



Yazı Ertan Murat Erişik
Kara Danışmanlık Genel Müdürü



SON yıllarda giderek daha fazla büyük veriden (big data) bahsedildiğini duyuyoruz. Kavram olarak büyük veri herkesin dilinde, ama herkes aynı şeyi mi anlıyor orası biraz tartışmalı. Neden büyük veri hakkında konuşuyoruz? Çünkü artık çevremizdeki hemen hemen her şey veri üretiyor. Sadece işletmelerde kullandığımız, günlük olarak fatura kesip müşteri kayıtları oluşturduğumuz uygulamalar değil, GPS sinyalleri üzerinden cihazın lokasyon, durum bilgisini paylaşan sensörler, bilgisayarda yapılan her işlemi otomatik olarak kayıt eden log dosyaları, web siteleri, GPS sinyal kayıtları, akıllı hale gelen giyilebilir teknolojiler dediğimiz ürünler, televizyon, buzdolapları gibi ev elektroniği cihazları, binek ve taşıma araçları kısacası etrafımızdaki pek çok şey veri üretiyor. Görünen o ki, veri üretme kapasitesine sahip 'şey'lerin sayısı yakın gelecekte artacak.

Peki ama bunun bize faydası ne?

Veriler, doğru kullanılabilirse, gerek çevremizde olup bitenler gerekse gelecekte neler olabileceğine dair ciddi anlamda fikir verebilir. Şirketler, veriye dayalı analizlerle kârlılıklarını artıracak stratejiler geliştirip uygulayabilir. Kurumlar hizmet kalitelerini arttıracak şekilde bu verilerden faydalanabilir.

BT endüstrisinde yeni bir alan

Tam olarak büyük veri deyince ne kastediyoruz?

İşletmelerde biriktirilen verileri kabaca iki kategoride toplamak mümkün. İlk kategoriyi yapısal veriler yani şirket ya da kurumlarda kullanılan yazılımlar aracılığıyla üretilen veriler olarak adlandırmak mümkün. ERP sistemleri, satış, arka ofis uygulamaları, finans, krediler, kredi kartları, stok, muhasebe... Her iş birimi için kullanılan modülleri düşünelim, bazıları ayrı ayrı sistemlerde, bazıları tek bir ERP paketinin içerisinde toplanmış farklı amaçlara yönelik yazılımların ürettiği veriler bu kategorinin tipik örneği olarak verilebilir. Veriler yapısal yani birbirleri ile kolayca eşlenebiliyorlar. Bir müşteri numaramız var, onunla müşterinin yaptığı işlemleri takip edebiliyoruz. Bu veriler aynı zamanda kontrolümüz altında, belli bir süre geçince sistem sorumluları siliyor ya da arşivliyor. İkinci tür ise yukarıda belirttiğimiz konvansiyonel uygulamalar dışında kalan ve son yıllarda büyük veri kavramını giderek popülerleştiren her türlü 'şey'in ürettiği veri. Bu kategoride giyilebilir teknolojilerden tutun da hareket ölçen sensörlerin ürettiği verilere kadar geniş bir yelpazeden bahsediyoruz. Yelpaze hem geniş hem de boyut çok büyük. Pek kontrolümüzde de değiller ya da çok hızlı büyümeden ötürü depolama kapasitelerini hızla işgal etme ihtimalleri var. Yapısal değil, elimizdeki diğer verilerle ilişkilendirmek o kadar da kolay değil. Bu tür verileri depolamaya kalksak veri tabanı şirketlerini zengin etmemiz gerekir. Belki de bu kategoriye 'çok büyük veri' demek, daha iyi bir olabilir.

İki kategoriyi ayrı ayrı mı değerlendirmek lazım?

Öncelikle burada elde edilecek fayda ile gerçekleştirilecek yatırım dengesini çok iyi



“SENSÖRLER, TELEVİZYONLAR, BUZDOLAPLARI, CEP TELEFONLARI VE ÇEVREMİZDEKİ DAHA PEK ÇOK CİHAZ VERİ ÜRETİYOR. ANALİZİ ZOR DEĞİL.”

gözetmek lazım. Özellikle ikinci kategorideki verilerin düşük maliyetle depolanıp işlenebilmesi için açık sistemler temelli teknolojilerin oluşturulduğunu bilmekte fayda var. Yani çok büyük veriler, aslında BT endüstrisinde yeni bir alanın doğmasına öncülük etti. İlk kategorimizde yer alan ve mevcutta kullanılan yazılımların ürettiği verilerden sonuç almak için veri ambarları oluşturuluyor genelde.

Veri ambarının faydaları

Farklı yazılımlara ait veri tabanlarında üretilen verileri, tek bir veri tabanında toplayıp temizliyorsunuz, entegre edip gerçeğin tek bir versiyonunu üretecek bir bilgi havuzuna dönüştürüyorsunuz. Bu veri havuzunun tarihsel derinliğini de mevcut yazılımlarınız ile kısıtlanmadan belirliyorsunuz. Mesela ERP yazılımında sene sonunda veriler silinirken ya da arşivlenirken, veri ambarınızda bütçeniz ve iş ihtiyaçlarınız doğrultusunda istediğiniz kadar tarihsel veri tutabiliyorsunuz. Böylece müşterilerinizin ya da sizden hizmet alan kişilerin/ kurumların geçmiş işlemlerini dilediğiniz detayda yeniden

düzenleme şansına sahip oluyorsunuz. Detay veriyi analiz etmenin trendleri ve müşteri davranışlarını anlama konusunda çok ciddi getirileri olduğunu bilmek lazım. Tüketiciler davranışları ile sürekli birtakım mesajlar veriyorlar aslında. Bizim yaptığımız, bu davranışları rasyonel iş kararlarına çevirirken veri ambarından faydalanmak. Mağazacılıktan bir örnek verelim. En çok satan ürünler listesi perakendecilikte genelde çok fazla değişmez. Ama dönemsel olarak çarpıcı farklılıklar olabilir. Tekel'in henüz özelleştirilmediği günlerde yaptığımız bir projede şöyle bir örnek hatırlıyorum. Bir rakı markası o zamanlar en çok satan ürünler listesinde 7. veya 8. sırada yer alıyordu. Ancak ramazan ayı yaklaştığında, ramazandan bir hafta önce bu ürünün satışları düşmeye başlıyor, Ramazan'ın ikinci haftasında bu ürün ilk 100 arasına zor giriyordu. Bu normal ve tahmin edilebilir bir gelişme. Ama tahmin edilemeyen şey şu: Aynı ürünün satışları ramazan ayının üçüncü ve dördüncü haftasında çok çarpıcı bir artışla 2. veya 3. sıraya yükseliyordu. Bu davranışı aylık ortalama satışlara bakarak tahmin etmek mümkün değil, hatta anket yaparak da tespit etmek mümkün değil. İnsanlar genelde anketlerde ideal cevaplar verme davranışı gösterirler. Ortalama rakamlara baktığımızda satışlarda değişen bir şey yok, ama detaya baktığımızda çok ciddi sonuçlar var. Tüketiciler, kasaya gelip parayı ödediklerinde bir anlamda bize gerçekleri söylüyorlar. Elbette bu örnek belli bir zaman diliminde ve belli bir müşteri portföyüne sahip bir mağazadaki örnek. Ama veri ambarı bize insan davranışları ile ilgili başka yerde bulamayacağımız