



R Programlama Diline Giriş

Ozancan Özdemir

Doktora Adayı, Groningen Üniversitesi

Ben Kimim? (Adettendir...)



- **Lisans:** ODTÜ İstatistik Bölümü, 2017
 - **Yüksek Lisans:** ODTÜ İstatistik Bölümü, 2020
Tez Başlığı: Performance Comparison of Machine Learning Methods and Traditional Time Series Methods for Forecasting. (Danışman: Prof.Dr.Ceylan Yozgatlıgil)
 - **Doktora:** University of Groningen, Bernoulli Institute, Devam ediyor
 - **Yarıda bırakılan Doktora 🤔:** ODTÜ İstatistik Bölümü, 2023
-
- **Araştırma Alanları:** Yapay ve Derin Öğrenme, Zaman Serileri, Ardışık Veri, Veri Görselleştirme, Seçim Tahmin Modelleri, R programlama, Açık Veri, Veri Bazlı Politika
-
- Araştırma Görevlisi: ODTÜ İstatistik Bölümü 2017-2023
 - Kurucu Ortak: VeriPie, 2020-2022
 - Doktora Adayı: University of Groningen, Bernoulli Institute, 2023-...

Ders İçeriği

- R ve R Studio Nedir?
- Temel fonksiyonlar
- Objeler
- Swirl paketi, R içinde R öğren.
- Kontrol Yapıları
- Döngüler
- Kendi Fonksiyonlarımızı Yazma
- Veri Okutma/Aktarma/Manipülasyonu
- Tanımlayıcı (Betimleyici) İstatistik Hesaplamaları
- Veri Görselleştirme (ggplot2 ve base)
- apply ailesi(?)

R Programlama Dili Tarihçesi



- R, S dili geliştirildikten bir süre sonra kullanıma girdi. S dilinin önemli bir kısıtlılığı, yalnızca ticari bir paket olan S-PLUS içinde mevcut olmasıydı.
- **R, 1991 yılında Auckland Üniversitesi İstatistik Bölümü'nde Ross Ihaka ve Robert Gentleman tarafından yaratıldı.**
- R'nin ilk kamuoyu duyurusu 1993 yılında yapıldı.
 - i. Ross Ihaka ve Robert Gentleman'ın R geliştirme deneyimleri, 1996 yılında Computational and Graphical Statistics dergisinde yayımlanan bir makalede belgelenmiştir.
- **Önemli Katkı - GNU Genel Kamu Lisansı: 1995 yılında Martin Mächler, R'ı ücretsiz yazılım yapmak için GNU Genel Kamu Lisansı'nı kullanmaları konusunda Ross ve Robert'ı ikna etti.**
- **R'nin Kamuoyuna Sunumu: 29 Şubat 2000 tarihinde R sürüm 1.0.0 kamuoyuna sunuldu.**

R Programlama Dili Nedir?

- R, son derece güçlü bir istatistiksel, grafik ve programlama aracıdır.
- R, kullanıcının algoritmaları programlamasına ve başkaları tarafından programlanmış araçları kullanmasına izin veren bir bilgisayar dilidir.
- R hayal ettiğiniz bir çok şeyi gerçekleştirmenize yarayan bir araçtır.
- Ücretsiz, açık kaynaklı ve sürekli büyüyen bir yazılımdır.
- İstatistikçiler, biyologlar, genetikçiler, ziraatçiler, ekonometristler ve veri bilimi ile ilgilenen daha birçok meslek grubu tarafından kullanılan interdisipliner bir programdır.

R Programlama Dili Nedir?

- R, koşullar, döngüler, kullanıcı tanımlı özyinelemeli işlevler ve giriş ve çıkış olanaklarını içeren iyi geliştirilmiş, basit ve etkili bir programlama dilidir.
- R etkin bir veri işleme ve depolama olanağına sahiptir.
- R diziler, listeler, vektörler ve matrisler üzerinde hesaplamalar için bir dizi operatör sağlar.
- R, veri analizi için geniş, tutarlı ve entegre bir araç koleksiyonu sağlar.
- R, veri analizi için grafiksel olanaklar sağlar veya doğrudan bilgisayarda ya da kağıtlara yazdırılır.
- Çoğu üniversite ve iş, R bilmenizi gerektirir veya en azından tercih eder.

Günümüzde R Programlama Dili



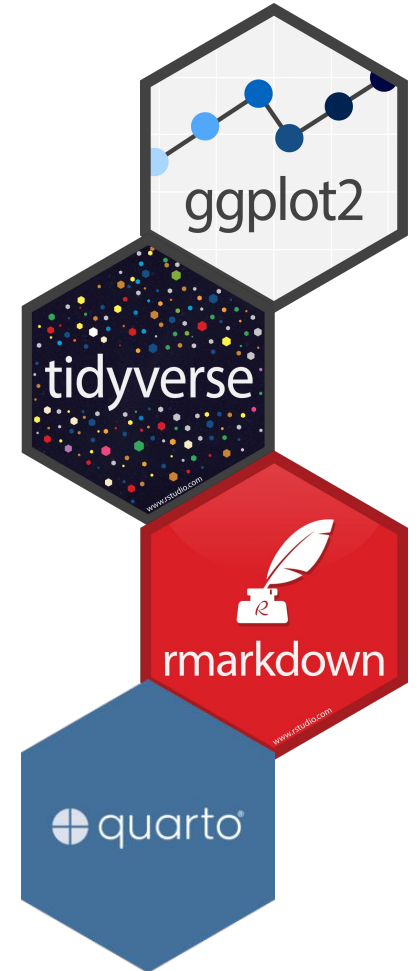
IEEE Spectrum 2023
Spectrum Sıralaması: IEEE kullanıcılarının genel ilgi alanı

| Companies That Use R | Who Uses R at This Company? | What Does This Company Use R For? |
|----------------------|--|---|
| Amazon | Senior data scientist, research scientist II, transportation analyst | Statistical analysis, exploratory data analysis, machine learning |
| Cognizant | Data engineer, data scientist, data analyst | Exploratory data analysis, statistical analysis |
| Deloitte Consulting | Data visualization developer, data scientist, backend developer | Machine learning, exploratory data analysis |
| Google | Data scientist, cloud AI engineer, analytical consultant | Exploratory data analysis, machine learning |
| Meta Platforms | Data science manager, people data solutions analyst, data Scientist | Statistical analysis, exploratory data analysis |
| Novartis | Senior principal real-world evidence (RWE) research analyst, senior data scientist, exploratory data analyst | Exploratory data analysis, machine learning, statistical analysis |
| Snapdeal | Product analyst, data scientist | Machine learning, statistical analysis |

CareerKarma

Kolaylıklar ve Avantajlar

- › Bedava! Bütçe dostu çözüm sağlaması
- › Zengin bir paket ekosistemine sahip olması.(CRAN ve GitHub)
- › İstatistiksel analizler için güçlü bir platform olması.
- › Veri görselleştirme paketleri; ggplot2, highcharts, plotly
- › Veri işleme (Data Handling; tidyverse paketi)
- › İşlenebilir veri tabanları sunar ve paketleri sunar (tibble, data-frame)
- › Topluluk Desteği, Güçlü bir R topluluğu.
 - X'de #rstats etiketi, Posit Community.
- › Tekrar üretilebilir sonuçları bir arada göstermeye yarayan paket ve araçlara sahip olması (R Markdown, Quarto)



Kolaylıklar ve Avantajlar

- › Parallel hesaplama yapabilmesi
- › Tekrarlanabilir araştırma imkanı sunması
- › Alan bazlı kapsamlı paketler sunması; bioconductor
- › Esneklik ve Genişlenebilirlik; C, python, Java kullanımı.
- › Akademik ve endüstriyel araştırmalarda yaygın olarak kullanılması
- › Birden fazla Grafiksel Kullanıcı Arayüzüne (Graphical User Interface) uygun olması; R Studio, Jupyter notebook, R Commander gibi
- › Zengin öğrenme materyali: bookdown (<https://bookdown.org/home/archive/>)
- › R Shiny 🥰



Bir Ek: Öğrenme Materyalleri



Zorluklar ve Dezavantajlar

- › Kolay unutulabilir olması
- › Dik öğrenme eğrisi
- › Yüksek hesaplama gücü gerektiren işlerde tam performans sağlayamaması
- › Bellek yönetimi yetersizliği
- › Kesişen paket ve fonksiyonların olması



LEARN PROGRAMMING IN JUST
2 WEEKS AND MAKE THOUSANDS OF
DOLLARS, BECOME GENIUS SOFTWARE
ENGINEER WITH MINIMAL EFFORT
AND DEVELOP YOUR OWN VIDEO
GAMES!!!



5 years later



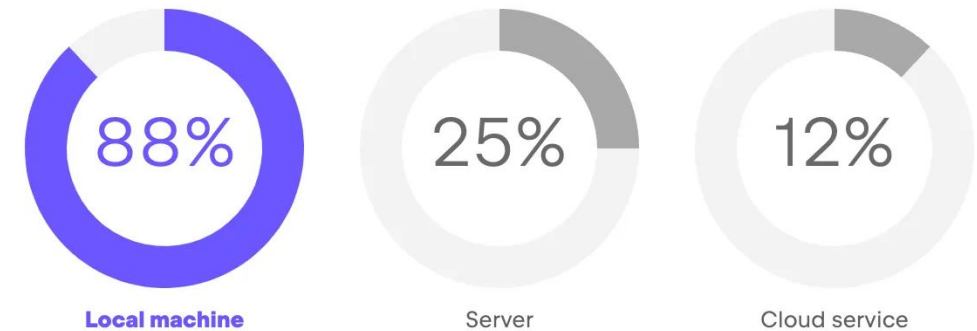
i dont remember how to
make an array



Zorluklar ve Dezavantajlar

- › Tam anlaşılır olmayan hata mesajları
- › Sınırlı nesne yönelimli programlama (Object Oriented Programming) desteği
- › Veri güvenliği
- › Paketlerin teorik olarak kontrol edilmemesi
- › Lokal olarak kullanımın ilk tercih olması
- › İntertet ya da Web uygulaması gibi kullanamama

Where do you execute R code?



The vast majority of users continue to execute R code on their local machines. The proportion of respondents using servers (25%) remains unchanged from last year.

JetBrains' State of the Developer Ecosystem



Ne zaman R kullanmalıyım?

Dış ses;

- › “Eğer yapacağınız işlem, veri temizleme ve düzenleme, veri görselleştirme , istatistiksel metot ve modellemeleri kapsayan (yapay öğrenme metotları), bir proje ise; veri miktarı R’ın belleğini zorlamayacak kapasitede ise R kullanabilirim.”
- › “İnteraktif, modellerin sonuçlarının gösterildiği, canlı veri ile güncellenen bir arayüz yapmak istersem R kullanabilirim.”

İç ses;

- › “Her zaman R kullanabilirim.” :)



İletişim



o.ozdemir@rug.nl



<https://ozancanozdemir.github.io>



<https://twitter.com/OzancanOzdemir>



<https://github.com/ozancanozdemir>



[www.linkedin.com/in/ozancan-özdemir-40b94768](https://www.linkedin.com/in/ozancan-ozdemir-40b94768)



Teşekkürler...

