

detayları veriyor. Bu detayları kullanarak stoklarımızı, personelimizi, pazarlama stratejilerimizi belirlemek ise bizim işimiz.

Veri ambarı konusunda neredeyiz?

Elbette veri ambarı sonuçlarını değerlendirerek son derece başarılı müşteri ilişkileri yönetimi yapan firmalar var. Perakende, telekomünikasyon ve bankalar için öncüler diyebiliriz. Bunların yanında pek çok farklı sektörde hâlâ yapılacak çok iş var, çünkü birtakım hatalar yapıyor. Her şeyden önce bu tip projeleri iş ihtiyaçlarına göre değerlendirmek lazım. Eğer ihtiyaçlara değil de teknoloji firmalarının çözümlerine odaklanılırsa istenilen sonuçlara varmak zorlaşabilir. Detay veri çoğu zaman önemsenmiyor, çünkü bu tip projelere statik rapor projeleri olarak bakılıyor ve veri özetlenmiş olarak tutulabiliyor. Oysa bu tip projeler analiz ortamları yaratırlar ve raporlama bu işin sadece bir parçasıdır. Yukarıdaki rakı satışları örneğinde olduğu gibi, aylık ortalamaya baksanız satışlarda değişim yok. Ama iki haftalık azalış, iki haftalık artış ile dengeleniyor. Eğer detay veriye bakmazsanız bunu göremezsiniz. Ama herhalde en büyük hatalardan biri, veri ambarı oluşturmak yerine bu entegrasyonu raporlama araçları ile yapmaya çalışmak.

Tekdüze verilerden küçük ama değerli bilgiler...

Her büyük veri projesi veri ambarı projesi midir? Hayır, ikinci kategori diye belirttiğimiz ve yapısal veriler dışında yer alan verilerin daha farklı ele alınması gerekir. Sensör verisini ele alalım mesela, bir marketler zincirindeki mağazalarda buzdolaplarının sıcaklığını belli periyotlarda ölçüp merkezi bir veri tabanına gönderen sensörler gibi... Mağaza sayısının çokluğunu ve ölçme periyodunun kısılalığını düşünürseniz, beş dakikada bir bütün mağazaların bütün buzdolaplarından gelen veriler devasa boyutlara ulaşacaktır. Bu tür verilerin tipik

özelliklerinden biri veri büyüklüğü ama bir diğeri ve en az büyüklük kadar baskın olan karakteri, bu verilerin çoğu tekdüzedir ve bize bir şey anlatmaz. Eğer gerçekten vahim bir durum yoksa, sıcaklık hep kontrol altındadır, sadece bazı ekstrem durumlarda incelenmeye degecek limit dışı veriler ürer, sıcaklık düşüşü (fazla enerji harcama riski) ya da yükselmesi (ürünlerin bozulma riski) gibi... Eğer bu tür bir veriyi konvansiyonel yaklaşımla pazarda satılan bir ilişkisel veri tabanını kullanarak analiz etmeye kalkarsanız çok büyük maliyetlere katlanmak zorunda kalırsınız. Büyük veri ya da çok büyük veri dediğimiz kavram burada devreye giriyor ve daha çok, çok büyük verinin içerisinde yer alan, küçük ama değerli miktardaki verilerin uygun maliyetli bir şekilde işlenmesini hedefliyor. Özellikle veri ambarı teknolojileri konusunda son yıllarda çok hızlı gelişmeler oluyor. Uygun maliyetli ve yüksek performans sunan yeni nesil çözümler var. Hafıza içi (in-memory) teknolojiye sahip ama aynı zamanda standartlara uygun ölçeklenebilir yeni nesil veri tabanları çok yüksek performansı oldukça uygun maliyetlerle sunabiliyorlar. Veri ambarı ve büyük veri projeleri birbirleri ile rekabette değiller o zaman? Hayır değiller! Bunları birbirlerinin tamamlayıcısı olarak düşünmek lazım. Bu iki ortamın zaman zaman birbirlerini tetikleyebileceğini düşünebiliriz. Büyük veri ortamındaki birtakım verilerin entegre edilerek daha hızlı bir sorgulama ortamına çekilebileceğini ya da veri ambarında tutulmasına gerek olmayan işlem bilgilerinin daha uygun maliyetli açık sistem tabanlı (hadoop) çözümlere kaydırılması gayet

mümkün. Bütün bunları yaparken maliyet ve fiyat performans dengesini hep akılda tutmak gerek elbette.

Nasıl başlamalı?

Temelde hangi sorunu çözeceğiniz konusunda çok iyi fikir sahibi olmanız lazım. Bu tür projeler, maliyetli, meşakkatli ve riskli projelerdir. Dolayısıyla başlamadan önce bir kapsam çalışması yapmak gayet mantıklı bir yaklaşım olacaktır. Bu çalışma ile projeye dayanak oluşturan iş ihtiyaçlarının neler olduğu, bu ihtiyaçların hangi adımlarda ve ne kadar sürelerde çözülebileceği, ne tür bir teknolojik altyapı kullanılacağı gibi bir yol haritasının belirlenmesi lazım. Elbette bu yol haritasının içereceği teknolojik bileşenler BT stratejileri ile uyumlu olmalı ve alternatifler içermeli. Veri ambarı projelerinde uyguladığımız bir prensip var: Proje adımlarını kullanıcılara elle tutulur bir sonuç üretecek şekilde içerik olarak zengin, ilgilerini azaltmayacak kadar da kısa tutmak. Eğer bir proje adımı anlamlı sonuç üretmiyorsa, kullanıcıların projeye olan desteği zayıflayacaktır. Ya da süresi çok uzun ise ilgileri azalıp başka konulara yönelme ihtimalleri olacaktır. İkisi arasında denge bulunmalı.

Sonuç:

Veriyi doğru kullanmak çevremizde olan biteni anlamak için en güçlü silahımız. Bu silahı verimliliği artırmak için iyi kullanmak lazım. Evet, veri boyutları büyüyor, çeşitleniyor, yönetmek zorlaşıyor, ama günümüzde bu sorunları aşabilecek teknolojiler ve yeni nesil yaklaşımlar elimizde var. Dolayısı ile doğru teknolojiyi doğru yerde kullanarak yüksek performanslı çözümleri uygun maliyetle üretmek mümkün. —



**“VERİYİ DOĞRU
KULLANMAK
ÇEVREMİZDE OLAN
BİTENİ ANLAMAK
İÇİN EN GÜÇLÜ
SİLAHIMIZ. BU
SİLAHI VERİMLİLİĞİ
ARTIRMAK İÇİN
İYİ KULLANMALI.”**

