



# Büyük Verinin İnteraktif Görselleştirilmesi: Tableau Üzerine Öğrenci Deneyimleri

Süleyman Eken<sup>1\*</sup>

<sup>1</sup> Kocaeli Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği Bölümü, Kocaeli, Türkiye (ORCID: 0000-0001-9488-908X)

(İlk Geliş Tarihi 16 Aralık 2019 ve Kabul Tarihi 7 Şubat 2020)

(DOI: 10.31590/ejosat.659823)

**ATIF/REFERENCE:** Eken, S. (2020). Büyük Verinin İnteraktif Görselleştirilmesi: Tableau Üzerine Öğrenci Deneyimleri. *Avrupa Bilim ve Teknoloji Dergisi*, (18), 262-271.

## Öz

Büyük veri görselleştirme, büyük veri kümelerinin önemini daha iyi ifade etmek ve veriler içindeki gizli desenleri ortaya çıkarmak için görsel öğelerin kullanımını tanımlar. Veri görselleştirme; grafikler, çizgeler, haritalar, tablolar veya çeşitli öğeler şeklinde olabilir. İnteraktif görselleştirme ise birden çok görsel arasında bağlantı kurmak ve bir görsel üzerinde direkt işlem yapılmasını sağlar. Veri görselleştirme; karar vericilerin özellikle istatistiksel analiz veya bilgisayar biliminde geçmişi olmayanların, analitik verileri hızlı ve etkili bir şekilde kavramasını mümkün kılar. Bu çalışmada 2018-2019 Bahar döneminde Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nde verilen "Büyük Veri Analizine Giriş" dersi kapsamında yapılan Tableau kullanarak büyük verilerin interaktif olarak görselleştirilmesi durum çalışmasından bahsedilmiştir. Farklı veri kümeleri üzerinde değerli bilgilerin çıkarılması, görsellerin birbirleriyle ilişkilendirilmesi, grafikler üzerinde gruplama, filtreleme ve aksiyon gerçekleştirme, pano üzerinde birden fazla görsel yerleştirme ve etkileşimli özgeçmiş hazırlama gibi tecrübeler elde edilmiştir. Ders sonrası yapılan tartışmaya göre bazı öğrencilerin iş zekâsı alanında çalışma yönelimleri olduğu görülmüştür.

**Anahtar Kelimeler:** Görsel analitik, interaktif görselleştirme, büyük veri analitiği, bilgi görselleştirme, interaktif özgeçmiş, coğrafi görselleştirme, metin görselleştirme, keşifsel veri analizi.

## Interactive Visualization of Big Data: Student Experiences on Tableau

### Abstract

Data visualization describes the use of visual elements to better express the significance of massive datasets and to uncover hidden data patterns. Data visualization can take the form of charts, graphs, maps, tables, or different elements. However, interactive data visualization enables direct actions on a plot to change elements and link between multiple plots. Data visualization makes it possible for decision makers, particularly those without a background in statistical analysis or computer science, to quickly and effectively comprehend analytical data. In this study, a case study on interactive visualization of big data by using Tableau within the scope of "Introduction to Big Data Analysis" course given in Computer Engineering Department at Kocaeli University in 2018-2019 Spring semester was discussed. Experiences such as extracting valuable information on different datasets, linking plots with each other, grouping, filtering and action implementation on graphics, multiple visual placement on the dashboard and preparing interactive resumes have been gained. According to the discussion after the lesson, it was seen that some students had the desire to work in the field of business intelligence.

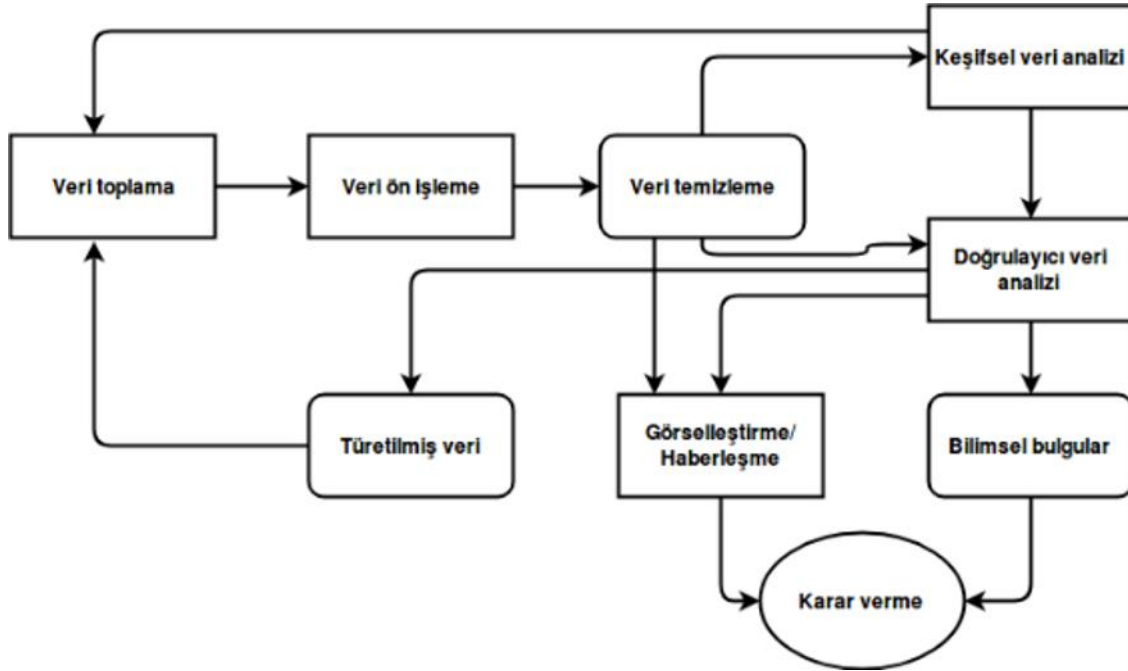
**Keywords:** Visual analytics, interactive visualization, big data analytics, information visualization, interactive resume, geovisualization, text visualization, exploratory data analysis.

\* Sorumlu Yazar: Kocaeli Üniversitesi, Teknoloji Fakültesi, Bilişim Sistemleri Mühendisliği Bölümü, Kocaeli, Türkiye, ORCID: 0000-0001-9488-908X, [suleyman.eken@kocaeli.edu.tr](mailto:suleyman.eken@kocaeli.edu.tr)

## 1. Giriş

Görselleştirme, klasik olarak verileri grafiksel olarak gösterme süreci olarak tanımlanır. Aynı verileri temsil etmenin çok farklı yolları vardır: dağılım (scatter) grafikleri, doğrusal (linear) grafikler, çubuk (bar) grafikler ve pasta (pie) grafikler. Ayrıca, aynı temsil tipini kullanan aynı veriler, şekle bakımlara bağlı olarak çok farklı algılanabilir. Görselleştirme asıl olarak insanlar ve veriler arasında grafiksel bir arayüz sağlamaktadır. Görselleştirme son yıllarda verileri güzelleştirme sanatı olarak algılanmaktadır. Veri görselleştirme uzmanları ve sanatçılar her gün veri tasarımı dünyasında şaşırtıcı şeyler ortaya koymaktadırlar. Veri görselleştirme yeni bir çalışma alanı da değildir ve birçok farklı alandaki verileri görselleştirmek için önemli bir araçtır (Friendly, 2008; Eken ve Kumru, 2014; Ulvi vd., 2013; Eken vd. 2012). Teknik olarak veri görselleştirme; veri sanatı, infografik ve veri panolarını (data dashboards) kapsar.

Şekil 1’de de görüldüğü üzere veri görselleştirme, sonuçların gösterilmesi açısından en efektif bir yoldur (Tuft, 2001). Son yıllarda araştırmacılar veri görselleştirme adımının son ürüne uygulanmasından ziyade veri bilimi sürecindeki her bir adıma uygulanabileceğini söylemişlerdir. Her bir veri bilimi süreci adımı ile ilgili görsel çıkarım ve yorumlama yapmak mümkün olup görselleştirme sonrası adımlar için karar alma ve kritik işlerin yapılması mümkün olmaktadır (Fox ve Hendler, 2011). Dağılım grafikleri, kutu grafikleri, histogramlar, gövde ve yaprak grafikleri ve daha fazlası gibi birçok görselleştirme tekniği geliştirilmiştir. Disiplinlerarası çalışmalarda keşifsel veri analizi (Exploratory Data Analysis, EDA) için başka görselleştirme teknikleri de uygulanabilir. Zihin haritaları (Buzan ve Griffiths, 2013), kavramsal haritalar (Novak ve Cañas, 2006) ve iş akışı görselleştirmeleri (Mou vd., 2017) bunlardan birkaçıdır. Bu görselleştirme tekniklerinin kullanılmasıyla farklı disiplinlerden gelen araştırmacılar; incelenen konuyu hızlı bir şekilde öğrenebilir, veri kümeleri hakkında fikir edinebilir ve incelenen konunun odaklandığı hipotezlerini tartışabilir, veri kümeleri hakkında öngörü kazanabilir ve hipotezleri tartışabilir (Tansley ve Tolle, 2009; Ma, 2017).



Şekil 1. Veri bilimi sürecindeki temel adımlar

Ayrıca veri görselleştirmede kullanıcının görselle etkileşime geçip geçmemesine göre statik ve interaktif/etkileşimli görselleştirmeden bahsetmek mümkündür. Statik görselleştirmeler genellikle internette yayınlanan veya bildiriler olarak basılan infografikler olarak görülür. Genellikle belirli bir veri hikayesine odaklanan kullanıcılar, önlerinde bulunanların ötesindeki ek hikayeleri keşfetmek için tek bir görünümün ötesine geçemezler. Hikaye, özellikle ilgi çekici bir tek sayfa düzeninde yakalanır. İnteraktif görselleştirmeler internette yaygın olarak yalnızca uygulamalar olarak görülür. Kullanıcılar seçtikleri görselleştirilmiş bir hikayeyi oluşturmak için belirli veri noktalarını seçebilirler. Veri görselleştirme uygulamaları olarak daha iyi bilinen bu görselleştirmeler, kullanıcının seçtikleri bir hikaye oluşturarak veri görselleştirme sürecinin bir parçası olmasına izin verir (Herman vd., 2018). Hedef müşteri, veri hikayesi ve yatırım getirisi/karlılığı gibi parametreler hangisinin kullanılacağını bulmaya yardımcı olabilecek önemli göstergelerdir. Somut olarak etkileşimli görselleştirme sistemleri, kullanıcıların aşağıdaki işlevler aracılığıyla bilgilerle etkileşime girmelerini sağlar (Zudilova-Seinstra vd., 2009; Murray, 2017):

- ◆ Şekil, renk, sınırlar ve örnekleme hızı gibi görselleştirme parametrelerini/niteliklerini iyileştirme (refine),
- ◆ Seçim, çeviri, döndürme, ölçeklendirme, toplama ve bir zaman yolunu takip etme gibi temel manipülasyon görevleri vasıtasıyla görselleştirilmiş nesnelere etkileşim kurma (interact),
- ◆ Kırpma ve ayırıştırma yoluyla veri kümesi içinden istenilen bilgi parçalarını çıkararak filtreleme (filter),
- ◆ Aynı veri kümesinin birden fazla alternatifi ve/veya temsili arasında dönüşüm yapma (transform).

İnteraktif görselleştirmenin mevcut uygulamaları, kullanıcıların temel grafik parametrelerini daha karmaşık çoklu-algılamalı görselleştirme sistemlerine yönlendirmelerini sağlayan basit araçlardan oluşmaktadır. Temel etkileşim, kullanıcıların seçim, ölçeklendirme, toplama, filtreleme vb. temel komutlarla büyük veri kümeleri ile etkileşime girdiği yayıncılık ve eğitim alanlarında yaygın olarak kullanılmaktadır (Washington Post, 2019; Gapminder, 2019). Daha gelişmiş İnteraktif görselleştirme uygulamaları, insan duyu sisteminin tam spektrumundan faydalanmasına izin verir ve sanal gerçeklik (VR) (Tamayo vd., 2018), çok biçimli (multimodal) görselleştirme sistemleri (MSS) (Srinivasan vd., 2018) ve artırılmış gerçeklik (AR) içerir (Büschel vd., 2018).

Makalenin geri kalan kısmı şu şekilde organize edilmiştir: 2. bölümde açık kaynak ve ticari interaktif büyük veri görselleştirme araçlarından bahsedilmiştir. 3. bölümde 2018-2019 Bahar döneminde Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nde verilen "Büyük Veri Analizine Giriş" dersi (Ders kaynakları, 2019) kapsamında yapılan Tableau kullanarak büyük verilerin interaktif olarak görselleştirilmesi durum çalışmalarından bahsedilmiştir. 4. bölümde ise sonuçlar tartışılmıştır.

## 2. Büyük Veri Görselleştirme Araçları

Büyük veri analitiği ve görselleştirme araçlarına girişimci ve araştırmacıların göstermiş olduğu büyük ilgi neticesinde bu bölümde, büyük veri kümesini görselleştirmek için kullanılacak doğru aracı bulmaya yönelik bir takım araçlardan bahsedilmiştir (Caldarola vd., 2015). Büyük veriyi görselleştirme için kullanılan araçlar şu ana başlıklar altında kategorize edilebilir: kapsam, görselleştirme yapısı, işletim sistemi, lisans, ölçeklenebilirlik, genişleyebilirlik (Caldarola ve Rinaldi, 2017).

Kapsam kategorisinde aracın daha çok tek başına masaüstü sunum veya ağ tabanlı uygulamalar için mi yoksa yazılım kütüphaneleri veya programlama dili modülleri için mi geliştirildiğine bakılır. Görselleştirme yapısında, aracın ne tür grafiksel nesnelere ve görsel bileşenler için uygun olduğu değerlendirilir. Bazı araçlar sadece bir grafiksel nesne için tasarlanmışken bazıları birden fazla görsel türünü destekleyecek şekilde gerçekleştirilmiştir. Aracın üzerinde çalıştığı işletim sistemi (örneğin; Linux, Windows, Mac OS X) de yine ele alınan konulardandır. Lisans kriterinde aracın ticari veya açık kaynak lisansları (Apache License, GNU GPL, vb.) ile lisanslanmasına bakılır. Ölçeklenebilirlik kriteri, çok büyük veri kümesini desteklemek için araçların yatay olarak ölçeklendirme (sisteme yeni düğüm/bilgisayar ekleme veya çıkarma) mekanizmaları hakkındadır. Bazı araçlar Apache Hadoop gibi çok büyük veri kümelerini işleme kapasitesine sahip uygulama çerçevesine bağlanabilme veya bulut sistemi kullanabilme kapasitesine sahiptir. Genişleyebilirlik altında araçların yeni eklentilerle genişletilebilmesi veya var olan depolama birimlerine bağlanabilmesi incelenir. Örneğin bazı araçlar spesifik yollarla Matlab ve R'a bağlanabilir. Görselleştirme araçları; grafik türlerine, ücretli veya ticari olmalarına ve programlama gerektirip gerektirmemesine göre Tablo 1'de özetlenmiştir.

2018-2019 Bahar döneminde Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nde verilen "Büyük Veri Analizine Giriş" dersi kapsamında, programlama gerektirmeyen ticari araçlardan Tableau ve geliştiriciler için tasarlanmış Seaborn ve matplotlib görselleştirme deneyimi için gösterilmiştir. Bu çalışmada ise interaktif görselleştirme sunan Tableau üzerinde durulmuştur. Tableau, hem ücretsiz sürümü hem de zengin ve interaktif veri görselleştirme olanakları ile veri görselleştirmeye ilgilenenler için vazgeçilmez araçlardan birisi. Bir business intelligence (iş zekâsı) yazılımı olarak Tableau, özel şirketlerden kamu kuruluşlarına bir çok alanda veri analizi ve veri görselleştirme amaçlı kullanılıyor.

Gartner analizleri, bilgi teknoloji ürünleri üzerinde araştırma ve tavsiyelerde bulunmaktadır. Büyük çaplı şirketlerin bilgi teknoloji yöneticileri, kamu kuruluşları, teknoloji şirketleri tarafından takip edilen bir değerlendirme kuruluşudur. Teknoloji ürünü geliştirilmesi, alımı ve kullanımı sürecinde karar verme aşamasında başvurulabilecek önemli bir otoritedir (Gartner, 2019). Analitik ve iş zekâsı platformları üzerine 2018'de verilen Gartner raporuna göre Tableau Desktop, Tableau Server ve Tableau Online gibi ürünleri ilk sıralarda yer almaktadır (Gartner BI Report, 2019). Ders içinde interaktif görselleştirme amaçlı Tableau seçilmesinin motivasyonu da budur. Özellikle kullanım kolaylığı ile öne çıkan, bir çok ihtiyaç için halihazırda modülleri bulunan ve ek ihtiyaçlar için hızlıca çözümler geliştiren oldukça genç markalardan biridir. Tableau, sunmuş olduğu kullanım kolaylığı ve öğrenim eğrisinin düşük olması bakımından birçok farklı çevrenin sıklıkla kullanılmakta olduğu bir program haline gelmiştir. Tableau ile sunumlarda ve hazırlanan raporlarda farklı görselleştirme yöntemleriyle panolar aracılığıyla fark oluşturulabilir çalışmalar yapılabilir. Tableau ile verileri keşfetmek (explore), analiz etmek (analyze), sunmak (present) ve paylaşmak (share) mümkündür.

Ders için "Tableau for Teaching" programı kapsamında ders süresi boyunca yazılım lisansları alınmıştır. Tableau Desktop 10 versiyonu kullanılmıştır. Ders öncesinde öğrencilere uygulamanın web sitesinde yer alan "Classroom Training" bölümlerini gözden geçirmeleri tavsiye edilmiştir (Tableau Classroom Training, 2019). Derste ise Tableau ile ilgili olarak şu konular anlatılmıştır: Başlangıç ekranı ve temel İşlemler, ayırık ve sürekli kavramları, grafik çeşitleri, veriyi düzenleme, yükleme ve birleştirme, analitik panel, hesaplamalar (calculations), parametreler, filtreleme, pano, panolar üzerinde etkileşim sağlama (actions), hikaye oluşturma, farklı dosya türlerine bağlanma, veri hazırlama, çalışmaları mobil uyumlu hale getirme ve nete yayınlama. Takip eden bölümde öğrencilerin yapmış oldukları çalışmalardan örnek durum çalışmaları verilmiştir.

Tablo 1. Görselleştirme araçları

Programlama gerektirmeyen araçlar		
Çizim ve grafikler	Ücretsiz	RAWGraphs, ChartBlocks
	Ticari - Bireyler veya iş için	Tableau, Power BI, QlikView, Datawrapper, Visme
	Ticari - İşletmeler için	Grow, iCharts
	İnfografik	Infogram, Visual.ly
	Haritalar	InstantAtlas
	Ağ grafikleri	Gephi
	Matematiksel grafik	Wolfram Alpha
Geliştiriciler için olan araçlar		
Çizim ve grafikler	Ücretsiz	ECharts, D3.js, Plot.ly, Chart.js, Google Charts, Ember Charts, Chartist.js, Seaborn, matplotlib
	Ticari	Highcharts, FusionCharts, ZingChart
	Haritalar	Leaflet, OpenLayers, Kartograph, CARTO
	Ağ grafikleri	Sigma
	Finansal grafikler	dygraphs

### 3. Örnek Durum Çalışmaları

Öğrencilerin herkese açık veri kümelerine (public dataset) erişimlerini sağlama ve bu veri kümelerinden bir/birkaçı üzerinde Tableau kullanarak yukarıda anlatılan konuların pekiştirilmesi ve interaktif görselleştirmenin daha iyi öğrenilmesi için bir geri bildirim aktivitesi verilerek kendilerini değerlendirmeleri sağlanmıştır. Ayrıca iş bulma sürecinde çok büyük bir etkiye sahip olan özgeçmiş dokümanının interaktif olarak oluşturulması ile ilgili bir deneyim de ders kapsamında sağlanmıştır.

#### 3.1. Herkese Açık Veri Kümesi Bulma ve İnteraktif Görselleştirme

Genelde lisans öğrencilerinin bir probleme yönelik herkese açık bir veri kümesi bulmaları çok zor olmakta ve buldukları veri kümesinin ise ne kadar işe yaracağını kestirememeleri bir problem olarak karşımıza çıkmaktadır. Yüksek kaliteli veri kümeleri ararken dikkat edilmesi gereken birkaç husus şu şekilde özetlenebilir:

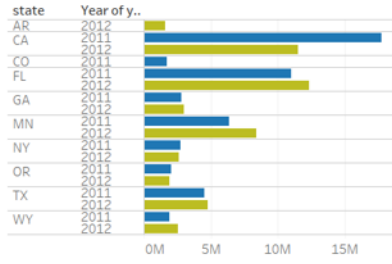
- ◆ Yüksek kalitede bir veri kümesi dağınık (messy) olmamalıdır; çünkü temizlik için çok fazla zaman harcamak gerekebilir.
- ◆ Yüksek kaliteli bir veri kümesinde çok fazla satır veya sütun olmamalıdır. Bu nedenle çalışması kolaydır.
- ◆ Veriler ne kadar temiz olursa büyük bir veri kümesinin temizlenmesi çok fazla zaman alabilir.
- ◆ Cevap aranan probleme çözümün elde edilecek veri kümesi ile olacağı unutulmamalıdır.

Ders kapsamında öğrencilere veri kümesi bulabilecekleri farklı web siteleri tanıtılmıştır: Google Dataset Search, Kaggle, UCI Machine Learning Repository, Data.gov, World Bank Open Data (The 50 best public dataset, 2019). İlgili kaynaklardan interaktif görselleştirme için en az bir tane veri kümesinin seçilmesi sağlanmıştır. İlgili veri kümeleri için Tableau üzerinde şu şekilde bir pano hazırlanması istenmiştir:

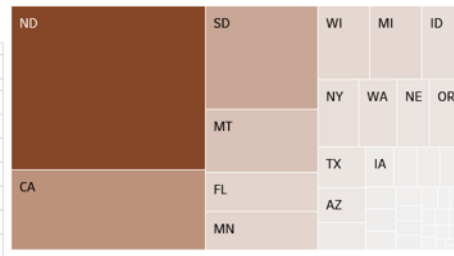
- a. En az dört farklı görselleştirme türü (örneğin, çubuk grafik, dağılım-scatterplot-grafiği, harita vb.) içersin.
- b. Her görselleştirme farklı ilişkileri göstermelidir.
- c. Her görselin mantıklı, değerli bilgileri içerecek şekilde olmasına dikkat ediniz.
- d. Grafiklerde grupta, filtreleme, aksiyonlar, hesaplamalar (veya daha fazlası) gibi bir takım konuların içermesini sağlayınız.

ABD’de bal üretimi ile ilgili bir veri kümesi (Kaggle, 2019) üzerinde yapılan interaktif görselleştirme sonucu Şekil 2’de verilmiştir. 1985-2016 yılları arasında intihar oranları üzerinde (Veri kümesi 1, 2019) yapılan interaktif görselleştirme sonucu Şekil 3’te verilmiştir. IMDB 5000 film veri kümesi üzerinde (Veri kümesi 2, 2019) yapılan interaktif görselleştirme sonucu Şekil 4’te verilmiştir.

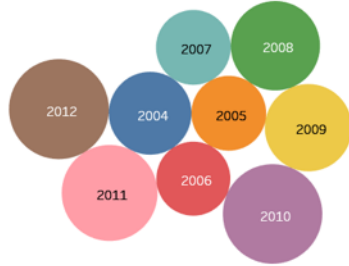
2011 ve 2012 yıllarında Amerikada 1.500.000 lbs üzerinde bal üretimi yapan 10 farklı eyalet



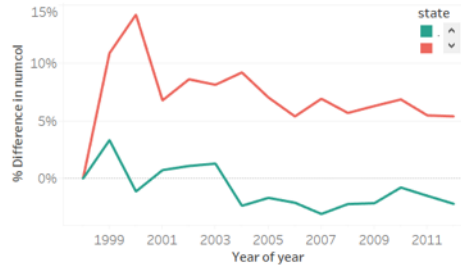
2010 yılında tüm eyaletlerdeki toplam stok miktarı



2004-2012 yıllarında ürünlerin tüm eyaletlerde ortalama ne kadara satıldığı



North Dakota ve California eyaletlerindeki toplam koloni sayısının yıllara göre karşılaştırması

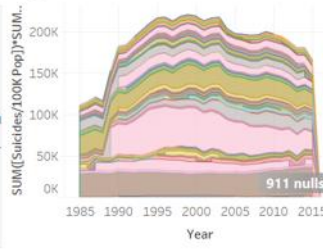


Şekil 2. ABD'de bal üretimi veri kümesi üzerinde interaktif görselleştirme

İntihar sayısı 80'den fazla olan ülkeler



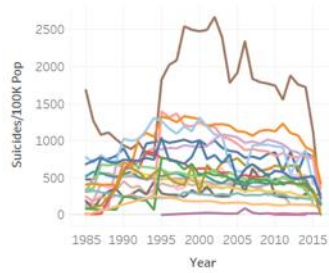
Yıllara göre ülkelerdeki intihar sayılarının ülke nüfusuna oranı



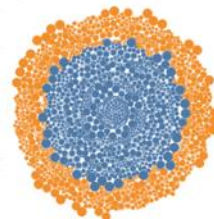
Count of Rows  
80 126

- Country (group)
- Albania, Antigu...
  - Bahamas, Bahra...
  - Cabo Verde, Can...
  - Denmark & Domi...
  - Ecuador, El Salv...
  - Fiji, Finland, Fra...
  - Georgia, Germa...
  - Hungary
  - Iceland, Ireland, ...
  - Jamaica & Japan
  - Kazakhstan, Kiri...
  - Latvia, Lithuani...
  - Macau, Maldive...
  - Netherlands, Ne...
  - Oman
  - Panama, Paragu...
  - Qatar
  - Republic of Kore...
  - Saint Kitts and ...
  - Thailand, Trinid...
  - Ukraine, United ...

Yıllara göre intihar oranları



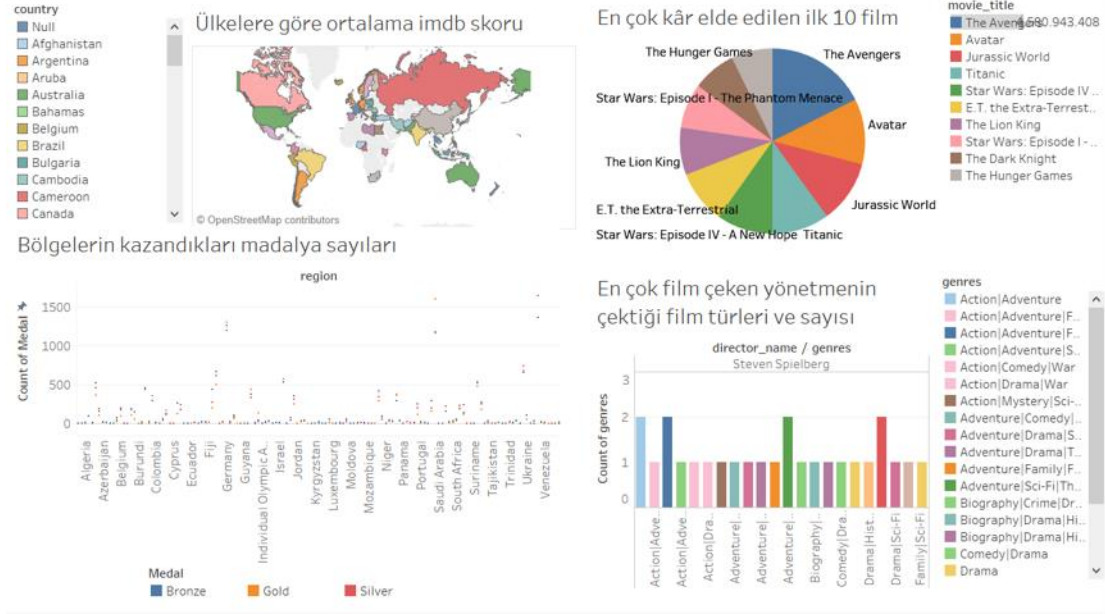
İntihar edenlerin cinsiyetlerine göre ayrımı



Highlight Country  
Highlight Country

Sex  
female  
male

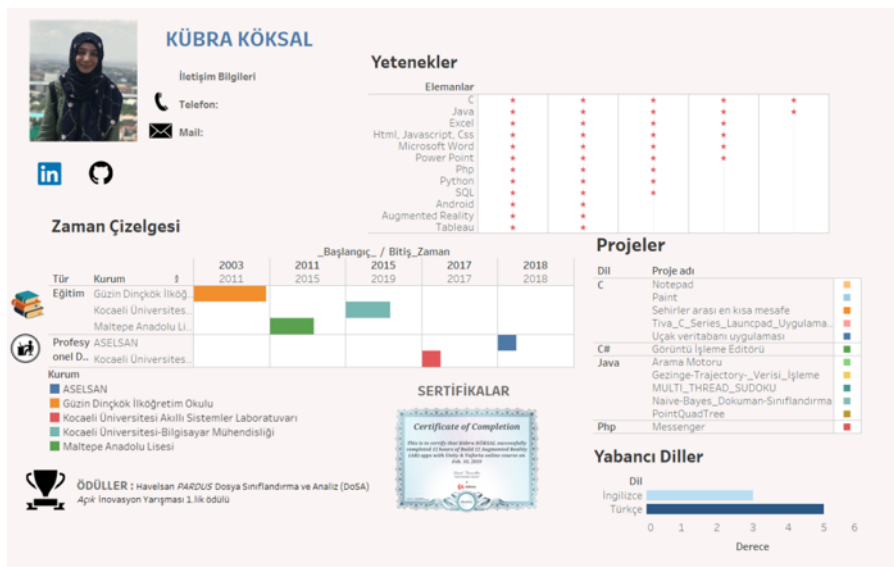
Şekil 3. İntihar oranları veri kümesi üzerinde interaktif görselleştirme

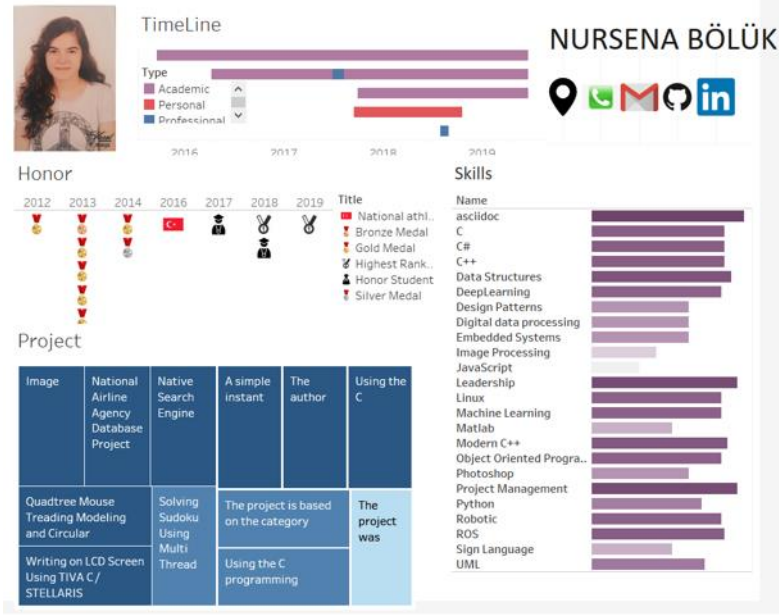


Şekil 4. IMDB 5000 veri kümesi üzerinde interaktif görselleştirme

### 3.2. İnteraktif Özgeçmiş Hazırlama

İşverenler her iş başvurusunu okumak için sadece altı saniye harcarlar. Bu nedenle özgeçmişin kalabalığın arasından sıyrılması önemlidir. Veri okuryazarlığının işyerinde en çok talep edilen becerilerden biri olduğu bir dünyada özgeçmişler bilgiyi net ve kullanışlı bir şekilde aktarma yeteneğini yansıtmalıdır. İnteraktif bir özgeçmişin potansiyel bir işveren için standart PDF veya Word belgesinden çok daha etkileyicidir. Eğitim veya tecrübeleri gösteren kronolojik bilgileri listelerden farklı olarak Gantt grafikleri ile zaman çizelgesi şeklinde vermek en önemli öğelere dikkat çekilmesini sağlamaktadır. İnceleyen kimse ilk bakışta kariyerinizin şeklini görebilir ve ardından detayları okumak için her bir çubuğun üzerinde gezinebilir. İnteraktif bir özgeçmiş ile daha sade bir tasarım ve detay arasında seçim yapmak zorunda kalınmamış olur. Tasarımı basit tutarak ve özgeçmişini inceleyen kimsenin istediğiniz gibi ayrıntılara dalmasına izin vermek için ipuçlarını kullanabilirsiniz. Ders kapsamında hazırlanan interaktif özgeçmişlerden bazıları Şekil 5'te verilmiştir. Tableau ile hazırlanmış diğer interaktif özgeçmişlere şuradan göz atabilirsiniz (Veri kümesi 3, 2019).



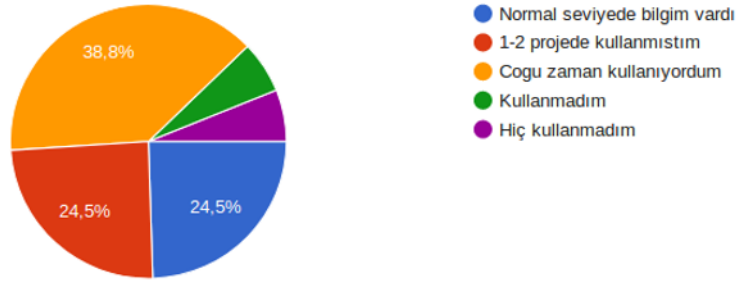


Şekil 5. (a) İnteraktif özgeçmiş örneği 1 (b) İnteraktif özgeçmiş örneği 2

### 3.3. Öğrenci Görüşleri

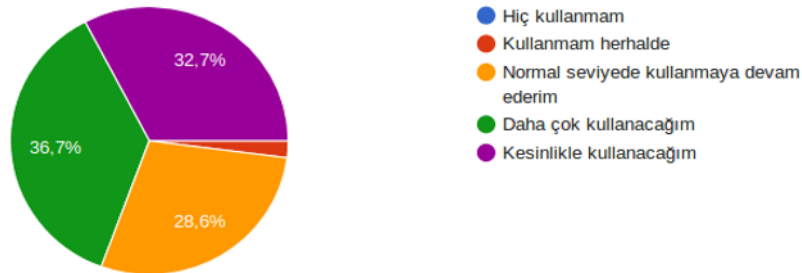
49 öğrenci ile ders sonrasında yapılan anket ve oranlar pasta grafik olarak aşağıda sunulmuştur.

Açık kaynak araçları/teknolojileri ders öncesindeki kullanma sıklığınız nedir?



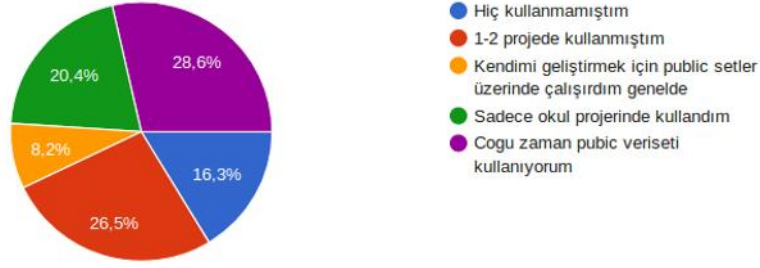
Şekil 6. Ders öncesi açık kaynak araç kullanım dağılımı

Açık kaynak araçları/teknolojileri ders sonrası kullanma sıklığınız ne olacak?



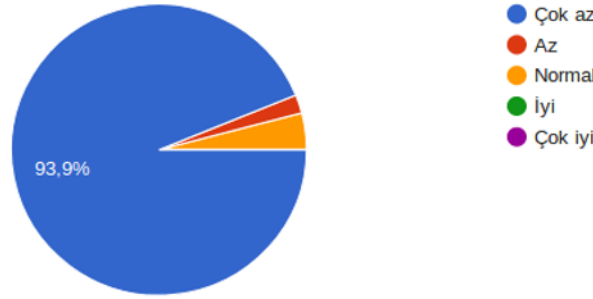
Şekil 7. Ders sonrası açık kaynak araç kullanım dağılımı

Ders öncesi herkese açık veri kümesi bulma ve kullanma seviyeniz ne idi?



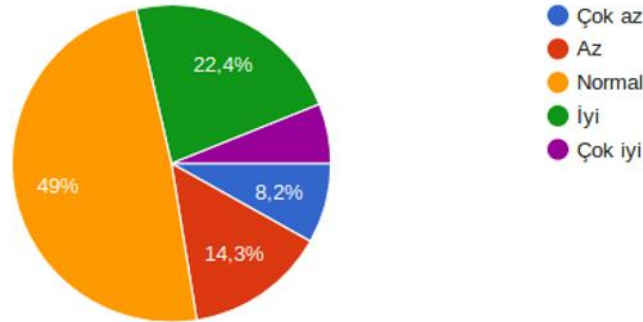
Şekil 8. Ders öncesi herkes açık veri kümesi üzerinde çalışma dağılımı

Ders öncesi Tableau kullanarak veri analizi ve görselleştirme yetkinliğiniz ne idi?



Şekil 9. Ders öncesi Tableau kullanma dağılımı

Ders sonrası Tableau kullanarak veri analizi ve görselleştirme yetkinliğiniz nedir?



Şekil 10. Ders sonrası Tableau kullanma dağılımı

Anket sonuçlarında da görüldüğü üzere öğrencilerin açık kaynak teknolojilerini kullanabilme, herkese açık veri kümesi bulabilme ve Tableau kullanarak veri analizi ve görselleştirme yetkinlikleri çok artmıştır. Özellikle görselleştirme ile ilgili yapılan öğrenci yorumlarından birkaçı aşağıdaki gibidir:

“Hocam ders kapsamı bakımından büyük veri ile ilgili birçok ihtiyacı karşılamaktadır. Ve her öğrencinin ilgisini çekebilecek konular vardır. Mesela Tableau benim çok hoşuma gitti. Boş kaldığım zamanlar Tableau ilgili çalışmalar yapmaya çalıştım.”

“Gerçekten öğrenmemiz sağlandı. Farklı farklı yerlerde görselleştirme yapabilmemiz, kod yazabilmemiz, dağıtık işlemler yapabilmemiz iş hayatı öncesi kendimize güvenimizi de artırdı. İnteraktif CV yapımı ise iş hayatımız için olumlu bir etken.”



## 4. Sonuç

Bu çalışmada, 2018-2019 Bahar döneminde Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nde verilen "Büyük Veri Analizine Giriş" dersi kapsamında yapılan interaktif görselleştirme üzerinde durulmuştur. Genel olarak öğrencilerin şu konular üzerinde kazanım sağladıkları görülmüştür: Tableau ile farklı veri kümeleri arasında iletişim kurabilmeleri, veri kümesinin görselleştirme öncesinde düzenlenmesi, verinin farklı türden birçok grafik türü ile nasıl görselleştirileceği, interaktif harita kullanımı, interaktif veri analizi, interaktif özgeçmiş hazırlama, pano oluşturma ve kullanıcıya pano üzerindeki elementlerle etkileşim sağlanması, filtreleme ve bazı alanların vurgusunun yapılmasını sağlama vb. Lisans seviyesinde hem bilimsel hem de iş amaçlı olarak verilerin analizinin görsel bir şekilde yapılabileceğini göstermesi açısından çalışma önemlidir.

İleriye dönük olarak ders kapsamında işlenen tüm modülleri içerecek şekilde büyük veri kümelerinin nasıl analiz edileceğini ve herhangi bir şirketin ve organizasyonun karar alma sürecini iyileştirecek kalıpları tanımlamanın öğrenilmesinde faydalı olacak kişisel öğretim programı ve öğrenciler üzerinde yansımaları üzerinde bir çalışma yapılması planlanmaktadır.

## Kaynakça

- Buzan, T., & Griffiths, C. (2013). *Mind Maps for Business 2nd edn: Using the ultimate thinking tool to revolutionise how you work*. Pearson UK.
- Büschel, W., Vogt, S., & Dachselt, R. (2018). Investigating link attributes of graph visualizations in mobile augmented reality. In *Proceedings of the CHI 2018 Workshop on Data Visualization on Mobile Devices. MobileVis* (Vol. 18).
- Caldarola, E. G., Picariello, A., & Castelluccia, D. (2015). Modern enterprises in the bubble: Why big data matters. *ACM SIGSOFT Software Engineering Notes*, 40(1), 1-4.
- Caldarola, E. G., & Rinaldi, A. M. (2017). Big Data Visualization Tools: A Survey-The New Paradigms, Methodologies and Tools for Large Data Sets Visualization. In *DATA* (pp. 296-305).
- Eken, S. (2019) Büyük Veri Analizine Giriş Ders kaynakları. Erişim Adresi: [https://piazza.com/kocaeli\\_university/spring2019/blm442/resources](https://piazza.com/kocaeli_university/spring2019/blm442/resources) (Erişim tarihi: 10 Aralık 2019)
- Eken, S., & Kumru, P. Y. (2014). Haritalar Üzerinde Suç Verilerinin Görüntülenmesi ve Analizinin Sağlanması: Kocaeli İli Örneği. *Karamanoğlu Mehmetbey Üniversitesi Sosyal Ve Ekonomik Araştırmalar Dergisi*, 2014(3), 67-72.
- Eken, S., Türkoğlu, S., & Sayar, A. (2012). Integration of OpenGL Graphic Libraries with Spatial Database as an Analysis and Visualization Tool. *Selçuk-Teknik Dergisi*, 11(3), 110-123.
- Fox, P., & Hendler, J. (2011). Changing the equation on scientific data visualization. *Science*, 331(6018), 705-708.
- Friendly, M. (2008). A brief history of data visualization. In *Handbook of data visualization* (pp. 15-56). Springer, Berlin, Heidelberg.
- Gapminder (2019). Erişim Adresi: <https://www.gapminder.org/tools/> (Erişim tarihi: 10 Aralık 2019)
- Gartner (2019). Erişim Adresi: <https://www.gartner.com/en> (Erişim tarihi: 10 Aralık 2019)
- Gartner BI Report (2019). Erişim Adresi: <https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms> (Erişim tarihi: 10 Aralık 2019)
- Herman, L., Juřík, V., Stachon, Z., Vrbík, D., Russnák, J., & Řezník, T. (2018). Evaluation of User Performance in Interactive and Static 3D Maps. *ISPRS International Journal of Geo-Information*, 7(11), 415.
- Kaggle (2019). Erişim Adresi: <https://www.kaggle.com/jessicali9530/honey-production> (Erişim tarihi: 10 Aralık 2019)
- Lee, B., Srinivasan, A., Stasko, J., Tory, M., & Setlur, V. (2018, May). Multimodal interaction for data visualization. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (p. 11). ACM.
- Ma, X. (2017). Linked Geoscience Data in practice: Where W3C standards meet domain knowledge, data visualization and OGC standards. *Earth Science Informatics*, 10(4), 429-441.
- Mou, X., Jamil, H. M., & Ma, X. (2017, April). Visflow: A visual database integration and workflow querying system. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)* (pp. 1421-1422). IEEE.
- Murray, S. (2017). *Interactive data visualization for the web: an introduction to designing with*. " O'Reilly Media, Inc."
- Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition*, 1, 2006-2001.
- Schutt, R., & O'Neil, C. (2013). *Doing data science: Straight talk from the frontline*. O'Reilly Media, Inc..
- Tableau Classroom Training (2019). Erişim Adresi: <https://www.tableau.com/learn/classroom> (Erişim tarihi: 10 Aralık 2019)
- Tamayo, J. L. R., Hernández, M. B., & Gómez, H. G. (2018). Digital Data Visualization with Interactive and Virtual Reality Tools. Review of Current State of the Art and Proposal of a Model. *Revista ICONO14 Revista científica de Comunicación y Tecnologías emergentes*, 16(2), 40-65.
- Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1). A. J. Hey (Ed.). Redmond, WA: Microsoft research.
- The 50 best public dataset (2019). Erişim Adresi: <https://medium.com/towards-artificial-intelligence/the-50-best-public-datasets-for-machine-learning-d80e9f030279> (Erişim tarihi: 10 Aralık 2019)
- Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.
- Ulvi, M., Eken, S., & Sayar, A. (2013). Service Oriented Visual Interpretation Tool for Times Series Data. *Anadolu University of Sciences & Technology-A: Applied Sciences & Engineering*, 14(3).

- Veri kümesi 1 (2019). Erişim Adresi: <https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016> (Erişim tarihi: 10 Aralık 2019)
- Veri kümesi 2 (2019). Erişim Adresi: <https://www.kaggle.com/carolzhangdc/imdb-5000-movie-dataset> (Erişim tarihi: 10 Aralık 2019)
- Veri kümesi 3 (2019). Erişim Adresi: <https://public.tableau.com/en-us/s/interactive-resume-gallery> (Erişim tarihi: 10 Aralık 2019)
- Washington Post, A World Apart (2019). Erişim Adresi: <https://www.washingtonpost.com/sf/local/2013/11/09/washington-a-world-apart> (Erişim tarihi: 10 Aralık 2019)
- Zudilova-Seinstra, E., Adriaansen, T., & Van Liere, R. (2009). Overview of interactive visualisation. In Trends in Interactive Visualization (pp. 3-15). Springer, London.