



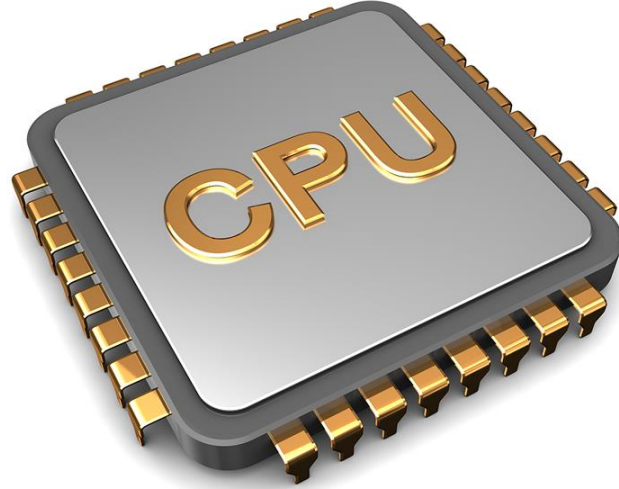
# Veri Kalitesi ve Bütünlüğü

# Veri Nedir?

Bilginin işlenmemiş saf hali



# Verinin Hafızada Tutulması





# Veri Kalitesi

Veriler, kullanım amacına göre gereklilikleri karşılıyorsa kalitelidir denilebilir.



# Veri Kalitesi



# Veri Kalitesi

Amerika Sağlık Bilgi Yönetimi Derneği tarafından geliştirilen  
“Veri Kalitesi Modeli” ne göre;

Doğruluk, Güvenilirlik, Tamlık, Zamanlılık, Kapsamlılık,  
Tutarlılık, Detaylılık.

# Veri Kalitesi

**Doğruluk:** Verilerin kullanılabilir olması için doğru ve geçerli olması son derece önemlidir.



# Veri Kalitesi

## Örnek (**Doğruluk**)

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	2	30
34	İstanbul	2011	NULL	4
...	...	...	...	...



# Veri Kalitesi

**Güvenilirlik:** Verilerin toplanması, işlenmesi, saklanması ve kullanılması aşamalarının sonuçlar aynı olmalıdır.



# Veri Kalitesi

## Örnek (**Güvenirlilik**)

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	2	30
34	İstanbul	2011	NULL	4
...	...	...	...	...

# Veri Kalitesi

**Tamlık:** Gerekli tüm veriler güncel, tam ve uygun şekilde tutulmalıdır



# Veri Kalitesi

## Örnek (**Tamlık**)

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık	Birim
34	İstanbul	2011	1	4	C
06	Ankara	2011	1	-2	C
23	Elazığ	2011	2	30	F
34	İstanbul	2011	NULL	4	C
...	...	...	...	...	...

# Veri Kalitesi

**Zamanlılık:** Veriye ait zaman bilgisi bazı durumlarda önemli olabilmektedir.





# Veri Kalitesi

## Örnek (**Zamanlılık**)

Numara	Adı	Soyadı	Yaş	Boy
1	Murat	Karabatak	23	175
2	Zeynep	Kaya	35	160
3	Gamze	Çetin	18	177
4	Ramazan	Güneş	42	190
...	...	...	...	...

# Veri Kalitesi

**Kapsamlılık:** Kayıtların kapsamlılığı, gerekli bütün veri elemanlarının ve verinin tümünün tanımlanan kapsama göre toplanması ile ölçülür.



# Veri Kalitesi

**Tutarlılık:** Verinin güvenilir olması ve tüm uygulamalarda aynı sonucu vermesi gerekir



# Veri Kalitesi

**Detaylılık:** Verinin detaylı olması, istenen seviyede alt bilgileri içermesi demektir.



# Veri Kalitesi

## Örnek (**Detaylılık**)

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	2	30
34	İstanbul	2011	NULL	4
...	...	...	...	...





Doç. Dr. Murat KARABATAK  
Yazılım Müh. Öğretim Üyesi

# Veri Kalitesi ve Bütünlüğü

# Veri Kalitesini Bozan Sebepler

İnsan Hatası

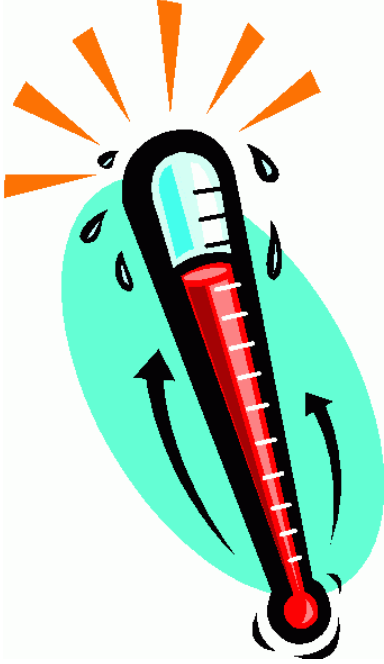


Sistemdeki Bozukluklar

**Sistemsel bir hata oluştu**

# Veri Kalitesini Bozan Sebepler

## Hatalı Veri Toplama Gereçleri



# Veri Kalitesini Bozan Sebepler

## Veri Giriş Problemleri

Telefon No: +90(5 \_\_ )\_ \_ \_ \_ \_

Cinsiyet : Erkek - Bayan  
Erkek - Kadın



# Veri Kalitesini Bozan Sebepler

## Veri Giriş Sırasında Kullanıcıların Hatalı Yorumları

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	2	30
46	K. Maraş	2011	5	14
...	...	...	...	...



# Veri Kalitesini Bozan Sebepler

## Tekrarlı Veri

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	2	30
34	İstanbul	2011	1	4
...	...	...	...	...

# Veri Kalitesini Bozan Sebepler

## Eksik Veri Girişleri

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	NULL	1	-2
23	NULL	2011	2	NULL
34	NULL	2011	1	4
...	...	...	...	...

# Veri Kalitesini Bozan Sebepler

## Geçerli Veri

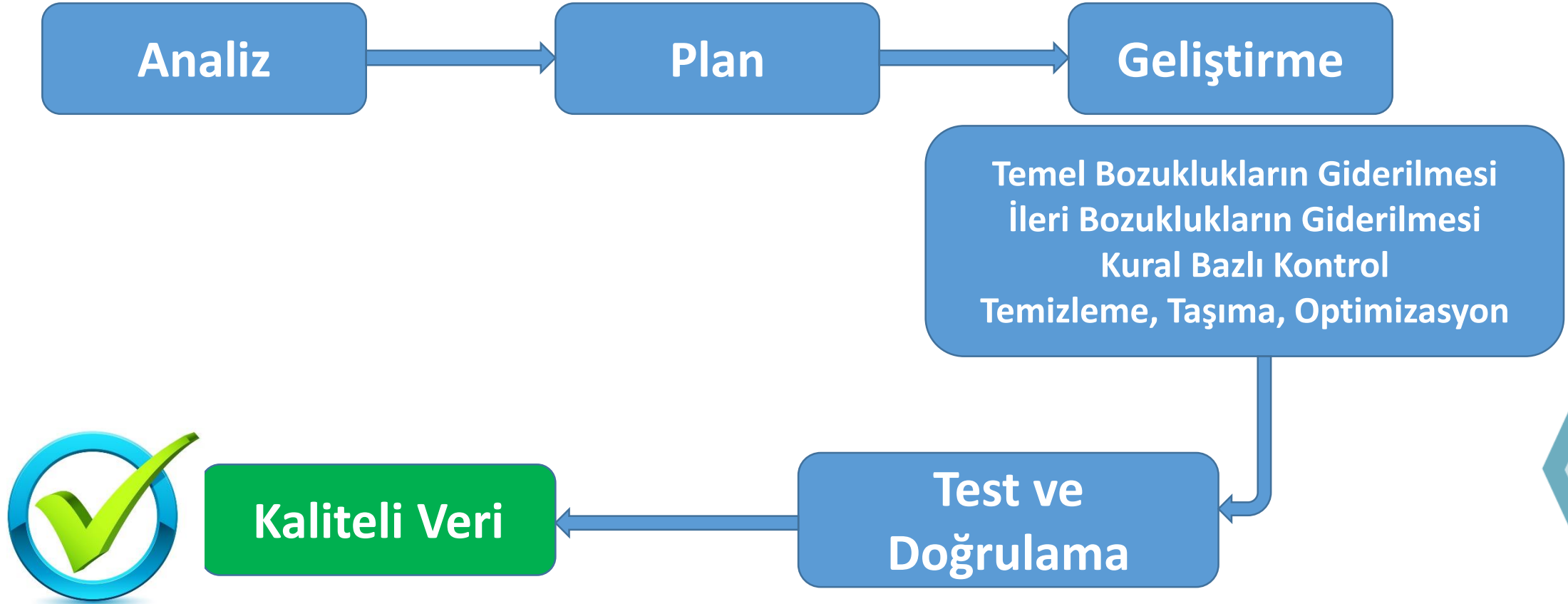
İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	16	154
06	Ankara	2000	1	-2
23	NULL	2011	2	NULL
34	NULL	2011	1	4
...	...	...	...	...



Doç. Dr. Murat KARABATAK  
Yazılım Müh. Öğretim Üyesi

# Veri Kalitesi ve Bütünlüğü

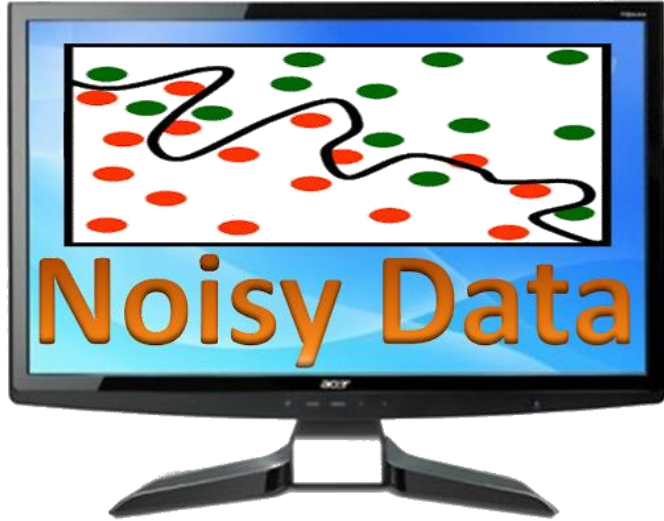
# Veri Kalitesini Arttırma Süreci





# Veri Temizleme

Hatalı, uyumsuz ve eksik kayıtların ortaya çıkarılması işlemi.

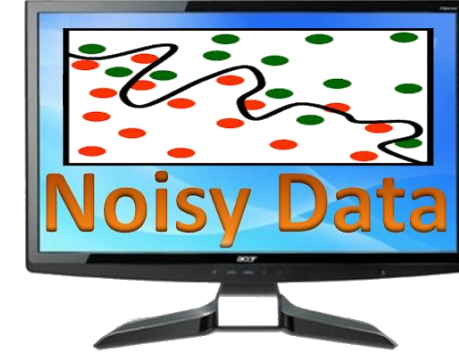


# Veri Temizleme

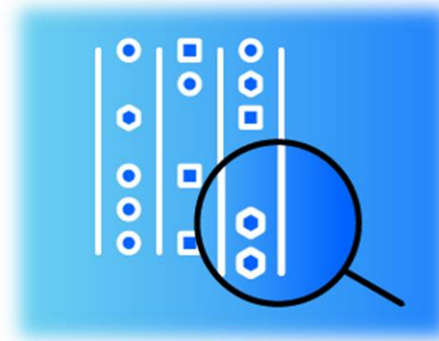
Eksik Veri



Gürültülü Veri



Tutarsız Veri



# Veri Temizleme

## Eksik Kayıtların Silinmesi

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	1	?
23	Elazığ	2012	1	5
...	...	...	...	...

# Veri Temizleme

## Genel Bir Sabit Kullanılması

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	4
06	Ankara	2011	1	-2
23	Elazığ	2011	1	Bilinmiyor
23	Elazığ	2012	1	5
...	...	...	...	...

# Veri Temizleme

Ortalama Bir Değer Girilmesi

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	5
06	Ankara	2011	1	-2
23	Elazığ	2011	1	3
23	Elazığ	2012	1	6
...	...	...	...	...

# Veri Temizleme

Sınıfa Ait Ortalama Bir Değer Girilmesi

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	5
06	Ankara	2011	1	-2
23	Elazığ	2011	1	6
23	Elazığ	2012	1	6
...	...	...	...	...

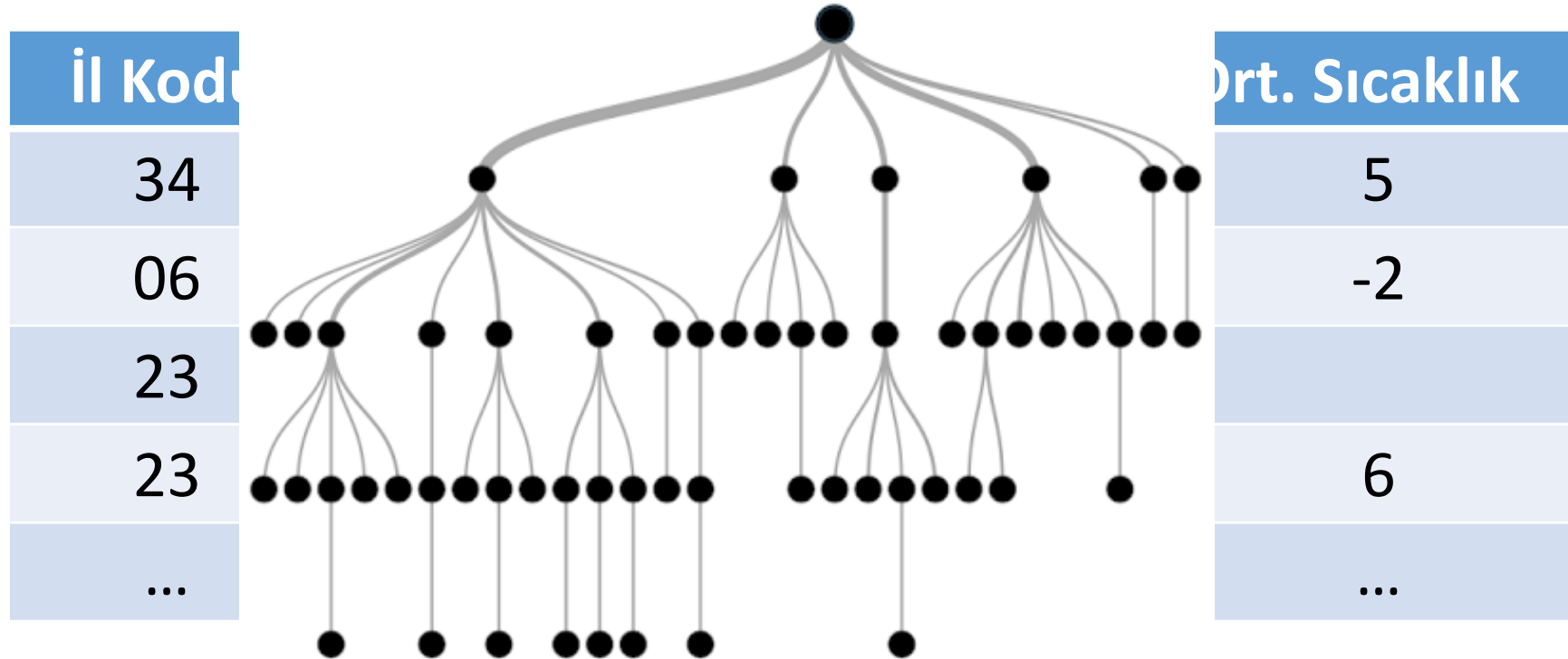
# Veri Temizleme

Regresyon veya Tahmin Modeli Oluşturmak

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	5
06	Ankara	2011	1	-2
23	Elazığ	2011	1	
23	Elazığ	2012	1	6
...	...	...	...	...

# Veri Temizleme

Regresyon veya Tahmin Modeli Oluşturmak





# Veri Temizleme

Regresyon veya Tahmin Modeli Oluşturmak

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	1	5
06	Ankara	2011	1	-2
23	Elazığ	2011	1	5
23	Elazığ	2012	1	6
...	...	...	...	...



# Veri Kalitesi ve Bütünlüğü

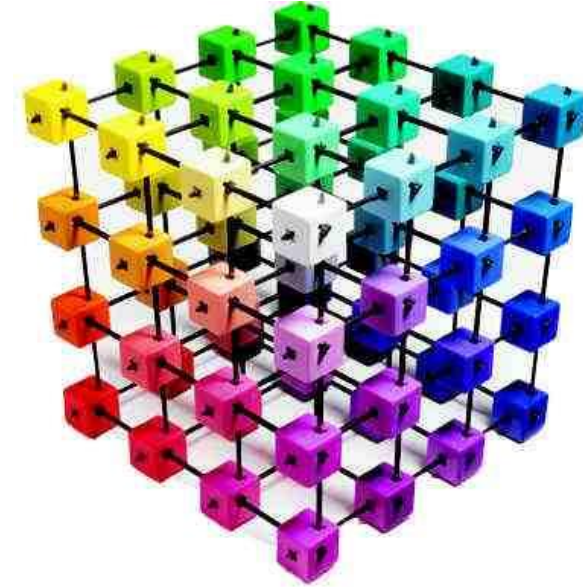
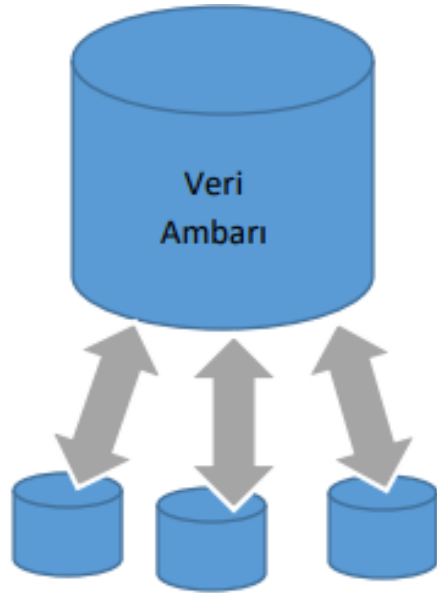
# Veri Optimizasyonu

Büyük tutarlarda verinin işlenmesi, zamansal kısıtlar nedeniyle optimizasyon gerektirir.



# Veri Optimizasyonu

## Veri Birleřtirme ve Veri K p 



# Veri Optimizasyonu

## Veri Birleştirme ve Veri Küpü

No	Ders Kodu	Vize1	Vize2	Final	Büt
1	YMT111	20	50	70	NULL
1	YMT112	60	70	20	80
2	YMT111	80	60	G	50
3	YMT111	50	50	50	NULL
...	...	...	...	...	...

# Veri Optimizasyonu

## Veri Birleştirme ve Veri Küpü

Numara	YMT111	YMT112	YMT113	...
1	53	73	...	...
2	60	...	...	...
3	50	...	...	...
4	...	...	...	...
...	...	...	...	...

# Veri Optimizasyonu

## Boyut İndirgeme

No	Ders Kodu	V1	V2	Fnl	Büt
1	YMT111	20	50	70	NULL
1	YMT112	60	70	20	80
2	YMT111	80	60	G	50
3	YMT111	50	50	50	NULL
...	...	...	...	...	...



No	Ders Kodu	Ort
1	YMT111	53
1	YMT112	73
2	YMT111	60
3	YMT111	50
...	...	...

# Veri Optimizasyonu

## Boyut Sıkıştırma



452 KB

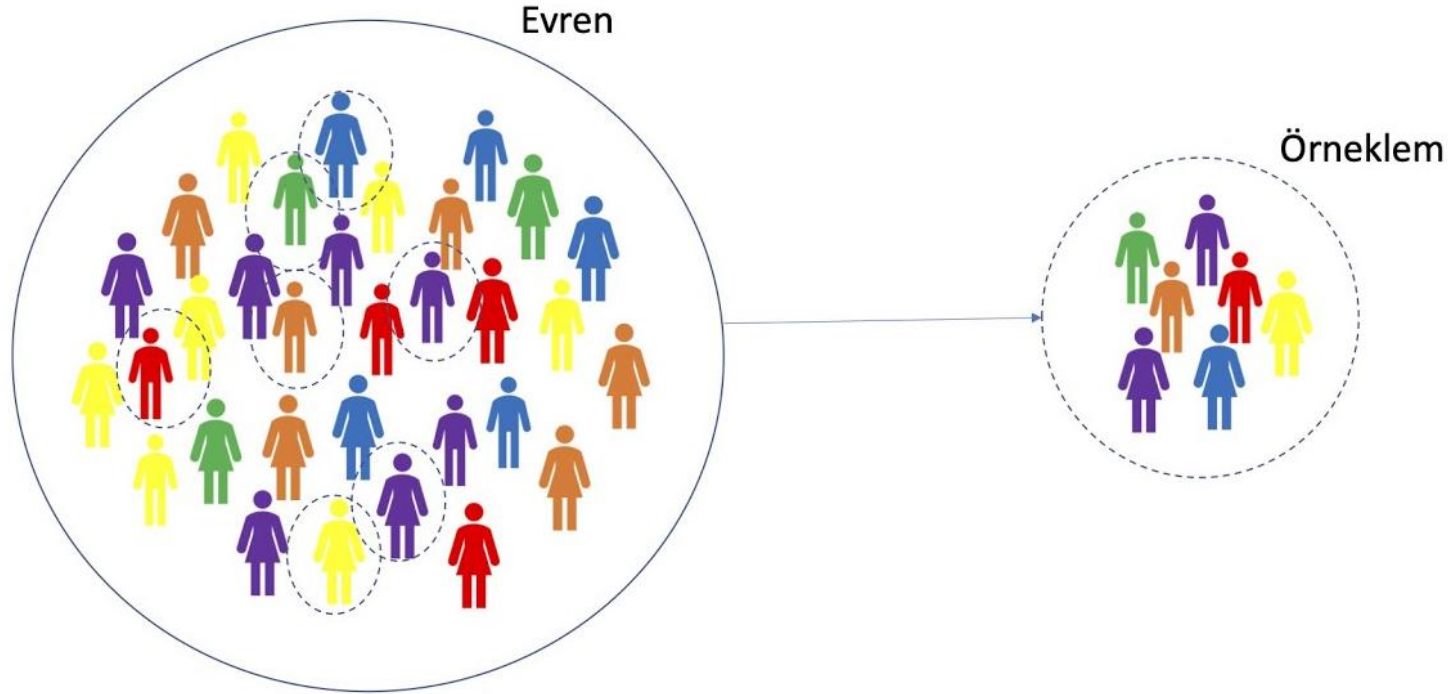


28 KB



# Veri Optimizasyonu

## Örnekleme





# Veri Kalitesi ve Bütünlüğü

# Veri Bütünlüğü

Verinin korunması ve tüm yaşam döngüsü boyunca verilerin doğruluğu ve tutarlılığının güvencesi anlamına gelir.



# Veri Bütünlüğünü Bozan Etkenler

Kötü Niyetli İşlem



İnsan Hatası



Donanım Hatası,



Tasarım Hatası



# Veri Bütünlüğünü Sağlama

Standart kural ve prosedürler kullanılmalı, tasarım aşamasında bir veri tabanı içerisinde bütünlük sağlanmalı.



# Veri Bütünlüğünü Sağlama

## Veri Temizleme ve Bakım





# Veri Bütünlüğünü Sağlama

## Veri Girişi Kontrolünü Sağlama



# Veri Bütünlüğünü Sağlama

## Veri Doğrulama Kuralları





# Veri Bütünlüğünü Sağlama

## Veri Doğrulama Kuralları

İl Kodu	İlin Adı	Yıl	Ay	Ort. Sıcaklık
34	İstanbul	2011	16	154
06	Ankara	2000	1	-2
23	NULL	2011	2	NULL
34	NULL	2011	1	4
...	...	...	...	...

# Veri Bütünlüğünü Sağlama

Veri bütünlüğü kurallarının sıkı bir şekilde uygulanması gerekmektedir.



© Can Stock Photo - csp11263712

# Veri Bütünlüğünü & Veri Kalitesi

Veri bütünlüğü kurallarının sıkı bir şekilde uygulanması gerekmektedir.



VS

