

Veri Ambarı - Data Warehouse

Veri ambarı zamanla birikmiş verilerden oluşan bir veri yığınıdır. Birden çok kaynaktan gelen büyük verileri birleştirir ve merkezileştirir. Kurumların sahip olduğu verileri karar destek amacı ile kullanılmasını sağlar. Raporlama ve veri analizi için kullanılan iş zekası sisteminin temel bileşenidir.

Veri ambarı, daha iyi karar vermeyi kolaylaştırmak için verilerin depolanmasına, analiz edilmesine ve yorumlanmasına izin veren merkezi bir depolama sistemidir . İşlem sistemleri, ilişkisel veritabanları ve diğer kaynaklar, veri ambarlarına düzenli olarak veri sağlar.

Veri ambarı , özellikle analiz olmak üzere iş zekası (BI) faaliyetlerini kolaylaştıran ve destekleyen bir tür veri yönetim sistemidir. Veri ambarları öncelikle aramaları ve analizleri kolaylaştırmak için tasarlanmıştır ve genellikle büyük miktarda geçmiş veri içerir.

Bir veri ambarı, operasyonel kaynaklardan ve harici veri kaynaklarından çıkarılan kurumsal veri ve bilgiler topluluğu olarak tanımlanabilir. Veriler, satış, pazarlama ve finans gibi çeşitli dahili uygulamalardan periyodik olarak alınır; müşteri arayüzü uygulamaları; yanı sıra harici ortak sistemleri. Bu veriler daha sonra karar vericilerin erişmesi ve analiz etmesi için hazır hale getirilir.

Veri Ambarı 'nın Temel Özellikleri

Veri ambarı karar verme sürecinde yöneticilere destek vermek üzere hazırlanmış,

- a. Konuya Yönelik
- b. Bütünleşik
- c. Zaman boyutu olan
- d. Sadece Okunabilen

veri topluluğudur.

Bir veri ambarının ana özellikleri aşağıdaki gibidir:

- **Konu Odaklı**

Bir veri ambarı, bir işletmenin genel süreçlerinden ziyade konuya dayalı bilgiler sağladığı için konu odaklıdır. Bu konular satış, promosyon, envanter vb. olabilir. Örneğin şirketinizin satış verilerini analiz etmek istiyorsanız satış odaklı bir veri ambarı oluşturmanız gerekir. Böyle bir depo, 'geçen yıl en iyi müşteriniz kimdi?' gibi değerli bilgiler sağlar. veya 'önümüzdeki yıl en iyi müşteriniz kim olabilir?'

- **Birleşik**

Bir veri ambarı, çeşitli kaynaklardan gelen verilerin tutarlı bir biçimde bütünleştirilmesiyle geliştirilir. Veriler, ambarda adlandırma, format ve kodlama açısından tutarlı ve evrensel olarak kabul edilebilir bir şekilde saklanmalıdır. Bu, etkili veri analizini kolaylaştırır .

Bütünleştirme farklı biçimlerde olabilir. Verinin kodlanmasında görüş birliğine varılması, Ölçü birimlerinin seçimin de tutarlılık, sayısal değerlerin fiziksel gösterimindeki tutarlılık gibi bütünleştirme kavramlarından söz edilebilir.

- **Uçucu Olmayan**

Veri ambarı veri yönetiminin gereksinimine yanıt vermek üzere tasarlandığı için günlük işlemlere tabi tutulmaz yani silinemez veya güncelleştirilemez. Veri ambarına girilen veriler değişmeden kalmalıdır. Tüm veriler salt okunurdur. Mevcut veriler girildiğinde önceki veriler silinmez. Bu, ne olduğunu ve ne zaman olduğunu analiz etmenize yardımcı olur.

Veri ambarında iki tür işlem den söz etmek mümkündür:

- a. Verinin yüklenmesi
- b. Veriye erişilmesi

- **Zaman Değişkeni**

Bir veri ambarında depolanan veriler, açık veya örtük olarak bir zaman unsuru ile belgelenir. Veri Ambarı 'ndaki bir zaman farkı örneği, gün, hafta veya ay gibi bir zaman unsuruna sahip olması gereken Birincil Anahtar 'da görüntülenir.

Veri ambarındaki veri, sadece o andaki değerlerle değil geçmişteki değerlerle de ilgilidir. Veri zaman içinde bir noktayla birleştirilerek değerlendirilir. Örneğin sömestre, mail yılı, ödeme dönemi gibi unsurlar göz önüne alınır.

Veri Ambarı Mimarisi

Genellikle veri ambarı mimarisi üç katmanlı bir yapıdan oluşur.

Alt Katman

Alt katman veya veri ambarı sunucusu genellikle ilişkisel bir veritabanı sistemini temsil eder. Verileri temizlemek, dönüştürmek ve bu katmana beslemek için arka uç araçları kullanılır.

Orta Katman

Orta katman, iki şekilde uygulanabilen bir OLAP sunucusunu temsil eder.

ROLAP veya ilişkisel OLAP modeli, çok boyutlu veri sürecini standart ilişkisel sürece eşleyen genişletilmiş bir ilişkisel veritabanı yönetim sistemidir.

MOLAP veya çok boyutlu OLAP, doğrudan çok boyutlu veri ve işlemler üzerinde hareket eder.

Üst Katman

Bu, veri ambarından veri alan ön uç istemci arabirimidir. Sorgu araçları, analiz araçları, raporlama araçları ve veri madenciliği araçları gibi çeşitli araçları içerir .

Veri Ambarı Nasıl Çalışır?

Veri Ambarı, çeşitli kaynaklardan toplanan verileri ve bilgileri tek bir kapsamlı veritabanında birleştirir. Örneğin, bir veri ambarı, bir kuruluşun satış noktası sistemlerinden, posta listelerinden, web sitesinden ve yorum kartlarından gelen müşteri bilgilerini birleştirebilir. Ayrıca çalışanlarla ilgili gizli bilgileri, maaş bilgilerini vb. içerebilir. İşletmeler, müşterileri analiz etmek için bu tür veri ambarı bileşenlerini kullanır.

Veri madenciliği, büyük hacimli verilerde anlamlı veri kalıpları aramayı ve artan satışlar ve karlar için yenilikçi stratejiler geliştirmeyi içeren bir veri ambarının özelliklerinden biridir.

Veri Ambarı Türleri

Üç ana veri ambarı türü vardır.

Kurumsal Veri Ambarı (EDW)

Bu tür depo, kuruluş genelinde karar destek hizmetlerini kolaylaştıran bir anahtar veya merkezi veritabanı işlevi görür. Bu tür ambarın avantajı, kuruluşlar arası bilgilere erişim sağlaması, veri temsiline birleşik bir yaklaşım sunması ve karmaşık sorguların çalıştırılmasına izin vermesidir.

Operasyonel Veri Deposu (ODS)

Bu tür veri ambarı gerçek zamanlı olarak yenilenir. Çalışan kayıtlarının saklanması gibi rutin faaliyetler için sıklıkla tercih edilir. Veri ambarı sistemlerinin işletmenin raporlama ihtiyaçlarını desteklemediği durumlarda gereklidir.

Data Mart

Data mart, belirli bir departmanı, bölgeyi veya iş birimini sürdürmek için oluşturulmuş bir veri ambarının alt kümesidir. Bir işletmenin her departmanı, verileri depolamak için merkezi bir havuza veya veri pazarına sahiptir. Data marttan gelen veriler periyodik olarak ODS 'de saklanır. ODS daha sonra verileri depolandığı ve kullanıldığı EDW 'ye gönderir.

Veri Ambarının İçerdiği Verilerin Sınıflandırılması

- Metadata
- Ayrıntı veri
- Eski ayrıntı veri
- Düşük düzeyde özetlenmiş veri
- Yüksek düzeyde özetlenmiş veri

Metadata

Metadata veri ambarı içindeki veriden ayrı bir yerdedir. Doğrudan işlemsel çevrede gelen veriyi içermez. Karar destek sistemleri analistlerine yardım etmek üzere yaratılan bir dizindir ve Veri ambarı içeriğinin neler olduğunu belirtir. *Kullanılan verinin yapısını ortaya koyan bilgiler içerir.* İşlemsel çevreden veri ambarına dönüştürülen verinin konumları hakkında bilgileri içeren bir kılavuzdur. İşlemsel çevreden alınan verinin hangi

algoritmaya göre düşük ya da yüksek seviyede özetlendiği hakkında bilgileri içeren bir kılavuzdur.

Ayrıntı Veri

Veri ambarı şuan ki ayrıntı veri ile ilgilendiğini varsayalım; Bu veri en son olayları içermektedir ve henüz işlenmediği için diğerlerine oranla daha büyük hacimlidir. Bu tür veri disk üzerinde saklandığından bunlara erişim ve yönetimi pahalıdır.

Ayrıntı veri denilince, şu andaki en son ayrıntı veri kastedilmektedir. Bu ayrıntılar belirli bir dönemi kapsayabilir. Örneğin satışlara ilişkin veri son beş yılın ayrıntılarını içerebilir.

Eski ayrıntı veri

Düşük düzeyde özetlenmiş veri

Yüksek düzeyde özetlenmiş veri

Veri Ambarı Mimarisi

Veri ambarı mimarisinin genel karakteristikleri şu şekilde sıralanabilir:

- Kaynaklardan alınan veri dönüştürülür.
- Veri ambarı oluşturulur.
- Kullanıcıların veri ambarına erişimi sağlanır.

Veri Ambarı Örneği Nedir?

Egzersiz ekipmanı yapan bir şirket düşünün. En çok satanı sabit bir bisiklet ve ürün gamını genişletmeyi ve onu desteklemek için yeni bir pazarlama kampanyası başlatmayı düşünüyor.

Mevcut müşterisini daha iyi anlamak için veri ambarına gider. Müşterilerinin ağırlıklı olarak 50 yaş üstü kadınlar mı yoksa 35 yaş altı erkekler mi olduğunu öğrenebilir. Bisikletlerini satmada en başarılı olan perakendeciler ve nerede bulundukları hakkında

daha fazla bilgi edinebilir. Şirket içi anket sonuçlarına erişebilir ve geçmiş müşterilerinin ürünleri hakkında neleri beğendiğini ve beğenmediğini öğrenebilir.

Tüm bu bilgiler, şirketin ne tür yeni model bisikletler yapmak istediğine ve bunları nasıl pazarlayıp reklamını yapacağına karar vermesine yardımcı olur.

SQL bir Veri Ambarı mı?

SQL veya Yapılandırılmış Sorgu Dili, anlayabileceği ve yanıt verebileceği bir veritabanıyla etkileşim kurmak için kullanılan bir bilgisayar dilidir. "Seç", "ekle" ve "güncelle" gibi bir dizi komut içerir. İlişkisel veritabanı yönetim sistemleri için standart dildir.⁶

Her ikisi de bilgi deposu olmasına rağmen, bir veritabanı veri ambarı ile aynı değildir. Veritabanı, organize edilmiş bir bilgi topluluğudur. Bir veri ambarı, birden çok kaynaktan sürekli olarak oluşturulan bir bilgi arşividir.⁷

Veri Ambarı 'nda ETL Nedir?

"ETL", "çıkarma, dönüştürme ve yükleme" anlamına gelir. ETL, birden çok kaynaktan gelen verileri tek bir veri depolama biriminde birleştiren ve daha sonra bir veri ambarına veya benzer bir veri sistemine yüklenen bir veri sürecidir. Veri analitiği ve makine öğreniminde kullanılır.