ve

üçüncü gruba X-B-C testlerinin uygulanmasıyla üç form deseninden yararlanarak veri toplanabilir.

Planlı kayıp veri desenleri araştırmacılara testlerin cevaplanması için gereken zaman, cevaplayıcı yükü ve kullanılan testlerin maliyeti vb. hususlarla ilgili sınırlıklara karşın avantaj sağlayabilir. Örneğin, farklı değişkenlerin modelde yer aldığı ve birkaç ölçeğin kullanılmasını gerektiren çalışmalarda, katılımcıların ölçeklerin tamamını cevaplamak için ihtiyaç duyacakları süre çok uzun olabilir. Her katılımcıdan bütün maddelerin cevaplanmasının istenmesi, özellikle son bölümlerde kalan maddeler için çok sayıda ve plansız kayıp verilerle karşılaşılmasına yol açabilir.

Araştırmalarda kullanılacak ölçme araçları belirlenirken, araçlarının psikometrik özelliklerinin yanında katılımcıların (yaş, eğitim düzeyi vb.) özelliklerinin de dikkate alınması gerekir. Bu bağlamda araştırmacı, düşük eğitim düzeyinde veya küçük yaş gruplarında yer alan bireylere uygulayacağı testlerdeki toplam madde sayısının fazla olmamasını göz önünde bulundurabilir. Fazla sayıda maddenin yer aldığı ölçme araçlarını kullanarak önceden planlanmamış kayıp veriler ile karşılaşmak yerine, iki ya da üç form deseninden faydalanarak cevaplayıcı yükü düşürülebilir. Zira cevaplayıcı yükü ile beraber kayıp verilerin de azalması beklenen bir durumdur.

Öte yandan, Çok Boyutlu Liderlik Ölçeği (Avolio ve Bass, 2004) gibi bazı ölçme araçları, çeşitli kurumlar tarafından ticari amaçlarla kullanılabilir. Bu tür testler, uygulanan kişi sayısına göre farklı şekillerde fiyatlandırılmaktadır. Testin, örneklemin tümü yerine belli bir kısmına uygulanması araştırmacıya daha az ekonomik yük getirecektir. Birden çok formun yer aldığı veri toplama desenleri araştırmacılara maddi açıdan da avantaj sağlayabilir.

Burada örnekleri verilen iki ve üç form desenleri, daha fazla sayıda ölçeğe veya formların oluşturulabileceği farklı kombinasyonlara uyarlanabilir. Anlaşılacağı üzere, bu tür veri toplama desenleri veri setinin bazı bölümlerinde kayıpların olmasını garanti eder. Bu gibi desenlerle elde edilen verilerin, ilerleyen bölümlerde bahsedilecek olan bazı kayıp veri baş etme teknikleri ile tamamlanması ve analiz edilmesi mümkündür. Planlı kayıp veri desenlerinin sağlayacağı temel avantaj, veri setindeki kayıpların nasıl bir örüntü sergileyeceğinin önceden bilinmesidir.

Kayıp Veriler Neden Sorun Oluşturur?

Kayıp verilerin yarattığı en temel sorun ulaşılabilen verinin, gerçekleştirilecek analizleri ve dolayısıyla bulgu ve sonuçları sınırlamasıdır. Pek çok araştırma için, veri toplamak kolay değildir. Dolayısıyla, veri kaybı istendik ve hatta bazı durumlarda telafi edilebilecek bir durum değildir. Ayrıca ilerleyen kısımlarda da