

## SOSYAL BİLİMCİLER İÇİN VERİ BİLİMİ ÜZERİNE AÇIK-KAYNAKLI KODLAMA ATÖLYESİ MODELİ\*

**Ömer Turan BAYRAKLI**, İstanbul Üniversitesi, İktisat Fakültesi,

[omerturanbayrakli@ogr.iu.edu.tr](mailto:omerturanbayrakli@ogr.iu.edu.tr), ORCID ID: <https://orcid.org/0000-0002-2867-458X>

**Onur KELEŞ**, Boğaziçi Üniversitesi, Sosyal Bilimler Enstitüsü, [onur.keles1@boun.edu.tr](mailto:onur.keles1@boun.edu.tr),

ORCID ID: <https://orcid.org/0000-0003-3157-822X>

**Talha GÖKTENTÜRK**, Yıldız Teknik Üniversitesi, Eğitim Fakültesi, [talhagok@yildiz.edu.tr](mailto:talhagok@yildiz.edu.tr),

ORCID ID: <https://orcid.org/0000-0002-8619-0698>

### GİRİŞ

Bir bilim araştırılmasında ortaya çıkarılmak istenen gerçeğin en mühim unsurlarından biri de problem cümlesinin oluşturulması, hipotez kurulması gibi araştırmanın çekirdeğini oluşturan adımların tamamlanmasıdır (Fraenkel ve diğerleri, 2012). İlgili adımların genellenebilirlik prensibinde çalışılması ise nicel araştırmaların özünü oluşturmaktadır (Creswell, 2009). Nicel araştırmaların gerçekleştirilebilmesi ise ancak araştırmanın amacı doğrultusunda ilgili analizlerin yürütülmesi ile mümkün olmaktadır. Parametrik veya parametrik olmayan veri setleri gibi sınıflamalarla gerçekleştirilen analizler için 20. yüzyılın sonuna doğru farklı programlar geliştirilmeye başlanmıştır.

Bilgisayar programlarındaki gelişime paralel olarak *International Business Machines*'in (IBM) öncülüğünü yaptığı başta *Statistical Package for the Social Sciences* (SPSS) ve *Analysis of Moment Structures* (AMOS) olmak üzere çok sayıda program bu doğrultuda geliştirilmiştir (Meyers ve diğerleri, 2013). Belirli sayıda analiz paketi içeren programlar düzenli aralıklarla güncellenmektedir (Field, 2009). Günümüzde ise kullanımı sürmekte ve araştırmalarda pratik faydası sebebiyle yer bulmaktadır. Bilhassa ek olarak kod yazmanın gerekli olmaması IBM'in SPSS ve AMOS gibi programlarının Türkiye'de yaygın olarak kullanımını kolaylaştırmaktadır. Ayrıca kullanım rehberi mahiyetindeki yayınların fazlalığı yaygın tercihin önünü açmaktadır (Uysal & Güyer, 2014). Bununla beraber ilgili programlar başta olmak üzere açık kaynaklı olmayan programlar sınırlı sayıda paket barındırması sebebiyle araştırmacıların daha geniş bir çerçevede çalışmalarını yürütmesine mâni olmaktadır.

Beklenebilecek bir gelişme olarak sosyal bilimlerdeki araştırmalar için klasik paket programları popülerliğini giderek kaybetmektedir (Bivand ve diğerleri, 2013). Sınırlı sayıda analiz metodunu barındıran SPSS, AMOS, Lisrel gibi programlar (Çokluk ve diğerleri, 2012) 21. yüzyılda gerçeğin keşfi için ihtiyacı karşılayamamaktadır. Bilim araştırmalarının esas amacı ise gerçeğin bütün yönleri ile açığa çıkarılmasıdır (Popper, 2013). Bu doğrultuda açık kaynaklı ve kodlama tabanlı R programı mühim bir alternatif olarak ortaya çıkmıştır.

R her ne kadar kod yazmayı öğrenme gerekliliği içerse de açık kaynaklıdır. İçindeki paketlerin sayısı yeni çalışmalarla üretilen paket sayısı arttıkça genişlemektedir. Dolayısıyla teorik olarak ihtiyaç duyulması hâlinde sınırsız sayıda yeni paket üretilerek analiz için daha geniş bir metodolojik seçeneçlilik sağlanmaktadır. Ayrıca program açık erişimdedir ve ücretsiz olarak kullanılabilir (Field ve diğerleri, 2012). Programın sosyal bilimlerde alanında çalışan araştırmacılar tarafından yaygın kullanımı Türkiye'deki çalışmaların kalitesinde ciddi bir artış

---

\* Bu çalışma ICOLDE2022-INTERNATIONAL CONGRESS ON OPEN LEARNING AND DISTANCE EDUCATION 2022 isimli kongrede sözlü bildiri olarak sunulmuştur.

ihtimali sunmaktadır. Sosyal bilimler alanı için ciddi bir potansiyel sunan kodlama tabanlı veri analiz süreçleri için de yol açıcı bir fonksiyona sahiptir. Python ise kodlama dilleri arasında son dönemde bu bakımdan öne çıkmaktadır.

Yapılan araştırmalar Python'un yükselen bir ilgi ile en sık kullanılan analiz aracı olduğunu ortaya koymaktadır (Stancin & Jovic, 2019). R ile birlikte günümüzde sık kullanılan bu yol 1991'de ilk ortaya çıkışından bu yana alternatif programlar yoluyla kullanım sıklığını artırmıştır (McKinney, 2012). Büyük veri ağlarında çalışmak, büyük veri kümelerini analiz etmek ve benzeri başlıklarda Python ciddi avantajlar sunmaktadır (Heiberger & Riebling, 2016). Büyük veri setleri ile çalışmanın yaygınlaşması sebebiyle araştırmacıların Python dilinde yetkinlik ihtiyacı kaçınılmaz olarak artmaktadır (Trilling, 2018).

Belirtilen avantajlara rağmen R programının ve Python dilinin yaygınlaşmadığı ve Lisrel, SPSS gibi programların Türkiye'de yaygın olarak kullanımının sürdüğü görülmektedir (Doğan & Uluman, 2016). Dünya'da yaygın kullanıma rağmen Türkiye'de böyle bir tablonun görülmesi ise (Arslan, 2021) ilgili seçeneklerin az bilinirliği ile açıklanabilir. Dolayısıyla Türkiye içinde R ve Python eğitimlerinin artması ve kullanımının yaygınlaştırılması ülke biliminin gelişmesinde mühim bir potansiyel arz etmektedir.

Potansiyel ile paralel olarak bu çalışma içinde sosyal bilimlerde bilim adamı adaylarına yönelik açık kaynaklı kodlama eğitimi verilmesi için bir model geliştirilmesine çalışılmıştır. Modelin bir yandan katılımcıların ihtiyaçlarını karşılamasına bir yandan da kodlama tabanlı analiz becerileri için yeterliliklerin kazanılmasına hizmet etmesi istenmiştir. Amaçlar doğrultusunda ise kurulan Kodlama Atölyesi'nin yeni çalışmalar ve açık kaynaklı ve kodlama tabanlı analizler için örnek bir rehber olması hedeflenmektedir.

## YÖNTEM

Araştırmanın amacı lisans seviyesindeki sosyal bilimler alanından bilim adayı adaylarına açık kaynaklı kodlama eğitimi verilmesine yönelik bir model oluşturmaktır. R ve Python eğitimlerinin verildiği bu atölyenin alt amaçları şu başlıklarda sıralanabilir:

- Açık kaynaklı kodlama tabanlı analize dair farkında olma seviyesinin artması.
- Sosyal bilim içerikli eğitim materyallerinin ve vazifelerin öz yeterlilik algılarına etkisinin belirlenmesi
- Katılımcıların kendi ihtiyaç alanlarının karşılanması

Araştırmacılar tarafından tutulan günlükler ve literatür taramasından elde edilen bulgular sosyal bilimlerde açık kaynaklı programlarda kodlama becerileri için ciddi bir yeterlilik problemi olduğuna işaret etmektedir (Akgün & Güntaş, 2018; Altıok ve diğerleri, 2018). İlgili problemde iyileşmenin gerçekleşmesi ise modelin arka planındaki motivasyon kaynağını oluşturmaktadır. Böylece araştırmanın metodunun uygulama odaklı aksiyon araştırması olarak belirlenmesi kararlaştırılmıştır (Creswell, 2012).

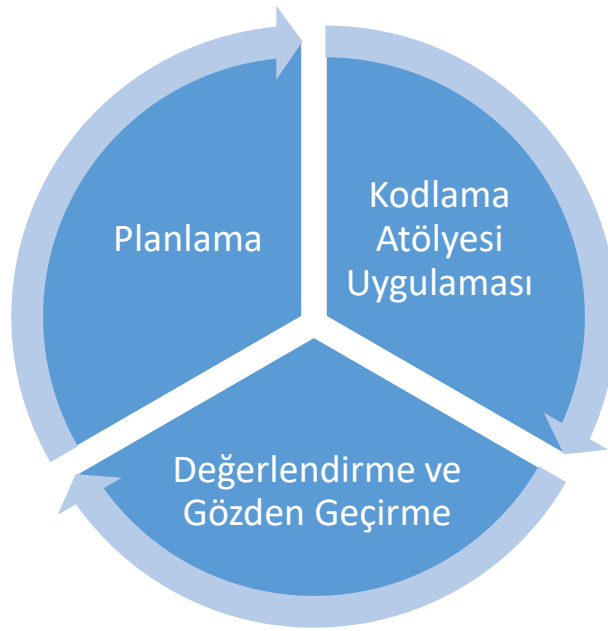
### Katılımcılar

Araştırma Yıldız Teknik Üniversitesi Dil Atölyesi Kulübü bünyesinde yürütülmüştür. Etkinliğin internet üzerinden duyurulmasının akabinde 74 başvuru alınmıştır. Katılımcılarının belirlenmesi için gönüllülük esası ve derse devam şartı aranmıştır. Bu şartı ise

30 katılımcı sağlamış ve 17 katılımcı başarı ile uygulama sürecini tamamlamıştır. Katılımcıların hepsi sosyal bilimler alanında lisans eğitimine devam eden üniversite öğrencileridir.

### Uygulama Süreci

Çalışma boyunca yapılan uygulamalar her katılımcının kendi öğrenme ihtiyaçları olduğu prensibinden hareket etmektedir. Dolayısıyla katılımcılardan alınan veriler ve araştırmacıların katılımcıların gelişim ihtiyaçlarına göre oluşturduğu yol haritası araştırmanın uygulama sürecini inşa etmiştir. Çağrı metni doğrultusunda Kodlama Atölyesi'ne başvuran lisans öğrencilerine kodlama becerilerine yönelik form uygulanmıştır. Ayrıca kodlama eğitimi sürecine yönelik görüşleri alınarak ön bilgi seviyeleri belirlenmiştir. Her haftalık uygulama sonrasında süreç gözden geçirilmiş ve varsa gerekli düzenlemeler yapılarak uygulamalar sürdürülmüştür.



**Resim 1.** *Kodlama Atölyesi'nin Haftalık Uygulama Süreci*

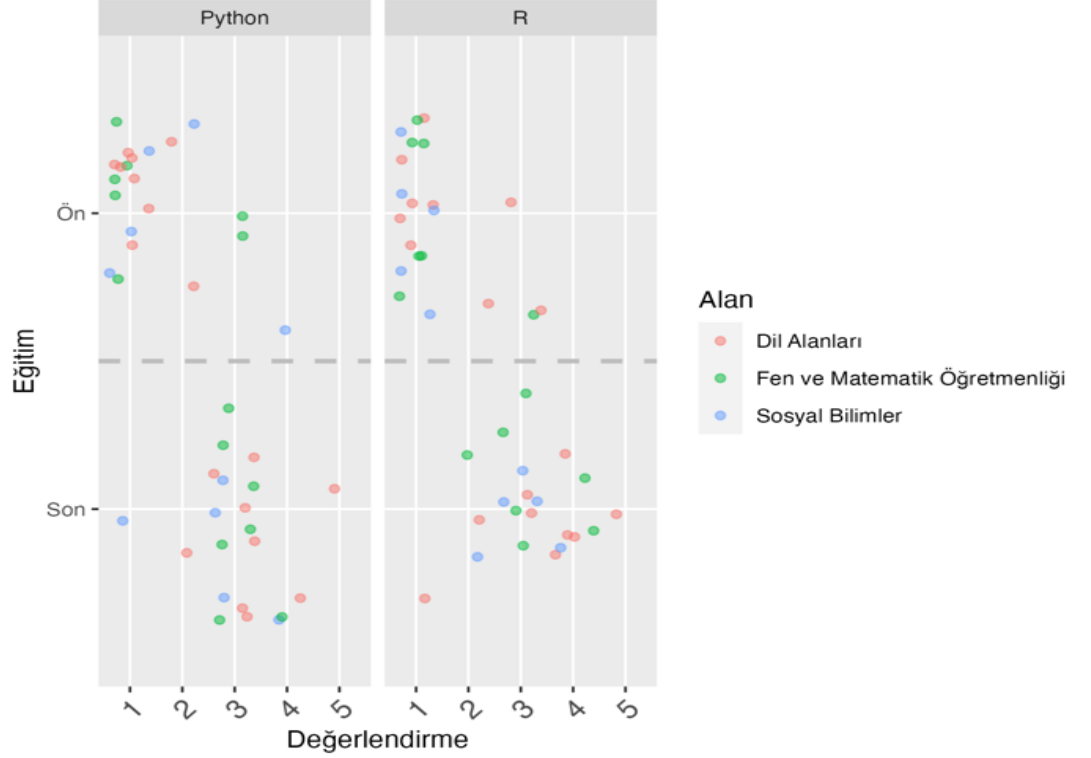
Böylece her modülde (R ve Python) 8 saat olmak üzere toplam 16 saatlik eğitim süreci tamamlanmıştır. Her hafta 2 ders olarak tamamlanan atölyede 1 teorik 1 de uygulamalı ders yapılarak hafta sonlandırılmıştır. Katılımcıların derse rahat erişebilmesi için hibrit sistem esas alınmıştır. Bir yandan sınıf içinde yüz yüze uygulama yapılırken bir yandan da Zoom üzerinden dersler canlı yayınlanmıştır. İletişim için Slack kullanılmıştır. Python ve R için kullanılan alan olarak Google Colaboratory seçilmiştir.

### Verilerin Analizi

Atölyenin etkisini anlamak için katılımcıların eğitimden önce ve sonra programlama becerisi değerlendirmeleri R'da MASS kütüphanesiyle analiz edilmiştir (Ripley ve diğerleri, 2013). Bağımlı değişken olarak öz değerlendirme puanları (1-5); sabit etkiler olarak uygulama (ön ve son) ve Dil (R ve Python) girilmiştir. Görüşme formu ve gözlemden elde edilen veriler ise tematik analiz doğrultusunda açık kodlama ile analiz edilmiştir (Miles & Huberman, 1994). Böylece veri havuzunun kendi iç dinamikleri ile süreci daha iyi bir şekilde tasvir etmesi amaçlanmıştır.

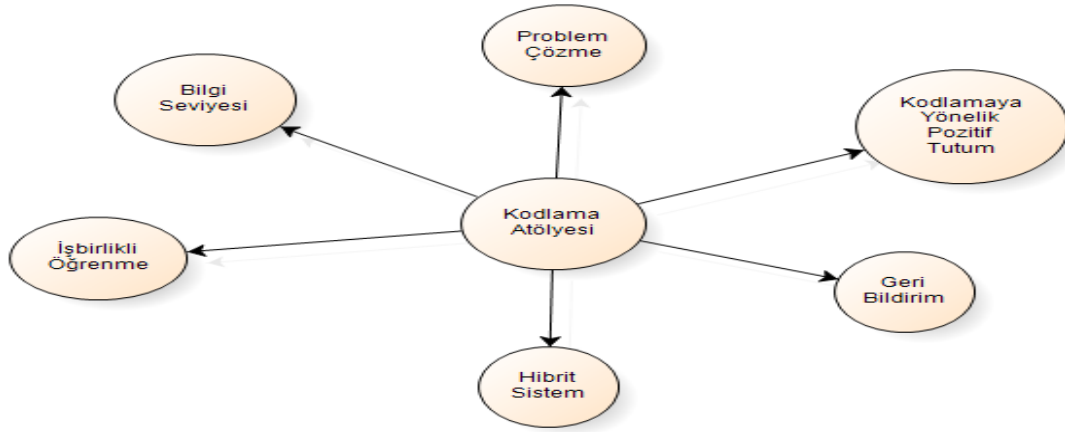
## BULGULAR

MASS paketinden elde edilen bulgular ve ilgili bulguları daha derinlikli anlamak için derlenen nitel verilerden elde edilen bulgular aşağıda sunulmuştur.



**Resim 2.** MASS Paketi ile Yapılan Analizden Elde Edilen Sonuçlar

Birinci resme göre gösterildiği gibi 5 haftalık eğitimi tamamladıktan sonra, dilden bağımsız olarak ( $p = .93$ ), katılımcılar R ve Python bilgilerinin anlamlı bir şekilde arttığı ve eğitimden yararlandıkları gözlemlenmiştir ( $p < .001$ ). MASS kütüphanesiyle normallik testinin ardından ordinal lojistik regresyon analizi ile elde edilen bu sonuç Kodlama Atölyesi'nin pozitif bir etki oluşturduğunu ortaya koymaktadır. Katılımcı görüşleri ve araştırmacı günlüklerinin nicel bulgularla örtüştüğü görülmektedir. Katılımcılar kolayca kodlamaya geçtiklerini ve belirttikleri ihtiyaçlara göre sürecin şekillenmesinin kendilerini pozitif etkilediğini belirtmişlerdir. Katılımcı görüşleri ve gözlem sonuçlarının harmanlanması ile oluşan nitel modelleme ise 3. şekilde sunulmaktadır.



**Resim 3.** Katılımcı Görüşleri ve Araştırmacı Gözlemlerinden Elde Edilen Sonuçlar

Uygulamayı başarıyla tamamlayarak süreç boyunca devamlılık gösteren bütün katılımcılardan elde edilen veriler ve araştırmacıların gözlemleri örtüşür görünmektedir. Buna göre hibrit sistem hem katılım kolaylığı sunması hem de çok yönlü seçenek sunma bakımından faydalı görülmüştür. Düzenli olarak geri bildirim almanın kodlama becerileri üzerinde faydalı görüldüğü belirlenmiştir. İşbirlikli öğrenmenin de dâhil olması ile bilgi seviyesi gelişmiş, problem çözme becerileri iyileşmiştir. Bunun tabii bir sonucu olarak da kodlamaya yönelik pozitif tutumun belirlendiği görülmektedir. Nihayetinde Kodlama Atölyesi uygulamasının amaçlarına ulaştığı ve temel seviyede kodlama becerisi kazandırmada iyileşme sunduğu söylenmelidir.

## SONUÇ VE ÖNERİLER

Araştırmadan elde edilen bulgular Kodlama Atölyesi'nin R ve Python kullanma becerileri yoluyla kodlama ve analiz problemlerini iyileştirdiğini ortaya koymaktadır. Alt başlıklarda ise uygulama sürecinin özellikleri için tartışmalar bilgi seviyesinin iyileşmesi, uygulama metodunun etkililiği, hibrit sistemin faydalarından bahsetmek gerekmektedir. İlgili başlıklar bir yandan Kodlama Atölyesi'nin ilk adımda keşfedilen pozitif yönleri olarak ön plana çıkarken gelecek araştırmalar içinde yol gösterme potansiyeli sunmaktadır.

Türkiye'de kodlama becerileri için yaşanan problemler düşünüldüğünde (Akgün & Güntaş, 2018; Altıok ve diğerleri, 2018) Kodlama Atölyesi mühim bir alternatiftir. Nicel bazda elde edilen sonuçlar bu noktada destekleyicidir. Gelecekte sosyal bilimlerde lisans müfredatlarında da kodlama eğitiminin daha fazla yer bulması faydalıdır. Bu noktada modelin kendine daha fazla yer bulması da beklenebilir. Nitekim araştırmalarda yaygın olarak R ve Python'un kullanılması (McKinney, 2012; Stancin & Jovic, 2019; Trilling, 2018) ihtiyacı artıran unsurlar arasındadır.

Uygulama süreci içinde düzenli geri bildirim verilmesi literatür ile uyumlu olarak (Jensen ve diğerleri, 2021; Kourgiantakis ve diğerleri, 2019) katılımcıların hatalarından ders çıkarmasını sağlamıştır. Sosyal bilimcilerin kodlama alanında Türkçe kaynak eksikliği de geri bildirim ile ilişkili olarak burada vurgulanmalıdır. Araştırmacılar artık klasik denebilecek SPSS ve AMOS gibi programlarını bir yandan da Türkçe kaynak zenginliği sebebiyle tercih etmektedir (Uysal & Güyer, 2014). Dolayısıyla eğitim modelinin geliştirilmesinin yanında mutlaka Türkçe kaynak zenginliği artmalıdır. Uygulamada geliştirilecek modellerin literatürün zenginleşmesine de katkı sağlayacağı söylenebilir.

Araştırmanın bir diğer yönü de Kodlama Atölyesi'nin kullandığı hibrit sistemin faydalarını ortaya çıkarmasıdır. Covid-19 salgını sonrası Dünya'da yaygın olarak kullanılmaya başlanan uzaktan eğitim imkânları (Adedoyin & Soykan, 2020) alternatif seçeneklerin fazlaşmasını sağlamıştır. Sürekli yeni binaların inşa edildiği kampüslerde artık inşaat maliyetleri ve alan yetmezliği mühim bir problem olarak öngörülebilir. Hibrit sistemler bir yandan derslerin kayıt altına alındığı bir yandan da sadece gerekli olması hâlinde yüz yüze katılım sağlandığı bir yapıdadır (Li ve diğerleri, 2021). Böylece sınıfların ihtiyacın üstünde yük taşıdığı yüz yüze uygulamaların önüne geçmektedir.

Tartışmanın son aşamasında bir müfredat karşılık olarak Kodlama Atölyesi'nin uygulanabilir olması için gerekliliklerden bahsetmek faydalıdır. Model araştırmacılar tarafından geliştirilmeye devam etmektedir. Dolayısıyla en az bir dönemlik bir uygulamada esnek sistemini koruyarak uygulanmaya devam etmelidir. Gelecek araştırmalarda bu noktada çalışmaların sürdürülmesine ve daha geniş kitlelere katkının teminine çalışılacaktır.

## KAYNAKLAR

- Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments*, 1–13.
- Akgün, Ö. E., & Güntaş, S. (2018). Lisansüstü öğrencilerinin bilimsel araştırma yeterliliklerinin danışman görüşlerine dayalı olarak incelenmesi: Kuzey Kıbrıs Türkiye Cumhuriyeti örneği. *Sakarya University Journal of Education*, 8(2), 131–144. <https://doi.org/10.19126/suje.365103>
- Altıok, S., Yükseltürk, E., & Üçgül, M. (2018). The investigation of postgraduate students' scientific research competencies and anxieties towards research in terms of various variables. *Journal of Higher Education and Science*, 8(2), 348–367. <https://doi.org/10.5961/jhes.2018.277>
- Arslan, İ. (2021). *Python ile veri bilimi* (4th ed.). Pusula.
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*. Springer New York. <https://doi.org/10.1007/978-1-4614-7618-4>
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (Vol. 2). Pegem Akademi Ankara.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches*. Sage Publications.
- Creswell, J. W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Pearson.
- Doğan, C. D., & Uluman, M. (2016). İstatiksel veri analizinde R yazılımı ve kullanımı. *İlköğretim Online*, 15(2), 615–634. <https://doi.org/10.17051/io.2016.24991>
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage Publications.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage Publications.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (7th ed.). McGraw-Hill.
- Heiberger, R. H., & Riebling, J. R. (2016). Installing computational social science: Facing the challenges of new information and communication technologies in social science. *Methodological Innovations*, 9, 1–11. <https://doi.org/10.1177/2059799115622763>

- Jensen, L. X., Bearman, M., & Boud, D. (2021). Understanding feedback in online learning – A critical review and metaphor analysis. *Computers & Education*, 173, 1–12. <https://doi.org/10.1016/j.compedu.2021.104271>
- Kourgiantakis, T., Sewell, K. M., & Bogo, M. (2019). The importance of feedback in preparing social work students for field education. *Clinical Social Work Journal*, 47(1), 124–133. <https://doi.org/10.1007/s10615-018-0671-8>
- Li, Q., Li, Z., & Han, J. (2021). A hybrid learning pedagogy for surmounting the challenges of the COVID-19 pandemic in the performing arts education. *Education and Information Technologies*, 26(6), 7635–7655. <https://doi.org/10.1007/s10639-021-10612-1>
- McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2013). *Performing data analysis using IBM SPSS*. John Wiley & Sons.
- Miles, M. B., & Huberman, A. M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Sage Publications.
- Popper, K. (2013). *Realism and the aim of science* (W. W. Bartley, Ed.). Routledge. <https://doi.org/10.4324/9780203713969>
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package ‘mass.’ *Cran r*, 538, 113–120.
- Stancin, I., & Jovic, A. (2019). An overview and comparison of free Python libraries for data mining and big data analysis. *42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 977–982. <https://doi.org/10.23919/MIPRO.2019.8757088>
- Trilling, D. (2018). *Doing computational social science with Python: An introduction*. Social Science Research Network. Trilling, Damian, Doing Computational Social Science with Python: An Introduction (January 21, 2018). Available at SSRN: <https://ssrn.com/abstract=2737682>
- Uysal, M., & Güyer, T. (2014). İstatistiksel veri analizine ilişkin genişleyebilir bir karar ağacı tasarımı. *Bilişim Teknolojileri Dergisi*, 7(3), 33–43. <https://doi.org/10.12973/bid.2022>