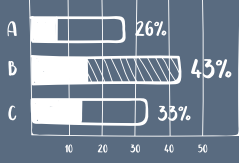


# SİVİL TOPLUM ÖRGÜTLERİ İÇİN TEMEL VERİ EĞİTİMİ KILAVUZU



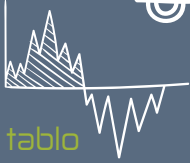
veri



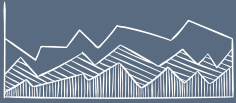
analiz



75%



tablo



rapor



grafik



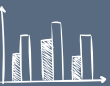
START



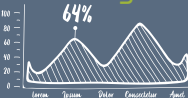
veri görselleştirme



formül



gözlem



anket



ESHİD

Eşit Haklar

İçin

İzleme Derneği

# SIVİL TOPLUM ÖRGÜTLERİ İÇİN TEMEL VERİ EĞİTİMİ KILAVUZU

Hazırlayan  
Pınar Dağ



“Bu yayın, The Fondation Internationale pour les D fenseurs des Droits de L’Homme (the FIDDH)’un maddi desteęiyle hazırlanmıř ve basılmıřtır. İerik tamamıyla Eřit Haklar İin İzleme Derneęi’nin sorumluluęu altındadır ve the FIDDH’in g r řlerini yansıtmaq zorunda deęildir.”



Eřit Haklar İin İzleme Derneęi  
G m řsuyu Mah. Aęa ıraęı Sok. No:7 Pamir Apt. D:1  
Beyoęlu/İstanbul  
Tel: 0212 293 63 77  
www.esihaklar.org  
email: esihaklar@gmail.com  
info@esihaklar.org

Hazırlayan: **Pınar Daę**  
Redaksiyon: **Cennet  zc mert Albayrak, Melih  zel, Nejat Tařtan**  
Kapak ve Sayfa Tasarımı: **Mehmet Ali elik**

Basım Yeri: Ceylan Matbaası- Tuncay Mat  
İSTANBUL - 2021

# İçindekiler

- 6 Şekil Listesi
- 9 Tablo Listesi
- 9 Kısaltmalar
  
- 11 Giriş
- 11 STÖ'ler için Verinin Önemi

## 1.BÖLÜM

- 12 Verileri Anlamlandırmak
- 12 Üst Veriyi Keşfedin (Metadata)
- 13 Veri Biyografisini İnceleyin
- 13 Veri Sözlüğü

## 2.BÖLÜM

- 15 Elektronik Tablo Becerilerini Güçlendirmek
- 16 Elektronik Tablodan ODS ya da CSV olarak Veri İndirmek
- 16 Elektronik Tablonun Kopyasını Oluşturmak
- 17 Elektronik Tabloyu Paylaşmak
- 17 Elektronik Tabloyu Kişilerle ve Gruplarla Paylaşmak
- 18 Elektronik Tablolara Veri Yükleme ve Veriyi Dönüştürmek
- 19 Elektronik Tablolardaki Coğrafi Kodlar
- 21 Verileri Sıralamak ve Filtrelemek
- 22 Formüllerle Hesaplama Yapmak
- 24 Verileri Pivot Tablolara Özetlemek

## 3. BÖLÜM

- 27 Veri Toplamının Önemi
- 28 Hayvan Hakları Odaklı Veri Toplama Örneği
- 32 Veri Toplama Yöntemleri
- 32 Gözlem
- 32 Görüşme
- 33 Anket
- 34 Ölçekler (Rakamsal Olarak Kodlama)
- 35 Bilgi Edinme Başvurusu ile Veri Toplamak
- 36 Kendi Verisini Toplayan STÖ Örnekleri
- 37 DataMiner ile Veri Toplamak
- 40 Elektronik Tablo ile HTML Sayfadan Veri Çekmek



- 41 Tableau Public ile PDF'den Veri Kazımak
- 43 Parseur ile Veri Toplama
- 44 Tabula ile PDF'den Veri Kazımak
- 45 Workbench Data ile Veri Toplama
- 47 DataBasic.io ile Veri Toplama
- 48 Github Hesabı Oluşturmanın Önemi

## 4. BÖLÜM

- 50 Veri Kalitesi ve Veri Doğrulama
- 50 Veri Seti Doğrulama
- 51 E-Tablo ile Veri Doğrulama

## 5. BÖLÜM

- 55 Veri Temizlemek ve Düzenlemek
- 56 Elektronik Tablo ile Veri Temizleme
- 59 Open Refine ile Dağınık Veri Temizleme
- 62 WorkBench Data ile Veri Temizleme
- 64 CSV Fingerprint ile Verideki Hataları Tespit Etme
- 66 CSV Link(?) ile Verilerdeki Hataları Tespit Etme
- 68 Breve ile Verideki Hataları Tespit Etme
- 69 Good Tables ile Verideki Hataları Tespit Etme

## 6. BÖLÜM

- 70 Veri Analizi
- 70 Veriyi Analiz Etmeye Nereden Başlanmalı?
- 70 Temel Ölçümler (İstatistik)
- 70 Veri Analizi Ölçüleri
- 77 Veri Analizi ve Hesaplama Araçları

## 7. BÖLÜM

- 79 Veri Çalışma Panosu Oluşturmak
- 81 Solo/Ekip Proje Taslağı
- 82 Araç Seçilirken Dikkate Alınacak Bazı Faktörler
- 83 Araç Öneri Tablosu

## 8. BÖLÜM

- 85 Veri Görselleştirme Sivil Toplum Örgütleri için Neden Önemlidir?
- 86 Keşfetmek ve Açıklamak İçin Görselleştiririz
- 87 Veri Görselleştirmenin Tanımı
- 89 Veri Görselleştirmenin Önemi
- 90 Veri Görselleştirme Tipleri

- 90 Veri Görselleştirme Süreci
- 90 Veri Görselleştirmenin Niteliğini Etkileyen Unsurlar
- 91 Veri Görselleştirme Yöntemleri Periyodik Tablosu
- 91 Veri Görselleştirme Prensipleri /Kuralları
- 93 Veri Görselleştirmeyi Geliştirme Yöntemleri
- 94 Veri Görselleştirme Elementleri
- 94 Veri Nasıl Sunulur?
- 95 Türkçe Veri Görselleştirme Kataloğunun Kullanımı
- 96 İyi Veri Görselleştirme Örnekleri
- 98 Kötü Veri Görselleştirme Örnekleri
- 99 Veri Görselleştirme Araçları
- 99 E-Tablo İle Veri Görselleştirme
- 102 Flourish İle Veri Görselleştirme
- 104 Tableau Public İle Veri Görselleştirme
- 112 Datawrapper ile Veri Görselleştirme

## **9. BÖLÜM**

- 115 Açık Veri ve Veri Paylaşımı Yöntemleri
- 131 Dünyadan Bazı Veri Kolektifleri
- 132 Bazı Uluslararası Veri Kaynakları
- 132 Bazı Ulusal Veri Kaynakları
- 132 Bazı Sivil Veri Kaynakları

## **10. BÖLÜM**

- 133 Kaynak
- 133 Temel Veri Eğitimi Çevrim İçi Kaynaklar
- 133 Temel Veri Eğitimi Basılı Kaynaklar
- 135 Kapanış

## Şekil Listesi

- 12 **Şekil 1** : 361 sağlık çalışanı ile yapılan pandemi anketi veri seti. (Kasım-Aralık-2020)
- 13 **Şekil 2** : TÜİK'in "Güvenlik Birimine Gelen veya Getirilen Çocuk İstatistikleri, 2015-2019" meta veri bölümünün ekran görüntüsüdür. Veri Setlerinin meta verisi detayları ile buradan incelenebilir. (<https://data.tuik.gov.tr/Bulten/Index?p=Güvenlik-Birimine-Gelen-veya-Getirilen-Cocuk-Istatistikleri-2015-2019-33632>)
- 13 **Şekil 3** : Veri Biyografisi Şablonu ([https://docs.google.com/spreadsheets/d/1bp-LWDFzTggkOXsnn6jefc6XoR55DRsQ\\_rqYmXqtG9Z8/edit#gid=0](https://docs.google.com/spreadsheets/d/1bp-LWDFzTggkOXsnn6jefc6XoR55DRsQ_rqYmXqtG9Z8/edit#gid=0))
- 14 **Şekil 4** : Ulusal Sağlık Veri Sözlüğü (USVS)
- 15 **Şekil 5** : Başlıklar, formülü görüntüleyen aktif hücre, birden çok sekmelerin görüntüsünün olduğu e-tablo ekran görüntüsüdür.
- 15 **Şekil 6** : veri.xlsx, veri.ods, veri.csv. Bilgisayarınızda yaygın olarak görülen üç veri biçimi xlsx, ods ve csv.
- 16 **Şekil 7** : Google E-Tablolar'da verilerinizi en yaygın dosya tiplerini seçerek dışa aktarabilirsiniz: Dosya>İndirme>Belge türü.
- 16 **Şekil 8** : Google e-tablonun kopyasını oluşturma. Dosya > Kopya oluşturma.
- 17 **Şekil 9** : Google e-tablo veri seti "Paylaş" düğmesi ekran görüntüsü.
- 17 **Şekil 10** : "Paylaş" düğmesine bastıktan sonra veri setini kişiler ve gruplarla paylaşma seçeneklerinin olduğu ekran görüntüsüdür.
- 17 **Şekil 11** : E-posta ekleyerek veri setinin paylaşılması ekran görüntüsüdür.
- 18 **Şekil 12** : Veri setini, linke sahip /internete sahip herkesle paylaşıp üzerinde düzenleme yapmaya izin verme ekran görüntüsüdür.
- 18 **Şekil 13** : Drive'ın sağ üst köşesinde yer alan dişli sembolünün ekran görüntüsüdür.
- 18 **Şekil 14** : Excel dosyasının yüklenmesi ve e-tabloya dönüşmesi ile Microsoft Excel, Virgülle Ayrılmış Değerler (CSV) dosya tipinin ekran görüntüsü yer almaktadır.
- 19 **Şekil 15** : Google E-Tablo Add-ons>Get add-ons ekran görüntüsüdür.
- 19 **Şekil 16** : Geocoding by SmartMonkey Install /eklenmesinin ekran görüntüsüdür.
- 19 **Şekil 17** : Eklenti bölümüne Geocoding by SmartMonkey'in eklendiğini ve sayfanın da açıldığını ekran görüntüsüdür.
- 20 **Şekil 18** : Geocode details ekran görüntüsüdür.
- 20 **Şekil 19** : Geocode details sonrası örnek veri setinin ekran görüntüsüdür.
- 20 **Şekil 20** : Ankara il sınırı haritası Google Map ekran görüntüsüdür. Enlem ve boylam koordinatlarını içermektedir.
- 21 **Şekil 21** : Veri setinin sıralamadan önce seçildiği ekran görüntüsüdür.
- 21 **Şekil 22** : Seçilen veri setinin veri bölümünden sıralama aralığı seçme ekran görüntüsüdür.
- 21 **Şekil 23** : Sıralama Aralığı seçildikten sonra hangi aralığın sıralandığını gösteren ekran görüntüsüdür. Bu görüntüde Koronavirüs'e yakalanmış ya da yakalanmamış olan sağlık çalışanlarının anket yanıtı sıralanmıştır.

- 
- 22 **Şekil 24** : Veri filtreleme ekran görüntüsüdür.
- 22 **Şekil 25** : Sağlık çalışanları ile yapılmış olan pandemi anketinde virüse yakalanmış olan kişileri “evet” yanıtına göre filtrelenmesinin ekran görüntüsüdür.
- 23 **Şekil 26** : Fonksiyonlar bölümünden ya da sütunu seçerek ya da formülü elle yazarak sütunda yer alan verileri toplama formülü ekran görüntüsüdür.
- 23 **Şekil 27** : Fonksiyonlar bölümünde ya da sütunu seçerek ya da formülü elle yazarak sütunda yer alan verilerin ortalama formülü ekran görüntüsüdür.
- 23 **Şekil 28** : Google E-tablolar fonksiyon listesi ekran görüntüsüdür.
- 24 **Şekil 29** : Sağlık Çalışanları Anket verisine “Pivot Table” fonksiyonunu uygulama ekran görüntüsüdür.
- 24 **Şekil 30** : Pivot Tablo oluşturma ve yeni sayfa açmak ekran görüntüsüdür.
- 25 **Şekil 31** : Veriyi özetlemek için açılan Pivot table sayfası ekran görüntüsüdür. Solda satır, sütun ve değerler, sağda ise pivot editörünün önerdiği satır, sütun ve değerler sekmelerinin yer aldığı ekran görüntüsüdür.
- 25 **Şekil 32** : Bölgelerine göre uzman hekimlerin sayılarının toplamı ekran görüntüsüdür.
- 26 **Şekil 33** : Bölgelerine göre uzman hekimlerin toplamlarının sütun grafik karşılaştırması ekran görüntüsüdür.
- 27 **Şekil 34** : Veri ile çalışma akış tablosu.
- 28 **Şekil 35** : New York Belediyesi’nin hayvanlara yönelik istismar verisini topladığı açık veri platformunun ekran görüntüsüdür. (<https://nycopendata.socrata.com/Social-Services/Animal-Abuse/vqss-uwk2>)
- 29 **Şekil 36** : <http://dadaviz.com/s/animal-abuse/> tarafından kategorisine göre hayvan istismarı grafiği.
- 29 **Şekil 37** : Google Alerts kullanılarak medyada yapılan hayvan haberleri için belirlenmiş kelimelerin ekran görüntüsüdür.
- 30 **Şekil 38** : Google Uyarı aracılığı ile e-mail adresine düşen hayvan haberleri ekran görüntüsü.
- 30 **Şekil 39** : Google e-tablo eklenti bölümüne eklenen Save Emails and Attachments’in ekran görüntüsüdür.
- 31 **Şekil 40** : Google Uyarılar ile toplanan hayvan haberleri verisinin eklenti ile e-tabloya aktarılmış olan ekran görüntüsüdür.
- 31 **Şekil 41** : Raporların linklerinin aylık olarak yapılandırıp gruplandığı verilerin ekran görüntüsüdür.
- 33 **Şekil 42** : Google Forms ekran görüntüsüdür.
- 35 **Şekil 43** : [cimer.gov.tr](http://cimer.gov.tr) ekran görüntüsüdür.
- 38 **Şekil 44** : TBMM 27. dönem milletvekilleri listesi kazıma ekran görüntüsüdür.
- 38 **Şekil 45** : TBMM 27. dönem milletvekilleri listesi kazıma ekranına aktarılmış veri görüntüsüdür.
- 39 **Şekil 46** : Kazınan veriler CSV, XLSX olarak indirilebilir ya da kopyalanarak e-tabloya ya da başka bir tabloya eklenebilir.
- 39 **Şekil 47** : TBMM sitesinden “Data Miner” ile kazınan verilerin e-tabloya aktarılarak, görselleştirilmesi sürecinin ekran görüntüsüdür.

- 40 **Şekil 48** : E-tabloda import fonksiyonu ile çekilecek tablo, liste, veri, feed vb. ekran görüntüsüdür.
- 40 **Şekil 49** : IMPORTHTML ile e-tabloya aktarılmış dünyada en çok satan kitap listesi ekran görüntüsüdür.
- 50 **Şekil 50** : DataONE Veri Yaşam Döngüsü.
- 56 **Şekil 51** : Şehirler sütununda yer alan boşluklar veri, boşlukları kaldırma ya da kırp ekran görüntüsüdür.
- 56 **Şekil 52** : İl isimlerinin tamamını büyük yapmayı sağlayan üst fonksiyon ekran görüntüsüdür.
- 57 **Şekil 53** : İl isimlerinin tamamını küçük yapmayı sağlayan alt fonksiyon ekran görüntüsüdür.
- 57 **Şekil 54** : İl isimlerinin tamamının ilk harfini büyük yapmayı sağlayan yazım düzeni fonksiyonu ekran görüntüsüdür.
- 58 **Şekil 55** : Hatalı, eksik, bitişik yazılmış olan bazı bölge isimlerinin “bul ve değiştir” fonksiyonu ile düzenlenme ekran görüntüsüdür.
- 58 **Şekil 56** : Bir sütunda yer alan isim ve adres verisini iki ayrı sütuna bölmeyi sağlayan fonksiyonun ekran görüntüsüdür.
- 59 **Şekil 57** : Dağınık veri temizleme aracı Open Refine’in işletim sistemlerine göre sürüm ekran görüntüsüdür.
- 61 **Şekil 58** : OpenRefine ile yapılan veri temizleme işlemlerinin fonksiyon ekran görüntüsüdür.
- 79 **Şekil 59** : Veri Çalışma Panosu: veri ile çalışırken çalışmanızı kolaylaştıran data storyboard taslak görüntüsüdür.
- 80 **Şekil 60** : Veri Çalışma Panosu ile yapılan hazırlık sonrası yine çevrim dışı olarak hazırlanmış bir veri görselleştirme panosu. / Fotoğraf: Pınar Dağ
- 81 **Şekil 61** : Örnek solo/ekip proje dosyası.
- 86 **Şekil 62** : Maarten Lambrechts’in sunumun ekran görüntüsüdür.
- 87 **Şekil 63** : <https://sarkac.org/2020/12/istanbulda-haftalik-vefat-sayilari/>
- 88 **Şekil 64** : (M.Ö 950-Avrupa) Bilinen ilk nicel gösterge - gezegenlerin ve ayların değişen konumunu gösteren bir çoklu zaman serisi grafiği.
- 88 **Şekil 65** : Harita (M.Ö. 6500) bilinen en eski harita-Çatalhöyük/Konya Müzesi-Türkiye

---

## Tablo Listesi

84 **Tablo 1** : Tablo: Veri ile çalışırken ihtiyaca göre kullanılacak araç önerileri. Bu araçların dışında da araçlar kullanılabilir. (Hazırlayan: Pınar Dağ)

## Kısaltmalar

**CSV** Virgülle Ayrılmış Değerler

**HDX** İnsani Yardım Veri Değişimi

**M.Ö.** Milattan Önce

**OCHA** Birleşmiş Milletler İnsani Yardım Koordinasyon Ofisi

**ODS** OpenDocument Elektronik Tablosu

**STÖ** Sivil Toplum Örgütü

**TBMM** Türkiye Büyük Millet Meclisi

**TDWI** The Data Warehousing Institute

**TÜİK** Türkiye İstatistik Kurumu

**USVS** Ulusal Sağlık Veri Sözlüğü

**VOYD** Veri Okuryazarlığı Derneği



# Giriş

"Sivil Toplum Örgütleri için Temel Veri Eğitimi Kılavuzu" örgütlerin veri odaklı çalışmalarını yürütürken daha sistematik şekilde ilerlemelerini sağlamak için hazırlanmıştır.<sup>1</sup> 10 bölümden oluşan Kılavuz, verinin Sivil Toplum Örgütleri (STÖ) için neden önemli olduğunu, veriyi anlamlandırmak için yapılması gereken ön hazırlıkları ve temel veri becerilerini geliştirmek için tasarlanmıştır. Bu amaçla Kılavuz'da veri toplama, veri doğrulama, veri temizleme, veri analizi, veri görselleştirme, açık veri ve sivil toplum ilişkisi başlıkları yer almaktadır.

## STÖ'ler için Verinin Önemi

21. yüzyılda hızla artan teknolojik ilerleme ve dijitalleşme günümüzde veri kullanımının önemini her geçen gün arttırmaktadır. Dijitalleşen içeriğin kullanılması verilerin daha iyi yönetim şartlarına ulaşması, dünyamızı daha iyi anlamak için ortaya konulan büyük gayretler olmakla beraber bu gayretlerin ötesine geçmek gerekmektedir. Çünkü veri tek başına böyle bir sorumluluk üstlenemez. Bunun için veri yönetimini ileriye taşıyacak değişimler, gelişmeler ve paradigma değişiklikleri şarttır.

Günümüzde, verinin nasıl olması gerektiğine dair birçok fikir/yöntem ortaya konmuşsa da hayatlarımızı nasıl etkilemesi gerektiğine dair gerçekçi bir yanıt hala verilmiş değildir. Nitekim [Tableau Public](#) tarafından her yıl hazırlanan veri eğilimleri raporunda, genel veri eğilimlerinin sürekli yer değiştirdiğini görüyoruz.

---

<sup>1</sup> Temel seviyede veriyle çalışmaya nereden başlanması gerektiğine yanıt vermeye çalışan online kılavuz çevrimiçi olarak ayrıca yayınlanacaktır.



# 1. BÖLÜM

Kılavuzun bu ilk bölümünde verinin anlamı, verinin ne olduğu ve temel veri becerilerinin neleri kapsadığı yer almaktadır. Özellikle veri türlerini, dosya tiplerini, temel bazı fonksiyonları anlamamanın, veriyle çalışmaya başlamada nasıl hızlı bir çalışma akışı sağladığını göreceğiz.

## Verileri Anlamlandırmak

Veri setleri ile çalışmadan önce veriyi anlamalıyız. Verinin bize hangi yanıtları verebileceğini, hangilerini veremeyeceğini tespit etmeliyiz. Örneğin, sağlık çalışanlarına<sup>2</sup> yönelik yapılan anketin verilerini analize hazırlarken, sağlık çalışanlarının hasta olma riskinin özel hastanede mi, devlet hastanesinde mi daha yüksek olduğunu yanıtını arıyorsunuz; ya da hastane doluluk oranlarının %100 dolu olduğunu söyleyen 200 sağlık çalışanını; Covid-19 pozitif hastalığını geçirenleri ya da yeterince tedbir alınmadığını belirtenlerle ilişkilendirmek istiyorsunuz. Bu ve benzeri soruların yanıtları için öncelikle veriyi anlamamız gerekmektedir.

	A	B	E	F	G	H	I	J	K	L
1	Timestamp	Kaç yaşındasınız	Virüslü ailenize bulaşma korkusu	Çalıştığınız sağlık kuruluşu	Göreviniz nedir? (Hekim Branşınız nedir?)	Koronavirüs	Yukarıdaki soruya cevabınız	Çalıştığınız hastanenin	Hasta kapı	
2	11/22/2020 18:08:02 -65 +	Fazlasıyla	Özel	Hekim	Kadın-doğum	Hayır	Yoğun bakım 2. Basamağı	Evet, çok o		
3	11/22/2020 18:27:57 -65 +	Hiç	Özel	Hekim	Çocuk	Hayır	Yüzde yüz	Hayır, hiç o		
4	11/28/2020 15:21:04 -65 +	Fazlasıyla	Özel	Hekim	Genel cerrahi	Evet	Kasım	100	Evet, çok o	
5	11/28/2020 15:24:27 -65 +	Biraz	Özel	Uzman hekim	Anesteziyoloji	Hayır			Hayır, hiç o	
6	11/29/2020 15:28:09 -65 +	Fazlasıyla	Özel	Hekim	Pediyatri	Hayır				
7	11/21/2020 18:17:20 -65 -65	Fazlasıyla	Özel	Hekim	Ortopedi ve travmatoloji	Hayır				
8	11/21/2020 18:40:22 -65 -65	Fazlasıyla	Özel	Hekim	Pediyatri	Evet	Ekim kasım	Şu anda %100	Hayır, hiç o	
9	11/21/2020 21:51:25 -65 -65	Biraz	Özel	Hekim	İç	Hayır		%50	Ordu ama c	

Şekil 1: 361 sağlık çalışanı ile yapılan pandemi anketi veri seti. (Kasım-Aralık-2020)

Pratikte genellikle veriyi anlamadan, veri analizine ve veri görselleştirmeye geçilmektedir. Bu yaklaşım nitelikli veri görselleştirmesi yapmanın önünde önemli bir engel teşkil etmektedir. Veri, doğru anlamlandırılmadığında kalitesi de maalesef belirsizleşir. Verinin dağınık olup olmadığını ya da verinin ne kadar doğru ve etkin bir şekilde toplandığını da anlayamayız. Bu nedenle veri setindeki veriyi iyi anlamak için üç yöntem incelenerek işe başlanmalıdır: Üst Veri ([Metadata](#)), Veri Biyografisi ([Data Biography](#)), Veri Sözlüğü ([Data Dictionary](#))

## Üst Veriyi Keşfedin (Metadata)

Üst veriler, veriler hakkındaki temel bilgilerdir. Yeni bir veri kümesini indirmeden önce üst verileri gözden geçirmek zaman kazandırır. İndirip kullanacağınız veride nelerin mevcut olduğunu ya da nelerin mevcut olmadığını anlamınıza yardımcı olur. Bir üst veri; veri hakkında aşağıda yer alan temel bilgileri sağlar:

<sup>2</sup> Bu kılavuzda örnek olarak kullanılan Sağlık Çalışanları Veri Seti, kılavuzun online versiyonu ile birlikte kullanıma sunulacaktır.

**Konum:** Veri setinin coğrafi kapsamı ve ayrıntı düzeyi hakkında bilgi sağlar.

**Metodoloji:** Verilerin nasıl toplandığına yönelik bilgileri açıklar.

**Kaynak:** Veri setini kimin yayınladığı hakkında bilgi sağlar.

**Tarih:** Veri setinin yayımlandığı ilk tarihi, güncellendiği son tarih, yayımlandığı tarih ve bir sonraki tahmini yayım tarihi için beklenen güncelleme sıklığını verir.

Metaveri	
Analitik Çerçeve, Kapsam, Tanımlar ve Sınıflamalar	▼
Verinin kapsamı	▼
Hesaplama Kuralları	▼
Temel Veri Kaynaklarının Niteliği	▼
Derleme uygulamaları	▼
Revizyonlar	▼
Diğer konular	▼

**Şekil 2:** TÜİK'in "Güvenlik Birimine Gelen veya Getirilen Çocuk İstatistikleri, 2015-2019" meta veri bölümünün ekran görüntüsüdür. Veri setlerinin meta verisi detayları ile buradan incelenebilir. (<https://data.tuik.gov.tr/Bulten/Index?p=Güvenlik-Birimine-Gelen-veya-Getirilen-Cocuk-İstatistikleri-2015-2019-33632>)

## Veri Biyografisini İnceleyin

Veri girişi yapmak oldukça zaman alan bir faaliyettir. Veri toplamada/girişinde, hem teknik hem de donanımsal sorunlar oluşabilir. Verilerin girişinde yaşanan bir tutarsızlık, veriyi yeniden doğrulama çalışması gerektirir. Yanlış girişi engellemek için yeni bir çalışma yürütmek gerekir. Kurumlar bu tür aksaklıkları engellemek adına veri biyografisi oluşturarak konuları belirgin şekilde açıklar, amacı belirler ve kaynakları için veri biyografisi oluştururlar. Ham veri, anket verileri, röportaj verileri biyografiye eklenerek verilere bir kimlik kazandırılır. Biyografileri incelemek bu sebeple önemlidir.

Veri Setinin Adı	Veri Seti Kaynağının Linki	Depolama Kaynağının Linki	Veriyi Kim Topladı	Verinin Sahibi Kim	Veri Nasıl Toplandı	Örnek Boyut	Kim Dahil Edildi
İnşaat Verileri 2020	<a href="http://www.tuik.gov.tr/Etiler/Tablo">http://www.tuik.gov.tr/Etiler/Tablo</a>	<a href="http://www.tuik.gov.tr/Etiler/Tablo">http://www.tuik.gov.tr/Etiler/Tablo</a>	Saim GÜRBUZ	TÜİK	1974-2011 yılları arasında İçişleri Bakanlığı Emniyet Genel Müdürlüğü ve Jandarma Genel Komutanlığı, 2012 yılından itibaren İçişleri Bakanlığı Emniyet Genel Müdürlüğü, Jandarma Genel Komutanlığı, Adalet Bakanlığı		

**Şekil 3:** Veri Biyografisi Şablonu ([https://docs.google.com/spreadsheets/d/1bpLWDFzTggkOXsnn6jefc6XoR55D-RsQ\\_rqYmXqtG9Z8/edit#gid=0](https://docs.google.com/spreadsheets/d/1bpLWDFzTggkOXsnn6jefc6XoR55D-RsQ_rqYmXqtG9Z8/edit#gid=0))

## Veri Sözlüğü

Bazı resmi ve resmi olmayan kurumlar, verilerle birlikte veri sözlüğü de yayımlar. Örneğin; aşağıda Sağlık Bakanlığı'na ait "[Ulusal Sağlık Veri Sözlüğü \(USVS\) 2.2](#)" sürümünde yer alan Hasta Kabul Veri Seti bilgileri gösterilmektedir. Veri sözlüğü, her bir değişkenin neyi temsil ettiğini anlamak için veri setindeki değişkenler için tanımlar barındırır, veri elemanlarının anlaşılmasına katkı sağlar.

Aşağıdaki örnek, Hasta Kabul Veri Seti vaka türü ve fazlası için hasta ile ilgili verilen bir tanım örneğinin yanı sıra veri tipini sayısal olarak da sağlamaktadır.

# ULUSAL SAĞLIK VERİ SÖZLÜĞÜ

## Sürüm 2.2



T.C. Sağlık Bakanlığı

<b>HASTA KABUL VERİ SETİ</b> .....	119
Veri Elemanları.....	119
Veri Elemanları Tanım ve İş Kuralları.....	119
Diğer Veri Setleri ile İlişkili Alanlar.....	121
HL7 Bilgileri.....	122
Veri Seti.....	122
Veri Korum.....	122
Veri Elemanları.....	122

<b>HASTA KABUL VERİ SETİ</b>						
Oluplama Tarihi	:	01.06.2007				
Sürüm Tarihi	:	01.06.2007				
Kaynak Kurum / Birim	:	Türkiye Kamu Hastaneleri Kurumu Türkiye Halk Sağlığı Kurumu Sağlık Hizmetleri Genel Müdürlüğü				
Kapsam	:	Kabul; sağlık kurum, kuruluş ve aile hekimi tarafından hastanın tedavi ve bakım sorumluluğundan üstlenilmesi olayını ifade eder. Kabulden sonra hastane ya da aile hekimi tarafından hastanın ihtiyaç duyduğu sağlık hizmetinin verilmesine başlanır. Bu veri seti, hastanın sağlık kurumuna veya aile hekimine kabul işlemlerine ait verileri içerir.				
Bağlam	:	Hastanın sağlık hizmeti almaya başlaması sırasında alınan bilgiler, verilen sağlık hizmetinin, kurullar arası ve kurum içi seviyelerin analizinde ve hizmet planlamasında kullanılan önemli veriler içermektedir.				
Bildirimi Yapacak Kurumlar	:	1.2.3. Basamak Sağlık Kurumları tarafından gönderilir.				
Bildirimi Yapılacak Paketler	:	Muayene Gönderim Paketi ile gönderilir.				
<b>Veri Elemanları</b>						
Veri Seti Elemanları	Tekrar	Durum	Veri Tipi	Alan B.	Format	SKRS Sistem Kodu
Kabul Zamanı*	1	Zorunlu	Tarih Saat	16	GGAAYYYY SS-DD	-
Kabul Şekli*	1	Zorunlu	Nümerik	2	N(2)	d100a73b-199b-4c6f-a03b-b175a22ba04
Sevki Tanımı*	1*	Koşullu	Alfanümerik	8	A(8)	SKRS İCD 10 Tablosu
Sevki Tarihi*	1	Koşullu	Tarih Saat	16	GGAAYYYY SS-DD	-
Geldiği (Poly)Klinik*	1	Koşullu	Alfanümerik	6	A(6)	e049ec57-e5d1-443d-e049-760af0146a3d
Vaka Türü	1	Zorunlu	Nümerik	2	N(2)	46380e82-4881-407d-9554-255d915d0959
Adli Vaka Geliş Şekli	1	Koşullu	Nümerik	1	N(1)	08ce62cb-14ac-42e7-9143-ed3771416d4f
Ambulans ile Geliş Durumu	1	Seçimli	Nümerik	1	N(1)	e5d9e3d6-b6d3-4846-9115-406a706b7e43
SGK Takip Numarası	1	Seçimli	Alfanümerik	8	A(8)	-
Protokol Numarası	1	Koşullu	Alfanümerik	50	A(50)	-

\* Veri Elemanları veri bütünlüğü için Tablolu ve Durum bilgileri için HL7 Bilgileri - Veri Korum bilgileri.

**Şekil 4:** Ulusal Sağlık Veri Sözlüğü (USVS), <https://e-saglik.gov.tr/TR.7141/usvs.html> ve <https://dosyamerkez.saglik.gov.tr/Eklenti/3703,usvs2220140507pdf.pdf?0>

## 2. BÖLÜM

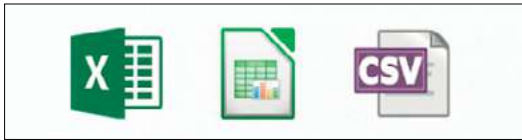
### Elektronik Tablo Becerilerini Güçlendirmek

Veri görselleştirmeye başlamadan önce e-tablo kullanma becerisini geliştirmek önemlidir. Günlük yaşantımızda işimizin önemli bir parçası olan e-tablolar, bütünüyle öğrenilmese de bizlere sadece zaman kazandırmakla kalmaz; ayrıca verileri çok daha iyi anlamamızı sağlar. Temel Veri Becerileri Listesi, kılavuz ekinde yer almaktadır.

**Şekil 5:** Başlıklar, formülü görüntüleyen aktif hücre, birden çok sekmelerin görüntüsünün olduğu e-tablo ekran görüntüsüdür.

Başlıklar				
İsim	İl	Tecrübe	Görev	Kurum
Zeynep	İzmir	30	Hekim	Özel
Yağmur	Ankara	45	Hekim	Özel
Pınar	İstanbul	32	Hekim	Devlet
Güler	Erzincan	44	Hekim	Devlet
Reyhan	Antalya	35	Hekim	Özel
Peruze	Balıkesir	12	Uzman hekim	Özel
Adalet	Gümüşhane	48	Hekim	Özel
Türkan	Sivas	33	Uzman hekim	Devlet
Güzella	Trabzon	12	Uzman hekim	Devlet
Selina	Eskişehir	29	Hekim	Devlet
		320		

Bu bölümde e-tablolarla çalışırken veri setini paylaşma, yükleme, veri çekme, indirme, temizleme, veri girişi, pivot ile özetleme, sıralama, filtreleme, dört işlem gibi temel işlemler hakkında bilgiler yer almaktadır. Elektronik tablolar verileri farklı şekilde düzenler, yapılandırır. E-tablo verilerini bilgisayara indirdiğimizde bir isim verip, ardından bir nokta ekleyip görselde görüldüğü gibi veri formatını temsil eden uzantıyı görmek mümkündür.



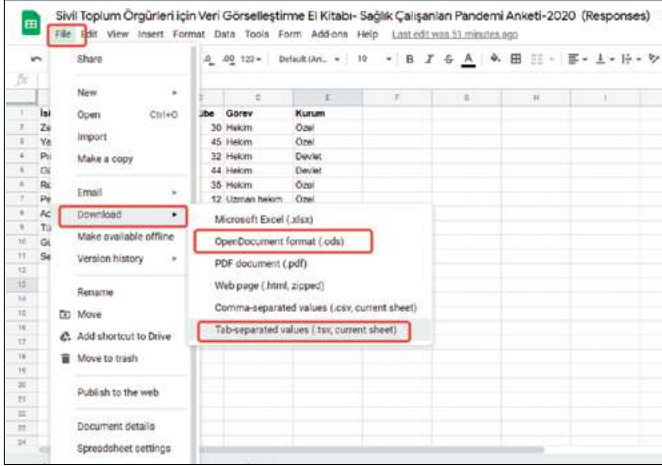
**Şekil 6:** Veri.xlsx, veri.ods, veri.csv. Bilgisayarımızda yaygın olarak görülen üç veri biçimi xlsx, ods ve csv.

- **.xlsx:** Verileri düzenlemek, analiz etmek ve depolamak için Microsoft Excel'de kullanılan dosya türüdür. Her hücre, matematiksel formüller dâhil olmak üzere metin veya sayısal veriler içerebilir.
- **.ods:** Açıklaması “Open Document Spreadsheet” yani çok sekmeli sayfaları, formülleri vb. kaydeden standartlaştırılmış bir açık format dosya türüdür.
- **.csv:** Microsoft Excel veya OpenOffice Calc gibi elektronik tablo programları tarafından yaygın olarak kullanılan virgülle ayrılmış değerler dosyasıdır.

Veri görselleştirme projelerinde sıklıkla yukarıda yer alan dosya tipleri kullanılmaktadır.

## Elektronik Tablodan ODS ya da CSV olarak Veri İndirmek

E-tablolar diğer platformlarla geçiş yaparak çalışmak Virgülle Ayrılmış Değerler (CSV) ve OpenDocument Elektronik Tablosu (ODS) gibi genel veya açık veri dosya formatlarında veri setini dışa aktarmamıza izin verir. Sırasıyla Dosya> İndir> CSV, Dosya> İndir> ODS adımlarını takip ederek veri istenilen dosya tiplerinde indirilebilir.



**Şekil 7:** Google E-Tablolar'da verilerinizi en yaygın dosya tiplerini seçerek dışa aktarabilirsiniz: Dosya>İndirme>Belge türü.

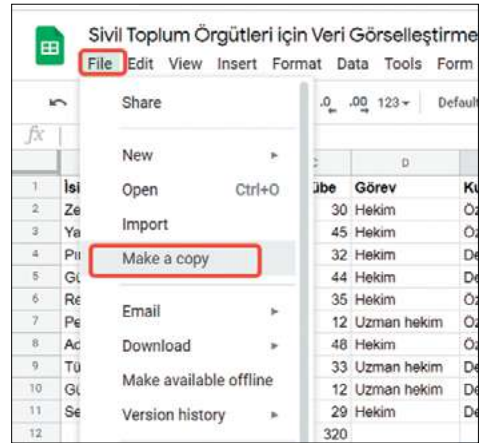
## Elektronik Tablonun Kopyasını Oluşturmak

E-tabloda çalışılan veri setinin kopyasını oluşturmak, veriyi korumak için önemlidir. Ayrıca veri dosyası üzerinde birden fazla kişi ortak çalışma yürütüyor ise, veriyi görselleştirmeye hazırlamadan önce orijinal olanı koruyup bir kopya daha oluşturarak çalışılmaktadır. Oluşturulan kopya yine aynı e-tablo sayfasında açılacaktır.

Kılavuzun online versiyonu örnek;

- 1) [Sivil Toplum Örgütleri için Temel Veri Eğitimi El Kitabı- Sağlık Çalışanları Pandemi Anketi-2020](#)
- 2) Anket verisini sadece görmeye (**view**) izin verilmiştir. **Gmail hesabınıza** giriş yaptıktan sonra veri setini çoğaltabilirsiniz.
- 3) **Dosya > Bir Kopyasını Oluşturma** seçeneği ile veri setinin kopyasını oluşturur ve dosya üzerinde rahatlıkla çalışabilirsiniz.

**Şekil 8:** Google e-tablonun kopyasını oluşturma. Dosya > Kopya oluşturma.



Veri setinin kopyasını oluşturduktan sonra kendi Drive hesabınızda düzenleyebilir, erişime açabilir ya da sadece size özel paylaşım seçenekleri belirleyebilirsiniz.



## Elektronik Tabloyu Paylaşmak

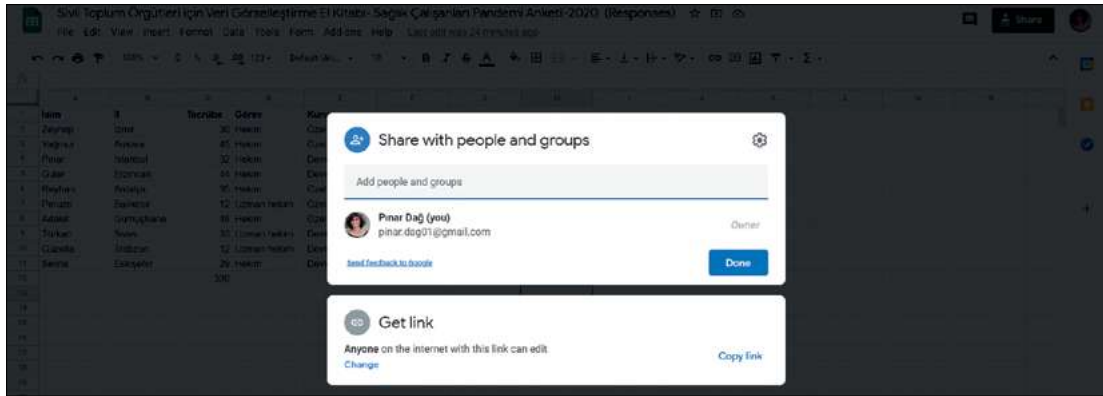
STÖ'lerin ortak çalışmaya açık yapısı veri paylaşımı için de geçerlidir. Google E-tablolar, verinizi Google hesabı sahibi olmayanlarla da çevrimiçi paylaşmamızı sağlar. Ayrıca kullanıcılara farklı alternatifler de sunar. Yeni bir sayfa oluşturduğumuzda varsayılan ayarı bize özeldir, yani içeriği sadece biz görüntüleyebilir ve düzenleyebiliriz. “Paylaş” sekmesi ni kullanarak tablo üzerindeki değişiklik ve görüntüleme seçeneklerini değiştirebiliriz.

İsim	İl	Tecrübe	Görev	Kurum
Zeynep	İzmir	30	Hekim	Özel
Yagmur	Ankara	45	Hekim	Özel
Pinar	İstanbul	32	Hekim	Devlet
Özler	Erzincan	44	Hekim	Devlet
Rayhan	Antalya	35	Hekim	Özel
Pınar	Balıkesir	12	Uzman hekim	Özel
Adilet	Gümüşhane	46	Hekim	Özel
Tarkan	Sivas	33	Uzman hekim	Devlet
Guzella	Trabzon	12	Uzman hekim	Devlet
Selma	Eskişehir	29	Hekim	Devlet
		320		

Şekil 9: Google E-tablo veri seti “Paylaş” düğmesi ekran görüntüsü.

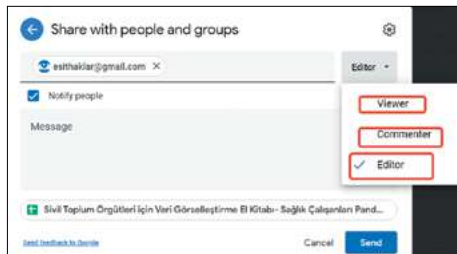
## Elektronik Tabloyu Kişilerle ve Gruplarla Paylaşmak

“Paylaş” sekmesi tıklandığında, kişilerle veya gruplarla paylaşım seçenekleri belirecektir. İlgili kişi ve kurumlara ait e-posta hesapları eklenerek veriyi görmelerini, veriye dair yorumlarını ve üzerinde çalışmalarını sağlanabilir.



Şekil 10: “Paylaş” sekmesi tıklandıktan sonra açılan ve veri setini kişiler ve gruplarla paylaşma seçeneklerinin ekran görüntüsüdür.

### Paylaşım Örneği:



Şekil 11: E-posta ekleyerek veri setinin paylaşılmasına ait ekran görüntüsüdür.

Herkese açık paylaşım örneği:



**Şekil 12:** Veri setini, linke sahip /internete sahip herkesle paylaşım üzerinde düzenleme yapmaya izin verme ekran görüntüsüdür.

Google E-Tablosu'nu paylaşmak için farklı seçenekler vardır: Ortak çalışmaya kişi davet edilebilir, belge üzerinde birlikte çalışma yürütülebilir.

## Elektronik Tablolara Veri Yükleme ve Veriyi Dönüştürme

Google E-tablolar veriyle çalışırken kullanılan en yaygın araçlardan biri olduğu için kılavuzda kapsamlı olarak yer almaktadır ancak başka araçları da veri görselleştirmek için kullanabiliriz. Verileri yüklemek ve dönüştürmek için;

1. Google Drive açınız.
2. Ayarlar (Dişli sembolüne) tıklayınız,



**Şekil 13:** Drive'ın sağ üst köşesinde yer alan dişli sembolünün ekran görüntüsüdür.

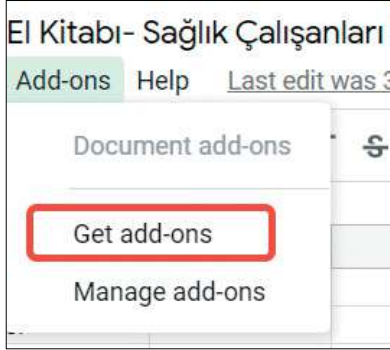
3. Ayarlar sayfasından, "Yüklemeleri Dönüştürün" işaretlenip "Tamam" seçildiğinde yüklenen her türlü içeriğin Microsoft Excel, Word, PowerPoint gibi biçimlere dönüştürülmesi sağlanır. Bu seçenek, işaretin kaldırılmasıyla tekrar pasif hale getirilebilir.



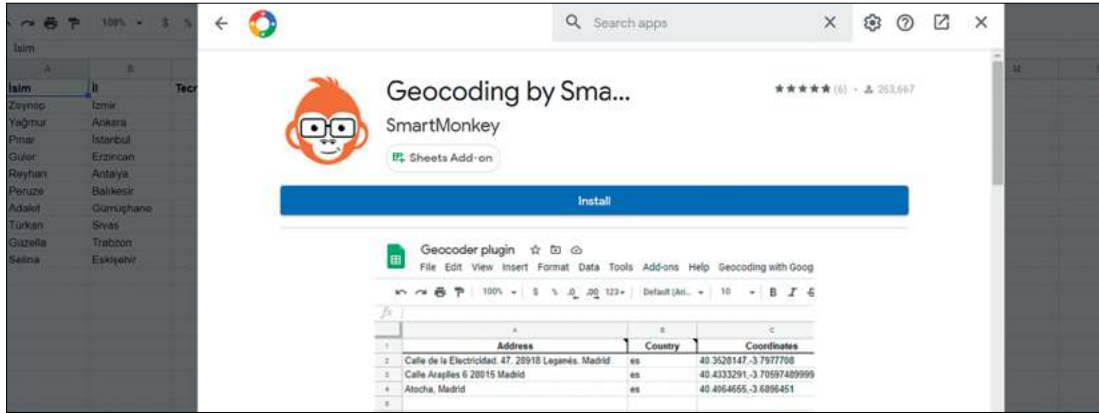
**Şekil 14:** Excel dosyasının yüklenmesi ve e-tabloya dönüştürülmesi ile Microsoft Excel, Virgülle Ayrılmış Değerler (CSV) dosya tipinin ekran görüntüsü yer almaktadır.

## Elektronik Tablolardaki Coğrafi Kodlar

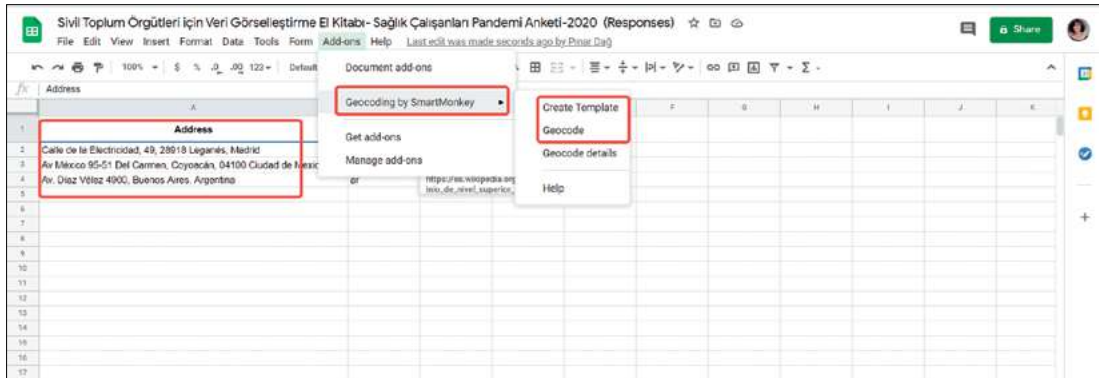
Coğrafi kodlama, adresleri veya konum adlarını harita üzerinde göstermeyi sağlayan coğrafi koordinatlara ( x ve y koordinatlarına) dönüştürebilmek anlamına gelir. İlki enlem, ikincisi ise boylamdır. Google e-tablonun ücretsiz eklentisi [Geocoding by SmartMonkey](#) e-tabloya ekliyoruz.



**Şekil 15:** Google E-Tablo Add-ons>Get add-ons ekran görüntüsüdür.



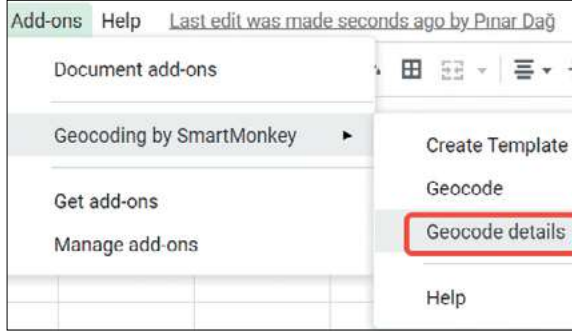
**Şekil 16:** Geocoding by SmartMonkey Install /eklenmesinin ekran görüntüsüdür.



**Şekil 17:** Eklenti bölümüne Geocoding by SmartMonkey'in eklendiğini ve sayfanın da açıldığını ekran görüntüsüdür.



Eklenti sonrası “Geocode details” seçildiğinde örnek verileri içeren yeni bir sayfa oluşacak ve üç yeni sütun; yani “Enlem”, “Boylam” ve “Adres” görülecektir.

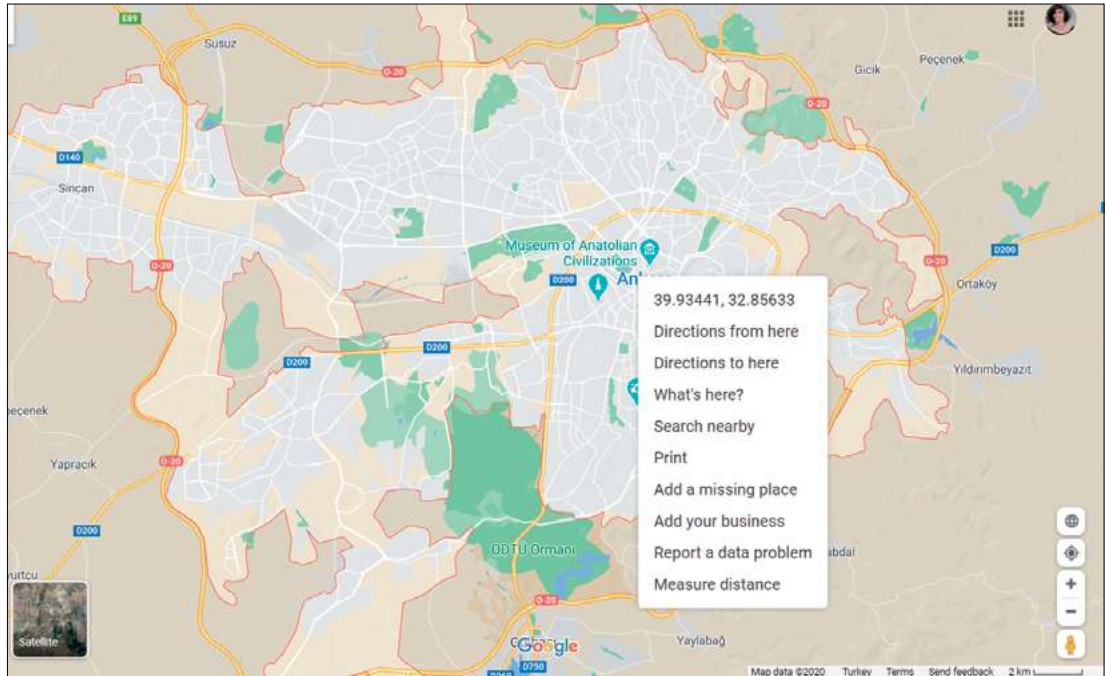


Şekil 18: Geocode details ekran görüntüsüdür.

Ardından örnek verilere enlem, boylam ve adres eklenebilir. Aşağıdaki görselde örnek verinin koordinasyonları ve adresler eklenmiş hali yer almaktadır.

Address	Country	Latitude	Longitude	Address found
Calle de la Electricidad, 49, 28918 Leganés, Madrid	es	40.3529951	-3.7976303	Calle de la Electricidad, 49, 28918 Leganés, Madrid, Spain
Av México 95-51 Del Carmen, Coyoacán, 04100 Ciudad de México	mx	19.3551609	-99.1679037	Av México 95, Del Carmen, Coyoacán, 04100 Ciudad de México, CDMX, Mexico
Av. Díaz Vélez 4900, Buenos Aires, Argentina	ar	-34.6087952	-58.4360388	Av. Díaz Vélez 4900, Buenos Aires, Argentina

Şekil 19: Geocode details sonrası örnek veri setinin ekran görüntüsüdür.



Şekil 20: Ankara il sınırı haritası Google Map ekran görüntüsüdür. Enlem ve boylam koordinatlarını içermektedir.

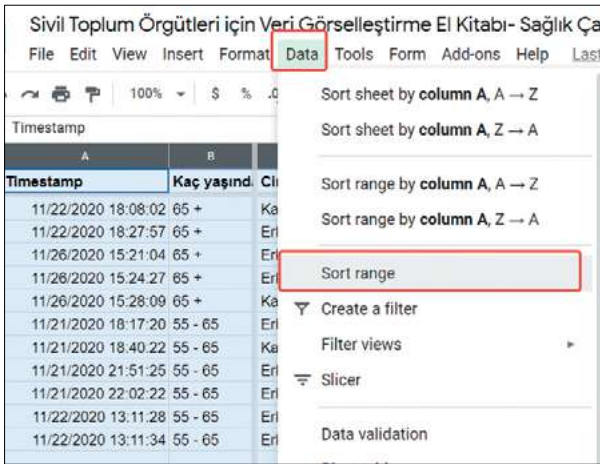
## Verileri Sıralamak ve Filtrelemek

E-tablolarda ya da Excel'de çalışırken verileri alfabetik ya da sayısal olarak sıralayabilir, görmek istemediğimiz verileri filtreleyebiliriz. Bunun için veri setinin tamamı görselde gösterilen hücre (kırmızı renkli) tıklanarak seçilir.

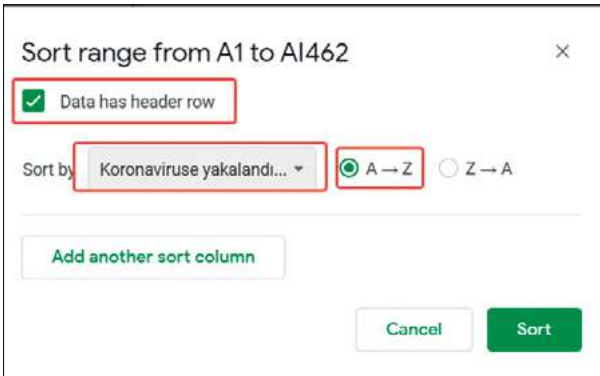
Timestamp	Kaç yaşınd	Cinsiyetiniz nedir?	Medeni durumunuz ?	Virüsü ailenize buluşturma korkus	Çalıştığınız sağlık kurum	Göreviniz nedir? (Hekim	Branşınız nedir?	Koronavirü Yukarıdaki soruya cev
11/22/2020 18:08:02 65 +	Kadın	Evlü	Façesıyla	Özali	Hekim	Kadın-doğum	Hayır	
11/22/2020 18:27:57 65 +	Erkek	Evlü	Hiç	Özali	Hekim	Genel	Hayır	
11/26/2020 15:21:04 65 +	Erkek	Evlü	Façesıyla	Özali	Hekim	Genel cerrahi	Evet	
11/26/2020 15:24:27 65 +	Erkek	Evlü	Biraz	Özali	Uzman hekim	Anesteziyoloji	Kasım	
11/26/2020 15:28:09 65 +	Kadın	Bekar	Façesıyla	Özali	Hekim	Pediyatri	Hayır	
11/21/2020 16:17:20 55 -65	Erkek	Evlü	Façesıyla	Özali	Hekim	Ortopedi ve travmatoloji	Hayır	
11/21/2020 18:40:22 55 -65	Kadın	Evlü	Façesıyla	Özali	Hekim	Pediyatri	Evet	
11/21/2020 21:51:25 55 -65	Erkek	Evlü	Biraz	Özali	Hekim	Kıv	Hayır	
11/21/2020 22:02:22 55 -65	Erkek	Evlü	Biraz	Özali	Hekim	KGB	Hayır	
11/22/2020 13:11:28 55 -65	Erkek	Evlü	Biraz	Özali	Hekim	Pediyatri	Hayır	

**Şekil 21:** Veri setinin sıralamadan önce seçildiği ekran görüntüsüdür.

Ardından veri bölümünden sıralama aralığı tıklanır.



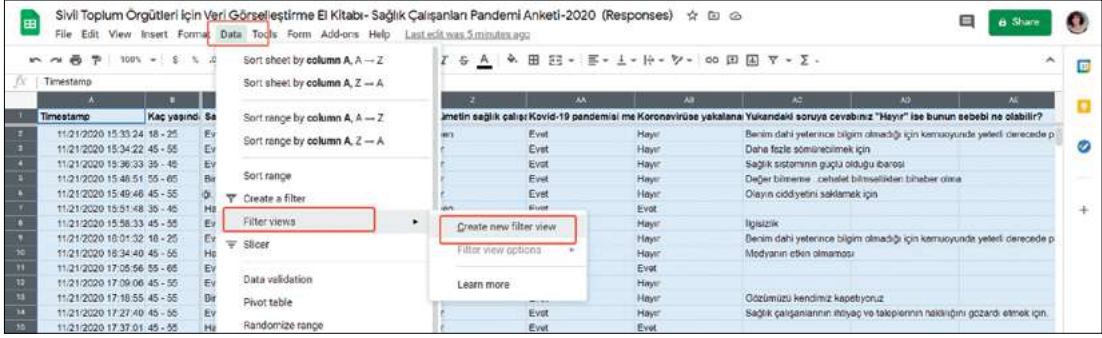
**Şekil 22:** Seçilen veri setinin veri bölümünden sıralama aralığı seçme ekran görüntüsüdür.



Örnek veri setimiz [Sağlık Çalışanları Pandemi Anketi](#)'nden devam edersek "Covid-19'a yakalandım." diyenler sıralanır. Bu sıralamayla diğer yanıtlara göre sağlık çalışanlarının yaşı, medeni durumu, cinsiyeti, hastane koşulları ve diğer tüm yanıtlar incelenerek, anketi yanıtlamış sağlık çalışanlarının Covid-19'a yakalanma nedenleri daha detaylı analiz edilebilir. Bu sıralama; ilişki kurulabilecek diğer boyutlar için de fikir verebilir.

**Şekil 23:** Sıralama Aralığı seçildikten sonra hangi aralığın sıralandığını gösteren ekran görüntüsüdür. Bu görüntüde Koronavirüs'e yakalanmış ya da yakalanmamış olan sağlık çalışanlarının anket yanıtı sıralanmıştır.

Veri setinin ilk satırı başlık olarak seçilir, yukarıda yer alan görselde de görüldüğü üzere sadece Koronavirüs’e yakalananların sıralanması sağlanır. Yani önce, sıralanmasını istediğiniz sütunu ve ardından bir sıralama düzeni seçmeniz gerekmektedir. Dilenen sayıda sütun sıralamaya eklenebilir. Veriyi filtrelemek için ise sırasıyla önce “Veri”, ardından “Filtre görünümüleri” ve son olarak “Yeni filtre görünümü oluştur” seçenekleri seçilir.



Şekil 24: Veri filtreleme ekran görüntüsüdür.

Örneğin, “Koronavirüs’e yakalandınız mı?” sorusuna “Evet” diyen sağlık çalışanları filtrelendi.



Şekil 25: Sağlık çalışanları ile yapılmış olan pandemi anketinde virüse yakalanmış olan kişileri “evet” yanıtına göre filtrelenmesinin ekran görüntüsüdür.

## Formüllerle Hesaplama Yapmak

E-tabloların satır ve sütunlarında otomatik olarak hesaplamalar yapılmak istendiğinde, basit formüller ve işlevler ekleyerek kullanıcıya çokça zaman kazandıran birçok temel işlem yapılabilir. Formüller, her zaman “eşittir” (=) işaretiyle başlar ve başka hücreleri toplamayı sağlar: =(C2+C11) gibi. Ayrıca bir aralığın tamamının toplamını hesaplamak için de kullanılabilir: =Toplam(C2:C50) gibi.

İsim	İli	Tecrübe	Görev	Kurum
Zeynep	Izmir	30	Hekim	Özel
Yağmur	Ankara	45	Hekim	Özel
Pınar	Istanbul	32	Hekim	Devlet
Güler	Erzincan	44	Hekim	Devlet
Reyhan	Antalya	35	Hekim	Özel
Peruze	Balikesir	12	Uzman hekim	Özel
Adalet	Gumüşhane	48	Hekim	Özel
Türkan	Sivas	33	Uzman hekim	Devlet
Guzella	Trabzon	12	Uzman hekim	Devlet
Selina	Eskişehir	29	Hekim	Devlet

**Şekil 26:** Fonksiyonlar bölümünden ya da sütunu seçerek ya da formülü elle yazarak sütunda yer alan verileri toplama formülü ekran görüntüsüdür.

İsim	İli	Tecrübe	Görev	Kurum
Zeynep	Izmir	30	Hekim	Özel
Yağmur	Ankara	45	Hekim	Özel
Pınar	Istanbul	32	Hekim	Devlet
Güler	Erzincan	44	Hekim	Devlet
Reyhan	Antalya	35	Hekim	Özel
Peruze	Balikesir	12	Uzman hekim	Özel
Adalet	Gumüşhane	48	Hekim	Özel
Türkan	Sivas	33	Uzman hekim	Devlet
Guzella	Trabzon	12	Uzman hekim	Devlet
Selina	Eskişehir	29	Hekim	Devlet

**Şekil 27:** Fonksiyonlar bölümünde ya da sütunu seçerek ya da formülü elle yazarak sütunda yer alan verilerin ortalama formülü ekran görüntüsüdür.

Tüm Google E-tablolar fonksiyon listesine,

<https://support.google.com/docs/table/25273?hl=tr> bölümünden ulaşabilirsiniz.

**Şekil 28:** Google E-tablolar fonksiyon listesi ekran görüntüsüdür.

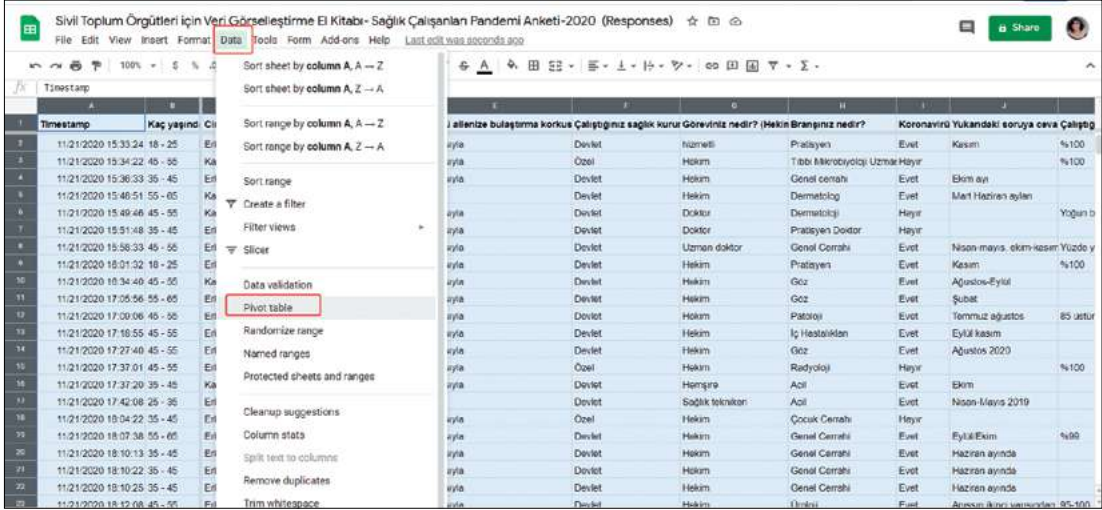


## Verileri Pivot Tablolarla Özetlemek

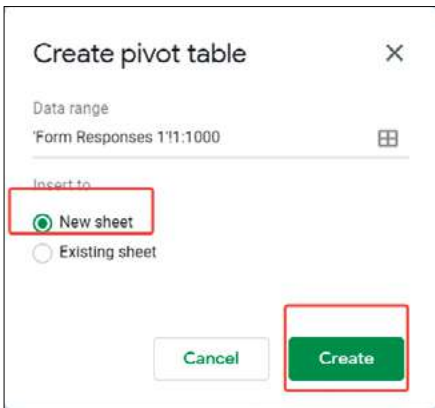
Özet tablolar, büyük miktardaki verileri analiz etmemize, veri noktaları arasındaki ilişkileri görmemize ve verileri daraltmamıza izin verir. Pivot tabloyu, verileri özetlemek ve e-tablodaki tüm bilgileri daha kolay anlamak için kullanırız.

Kılavuzun önceki bölümlerinde sıkça başvurduğumuz [Sağlık Çalışanları Pandemi Anketi-2020](#) anket verisi üzerinden örnek vermek gerekirse; anketi yanıtlayanlar arasında bölgelere göre uzman hekimlerin sayısını özetliyor ve ilk önce özetlenecek veri setini seçiyoruz.

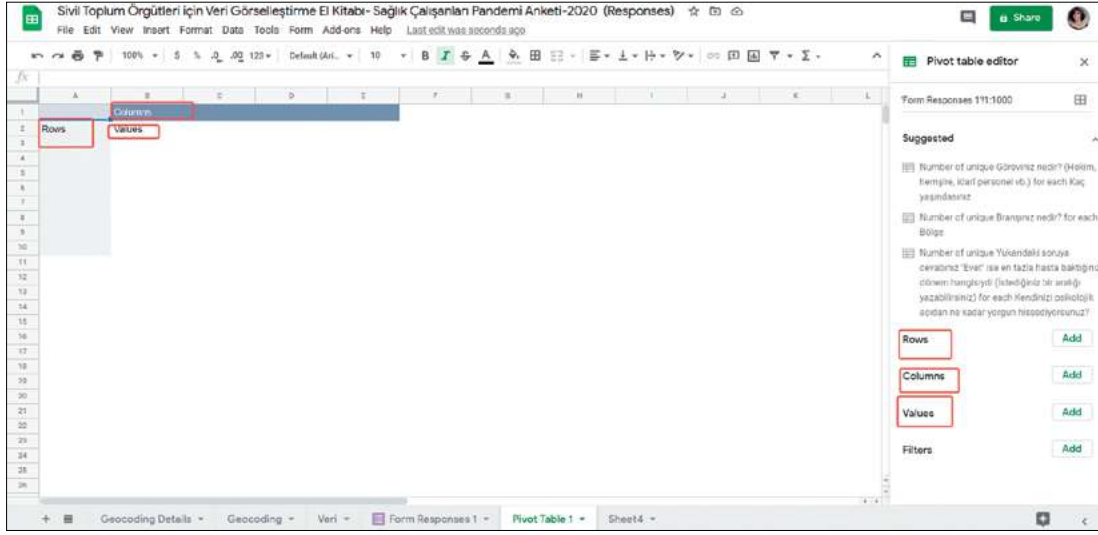
Bunun için ilk olarak veri bölümünde listelenen başlıklardan “Pivot Tablo”yu seçmek gereklidir. Ardından açılan menüde “Yeni Sayfa”yı seçerek özet tablolar bölümündeki değişkenleri hedeflediğimiz sonuca (bölgeler) göre işaretleyerek veriyi bir aşama özetlemek ve oluşan grafikle bölgelere göre hekim sayılarını karşılaştırmak mümkündür.



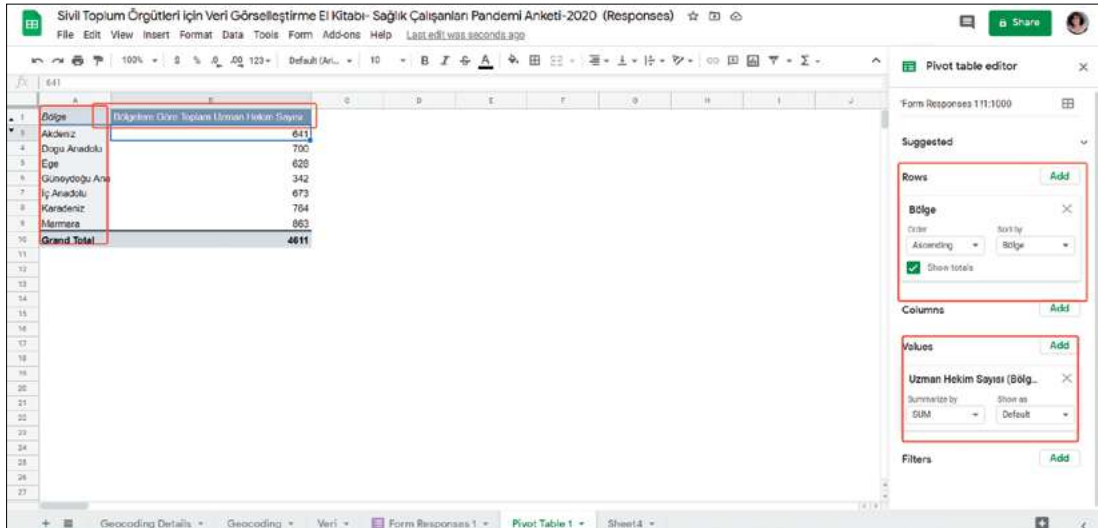
Şekil 29: Sağlık Çalışanları Anket verisine “Pivot Table” fonksiyonunu uygulama ekran görüntüsüdür.



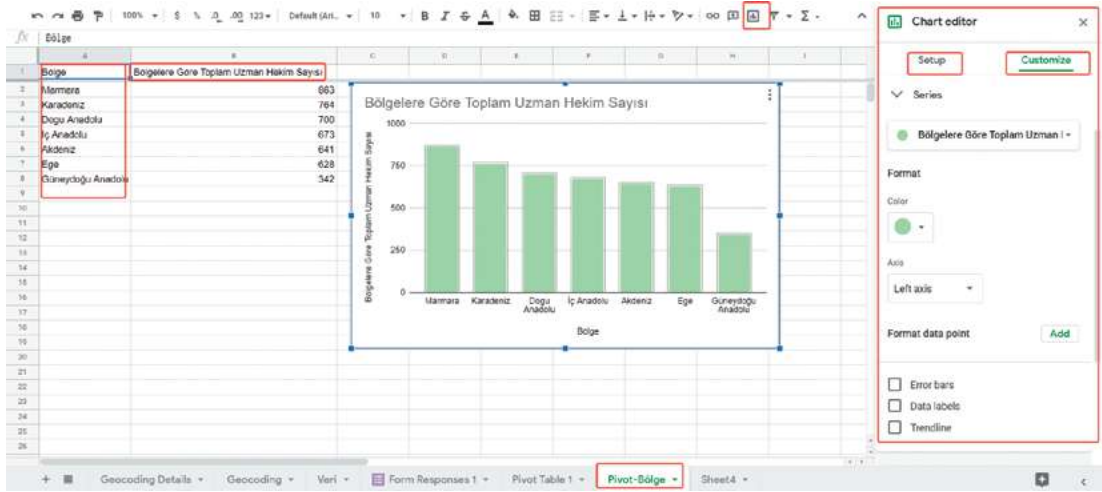
Şekil 30: Pivot Tablo oluşturma ve yeni sayfa açmak ekran görüntüsüdür.



**Şekil 31:** Veriyi özetlemek için açılan “Pivot table” sayfası ekran görüntüsüdür. Solda satır, sütun ve değerler, sağda ise Pivot editörünün önerdiği satır, sütun ve değerler sekmelerinin yer aldığı ekran görüntüsüdür.



**Şekil 32:** Bölgelerine göre uzman hekimlerin sayılarının toplamı ekran görüntüsüdür. (Bu bölümde yer alan veriler gerçeği yansıtmamaktadır, veriler temsilen eklenmiştir.)



**Şekil 33:** Bölgelerine göre uzman hekimlerin toplamlarının sütun grafik karşılaştırması ekran görüntüsüdür. (Bu bölümde yer alan veriler gerçeği yansıtmamaktadır. Veriler temsilen eklenmiştir. Ankette bölge, şehir ya da il bilgisi toplanmamıştır. Özet tablo fonksiyonunu (pivot table) göstermek için hazırlanmıştır.)

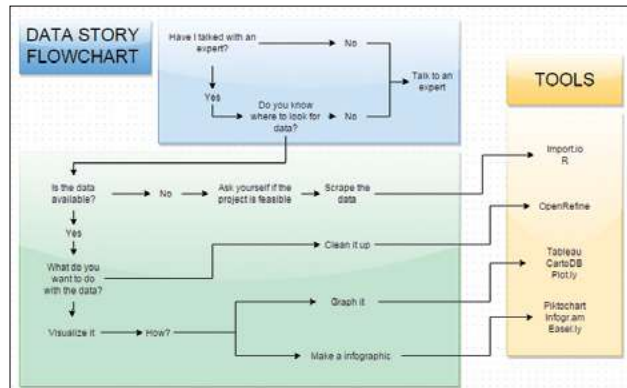
## 3. BÖLÜM

### Veri Toplamının Önemi

Günümüz küresel dijital ikliminde akıllı telefonlar şiddetin görüntülerini belgeleyebilir, dağıtabilir, çoğaltabilir. Uydular ve dronlar uzak mesafelerdeki afetlere yönelik değerlendirmeler yapıp, raporlar verebilir. Toplanan sosyal medya verileri, siyasi protestolar hakkında gerçek zamanlı bir farkındalık sağlayabilir. Tüm bu olumlu tabloya rağmen; teknolojik fayda, riskler/zararlar açısından bakıldığında kendi içinde zorluklar ve fırsatlar yaratabilmektedir. Bu teknolojiler, ihtiyaç sahibi insanlara yardım etmek ve yetkinin kötüye kullanımını önlemek için sorumlu bir şekilde nasıl kullanılabilir? Hak odaklı veri sistemi ve kullanımı nasıl sağlar?

Sivil toplum örgütleri ve kalkınma ajansları, artan kurumsal ihtiyaçlara karşılık yeni veri odaklı yaklaşımlar geliştirmektedirler. Veriler her zaman bu sektörler için önem taşıırken aynı zamanda dijital ve ağ teknolojilerinin yapısı veriyi toplama, depolama, analiz etme yönünden giderek daha da hızlanmaktadır. Ancak bu hız, kullanıcıları için olumlu birçok getirisinin yanında hızla yenilenen teknolojilerin, insan ve toplum hayatını nasıl etkileyeceğine dair henüz bir netice sunmamaktadır. Örneğin; veri görselleştirme kapsamında yapılan tüm bu çabaların, çevremizi nasıl etkileyebileceğini henüz bilmiyoruz. Hayatımızın sosyal, finansal yönlerini değiştiren teknolojilerin aynı zamanda adalet hitap etme potansiyeline sahip olup olmadığı da bilinmemektedir. Bu kapsamda; hak odaklı veri toplamak ve görselleştirmek, kesişen endişeler arasında dünyada/Türkiye’de hak odaklı veri politikaları geliştiren kurumların nasıl çalıştıklarını görmemizi sağlayabilir.

Bu noktada; veri görselleştirme projeleri yürütülürken yapılan hazırlıklar, çalışmanın genelini etkileyen bir öneme sahiptir. Örneğin; hayvan hakları odaklı yürüttüğünüz bir çalışmada hayvanlara yönelik istismarı, şiddeti görünür kılmak ve aynı zamanda Türkiye genelinde bölgelere göre nasıl bir değişim gösterdiğini yıl yıl ortaya koymak isteyebilirsiniz. İlk bakışta hak odaklı bu veriyi çalışmak çok önemli ve belki kolay gözükecektir belki, ancak böylesi bir çalışmayı yapabilecek hijyenik bir veri setine sahip olmadığınızı, yani tutarlı ve hatasız olmadığını küçük bir araştırmayla hemen fark edeceksiniz. Peki, böyle bir durumda görselleştirmek istediğiniz konuyu siz seçemiyorsanız ve veriyi kendiniz toplamanız gerektiğinde nasıl bir yol izlemelisiniz? Öncelikle bu duruma veri ile çalışan [James Gaines](#)’in hazırladığı akış tablosu ile yanıt vermek istiyorum.



Şekil 34: Veri ile çalışma akış tablosu. Tasarım, <https://dataimp.wordpress.com/>'a aittir.



Yukarıdaki akış tablosu, veri odaklı yürütülen projelerde çeşitli taktikler geliştirerek neye sahip olduğunu ve neye sahip olunmadığını anlayarak, çalışmanın daha nitelikli şekilde ilerlemesine yardımcı olmaktadır. Yukarıdaki akış tablosu örneği kullanıcıya aşama aşama nasıl ilerleyeceğini göstermektedir.

“Hayvan hakları odaklı veri mevcut mu?”

Cevap “Hayır” ise, o zaman bu konunun bir uzmanı varsa görüşün. Eğer uzman, hayvan hakları odaklı toplanmış bir verinin olup olmadığından emin değilse, veriye nasıl erişebileceği konusunda yol gösterecektir. STÖ’ler bu konuda veri toplamış mı? Resmi kaynaklarda mevcut mu? Açık veri portalları sunuyor mu? İndirmeye uygun mu? Yoksa bir yerden araçlarla kazıyabilir misiniz?

“Kazıdıklarınızı temizlemek için kullanmanız gereken araçları biliyor musunuz?”

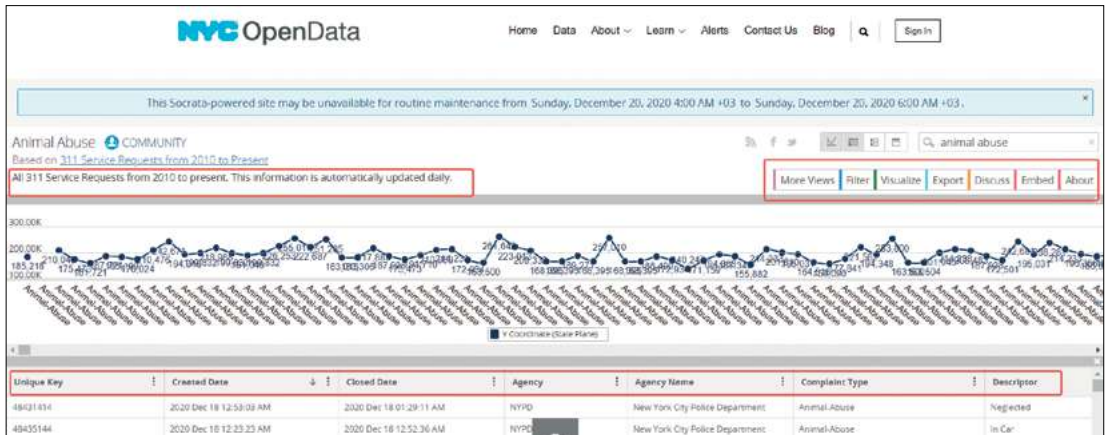
Cevap “Evet” ise, o zaman veri setini hangi araçlarla analiz edeceksiniz ya da hangi araçlarla görselleştireceksiniz?”

Daha somut bir örnek üzerinden aşama aşama inceleyelim:

## Hayvan Hakları Odaklı Veri Toplama Örneği

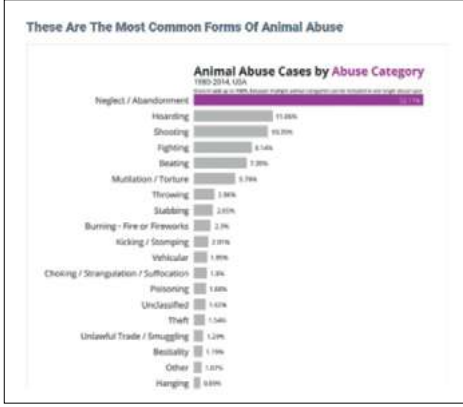
**Yanıt aranan sorular:** Hayvan Hakları odaklı veri analizleri karar alma süreçlerinde nasıl kullanılmalı? Veri toplama, analiz etme hak savunucularının uygulamalarını nasıl şekillendirir? Belirli riskleri barındıran veri araçları / teknikleriyle veri aktivistleri nasıl bir etkileşim içinde olmalı? Hayvan Hakları odaklı veride genel ilkeler oluşturulabilir mi?

Bu soruların yanıtını bulabilmek için öncelikle hayvanlara yönelik şiddetin boyutlarını görünür kılan veriler toplamak gerekiyor. Örneğin; [New York Belediyesi](https://nycopendata.socrata.com/Social-Services/Animal-Abuse/vqss-uwk2) düzenli olarak hayvanlara yönelik istismar verisini toplayarak topladığı verileri herkesin erişimine açıyor. 2010 yılından beri toplanan verilerde hayvanların öldürülme biçimleri detaylı olarak yer alıyor.



**Şekil 35:** New York Belediyesi'nin hayvanlara yönelik istismar verisini topladığı açık veri platformunun ekran görüntüsüdür. (<https://nycopendata.socrata.com/Social-Services/Animal-Abuse/vqss-uwk2>)

Aşağıda, dadaviz.com tarafından 1990-2014 yılları arasında hayvanların ne tür bir istismara uğradığını kategorik olarak sunan bir grafik görüyorsunuz. Veriler, 1990-2014 yıllarını kapsadığı için güncel olmayabilir ancak günlük güncellenen bir veri setine erişmek hala mümkündür. Günlük raporlanan verilere göre hayvanların çoğunlukla ihmalinden, bakımsızlıktan hayatlarını kaybettiğini görüyoruz. Bu bilgiyi öğrenmek ne işe yarar? Öncelikle hayvan hakları odaklı alınan ya da alınacak olan politika ve kararları etkiler.

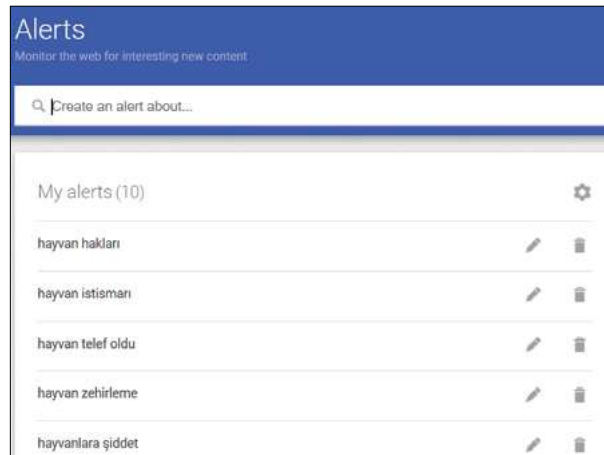


**Şekil 36:** <http://dadaviz.com/s/animal-abuse/> tarafından kategorisine göre hayvan istismarı grafiği.

Türkiye’de bu alanda hayvanlara yönelik sistematik şiddet verisini toplayan [Hayvan Hakları İzleme Komitesi](#) mevcuttur ancak veriler tüm gayretlere rağmen yeterince güncel değildir. Bu verilerin sistematik olarak toplanabilmesi için yerel yönetimlerin, açık veri platformlarının devreye girmesi gerekmektedir. Hak odaklı çalışmaların, tür ayrımı yapılmaksızın her türlü şiddet verisini toplayan bir yöntemle geliştirilmesi gerekmektedir.

Veri toplamak, hangi verinin toplanacağına karar vermek ve o veriyi herkesin erişimine açmak kesinlikle politik bir süreçtir. Bu sebeple STÖ’lerinin çağımız için önemli motivasyonlarından biri de çalıştıkları alanla ilgili özgün, hijyenik veri toplama motivasyonunu yaratmaları olmalıdır. Örneğin; bir yıl boyunca hayvanların medyada nasıl haberleştirildiğini raporlamak, ay ay incelemek ve verileri analiz etmek istiyorsunuz.

1) Google uyarı sistemini kullanarak anahtar kelimeler belirleyerek hayvanlara yönelik yapılan haberlerin anlık verisini toplayabilirsiniz <https://www.google.com/alerts>



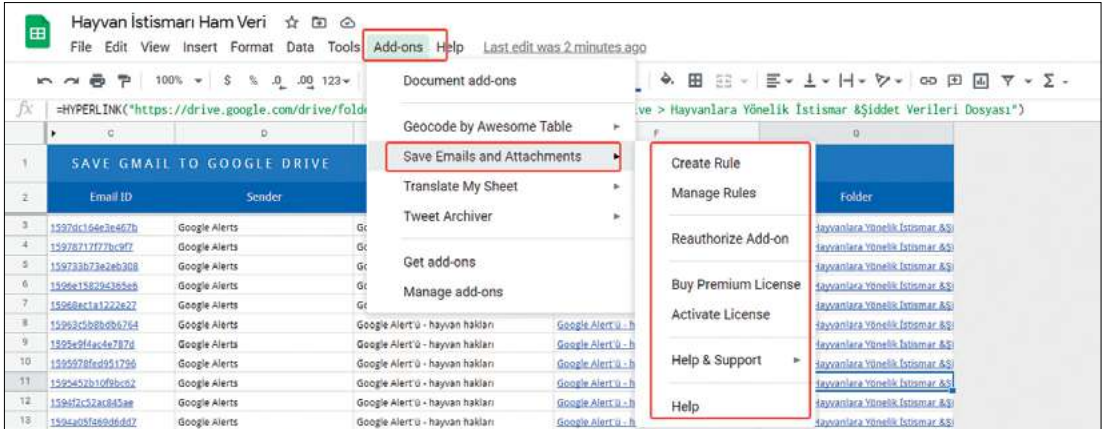
**Şekil 37:** Google Alerts kullanılarak medyada yapılan hayvan haberleri için belirlenmiş kelimelerin ekran görüntüsüdür.

2) Ardından e-postanıza düşen haberleri add-ons (eklentiler) bölümüne ekleyeceğiniz Save Emails and Attachment ile e-tabloya aktararak yapılandırmaya, kategorilerine ayırmaya başlayabilirsiniz. Yöntem zahmetli olabilir ancak bugüne kadar hiç toplanmamış bir veriyi en az teknik bilgi ile toplamaya başlayabilirsiniz.

<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - hayvan telef oldu - hayvan telef oldu Günde bir güncelleme - 19 Aralık 2020 HABERLER Hayvanları telef olan çift...	6:16 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - kuzu - kuzu Günde bir güncelleme - 19 Aralık 2020 HABERLER İrmakta bacakları bağlı 2 kuzu ölüsü bulundu Gaz...	3:42 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - hayvanlara şiddet - hayvanlara şiddet Günde bir güncelleme - 19 Aralık 2020 HABERLER Çevreye ve hayvanlara z...	3:36 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - koyun - koyun Günde bir güncelleme - 19 Aralık 2020 HABERLER Ağrı'da 5 Yılda 750 Bin Koyun Projesi ile çiftçil...	3:01 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - inek - inek Günde bir güncelleme - 19 Aralık 2020 HABERLER İnek otomobile çarptı Hürriyet Kazayı, İnek yara alm...	3:01 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - kedi - kedi Günde bir güncelleme - 19 Aralık 2020 HABERLER Antalya'da yavru bir kedi, bacakları ve kuyruğu kesil...	3:01 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - köpek - köpek Günde bir güncelleme - 19 Aralık 2020 HABERLER Zincirle bağlanıp evde bırakılan köpeğin sevindi...	3:01 PM
<input type="checkbox"/>	☆	Google Alerts	Google Uyarı - hayvan hakları - hayvan hakları Günde bir güncelleme - 19 Aralık 2020 HABERLER Hayvanlara eziyeti ancak hapi...	1:01 PM

Şekil 38: Google Uyarı aracılığı ile e-mail adresine düşen hayvan haberleri ekran görüntüsü.

3) "Save Emails and Attachments"i eklentiler bölümüne aşağıda gösterildiği gibi eklemelisiniz:



Şekil 39: Google E-tablo eklenti bölümüne eklenen "Save Emails and Attachments"ın ekran görüntüsüdür.

Uygulama eklendikten sonra tüm uyarılar PDF olarak e-tablonuza aktarılacaktır. Ardından bu verileri yapılandırmanız gerekmektedir.

1	SAVE GMAIL TO GOOGLE DRIVE				
2	Email ID	Sender	Subject	File	Folder
3	15979c15de3e667b	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
4	15979c15de3e667b	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
5	159733b73e2eb308	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
6	1595e15929c365e6	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
7	15958ec1a1222e27	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
8	15953c358b0b0764	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
9	15959ef5ac4e787d	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
10	1595978fd951796	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
11	1595e52b10f90e82	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
12	1595f2c72ac845ae	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
13	1594a0586606d4d7	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
14	1594a0f5956537b2	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası
15	15925e027e0b6124	Google Alerts	Google Alert Ü - hayvan hakları	Google Alert Ü - hayvan hakları.pdf	My Drive - Hayvanlara Yönelik İstismalar &Şiddet Verileri Dosyası

**Şekil 40:** Google Uyarılar ile toplanan hayvan haberleri verisinin eklenti ile e-tabloya aktarılmış olan ekran görüntüsüdür.

Ardından, e-tabloya atılıp yapılandırılan verileri görselleştirip, aylık olarak hayvanların nasıl öldüğünü ya da ne tür bir şiddet sonucu hayatını kaybettiğini yeterli olmasa da görünür kılmannın teknik olarak yöntemini geliştirmiş olursunuz. Örneğin; aşağıda 2020 yılında hayvanlara yönelik medyada çıkan haberlerde, sadece Temmuz ayında şiddet türüne göre öldürülen/ölen hayvan grafiği (ham veriler) bulunmaktadır:

1	İl	İlçe	Hayvan Türü	Hayvan Sayısı	Şiddet Tanımı	Ölü hayvan sayısı	Yaralı Hayvan Sayısı	Açıklama
2	Aydın	Çine	Büyükbaş	-	Çiçek Hastalığı	-	-	-
3	Kastamonu	Taşköprü	Koçak	1	Öldürme	1	-	-
4	Aydın	İncirliova	Domuz	10	Öldürme	10	-	-
5	Adana	-	At	1	Kaza	1	-	-
6	Antalya	-	Kuş	4	Kaza	4	-	-
7	Bursa	Karacabey	Koyun	9	Kaza	9	-	-
8	Antalya	Finike	Koçak	1	Kaza	1	-	-
9	Batman	Kozluk	Kedi	1	Kaza	1	-	-
10	Ordu	Fatsa	Büyükbaş	8	Doğal Afet	8	-	-
11	Van	Gürpınar	-	4	Kaza	4	-	-
12	Sakarya	Karasu	Büyükbaş	5	Kaza	5	-	-
13	Tekirdağ	-	Domuz	1	Kaza	1	-	-
14	Bursa	-	Büyükbaş	6	Kaza	6	-	-
15	Eskişehir	-	Kuş	10	Kaza	10	-	-
16	Tokat	Eriş	Büyükbaş - Arı	300 001	Kaza	300 001+	-	-
17	Eskişehir	Odunpazarı	Arı	500	Kaza	500 000+	-	-
18	Bolu	Gerede	Büyükbaş	16	Kaza	16	-	-
19	Antalya	Konyaaltı	At	9	Öldürme	9	-	-
20	Bartın	Kemerköprü	Balık	100+	Zehirlenme	100+	-	-
21	Antalya	Demre	Balık	100+	Kuraklık	100+	-	-
22	İğdir	Oba Köyü	Büyükbaş	1	Kaza	1	-	-
23	Kırklareli	Lüleburgaz	Kaplımbağ	1	Kaza	1	-	-
24	Manisa	Yunusemre	Leylak	2	Kaza	2	-	-
25	Diyarbakır	Kulp	Büyükbaş	5	Kaza	5	-	-
26	Van	Gürpınar	Koyun	10+	Kaza	10+	-	-

**Şekil 41:** Raporların linklerinin aylık olarak yapılandırıp gruplandırıldığı verilerin ekran görüntüsüdür. Temmuz 2020 şiddet türü ve ölü sayısı grafiği de yer almaktadır. (Ham veri)

Bu örnek üzerinden değerlendirme yapmak gerekir ise, hayvanlara yönelik şiddet verisi toplamak; toplum üzerinde farkındalık yaratmaya aracılık eder, hayvanların karşılaştıkları şiddet ve zulmün görünür kılınması sağlanır; bu şiddet ve zulmün ırkçılık ve cinsiyetçilik kadar benzerlik gösterdiğinin altı çizilir / saptanır; hayvanların dünyayı paylaştığımız duyarlı canlılar olarak yaşam haklarının tanınmasına ve bu hakların savunma tekniklerinin geliştirilmesine yardımcı olur. Sonuç olarak; veriler her zaman mevcut olmayabilir. Onları toplamak, bir araya getirmek, yorumlamak ve ilişkilendirmek için önce toplanmasını sağlamak gerekir. Aynı zamanda ölçekleri kullanarak ya da anket, gözlem, görüşme yaparak da veri toplayabilirsiniz. Yukarıda verilen örnek hiç toplanmamış ya da az toplanmış ya da eksik toplanmış bir veri çalışmasında, az teknik bilgi ile nasıl veri toplanır sorusuna yanıt niteliği taşımaktadır. Elbette daha pratik yöntemler de mevcuttur.

Ancak kurumunuz için veri topluyorsanız;

- Veri toplama yöntemini amaca uygun seçmelisiniz.
- Veri toplarken sistematik olmalısınız.
- Verileri toplarken ve topladıktan sonra uygun şekilde sınıflandırmalısınız.

## Veri Toplama Yöntemleri

Gözlem, görüşme, anket, yazılı kaynaklar, anket ve ölçekler (rakamsal olarak kodlama) veri toplama yöntemleridir.

### Gözlem

**Gözlem ve Türleri:** Gözlem, herhangi bir ortamda oluşan davranışları ayrıntılı olarak tanımlamak amacıyla kullanılan bir veri toplama tekniğidir.

**Süresine Göre Gözlem Çeşitleri:** Sürekli gözlem, zamanı gelince gözlem, aralıklı gözlem.

**Yapısına Göre Gözlem Çeşitleri:** Doğrudan doğruya gözlem ve araçla yapılan gözlem.

### Görüşme

**Görüşme ve Türleri:** Verilerin kaydedilmesini kolaylaştırmak, standart oluşturmak amacıyla geliştirilen gözlem boyunca kullanılan araçtır. Farklı gözlemcilerin aynı şeyleri gözlemlenmesini ve gözlemciler arasında tutarlılığın kontrol edilmesini sağlar. Araştırmada elde edilmek istenen amaca yönelik veri toplamada hangi yöntem uygunsa o seçilir.

- Grup görüşmesi
- Bireysel görüşme



### Yapılandırılmış Görüşmeler:

- Bu görüşme türünde görüşme soruları önceden hazırlanır. Görüşmeci bu plana göre görüşme yapar.
- Yanıtların çözümlenmesi kolay ancak derinlik sağlamada sınırlıdır.

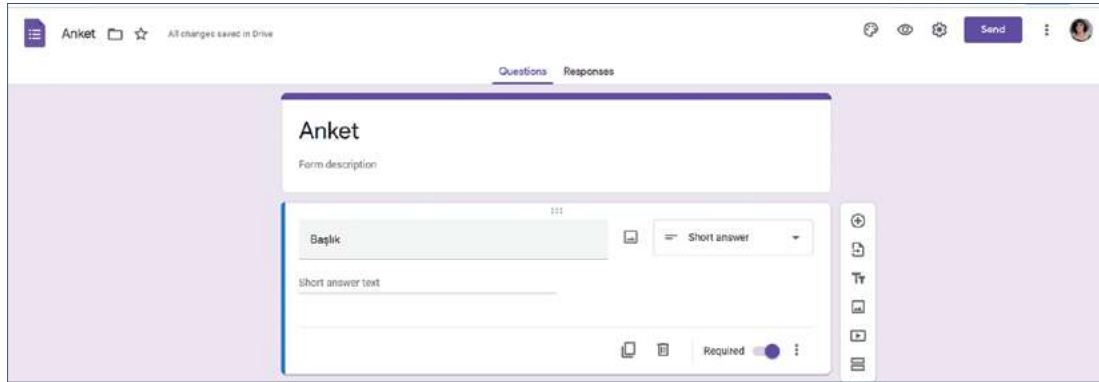
### Görüşmenin Artı ve Eksileri:

Cevaplama oranının yüksekliği, anket veya ölçeklerle elde edilemeyecek derinlikte verilerin elde edilebilmesi ve esneklik olanağı yaratır. Ancak, maliyeti fazla ve zaman alıcıdır.

## Anket

Anket, belli bir araştırmanın amacına uygun düzenlenmiş soru listesidir. Anket tekniği, araştırma yöntemleri arasında en çok kullanılanlardan biridir. Bunun sebebi kolay, ucuz ve bilgileri doğrudan doğruya toplamaya uygun oluşudur.

Google'ın "Drive" bölümünde yer alan çevrimiçi anket özelliği kullanılarak ücretsiz anket soruları hazırlamak mümkündür. Ayrıca, tüm yanıtların verisi eşzamanlı olarak e-tabloya da aktarılmaktadır. Büyük ya da küçük her çeşit veriyi organize etmenize yardımcı olur. İstedığınız şekilde soru ve cevap bölümleri hazırlanabilir. Çoktan seçmeliden, açılır listelere ve doğrusal ölçeğe kadar bir dizi soru seçeneği mevcuttur. Resimler ve YouTube videoları eklenebilir.



Şekil 42: Google Forms ekran görüntüsüdür.

Anketlerinize verilen yanıtlar, gerçek zamanlı yanıt verileri, grafiklerle birlikte otomatik ve düzenlenmiş olarak formda toplanır. İsterseniz tüm verilerinizi E-Tablolar'da görüntüleyerek daha kapsamlı analizler yapabilirsiniz.

### Anket Geliştirme Süreci

1. Amaç ve soruları belirlemek
2. Madde yazmak,
3. Taslak form oluşturmak,

4. Uzman görüşü almak ve ön uygulama formu oluşturmak,
5. Ön uygulama,
6. Sonuçları değerlendirme
7. Ankete son şeklini vermek

### **Anket Hazırlanırken Dikkat Edilecekler**

Anket formunun başında, katılımcılara yönelik kısa bir açıklama bulunmalıdır. Bu açıklamada anketin amacı, soruların nasıl cevaplandırılacağı belirtilerek, görüşmecilere çalışmaya katkıda bulduklarından ötürü teşekkür edilmelidir.

### **Anket Sorularının Hazırlanması**

Sorular katılımcıları/görüşmecileri rahatsız eden nitelikte olmamalıdır. Soru tipi anketin uygulanacağı örnekleme uyumlu olmalıdır. En kolay cevaplanabilecek sorular ve hassasiyeti görece daha düşük sorulara anketin ilk kısmında yer verilmelidir.

### **Anketin Olumlu Yönleri**

Ekonomiktir, kolay uygulanabilir, büyük örneklemelerden veri toplanabilir, standartlaştırılmış cevaplar verir, hazır cevapları sınıflandırmak kolaydır, araştırmacının görüşmeciyi etkileme olasılığı düşüktür.

## **Ölçekler (Rakamsal Olarak Kodlama)**

Ölçekler; bireylerin tutum, davranış ve eğilimleri gibi (yargısal verilerin) soyut kavramların ölçülebilmesi için kullanılmaktadır. (Tutum, eğilim)

### **Likert Tipi Ölçekler**

En sık kullanılan ölçekler Likert tipi derecelendirme ölçekleridir. Ölçek maddeleri, olumsuzdan olumsuza veya olumsuzdan olumluya olmak üzere 3- 5-7 veya 9 nolu olabilir. Bireyden, belirli bir durum karşısında nasıl davranacağını seçmesi istenir.



## **Likert Ölçeği**

- Likert ölçeği psikolojide, çalışan anketlerinde yaygın olarak yer alır.
- Anket çalışmalarında, en yaygın kullanılan yanıt uygulamalarından biridir.
- Ölçek adını, mucidi psikolog Rensis Likert'ten almıştır.

**Tablo 1.** Likert'in Tezinde Kullandığı 5 Seçenekli Ölçüm Sistemi.

Kesinlikle Onaylıyorum	Onaylıyorum	Kararsızım	Onaylamıyorum	Kesinlikle Onaylamıyorum

## Yapısal Nitelikler

**Geçerlilik:** Ölçülmek istenen özelliği tam ve doğru bir şekilde, bir başka özellikle karıştırmadan ölçebilme derecesidir.

**Güvenilirlik:** Bir ölçeği, ölçtüğü bir şeyi her defasında aynı derecede ölçebilmesidir. Her ölçümde, ölçeği dolduran bireyin durumunda, ölçülen konu ile ilgili bir değişiklik olmamışsa ölçmenin aynı sonucu vermesi beklenir.

## Güvenirliği Sağlama Yöntemleri:

**Cronbach Alpha ( $\alpha$ ) Katsayısı:** İç tutarlılık katsayısıdır. Alpha değerinin +1.00'e yaklaşması testin kendi içinde tutarlılığının arttığını gösterir. Cronbach alfa katsayısı 70 ve üzerinde olduğunda kabul edilebilir olduğu belirtilmektedir.

Sonuç olarak; veri toplama yöntemi amaca uygun seçilmelidir, veri toplarken sistematik olunmalıdır. Verileri toplarken ve verileri topladıktan sonra uygun şekilde sınıflandırma yapılmalıdır. Bütün bunlar toplanan verileri analiz etme aşamasında kolaylık sağlayacaktır.

## Bilgi Edinme Başvurusu ile Veri Toplamak

Türkiye'de 2003 yılında yürürlüğe giren Bilgi Edinme Hakkı Kanunu, resmi kaynaklardan veri talep etmek için sık kullanılan bir hizmettir. Bu kanundan yararlanmak için [CİMER](http://cimer.gov.tr)'e başvuru yapmak gerekmektedir.



**Şekil 43:** cimer.gov.tr ekran görüntüsüdür.

Doğrudan kimlik bilgilerinizi girerek ya da e-devlet üzerinden giriş yapıp sorularınızı ekleyebilir ya da yükleyebilirsiniz. Eklenebilen dosya türleri JPG, PNG, TIFF, GIF'tir ve maksimum dosya boyutu ise 3 MB'dir. Günlük başvuru sınırlandırılmış olup, en fazla bir adet başvuru yapabilirsiniz. Elektronik ortamda yapılan başvurularda, başvuru metni 1.500 karakterle sınırlıdır.



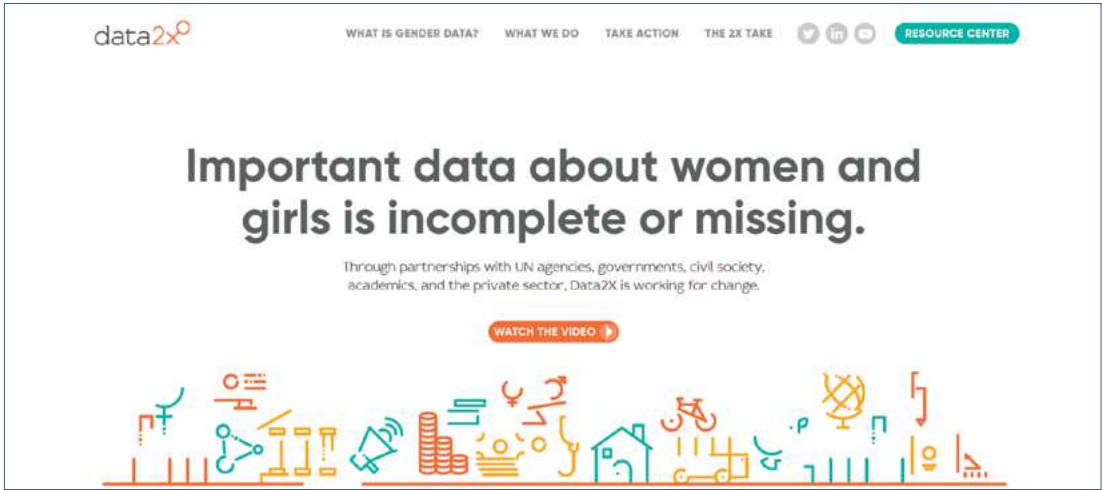
Bilgi Edinme Hakkı Kanunu kapsamında yapılan başvuruya ilişkin kurumun yanıtından memnun olunmadığında, cevabın tebliğ edildiği tarihten itibaren 15 gün içerisinde Bilgi Edinme Değerlendirme Kurulu'na itiraz edebilir ya da idari yargıya başvuruda bulunabilirsiniz.

Ayrıca, ilgili sitede yer alan “50 Soruda CİMER” başlıklı rehber de veri talebinde bulunurken dikkat edilmesi gereken konulara yönelik detayları içermektedir.

<https://www.cimer.gov.tr/50sorudacimer.pdf>

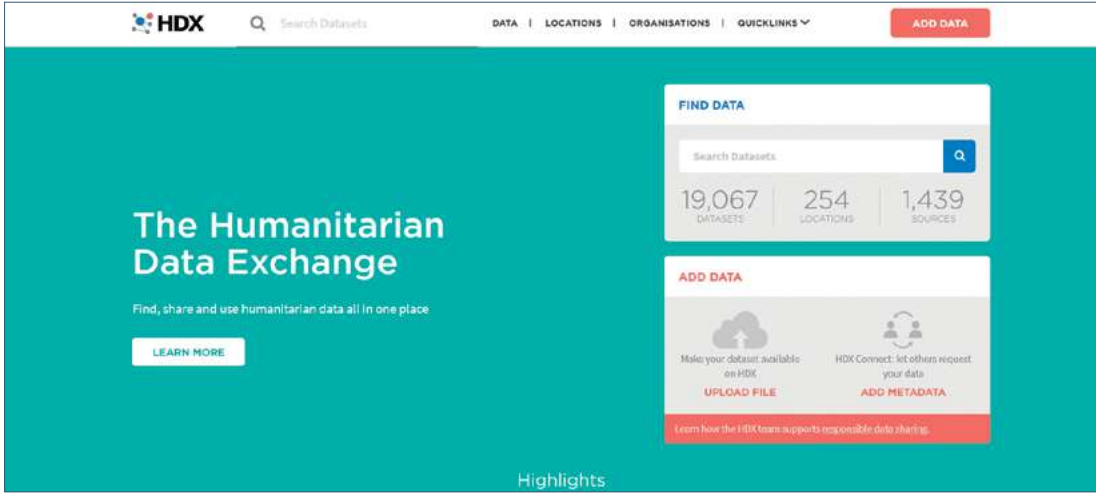
## Kendi Verisini Toplayan STÖ Örnekleri

1) <https://data2x.org/> : Cinsiyet odaklı veri toplayan bu STÖ, topladığı verilerle özellikle eşitsizlikleri görünür kılmaktadır. Toplanan veriler, kız çocukları ve kadınların yaşadığı sorunları çözmeye yönelik politika geliştirmelerinin yolunu açmaktadır. Ancak tarafsız veriye erişim hala pek mümkün olmadığı için bu örgüt özellikle toplumsal cinsiyet verisi üzerinden bu sorunu görünür kılmayı hedeflemektedir.



Bu örnek üzerinden incelendiğinde görülüyor ki, insanların yaşamlarındaki ilerlemeyi izlemek için tarafsız veriler her zaman çok önemlidir; ancak bu verilerde dünya genelinde eksikler de mevcuttur. Dolayısıyla kadınların ve kız çocuklarının hayatlarının ve karşılaştıkları kısıtlamaların yalnızca bir kısmını görebiliyoruz. Neticede toplumsal cinsiyet verilerinde hala daha, gözle görülür oranda boşluklar mevcuttur. <https://data2x.org/> de bu amaç için veri toplamakta ve verilerini erişime açmaktadır.

2) <https://data.humdata.org/> OCHA tarafından geliştirilen İnsani Yardım Veri Değişimi (HDX) platformu, kriz durumunda ve kuruluşlar arasında veri paylaşımı için oluşturulmuş, veri toplayan açık bir platformdur. Platform, insani yardım verilerini bulmayı ve analiz etmeyi amaçlamaktadır. İnsani Yardım Veri Değişimi Platformu, 200'den fazla ülke ve bölgeye dair veri sağlamaktadır.



3) <https://www.deneyehayir.org/> Hayvan deneylerinin sonlandırılarak, alternatif yöntemlere geçiş amacıyla kurulan Türkiye'deki ilk anti-viviseksiyonist sivil toplum örgütüdür. Örgüt, hayvan deneylerinin sonlandırılması ve hayvanların sistematik olarak nasıl öldürüldüklerini görünür kılmak için yıl bazında, Tarım ve Orman Bakanlığı verilerini derleyerek topluyor.



## Data Miner ile Veri Toplamak

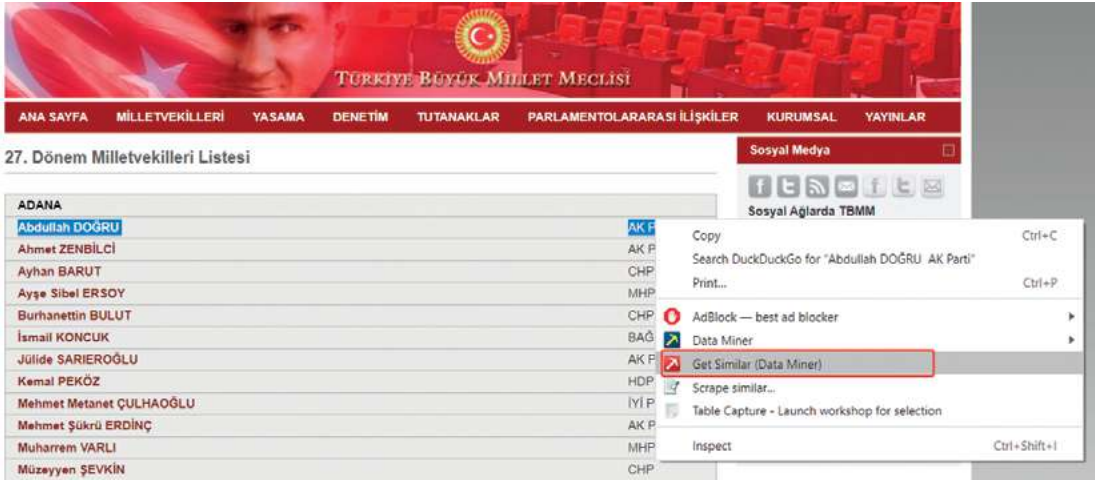
**Veri kazıma**; webden veri toplamak, araştırmak; verileri bir araya getirmek demektir. Veri kazıyanlara da çoğunlukla kazıyıcı adı verilir. Kazımak (veri madenciliği) görselleştirmek istediğiniz veriye ulaşamadığınızda alternatif bir teknolojidir. Veri madenciliği; veriye erişim sorunu, verinin doğrulanması ihtiyacı, doğru bilgiye ulaşma ihtiyacı gibi ve özellik görselleştirmek/analiz etmek istenilen veriye ulaşılamadığında kullanılan alternatif bir yöntemdir. Kitabın bu bölümünde **Data Miner** gibi araçlarla pratik bir şekilde nasıl veri kazıyacağınız yer alacak.

Data Miner bir Google Chrome eklentisidir. Web sitelerinden veri kazıma ve kazınan verinin CSV dosyası ya da Excel dosyası olarak e-tabloya aktarılmasını sağlar. 50,000+ üstü ücretsiz ön sorgu yapabilirsiniz. Herhangi bir koda ihtiyaç duymadan 15,000 + üstü popüler site için kullanılabilir.

### Örnek Data Miner ile Kazıma

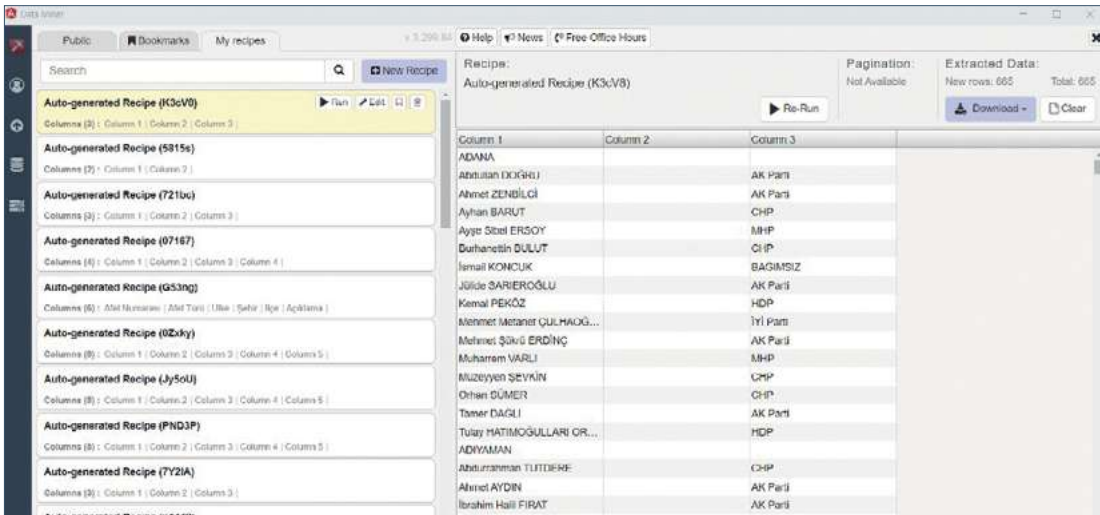
- Meclis 27. dönem vekil verilerini bu [https://www.tbmm.gov.tr/develop/owa/milletvekillerimiz\\_sd.liste](https://www.tbmm.gov.tr/develop/owa/milletvekillerimiz_sd.liste) adresten kazıyorsunuz.

İlk ya da ikinci ismi seçiniz. Fareyi sağ tıklayıp **Get Similar (Data Miner)**'ı seçiniz.

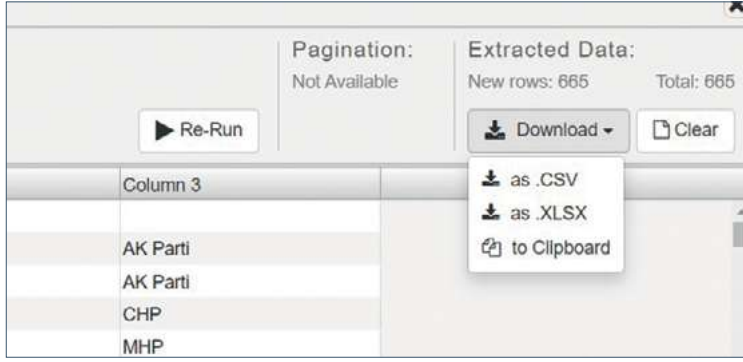


Şekil 44: TBMM 27. dönem milletvekilleri listesi kazıma ekran görüntüsüdür.

- Tüm HTML sayfası “Data Miner” veri kazıma sayfasına yüklenecektir.

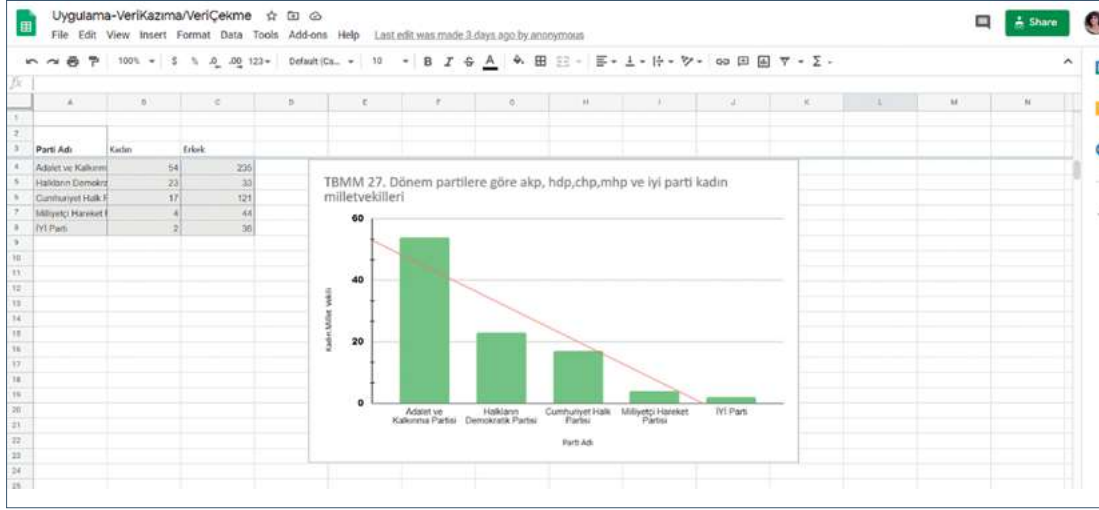


Şekil 45: TBMM 27. dönem milletvekilleri listesi kazıma ekranına aktarılmış veri görüntüsüdür.



**Şekil 46:** Kazınan veriler CSV, XLSX olarak indirilebilir ya da kopyalanarak e-tabloya ya da başka bir tabloya eklenebilir.

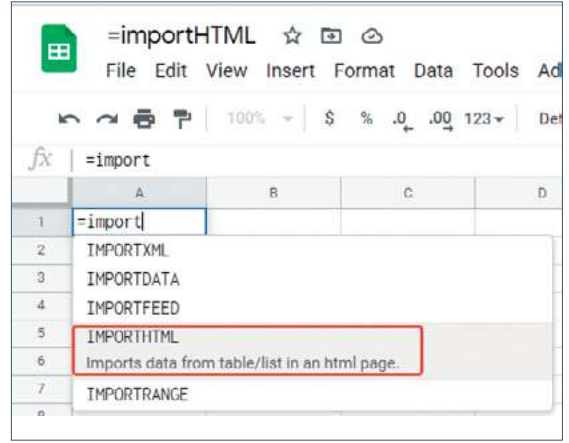
- Partilere ait vekiller isimleri, bağlı oldukları partiler ve iller verisi yer almakta.
- Download bölümünden veriler **csv**, **xlsx** ya da kopyalanarak analiz edilecek bölüme yüklenebilir/kopyalanabilir.
- Ayrıca aynı sayfaya otomatik olarak veri çekmesi için komut da verilebilir.



**Şekil 47:** TBMM sitesinden “Data Miner” ile kazınan veriler e-tabloya aktarılıp temizlenir / filtrelenip analiz edilip, bu şekilde görselleştirilebilir.

## Elektronik Tablo ile HTML Sayfadan Veri Çekmek

Öncelikle bir Gmail hesabına giriş yapmalı, ardından Google Drive (drive.google.com) adresinden Yeni Menü>Google e-tablolara ulaşmalıyız. Sonrasında karşınıza boş bir e-tablo dosyası açılacak, dosyanın adını değiştirerek başlayabilirsiniz. İlk hücreye eşittir, =IMPORTHTML şeklinde yazarak HTML sayfasındaki bir tablodan veya listeden verileri çekebilir ve içe aktarabilirsiniz.



Şekil 48: E-tabloda import fonksiyonu ile çekilecek tablo, liste, veri, feed vb. ekran görüntüsüdür.

Örneğin, dünya çapında en çok satan kitaplar listesi verisini e-tabloya aktaracağım. Bunun için IMPORTHTML(url; sorgu; izin) şeklinde söz dizimini oluşturmalısınız.

Formül: =IMPORTHTML("https://tr.wikipedia.org/wiki/D%C3%BCnya\_%C3%A7ap%C4%B1nda\_en\_%C3%A7ok\_satan\_kitaplar\_listesi";"table";3)

Türkçe adı	Özgün adı	Yazar(lar)	Orjinal dili	İlk basım yılı	Yaklaşık satış
"Aslan, Cadi ve Dolap"	"The Lion, the Witch and the Wardrobe"	C. S. Lewis	İngilizce	1950	85 milyon(7)
"Ayşe"	"Sha"	H. Rider Haggard	İngilizce	1807	83 milyon(8)
"Da Vinci Şifresi"	"The Da Vinci Code"	Dan Brown	İngilizce	2003	80 milyon(9)
"Dağın ve Zengin Çi Yeşil"	"The Catcher in the Rye"	J. D. Salinger	İngilizce	1951	70 milyon(10)
"Donkişot"	"The Catcher in the Rye"	Paulo Coelho	Portokizce	1996	65 milyon(11)
"Sımsıkı"	"Steps to Christ"	Ellen G. White	İngilizce	1922	60 milyon(12)
"Lolita"	"Lolita"	Vladimir Nabokov	İngilizce	1955	50 milyon(14)
"Hadi"	"Havds Lehr: and Wandersjörns"	Johanna Spyri	Almanca	1880	50 milyon(15)
"Bebek Bakımı ve Çocuk"	"The Common Sense Book of Baby"	Dr. Benjamin Sp	İngilizce	1948	50 milyon
"Yeşilin Kız Anne"	"Anne of Green Gables"	Lucy Maud Mont	İngilizce	1906	50 milyon(16)
"Siyah İnci"	"Black Beauty: His Grooms and Con"	Anna Sewell	İngilizce	1877	50 milyon(17)
"Gülen Ad"	"Il Nome della Rosa"	Umberto Eco	İtalyanca	1980	50 milyon(18)
"Kartal"	"The Eagle Has Landed"	Jack Higgins	İngilizce	1975	50 milyon(19)
"Watership Toposu"	"Watership Down"	Richard Adams	İngilizce	1972	50 milyon(20)

Şekil 49: IMPORTHTML ile e-tabloya aktarılmış dünyada en çok satan kitap listesi ekran görüntüsüdür.

Gördüğünüz üzere dünyada en çok satan kitaplar listesi e-tabloya aktarıldı. Bu yöntem sadece HTML kazıma sorgusu ile sınırlı değildir. XML, RANGE, FEED, DATA sorgusu yapmak mümkündür.

**IMPORTXML:** Çeşitli yapılandırılmış veri türlerinden (XML, HTML, CSV, TSV ve RSS / ATOM XML (feed'leri dâhil) verileri çeker.

**IMPORTRANGE:** Belirtilen bir e-tablodan bir hücre aralığını e-tabloya çeker.

**IMPORTFEED:** RSS veya ATOM feed'ini çeker.

**IMPORTDATA:** Belirli bir url'deki verileri .csv (virgülle ayrılmış değer) veya .tsv (sekmeyle ayrılmış değer) biçiminde olanları çeker.

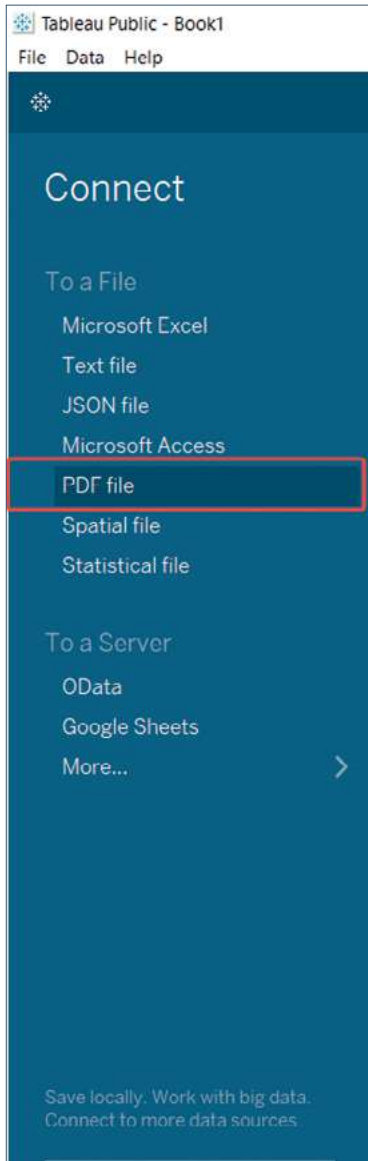


## Tableau Public ile PDF'den Veri Kazımak

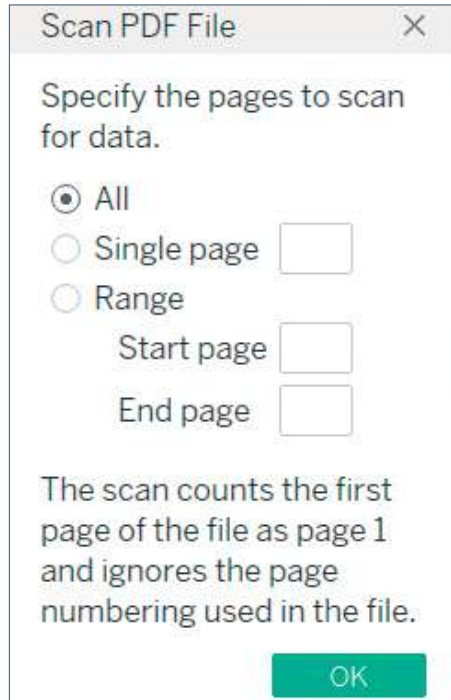
<https://public.tableau.com/en-us/s/> İş analitiği ve veri görselleştirme aracı olmasının yanında PDF'leri CSV dosya tipine dönüştürüp veri toplamanıza yardımcı olur. Örneğin; 2020 Kasım-Aralık tarihi aralığında raporlanan 71 sayfalık merkezi yönetim bütçe yükümlüklerini, (2006-2023) pdf dosyasında yer alan verileri, açık format bir dosya tipine dönüştürmek istiyorsunuz. İlgili PDF dosyasını, masaüstüne indirmelisiniz öncelikle.

<https://ms.hmb.gov.tr/uploads/2020/12/2020-Kasim-Aylik-Butce-Bulteni-min.pdf>

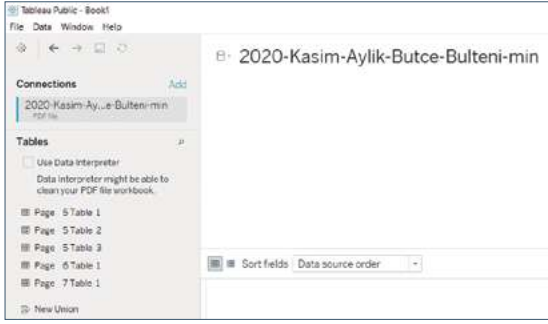
“Tableau Public” aracını açtıktan sonra “PDF file” bölümünü seçip indirdiğiniz bütçe dosyasını yüklemelisiniz.



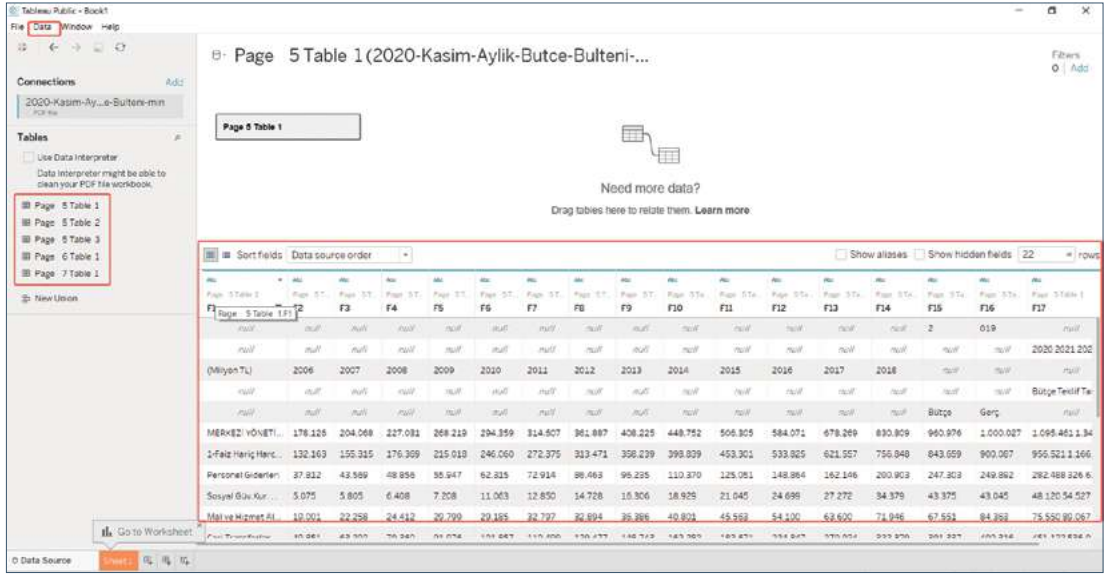
Yükledikten sonra editör size aralık girerek ya da sayfa belirleyerek PDF'te ilgili sayfaları seçmenizi önerecek. Bizim örneğimizde, 5. sayfadan başlayan ve 7. sayfada biten bölümü kazımamızı istiyoruz.



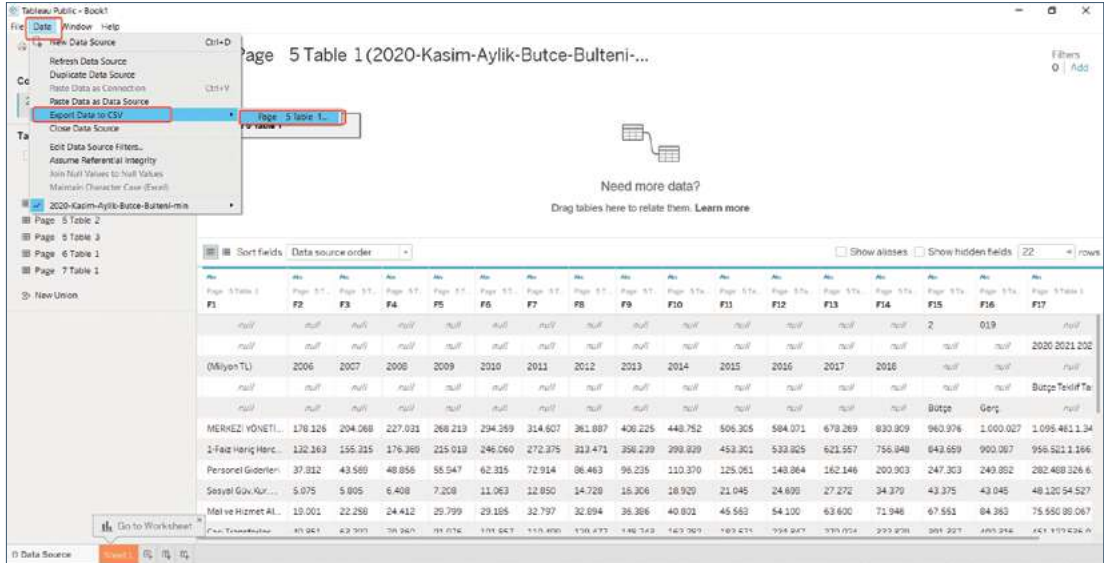
Ardından DataSource yani veri kaynağına sayfaların ak-



tarıldığı bir bölümle karşılaşıyoruz. Görüldüğü üzere Tables bölümünde seçilen sayfalar yer alıyor. Ardından ilk tablo tıkladığında “sort fields” bölümüne eklenecektir.



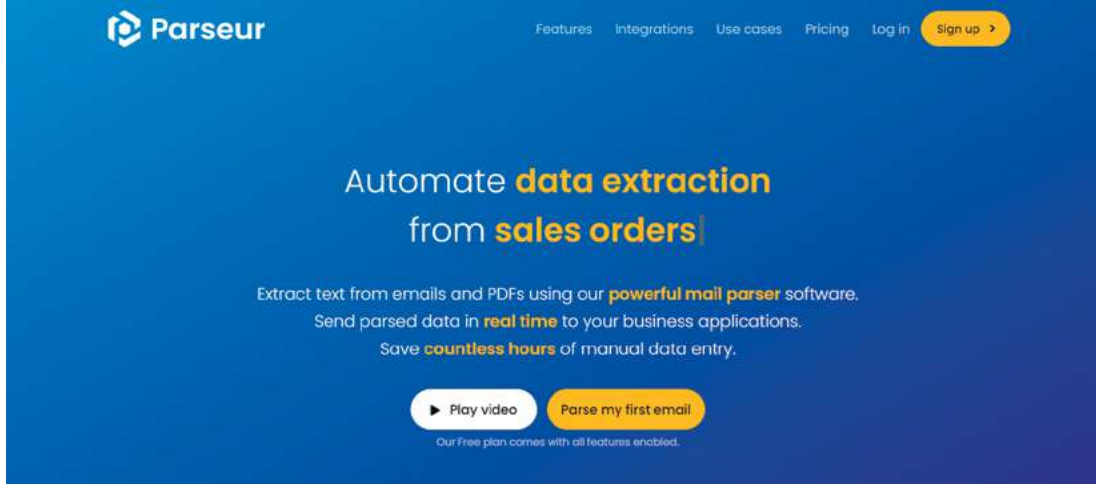
Data bölümü tıkladığında, verileri CSV olarak indir seçilip kullanılabilir. Bu şekilde farklı PDF'lerden de veri toplayabilir, topladığımız verileri yapılandırabilirsiniz.



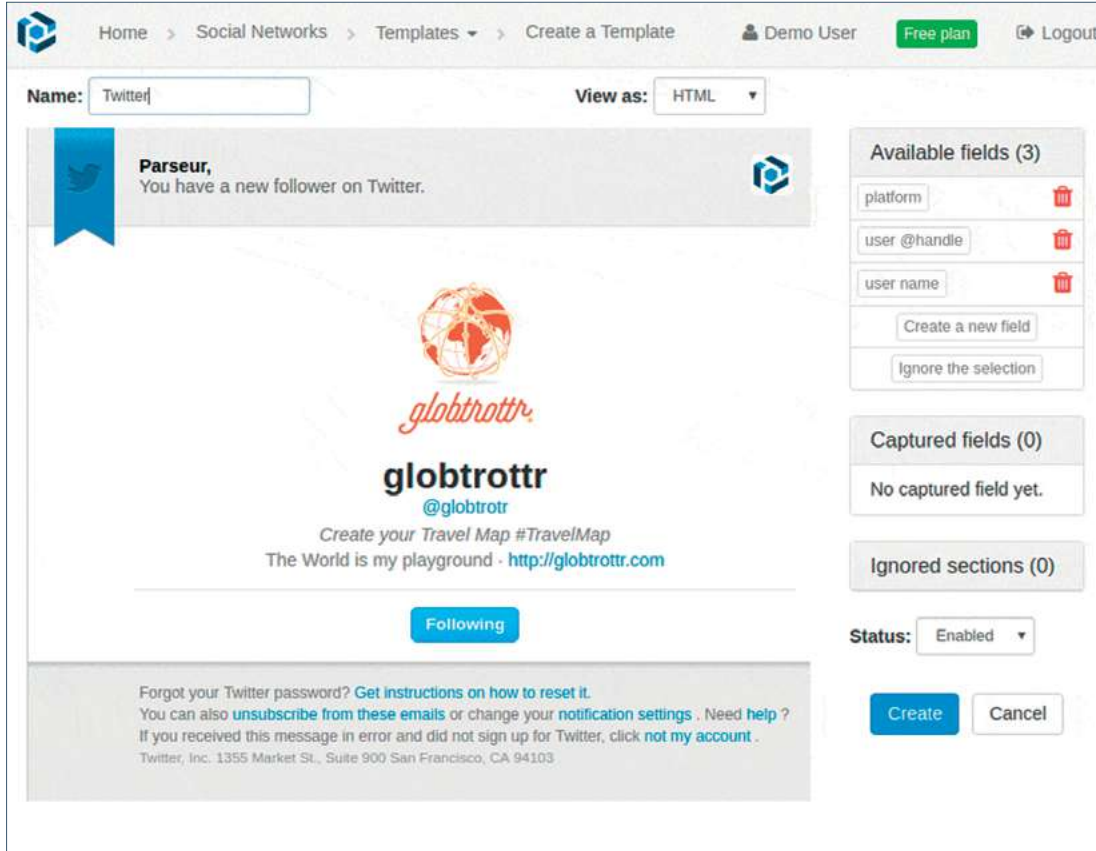


## Parseur ile Veri Toplama

<https://parseur.com/> : Verileri ayrıştırılan araç. E-postalardan ve PDF'lerden metin çıkarmayı ve yapılandırılabilir bir formata dönüştürmeyi sağlar. Ayrıştırılmış verileri gerçek zamanlı olarak size yollar.



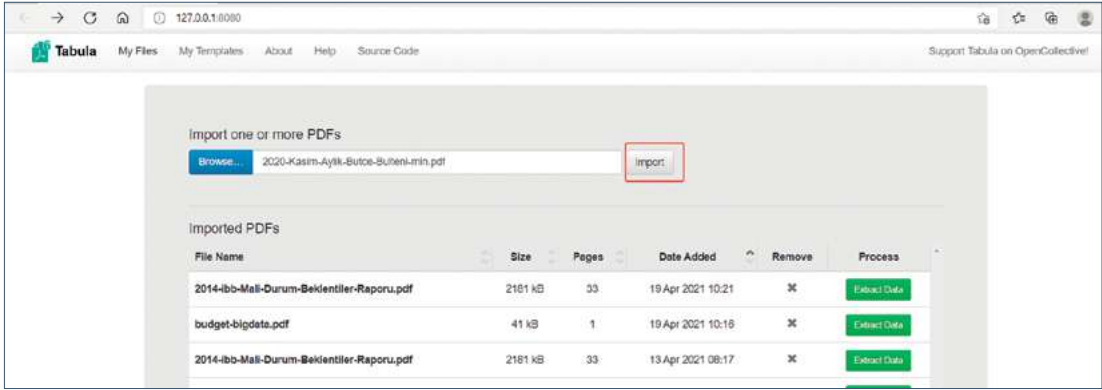
The image shows the Parseur website landing page. The header includes the Parseur logo and navigation links: Features, Integrations, Use cases, Pricing, Log in, and a Sign up button. The main headline reads "Automate data extraction from sales orders". Below this, it states: "Extract text from emails and PDFs using our powerful mail parser software. Send parsed data in real time to your business applications. Save countless hours of manual data entry." There are two buttons: "Play video" and "Parse my first email". A small note at the bottom says "Our free plan comes with all features enabled."



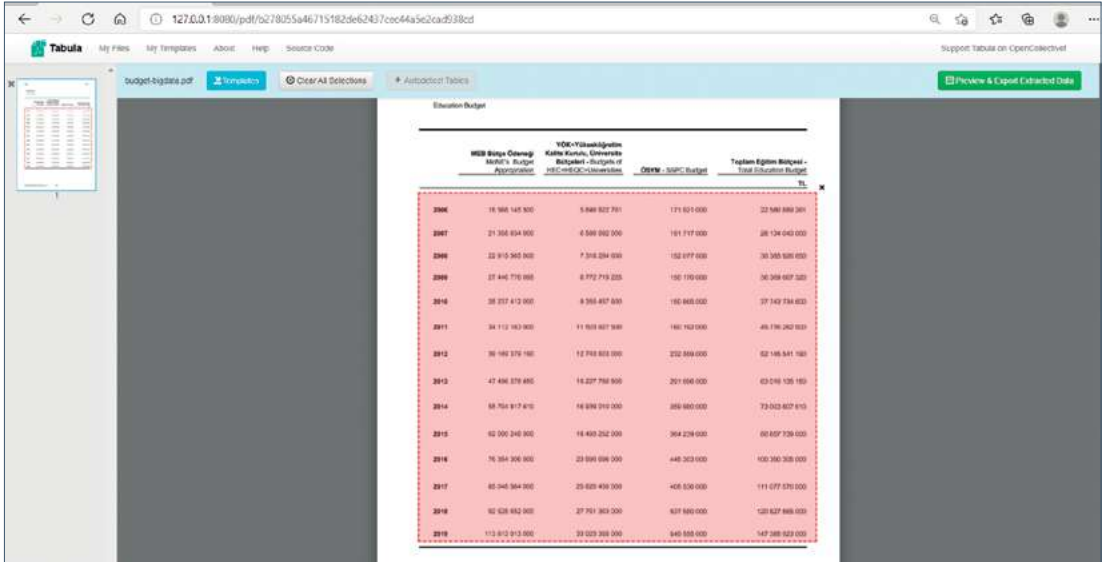
The image shows the Parseur interface for creating a template. The top navigation bar includes Home, Social Networks, Templates, and Create a Template. The user is logged in as Demo User on a Free plan. The main area shows a Twitter notification: "Parseur, You have a new follower on Twitter." Below the notification is a profile card for "globtrottr" (@globtrottr) with a "Following" button. The right sidebar shows "Available fields (3)" with a list: platform, user @handle, and user name. There are buttons for "Create a new field" and "Ignore the selection". Below that, "Captured fields (0)" shows "No captured field yet." and "Ignored sections (0)". The "Status" is set to "Enabled". At the bottom, there are "Create" and "Cancel" buttons.

## Tabula ile PDF'den Veri Kazımak

<https://tabula.technology/> PDF dosyalarındaki verileri kazımanıza yardımcı olur. Kullanımı son derece kolaydır. Verileri csv, json gibi birçok formatta indirmenizi sağlar. Adobe Acrobat'tan farklı olarak gazeteciler, STÖ'ler için geliştirilmiş açık kaynaklı ve ücretsiz bir araçtır. Tabula sitesinden bilgisayarınıza uygun olan sürümü indiriniz. Kurulumu yaptıktan sonra Tabula, tarayıcınızda sekme açarak çalışacaktır. Veri kazımak istediğiniz pdf dosyasını seçtikten sonra "Process" edilmesini beklemek gerekiyor. Pdf dosyasının büyüklüğüne göre bu işlem biraz zaman alabilir. Açılan sekmede, "Browse" ile bilgisayarınızdaki pdf'leri seçmeniz gerekiyor.



Kazımak istediğiniz verileri, fare ile seçtikten sonra "Preview & Export Extracted Data" butonuna tıklayıp tablo haline getirilmiş verinizi görebilirsiniz.



Karşınıza *default* olarak CSV formatında çıkacaktır; ancak "Export Format" seçeneğinden json gibi farklı formatlar ya da çok fazla veriyle çalıştığınız zaman işinize yarayabilecek zip seçeneğini kullanabilirsiniz. Export'a tıklayarak verinizi indirebilirsiniz.

The screenshot shows the Tabula web interface. The 'Export Format' dropdown menu is set to 'CSV'. The 'Export' button is visible. Below the menu, there is a 'Preview of Extracted Tabular Data' table with 5 columns and 14 rows of data.

Year	Column 1	Column 2	Column 3	Column 4	Column 5
2006	16 558 145 500		6 846 822 761	171 921 000	22 586 685 261
2007	21 356 634 000		6 566 662 000	191 717 000	28 134 042 000
2008	22 910 560 900		7 318 284 800	132 077 000	30 360 926 600
2009	27 446 778 095		6 772 719 225	150 170 000	36 369 667 320
2010	26 237 412 000		6 355 457 600	150 866 000	37 743 734 600
2011	34 112 163 000		11 503 927 600	160 192 000	45 776 282 500
2012	39 139 379 190		12 743 603 000	232 059 000	52 143 541 190
2013	47 496 378 650		15 227 760 500	291 596 000	63 016 135 150
2014	55 734 817 510		16 939 010 000	359 960 000	73 003 807 610
2015	62 000 248 000		18 493 252 000	364 239 000	80 857 739 000
2016	76 334 306 000		23 590 696 000	445 303 000	100 390 305 000
2017	86 046 584 000		25 620 450 000	408 536 000	111 077 570 000
2018	92 526 652 000		27 761 363 000	537 500 000	120 827 565 000
2019	113 813 012 000		33 023 355 000	549 555 000	147 385 923 000

The screenshot shows the Tabula web interface with the 'Export Format' dropdown menu open. The menu options are: CSV, TSV, JSON (tab-separated), JSON (table), and ZIP of CSVs. The 'CSV' option is highlighted.

Dosyanızı Google E-Tablolar'ı kullanarak Excel formatına dönüştürebilirsiniz. E-tablolara girerek Dosya > İçerik Aktar şeklinde ilerleyip csv verisini kopyalayabilirsiniz.

## Workbench Data ile Veri Toplama

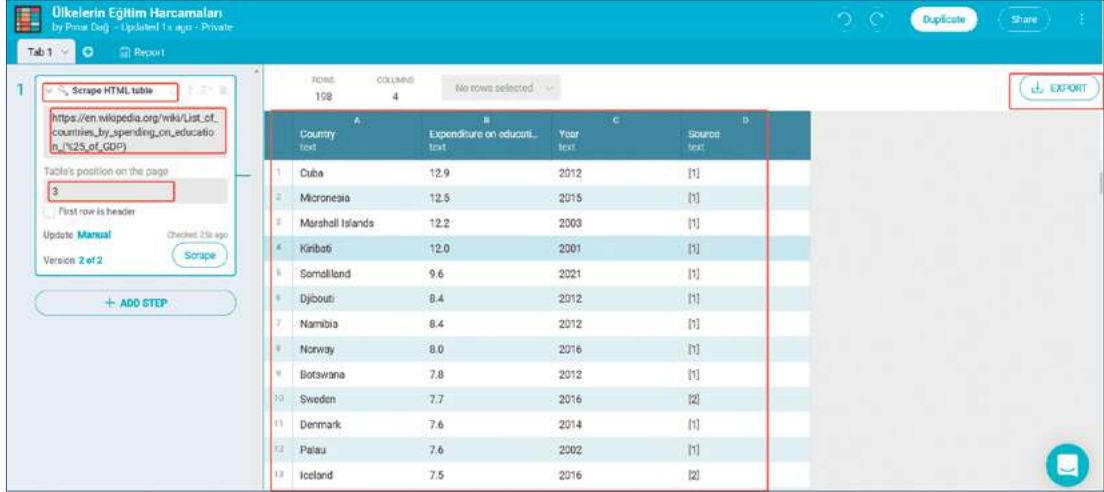
<https://workbenchdata.com/> aracı ile URL'den html veya html tablolardan veri çekmek/mümkündür.

The screenshot shows the Workbench Data web application interface. The 'CHOOSE A DATA SOURCE' section is active, with a search bar containing 'html'. Two options are displayed:

- Load HTML from URL**: Scrape the entire HTML from a webpage to be able to search its content or monitor changes.
- Scrape HTML table**: Load a table from a webpage and update it when the webpage changes.

Örneğin; ülkelerin eğitime harcadığı bütçenin listesini CVS olarak indirmek ya da çalışmak istiyorsunuz. İlk önce çalışma bölümünü açıp linki paylaşmalısınız.

[https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_spending\\_on\\_education\\_\(%25\\_of\\_GDP\)](https://en.wikipedia.org/wiki/List_of_countries_by_spending_on_education_(%25_of_GDP)) Uygulamanın/sayfanın/sitenin hangi sayfada yer alan tabloyu çekmesini istiyorsanız onu yazmalısınız. Bu örnekte, 3. Sayfanın verisini çekmek istiyoruz. Ardından o sayfanın verileri çalışma ekranına aktarılmış durumda ekrana geliyor.




Country	Expenditure on education	Year	Source
Cuba	12.9	2012	[1]
Micronesia	12.5	2015	[1]
Marshall Islands	12.2	2003	[1]
Kiribati	12.0	2001	[1]
Somalia	9.6	2021	[1]
Djibouti	8.4	2012	[1]
Namibia	8.4	2012	[1]
Norway	8.0	2016	[1]
Botswana	7.8	2012	[1]
Sweden	7.7	2016	[2]
Denmark	7.6	2014	[1]
Palau	7.6	2002	[1]
Iceland	7.5	2016	[2]

Export bölümü tıklandığında, görüldüğü üzere eğitim harcamaları listesi CSV ya da Json Feed olarak indirilebilir ve veri olarak kullanıma hazırdır.




## DataBasic.io ile Veri Toplama


<https://www.databasic.io/en/>: Veriyle temel seviyede çalışmak için geliştirilmiş olan araç, özellikle <https://www.databasic.io/en/wordcounter/> bölümüyle metin analizi yapip veri oluşturmanıza olanak sağlıyor. Verileri CSV olarak da indirebiliyorsunuz.




DataBasic is a suite of easy-to-use web tools for beginners that introduce concepts of working with data. These simple tools make it easy to work with data in fun ways, so you can learn how to find great stories to tell.




WordCounter analyzes your text and tells you the most common words and phrases.




WTFcsv tells you WTF is going on with your .csv file.



SameDiff compares two or more text files and tells you how similar or different they are.



ConnectTheDots shows you how your data is connected by analyzing it as a network.



**word COUNTER**

Words used in Metallica's lyrics

These are lyrics from many of Metallica's songs. They were scraped from lyrics2Live by Rahul Bhargava in 2014.

Here is a picture of the words used most often in your document. Words used more often are bigger, and ones used less often are smaller. This picture, called a "word cloud", is helpful to get a sense of the most used words in a document.

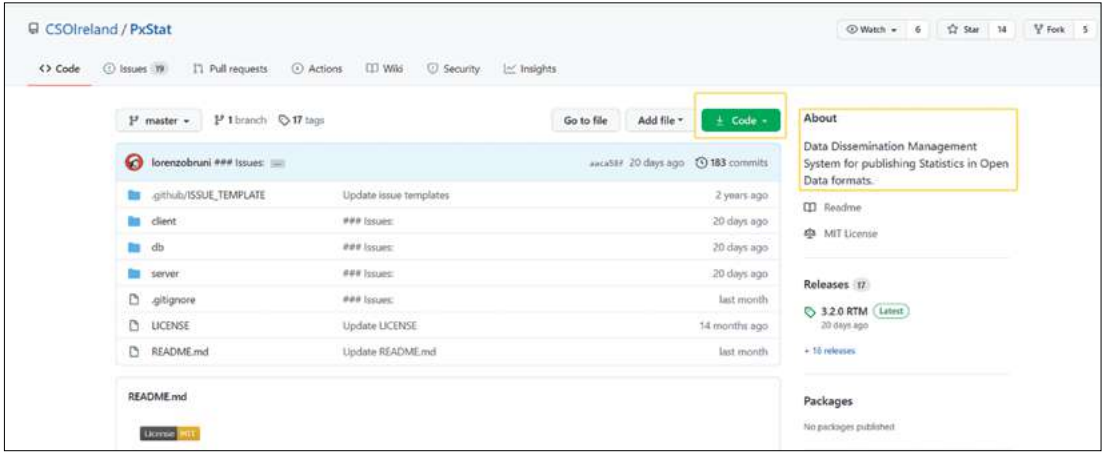
TOP WORDS		BIGRAMS		TRIGRAMS	
Word	Frequency	bigram	Frequency	trigram	Frequency
yeah	373	I am	112	na na na	69
I'm	307	and I	99	hey hey hey	39
never	246	in the	81	tick tick tick	34
see	219	I can't	78	die die die	28
time	172	na na	72	yeah yeah yeah	26
take	153	of the	71	within my hands	28
oh	148	yeah yeah	66	all within my	27
can't	138	to me	64	this is the	26
one	138	what you	62	better than you	26
away	132	to die	60	I do believe	26
do	129	we for	59	I won't believe	25
die	119	to be	68	don't say so	24
make	118	what do	59	try to me	24
up	112	the end	58	to me oh	24
hey	112	hey hey	55	see oh baby	24
but	110	oh you	54	it's my world	24
come	108	to be	48	the end of	24
new	104	tick tick	47	shoot me again	24
in	104	what I	44	no and so	23

## Github Hesabı Oluşturmanın Önemi

Github (<https://github.com/>), 2008 yılında kurulmuş olan web tabanlı bir depolama servisidir. Ortak çalışmalarda en çok tercih edilen platformdur. STÖ'ler geliştirdikleri projeleri ortak çalışmaya açmak istediklerinde ya da topladıkları verileri erişime açmak istediklerinde sıkça kullanılmaktadırlar. Açık kaynaklı projeler için tüm servislerin ücretsiz ve şeffaf olması kullanımı arttırmaktadır.



## Örnek çalışma



<https://github.com/CSOIreland/PxStat>

## Örnek Kurum ve Kişi Github Profili UNDP Github sayfası



<https://github.com/undp>



Search or jump to...

Full requests Issues Marketplace Explore

Overview **Repositories** Projects Packages

Popular repositories Customize your pins

**detective.io**  
Forked from jshugart/detective.io  
Detective.io is a platform that forks your investigation and lets you make powerful queries to mine it. Simply describe your field of study and detective.io builds the input interface as well as a ...  
Python

**community-resources**  
Forked from school-of-data/community-resources  
Templates for sharing knowledge within and outside the community

**eu-structural-funds**  
Forked from es-data/eu-structural-funds  
The EU structural funds datasets on regional and national level (in progress).  
Jupyter Notebook

**gazetecileriniazima**  
Dost web tabanı arşivden tam programına arşivine kadar erişilebilirlik ve baskılarla kullanılarak online katta yedince web sayfası belge veya elektronik tablodan nasıl bilgi toplayabileceğin...

**firstpythannotbook**  
Jupyter Notebook

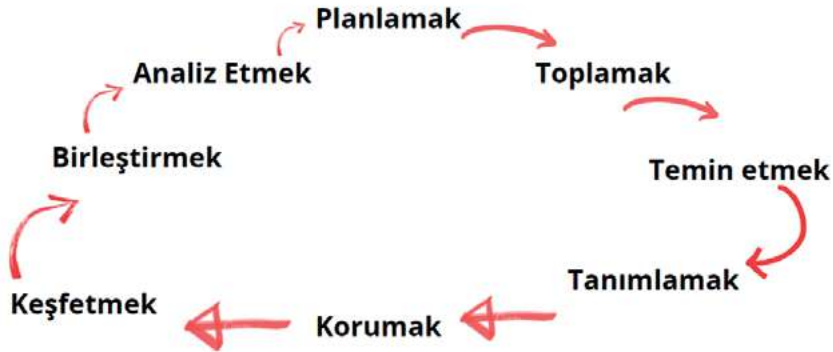
**pinardag**  
pinardag  
Lecturer  
Edit profile  
As 11 followers · 16 following · 19 stars  
pinardag@khas.edu.tr  
@pinardag

<https://github.com/pinardag>

## 4. BÖLÜM

### Veri Kalitesi ve Veri Doğrulama

[The Data Warehousing Institute](#) (TDWI) araştırmasına göre, sadece Britanya’da veri kalitesindeki problemlerin maliyeti yılda en az 600 milyar dolar olarak rapor edilmiştir. Verilerin kalitesinden dolayı alınamayan doğru kararlar neticesinde oluşan mağduriyet, fırsatları yakalamayı güçleştirmekle beraber kurumlar için mağduriyet de yaratmaktadır. Ancak çok iyi bir veri yönetiminiz olsa bile, veri kalitesine yönelik sorunların önünü almak çok mümkün olmayabilir. Bu doğrultuda, öncelikle yapılması gereken veri kalitesini etkileyen olguları anlamaktır. Yaygın olarak insan kaynaklı sorunların başında; verilerde oluşan format sorunları, standart belirleme eksikliği, tutarsızlık, eksik veri, tekrarlı içeriğin olması gibi detaylar bildirilmektedir. Ancak sadece insan kaynaklı değil, teknoloji kaynaklı gelişen veya süreç aşamasında da gelişen birtakım sorunlar sebebiyle veri kalitesinin bozulması mümkündür. Bu sebeple STÖ’ler topladıkları ya da toplamayı planladıkları verilerin öncelikle yönetim planını iyi kurgulamalıdır. Kurum, verilerin üretiminden sorumlu yetkin kişileri belirlemelidir. Verilerin toplanma ya da işlenme aşamasında sorumluluklar iyi belirlenmeli ve aşamalı kontrol mekanizması sağlanmalıdır. STÖ hangi alanda veri toplamayı ya da üretmeyi hedefliyor ise bu konudaki politikası açık olmalıdır. Kullanılacak teknoloji, pratik ve kullanıcı dostu olmalıdır.



Şekil 50: DataONE Veri Yaşam Döngüsü (<https://old.dataone.org/data-life-cycle>)

### Veri Seti Doğrulama

Kullanım sürecinde ise veri setlerinin doğruluğu önem taşır. Hızlı şekilde veri setinin doğrulanması şöyle yapılabilir:

- Diğer tüm kaynaklar gibi veri setlerine de aynı şekilde yaklaşın. Hiçbir kaynağa körü körüne güvenilmemelidir. Röportaj yaptığınız bir kişinin verdiği bilgileri ve ifade ettiklerini nasıl doğruluyorsanız, aynı şeyi veri setleri için de yapmalısınız. Verinin nasıl ve kim/hangi kurumlar tarafından oluşturulduğunu doğrulamalısınız.

- Üst veriyi okuyun. Veri setiyle çalışmadan önce, genelde insanların pek okumadığı, o ince, gizli, küçük bölümleri okuyun. Üst veri ve metodolojiler, eğer varsa, veriler eksik veya tahmin edilen verilere yönelik bilgileri belirtebilir. Bazı tahminler hatalı varsayımlara dayanabilir.
- Veri toplama işlemini yeniden yapılandırın. Metodolojiyi okuduktan sonra, belirli göstergenin sosyal ve politik bağlamı ışığında veri toplama sürecinin güvenilirliğini değerlendirin. Ülke çapında veri setleri için tek bir organizasyonun verileri toplamış olup olmadığını veya farklı ofislerin bilgi paylaşım paylaşımadığını kontrol edin.
- Elektronik tabloları test edin. En düşük ve en yüksek girişleri inceleyin ve mantıklı mı şüpheli mi geliyor inceleyin, kontrol edin. Neyin eksik olduğunu değerlendirin: Boş olması gereken boş satırlar var mı? Yalnızca örnek verileri görüyorsanız, neyin dışarıda bırakıldığı ve neden net olduğu açık mı? Tek tek girişleri rastgele doğrulayın: Bir veya iki kayıt seçin ve veri grubunun dışındaki bir arama yoluyla bunları ayrıca doğrulayın.
- Güncel araçları kullanarak kendinizi koruma altına alın. Doğrulama yapanlar kullanıcı dostu olmayan araçları kullanabiliyorlar. Örneğin CSV / Excel dosyaları yerine PDF olarak yüklenen e-tablolar veya bir seferde sadece bir sorgu ile veri tabanlarına yüklenen veriler. Bu durum bazen hatalar doğurabiliyor. Bu sebeple Import.io veya Dataminer gibi Chrome'un veri kazıma araçlarını kullanmayı öğrenin.

## E-Tablo ile Veri Kaydını Doğru Oluşturmak/Verinin Doğruluğunu Sağlamak

Toplanan verilerin girişini yaparken, verilerin doğru girilmesini sağlamak önemlidir. Doğrulama, orijinal kaynak belgedeki ya da veritabanındaki verilerin sisteme girdiğiniz verilerle tamamen aynı olup olmadığını kontrol etmek anlamına gelmektedir. Özellikle veri topluyorsanız ya da sık sık kazıdığınız veri tabanları varsa buna dikkat etmeniz gerekmektedir. Veri taşıma işleminden sonra doğruluk ve tutarsızlıklar açısından farklı veri türlerinin kontrol edilmesidir. Verilerin mümkün olduğunca az hata içerdiğinden emin olmak için doğrulama yapılabilir. Verinin kaydını hatasız yapmak yani verinin doğru kayıt altına alındığını sağlamak önemlidir. Çoğunlukta bu süreçte veri kalitesinde sorun oluşur, çok zaman alır. Yani, verilerin doğru girilmesini sağlamak da verileri girerken yürütülen süreçler de maliyetlidir. Peki, sivil toplum örgütü çalışanı veri doğrulama yöntemlerini bilmeli midir? STÖ çalışanları verileri doğrulayarak çalışmak zorunda değildir (bu bazen mümkün de olmayabilir) / bilmek zorunda değildir ancak kendi veritabanınızı oluşturursanız (**propública (nefret endeksi)**, **538 (futbol veritabanı)**, **bianet (kadın cinayetleri)** örneklerinde olduğu gibi veri toplarsanız) mutlaka yöntemleri bilmeniz gerekir.

Veri girişi yaparken sıklıkla karşılaşılan birkaç standart hata daha vardır: Bunlardan başlıcaları transkripsiyon (kayıt) hataları ve transpozisyon (aktarma) hatalarıdır.

**Kayıt Hataları:** Verilerin manuel olarak girilmesinde insan kaynaklı hatalar artmaktadır (Seçim örnekleri, sandık başında çalışanlar vb.). Bu hatalar, kullanıcının ne yazıldığını veya ne söylendiğini yanlış anlaması, acele etmesi ve detaylara yeterince dikkat etmemesi gibi sebeplerden kaynaklanır.

**Aktarma Hataları:** Rakamların veya harflerin sırası karıştırıldığında meydana gelir. Örneğin; 78 yerine 87 yazmak gibi.

### Veri Doğrulama Yöntemleri

**Çift giriş:** Bu, verilerin iki kez girilmesi ve iki girişin karşılaştırılması anlamına gelir.

**Verilerin düzeltilmesi:** Bu işlem, başka bir kişinin orijinal belgeye göre veri girişini kontrol etmesini gerektirir. Hem sıkıcı hem de pahalı bir işlemdir.

**Orijinal dokümandan verilerin kontrol edilmesi yerine ekrandan bu kontrolün yapılması:** Kayıt ve aktarma hatalarının tanımlanmasına yardımcı olur. Ayrıca, çift giriş tekniğine kıyasla zaman kazandırır. Göz yorgunluğu yaratır.

**Verinin kopyasının basılması ve orijinal doküman ile karşılaştırılması:** Bu en basit doğrulama yöntemidir. Çünkü her iki kopyayı yan yana koyabilir ve her ikisini de hatalar için tarayabilirsiniz. Bununla birlikte, kontrol edilmesi gereken büyük miktarda veri olması, durumu zorlaştırabilir. Ayrıca, çok hızlı bir şekilde taranırsa hatalar göz ardı edilebilir.

**Destek almak:** Zaman alıcı bir yöntemdir ancak ekip olarak yapılır ise hataları yakalama imkânı artar. Verilerin girişi/kaydı sizin tarafınızdan yapıldıysa, kontrolünün başkası tarafından yapılması sağlanmalıdır.

Ancak, bir gazeteci düzenli olarak bir veri girişi yapıyor ve veri editörleriyle ortak çalışıyorsa, e-tablonun veri doğrulama fonksiyonunu kullanarak yanlış veri girişini engelleyebilir ve bir uyarı da sağlayabilir.

Örnek bir uygulama üzerinden gitmek gerekirse; “En İyi 100 Türk Filmi” verisini e-tabloya aktardınız ve yapım yılını seçerek 1963 ile en yüksek değer 2030 aralığını belirlediniz. Veri girişleri bu aralık dışına çıktığında, giriş reddi uyarısı alacak ve geçersiz kayıt olduğu belirtilecektir. Bu işlemler otomatik kayıt araçları kullanılarak da yapılabilir. E-tablo üzerinden altı aşama üzerinden nasıl yapıldığını görebilirsiniz:

**Google  
E-Tabloyu  
Kullanarak  
Yanlış Veri  
Girişini  
Engelleme**

E-tablo:  
[https://docs.google.com/spreadsheets/d/1Kh101Hyh1139hXHje7eP2pRPYCs-3J\\_9eWCfwR9mN9c/edit#gid=0](https://docs.google.com/spreadsheets/d/1Kh101Hyh1139hXHje7eP2pRPYCs-3J_9eWCfwR9mN9c/edit#gid=0)

Kullanılacak veri seti:  
[https://tr.wikipedia.org/wiki/En\\_%C4%B0yi\\_100\\_T%C3%BCrk\\_Filmi](https://tr.wikipedia.org/wiki/En_%C4%B0yi_100_T%C3%BCrk_Filmi)

## 1-Sınıflandırılacak sütunu seçiniz “Yapım Yılı”

Film	Yayıncı	Yapım Yılı
1. Subu Yaz	Metin Erkan	1963
2. Tuzlu Omer	Halit Sancar	1964
3. Gurmet Kuyumcu	Halit Sancar	1964
4. Sarmal Zaman	Metin Erkan	1965
5. Ah Güne İstanbul	Ali Yılmaz	1966
6. Mahcuplu	Süreyya Duru	1969
7. Umud	Yılmaz Güney	1970
8. Ah Mıgazan Ah	Mehmet Öner	1970
9. Babo	Yılmaz Güney	1971
10. Dönüş	Tunçel Şenel	1972
11. Çarın Kardeşi	Ertan Eğilmez	1973
12. Owin	Omer Lütfi Aras	1973
13. Hırsızın Sirkisi	Ertan Eğilmez	1974
14. Mavi Boncuk	Ertan Eğilmez	1974
15. Kıydan İndim Şehere	Ertan Eğilmez	1974
16. Akademi	Yılmaz Güney	1974
17. Ergin Aile	Ergin Özyer	1975
18. Tosun Paşa	Karke Tektel	1976
19. Süi Kardeşler	Ertan Eğilmez	1976
20. Kızılder Kız	Zeki Ökten	1976
21. Seni Bilmem Al Yazmam	Ali Yılmaz	1977
22. Çocuklar Kız	Zeki Ökten	1977

## 2. “Veri” yi ve ardından “Veri Doğrulama” yı tıklayın.

The screenshot shows the 'Data' menu open in Google Sheets. The 'Data Validation' option is highlighted with a red box and a yellow arrow. Other options in the menu include 'Sort range', 'Create a filter', 'Filter views', 'Pivot table', 'Named ranges', 'Protected sheets and ranges', 'Split text to columns', 'Remove duplicates', 'Trim whitespace', 'Group', and 'Ungroup'.

3. Açılan veri doğrulama penceresinde, “Kriter/ölçüt” in yanındaki aşağı açılır menüyü tıklayınız. Seçilen hürelere izin vermek için bir giriş türü belirlememiz gerekiyor. Seçtiğimiz satır için örnek; “C1:C998” bir filmin yayınlandığı yıl için dört basamaklı bir sayının işlendiğinden emin olmak için “Sayı/Number” seçeneğini seçiniz. “Yalnızca metin, tarihler, önceden tanımlanmış bir seçenek listesi, belirtilen aralıktaki öğeler veya özel doğrulama formülü” gibi diğer ölçütleri de seçebilirsiniz.

### Data validation

Cell range: veri doğrulama'E

Criteria: **List from a range** Sheet1!A2:D

On invalid data: **Number** Text input

Appearance: Text, Date, Custom formula is, Checkbox

Buttons: Remove validation, Save

**4-Ölçütler belirlenecek değerlere bağlıdır** yani her bir hücreye yazılan verileri doğrulamak için bu özelliği doğru olarak belirtmek gerekiyor. Bir filmin yalnızca yayımlandığı yılı belirlediğimizden, "Arasında" doğrulamayı kullandık. Bu örnek kapsamında en az değer **1963 ile en yüksek değer 2030** aralığını belirledik. Veri ekleme süreci devam edeceği için 2030 eklendi.

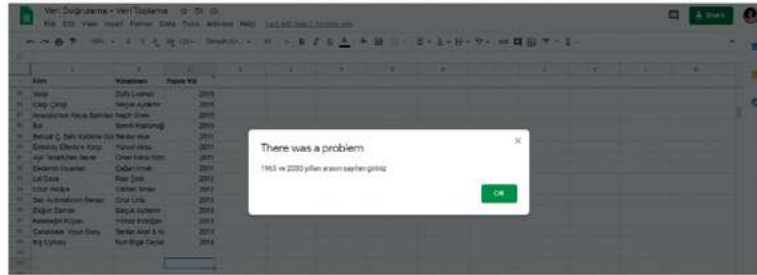
.Ardından, sayfaların bir uyarı mesajı gösterip göstermemesini eya yazılan herhangi bir şeyi imamen reddetmesini ve veriler geçersiz olduğunda bir hata mesajı gösterip göstermemesini seçin.

lyarı mesajı, numaranın belirli bir rakık arasında olması gerektiğini ildirir ve geçersiz verileri hücrede ırmızı bir bildirim ile tutar.

**Girişi Reddet** i tıklayın, Doğrulama Yardım Metnini iöster" onay kutusunu tıklayın ve rdından yardımcı bir uyarı mesajı azın. Doğrulama aracını apatmak için "Kaydet"

	A	B	C	D	E	F
1	Film	Yönetmen	Yapım Yılı			
2	Suzut Yaz	Metin Erişan	1963	Invalid:		
3	Tunisi Ömer	Hüli Saner	1964	Input must be a number between		
4	Gurbet Kuşlan	Halit Refiğ	1964	1963 and 2030		
5	Sıvmet Zamanı	Metin Erişan	1965			
6	An Güzel İstanbul	Atif Yılmaz	1966			
7	Malkeçoğlu	Sureyya Duru	1969			
8	Umut	Yılmaz Güney	1970			
9	An Müğah Ah	Mehmet Dişer	1970			
10	Baba	Yılmaz Güney	1971			
11	Donuğ	Tunçin Şoray	1972			

**6.Şimdi, birisi geçersiz veri girmeye çalıştığında, verideki hatayı düzeltebilmesi için yardımcı bir mesaj verildi.**



*Sonuç olarak, e-tabloya girilen verilerin doğruluğundan emin olmak gerekiyor, bilgiyi toplarken yapılabilecek hataları öngörmek gerekiyor.*



## 5.BÖLÜM

### Veri Temizlemek ve Düzenlemek

Veriler her zaman düzenli olmayabilir. Bu sebeple verilerde oluşabilecek hataları, dağınıklıkları anlamak ve düzeltmek gerekmektedir. Veriyi temizlemek demek, verinin neden dağınık olduğunu tespit edebilmektir. Verilerin çoğu otomatik olarak toplansa dahi elle girilen verilerde ya da verilerin aktarılmasında sorunlar oluşabilmektedir. Veriler dağınık olmasa bile verileri düzenlemek gerekebilir. Sayıca fazla veri seti ile çalışıldığında ise veri setlerinin birleştirilerek yapılandırılması önemli olabilir. Örneğin; temizlemek istediğiniz veri setinde eksik değerler, yazım hataları, gizemli değerler, kısaltmalar, gömülü değerler, kelime aktarımı, çift kayıtlar, çelişen kayıtlar gibi birçok sorun tespit edilirse, verinin doğruluğu farklı kaynaklarla karşılaştırılıp analiz edilmelidir. Veri görselleştirmesi de bu sürecin olumlu sonuçlandırılmasıyla başlatılmalıdır. Yeni başlayanlar için Quartz tarafından hazırlanmış ve Türkçeye de çevrilmiş olan Kötü Veri Rehberi <https://github.com/Quartz/bad-data-guide> incelenebilir.

README.md

### The Quartz guide to bad data

An exhaustive reference to problems seen in real-world data along with suggestions on how to resolve them.

As a reporter your world is full of data. And those data are full of problems. This guide presents thorough descriptions and suggested solutions to many of the kinds of problems that you will encounter when working with data.

Most of these problems can be solved. Some of them can't be solved and that means you should not use the data. Others can't be solved, but with precautions you can continue using the data. In order to allow for these ambiguities, this guide is organized by who is best equipped to solve the problem: you, your source, an expert, etc. In the description of each problem you may also find suggestions for what to do if that person can't help you.

You cannot possibly review every dataset you encounter for all of these problems. If you try to do that you will never get anything published. However, by familiarizing yourself with the kinds of issues you are likely to encounter you will have a better chance of identifying an issue before it causes you to make a mistake.

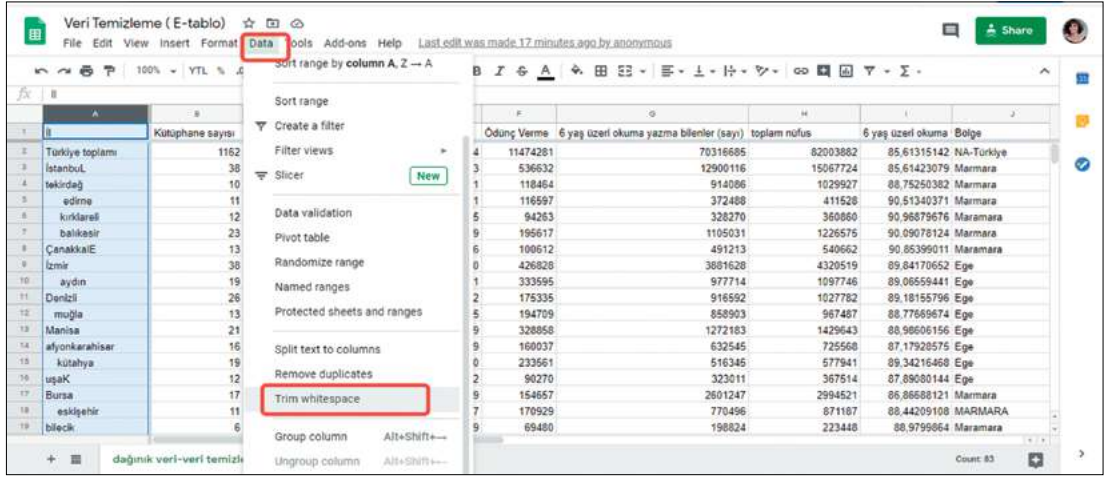
If you have questions about this guide please email Chris. Good luck!

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). Send your pull requests!

## Elektronik Tablo ile Veri Temizleme

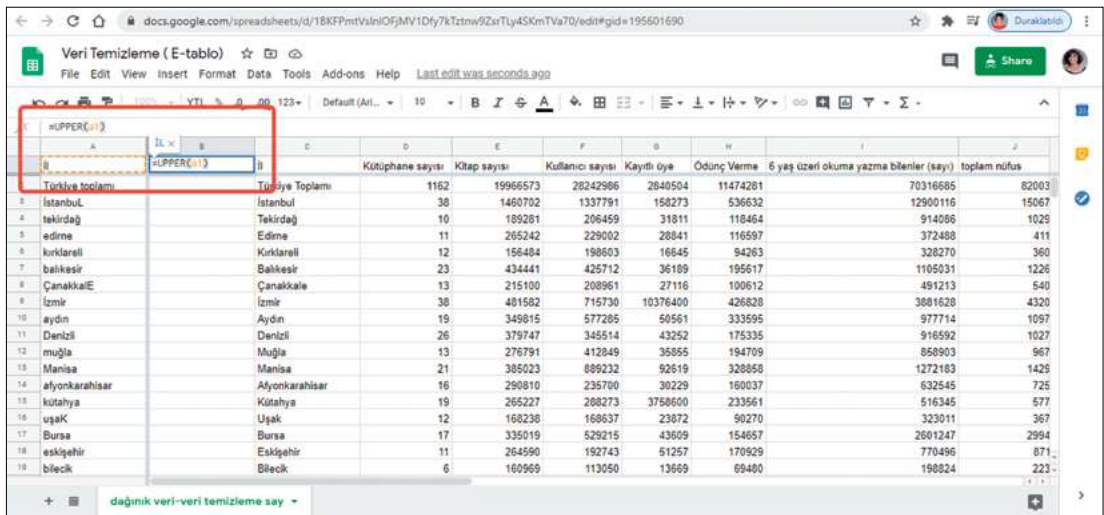
Google'ın [doküman düzenleyici yardım bölümünde](#) veriyi temizlerken veya düzenlerken kullanabileceğiniz çok zengin alternatifler mevcuttur. Biz mevcut veri setinde yer alan boşlukları kaldıracamız, üst ve alt, yazım düzeni, bul ve değiştir, metni sütuna böl fonksiyonlarını göreceğiz.

### Boşlukları Kaldırma



Şekil 51: Şehirler sütununda yer alan boşluklar veri, boşlukları kaldırma ya da kırıp ekran görüntüsüdür.

### Üst Fonksiyon



Şekil 52: İl isimlerinin tamamını büyük yapmayı sağlayan üst fonksiyon ekran görüntüsüdür.

## Alt Fonksiyon

	A	B	C	D	E	F	G	H	I	J
	=LOKER(A1)	İstanbul		Kütüphane sayısı	Kıtap sayısı	Kullanıcı sayısı	Kayıtlı üye	Ödünç Verme	6 yaş üzeri okuma yazma bilenler (sayı)	toplam nüfus
1	Türkiye toplamı	Türkiye Toplamı		1162	19966573	28242966	2840504	11474281		70316685
2	İstanbul	İstanbul		38	1460702	1337791	158273	536632		12900116
3	tekirdağ	Tekirdağ		10	189281	206459	31811	118464		914086
4	edime	Edime		11	265242	225002	28841	116597		372488
5	kirkirelli	Kirkirelli		12	156484	198603	16645	94263		328270
6	balikesir	Balikesir		23	434441	425712	36189	195617		1105031
7	ÇanakkaleE	Çanakkale		13	215100	208961	27116	100612		491213
8	İzmir	İzmir		38	481582	715730	10376400	426828		3881628
9	aydin	Aydın		19	349815	577285	50561	333595		977714
10	Denizli	Denizli		26	379747	345514	43252	175335		916592
11	muğla	Muğla		13	276791	412849	35855	194709		858903
12	Manisa	Manisa		21	385023	889232	92619	328858		1272183
13	afyonkarahisar	Afyonkarahisar		16	290810	235700	30229	160037		632545
14	kütahya	Kütahya		19	265227	288273	3758600	233561		516345
15	uşak	Uşak		12	168238	168637	23872	90270		323011
16	Bursa	Bursa		17	335019	529215	43609	154657		2601247
17	eskişehir	Eskişehir		11	264590	192743	51257	170929		770496
18	bilecik	Bilecik		6	160969	113050	13669	65480		198824

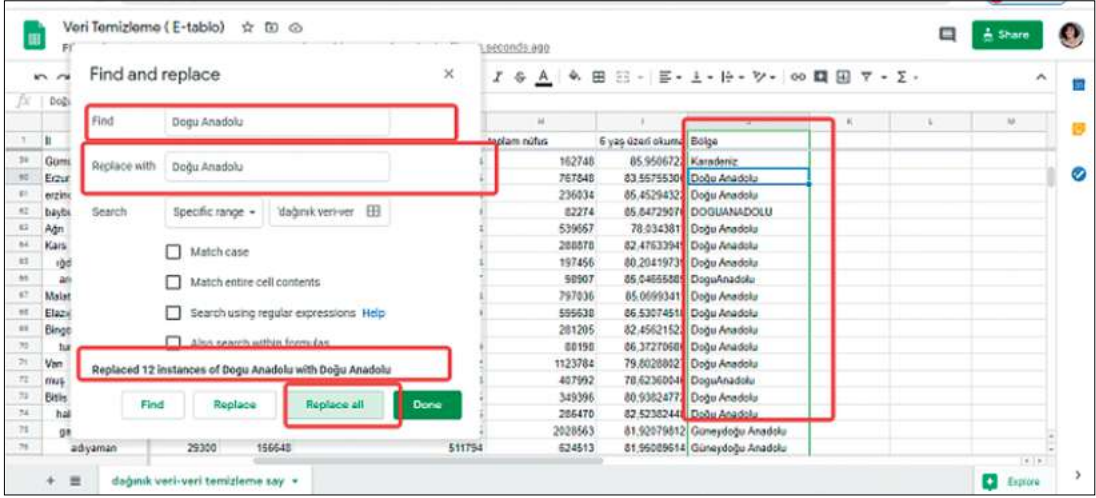
Şekil 53: İl isimlerinin tamamını küçük yapmayı sağlayan alt fonksiyon ekran görüntüsüdür.

## Yazım Düzeni

	A	B	C	D	E	F	G	H	I	J
	=proper(A1)	İstanbul		Kütüphane sayısı	Kıtap sayısı	Kullanıcı sayısı	Kayıtlı üye	Ödünç Verme	6 yaş üzeri okuma yazma bilenler (sayı)	toplam nüfus
1	Türkiye toplamı	Türkiye Toplamı		1162	19966573	28242966	2840504	11474281		70316685
2	İstanbul	İstanbul		38	1460702	1337791	158273	536632		12900116
3	tekirdağ	Tekirdağ		10	189281	206459	31811	118464		914086
4	edime	Edime		11	265242	225002	28841	116597		372488
5	kirkirelli	Kirkirelli		12	156484	198603	16645	94263		328270
6	balikesir	Balikesir		23	434441	425712	36189	195617		1105031
7	ÇanakkaleE	Çanakkale		13	215100	208961	27116	100612		491213
8	İzmir	İzmir		38	481582	715730	10376400	426828		3881628
9	aydin	Aydın		19	349815	577285	50561	333595		977714
10	Denizli	Denizli		26	379747	345514	43252	175335		916592
11	muğla	Muğla		13	276791	412849	35855	194709		858903
12	Manisa	Manisa		21	385023	889232	92619	328858		1272183
13	afyonkarahisar	Afyonkarahisar		16	290810	235700	30229	160037		632545
14	kütahya	Kütahya		19	265227	288273	3758600	233561		516345
15	uşak	Uşak		12	168238	168637	23872	90270		323011
16	Bursa	Bursa		17	335019	529215	43609	154657		2601247
17	eskişehir	Eskişehir		11	264590	192743	51257	170929		770496
18	bilecik	Bilecik		6	160969	113050	13669	65480		198824

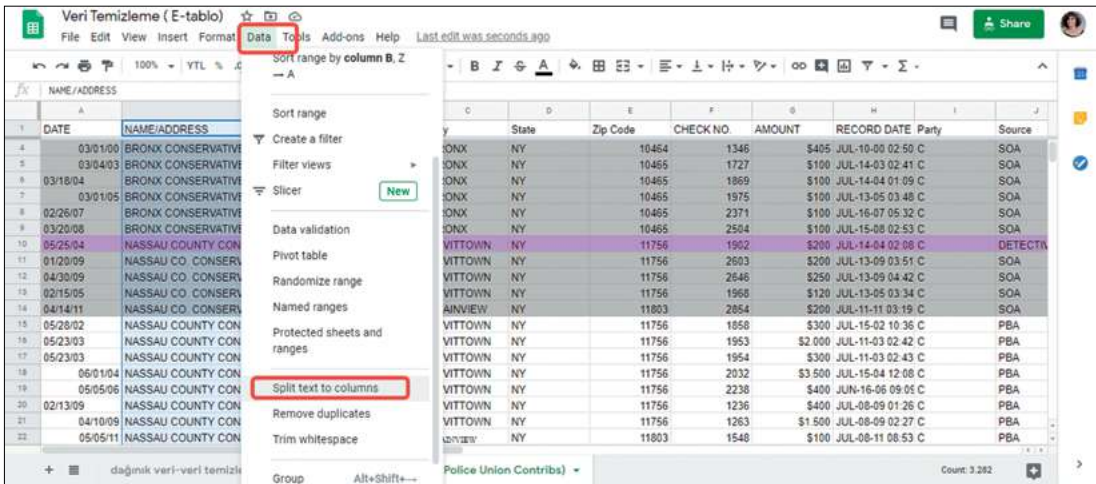
Şekil 54: İl isimlerinin tamamının ilk harfini büyük yapmayı sağlayan yazım düzeni fonksiyonu ekran görüntüsüdür.

## Bul ve Değiştir



**Şekil 55:** Hatalı, eksik, bitişik yazılmış olan bazı bölge isimlerinin “bul ve değiştir” fonksiyonu ile düzenlenme ekran görüntüsüdür.

## Metni Sütunlara Böl



**Şekil 56:** Bir sütunda yer alan isim ve adres verisini iki ayrı sütuna bölmeyi sağlayan fonksiyonun ekran görüntüsüdür.



## Open Refine ile Dağınık Veri Temizleme

Hadley Wickham, 2014 yılında tanıttığı “Tidy Data” kavramını 3 maddede açıklamıştır: Her sütun bir değişkeni temsil eder, her satır bir gözlemi temsil eder, her bir gözlem birimi bir tablo oluşturur. Open Refine aracı da diğer tüm veri temizleme ve birleştirme araçlarında olduğu gibi sütunları, satırları, yazım hatalarını ve daha pek çok hatayı yani veride tespit edilen dağınıkları düzenlemeyi sağlar. Veriyi istediğiniz yönde manipüle etmenizi sağlayan, kullanımı kolay ve açık kaynaklı bir araçtır.

Aşağıdaki linki kullanarak bilgisayarınızda Java JRE yüklü olduğundan emin olarak bilgisayarınız için uygun formatı seçip indirebilirsiniz. Kurulum talimatlarını takip ederek kolaylıkla kurulumu tamamlayabilirsiniz.

Tüm sürümlerini de buradan inceleyip indirebilirsiniz.

<https://openrefine.org/download.html>

- **Windows kit**, This requires Java to be installed on your computer. Download, unzip, and double-click on *openrefine.exe* or *refine.bat* if the former does not work.
- **Windows kit with embedded Java**, includes OpenJDK Java, available under the GPLv2+CE license. Download, unzip, and double-click on *openrefine.exe* or *refine.bat* if the former does not work.
- **Mac kit**, Download, open, drag icon into the Applications folder and double click on it. You do not need to install Java separately.
- **Linux kit**, Download, extract, then type `./refine` to start. This requires Java to be installed on your computer.

**Şekil 57:** Dağınık veri temizleme aracı Open Refine’in işletim sistemlerine göre sürüm ekran görüntüsüdür.

Open Refine’ı kullanarak verinizde yer alan boşlukları kaldırabilir, sütunları bölebilir, verileri sıralayabilir, kümeleyebilir ve yeniden adlandırabilirsiniz. Ayrıca temizlediğiniz verilerinizi beş farklı tipte açık dosya olarak da indirebilirsiniz.

## Remove White Spaces / Boşlukları Kaldırmak

The screenshot shows a spreadsheet with columns for City, State, Zip Code, CHECK NO., AMOUNT (\$), RECORD DATE, and Party. The 'Edit column' menu is open, and the 'Common transforms' sub-menu is selected, with 'Trim leading and trailing whitespace' highlighted.

## Rename Columns / Sütunlara İsim Vermek

The screenshot shows a spreadsheet with columns for NAME OF INITIATIVE/ADDRESS 1 and NAME OF INITIATIVE. The 'Edit column' menu is open, and 'Rename this column' is selected.

## Split Columns / Sütunları Bölmek

The screenshot shows a spreadsheet with a column for NAME OF INITIATIVE/ADDRESS. The 'Edit column' menu is open, and 'Split into several columns...' is selected.

The dialog box is titled 'Split column NAME OF INITIATIVE/ADDRESS into several columns'. It has two main sections: 'How to Split Column' and 'After Splitting'. In 'How to Split Column', 'by separator' is selected. The separator is set to a comma (,). There are checkboxes for 'regular expression' and 'Split into [ ] columns at most (leave blank for no limit)'. In 'After Splitting', 'Guess cell type' and 'Remove this column' are checked. There are 'OK' and 'Cancel' buttons at the bottom.

## Cluster / Kümelemek

The screenshot shows a spreadsheet with columns for City, State, Zip Code, and CHECK NO. The 'Cluster & Edit column' dialog box is open for the 'City' column. The 'Method' is set to 'by collapse' and the 'Keying Function' is 'fingerprint'. The dialog shows a table of clusters with columns for Cluster Size, Row Count, Values in Cluster, Merge?, New Cell Value, and # Rows in Cluster. The table lists several clusters of city names like 'E ISLIP', 'ST JAMES', 'W. BABYLON', 'N. MERRICK', 'N. VALLEY STREAM', 'N VALLEY', and 'W. ISLIP'. There are 'Select All', 'Deselect All', 'Merge Selected & Re-Cluster', 'Merge Selected & Close', and 'Close' buttons at the bottom.



## Sorting / Sıralama

PK NO.	AMOUNT (\$)	RECORD ID
1562	Facet	-07-03T10
2628	Text filter	-07-13T16
1348	Edit cells	-07-10T14
1727	Edit column	-07-14T14
1869	Transpose	-07-14T13
1975	Sort...	-07-13T15
2371	View	-07-16T17
2504	Reconcile	-07-15T14
1902		-07-24T22
2603		-07-13T15

**Sort by AMOUNT (\$)**

Sort cell values as

text  case-sensitive

numbers

dates

booleans

Position blanks and errors

Valid values

Errors

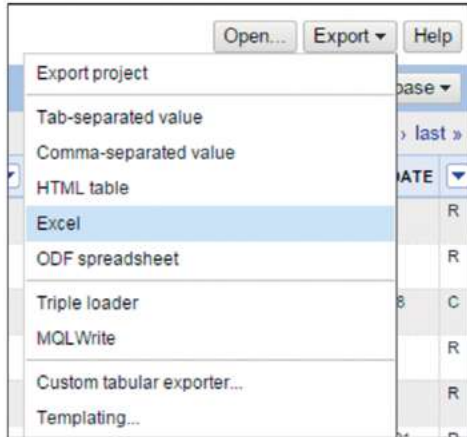
Blanks

Drag and drop to re-order

smallest first  largest first

OK Cancel

## Export / Dışa Aktarmak (csv, tsv..)



## Correct Formats / Formatı Düzenlemek

DATE	NAME OF INITIATIVE	ADDRESS 2	City
Facet	NSERVATIVE COMM	415 MINNIEFORD AVE	BRONX N.Y
Text filter	CONSERVATIVE PARTY	1 SYDNEY ST	PLAINVIEW NY
Edit cells	TRANSFORM...	105 BOBOLINK LANE	LEVITOWN N.Y
Edit column	COMMON TRANSFORMS	FEE 105 BOBOLINK LANE	LEVITOWN N.Y
Transpose			
Sort...			
View			
Reconcile			

- Transform...
- Common transforms
  - Trim leading and trailing whitespace
  - Collapse consecutive whitespace
  - Unescape HTML entities
- Blank down
- Split multi-valued cells...
- Join multi-valued cells...
- Cluster and edit...
- To titlecase
- To uppercase
- To lowercase
- To number
- To date
- To text
- Blank out cells

**Şekil 58:** OpenRefine ile yapılan veri temizleme işlemlerinin fonksiyon ekran görüntüsüdür.

Elbette yapabileceğiniz bunlarla da sınırlı değildir. 21'den fazla eklenti [Open Refine](#)'a kurarak daha gelişmiş veri temizleme işlemi yürütebilirsiniz. [OpenRefineAPI](#) ile 8 kütüphaneyle de çalışabilirsiniz.

- python
  - [fork with extended CLI](#)
- <https://github.com/maxogden/refine-python>
- ruby
- javascript - nodejs
- R
- php
- java
- bash

## WorkBench Data ile Veri Temizleme

WorkBench Data (<https://app.workbenchdata.com/lessons/en/clean-and-standardize>) ile web üzerinde, dağınık verinizi yükleyip temizleyebilir ya da kazıdığınız veriyi temizleyebilirsiniz.

Overview

In this tutorial you will work with **messy data** to compare population growth in major U.S. metro areas.

You will clean columns and rows, standardize values to make the dataset actionable, and reshape the table from long to wide format to gain insight and build a chart series.

Training II. Clean dirty data

Tab 1 Tab 2 Report

1 Load from URL

[https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population\\_growth\\_data.csv](https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population_growth_data.csv)

Has header row

Update Manual Checked 29m ago

Version 1 of 1 Update

ROWS	COLUMNS	No rows selected	EXPORT		
24	5				
Date text	A	MetroArea text	B	Population text	C
1	2012-01-01			8,523,132	
2	2011-01-01			7,122,446	
3	2012-01-01	Austin			
4	2012-01-01	Seattle			

hosted on Workbench.

Instructions

- ✓ Add the data source Load from URL.
- ✓ Paste in the following URL and click Update to load the data.

[https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population\\_growth\\_data.csv](https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population_growth_data.csv)

The two columns at the far right of the dataset are empty.

- ✓ Add the step 'Drop empty columns'
- ✓ Add the step 'Drop empty rows'.
- ✓ In Drop empty rows, select columns **MetroArea** and **Population**. The rows missing values in those columns will be deleted.

Previous 2 of 5 Next

Training II. Clean dirty data

Tab 1 Tab 2 Report

1 Load from URL

[https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population\\_growth\\_data.csv](https://storage.googleapis.com/production-static.workbenchdata.com/lessons/en/clean-and-standardize/population_growth_data.csv)

Has header row

Update Manual Checked 29m ago

Version 1 of 1 Update

2 Drop empty columns

Drop columns with no values

3 Drop empty rows

Drop row if any of these columns is empty

MetroArea Population

ROWS	COLUMNS	No rows selected	EXPORT		
24	5				
Date text	A	MetroArea text	B	Population text	C
1	2012-01-01			8,523,132	
2	2011-01-01			7,122,446	
3	2012-01-01	Austin			
4	2012-01-01	Seattle			
5	1990-01-01	San Jose-San Francisco...		6,794,848	
6	2000-01-01	San Jose-San Francisco...		7,656,428	
7	2010-01-01	San Jose-San Francisco...		8,153,696	
8	2016-01-01	San Jose-San Francisco...		8,751,807	
9	1990-01-01	Austin - Round Rock		846,227	
10	2000-01-01	Austin - Round Rock		1,249,963	
11	2010-01-01	Austin		1,716,320	

Convert types

Data types are used to tell computers how to interpret and use values. The columns **Date** and **Population** are text, so Workbench can't do math with them. They should be **date & time** and **number**, respectively.

Instructions

- ✓ In the table header, click on the menu of the column **Date** and select **Convert to date & time**.
- ✓ Likewise, use the menu of the column **Population** to select **Convert to number**.

Previous 3 of 5 Next

Training II. Clean dirty data

Tab 1 Tab 2 Report

4 Convert to date & time

Date

Input format: AUTO

Convert non-dates to null

5 Convert to number

Population

Integer or decimal: Any number

Thousand and decimal separators: United States: 1,500.00

Ignore non-numeric characters

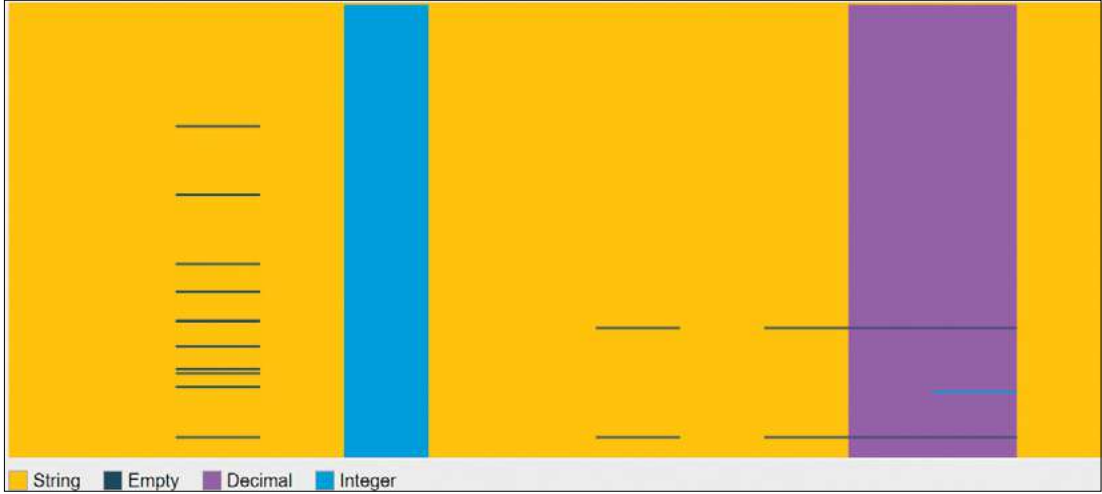
Convert errors to null

ROWS	COLUMNS	No rows selected	EXPORT		
20	3				
Date text	A	MetroArea text	B	Population text	C
1	1990-01-01	San Jose-San Francisco...		6,794,848	
2	2000-01-01	San Jose-San Francisco...		7,656,428	
3	2010-01-01	San Jose-San Francisco...		8,153,696	
4	2016-01-01	San Jose-San Francisco...		8,751,807	
5	1990-01-01	Austin - Round Rock		846,227	
6	2000-01-01	Austin - Round Rock		1,249,963	
7	2010-01-01	Austin		1,716,320	
8	2016-01-01	Austin - Round Rock		2,056,405	
9	1990-01-01	Dallas - Fort Worth		4,353,824	
10	2000-01-01	Dallas - Fort Worth		5,596,513	
11	2010-01-01	DallasFORTHWorth		6,851,525	



## CSV Fingerprint ile Verideki Hataları Tespit Etme

<https://setosa.io/blog/2014/08/03/csv-fingerprints/> Çalıştığınız veri setinizde çok fazla detaya girmeden, genel bir denetimle csv dosyasının görüntüsüne yer verir ve veri setinizdeki hataları bulmayı kolaylaştırır. Belirli bir biçime uyan bir CSV dosyası oluşturmaya çalıştığınızda hata yapmak kolaydır. Örneğin; eşleşmeyen ondalık sayılar, eşleşmeyen sayı türü (ondalık yerine tam sayı, vb.), boşluk, dizin gibi.



CSV veri setini kopyalayıp yapıştırdıktan sonra “generate” kısmına basılarak verinin kalitesi ve verideki hatalar incelenebilmektedir. PISA’da<sup>3</sup> dünya genelinde kız ve erkek bireylerin matematik performanslarını içeren CSV veri seti dosyası [kullanılmıştır](#).

### 1.aşama

example1.csv [example2.csv](#)'. Below this is a large grey text area with the placeholder text 'copy and paste your CSV here.'. At the bottom left, there is a dark blue button with the text 'generate'. At the bottom right, there is a small text: 'built with love by [setosa](#)'."/>

3 <https://data.oecd.org/pisa/mathematics-performance-pisa.htm>

## CSV fingerprint

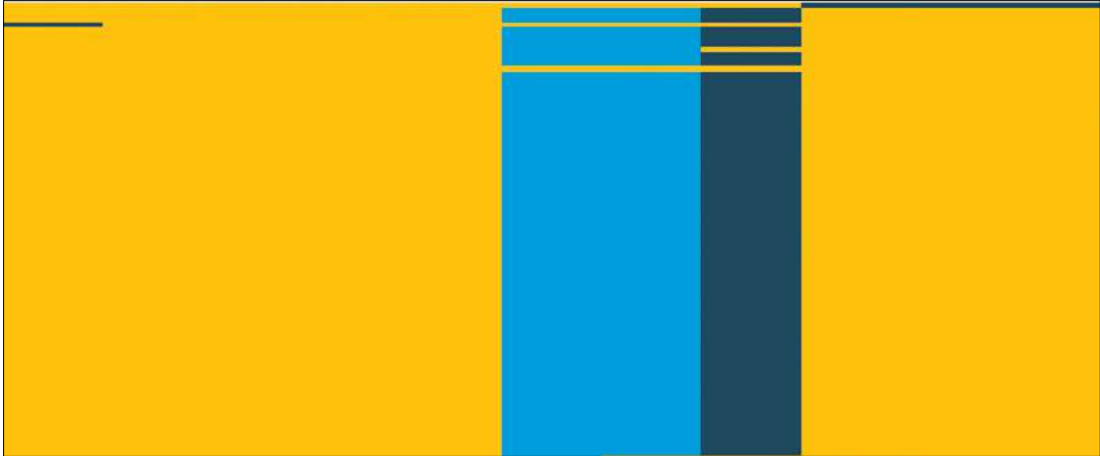
Here are a few examples: [example1.csv](#) [example2.csv](#)

```
LOCATION,"INDICATOR","SUBJECT","MEASURE","FREQUENCY","TIME","Value","Flag Codes"
AUS,"PISAMATH","BOY","MEANSORE","A","2015",497,
AUS,"PISAMATH","GIRL","MEANSORE","A","2015",491,
AUT,"PISAMATH","BOY","MEANSORE","A","2015",510,
AUT,"PISAMATH","GIRL","MEANSORE","A","2015",483,
BEL,"PISAMATH","BOY","MEANSORE","A","2015",514,
BEL,"PISAMATH","GIRL","MEANSORE","A","2015",500,
CAN,"PISAMATH","BOY","MEANSORE","A","2015",520,
CAN,"PISAMATH","GIRL","MEANSORE","A","2015",511,
CZE,"PISAMATH","BOY","MEANSORE","A","2015",496,
```

**generate**

built with love by [setosa](#)

Yüklenen ve incelenen veri setinde boşluk, dizin sorunu görülüyor.





## CSV Lint ile Verilerdeki Hataları Tespit Etme

<https://csvlint.io/> ile indirdiğiniz CSV dosyalarındaki hataları kontrol etmek için [CSV Lint](#) aracını kullanabilirsiniz. Dosyanızı yükledikten sonra araç, sizin için csv dosyasındaki hataları veya eksik verileri tespit eder. Link yapıştırmak veya bilgisayarınızdaki dosyayı yükleyip “Validate” butonuna basmanız yeterlidir.

### Kullanım Aşamaları:

Enter URL kısmının sağ tarafında yer alan + işaretine basarak birden fazla CSV URL girerek kontrol edebilirsiniz.



URL CSV konulduktan sonra “Validate” işaretleyip dosyanızın kontrol edilmesini beklemelisiniz.

Enter a link to your CSV:

+

Submitted urls are recorded in a public [list of validation reports](#). If you want to validate private data then upload a file from your computer, using the Browse button below.

Or upload a file:

browse

Add optional schema (in .json format)

✓ Validate

Aşağıdaki görselde, URL yüklendikten sonra aracın CSV dosyası okunabilirliğinin nasıl olduğuna dair sonuçlar görülmektedir. Görsel; bu CSV dosyasının kontrolü sonrası, dosyanın okunabilirliği ile ilgili hatalar olduğunu göstermektedir.

### Validation Results

CSV invalid

<https://docs.google.com/spreadsheets/d/1pXcqV-2Trwg-TUzZsED1mopXokJeXgDvbYmj9q-IAE8/edit#gid=0>

**Sorry, your CSV did not pass validation.**  
Please review the errors and warnings below:

Download Standardised CSV

Validate again

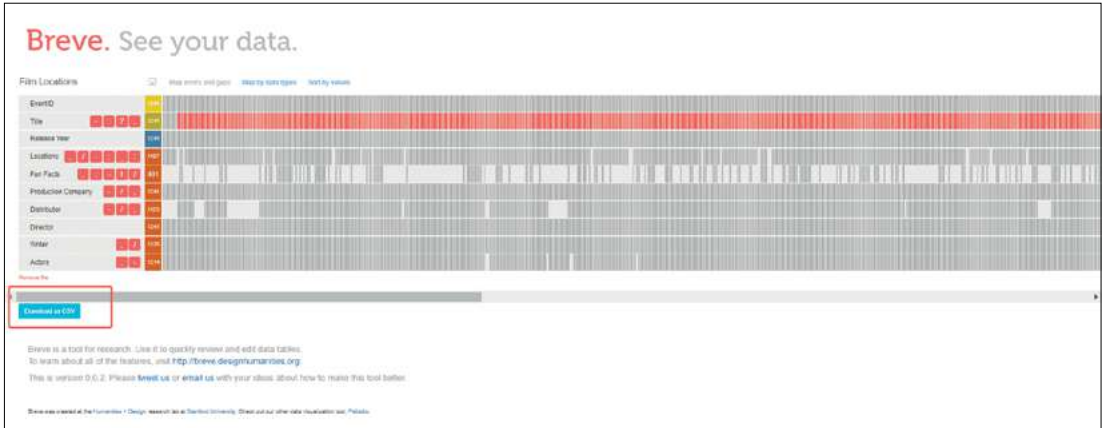
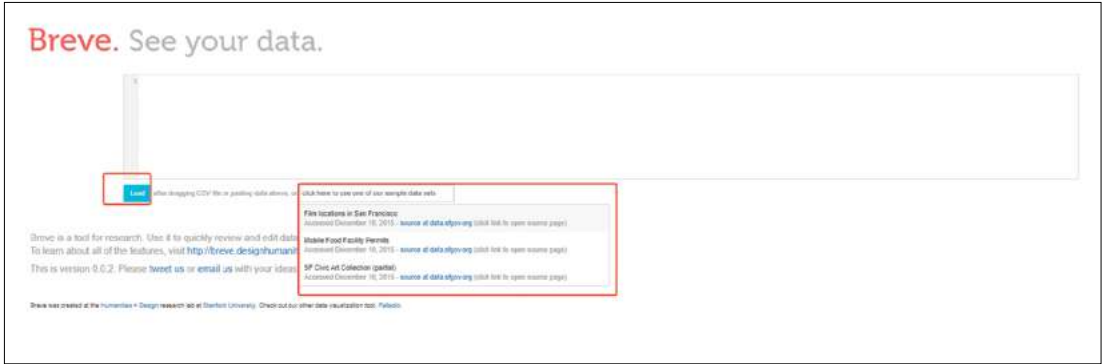
	273 Errors	4 Warnings	2 Messages
Structure	272	2	2
Schema	0	2	0
Context	1	0	0

## Breve ile Verideki Hataları Tespit Etme

Breve <http://hdlab.stanford.edu/breve/>; eksik ve dağınık verilerle çalışmak zorunda olanlar için tasarlanmış olan bir veri aracıdır. Veriler; görülmesi zor olabilecek tutarsızlıklar ve hatalarla doludur. Breve, tablo şeklindeki verilerin meta görünümünü verir. Ayrıca kayıtları ve sütunları ayrıntılı olarak incelemenizi ve değerleri düzenlemenizi sağlar.

Bilgisayarınıza indirerek ya da web üzerinden kullanmak mümkündür.

<http://hdlab.stanford.edu/breve/platforms/browser/www/>



*Breve eksik ve dağınık verilerle çalışmak zorunda olanlar için tasarlanmış olan bir veri aracıdır.*

## Good Tables ile Verideki Hataları Tespit Etme

Good Tables <https://goodtables.io/>: Verilerin iyi durumda olup olmadığını öğrenmenize yardımcı olur. Açık kaynak olarak sunulan araç ücretsizdir. Araç; boş satırlar ve sütunlar gibi yapısal sorunların olup olmadığını kontrol ederken; ayrıca verilerin belirlenen şemaya uyduğundan emin olmayı sağlamaktadır. Araç, özellikle açık veriyle çalışan gazeteciler için oldukça yararlıdır.

Çoğu zaman veri kümeleri CSV ve Excel formatında oluşturulmaktadır. Dolayısıyla bu durum kötü veri ihtimalini de arttırmaktadır. Bu aracın kullanımı oldukça kolaydır. Doğrudan dosyanızı ya da web tabanlı veri setinizi Good Tables arayüzüne koyduğunuzda süreci başlatmış oluyor; aracın veri setini kontrol ederek, hataları bulup tamir etmesini sağlıyorsunuz.

**Publish data with confidence**  
Validate your data on every change, so you catch errors before your users

Validate your data for free  
Sign in with GitHub

github/okfn/test-data/master/some-file.csv Table 1 of 3

1x **Missing Header**

1	id	name	name
2			

1x **Duplicate Header**

1	id	name	name
2			

Pushed on master 7 minutes ago 292801  
Pushed on master 2 days ago 62087  
Pushed on master 3 days ago 756648  
Pushed on master 5 days ago 278848

**Invalid** Share these results

Results as JSON: [Link](#)  
Data File Preview: [Show/Hide / https://raw.githubusercontent.com/okfn/goodtables-web/master/examples/invalid-on-structure.csv](https://raw.githubusercontent.com/okfn/goodtables-web/master/examples/invalid-on-structure.csv)

A total of 4 rows of data were processed.

Below we have a sample of 4 errors found when validating the data. Use this to identify problems, fix them in the source, and then upload the data to validate again!

1 **ERROR** [Structure](#) Missing Header : Header column number 1 is empty. Read more

2 **ERROR** [Structure](#) Defective Row : The row dimensions are incorrect compared to headers. Read more

**HEADERS**

Department	Entity	Date of Payment	Expense Type	Expense Area	Supplier	Transaction Number	Amount	Description
------------	--------	-----------------	--------------	--------------	----------	--------------------	--------	-------------

## 6.BÖLÜM

### Veri Analizi

Sivil toplum örgütlerinin (STÖ), çalışma alanları kapsamında; ham veri paylaşımları ne kadar önemliyse, verileri analiz edip yorumlamaları da o denli önemlidir.

Veri analizi; yararlı bilgileri keşfetmek, sonuçları bildirmek ve karar vermeyi desteklemek amacıyla; verileri inceleme, temizleme, dönüştürme ve modelleme sürecidir. Veri analizinin çeşitli isimler altında birden fazla yaklaşımı vardır ama günümüzde veri analizi kararları daha bilimsel hale getirmede, soruları daha detaylı görmede etkili olmaktadır.

### Veriyi Analiz Etmeye Nereden Başlanmalı?

Veri setleri hikâyeleri gizler, onları bulmalıyız. Veri setine herhangi bir kaynağa davranır gibi davranmalısınız. Sormak istediğiniz soruların listesini yapın. Size yardımcı olması için veri analizi araçlarını kullanın (LibreOffice vb.). Kusursuz olun, hata yapmayın. Adımlarınızı belgeleyin, tekrar kontrol edin. Çalışma arkadaşınızın analiz sonucunuzu incelemesini isteyin.

### Temel Ölçümler (İstatistik)



**Merkezi Eğilim Ölçüleri ve Önemi:** Veri üzerinde sadece frekans dağılımlarına ve grafiklere bakmayız, genel durumu yansıtacak ölçülere de gereksinim duyarız. Yani verilerin toplandığı ham haline değil, üzerinde yapılacak karşılaştırmalara, genellemelere ve yorumlara olanak sağlaması için bu ölçülere ihtiyaç duyarız.

### Veri Analizi Ölçüleri

**Merkezi Eğilim:** Bir değişkenin bütün farklı değerlerinin çevresinde toplandığı merkezi değerleri gösterir. Sıklıkla aritmetik ortalama, mod, medyan kullanılır.

Aritmetik Ortalama (Ortalama/Mean)

Tepe Değer (Mod/Mode)

Ortanca (Medyan)

**Merkezi Dağılım:** Verilerin ortalamasının etrafında mı uzağında mı dağıldığı, minimum, maksimum, standart sapma, ranj, çeyrek kayma ve çeyrek arası açıklık ile incelenir.

Ranj (Genişlik-Aralık)

Çeyrek Kayma (Sapma)

Çeyrekler Arası Açıklık

Standart Sapma

Varyans

**Aritmetik Ortalama (Ortalama), Tepe Değer (Mod), Ortanca (Medyan):**

- 1-Verinin değişkenliğini göstermemize yardımcı olur.
- 2-Verileri gruplamak için kullanabileceğiniz farklı yolları gösterir.
- 3-Farklı değişkenlerin bir spesifik veri setinin veya bir grubun içerisinde nasıl farklılık gösterdiğini bulmayı sağlar.
- 4-Çıkarımlar yaparken dikkate almamız gereken yanılma payını belirler.

## Aritmetik Ortalama (Ortalama/Mean)

Aritmetik ortalama veri setindeki tüm veri noktalarının toplamının toplam veri noktasına bölümü ile edilen bir "ortalama" sayıdır. Yani bir serideki verilerin toplamının veri adedine bölünmesi ile bulunur. Bizim veri setimizde toplam 10 adet kayıt bulunmaktadır.

< [45,65,68,45,21,73,45,35,19,20]>

Ortalama  $= (45+65+68+45+21+73+45+35+19+20) = 436 / 10 = 43,6$ 'dır.

$$\bar{x} = \frac{\sum x}{N}$$

$\sum x$  = the sum of  $x$

$N$  = number of data

Mean (Ortalama) değer formülü

<https://docs.google.com/spreadsheets/d/1wvD5wbHQqMGN9lhM3g-HbUmF4qZ3r6Ap4LMZToly1vY/edit#gid=0>

## Aritmetik Ortalama(Ortalama/Mean) Özellikleri

- En fazla kullanılan merkezi eğilim ölçüsüdür.
- Aritmetik ortalama serideki bütün değerlerden etkilenir.
- Uç değerlerin fazla olması aritmetik ortalamayı etkiler.

günlük-yaş
1615
1318
1082
1003
1121
1491
1082
1321
1492

Sum: 121507  
 Avg: 1992  
 Min: 28

## Teper Değer (Mod/Mode)

Tepe Değer bir veri setinde en sık gözlemlenen değeri göstermektedir. Veri setinin Tepe Değer'i **olamayacağı** gibi **birden fazla** Tepe Değeri de olabilir. Veri setimizde 10 farklı kişiye ait yaş değerlerinin olduğunu varsayalım.

< [45,65,68,45,21,73,45,35,19,20]>

Veri setimizde 45 değeri 3 kez tekrarlamıştır, diğer veriler ise sadece 1 kez kaydedilmiştir.

Tepe değer **45**'dir.

Tepe Değeri
45

## Ortanca Değer (Medyan/Median)

Ortanca Değer sıralı bir veri setindeki orta değeri göstermektedir. Ortanca değeri bulabilmek için mutlaka veri setimizi **küçükten büyüğe** sıralamamız gerekmektedir. Veri sayısının **tek olduğu durumlarda** Ortanca (Medyan) tam ortadaki değer iken, veri sayısının çift olduğu durumlarda Ortanca (Medyan) tam ortadaki 2 sayının aritmetik ortalamasıdır.

< [45,65,68,45,21,73,45,35,19,20]>

Öncelikle veriyi küçükten büyüğe doğru sıralıyoruz: 19,20,21,35, **45**,45,45,65,68,73

Ortanca değer: 45



Veri setimiz çok büyük olduğunda daha kolay şekilde ortanca değeri bulmak için aşağıdaki formülü kullanılmaktadır.

$$\frac{(n+1)}{2}$$

Median değerinin konum formülü

10 adet değer olan veri setinde ortanca değerinin konumu  $(10+1)/2= 5.5$  sayıdır. ( [5.6.](#) sıra)

## Ortanca Değer (Medyan/Median)

- Veride aykırı/uç değerler olduğu zaman aritmetik ortalamadan daha sağlıklı bilgi verir.
- Basit bir sıralama işlemi gerektirdiği için pratiktir.

## Merkezi Dağılım Ölçüsü

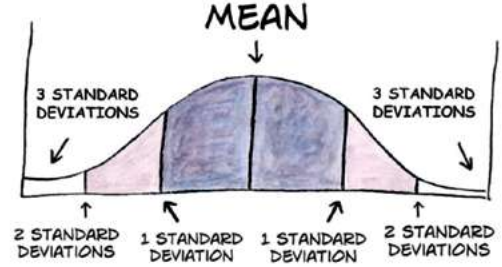
### Standart Sapma(Standard Deviation)

**Veriler tutarlı mı, değil mi anlamak için kullanılır.** Yani bir veri setinde bulunan gözlemlerin birbirine olan yakınlığını ve uyumunu ölçen yönteme diyoruz. Veri grubundaki sayıların aritmetik ortalamaya göre nasıl bir **yayılım** gösterdiğini anlamamızı sağlıyor.

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

## Standart Sapma(Standard Deviation)

- Eğer bir çok veri ortalamaya yakın ise Standart Sapma da düşük olacaktır.
- Eğer ortalamaya uzakta bir yaygın dağılım söz konusuysa standart sapma değeri de büyük olacaktır.
- Bütün verilerin tıpatıp aynı olması gibi bir durum söz konusuysa, standart sapma 0 olacaktır.
- Veri setinin özetlenmesinde çoğunlukla standart sapma, ortalama ile birlikte kullanılır. **Ortalama merkezi değerin ölçümü, standart sapma ise dağılımın ölçüsüdür.**



## Standart Sapma(Standard Deviation)

Veri setinin standart sapma değerini bulabilmek için aşağıdaki adımları sırasıyla uygulamak gerekmektedir.

- |   |     |
|---|-----|
| 1)Veri setinin Mean (Ortalama) değeri bulunur.                                  | Yaş |
| 2)Her bir veri noktası ile Mean (Ortalama) değer arasındaki uzaklık hesaplanır. | 23  |
| 3)Bulunan uzaklık değerlerinin kareleri alınır .                                | 23  |
| 4)Farkların kareleri toplanır.  | 34  |
| 5)Elde edilen toplam, serinin eleman sayısının bir eksiğine bölünür.            | 35  |
| 6) Bulunan sayının karekökü alınır.   | 35  |
|   | 35  |
|   | 43  |
|   | 45  |
|   | 67  |

	x(i)	x(i) - Mean	(x(i) - Mean ) <sup>2</sup>
1	23	15.22	231.66
2	27	11.22	125.89
3	34	4.22	17.81
4	35	3.22	10.37
5	35	3.22	10.37
6	35	3.22	10.37
7	43	-4.78	22.85
8	45	-6.78	45.97
9	67	-28.78	828.29
	Total		1,303.56

Veri setinin Standart Sapma değeri :  $\sqrt{1,303.56/(9-1)}=12.727$

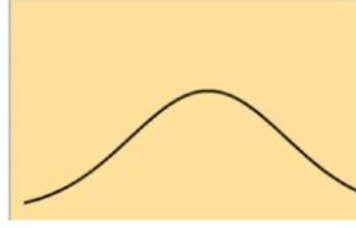
## Standart Sapma(Standard Deviation)

Hangi grafikte standart sapma büyüktür hangisinde küçüktür?

A

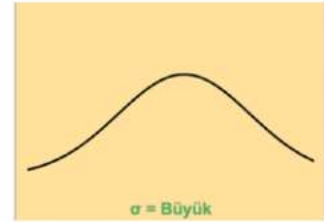
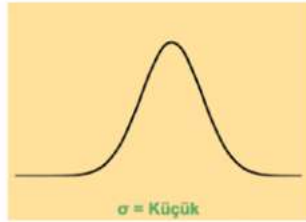


B



## Açıklama

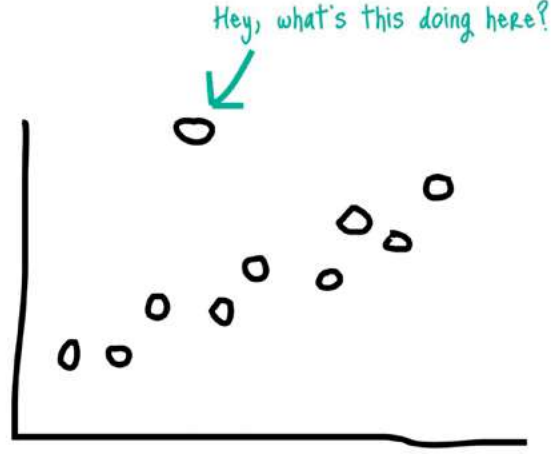
Standart sapma ( $\sigma$ ) küçüldükçe eğri dikleşmeye, standart sapma büyüldükçe eğri düzleşmeye başlamaktadır. Bunun temel sebebi ise **normal dağılım bir yoğunluk eğrisidir** ve toplam alanı 1'e eşittir. Standart sapma arttıkça yükseklik azalmakta genişlik artmakta ancak toplam alan değişmemektedir.



<https://veribulteni.voyd.org.tr/bir-satginin-simulasyonu-the-washington-postun-en-cok-okunan-makalesinin-arka-plan/>

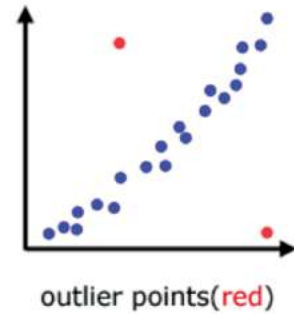
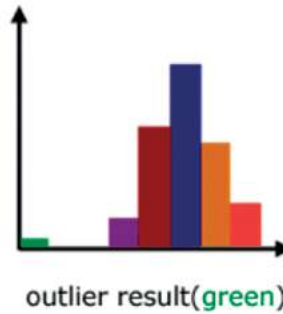
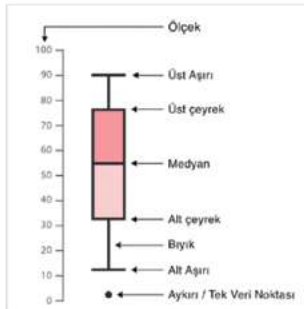
## Uç/Aykırı Değer (Outlier)

Bir dağılımda ortalamaya ya da diğer merkezi eğilim ölçülerine göre aşırı sapma gösteren en uç noktada yer alan değerlerdir. Uç değerler, diğer verilerle karşılaştırıldığında veri setine uygun olmadığı düşünülen aşırı değerlerdir. Bu aşırı değerler hatalı olabileceği gibi gerçeği de yansıtabilir. Bu nedenle hata ayıklaması yapılmalı ve doğruluğu kontrol edilmelidir.



## Uç/Aykırı Değer

Anatomi



[https://datavizcatalogue.com/TR/yonemleri/utu\\_biyiki\\_grafikleri.html](https://datavizcatalogue.com/TR/yonemleri/utu_biyiki_grafikleri.html)

## Veri Analizi ve Hesaplama Araçları

### Hesaplama Araçları

<https://standart-sapma-hesaplama.hesabet.com/>

<https://www.calculator.net/mean-median-mode-range-calculator.html>

71
64
98
88
64
80
66
96
87
62

**Calculator.net** FIN

home / math / mean, median, mode, range calculator

### Mean, Median, Mode, Range Calculator

**Result**

<b>Mean (Average)</b>	77.6
<b>Median</b>	75.5
<b>Range</b>	36
<b>Mode</b>	64, appeared 2 times
<b>Geometric Mean</b>	76.491700203732
<b>Largest</b>	98
<b>Smallest</b>	62
<b>Sum</b>	776
<b>Count</b>	10

**Sorted Data Set:** 62, 64, 64, 66, 71, 80, 87, 88, 96, 98

Mod, Medyan ve Standart Sapma Hesaplama

Diğer istatistikler: ortalama, aritmetik ortama, ortanca (ortanca değeri) ve aritmetik ortama (aritmetik ortama) hesaplayabilirsiniz.

Veriler (0-10):

0-10: 2 veriler

Yeni Veriler Ekle:

### E-tablo / Excel

tgs-uygulama ☆ 🔍

File Edit View Insert Format Data Tools Add-ons Help Last edit was 19 minutes ago

100% 123- Default (Arial) 10

	Kısa	Boy	
1			
2	71	171	
3	64	172	
4	98	190	
5	88	183	
6	64	158	
7	80	173	
8	66	188	
9	96	185	
10	87	177	
11	62	159	
12	Aritmetik Ortalama	77.6	175.0
13	Ortanca (Medyan)	75.5	175
14	Tape Değer (Mod)	64	mod yok
15	Ranji (Genişlik) - fark	13.9279517	34
16	Standart Sapma	=STDEV.S(B2:B13)	11.5161722
17	Varyans	194.2666667	132.6222222
18	Korelasyon	0.7096738351	

<https://docs.google.com/spreadsheets/d/1wvD5wbHQqMGN9lhM3g-HbUmF4oZ3r6Ap4LMZToly1vY/edit#gid=0>

Sheet1

<https://support.google.com/docs/answer/3093990?hl=en>

The image shows two screenshots of the Google Docs interface. The left screenshot shows the 'Insert' menu with 'Function' selected, and the 'All' option highlighted in red. The right screenshot shows the 'MODE' function selected, with its definition: 'Most commonly occurring value in a dataset.'

## R/Python

```

import numpy as np
import pandas as pd
from scipy import stats
np_array=np.array([23,27,34,35,35,43,49,67])
np_array
array([23, 27, 34, 35, 35, 43, 49, 67])

print('Mean', np.mean(np_array));
print('Median', np.median(np_array));
print('Std', np.std(np_array));
print('Mode', stats.mode(np_array))

```

Mean 38.22222222222222  
Median 35.0  
Std 12.03492859026756  
Mode ModeResult(mode=array([35]), count=array([3]))



## 7. BÖLÜM

Hızla dijitalleşen dünyada dijital araçların niteliği, etki alanı ve yapısı hızla değişip gelişmektedir. Günümüzde veriyle çalışma yürütürken kullanılacak aracın belirlenmesi, sürecin hızlı ve zahmetsiz ilerlemesini sağlamak adına verilecek en önemli kararlarından birisidir. Doğru aracı seçmek veya denemek için uygulayarak ilerlemek gerekir ancak her araca bu zamanı ayırmak sanıldığı kadar pratik bir çözüm değildir. Bu bölümde özellikle verinizi anladıktan sonra hangi araçları neden seçmeniz gerektiği yer almaktadır. Bunu yaparken ise en önemli aşamayla başlamak, yani kullanacağınız veriyi doğru şekilde aktarmak için ‘**veri çalışma panosu ve solo/ekip proje taslağı**’ örneklerini göreceğiz. Ardından, dijital araçları seçerken dikkate alınması gereken nedenler incelenecektir.

### Veri Çalışma Panosu Oluşturmak

Veri görselleştirmesi yapmadan önce içeriği nasıl geliştirmek gerektiğine karar verilmelidir. Bir araya getirilen veri setlerinin dikkat çekici şekilde aktarılabilmesi için bu ön çalışmayı yürütmek, konunun etkin şekilde görülmesine yardımcı olur. Ancak yaygın olarak karşılaşılan durum neyi anlatmanın önemli olacağı ya da önemsiz olacağını nasıl tespit edileceği sorunudur. Bu oldukça normal bir durumdur çünkü bir iz takip etmezseniz veri setleriniz size pek cömert davranmaz. Örneğin; “Hangi veri seti Covid-19 pandemisi sürecinde sağlık çalışanlarının çalışma koşullarını en etkili şekilde anlatır?” ya da “Özel hastanede çalışan sağlık çalışanlarıyla, devlet hastanesinde çalışan sağlık çalışanlarının Covid-19 ‘a yakalanma oranını gösterebilir mi?”



Şekil 59:Veri Çalışma Panosu: veri ile çalışırken çalışmanızı kolaylaştıran data storyboard taslak görüntüsüdür.

Farklı örnekler de eklenebilir ancak yukarıda gördüğümüz görsel taslak bir veri çalışma panosudur. Siz ve ekip arkadaşlarınız veri setleri içinde kaybolmadan önce böyle bir çalışma yaparak verileri net görebilirsiniz. Panoyla çalışırken ise aşağıda yer alan çok temel sorulara odaklanarak pratik yapmaya başlayabilirsiniz:

- Problem nedir?
- Problemlle ilgili hangi soruları sorabilirsiniz?
- Veri nerede ya da mevcut veride ne önemli?
- Sorunu etkin şekilde görünür kılmak için hangi görsel yöntemi kullanmalıyız ?



**Şekil 60:** Veri Çalışma Panosu ile yapılan hazırlık sonrası yine çevrim dışı olarak hazırlanmış bir veri görselleştirme panosu. / Fotoğraf: Pınar Dağ

Böylece ekip çalışmalarında verimi arttıran bu yöntemle fikirlerin kağıda taşınmasıyla hızlı bir araştırma yapılmış olur, **Veri Çalışma Panosu** oluşturmak için özel bir taslağa, özel bir araç ve gerece ihtiyaç duymadan sadece bir kağıt ve renkli kalemlerle deneyebilirsiniz.

## Solo/Ekip Proje Taslađı

Diđer basit ve oldukça pratik yöntem ise, üzerinde çalışılan verinin görselleştirme aşamasına giderken, doğru sorular sorup yanıt almayı kolaylaştıran solo ya da ekip proje taslađı ile çalışmaktır.

**HER ZAMAN ŞÜPHECİ OLUNUZ, VERİYE GÜVENMEYİNİZ**

**Başlık :**

**Konu:**

**Projenizin Amacı:**

**Kaynaklarınız ( en az 15)**

**Özetleyiniz (hedef, sonuç, çıktılar)**

**Bilgi Edinme Başvurusu**  
(En az 15 soru konu ile ilgili)

**Uygulanacak Method**

**Şekil 61:** Örnek solo/ekip proje dosyası. Taslađa buradan ulaşabilirsiniz:  
[https://docs.google.com/document/d/1-jK\\_9Jwojtd\\_Y9suCTMk89Ml\\_NzOq0QEzn0FSBm7O4/edit](https://docs.google.com/document/d/1-jK_9Jwojtd_Y9suCTMk89Ml_NzOq0QEzn0FSBm7O4/edit)

Taslak dosya, veri toplayarak yürüttüğünüz her türlü projenin niteliğini ve niceliğini geliştirmede yardımcı olacaktır. Çalışmaya başlamadan önce ise dosyada yer alan şablon sorular, veride önemli ve önemsiz olan detayları daha hızlı kavramanızı sağlar. Ayrıca verinizde eksik olanı da görmenizi sağlayacaktır. Kriter alınabilecek bazı sorular:

- Açıklama/Hedef
- Neden bu proje/haber/içerik önemli?
- Veri kaynakları güncel mi?
- Bilgi Edinme Hakkından yararlanabilir misiniz? Veriyi CİMER'den edinebilir misiniz?
- Proje kamu yararı taşıyor mu? Eğer öyle ise aktarınız/detaylı şekilde gösteriniz.
- İnsanlar neden sizin veri güdümlü projenizle/çalışmanızla ilgilensin?
- Hangi araçları kullanmayı düşünüyorsunuz? Neden?
- Veri seti temini kolayca mümkün mü? Veri seti yapılandırılmış mı? Temizlemek gerekiyor mu?

### Araç Seçilirken Dikkate Alınacak Bazı Faktörler

Dijital araçların sayısının çok fazla olduğundan bahsetmiştik ancak daha önemlisi; araçları neye göre seçmeliyiz, tecrübeler yeterli olabilir mi? Aşağıda daha fazlasını da ekleyebileceğimiz bazı faktörleri dikkate alarak araç seçimi yapabilirsiniz.

- 1. Öğrenmesi Kolay:** Aracı öğrenmek için ne kadar zaman harcamak gerekiyor?
- 2. Ücretsiz veya Uygun Fiyatlı:** Araç ücretsiz mi? Ücretli ise neler sunuyor ve üyelik hizmetleri aylık mı, yıllık mı? Uygun fiyatlı mı?
- 3. Güçlü Olması:** Araç, beklentilerinizi karşılıyor mu? İhtiyacınız olan tüm özellikleri sunuyor mu? Örneğin, veri projeniz için yeterli türde veri görselleştirme yöntemlerini sunuyor mu? Ya da sizin kullanmak istediğiniz araç [Ark Diyagramı](#) ya da öncellediğiniz bir grafiği çalışmaya olanak sunuyor mu?
- 4. Destek/Yardım Sağlama:** Aracı geliştirenlere hızlı ulaşabiliyor musunuz? Araç güncelleniyor mu? Araçla ilgili bir forum ortamı var mı? Örneğin, [workbench data](#) ile çalışırken hata yaptığımızda ekranın sağ tarafında bir yardım paneli açılarak sunucu tarafından sunulan çevrimiçi çözümlere süratle ulaşabiliyorsunuz.
- 5. Taşınabilir Olma:** Verilerinizi seçtiğiniz araca kolayca yükleyebiliyor musunuz? Ya da kolaylıkla indirebiliyor musunuz? Örneğin, [Flourish](#) aracı veri setini yüklemek ya da yapıştırmak için birçok alternatif sunmanın yanında üstelik verileri indirmek için kullanıcılara olanak sunuyor.

**6. Güvenli ve Özel Hesap Sunma:** Araç kişisel verilerinizi koruyor mu? Araç kullanıcıları ile ilgili ya da aracı geliştiren kurumla ilgili herhangi bir hack olayı yaşanmış mı? Yaşanmış ise şirket bu süreci nasıl yürütmüş?

**7. İş Birliği Sağlama:** Araç, ekip olarak ortak şekilde çalışmaya olanak sunuyor mu? Birlikte veri görselleştirme çalışması yapmak mümkün oluyor mu? Olanak sunuyor ise ayrıca yetki atama, erişim çeşitliliği ya da üzerinde çalışma hizmetleri de sağlıyor mu? Örneğin, e-tabloda veri görselleştirme yaparken çalışma arkadaşlarınızı yorum, üzerinde birlikte çalışma ya da sadece çalışmayı görme alternatiflerini kullanarak çalışmaya dâhil edebilirsiniz. Nitekim bu yolla hataları hızlıca engelleyebilirsiniz.

**8. Farklı /Çapraz Platformlarda Çalışma:** Araç farklı işletim sistemleri ile uyumlu mu? Örneğin, Windows, Mac, Chromebook veya Linux ile çalışır mı? Ayrıca araç, farklı ekran boyutlarına uygun veri görselleştirmeleri sağlıyor mu? Akıllı telefonlara, tabletlere ya da daha küçük cihazlara uyumlu mu? Haritalar, grafikler tatmin edici oluyor mu?

**9. Açık Kaynak:** Aracın yazılım kodu herkes tarafından ulaşılabilir mi? Farklı yazılımcılar geliştirebilir mi ya da değiştirip tekrar kullanabilir mi?

**10. Görme Engelli Okuyucular için Erişilebilir Olması:** Bugüne değin üzerinde en az durulan etken olan, renkleri algılamada zorluk çekenler / görme bozukluğu hastaları için de veri görselleştirmelerini yeniden düşünmek ve bu duruma uygun geliştirmeler yapmak gerekmektedir. Ayrıca; veri görselleştirme aracı, görme engelli okuyucular için erişilebilir görselleştirmeler oluşturmaya olanak sunuyor mu?

## Araç Öneri Tablosu

Kitabın bir önceki bölümünde, seçilecek araçların bazı faktörler dikkate alınarak seçilmesinin çok daha faydalı olabileceğinden bahsetmiştik. Ayrıca yürütülecek projeye göre de şekillendiğini anlamıştık. Kullanmanın kolay olması, destek hizmeti sunması, açık kaynak olması, güçlü olması, ortak çalışmaya uygun olması ve fazlasından söz edildi. Bu bölümde ise araçlar yer alacaktır. Hangi araçların veri görselleştirme projelerinde ne tür faydalar sağladığından söz edeceğiz. STÖ'ler, büyük-küçük fark etmeksizin birçok veriyle çalışabileceğinden, veri görselleştirme projelerinizde aşağıda yer alan araç setinden belki bir tanesi ile projenizi tamamlayabilirsiniz ancak farklı araçları tanımak aksayan bir projede belki ilk kurtarıcınız olacaktır.

Şüphesiz; temel ve ileri seviye fark etmeksizin tüm kullanıcılar için çoğu açık kaynaklı olmak üzere daha fazla aracı eklemek mümkündür. Aşağıdaki tablo, veri görselleştirme projenize başlarken hangisine ihtiyacınız olduğuna yanıt vermesi için hazırlanmıştır. Ayrıca, araç seçimi yaparken, Hollandalı veri gazetecisi Winny de Jong tarafından geliştirilen <https://datajournalism.tools/> sitesinden de yararlanabilirsiniz.

Araç	Veri Toplama	Veri Temizleme	Veri Analizi	Grafik	Geocode	Harita	Tablo	Kod	Dönüştürme
<a href="#">LibreOffice Excel</a>	X	X	X	X	X	X	X		
Google E-tablo	X	X	X	X	X	X	X		
<a href="#">Tabula</a> PDF'den veri çekme		X							
<a href="#">Open Refine</a> Dağınık Veri Temizleme		X							
<a href="#">DataMiner</a> Veri kazıyıcı temizleyici eklenti	X	X							
<a href="#">Workbench Data</a> Veri çekme, temizleme, analiz, grafik, tablo	X	X	X	X			X		
<a href="#">Datawrapper</a> Grafik/harita/tablo		X		X	X	X	X		
<a href="#">Flourish</a> Grafik, harita, tablo				X	X	X	X		
<a href="#">Tableau Public</a> Veri kazıma, temizleme, grafik, harita, tablo	X	X		X	X	X	X		
<a href="#">Highcharts</a> kod şablonları grafik				X					
<a href="#">Google My Maps</a> geocode/harita					X	X			
<a href="#">GitHub</a> düzenleme ve kod								X	
<a href="#">GitHub Desktop &amp; Atom</a> kod editörü								X	
<a href="#">Mapshaper</a> Coğrafi verileri düzenler ve birleştirir									X
<a href="#">R</a> *	X	X	X	X	X	X	X	X	X
<a href="#">Python</a>	X	X	X	X	X	X	X	X	X

*Araç Önerileri Tablosu: Veri ile çalışırken ihtiyaca göre kullanılacak araç önerileri. Bu araçların dışında da araçlar kullanılabilir. (Hazırlayan: Pınar Dağ)*

\* <https://www.r-project.org/>



## 8. BÖLÜM

### Veri Görselleştirme Sivil Toplum Örgütleri (STÖ) için Neden Önemlidir?

Veri görselleştirme, veriyi aktarmada önemli bir araçtır. Temel amacı; çok fazla verinin daha iyi anlatılmasını sağlamaktır. Sivil toplum örgütleri (STÖ) kendi verisini toplayan kurumlar olup, sıklıkla rapor hazırlar ve paylaşırlar. Veriyle çalışmaları, STÖ'lerin doğasında vardır ancak nasıl çalıştıkları ya da nasıl çalışmaları gerektiği konusu önemlidir. *“Veri ne söylüyor, söylediği şey hangi soruna dikkat çekiyor ve dikkat çektiği noktada hangi mekanizmayı harekete geçirebiliyor?”* zincirinin takip edilmesi, önemli bir veri görselleştirme çıktısıdır. Veri görselleştirmeyi iyi kılan şey de tam olarak budur; neden-sonuç ilişkisini iyi kalitede veri ile sunup bir şey söylemesini sağlamaktır. Analiz, tasarım ve iyi hikâyeye olarak, üç ayağı bulunan veri görselleştirme; sivil toplumla kurulan dili güçlendirir.

Bu kılavuzda; veriyi anlama, temizleme, görselleştirme gibi yöntemlerle daha etkili çalışabilmenin yolu aşama aşama yer almaktadır. Metinleri, zengin sözcüklerle daha ilgi çekici hale getirebildiğinizde, bu durum yazılan metnin okunurluğunu artırır ve metni kendi başına güçlü bir içeriğe dönüştürür. Kaliteli veri kullanılarak yapılan veri görselleştirmeleri de aynı amacı taşırlar. Yani nicel, nitel, kategorik verileri kullandığımızda; nicel, ilişkisel veya uzamsal kalıpları görüntülere dönüştürerek, veri hikâyeleri göstermiş olursunuz. Sayısız kaynak, veri görselleştirme için en ideal araçları ve kaynakları sunmaktadır; ancak bu kılavuz, özellikle interaktif araçlarla nasıl çalışabileceğinizi, bu görselleştirmeleri nasıl sitenize ekleyebileceğinizi ve doğru bilgiyle mücadelede veri görselleştirmenin nasıl etkili bir araç olduğunu püf noktalarıyla ele almaktadır.

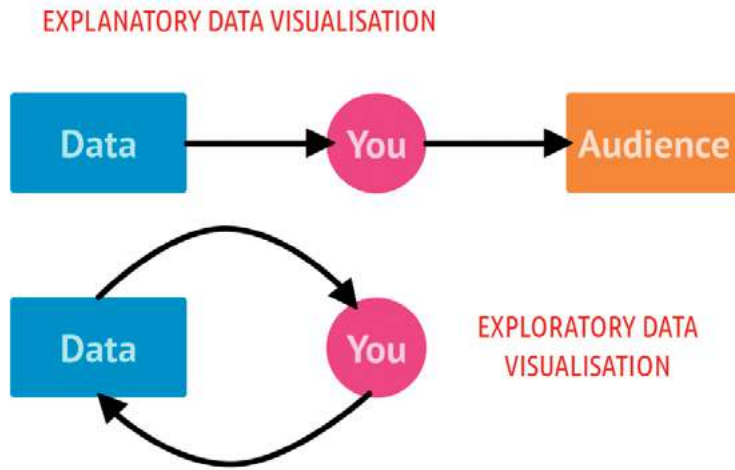
Son 20 yılda artan veri, görselleştirmenin de daha etkin bir şekilde kullanılmasını sağladı. Önceleri sadece basılı örneğiyle ilişki kurabildiğimiz tasarımlara artık web üzerinden hızlıca erişebiliyoruz. Aynı zamanda teknolojinin yardımıyla, kendi üretimimizi de yapabileceğimiz bir dönemi yaşamaktayız. Bu yöntemlerin geliştiğini görmek önemli olmakla beraber, artık herkes çok sık içerik paylaşabilmektedir. Bu durum, yanlış bilgiyle mücadele etmede veri görselleştirmenin kullanımında önemli bir rol üstlenmesine rağmen, artan içerik paylaşımı, üretilen içeriklere dair güven problemleri yaratmaktadır. Tam bu noktada nasıl akıllıca kararlar verebiliriz? İnsan hakları, derin yoksulluk, işsizlik, hayvan hakları, iklim krizi ve nicesi gibi gerçek sorunlar hakkında çelişkili görseller, veriler sunulduğunda hangisine inanmalıyız? Kılavuzun ilerleyen bölümlerinde, veri görselleştirmenin ve kullanılan verinin doğruluğu, örneklerle yer alacak. Veri görselleştirme dünyasının kirli taraflarını anlamadığımız sürece, etkileşimli araçların gölgesinde gerçeği unutmamız ve bu unutuşun bir parçası olmamız mümkün. Bu pratik bilgilerin ve ipuçlarının olduğu kılavuz tam da bu amaçla hazırlandı.

## STÖ'ler kendi verilerini nasıl toplar, anlar, yazar ve anlatır? Nereden başlamalıdır?

### Keşfetmek ve Açıklamak İçin Görselleştiririz

Görselleştirme yöntemleri farklılık göstermektedir ancak bu kılavuz kapsamında biz, veride olanı keşfetme ve veride olanı açıklamaya odaklanıyoruz. Verileri keşfetmek için görselleştirdiğimizde, özellikle **keşifsel veri görselleştirme**, farklı ilişkileri ortaya çıkarır, daha fazla vurgu veya analiz için en anlamlı bilgileri keşfetmenin yolunu aralar.

Diğer önemli bir neden ise; açıklama yapmak için verileri görselleştiririz. **Açıklayıcı veri görselleştirme** hedef kitemizi bilgilendirmek, ikna etmek, sorunu görünür kılmak ve merak etmelerini sağlamak için kullanılabilir. Kılavuz'un bu bölümü özellikle açıklayıcı veri görselleştirmeye odaklanmaktadır.



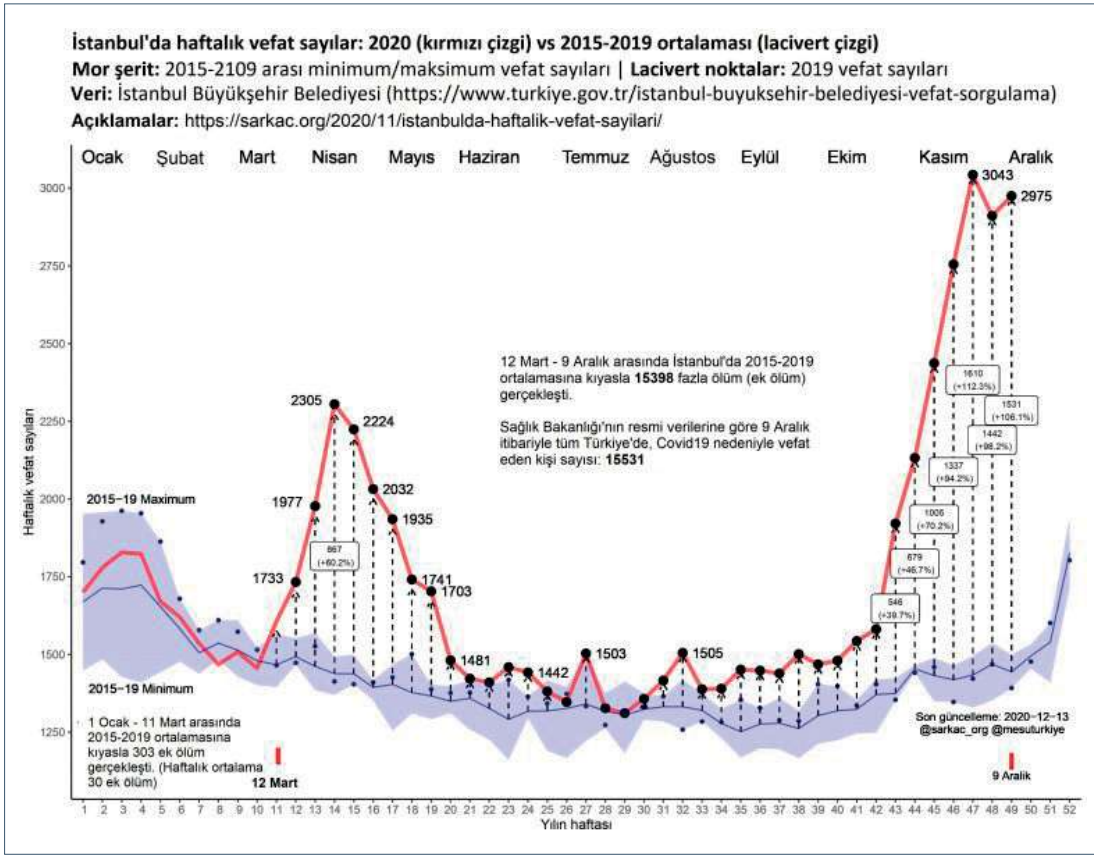
**Şekil 62:** Veri > Siz > Seyirci /Okuyucu/İzleyici | Veri >Siz -Siz >Veri (Açıklayıcı Veri Görselleştirme)  
Maarten Lambrechts'in sunumunun ekran görüntüsüdür.

Bizim üzerinde yoğunlaştığımız yöntem, açıklayıcı veri görselleştirmedir. Bunun için öncelikle, net bir hedef kitle ve net bir hedef belirlenmelidir. Bu çalışmaya başlarken de kendinize şu soruyu sorarak başlamalısınız:

**Ne demeye çalışıyoruz?:** Yaptığımız görsel seçim, aktarmaya çalıştığımız mesajı aktarmaya rehberlik etmelidir.

**Bunu kime söylemeye çalışıyoruz?:** Hedef kitlenizin kim olduğunu ve onların görselleştirmenizi nasıl kullanacaklarını düşünüp, asıl konunuza odaklanmanıza yardımcı olacaktır.

## Örnek:



Şekil 63: <https://sarkac.org/2020/12/istanbulda-haftalik-vefat-sayilari/>

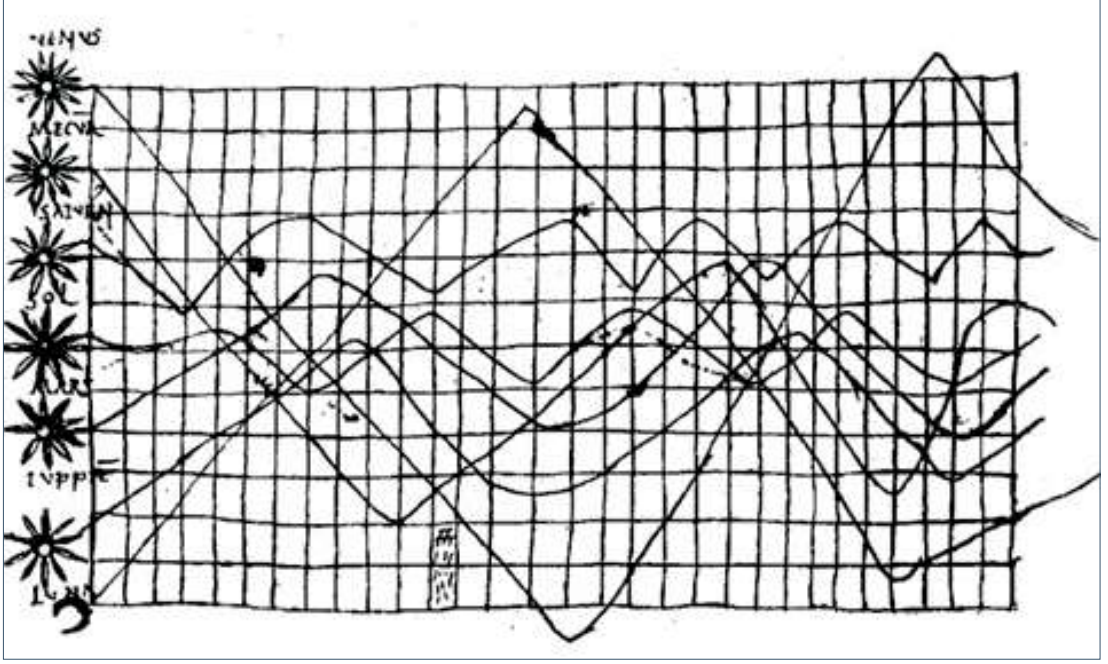
**Ne demeye çalışıyorlar?:** İstanbul'da ek ölümler incelendiğinde, haftalık ölümler daha fazla artmaktadır.

**Bunu kime söylemeye çalışıyorlar?:** Sarkaç okuyucularına.

### Veri Görselleştirmenin Tanımı

Ham verilerin, istatistik ve tasarım bilgisi kullanılarak görselleştirilmesi ve anlamlı hale getirilmesi olarak tanımlanan veri görselleştirme; verinin anlatılması ve anlaşılmasını sağlayan önemli bir araçtır. Bir neden-sonuç ilişkisi sunar, anlatmak isteneni yalın bir şekilde aktarır ve sorunu görünür kılar. STÖ'lerin üyelerine, çalışma alanları ile ilgili yürüttükleri faaliyetleri daha iyi anlatmalarını sağlar, hızlı ve etkin karar almaya yardımcı olur. Sahip olunan verileri daha iyi anlamak, veri setini farklı açıdan görebilmek, ilişkileri /bağlantıları daha etkili şekilde kurmak ve bulgulamak için veri görselleştirme önemlidir.

## 1. Örnek:



**Şekil 64:** (M.Ö 950-Avrupa) Bilinen ilk nicel gösterge - gezegenlerin ve ayların değişen konumunu gösteren bir çoklu zaman serisi grafiği.

## 2. Örnek:



**Şekil 65:** Harita (M.Ö. 6500) bilinen en eski harita-Çatalhöyük/Konya Müzesi-Türkiye

## Veri Görselleştirmenin Önemi

Veri görselleştirmenin; estetik, etkileyici ve dikkat çekici yönlerinin dışında, neden günümüz için önemli olduğunu, nasıl yapılması gerektiğini, geliştirmenin yöntemlerini, kurallarını, görsel etiğin ne olduğunu ve sınırlarını kapsamlı olarak bilmek önemlidir. Görselleştireceğiniz veriye; belki aşına, kaynaklarını biliyor ya da veriyi toplayan siz olabilirsiniz. Belki çoktan, ne hakkında olduğunu ve neyin analiz edileceğini de biliyorsunuz. Yani, elinizde mevcut olan veriye yabancı değilsiniz. Bu durumda zaman kaybetmeden biraz daha hızlı ilerleyebilirsiniz. Ancak tersi durumda ne yapılabilir? Ya da tersi durumda ne olur? Veri setlerine dair herhangi bir fikriniz yok, başka bir ekip/kaynak tarafından oluşturulmuş, güvenilirliği ve yapısına dair bilginiz de sınırlı ise; bu durumda veri görselleştirme çalışması yürütürken, ön işleme sürecine daha fazla odaklanmanız; yani veri kalitesinden emin olarak çalışmayı yürütmeniz gerekmektedir. Nihayetinde, her ne durumda olursanız olun kullanacağınız veri; yürüttüğünüz faaliyetlerinizde, karar alma süreçlerine etki etmektedir. Kılavuz'un ilk bölümünde yer alan veriyi anlama bölümü, veriyi görselleştirmeye başlamadan önce, bu yüzden önemli bir noktadır. Bu sebeple veriyi kullanmadan, ne anlattığını anlamak; veri görselleştirmeye neden ihtiyaç duyulduğunu da açıklayan sebeplerden biridir. Bilgiyi nasıl etkin şekilde kullanıp, karar alma sürecine etkisini sağlayabilirsiniz? Bu sorunun önemli cevaplarından biri; veri görselleştirme yöntemleridir. Bu yöntemler, karar almayı sağlar; yürütülen her çalışmanın daha etkin hazırlanmasına, aktarılmasına yardım eder. Ancak, her görselleştirme dikkatli olunmadığında yarardan çok ciddi bir yanılgıya neden olabilir. Bazen, en iyi kafa karıştıran kaynağa da dönüşebilir. Ama özenle üretilmiş görselleştirmeler, gerçek bir çözüm üretebilir. Bu sebeple, belirtmek gerekir ki eğer görselleştirme doğru yapılırsa, minik bir veri seti bile olsa, çok şey ifade edebilir.

20. yüzyılın başlarında Gestalt Psikologları, elementlerin bir şekle/sembole dönüştürüldüğünde; algısal olarak çıktısının, beklenenden daha etkin olduğunu gözlemlemiştir. Bilgisayar tabanlı **görselleştirme** çalışmalarının da 1984'den itibaren hız kazandığı belirtilir.

### Veri Görselleştirmenin Kökleri

- [Kartografya/Haritacılık](#)
- [İstatistik](#)
- [Veri](#)
- [Görsel Düşünme](#)
- [Teknoloji](#)

### Veri Görselleştirme Tipleri

- Doğrusal Görselleştirme
- Düzlemsel Görselleştirme
- Hacimsel Görselleştirme
- Çok-Boyutlu Görselleştirme
- Hiyerarşik Görselleştirme
- Zamansal Görselleştirme
- İletimsel Görselleştirme

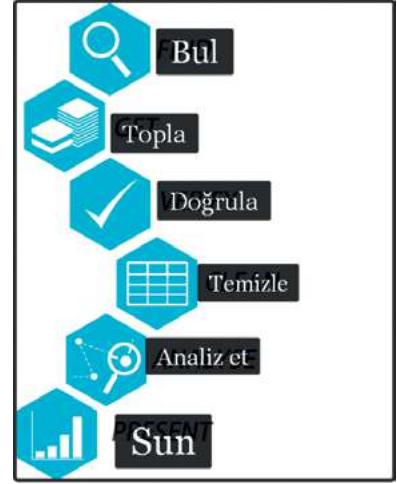


## Veri Görselleştirme Süreci

Ham veriyi bulma, toplama, doğrulama/ güncelliğinden emin olma, temizleme, analiz etme ve veri setinin özelliğine göre görselleştirme/sunma süreçlerini kapsar.

## Veri Görselleştirmenin Niteliğini Etkileyen Unsurlar

**Veri Kalitesi:** Görselleştirme için kullanılan verinin kalitesi, iyi bir görselleştirme çalışmasının öncelikli konusudur. Ham veri yeterince iyi kalitede değilse, görselleştirmek de mümkün olmayabilir. Dağınık veri bölümünde de yer aldığı gibi veriyi temizlemek ve kullanılabilir hale getirmek görselleştirmeyi doğrudan etkilemektedir.



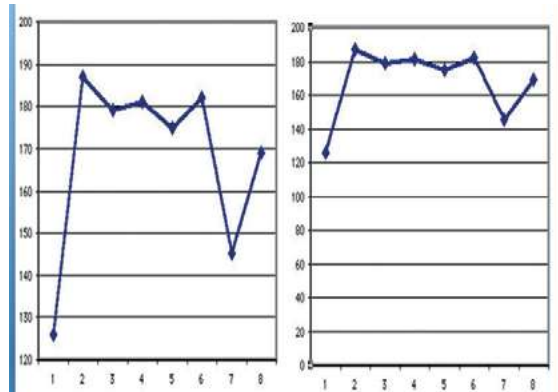
**Neden-Sonuç İlişkisi:** Veri görselleştirme sıklıkla ilişkileri göstermede kullanılır.

**Doğrulanabilir-Kanıtlanabilir Olma:** Veri görselleştirme; anlamlandırma, ham veriyi toplama, doğrulama, analiz etme, sunma sürecini kapsıyor mu? Kapsamıyor ise ne işe yaradıkları sorgulanmalıdır. Hangi açıdan ölçme / analiz etme / karar verme gibi detayları içinde barındırıyor? Yoksa sadece güzel görsellerden mi ibaret? Bunları bilmek önemlidir.

**Görsel Etik:** Görseller ve fotoğraflar; düşüncemizi, duygularımızı ve davranışlarımızı, iyi ve kötü yönde şekillenmesine etki etmektedir. Görsel etik, sadece dijital etik ya da görsel gazetecilik, aktivizm, reklamcılık, halkla ilişkiler ya da TV gazeteciliği değildir. Görsel etik; iletişimin sinir sistemidir, iletişimin "ruhudur."

**Görsel Etik Neden Önemlidir:** Görmek; inanmaktır, görselleştirilen şey kelimelerden daha güçlüdür, daha fazla alternatif sunar.

## Eksene Dikkat!





## Veri Görselleştirme Yöntemleri Periyodik Tablosu

A PERIODIC TABLE OF VISUALIZATION METHODS																																			
☼☼☼ <b>C</b> cartesian		☼☼☼ <b>Data Visualization</b> Visual representations of quantitative data in schematic form (either with or without axes)				☼☼☼ <b>Strategy Visualization</b> The systematic use of complementary visual representations in the analysis, development, formulation, communication, and implementation of strategies in organizations.				☼☼☼ <b>G</b> graphic facilitation																									
☼☼☼ <b>Tb</b> table		☼☼☼ <b>Ga</b> cartesian coordinates		☼☼☼ <b>Information Visualization</b> The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image; it is mapped to screen space. The image can be changed by users as they proceed working with it.				☼☼☼ <b>Metaphor Visualization</b> Visual Metaphors position information graphically to organize and structure information. They also carry an insight about the represented information through the key characteristics of the metaphor that is employed.				☼☼☼ <b>Me</b> meeting trail		☼☼☼ <b>Mm</b> metro map		☼☼☼ <b>Tm</b> templ		☼☼☼ <b>St</b> story template		☼☼☼ <b>Tr</b> tree		☼☼☼ <b>Ct</b> cartoon													
☼☼☼ <b>Pi</b> pie chart		☼☼☼ <b>L</b> line chart		☼☼☼ <b>Concept Visualization</b> Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.				☼☼☼ <b>Compound Visualization</b> The complementary use of different graphic representation formats in one single scheme or frame.				☼☼☼ <b>Go</b> communication diagram		☼☼☼ <b>Fp</b> flight plan		☼☼☼ <b>Cs</b> concept sections		☼☼☼ <b>Br</b> bridge		☼☼☼ <b>Fu</b> funnel		☼☼☼ <b>Ri</b> risk picture													
☼☼☼ <b>B</b> bar chart		☼☼☼ <b>Ae</b> area chart		☼☼☼ <b>R</b> radar chart cobweb		☼☼☼ <b>Pa</b> parallel coordinates		☼☼☼ <b>Hy</b> hypertoric tree		☼☼☼ <b>Cy</b> cycle diagram		☼☼☼ <b>T</b> timeline		☼☼☼ <b>Ve</b> veas diagram		☼☼☼ <b>Mi</b> mindmap		☼☼☼ <b>Sq</b> square 90 opposition		☼☼☼ <b>Cc</b> concentric circles		☼☼☼ <b>Ar</b> argument side		☼☼☼ <b>Sw</b> swiss cheese diagram		☼☼☼ <b>Gc</b> gantt chart		☼☼☼ <b>Pm</b> perspectives diagram		☼☼☼ <b>D</b> diagrama diagram		☼☼☼ <b>Pr</b> parameter ruler		☼☼☼ <b>Kn</b> knowledge map	
☼☼☼ <b>Hi</b> histogram		☼☼☼ <b>Sc</b> scatterplot		☼☼☼ <b>Sa</b> safety diagram		☼☼☼ <b>In</b> information lines		☼☼☼ <b>E</b> every relationship diagram		☼☼☼ <b>Pt</b> path net		☼☼☼ <b>Fl</b> flow chart		☼☼☼ <b>Cl</b> clustering		☼☼☼ <b>Lc</b> logic chart		☼☼☼ <b>Py</b> pyramid technique		☼☼☼ <b>Ce</b> cause-effect chains		☼☼☼ <b>Tl</b> timeline map		☼☼☼ <b>Dt</b> decision tree		☼☼☼ <b>Ep</b> epic critical path method		☼☼☼ <b>Ef</b> escape list		☼☼☼ <b>Co</b> concept map		☼☼☼ <b>lc</b> iceberg		☼☼☼ <b>Lm</b> learning map	
☼☼☼ <b>Tk</b> tabley box plot		☼☼☼ <b>Sp</b> spirogram		☼☼☼ <b>Da</b> data map		☼☼☼ <b>Tp</b> tremap		☼☼☼ <b>Cn</b> connet		☼☼☼ <b>Sy</b> system type / simulation		☼☼☼ <b>Df</b> data flow diagram		☼☼☼ <b>Se</b> semantic network		☼☼☼ <b>So</b> soft system modeling		☼☼☼ <b>Sn</b> energy map		☼☼☼ <b>Fo</b> force field diagram		☼☼☼ <b>Ib</b> in-organization map		☼☼☼ <b>Pr</b> process event chain		☼☼☼ <b>Pe</b> pert chart		☼☼☼ <b>Ev</b> evacuation knowledge map		☼☼☼ <b>V</b> tree diagram		☼☼☼ <b>Hh</b> heaven 'x' bell chart		☼☼☼ <b>I</b> infomural	
☼☼☼ <b>Cy</b> Process Visualization		☼☼☼ <b>Hy</b> Structure Visualization		☼☼☼ <b>Su</b> supply demand curve		☼☼☼ <b>Pe</b> performance charting		☼☼☼ <b>St</b> strategy map		☼☼☼ <b>Oc</b> organisation chart		☼☼☼ <b>Ho</b> house of quality		☼☼☼ <b>Fd</b> feedback diagram		☼☼☼ <b>Ft</b> future tree		☼☼☼ <b>Mq</b> map, quadrant		☼☼☼ <b>Ld</b> life-cycle diagram		☼☼☼ <b>Po</b> porter's five forces		☼☼☼ <b>S</b> s-cycle		☼☼☼ <b>Sm</b> stakeholder map		☼☼☼ <b>Is</b> ishikawa diagram		☼☼☼ <b>Tc</b> technology roadmap					
☼☼☼ <b>Ed</b> edgeworth box		☼☼☼ <b>Pf</b> portfolio diagram		☼☼☼ <b>Sg</b> strategic game board		☼☼☼ <b>Mz</b> marketing's organograph		☼☼☼ <b>Z</b> zwoy's morphological box		☼☼☼ <b>Ad</b> advertising diagram		☼☼☼ <b>De</b> decision discovery diagram		☼☼☼ <b>Bm</b> bcg matrix		☼☼☼ <b>Stc</b> strategic canvas		☼☼☼ <b>Vc</b> value chain		☼☼☼ <b>Hy</b> hype-cycle		☼☼☼ <b>Sr</b> stakeholder rating map		☼☼☼ <b>Ta</b> tag		☼☼☼ <b>Sd</b> spiral diagram									
☼☼☼ <b>Overview</b>		☼☼☼ <b>Detail</b>		☼☼☼ <b>Detail AND Overview</b>		☼☼☼ <b>Divergent thinking</b>		☼☼☼ <b>Convergent thinking</b>																											

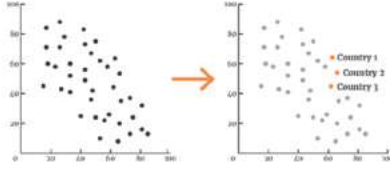
Üretilen görselleştirmelerin pek çoğu, periyodik tabloda yer alan yöntemlerden yararlanılarak geliştirilmektedir. Tablo, hangi veri seti için hangi görsel seçilmeli, yöntem ne olmalı gibi sorulara yanıtlar vermektedir. İşlediğimiz veri görselleştirme bölümünün tablodaki konumu; "sarı renkli" olanları temsil etmektedir. Tablo interaktif olarak buradan incelenmelidir: [https://www.visual-literacy.org/periodic\\_table/periodic\\_table.html](https://www.visual-literacy.org/periodic_table/periodic_table.html)

Örneğin; periyodik tabloda sarı bölümlerde yer alan bazı harfler; C (süreç), T (tablo), Pi (Yuvarlak Diyagram), L (Çizgi Grafik), B (Çubuk Grafik) vb. gibi, veri ile görselleştirme yaparken önerilen grafikleri kapsar. Cy, Hy ve harfler üzerinde yer alan semboller ise, yapılan görselleştirmenin neye yanıt verebileceğini içerir. Tablo, henüz Türkçe'ye kazandırılmış olmasa da, link üzerinden interaktif kısmını inceleyerek ve üzerinde biraz zaman geçirerek, görselleştirme yöntem tablosunu daha iyi anlama şansı bulabilirsiniz.

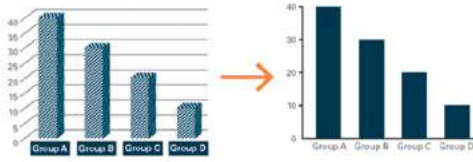
## Veri Görselleştirme Prensipleri /Kuralları

Kullanılan verinin grafikte net bir şekilde gösterilmesi, gereksiz biçimlendirmelerden uzak durulması, grafiklere başlık, eksenlere isim vermek, lejant, veri etiketi, yatay-dikey eksen açıklamaları, kaynak belirtmek, statik görsellerde veri etiketi kullanmak ve doğru font seçmek, öncelikli veri görselleştirme kurallarındandır.

## Veriyi göstermek



## Gereksiz biçimlendirmeleri azaltmak

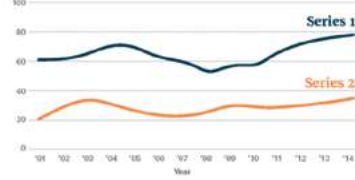


## Grafiklere başlık, eksenlere isim vermek

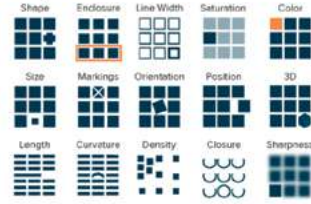
(İçerik, veri etiketi, yatay-dikey eksen açıklamaları, kaynak belirtme, statik görsellerde veri etiketi kullanma, doğru font seçmek)

## Chart Title Here

(Y axis label here)

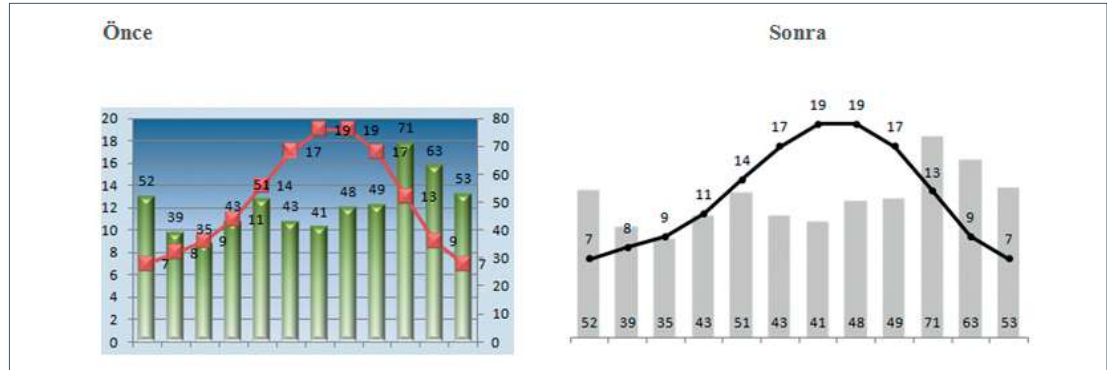


## Doğru Renk Seçmek

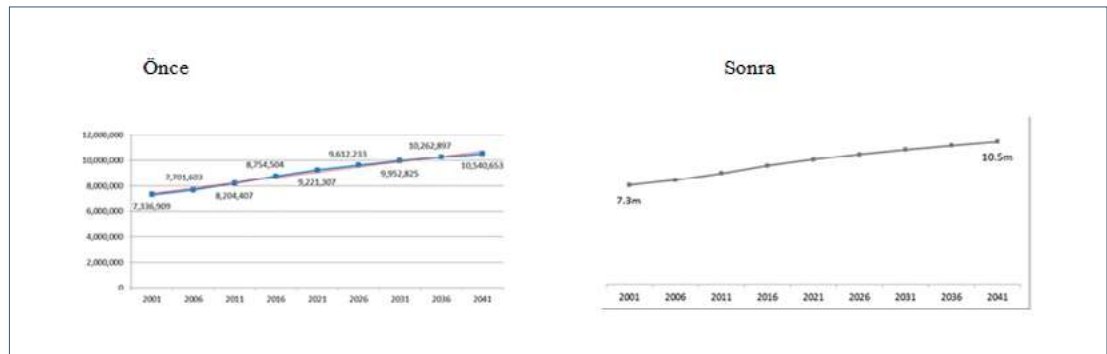


Kaynak

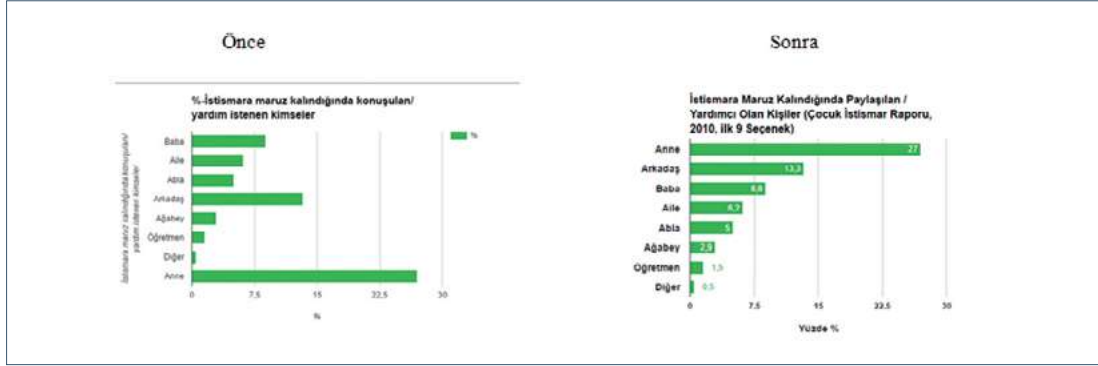
## 1) Gereksiz Biçimlendirmeden Uzak Durma



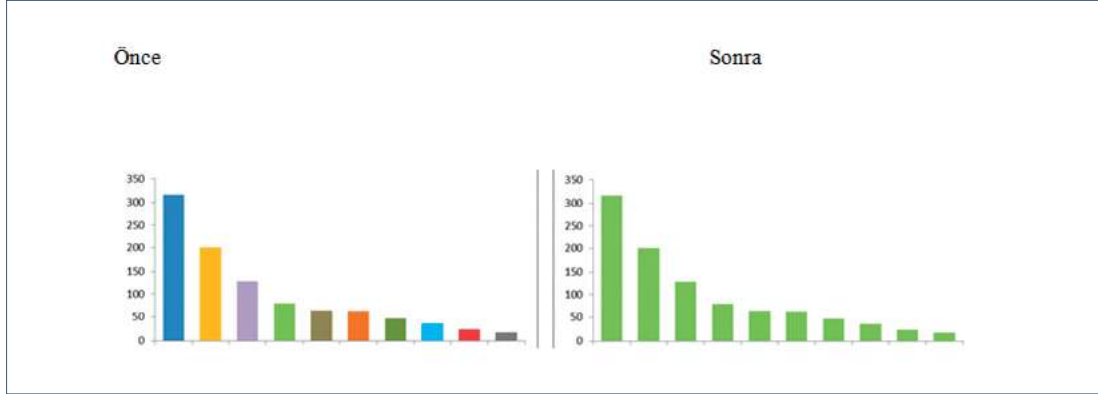
## 2) Gereksiz Dağınıklıkların Kaldırılması



### 3) Grafik Yapmadan Önce Filtreleme



### 4) Gereksiz Renk Kullanılmaması

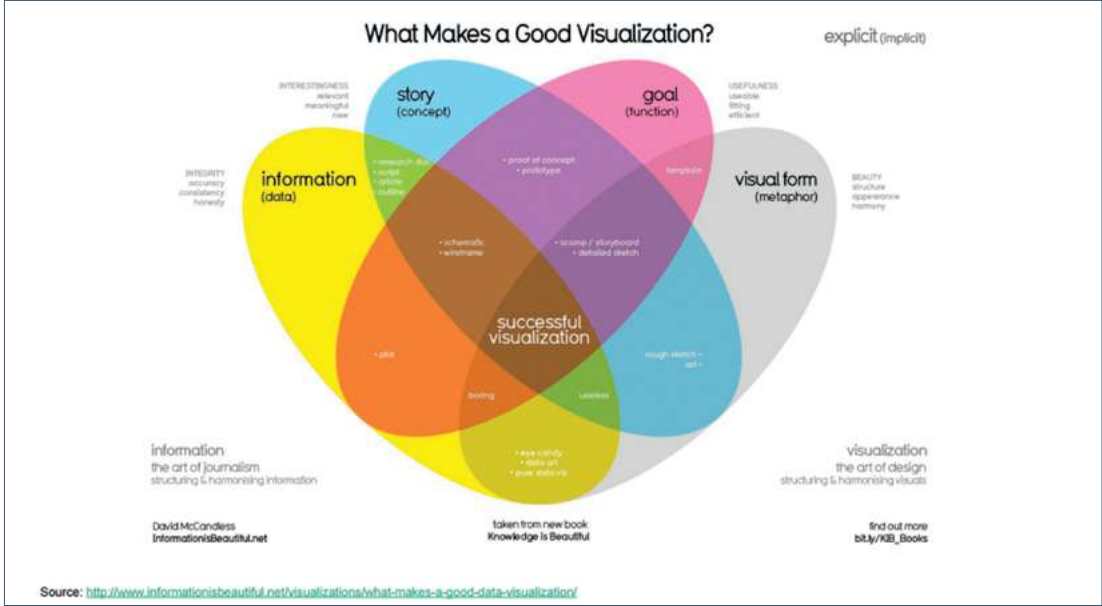


### Veri Görselleştirmeyi Geliştirme Yöntemleri

- Veri seti için, doğru görselleştirme aracı seçildiğinden emin olunmalıdır. Rahat kullanılan, pratik yapma şansının en fazla olduğu araç seçilmelidir.
- Bilginin doğru sunulduğundan emin olunmalıdır, şüpheli yaklaşılmalıdır. (Görselleştirme aracının; veride çarpıklıklar oluşturması, verinin yanlış sunulmasına/yorumlanmasına neden olması gibi).
- Bilgiyi mantık çerçevesinde analiz etmek/sıralamak/kullanmak (Kategoriye dikkat etmek, alfabetik sıralamak vb. ) gerekmektedir. Bunu yapmak, okuyucunun veri ile daha fazla iletişime geçmesini sağlar.
- Karşılaştırılması kolay olmalıdır. (İzleyen/okuyan için kolay filtrelenebilir/ karşılaştırılmalıdır, kolay anlam çıkartılmalıdır ve bilgi alınmalıdır.)
- Verinin gereken dikkati çekmesi sağlanmalıdır. (Gereksiz ikonlar, semboller, ek görselleştirmeler ile okuyucuların dikkati dağıtılmamalıdır).
- Uygun renk paletleri ve simgeler kullanılmalıdır. (Anlamı artırır. Gereksiz etiketlemeleri kaldırmak önemlidir.) Örneğin; ColorTab kullanılabilir.

## Veri Görselleştirme Elementleri

Veri görselleştirme beş elementten oluşur: **Veri** > **Görünüm/Tasarım** > **Bulgu** > **Hikâye** > **Etkileşim**. Bu elementler, görselleştirmenin iyi olmasını da sağlayan önemli ana başlıklardır.



**Veri:** Verisiz hiçbir şey sunamazsınız. (Etkileşim için veri gereklidir.)

**Görünüm:** Veriyi analiz etmenin en iyi yöntemidir; görünüm/tasarım. (Görmek inanmaktır.)

**Bulgular:** İyi bir haber/hikâye için gerekli olabilecek görsel trendleri, farklılıkları, vb. keşfetmek.

**Hikâye/Haber/İçerik/Rapor:** Veriden çıkan bulgular, hikâyenin/raporun/araştırmanın gelişmesini güçlendirir. Bulunan hikâye/çıkarmı, verinin önemini de gösterir.

**Etki:** Görselleştirilen verinin etki yaratması, farklı olanı ortaya koyabilmesidir.

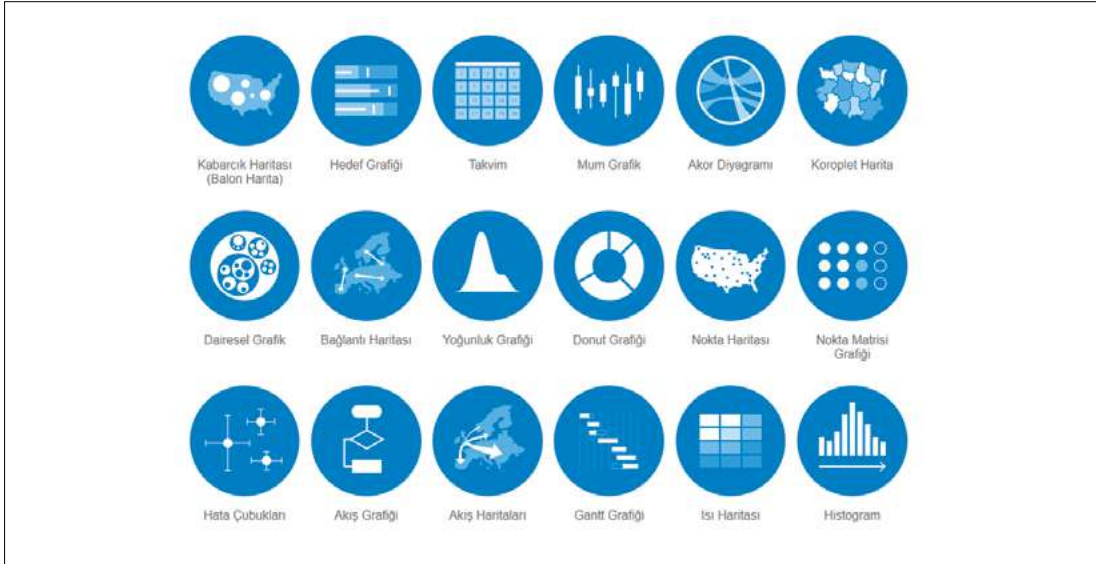
## Veri Nasıl Sunulur?

Veri seti ile değişik kategorilerden değerleri karşılaştırmak için çubuk grafik; bir değeri zaman içinde takip etmek için çizgi grafik; iki veri arasındaki etkileşimi göstermek için serpilme grafik; coğrafyayla ilgili veriyi görselleştirmek için ise haritalandırma çalışması yapabilirsiniz.

Yapmak İsteddiğiniz:	Seçilen:
Değişik kategorilerden değerleri karşılaştırmak	Çubuk grafik
Bir değeri zaman içinde takip etmek (zaman dizisi)	Çizgi grafik
İki veri arasındaki etkileşimi göstermek	Serpilme grafik
Coğrafyayla ilgili veri göstermek	Harita

## Türkçe Veri Görselleştirme Kataloğunun Kullanımı

Veri Görselleştirme Kataloğu; veri ile çalışan kişilerin sıklıkla yararlandığı önemli bir araçtır. Türkçesi [datavizcatalogue.com/tr](http://datavizcatalogue.com/tr) olarak orijinal sitede subdomain olarak yer almaktadır. 60 görselleştirmenin çevrildiği katalogda, 'Fonksiyona Göre Ara', 'Liste Olarak Gör' bölümleri yer aldığı gibi, ana sayfada yer alan her bir grafik tıkladığında tanım, kullanım fonksiyonu, örnekler, benzer grafikler, görselleştirme sağlayan araçlar, anatomi de yer almaktadır.



### Türkçe Veri Görselleştirme Kataloğu Ekran Görüntüsü

Veri görselleştirme projelerine ilk kez başlayanlar için önemli bir rehber niteliği taşıyan katalog, temel grafiklerden, programlama gerektiren grafiklere kadar çok sayıda görsel yöntem sunuyor.

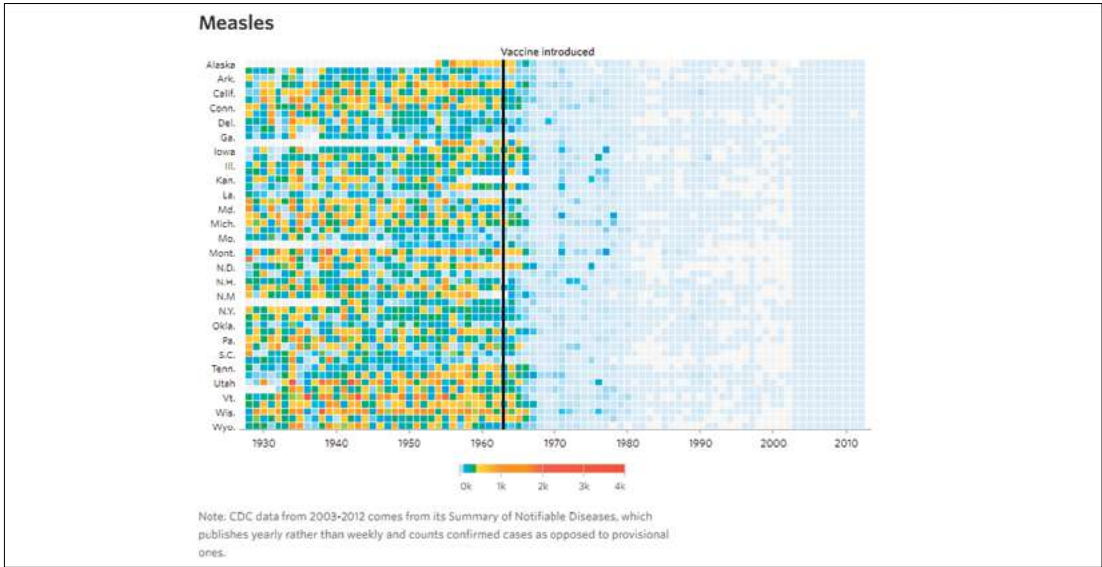


## İyi Veri Görselleştirme Örnekleri

### 1. Örnek

#### 20. Yüzyılda Bulaşıcı Hastalıklarla Mücadele: Aşıların Etkisi

Çalışma, 2015 yılında veri görselleştirme kategorisinde ödül almıştır. Yaklaşık 70 yıl içinde ve 50 eyaletin tamamında, enfekte insan sayısının aşı sonrası nasıl değiştiğini göstermektedir. Isı haritaları 100.000 bin kişi başına düşen vaka sayısını göstermektedir. Aşının önemine dikkat çeken veri görselleştirme çalışmasında; kızamık, kabakulak, boğmaca, çocuk felci, kızamıkçık ve çiçek hastalığı ile mücadelede aşı sonrası vaka düşüşlerini görüyorsunuz. Rahat okunan grafikleri, açık başlıkları, verilerin herkesin erişimine açık olması, hem tasarım hem de renk seçimi, görselleştirmeyi anlamayı kolaylaştırmaktadır. Aşağıda bulunan kızamık örneğiyle 1980 sonrası kızamık vakasına rastlanmadığını dahi görebilirsiniz.



#### 20. Yüzyılda Bulaşıcı Hastalıklarla Mücadele: Aşıların Etkisi

Diğer tüm hastalıkların veri görselleştirmelerine bu linkten erişebilirsiniz:

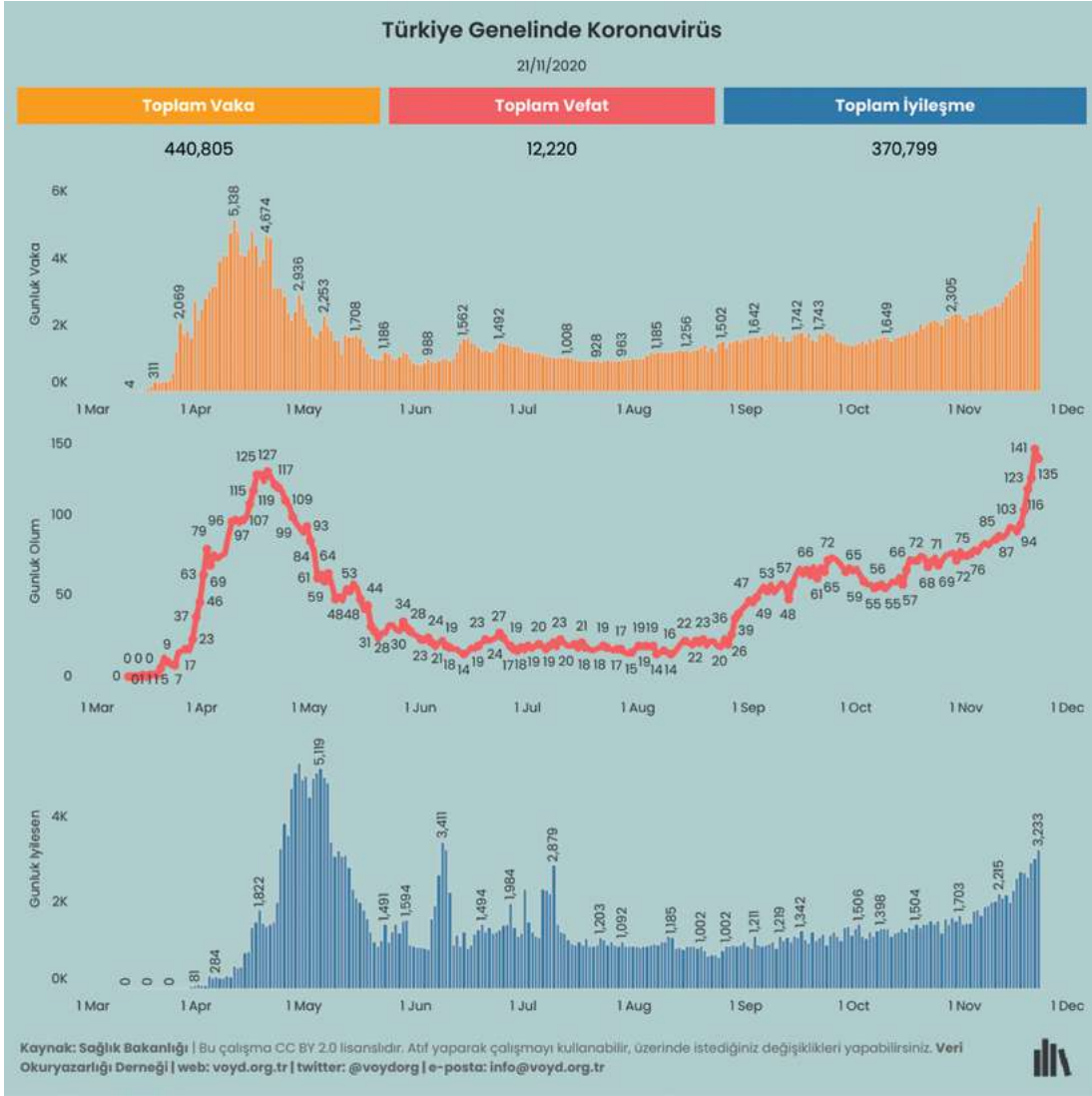
<http://graphics.wsj.com/infectious-diseases-and-vaccines/>

### 2. Örnek

#### Koronavirüs Takip Sayfası [Güncelleniyor...]

[Veri Okuryazarlığı Derneği \(VOYD\)](http://veribulteni.voyd.org.tr)'nin [veribulteni.voyd.org.tr](http://veribulteni.voyd.org.tr)'de düzenli olarak güncellediği Koronavirüs takip sayfası, çoklu veri görselleştirme panosu da iyi örneklerden biridir. Verilere Github sayfasından ulaşmak mümkün, tüm veriler olmasa da günlük açıklanan vaka, vefat, iyileşme verilerinin zaman içinde nasıl bir değişim gösterdiği rahat bir şekilde okunmaktadır. Görselleştirmeyi indirmek ve interaktif incelemek mümkündür.





### Türkiye Geninde Koronavirüs Grafiği / Voyd -2020

Sayfaya buradan ulaşabilirsiniz: <https://veribulteni.voyd.org.tr/koronavirus-takip/>

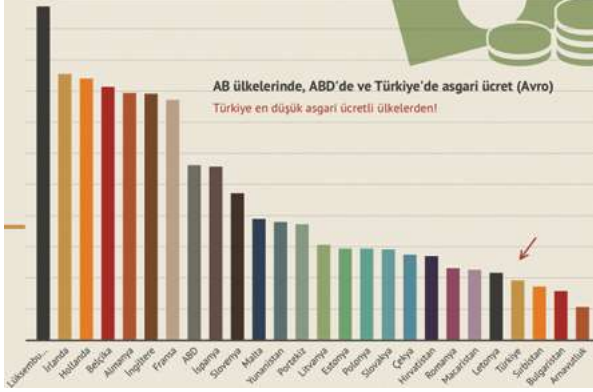
Bir veri görselleştirme çalışmasının; etkin, güçlü ya da iyi olup olmadığını en doğru şekilde anlamanın ya da üretmenin yolu daha sık uygulama yapmaktır ve daha fazla iyi örnek incelemektir. İyi bir veri görsel okuryazarı olmanın, ilk ve en etkili yolu da budur: uygulamak, örnek incelemek ve grafikleri okumaktır. Daha fazla iyi örnekleri, buradan takip edebilirsiniz: <https://informationisbeautiful.net/beautifulnews/>

## Kötü Veri Görselleştirme Örnekleri

### 1.Örnek

#### Farklı ülkelerdeki asgari ücretlerin kıyaslaması

Sütun ve çubuk grafikler kategorik verilerin karşılaştırılmasında kullanılır. AB ülkelerinde, ABD’de ve Türkiye’de asgari ücret (Avro) grafiği, Türkiye’nin en düşük asgari ücretli ülkelerden biri olduğunu açıklamak için hazırlanmış ancak renklerin neyi temsil ettiği açık değildir. Aynı renk seçiminin yapıldığı ülkeler karşılaştırma yapmaya olanak sağlamamaktadır. Grafikte, renklerin neyi temsil ettiğine dair bir açıklama veya ipucu olmadığından karşılaştırma yapmak olanaksızdır.



Grafik buradan incelenebilir: Aposto <https://apos.to/s/5fd2115c0952ab0008691609>

### 2. Örnek

<https://www.nature.com/>’da yer alan bu harita, deniz canlıları ile ilgili bir çalışmayı temsil etmektedir; ancak kullanılan veri etiketlerini, renkleri ve karşılaştırmaları anlamak mümkün değildir. Bazı verilerin lejant olmadığı da açıktır. Haritanın statik olması nedeniyle de okunması oldukça zor bir veri görselleştirme çalışmasıdır.



Haritaya bu adresten ulaşabilirsiniz: <https://www.nature.com/articles/s41396-019-0482-0/figures/7>

Dünyada ve Türkiye’de yapılan veri görselleştirmelerde, mantık ve tasarım hatalarına rastlanabilmektedir. Bu hataları azaltmanın en iyi yöntemi, şüphesiz, daha fazla veri gör-

selleştirme örneklerini, incelemek ve okumaktır. Kötü veri görselleştirme örneklerini buradan <https://viz.wtf/> takip edebilir, yapılan çalışmanın neden hatalı olduğunu kendinizi test ederek geliştirebilirsiniz.

## Veri Görselleştirme Araçları

Kılavuzlar ve el kitapları uygulama gerektiren çalışmalar için her zaman çok faydalı olmayabilir. Özellikle veri görselleştirme projeleri geliştirirken, basılı kaynaklardan çok video-ders kayıtları ve çevrimiçi öğretici rehberler, daha etkin şekilde veri görselleştirme araçlarını kullanmanızı sağlayacaktır.

Bazı veri görselleştirme araçları şöyle:

## E-Tablo İle Veri Görselleştirme

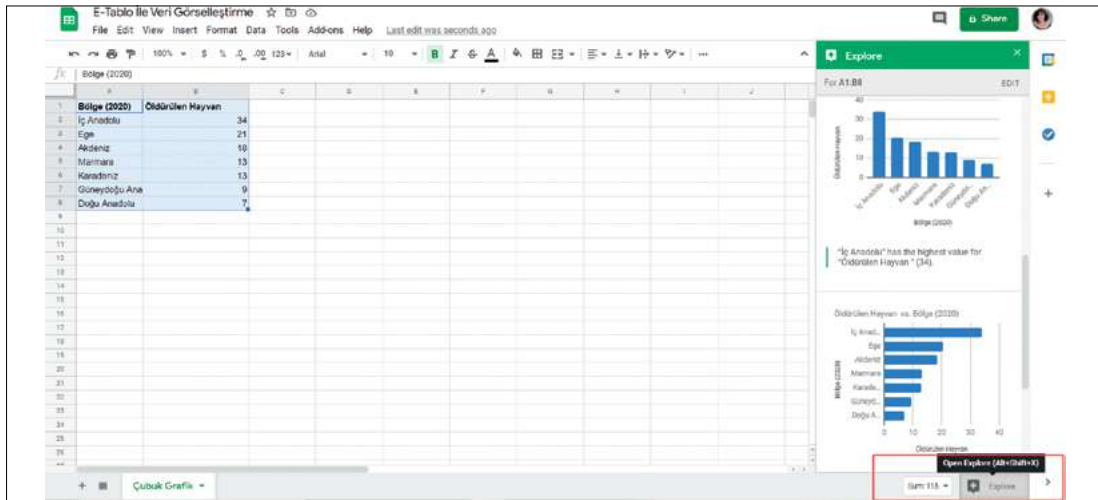
<https://www.google.com/sheets/about/>, Google E-tablolar verilerinizi grafiklerle sunmanızı sağlar. 30'dan fazla veri görselleştirme tekniği kullanmanıza olanak sunar ve interaktif olarak sitenize de ekleme imkânı sağlar.

## Örnek Çubuk Grafik Oluşturmak

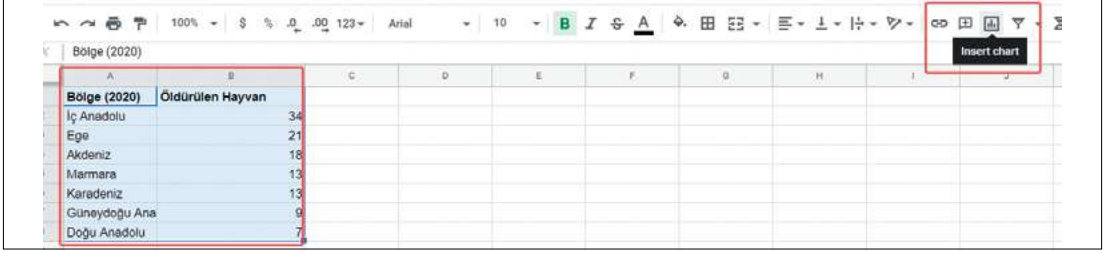
Kullanılan veri 20 Aralık 2020 tarihine kadar medyaya yansıyan öldürülen hayvanların verisidir.

Sütunları seçip sağ alt köşede yer alan *Hızlı Grafik* bölümünü kullanarak *Çubuk Grafik* çalışabilirsiniz.

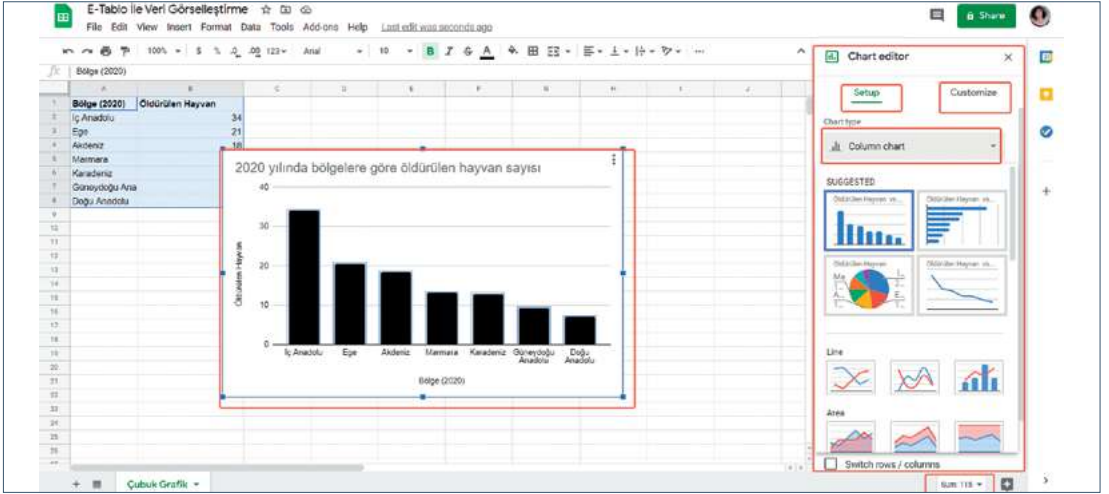
fx	Bölge (2020)		
	A	B	C
1	<b>Bölge (2020)</b>	<b>Öldürülen Hayvan</b>	
2	İç Anadolu		34
3	Ege		21
4	Akdeniz		18
5	Marmara		13
6	Karadeniz		13
7	Güneydoğu Ana		9
8	Doğu Anadolu		7
9			



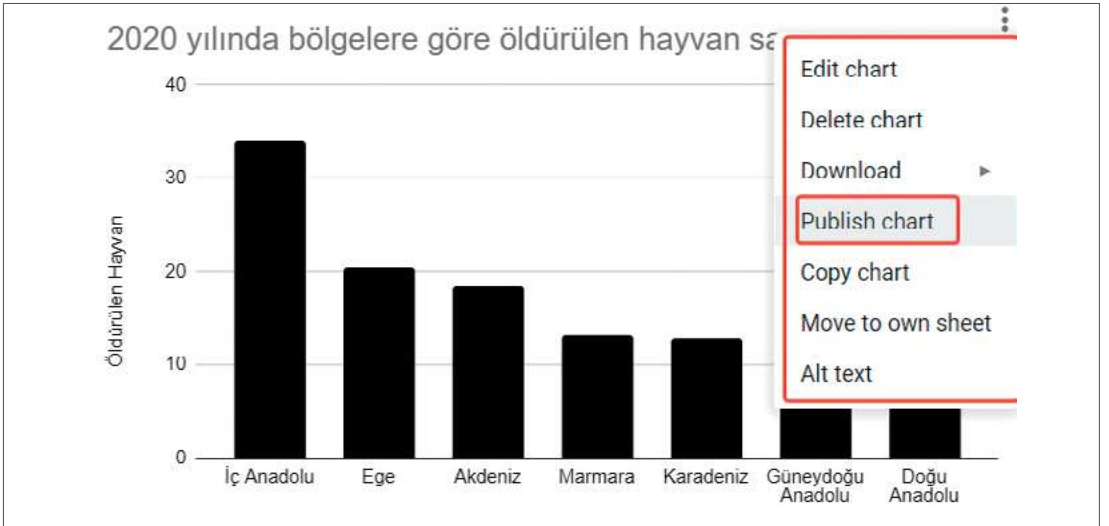
Ya da üst bölümde yer alan grafik simbolünü tıklayıp, 'grafik ekle'yi seçerek; verinize en uygun veri görselleştirme türünü seçebilirsiniz.

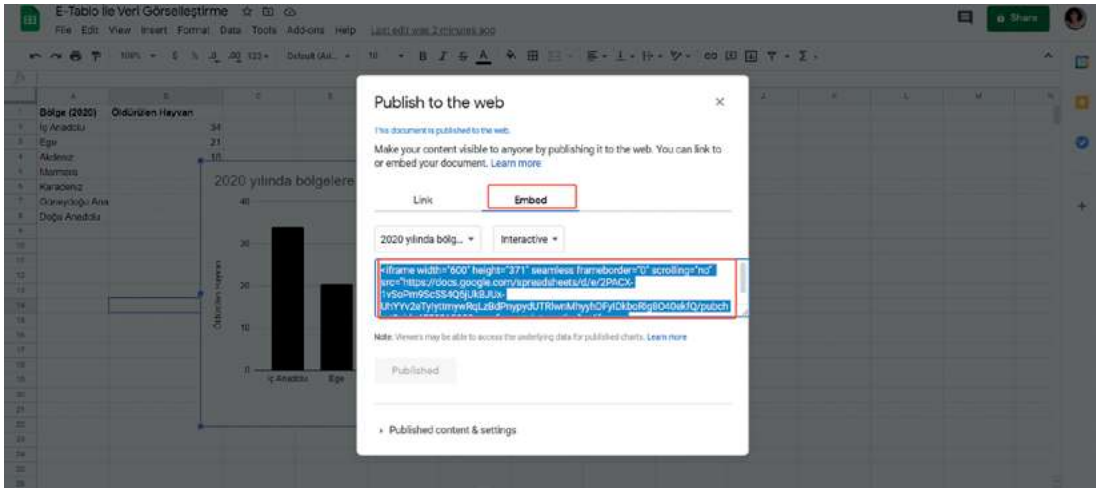
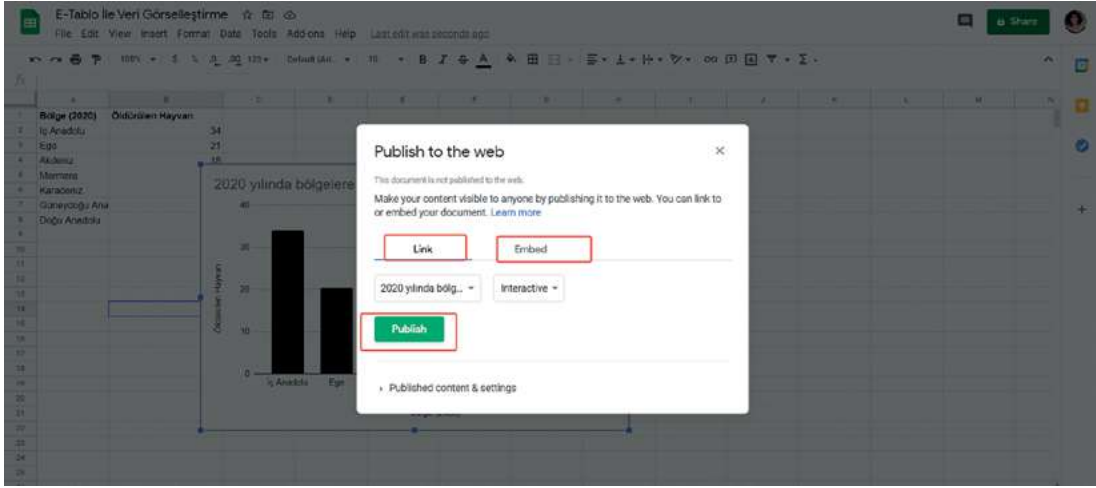


Ardından verinizi kullanmadan önce renk, boyut, başlık, alt başlık ve daha pek çok detay üzerinden özelleştirebilirsiniz.

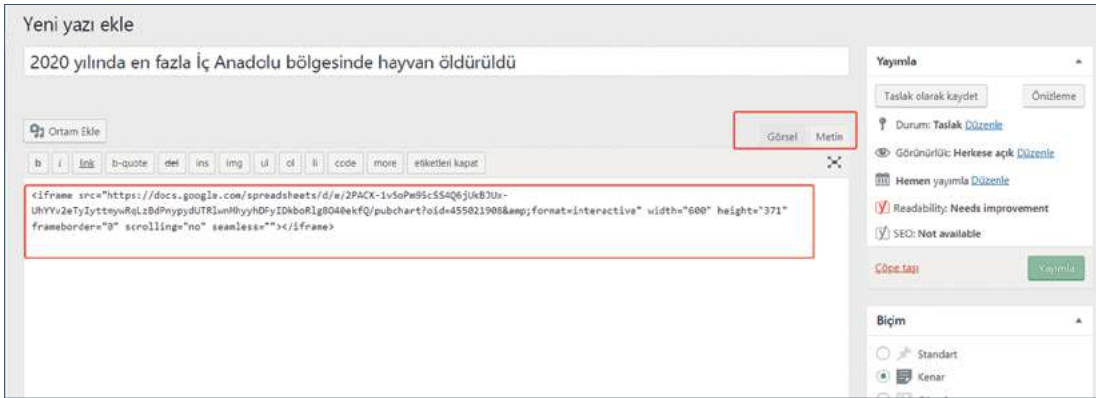


Çubuk grafiği tıklayarak, yayımlayabilir ya da kodunu sitenize ekleyerek kullanabilirsiniz.





Yayınladıktan sonra WordPress paneline ekleyebilirsiniz.









## Veri yükleme ya da yapıştırma işlemi

Ay	Ölen Sağlık Çalışanı
1 Mart	5
2 Nisan	31
3 Mayıs	10
4 Haziran	7
5 Temmuz	5
6 Ağustos	18
7 Eylül	32
8 Ekim	29
9 Kasım	64
10 Aralık	107
11 Ocak 2021	

## Veriyi görselleştirme ve erişime açma işlemi

**2020 yılında Covid-19 nedeniyle hayatını kaybeden sağlık çalışanları**  
10 ayda 332 sağlık çalışanı hayatını kaybetti. En fazla ölüm 107 kişiyle aralık ayında gerçekleşti.

Ay	Ölen Sağlık Çalışanı
1 Mart	5
2 Nisan	31
3 Mayıs	10
4 Haziran	7
5 Temmuz	5
6 Ağustos	18
7 Eylül	32
8 Ekim	29
9 Kasım	64
10 Aralık	107
11 Ocak 2021	107

Kaynak: Türk Tabipleri Birliği - Hazırlayan: Pınar Dağ

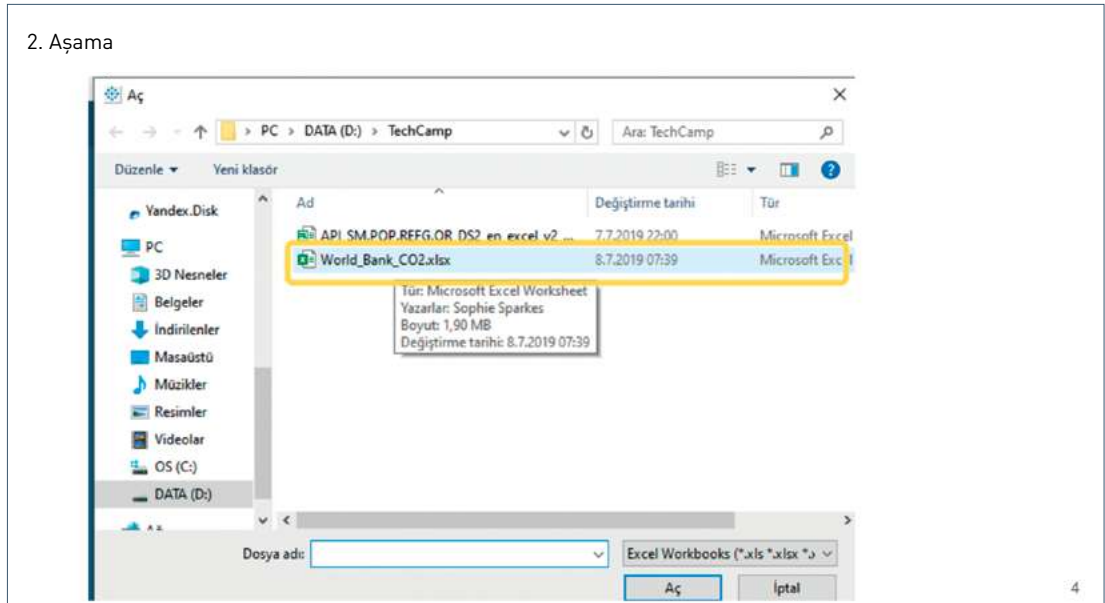
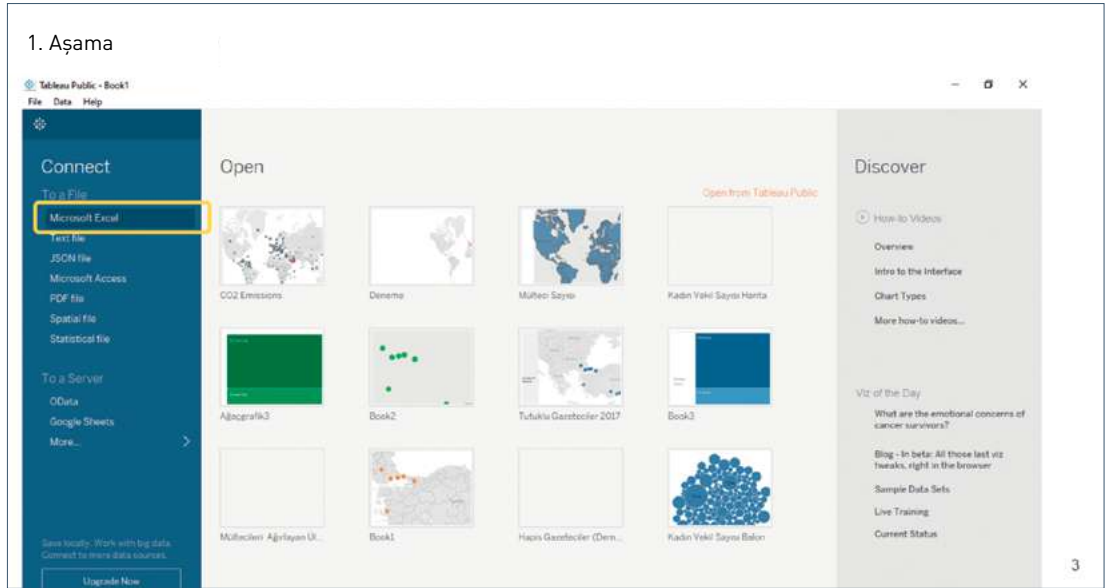
<https://public.flourish.studio/visualisation/4955925/>

## Tableau Public İle Veri Görselleştirme

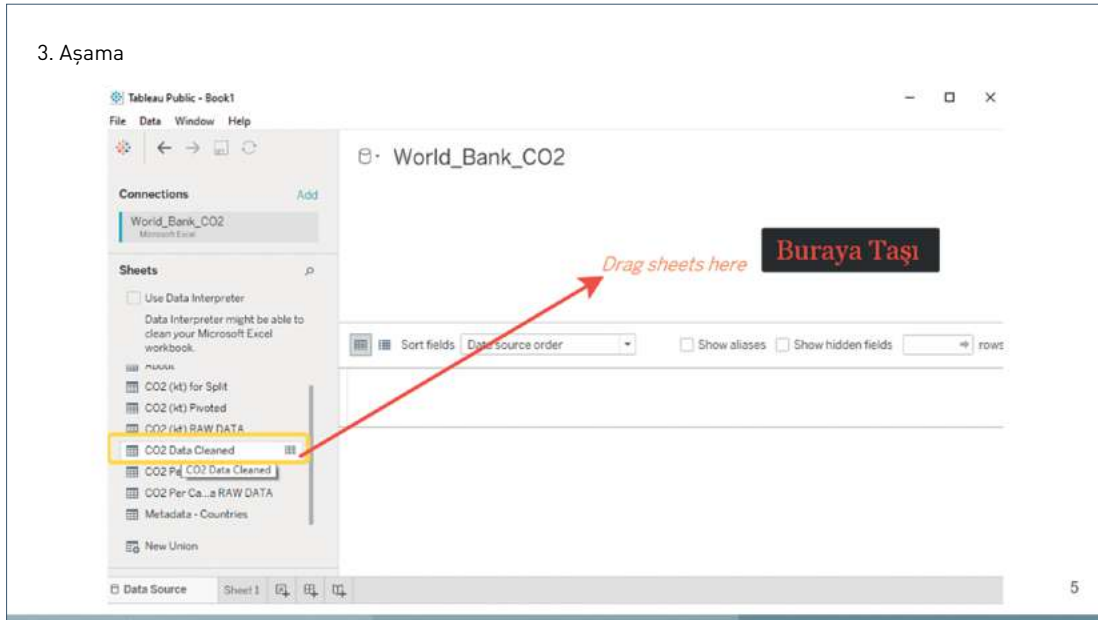
<https://public.tableau.com/en-us/s/>, Tableau Public bir veri analitik ve görselleştirme aracıdır. 10 Gb limitli araç, sürükle-bırak mantığı ile çalışmaktadır. Araç, tüm veri görselleştirme tekniklerini barındırmaktadır. Çalışmayı paylaşma ve embed etme imkânı sunan araç, görselleştirmelerinizi pdf ve fotoğraf (png) olarak dışarı aktarmanızı da sağlamaktadır.

### Örnek çalışma üzerinden inceleyelim:

Dünya Bankası'na ait olan, karbon emisyon verileri üzerinden aşama aşama harita çalışması, çizgi grafikleri, dashboard paneli çalışmalarını göreceksiniz.

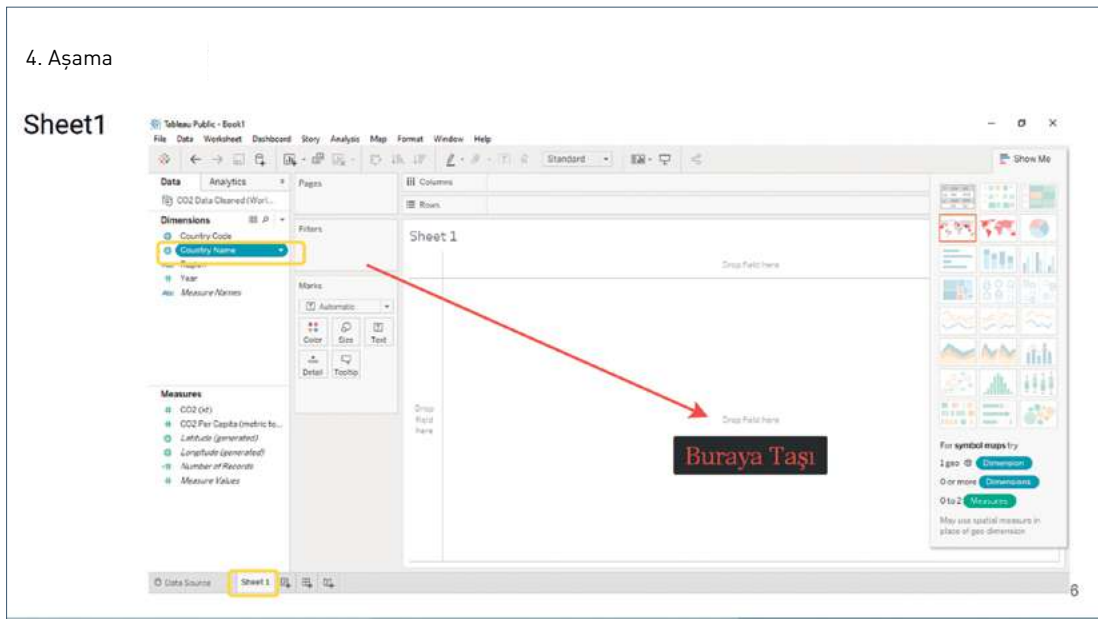


## 3. Aşama



## 4. Aşama

## Sheet1



### 5.Aşama

The screenshot shows the Tableau Public interface. The main view is a world map with blue dots representing data points. A red arrow points to a specific dot in the Middle East region, which is labeled "Buraya Taşı" (Move Here). The interface includes a sidebar with Dimensions (Country Code, Country Name, Region, Year, Measure Names) and Measures (CO2 (t/capita), Latitude (generated), Longitude (generated), Number of Records, Measure Values). The main view shows a map with a grid and a "Show Me" panel on the right.

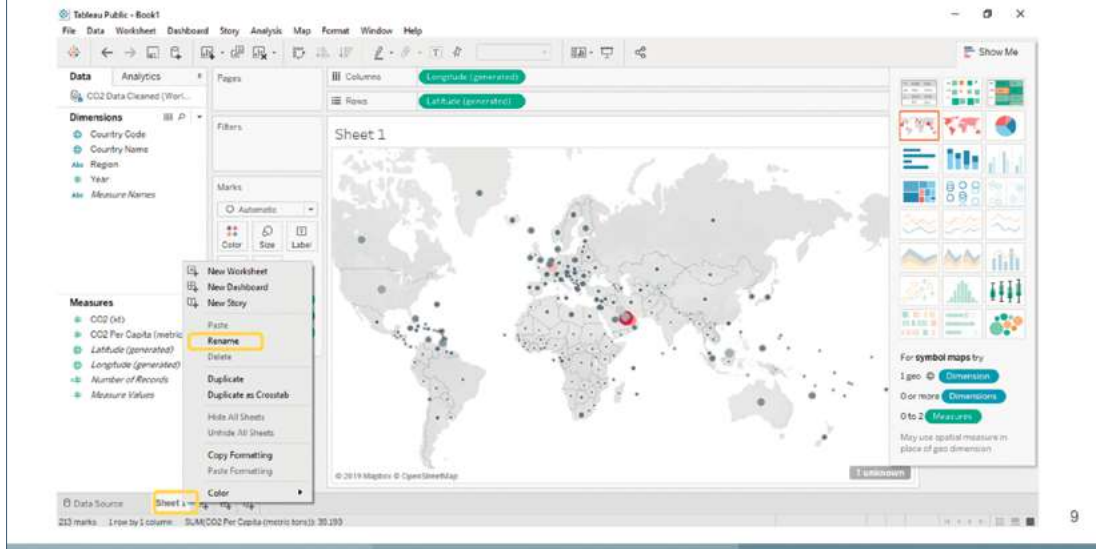
7

### 6.Aşama ( Renk ekleme)

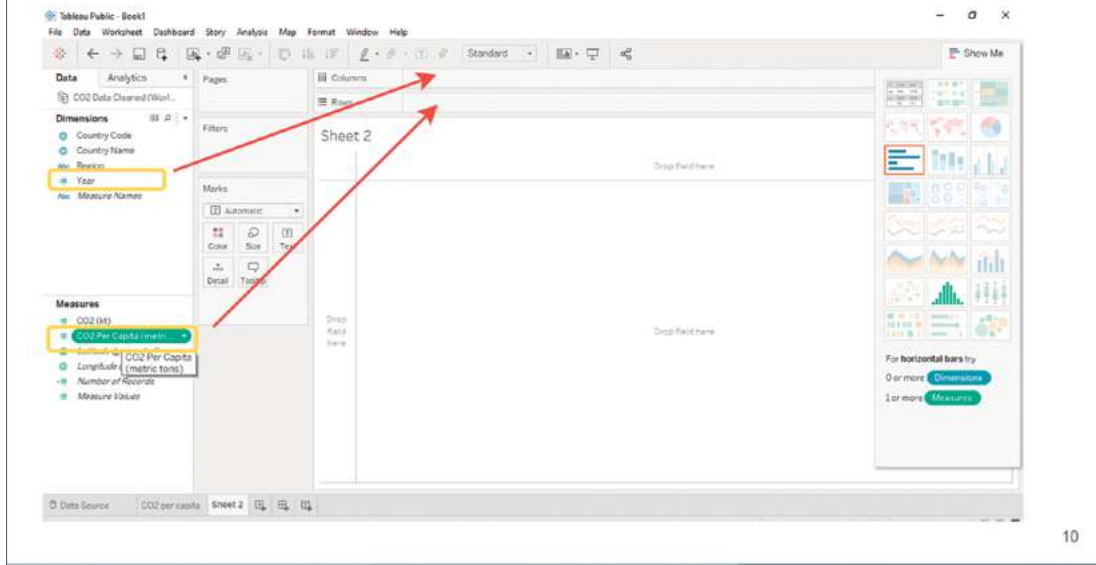
The screenshot shows the Tableau Public interface with the "Edit Colors" dialog box open for the "CO2 Per Capita (metric tons)" measure. The dialog box is open, showing a color palette with a gradient from dark blue to red. The "Reversed" checkbox is checked, and the "Use Full Color Range" checkbox is also checked. The "Include Totals" checkbox is unchecked. The "Advanced >>" button is visible. The background shows the same world map as in the previous screenshot.

8

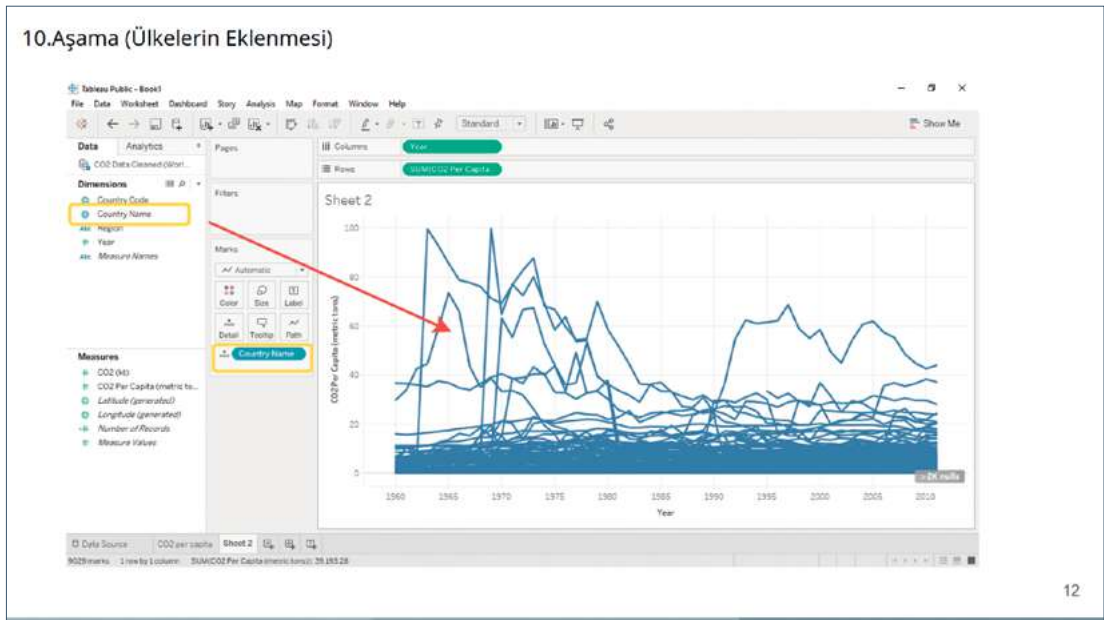
## 7.Aşama (Rename -Ad ekleme) -CO2 per capita



## 8.Aşama Sheet 2 (CO2'nin zaman içindeki değişimi için ikinci grafik)

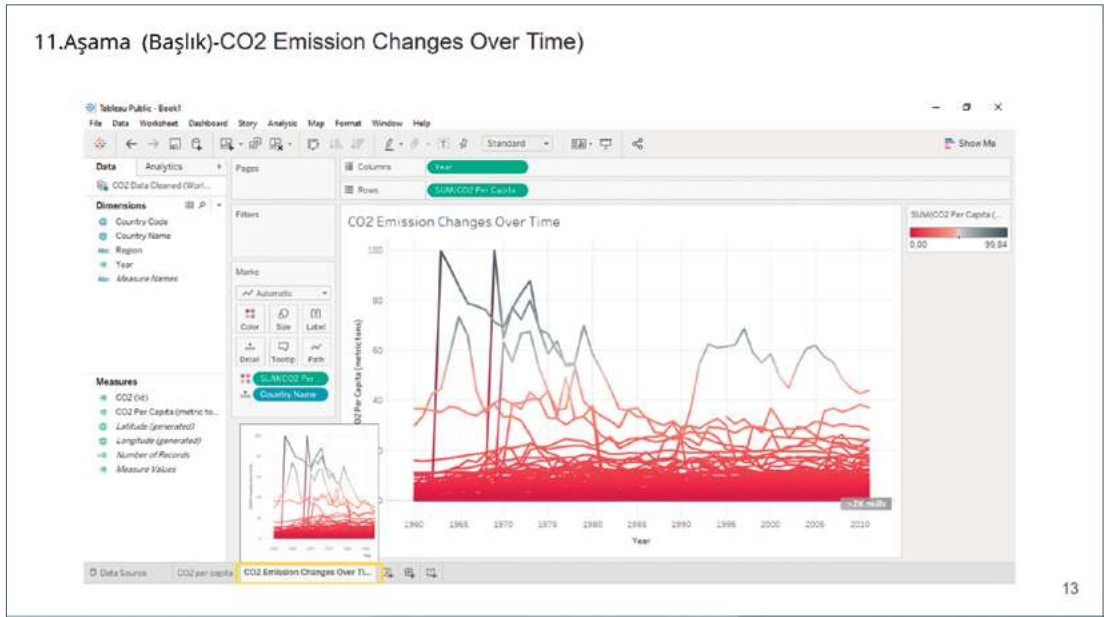


## 10.Aşama (Ülkelerin Eklenmesi)



12

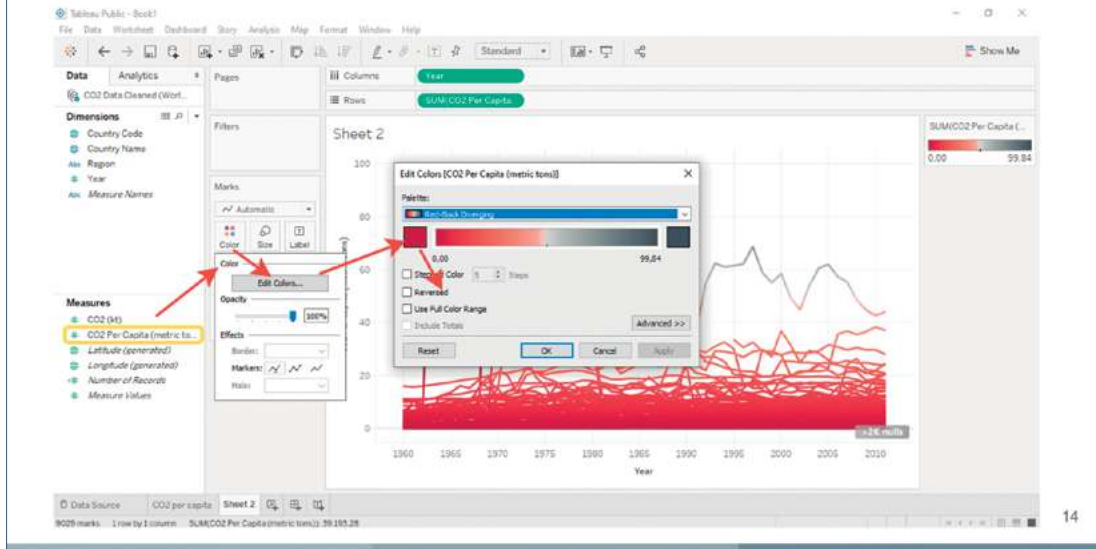
## 11.Aşama (Başlık)-CO2 Emission Changes Over Time



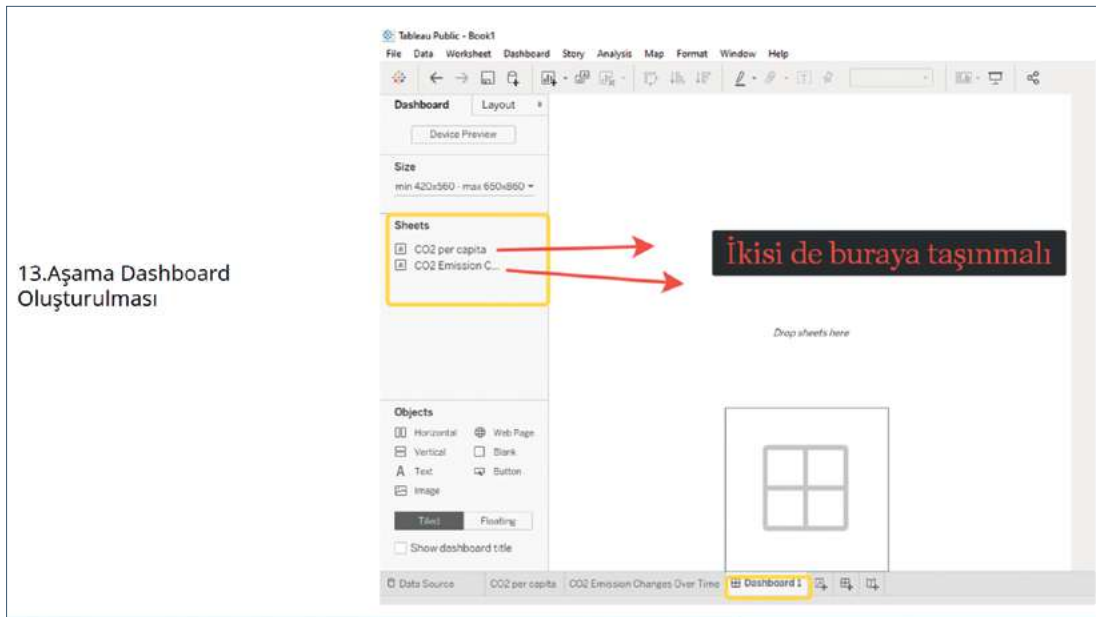
13



## 12. Aşama (Renk -Lejant )



## 13. Aşama Dashboard Oluşturulması



## 14.Aşama

The screenshot shows the Tableau Public interface for a dashboard titled 'CO2 per capita'. The dashboard contains two visualizations: a world map showing CO2 emissions by country and a line chart titled 'CO2 Emission Changes Over Time' showing CO2 per capita (metric tons) from 1960 to 2012. The interface includes a menu bar (File, Data, Worksheet, Dashboard, Story, Map, Format, Window, Help), a toolbar, and a sidebar with 'Size' (min 420x560, max 650x860), 'Sheets' (CO2 per capita, CO2 Emission C...), and 'Objects' (Horizontal, Vertical, Text, Image, Web Page, Blank, Button, Floating). Annotations in red boxes point to the 'Başlık' (Title) of the map, the 'CO2 per capita' field in the 'Size' section, and the legend area of the map.

16

## 15.Aşama (Başlık Ekleme)-Text tut /yukarı taşı

(Which Countries Has The Highest CO2 Carbon Emission Per Capita? Are They The Countries You Expected?)

The screenshot shows the same Tableau Public interface as in the previous step, but with an 'Edit Text' dialog box open. The dialog box contains the text: 'Which Countries Has The Highest CO2 Carbon Emission Per Capita? Are They The Countries You Expected?'. The dialog box has a title bar 'Edit Text' and buttons for 'OK' and 'Cancel'.

17

## 16.Aşama

Web'de  
yayınlama)

The screenshot shows the Tableau Public interface. The 'File' menu is open, and 'Save to Tableau Public Account' is highlighted. The dashboard contains two views: a world map titled 'Which Countries Has The Highest CO2 Carbon Emission Per Capita?' and a line chart titled 'Emission Changes Over Time' showing CO2 per capita from 1960 to 2010. The status bar at the bottom indicates 'Data Source', 'CO2 per capita', 'CO2 Emission Changes Over Time', and 'Dashboard 1'.

17.Aşama (webde yayınlamak için giriş yapılmalı ve çalışmaya isim  
verilmeli)

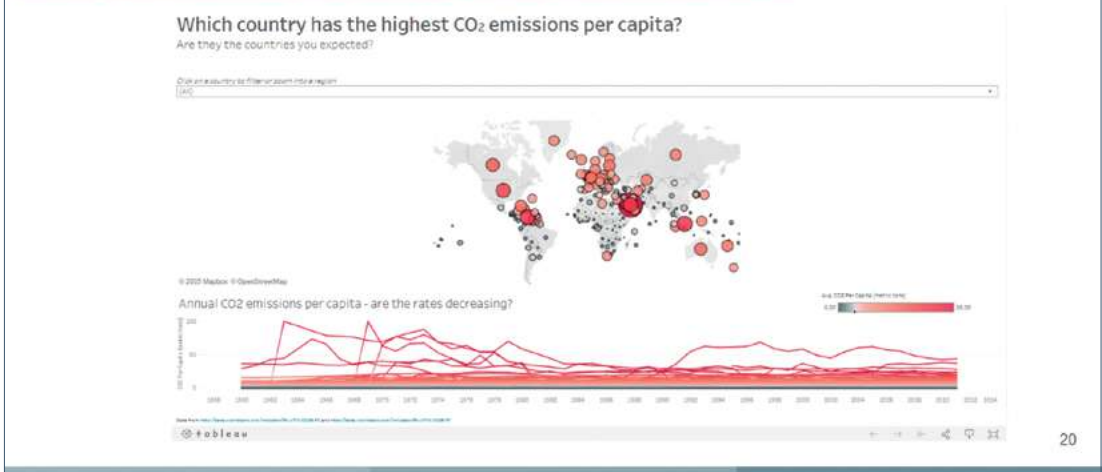
The left screenshot shows the 'Tableau Public Sign In' dialog box with the email 'pinardag@arabonemal.com' and a password field. The right screenshot shows the 'Save Workbook to Tableau Public' dialog box with the workbook name 'CO2 Emission C...' and a 'Save' button highlighted. The dashboard content is visible in the background of both screenshots.

19

## 18.Aşama

Viz: <https://public.tableau.com/profile/pinardag#!/vizhome/CarbonEmission1960-2010/Dashboard1?publish=yes>

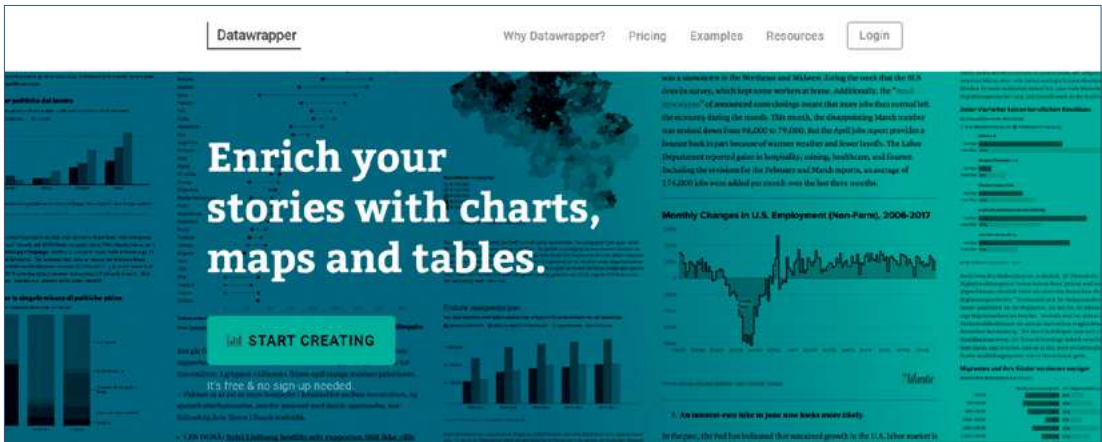
[https://public.tableau.com/profile/pinardag#!/vizhome/Demo\\_15626312462510/Dashboard1?publish=yes](https://public.tableau.com/profile/pinardag#!/vizhome/Demo_15626312462510/Dashboard1?publish=yes)



[https://public.tableau.com/profile/pinardag#!/vizhome/CO2Emissions\\_15625623981380/CO2Emissions](https://public.tableau.com/profile/pinardag#!/vizhome/CO2Emissions_15625623981380/CO2Emissions)

## Datawrapper ile Veri Görselleştirme

<https://www.datawrapper.de/>, Datawrapper, kullanımı oldukça kolay bir veri görselleştirme aracıdır. Sayısız veri grafiği oluşturulmasını sağlayan araç, kullanıcı dostu yapıyla veri görselleştirme araçları arasında en kullanışlı olanlardandır. Dört farklı veri yükleme seçeneği sunan Datawrapper, aynı zamanda belli ölçülerde verinizi düzenleyebilmenizi sağlamaktadır. Bu araçta oluşturduğunuz veri grafiğine ekleyebileceğiniz başlık, açıklama, kaynak ve yayıncı bilgisi sayesinde ürettiğiniz veri grafiği, kullanıcı tarafından anlaşılabilir ve teyit edilebilir olacaktır.



Aracı, örnek emisyon verisi üzerinden aşama aşama inceleyelim.

**Datavrapper** + New Chart New Map New Table River Login / Sign Up Language

1 Upload Data 2 Check & Describe 3 Visualize 4 Publish & Embed

How do you want to upload your data?

Copy & paste data table   
 XLS/CSV upload   
 Import Google Spreadsheet   
 Link external dataset

**Copy & paste your data**

Select your data (including header row/column) in Excel or LibreOffice and paste it in the text field on the right. You can also upload a CSV or Excel file from your computer.

If you just want to try Datavrapper, here's a list of some example datasets you can use.

Select a sample dataset

Paste your copied data here...

**Proceed**

1 Upload Data 2 Check & Describe 3 Visualize 4 Publish & Embed

How do you want to upload your data?

Copy & paste data table   
 XLS/CSV upload   
 Import Google Spreadsheet   
 Link external dataset

**Copy & paste your data**

Select your data (including header row/column) in Excel or LibreOffice and paste it in the text field on the right. You can also upload a CSV or Excel file from your computer.

If you just want to try Datavrapper, here's a list of some example datasets you can use.

Global CO2 Emissions

```
Year;Liquid fuel;Solid fuel;Gas fuel;Cement production;Gas flaring
1890;0;54;0;0;0
1891;0;54;0;0;0
1892;0;57;0;0;0
1893;0;59;0;0;0
1894;0;69;0;0;0
1895;0;71;0;0;0
1896;0;76;0;0;0
1897;0;77;0;0;0
1898;0;78;0;0;0
1899;0;83;0;0;0
1900;0;91;0;0;0
1901;0;95;0;0;0
```

**Proceed**

1 Upload Data 2 Check & Describe 3 Visualize 4 Publish & Embed

**Make sure the data looks right**

Please make sure that Datavrapper interprets your data correctly. In the table, **number** columns should be shown in blue, **dates** in green and **text** in black. A **red** cell indicates a problem in your dataset that needs to be fixed.  cells contain no data.

First row as label

**Output locale**

Defines decimal and thousand separators as well as translation of month and weekday names.

English (en-US)

Click on table header to edit column properties

Sort view by... Search data table

#	A	B	C	D	E	F
1	Year	Liquid fuel	Solid fuel	Gas fuel	Cement production	Gas flaring
2	1890	0	54	0	0	0
3	1891	0	54	0	0	0
4	1892	0	57	0	0	0
5	1893	0	59	0	0	0
6	1894	0	69	0	0	0
7	1895	0	71	0	0	0
8	1896	0	76	0	0	0
9	1897	0	77	0	0	0
10	1898	0	78	0	0	0
11	1899	0	83	0	0	0
12	1900	0	91	0	0	0
13	1861	0	95	0	0	0
14	1862	0	96	0	0	0
15	1863	0	103	0	0	0
16	1864	0	112	0	0	0
17	1865	0	119	0	0	0

Swap rows and columns (transpose)   
 Add column...   
 Revert changes...





1 Upload Data ✓ 2 Check & Describe ✓ 3 Visualize 4 Publish & Embed

**Chart type** Refine Annotate Design

Bar Chart Split Bars Stacked Bars Grouped Bars

Bullet Bars Dot Plot Range Plot Arrow Plot

Column Chart Grouped Column Chart Stacked Column Chart Lines

Area Chart Pie Chart Donut Chart Multiple Pies

Multiple Donuts Scatter Plot Election Donut Table

Archived chart types: —

**Hint:** In case the visualization doesn't look like you expected, you should try to transpose the data

**[Insert title here]**  
Global CO2 Emissions from Fossil-Fuel Burning, Cement Manufacture, and Gas Flaring, 1751-2014.

■ Liquid fuel ■ Solid fuel ■ Gas fuel ■ Cement production ■ Gas flaring

Source: Carbon Dioxide Information Analysis Center - Get the data - Created with Datawrapper

CHART SIZE: 600 x 400

COLORBLIND CHECK: Norm, Deut, Prot, Trit, Achr

< Back Proceed >

1 Upload Data ✓ 2 Check & Describe ✓ 3 Visualize 4 Publish & Embed

**Chart type** Refine Annotate Design

**Vertical axis**

Custom range: 0 - 100

Show grid lines

Show axis labels

Number format: (Automatic)

Custom ticks: e.g. 10,20,30

Position:  outside  inside

Alignment:  left  right

Grid labels will always be displayed on the left when Category labels are displayed directly.

**Horizontal axis**

Automatically sort columns

Reverse order

Rotate labels:  auto  always  never

**Appearance**

Column color: e91e63 customize colors

Show values

on hover  always

Space between columns:  47%

**İskemik Kalp Hastalığı**

Get the data - Created with Datawrapper

CHART SIZE: 400 x 312

COLORBLIND CHECK: Norm, Deut, Prot, Trit, Achr

Deuteranopia is a form of red-green blindness (1% of males).

?

1 Upload Data ✓ 2 Check & Describe ✓ 3 Visualize ✓ 4 Publish & Embed

**Embed your chart or download it as PNG.**

You can embed your chart in your website for free with Datawrapper. Simply enter your e-mail address and we'll send you an email with a link to embed your chart. Already have an account? Log in to embed your chart.

Enter your e-mail:   
e.g. your.email@gmail.com [Send my embed code](#)

**Global CO2**

Global CO2 Emissions from Fossil-Fuel Burning, Cement Manufacture, and Gas Flaring, 1751-2014.

■ Liquid fuel ■ Solid fuel ■ Gas fuel ■ Cement production ■ Gas flaring

Source: Carbon Dioxide Information Analysis Center - Get the data - Created with Datawrapper



## 9.BÖLÜM

### Açık Veri ve Veri Paylaşımı Yöntemleri

[OpenDefinition.org](http://OpenDefinition.org)'un tanımına göre, açık veri herhangi bir telif hakkı, patent ya da diğer kontrol mekanizmalarına tabi olmaksızın, herkes tarafından ücretsiz, özgürce, tekrar kullanılabilen ve dağıtılabilen veridir.

### Açık Veri ve Açık Bilgi

Açık bilgi, açık verinin erişilebilir, anlaşılabilir, anlamlı ve birinin gerçek sorununu çözmeye yardımcı olmasıdır. Açık veriden değer yaratmak; sorun çözer, hayat kalitesine katkı sağlar. Verinin açılması, 2005'li yılların başında güçlü bir şekilde dünya gündemine taşındı; üzerine konferanslar, çalıştaylar yapıldı, açık kaynak projeleri ve açık kaynak araçlar geliştirilmesi teşvik edildi. Devletler, açık veri politikaları geliştirerek açık veri portalları açtı. Ancak açıklık tanımına göre, maalesef dünyada yayımlanan veri setlerinin %89'u açık değil. Bir başka ifadeyle, veri setleri açıklık tanımına göre sadece %11 oranında açık.

### Açık Veri Türleri Neler Olabilir, Özellikleri Nedir?

**Açık verinin temel özellikleri:** Kullanılabilirlik-erişim, tekrar kullanım-dağıtım, evrensel katılım ve makine okunabilirliği.

**Veri hangi durumda açıktır? :** Teknik olarak açıklık ve yasal olarak açıklık

**Teknik olarak açıklık:** Makineler tarafından okunabilen (machine-readable) standart yapıya uygunluk sağlamalıdır. Bilgisayar uygulamaları tarafından alınabilmeli ve anlamlı bir şekilde işleme tabi tutulabilmelidir.

**Yasal olarak açıklık:** Açık bir şekilde lisanslanmalıdır. Herhangi bir sınırlama olmaksızın, ticari ya da ticari olmayan kullanım ve tekrar kullanıma izin vermelidir.

**Veri neden açılmalıdır?:** Şeffaflık, sosyal ve ticari değerlerin serbestleşmesi ve katılımın sağlanması için veriler açılmalıdır.

### Veri Nasıl Açılır?:

**Basit tutularak:** Verinin küçük, basit ve hızlı şekilde işlenmesidir. Tüm veri setlerinin, hemen açılması gerekmez.

**Açık lisans uygulayarak:** Veride, hangi fikri mülkiyet haklarının bulunduğu belirlenmesi ve en uygun açık lisansın uygulanması.

**Veriyi kullanılabilir yapılarak:** Verinin herkes tarafından kullanılabilir, yani analiz edilebilir açık format dosya tipine/yapısına dönüştürülmesinin sağlanması.

*Veriyi bulunabilir yapılarak:* Açılan veri setlerinin bir liste halinde internet üzerinden yayımlanarak bulunması sağlanarak.

### **Açık Devlet Verisi, Açık Devlet, Açık Yönetim Ortaklığı**

**Açık Devlet Verisi:** Devlet ya da devlet kontrolündeki birimlerin ürettiği, herkes tarafından kullanılabilir, tekrar kullanılabilir ve dağıtılabılır olan veridir.

**Açık Devlet:** Verilerin açıkça paylaşılmasının ardından, açık devlet kavramı ortaya çıkmıştır. Halkın, bilgilere erişebilmesinin yanı sıra katılımına ve iş birliğine önem verilmesi de bu sürecin önemli unsurları arasındadır. Açık devletin bileşenleri ise şeffaflık, katılım ve iş birliğidir.

**Açık Yönetim Ortaklığı:** 20 Eylül 2011 tarihinde kurulmuş olan bir girişimdir. Girişimin temel hedefi; saydamlık ve hesap verilebilirliğin artırılması ve yolsuzlukla etkin bir biçimde mücadele edilmesi, vatandaşların ve sivil toplumun kamusal karar alma ve uygulama süreçlerine daha fazla katılımının sağlanması, açık ve etkin bir kamu yönetimi için teknolojik imkânların daha fazla kullanılması ve vatandaşların devlet karşısında daha güçlü konuma getirilmesidir. Açık Yönetim Ortaklığı'na dâhil olmak isteyen ülkelerin, öncelikle açık yönetimle ilgili asgari yeterlilikleri karşılama ve Açık Yönetim Ortaklığı Bildirisi'ni kabul etmeleri gerekmektedir. Ortaklık için ilk adım olan asgari yeterliliklerin karşılanmasıyla kastedilen ise hükümetlerin dört temel alanda (Mali Şeffaflık, Bilgi Edinme Hakkı, Gelir ve Mal Bildirimi, Vatandaş Katılımı) belirlenen kriterlere uyum sağlamasıdır. Türkiye, 20 Eylül 2011'deki resmi açılış töreninde Açık Yönetim Ortaklığı Bildirisi'ni kabul ederek ortaklığın bir parçası olmuştur. Bu süreçte, yani 2011'den günümüze Türkiye, söz verdiği <http://www.saydamlik.gov.tr>, [www.harcama.gov.tr](http://www.harcama.gov.tr) ve [www.duzenleme.gov.tr](http://www.duzenleme.gov.tr) sitelerini eylem plan takvimine göre henüz hayata geçirmemiştir.

22 Eylül 2016 tarihinde Türkiye'nin Açık Yönetim Ortaklığı Statüsü pasif duruma alındı. Yolsuzlukla mücadele ve şeffaf yönetim anlayışının geliştirilmesi kapsamında, Türkiye'nin de üyesi olduğu Açık Yönetim Ortaklığı'nın Yürütme Komitesi, Türkiye'nin 2014'ten bu yana bir eylem planı sunmamasına bağlı olarak, üyeliğinin pasif duruma alındığını açıklamıştı.

## Açık Lisans Nedir?

- Veriler yasal olarak ücretsiz kullanıma izin veriyorsa açıktır .

## Açık Lisanslar...

- Bazı hükümetler kendi lisanslarını yazmayı veya yaygın olarak kabul edilen açık lisanslara benzer kullanım hakları veren 'kullanım şartlarını' formüle etmeyi seçmiştir.
- Bu durum, hükümetlerin kullanıcılar için gereksiz bilgiler eklemesine, uyumsuzluğa neden olmasına ve verilerin tekrar kullanılmasını önemli ölçüde azaltabilir.

## Açık Lisanslar...

---

- Genel lisanslar, kullanım haklarını açıkça bildirerek bu sorunları azaltmak ve verilerin paylaşımını ve yeniden kullanımını mümkün kılmak için tasarlanmıştır.

## Açık/Genel Lisanslar

---

- Telif hakkı lisansları, kanunun oluşturduğu geleneksel "tüm hakları saklıdır" yapısı içinde bir denge oluşturur.
- Bireysel yaratıcılardan büyük şirket ve kurumlara kadar herkese basit, standartlaştırılmış bir yol ile telif izinleri vermesini sağlar.
- Telif hakları çerçevesinde kopyalanabilir, dağıtılabilir, düzenlenebilir, karıştırılabilir ve üzerine inşa edilebilir bir içerik havuzu olabilir.



## CC Lisansları ve Örnekleri

- Attribution CC BY
- Attribution ShareAlike CC BY-SA
- Attribution-NoDerivs CC BY-ND
- Attribution-NonCommercial CC BY-NC
- Attribution-NonCommercial-ShareAlike CC BY-NC-SA
- Attribution-NonCommercial-NoDerivs CC BY-NC-ND



## Attribution CC BY

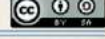


- Bu lisans, orijinal dosyanın oluşturulması için sizi destekleri sürece ticari olarak bile çalışmanızı başkalarının dağıtmasına, yeniden düzenlemesine, değiştirmesine ve geliştirmesine olanak tanır. Bu, sunulan lisansların en uygunudur. Lisanslı materyallerin maksimum dağıtımını ve kullanımını için önerilir.

### Örnekler:

- PLOS
- Saylor.org
- Chris Zabriskie

# Attribution ShareAlike CC BY-SA

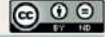


- Bu lisans, başkalarının sizi destekledikleri ve yeni tasarımlarını aynı şartlar altında lisansladıkları sürece, ticari amaçlarla bile çalışmanızı yeniden düzenlemelerine, değiştirmelerine ve geliştirmelerine olanak tanır. Sizininkine dayanan tüm yeni işler aynı lisansı taşıyacaktır, bu nedenle tüm türevler ticari kullanıma da izin verecektir.

## Örnekler:

- Wikipedia ve Wikimedia Commons
- Arduino
- Peer 2 Peer Üniversitesi (P2PU)

# Attribution-NoDerivs CC BY-ND



- Bu lisans, başkalarının işi ticari olarak da dahil olmak üzere herhangi bir amaç için yeniden kullanmalarını sağlar; ancak başkalarıyla uyarlanmış biçimde paylaşamaz ve size bir destek sağlanmalıdır.

## Örnekler:

- Drupal Güvenlik Raporu
- Behance
- GNU / Linux ve Özgür Yazılım Vakfı (FSF)



## Attribution-NonCommercial CC BY-NC



- Bu lisans, başkalarının işinizi ticari olmayan şekilde yeniden düzenlemesine, değiştirmesine ve geliştirmesine olanak tanır. Yeni çalışmalarında size atıfta bulunulmasına, bilgilendirilmenize ve çalışmaların ticari olmamasına rağmen türev çalışmalarını aynı koşullarda lisanslamak zorunda değildir.

### Örnekler:

- Brooklyn Müzesi
- Jonathan Coulton
- Wired.com fotoğrafçılığı

## Attribution-NonCommercial-Share Alike CC BY-NC-SA



- Bu lisans, başkalarının sizi destekledikleri ve yeni tasarımlarını aynı şartlar altında lisansladıkları sürece, ticari olmayan şekilde çalışmalarınızı yeniden düzenlemelerine, değiştirmelerine ve geliştirmelerine olanak tanır.

### Örnekler:

- MIT Open CourseWare
- Cory Doctorow
- Jonathan Worth

# https://data.ibb.gov.tr/license

**Lisans**

**İstanbul Büyükşehir Belediyesi Açık Veri Lisansı**

Bu lisans kapsamında bulunan Bilgileri yalnızca birkaç koşulda, özgürce ve esnek bir şekilde kullanmanız ve yeniden kullanmanız önerilir.

**Bu lisans altındaki Bilgileri kullanma**

Bu lisans kapsamında açıkça sunulan telif hakkı ve veritabanı hakkı materyallerinin ('Bilgi') kullanımı, aşağıdaki şartları ve koşulları kabul, beyan ve taahhüt ettiğinizi gösterir.

Lisans Veren, aşağıdaki koşullara tabi Bilgileri kullanmanız için size dünya çapında, telifsiz, daimi, münhasır olmayan bir lisans sunar.

Bu lisans, adil işlem veya adil kullanım veya başka herhangi bir telif hakkı veya veritabanı hakkı istisnası ve sınırlaması altındaki özgürlüğünüzü etkilemez.

# https://data.ibb.gov.tr/license

**Şunları yapmaktasınız özgürsünüz:**

- Bilgileri kopyalamak, yayınlamak, dağıtmak ve iletmek
- Bilgiyi uyarlamak
- Bilgiyi, ticari olarak ve ticari olmayan şekilde, örneğin diğer Bilgilerle birleştirerek veya kendi ürününüze veya uygulamanıza ekleyerek kullanabilirsiniz.

**Şunları yapmalısınız (yukarıdakilerden herhangi birini yaptığınız yerde):**

- Bilgi Sağlayıcı(lar) tarafından belirtilen herhangi bir atıf bildirimini ekleyerek veya bunlara bağlantı vererek ürününüzdeki veya uygulamanızdaki Bilgilerin kaynağını referans gösterin ve mümkünse bu lisansa bir bağlantı sağlayın.
- Bilgi Sağlayıcı belirli bir ilişkilendirme ifadesi sağlamazsa, aşağıdaki ifadeyi kullanmanız gerekir:

**Atıf 4.0 Uluslararası (CC BY 4.0) kapsamında lisanslanan kamu sektörü bilgilerini içerir.**

Birden fazla Bilgi Sağlayıcıdan gelen Bilgileri kullanıyorsanız ve birden fazla özelliği listelemek ürününüzde veya uygulamanızda pratik değilse, gerekli ilişkilendirme ifadelerini içeren bir kaynağa URI veya köprü ekleyebilirsiniz.

Bunlar, bu lisansın önemli koşullarıdır ve bunlara uymazsanız, bu lisans kapsamında size verilen haklar veya Lisans Veren tarafından verilen benzer bir lisans otomatik olarak sona erecektir.

# İzmir Açık Veri Portalı



İzmir

## Attribution-NonCommercial-NoDerivs CC BY-NC-ND



- Bu lisans, altı ana lisansımızın en kısıtlayıcısıdır ve yalnızca başkalarının çalışmalarını indirmesine ve sizi desteklediği sürece başkalarıyla paylaşmasına izin verir, ancak bunları hiçbir şekilde değiştiremez veya ticari olarak kullanamazlar.

### Örnekler:

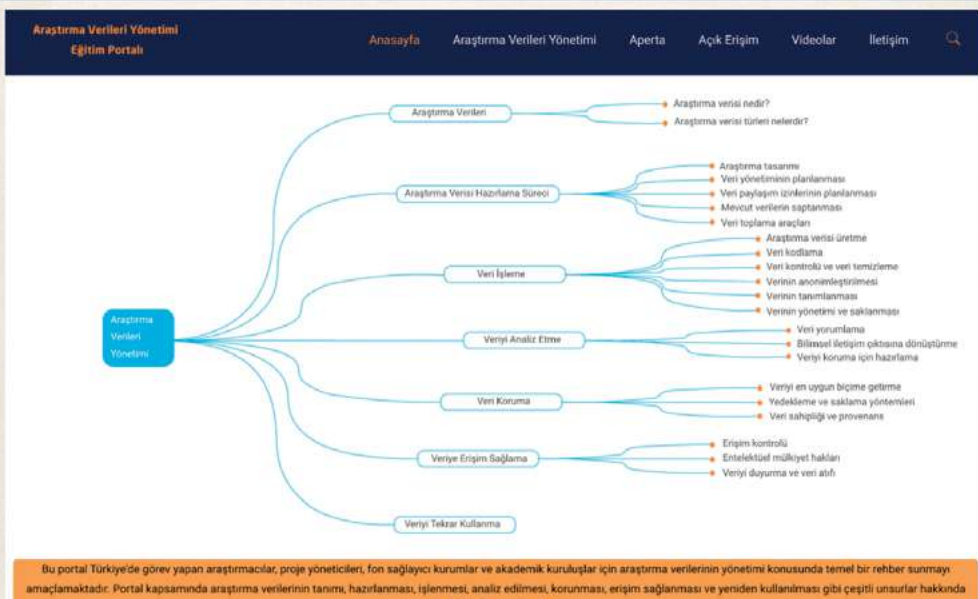
- TPB AFK
- TED Sohbetleri videoları
- Propublica



# Balıkesir Açık Veri Portalı



<https://acikveri.ulakbim.gov.tr/>



Bu portal Türkiye'de görev yapan araştırmacılar, proje yöneticileri, fon sağlayıcı kurumlar ve akademik kuruluşlar için araştırma verilerinin yönetimi konusunda temel bir rehber sunmayı amaçlamaktadır. Portal kapsamında araştırma verilerinin tanımı, hazırlanması, işlenmesi, analizi edilmesi, korunması, erişim sağlanması ve yeniden kullanılması gibi çeşitli unsurlar hakkında

# <https://data.gov.sg/open-data-licence>

data.gov.sg/open-data-licence

## Singapore Open Data Licence

version 1.0

**Jump to**

- Acceptance & Grant of Licence
- What You Can Do
- Additional Conditions
- Limitations
- Definitions and Interpretations

**Acceptance & Grant of Licence**

By Using (as defined below) the datasets on the Relevant Websites, you agree to the terms of this Licence.

The Agency grants you a worldwide, perpetual, royalty-free, non-exclusive licence to Use the datasets, subject to the terms of this Licence.

**What You Can Do:**

- ✓ You can use, access, download, copy, distribute, transmit, modify and adapt the datasets, or any derived analyses or applications, whether commercially or non-commercially ("Use").
- ✓ You may grant a sub-licence of the datasets if this is necessary to enable users of your application and/or website ("your Sub-Licensees"), to use your application or access your website.
- ✓ If your Sub-Licensees require additional rights, your application and/or website should direct your Sub-Licensees to

## Dikkate Alınması Gereken Öneriler...

- Genellikle hükümet verileri telif hakkı koruması altına girmez ve bu şekilde sunulmamalıdır. Hükümetler, fikri mülkiyet (IP) korumasının kapsamı hakkında bilinçli ve açık olmalıdır.
- Standartlaştırılmış açık lisans kullanımı önemlidir ve ilk tercih olmalıdır. Açık tanım, birbiriyle birlikte çalışabilen uyumlu lisanslar sağlar.
- Bazı durumlarda, hükümetler özelleştirilmiş bir açık hükümet lisansı kullanmak isteyebilirler. Bunlar , mümkün olan en az kısıtlamayla ve mümkün olduğunca açık olmalıdır.

## Dikkate Alınması Gereken Öneriler...

- Verilerin yanında açık lisanslama ayrıntıları yayınlanmalı. Lisans verilere açıkça eklenmeli ve hem insan hem de makine tarafından okunabilir olmalıdır.
- Lisans sürümü vurgulanmalı ve verilerin nasıl kullanılacağına dair bağlam olmalı.
- Mümkün olduğunca, standart lisanslarda bulunmayan kısıtlayıcı maddelerden kaçınılmalı

**Kaynak:** [https://docs.google.com/presentation/d/1laWnVcW7tq9MnXXVNCdhd6hc84upIw0\\_VrKzijut-8co/edit#slide=id.p19](https://docs.google.com/presentation/d/1laWnVcW7tq9MnXXVNCdhd6hc84upIw0_VrKzijut-8co/edit#slide=id.p19) / İrem Oran /VOYD



## Dünyadan Açık Veri Örnekleri

### Projeler

Farklı kategorilerde açık veri projelerine rastlanıyor ama genelde amaçlanan etkiler:

1. Yönetişimi Geliştirme
2. Vatandaş Güçlendirme
3. Fırsat ve İş Olanakları Yaratma
4. Sosyal Problemleri Çözme

### İsveç'te Open Aid Projesi

Open Aid projesi devletin şeffaflığını ve hesapverebilirliğini arttırmak amacıyla İsveç'in uluslararası finansal desteklerini & yardımlarını izlemek ve paylaşmak için başlatıldı.



### İsveç'te Open Aid Projesi

- 2012 yılında açılan openaid.se web portalı açık devlet verisini kullanarak İsveç tarafından desteklenen projeleri ve girişimleri gösteriyor.
- Yardımların hangi ülkede ve hangi sektörlerde ve alanlarda yapıldığını ve ne kadar miktarda fon aktarıldığını bu projede görebiliyoruz.  
<https://openaid.se/aid/sweden/all-countries/2019/>
- [Portali inceleyerek](#) İsveç'in Türkiye'ye yaptığı fon yardımları ve fon yardımlarının alanlarını görebiliriz.
- Bu portalda son 20 yılın yardım verileri kolaylıkla indirilebiliyor
- Projede kullanılan portal tam anlamıyla açık kaynak, sivil toplum tarafından farklı temalar ve verilerle kullanılabilir. Projenin API desteği bulunuyor.
- Openaid portalı partner ve komşu ülkeler için örnek oluşturdu. Örneğin Danimarka'da benzer bir proje faaliyete geçti: <https://openaid.um.dk/>



## İngiltere'de Kurumsal Şeffalık İçin Open Corporates

2013 yılında dönemin İngiltere Başbakanı Cameron tarafından onaylanan Şeffaf Sahiplik yasası sonrasında İngiltere'de kurumsal şirket ve holding sahipliklerini daha açık ve şeffaf kılmak için kurulmuştur.

Kar amacı gütmeyen kuruluşların katkılarıyla oluşturulan veri tabanı özellikle finansal sektördeki yolsuzlukları ve göstermelik gayrimenkul ve şirket sahipliklerini ortaya çıkarma amacı taşıyor.

**opencorporates**  
The Open Database Of The Corporate World

Company name or number

Companies Offices Only in Sweden

**ICA AB**

Company Number: 55620-1509  
Company Type: Aktiebolag  
Jurisdiction: Sweden

Source: Bolagsverket (Swedish Companies Registration Office), <http://www.bolagsverket.se/>  
9 Jun 2017

[ADD DATA \(WEBSITE, ADDRESS ETC\)](#)

Company Addresses

COMPANY ADDRESS  
Svetsarvägen 16, SE-171 93 Stockholm, SE

COMPANY ADDRESS  
SE-173 85 SOLNA, SE

COMPANY ADDRESS  
SE-173 85 SOLNA, SE

COMPANY ADDRESS  
Svetsarvägen 16, SE-171 93 Stockholm, SE



## İngiltere'de Kurumsal Şeffalık İçin Open Corporates

- Open Corporates'i kullanarak şirketlerin bilgilerine ve çalışanlarına ulaşılabilir
- Open Corporates'da 105 bölge ve ülkede 85 milyon şirketin verisi açık olarak yer alıyor.
- Kısaca kurumsal şeffaflık için şirket ve sahiplik verilerine dair bir açık veri tabanı (açık veri lisanslı)
- Veri tabanının kapsadığı ülkeler arasında Türkiye bulunmuyor

## Meksika'da Eğitim Verileri Paylaşıldı

Meksika Rekabet Enstitüsü (IMCO) ve tarafından Mejora Tu Escuela (Okulunu Geliştir Projesi) (<http://mejoratuescuola.org>) vatandaşlara okul performansı hakkında bilgi veren çevrimiçi bir platform.

Ebeveynlerin çocukları için en iyi seçeneği seçmelerine yardımcı olur, daha kaliteli eğitim talep etmelerini sağlar ve onlara çocuklarının eğitimine katılmaları için araçlar sağlar.

Ayrıca, Meksika'daki genel eğitim kalitesini yükselten süreçte, okul yöneticilerine, politika yapıcılara ve STK'lara, iyileştirme ve yolsuzlukla mücadele için gereken alanları tespit etmek için veriler sağlar.

## Meksika'da Eğitim Verileri Paylaşıldı

Platformun ailelerin de dahil edildiği 4 adima sahip

Okulunu tanı, Okulunu Karşılaştır, Okuluna Puan Ver ve Okulunu Geliştir.

Bu adımlarla ebeveynlerin okulları daha iyi tanınması ve çocukların eğitimlerinde aktif olarak rol almasının yanı sıra mevcut okulların eğitim kalitesini de vatandaşın katkısıyla arttırmak planlanıyor.

Platform okullara dair veriler barındırıyor, şimdiye 25 veri tabanına sahip



## Meksika'da Eğitim Verilerinin Paylaşıldı

Projenin etkisine gelirse nicel olarak günlük 50 bin ziyaretçi almış, bunu sadece topluluklarla birebir iletişim ve ağızdan ağıza iletişimle sağlamışlar. (Pazarlama, reklam kampanyası kullanılmamış)

Aileler ve Ebeveynler daha eğitim konusunda aktif kılındı ve eğitim olanakları ve koşullarına dair bilgilendirildi

Meksika eğitim sisteminde hesap verebilirliği arttırdı, özellikle yolsuzlukların, eğitim fonlarının gereksiz kullanılması gibi sorunlara ışık tuttu engelledi.

## ABD Hava Durumu Verilerini Paylaştı

ABD Ulusal Okyanus ve Atmosfer İdaresi tarafından paylaşılan veriler ABD'de ekonomik değer ve iş kolları yaratılmasını sağladı.

Toplanan verilerin açılmasıyla hizmet ve teknik alanlardaki iş alanlarına bilgi sağlıyor. Daha iyi hava tahmini ve sıcaklık & yağış tahminleri yapılarak elektrik üretiminden, tarımsal faaliyetlere kadar birçok sektörde katkı sunuluyor



## ABD Hava Durumu Verilerini Paylaştı

### Etkileri

NOAA uydu verilerinden dünya gemi yönlendirmesinin faydalarının yıllık 95 milyon \$ olduğu tahmin edilmektedir.

Kuruluşun bünyesinde NWS Doppler radarlarının kurulması, kasırga ölümlerini ve yaralanmalarını 80'lerin sonlarında ve 90'ların başlarındaki seviyelere göre yüzde 40 azalttı.

ABD tarım endüstrisinde daha iyi ekim zamanlaması sağladığı için yaklaşık 460 milyon bir katma değere karşılık geliyor.

## ABD Hava Durumu Verilerini Paylaştı

### Etkileri

NOAA'nın gerçek zamanlı verileri, yıllık 700milyon \$ 'ın üzerinde katma değere sahip ve özel hava durumu hizmet endüstrisine veri tedarik diyor

ABD'nin 8-10 milyar dolarlık hava durumu endüstrisi NOAA'nın mevsimsel hava durumu verilerine ve kayıtlarına dayanmaktadır.

ABD elektrik jeneratörleri 24 saatlik sıcaklık tahminlerini kullanarak yılda 166 milyon dolar tasarruf sağlıyor.



## Sonuç

- Bir çok alanda açık veri projeleri ve değer üretimi görmek mümkün
- Gerçekleştirilen projeler sadece kamu kuruluşları tarafından tek taraflı bir aksiyon değil aynı zamanda vatandaşları, sivil toplumu, girişimcileri ve iş dünyasını da kapsayan katılımcı bir eylem planına, politikalara dayanıyor.
- Türkiye'de açık veri ekosistemi oluşturmak için bu derece kapsamlı ve katılımcı bir bakış açısına ihtiyacımı var.

## Kaynaklar

- <https://www.opendataimpactmap.org/usecases>
- <http://odimpact.org/index.html#explore>
- <http://opendatatoolkit.worldbank.org/en/essentials.html>
- <https://carto.com/blog/forty-brilliant-open-data-projects-preparing-smart-cities-2018/>

Kaynak:[https://docs.google.com/presentation/d/1-JCqBmysChG-PQDiWWEgy\\_wFG\\_EWVwWkPgn5Wd9Qt08/edit#slide=id.g52cfc7e9dd\\_0\\_47/Sadettin Demirel /Voyd](https://docs.google.com/presentation/d/1-JCqBmysChG-PQDiWWEgy_wFG_EWVwWkPgn5Wd9Qt08/edit#slide=id.g52cfc7e9dd_0_47/Sadettin%20Demirel/Voyd)

## Dünyadan Bazı Veri Kolektifleri

- <https://www.dvcv.com/>
- <https://datacollective.org/>
- <https://itforchange.net/digital-new-deal/2020/10/22/towards-workers-data-collectives/>
- <https://www.crunchbase.com/organization/data-collective>
- <https://collectivedata.com/>
- <https://www.smartdatacollective.com/>
- <https://berringdatacollective.com/>
- <https://amsterdamdatacollective.com/>

## Bazı Uluslararası Veri Kaynakları

Dünya Bankası Veri <http://data.worldbank.org/>

Dünya Geneline Tüm Şirketlerin Bilgilerinin Toplandığı Açık Veritabanı: OpenCorporates <https://opencorporates.com/>

BM Veri <http://data.un.org/>

Google Halka Açık Veri Arama Bölümü: Public Data Explorer <https://www.google.com/publicdata/directory>

Ekonomik Kalkınma ve İşbirliği Örgütü Veri: <https://data.oecd.org>

Önemli bir uluslararası veri kaynağı Data Catalogs: [Data Catalogs](#)

World Health Organization: <https://www.who.int/>

United Nations: <https://www.un.org/en/>

DataCatalogs.org: <https://datacatalogs.org/>

The Guardian's world government data portal: <https://www.theguardian.com/politics/government-data>

Google's public data directory: <https://www.google.com.tr/publicdata/directory>

Voyd veri bültenişi veri seti bölümü: <https://us19.campaign-archive.com/home/?u=b-2c7e2c0521616bfb8413e46&id=12f6ce7cf6>

## Bazı Ulusal Veri Kaynakları

Türkiye İstatistik Kurumu (TÜİK) <http://www.tuik.gov.tr>

Resmi Gazete <http://www.resmigazete.gov.tr/default.aspx#>

Yüksek Seçim Kurulu <http://www.ysk.gov.tr>

Türkiye Odalar ve Borsalar Birliği <http://www.tobb.org.tr>

T. C. Sosyal Güvenlik Kurumu <http://www.sgk.gov.tr>

Türkiye Bankalar Birliği <http://www.tbb.org.tr>

Bankacılık Düzenleme ve Denetleme Kurulu <http://www.bddk.org.tr>

T.C. Başbakanlık Sermaye Piyasası Kurulu <http://www.spk.gov.tr>

T.C. Başbakanlık Hazine Müsteşarlığı <http://www.hazine.gov.tr>

Kalkınma Bakanlığı <http://www.kalkinma.gov.tr>

T.C. Merkez Bankası <http://www.tcmb.gov.tr>

Maliye Bakanlığı <http://www.bumko.gov.tr/>

## Bazı Sivil Veri Kaynakları

Oy ve Ötesi <http://oyveotesi.org/>

İş Cinayetleri <http://isigmeclisi.org/>

Erkek Şiddeti Verisi <https://bianet.org/kadin/bianet/133354-bianet-siddet-taciz-teca-vuz-cetelesi-tutuyor>

Covid19 sebebi ile ölen sağlık çalışanları <https://siyahkurdelve.com/>



## 10.BÖLÜM

### Kaynak

#### Temel Veri Eğitimi Çevrim İçi Kaynaklar

- <http://verigazeteciligi.com/>
- <https://veribulteni.voyd.org.tr/>
- <https://veribulteni.voyd.org.tr/e-book/>
- <https://veribulteni.voyd.org.tr/temel-veri-gazeteciligi-egitim-seti/>
- <https://handsondataviz.org/sketch.html>
- [https://ocha-dap.github.io/quick-tips-for-visualising-data/#/relationship\\_types](https://ocha-dap.github.io/quick-tips-for-visualising-data/#/relationship_types)
- <https://sites.google.com/view/verigazeteciligi/vg/ders-mod%C3%BClleri/iii-mod%C3%BCl>
- <https://gazetecilikokulu.net/index.php/egitimler/>
- <https://p24akademi.org/kurslar/>

#### Temel Veri Eğitimi Basılı Kaynaklar

- <https://www.datapine.com/blog/best-data-visualization-books/>
- [http://www.verigazeteciligi.com/wp-content/uploads/2020/12/2020-12-27\\_20-46-14.png](http://www.verigazeteciligi.com/wp-content/uploads/2020/12/2020-12-27_20-46-14.png)
- <https://gazetecilikokulu.net/index.php/2021/01/16/sivil-toplum-orgutleri-icin-hak-temelli-gazetecilik-kilavuzu/>

### Veri Eğitim Becerileri Listesi

#### Google Tablo Becerileri:

- Nasıl veri alınır, farklı dosya biçimleri arasındaki fark, Excel ve Google E-Tablolar ne zaman kullanılır? /Temel Konular:
  - Araştırma Searching
  - Sıralama/Sorting
  - Filtreleme/Filtering
  - Dondurma/Freezing panes
  - Ortalama, medyan, toplam, maksimum, minimum ve yüzde değişimi gibi temel matematik
    - Formülleri \$ ile sabitleme
    - Hücre ve Sütun Biçimlendirme
    - Koşullu Biçimlendirme
- Özet Tablolar
- Veri Temizleme
  - Sütun Ekle/Sil
  - Özel yapıştır

- Yerini değiştirmek (devrik)
- Bul ve değiştir (joker karakter)
- Temizle ve kırp
- Sütunları ayırma ve sütun içindekileri sağ/sol/orta hizalama
- Tarih ve Zaman Manipülasyonu
- Vlookup(Düşeyara) ile verileri birleştirme ya da temizleme
- Eğer/o zaman ile veri temizleme ya da kategorileştirme

### **Diğer Beceriler:**

- Yaptığınız araştırmaları endekslemek/ kendi araştırmanızı sağlayan veri biyografisi oluşturmak (kayıtlarınız, çalışmalarınız vb.)
- Soru sormak için kiminle iletişim kurulmalıdır ya da analizde yardım almak için kiminle görüşmeli /ların muhatabı kim.
- Veriyi PDF'den çekmeyi /çıkarmayı bilmek (Tabula ya da Cometdocs )  
Getting data out of a PDF with Cometdocs or Tabula
- Organize olmak için e-tablo kullanmak

### **Veri setleri:**

- Yeni bir veri seti açtığınızda kendinize sormanız gereken sorular
- Her muhabirin Nüfus Sayımı gibi bilmesi gereken temel veri setleri ve bunların nasıl kullanıldığı.
- Veri sözlüğü nedir ve nasıl kullanılır
- Fibs gibi ortak kodlama
- Veri kazıma nedir /kiminle bu konuda konuşulur, R, Python programları ile neler yapabilirim?
- Kendi veri setinizi oluştururken en iyi araçlar/yöntemler

### **Haberlerde Veri Kullanma:**

- 7 çeşit veri hikayesi
- Haber keşfi için veri kullanımı
- Varsayımlarınızı sayılarla kontrol etme - hikayenizdeki sayılara ne dönüştürülebilir?
- Veriler alternatif hikaye formlarında nasıl kullanılabilir?
- Grafik editörleri ile nasıl iş birliği yapılır?
- Etik ve kötü veri ile başa çıkma
  - Korelasyon nedenselliğe eşit değildir
  - Kiraz toplama
  - Bağlamın önemi
  - Hata payı, düzeltilmiş ve düzeltilmemiş

## Kapanış

**STÖ'ler için Temel Veri Eğitimi Kılavuzu** veriyle çalışma yöntemlerini daha iyi anlamak ve ilerletmek için hazırlandı. Günümüz dünyası için temel seviyede veriyle çalışma becerisi artık bir ihtiyaç. Bu kılavuz tam da bu sebeple; veriyle çalışma yöntemlerine, araçlarına ve örneklerine yönelik temel seviyede bir farkındalığı yaratmayı amaçlamıştır. Ancak unutulmamalıdır ki, STÖ'ler ürettikleri verileri daha iyi anlamlandırmak ve etkili karar almak için düzenli olarak uygulama yapmalıdırlar.



SİVİL TOPLUM  
ÖRGÜTLERİ İÇİN  
**TEMEL VERİ  
EĞİTİMİ**  
KILAVUZU



**Eşit Haklar İçin İzleme Derneği**

Gümüşsuyu Mah. Ağa Çırağı Sok.  
No:7 Pamir Apt. D:1  
Beyoğlu/İstanbul  
Tel: 0212 293 63 77  
[www.esithaklar.org](http://www.esithaklar.org)  
email: [esithaklar@gmail.com](mailto:esithaklar@gmail.com)  
[info@esithaklar.org](mailto:info@esithaklar.org)