



Araştırma Makalesi  
Geliş Tarihi: 06.05.2020  
Kabul Tarihi: 22.03.2021

Research Article  
Received: 06.05.2020  
Accepted: 22.03.2021

## Bir büyük veri görselleştirme uygulaması olarak konut tercih infografikleri<sup>1\*</sup>

**Elif Erkurt<sup>2</sup>**

Dr.  
elifsntrk@hotmail.com  
 0000-0001-5665-9507

**İlknur Esen Yıldırım**

Prof. Dr., Marmara Üniversitesi  
eyildirim@marmara.edu.tr  
 0000-0003-2574-4340

### ÖZ

Günümüz dünyasında büyük veri olgusunun gittikçe önem kazanması ile birlikte istatistikçilere, veri analistlerine ve bu analizleri mümkün kılacak teknolojik alt yapının geliştirilmesi için de yazılımcılara duyulan ihtiyaç artmaktadır. İlgili alan uzmanları için temel hedef, ele alınan büyük veriden rafine edilebilecek enformasyonu ortaya çıkarmak ve mümkünse bulguları kapsamlı, işlevsel ve sade görsellere yansıtılabilmektir. Bu çalışmada, emlak sektörüne ilişkin bir uygulama üzerinde, büyük veri analizi ile üretilebilecek görsellerin, ilgili hedef kitleye sunabileceği önemli, hatta kimi zaman alternatifsiz kolaylığa dikkat çekilmeye çalışılmıştır. Bu amaçla; R programı yardımıyla, konut tercihini kolaylaştırıcı bir araç olarak, kullanımı kolay ve dinamik yapıda infografikler üretilmiştir. Böylelikle, konut satın almayı planlayan bireylerin konutun yaşı, yüzölçümü, oda sayısı, fiyatı gibi önemli buldukları kriterleri dikkate alan, veri değişimine duyarlı, iller/ilçeler arasında karşılaştırma yapılmasına imkân veren infografiklerden yararlanarak tercih yapabilmelerine olanak sağlanmıştır.

### Anahtar Kelimeler:

Büyük Veri,  
Görselleştirme,  
İnfoğrafik, Konut,  
Emlak.

**JEL Kodları:** C19, C49,  
C89.

## Real estate preference infographics as big data visualization application

### ABSTRACT

Statisticians, data scientists and software programmers/engineers who develop technological infrastructure to enable performing analytics are in high demand with the increasing importance and usage of big data in the last decades. The main goal of subject matter experts is to extract actionable insight from big data, and if possible, to deliver these insights with effective and simple visualizations. This study aims to draw attention to important, even sometimes unrivaled convenience offered to the relevant target audience by visualizations, which were created with big data analysis on an application related to real estate sector. For this purpose, by means of R program, and as a tool to ease the housing selections; user-friendly and dynamic-structured infographics were created. Thus, individuals, who are planning to buy houses, are given the opportunity to make selections by using infographics which enable them to take essential criteria into consideration - such as the age of the property, area measures, number of rooms, price; are sensitive to data variations and allow to make comparison between cities/districts.

**Keywords:** Big Data,  
Visualization,  
Infographics, Housing,  
Real Estate.

**JEL Codes:** C19, C49,  
C89

<sup>1</sup> Bu çalışma 2020 yılında Marmara Üniversitesinde İlknur Esen Yıldırım danışmanlığında Elif Erkurt tarafından hazırlanan "Büyük Veri Görselleştirme ve Türkiye'de Konut Sektörüne İlişkin İnfografikler" başlıklı doktora tezinden türetilmiştir.

\* Atıf vermek için/To cite: "Erkurt, E. ve Yıldırım İ. E. (2021). Bir büyük veri görselleştirme uygulaması olarak konut tercih infografikleri. *KOCATEPEİİBFD*, 23(1), 36-52. <https://doi.org/10.33707/akuiibfd.733379>

<sup>2</sup> Sorumlu Yazar/Corresponding Author



## Extended Summary

The big data phenomenon has recently emerged as a result of technological evolution and widespread use of social media. Big data mainly refers to unstructured, very large volume data types. As accessing information hidden in such complex types of data sets is extremely hard; subject matter experts aim to develop new methods. Visualization is a frequently used approach to effectively organize, analyse and present the information it carries to the user.

Visualization is demystifying the information hidden in big data through various graphics or videos in a static or dynamic way. Complex data sets turn into an output that users can easily and quickly benefit from. Examples of this new visualization language used for exploratory or explanatory purposes are particularly found in business analysis, governments' open data systems, and media.

The necessity for big data analysis is felt intensely in the real estate sector. Limited availability of papers about big data visualization in this sector has become one of the motivation factors of this research. The number of advertisements for sale or rental is quite high, and significant changes may be observed on a daily basis. Therefore, visuals are required to analyse the current real estate ads practically and entirely.

This study aims to develop visual tools by which users can easily evaluate the house best fits their preferences among all alternatives. As the number of houses may reach up to hundreds of thousands, researching the most suitable house seems frustrating and tiresome. The major goal of service providers is to bring the existing housing portfolios to the right customers. Based on these facts, the infographics created within the scope of this study will allow individuals to find their dream houses after satisfactory, pleasant, and effortless research. The goal is to benefit from the data not only with maximum efficiency but also with minimum effort and time.

Housing advertisements in the cities of Istanbul, which has the highest number of housing in Turkey, and Ankara and Izmir, two of the metropolitan cities in Turkey to which people migrate from Istanbul the most, have been analysed comparatively. The data set used in the study has been acquired from zingat.com, and it was compiled from 79,632 residential for sale ads in Istanbul, Ankara, and Izmir in 2018.

The variables "net area, number of rooms + living room, age of the building, and price" have been included, which are among the main criteria considered when purchasing a house. Statistically significant relationships were found between all variables.

The R-3.5.2 programming language was used for the analyses. "Dplyr," "ggplot2," and "plotly" packages were used for performing related analysis and visualization. Median values were used as a central tendency measure to prevent the effect of extreme values.

The developed animation graphics allow a comparison of all for-sale residential advertisements in the provinces of Istanbul-Ankara and Istanbul-Izmir in terms of selected criteria. All data can be examined on the animated graph based on the selected criteria. Video formatted graphics can be paused and details of any advertisement can be seen in a colourful box. Using these infographics makes it possible to evaluate all the advertisements meeting the selected criteria, make comparisons between provinces, and examine the details of a specific advertisement. The visuals presented in the text have been created based on the age of the building, and the animation was paused in certain years to take images.

Overall, the net area of residences increases as the age of the building decreases in all three provinces. The number of houses for sale along with their housing prices was found to be higher in Istanbul than those in the other two provinces. Additionally, while there are more residential alternatives for sale in Istanbul than in Ankara for all building ages, the number and prices of houses in Izmir have a similar trend as in Istanbul in recent years.

Although there are many housing ads out there, a system with visuals to compare ads at the provincial, district level and on the basis of certain criteria does not yet exist. With the user-friendly infographics presented in this study, it is possible to see current ads on a functional visual and to obtain the details of a particular advertisement with faster and easier research.

Visualizations, similar to the ones developed in our study, can be adapted to lands, workplaces, or ad services for cars, job search, and many others. These types of infographics are big-time savers, especially when dealing with large data sets exhibiting instantaneous change.

If the data has a geographical basis, infographics can be enriched with relevant maps. Communication bridges can also be established in the infographics that connect service providers and receivers.

While transferring big data to visuals, we have come across a missing data problem. It is a common issue across many sectors and needs to be handled.

## I. Giriş

Büyük veri çağı olarak adlandırılan yaşadığımız dönem, özellikle işgücü piyasası için altın çağdır. Ancak karmaşık yapıdaki ve büyük miktardaki bu veri yığınına bilgiye dönüştürmek ve rekabet piyasasında yer edinebilmek için güçlü bir teknolojik alt yapıya, yetenekli ve iyi eğitilmiş işgücüne sahip olmak, ayrıca yeterince hızlı olmak gerekmektedir. Bu kriterleri bir arada sağlayabilen şirketler dünya çapında başarı elde etmektedir. Google, Twitter, Amazon; büyük veri alt yapısını kurarak kazanca dönüştürmeyi başaran şirketlerden bazılarıdır.

Büyük verinin çözümlenmesinin yanı sıra, elde edilen bulguların bu veriden bilgi edinecek kesime hitap edilebilmesi, anlaşılır ve yararlanılabilir bir formda sunulabilmesi de önemli bir sorundur. Bu kapsamda uygun grafiklerden yararlanmak, bilgiyi ilgi çekici ve sade görseller aracılığıyla özetlemek en etkili yollardan biridir. Bu nedenle büyük veri ve görselleştirme kavramlarının birlikte anılması tesadüf değildir.

Bu çalışmada, emlak sektörü verisi üzerinde bir büyük veri görselleştirme uygulaması yapılmıştır. Literatür kısmında detaylı verildiği üzere, Türkiye ölçeğinde konut sektörünü ele alan çalışma sayısının az sayıda olması, bu çalışmanın motivasyon faktörlerinden biridir. Çalışma kapsamında, Türkiye’de il bazında konut tercihinde kullanılacak ve ilgilenilen kriterlere göre ilanlar arasında sorgulama yapılmasına imkân verecek nitelikte infografikler üretilmiştir. Bu infografikler üzerinde, tüm mevcut ilan datasının görsel üzerinde incelenmesi ve ikiye bölünmüş olarak iller arası karşılaştırma yapılması mümkündür.

Çalışmanın birinci bölümünde literatür özetine yer verilecek, ikinci bölümünde büyük veri, üçüncü bölümde büyük veri teknolojisi ve dördüncü bölümünde büyük veri görselleştirme hakkında bilgiler verilerek, “Türkiye’de Emlak Sektörüne İlişkin Infografikler” başlığı altında yapılan ampirik çalışmanın veri seti, gerçekleştirilen analizler, bulgular ve yorumlar sunulacaktır. Bu çalışma, “Büyük Veri Görselleştirme ve Türkiye’de Konut Sektörüne İlişkin Infografikler” başlıklı yayımlanmamış doktora tezinden üretilmiştir (Erkurt, 2020).

## II. Literatür

Büyük veri kavramı ortaya atıldığı günden itibaren yayınlanan araştırmalar incelendiğinde, sıklıkla sosyal medya verisi üzerinde çalışıldığı, büyük veri analizi öncesinde ve sonrasında karşılaşılan yapısal ve teknik sorunların ele alındığı görülmektedir. Bu bağlamda, öncelikle büyük veri analizindeki güçlükleri konu alan bazı çalışmalardan bahsedilecektir.

Tole (2013), “Big Data Challenges” başlıklı makalesinde büyük veri kullanımında sıklıkla karşılaşılan veri depolama, gizlilik ve güvenlik, veri paylaşımı gibi problemleri ele almıştır. Ayrıca pahalı donanım ekipmanları, alt yapı yetersizliği ve analiz esnasında karşılaşılan zorluklar gibi çeşitli sorunlara ve bu sorunların en hızlı ve uygun şekilde nasıl çözüme kavuşturulabileceğine değinmiştir (Tole, 2013).

Sakı (2016), “Big Data: Understanding Big Data” adlı çalışmasında büyük veriyi detaylı olarak incelemiş, büyük veride sorun olabilen depolama alanı ve veri işleme gibi konuları, ayrıca büyük verinin işlenmesi aşamasında kullanılan MapReduce uygulamasını detaylı bir şekilde ele almıştır. Bunun yanında sağlık ve genetik gibi çeşitli konularda bazı senaryolar üreterek büyük veri analitiği ile ilgili kavramsal birtakım yaklaşımları açıklamış, kişilerin hücreleri göz önüne alındığında bile büyük veri ile karşılaşıldığını vurgulamıştır. Bu şekilde farklı sektörlerde büyük veri kullanımının ne gibi avantajlar sağlayabileceğine değinmiştir (Sakı, 2016).

Soto ve diğerlerinin (2018), 51. Hawaii Uluslararası Sistem Bilimleri Konferansında sundukları “Data Quality Challenges in Twitter Content Analysis for Informing Policy Making in Health Care” isimli çalışmalarında, kişilerin Twitter gibi çeşitli sosyal medya platformlarından veri derlenmesi ve analiz edilmesi gibi çeşitli aşamalarda karşılaştıkları zorlukları açıklamışlardır. Bu çalışma ile bu tarz analizlerle çeşitli politikalar oluşturmak, sosyal medya ve gerçek dünya arasındaki temel farklılıkların açıklanabileceği bir değerlendirme yapmak amaçlanmıştır. Sosyal medya analizi esnasında karşılaşılan zorlukların aşılmasına ilişkin bazı stratejiler sunularak metin madenciliği üzerinde çeşitli değerlendirmeler yapılmıştır (Soto vd., 2018).

Hewage ve diğerleri (2018), “Review: Big Data Techniques of Google, Amazon, Facebook and Twitter” adlı makalelerinde Google, Amazon, Facebook ve Twitter gibi sosyal medya kuruluşları tarafından kullanılan büyük veri tekniklerini karşılaştırmışlardır. Ele alınan veriler, veri depolama özellikleri ve tekniklerine ilişkin olup 2007-2015 yılları arasında 16 hakemli bilimsel yayından elde edilmiştir. Çalışma literatürdeki teknikler ve karşılaştırmalar aracılığıyla uygulanan farklı deneysel analizlerin açıklanmasını içermektedir. Bu tarz sosyal medya platformlarının teknolojileri, açık kaynak kodlu olup olmamaları, depolama yöntemleri ve alt yapıları gibi çeşitli açılardan farkları ortaya konmuştur (Hewage vd., 2018).

Büyük veri 2012 yılından bu yana dünyada oldukça popüler hale gelmiş ve akademik alanda çok sayıda çalışma yapılmış olup; ülkemizde henüz yeterli düzeyde olmasa da gittikçe artan sayıda yayın ve etkinlik gerçekleştirildiği tespit edilmiştir.

Demirtaş ve Argan'ın (2015), "Büyük Veri Ve Pazarlamadaki Dönüşüm: Kuramsal Bir Yaklaşım" başlıklı çalışmaları büyük veri konusunda yapılan çalışmalara örnek olarak verilebilir. Bu çalışmada büyük veri kavramı incelenmiş, pazarlama alanında yapılan uygulamalar ve pazarlamada büyük veri kullanımının avantajları ele alınmış ve diğer sektörlerdeki işleyişine yer verilmiştir. (Demirtaş & Argan, 2015)

Aktan'ın (2018), "Büyük Veri: Uygulama Alanları, Analitiği ve Güvenlik Boyutu" adlı çalışmasında büyük veri bileşenleri incelenmiş, büyük veri analitiği süreçleri ve Hadoop detaylı şekilde ele alınmış, büyük veride güvenlik sorunu ve gereken önlemlere yer verilmiştir. (Aktan, 2018)

Çelik ve Akdamar'ın (2018), "Büyük Veri Ve Veri Görselleştirme" adlı makalelerinde büyük veri ve veri görselleştirmenin önemine yer verilmiş, görselleştirme teknikleri detaylı olarak ele alınmıştır. (Çelik & Akdamar, 2018)

Türkiye'de "e-devlet" hizmeti, en kapsamlı büyük veri uygulamalarından biridir. Ayrıca, büyük veri kullanan devlet kurumlarına örnek olarak, Sağlık Bakanlığı ve Milli Eğitim Bakanlığı gösterilebilir. Sağlık Bakanlığı; e-Nabız, MHRS (Merkezi Hekim Randevu Sistemi) gibi uygulamalarla hastaların tetkik sonuçlarına ve tıbbi özgeçmişlerine ulaşılmasına, randevu alınabilmesine imkân sağlamıştır. Milli Eğitim Bakanlığı bünyesindeki e-okul uygulaması ile de öğrencilere ait veriler tüm öğrencilik hayatları boyunca kaydedilebilmekte; bu sayede öğrencilerin bireysel ve kurumsal takibi yapılabilmektedir (Erbay & Kör, 2016, s. 21).

2017 yılında "Big Data Çalıştayı", "Türk Keneşi<sup>3</sup> ve Dünya Komünikasyon ve Bilgi Toplumu Günü" gibi çeşitli etkinliklerle büyük veri gündeme gelmiş; ülkemizdeki kullanımı değerlendirilmiştir. Bu toplantılarda, daha çok halihazırda en fazla büyük veri kullanan telekomünikasyon şirketlerinin kullanım alanları ve ilgili ürünleri tanıtılmıştır ("Türk Keneşi & Dünya Telekomünikasyon ve Bilgi Toplumu Günü Etkinliği", 2017).

Konut sektörü üzerine yapılan büyük veri literatürü incelendiğinde, bireylerin konut arama davranışları, konut ilan kriterleri, konut fiyatı tahmini gibi başlıklarda ele alındığı görülmektedir.

Du, Li ve Zhang (2014), "Survey on the Applications of Big Data in Chinese RealEstate Enterprise" adlı çalışmalarında Çin'deki konut satış şirketlerindeki büyük veri uygulamaları üzerine bir uygulama gerçekleştirmişlerdir. Çalışmada, konut sektöründe yaşanan problemler ve çözümleri ele alınmış; gizlilik, veri işleme ve konutların kendine özgü özelliklerinden kaynaklanan zorluklar gibi pek çok sorun tanımlanmıştır. Konut sektörü üzerine yapılan büyük veri uygulamalarının, henüz yeterli düzeyde olmasa da, konut satışlarını destekleyici mahiyette olduğu kanaatine ulaşılmıştır (Du vd., 2014).

Sun ve diğerlerinin (2015), "Combining Online News Articles and Web Search to Predict the Fluctuation of Real Estate Market in Big Data Context" başlıklı makalelerinde yine Çin'deki konut sektörü üzerine yapılmış bir büyük veri uygulaması görülmektedir. Arama motoru sorgu verileri ve çevrimiçi kullanıcılar tarafından aranan kelimelerin sıklıklarının analiz edildiği ve internet kullanıcılarının duygularının yansıtılmasının hedeflendiği bu çalışmada haber ve sorgu faktörleri modele dahil edilerek, Çin'deki konut fiyatlarının öngörülmesi hedeflenmiştir. Önerilen entegre model, temsili olarak seçilen dört şehir için uygulanmış; elde edilen bulgular modeli doğruladığından, Çin'de şehir bazında konut arama davranışlarının etkisinin güçlü olduğu hipotezi kanıtlanmıştır (Sun vd., 2015).

Loberto ve diğerleri tarafından (2018), "The Potential Of Big Housing Data: An Application To The Italian Real-Estate Market" adıyla yayınlanan bu çalışmada veri seti olarak İtalya'da konut satış ve kiralama hizmeti veren çevrimiçi bir portal olan "Immobiliare.it" aracılığıyla elde edilen konut satış ilanları kullanılmıştır. Sitenin amacı konut piyasasındaki alıcı ve satıcıların eşleşebilmesini kolaylaştırabilmektir. Bu sitedeki veri setinin temel sorunu ilanlar ile gerçek konutların birbiriyle uyumsuzluğu olup, makine öğrenmesi araçları kullanılarak bu uyumsuzluğun önüne geçilmesi amaçlanmıştır. Veri setindeki, bir başka ifadeyle ilanlardaki verilen bilgilerin mevcut resmi konut istatistikleri ile tutarlı olduğunun gösterilmesi elde edilen sonuçlar arasındadır (Loberto vd., 2018).

Büyük veri görselleştirme üzerine yapılan çalışmalar incelendiğinde yine oldukça ilgi çeken bir konu olduğu anlaşılmaktadır. İlgili konuya ilişkin bazı çalışmalara yer verilecektir.

Maçs ve diğerlerinin (2015), "Time-Series Application on Big Data Visualization of Consumption in Supermarkets" adlı makalelerinde tüketimin evriminin görselleştirilmesi ve dönemsel davranışların tespiti amaçlanmış, büyük veri aracılığıyla ürün tüketiminde bilgi görselleştirme uygulanmıştır. Çalışma için 730

<sup>3</sup> Türk Dili Konuşan Ülkeler Birliği Konseyi

gün sürecinde 2.86 milyar işlem gerçekleştirilmiştir. Elde edilen bulgulara göre haftanın çoğu gününde tüketicilerde davranış tekrarı gözlemlenmiş, tüketimin genellikle 16:00-20:00 saat aralığında yüksek olduğu tespit edilmiştir. Gelecekte araştırmadaki sapmaları daha yüksek oranda belirleyebilecek bir araç geliştirmek ve kullanıcılara belirleyici bir alışveriş takvimi sunmak amaçlanmıştır (Maças vd., 2015).

Zhang, Zhou ve Zhang'ın (2017), "Quantifying and Visualizing Jobs-Housing Balance with Big Data: A Case Study of Shanghai" isimli çalışmalarında hücresele ağ verilerinin işyeri ve barınma amaçlı konutların nasıl işlenebileceği üzerinde durulmuştur. Şangay'a ilişkin bir vaka analizi içeren ilgili çalışmada bu veriler ile Baz Alıcı-Verici İstasyonları ile yerel yolcular tanımlanmış ve coğrafi olarak kodlanmıştır. Bireylerin yaklaşık olarak işe gidiş geliş yollarındaki trafik akışları da görselleştirilmiş ve Şangay'daki işe gidip gelme amacıyla kullanılan yollarının optimize edilebilmesi için 12.5 kilometreden daha yakın mesafeye sahip olmaları gerektiği bulgusuna ulaşılmıştır (Zhang vd., 2017).

Su, Liu ve Wang'ın (2018), Journal of Ambient Intelligence and Humanized Computing'de yayınlanan "A Principle of Designing Infographic for Visualization Representation of Tourism Social Big Data" isimli çalışmalarında Octopus web veri toplama yazılımı aracılığıyla bir Facebook hayran sayfasının seyahatle ilgili büyük boyutlu sosyal verilerini toplamayı ve Tableau aracılığıyla 6 farklı grafik türü oluşturmayı amaçlamışlardır. Elde edilen bulgulara göre grafik boyutu, rengi, renk gradyanları, verileri, sıralaması gibi bilgilerin kullanıcıların grafiği izleme sonuçlarını etkilemektedir. Bu sebeple turizm verileri ele alınarak yapılan çalışma sonucunda infografiği göstermek için uygun renk kullanımı, hızlı bir şekilde algılanabilmesi için verileri işaretleme, öğelerin boyutlarının sıralanması ve farklı boyutlardan yararlanarak oranların arasındaki farkı algılayabilmesini sağlamak fikirleri önerilmiştir. Bu sayede elde edilen bulguların kullanıcılara en doğru biçimde iletilmesi sağlanmaktadır (Su vd., 2018).

Lavalle, Teruel, Maté ve Trujillo'nun (2020), "Improving Sustainability of Smart Cities through Visualization Techniques for Big Data from IoT Devices" isimli çalışmalarında IoT cihazlarından gelen büyük veri artışından yararlanan akıllı şehirlerin yangın gibi olası tehlikeleri önlemek ve kaynakları optimize etmek için uygun görselleştirme tekniklerine dayalı bir metodoloji sunmaları ele alınmıştır. Kullanıcıların hedeflerini tanımlamalarına ve en iyi görselleştirme yöntemini bulmalarına yardımcı olmak amacıyla akıllı şehirler için gerçek veriler kullanılmıştır. Bu sayede stratejik kararlar almak için gerekli verilere ulaşabilen kullanıcılar, IoT sensörlerinden yararlanabilirler. Bu çalışmada itfaiye departmanının çağrı hizmetine dayalı bir vaka çalışmasına yer verilmiş, gerçek veri olarak vatandaş çağrıları kullanılmıştır. Görselleştirmenin kullanıcıların bilgi ihtiyaçlarını daha iyi karşıladığı bulgusu elde edilmiştir. Bununla birlikte ilgili çalışmanın diğer akıllı şehirlerin süreçlerine de uygulanma amacı, gelecek çalışmaların konusunu oluşturmaktadır (Lavalle vd., 2020).

Büyük veri görselleştirmeye ilişkin Türkiye adresli çalışmalar tespit edilmiştir.

M. Anuşlu ve T. Anuşlu'nun (2018), Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi'nde yayınlanan "Application of Big Data Visualization with Agricultural Value-Added of All Countries" isimli makalelerinde ülkelerin 1960-2016 yılları arasındaki tarımsal katma değer yüzdelerinin R yazılım programı aracılığıyla görselleştirilmesi yapılmıştır. Bu araştırma sonucunda tarımda etkin bir yere sahip olan veya herhangi bir etkisi olmayan ülkelerin katma değer yüzdelerinin yıllar içindeki değişimi interaktif harita (korople) aracılığıyla okuyucuya sunulmuştur. Elde edilen bulgularda Çin'in özellikle son yıllarda tarımsal katma değerde dünya lideri olduğu, ABD'nin ise dönemler boyunca iyi bir konuma sahip olduğu görülmüştür. Türkiye'de ise yıllar içinde tarımsal katma değerde önemli bir değişim görülmediği ifade edilmiştir (Anuşlu & Anuşlu, 2018).

Eken'in (2020), Avrupa Bilim ve Teknoloji Dergisi'nde yayınlanan "Büyük Verinin İnteraktif Görselleştirilmesi: Tableau Üzerine Öğrenci Deneyimleri" isimli çalışmada interaktif görselleştirme tanımlanmış ve Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü'nde verilen "Büyük Veri Analizine Giriş" ders içeriğinde yer alan Tableau ile büyük verinin interaktif görselleştirilmesi üzerinde durulmuştur. Çalışma ile büyük verinin sahip olduğu bilginin özünün elde edilmesi, grafikler ve çeşitli görseller aracılığıyla veri kümelerini ilişkilendirme gibi çeşitli çıkarımlarda bulunulmuştur (Eken, 2020).

Uğur ve Akbıyık (2020), "Tourism Management Perspectives" dergisinde yayınlanan "Impacts of COVID-19 on Global Tourism Industry: A Cross-Regional Comparison" adlı çalışmalarında metin madenciliği yöntemi aracılığıyla turizm endüstrisinin COVID-19 salgınından etkilenme durumunu incelemişlerdir. Bu amaçla 30.12.2019-15.03.2020 tarihleri arasında TripAdvisor forumlarından yaklaşık 75.000 yorumu analiz etmiş ve gezginlerin seyahatlerini iptal etmeleri veya ertelemeleri, seyahat sigortası, para iadeleri gibi pek çok olumsuz durum sebebiyle ilgili sektörün küresel krizlerden kolaylıkla etkilendiği bulgusunu elde etmişlerdir (Uğur & Akbıyık, 2020).

Türkiye'de büyük veriyle ilgili çok sayıda yayın mevcut olsa da, konut sektörüne ilişkin çok az sayıda çalışma mevcuttur. "Türkiye'de Satış Bulunan Konutların İl ve Bölgeler Bazında Dağılımının Büyük Veri Teknolojisi

İle İncelenmesi” başlıklı çalışmada, 2017 yılının ilk altı ayında internet sitelerinde satış ilanı mevcut olan 81 ile ilişkin yaklaşık 400.000 dairenin oda sayısı, fiyatı gibi çeşitli özelliklerinin il ve bölgelere göre dağılımı incelenmiştir. Araştırmanın amacı, satış fiyatlarının bölgeler bazında medyan ve ortalama değerlerinin tespiti ile bölgeler arası farklılıkların ortaya konulmasıdır. Araştırmadan elde edilen bulgulara göre en yüksek ortalama satış fiyatına sahip olan bölge Marmara, en düşük fiyatlara sahip olan bölge ise Doğu Anadolu Bölgesidir (Yılmaz vd., 2017).

### III. Büyük Veri Kavramı

“Büyük veri”, barındırdığı anlamları tam olarak yansıtmadığı eleştirilerine konu olan, tanımlanması zor bir kavramdır. Büyük veri ifadesinin miktarı vurguladığı, ancak miktar algısının bir standardının olmayıp kişiden kişiye göre değiştiği ve verinin karmaşık yapısını aktarmadaki yetersizliği sıkça dile getirilmektedir. Dolayısıyla esasında büyük veri, anlam olarak ifade edilenden çok daha fazlasını nitelendirir (Davenport, 2014, s. 13). Büyük miktarda veri toplamak, depolayarak saklamak, analiz etmek ve elde edilen sonuçlarla gerekli tahminlerde bulunmak gibi oldukça detaylı ve uzun zamana yayılan sürecin tamamı “büyük veri” olarak tanımlanabilir (Feinleib, 2013, s. 21).

Büyük veri kavramı, ilk olarak 2000 yılında Seattle’da gerçekleştirilen 8. Dünya Ekonometri Kongresi’nde gündeme gelmiştir. Franchis X. Diebold tarafından sunulan “Big Data Dynamic Factor Models sor Macroeconomic Measurement and Forecasting” adlı makale, büyük veri teriminin geçtiği ilk akademik referans olarak kabul edilmektedir (Lohr, 2013; Gürsakal, 2014, s. 19,20). Ancak büyük veri için dönüm noktası, 2012 yılıdır. 2012’de İsviçre’de düzenlenen Dünya Ekonomik Forumu’nda hazırlanan “Büyük Veri, Büyük Etki” adlı raporda büyük veri, “öncelikli ve önemli konu” olarak takdim edilmiştir. Aynı yıl ABD’den, 200 milyon dolar gibi oldukça büyük bir bütçenin büyük veri çalışmaları için kullanmasının planlandığı açıklaması gelmiştir (Lohr, 2012).

Büyük veri, geleneksel olarak 5V olarak isimlendirilen boyutlar ile tanımlanmaktadır. Bu boyutlar, hacim (volume), hız (velocity), çeşitlilik (variety), doğruluk (veracity) ve değerdir (value). Kurumlar açısından sürecin ilk zorlu kısmının söz konusu boyutlar olduğu söylenebilir. Devasa boyutlara ulaşan (hacim), heterojen yapıda bulunan (çeşitlilik), sürekli ve son derece hızla artan (hız), dolayısıyla eksik veya güvenilirliği düşük olabilen (doğruluk) ve analizi gerçekleştirilerek elde edilen çıktıdan enformasyon üretilebilen büyük veri, oldukça kapsamlı ve zorlu şartlar barındırır. Bunun yanı sıra yapılan büyük veri çalışmaları verilere bakış açısını genişletmiş ve farklı boyutların dahil edilmesine aracı olmuştur. Günümüzde büyük veri boyutları 7V, hatta 10V ile anılmakta olup bu boyutlar verilerin doğruluğunun yanı sıra geçerli olması gereğini ifade eden geçerlilik (validity), verilerin yaşam süresinin önemini belirten oynaklık (volatility), tutarsız veri akışını dikkate alan değişkenlik (variability), büyük verinin gerektiğinde daha fazla veri üretme yeteneğine sahip olması anlamına gelen yaşayabilirlik-canlılık (viability) ve büyük verinin karmaşık yapısı içerisinde küçük bir değişimin bile önemli bir etkiye sahip olabileceğini ifade eden akışkanlık (viscosity) şeklinde tanımlanmaktadır (Khan vd., 2018, ss. 53,54).

Büyük veri, 100 terabayttan 2 petabayta kadar hacme sahip olabilen yapılandırılmamış veri olup, bu formdaki veriye başa çıkılabilmesi için uygun analiz programlarının geliştirilmesi ve bu sayede verinin değere dönüştürülmesi gerekmektedir (Davenport, 2014, s. 7-10). Bununla birlikte; karmaşık ve kaotik bir yapıya sahip olan, miktarı hızla büyüyen bu veri yapısını analiz edebilecek programları geliştirmek aynı hızla gerçekleşmemektedir (Gürsakal, 2014, ss. 9-35).

Büyük veriden elde edilen bilgilerin anlaşılır ve kapsamlı bir biçimde sunulması ve verilmek istenen mesajların iletebilmesinde, görselleştirmeden yararlanmak en etkin yaklaşımdır. Arka planında istatistik teknikler, yazılımlar ve grafikleme bilgisi bulunan infografikler, günümüzde analistler ve kullanıcılar için işlerliği yüksek araçlar haline gelmiştir.

Büyük veri analizleri, verinin bilgiye, yani değere dönüşmesine; böylelikle şirketlerin rekabet ortamında avantaj sahibi olmasına imkân verir (Watson, 2014, s. 1252). Bununla birlikte, her ölçekte şirketin büyük verinin boyutları ile ilgili, kullanılan teknik alt yapının yetersizliği, veri kalitesi, geçerliliği, gizlilik ve güvenlik, veri paylaşımındaki isteksizlik ve zorluklar gibi pek çok unsurdan dolayı sorunlarla karşı karşıya kaldığı bilinmektedir.

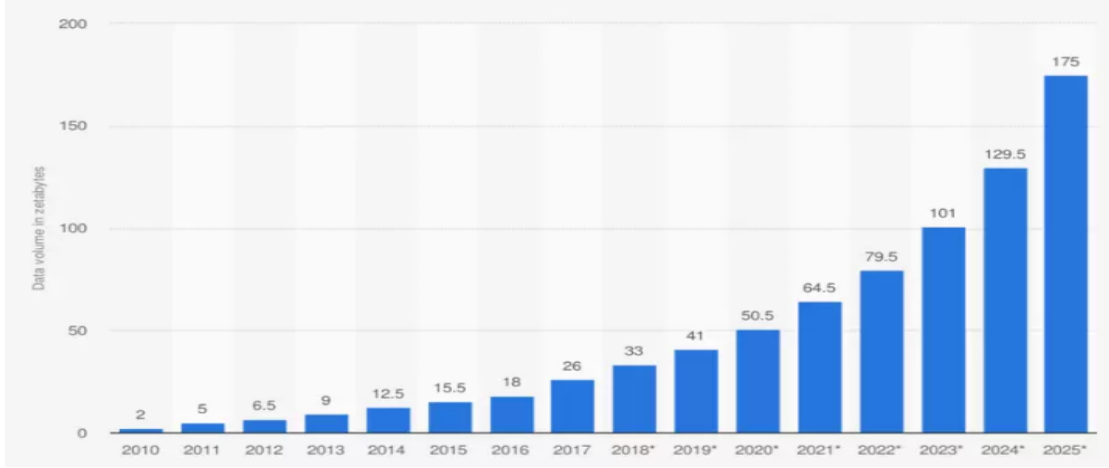
### IV. Büyük Veri Teknolojisi

Büyük veri tipleri genel olarak iki gruba ayrılır: Birincisi; finansal, kurumsal vb. içeriğe sahip olan düzenli ve yapılandırılmış formdaki verileri kapsar. İkincisi ise; sosyal medyanın hayatımızdaki yadsınamaz yeri ile birlikte ortaya çıkan; e-posta, resim, metin, video gibi yarı yapılandırılmış veya yapılandırılmamış verileri içerir IDC (International Data Corporation)’nin yayınladığı rapor incelendiğinde, veri miktarının 2020 yılında 35 zettabayt ile, 2009 yılındakinden 44 kat fazla olacağını öngörüldüğü anlaşılmaktadır. Ancak bu rakama 2018 yılında ulaşılmış ve yaklaşık 33 zettabayt veri üretimiyle bir rekor kırılmıştır. Yeni veriler

doğrultusunda IDC'nin 2025'de ulaşılacak veri miktarı tahmini güncellenerek 175 zettabayta olarak değerlendirilmiştir (Press, 2020). Bu miktardaki veri ile 23 defa aya gidip gelenebilir veya verinin 222 defa dünyayı çevreleyebilir olduğunu bilmek, miktarın büyüklüğünün daha anlaşılır olmasını sağlamaktadır (Reinsel vd., 2018, s. 7).

Büyük veri miktarının yıllara göre artış eğilimi Şekil 1'de açıkça görülmektedir.

**Şekil 1:** Dünya Genelindeki Veri Miktarının Yıllara Göre Artışı (Zettabayt)



**Kaynak:** <https://www.weforum.org/agenda/2020/01/privacy-in-a-world-of-ai-and-big-data/>

Şekil 1'e göre 2010 yılında 2 zettabayt (2 trilyon gigabayt) olan veri miktarı, 2017 yılında 26 zettabayta ulaşmıştır. Verinin artış hızı ve teknolojik gelişmeler de göz önüne alınarak 2024 yılında 129.5 zettabayta ulaşacağı tahmin edilmiştir.

Ekonomik açıdan ele alındığında; 2019 yılında Cloudera ve Hortonworks şirketlerinin birleşmesiyle büyük veri kullanımında dünya üzerinde ciddi bir artış meydana geldiği görülmüştür. Büyük veri endüstrisi 2019'da 2018 yılına göre 20 milyar doların üzerinde büyüyerek 189 milyar dolar değerine ulaşmıştır. 2022 yılında ise bu rakamın 247 milyar dolar olması beklenmektedir (Dialani, 2020). Büyük veri miktarındaki artış, büyük veri endüstrisindeki artış eğilimini açıklamaktadır.

Büyük verinin depolanması ve analizi için, yeni teknolojiler ve yeni dosyalama sistemleri geliştirilmektedir. Veri depolama ve yönetim sistemleri arasında ilk akla gelen ve en çok tercih edilenler, NoSQL, MapReduce, Hadoop'tur (Padmavalli, 2016, s. 16). NoSQL, ilişkisel olmayan bir veri tabanıdır; herhangi bir yapının XML, metin, ses, video gibi her türlü verisini depolayabilir. NoSQL, belge ve grafikler gibi veriler için tasarlanmıştır; kendi veri depolama ve alma yöntemlerini kullanır (Watson, 2014, s. 1258).

MapReduce ("Eşleldirge" şeklinde Türkçeleştirilmiştir), büyük miktardaki veriyi farklı bilgisayarlara dağıtarak işlemeyi hedefleyen bir tekniktir; büyük veri, analizden önce küçük ve yönetilebilir parçalara bölünür ve birbirine paralel olarak çalışabilen farklı sunuculara dağıtılır. Bu sayede; çok sayıda eşanlı işlem gerçekleştirilirken, olası bir aksaklığın diğer veri gruplarını etkilemesinin önüne geçilir. İşlemlerin sonunda elde edilen sonuçlar birleştirilir (Gürsakar, 2014, ss. 211-213).

Hadoop, Facebook ve Google gibi birçok büyük şirket tarafından da tercih edilen bir açık kaynak yazılımıdır. Birden fazla veri kaynağıyla çalışılır ve birden fazla sistemi bir araya getirerek büyük ölçekli işlemler yapılabilir. HDSF (Hadoop Dağıtılmış Dosya Sistemi) ile verileri daha küçük bloklara böler ve kümelere dağıtır (Tole, 2013, s. 37). Hadoop, özünde MapReduce teknolojisini kullanır (Gürsakar, 2014, s. 217).

Özellikle Hadoop ve bulut alt yapıları teknolojik gelişmeler neticesinde veri depolama ve işleme maliyetinin düşmesi ve tüketici talepleri göz önüne alınarak yeni ürün ve hizmetler geliştirilmesine imkân sağlaması ile birlikte, artık büyük veriye yapılan yatırımlar şirketler için standart bir bütçe kalemi haline gelmiştir (Davenport, 2012).

## V. Büyük Veri Görselleştirme

2020 yılı başında dijital evrendeki bayt sayısı, gözlemlenebilir evrendeki yıldız sayısından 40 kat daha büyüktü (Vuleta, 2021). Bu miktardaki veriden üretilen enformasyondan en hızlı biçimde yararlanabilmek

için yeni tekniklere ihtiyaç duyulur. Bu teknikler “yeni veri görselleştirme dili” olarak ifade edilebilir (Techopedia, 2018).

Büyük veri görselleştirmesi hemen hemen her türden veriyi grafiksel bir biçimde özetlemeyi amaçlar. Veri görselleştirmesi “keşifsel” (exploration) ve/veya “açıklayıcı” (explanation) amaçlarla gerçekleştirilebilir (Iliinsky & Steele, 2011, s. 7). Söz konusu görseller, kullanıcılara mevcut durumu, ilişki yapısını veya analiz bulgularını grafik üzerinde değerlendirerek, karmaşık kavramların ve kavramlar arası ağların anlaşılmasını sağlayacak yeni modellerin geliştirilmesine imkân sağlarlar.

Teknolojinin görselleştirmeye doğrudan katkısı olup, özellikle son birkaç yıldır bu yönde önemli gelişmeler yaşanmaktadır. Birçok veri görselleştirme tekniği zamanla interaktif hale gelmiş, açık programlama ara yüzleri ve interaktif web tabanlı görselleştirme aracı kitlerinin yaygınlaşmasıyla üretimi kolaylaşmıştır. Ayrıca web tabanlı bağlantı teknolojisi sayesinde, verilere doğrudan ulaşılarak, görseller verideki değişime eş zamanlı adapte edilebilmektedir. Bu görselleştirme örneklerine, özellikle iş analizlerinde, hükümetlerin açık veri sistemlerinde ve medya infografiklerinde rastlanmaktadır (Fox & Hendler, 2011, s. 706).

## VI. Türkiye’de Konut Sektörüne İlişkin İnfografikler

Bu çalışmada, Türkiye’de emlak sektörüne ilişkin bir büyük veri görselleştirme uygulaması yapılmıştır. Çalışma kapsamında doğrudan değerlendirilemeyecek kadar çok sayıda ve yapılandırılmamış formda ilan içeren emlak sektörü ele alınarak, bireylerin konut tercihini kolaylaştıran, tüm alternatiflerin değerlendirilmesine imkân sağlayan, kullanıcı dostu infografikler üretilmiştir. Böylelikle, konut satın almayı planlayan bireylerin konutun yaşı, yüzölçümü, oda sayısı, fiyatı gibi önemli buldukları kriterleri dikkate alan, iller/ilçeler arasında karşılaştırma yapılmasına imkân veren ve hatta veri değişimine duyarlı söz konusu infografiklerden yararlanarak kendilerine en uygun tercihi yapabilmelerine imkân sağlanmıştır.

### VI.I. Çalışmanın Veri Seti ve Metodu

Çalışmada kullanılan veri seti Zingat Gayrimenkul Bilgi Sistemleri A.Ş.’den temin edilen 2018 yılı satılık ilan verilerini içermektedir. Çalışma kapsamında, İstanbul ve İstanbul’un 2018 yılında göç verdiği Ankara ve İzmir illeri ele alınarak analiz edilmiştir.

79.632 kayıt içeren veri setinde, konut satın alma düşüncesi olan bireylerin dikkate alacakları belli başlı kriterler 6 adet değişken biçiminde tanımlanmıştır: Bunlar; “oda+salon sayısı”, “bina yaşı”, “net alan”, “adres”, “fiyat” ve “para birimi”dir. Elde edilen veri setinde para birimi değişkeni “TL”, “Euro”, “Dolar” ve “Sterlin” olmak üzere dört kategoriden oluşmakta ve fiyat değişkeni bu para birimlerinin değerlerini içermektedir. Fiyat değişkeninin sahip olduğu bilginin standartlaştırılarak çalışmaya uygun hale getirilmesi amacıyla diğer para birimleri ilgili dönemdeki döviz kuru ele alınarak TL’ye çevrilmiştir.

Değişkenlerin kategori etiketleri Tablo 1’de verilmiştir. Adres değişkeni ise İstanbul, Ankara ve İzmir’e ilişkin ilçe ve mahalle bilgilerini içerdiği için tabloda yer almamıştır.

**Tablo 1.** Araştırmada Kullanılan Değişkenler ve Kategorileri

Fiyat	Net Alan	Oda-Salon	Bina Yaşı	Para Birimi
200bin altı	75 metrekare altı	1+0 ve 1+1	0	Euro
200-300bin	75-100	2+1	1-5	Dolar
300-500 bin	100-120	3+1	6-10	Türk Lirası
500bin-1milyon	120-150	4+1	11-20	Sterlin
1 milyon-5milyon	150-250	5+1 ve üstü	21-30	
5milyon +	250+		30+	

Merkezi eğilim ölçüsü olarak, değişkenlerin medyan değerleri kullanılmıştır. Bu sayede uç değer problemi de giderilmiştir.

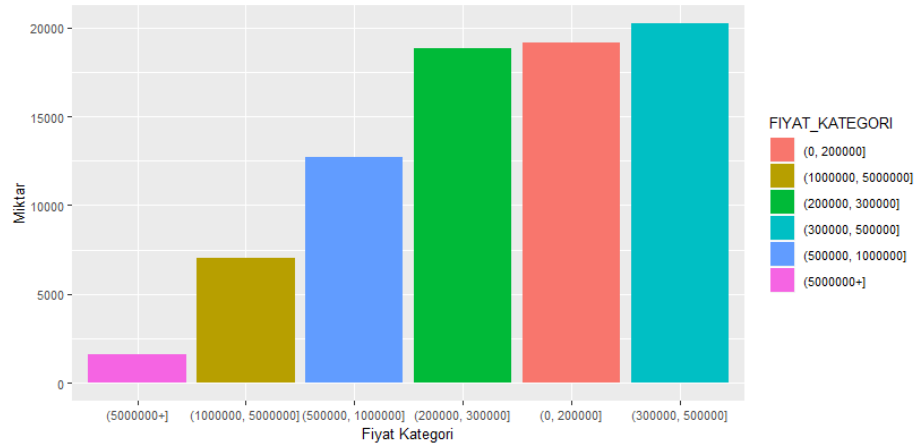
Analizler, R programlama dilinin 3.5.2 versiyonu aracılığıyla gerçekleştirilmiştir. R’ın “dplyr” ve “ggplot2” paketleri kullanılmıştır.

### VI.II. Değişkenlere İlişkin Tanımlayıcı İstatistikler ve Testler

Bu kısımda, ele alınan değişkenler bazında tanımlayıcı istatistikler sunulmuştur.

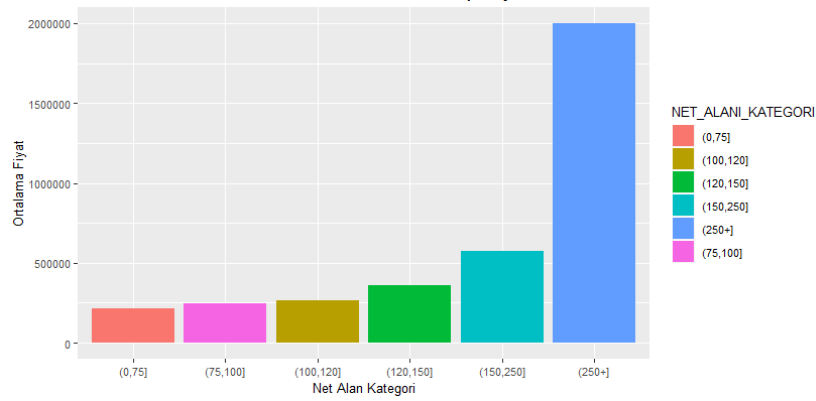
Veri setindeki konutların satış fiyatlarına göre dağılımı Şekil 2’de verilmiştir.



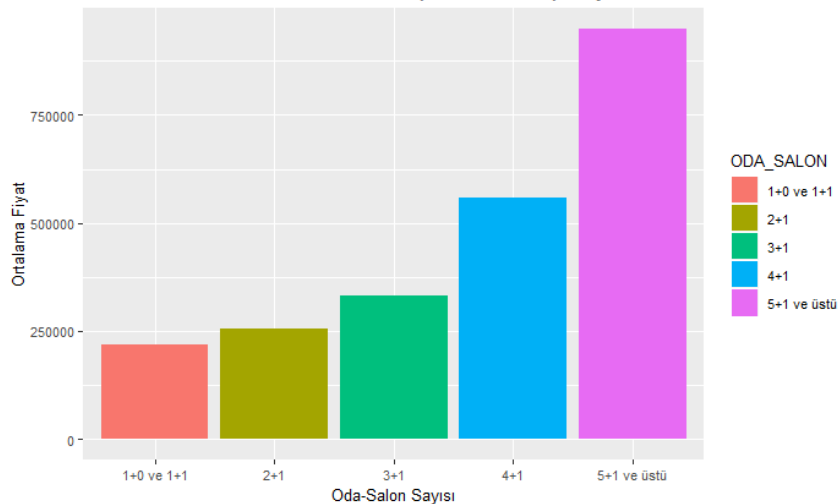
**Şekil 2:** Konutların Satış Fiyatlarına Göre Dağılımı

Şekil 2'ye göre satılık konutlardan fiyatları 5 milyon TL ve üzerinde olanlar veri setinde en az bulunan konutlardır. Veri setinde en fazla bulunan konutlar, 300.000-500.000 TL arası fiyatla ilana verilen konutlar olup satılık konutların yaklaşık %25'ini oluşturur. Miktar olarak takip eden diğer iki kategori ise 0-200.000 TL ve 200.000-300.000 TL arası konutları kapsamaktadır ve %24 ile yaklaşık aynı yüzdeye sahiptir.

Net alan değişkenininin satış fiyatına göre dağılımı Şekil 3'te verilmiştir.

**Şekil 3:** Konutların Net Alanlarına Göre Fiyat Dağılımı

Şekil 3'e göre konutların alan metrekaresi arttıkça fiyatları da artmaktadır. Ayrıca; ilk 3 kategorideki (0-75 metrekare, 75-100 metrekare, 100-120 metrekare) konutların fiyat medyanları birbirine yakın olup, 210-265 bin TL aralığında iken, 250 ve üzeri metrekareye sahip konutların fiyatları diğer kategorilere göre açık farkla yüksektir.

**Şekil 4:** Konutların Oda+Salon Sayısına Göre Fiyat Dağılımı

Şekil 4 incelendiğinde; ilk iki kategoride (1+0 ve 1+1 ile 2+1 oda/salon) önemli bir fiyat farkı bulunmazken, sonraki kategorilerde oda sayısı arttıkça fiyat farkının dikkat çekici biçimde arttığı görülebilir. 3+1 konutların fiyat medyanları 332.000 TL iken, 4+1 konutların 559.500 TL, 5+1 ve üstü konutların 950.000 TL olarak hesaplanmıştır.

Kullanılan değişkenler arasında ilişki olup olmadığının tespiti için yapılan Ki-kare Testi sonuçları Tablo 2’de özetlenmiştir.

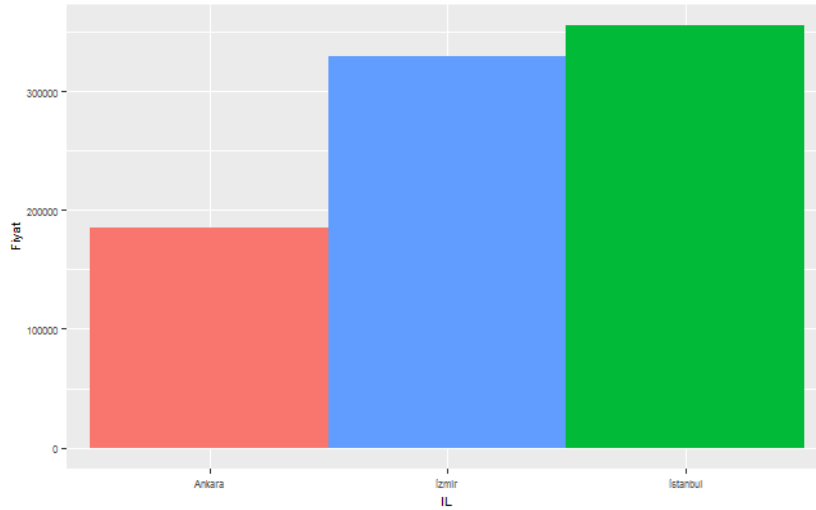
**Tablo 2.** Değişkenler Arası İlişki Testi Sonuçları

Kriterler	S.D.	Ki-kare İstatistiği	P
BINA_YASI-FİYAT_KATEGORI	25	2611.06	~0
BINA_YASI-NET_ALANI_KATEGORI	25	3694.38	~0
BINA_YASI-ODA_SALON	20	5031.196	~0
FIYAT_KATEGORI-NET_ALANI_KATEGORI	25	43183.41	~0
FIYAT_KATEGORI-ODA_SALON	20	25427.13	~0
NET_ALANI_KATEGORI-PARA_BIRIMI	15	12892.73	~0
NET_ALANI_KATEGORI-ODA_SALON	20	133506.1	~0

Tablo 2’de görüldüğü gibi tüm kriter çiftleri (binanın yaşı ile fiyatı, binanın yaşı ile net alanı, binanın yaşı ile oda salon sayısı, fiyatı ile net alanı, fiyatı ile oda salon sayısı, net alanı ile para birimi ve net alanı ile oda salon sayısı) arasında istatistiksel olarak anlamlı ilişki vardır.

İstanbul ili ile diğer iki ilin konut fiyatları açısından karşılaştırılmasının, tüketicilere temel bir bilgi sağlayacağı düşüncesi ile Şekil 4 oluşturulmuştur. Grafiğin çizimi R’in “dplyr” ve “ggplot2” paketleri aracılığıyla gerçekleştirilmiştir.

**Şekil 5:** İllere Göre Konut Satış Fiyatları (Medyan)



Şekil 5’e göre ele alınan 3 il arasında konut satış fiyatlarının en düşük olduğu il olan Ankara’da, satılık konutların fiyat medyanı yaklaşık 185.000 TL’dir.

Medyan fiyat İzmir’de 330.000 TL, İstanbul’da ise 355.000 TL olarak elde edilmiştir. Görüldüğü üzere iki ildeki satılık konutların fiyat ortalamaları (medyan) birbirine yakın olup; Ankara’daki satılık konutların fiyat ortalamaları ile aralarında ciddi bir fark tespit edilmiştir.

Mann Whitney U Testi aracılığıyla İstanbul ile Ankara ve İzmir’in konut fiyatları arasında anlamlı bir fark olup olmadığı incelenmiştir. Test sonuçları Tablo 3’te verilmiştir. Test sonuçlarına göre İstanbul ve diğer iller için konut fiyatlarının birbirinden anlamlı şekilde farklı olduğu tespit edilmiştir.

**Tablo 3.** Konut Fiyatları İçin Mann Whitney U Testi Sonucu

İLLER	Mann Whitney U Test Sonucu	
	W	P
İstanbul-Ankara	567320000	~0
İstanbul-İzmir	396790000	~0

### VI.III. Satılık Konut İlanlarının İller Bazında İnfografikler Üzerinde Karşılaştırılması

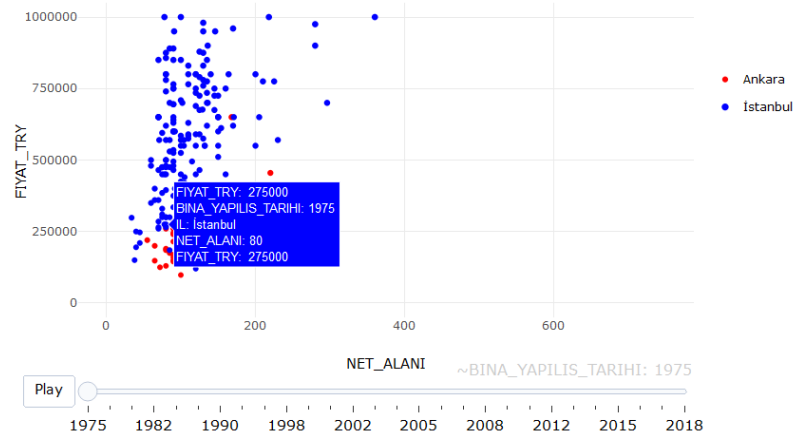
Çalışma kapsamında, satılık konut ilanları yönünden İstanbul ili ile Ankara ve İzmir'in karşılaştırılabilmesi amacıyla, tüm mümkün kriterler yönünden animasyon görselleri üretilmiştir. İlgili görseller, R'in "plotly" kütüphanesi aracılığıyla oluşturulan dinamik grafiklerdir. Daha net bir ifade ile, seçilen baz kriterin değerlerine bağlı olarak akan bir grafik üzerinde illere ait konut fiyatları, net alanlar vb. bilgiler görülebilmektedir. Her grafik için mavi renk İstanbul'u, kırmızı renk karşılaştırılan diğer ili temsil etmektedir. Ayrıca grafik durdurularak gözlem değerlerinin üzerine gelindiğinde renkli bir kutucuk üzerinde ilgili ilanın detayları görülebilmektedir. Böylece sadece animasyon grafikleri kullanılarak, seçilen kritere göre iller ya da ilçeler arası karşılaştırma yapılabilmesi ve birebir ilgilenilen konut ilanlarına ilişkin gözlem değerlerinin incelenmesi mümkün olmaktadır.

Bu çalışma kapsamında elde edilen grafikler dinamik yapıda olduğundan, aslen video biçiminde oynatılarak sunulmaktadır. Bu makalede örnek görüntü olarak, bina yapılış tarihi bazında bazı yıllarda video durdurularak, elde edilen görüntüler paylaşılacaktır.

#### VI.III.I. İstanbul ve Ankara İçin Satılık Konut İlanlarının Karşılaştırılması

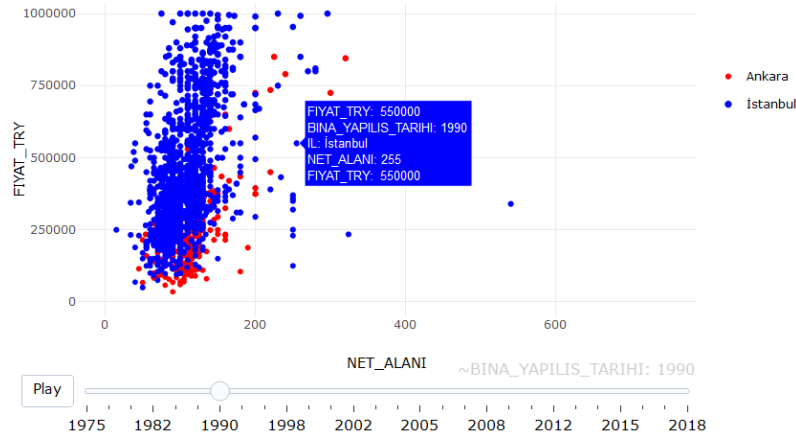
İstanbul ve Ankara illeri için satılık konutlar, net alan ve fiyat kriterleri dikkate alınarak bina yapılış tarihine göre oynatılan infografikte izlendiğinde, bina yapılış tarihi 1975 olan ilanlar için aşağıdaki görüntü elde edilmiştir.

Şekil 6: İstanbul ve Ankara Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 1975



Şekil 6 genel olarak incelendiğinde, 1975 yapım yılına sahip satıştaki konutların Ankara'da en fazla 200 metrekare büyüklüğünde olduğu görülmekte iken, İstanbul'da ise 400 metrekareye kadar konut bulunabilmektedir. Ayrıca, Ankara'daki konut fiyatlarının daha homojen olduğu anlaşılmaktadır. İstanbul'daki fiyatların Ankara'daki fiyatlarla kıyaslandığında daha yüksek olduğu da göze çarpmaktadır. Örneğin Ankara'da 90 metrekarelik bir konut 260.000 TL'den satılmak üzere ilan verilmişken aynı metrekareye sahip konutların fiyat aralığı 185.000-260.000 TL'dir. İstanbul'daki 80 metrekarelik bir konut için ise 275.000 TL istenmektedir. Yine 80 metrekarelik konutların İstanbul için fiyat aralığı ise 265.000-875.000 TL'dir.

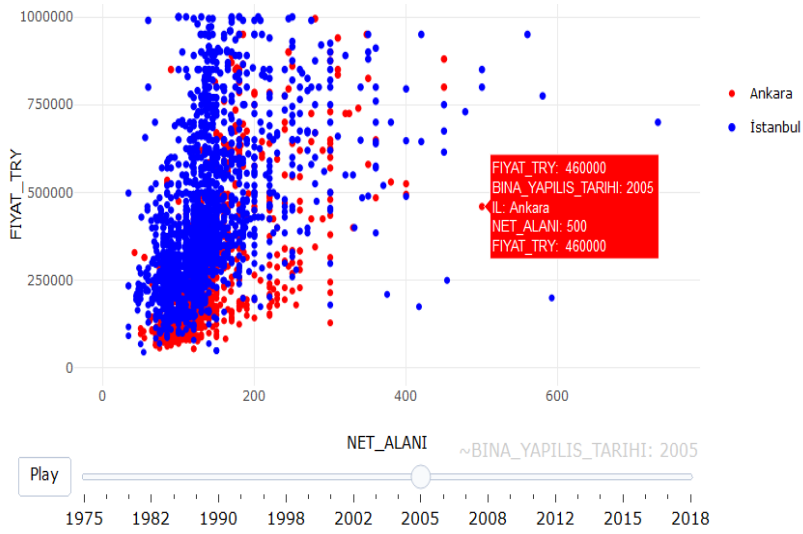
Şekil 7: İstanbul ve Ankara Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 1990



1990 yılında yapılmış olan konutlar incelendiğinde, 1975 yılına kıyasla özellikle İstanbul'daki konut sayısı artışı dikkat çekicidir. Ayrıca, Ankara'daki 200 metrekareden büyük konutların fiyatlarının birbirine yakın olduğu, İstanbul'da konut fiyatlarının heterojen olduğu anlaşılmaktadır. Örneğin yaklaşık 250 metrekare büyüklüğe sahip konutlar incelendiğinde Ankara'da uygun kriterde konut bulunmadığı, İstanbul'da ise 125.000-950.000 TL aralığında oldukça geniş bir fiyat aralığında konut bulunabileceği dikkat çekmektedir. 300 metrekarenin üzerinde konut nadirdir. Net alan açısından bakıldığında da genel olarak İstanbul'daki konutlar 60-160 metrekare aralığında yer alsa da her büyüklükte konut mevcuttur; Ankara'daki satılık konutların ise 70-125 metrekare aralığında yoğunlaştığı tespit edilmiştir.

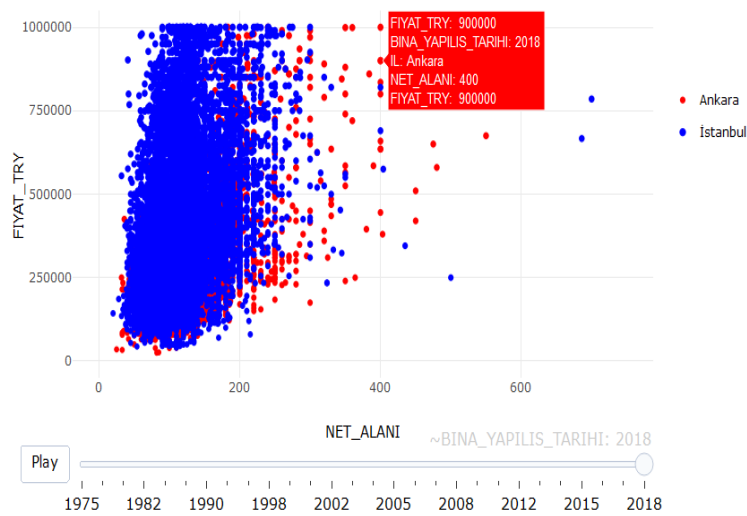
Grafik üzerinde iki ilin aynı alana sahip olan konutları incelendiğinde, tercih alternatifleri açıkça görülebilir. Örneğin, Ankara'daki 45 metrekarelik bir konutun 115.000 TL, İstanbul'da ise 40 metrekarelik bir konutun 189.000 TL olduğu, 45 metrekarelik bir diğer konutun ise 230.000 TL'ye satışa sunulduğu anlaşılmaktadır.

**Şekil 8:** İstanbul ve Ankara Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 2005



Yapım yılı 2005 olan satılık konutlara ilişkin Şekil 8 incelendiğinde, her iki ilin satılık konut sayılarında ve büyük metrekareli konut sayılarında artış olduğu görülmektedir. Aynı alan büyüklüğüne sahip konutlar için İstanbul'da fiyatların nispi olarak daha yüksek olduğu söylenebilir. Örneğin 500 metrekare büyüklüğündeki konutlar incelendiğinde İstanbul'daki fiyatların Ankara'daki benzer bir konutun fiyatına göre ortalama 375.000 TL daha yüksek olduğu dikkat çekici bir bulgudur.

**Şekil 9:** İstanbul ve Ankara Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 2018



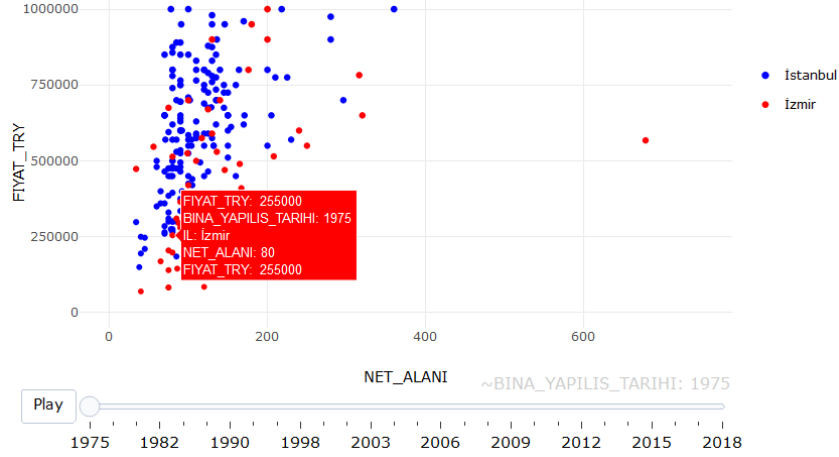
2018 tarihli konutlara bakıldığında, Şekil 9'da görüldüğü gibi İstanbul'un satılık konut sayısı açısından Ankara'ya üstünlüğü aşikârdır. Her iki ilde de konut fiyatlarının heterojenliği göze çarparken; Ankara'da 300 metrekare üzerindeki konutların fiyatları nispi olarak daha yüksektir. Örneğin Ankara'daki 400 metrekare büyüklüğündeki konutların fiyat aralığı 380.000-1.000.000 TL iken İstanbul'daki benzer

konutlar 575.000-820.000 TL aralığında fiyatlara sahiptir. Ayrıca yeni yapım tarihli konutlar genel olarak 75-250 metrekare aralığındadır. 400 metrekare üzerindeki konut sayısı her iki şehir için de oldukça az iken Ankara'da 600 metrekare üzeri satılık konut bulunmamaktadır.

### VI.III.II. İstanbul ve İzmir İçin Satılık Konutların Karşılaştırılması

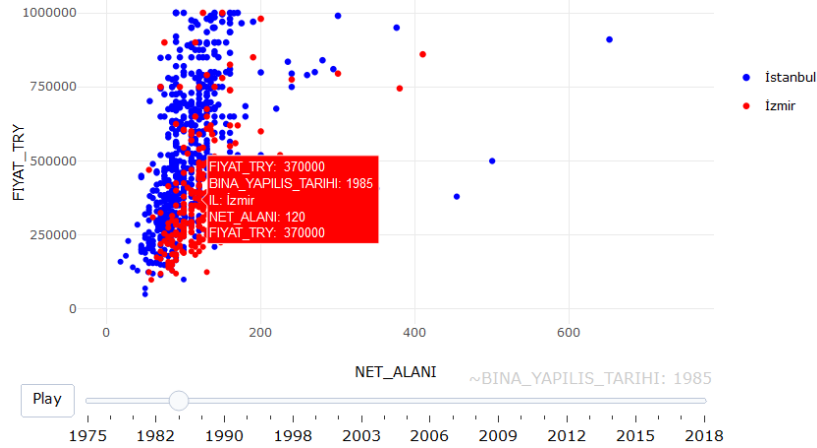
İstanbul ve İzmir illerinde de ilk olarak yine 1975 yılında yapılmış konutlar için karşılaştırma yapılmıştır.

**Şekil 10:** İstanbul ve İzmir Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 1975

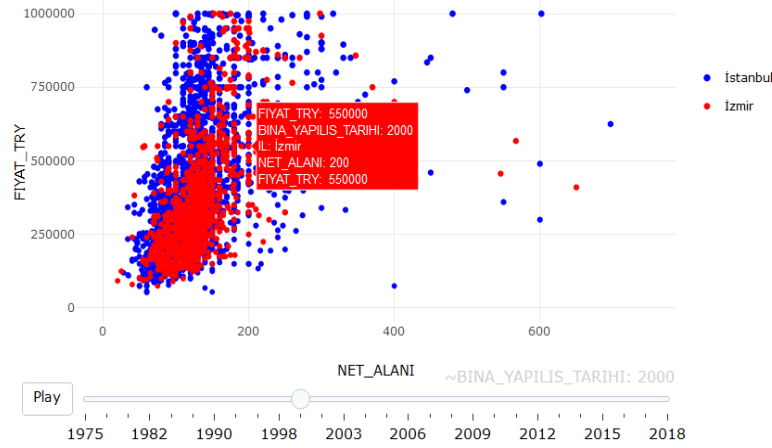


Veri setindeki en eski tarihli binalar karşılaştırıldığında, Şekil 10'da görüldüğü gibi İstanbul ile İzmir'deki konut yapılarının benzerliğinden bahsedilebilir. Ancak yine de İstanbul'daki konut sayısı İzmir'deki konut sayısından fazladır. 300 metrekarenin üzerinde konut ise nadirdir. Genel olarak, aynı fiyata sahip konutların İzmir'de daha büyük metrekareye sahip olduğu söylenebilir. Örneğin İzmir'de 80 metrekarelik bir konut için 255.000 TL, İstanbul'da 70 metrekarelik bir konut için 260.000 TL, 45 metrekarelik bir konut için ise 247.000 TL talep edilmektedir.

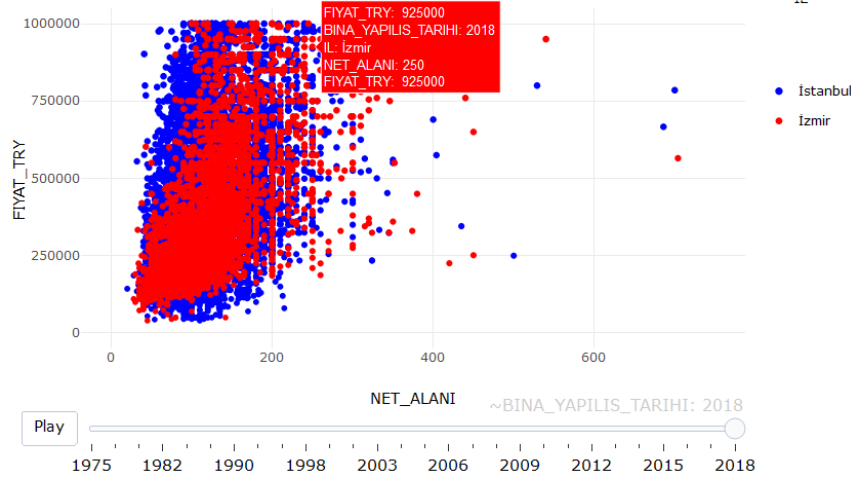
**Şekil 11:** İstanbul ve İzmir Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 1985



Şekil 11 incelendiğinde, yapım yılı 1985 olan konutların sayıca fazlaştığı, iki şehirdeki satılık konut fiyatlarının benzerliği göze çarpmaktadır. İzmir'deki konutlar genellikle 70-150 metrekare arasında olup 180.000-500.000 TL fiyat aralığında satışa sunulmuşken İstanbul'daki konutlar 50-160 metrekare arasındadır. Ancak fiyatların heterojen yapıları gereği İstanbul için belirgin bir fiyat aralığı verilememektedir. Her iki il için de 350'den büyük metrekareye sahip konutların sayısı ise azdır ve bu büyüklükteki konutlar için anlamlı bir farklılık yoktur.

**Şekil 12:** İstanbul ve İzmir Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 2000

Şekil 12'ye göre İzmir ve İstanbul'un genel olarak konut fiyatları ve yüzölçümleri yönünden benzer olduğu söylenebilir. Görüldüğü gibi, İstanbul veya İzmir'den konut almak isteyen kişilerin yapım yılı yakın tarihli olan konutlar için benzer ücretleri ödeyecekleri yorumu yapılabilir. Genel olarak, İzmir'de daha fazla satılık konut olduğu söylenebilir.

**Şekil 13:** İstanbul ve İzmir Satılık Konutlar (Net Alan-Fiyat) -Yapım Yılı 2018

Şekil 13 incelendiğinde hem İstanbul'da hem de İzmir'de 2018 yapım yılı olan her metrekarede ve her fiyat düzeyinde satılık konut bulunduğu söylenebilir. Genel olarak, İzmir'de daha fazla satılık konut olduğu anlaşılmaktadır. Ayrıca, hemen hemen her alan büyüklüğü için en pahalı evlerin İzmir'de olduğu görülmektedir. 500 metrekareden büyük konutlar için ise net bir ayırım yapılamamaktadır.

## VII. Sonuç

Bu çalışmada, İstanbul, Ankara ve İzmir illerine ait yüzbinlerce konut ilanını içeren bir veri setinden infografikler üretilmiştir. Bu infografikler yardımıyla bireylere, konut tercihinde dikkate aldıkları kriterler bazında, karşılaştırma yapmak istedikleri il ve ilçelerdeki ilanları veri setine ihtiyaç duymadan inceleyebilecekleri, ilgilerini çeken ilanların detaylarını görsel üzerinde görebilme şansı bulabilecekleri, kullanımı kolay bir araç sunulmaktadır. Söz konusu infografiklerin, verideki değişime adapte edilerek sürekli güncellenebilme imkânı da bulunmaktadır.

Genel bir değerlendirme yapılacak olursa, ele alınan üç ilde de bina yaşı azaldıkça konutların yüzölçümlerinin büyüdüğü görülmektedir. İstanbul'da satılık konut sayısının daha fazla olduğu, Ankara ve İzmir'e kıyasla ev fiyatlarının daha yüksek olduğu söylenebilir. Bunun yanı sıra, tüm yapım yılları için İstanbul'da Ankara'ya göre daha fazla satılık konut alternatifi varken, İzmir'de son yıllarda yapılan konut sayısının ve fiyatlarının İstanbul ile benzerlik gösterdiği anlaşılmaktadır.

Bu tür kullanıcı dostu infografiklerin emlak sektörü dışındaki sektörlerde de adapte edilmesi mümkündür. Böylece veriden bilgi üretmek; bilgiyi, sade ama kapsayıcı biçimde sunmak ve böylece büyük veriden belirsizlikleri azaltabilecek enformasyonlar elde etmek mümkündür.

---

Tüketicileri en az düzeyde rahatsız ederek ürün tanıtmanın, vatandaşları en az düzeyde yorarak hizmet vermenin, zamanı en etkin biçimde kullanarak iş yapmanın ve tabii ki veriden en üst düzeyde istifade edebilmenin sıkça vurgulandığı bu yıllarda, infografikler bireyler ve kurumlar için elzemdir. Günümüzde kurumlardan beklenen, verdikleri hizmetler ile ilgili bu tür infografikleri hitap ettikleri kitlenin hizmetine sunması ve bu hizmeti sunabilmek amacıyla da günümüz teknolojisinin ve rekabet koşullarının bir gereği olarak, büyük veri alt yapısı ve iş gücü tedarikine öncelikli olarak önem vermesidir.

Bir diğer yönden; veri görselleştirme uygulamalarının nicel ve nitel olarak artışı, karar vericilere fayda sağlamanın yanı sıra; istatistik, veri madenciliği, veri bilimi, veri analitiği alanlarının önemine ışık tutması yönüyle de önemlidir.

---

**Araştırmacıların Katkı Oranı Beyanı:** Yazarlar, makaleye katkılarının eşit olduğunu beyan etmişlerdir.

**Destek ve Teşekkür Beyanı:** Yazarlar, destek ve teşekkür beyanında bulunmamışlardır.

**Çatışma Beyanı:** Yazarlar, kendileri ve üçüncü taraflar açısından herhangi bir çıkar çatışmasının olmadığını beyan etmişlerdir.

**Etik Kurul Raporu Gerekliliği Beyanı:** Yazarlar, makalede etik kurul raporuna ihtiyaç duyulmadığını beyan etmişlerdir.

---

## Kaynakça

- Aktan, E. (2018), Büyük veri: uygulama alanları, analitiği ve güvenlik boyutu, *Bilgi Yönetimi Dergisi*, 1(1), 1-22. <https://dergipark.org.tr/tr/download/article-file/482194>
- Anuşlu, M., & Anuşlu, T. (2018). Application of big data visualization with agricultural value-added of all countries, *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 10(2), 94-98.
- Çelik, S., & Akdamar, E. (2018). Büyük veri ve veri görselleştirme, *Akademik Bakış Dergisi*, S.65, 253-264. [https://www.researchgate.net/publication/325604426\\_BUYUK\\_VERI\\_VE\\_VERI\\_GORSELLESTIRME](https://www.researchgate.net/publication/325604426_BUYUK_VERI_VE_VERI_GORSELLESTIRME)
- Davenport, T. (2014). *Big data@work*. İstanbul: Türk Hava Yolları Yayınları.
- Davenport, T. (2012). *Three big benefits of big data analytics*. [http://book.itep.ru/depositary/big\\_data/AST-0147176\\_Three\\_Big\\_Benefits\\_of\\_Big\\_Data\\_Analytics.pdf](http://book.itep.ru/depositary/big_data/AST-0147176_Three_Big_Benefits_of_Big_Data_Analytics.pdf)
- Demirtaş, B., & Argan, M. (2015). Büyük veri ve pazarlamadaki dönüşüm: kuramsal bir yaklaşım. *Pazarlama ve Pazarlama Araştırmaları Dergisi*, S.15, 1-21. <https://dergipark.org.tr/en/download/article-file/1672065>
- Dialani, P. (2020, September 24). *Top 10 Big Data Trends Of 2020*. <https://www.analyticsinsight.net/top-10-big-data-trends-2020/>
- Du, D., Li, A., & Zhang, L. (2014). Survey on the applications of big data in chinese real estate enterprise. *Procedia Computer Science*, 30, 24-33. <https://www.sciencedirect.com/science/article/pii/S1877050914005547>
- Eken, S. (2020). Büyük verinin interaktif görselleştirilmesi: tableau üzerine öğrenci deneyimleri, *Avrupa Bilim ve Teknoloji Dergisi*, S.18, 262,271. <https://dergipark.org.tr/tr/pub/ejosat/issue/52599/659823>
- Erbay, H, ve Kör, H., (2016, 3-6 Ekim), *Büyük veri ve büyük verinin analizi*, Uluslararası Bilim ve Teknoloji Konferansı, Ankara.
- Erkurt, E. (2020). *Büyük veri görselleştirme ve türkiye'de konut sektörüne ilişkin infografikler*, [Yayımlanmamış doktora tezi]. Marmara Üniversitesi.
- Feinleib, D. (2013). *Big Data Demystified: How Big Data is Changing The Way We Live, Love and Learn*. Big Data Group.
- Fox, P., & Hendler, J. (2011, February). Changing the equation on scientific data visualization. *Science*, V.331, Issue:6018, 705-709.
- Gast, A. (2020, January 9). *Why We Need to Talk About Big Data*. <https://www.weforum.org/agenda/2020/01/privacy-in-a-world-of-ai-and-big-data/>.
- Gürsakal, N. (2014). *Büyük veri (Genişletilmiş 2. Baskı)*. Dora Yayınevi.
- Hewage, T. N., Halgamuge, M. N., Syed, A., & Ekici, G. (2018, February). Review: big data techniques of google, amazon, facebook and twitter. *Journal of Communications*, 13(2), 94-100.
- Iliinsky, N. & Steele, J. (2011). *Designing Data Visualizations*. O'Reilly.
- Khan, N., Alsaquer, M., Shah, H., Badsha, G., Abbasi, A.A., ve Salehian, S. (2018). *The 10 vs, issues and challenges of big data*. ICBDE '18: Proceedings of the 2018 International Conference on Big Data and Education, 52-56.
- Lavalle, A., Teruel, M., Maté, A., & Trujillo, J. (2020). Improving sustainability of smart cities through visualization techniques for big data from IoT devices. *Sustainability*, 12(14),1-17.
- Loberto, M., Luciani, A., & Pangallo, M. (2018, April). The potential of big housing data: an application to the italian real-estate market. *Banca D'Italia, No.1171*. [https://www.bancaditalia.it/pubblicazioni/temi-discussione/2018/2018-1171/en\\_tema\\_1171.pdf?language\\_id=1](https://www.bancaditalia.it/pubblicazioni/temi-discussione/2018/2018-1171/en_tema_1171.pdf?language_id=1)
- Lohr, S. (2012, August 11). *How big data became so big*. [http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html?\\_r=0](http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html?_r=0)
- Lohr, S. (2013, February 1). *The origins of big data: an etymological detective story*. [https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/?\\_r=0](https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/?_r=0)



- Maçãs, C., Cruz, P., Amaro, H., Polisciuci E., Carvalho, T., Santos, F., ve Machado, P. (2015). *Time-series application on big data visualization of consumption in supermarkets*. 6th International Conference on Information Visualization Theory and Applications. [https://www.researchgate.net/publication/280039332\\_Time-Series\\_Application\\_on\\_Big\\_Data\\_Visualization\\_of\\_Consumption\\_in\\_Supermarkets](https://www.researchgate.net/publication/280039332_Time-Series_Application_on_Big_Data_Visualization_of_Consumption_in_Supermarkets)
- Padmavalli, M. (2016, November-December). Big data: emerging challenges of big data and techniques for handling. *IOSR Journal of Computer Engineering (IOSR-JCE)*, V.18, I.6, 13-18.
- Press, G. (2020, January 6). *6 predictions about data in 2020 and the coming decade*. <https://www.forbes.com/sites/gilpress/2020/01/06/6-predictions-about-data-in-2020-and-the-coming-decade/?sh=122605c74fc3>
- Reinsel, D., Gantz, J., & Rydning, J. (2018, November). The digitization of the world from edge to core. *IDC White Paper*. <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>
- Sakyi, K. T. (2016). Big data: understanding big data. <https://arxiv.org/ftp/arxiv/papers/1601/1601.04602.pdf>
- Soto, A. J., Ryan, C., Silva, F. P., Das, T., Wolkowicz, J., Milios, E. E., ve Brooks, S. (2018). *Data quality challenges in twitter content analysis for informing policy making in health care*. Proceedings of the 51st Hawaii International Conference on System Sciences, 760-769. <https://pdfs.semanticscholar.org/009b/ea25c5e4712fa6e9b266380511f807891bde.pdf>
- Su, K., Liu, C., & Wang, Y. (2018). A principle of designing infographic for visualization representation of tourism social big data. *Journal of Ambient Intelligence and Humanized Computing*. <https://link.springer.com/article/10.1007/s12652-018-1104-9>
- Sun, D., Du, Y., Xu, W., Zuo, M., Zhang, C., & Zhou, J. (2015). Combining online news articles and web search to predict the fluctuation of real estate market in big data context. *Pacific Asia Journal of the Association for Information Systems*, 6(4), 19-37.
- Techopedia, (2018). *Data Visulization*. <https://www.techopedia.com/definition/30180/data-visualization>
- Tole, A. A. (2013). Big data challenges. *Database Systems Journal*, V. IV, No. 3, 31-40.
- Türk Kenesi & Dünya Telekomünikasyon ve Bilgi Toplumu Günü Etkinliđi. (2017, 17 Mayıs). <https://www.timeturk.com/turk-kenesi-dunya-telekomunikasyon-ve-bilgi-toplumu-gunu-etkinligi/haber-637200>
- Uđur, N. G., & Akbıyık, A. (2020). Impacts of COVID-19 on global tourism industry: a cross-regional comparison. *Tourism Management Perspectives*, 36, 1-13.
- Vuleta, B. (2021, January 28). *How much data is created every day? [27 Staggering Stats]*. <https://seedscientific.com/how-much-data-is-created-every-day/>
- Watson, H. J. (2014). Tutorial: big data analytics: concepts, technologies, and applications. *Communications of the Association for Information Systems*, V.34, A.65, 1244-1268.
- Yılmazel, S., Afşar, A., & Yılmazel, Ö. (2017). Türkiye’de satışı bulunan konutların il ve bölgeler bazında dağılımının büyük veri teknolojisi ile incelenmesi. *Sakarya İktisat Dergisi*, 6(3), 1-21.
- Zhang, P., Zhou, J., & Zhang, T. (2017). Quantifying and visualizing jobs-housing balance with big data: a case study of shanghai. *Cities*, V:66, 10-22.