

VERİSETİ ORGANİZASYONU

Dr. Cahit KARAKUŞ

VERİ Seti nedir?

- Bilgi topluluğudur.
- Bilgiyi depolayan, düzenleyen ve işleyen bir matematiksel modeldir.
- Veri seti, en geniş anlamıyla; birbiriyle ilişkili verilerin tekrara yer vermeden, çok amaçlı kullanımına olanak sağlayacak şekilde depolanması olarak tanımlanabilir.
- Tablolardan oluşmak zorundadır.

Veriseti

- Bir veri yığını içerisinde veriseti oluşturmak, hata ve belirsizlikleri bulmak, düzenlemek, korumak, kullanıma hazır hale getirmek, analiz etmek, simülasyon ve optimize etmek, geliştirmek, yeni girişin davranış sapma değişimini ve etkileşimini bulmak, sürdürülebilir kılmak gibi çeşitli karmaşık işlemlerin gerçekleştirildiği bir yazılım sistemidir.
- $F(x)$ = iskelet yapısı oluşturma
 - Giren her yeni veriden tepki ve etki bulma
 - Yeniden güncelleme
 - Dağılma özelliği ile multi fonksiyonlar elde etme.

Modelleme, Ürünleştirme: Analiz,optimizasyon, simülasyon...

VERİSETİ ORGANİZASYONU

İşlevsel veritabanları oluşturma, güvenlik, iletişim, erişim, sınıflandırma, veri yığınınından analiz işlemleri ve analiz sonuçlarından raporlama, sunum, denetleme ve yönetme olanağı sağlayan süreçler bütünüdür.

Veriseti Bileşenleri

Veriseti organizasyonu tanımlama, biçimleme, değiştirme, paylaşım ve koruma işlemlerini yürüten genel amaçlı yazılım sistemidir.

- **Tanımlama:** Veri tipleri, yapıları ve uygulamaları
- **Biçimleme:** Kayıt biriminde saklama
- **Değiştirme:** Saklanan bilginin sorgulanması, gerçek hayatta temsil ettiği örneğin değişimlerine göre güncelleme, rapor üretilmesi
- **Paylaşım:** Aynı anda birden çok kullanıcı ve uygulamanın erişimine izin verme
- **Koruma:** Yazılım ve donanım hatalarına, yetkisiz ve zararlı erişime karşı

Veriseti örnekleri

Veristi, birbirleriyle bağlantılı bilgi kümeleridir.

Bir veriseti;

- Telemetrik ölçüleri yansıtır. Elektriksel: amper, gerilim, şiddet, güç, Algılama,
- Özel içsel bütünlüğe sahip, mantıklı veri yığını; muhasebe girdileri; stok, satış, lojistik, banka...
- Özel amaçlı veri yığınları; bilimsel

Veriseti Örnekleri

- Üniversite- Öğrenci işleri bilgi sistemi
- Hastane-Hasta, doktor, tedavi, araç-gereç, mali bilgiler
- Ticari bir şirket- Müsteri, ürün, satış, ödeme, teslimat bilgileri
- Banka-Müşteri, mevduat, kredi kartı, kredi bilgileri
- Üretim, enerji
- İstatistiksel veri yığını
- Performans, verim, hata bulma, ...

Neden verisetleri kullanılır?

- Verilerin tutulması, saklanması ve erişilmesinde geleneksel yaklaşım verilerin ayrı ayrı dosyalarda gruplanması yaklaşımını kullanmaktadır.
- Verilerin artması, verilere aynı anda erişme ve düzenlenme ihtiyacı ile geleneksel yaklaşım yetersiz kalmıştır.

Veriseti organizasyonunun yararları

- Minimum veri tekrarı
- Veri tutarlılığı
- Veri bütünleştirme
- Erişim paylaşımı
- Uygulama geliştirme kolaylığı
- Tekbiçim güvenlik, mahremiyet ve tutarlılık denetimleri
- Veri erişilebilirliği ve cevap verebilme
- Veri bağımsızlığı
- Azaltılmış program yaşatımı

Veri seti oluřturmanın avantajları

- Ortak verilerin tekrarını önlenmesi;
- Verilerin merkezi denetiminin ve tutarlılıđının sađlanması
- Güvenli veri paylaşımının sađlanması
- Fiziksel yapı ve erişim yöntemi karmaşıklıklarının, çok katmanlı mimarilerle kullanıcıdan gizlenmesi,
- Her kullanıcıya yalnız ilgilendiđi verilerin, alışık olduđu kolay, anlaşılır yapılarda sunulması

Veriseti uygulamalarının avantajları

- Çözümleme, tasarım ve geliştirme araçları ile uygulama yazılımı geliştirmenin kolaylaşması.
- Veri bütünlüğü,
- Güvenlik ve gizlilik istenilen düzeyde
- Yedekleme, yeniden başlatma, onarma gibi süreç yönetim, risk analizi ve işletim sorunlarına çözüm getirilmesi

Veriseti organizasyonun avantajları

- “*Veri Tekrarı*” engellenir: Aynı veri farklı dosyalarda tekrar tekrar yer almaz. Bu verinin daha az yer kaplamasını sağlar.
- “*Çoklu Güncelleme*” yapılabilir: Birden fazla dosyada tekrarlanan verinin herhangi birini değiştirdiğimizde diğer dosyalardaki veri de aynı anda değişir.
- “*Gereksiz Bellek Kullanımı*” engellenir: Aynı veriler defalarca tekrarlanmadığı için bilgisayar belleğinde gereksiz yer işgal etmez.
- “*Erişim Dili*” standartlaşır: Veritabanına erişim dili uygulanmadan uygulamaya geçmez. Standart bir dil kullanımı vardır.

Veriseti organizasyonun avantajları

- Herhangi bir evrak saklamaya gerek yoktur.
- Bilgiler istenildiği zaman görülebilir.
- Bilgilerin davranışın kontrolü, denetlenmesi ve izlenmesi sağlanır.
- Verinin paylaşımı sağlanır.
- Raporlar alınabilir.
- Verilerin tutarlı olmasını sağlar (Veri Bütünlüğü = Data Integrity).
- Aynı andaki erişimlerde tutarsızlıkların ortaya çıkmasını önler.
- Verilerin güvenliğini sağlar (Kullanıcıların her alana erişememesi iyi bir özelliktir. Bunun için çeşitli yetkiler atanır ve verilerle birlikte bu yetkiler de saklanır).

Veri yığınının tepkisel davranışı

- Bellek alanını kullanım
- Veriyi saklama
- Veriye erişim
- Donuklaşma
- Panik ve acil durum
- Enerji kesilmesi
- Şüpheli davranışlar: Virüs, Veri hırsızlığı
- Gruplaşma, etkileşim

Veritabanı yönetim sistemleri

- Oracle database
- IBM DB/2
- Adaptive Server Enterprise
- Informix
- Microsoft Access
- Microsoft SQL Server
- Microsoft Visual FoxPro
- MySQL
- PostgreSQL
- Progress
- SQLite
- Teradata
- CSQL
- OpenLink Virtuoso

Veri seti organizasyonu türleri

- PC veri tabanları
- Merkezi veri tabanları
- İstemci/sunucu veri tabanları
- Dağıtık veri tabanları
- Veri tabanı modelleri

Bilgi: Data – Veri

Veri tipleri:

- Analog; peryodik, eşit aralıklı; genlik, frekans, faz
- sayısal
- Tetiklemeli
- Tablo, grafik
- Resim, video

Bilgi: Data – Veri

Veri akışı:

- Veri kaynakları,
- Depolama,
- Transfer, iletim
- Sınıflandırma

Bilgi: Data – Veri

Veri kaynakları özellikleri:

- Mevcut kaynaklardan (eski kayıtlar, arşivler, raporlar, yıllıklar)
- Süreç girdi ve çıktılarından
- İzleme – gözetleme
- Anket - deney

Bilgi: Data – Veri

Veri kaynakları:

- Sağlık,
- Ekonomi,
- Güvenlik: Görüntüleme - izleme, Savunma – saldırı – korunma, Arama - kurtarma

Bilgi: Data – Veri

Veri doğrulama:

- Hatalı
- Belirsiz
- Tutarsız
- Standartlara uygunsuzluk
- Manipule
- Gizleme

Analiz

- Bir madde içerisindeki bileşiklerin hepsini veya bir kaçının miktarını ve neler olduğunu ortaya koyma. Çözümleme, Tahlil...
- İnsanların yaşamları boyunca karşılaştıkları sorunları çözüm arayışları zamanla bu çözümleri modeller üzerinde arama yaklaşımını doğurmuştur. Matematik ve bilgisayardaki gelişmeleri dış dünyanın problemlerini matematiksel olarak problemleyip bu çözümleri modelleyip çözerek bu çözümleri gerçek hayata yansıtma olanağı vermiştir.
- Matematiksel modelleme tekniğinde doğrusal ve doğrusal olmayan modeller.

Hata

Örnekleme Hatası

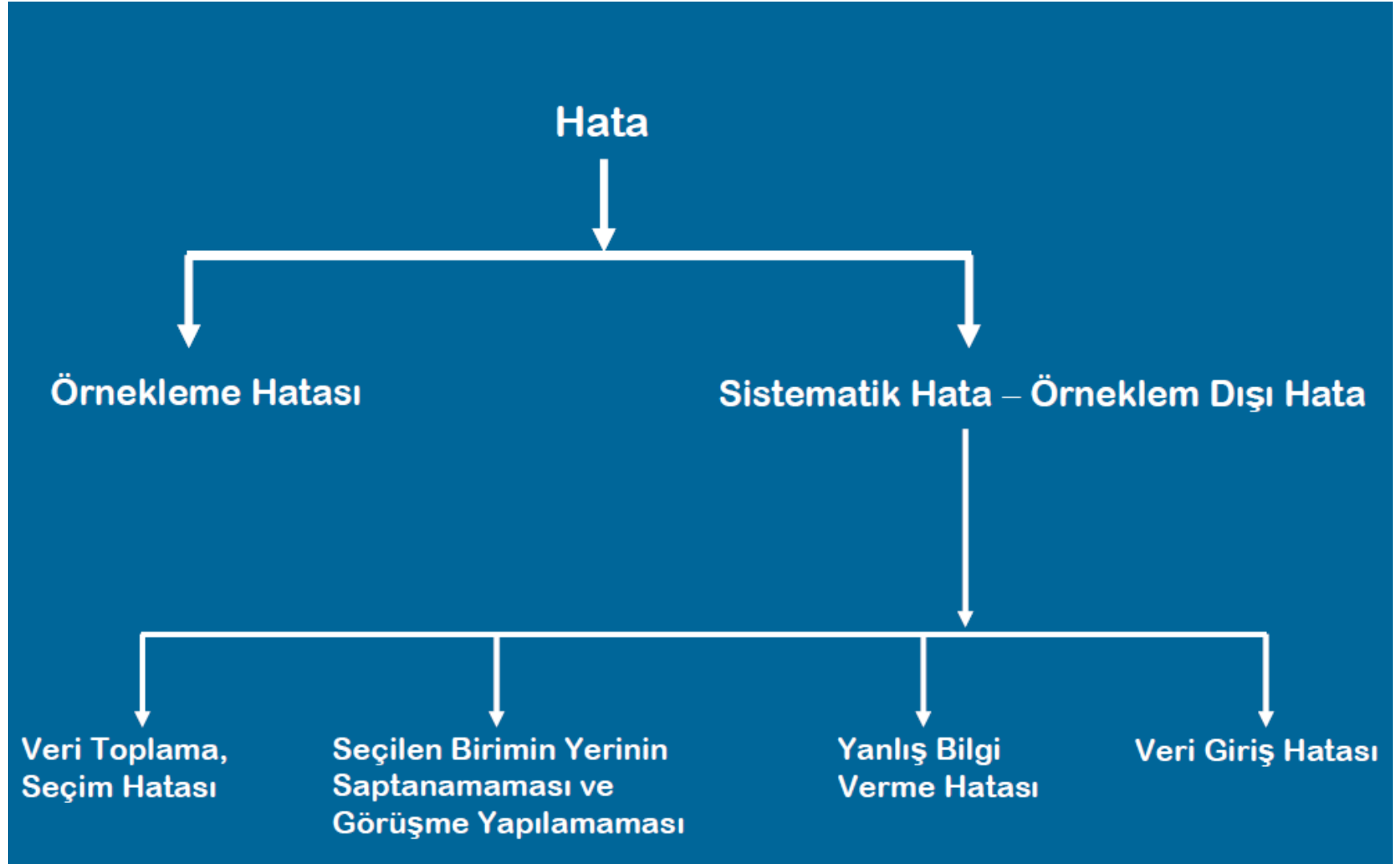
Sistemik Hata – Örneklem Dışı Hata

**Veri Toplama,
Seçim Hatası**

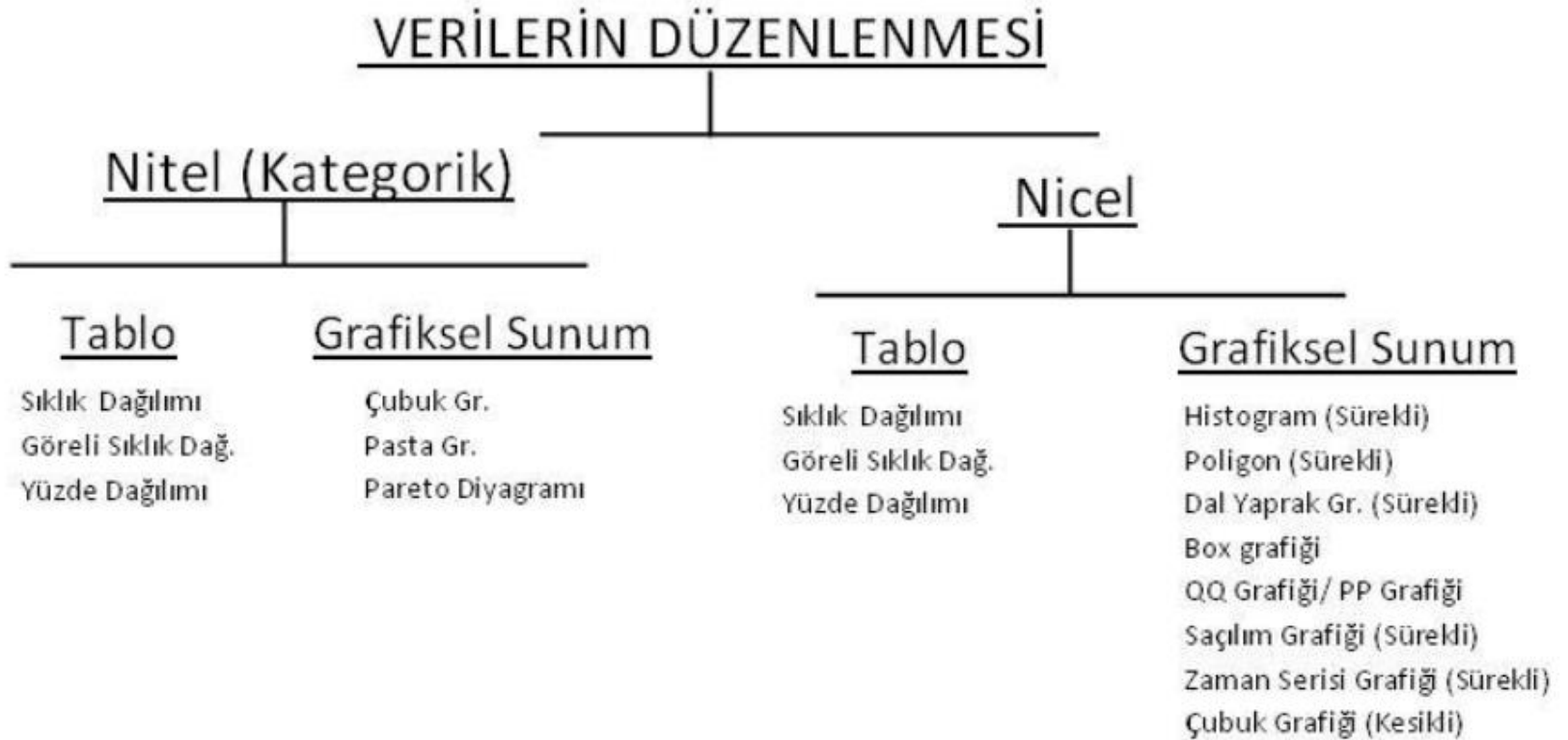
**Seçilen Birimin Yerinin
Saptanamaması ve
Görüşme Yapılamaması**

**Yanlış Bilgi
Verme Hatası**

Veri Giriş Hatası



Verilerin düzenlenmesi



Veri Analiz

- Analog işaret başta olmak üzere tüm giriş bilgilerinin sayısal veri haline dönüştürülmesi aşamasında örnekleme,
- Değerlendirme
- Risk analizi
- Sapma analizi
- Regresyon analizi, verileri denklemeleştirme
- Korelasyon
- Olasılık
- Kestirim, öngörü
- İstatistiksel analiz
- İşaret işleme

Veri Analiz, veri toplama

Ölçerseniz, yönetirsiniz

- Bilgi toplama kaynaklarına çevresel etkiler ve riskler
- Donanım kalitesi: kalibre, onay, yetkilendirme, verim, kayıp, ısısız davranış
- Veri toplama, transferi, sınıflandırma ve saklama

Verileri denklemeleştirme, Korelasyon

- Korelasyon, bir deęişkenin deęeri deęişirken dięer bir deęişken bununla lineer(doęrusal) lineer olmayan ilişkili olarak deęişiyorsa korelasyon vardır.
- Varsayım, kestirim, olasılık
- Sebep sonuç ilişkisi
- Test
- Ret bölgeleri

Simülasyon

- Bir sistemi temsil edebilecek bir matematiksel model oluşturma işlemidir.
- Gerçek sistemin modelinin tasarlanması ve bu model ile sistemin işletilmesi amacıyla yönelik olarak , sistemin davranışını anlayabilmek veya değişik stratejileri değerlendirebilmek için deneyler yürütülmesi sürecidir.
- Geliştirilen veya yeniden düzenlenen süreçleri tamamlamada ve deneme çalışmalarını yürütmede ve süreçlerin hata zamanlarını tahmin etmek için yapılan deneysel çalışmadır .Yeni sürecin değişikliklere gösterdiği olası reaksiyonları da anlayabiliriz.

Optimizasyon

- Her hangi bir dalda (örneğin bilgisayar yada telefonda) performansı artırmak ve daha iyi verim alabilmek için yani sistemi daha iyi bir hale getirmek için yapılan işlemlerin tümüne optimize yada optimize etmek denir.
- Yeni girdiye aşırı duyarlılık gösteren parametreler; kaotik davranış.
- Giriş değerinde en küçük değişim, çıkışta çok büyük sapmaya neden olursa
- Bir sistemin işlevini tanımlayan çok sayıda değişkenleri olan bir denklemde, değişken değerlerinin belirlenmesidir.

Sentez

- Bir sistemin öğelerini mantıksal bir tarzda bir araya getirme işlemi.
- Birleştirme faaliyetleri
- Bir madde ya da sistemi oluşturmaktır.
- Element veya başka maddeleri bir araya getirerek yapay olarak bileşik cisimler oluşturma, bireşim.
- Varılan sonuca gidilme, bireşim

İlişkisel verisetleri

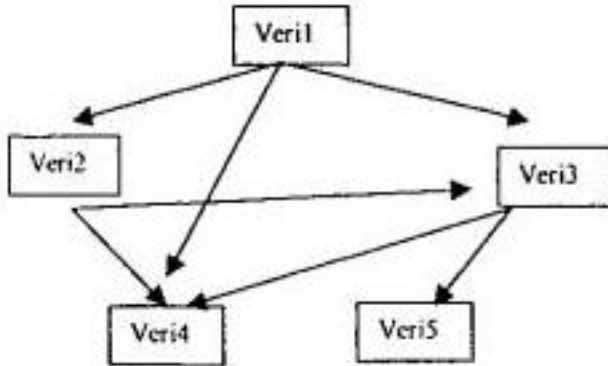
- Tablolar arasındaki bağlantılar matematiksel ilişkilerle gösterilir.

Nesneye yönelik veri setleri

- Benzer davranış için tek model oluşturma...; modelin uygunluğu test edilmelidir.
- Günümüzdeki pek çok kelime işlemci ve hesap tablosu programında kullanılan nesnelere artık veri setlerinde kullanılmaktadır.
- Nesneye yönelik veri seti Visual Basic, Java, C++, C-Sharp gibi nesneye yönelik bir yazılım dille oluşturulur.

Veri organizasyonunda iletişim alt yapısı

- Hiyerarşik veriseti ağaçların daha gelişmiş hali olan dallanmaların olduğu yapı ortaya çıkmıştır.



Hiyerarşik veriseti

- Hiyerarşik verisetlerinde bilgiler bir ağaç yağısında saklanır.

