

Teaching Practice/Öğretim Uygulaması

Evaluating KIP against SPSS: A Reliable and Valid Statistical Tool for Academic Research

Süleyman ULUPINAR¹  Serhat ÖZBAY²  Fadime ULUPINAR³  Selim ASAN^{4*} 
İzzet İNCE⁵  Deniz BEDİR⁶  Cebail GENÇOĞLU⁷  Salih ÇABUK⁸ 

^{1,2,3,4,6,7,8} Erzurum Technical University, Faculty of Sports Sciences, Erzurum, Türkiye

¹ suleyman.ulupinar@erzurum.edu.tr; ² serhat.ozbay@erzurum.edu.tr; ³ fadime.ulupinar@erzurum.edu.tr; ⁴ selim.asan@erzurum.edu.tr;

⁶ deniz.bedir@erzurum.edu.tr; ⁷ cebrail.gencoglu@erzurum.edu.tr; ⁸ salih.cabuk@erzurum.edu.tr

⁵ Ankara Yıldırım Beyazıt University, Faculty of Sports Sciences, Ankara, Türkiye, izzetince43@gmail.com


* Corresponding Author: selim.asan@erzurum.edu.tr

Article Info

Received: 19 January 2025

Accepted: 08 March 2025

Keywords: KIP, parametric tests, non-parametric tests, data analysis

 10.18009/jcer.1622772

Publication Language: Turkish

Abstract

This study aims to evaluate the “Simplified Statistical Program (KIP),” a statistical tool designed to simplify analysis and reporting for academic research, by comparing it with the reference software, SPSS (version 27.0). The study assesses KIP’s performance in parametric and non-parametric tests, assumption checks, graphical capabilities, and Word-format output generation using four publicly available datasets selected for their suitability to test a range of statistical analyses. The program demonstrated high accuracy and reliability, with results largely consistent with SPSS across these datasets, showing identical *p*-values, statistical test values, and effect sizes in both parametric and non-parametric tests, as well as assumption checks. In conclusion, KIP is a reliable and user-friendly tool for academic and professional use, but it has limitations (e.g. lack of multi-platform support).



To cite this article: Ulupinar, S., Özbay, S., Ulupinar, F., Asan, S., İnce, İ., Bedir, D., ... & Çabuk, S. (2025). KIP ve SPSS karşılaştırması: Akademik araştırmalar için güvenilir ve geçerli bir istatistik yazılımı. *Journal of Computer and Education Research*, 13 (25), 602-627. <https://doi.org/10.18009/jcer.1622772>

KIP ve SPSS Karşılaştırması: Akademik Araştırmalar için Güvenilir ve Geçerli Bir İstatistik Yazılımı

Makale Bilgisi

Geliş: 19 Ocak 2025

Kabul: 08 Mart 2025

Anahtar kelimeler: KIP, parametrik testler, non-parametrik testler, veri analizi

 10.18009/jcer.1622772

Yayın Dili: Türkçe

Öz

Bu çalışma, akademik araştırmalar için analiz ve raporlama süreçlerini kolaylaştırmak amacıyla geliştirilen “Kolaylaştırılmış İstatistik Programı (KİP)”in, referans yazılım SPSS (sürüm 27.0) ile karşılaştırmalı olarak değerlendirilmesini amaçlamaktadır. Çalışma, KİP’in parametrik ve non-parametrik test performansını, varsayım kontrollerindeki başarısını, grafik oluşturma yeteneklerini ve Word formatında çıktı üretim kapasitesini, analiz çeşitliliğini desteklemek amacıyla seçilen dört açık kaynaklı veri seti kullanılarak incelemektedir. Bulgular, KİP’in bu veri setlerinde SPSS ile büyük ölçüde tutarlı olduğunu, parametrik ve non-parametrik testlerde *p* değerleri, istatistiksel test değerleri ve etki büyüklüklerinin eşleştiğini, varsayım kontrollerinde ise tam uyum sağladığını göstermiştir. Sonuç olarak, KİP, akademik ve profesyonel kullanım için güvenilir ve kullanıcı dostu bir araç olarak öne çıkmakla birlikte, çoklu platform desteği eksikliği gibi sınırlamaları bulunmaktadır.

Summary

Evaluating KIP against SPSS: A Reliable and Valid Statistical Tool for Academic Research

Süleyman ULUPINAR ¹  Serhat ÖZBAY ²  Fadime ULUPINAR ³  Selim ASAN ^{4*} 
İzzet İNCE ⁵  Deniz BEDİR ⁶  Cebail GENÇOĞLU ⁷  Salih ÇABUK ⁸ 

^{1,2,3,4,6,7,8} Erzurum Technical University, Erzurum, Türkiye,

¹ suleyman.ulupinar@erzurum.edu.tr; ² serhat.ozbay@erzurum.edu.tr; ³ fadime.ulupinar@erzurum.edu.tr; ⁴ selim.asan@erzurum.edu.tr;

⁶ deniz.bedir@erzurum.edu.tr; ⁷ cebrail.gencoglu@erzurum.edu.tr; ⁸ salih.cabuk@erzurum.edu.tr

⁵ Ankara Yıldırım Beyazıt University, Ankara, Türkiye,

* Corresponding Author: selim.asan@erzurum.edu.tr

Introduction

Statistical analyses play a critical role in scientific research, helping to test hypotheses and derive meaningful insights from complex datasets. While tools like SPSS are widely used for these purposes due to their extensive features, they often present challenges such as complexity, high costs, and steep learning curves, particularly for users with limited statistical knowledge. This has driven the need for more accessible and user-friendly tools that can simplify analysis and reporting processes.

Existing alternatives like Jamovi and JASP offer free and open-source solutions but often lack essential features, such as generating tables suitable for academic reporting and simultaneously conducting multiple analyses. These limitations underscore the demand for a practical tool that combines ease of use, functionality, and compatibility with academic standards.

The Simplified Statistical Program (KİP) was developed to address these needs, offering features such as Turkish language support, a user-friendly interface, automated assumption checks, and the ability to report parametric and non-parametric results simultaneously. Additionally, KİP generates Word-formatted outputs tailored for academic use, enhancing the efficiency of thesis and manuscript preparation. The program's ability to calculate effect sizes and provide basic visualizations further enriches the user experience.

This study aims to evaluate KİP's reliability and validity by comparing its results with SPSS across various datasets and statistical tests. It also seeks to demonstrate the program's potential to make statistical analysis more accessible and practical for researchers,

thereby contributing to the development of more user-centered statistical software solutions in the future.

Method

This study adopts a validation and comparison research model to evaluate the accuracy and reliability of the Simplified Statistical Program (KİP) in comparison with SPSS. The program was developed using Python and PyQt5, creating a user-friendly interface optimized for academic use. Statistical functions, such as *t*-tests, ANOVA, correlation, regression, and assumption checks, were integrated with libraries like SciPy and Pingouin, while outputs were designed to meet academic standards using python-docx.

A total of 31 participants, including 6 graduate students and 25 academics experienced in SPSS, tested KİP. Participants analyzed datasets tailored to their expertise and provided feedback on ease of use, functionality, and output quality.

Four publicly available datasets (New Drug Data, Body Fat Data, Exam Data, and Disease Data) were analyzed to compare KİP's results with SPSS (http://calcnet.mth.cmich.edu/org/spss/prjs_datasets.htm). Both parametric and non-parametric tests, assumption checks, and post-hoc analyses were conducted. KİP's ability to generate descriptive statistics, visualizations, and Word-formatted outputs was also assessed.

Results

The independent *t*-test analysis using the "Exam Data" dataset showed complete consistency between the Simplified Statistical Program (KİP) and SPSS for parametric analyses, with identical means, standard deviations, *t*-values, and *p*-values. Variance homogeneity results (Levene's test) were also identical. For non-parametric analyses, Welch *t*-test and Mann-Whitney U test results matched closely, with a minor discrepancy in the *p*-value for the Mann-Whitney U test (KİP: 0.970; SPSS: 0.960), attributed to methodological differences between the programs.

Paired sample *t*-test and Wilcoxon Signed-Rank test with the "New Drug Data" dataset showed perfect agreement between KİP and SPSS in all statistical measures, including *p*-values. Similarly, one-way ANOVA and Kruskal-Wallis tests using the "Disease Data" dataset yielded identical results for *F*-values, *p*-values, and variance homogeneity.

Repeated measures ANOVA with the "Body Fat Data" dataset demonstrated full alignment in F -values, p -values, and partial eta-squared values. Mauchly's test for sphericity and Bonferroni post-hoc analyses also matched exactly. Non-parametric Friedman test results were identical.

Two-way mixed ANOVA with the "New Drug Data" dataset showed full consistency in group effects and interaction effects, with a minor discrepancy in the F -value for the main effect of time (KİP: $F = 120.721$; SPSS: $F = 120.070$). This difference arose from differences in the handling of variance components by KİP and SPSS but did not affect the interpretability of results.

Correlation analysis using the "Body Fat Data" dataset produced identical Pearson and Spearman correlation results, as well as Cronbach's alpha values. Intraclass correlation coefficients (ICC) for various models demonstrated strong alignment, further supporting the reliability of KİP.

Simple and multiple regression analyses with the "Body Fat Data" dataset showed complete agreement between KİP and SPSS for R -squared values and parameter estimates (B values, Beta coefficients, t -values, and p -values), affirming the program's accuracy in regression analyses.

Discussion and Conclusion

The Simplified Statistical Program (KİP) demonstrated high reliability and validity when compared to SPSS, as results across various statistical analyses showed strong agreement. Parametric tests, including t -tests, ANOVA, and regression, as well as non-parametric tests like Mann-Whitney U, Kruskal-Wallis, and Friedman tests, exhibited near-perfect alignment in test statistics (t , F values), p -values, and effect sizes. This highlights KİP's robustness and compatibility with academic standards while emphasizing its applicability across disciplines like health, sports, and social sciences.

While SPSS remains a trusted tool for advanced statistical analyses, its complexity and cost are often cited as barriers for novice users. KİP bridges this gap by offering a user-friendly interface, Turkish language support, and Word-formatted outputs optimized for academic reporting. These features make it especially appealing to postgraduate students and researchers with limited statistical expertise. Furthermore, its compact design and ability

to perform comprehensive parametric and non-parametric analyses, along with assumption checks and effect size calculations, position KIP as a practical alternative to established tools.

Despite its strong performance, minor discrepancies were observed in certain analyses, such as the Mann-Whitney U test and mixed ANOVA. These differences were attributed to methodological variations between KIP's underlying Python libraries and SPSS. However, these discrepancies had no significant impact on result interpretation, reinforcing KIP's reliability.

KIP's intuitive design and emphasis on accessibility offer significant advantages for researchers, especially in practical fields. However, limitations such as lack of multi-platform compatibility, limited graphical options, and absence of multilingual support provide avenues for future improvement. Addressing these constraints could enhance KIP's usability and extend its reach to a broader global audience. Overall, KIP represents a valuable contribution to accessible and efficient statistical analysis tools, with the potential for widespread adoption in both academic and professional settings.

Giriş

Bilimsel arařtırmalar, gözlemlenen olgular arasındaki iliřkileri anlamlandırmak ve bu iliřkilerin genellenebilirliđini sađlamak için istatistiksel analizlere dayanır (Alpar, 2014; Özdamar, 2018; Shepherd & Richardson, 2024). Bu analizler, yalnızca hipotezlerin test edilmesinde deđil, aynı zamanda karmařık veri setlerinden anlamlı sonuçlar çıkarılmasında da kritik bir rol oynar (Akbulut, 2022; Ulupınar & İnce, 2021). Özellikle uygulamalı alanlarda, istatistiksel yöntemler dođru kararların alınması ve bilimsel bulguların geçerliliđinin sađlanması temel araçlar olarak öne çıkar (Hopkins ve diğ., 2009; Verma, 2016). Örneđin, bir egzersiz programının etkinliđini ölçmek, tedavi yöntemlerini karşılařtırmak ya da sosyal bir müdahalenin etkisini deđerlendirmek gibi arařtırma soruları, ancak dođru istatistiksel analizlerle yanıtlanabilir (Özbyay ve diğ., 2024; Ulupınar & İnce, 2021). Bununla birlikte, istatistiksel süreçlerin dođruluđu sadece kullanılan yöntemlerin uygunluđuna deđil, aynı zamanda analizlerin gerçekleştirildiđi araçların güvenilirliđine de bađlıdır (Bademci, 2022; Loerts ve diğ., 2020; Ulupınar, 2022). Bu nedenle, arařtırma süreçlerinde kullanılan yazılımlar, hem dođru hem de kolay erişilebilir olmalıdır.

İstatistiksel analizlerin önemi akademik çalışmalar için oldukça büyük olmasına rađmen, bu süreçleri kolaylařtıran mevcut araçlar çođu zaman kullanıcılar için karmařık ve yorucu bir deneyim sunar (Lino Calle ve diğ., 2024; Love ve diğ., 2019). Örneđin, Statistical Package for the Social Sciences (SPSS) gibi popüler yazılımlar, geniş özellik yelpazesi ve güçlü analiz yetenekleri ile dikkat çekse de, özellikle istatistik bilgisi sınırlı kullanıcılar için öğrenmesi zor bir yapıya sahiptir (Alpar, 2014; Özdamar, 2018). Menülerin karmařıklıđı, teknik terimlerin fazlalıđı, veri giriř süreçlerinin ve sonuç çıktılarının raporlanmasının kullanıcı dostu olmaması, temel analizler için bile bu tür araçların kullanımını zorlařtırmaktadır (Jain & Sengar, 2024; Stehlik-Barry & Babinec, 2017). Ayrıca, SPSS gibi programların lisans maliyetleri, lisansüstü öğrenciler ve bađımsız arařtırmacılar için ciddi bir engel oluřtırmaktadır. Bu durum, pratik uygulamalara odaklanan disiplinlerde, arařtırmacıları daha ulařılabilir ve kullanıcı dostu çözümler arayıřına yönlendirmektedir. Bu bağlamda, KİP, SPSS'e kıyasla daha basit bir arayüz ve ücretsiz erişim sunarak bu engelleri aşmayı hedefler; ancak, SPSS'in sunduđu ileri analiz seçenekleri ve geniş veri işleme kapasitesine kıyasla KİP'in henüz sınırlı işlevselliđi (örneğin, grafik çeřitliliđi ve çoklu platform desteđi eksikliđi) bir dezavantaj olarak öne çıkmaktadır. Jamovi ve JASP gibi

ücretsiz ve açık kaynaklı yazılımlar bazı kullanıcılar için alternatif olarak öne çıksa da, bu araçlar çoğu zaman kullanıcıların pratik ihtiyaçlarına tam olarak yanıt verememektedir (Love ve diğ., 2019; Şahin & Aybek, 2019). Örneğin, bu programlar genellikle akademik raporlamalara uygun tablo oluşturma konusunda ve birçok analizi aynı anda gerçekleştirme ve raporlama konusunda tercih edilmeyebilirler. Buna karşın, KİP, Jamovi ve JASP'e kıyasla Türkçe dil desteği ve Word formatında otomatik raporlama gibi özelliklerle avantaj sağlar; ancak, bu araçların sunduğu gelişmiş görselleştirme seçenekleri KİP'te henüz mevcut değildir. Bu tür eksiklikler, özellikle araştırma süreçlerini hızlandırmak ve çıktıları doğrudan akademik çalışmalarında kullanmak isteyen kullanıcılar için önemli bir dezavantaj oluşturmaktadır. Dolayısıyla, hem kullanım kolaylığı hem de işlevselliği bir arada sunabilen bir araca duyulan ihtiyaç devam etmektedir.

Türkçe dil desteğiyle, kullanıcı dostu bir arayüze sahip, pratik bir istatistik aracı geliştirme fikri, özellikle lisansüstü öğrenciler ve temel istatistik analizlerine ihtiyaç duyan araştırmacılar için önemli bir gereksinimden doğmuştur. Böyle bir araç, yalnızca istatistiksel analizlerin karmaşıklığını azaltmakla kalmayıp, aynı zamanda kullanıcıların raporlama süreçlerini hızlandırmayı hedefler. Özellikle tez ve makale gibi akademik raporlamalarda kullanılmak üzere Word formatında çıktılar üretebilmek, mevcut araçlarla karşılaştırıldığında büyük bir avantaj sağlar. Bunun yanı sıra, analiz sürecini desteklemek için istatistiksel varsayımların kontrol edilmesi, tanımlayıcı istatistiklerin otomatik sunulması ve parametrik ile non-parametrik sonuçların aynı anda raporlanabilmesi, bu tür bir aracın sağladığı işlevselliği daha da ileriye taşıyabilir. Ek olarak, pratik öneme sahip sonuçların (örneğin, etki büyüklüğü) otomatik hesaplanarak raporlara entegre edilmesi, kullanıcıların analizlerinin derinliğini artırırken, sonuçların daha kolay yorumlanmasını sağlar. Temel grafikler ile analiz sonuçlarını görselleştirme süreci, araştırmacıların veri yorumlamasını daha etkili hale getirebilir. Aynı zamanda, düşük dosya boyutu ve yüksek erişilebilirlik sayesinde, farklı cihazlarda kolayca kullanılabilir ve herhangi bir ek yazılım maliyeti gerektirmez. Bu özellikler, yalnızca akademik çalışmalar için değil, aynı zamanda istatistik süreçlerini öğrenmek isteyen bireyler için de büyük bir avantaj sağlar. Bu özelliklere sahip bir analiz aracının sağlayacağı faydalar, yalnızca analiz süreçlerini kolaylaştırmakla kalmaz, aynı zamanda istatistiksel bilgiye erişimi demokratikleştirerek geniş bir kullanıcı kitlesine hitap eder.

Bu çalışma, akademik arařtırmalar için analiz sürecini ve raporlamayı basitleřtirmek üzere geliřtirilen “Kolaylařtırılmıř İstatistik Programı (KİP)” isimli istatistik aracının referans bir araçla olan tutarlılıđını bilimsel olarak deđerlendirmeyi amaçlamaktadır. Programın farklı veri setleri ve analiz türlerinde sađladığı sonuçların dođruluđu, güvenilirliđi ve tekrarlanabilirliđi incelenecektir. Aynı zamanda, akademik kullanım için sunduđu avantajlar, kullanıcı deneyimi ve raporlama süreçlerindeki kolaylıklar detaylı bir řekilde ele alınacaktır. Çalışma kapsamında, programın parametrik ve non-parametrik testlerdeki performansı, varsayım kontrollerindeki başarısı, grafik ve Word formatında çıktı üretme gibi özellikleri deđerlendirilecektir. Bu analizler, aracın akademik standartlara uygunluđunu ve pratik kullanımda sađladığı faydaları ortaya koymayı hedeflemektedir. Sonuç olarak, bu çalışmanın, KİP programının istatistiksel analiz süreçlerindeki etkinliđini ve kullanıcı dostu özelliklerini kanıtlayarak, akademik ve profesyonel dünyada yaygın kullanım potansiyelini ortaya koyması beklenmektedir. Ayrıca, bu tür bir araç geliřtirme fikrinin, gelecekte istatistik yazılımlarına yönelik daha erişilebilir, etkili ve kullanıcı odaklı yaklaşımlara ilham verebileceđi düşünölmektedir.

Yöntem

Arařtırma Modeli

Bu çalışma, geliřtirilen istatistik aracının (KİP) analiz sonuçlarının SPSS ile uyumunu deđerlendirmek amacıyla tasarlanmış bir dođrulama ve karşılařtırma arařtırması modeline dayanmaktadır. Arařtırma, programın sonuçlarının dođruluđunu ve tekrarlanabilirliđini ölçmek için farklı veri setleri kullanılarak gerçekeřtirilmiřtir.

Programın Geliřtirme Süreci

Bu çalışma kapsamında istatistiksel analizleri kolaylařtırmayı hedefleyen bir araç geliřtirilmiřtir. Programın geliřtirilmesi sırasında Python programlama dili tercih edilmiř ve özellikle PyQt5 kütüphanesi ile kullanıcı dostu bir arayüz tasarlanmıřtır (PyQt5, grafik arayüz oluřturmada yaygın bir araçtır ve teknik kullanıcılar için esneklik sađlar). Geliřtirme ortamı olarak PyCharm IDE kullanılmıř, bu da kodun düzenli, modöler ve geniřletilebilir bir yapıda oluřturulmasını sađlamıřtır. PyQt5, kullanıcı etkileřimlerini yönetmek için QMainWindow, QTableWidgetItem, QMessageBox, QShortcut gibi sınıflar kullanılarak analize yönelik dinamik bir arayüz oluřturulmasına olanak tanımıřtır; bu sınıflar, teknik olarak

arayüzün işlevselliğini artırmak için temel bileşenlerdir. QFileDialog gibi bileşenler, dosya içe ve dışa aktarma süreçlerini kolaylaştırmıştır (Şekil 1). İstatistiksel analiz fonksiyonlarının doğruluğunu sağlamak için SciPy kütüphanesi ile bağımsız örneklemeler *t*-test (ttest_ind), bağımlı örneklemeler *t*-test (ttest_rel), ANOVA (f_oneway), normallik testi (shapiro), ve varyans homojenliği (levene) gibi analizler entegre edilmiştir (SciPy, matematiksel ve istatistiksel hesaplamalar için kullanılan bir kütüphanedir ve bu testler, analizlerin temel yapı taşlarını oluşturur). Ek olarak, Pingouin kütüphanesi, daha ileri istatistiksel analizler için kullanılmıştır; örneğin, ANOVA, sphericity ve intraclass correlation (ICC) hesaplamaları gerçekleştirilmiştir (Pingouin, ileri düzey istatistiksel analizler için kullanıcı dostu bir alternatif sunar). Pandas ve NumPy kütüphaneleri ise veri işleme ve hesaplama süreçlerini desteklemiştir (Pandas ve NumPy, veri manipülasyonu ve sayısal hesaplamalar için temel araçlardır). Grafiklerin oluşturulması ve görselleştirme süreçlerinde, Matplotlib kullanılarak kolay anlaşılır ve özelleştirilebilir grafikler tasarlanmıştır (Matplotlib, veri görselleştirme için yaygın bir araçtır). Analiz sonuçlarının akademik standartlara uygun şekilde raporlanmasını sağlamak için python-docx kütüphanesi entegre edilmiştir; bu, Word formatında tablolar ve grafiklerle düzenlenebilir raporlar oluşturulmasını sağlamıştır (Şekil 2).

Geliştirilen programın, kullanıcıların doğrudan çalıştırabileceği bir setup dosyasına dönüştürülmesi amacıyla NSIS (Nullsoft Scriptable Install System) kullanılmıştır; NSIS, teknik olarak kurulum sürecini basitleştiren bir araçtır. NSIS, hafif ve hızlı bir kurulum dosyası oluşturularak programın kolayca dağıtılmasını sağlamıştır. Ayrıca, kurulum dosyasının sistemde düşük yer kaplaması, programın erişilebilirliğini artırmıştır. Program, düşük sistem gereksinimleriyle çalışacak şekilde optimize edilmiş ve Türkçe dil desteğiyle geniş bir kullanıcı kitlesine hitap edecek şekilde tasarlanmıştır. Program geliştirilirken, kullanıcı geri bildirimleri alınarak üç farklı aşamada düzeltmeler ve iyileştirmeler yapılmıştır. Her geri bildirim aşamasında programın arayüzü ve işlevselliği optimize edilerek nihai hali oluşturulmuştur.

KİP

Bu istatistik programı Doç. Dr. Süleyman ULUPINAR ve Doç. Dr. Serhat ÖZBAY tarafından geliştirilmiştir.

Bağımsız Örneklem t-Test

	GrNo	Age	height	body.mass	BMI	train.age	kastbest	kast1	kast2	kast3	kast4	kast5	
1	ERKEK	18,00	169,00	59,00	20,70	9,00	6,21	6,53	6,32	6,21	6,61	6,86	32,5
2	ERKEK	21,00	173,00	66,00	22,10	10,00	6,16	6,16	6,50	6,73	6,78	7,06	33,2
3	ERKEK	22,00	174,00	70,00	23,10	8,00	5,60	5,60	5,85	5,82	5,72	5,67	28,6
4	ERKEK	17,00	172,00	67,00	22,60	11,00	5,70	5,93	5,70	5,82	6,09	6,02	29,5
5	ERKEK	22,00	183,00	79,00	23,60	6,00	6,30	6,79	6,70	6,30	6,65	6,58	33,0
6	ERKEK	25,00	174,00	70,00	23,10	7,00	5,77	5,82	5,77	6,07	5,83	6,23	29,7
7	ERKEK	23,00	169,00	67,00	23,50	9,00	5,10	5,36	5,10	5,13	5,66	7,00	28,2
8	ERKEK	19,00	177,00	64,00	20,40	10,00	7,14	7,14	7,20	7,50	8,03	8,39	38,4
9	ERKEK	19,00	175,00	69,00	22,50	7,00	6,16	6,16	6,16	6,16	6,68	6,47	32,3
10	ERKEK	23,00	174,00	70,00	23,10	7,00	6,16	6,16	6,16	6,16	6,54	7,47	34,1
11	ERKEK	24,00	172,00	61,00	20,60	9,00	6,16	6,16	6,16	6,16	6,78	7,05	33,2
12	ERKEK	25,00	170,00	57,00	19,70	8,00	6,16	6,16	6,16	6,16	6,34	6,72	31,0
13	KADIN	21,00	168,00	59,00	20,90	3,00	8,63	8,75	8,63	8,11	8,56	9,16	43,2
14	KADIN	19,00	169,00	67,00	23,50	2,00	6,59	6,59	7,44	7,81	7,51	8,28	37,6

Analiz Tamamlandı

Bağımsız t-Test analizleri başarıyla tamamlandı.

OK

	Değişken	ERKEK	KADIN	t-İstatistiği	p-Değeri	Hedges' g
1	Age	21.5 ± 2.71	20.83 ± 2.04	0.681	0.503	0.27
2	height	173.5 ± 3.85	172.67 ± 4.23	0.505	0.619	0.20
3	body.mass	66.58 ± 5.9	67.0 ± 6.56	-0.164	0.872	0.07
4	BMI	22.08 ± 1.36	22.44 ± 1.41	-0.634	0.533	0.25
5	train.age	8.42 ± 1.51	2.75 ± 1.14	10.403	0.000	4.09

Analiz yap Sonuçları kaydet Temizle

Şekil 1. KİP programı arayüzünde veri girişi ve analiz süreci

Bağımsız Örneklem t-Test Sonuçları

Değişken	ERKEK (Ort ± SS)	KADIN (Ort ± SS)	t-İstatistiği	p-Değeri	Hedges' g
kast1	6.26 ± 0.55	8.15 ± 0.83	-6.533	0.000	2.59
kast2	6.2 ± 0.58	7.99 ± 0.8	-6.310	0.000	2.47
kast3	6.28 ± 0.59	8.11 ± 0.78	-6.490	0.000	2.56
kast4	6.48 ± 0.64	8.47 ± 0.83	-6.590	0.000	2.60
kast5	6.79 ± 0.71	9.22 ± 1.01	-6.833	0.000	2.68
kasttotal	32.01 ± 2.83	41.74 ± 4.0	-6.879	0.000	2.71

Welch t-Test Sonuçları

Değişken	ERKEK (Ort ± SS)	KADIN (Ort ± SS)	t-İstatistiği	p-Değeri	Hedges' g
kast1	6.26 ± 0.55	8.15 ± 0.83	-6.533	0.000	2.59
kast2	6.2 ± 0.58	7.99 ± 0.8	-6.310	0.000	2.47
kast3	6.28 ± 0.59	8.11 ± 0.78	-6.490	0.000	2.56
kast4	6.48 ± 0.64	8.47 ± 0.83	-6.590	0.000	2.60
kast5	6.79 ± 0.71	9.22 ± 1.01	-6.833	0.000	2.68
kasttotal	32.01 ± 2.83	41.74 ± 4.0	-6.879	0.000	2.71

Non-parametrik Mann-Whitney U Testi Sonuçları

Değişken	ERKEK (Median [ÇAÇ])	KADIN (Median [ÇAÇ])	p-Değeri
kast1	6.34 [0.74]	8.22 [0.75]	0.000
kast2	6.29 [0.71]	8.1 [0.97]	0.000
kast3	6.24 [0.59]	8.04 [0.94]	0.000
kast4	6.58 [0.68]	8.54 [0.91]	0.000
kast5	6.79 [0.64]	9.12 [1.67]	0.000
kasttotal	32.41 [3.54]	42.19 [5.73]	0.000

Normallik Testi Notları

kast1: ERKEK p-değeri=0.911, KADIN p-değeri=0.662

kast2: ERKEK p-değeri=0.970, KADIN p-değeri=0.847

kast3: ERKEK p-değeri=0.777, KADIN p-değeri=0.818

kast4: ERKEK p-değeri=0.097, KADIN p-değeri=0.617

kast5: ERKEK p-değeri=0.763, KADIN p-değeri=0.721

kasttotal: ERKEK p-değeri=0.314, KADIN p-değeri=0.691

Varyans Homojenliği Testi Notları

kast1: Levene p-değeri=0.416

kast2: Levene p-değeri=0.309

kast3: Levene p-değeri=0.447

kast4: Levene p-değeri=0.452

kast5: Levene p-değeri=0.214

kasttotal: Levene p-değeri=0.232

Etki Büyüklüğü Notları

kast1: Cohen's d=2.68, Hedges' g=2.59

kast2: Cohen's d=2.56, Hedges' g=2.47

kast3: Cohen's d=2.65, Hedges' g=2.56

kast4: Cohen's d=2.69, Hedges' g=2.60

kast5: Cohen's d=2.78, Hedges' g=2.68

kasttotal: Cohen's d=2.81, Hedges' g=2.71

Grup Boyutları ve Yüzdeleri

Grup	Sayı (n)	Yüzde (%)
ERKEK	12	50.00
KADIN	12	50.00

Şekil 2. KİP programında gerçekleştirilen analizlerin word dosyasına aktarılması

Katılımcılar/Kullanıcılar

Bu çalışma kapsamında, geliştirilen istatistik aracının kullanım kolaylığı, işlevselliği ve doğruluğunu değerlendirmek amacıyla toplamda 6 lisansüstü öğrenci ve 25 akademisyen, programın kullanımına ilişkin yapılandırılmış görevleri tamamlamıştır. Katılımcıların tamamı, Statistical Package for the Social Sciences (SPSS) kullanımı konusunda deneyime sahip olup, KİP'te yer alan analizlerde de tecrübe sahibidirler. Bu durum, geliştirme sürecinin istatistiksel araçlara hâkim ve bilinçli bir kullanıcı kitlesi tarafından gerçekleştirilmesini sağlamıştır. Katılımcılar, programın farklı özelliklerini (parametrik ve non-parametrik analizler, grafikler, Word çıktısı, hız ve kullanım kolaylığı) değerlendirerek geri dönüşlerde bulunmuşlardır. Kullanıcı geri bildirimleri, programın nihai versiyonuna ulaşmadan önce toplanmış; katılımcılar, karşılaştıkları hataları ve geliştirilmesi gereken noktaları detaylı raporlar halinde iletmışlerdir. Bu raporlar, arayüzdeki iyileştirmelerden analiz fonksiyonlarının optimizasyonuna kadar geniş bir yelpazeyi kapsayarak, geliştirme sürecine yön vermiştir. Her katılımcı, programı kendi uzmanlık alanına uygun veri setleriyle test ederek geri bildirim sağlamıştır. Bu tür yazılım geliştirme süreçlerinde, kullanıcı geri bildirimleri genellikle iteratif bir yaklaşımla toplanır; hataların tespiti, önerilerin değerlendirilmesi ve uygulamaya yönelik adımların atılmasıyla nihai ürün şekillenir. KİP'in durumunda da bu yöntem izlenmiş, üç farklı aşamada alınan geri bildirimler doğrultusunda program optimize edilerek son hali verilmiştir. Bu yaklaşım, kullanıcı deneyimini değerlendirmede hataların giderilmesi ve işlevselliğin artırılması gibi kriterleri temel almıştır.

Veri Setleri

Kullanıcı geri dönüşleri süreci için her kullanıcının kendi alanında istediği veri setini analiz etmesine izin verilmiştir. Bu çalışmada yer alan raporlamalar için ise Central Michigan University'nin Matematik Bölümü tarafından sağlanan açık kaynaklı veri setleri kullanılmıştır (http://calcnet.mth.cmich.edu/org/spss/prjs_datasets.htm). SPSS eğitimine yönelik olarak oluşturulan bu veri setleri, farklı analiz türlerini test etmek ve sonuçların doğruluğunu değerlendirmek amacıyla seçilmiştir. Veri setlerinin seçimi, KİP'te yer alan tüm istatistiksel analizleri (parametrik ve non-parametrik testler, korelasyon, regresyon, varsayım kontrolleri ve grafik çıktıları) kapsamlı bir şekilde karşılamak ve programın çok

yönlülüğünü değerlendirmek için yapılmıştır. Çalışmada kullanılan veri setleri aşağıda yer almaktadır:

New Drug Data: Bir ilaç denemesine ait verilerden oluşan bu veri seti, iki grup arasında yapılan karşılaştırmalar için bağımsız örneklem t -test ve varyans analizi gibi analizlerin uygulanmasına olanak tanımıştır.

Body Fat Data: Vücut yağ oranlarının farklı bireyler arasındaki varyasyonlarını inceleyen bu veri seti, korelasyon ve regresyon analizlerinde kullanılmıştır.

Exam Data: Sınav sonuçlarına dayalı bu veri seti, akademik performansı ölçmek ve parametrik ile non-parametrik testlerin performansını karşılaştırmak için kullanılmıştır.

Disease Data: Farklı sektörel grupların ve sınıfların yer aldığı bu veri seti, özellikle ikiden fazla grup barındırdığı için varyans analizlerinde kullanılmak üzere tercih edilmiştir.

Bu veri setleri, KIP'in sunduğu istatistiksel yöntemlerin tamamını test edebilecek çeşitlilikte ve yapıdadır. Veri setlerinin her biri, istatistiksel varsayımların (örneğin, normallik testi, varyans homojenliği) kontrol edilmesine ve temel analiz yöntemlerinin doğruluğunun test edilmesine olanak tanımıştır. Ayrıca, programın sonuçları SPSS ile karşılaştırılarak, sonuçların tutarlılığı ve doğruluğu değerlendirilmiştir. Bu veri setleri, parametrik ve non-parametrik analizlerin yanı sıra tanımlayıcı istatistikler ve grafik çıktısı gibi özelliklerin test edilmesinde etkili olmuştur.

Veri Analizi Yöntemleri

Bu çalışmada, geliştirilen programın farklı istatistiksel testlerdeki performansı detaylı bir şekilde incelenmiş ve SPSS ile karşılaştırılmıştır. Her bir test türünde tanımlayıcı istatistikler (ortalama, standart sapma, medyan vb.), varsayımlar (normallik, varyans homojenliği, küresellik, çoklu doğrusal bağlantı vb.), test istatistikleri (t , F vb.) ve p -değerleri karşılaştırılmıştır. Aşağıda, her testin analizine ilişkin detaylar açıklanmıştır: Bağımsız iki grup arasındaki karşılaştırmalar için "Exam Data" veri seti kullanılmıştır. Bu veri setinde, kadın ve erkek katılımcıların sınav sonuçları karşılaştırılmıştır. Analizlerde standart bağımsız örneklem t -test, Welch t -test ve non-parametrik Mann-Whitney U testi uygulanmıştır. Eşleştirilmiş ölçümler için "New Drug Data" veri seti kullanılmıştır. Katılımcıların tedavi öncesi ve sonrası değerleri karşılaştırılmıştır. Bu analizde, standart

bağımlı örneklem *t*-test (paired *t*-test) ve non-parametrik Wilcoxon Signed-Rank testi uygulanmıştır. Üç veya daha fazla grup ortalamasını karşılaştırmak için “Disease Data” veri seti kullanılmıştır. Bu veri seti üzerinde tek yönlü ANOVA ve non-parametrik Kruskal-Wallis testi uygulanmıştır. Ayrıca, Tukey post-hoc testi ile gruplar arasında ikili karşılaştırmalar gerçekleştirilmiştir. Tukey testi, varyansların homojen olduğu ve grupların örneklem büyüklüklerinin dengeli olduğu durumlarda en uygun yöntemlerden biridir. Bu koşullar, tek yönlü ANOVA'nın varsayımlarına da uyduğundan, Tukey testi sonuçların daha tutarlı ve güvenilir olmasını sağlar. Tukey testi, ailewise hata oranını (family-wise error rate) kontrol altında tutarak birden fazla karşılaştırma sırasında tip I hata olasılığını minimize eder. Aynı katılımcıların farklı zaman noktalarındaki ölçümlerini değerlendirmek için “Body Fat Data” veri seti kullanılmıştır. Parametrik tekrarlı ölçümler ANOVA ve non-parametrik Friedman testi sonuçları karşılaştırılmıştır. Ayrıca Bonferroni düzeltmeli post-hoc analiz sonuçları SPSS ile uyumu incelenmiştir. Bonferroni düzeltmesi, post-hoc analizlerde en muhafazakâr yöntemlerden biridir. Çok sayıda karşılaştırma yapıldığında tip I hata riskini kontrol etmek için tercih edilir. Bu özellik, özellikle tekrarlı ölçümler ANOVA gibi çoklu karşılaştırmalar içeren durumlarda önemlidir.

İki-yönlü karma ANOVA için cinsiyet (kadın ve erkek) ile zaman noktaları (önce ve sonra) arasındaki etkileşimi değerlendirmek amacıyla New Drug Data veri seti kullanılmıştır. İki faktörlü (cinsiyet*zaman) analizler gerçekleştirilmiş ve parametrik sonuçlar değerlendirilmiştir. Korelasyon analizlerinde Body Fat Data veri seti kullanılmıştır. Pearson ve Spearman korelasyon katsayıları, ICC (Intraclass Correlation Coefficient) ve Cronbach Alpha değerleri karşılaştırılmıştır. Basit ve çoklu regresyon analizleri için Body Fat Data veri seti kullanılmıştır. Bağımlı değişken olarak *y* sütunu, bağımsız değişkenler olarak *x* sütunları kullanılmıştır. Regresyon analizlerinde, *R*-kare katsayıları, sabit (constant), Beta ve B değerleri karşılaştırılmıştır.

Bulgular

Bağımsız örneklem *t*-testi analizleri, “Exam Data” veri seti kullanılarak cinsiyet grupları arasındaki sınav puanlarının karşılaştırılması amacıyla gerçekleştirilmiştir (Tablo 1). Parametrik analizlerde, bağımsız örneklem *t*-testi uygulanmış ve ortalamalar, standart sapmalar, *t* değerleri ve *p* değerleri açısından geliştirilen program ile SPSS yazılımı arasında

tam bir uyum gözlenmiştir (Tablo 1.A.). Ayrıca, varyans homojenliği testi (Levene testi) sonuçları da her iki yazılım arasında birebir uyum göstermiştir. Non-parametrik analizlerde, aynı veri seti üzerinde Welch *t*-Testi (Tablo 1B.) ve Mann-Whitney U testi (Tablo 1C.) uygulanmıştır. Mann-Whitney U testinden elde edilen *p*-değeri 0.970 olarak hesaplanırken, SPSS yazılımında bu değer 0.960 olarak bulunmuştur (Tablo 1C.) Her iki yazılım arasında hesaplanan *p*-değerlerinde gözlenen bu küçük fark, sonuçların genel yorumlanabilirliğini etkilememiştir.

Bu küçük farkın olası nedeni, verilerdeki aynı değerlere sahip gözlemlerin (bağılıklar) test sırasında nasıl ele alındığı olabilir. SPSS, bu bağılıklar için özel bir düzeltme uygular ve bu durum *p*-değerinde küçük farklılıklar yaratabilir. Bunun yanı sıra, SPSS genellikle küçük örneklem büyüklüklerinde kesin (exact) *p* değerini hesaplarken, SciPy varsayılan olarak asimptotik bir yöntem kullanır. Kesin hesaplama, özellikle küçük örneklerde daha doğru sonuçlar ürettiği için bu yöntem farkı *p* değerinde hafif sapmalara neden olabilir. Bu farklılıklar, yazılımlar arasında varsayılan yöntemlerin farklı olmasından kaynaklanır ve genellikle analiz sonuçlarının genel yorumlanmasını etkilemeyecek kadar küçüktür.

Tablo 1A. Bağımsız örneklem *t*-test sonuçları

Değişken	Kadın (Ort ± SS)	Erkek (Ort ± SS)	t-İstatistiği	p-Değeri	Hedges' g
Puanlar	39.26 ± 7.84	40.5 ± 5.15	-0.605	0.549	0.19

Tablo 1B. Welch *t*-test sonuçları

Değişken	Kadın (Ort ± SS)	Erkek (Ort ± SS)	t-İstatistiği	p-Değeri	Hedges' g
Puanlar	39.26 ± 7.84	40.5 ± 5.15	-0.605	0.549	0.19

Tablo 1C. Non-parametrik Mann-Whitney U testi sonuçları

Değişken	Kadın (Median [ÇAÇ])	Erkek (Median [ÇAÇ])	p-Değeri
Puanlar	42.5 [12.0]	40.0 [6.5]	0.970

ÇAÇ: Çeyrekler arası açıklık

Bağımlı örneklem *t*-testi ve Wilcoxon Signed-Rank testi, “New Drug Data” veri seti kullanılarak tedavi öncesi ve sonrası kan basıncı değerlerinin (Before_exp_BP ve After_exp_BP) karşılaştırılması amacıyla uygulanmıştır (Tablo 2). Parametrik analizlerde bağımlı örneklem *t*-testi sonuçları, ortalamalar, standart sapmalar, *t* değerleri ve *p*

değerleri açısından geliştirilen program ile SPSS arasında tam bir uyum göstermiştir (Tablo 2A.). Non-parametrik analizlerde Wilcoxon Signed-Rank testi uygulanmış ve elde edilen p -değeri her iki yazılım arasında tamamen aynı bulunmuştur (Tablo 2B.).

Tablo 2A. Bağımlı örneklem t -test sonuçları

Değişken	Ön-test (Ort ± SS)	Son-test (Ort ± SS)	t-İstatistiği	p-Değeri	Hedges' g
Before_exp_BP	98.3 ± 5.17	88.6 ± 4.56	11.070	0.000	-1.55

Tablo 2B. Wilcoxon Signed-Rank test sonuçları

Değişken	Ön-test (Median [ÇAÇ])	Son-test (Median [ÇAÇ])	Wilcoxon W	p-Değeri
Before_exp_BP	97.7 [3.78]	88.15 [6.88]	0.0	0.000

ÇAÇ: Çeyrekler arası açıklık

Çok gruplu ANOVA analizi, "Disease Data" veri seti kullanılarak üç farklı sınıf (Class) arasında yaş değerlerinin (Age) karşılaştırılması amacıyla gerçekleştirilmiştir (Tablo 3). Parametrik analizlerde, tek yönlü ANOVA sonuçları ortalamalar, standart sapmalar, F değerleri ve p değerleri açısından geliştirilen program ile SPSS arasında tam uyum göstermiştir (Tablo 3A.). Non-parametrik analizlerde Kruskal-Wallis testi uygulanmış ve p -değerleri her iki yazılımda da aynı bulunmuştur (Tablo 3B.). Ayrıca, varyans homojenliği testi (Levene testi) sonuçları da her iki yazılım arasında birebir uyum göstermiştir.

Tablo 3A. Çok gruplu ANOVA sonuçları

Değişken	Grup-1	Grup-2	Grup-3	F-İstatistiği	p-Değeri
Class	27.78 ± 19.16	23.02 ± 16.18	23.83 ± 20.27	1.229	0.295

Tablo 3B. Kruskal-Wallis testi sonuçları

Değişken	Grup-1 (Median [ÇAÇ])	Grup-2 (Median [ÇAÇ])	Grup-3 (Median [ÇAÇ])	Kruskal-Wallis İstatistiği	p-Değeri
Class	25.0 [14.0-35.0]	18.0 [9.0-33.0]	17.0 [8.25-34.5]	3.224	0.200

ÇAÇ: Çeyrekler arası açıklık

Tekrarlı ölçümler ANOVA analizi, "Body Fat Data" veri seti kullanılarak farklı zaman noktalarındaki ölçümlerin karşılaştırılması amacıyla gerçekleştirilmiştir (Tablo 4). Parametrik analizlerde tekrarlı ölçümler ANOVA sonuçları, ortalamalar, standart sapmalar,

F istatistikleri, p değerleri ve kısmi eta kare değerleri açısından geliştirilen program ile SPSS arasında tam uyum göstermiştir (Tablo 4A). Küresellik varsayımını değerlendirmek için Mauchly Küresellik Testi uygulanmış ve analizde elde edilen *p*-değeri, Mauchly's *W* ve Approx. Chi-square değerleri her iki yazılımda da aynı bulunmuştur. Post-hoc analizlerde, Bonferroni düzeltmeli karşılaştırma sonuçları detaylı bir şekilde incelenmiş ve p değerlerinin SPSS ile tam bir uyum gösterdiği gözlenmiştir. Non-parametrik analizlerde ise Friedman testi uygulanmış ve *p*-değeri açısından tam bir tutarlılık sağlanmıştır (Tablo 4B.).

Tablo 4A. Tekrarlı ölçümlerde ANOVA sonuçları

Değişken	Zaman 1 (Ort ± SS)	Zaman 2 (Ort ± SS)	Zaman 3 (Ort ± SS)	F-İstatistiği	p-Değeri	Kısmi Eta Kare
Değişken 1	51.17 ± 5.23	27.62 ± 3.65	20.2 ± 5.11	401.979	0.000	0.955

Tablo 4B. Friedman testi sonuçları

Zaman 1 (Median [ÇAÇ])	Zaman 2 (Median [ÇAÇ])	Zaman 3 (Median [ÇAÇ])	Friedman İstatistiği	p-Değeri
52.0 [47.78-54.62]	27.9 [24.75-30.02]	21.2 [17.05-24.28]	38.1	0.0

ÇAÇ: Çeyrekler arası açıklık

İki yönlü karma ANOVA analizi, "New Drug Data" veri seti kullanılarak grupların (Gender) ve zaman noktalarının (Before_exp_BP ve After_exp_BP) ana etkilerinin yanı sıra grup*zaman etkileşimini saptamak amacıyla gerçekleştirilmiştir (Tablo 5). Parametrik analizlerde, ortalamalar ve standart sapmalar açısından geliştirilen program ile SPSS arasında tam uyum sağlanmıştır (Tablo 5A). Ayrıca varyans homojenliği testi sonuçları da her iki yazılımda birebir uyum göstermiştir. Analiz sonuçlarında, grupların ana etkisi ve grup*zaman etkileşim etkileri için elde edilen F istatistikleri, p değerleri ve kısmi eta kare değerleri, geliştirilen program ve SPSS arasında tam bir uyum göstermiştir (Tablo 5B). Ancak, zamanın ana etkisine ilişkin F istatistiği değerinde küçük bir fark gözlenmiştir. Örneğin, geliştirilen programda $F = 120.721$, SPSS'te ise $F = 120.070$ olarak hesaplanmıştır. Bu farkın olası nedeni, geliştirilen programda kullanılan Pingouin kütüphanesinin, zamanın ana etkisini hesaplarken grupları tek bir bütün olarak ele alması ve yalnızca zaman faktörüne odaklanmasıdır. Bu yaklaşım, gruplar arasındaki farklılıkları ve etkileşimleri doğrudan modele dahil etmeden, zamanın genel etkisini test etmeye odaklanır. Buna karşın, SPSS zamanın ana etkisini hesaplarken, grup*zaman etkileşiminden kaynaklanan varyansı da göz

önünde bulundurmaktadır. Ancak bu fark, p değerleri ve genel sonuçların yorumlanabilirliğini etkilemeyecek kadar küçük bir sapma olarak değerlendirilmiştir.

Tablo 5A. Karma ANOVA tanımlayıcı istatistikler

Gruplar / Zamanlar	Ön-test	Son-test
Kadın	98.01 ± 5.15	87.87 ± 4.62
Erkek	98.62 ± 5.28	89.39 ± 4.45

Tablo 5B. Karma ANOVA sonuçları

Etki	F-İstatistiği	p-Değeri	Kısmi Eta Kare
Time	120.721	0.000	0.716
Group	1.002	0.322	0.020
Interaction	0.272	0.605	0.006

Korelasyon analizi, "Body Fat Data" veri seti kullanılarak gerçekleştirilmiş ve veri setindeki dört değişkenin tamamı kullanılarak bir korelasyon matrisi oluşturulmuştur (Tablo 6). Pearson korelasyon analizinde tüm r ve p değerleri, geliştirilen program ile SPSS arasında tam uyum göstermiştir (Tablo 6A-6B). Aynı şekilde, Spearman korelasyon analizinde de sonuçlar birebir uyumlu bulunmuştur (Tablo 6C). Güvenilirlik analizinde hesaplanan Cronbach alpha değeri, her iki yazılımda da aynı sonuçları vermiştir. Sınıf içi korelasyon katsayıları (ICC) farklı modeller ve türler için hesaplanmıştır (Tablo 6D). ICC_{1k} modeli için hesaplanan değer, mutlak uyumu (absolute agreement) esas almıştır. Benzer şekilde ICC_{2k} modeli de mutlak uyumu değerlendirmiştir. ICC_{3k} modeli, tutarlılığı (consistency) temel almış ve yüksek bir ICC değeri ile dar bir güven aralığı sunmuştur. Bu sonuçlar, geliştirilen programın SPSS ile uyumlu çalıştığını ve farklı ICC türlerinde doğru hesaplamalar yaptığını ortaya koymaktadır.

Tablo 6A. Pearson korelasyon matrisi

	y1	x1	x2	x3
y1	1.000			
x1	0.924	1.000		
x2	0.458	0.085	1.000	
x3	0.843	0.878	0.142	1.000

Tablo 6B. Pearson korelasyon katsayıları ve *p*-değerleri

Değişken 1	Değişken 2	r (Korelasyon Katsayısı)	p-Değeri
y1	x1	0.924	0.000
y1	x2	0.458	0.042
y1	x3	0.843	0.000
x1	x2	0.085	0.723
x1	x3	0.878	0.000
x2	x3	0.142	0.549

Tablo 6C. Spearman korelasyon katsayıları ve *p*-değerleri

Değişken 1	Değişken 2	r (Spearman Katsayısı)	p-Değeri
y1	x1	0.896	0.000
y1	x2	0.398	0.082
y1	x3	0.795	0.000
x1	x2	0.059	0.806
x1	x3	0.849	0.000
x2	x3	0.139	0.560

Tablo 6D. ICC (Sınıf içi korelasyon) sonuçları

Model	ICC	Güven Aralığı (CI95%)	Tür (Agreement/Consistency)
ICC _{1k}	-2.092	-5.07 - -0.36	Absolute
ICC _{2k}	0.216	-0.0 - 0.52	Absolute
ICC _{3k}	0.854	0.71 - 0.94	Consistency

*Cronbach Alfa Sonuçları**Cronbach Alfa:* 0.854

Basit regresyon analizi, "Body Fat Data" veri seti kullanılarak gerçekleştirilmiştir (Tablo 7). Bu analizde, y1 değişkeni bağımlı, x1 değişkeni ise bağımsız değişken olarak tanımlanmıştır. Analiz sonuçlarında, geliştirilen program ile SPSS arasında tam uyum sağlandığı gözlenmiştir. Regresyon modeline ilişkin *R*-kare ve Adjusted *R*-kare değerleri her iki yazılımda da aynı bulunmuştur. Ayrıca, modelin parametrelerine ilişkin hesaplamalarda, hem sabit terim (constant) hem de bağımsız değişken için elde edilen B değerleri, t değerleri ve p değerleri tamamen uyumlu sonuçlar vermiştir.

Tablo 7. Basit regresyon analizi sonuçları
Regresyon katsayı tablosu

Değişken	B Değeri	Standart Hata	t-Değeri	p-Değeri	%95 Güven Aralığı (Alt)	%95 Güven Aralığı (Üst)	Beta Değeri
constant	-20.059	4.452	-4.506	0.000	-29.413	-10.706	-20.903
x1	0.887	0.087	10.240	0.000	0.705	1.068	0.924

Ek Bilgiler

R^2 : 0.853 (Modelin açıklayıcılığı)

*Adjusted R*²: 0.845 (Düzeltilmiş açıklayıcılık)

Regresyon Denklemi: $y = -20.059 + 0.000 * x$ (Model)

Çoklu regresyon analizi, Body Fat Data veri seti kullanılarak gerçekleştirilmiştir (Tablo 8). Bu analizde, y1 değişkeni bağımlı değişken olarak tanımlanırken, x1, x2 ve x3 değişkenleri bağımsız değişken olarak modellenmiştir. Regresyon modeline ilişkin R-kare değerleri, geliştirilen program ile SPSS arasında tam uyum göstermiştir. Modelin parametrelerine yönelik olarak sabit terim (constant) ve bağımsız değişkenler için hesaplanan B değerleri, Beta değerleri, t değerleri ve p değerleri her iki yazılımda da birebir aynı bulunmuştur.

Tablo 8. Çoklu regresyon analizi sonuçları
Regresyon Katsayıları

Değişken	Katsayı	Std Hata	t-Değeri	p-Değeri
Sabit	-32.327	0.713	-45.348	0.000
x1	0.833	0.018	46.833	0.000
x2	0.524	0.012	42.459	0.000
x3	0.026	0.018	1.437	0.170

Regresyon Detayları

R^2 (Belirleme Katsayısı): 0.999

Standardize Edilmiş Katsayılar (Beta):

Beta x1: 0.868, *Beta x2*: 0.380

Beta x3: 0.027, *Beta Sabit*: -32.327

Regresyon Denklemi

$Y = -32.327 + 0.833 * x1 + 0.524 * x2 + 0.026 * x3$

Tartışma ve Sonuç

Bu çalışmada geliştirilen KİP programının, SPSS gibi yaygın olarak kullanılan bir yazılımla istatistiksel analiz sonuçları açısından gösterdiği uyum, programın güvenilirliği ve geçerliliği konusunda güçlü kanıtlar sunmaktadır. Bulgular, hem parametrik (örneğin, bağımsız *t*-test, bağımlı örneklemeler *t*-test, ANOVA ve regresyon) ve hem de non-parametrik analizlerde (örneğin, Mann-Whitney U testi, Kruskal-Wallis testi ve Friedman testi) istatistiksel ölçütlerin (*t* değerleri, *F* istatistikleri, *p* değerleri ve etki büyüklükleri) neredeyse birebir uyum gösterdiğini ortaya koymuştur. Bu da çalışmada geliştirilen programın, istatistiksel hesaplama algoritmalarının doğruluğunu ve güvenilirliğini desteklemektedir. Parametrik ve non-parametrik testler arasındaki bu uyum, yalnızca analiz sonuçlarının güvenilirliğini değil, aynı zamanda geliştirilen programın farklı araştırma türlerine ve veri yapısına sahip çalışmalar için uygulanabilirliğini göstermektedir. Bu durum, uygulamalı veri işleme gerektiren birçok disiplin için geliştirilen programın potansiyelini öne çıkarmaktadır (Hopkins ve diğ., 2009; Ulupınar & İnce, 2021; Verma, 2016).

İstatistiksel analizlerde yaygın olarak kullanılan yazılımlar, kullanıcıların ihtiyaçlarına ve analiz türlerine göre farklı avantajlar ve sınırlamalar sunmaktadır. SPSS, özellikle akademik ve profesyonel araştırmalarda güvenilirliği ile öne çıkan bir araçtır; ancak geniş kapsamlı özellikleri nedeniyle öğrenilmesi zaman alabilmekte ve kullanıcı dostu bir deneyim sunma konusunda eleştirilere maruz kalmaktadır (Özdamar, 2018; Stehlik-Barry & Babinec, 2017; Verma, 2016). Öte yandan, R ve Python gibi açık kaynaklı yazılımlar, esneklik ve özelleştirilebilirlik sağlarken, bu yazılımların da kullanıcıların temel analizleri hızlıca gerçekleştirmesine yönelik kolaylıklar sunmada yetersiz kalabileceği belirtilmiştir (Ramachandran & Tsokos, 2020; Vallat, 2018). Jamovi ve JASP gibi yeni nesil istatistik yazılımları, kullanıcı dostu arayüzleriyle dikkat çekmekte ve akademik raporlama süreçlerini hızlandırmayı hedeflemektedir (Love ve diğ., 2019; Şahin & Aybek, 2019). Bu tür yeni nesil yazılımların yaklaşımını benimseyen KİP programı, hem kullanıcı dostu arayüzü hem de parametrik ve non-parametrik testlerde geniş bir analiz yelpazesi sunması ile bu yazılımlara bir alternatif oluşturabilecek niteliktedir. Türkçe dil desteği ve akademik standartlara uygun Word çıktıları sağlayan bu program, pratik uygulamalara odaklanan disiplinler için önemli bir avantaj sunmaktadır. Ayrıca, SPSS gibi ücretli yazılımların aksine, açık kaynaklı araçların erişilebilirliği ile pratikliği birleştiren bu yaklaşım, lisansüstü

öğrenciler ve istatistik deneyimi sınırlı olan kullanıcılar için güçlü bir seçenek olarak değerlendirilebilir (Lino Calle ve diğ., 2024; Şahin & Aybek, 2019).

Bu çalışmada geliştirilen program, parametrik ve non-parametrik testlerde yüksek bir doğruluk göstermiştir. Bağımsız örneklemeler *t*-test, bağımlı örneklemeler *t*-test ve ANOVA gibi parametrik analizlerde, ortalamalar, standart sapmalar, *t* ve *F* değerleri ile *p* değerleri açısından tam bir uyum sağlanmıştır. Aynı şekilde, Mann-Whitney U testi, Kruskal-Wallis testi ve Friedman testi gibi non-parametrik analizlerde de programın sonuçları büyük ölçüde SPSS ile uyumlu bulunmuştur. Ancak, birkaç analizde küçük farklılıklar gözlemlenmiştir. Mann-Whitney U testi analizinde, elde edilen *p* değerinde program ve SPSS arasında küçük bir fark tespit edilmiştir (örneğin, $p = 0.970$ vs. $p = 0.960$). Bu farkın nedeni, SPSS'in aynı değerlere sahip gözlemleri test sırasında farklı bir yöntemle ele alması ve bu gözlemlere yönelik özel düzeltmeler uygulamasıdır (Bressert, 2012; Virtanen ve diğ., 2020). Buna karşın, Python'un SciPy kütüphanesi varsayılan olarak asimptotik bir yaklaşım kullanır ve kesin (exact) *p* değeri hesaplama yerine bu yöntemi benimser. Bu yöntem farklılığı, özellikle küçük örneklem büyüklüklerinde hafif sapmalara neden olabilir ancak bu sapmalar genel analiz sonuçlarını etkileyecek boyutta değildir (REF1). Ayrıca bu yaklaşımların birinin diğerine göre kesin olarak üstünlüğü konusunda bir netlik bulunmamaktadır.

Karma ANOVA analizinde ise zamanın ana etkisi hesaplamasında küçük bir farklılık gözlemlenmiştir. SPSS'te bu hesaplama, grup*zaman etkileşiminden kaynaklanan varyansı da modele dahil ederek yapılırken, geliştirilen programın kullandığı Pingouin kütüphanesi zamanın ana etkisini hesaplarken grupları tek bir bütün olarak ele alır ve yalnızca zaman faktörüne odaklanır (Hill ve diğ., 2024; Vallat, 2018). Bu farklılık, *F* istatistiğinde küçük sapmalara yol açabilmektedir (örneğin, KİP'te $F = 120.721$; SPSS'te $F = 120.070$). Ancak, bu farkın *p* değerleri üzerinde hiçbir etkisi olmamış ve analiz sonuçlarının genel yorumlanabilirliğini değiştirmemiştir. Genel olarak, bu küçük farklılıklara rağmen program, parametrik ve non-parametrik testlerde yüksek bir tutarlılık göstermiştir. Özellikle normallik testleri (Shapiro-Wilk testi), varyans homojenliği (Levene testi) ve küresellik (Mauchly testi) gibi varsayımların değerlendirilmesinde sağlanan uyum, programın güvenilir bir analiz aracı olduğunu desteklemektedir. Bonferroni ve Tukey gibi post-hoc analizlerin doğru ve tutarlı bir şekilde uygulanması, programın daha karmaşık analizlerde de başarılı bir performans sergilediğini ortaya koymaktadır. Ayrıca, korelasyon analizlerinde Pearson ve

Spearman katsayılarının yanı sıra sınıf içi korelasyon katsayılarının (ICC) ve Cronbach Alpha güvenilirlik değerlerinin SPSS ile birebir uyum göstermesi, programın güvenilirlik ve doğruluk açısından güçlü bir performans sergilediğini ortaya koymaktadır. Benzer şekilde, regresyon analizlerinde, R-kare ve Adjusted R-kare değerleri ile sabit terim ve bağımsız değişkenlere ilişkin B, Beta, t ve p değerlerinin birebir uyumlu olması, programın modelleme süreçlerinde de başarılı bir araç olduğunu kanıtlamaktadır.

Geliştirilen program, kullanıcı dostu özellikleriyle istatistiksel analiz süreçlerini hem teknik hem de pratik açıdan daha erişilebilir hale getirmiştir. Türkçe dil desteği sunması, özellikle SPSS gibi yabancı dil ağırlıklı yazılımları kullanmakta zorluk çeken kullanıcılar için önemli bir avantajdır. Kullanıcı arayüzünün basit ve anlaşılır olması, lisansüstü öğrencilerin ve temel analizleri gerçekleştiren araştırmacıların istatistiksel süreçlerde hız ve kolaylık kazanmasını sağlamaktadır (Love ve diğ., 2019; Şahin & Aybek, 2019). Bu durum, özellikle uygulamalı alanlarda, karmaşık yazılımlar yerine pratik çözümlere olan ihtiyacı karşılamaktadır. Programın en dikkat çekici avantajlarından biri, tüm varsayımları (örneğin, normallik, varyans homojenliği, küresellik) ve parametrik ile non-parametrik analizleri aynı anda sunabilmesidir. Bu özellik, kullanıcıların analiz süreçlerini daha kapsamlı bir şekilde yürütmelerine olanak tanımaktadır. Özellikle, pratik öneme vurgu yapan etki büyüklüklerinin (ör. Cohen'in d, Hedge's g, eta kare) hesaplanması, programın güçlü yönlerinden biridir. SPSS, son sürümlerine kadar bu tür etki büyüklüklerini otomatik olarak hesaplamazken, geliştirilen programın bu işlevi kullanıcılara önemli bir kolaylık sağlamaktadır (Hopkins ve diğ., 2009; Ulupınar & İnce, 2021).

Programın Word formatında doğrudan akademik raporlama standartlarına uygun tablo çıktıları sunması, kullanıcılar için büyük bir zaman tasarrufu yaratmaktadır. Çıktılar, virgülden sonraki rakamları sınırlı tutarak net ve düzenli bir görünüm sağlamaktadır. SPSS gibi yazılımlarda tabloların genellikle manuel olarak düzenlenmesi gerektiği göz önüne alındığında, bu özellik analiz sonuçlarının doğrudan raporlara entegre edilmesini kolaylaştırmaktadır (Gotelli & Ulrich, 2012; Stehlik-Barry & Babinec, 2017). Grafik çıktısı özelliklerinin kullanıcı dostu yapısı da programın işlevselliğini artırmaktadır. Programın bir diğer önemli avantajı, kullanıcı odaklı öğretici bir yaklaşımı benimsemesidir. Her analiz sonunda, hedef kitleye (lisansüstü öğrenciler ve temel düzeyde bilgi sahibi araştırmacılar) yönelik rehber bilgiler sunulmakta ve bu bilgiler analiz sonuçlarının nasıl yorumlanacağına

dair yol gösterici olmaktadır. Bu, programın yalnızca bir analiz aracı olarak değil, aynı zamanda eğitim sürecine katkı sağlayan bir araç olarak da değer taşıdığını göstermektedir. Ayrıca, teknik yapısı itibarıyla programın hafif ve kompakt bir tasarıma sahip olması, sistem kaynaklarını zorlamadan çalışmasına olanak tanımaktadır. SPSS gibi daha kapsamlı yazılımlar bilgisayarda büyük bir depolama alanı gerektirirken, geliştirilen program 130 MB depolama alanı ile önemli bir avantaj sunmaktadır. SPSS'in son sürümleri ise bilgisayarda genellikle 3 GB veya daha fazla depolama alanı gerektirmektedir. Sonuç olarak, bu teknik ve işlevsel avantajlar, programın yalnızca analizlerde değil, aynı zamanda kullanıcı deneyimi açısından da güçlü bir alternatif sunduğunu göstermektedir. Özellikle kullanıcı dostu tasarımı, akademik raporlamaya uygun çıktıları ve öğretici rehberlik özelliğiyle, programın istatistik bilgisi sınırlı olan bireyler için etkili bir çözüm sunduğu söylenebilir.

Sınırlılıklar

Bu çalışmada geliştirilen program, temel istatistiksel analizlerde yüksek doğruluk ve kullanıcı dostu özellikler sunmasına rağmen, bazı sınırlamalara sahiptir. İlk olarak, programın sonuçlarının güvenilirliği analizlerde doğrulanmış olsa da, herkesçe kabul edilen bir dijital imza veya sertifika taşımamaktadır. Bu durum, özellikle kurumsal kullanıcılar için güvenilirlik algısını sınırlayabilir. Gelecekte, programın uluslararası standartlarda bir dijital imzaya sahip olması, bu sınırlamayı giderebilir. Program şu anda yalnızca Windows işletim sistemi için geliştirilmiştir. Ancak, Linux ve MacOS kullanıcılarının da erişim sağlaması için özelleştirilmiş versiyonlara ihtiyaç duyulmaktadır. Bu durum, programın erişilebilirliğini sınırlamakta ve farklı platformlardaki kullanıcılar için bir engel oluşturmaktadır. Gelecekteki çalışmalar, platformlar arası bir uyumluluk sağlayarak bu eksikliği giderebilir. Dil desteği açısından program yalnızca Türkçe olarak tasarlanmıştır. Türkçe dil desteği, yerel kullanıcılar için büyük bir avantaj sağlasa da, uluslararası araştırmacılar için bir sınırlama oluşturmaktadır. Çoklu dil desteğinin eklenmesi, programın küresel ölçekte kullanılabilirliğini artıracak önemli bir geliştirme alanıdır.

Grafiksel analizler açısından program, temel düzeyde grafikler sunmasına rağmen, grafik çeşitliliği sınırlıdır. Örneğin, etkileşimli görseller, zaman serisi grafikleri ve diğer ileri düzey grafik türleri programa henüz entegre edilmemiştir. Kullanıcıların analiz sonuçlarını daha etkili bir şekilde görselleştirmesi için grafik seçeneklerinin artırılması faydalı olacaktır. Bunun yanı sıra, programın çoklu dil desteği bulunmaması ve yalnızca Türkçe olarak

çalışması, uluslararası kullanıcılar için önemli bir kısıtlama yaratmaktadır. Ayrıca, Mac ve Linux işletim sistemlerinde çalışmaması, bu platformları kullanan araştırmacılar için erişimi sınırlandırmaktadır. Programın gelecekte çoklu dil desteği ve çapraz platform uyumluluğu ile güncellenmesi, daha geniş bir kullanıcı kitlesine ulaşmasını sağlayacaktır. Son olarak, programın mobil uygulaması bulunmamaktadır. Mobil uyumluluk, özellikle saha çalışmaları yapan araştırmacılar ve hızlı analiz ihtiyacı duyan kullanıcılar için önemli bir özellik olabilirdi. Gelecekte, programın hem Android hem de iOS platformları için optimize edilmiş bir mobil uygulamasının geliştirilmesi, daha geniş bir kullanıcı kitlesine ulaşmasını sağlayabilir. Bu sınırlamalar, programın mevcut işlevselliğini gölgelemese de, gelecekteki geliştirmeler için önemli bir yol haritası sunmaktadır. Program, bu eksikliklerin giderilmesiyle daha kapsamlı ve erişilebilir bir analiz aracı haline gelebilir.

Bilgilendirme

Bu çalışmada insan veya hayvan deneklerinden veri toplanmamıştır. Bu nedenle çalışma, etik kurul onayı gerektiren çalışmalar kapsamında yer almadığından etik kurul onayı alınmamıştır.

Erişim Linki

[Kolaylaştırılmış İstatistik Programı \(KİP\)](#) erişimine ulaşmak için tıklayabilirsiniz.

Yazar Katkı Beyanı

Süleyman ULUPINAR: *Literatür taraması, program geliştirme, kavramsallaştırma, uygulama, verilerin toplanması, işlenmesi, analizi, yorumlanması, inceleme-yazma, düzenleme.*

Serhat ÖZBAY: *Literatür taraması, program geliştirme, kavramsallaştırma, uygulama, verilerin toplanması.*

Fadime ULUPINAR: *Literatür taraması, verilerin toplanması, işlenmesi, analizi.*

Selim ASAN: *Verilerin işlenmesi, analizi, yorumlanması, inceleme-yazma, düzenleme.*

İzzet İNCE: *Verilerin işlenmesi, analizi, yorumlanması, inceleme-yazma.*

Deniz BEDİR: *Verilerin toplanması, işlenmesi, analizi, yorumlanması, inceleme-yazma, düzenleme.*

Cebrail GENÇOĞLU: *Literatür taraması, kavramsallaştırma, verilerin toplanması, işlenmesi, analizi, yorumlanması*

Salih ÇABUK: *Literatür taraması, kavramsallaştırma, inceleme-yazma, düzenleme.*

Kaynaklar

- Akbulut, Ö. (2022). Bilimsel arařtırmalarda istatistiksel anlamlılığın raporlanmasında güncel yaklaşımlar: Hatalar ve dođrular. *International Journal of Eastern Mediterranean Agricultural Research*, 5(1), 1-19.
- Alpar, R. (2014). *Uygulamalı istatistik ve geđerlilik-güvenirlilik: SPSS'de çözümleme adımları ile birlikte* (3. baskı). Detay Yayıncılık.
- Bademci, V. (2022). Correcting fallacies about validity as the most fundamental concept in educational and psychological measurement. *International e-Journal of Educational Studies*, 6 (12), 148-154. <https://doi.org/10.31458/iejes.1140672>
- Bressert, E. (2012). *Scipy and numpy: An overview for developers*. O'Reilly Media.
- Fagerland, M. W. (2012). t-tests, non-parametric tests, and large studies—a paradox of statistical practice? *BMC Medical Research Methodology*, 12, 1-7. <https://doi.org/10.1186/1471-2288-12-78>
- Gotelli, N. J., & Ulrich, W. (2012). Statistical challenges in null model analysis. *Oikos*, 121(2), 171-180. <https://doi.org/10.1111/j.1600-0706.2011.19731.x>
- Hill, C., Du, L., Johnson, M., & McCullough, B. (2024). Comparing programming languages for data analytics: Accuracy of estimation in Python and R. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, e1531. <https://doi.org/10.1002/widm.1531>
- Hopkins, W., Marshall, S., Batterham, A., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise*, 41(1), 3-13. <https://doi.org/10.1249/MSS.0b013e31818cb278>
- Jain, P., & Sengar, S. (2024). Unraveling the role of IBM SPSS: A comprehensive examination of usage patterns, perceived benefits, and challenges in research practice. *Educational Administration: Theory and Practice*, 30(5), 9523-9530. <https://doi.org/10.53555/kuvey.v30i5.4609>
- Lino Calle, V. A., Carvajal Rivadeneira, D. D., Sornoza Parrales, D., Vergara Ibarra, J. L., & Intriago Delgado, Y. M. (2024). Jamovi, the technological tool for analyzing and interpreting data in civil engineering projects. *Revista Innovaciones Educativas*, 26(41), 151-165. <http://ddoi.org/10.22458/ie.v26i41.5145>
- Loerts, H., Lowie, W., & Seton, B. (2020). *Essential statistics for applied linguistics: Using R or JASP*. Bloomsbury Publishing.
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., ... Epskamp, S. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88(1), 1-17. <https://doi.org/10.18637/jss.v088.i02>
- Özbay, S., Ulupınar, S., Gençođlu, C., Ouergui, I., Öget, F., Yılmaz, H. H., ... Uçan, İ. (2024). Effects of Ramadan intermittent fasting on performance, physiological responses, and bioenergetic pathway contributions during repeated sprint exercise. *Frontiers in Nutrition*, 11, 1322128. <https://doi.org/10.3389/fnut.2024.1322128>

- Ramachandran, K. M., & Tsokos, C. P. (2020). *Mathematical statistics with applications in R*. Academic Press.
- Shepherd, M. A., & Richardson, E. J. (2024). Opting for open-source? A review of free statistical software programs. *Teaching Statistics*, 46(1), 53-63. <https://doi.org/10.1111/test.12360>
- Stehlik-Barry, K., & Babinec, A. J. (2017). *Data analysis with IBM SPSS statistics*. Packt Publishing.
- Şahin, M., & Aybek, E. (2019). Jamovi: An easy to use statistical software for the social scientists. *International Journal of Assessment Tools in Education*, 6(4), 670-692. <https://doi.org/10.21449/ijate.661803>
- Ulupınar, S. (2022). Atletik performans ölçümlerinde test-tekrar test güvenilirliği analizleri. *İstanbul Gelişim Üniversitesi Sosyal Bilimler Dergisi*, 9(2), 738-747. <https://doi.org/10.17336/igusbd.809612>
- Ulupınar, S., & İnce, İ. (2021). Spor bilimlerinde etki büyüklüğü ve alternatif istatistik yaklaşımları. *Spormetre Beden Eğitimi ve Spor Bilimleri Dergisi*, 19(1), 1-17. <https://doi.org/10.33689/spormetre.794015>
- Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, 3(31), 1026. <https://doi.org/10.21105/joss.01026>
- Verma, J. (2016). *Sports research with analytical solution using SPSS*. John Wiley & Sons.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & Bright, J. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261-272. <https://doi.org/10.1038/s41592-019-0686-2>