

Lojistik Regresyon Analizi ile Pisa Araştırmasında Öğrenci Başarısının Modellenmesi

Recep Bindak¹

Özet

Bu çalışmanın amacı Türkiye örneklemindeki öğrencilerin PISA sınavındaki başarılarını binary lojistik regresyon ile modellemektir. Açıklayıcı değişken olarak bazı sosyo-kültürel özellikler kullanılmış olup Bağımlı değişken iki kategorilidir ve öğrenci puanının OECD ortalaması üzerinde olup olmamasını belirtmektedir. Veriler PISA-2009 Türkiye örneklemine aittir. PISA 2009 Türkiye örneklemini, okul türlerine göre tabakalı rastgele yöntemle belirlenen toplam 170 okuldan 4996 öğrenciden oluşmaktadır. Katılımcıların "başarı" grubuna girme ihtimali üzerine cinsiyet, evde konuşulan dil, evde kitap sayısı, bölge, ebeveyn eğitim düzeyi, bilgisayara yönelik tutum, okula yönelik tutum ve varlık indeksinin etkilerini saptamak için lojistik regresyon analizi uygulanmıştır. Model istatistiksel olarak anlamlı bulunmuştur. Modelin açıklanabilen değişkenliği %23.8 (Nagelkerke R²), doğru sınıflandırma oranı %67,9'dir. Analiz sonuçlarına göre kızların başarılı grupta yer alması erkeklere göre 1,71 kat daha olasıdır. Evde konuştuğu dil Türkçe olan öğrencinin başarılı grupta yer alması diğerlerine göre 1,65 kat daha olasıdır. Evde bulunan kitap sayısının görece yüksek olması, ebeveyn eğitim düzeyi, bilgisayar tutumu ve varlık indeksi (WEALTH) başarılı grupta yer alma olasılığının yükselmesi ile ilişkili bulunmuştur. Okula yönelik tutum anlamlı bulunamadı.

Anahtar kelimeler

Lojistik regresyon • PISA araştırması • Öğrenci başarıları.

Jel Sınıflaması

C02 • C13 • I21

Modeling Students' Achievement in PISA Research with Logistic Regression Analysis

Abstract

The aim of this study is to the model the achievements of students in the Turkey sample on the PISA exam with binary logistic regression. Some socio-cultural variables were used as explanatory variables. The dependent variable is dichotomous and indicates whether the student score is above the OECD average. The data were from PISA 2009. The PISA 2009 Turkey sample consists of 4,996 students from 170 schools selected by stratified randomized sampling according to school types. Logistic regression analysis was performed to ascertain the effects of gender, home language, number of books at home, statistical region, parents' education level, attitude toward computer, attitude toward school, and family wealth index on the likelihood that participants belong to the "success" group. The model was statistically significant. The model explained 23.8% (Nagelkerke R²) of the variance in success and correctly classified 67.9% of cases. According to analysis, girls were 1.71 times more likely to exhibit success than boys. Students whose home language was Turkish were 1.65 times more likely to exhibit success than students with other home languages. Increase in the number of books at home, parents' educational level, attitude toward computer, and family wealth (WEALTH) index were associated with an increase in the likelihood of belonging to success group. Attitude toward school was not statistically significant.

Keywords

Logistic regression • PISA survey • Student achievement

Jel Classification

C02 • C13 • I21

1 Sorumlu Yazar: Recep Bindak (Dr. Öğr. Üyesi). Gaziantep Üniversitesi Teknik Bilimler Meslek Yüksek Okulu, Gaziantep, Türkiye. Eposta: bindak@gantep.edu.tr

Atf: Bindak, R. (2018). Lojistik Regresyon Analizi ile Pisa Araştırmasında öğrenci başarısının modellenmesi. *Ekonometri ve İstatistik e-Dergisi*, 14(28), 57-74. <https://dx.doi.org/10.26650/ekoist.2018.14.28.0010>

Extended Summary

Program for International Student Assessment (PISA) is an international research project organized by the OECD, in which the knowledge and skills of students are evaluated. This research, which is performed in countries covering approximately 90% of the world economy, is carried out every three years. The PISA survey assesses the extent to which 15-year-old students in OECD member and non-member countries have the basic knowledge and skills needed to be relevant in modern society. In the PISA survey, information pertaining to the reading skills and mathematical and science literacies of the students as well as their views about themselves, their motivation, families, and schools, are collected using measuring tools containing various types of questions. PISA is the most comprehensive and detailed international program that assesses students' performance and collects the data necessary to explain the differences in students' performances. The results of the PISA project significantly contribute to understanding both the international and domestic educational outcomes. Therefore, the in-depth examination of the data obtained from the PISA project has become an important research area; this data include records of the obtained results, and it is shared with the scientific community. The degrees of reliability and validity of PISA results are high, and the results significantly aids the understanding of educational outcomes.

The purpose of this study is to examine the relationship of students' achievements in mathematics, reading, and science with psychological, sociological, and cultural variables using the data from PISA 2009 Turkey sample. The data obtained from the questionnaires and tests administered to the students, within the Turkey sample, under the PISA 2009 program are used in this study. This data was obtained from the database at www.pisa2009.acer.edu.au, which is available to all researchers. PISA 2009 Turkey sample consists of 56 provinces from 12 units of statistical regions (NUTS) and 4,996 students from 170 schools, which were selected by stratified random sampling according to the type of school.

Method

The relationship between students' achievement in mathematics, reading, and science and the psychological, sociological, and cultural variables was examined by forming logistic regression models. Logistic regression analysis is a regression method that is used to develop classification and assignment processes.

Dependent variables:

- (i) 1 = group (1). At least one of the mathematics or reading or science score is equal to or above the OECD average;

- (ii) 0 = group (0). Mathematics or reading or science score is below the OECD average.
- (iii) The independent variables used in the study are as follows:
- (iv) X1 = Gender (1 = female; reference category = male)
- (v) X2 = Home language (1= Turkish; reference category = other languages)
- (vi) X3 = Statistical region (reference category = TRC)
- (vii) X4 = Number of books at home
- (viii) X5 = Levels of parents' education (total in years)
- (ix) X6 = Attitude toward computer
- (x) X7 = Attitude toward school
- (xi) X8 = Wealth index

The relationship between the dependent and independent variables was demonstrated by binary logistic regression analysis.

Results

The omnibus test statistics of the models were significant ($c2 = 849.01$; $p < 0.01$). In addition, the models were observed to have a good fit in terms of determination (Hosmer - Lemeshow test statistic was 10.247; $df = 8$; $p = 0.248$). Table 8 contains the coefficients, which show the relationship of all the variables in the analysis with dependent variables; the p-values of Wald test statistics, which determine whether standard errors of coefficients and coefficients are statistically significant; and the odds ratios.

According to the logistic regression analysis results, the X1, X2, X3, X4, X5, X6, and X8 variables had a significant effect in determining the "success" group. The gender variable (X1) was found to be associated with the success group, while the other variables were held constant. The possibility of the female students to be in the success group was 1.711 times higher than that for the male students. The home language variable (X2) was found to be associated with the success group. The possibility of the students who spoke Turkish at home to be in the success group was 1.654 times higher than that for students who spoke other languages. The statistical region variable (X3) was found to be associated with the success group. Additionally the independent variables such as the number of the books at home (X4), the parents' education levels (X5), the attitude toward computer (X6), and the wealth index (X8) were associated with the possibility of the student to be in the success group. However, the variable of attitude toward school (X7) did not significantly affect the mentioned dependent variable.

A correct classification rate of the models was created with a cut-off value of 0.5. The correct classification rate was obtained as 67.9%. Accordingly, using the independent variables in the model, whichever of the groups (pass or fail) the student will be assigned to in these areas could be estimated with an accuracy of 67.9%.

It is clear that PISA is quite an extensive application in terms of both the measured schools, environment, student characteristics, and the sample in which it is applied anywhere in the world. Therefore, it can be said that the output of the program depicts the educational outcomes of the participating countries.

Lojistik Regresyon Analizi ile Pisa Araştırmasında Öğrenci Başarısının Modellenmesi

Uluslararası Öğrenci Değerlendirme Programı [PISA], Ekonomik İşbirliği ve Kalkınma Teşkilatı [OECD] tarafından düzenlenen, öğrencilerin, matematik, fen ve okuma becerileri alanlarındaki bilgi ve becerilerinin değerlendirildiği uluslararası bir araştırma projesidir. Dünya ekonomisinin yaklaşık olarak %90'ını kapsayan ülkelerde yapılan bu araştırma üç yılda bir yapılmaktadır. PISA araştırmasında OECD üyesi ülkeler ve diğer katılımcı ülkelerdeki 15 yaş grubu öğrencilerin modern toplumda yerlerini alabilmeleri için gereken temel bilgi ve becerilere ne ölçüde sahip oldukları değerlendirilmektedir (MEB EARGED, 2010).

PISA araştırmasında değişik soru türleri bulunan ölçme araçları ile matematik okuryazarlığı, okuma becerileri ve fen okuryazarlığı konu alanları ile öğrencilerin kendileri hakkındaki görüşleri, motivasyonları, aileleri ve okulları hakkında bilgi toplanmaktadır. PISA, öğrenci performansını değerlendiren ve öğrenci performanslarının farklılığını açıklamak için gerekli verileri toplayan en kapsamlı ve en detaylı uluslararası programdır. Değerlendirmenin yapısı ve kapsamı ve anketlerle toplanacak bilgiyle ilgili kararlar, katılan ülkelerdeki uzmanlar tarafından verilmekte, çeviri, örneklem oluşturma ve veri toplama işlemlerinde uyulması zorunlu olan kalite güvence mekanizmaları uygulanmaktadır. Sonuç olarak, PISA sonuçlarının geçerlik ve güvenilirlik derecesi yüksektir ve bu sonuçlar eğitim çıktılarının anlaşılmasına önemli derecede katkı sağlamaktadır (MEB EARGED. 2007).

Tablo 1

Pisa Sonuçları Ortalama ve Standart Sapmaları ile Türkiye Örnekleminde Puanı OECD Ortalaması Üzerinde Olan Öğrencilerin Oranı

Alan	Türkiye ort.±st.sapma	OECD ort±s.sapma	OECD ortalaması üzerinde olanların oranı
Matematik	445±93	496±92	%26.9
Okuma	464±82	493±93	%37.4
Fen Bilimleri	454±81	501±94	%27.9

Kaynak: <http://stats.oecd.org/PISA2009Profiles> (23.05.2013)

PISA 2009 sonuçlarına göre matematik okuryazarlığı, okuma becerileri ve fen alanında en başarılı ülkeler sırasıyla 546 ortalama ile Kore, 556 ortalama ile Çin-Şanghay ve 554 ortalama ile Finlandiya olmuştur. Her bir alan için Türkiye örnekleminde puanı OECD ortalaması üzerinde olan öğrencilerin oranı Tablo 3'te verilmektedir (MEB EARGED, 2007). Türkiye örnekleminde en az bir puanı OECD ortalaması üzerinde olan öğrencilerin oranı %45,93 olduğu belirlenmiştir (bakınız Tablo 1).

Lojistik Regresyon

Lojistik regresyon, makine öğrenimi, veri madenciliği ve istatistikler de dahil olmak üzere birçok alanda sıklıkla kullanılan ve sağlam istatistiksel bir arka plana sahip çok popüler bir sınıflandırma algoritmasıdır. Lojistik regresyonun popülerliği basitlik ve

model parametrelerinin yorumlanabilirliği ile ilişkilendirilebilir (Ngufor ve Wojtusiak, 2016; Buzas, 2009). Lojistik regresyon analizi, sınıflama ve atama işlemi yapmaya yardımcı olan bir regresyon yöntemidir. Normal dağılım varsayımı, süreklilik varsayımı ön koşulları bulunmamaktadır. Lojistik regresyon ile bağımlı değişken üzerinde açıklayıcı değişkenlerin etkileri olasılık olarak belirlenerek risk faktörlerinin olasılık olarak belirlenmesi sağlanır. Lojistik regresyon; kategorik ve ikili (binary, dichotomous), üçlü ve çoklu kategorilerde gözlemlendiği durumlarda yanıt değişkeninin açıklayıcı değişkenlerle neden sonuç ilişkisini belirlemede yararlanılan bir yöntemdir (Özdamar, 1997).

Lojistik regresyon metodunun kullanım amacı herhangi bir model inşa tekniğinin amacı ile aynıdır. Bu da bir bağımlı değişken ile bir açıklayıcı değişkenler takımı arasındaki ilişkiyi tanımlayan en iyi matematiksel modeli bulmaktır. En yaygın bilinen model bağımlı değişkenin sürekli olduğu lineer regresyondur. Lojistik regresyon modelini lineer regresyon modelinden ayıran şey bağımlı değişkenin iki sonuçlu (veya kesikli) olmasıdır. Lojistik ve lineer regresyon arasındaki bu fark, hem parametrik model seçimini hem de varsayımları etkiler (Hosmer ve Lemeshow, 2000).

İki değerli (binary) y rastgele cevap değişkeni π olasılıklı bir Bernoulli dağılımına sahip ve p tane açıklayıcı değişkenin ikinci gözlem vektörü $\mathbf{X} = (x_{1i}, x_{2i}, \dots, x_{pi})$ olsun. X verilmişken Y nin koşullu ortalaması olan $E(Y | X) = \pi(x_i)$ denklem (1) ile tanımlanır:

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi})} \quad (1)$$

Denklem(1) e Lojistik regresyon fonksiyonu adı verilir. Burada β_i ler bilinmeyen parametrelerdir, ayrıca $E(y_i) = \pi(x_i)$ olduğundan x_i nin tüm değerleri için koşullu ortalama $[0, 1]$ aralığında kalır. Lojistik regresyon fonksiyonuna Lojit adı verilen dönüşüm uygulanırsa parametrelere göre doğrusal olan yeni bir değişken elde edilir:

$$g(x_i) = \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} \quad (2)$$

Bu durumda yeni değişken $g(x_i)$, (Lojit) modeldeki parametrelerle doğrusaldır, üstelik sürekli olup $(-\infty, +\infty)$ aralığında değerler alabilir. π ile $g(x)$ aynı yönde artar ve eğer $\pi(x) < 0.5$ ise $g(x)$ negatif, $\pi(x) > 0.5$ ise $g(x)$ pozitif değer alır. Böylece x açıklayıcı değişkeni sonuçta $\pi(x) < 0.5$ ise bir bağımlı değişkenin bir kategorisine ve eğer $\pi(x) > 0.5$ ise diğer kategorisine atanmış olur. Lojistik regresyon modelinde $Y = \pi(x) + \varepsilon$ hata terimi sadece iki farklı değer aldığından normal dağılıma uymaz. Bu nedenle parametreleri tahmin etmede en küçük kareler yerine iteratif bir metod olan maksimum olabilirlik tahmin metodu kullanılır (Hosmer ve Lemeshow, 2000; Beyaztaş ve Alın, 2014). Maksimum olabilirlik metodunda, açıklayıcı değişken de-

ğerleri ve $\beta_0, \beta_1, \dots, \beta_p$ verildiğinde gözlenen Y değerlerini elde etme olasılığını en yüksek yapan fonksiyonun (yani log-olabilirlik fonksiyonunun) değeri maksimize edilir (Menard, 2001).

Lojistik regresyonda parametre seçim kriteri Log Olabilirlik (log likelihood) ile yapılmaktadır. Paket programlarında -2 ile çarpılmış şekli verilmektedir (kısaca -2LL). -2LL değerinin yüksek olması bağımlı değişkenin kötü yordandığını belirtir. Modelde sadece sabit terim varken ki başlangıç -2LL değeri D_0 ve modelde tüm açıklayıcı değişkenler eklendikten sonraki -2LL değeri D_M olmak üzere $D_0 - D_M$ farkı Model Ki-Kare olarak adlandırılır (Omnibus test tablosunda yer alır). Lojistik regresyonda iki -2LL değeri arasındaki fark χ^2 dağılımına uyar ve modellerin birbirinden farklı olup olmadığını test etmede kullanılır. Testin sıfır hipotezi $H_0: \beta_1 = \beta_2 = \dots = \beta_p$ katsayılarının hepsinin aynı anda 0 olduğu biçiminde ifade edilmektedir. Dolayısıyla H_0 in ret edilmesi modele açıklayıcı değişken eklemenin modeli daha da iyileştireceğini belirtir.

Katsayıların bireysel olarak anlamlı olup olmadığı genellikle Wald istatistiği ile belirlenir. $W^2 = [\beta_k / se(\beta_k)]^2$, bu asimptotik olarak χ^2 dağılımına uyar. Veya karekök alınırsa standart normal dağılıma uyar. Cox & Snell R^2 ve Nagelkerke R^2 , bağımsız değişkenlerin modele dahil edilmesinin varyasyonu ne kadar azalttığının bir ölçüsüdür. Nagelkerke R^2 değeri Cox&Snell R^2 değerinin standart biçimi olup 0-1 arasında değer alır ve 1'e yakın R^2 değeri cevap değişkeni ile açıklayıcı değişkenler arasında yüksek ilişki olduğunu belirtmektedir (Menard, 2001).

Literatürde lojistik regresyon analizinin farklı disiplinlerde çok sayıda çalışmada kullanıldığını görmek mümkündür. Tıptan (Bircan, 2004; Coşkun ve ark., 2004; Sağıroğlu ve ark., 2014; Karaman ve arkadaşları, 2015), iktisada (Ege ve Bayrakdaroğlu, 2009; Yörük ve Dünder, 2012); davranış bilimlerinden (Girginer ve Cankuş, 2008; Aktaş, 2009; Tokmak ve Aksoy, 2016), mühendisliğe (Aktaş ve Erkuş, 2009; Takma, İşçi-Güneri ve Gevrekçi, 2016); spordan (Ergül, Yavuz ve Yavuz, 2014) eğitim bilimlerine (Güneri ve Apaydın, 2004; Çokluk, 2010; Alkharusi, 2012) daha çok sınıflama amacıyla lojistik regresyonun kullanıldığı görülmektedir.

Türkiye örnekleminin PISA sınavlarındaki başarılarını çeşitli değişkenler açısından inceleyen çalışmalara rastlamak mümkündür. Örneğin öğrencilerin puanlarını etkileyen faktörler (Anıl, 2009; Alacacı ve Erbas, 2010; Acar, 2012; Aydın, Sarier ve Uysal, 2012; Gürsakal, 2012; Yılmaz ve Aztekin, 2012), okuma becerileri alanında başarılı-başarısız okulları ayırt eden okul içi etmenleri ve sosyo-ekonomik faktörler (Güzle-Kayır, 2012), (Özer-Özkan ve Doğan, 2013), öğrencilerin duyuşsal özelliklerini sınıflandıran değişkenlerin belirlenmesi (Tatlı, Ergin ve Demir, 2016), matematik başarısını etkileyen sosyo-demografik özellikler (Gürsakal, Murat ve Gürsakal, 2016) incelenmiştir. Ayrıca Yurtdışında yapılmış bazı çalışmalarda araştırmacılar kendi ulusal verilerini değerlendirmiş ve ülke içi bölge farklılıklarını ele almıştır (ör-

neğin Bradshaw ve ark., 2007; Cordero Ferrera ve ark., 2011; Gamboa ve Waltenberg, 2012; Sun ve Bradley, 2012; Yuan-Hsuan ve Jiun-Yu, 2012; Cornoldi, Giofrè ve Martini, 2013).

Bu çalışmanın temel amacı lojistik regresyon analizi ile PISA sınavına ait Türkiye örnekleme verilerinin derinlemesine incelenmesidir. PISA projesinin sonuçları gerek ülkeler arası gerek ülke içi eğitim çıktılarının anlaşılmasına önemli derecede katkı sağlamaktadır. Bu nedenle PISA projesinde ortaya çıkan verilerin derinlemesine incelenmesi, elde edilen sonuçların ortaya konulması ve bilim dünyası ile paylaşılması oldukça önemli bir araştırma alanı olarak karşımıza çıkmaktadır. PISA projesinin bu önemi dikkate alınarak bu çalışmada 2009 Türkiye örnekleminde elde edilen verilere göre öğrencilerin performansları, sosyo-ekonomik değişkenlere göre analiz edilmiştir. Sonuçta lojistik regresyon modeli ile “başarılı” öğrencilerden oluşan grubun sosyo - kültürel özelliklerinin ortaya konulması hedeflenmiştir. Bu yönüyle çalışmanın literatüre bir katkı sağlaması hedeflenmiştir.

Materyal ve Metot

PISA projesinde elde edilen veriler üzerinde yapılmış çalışmalar bir bütün olarak incelendiğinde “başarı” için standart bir ölçütün olmadığı, daha çok ulusal ortalama puan, ülke sıralaması, veya en başarılı ülkeye göre durumun başarı ölçütü olarak alındığı görülmektedir. Bu çalışmada başarı için farklı bir ölçüt ele alınmış ve Türkiye örnekleme için başarı ölçütü olarak en az bir puanı OECD ortalaması üzerinde olanlar “başarılı” grupta kabul edilmiştir.

Çalışmada Türkiye örnekleminde yer alan öğrencilere PISA 2009 programı kapsamında uygulanan anket ve testlerden elde edilen veriler kullanılmıştır. Bu veriler tüm araştırmacıların kullanımına açık olan veritabanından (www.pisa2009.acer.edu.au) elde edilmiştir. PISA2009 Türkiye örnekleme, 12 istatistiki bölge biriminden (NUTS) 56 il ve okul türlerine göre tabakalı rastgele yöntemle belirlenen toplam 170 okuldan 4996 öğrenciden oluşmuştur.

Öğrenci başarısının açıklayıcı değişkenlerle ilişkisi için lojistik model oluşturulurken öğrenciler “başarılı grup” ve “başarısız grup” olarak iki kategoriye ayrılmıştır. İki kategoriye ayırmada OECD ortalama puanları esas alınmış, eğer öğrencinin matematik, okuma ve fen puanlarından en az biri OECD ortalamasına eşit veya üzerinde ise öğrenci “başarılı” gruba, aksi halde “başarısız” gruba atanmıştır. Bu şekilde oluşturulan iki grup lojistik regresyon analizinde bağımlı değişken olarak kullanılmıştır. Çalışmada sosyo-kültürel yapı ile ilgili açıklayıcı değişkenler ele alınmıştır. Bağımlı değişken kodları ile açıklayıcı değişkenler aşağıdaki şekilde tanımlanmıştır:

Bağımlı Değişken: 1 = grup(1) Matematik/Okuma /Fen puanlarından en az biri OECD ortalamasına eşit veya üstünde, 0 = grup(0) Matematik/Okuma /Fen puanı

OECD ortalaması altında

Çalışmada kullanılan açıklayıcı değişkenler şunlardır:

- (i) X1 = Öğrencinin cinsiyeti. (1=Kız, Referans kategorisi = Erkek)
- (ii) X2 = Evde konuşulan dil (1= Türkçe, referans kategorisi=başka dil)
- (iii) X3 = İstatistiki Bölge. Türkiye İstatistik Kurumu tarafından kullanılan bu sınıflama İstatistiki Bölge Birimleri Sınıflaması Düzey 1 olarak adlandırılmaktadır. Bu sınıflamaya göre Türkiye 12 bölgeye ayrılmıştır. TR1 İstanbul, TR2 Batı Marmara, TR3 Ege, TR4 Doğu Marmara, TR5 Batı Anadolu, TR6 Akdeniz, TR7 Orta Anadolu, TR8 Batı Karadeniz, TR9 Doğu Karadeniz, TRA Kuzeydoğu Anadolu, TRB Ortadoğu Anadolu ve TRC Güneydoğu Anadolu biçiminde tanımlanmıştır. Bu çalışmada Türkiye'nin 12 istatistiki bölgesi kukla değişken olarak kodlanmıştır (referans kategorisi= TRC)
- (iv) X4 = Evde kitap sayısı, kişinin beyanına göre evdeki kitap sayısını göstermektedir.
- (v) X5 = Anne baba eğitim durumu, anne ile babanın eğitim durumlarının (yıl olarak en son bitirdikleri okul) toplamı alınmıştır.
- (vi) X6 = Bilgisayara yönelik tutum; öğrencilerin bilgisayara yönelik tutumları 4 maddeden oluşan 4'lü Likert tipinde bir ölçek ile belirlenmiştir. Teorik puan ranjı 1-4 olup yüksek puan olumlu tutumu ifade etmektedir.
- (vii) X7 = Okul tutumu; öğrencilerin okula karşı tutumları okul ile ilgili anket maddelerine verdikleri cevapların ortalaması 0 varyansı 1 olan puanlara dönüştürülmesi ile belirlenmektedir. Yüksek puan okula yönelik olumlu tutumu göstermektedir.
- (viii) X8 = Aile Varlık indeksi (WEALTH index); öğrencinin ailesinin veya kendisinin sahip olduğu varlıklardan hesaplanmış ve standartlaştırılmış bir indekstir (örneğin kendine ait oda, internet, bulaşık makinesi, klima, video kamera, ile evde sahip olunan cep telefonu, tv, bilgisayar araba ve banyo sayısı).

Verilerin analizi

PISA anketlerinde öğrencinin sosyo-kültürel yapısını ifade edebilecek çok sayıda madde bulunmaktadır. Bunlardan birçoğu birleştirilerek bir indeks ile ifade edilmiştir (Bu indekslerin hesaplanması için bakınız: (OECD, 2013). Bu indekslerden çoğu birbirleri ile yüksek ilişki gösterdiğinden bu çalışmada yukarıda kısaca açıklanan aile varlık indeksi (X8) ve okul tutumu (X7) açıklayıcı değişken olarak kullanılmıştır.

Açıklayıcı değişken değerleri herhangi bir dönüşüme tabi tutulmamış, orijinal değerler olduğu gibi kullanılmıştır. Yukarıda tanımlandığı gibi açıklayıcı değişkenlerden X1, X2 ve X3 değişkenleri kategorik değişken olarak işlem görmüş; X4, X5, X6, X7 ve X8 değişkenleri ise sayısal değişken olarak işlem görmüştür. Tablo-2’de kategorik değişkenlerin bağımlı değişken gruplarına göre frekansları ile Ki-kare bağımsızlık test sonuçları verilmiştir. Tüm kategorik değişkenler için test sonucu anlamlıdır ($p < 0.01$). Dolayısıyla bütün kategorik değişkenlerin lojistik model için birer aday açıklayıcı değişken olabilecekleri değerlendirilmiştir.

Tablo 2
Kategorik Açıklayıcı Değişkenlerin Bağımlı Değişken Kategorilerine Göre Frekansları

	n	Grup(0)	Grup(1)	Ki-Kare
<i>Cinsiyet</i>				
Kız	2432	%49.8	%50.2	77.48**
Erkek	2534	%62.2	%37.8	
<i>Evde Konuşulan Dil</i>				
Türkçe	4740	%55.0	%45.0	56.69**
Diğer	206	%81.6	%18.4	
<i>Bölge</i>				
TR1	798	%56.0	%44.0	133.50**
TR2	244	%52.0	%48.0	
TR3	618	%53.7	%46.3	
TR4	523	%56.0	%44.0	
TR5	478	%47.7	%52.3	
TR6	631	%54.4	%45.6	
TR7	295	%49.5	%50.5	
TR8	474	%55.6	%44.4	
TR9	212	%46.2	%53.8	
TRA	141	%52.5	%47.5	
TRB	217	%73.3	%26.7	
TRC	435	%76.8	%23.2	

** : $p < 0.01$

Sürekli değişken değerleri içinde uçdeğer olup olmadığını belirlemek için sadece sürekli değişkenlerin yer aldığı bir regresyon modeli oluşturulmuş ve Mahalanobis uzaklıkları incelenmiştir. Mahalanobis uzaklık değerleri için $p < 0.001$ olan toplam 54 verinin uçdeğer olduğuna karar verilmiş ve sonraki analizlerde bunlar hariç tutulmuştur. Açıklayıcı değişkenlerin birbirileri ile korelasyonlarına bakılmış (Tablo 4), değişkenlerin birbirileri ile çok yüksek düzeyde ilişkili olmadıkları görülmüş ve çoklu bağlantı probleminin olmadığı sonucuna ulaşılmıştır. Araştırmanın başında açıklayıcı değişkenler ile bağımlı değişken arasında var olan herhangi teorik bir model önerisi bulunmadığından lojistik regresyon analizi Enter metodu kullanılarak gerçekleştirilmiştir. İstatistiksel analizler SPSS programı ile yapılmıştır. İstatistiksel anlamlılık düzeyi 0.05 olarak alınmıştır. Analiz sonucunda tahmin edilen katsayıların istatistiksel olarak anlamlı olup olmadıklarına Wald istatistiğine bakılarak karar verilmiş, eğer katsayılar istatistiksel olarak anlamlı ise Odds [Exp(B)] oranlarına bakılarak yorum yapılmıştır.

Bulgular ve Tartışma

Araştırmada kullanılan değişkenlere ait betimsel istatistikler Tablo-2 ve Tablo-3'te sunulmuştur. Açıklayıcı değişkenlerden sürekli olanların bağımlı değişken kategorilerine göre ortalama ve standart sapmaları Tablo -3'te özetlenmiştir.

Evdeki kitap sayısı (X4) değişkenine ait veriler sınıflandırılmış seri biçimindedir. Sınıf frekansları ve sınıf orta noktaları kullanılarak seri sürekli değişkene dönüştürülmüştür. X4 ile birlikte diğer sürekli değişkenlerin lojistik regresyon modeli için birer aday açıklayıcı değişken olup olmadıklarına karar vermek amacıyla söz konusu değişkenlerin ortalamaları gruplara göre karşılaştırılmış, sonuçlar Tablo-3'te verilmiştir. T-test sonuçlarına göre tüm sürekli değişken ortalamalarının gruplara göre farklılık gösterdiği anlaşılmış ve tümünün analize alınmasına karar verilmiştir. Değişkenlerin birbirleri ile korelasyonları Tablo-4'te verilmiştir. Tablo 4'ten, açıklayıcı değişkenler arasındaki en yüksek ilişkinin baba eğitim düzeyi ile ekonomik sosyal kültürel statü indeksi arasında ve en düşük ilişkinin baba eğitim durumu ile kültürel varlık indeksi arasında olduğu görülmektedir.

Tablo 3

Nümerik Açıklayıcı Değişkenlerin Bağımlı Değişken Kategorilerine Göre Betimsel İstatistikleri ()

Değişken	n	X ± Ss	t-test
<i>X4, Evde Kitap sayısı^(a)</i>			
Grup(0)	2741	53.22±93.16	17.840**
Grup(1)	2174	114.3±145.7	
<i>X5, Annebaba Eğitim durumu</i>			
Grup(0)	2789	4.206±3.626	14.905**
Grup(1)	2177	5.742±3.574	
<i>X6, Bilgisayar Tutum</i>			
Grup(0)	2603	2.813±0.819	10.814**
Grup(1)	2123	3.061±0.735	
<i>X7, Okul Tutum</i>			
Grup(0)	2596	0.257±1.021	3.495**
Grup(1)	2102	0.150±1.067	
<i>X8, Aile Varlık İndeksi</i>			
Grup(0)	2781	-1.334±1.212	21.401**
Grup(1)	2177	-0.611±1.139	

** : $p < 0.01$ ^(a): sınıflandırılmış seri sayısal değerlere dönüştürüldü

grup(0): puanı OECD ortalamasının altında olan öğrenciler; grup(1): puanı OECD ortalaması üzerinde olan öğrenciler

Tablo 4

Modele sayısal değişken olarak alınan açıklayıcı değişkenlerin birbirleri ile korelasyonları

	Ebeveyn Eğitim d.	Bilgisayar tutumu	Okul tutumu	Kültürel Varlık indeksi
Evde kitap sayısı	.316**	.136**	-.071**	.445**
Ebeveyn eğitim d.		.084**	-.106**	.325**
Bilgisayar tutumu.			-.038*	.272**
Okul tutumu				-.144**

** : $p < 0.01$

Tablo 5

Başlangıç Modeli için Sınıflandırma Tablosu

		Tahmin edilen		Doğru sınıflandırma
		Grup (0)	Grup (1)	
Gözlenen	Grup(0)	2314	0	%100
	Grup(1)	2017	0	%0
Toplam				%53.4

Tablo 6

Model Özeti ve Modelin Açıklanabilen Değişkenliği

Step	-2 LL	Cox & Snell R ²	Nagelkerke R ²
1	5134.647	.178	.238

Başlangıç modelinde sadece sabit terim bulunmakta, dolayısıyla sadece bir kategoriye göre doğru sınıflama olmaktadır (Tablo 5). Modelde olmayan açıklayıcı değişkenlerin modele eklenmesi durumunda tahmin gücünü istatistiksel açıdan anlamlı derecede artıracığı genel istatistik skorundan (Overall statistic $\chi^2=775,95$; $sd=18$; $p < 0.01$) anlaşılmıştır.

Analiz sonuçları incelendiğinde Omnibus Ki-kare istatistiğinin model için anlamlı ($\chi^2 = 849.01$ $p < 0.01$) olduğu anlaşılmaktadır. Omnibus Ki-kare testi için sıfır hipotez “ H_0 : Lojistik regresyon katsayılarının hepsi aynı anda sıfıra eşittir” şeklinde ifade edildiğinden modele ilave edilen değişkenlerin modele katkılarının önemli olduğu anlaşılmaktadır. Ayrıca lojistik regresyon modelinin bir bütün olarak uyumunu değerlendiren Hosmer Lemeshow Ki-kare uyum iyiliği testi sonuçlarına göre “ H_0 : parametreler belirleyicilik açısından iyi bir ayrımcılığa sahiptir” şeklindeki sıfır hipotezi kabul edilmiştir ($\chi^2 = 10.247$, $sd = 8$ ve $p = 0.248$). Bu sonuçlara göre elde edilen modelin verilere uygun olduğu söylenebilir.

Modelde yer alan her bir açıklayıcı değişkenle ilgili lojistik regresyon katsayısı, standart hata, Wald istatistiği, serbestlik derecesi, Wald istatistiğinin anlamlılık düzeyini gösteren p değeri, Odds oranı ile Odds oranı için %95 güven aralığı Tablo 8’da sunulmuştur.

Wald istatistiğine göre öğrencinin başarılı gruba atanmasında X7 dışındaki tüm değişkenler anlamlı katkı sağlamaktadır. Bir başka deyişle Lojistik regresyon analizi sonuçlarına göre başarı gruplarını belirlemede X1, X2, X3, X4, X5, X6 ve X8 değişkenleri anlamlı etkiye sahiptir.

Tablo 7
Hosmer-Lemeshow Testi için Kontenjans Tablosu

Grup	Y = 0		Y = 1	
	Gözlenen	Beklenen	Gözlenen	Beklenen
1	378	372.44	55	60.56
2	325	333.54	108	99.46
3	295	306.13	138	126.87
4	269	278.90	164	154.10
5	271	252.53	162	180.47
6	218	223.82	215	209.18
7	201	193.79	232	239.21
8	168	161.70	265	271.30
9	129	122.20	304	310.80
10	60	68.93	374	365.07

$\chi^2 = 10.247$; sd = 8; $p = 0.248$

Tablo 8
Lojistik Regresyon Modeli için Bulunan Katsayılar, Katsayıların Standart Hataları, Wald İstatistikleri, Odds Oranları ve Odds Oranları İçin %95 Güven Aralıkları

	B	S.hata	Wald	sd	p	Odds	Exp(B) için %95 G.A.	
							alt	üst
Sabit	-3.368	.308	119.97	1	.000	.034		
X1.Cinsiyet(1=kız)	.537	.068	61.56	1	.000	1.711	1.496	1.957
X2.Evde k. Dil (1=Tür.)	.503	.247	4.16	1	.041	1.654	1.020	2.683
X3.Bölge (Referans=TRC)			81.46	11	.000			
TR1	.178	.164	1.18	1	.277	1.195	.866	1.648
TR2	.608	.202	9.09	1	.003	1.837	1.237	2.729
TR3	.365	.170	4.61	1	.032	1.441	1.032	2.012
TR4	.329	.174	3.57	1	.059	1.389	.988	1.954
TR5	.664	.175	14.34	1	.000	1.942	1.378	2.739
TR6	.722	.169	18.35	1	.000	2.059	1.479	2.865
TR7	.850	.195	19.10	1	.000	2.340	1.598	3.426
TR8	.538	.186	8.33	1	.004	1.713	1.189	2.468
TR9	1.191	.212	31.52	1	.000	3.290	2.171	4.986
TRA	1.003	.242	17.18	1	.000	2.727	1.697	4.383
TRB	-.078	.228	.12	1	.734	.925	.591	1.447
X4.Evde Kitap Sayısı	.376	.031	147.05	1	.000	1.457	1.371	1.548
X5.Anebababa Eğitim d.	.108	.012	76.21	1	.000	1.114	1.087	1.141
X6.Bilgisayar Tutum	.239	.046	27.32	1	.000	1.270	1.161	1.388
X7.Okul Tutum	-.023	.033	.47	1	.493	.978	.916	1.043
X8.Aile Varlık İndeksi	.208	.036	33.787	1	.000	1.231	1.148	1.321

Bağımlı değişken; (Y = 0; tüm puanları OECD ortalaması altı, Y = 1; en az bir puanı OECD ortalaması üstü)

Açıklayıcı değişkenlerden X7 (Okul tutumu) dışındaki değişkenlerin modele katkıları istatistiksel olarak anlamlıdır. X7 değişkeninin modele katkısı ihmal edilebilecek kadar küçüktür ($p > 0.05$). Lojistik regresyon katsayıları ve odds oranları göz önüne alındığında her bir açıklayıcı değişken için aşağıdaki sonuçlar ifade edilebilir:

Cinsiyet değişkeni için elde edilen odds oranına göre, diğer tüm değişkenler sabit iken kız öğrencinin başarılı grupta yer alması olasılığı erkeklere göre 1.711 kat daha yüksektir.

Tablo 9
Modelin Doğru Sınıflandırma Oranı

		Tahmin edilen		Doğru sınıflandırma ^a
		Grup (0)	Grup (1)	
Gözlenen	Grup(0)	1708	606	%73.8
	Grup(1)	784	1233	%61.1
Toplam				%67.9

^a: Kesme değeri 0.500 alınmıştır.

Evde konuşulan dil de başarılı gruba atamada istatistiksel olarak anlamlı etkiye sahiptir. Evde test dili (yani Türkçe) konuşan bir öğrencinin başarılı grupta yer alması evde başka bir dil konuşan öğrenciye göre 1.645 kat daha yüksektir.

Evde bulunan kitap sayısı sıralama değişkeni olduğu göz önüne alınarak modele eklenmiştir. Modelde pozitif ve anlamlı katsayıya sahiptir ($B = 0.376$; $Wald = 147.05$; $p < 0.001$). Evde bulunan kitap sayısına göre 1 kategori üst sırada olmak başarılı gruba atanmayı 1.457 kat arttırmaktadır.

Bölge değişkeni lojistik modele istatistiksel olarak anlamlı katkı sağlamaktadır ($Wald = 81.460$; $sd = 11$; $p < 0.01$). Bölge değişkeni göz önüne alındığında referans kategorisi olan TRC bölgesi başarı yönünden TR1 ve TRB ile benzer bulunmuştur. Diğer bölgelerde yaşayan öğrencilerin başarılı grupta yer alma olasılığı TRC bölgesinde yaşayan öğrencilere göre anlamlı derecede daha yüksek olduğu belirlenmiştir. Söz gelimi diğer bütün değişkenler bakımından aynı olan iki öğrenciden TR9 bölgesindeki öğrencinin başarılı gruba atanması olasılığı TRC bölgesindeki öğrenciye göre 3.29 kat daha yüksektir.

Anne-baba eğitim seviyesi değişkeni başarılı gruba atanmada istatistiksel olarak anlamlı etkiye sahiptir ($B = 0.108$; $Wald = 76.21$; $p < 0.001$). Anne-baba eğitim seviyesinin 1 puan yüksek olması başarılı grupta atanmada 1.114 kat etkiye sahiptir.

Okula yönelik tutum ile öğrencinin başarı gruplarından birine atanması arasında anlamlı ilişki bulunamamıştır ($B = -0.23$; $Wald = 0.470$; $p > 0.05$).

Aile varlık indeks puanı öğrencinin başarılı gruba atanmasında anlamlı etkiye sahiptir ($B = 0.208$; $Wald = 33.787$; $p < 0.001$). Varlık indeks puanının 1 puan yüksek olması öğrencinin başarılı grupta yer alma olasılığını 1.231 kat yükseltmektedir.

Modellerin doğru sınıflandırma oranı kesme değeri 0.5 alınarak oluşturulmuştur. Tablo 9'dan modelin doğru sınıflandırma oranının %67,9 olduğu görülmektedir. Buna göre modele alınan açıklayıcı değişkenler kullanılarak öğrencinin söz konusu alanlarında başarılı/başarısız gruplarından hangisine atanacağı %67,9 doğrulukla tahmin edilebilir. Bu değer şanstaki kaynaklı başarı olan %50,23'ün $((2314/4331)^2 + (2017/4331)^2 = .2854 + .2169 = .5023)$ 1.25 katı olan %62.79'dan ($\%50.23 \cdot 1.25 = .6279$) daha büyüktür. Dolayısıyla modelin başarılı bir sınıflama yaptığı söylenebilir.

Sonuç ve Öneriler

Analiz sonucu elde edilen sonuçlara göre Öğrencinin başarılı grupta yer alması, cinsiyet (kız), evde konuşulan dil (Türkçe), bölge, evde bulunan kitap sayısı, ebeveyn eğitim düzeyi, bilgisayar tutumu ve aile varlık indeksi değişkenlerinin başarılı-başarısız grup kategorilerini anlamlı biçimde yordadığı görülmüştür (Tablo 6). Ayrıca okula yönelik tutum değişkeninin başarılı-başarısız grup kategorileri ile anlamlı ilişki göstermedikleri ortaya çıkmıştır. Analiz sonucu elde edilen modele göre öğrencinin başarılı grupta yer alması, cinsiyet (kız), evde konuşulan dil (Türkçe), bölge, evde bulunan kitap sayısı, ebeveyn eğitim düzeyi, bilgisayar tutumu ve aile varlık indeksi değişkenlerinin başarılı-başarısız grup kategorilerini anlamlı biçimde yordadığı görülmüştür (Tablo 6). Ayrıca okula yönelik tutum değişkeninin grup kategorilerine atamada anlamlı etkiye sahip olmadığı ortaya çıkmıştır.

Literatürde bu sonuçları destekleyen bağımsız araştırmaların bulgularına rastlamak mümkündür. Sözelimi, cinsiyet değişkeni söz konusu olduğunda genellikle matematikte erkekler lehine ve okumada kızlar lehine anlamlı farklılar elde edildiği rapor edilmiştir (Gürsakal, 2012; Özer-Özkan ve Doğan, 2013; Kılıç, Çene ve Demir, 2012; Stoeet ve Geary, 2013). Bununla birlikte ebeveyn eğitim durumunun görece yüksek olması (Akyüz ve Pala, 2010; Gamboa ve Waltenberg, 2012; Gürsakal, Murat ve Gürsakal, 2016); refah düzeyi yüksek çevrenin (Alacacı ve Erbaş, 2010; Güzle-Kayır, 2012; Anıl, 2011; Demir ve Kılıç, 2010); sosyo-ekonomik düzeyin yüksekliğinin (Yıldırım, 2012; Aydın, Erdağ ve Taş, 2011; Gürsakal, Murat ve Gürsakal, 2016) başarı ile pozitif ilişkili olduğu rapor edilmiştir. Ayrıca test dili ile evde konuşulan dilin etkisi farklı olması başarı için bir dezavantaj olduğu (Sökmen, 2016) ve okula yönelik tutumun başarı ile ilişkili olmadığına dair literatürde (Lee, 2016) bulgular bulunmaktadır.

Uyum istatistikleri, doğru sınıflandırma oranı ve model kullanılarak yapılan sınıflandırmanın yapılmış bağımsız çalışmaların bulguları ile örtüşmesi, bu çalışmada elde edilen lojistik modelin başarılı olduğunu ve ekolojik olarak geçerli olduğunu işaret etmektedir.

Çalışma 2009 yılı PISA sınavına dayanmaktadır. Bu da bir sınırlılık olup çalışmanın zayıf tarafını oluşturmaktadır. Çünkü bu sınavdan sonra PISA tarafından iki sınav daha yapılmıştır. 2012 yılı sınavı rapor ve sonuçları açıklanmış, 2015 yılında yapılan sınavın henüz tümü olmasa da, belli raporları nisan ayında açıklanmıştır. Bununla birlikte bu çalışma, çıkarsamalı istatistikte önemli bir yeri olan binary lojistik regresyon metodunun gerçek verilere bir uygulamasını ele aldığından literatüre katkı sağladığı söylenebilir.

Kaynakça/References

- Acar, T. (2012). 2009 yılı uluslararası öğrenci başarılarını değerlendirme programında Türk öğrencilerin başarılarını etkileyen faktörler. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(2), 309–314.
- Aktaş, C. (2009). Lojistik regresyon analizi: öğrencilerin sigara içme alışkanlığı üzerine bir uygulama. *Erciyes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 1(26), 107–122.
- Aktaş, C. ve Erkuş, O. (2009). Lojistik regresyon analizi ile Eskişehir'in sis kestiriminin incelenmesi. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 8(16), 47–59.
- Akyüz, G. ve Pala, N. M. (2010). Pisa 2003 sonuçlarına göre öğrenci ve sınıf özelliklerinin matematik okuryazarlığına ve problem çözme becerilerine etkisi. *İlköğretim Online*, 9(2), 668–678. <http://ilkogretim-online.org.tr> adresinden edinilmiştir.
- Alacacı, C. & Erbas, A. K. (2010). Unpacking the inequality among Turkish schools: Findings from PISA 2006. *International Journal of Educational Development*, 30, 182–192.
- Alkharusi, H. (2012). A logistic regression model predicting assessment literacy among in-service teachers. *Eğitimde Kuram ve Uygulama*, 7(2), 280–291
- Anıl, D. (2009). Uluslararası öğrenci başarılarını değerlendirme programında (Pisa) Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler. *Eğitim ve Bilim*, 34(152), 87–100.
- Anıl, D. (2011). Türkiye'nin Pisa 2006 Fen Bilimleri başarısını etkileyen faktörlerin yapısal eşitlik modeli ile incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(3), 1253–1266.
- Aydın, A., Erdağ, C. ve Taş, N. (2011). 2003-2006 Pisa okuma becerileri sonuçlarının karşılaştırılması olarak değerlendirilmesi: En başarılı beş ülke ve Türkiye. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(2), 651–673.
- Aydın, A., Sarier, Y., & Uysal, Ş. (2012). Sosyoekonomik ve sosyokültürel değişkenler açısından PISA matematik sonuçlarının karşılaştırılması. *Eğitim ve Bilim*, 37(164), 20–30.
- Beyaztaş, U., & Alın, A. (2014). Jackknife-after-Bootstrap as logistic regression diagnostic tool. *Communications in Statistics Simulation and Computation*, 43, 2047–2060.
- Bircan, H. (2004). Lojistik regresyon analizi: Tıp verileri üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2, 185-208.
- Bradshaw, J., Sturman, L., Vappula, H., Ager, R. & Wheeler, R. (2007). *Achievement of 15-year-olds in England: PISA 2006 National Report*. Slough: NFER. Retrieved from: <http://www.nfer.ac.uk/>
- Buzas, J. S. (2009). A note on corrected scores for logistic regression. *Statistics and Probability Letters*, 79, 2351–2358.
- Cordero Ferrera, J. M., Cebada, E. C., Chaparro, F. P. & González, D. S. (2011). Exploring educational efficiency divergences across Spanish regions in Pisa 2006. *Revista de Economía Aplicada*, XIX(57), 117–145.
- Cornoldi, C., Giofrè, D. & Martini, A. (2013). Problems in deriving Italian regional differences in intelligence from 2009 PISA data. *Intelligence*, 41, 25–33.
- Coşkun, S., Kartal, M., Coşkun, A. ve Bircan, H. (2004). Lojistik regresyon analizinin incelenmesi ve dış hekimliğinde bir uygulaması. *Cumhuriyet Üniversitesi Dış Hekimliği Fakültesi Dergisi*, 7(1), 41–50.
- Çokluk, Ö. (2010). Lojistik regresyon analizi Kavram ve uygulama. *Kuram ve Uygulamada Eğitim Bilimleri*, 10(3), 1357–1407.

- Demir, İ., & Kılıç, S. (2010). Using PISA 2003, examining the factors affecting students' mathematics achievement. *Hacettepe University Journal of Education*, 38, 44–54.
- Ege, İ., & Bayrakdaroğlu, A. (2009). İMKB şirketlerinin hisse senedi getiri başarılarının lojistik regresyon tekniği ile analizi. *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 5(10), 139–158.
- Ergül, B., Yavuz, A. A. & Yavuz, H. S. (2014). Classification of NBA league teams using discriminant and logistic regression analyses. *Pamukkale Journal of Sport Sciences*, 5(1), 48–60.
- Gamboa, L. F. & Waltenberg, F. D. (2012). Inequality of opportunity for educational achievement in Latin America: Evidence from PISA 2006–2009. *Economics of Education Review*, 31, 694–708.
- Girginer, B. & Cankuş, B. (2008). Tramvay yolcu memnuniyetinin lojistik regresyon analiziyle ölçülmesi Estram örneği. *CBÜ İİBF Yönetim ve Ekonomi Dergisi*, 15(1), 181–193.
- Güneri, N. & Apaydın, A. (2004). Öğrenci başarılarının sınıflandırılmasında lojistik regresyon analizi ve yapay sinir ağları yaklaşımı. *Gazi Üniversitesi Ticaret ve Turizm Fakültesi Dergisi*, 1, 170–188.
- Gürsakal, S. (2012). PISA 2009 Öğrenci başarı düzeylerini etkileyen faktörlerin değerlendirilmesi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 17(1), 441–452.
- Gürsakal, S., Murat, D. & Gürsakal, N. (2016). Assessment of Pisa 2012 results with quantile regression analysis within the context of inequality in educational opportunity. *Alphanumeric Journal*, 4(2), 41–54
- Güzle Kayır, Ç. (2012). Pisa 2009-Türkiye verilerine dayanarak okuma becerileri alanında başarılı okullar ile başarısız okulları ayırt eden okul içi etmenler ve sosyo-ekonomik faktörler. (Yüksek lisans tezi, Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Antalya). <https://tez.yok.gov.tr/UlusalTezMerkezi/> adresinden edinilmiştir.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (Second edition). New York: John Wiley & Sons, Inc.
- Karaman, E., Turtay, M., Çolak, C., Oğuztürk, H., Gürbüz, Ş., Kamışlı, S. ve Çelik, E. (2015). İşlem risk faktörleri ve etkileri. *Turgut Özal Tıp Merkezi Dergisi*, 22(4), 225–230
- Kılıç, S., Çene, E. ve Demir, İ. (2012). Türkiye'deki matematik başarısının öğrenme stratejileri açısından 8 ülkeyle karşılaştırılması. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(4), 2585–2598
- Lee, J. (2016). Attitude toward school does not predict academic achievement. *Learning and Individual Differences*, 52, 1–9.
- Lee, Y. H. & Wu, J.-Y. (2012). The effect of individual differences in the inner and outer states of ICT on engagement in online reading activities and Pisa 2009 reading literacy: Exploring the relationship between the old and new reading literacy. *Learning and Individual Differences*, 22, 336–342. Retrieved from: <https://ir.nctu.edu.tw/>
- Menard, S. (2001). *Applied logistic regression analysis* (Second edition). Thousand Oaks: Sage, CA.
- Milli Eğitim Bakanlığı. (2007). *Pisa 2006 ulusal ön raporu*. Ankara: Milli Eğitim Bakanlığı.
- Milli Eğitim Bakanlığı. (2010). *Pisa 2009 ulusal ön raporu*. Ankara, Milli Eğitim Bakanlığı.
- Ngufor, C. & Wojtusiak, J. (2016). Extreme logistic regression. *Advances in Data Analysis and Classification*, 10, 27–52.
- OECD. (2012). *PISA 2009 technical report*. OECD Publishing. Retrieved from: <https://www.oecd.org/pisa/>
- Özdamar, K. (1997). *Paket programlar ile istatistiksel veri analizi I*. Eskişehir: Kaan Kitabevi
- Özer Özkan, Y. ve Doğan, B. (2013). İlköğretim 8. sınıf öğrencilerinin okuma becerilerinin kestirilmesinde etkili olan değişkenlerin belirlenmesi. *International Journal of Social Science*, 6(4), 667–680.

- Sađırođlu, Ő., olak, C., olak, M., Atıcı, M. ve Alasulu, N. (2014). Koroner arter hastalıđının sınıflanmasında Radial Basis fonksiyonu sinir ađı ve lojistik regresyon analizi. Ankara Üniversitesi Tıp Fakültesi Mecmuası, 60(3), 97–102.
- Sökmen, A. (2016). Anadilde eğitim ve başarı üzerine etkileri: Almanya örneđi. <http://www.tepav.org.tr> adresinden edinilmiştir.
- Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA Data. PLoS ONE 8(3), e57988
- Sun, L. & Bradley, K. D. (2012). Using the U S Pisa results to investigate the relationship between school computer use and student academic performance. Retrieved from: <http://www.uky.edu/>
- Takma, ., İşçi-Güneri, Ö. & Gevrekçi, Y. (2016). Investigation of stillbirth rate using logistic regression analysis in Holstein Friesian calves. *Ege Üniversitesi Ziraat Fakültesi Dergisi*, 53(3), 245–250.
- Tatlı, C. E., Ergin, D. A. ve Demir, E. (2016). Pisa 2012 Türkiye verilerine göre öğrencilerin matematik kaygısının sınıflandırıcıları. *İlköğretim Online*, 15(2), 696–707. <http://ilkogretim-online.org.tr> adresinden edinilmiştir.
- Tokmak, G. ve Aksoy, R. (2016). Futbol taraftarlarının takımın forma sponsor markasına olan satın alma niyetlerini etkileyen faktörlerin lojistik regresyon analizi ile tespit edilmesi. *Spor Bilimleri Dergisi*, 6(2), 1303–1414.
- Yıldırım, K. (2012). Pisa 2006 verilerine göre Türkiye’de eğitimin kalitesini belirleyen temel faktörler. *Türk Eğitim Bilimleri Dergisi*, 10(2), 229–255.
- Yılmaz, H. B. ve Aztekin, S. (2012, Haziran). *Türkiye’deki 15 yaş grubu öğrencilerin matematik okuryazarlığı başarılarını etkileyen bazı faktörlerin okul ve öğrenci düzeyine göre incelenmesi*. X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Niğde.
- Yörük, D. ve Dündar, S. (2012). Tüketicilerin internette alışveriş yapma olasılıklarının lojistik regresyon yöntemiyle tahmini. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 25, 451–462.