



HACETTEPE
ÜNİVERSİTESİ
İSTATİSTİK BÖLÜMÜ

İST155 İSTATİSTİK GİRİŞ I

DERS 2 (TEMEL TANIM ve KAVRAMLAR)

**Ders sorumluları: Prof. Dr. Serpil AKTAŞ ALTUNAY (01 Şubesi)
Doç. Dr. Ayten YİĞİTER (02 Şubesi)**

İSTATİSTİK NEDİR?

İstatistik denildiğinde ilk akla gelen belli bir konuda periyodik olarak toplanmış bilgilerdir.

Örneğin,

The screenshot shows the homepage of the TÜİK (Turkish Statistical Institute) Data Portal. The header features the TÜİK logo, a bell curve icon, and the text "İSTATİSTİK VERİ PORTALI". Below the header is a search bar with the placeholder "Ürünlerde Ara" and a button labeled "Ara". Underneath the search bar are popular search terms: "nüfus", "ihraçat", "tüfe", and "göç". The main navigation menu includes "Ana Sayfa", "Coğrafi İstatistik Portalı", and "Ulusal Veri Sayfası". On the right side of the header, there are social media icons for Facebook, Twitter, YouTube, and a "T" icon. The main content area is titled "İSTATİSTİK KONULARI" and displays eight categories with corresponding icons: Adalet ve Seçim, Bilim, Teknoloji ve Bilgi Toplumu, Çevre ve Enerji, Dış Ticaret, Eğitim, Kültür, Spor ve Turizm, Ekonomik Güven, Enflasyon ve Fiyat, Gelir, Yaşam, Tüketim ve Yoksulluk. Below this, there are four more categories: İnşaat ve Konut, İstihdam, İşsizlik ve Ücret, Nüfus ve Demografi, Sağlık ve Sosyal Koruma, Sanayi, Tarım, Ticaret ve Hizmet, and Ulaştırma ve Haberleşme. At the bottom center is a link to "Ulusal Hesaplar".

- Demografik istatistikler
- Eğitim istatistikleri
- Turizm istatistikleri
- Çevre istatistikleri
- İşgücü istatistikleri
- Milli Eğitim istatistikleri
- Sağlık İstatistikleri vb...

<https://data.tuik.gov.tr/>, 29.09.2024

İstatistik genel olarak çok sayıda sebebe bağlı ve değişkenlik gösteren olayları inceleyen bir bilim dalıdır. İstatistiğin konusu **yığın** olaylardır denir.

Genç nüfusta işsizlik oranı %26,1, istihdam oranı %28,2 oldu

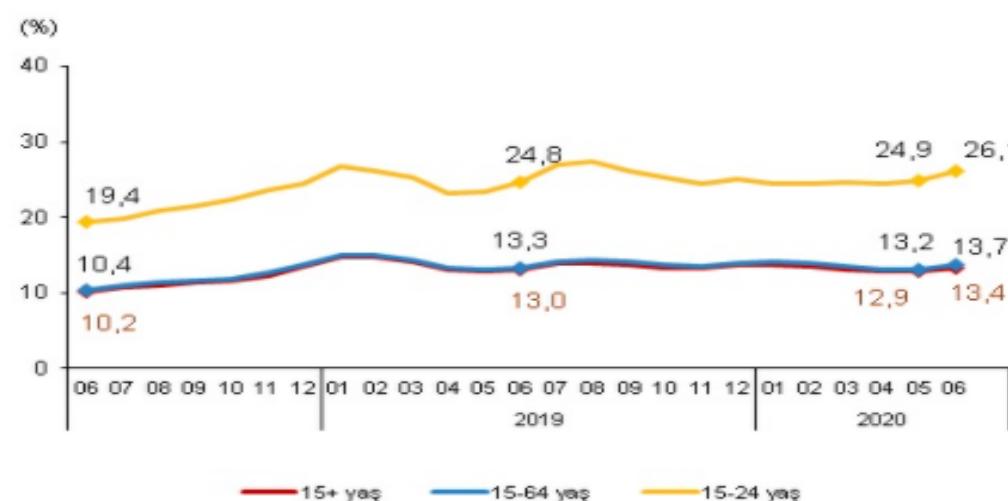
15-24 yaş grubunu kapsayan genç nüfusta işsizlik oranı bir önceki yılın aynı dönemine göre 1,3 puanlık artışla %26,1, istihdam oranı ise 6,0 puan azalarak %28,2 oldu. Aynı dönemde işgücüne katılma oranı 7,4 puanlık azalışla %38,1 seviyesinde gerçekleşti. Ne eğitimde ne de istihdamda olanların oranı ise bir önceki yılın aynı dönemine göre 3,1 puanlık artışla %29,3 seviyesinde gerçekleşti.

Mevsim etkisinden arındırılmış temel işgücü göstergeleri, 15-24 yaş, Haziran 2019, Haziran 2020

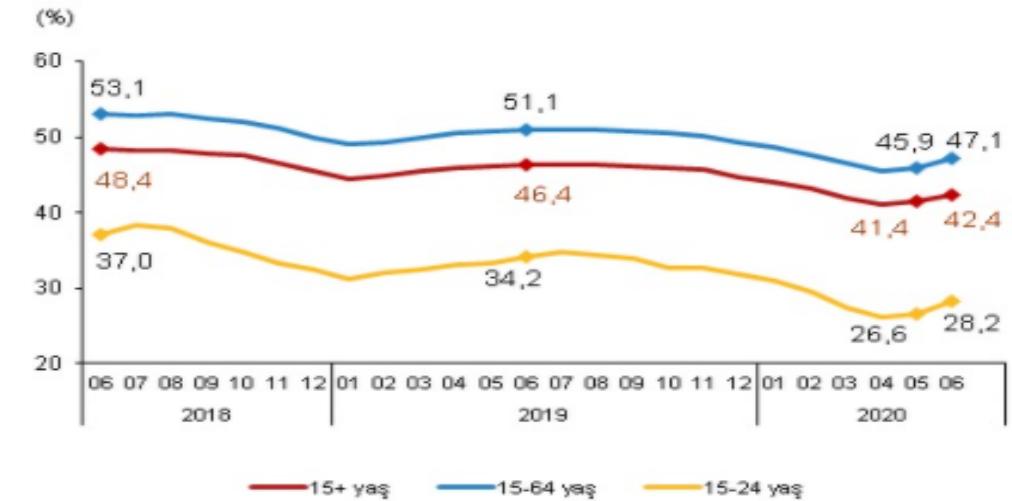
| | Toplam | | | Erkek | | | Kadın | | |
|--|--------|------|------|-------|------|------|-------|------|------|
| | 2019 | 2020 | Fark | 2019 | 2020 | Fark | 2019 | 2020 | Fark |
| İşgücüne katılma oranı | 45,5 | 38,1 | -7,4 | 57,0 | 49,3 | -7,7 | 33,6 | 26,5 | -7,1 |
| İstihdam oranı | 34,2 | 28,2 | -6,0 | 44,6 | 37,7 | -6,9 | 23,6 | 18,3 | -5,3 |
| İssizlik oranı | 24,8 | 26,1 | 1,3 | 21,9 | 23,6 | 1,7 | 29,8 | 31,1 | 1,3 |
| Ne eğitimde ne istihdamda olanlar ⁽¹⁾ | 26,2 | 29,3 | 3,1 | 19,0 | 22,2 | 3,2 | 33,6 | 36,6 | 3,0 |

(1) Çalışmayan ve eğitimde (örgün ve yaygın) olmayan gençlerin, toplam genç nüfus içindeki oranıdır.

İssizlik oranı, Haziran 2018-Haziran 2020



İstihdam oranı, Haziran 2018-Haziran 2020



Mevsim etkisinden arındırılmış işsizlik oranı %14,3, istihdam oranı %41,6 oldu

Mevsim etkisinden arındırılmış işsizlik oranı bir önceki döneme göre 0,2 puan artarak %14,3 oldu. İşsiz sayısı bir önceki döneme göre 136 bin kişi artarak 4 milyon 330 bin kişi olarak gerçekleşti.

Mevsim etkisinden arındırılmış istihdam oranı bir önceki döneme göre 0,6 puan artarak %41,6 oldu. İstihdam edilenlerin sayısı 431 bin kişi artarak 26 milyon 36 bin kişi olarak tahmin edildi.

İşgücüne katılma oranı %53,3 olarak gerçekleşti

İşgücü 2023 yılında bir önceki yıla göre 562 bin kişi artarak 34 milyon 896 bin kişi, işgücüne katılma oranı ise 0,2 puan artarak %53,3 olarak gerçekleşti. İşgücüne katılma oranı erkeklerde %71,2, kadınlarda ise %35,8 oldu.

Genç nüfusta işsizlik oranı %17,4 oldu

15-24 yaş grubunu kapsayan genç nüfusta işsizlik oranı 2023 yılında bir önceki yıla göre 2,0 puan azalarak %17,4 oldu. Bu yaş grubunda işsizlik oranı; erkeklerde %14,3, kadınlarda ise %23,2 olarak tahmin edildi.

Temel işgücü göstergeleri, 15+ yaş, 2022, 2023

| | 2023 | | | 2022 | | | Bir önceki yıla göre fark | | |
|---|--------|--------|--------|--------|--------|--------|---------------------------|-------|-------|
| | Toplam | Erkek | Kadın | Toplam | Erkek | Kadın | Toplam | Erkek | Kadın |
| (Bin kişi) | | | | | | | | | |
| 15 ve daha yukarı yaşındaki nüfus | 65 425 | 32 383 | 33 042 | 64 679 | 32 006 | 32 673 | 746 | 377 | 369 |
| İşgücü | 34 896 | 23 052 | 11 844 | 34 334 | 22 862 | 11 473 | 562 | 190 | 371 |
| İstihdam | 31 632 | 21 286 | 10 346 | 30 752 | 20 818 | 9 935 | 880 | 468 | 411 |
| İşsiz | 3 264 | 1 766 | 1 498 | 3 582 | 2 044 | 1 538 | - 318 | - 278 | - 40 |
| İşgücüne dahil olmayanlar | 30 528 | 9 330 | 21 198 | 30 345 | 9 144 | 21 200 | 183 | 186 | - 2 |
| (%) | | | | | | | | | |
| İşgücüne katılma oranı | 53,3 | 71,2 | 35,8 | 53,1 | 71,4 | 35,1 | 0,2 | -0,2 | 0,7 |
| İstihdam oranı | 48,3 | 65,7 | 31,3 | 47,5 | 65,0 | 30,4 | 0,8 | 0,7 | 0,9 |
| İşsizlik oranı | 9,4 | 7,7 | 12,6 | 10,4 | 8,9 | 13,4 | -1,0 | -1,2 | -0,8 |
| Genç nüfusta işsizlik oranı (15-24 yaş) | 17,4 | 14,3 | 23,2 | 19,4 | 16,4 | 25,2 | -2,0 | -2,1 | -2,0 |

Tablodaki rakamlar yuvarlanmadan dolayı toplamı vermeyebilir.

İstihdamın %57,6'sı hizmet sektöründe yer aldı

İstihdam edilenlerin %14,8'i tarım, %21,2'si sanayi, %6,3'ü inşaat, %57,6'sı ise hizmet sektöründe yer aldı. Bir önceki yıl ile karşılaştırıldığında hizmet sektörünün istihdam edilenler içindeki payı 1,1 puan, inşaat sektörünün payı 0,3 puan artarken, tarım sektörünün payı 1,0 puan, sanayi sektörünün payı 0,5 puan azaldı.

İstihdamın sektörel dağılımı, (%), 2022, 2023



<https://data.tuik.gov.tr/Bulton/Index?p=%C4%B0%C5%9Fg%C3%BCc%C3%BC%C-%C4%B0statistikleri-2023-53521&dil=1>, 29.09.2024

Örneğin boşanma bir **yığın** olaydır. Bu kapsamda alınan tüm bireylerin ortak özelliği, boşanmış olmalarıdır. Ancak bireylerin cinsiyeti, eğitim durumu, boşanma sebepleri, yaşıları vb. özellikleri bireyden bireye farklılık gösterir.

İşsizlik olayı bir **yığın** olaydır. İşsiz tanımına uygun tüm bireylerin, cinsiyet, yaş, eğitim, meslek, daha önce istihdam edilmiş edilmediği vb. özellikleri yine bireyden bireye farklılık gösterir.

İssiz: Referans dönemi içinde istihdam halinde olmayan (kâr karşılığı, yevmiyeli, ücretli ya da ücretsiz olarak hiç bir işte çalışmamış ve böyle bir iş ile bağlantısı da olmayan) kişilerden iş aramak için son 4 hafta içinde iş arama kanallarından en az birini kullanmış ve 2 hafta içinde işbaşı yapabilecek durumda olan 15 ve daha yukarı yaştaki fertler işsiz nüfusa dahildirler (TÜİK, İşgücü İstatistikleri, Haziran 2020, Sayı: 33790, 10 Eylül 2020, Saat: 10:00, Haber Bülteni).

İstatistik

Tanımlayıcı istatistik (Descriptive Statistics)

Verilerin derlenmesi, tablolar ve grafikler halinde özetlenmesi, konum ve değişim ölçülerinin hesaplanması

Çıkarsamalı istatistik (Inferential Statistics)

Kitle parametrelerine ilişkin güven aralıklarının elde edilmesi, hipotezlerin testi, neden sonuç ilişkilerin belirlenmesi, istatistiksel modellerin kurulması, geleceğe yönelik tahmin yapılması



İST155 İSTATİSTİĞE GİRİŞ I



İST156 İSTATİSTİĞE GİRİŞ II

Bazı Tanımlar ve Kavramlar

Araştırma: İlgiilenilen bir konuda sorunların belirlenmesi, incelenmesi, çözüm yollarının planlanması, uygulamaya konulması ve sonuçların değerlendirilmesine yönelik yapılan bir çalışmadır. Bir araştırmanın çeşitli aşamaları vardır:

1. Araştırma konusunun saptanması:

Araştırmacı konu hakkında yeterince bilgi sahibi olması gereklidir. Zaman, maliyet vb kısıtlara göre, araştırmanın konusu sınırlı olmalıdır. Araştırmanın sonucunda elde edilen bilgiler doğu biçimde değerlendirilmelidir.

2. Araştırmanın Planlanması:

Araştırmanın amacı iyi belirlenmeli, konu ile ilgili yeni kavramlar tanıtılmalı, araştırmanın nasıl yapılacağı ve hangi yöntemin uygulanacağı açıklanmalıdır.

3. Araştırmanın uygulanması:

Araştırmanın uygulama aşamasında araştırma birimleri saptanmalıdır. Araştırma birimi araştırmanın konusuna göre değişir. Örneğin bir şehirdeki hane halkı gelirleri ile ilgili bir araştırma yapılyorsa araştırma birimi hane halkı, bir bakterinin üreme hızı ile ilgili bir araştırmada, araştırma birimi her bir deney tüpü olacaktır. Araştırma konusuna göre, ilgiilenilen sorunu tespit etmeye yönelik hangi sorular (bir anket çalışmasında) sorulacak ya da hangi ölçümler alınacak (bir deney çalışmasında) belirlenir. Sistematik olarak elde edilecek bilgiler kayıt edilir.

4. Araştırmanın sunulması:

Araştırmanın uygulama aşamasında elde edilen veriler analiz edilir ve sonuçları rapor haline getirilir.

Araştırmanın aşamaları:

1. Araştırma konusunun saptanması
2. Araştırma probleminin ve amaçlarının tanımlanması
3. Araştırma modelinin ve hipotezlerinin belirlenmesi
4. Araştırma yönteminin tasarıımı
5. Kitle(anakütle) ve örneklemın belirlenmesi
6. Veri toplama araçlarının geliştirilmesi
7. Saha çalışması ve veri toplama
8. Veri hazırlama ve analizi
9. Araştırma raporunun hazırlanması ve sunulması

Geriye Dönük (Retrospektif) Araştırma:

Araştırmmanın yönü zaman boyutunda geriye yönelik olup “ne oldu?” sorusuna yanıt aranır. Bu araştırma yöntemi bir anlamda “sonuçtan başlayarak nedeni bulmaya yönelik çalışma yöntemi” olarak da tanımlanabilir.

Örneğin, akciğer kanserine yakalananlar ile sağlıklı bireylerden (kanser ve benzeri hastalığı olmayan) oluşan gruplara geçmişte sigara içip içmedikleri, ne sıklıkta içtikleri, vb. sorular yöneltilir ve her iki grup için sigara içenlerin yüzdesi saptanarak aralarında anlamlı bir fark olup olmadığı istatistiksel yöntemlerle saptanır.

Kesitsel Araştırma:

Bu tür çalışmalarla, şu anda “ne oluyor” sorusuna yanıt aranır. Bu amaçla, kısa bir zaman diliminde bir grup denekten veriler elde edilir. Zaman diliminde bir noktaya odaklanan çalışmalar olduğu için bazen prevalans çalışmaları olarak da adlandırılır. Örneğin, bir toplumda bir hastalığın ya da olayın bir zaman diliminde bulunma sıklığının (prevalansının) belirlenmesi kesitsel bir çalışmadır.

İleriye Dönük (Prospektif) Araştırma:

Bu araştırmalarda “ne olacak?” sorusuna yanıt aranır. İleriye dönük araştırmalar, neden-sonuç ilişkisini ortaya koymada geriye dönük araştırmalardan daha güçlündür. Araştırmada ilgili değişkenlerin elde edebilmesi ve ölçülmesi avantaj olmakla birlikte, bu araştırmaların uzun süre gerektirmesi maliyetin artması ve güncellliğinin yitirmesi bakımından dezavantajdır. Gözlem süresinin uzun olması kayıp gözlemlerin ortaya çıkmasına neden olabilmektedir.

Örneğin, tam gün okula giden çocukların yarım gün okula giden çocukların ileriye doğru izlenip gribi yakalanma sıklıklarının incelenmesi.

https://biyoistatistik.hacettepe.edu.tr/lisans/SBE/arastirma_turleri.pdf, 06.10.2023

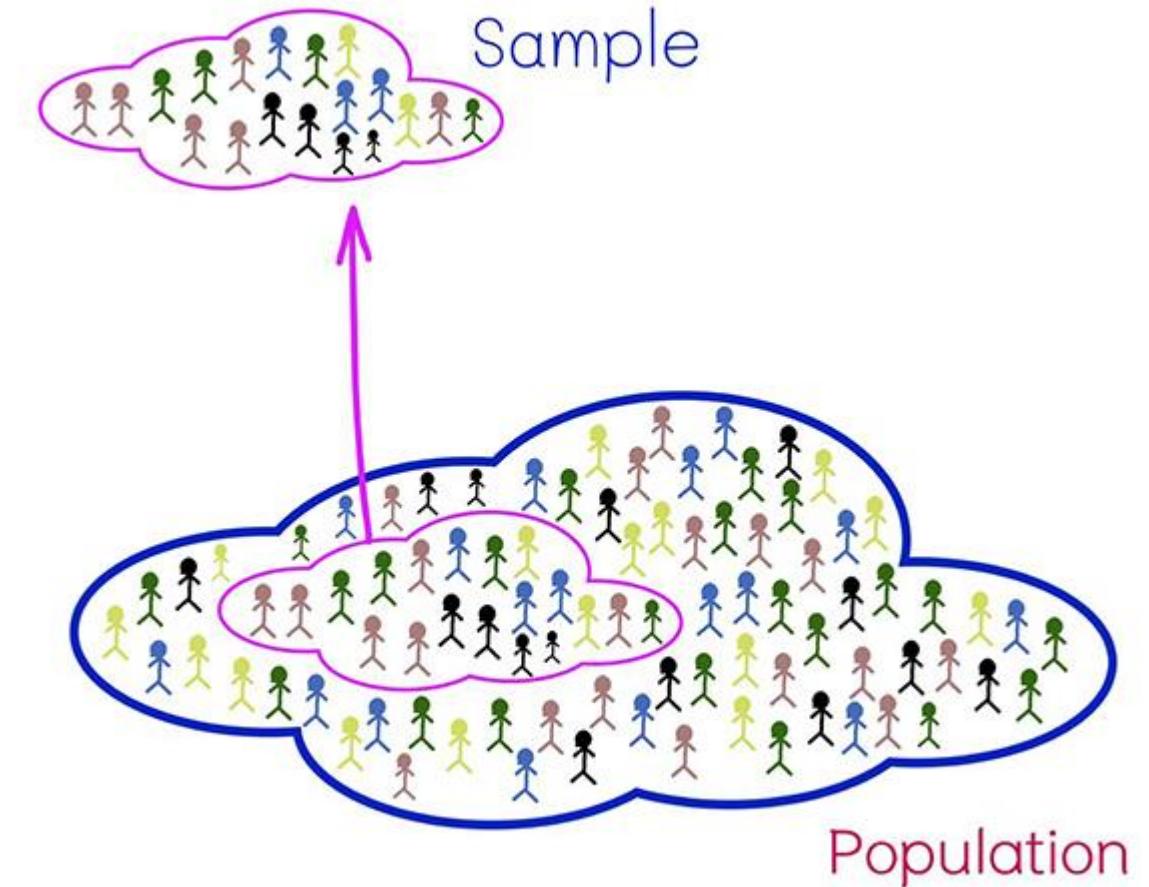
KİTLE VE ÖRNEKLEM KAVRAMLARI

Pastanın tadını
anlamak için
pastanın
tamamını mı
yemeliyiz ?



Kitle (population): Araştırma kapsamına giren aynı özelliğe sahip birimlerin tümüne kitle denir. Örneğin H. Ü. okuyan öğrenciler üzerinde yapılan bir araştırmada kitle, H. Ü. okuyan tüm öğrenciler, Türkiye'de gençlerin bağımlı madde kullanımları ile ilgili bir araştırmada, kitle 15-25 yaşları arasındaki gençlerin tümüdür.

Ancak çoğu araştırmada kitle birimlerin tümüne ulaşmak olası değildir (zaman, maliyet, personel, ulaşım vb.). Bu nedenle araştırma, kitleden çekilen bir grup birimin oluşturduğu topluluk üzerinden yürütülür. Bu gruba örneklem adı verilir.



<http://analytics-magazine.org/survey-sampling/>, 09.10.2020

Örneklem (sample): Bir kitleden belirli kurallara göre (ki bu kurallara örnekleme seçme yöntemleri adı verilir) seçilen aynı özelliğe sahip birimlerin oluşturduğu topluluğa örneklem (ya da örnek) denir. Örneklem kitle birimlerinin gözlenen bir alt kümesidir.

Kitleden onu temsil eden örneklemi nasıl seçeceğiz?

Örnekleme (sampling): Kitleden örneklem seçme işlemine «Örnekleme» denir.

Örnekleme bize ne kazandırır ?

Zaman, işgücü ve maliyet yönünden tasarruf sağlar.

Örneklemden elde edilen sonuçlar kitleyi temsil edecek kadar güvenilirdir.

ÖRNEKLEME YÖNTEMLERİ

Olasılığa
bağlı

Basit rasgele
örneklemme

Tabakalı
örneklemme

Sistematik
örneklemme

Küme
örneklemesi

Olasılığa bağlı
olmayan

Kota
Örneklemesi

Kartopu
örneklemesi

Amaca bağlı
örneklemme

Olasılığa Bağlı Olmayan Örnekleme Yöntemleri

Amaca bağlı Örnekleme: N tane birim içinden n tanesinin araştırıcıın kendi görüş ve tecrübelerine dayanarak seçilmesidir.

Kota Örneklemesi: Kitle nitelik bakımından farklı kısımlardan oluşuyorsa, örneklemenin temsil gücünü artırmak için her kısımdan önemi ile orantılı olarak keyfi örnekleme ile birim seçme işleminin yapıldığı örnekleme yöntemidir. Kitle yaş cinsiyet, öğrenim düzeyi gibi değişkenlere göre tabakalara ayrılır. Bu tabakalar homojendir. Her tabakayı temsil edecek örneklem sayısı belirlenir. Her tabakadan belirlenen kota kadar kişiler seçilir.

Kartopu Örneklemesi: Kartopu örneklemde öncelikle kitleye ait birimlerden birisi ile temas kurulur. Temas kurulan birimin yardımıyla ikinci birime, ikinci birimin yardımıyla üçüncü birime gidilir. Bu şekilde, sanki bir kartopunun büyümlesi gibi örneklem büyüklüğü genişler.

Olasılığa Bağlı Örnekleme Yöntemleri

Basit Rasgele Örnekleme: N tane birimden yerine koyarak (eşit olasılıkla), ya da yerine koymadan n tane ($n < N$) deneğin seçilmesidir. Örneklem seçme işlemi için rasgele sayılar tablosundan yararlanılır.

Rasgele sayılar tablosu her yöne okunabilecek şekilde düzenlenmiştir. N birimden n tanesinin seçilme işlemi için, önce rasgele sayılar tablosundan bir başlangıç rakamı seçilir. N sayısının basamak sayısı r ise, seçilen başlangıç rakamı ilk basamak olacak şekilde sağa doğru, aynı satırda r basamaklı sayı oluşturulur. Bu ilk rasgele sayı olmak üzere, bu sayının bittiği yerden başlayarak sağa doğru r basamaklı ikinci sayı oluşturulur. Bu işlem n tane birime ulaşınca dek sürer. n sayısını aşan sayılar dikkate alınmaz.

Rasgele Rakamlar Tablosu (RRT) tarihte ilk olarak 1927 yılında Trippett tarafından hazırlanmıştır. Trippett 41600 rakam içeren 26 sayfalık "Random Sampling Numbers" isimli bir kitap yayınlamıştır.

| Part of page XIV of Tippett's random sample numbers | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|----|----|----|----|----|----|----|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|
| 44 | 03 | 22 | 46 | 10 | 00 | 17 | 00 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | |
| 7 | 8 | 1 | 6 | 6 | 5 | 7 | 2 | 0 | 8 | 0 | 2 | 6 | 3 | 1 | 4 | 0 | 7 | 0 | 2 | 4 | 3 | 6 | 9 | 9 | 7 | 2 | 8 | 0 | 1 | 9 | 8 | |
| 3 | 2 | 0 | 4 | 9 | 2 | 4 | 3 | 4 | 9 | 3 | 5 | 8 | 2 | 0 | 0 | 3 | 6 | 2 | 3 | 4 | 8 | 6 | 9 | 6 | 9 | 3 | 8 | 7 | 4 | 8 | 1 | |
| 2 | 9 | 7 | 6 | 3 | 4 | 1 | 3 | 2 | 8 | 4 | 1 | 4 | 2 | 4 | 1 | 2 | 4 | 2 | 4 | 1 | 9 | 8 | 5 | 9 | 3 | 1 | 3 | 2 | 3 | 2 | 2 | |
| 8 | 3 | 0 | 3 | 9 | 8 | 2 | 2 | 5 | 8 | 8 | 8 | 2 | 4 | 1 | 0 | 1 | 1 | 5 | 8 | 2 | 7 | 2 | 9 | 6 | 4 | 4 | 3 | 2 | 9 | 4 | 3 | |
| 5 | 5 | 5 | 6 | 8 | 5 | 2 | 6 | 6 | 1 | 6 | 6 | 8 | 2 | 3 | 1 | 2 | 4 | 3 | 8 | 8 | 4 | 5 | 5 | 4 | 6 | 1 | 8 | 4 | 4 | 4 | 5 | |
| 2 | 6 | 3 | 5 | 7 | 9 | 0 | 0 | 3 | 3 | 7 | 0 | 9 | 1 | 6 | 0 | 1 | 6 | 2 | 0 | 3 | 8 | 8 | 2 | 7 | 7 | 5 | 7 | 4 | 9 | 5 | 0 | |
| 3 | 2 | 1 | 1 | 4 | 9 | 1 | 9 | 7 | 3 | 0 | 6 | 4 | 9 | 1 | 6 | 7 | 6 | 7 | 7 | 8 | 7 | 3 | 3 | 9 | 9 | 7 | 4 | 6 | 7 | 3 | 2 | |
| 2 | 7 | 4 | 8 | 6 | 1 | 9 | 8 | 7 | 1 | 6 | 4 | 4 | 1 | 4 | 8 | 7 | 0 | 8 | 6 | 2 | 8 | 8 | 8 | 8 | 5 | 1 | 9 | 1 | 6 | 2 | 0 | |
| 7 | 4 | 7 | 7 | 0 | 1 | 1 | 1 | 1 | 6 | 3 | 0 | 2 | 4 | 0 | 4 | 2 | 9 | 7 | 9 | 7 | 9 | 9 | 1 | 9 | 6 | 8 | 3 | 5 | 1 | 2 | 5 | |
| 5 | 3 | 7 | 9 | 7 | 0 | 7 | 6 | 2 | 6 | 9 | 4 | 2 | 9 | 2 | 7 | 4 | 3 | 9 | 9 | 5 | 5 | 1 | 9 | 8 | 1 | 0 | 6 | 8 | 5 | 0 | 1 | |
| 9 | 2 | 6 | 4 | 4 | 6 | 0 | 7 | 2 | 0 | 1 | 2 | 3 | 9 | 2 | 0 | 7 | 7 | 6 | 6 | 3 | 8 | 1 | 7 | 3 | 2 | 5 | 6 | 1 | 6 | 4 | 0 | |
| 5 | 8 | 5 | 8 | 7 | 7 | 6 | 6 | 3 | 1 | 7 | 0 | 0 | 5 | 0 | 0 | 2 | 5 | 9 | 3 | 0 | 5 | 4 | 5 | 5 | 3 | 7 | 0 | 7 | 8 | 1 | 4 | |
| 2 | 8 | 8 | 9 | 6 | 6 | 2 | 8 | 6 | 7 | 5 | 7 | 8 | 2 | 3 | 1 | 1 | 5 | 8 | 9 | 0 | 0 | 6 | 2 | 0 | 0 | 4 | 7 | 3 | 8 | 1 | 5 | |
| 5 | 1 | 3 | 1 | 8 | 1 | 8 | 6 | 3 | 7 | 0 | 9 | 4 | 5 | 2 | 1 | 6 | 6 | 6 | 5 | 5 | 3 | 2 | 5 | 5 | 3 | 8 | 3 | 2 | 7 | 0 | 2 | |
| 9 | 0 | 5 | 5 | 7 | 1 | 9 | 6 | 2 | 1 | 7 | 2 | 3 | 2 | 0 | 7 | 1 | 1 | 1 | 4 | 1 | 3 | 8 | 4 | 4 | 3 | 5 | 9 | 4 | 4 | 8 | 8 | |
| 7 | 9 | 0 | 0 | 5 | 8 | 7 | 0 | 2 | 6 | 0 | 6 | 8 | 8 | 1 | 3 | 5 | 5 | 0 | 9 | 4 | 3 | 2 | 4 | 0 | 0 | 3 | 0 | 4 | 7 | 5 | 0 | |
| 3 | 6 | 9 | 3 | 9 | 2 | 1 | 2 | 0 | 5 | 5 | 7 | 7 | 3 | 6 | 9 | 0 | 7 | 1 | 6 | 2 | 9 | 5 | 6 | 8 | 1 | 3 | 1 | 2 | 9 | 4 | 3 | 8 |
| 0 | 3 | 8 | 0 | 3 | 3 | 3 | 8 | 0 | 1 | 3 | 8 | 4 | 5 | 6 | 0 | 4 | 2 | 0 | 3 | 6 | 4 | 9 | 6 | 3 | 8 | 0 | 6 | 0 | 3 | 4 | 7 | |
| 0 | 2 | 4 | 6 | 4 | 4 | 6 | 9 | 9 | 7 | 1 | 9 | 8 | 3 | 1 | 6 | 1 | 2 | 8 | 5 | 0 | 3 | 5 | 7 | 2 | 3 | 8 | 0 | 2 | 3 | 9 | 0 | |
| 7 | 2 | 6 | 6 | 0 | 0 | 8 | 1 | 6 | 8 | 9 | 7 | 2 | 8 | 5 | 1 | 4 | 6 | 6 | 6 | 0 | 6 | 2 | 0 | 4 | 5 | 9 | 6 | 3 | 4 | 0 | 0 | |
| 9 | 3 | 1 | 2 | 4 | 7 | 7 | 9 | 5 | 7 | 3 | 7 | 8 | 9 | 1 | 8 | 4 | 5 | 5 | 0 | 3 | 9 | 0 | 4 | 5 | 5 | 7 | 3 | 0 | 2 | 2 | 9 | |
| 6 | 1 | 1 | 1 | 6 | 0 | 9 | 8 | 0 | 9 | 6 | 5 | 7 | 3 | 5 | 2 | 6 | 8 | 4 | 7 | 3 | 0 | 3 | 4 | 4 | 9 | 7 | 7 | 0 | 3 | 7 | 7 | 0 |
| 2 | 3 | 1 | 0 | 4 | 4 | 7 | 6 | 9 | 1 | 4 | 8 | 0 | 6 | 7 | 9 | 2 | 6 | 6 | 2 | 2 | 0 | 6 | 2 | 0 | 5 | 2 | 2 | 9 | 2 | 3 | 4 | |
| 9 | 8 | 2 | 6 | 8 | 8 | 5 | 7 | 8 | 6 | 7 | 5 | 6 | 6 | 4 | 2 | 5 | 4 | 7 | 1 | 8 | 8 | 2 | 0 | 4 | 3 | 0 | 8 | 2 | 1 | 0 | 5 | |
| 6 | 7 | 0 | 3 | 8 | 2 | 4 | 8 | 6 | 0 | 6 | 4 | 6 | 9 | 6 | 2 | 0 | 0 | 5 | 3 | 8 | 1 | 8 | 8 | 6 | 4 | 9 | 4 | 4 | 5 | 0 | 9 | |

<https://docplayer.biz.tr/1809758-Rasgelelik-ragele-rakamlar-tablosu.html>, 09.10.2020

N=100, n=10 olsun.

Bir başlangıç noktası belirlenir. Sonra 3 basamaklı sayılar sağa doğru seçilir. 64, 86, 50, 25, 78, 47, 94, ...

| Part of page XIV of Tippett's random sample numbers | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 7 | 8 | 1 | 6 | 6 | 5 | 7 | 2 | 0 | 8 |
| 3 | 2 | 0 | 4 | 9 | 2 | 4 | 3 | 5 | 0 |
| 2 | 9 | 7 | 6 | 3 | 4 | 1 | 3 | 2 | 8 |
| 8 | 3 | 0 | 3 | 9 | 8 | 2 | 2 | 2 | 8 |
| 5 | 5 | 5 | 6 | 8 | 5 | 2 | 6 | 6 | 6 |
| 2 | 6 | 3 | 5 | 7 | 9 | 0 | 0 | 3 | 7 |
| 3 | 2 | 1 | 1 | 4 | 9 | 1 | 9 | 7 | 0 |
| 2 | 7 | 4 | 8 | 7 | 1 | 6 | 4 | 8 | 6 |
| 7 | 4 | 7 | 7 | 0 | 1 | 1 | 1 | 1 | 1 |
| 5 | 3 | 7 | 9 | 7 | 0 | 7 | 6 | 6 | 6 |
| 9 | 2 | 6 | 4 | 4 | 6 | 0 | 7 | 2 | 1 |
| 5 | 8 | 5 | 8 | 7 | 7 | 6 | 6 | 3 | 0 |
| 2 | 8 | 8 | 9 | 6 | 6 | 2 | 8 | 1 | 2 |
| 5 | 1 | 3 | 1 | 8 | 1 | 8 | 6 | 3 | 1 |
| 9 | 0 | 5 | 5 | 7 | 1 | 9 | 6 | 2 | 1 |
| 7 | 9 | 0 | 0 | 5 | 8 | 7 | 0 | 0 | 3 |
| 3 | 6 | 9 | 3 | 9 | 2 | 1 | 2 | 0 | 6 |
| 0 | 3 | 8 | 0 | 3 | 3 | 3 | 8 | 3 | 8 |
| 0 | 2 | 4 | 6 | 4 | 4 | 6 | 9 | 7 | 1 |
| 7 | 2 | 6 | 6 | 0 | 0 | 8 | 1 | 9 | 7 |
| 9 | 3 | 1 | 2 | 4 | 7 | 7 | 9 | 5 | 7 |
| 6 | 1 | 1 | 1 | 6 | 0 | 9 | 8 | 0 | 9 |
| 2 | 3 | 1 | 0 | 4 | 4 | 7 | 6 | 9 | 1 |
| 9 | 8 | 2 | 6 | 8 | 8 | 5 | 7 | 7 | 5 |
| 6 | 7 | 0 | 3 | 8 | 2 | 4 | 8 | 6 | 4 |

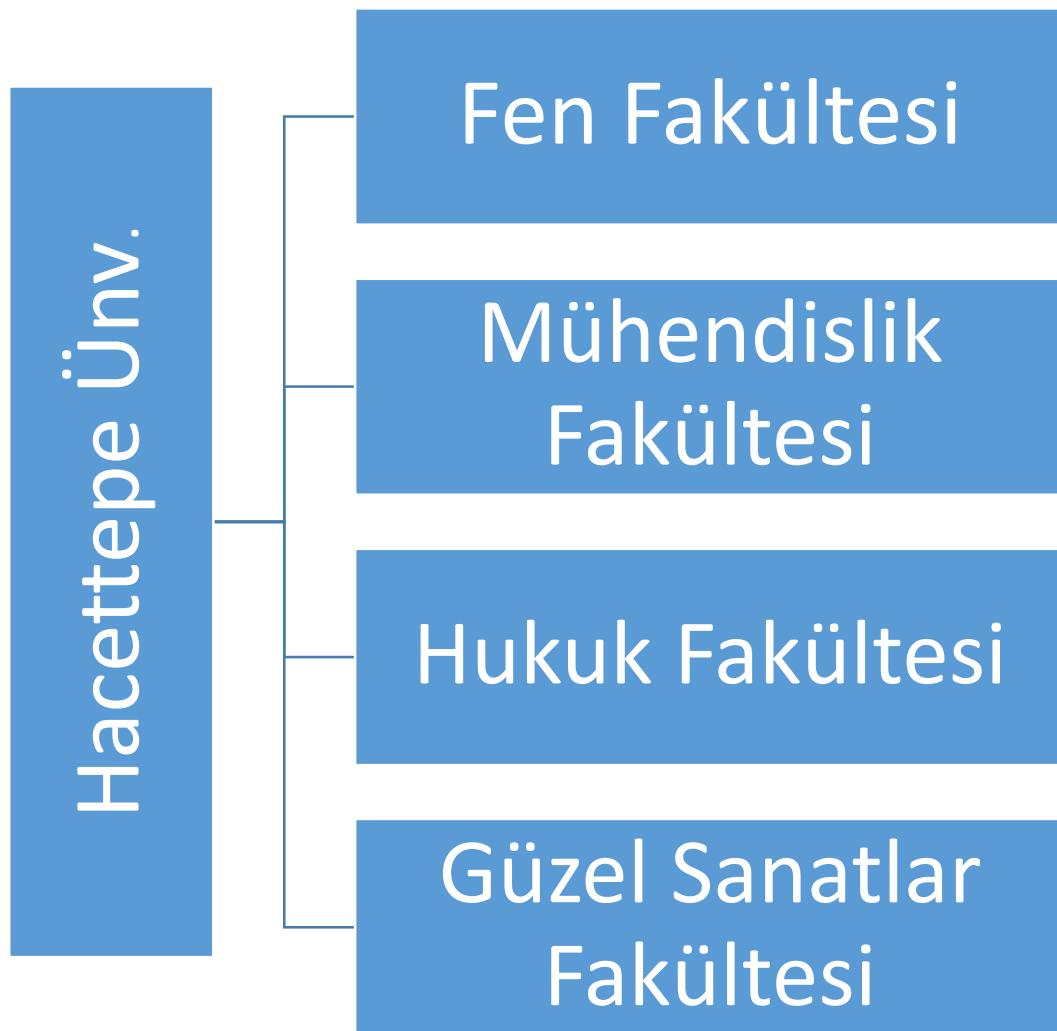
Sistematik Örnekleme: Bu örneklemde kitle birimleri bir seri olarak numaralandırılır. Numaralandırılmış seri içinde bir başlangıç sayısı keyfi olarak seçilir. Bu sayıdan başlayarak örneklem büyülüğüne göre, sistematik olarak seçim yapılır. $N=100$ olsun ve 10 birimlik örneklem seçelim. $k=100/10=10$ olur. Başlangıç noktası 1 ile 10 arasından 8 olsun. 8'e 10 ve katları eklenerek 10 tane örnekleme ait numara çekilmiş olur.

1,2,3,4,5,6,7,**8**,9,10,11,12,13,14,15,16,17,**18**,19,20,21,22,23,24,25,26,27,**28**,29,30,31,3
2,33,34,35,36,37,**38**,39,40,41,42,43,44,45,46,...

Seçilecek birimler: **8, 18, 28, 38, 48, 58, 68, 78, 88, 98**

Tabakalı Örnekleme: Kitle tekdüze(homojen) yapıda değilse, ancak kendi içinde tekdüze alt parçalara ayrılabilirse tabakalı örnekleme kullanılır. Örneklem büyülüğü tabakaların büyülüğüne orantılı olarak dağıtıılır.

| Tabakalar | Tabaka büyülüğü | Örneklem büyülüğü |
|-----------|-----------------|-------------------|
| 200 | $200/1000=0.20$ | $0.20*100=20$ |
| 300 | $300/1000=0.30$ | $0.30*100=30$ |
| 500 | $500/1000=0.50$ | $0.50*100=50$ |



Küme örneklemesi: Örneklemeye birimi tek kişi ya da aile değil bir grup, demet ya da kümedir. Araştırma yapılacak bireyler geniş bir alana dağılmış durumda iseler, basit rastgele ve tabakalı rasgele örneklemeye yöntemiyle yapılan seçim ile örneklemeye çıkan bireylere ulaşmak kolay olmayabilir. Böyle bir durumda küme örneklemeye yöntemi uygulama kolaylığı sağlar. Bir bankanın geniş şubelerinde çalışanların memnuniyeti üzerine yapılan bir çalışmada, her bir şube küme alınıp, kümeler rasgele seçilir.

ÖLÇEKLER (SCALES)

Ölçme ve Ölçek (scale)

Araştırma konusunu ile ilgili olarak sayısal değerler elde etme işlemine ölçme (measure), ölçme sonunda elde edilen bilgiye ölçüm (measurement) denir.

Ölçme işleminde önemli olan ölçektir.



1) Nominal (sınıflayıcı) ölçü (Nominal scale)

Bu ölçüde temel özellik eşitliktir. Ölçme işlemi eşit olup olmamaya göre yapılır. Kişileri cinsiyete göre sınıflama, kitapları içeriklerine göre sınıflama, hayvanları üreme türlerine göre sınıflama vb.

2) Ordinal (sıralanabilir) ölçü (Ordinal scale)

Bu ölçüde temel özellik sıralamadır. Sıralamada üstünlük üstün olmama ya da eşitlik ölçüsü bir arada kullanılır. Üniversite akademik personelin unvanlarına, kişilerin eğitim durumlarına göre sıralama vb.

3) Aralıklı ölçü (Interval scale)

Bu ölçünün temel özelliği, bağıl (göreceli) bir ölçü biriminin kullanılmasıdır. Bu ölçünün bir başlangıç bir de bitiş noktası vardır. Sınavlarda kullandığımız 10 ya da 100 lük ölçü, sıcaklık ölçüleri, takvim ölçüleri aralık ölçüsüne örnek olarak verilebilir.

4) Oranlı ölçü (Ratio scale)

Bu ölçünün temel özelliği bir başlangıç notasına sahip olmasıdır. Oranlı ölçüde standart ölçü birimleri kullanılır. Başlangıç noktasının değişmez olması nedeniyle karşılaştırma yapılabilir. Uzaklık ölçüsü, hacim ölçüsü, ağırlık ölçüsü vb. oranlı ölçüye örnek olarak verilebilir.

Sınıflayıcı Ölçek :

Cinsiyet
Saç rengi
Tuttuğu takım
Hastalık türü

Sıralayıcı Ölçek :

Memnuniyet ölçen sorular
Eğitim düzeyi
Harf notları

Aralıklı Ölçek:

Celcius
Fahrenayt
Notlar(10luk ya da 100lük)

Oransal Ölçek:

Hacim
Ağırlık
Uzunluk

«Aylık doğalgaz tüketimi hangisine örnektir?»

Değişken (Variable): Araştırma kapsamında yer alan birimlerde ya da gözlemlerde birimden birime ya da gözlemden gözleme değişen nicelik ya da niteliktir.

Sürekli Değişken (Continuous variable): Değişkenler ölçüлerek ya da sıralanarak elde edilir. İki ölçüm arası sonsuz sayıda noktaya bölünebilir. Aralık biçiminde ifade edilebilirler. Örneğin; boy uzunluğu, kilo, fiyat, gelir vb.

Kesikli Değişken (Discrete variable): Ölçümler 0, 1, 2 gibi kesin değerler alır. Ara değerler söz konusu değildir. Ailedeki çocuk sayısı, bir fabrikada çalışan işçi sayısı, vb.

Cinsiyet, yazı-tura, ülke kodları, şirket türü vb. nitel değişkenler, genellikle tamsayı değerlerle sayısallaştırılır.

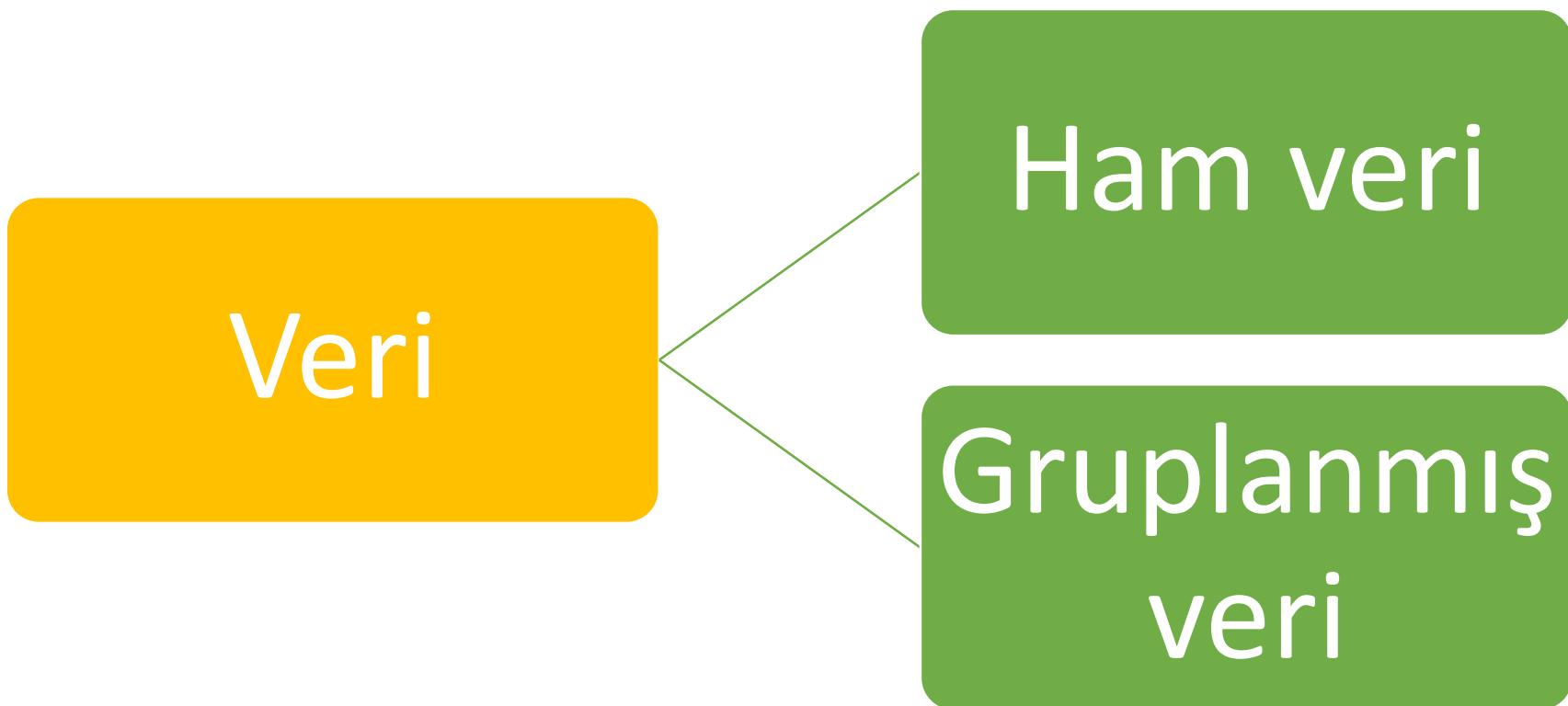
Veri (Data): Araştırma kapsamında incelenenek değişkenlerin ölçülmüş değerlerin tümüdür. Bir başka ifadeyle araştırmada gözlemlerden elde edilen sayısal olan ya da sayısal olmayan sonuçlara veri denir. Verilerin elde ediliş süreci önemlidir. Örneğin araştırmamızda ilgilendığımız değişkene ilişkin ölçüm bir deney ile elde edilecekse, deney koşullarının aynı tutulması gereklidir.

Ham veri (raw data): Verilerin düzenlenmemiş haline ham veri denir.

İstatistikte verilerin elde edilmesi (toplanması) en temel bölümünü oluşturur.

Veriler genellikle dört farklı yolla toplanır:

- 1) Yayınlanmış kaynaklardan ya da kayıtlardan: Türkiye İstatistik Kurumu'nun yayınlarından, hastane kayıtlarından vb.
- 2) Tasarlanmış bir deney ya da denemeden: Yiyecekleri buzdolabında saklanmanın besin değerini koruması ile ilgili bir deneyden, yeni geliştirilen bir ilaçın deney hayvanlarında test edilmesinden vb.
- 3) Anket sonuçlarından: Bir reklam şirketinin belirli bir markanın reklamı ile ilgili yaptığı kamuoyu araştırmalarından vb.
- 4) Gözlem sonuçlarından: Bir bölgede yaşayan hayvanların beslenmesi ile ilgili bir araştırmadan, bir karayolunda trafik yoğunluğu ile ilgili bir araştırmadan vb.



Yaş: 21, 22, 35, 55, 67, 87, 32, 36 (Ham veri)

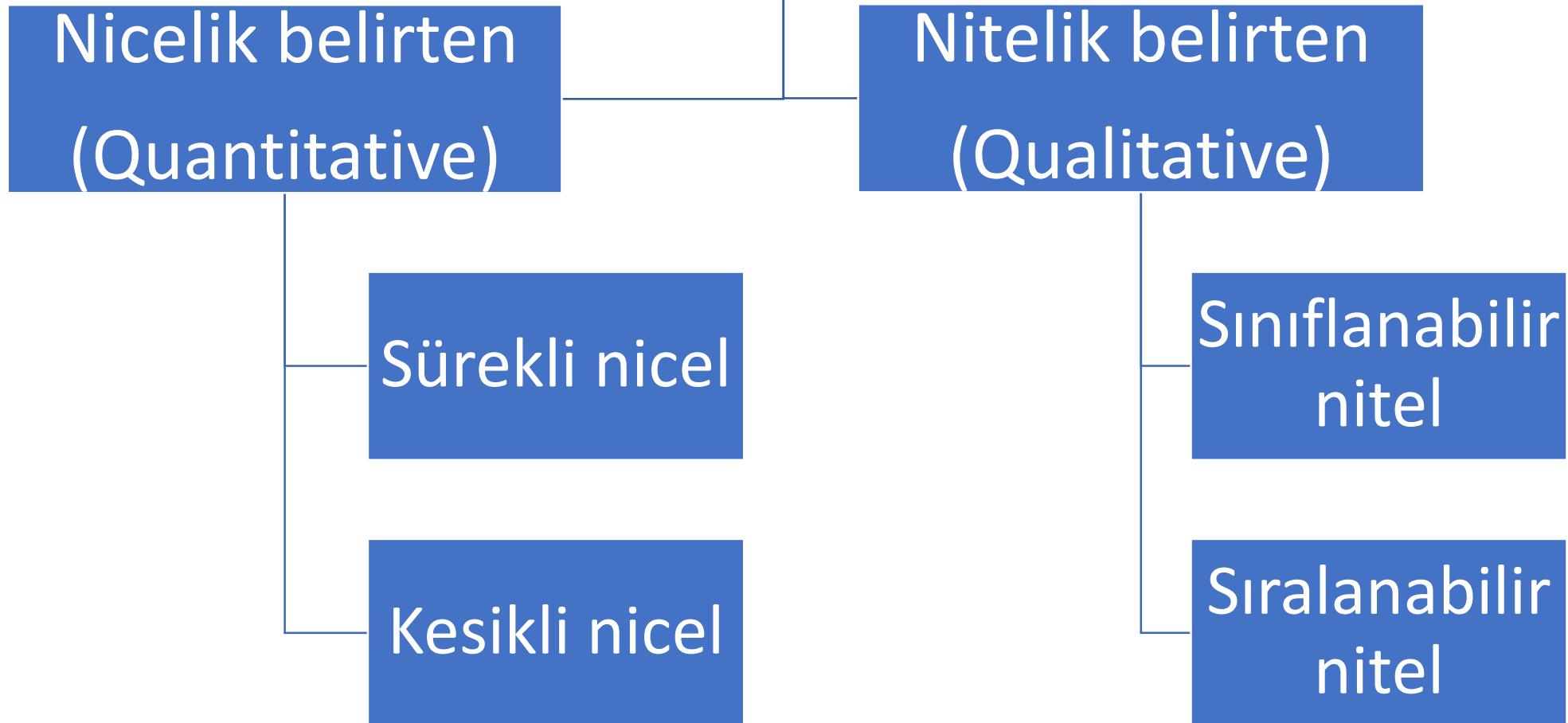
Yaş: 20-21 , 22-30 , 31-45, (Gruplanılmış veri)

Veriler nicel ve nitel olarak iki gruba ayrılır:

Nicel veriler: Bir deneğin (gözlemin ya da birimin) sayılabilir ölçülebilir özelliğini veren değerlerdir. Ondalıklı değer alabilen nicel veriler, sürekli nicel verilerdir. Örneğin boy uzunluğu, kilo sürekli nicel verilerdir. Sayılarak elde edilen ve tam sayılı değerler alabilen nicel veriler, kesikli nicel verilerdir. Örneğin çocuk sayısı, kitap sayısı, hane halkı sayısı vb.

Nitel veriler: Bir deneğin niteliğini belirten veriler nitel verilerdir. Birbirinden bağımsız isim bildiren, kod ve numara ile gösterilebilen, sınıflara ayrılabilen nitel veriler sınıflanabilir nitel verilerdir. Örneğin doğum yerleri, ülkeler, fakülte türleri vb. sınıflanabilir nitel verilerdir. Belirli bir miktar belirtmeyen, sıra ya da dereceye göre elde edilen nitel veriler sıralanabilir nitel verilerdir. Örneğin başarı durumu, akademik unvan, eğitim düzeyi vb. sıralanabilir nitel verilerdir.

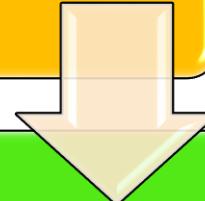
Değişken Türleri



ÖRNEK: yaş, cinsiyet, medeni durum, istihdam durumu, eğitim durumu,...birer değişkendir.

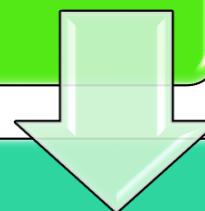
YAŞ, GELİR, AĞIRLIK

- Sürekli Nicel



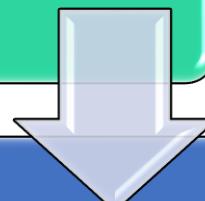
Eğitim durumu, Akademik unvan, Memnuniyet

- Sıralanabilir Nitel



Cinsiyet, Uyruk, Yaşadığı şehir

- Sınıflanabilir Nitel



Ailedeki çocuk sayısı, Kitap sayısı

- Kesikli Nicel

Derse devamsız
olduğu gün sayısı

Tansiyon

Diyabet (Var, Yok)

Acil servise
gelenlerin sayısı

Memnuniyet (1-5ölçeğinde)

Harf Notu

Vücut Kitle
indeksi

Gelen Turistlerin
Milliyeti

Yarışı bitirme
derecesi (dk)

Tamsayım (census): Bir araştırma kapsamında, kitledeki tüm birimlerine ulaşarak istenen bilginin elde edilmesi işlemidir. Bunun yapılabilmesi için incelenenecek kitlenin büyüklüğünün, belirlenen maliyet ve zaman gibi kısıtlara uygun olması gereklidir. Bazı durumlarda (nüfus sayımları gibi) kitle büyük olsa bile tam sayılmaması zorunlu olmaktadır. Gelişen teknoloji ile birlikte bu tür tam sayımlar daha kolay yapılabilir hale gelmiştir. En son nüfus sayımı 2000 yılında yapılmıştır. 2007 yılından itibaren ise Adrese Dayalı Nüfus Kayıt Sistemi (ADNKS) kapsamında nüfus sayımı yapılmaktadır.

Gözlem (denek) (observation): Kitle ya da örneklemde yer alan her bir birime gözlem ya da denek denir. Genellikle kitledeki denek sayısı “ N ”, örneklemdeki denek sayısı “ n ” ile gösterilmektedir.

Parametre ve istatistik: Kitleye ilişkin özelliklerin sayısal ifadeler parametre, örneklemeye ilişkin sayısal ifadeler istatistik olarak adlandırılır. Araştırma daha çok örneklemeler üzerinden yapıldığı için, örneklemden elde edilen istatistikler kitle parametrelerinin tahmini için kullanılıyorsa tahmin adını alır.

Parametre (parameter)

Kitle ortalaması : μ

Kitle Varyansı: σ^2

Kitle Büyüklüğü : N

Istatistik (statistic)

Örneklem ortalaması: \bar{X}

Örneklem varyansı: S^2

Örneklem büyülüğu :n

BÜYÜK VERİ (BIG DATA) NE DEMEK ?

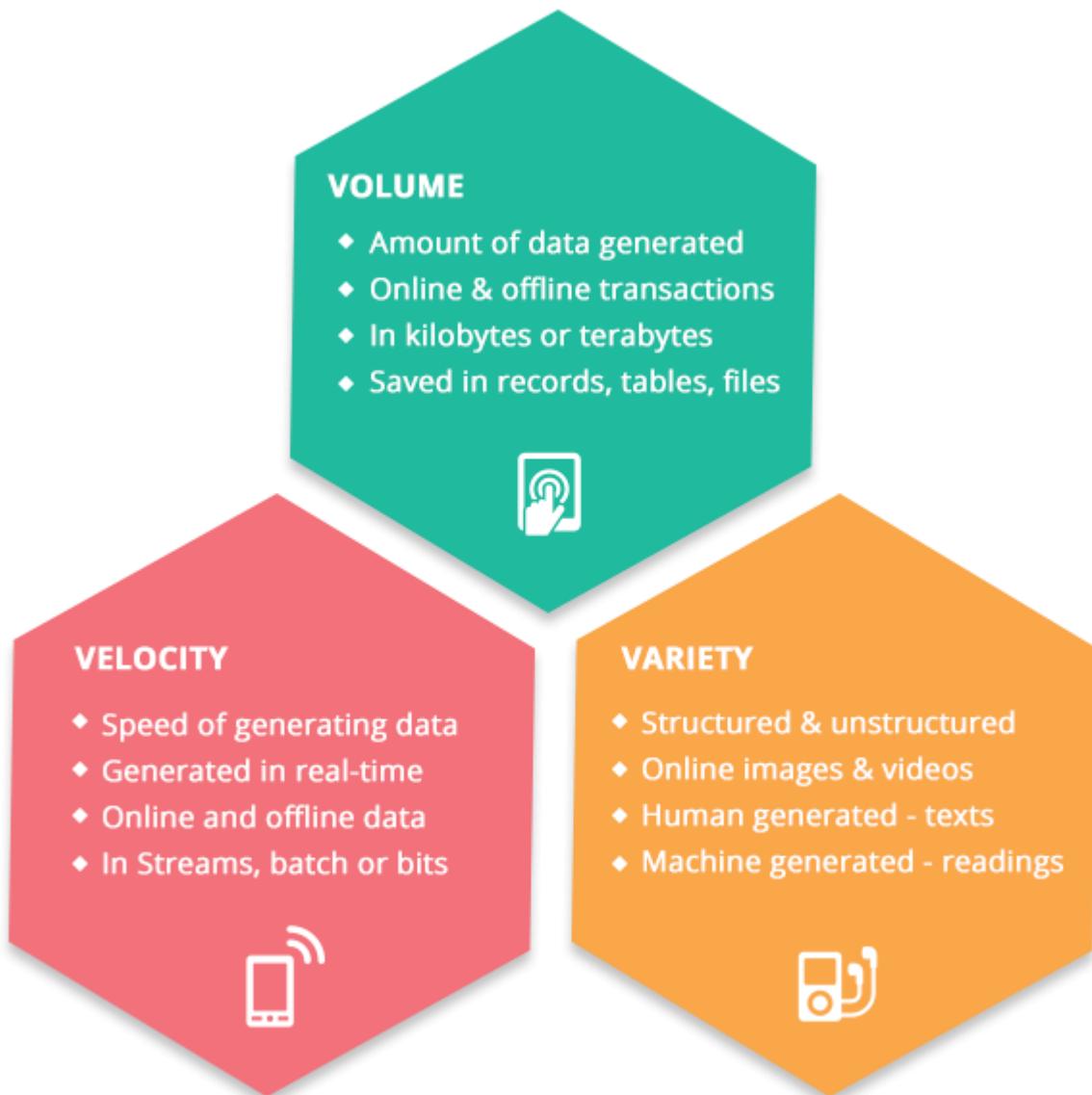
Sosyal medya araçları, arama motorları , internet alışveriş siteleri gibi kaynaklardan elde edilen büyük hacimli verilerdir. Bu veriler işlenerek bu verilerden anlamlı bilgiler elde etme amaçlanır.

Büyük veriye yatırım yapılarak satış, ürün geliştirme ve daha başka farklı alanlarda etkin bir fayda sağlanmaktadır.

Büyük veriye yatırım yapan şirketler:

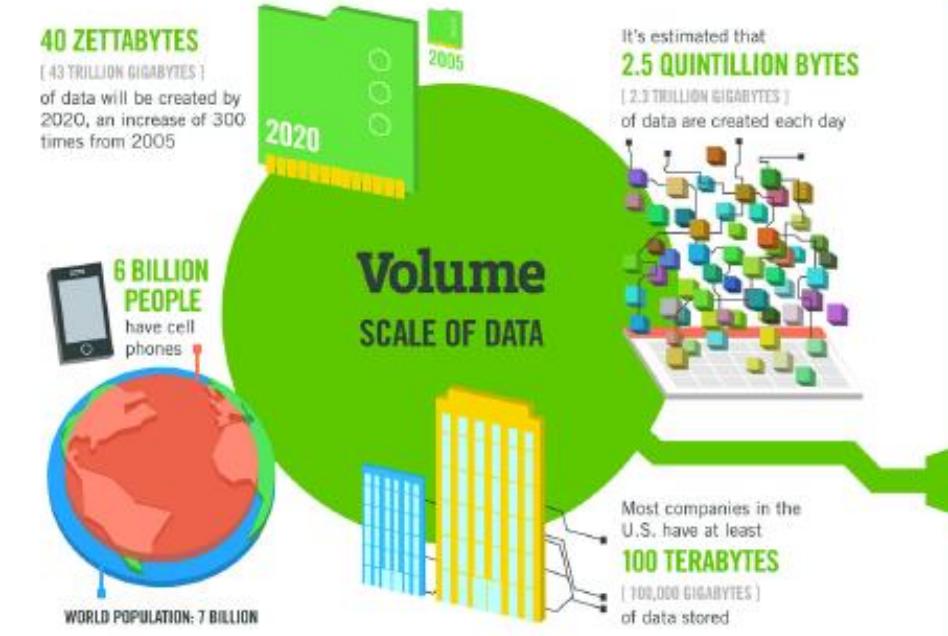
Facebook, Instagram, Amazon, Netflix, Starbucks, General Electrics, Whatsapp

THE 3Vs OF BIG DATA



www.whishworks.com

<https://www.whishworks.com/2017/09/08/understanding-the-3-vs-of-big-data-volume-velocity-and-variety/>, 09.10.2020



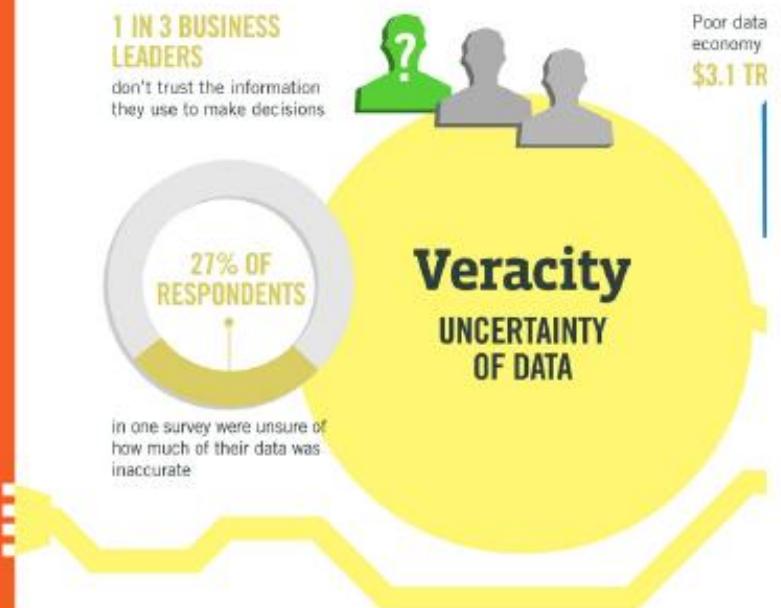
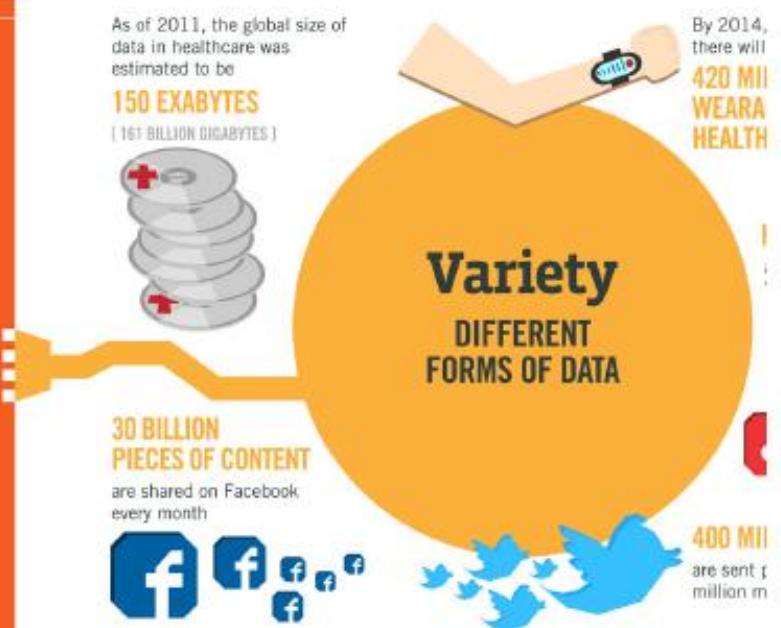
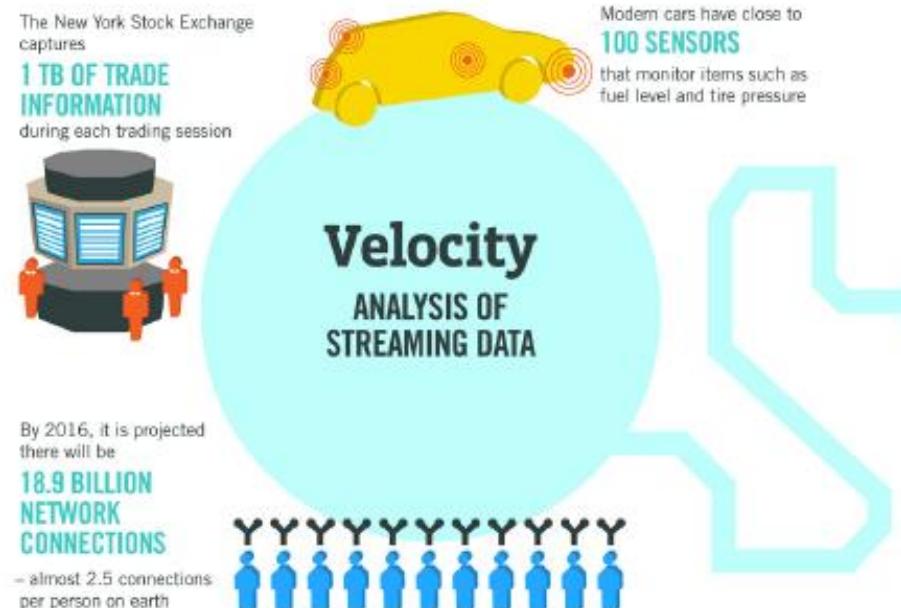
The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume**, **Velocity**, **Variety** and **Veracity**.

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

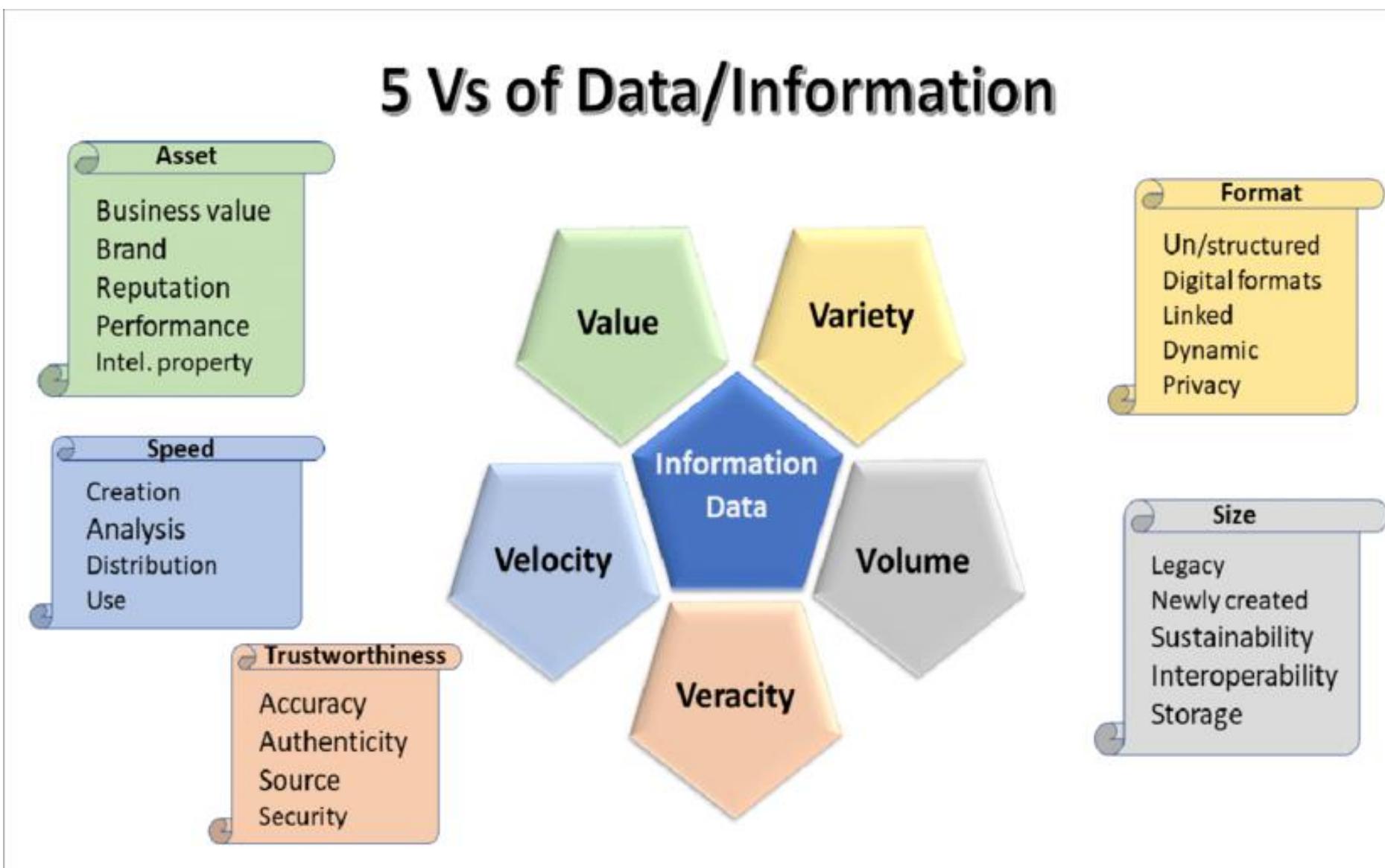
By 2015, **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States.



Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

<https://www.ibmbigdatahub.com/infographic/four-vs-big-data>, 09.10.2020

5 Vs of Data/Information

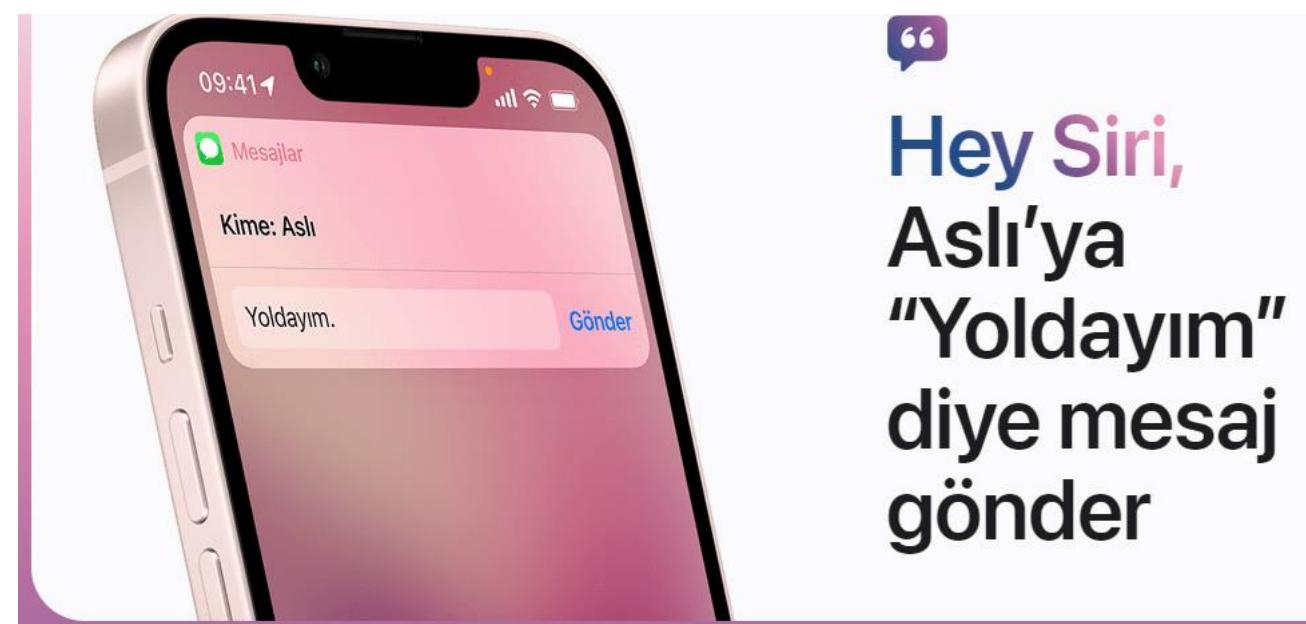


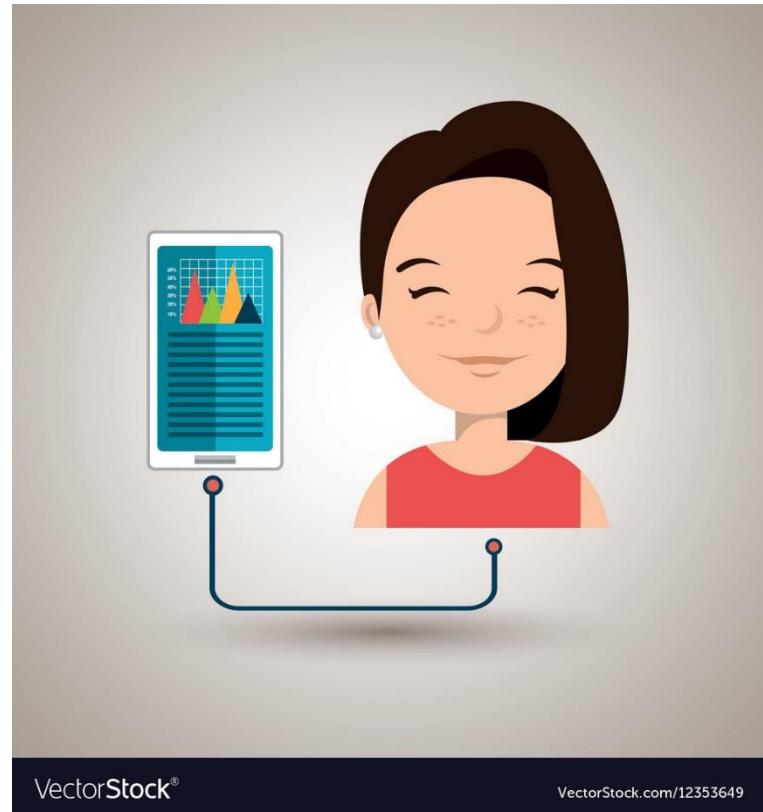
Savić, Dobrica. (2019). Are we ready for the future? Impact of Artificial Intelligence on Grey Literature Management. Grey Journal. 15. 7-15.

Derin Öğrenme (Deep Learning) Nedir?

Derin öğrenme, insan beynini taklit eden yapay sinir ağları algoritmaları kullanılarak veriden öğrenmeyi sağlayan makine öğreniminin bir alt kümesidir. Derin öğrenme algoritmaları, öğrenmek için çok fazla veri gerektirdiğinden, veri akışındaki hız, derin öğrenme yeteneklerinin son yıllarda artmasına neden olmuştur.

Siri
Dil çevirimcileri
Sürücüsüz araçlar
Yüz tanıma
Kişiselleştirilmiş ürünler...





Bir sonraki derste sıklık dağılımları, grafikler inclenecek.



KAYNAKLAR

- 1-) H.Demirhan, C.Hamurkaroğlu ,“İstatistiksel Yöntemlere Giriş”, H.Ü.Yayınları, 2011.
- 2-) Serpil Cula, Zehra Muluk, “Temel İstatistik Yöntemler”, Başkent Üniversitesi yayınları,2006.
- 3-) Levent Özbeğ, Esin Köksal Babacan, “İstatistiğe Giriş”, TÜBİTAK e-kitap.
- 4-) Birdal Şenoğlu, Mehmet Yılmaz, Sibel Açık Kemaloğlu, İstatistiğe Giriş, TÜBİTAK e-kitap.
- 5-) <http://analytics-magazine.org/survey-sampling/>
- 6-) <https://www.whishworks.com/2017/09/08/understanding-the-3-vs-of-big-data-volume-velocity-and-variety/>
- 7-) <https://www.ibmbigdatahub.com/infographic/four-vs-big-data>
- 8-) <http://www.anketis.com/blog/detay/evren-orneklem>
- 9-) <https://stattrek.com/statistics/random-number-generator.aspx#error>
- 10-) https://acikders.ankara.edu.tr/pluginfile.php/149104/mod_resource/content/0/EP%C4%B0DEM%C4%BOYOLOJ%C4%B0DE%20ARA%C5%9ETIRMA%20Y%C3%96NTEMLER%C4%B0.pdf